Irena Peeva *Editor*

Commutative Algebra

Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday

Springer

Commutative Algebra

Irena Peeva
Editor

Commutative Algebra

Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday

Springer

Editor
Irena Peeva
Cornell University
Malott Hall
Ithaca, New York
USA

ISBN 978-1-4614-5291-1 ISBN 978-1-4614-5292-8 (eBook)
DOI 10.1007/978-1-4614-5292-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2012955368

Mathematics Subject Classification (2010): 13-02

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Commutative algebra is a vibrant field with activity on many fronts and lively interactions with other fields such as algebraic geometry, algebraic combinatorics, computational algebra, invariant theory, mathematical physics, noncommutative algebra, representation theory, singularity theory, and subspace arrangements. There have been truly exciting recent developments both in core commutative algebra and at the interface with the above listed fields.

The main goal of this book is to showcase the field of commutative algebra in expository papers, especially for the benefit of young mathematicians. This book will aid the readers to broaden their background and gain deeper understanding of the current research in the area.

All papers are dedicated to David Eisenbud in celebration of his many and inspiring contributions to a broad range of topics.

Currently, David Eisenbud is a professor of mathematics at the University of California, Berkeley. He received his Ph.D. in mathematics in 1970 at the University of Chicago under Saunders MacLane and Chris Robson. He was director of the Mathematical Sciences Research Institute (MSRI) from 1997 to 2007 and vice president for mathematics and the physical sciences at the Simons Foundation from 2009 to 2012. From 2003 to 2005 David Eisenbud was president of the American Mathematical Society. In 2006 he was elected a fellow of the American Academy of Arts and Sciences.

Ithaca, NY, USA — Irena Peeva

Peeva was partially supported by NSF grant DMS-1100046.

Contents

Lazarsfeld–Mukai Bundles and Applications

Marian Aprodu

Introduction

Lazarsfeld–Mukai bundles appeared naturally in connection with two completely different important problems in algebraic geometry from the 1980s. The first problem, solved by Lazarsfeld, was to find explicit examples of smooth curves which are generic in the sense of Brill–Noether–Petri [18]. The second problem was the classification of prime Fano manifolds of coindex 3 [23]. More recently, Lazarsfeld–Mukai bundles have found applications to syzygies and higher-rank Brill–Noether theory.

The common feature of all these research topics is the central role played by $K3$ surfaces and their hyperplane sections. For the Brill–Noether–Petri genericity, Lazarsfeld proves that a general curve in a linear system that generates the Picard group of a $K3$ surface satisfies this condition. For the classification of prime Fano manifolds of coindex 3, after having proved the existence of smooth fundamental divisors, one uses the geometry of a two-dimensional linear section which is a very general $K3$ surface.

The idea behind this definition is that the Brill–Noether theory of smooth curves on a $K3$ surface, also called $K3$ *sections*, is governed by higher-rank vector bundles on the surface. To be more precise, consider S a $K3$ surface (considered always to be smooth, complex, projective), C a smooth curve on S of genus ≥ 2, and $|A|$ a base-point-free pencil on C. If we attempt to lift the linear system $|A|$ to the surface S, in most cases, we will fail. For instance, $|A|$ cannot lift to a pencil on S if C generates $\mathrm{Pic}(S)$ or if S does not contain any elliptic curve at

M. Aprodu (✉)
Romanian Academy, Institute of Mathematics "Simion Stoilow" P.O. Box 1-764,
RO 014700, Bucharest, Romania

Şcoala Normală Superioară Bucureşti, Calea Griviţei 21, RO-010702, Bucharest, Romania
e-mail: marian.aprodu@imar.ro

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_1,

all. However, interpreting a general divisor in $|A|$ as a zero-dimensional subscheme of S, it is natural to try and find a rank-two bundle E on S and a global section of E whose scheme of zeros coincides with the divisor in question. Varying the divisor, one should exhibit in fact a two-dimensional space of global sections of E. The effective construction of E is realized through elementary modifications, see Sect. 1, and this is precisely a Lazarsfeld–Mukai bundle of rank two. The passage to higher ranks is natural, if we start with a complete, higher-dimensional, base-point-free linear system on C. At the end, we obtain vector bundles with unusually high number of global sections, which provide us with a rich geometric environment.

The structure of this chapter is as follows. In the first section, we recall the definition of Lazarsfeld–Mukai bundles and its first properties. We note equivalent conditions for a bundle to be Lazarsfeld–Mukai in Sect. 1.1, and we discuss simplicity in the rank-two case in Sect. 1.2. The relation with the Petri conjecture and the classification of Mukai manifolds, the original motivating problems for the definition, are considered in Sects. 1.3 and 1.4, respectively. In Sect. 2 we treat the problem of constancy of invariants in a given linear system. For small gonalities, Saint-Donat and Reid proved that minimal pencils on $K3$ sections are induced from elliptic pencils on the $K3$ surface; we present a short proof using Lazarsfeld–Mukai bundles in Sect. 2.1. Harris and Mumford conjectured that the gonality should always be constant. We discuss the evolution of this conjecture, from Donagi–Morrison's counterexample, Sect. 2.1, to Green–Lazarsfeld's reformulation in terms of Clifford index, Sect. 2.2 and to Ciliberto–Pareschi's results on the subject, Sect. 2.3. The works around this problem emphasized the importance of parameter spaces of Lazarsfeld–Mukai bundles. We conclude the section with a discussion of dimension calculations of these spaces, Sect. 2.4, which are applied afterwards to Green's conjecture. Sect. 3 is devoted to Koszul cohomology and notably to Green's conjecture for $K3$ sections. After recalling the definition and the motivations that led to the definition, we discuss the statement of Green's conjecture, and we sketch the proof for $K3$ sections. Voisin's approach using punctual Hilbert schemes, which is an essential ingredient, is examined in Sect. 3.3. Lazarsfeld–Mukai bundles are fundamental objects in this topic, and their role is outlined in Sect. 3.4. The final step in the solution of Green's conjecture for $K3$ sections is tackled in Sect. 3.5. We conclude this chapter with a short discussion on Farkas–Ortega's new applications of Lazarsfeld–Mukai bundles to Mercat's conjecture (which belongs to the rapidly developing higher-dimensional Brill–Noether theory), Sect. 4.

Notation. The additive and the multiplicative notation for divisors and line bundles will be mixed sometimes. If E is a vector bundle on X and $L \in \mathrm{Pic}(X)$, we set $E(-L) := E \otimes L^*$; this notation will be used especially when E is replaced by the canonical bundle K_C of a curve C.

1 Definition, Properties, the First Applications

1.1 Definition and First Properties

We fix S a smooth, complex, projective $K3$ surface and L a globally generated line bundle on S with $L^2 = 2g-2$. Let $C \in |L|$ be a smooth curve and A be a base-point-free line bundle in $W^r_d(C) \setminus W^{r+1}_d(C)$. As mentioned in the Introduction, the definition of Lazarsfeld–Mukai bundles emerged from the attempt to lift the linear system A to the surface S. Since it is virtually impossible to lift it to another linear system, a higher-rank vector bundle is constructed such that $H^0(C,A)$ corresponds to an $(r+1)$-dimensional space of global sections. Hence $|A|$ lifts to a higher-rank analogue of a linear system.

The kernel of the evaluation of sections of A

$$0 \to F_{C,A} \to H^0(C,A) \otimes \mathcal{O}_S \xrightarrow{\text{ev}} A \to 0 \tag{1}$$

is a vector bundle of rank $(r+1)$.

Definition 1.1 (Lazarsfeld [18], Mukai [23]). The *Lazarsfeld–Mukai bundle* $E_{C,A}$ associated to the pair (C,A) is the dual of $F_{C,A}$.

By dualizing the sequence (1) we obtain the short exact sequence

$$0 \to H^0(C,A)^* \otimes \mathcal{O}_S \to E_{C,A} \to K_C(-A) \to 0, \tag{2}$$

and hence $E_{C,A}$ is obtained from the trivial bundle by modifying it along the curve C and comes equipped with a natural $(r+1)$-dimensional space of global sections as planned.

We note here the first properties of $E_{C,A}$:

Proposition 1.2 (Lazarsfeld). *The invariants of E are the following:*

(1) $\det(E_{C,A}) = L$.
(2) $c_2(E_{C,A}) = d$.
(3) $h^0(S,E_{C,A}) = h^0(C,A) + h^1(C,A) = 2r - d + 1 + g$.
(4) $h^1(S,E_{C,A}) = h^2(S,E_{C,A}) = 0$.
(5) $\chi(S, E_{C,A} \otimes F_{C,A}) = 2(1-\rho(g,r,d))$. *where* $\rho(g,r,d) = g-(r+1)(g-d+r)$.
(6) *$E_{C,A}$ is globally generated off the base locus of $K_C(-A)$; in particular, $E_{C,A}$ is globally generated if $K_C(-A)$ is globally generated.*

It is natural to ask conversely if given E a vector bundle on S with $\text{rk}(E) = r+1$, $h^1(S,E) = h^2(S,E) = 0$, and $\det(E) = L$, E is the Lazarsfeld–Mukai bundle associated to a pair (C,A). To this end, note that there is a rational map

$$h_E : G(r+1, H^0(S,E)) \dashrightarrow |L|$$

defined in the following way. A general subspace $\Lambda \in G(r+1, H^0(S,E))$ is mapped to the degeneracy locus of the evaluation map: $\text{ev}_\Lambda : \Lambda \otimes \mathcal{O}_S \to E$.

If the image $h_E(\Lambda)$ is a smooth curve $C \in |L|$, we set $\mathrm{Coker}(\mathrm{ev}_\Lambda) := K_C(-A)$, where $A \in \mathrm{Pic}(C)$ and $\deg(A) = c_2(E)$, and observe that $E = E_{C,A}$. Indeed, since $h^1(S,E) = 0$, A is globally generated, and from $h^2(S,E) = 0$ it follows that $\Lambda \cong H^0(C,A)^*$. The conclusion is that:

Proposition 1.3. *A rank-$(r+1)$ vector bundle E on S is a Lazarsfeld–Mukai bundle if and only if $H^1(S,E) = H^2(S,E) = 0$ and there exists an $(r+1)$-dimensional subspace of sections $\Lambda \subset H^0(S,E)$, such that the degeneracy locus of the morphism ev_Λ is a smooth curve. In particular, being a Lazarsfeld–Mukai vector bundle is an open condition.*

Note that there might be different pairs with the same Lazarsfeld–Mukai bundles, the difference being given by the corresponding spaces of global sections.

1.2 Simple and Non-simple Lazarsfeld–Mukai Bundles

We keep the notation from the previous subsection. In the original situation, the bundles used by Lazarsfeld [18] and Mukai [23] are simple. The non-simple Lazarsfeld–Mukai bundles are, however, equally useful [3, 5]. For instance, Lazarsfeld's argument is partly based on an analysis of the non-simple bundles.

Proposition 1.2 already shows that for $\rho(g,r,d) < 0$ the associated Lazarsfeld–Mukai bundle cannot be simple. The necessity of making a distinction between simple and non-simple bundles for nonnegative ρ will become more evident in the next sections.

In the rank-two case, one can give a precise description [6] of non-simple Lazarsfeld–Mukai bundles, see also [5] Lemma 2.1:

Lemma 1.4 (Donagi–Morrison). *Let $E_{C,A}$ be a non-simple Lazarsfeld–Mukai bundle. Then there exist line bundles $M, N \in \mathrm{Pic}(S)$ such that $h^0(S,M)$, $h^0(S,N) \geq 2$, N is globally generated, and there exists a locally complete intersection subscheme ξ of S, either of dimension zero or the empty set, such that $E_{C,A}$ is expressed as an extension*

$$0 \to M \to E_{C,A} \to N \otimes I_\xi \to 0. \tag{3}$$

Moreover, if $h^0(S, M \otimes N^*) = 0$, then $\xi = \emptyset$ and the extension splits.

One can prove furthermore that $h^1(S,N) = 0$, [3] Remark 3.6.

We say that (3) is the *Donagi–Morrison extension* associated to $E_{C,A}$. This notion makes perfect sense as this extension is uniquely determined by the vector bundle, if it is indecomposable [3]. Actually, a *decomposable* Lazarsfeld–Mukai bundle E cannot be expressed as an extension (3) with $\xi \neq \emptyset$, and hence a Donagi–Morrison extension is always unique, up to a permutation of factors in the decomposable case. Moreover, a Lazarsfeld–Mukai bundle is decomposable if and only if the corresponding Donagi–Morrison extension is trivial.

In the higher-rank case, we do not have such a precise description.[1] However, a similar sufficiently strong statement is still valid [18, 19, 26].

Proposition 1.5 (Lazarsfeld). *Notation as above. If $E_{C,A}$ is not simple, then the linear system $|L|$ contains a reducible or a multiple curve.*

In the rank-two case, this statement comes from the decomposition $L \cong M \otimes N$.

1.3 The Petri Conjecture Without Degenerations

A smooth curve of genus g is said to satisfy *Petri's condition*, or to be *Brill–Noether–Petri generic*, if the multiplication map (the Petri map)

$$\mu_{0,A} : H^0(C, A) \otimes H^0(C, K_C(-A)) \to H^0(C, K_C),$$

is injective for any line bundle A on C. One consequence of this condition is that all the Brill–Noether loci $W^r_d(C)$ have the expected dimension and are smooth away from $W^{r+1}_d(C)$; recall that the tangent space at the point $[A]$ to $W^r_d(C)$ is naturally isomorphic to the dual of $\text{Coker}(\mu_{0,A})$ [4]. The Petri conjecture, proved by degenerations by Gieseker, states that a general curve satisfies Petri's condition. Lazarsfeld [18] found a simpler and elegant proof without degenerations by analyzing curves on very general $K3$ surfaces.

Lazarsfeld's idea is to relate the Petri maps to the Lazarsfeld–Mukai bundles; this relation is valid in general and has many other applications. Suppose, as in the previous subsections, that S is a $K3$ surface and L is a globally generated line bundle on S. For the moment, we do not need to assume that L generates the Picard group. E. Arbarello and M. Cornalba constructed a scheme $\mathcal{W}^r_d(|L|)$ parameterizing pairs (C, A) with $C \in |L|$ smooth and $A \in W^r_d(C)$ and a morphism

$$\pi_S : \mathcal{W}^r_d(|L|) \to |L|.$$

Assume that $A \in W^r_d(C) \setminus W^{r+1}_d(C)$ is globally generated, and consider M_A the vector bundle of rank r on C defined as the kernel of the evaluation map

$$0 \to M_A \to H^0(C, A) \otimes \mathcal{O}_C \xrightarrow{\text{ev}} A \to 0. \tag{4}$$

Twisting (4) with $K_C \otimes A^*$, we obtain the following description of the kernel of the Petri map:[2]

$$\text{Ker}(\mu_{0,A}) = H^0(C, M_A \otimes K_C \otimes A^*).$$

---

[1] In fact, we do have a Harder–Narasimhan filtration, but we cannot control all the factors.

[2] This ingenious procedure is an efficient replacement of the base-point-free pencil trick; "it has killed the base-point-free pencil trick," to quote Enrico Arbarello.

There is another exact sequence on C

$$0 \to \mathcal{O}_C \to F_{C,A}|_C \otimes K_C \otimes A^* \to M_A \otimes K_C \otimes A^* \to 0,$$

and from the defining sequence of $E_{C,A}$ one obtains the exact sequence on S

$$0 \to H^0(C,A)^* \otimes F_{C,A} \to E_{C,A} \otimes F_{C,A} \to F_{C,A}|_C \otimes K_C \otimes A^* \to 0.$$

From the vanishing of $h^0(C, F_{C,A})$ and of $h^1(C, F_{C,A})$, we obtain

$$H^0(C, E_{C,A} \otimes F_{C,A}) = H^0(C, F_{C,A}|_C \otimes K_C \otimes A^*).$$

Suppose that $\mathcal{W} \subset \mathcal{W}^r_d(|L|)$ is a dominating component and $(C, A) \in \mathcal{W}$ is an element such that A is globally generated and $h^0(C, A) = r + 1$. A deformation-theoretic argument shows that if the Lazarsfeld–Mukai bundle $E_{C,A}$ is simple, then the coboundary map $H^0(C, M_A \otimes K_C \otimes A^*) \to H^1(C, \mathcal{O}_C)$ is zero [26], which eventually implies the injectivity of $\mu_{0,A}$.

By reduction to complete base-point-free bundles on the curve [18, 26] this analysis yields:

Theorem 1.6 (Lazarsfeld). *Let C be a smooth curve of genus $g \geq 2$ on a $K3$ surface S, and assume that any divisor in the linear system $|C|$ is reduced and irreducible. Then a generic element in the linear system $|C|$ is Brill–Noether–Petri generic.*

A particularly interesting case is when the Picard group of S is generated by L and $\rho(g, r, d) = 0$. Obviously, the condition $\rho = 0$ can be realized only for composite genera, as $g = (r+1)(g-d+r)$, for example, $r = 1$ and g even. Under these assumptions, there is a unique Lazarsfeld–Mukai bundle E with $c_1(E) = L$ and $c_2(E) = d$, and different pairs (C, A) correspond to different $\Lambda \in G(r+1, H^0(S, E))$; in other words the natural rational map $G(r+1, H^0(S, E)) \dashrightarrow \mathcal{W}^r_d(|L|)$ is dominating. Note that E must be stable and globally generated.

1.4 Mukai Manifolds of Picard Number One

A Fano manifold X of dimension $n \geq 3$ and index $n-2$ (i.e., of coindex 3) is called a *Mukai manifold*.[3] In the classification, special attention is given to prime Fano manifolds: note that if $n \geq 7$, X is automatically prime as shown by Wisniewski; see, for example, [16].

Assume that the Picard group of X is generated by an ample line bundle L, and let the sectional genus g be the integer $(L^n)/2 + 1$. Mukai and Gushel used vector bundle techniques to obtain a complete classification of these manifolds. A first major obstacle is to prove that the fundamental linear system contains indeed a

[3] Some authors consider that Mukai manifolds have dimension four or more.

smooth element, aspect which is settled by Shokurov and Mella; see, for example, [16]. Then the $(g+n-2)$-dimensional linear system $|L|$ is base-point-free, and a general linear section with respect to the generator of the Picard group is a $K3$ surface. More precisely, if $\mathrm{Pic}(X) = \mathbb{Z}\cdot L$, then for $H_1, \cdots, H_{n-2}$ general elements in the fundamental linear system $|L|$, $S := H_1 \cap \cdots \cap H_{n-2}$ is scheme-theoretically a $K3$ surface. Note that if $n \geq 4$ and $i \geq 3$, the intersection $H_1 \cap \cdots \cap H_{n-i}$ is again a Fano manifold of coindex 3.

Mukai noticed that the fundamental linear system either is very ample, and the image of X is projectively normal or is associated to a double covering of $\mathbb{P}^n$ $(g=2)$ or of the hyper-quadric $Q^n \subset \mathbb{P}^{n+1}$ $(g=3)$. The difficulty of the problem is thus to classify all the possible cases where $|L|$ is normally generated, called *of the first species*. Taking linear sections one reduces (not quite immediately) to the case $n=3$ [16] p.110.

For simplicity, let us assume that X is a prime Fano 3-fold of index 1. If $g=4$ and $g=5$, X is a complete intersection; hence the hard cases begin with genus 6. A hyperplane section S is a $K3$ surface, and, by a result of Moishezon, $\mathrm{Pic}(S)$ is generated by $L|_S$.

Let us denote by $\mathcal{F}_g$ the moduli space of polarized $K3$ surfaces of degree $2g-2$ and $\mathcal{M}_g$ the moduli space of genus-g curves. There are two nice facts in Mukai's proof involving these two moduli spaces. His first observation is that if there exists a prime Fano 3-fold X of the first species of genus $g \geq 6$ and index 1, the rational map $\phi_g : \mathcal{F}_g \dashrightarrow \mathcal{M}_g$ is *not* generically finite [24]. The second nice fact is that ϕ_g is generically finite if and only if $g=11$ or $g \geq 13$ [24].[4] Hence, one is reduced to study the genera $6 \leq g \leq 12$ with $g \neq 11$. At this point, Lazarsfeld–Mukai bundles are employed. By the discussion from Sect. 1.3, for any decomposition $g = (r+1)(g-d+r)$, with $r \geq 1$, $d \leq g-1$, there exists a unique Lazarsfeld–Mukai bundle E of rank $(r+1)$. It has already been noticed that the bundle E is stable and globally generated. Moreover, the determinant map

$$\det : \wedge^{r+1} H^0(S,E) \to H^0(S,L)$$

is surjective [23], and hence it induces a linear embedding

$$\mathbb{P}H^0(S,L)^* \hookrightarrow \mathbb{P}(\wedge^{r+1} H^0(S,E)^*).$$

Following [23], we have a commutative diagram

$$\begin{array}{ccc} S & \xrightarrow{\phi_E} & G \\ \Big\downarrow \phi_{|L|} & & \Big\downarrow \text{Pluecker} \\ \mathbb{P}H^0(L)^* & \hookrightarrow & \mathbb{P}(\wedge^{r+1} H^0(E)^*) \end{array}$$

[4] In genus 11, it is actually birational [25].

where $G := G(r+1, H^0(S, E)^*)$ and ϕ_E is given by E. This diagram shows that S is embedded in a suitable linear section of the Grassmannian G. Moreover, this diagram extends over X: by a result of Fujita, E extends to a stable vector bundle on X, and the diagram over X is obtained for similar reasons. Hence X is a linear section of a Grassmannian. By induction on the dimension, X is contained in a *maximal* Mukai manifold, which is also a linear section of the Grassmannian. A complete list of maximal Mukai manifolds is given in [23]. Notice that in genus 12, the maximal Mukai manifolds are threefold already.

2 Constancy of Invariants of $K3$–Sections

2.1 Constancy of the Gonality. I

In his analysis of linear systems on $K3$ surfaces Saint–Donat [28] shows that any smooth curve which is linearly equivalent to a hyperelliptic or trigonal curve is also hyperelliptic, respectively trigonal. The idea was to prove that the minimal pencils are induced by elliptic pencils defined on the surface. This result was sensibly extended by Reid [27] who proved the following existence result:

Theorem 2.1 (Reid). *Let C be a smooth curve of genus g on a $K3$ surface S and A be a complete, base-point-free g^1_d on C. If*

$$\frac{d^2}{4} + d + 2 < g,$$

then A is the restriction of an elliptic pencil on S.

It is a good occasion to present here, as a direct application of techniques involving Lazarsfeld–Mukai bundles, an alternate shorter proof of Reid's theorem.

Proof. We use the notation of previous sections. By the hypothesis, the Lazarsfeld–Mukai bundle E is not simple, and hence we have a unique Donagi–Morrison extension

$$0 \to M \to E \to N \otimes I_\xi \to 0,$$

with ξ of length ℓ. Note that $M \cdot N = d - \ell \leq d$. By the Hodge index theorem, we have $(M^2) \cdot (N^2) \leq (M \cdot N)^2 \leq d^2$, whereas from $M + N = C$ we obtain $(M^2) = 2(g-1-d) - (N^2)$, hence

$$(N^2) \leq \frac{d^2}{2(g-1-d) - (N^2)}.$$

Therefore, the even integer $x := (N^2)$ satisfies the following inequality $x^2 - 2x(g-1-d) + d^2 \geq 0$. The hypothesis shows that the above inequality fails for $x \geq 2$, and hence N must be an elliptic pencil. □

In conclusion, for small values, the gonality[5] is constant in the linear system. Motivated by these facts, Harris and Mumford conjectured that *the gonality of* $K3$*-sections should always be constant* [14].

This conjecture is unfortunately wrong as stated: Donagi and Morrison [6] gave the following counterexample:

Example 2.2. Let $S \to \mathbb{P}^2$ be a double cover branched along a smooth sextic and L be the pull-back of $\mathcal{O}_{\mathbb{P}^2}(3)$. The curves in $|L|$ have all genus 10. The general curve $C \in |L|$ is isomorphic to a smooth plane sextic, and hence it is pentagonal. On the other hand, the pull-back of a general smooth plane cubic Γ is a double cover of Γ, and thus it is tetragonal.

2.2 Constancy of the Clifford Index

Building on his work on Koszul cohomology and its relations with geometry, M. Green proposed a reformulation of the Harris-Mumford conjecture replacing the gonality by the Clifford index.

Recall that the *Clifford index* of a nonempty linear system $|A|$ on a smooth curve C is the codimension of the image of the natural addition map $|A| \times |K_C(-A)| \to |K_C|$. This definition is nontrivial only for relevant linear systems $|A|$, i.e., such that both $|A|$ and $|K_C(-A)|$ are at least one-dimensional; such an A is said to *contribute to the Clifford index*. The *Clifford index of* C is the minimum of all the Clifford indices taken over the linear systems that contribute to the Clifford index and is denoted by $\mathrm{Cliff}(C)$. The Clifford index is related to the gonality by the following inequalities

$$\mathrm{gon}(C) - 3 \leq \mathrm{Cliff}(C) \leq \mathrm{gon}(C) - 2,$$

and curves with $\mathrm{gon}(C) - 3 = \mathrm{Cliff}(C)$ are very rare: typical examples are plane curves and Eisenbud–Lange–Martens–Schreyer curves [8, 17].[6]

From the Brill–Noether theory, we obtain the bound $\mathrm{Cliff}(C) \leq [(g-1)/2]$ (and, likewise, $\mathrm{gon}(C) \leq [(g+3)/2]$), and it is known that the equality is achieved for general curves. The Clifford index is in fact a measure of how special a curve is in the moduli space.

The precise statement obtained by Green and Lazarsfeld is the following [12]:

Theorem 2.3 (Green–Lazarsfeld). *Let* S *be a* $K3$ *surface and* $C \subset S$ *be a smooth irreducible curve of genus* $g \geq 2$*. Then* $\mathrm{Cliff}(C') = \mathrm{Cliff}(C)$ *for every smooth curve* $C' \in |C|$*. Furthermore, if* $\mathrm{Cliff}(C)$ *is strictly less than the generic value*

[5] The gonality $\mathrm{gon}(C)$ of a curve C is the minimal degree of a morphism from C to the projective line.

[6] It is conjectured that the only other examples should be some half-canonical curves of even genus and maximal gonality [8]; however, this conjecture seems to be very difficult.

$[(g-1)/2]$, *then there exists a line bundle* M *on* S *whose restriction to any smooth curve* $C' \in |C|$ *computes the Clifford index of* C'.

The proof strategy is based on a reduction method of the associated Lazarsfeld–Mukai bundles. The bundle M is obtained from the properties of the reductions; we refer to [12] for details.

From the Clifford index viewpoint, Donagi–Morrison's example is not different from the other cases. Indeed, all smooth curves in $|L|$ have Clifford index 2. We shall see in the next subsection that Donagi–Morrison's example is truly an isolated exception for the constancy of the gonality.

2.3 Constancy of the Gonality. II

As discussed above, the Green–Lazarsfeld proof of the constancy of the Clifford index was mainly based on the analysis of Lazarsfeld–Mukai bundles. It is natural to try and explain the peculiarity of Donagi–Morrison's example from this point of view. This was done in [5]. The surprising answer found by Ciliberto and Pareschi [5] (see also [6]) is the following:

Theorem 2.4 (Ciliberto–Pareschi). *Let* S *be a* $K3$ *surface and* L *be an ample line bundle on* S. *If the gonality of the smooth curves in* $|L|$ *is not constant, then* S *and* L *are as in Donagi–Morrison's example.*

Theorem 2.4 was refined by Knutsen [17] who replaced ampleness by the more general condition that L be globally generated. The extended setup covers also the case of exceptional curves, as introduced by Eisenbud, Lange, Martens, and Schreyer [8].

The proof of Theorem 2.4 consists of a thorough analysis of the loci $\mathcal{W}^1_d(|L|)$, where d is the minimal gonality of smooth curves in $|L|$, through the associated Lazarsfeld–Mukai bundles. The authors identify Donagi–Morrison's example in the following way:

Theorem 2.5 (Ciliberto–Pareschi). *Let* S *be a* $K3$ *surface and* L *be an ample line bundle on* S. *If the gonality of smooth curves in* $|L|$ *is not constant and if there is a pair* $(C, A) \in \mathcal{W}^1_d(|L|)$ *such that* $h^1(S, E_{C,A} \otimes F_{C,A}) = 0$, *then* S *and* L *are as in Donagi–Morrison's example.*

To conclude the proof of Theorem 2.4, Ciliberto and Pareschi prove that non-constancy of the gonality implies the existence of a pair (C, A) with $h^1(S, E_{C,A} \otimes F_{C,A}) = 0$; see [5] Proposition 2.4.

It is worth to notice that, in Example 2.2, if C is the inverse image of a plane cubic and A is a g^1_4 (the pull-back of an involution), then $E_{C,A}$ is the pull-back of $\mathcal{O}_{\mathbb{P}^2}(1) \oplus \mathcal{O}_{\mathbb{P}^2}(2)$ [5], and hence the vanishing of $h^1(S, E_{C,A} \otimes F_{C,A})$ is guaranteed in this case.

2.4 Parameter Spaces of Lazarsfeld–Mukai Bundles and Dimension of Brill–Noether Loci

We have already seen that the Brill–Noether loci are smooth of expected dimension at pairs corresponding to simple Lazarsfeld–Mukai bundles. It is interesting to know what is the dimension of these loci at other points as well. Precisely, we look for a uniform bound on the dimension of Brill–Noether loci of general curves in a linear system.

A first step was made by Ciliberto and Pareschi [5] who proved, as a necessary step in Theorem 2.4, that an ample curve of gonality strictly less than the generic value, general in its linear system, carries finitely many minimal pencils. This result was extended to other Brill–Noether loci [3], proving a phenomenon of *linear growth* with the degree; see below. Let us mention that, for the moment, the only results in this direction are known to hold for pencils [3] and nets [20].

As before, we consider S a $K3$ surface and L a globally generated line bundle on S. In order to parameterize all pairs (C, A) with non-simple Lazarsfeld–Mukai bundles, we need a global construction. We fix a nontrivial globally generated line bundle N on S with $H^0(L(-2N)) \neq 0$ and an integer $\ell \geq 0$. We set $M := L(-N)$ and $g := 1 + L^2/2$. Define $\widetilde{\mathcal{P}}_{N,\ell}$ to be the family of *vector bundles* of rank 2 on S given by nontrivial extensions

$$0 \to M \to E \to N \otimes I_\xi \to 0, \tag{5}$$

where ξ is a zero-dimensional locally complete intersection subscheme (or the empty set) of S of length ℓ, and set

$$\mathcal{P}_{N,\ell} := \{[E] \in \widetilde{\mathcal{P}}_{N,\ell} :\ h^1(S, E) = h^2(S, E) = 0\}.$$

Equivalently (by Riemann–Roch), $[E] \in \mathcal{P}_{N,\ell}$ if and only if $h^0(S, E) = g - c_2(E) + 3$ and $h^1(S, E) = 0$. Note that any non-simple Lazarsfeld–Mukai bundle on S with determinant L belongs to some family $\mathcal{P}_{N,\ell}$, from Lemma 1.4. The family $\mathcal{P}_{N,\ell}$, which, a priori, might be the empty set, is an open Zariski subset of a projective bundle of the Hilbert scheme $S^{[\ell]}$.

Assuming that $\mathcal{P}_{N,\ell} \neq \emptyset$, we consider the Grassmann bundle $\mathcal{G}_{N,\ell}$ over $\mathcal{P}_{N,\ell}$ classifying pairs (E, Λ) with $[E] \in \mathcal{P}_{N,\ell}$ and $\Lambda \in \mathrm{G}(2, H^0(S, E))$. If $d := c_2(E)$ we define the rational map $h_{N,\ell} : \mathcal{G}_{N,\ell} \dashrightarrow \mathcal{W}^1_d(|L|)$, by setting $h_{N,\ell}(E, \Lambda) := (C_\Lambda, A_\Lambda)$, where $A_\Lambda \in \mathrm{Pic}^d(C_\Lambda)$ is such that the following exact sequence on S holds:

$$0 \to \Lambda \otimes \mathcal{O}_S \stackrel{\mathrm{ev}_\Lambda}{\to} E \to K_{C_\Lambda} \otimes A^*_\Lambda \to 0.$$

One computes $\dim\ \mathcal{G}_{N,\ell} = g + \ell + h^0(S, M \otimes N^*)$. If we assume furthermore that $\mathcal{P}_{N,\ell}$ contains a Lazarsfeld–Mukai vector bundle E on S with $c_2(E) = d$ and consider $\mathcal{W} \subset \mathcal{W}^1_d(|L|)$ the closure of the image of the rational map $h_{N,\ell} : \mathcal{G}_{N,\ell} \dashrightarrow \mathcal{W}^1_d(|L|)$, then we find $\dim\ \mathcal{W} = g + d - M \cdot N = g + \ell$.

On the other hand, if $C \in |L|$ has Clifford dimension one and A is a globally generated line bundle on C with $h^0(C, A) = 2$ and $[E_{C,A}] \in \mathcal{P}_{N,\ell}$, then $M \cdot N \geq \text{gon}(C)$.

These considerations on the indecomposable case, together with a simpler analysis of decomposable bundles, yield finally [3]:

Theorem 2.6. *Let S be a $K3$ surface and L a globally generated line bundle on S, such that general curves in $|L|$ are of Clifford dimension one. Suppose that $\rho(g, 1, k) \leq 0$, where $L^2 = 2g - 2$ and k is the (constant) gonality of all smooth curves in $|L|$. Then for a general curve $C \in |L|$, we have*

$$\dim W^1_{k+d}(C) = d \textit{ for all } 0 \leq d \leq g - 2k + 2. \tag{6}$$

The condition (6) is called the *linear growth condition*. It is equivalent to

$$\dim W^1_{g-k+2}(C) = \rho(g, 1, g - k + 2) = g - 2k + 2.$$

Note that the condition that C carry finitely many minimal pencils, which is a part of (6), appears explicitly in [5]. It is directly related to the constancy of the gonality discussed before.

3 Green's Conjecture for Curves on $K3$ Surfaces

3.1 Koszul Cohomology

Let X be a (not necessarily smooth) complex, irreducible, projective variety and $L \in \text{Pic}(X)$ globally generated. The Euler sequence on the projective space $\mathbb{P}(H^0(X, L)^*)$ pulls back to a short exact sequence of vector bundles on X

$$0 \to M_L \to H^0(X, L) \otimes \mathcal{O}_X \to L \to 0. \tag{7}$$

After taking exterior powers in the sequence (7), twisting with multiples of L and going to global sections, we obtain an exact sequence for any nonnegative p and q:

$$0 \to H^0(\wedge^{p+1} M_L \otimes L^{q-1}) \to \wedge^{p+1} H^0(L) \otimes H^0(L^{q-1}) \xrightarrow{\delta} H^0(\wedge^p M_L \otimes L^q). \tag{8}$$

The finite-dimensional vector space $K_{p,q}(X, L) := \text{Coker}(\delta)$ is called the *Koszul cohomology space*[7] of X with values in L [11, 13, 19]. Observe that $K_{p,q}$ can be defined alternatively as:

[7] The indices p and q are usually forgotten when defining Koszul cohomology.

$$K_{p,q}(X,L) = \mathrm{Ker}\left(H^1(\wedge^{p+1}M_L \otimes L^{q-1}) \to \wedge^{p+1}H^0(L) \otimes H^1(L^{q-1})\right),$$

description which is particularly useful when X is a curve.

Several versions are used in practice, for example, replace $H^0(L)$ in (7) by a subspace that generates L or twist (8) by $\mathcal{F} \otimes L^{q-1}$ where $\mathcal{F}$ is a coherent sheaf. For our presentation, however, we do not need to discuss these natural generalizations.

Composing the maps

$$\wedge^{p+1}H^0(L) \otimes H^0(L^{q-1}) \xrightarrow{\delta} H^0(\wedge^p M_L \otimes L^q) \hookrightarrow \wedge^p H^0(L) \otimes H^0(L^q)$$

we obtain, by iteration, a complex

$$\wedge^{p+1}H^0(L) \otimes H^0(L^{q-1}) \to \wedge^p H^0(L) \otimes H^0(L^q) \to \wedge^{p-1}H^0(L) \otimes H^0(L^{q+1})$$

whose cohomology at the middle is $K_{p,q}(X,L)$, and this is the definition given by Green [11].

An important property of Koszul cohomology is upper-semicontinuity in flat families with constant cohomology; in particular, vanishing of Koszul cohomology is an open property in such families. For curves, constancy of h^1 is a consequence of flatness and of constancy of h^0, as shown by the Riemann–Roch theorem.

The original motivation for studying Koszul cohomology spaces was given by the relation with minimal resolutions over the polynomial ring. More precisely, if L is very ample, then the Koszul cohomology computes the minimal resolution of the graded module

$$R(X,L) := \bigoplus_q H^0(X, L^q)$$

over the polynomial ring [11, 13]; see also [2, 7], in the sense that any graded piece that appears in the minimal resolution is (non-canonically) isomorphic to a $K_{p,q}$. If the image of X is projectively normal, this module coincides with the homogeneous coordinate ring of X. The projective normality of X can also be read off Koszul cohomology, being characterized by the vanishing condition $K_{0,q}(X,L) = 0$ for all $q \geq 2$. Furthermore, for a projectively normal X, the homogeneous ideal is generated by quadrics if and only if $K_{1,q}(X,L) = 0$ for all $q \geq 2$.[8] The phenomenon continues as follows: if X is projectively normal and the homogeneous ideal is generated by quadrics, then the relations between the generators are linear if and only if $K_{2,q}(X,L) = 0$ for all $q \geq 2$ etc, whence the relation with syzygies [11].

Other notable application of Koszul cohomology is the description of Castelnuovo-Mumford regularity, which coincides with, [2, 11]

$$\min_q\{K_{p,q}(X,L) = 0, \text{ for all } p\}.$$

[8]The dimension of $K_{1,q}$ indicates the number of generators of degree $(q+1)$ in the homogeneous ideal.

Perhaps the most striking property of Koszul cohomology, discovered by Green and Lazarsfeld [11, Appendix], is a consequence of a nonvanishing result:

Theorem 3.1 (Green–Lazarsfeld). *Suppose X is smooth and $L = L_1 \otimes L_2$ with $r_i := h^0(X, L_i) - 1 \geq 1$. Then $K_{r_1+r_2-1,1}(X, L) \neq 0$.*

Note that the spaces $K_{p,1}$ have the following particular attribute: if $K_{p,1} \neq 0$ for some $p \geq 1$ then $K_{p',1} \neq 0$ for all $1 \leq p' \leq p$. This is obviously false for $K_{p,q}$ with $q \geq 2$.

Theorem 3.1 shows that the existence of nontrivial decompositions of L reflects onto the existence of nontrivial Koszul classes in some space $K_{p,1}$. Its most important applications are for curves, in particular for canonical curves, case which is discussed in the next subsection. In the higher-dimensional cases, for surfaces, for instance, the meaning of Theorem 3.1 becomes more transparent if it is accompanied by a restriction theorem which compares the Koszul cohomology of X with the Koszul cohomology of the linear sections [11]:

Theorem 3.2 (Green). *Suppose X is smooth and $h^1(X, L^q) = 0$ for all $q \geq 1$. Then for any connected reduced divisor $Y \in |L|$, the restriction map induces an isomorphism*

$$K_{p,q}(X, L) \xrightarrow{\sim} K_{p,q}(Y, L|_Y),$$

for all p and q.

The vanishing of $h^1(X, \mathcal{O}_X)$ suffices to prove that the restriction is an isomorphism between the spaces $K_{p,1}$ [2].

In the next subsections, we shall apply Theorem 3.2 for $K3$ sections.

Corollary 3.3. *Let C be a smooth connected curve on a $K3$ surface S. Then*

$$K_{p,q}(S, \mathcal{O}_S(C)) \cong K_{p,q}(C, K_C)$$

for all p and q.

One direct consequence is a duality theorem for Koszul cohomology of $K3$ surfaces.[9] It shows the symmetry of the table containing the dimensions of the spaces $K_{p,q}$, called *the Betti table*.

3.2 Statement of Green's Conjecture

Let us particularize Theorem 3.1 for a canonical curve. Consider C a smooth curve and choose a decomposition $K_C = A \otimes K_C(-A)$. Theorem 3.1 applies only if $h^0(C, A) \geq 2$ and $h^1(C, A) \geq 2$, i.e., if A contributes to the Clifford index.

[9]Duality for Koszul cohomology of curves follows from Serre's duality. For higher-dimensional manifolds, some supplementary vanishing conditions are required [11, 13].

The quantity $r_1 + r_2 - 1$ which appears in the statement equals $g - \mathrm{Cliff}(A) - 2$, and hence, if A *computes* the Clifford index, we obtain the following:

Theorem 3.4 (Green–Lazarsfeld). *For any smooth curve C of genus g Clifford index c we have $K_{g-c-2,1}(C, K_C) \neq 0$.*

It is natural to determine whether or not this result is sharp, question which is addressed in the statement Green's conjecture:

Conjecture 3.5 (Green). Let C be a smooth curve. For all $p \geq g - c - 1$, we have $K_{p,1}(C, K_C) = 0$.

For the moment, Green's conjecture remains a hard open problem. At the same time, strong evidence has been discovered. For instance, it is known to hold for general curves [31,32], for curves of odd genus and maximal Clifford index [15,32], for general curves of given gonality [30, 31],[10] [29], for curves with small Brill–Noether loci [1], for plane curves [21], for curves on $K3$ surfaces [3, 31, 32], etc.; see also [2] for a discussion.

We shall consider in the sequel the case of curves on $K3$ surfaces with emphasis on Voisin's approach to the problem and the role played by Lazarsfeld–Mukai bundles. It is interesting to notice that Green's conjecture for $K3$ sections can be formulated directly in the $K3$ setup, as a vanishing result on the moduli space $\mathcal{F}_g$ of polarized $K3$ surfaces. However, in the proof of this statement, as it usually happens in mathematics, we have to exit the $K3$ world, prove a more general result in the extended setup, and return to $K3$ surfaces. The steps we have to take, ordered logically and not chronologically, are the following. In the first, most elaborated step, one finds an example for odd genus [31,32]. At this stage, we are placed in the moduli space $\mathcal{F}_{2k+1}$. Secondly, we exit the $K3$ world, land in $\mathcal{M}_{2k+1}$, and prove the equality of two divisors [15, 31]. The first step is used, and the identification of the divisors extends to their closure over the component Δ_0 of the boundary [1]. In the third step, we jump from a gonality stratum $\mathcal{M}^1_{g,d}$ in a moduli space $\mathcal{M}_g$ to the boundary of another moduli space of stable curves $\overline{\mathcal{M}}_{2k+1}$, where $k = g - d + 1$ [1]. The second step reflects into a vanishing result on an explicit open subset of $\mathcal{M}^1_{g,d}$. Finally one goes back to $K3$ surfaces and applies the latter vanishing result [3] on $\mathcal{F}_g$. In the steps concerned with $K3$ surfaces (first and last), the Lazarsfeld–Mukai bundles are central objects.

3.3 Voisin's Approach

The proof of the generic Green conjecture was achieved by Voisin in two papers [31, 32], using a completely different approach to Koszul cohomology via Hilbert scheme of points.

[10]Voisin's and Teixidor's cases complete each other quite remarkably.

Let X be a complex connected projective manifold and L a line bundle on X. It is obvious that any global section σ is uniquely determined by the collection $\{\sigma(x)\}_x$, where $\sigma(x) \in L|_x \cong \mathbb{C}$ and x belongs to a nonempty open subset of X. One tries to find a similar fact for multisections in $\wedge^n H^0(X, L)$.

Let $\sigma_1 \wedge \cdots \wedge \sigma_n$ be a decomposable element in $\wedge^n H^0(X, L)$ with $n \geq 1$. By analogy with the case $n = 1$, we have to look at the restriction $\sigma_1|_\xi \wedge \cdots \wedge \sigma_n|_\xi \in \wedge^n L|_\xi$ where ξ is now a zero-dimensional subscheme, and it is clear that we need n points for otherwise this restriction would be zero. Note that a zero-dimensional subscheme of length n defines a point in the punctual Hilbert scheme $X^{[n]}$. For technical reasons, we shall restrict to curvilinear subschemes[11] which form a large open subset $X_c^{[n]}$ in a connected component of the Hilbert scheme.[12] Varying $\xi \in X_c^{[n]}$, the collection $\{\sigma_1|_\xi \wedge \cdots \wedge \sigma_n|_\xi\}_\xi$ represents a section in a line bundle described as follows. Put $\Xi_n \subset X_c^{[n]} \times X$ the incidence variety and denote by q and p the projections on the two factors; note that q is finite of degree n. Then $L^{[n]} := q_* p^*(L)$ is a vector bundle of rank n on $X_c^{[n]}$, and the fibre at a point $\xi \in X_c^{[n]}$ is $L^{[n]}|_\xi \cong L|_\xi$. In conclusion, the collection $\{\sigma|_\xi \wedge \cdots \wedge \sigma|_\xi\}_\xi$ defines a section in the line bundle $\det(L^{[n]})$. The map we are looking at $\wedge^n H^0(L) \to H^0(\det(L^{[n]}))$ is deduced from the evaluation map $\mathrm{ev}_n : H^0(L) \otimes \mathcal{O}_{X_c^{[n]}} \to L^{[n]}$, taking $\wedge^n \mathrm{ev}_n$ and applying H^0. It is remarkable that [9, 31, 32]:

Theorem 3.6 (Voisin, Ellingsrud–Göttsche–Lehn). *The map*

$$H^0(\wedge^n \mathrm{ev}_n) : \wedge^n H^0(X, L) \to H^0\left(X_c^{[n]}, \det(L^{[n]})\right)$$

is an isomorphism.

Since the exterior powers of $H^0(L)$ are building blocks for Koszul cohomology, it is natural to believe that the isomorphism above yields a relation between the Koszul cohomology and the Hilbert scheme. To this end, the Koszul differentials must be reinterpreted in the new context.

There is a natural birational morphism[13]

$$\tau : \Xi_{n+1} \to X_c^{[n]} \times X,\ (\xi, x) \mapsto (\xi - x, x)$$

presenting Ξ_{n+1} as the blowup of $X_c^{[n]} \times X$ along Ξ_n. If we denote by D_τ the exceptional locus, we obtain an inclusion [31]

$$q^* \det(L^{[n+1]}) \cong \tau^*(\det(L^{[n]}) \boxtimes L)(-D_\tau) \hookrightarrow \tau^*(\det(L^{[n]}) \boxtimes L)$$

---

[11] A curvilinear subscheme is defined locally, in the classical topology, by $x_1 = \cdots = x_{s-1} = x_s^k = 0$; equivalently, it is locally embedded in a smooth curve.

[12] The connectedness of $X_c^{[n]}$ follows from the observation that a curvilinear subscheme is a deformation of a reduced subscheme.

[13] We see one advantage of working on $X_c^{[n]}$: subtraction makes sense only for curvilinear subschemes.

whence

$$H^0\left(X_c^{[n+1]}, \det(L^{[n+1]})\right) \hookrightarrow H^0(X_c^{[n]} \times X, \det(L^{[n]}) \boxtimes L),$$

identifying the left-hand member with the kernel of a Koszul differential [31]. A version of this identification leads us to [31, 32]:

Theorem 3.7 (Voisin). *For any integers m and n, $K_{n,m}(X, L)$ is isomorphic to the cokernel of the restriction map:*

$$H^0\left(X_c^{[n+1]} \times X, \det(L^{[n+1]}) \boxtimes L^{m-1}\right) \to H^0\left(\Xi_{n+1}, \det(L^{[n+1]}) \boxtimes L^{m-1}|_{\Xi_{n+1}}\right).$$

The vanishing of Koszul cohomology is thus reduced to proving surjectivity of the restriction map above. In general, it is very hard to prove surjectivity directly, and one has to make a suitable base-change [31].

3.4 The Role of Lazarsfeld–Mukai Bundles in the Generic Green Conjecture and Consequences

In order to prove Green's conjecture for general curves, it suffices to exhibit one example of a curve of maximal Clifford index, which verifies the predicted vanishing. Afterwards, the vanishing of Koszul cohomology propagates by semicontinuity. Even so, finding one single example is a task of major difficulty. The curves used by Voisin in [31, 32] are $K3$ sections, and the setups change slightly, according to the parity of the genus. For even genus, we have [31]:

Theorem 3.8 (Voisin). *Suppose that $g = 2k$. Consider S a $K3$ surface with* $\text{Pic}(S) \cong \mathbb{Z} \cdot L$, *$L^2 = 2g-2$, and $C \in |L|$ a smooth curve. Then $K_{k,1}(C, K_C) = 0$.*

For odd genus, the result is [32]:

Theorem 3.9 (Voisin). *Suppose that $g = 2k + 1$. Consider S a $K3$ surface with* $\text{Pic}(S) \cong \mathbb{Z} \cdot L \oplus \mathbb{Z} \cdot \Gamma$, *$L^2 = 2g - 2$, Γ a smooth rational curve. $L \cdot \Gamma = 2$ and $C \in |L|$ a smooth curve. Then $K_{k,1}(C, K_C) = 0$.*

Note that the generic value for the Clifford index in genus g is $[(g-1)/2]$, and hence, in both cases, the prediction made by Green's conjecture for general curve C is precisely $K_{k,1}(C, K_C) = 0$.

There are several reasons for making these choices: the curves have maximal Clifford index, by Theorem 2.3 (and the Clifford dimension is one), the Lazarsfeld–Mukai bundles associated to minimal pencils are L-stable, the hyperplane section theorem applies, etc.

We outline here the role played by Lazarsfeld–Mukai bundles in Voisin's proof and, for simplicity, we restrict to the even-genus case. By the hyperplane section Theorem 3.2, the required vanishing on the curve is equivalent to $K_{k,1}(S, L) = 0$.

From the description of Koszul cohomology in terms of Hilbert schemes, Theorem 3.7, adapting the notation from the previous subsection, one has to prove the surjectivity of the map

$$q^* : H^0\left(S_c^{[n+1]}, \det(L^{[n+1]})\right) \to H^0\left(\Xi_{n+1}, q^*\det(L^{[n+1]})|_{\Xi_{n+1}}\right).$$

The surjectivity is proved after performing a suitable base-change.

We are in the case $\rho(g,1,k+1)=0$; hence there is a unique Lazarsfeld–Mukai bundle E on S associated to all g^1_{k+1} on curves in $|L|$. The uniqueness yields an alternate description of E as extension

$$0 \to \mathcal{O}_S \to E \to L \otimes I_\xi \to 0,$$

where ξ varies in $S_c^{[k+1]}$.

There exists a morphism $\mathbb{P}H^0(S,E) \to S^{[k+1]}$ that sends a global section $s \in H^0(S,E)$ to its zero set $Z(s)$. By restriction to an open subset $\mathbb{P} \subset \mathbb{P}H^0(S,E)$, we obtain a morphism $\mathbb{P} \to S_c^{[k+1]}$, inducing a commutative diagram

$$\begin{array}{ccc} \mathbb{P}' = \mathbb{P} \times_{S_c^{[k+1]}} \Xi_{k+1} & \longrightarrow & \Xi_{k+1} \\ \downarrow{\scriptstyle q'} & & \downarrow{\scriptstyle q} \\ \mathbb{P} & \longrightarrow & S_c^{[k+1]}. \end{array}$$

Set-theoretically

$$\mathbb{P}' = \{(Z(s), x) | s \in H^0(S,E), x \in Z(s)\}.$$

Unfortunately, this very natural base-change does not satisfy the necessary conditions that imply the surjectivity of q^*, [31]. Voisin modifies slightly this construction and replaces $\mathbb{P}$ with another variety related to $\mathbb{P}$ which parameterizes zero-cycles of the form $Z(s) - x + y$ with $[s] \in \mathbb{P}$, $x \in \mathrm{Supp}(Z(s))$ and $y \in S$. It turns out, after numerous elaborated calculations using the rich geometric framework provided by the Lazarsfeld–Mukai bundle, that the new base-change is suitable and the surjectivity of q^* follows from vanishing results on the Grassmannian [31].

In the odd-genus case, Voisin proves first Green's conjecture for smooth curves in $|L+\Gamma|$, which are easily seen to be of maximal Clifford index. The situation on $|L+\Gamma|$ is somewhat close to the setup of Theorem 3.8, and the proof is similar. The next hard part is to descend from the vanishing of $K_{k+1,1}(S, L \otimes \mathcal{O}_S(\Gamma))$ to the vanishing of $K_{k,1}(S,L)$. This step uses again intensively the unique Lazarsfeld–Mukai bundle associated to any g^1_{k+2} on curves in $|L+\Gamma|$.

The odd-genus case is of maximal interest: mixed with Hirschowitz-Ramanan result [15], Theorem 3.9 gives a solution to Green's conjecture for *any* curve of odd genus and maximal Clifford index:

Theorem 3.10 (Hirschowitz–Ramanan, Voisin). *Let C be a smooth curve of odd genus $2k+1 \geq 5$ and Clifford index k. Then $K_{k,1}(C, K_C) = 0$.*

Note that Theorem 3.10 implies the following statement:

Corollary 3.11. *A smooth curve of odd genus and maximal Clifford index has Clifford dimension one.*

The proof of Theorem 3.10 relies on the comparison of two effective divisors on the moduli space of curves $\mathcal{M}_{2k+1}$, one given by the condition $\mathrm{gon}(C) \leq k+1$, which is known to be a divisor from [14], and the second given by $K_{k,1}(C, K_C) \neq 0$. By duality $K_{k,1}(C, K_C) \cong K_{k-2,2}(C, K_C)$. Note that $K_{k-2,2}(C, K_C)$ is isomorphic to

$$\mathrm{Coker}\left(\wedge^k H^0(K_C) \otimes H^0(K_C)/\wedge^{k+1} H^0(K_C) \to H^0(\wedge^{k-1} M_{K_C} \otimes K_C^2)\right)$$

and the two members have the same dimension. The locus of curves with $K_{k,1} \neq 0$ can be described as the degeneracy locus of a morphism between vector bundles of the same dimension, and hence it is a virtual divisor. Theorem 3.9 implies that this locus is not the whole space, and in conclusion it must be an effective divisor. Theorem 3.1 already gives an inclusion between the supports of two divisors in question, and the set-theoretic equality is obtained from a divisor class calculation [15].

3.5 Green's Conjecture for Curves on $K3$ Surfaces

We have already seen that general $K3$ sections have a mild behavior from the Brill–Noether theory viewpoint. In some sense, they behave like general curves in any gonality stratum of the moduli space of curves.

As in the previous subsections, fix a $K3$ surface S and a globally generated line bundle L with $L^2 = 2g-2$ on S, and denote by k the gonality of a general smooth curve in the linear system $|L|$. Suppose that $\rho(g,1,k) \leq 0$ to exclude the case $g = 2k-3$ (when $\rho(g,1,k) = 1$). If in addition the curves in $|L|$ have Clifford dimension one, Theorem 2.6 shows that

$$\dim W^1_{g-k+2}(C) = \rho(g,1,g-k+2) = g-2k+2,$$

property which was called the *linear growth condition*.

This property appears in connection with Green's conjecture [1] for a much larger class of curves:

Theorem 3.12. *If C is any smooth curve of genus $g \geq 6$ and gonality $3 \leq k < [g/2]+2$ with $\dim W^1_{g-k+2}(C) = \rho(g,1,g-k+2)$, then $K_{g-k+1,1}(C, K_C) = 0$.*

One effect of Theorems 3.12 and 3.1 is that an arbitrary curve that satisfies the linear growth condition is automatically of Clifford dimension one and verifies Green's conjecture.

Theorem 3.12 is a consequence of Theorem 3.10 extended over the boundary of the moduli space. Starting from a k-gonal smooth curve $[C] \in \mathcal{M}_g$, by identifying pairs of general points $\{x_i, y_i\} \subset C$ for $i = 0, \dots, g-2k+2$ we produce a stable irreducible curve

$$\big[X := C/(x_0 \sim y_0, \dots, x_{g-2k+2} \sim y_{g-2k+2})\big] \in \overline{\mathcal{M}}_{2(g-k+1)+1},$$

and the Koszul cohomology of C and of X are related by the inclusion $K_{p,1}(C, K_C) \subset K_{p,1}(X, \omega_X)$ for all $p \geq 1$, [31]. If C satisfies the linear growth condition then X has maximal gonality[14] $\mathrm{gon}(X) = g - k + 3$, i.e., X lies outside the closure of the divisor $\mathcal{M}^1_{2(g-k+1)+1, g-k+2}$ consisting of curves with a pencil g^1_{g-k+2}. The class of the failure locus of Green's conjecture on $\overline{\mathcal{M}}_{2(g-k+1)+1}$ is a multiple of the divisor $\overline{\mathcal{M}}^1_{2(g-k+1)+1, g-k+2}$; hence Theorem 3.10 extends to irreducible stable curves of genus $2(g-k+1)+1$ of maximal gonality $(g-k+3)$. In particular, $K_{g-k+1,1}(X, \omega_X) = 0$, implying $K_{g-k+1,1}(C, K_C) = 0$.

Coming back to the original situation, we conclude from Theorems 3.12 and 2.6 and Corollary 3.3 that Green's conjecture holds for a $K3$ section C having Clifford dimension one. If $\mathrm{Cliff}(C) = \mathrm{gon}(C) - 3$, either C is a smooth plane curve or else there exist smooth curves $D, \Gamma \subset S$, with $\Gamma^2 = -2, \Gamma \cdot D = 1$ and $D^2 \geq 2$, such that $C \equiv 2D + \Gamma$ and $\mathrm{Cliff}(C) = \mathrm{Cliff}(\mathcal{O}_C(D))$ [5, 17]. The linear growth condition is no longer satisfied, and this case is treated differently, by degeneration to a reduced curve with two irreducible components [3].

The outcome of this analysis of the Brill–Noether loci is the following [3,31,32]:

Theorem 3.13. *Green's conjecture is valid for any smooth curve on a $K3$ surface.*

Applying Theorem 3.13, Theorem 3.2, and the duality, we obtain a full description of the situations when Koszul cohomology of a $K3$ surface is zero [3]:

Theorem 3.14. *Let S be a $K3$ surface and L a globally generated line bundle with $L^2 = 2g - 2 \geq 2$. The Koszul cohomology group $K_{p,q}(S, L)$ is nonzero if and only if one of the following cases occurs:*

(1) $q = 0$ and $p = 0$, or
(2) $q = 1$, $1 \leq p \leq g - c - 2$, or
(3) $q = 2$ and $c \leq p \leq g - 1$, or
(4) $q = 3$ and $p = g - 2$.

The moral is that the shape of the Betti table, i.e., the distribution of zeros in the table, of a polarized $K3$ surface is completely determined by the geometry of hyperplane sections; this is one of the many situations where algebra and geometry are intricately related.

[14] The gonality for a singular stable curve is defined in terms of admissible covers [14].

4 Counterexamples to Mercat's Conjecture in Rank Two

Starting from Mukai's works, experts tried to generalize the classical Brill–Noether theory to higher-rank vector bundles on curves. Within these extended theories,[15] we note the attempt to find a proper generalization of the Clifford index. H. Lange and P. Newstead proposed the following definition. Let E be a semistable vector bundle of rank n of degree d on a smooth curve C. Put

$$\gamma(E) := \mu(E) - 2\frac{h^0(E)}{n} + 2.$$

Definition 4.15 (Lange–Newstead). The *Clifford index* of rank n of C is

$$\mathrm{Cliff}_n(C) := \min\{\gamma(E) : \ \mu(E) \leq g-1, \ h^0(E) \geq 2n\}.$$

From the definition, it is clear that $\mathrm{Cliff}_1(C) = \mathrm{Cliff}(C)$ and $\mathrm{Cliff}_n(C) \leq \mathrm{Cliff}(C)$ for all n.[16]

Mercat conjectured [22] that $\mathrm{Cliff}_n(C) = \mathrm{Cliff}(C)$. In rank two, the conjecture is known to hold in a number of cases: for general curves of small gonality, i.e., corresponding to a general point in a gonality stratum $\mathcal{M}^1_{g,k}$ for small k (Lange-Newstead), for plane curves (Lange–Newstead), for general curves of genus ≤ 16 (Farkas–Ortega), etc. However, even in rank two, the conjecture is false. It is remarkable that counterexamples are found for curves of maximal Clifford index [10]:

Theorem 4.16 (Farkas–Ortega). *Fix $p \geq 1$, $a \geq 2p+3$. Then there exists a smooth curve of genus $2a+1$ of maximal Clifford index lying on a smooth $K3$ surface S with $\mathrm{Pic}(S) = \mathbb{Z}\cdot C \oplus \mathbb{Z}\cdot H$, $H^2 = 2p+2$, $C^2 = 2g-2$, $C \cdot H = 2a + 2p + 1$, and there exists a stable rank-two vector bundle E with $\det(E) = \mathcal{O}_S(H)$ with $h^0(E) = p+3$, $\gamma(E) = a - \frac{1}{2} < a = \mathrm{Cliff}(A)$, and hence Mercat's conjecture in rank two fails for C.*

The proof uses restriction of Lazarsfeld–Mukai bundles. However, it is interesting that the bundles are not restricted to the same curves to which they are associated. More precisely, the genus of H is $2p+2$ and H has maximal gonality $p+2$. Consider A a minimal pencil on H, and take $E = E_{H,A}$ the associated Lazarsfeld–Mukai bundle. The restriction of E to C is stable and verifies all the required properties.

A particularly interesting case is $g = 11$. In this case, as shown by Mukai [25], a general curve C lies on a unique $K3$ surface S such that C generates $\mathrm{Pic}(S)$.

---

[15] Higher-rank Brill–Noether theory is a major, rapidly growing research field, and it deserves a separate dedicated survey.

[16] For any line bundle A, we have $\gamma(A^{\oplus n}) = \mathrm{Cliff}(A)$.

It is remarkable that the failure locus of Mercat's conjecture in rank two *coincides* with the Noether-Lefschetz divisor

$$\mathcal{NL}^4_{11,13} := \left\{ [C] \in \mathcal{M}_{11} : \begin{array}{l} C \text{ lies on a } K3 \text{ surface } S,\ \mathrm{Pic}(S) \supset \mathbb{Z}\cdot C \oplus \mathbb{Z}\cdot H, \\ H \in \mathrm{Pic}(S) \text{ is nef}, H^2 = 6,\ C\cdot H = 13,\ C^2 = 20 \end{array} \right\}$$

inside the moduli space $\mathcal{M}_{11}$. We refer to [10] for details.

Acknowledgements I am grateful to S. Druel, G. Farkas, and A. Ortega for useful discussions on this subject. This work was partly supported by the grant "Vector bundle techniques in the geometry of complex varieties," PN-II-ID-PCE-2011-3-0288, contract no. 132/05.10.2011.

References

1. Aprodu, M.: Remarks on syzygies of d-gonal curves. Math. Res. Lett. **12**(3), 387–400 (2005)
2. Aprodu, M., Nagel, J.: Koszul cohomology and algebraic geometry. University Lecture Series, vol. 52. American Mathematical Society, Providence (2010)
3. Aprodu, M., Farkas, G.: Green's conjecture for curves on arbitrary K3 surfaces. Compositio Math. **147**, 839–851 (2011)
4. Arbarello, E., Cornalba, M., Griffiths, P.A., Harris, J.: Geometry of algebraic curves, vol. I. Grundlehren der Mathematischen Wissenschaften, vol. 267. Springer, New York (1985)
5. Ciliberto, C., Pareschi, G.: Pencils of minimal degree on curves on a $K3$ surface. J. Reine Angew. Math. **460**, 15–36 (1995)
6. Donagi, R., Morrison, D.R.: Linear systems on $K3$ sections. Diff. J. Geom. **29** 49–64 (1989)
7. Eisenbud, D.: Geometry of syzygies. Graduate Texts in Mathematics, vol. 229. Springer, New York (2005)
8. Eisenbud, D., Lange, H., Martens, G., Schreyer, F.-O.: The Clifford dimension of a projective curve. Compositio Math. **72**, 173–204 (1989)
9. Ellingsrud, G., Göttsche, L., Lehn, M.: On the cobordism class of the Hilbert scheme of a surface. Alg. J.: Geom. **10**, 81–100 (2001)
10. Farkas, G., Ortega, A.: Higher-rank Brill–Noether theory on sections of $K3$ surfaces. Internat. J. Math. **23**(7), 1250075, 18 (2012)
11. Green, M.: Koszul cohomology and the geometry of projective varieties. Diff. J. Geom. **19**, 125–171 (1984)
12. Green, M., Lazarsfeld, R.: Special divisors on curves on a $K3$ surface. Inventiones Math. **89**, 73–90 (1987)
13. Green, M.: Koszul cohomology and geometry. In: Cornalba, M., et al. (eds.) Proceedings of the First College on Riemann Surfaces Held in Trieste, Italy, November 1987, pp. 177–200. World Scientific, Singapore (1989)
14. Harris, J., Mumford, D.: On the Kodaira dimension of the moduli space of curves. Inventiones Math. **67**, 23–86 (1982)
15. Hirschowitz, A., Ramanan, S.: New evidence for Greens conjecture on syzygies of canonical curves. Ann. Sci. École Norm. Sup. **31**, 145–152 (1998)
16. Iskovskih, V.A., Prokhorov, Yu.: Fano varieties. In: Parshin, A.N., Shafarevich, I.R. (eds.) Algebraic Geometry V. Encyclopedia of Mathematical Science. Springer, New York (1999)
17. Knutsen, A.: On two conjectures for curves on $K3$ surfaces. Internat. Math. J. **20**, 1547–1560 (2009)
18. Lazarsfeld, R.: Brill–Noether–Petri without degenerations. Diff. J. Geom. **23**, 299–307 (1986)

19. Lazarsfeld, R.: A sampling of vector bundle techniques in the study of linear series. In: Cornalba, M., et al. (eds.) Proceedings of the first college on Riemann surfaces held in Trieste, Italy, November 1987, pp. 500–559. World Scientific, Singapore (1989)
20. Lelli-Chiesa, M.: Stability of rank-3 Lazarsfeld–Mukai bundles on K3 surfaces. Preprint arXiv:1112.2938.
21. Loose, F.: On the graded Betti numbers of plane algebraic curves. Manuscr. Math. **64**, 503–514 (1989)
22. Mercat, V.: Clifford's theorem and higher rank vector bundles. Internat. Math. J.: **13**, 785–796 (2002)
23. Mukai, S.: Biregular classification of Fano 3-folds and Fano manifolds of coindex 3. Proc. Natl. Acad. Sci. USA **86**, 3000–3002 (1989)
24. Mukai, S.: Fano 3-folds. London Math. Soc. Lect. Note Ser. **179**, 255–263 (1992)
25. Mukai, S.: Curves and $K3$ surfaces of genus eleven. In: Moduli of Vector Bundles. Lecture Notes in Pure Applied Mathematics, vol. 179, pp. 189–197. Dekker, New York (1996)
26. Pareschi, G.: A proof of Lazarsfeld's theorem on curves on $K3$ surfaces. Alg. J. Geom. **4**, 195–200 (1995)
27. Reid, M.: Special linear systems on curves lying on $K3$ surfaces. J. London Math. Soc. (2) **13**, 454–458 (1976)
28. Saint-Donat, B.: Projective models of $K3$ surfaces. American J. Math. **96**(4), 602–639 (1974)
29. Schreyer, F.-O.: Green's conjecture for general p-gonal curves of large genus. Algebraic curves and projective geometry (Trento, 1988), pp. 254–260. Lecture Notes in Mathematics, vol. 1389. Springer, Berlin (1989)
30. Teixidor i Bigas, M.: Green's conjecture for the generic r-gonal curve of genus $g \geq 3r - 7$. Duke Math. J. **111**, 195–222 (2002)
31. Voisin, C.: Green's generic syzygy conjecture for curves of even genus lying on a $K3$ surface. J. European Math. Soc. **4**, 363–404 (2002)
32. Voisin, C.: Green's canonical syzygy conjecture for generic curves of odd genus. Compositio Math. **141**, 1163–1190 (2005)

Some Applications of Commutative Algebra to String Theory

Paul S. Aspinwall

1 Introduction

String theory was first introduced as a model for strong nuclear interactions, then reinterpreted as a model for quantum gravity, and then all fundamental physics. However, one might argue that its most successful applications to date have been in the realm of pure mathematics and geometry. The superstring is most easily understood in ten dimensions. In order to make contact with the observed physical world of four spacetime dimensions, one compactifies on a six-dimensional manifold. This is most easily analyzed in the case where this manifold is a Calabi–Yau threefold. Fortuitously, such varieties happen to be of great mathematical interest.

Historically, geometry, as used by physicists, has generally been differential geometry because of its role in general relativity. More recently, especially because of the use of supersymmetry, string theory has come to rely more heavily on algebraic geometry. Thus tools in commutative algebra have become more useful in recent years. A simplified model of great interest in string theory, known as topological field theory, is where the connections to commutative algebra become most manifest.

The purpose of this chapter is to review some particular applications of commutative algebra to string theory. Developments in recent years in computer packages for commutative algebra, such as Macaulay 2 [1], mean that it is of great practical value if a problem can be translated into a question in commutative algebra. We will discuss three examples where this happens.

The first two applications are closely related and involve the structure of topological field theory itself. Certain products in the theory can be interpreted as

P.S. Aspinwall (✉)
Mathematics Department, Duke University
e-mail: psa@cgtp.duke.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_2,

Ext computations for sheaves on the Calabi–Yau or in terms of matrix factorizations that are very amenable to computer algebra. This may also be viewed intrinsically as an efficient way to compute certain Ext groups. As a by-product, we are also able to see some elements of Hochschild cohomology that are relevant for open-to-closed string transitions.

The final application is related to monodromy. This can be viewed as monodromy of integral 3-cycles in a Calabi–Yau threefold under loops in the moduli space of complex structures or, via mirror symmetry, as automorphisms of the derived category induced by varying the complexified Kähler form. This monodromy is also related to solutions of the well-studied GKZ system of differential equations. We show how this monodromy can be stated in terms of a ring which we compute in a fairly nontrivial example.

2 Categorical Topological Field Theory

2.1 Closed String Theories

Before we look at the applications, we will review a contemporary mathematical picture of string theory. As a string moves through space and time, it sweeps out a surface known as the worldsheet. A central idea in string theory is that one "pulls back" physics from spacetime to the worldsheet, thus reducing physics problems to problems in two-dimensional field theory. Of particular interest are field theories on the worldsheet which are invariant under two-dimensional conformal transformations, i.e., "conformal field theories." Some of the structure of conformal field theories can be determined by much simpler "topological field theories." Fortunately, all the mess and non-rigor of quantum field theory can be avoided in topological field theories, as they have a nice direct categorical description thanks to Atiyah [2] based on some then-unpublished ideas by Segal. We will briefly review this picture here, but we refer to [3], Chap. 2 in [4] and references therein for a more complete treatment.

Let $\mathbf{Cob}_c$ be a category whose objects are closed oriented 1-dimensional manifolds, i.e., disjoint unions of circles. Given a pair of objects, M and N, a morphism is a cobordism from M to N. Diffeomorphic cobordisms are considered equivalent. Note that we preserve orientations in the sense that the boundary of a cobordism is $\overline{M} \coprod N$, where $\overline{M}$ is the orientation-reversed M.

Composing cobordisms gives an obvious category structure on $\mathbf{Cob}_c$, where the identity morphism may be taken to be the cylinder $M \times [0,1]$. $\mathbf{Cob}_c$ is a monoidal category, i.e., it has a tensor product on objects and a unit object. In this case the tensor product is disjoint union, and the unit object is empty.

Let **Vect** be the monoidal category of vector spaces over a field k. Here tensor product is the usual tensor product of vector spaces and the unit object is the one-dimensional space k.

Definition 1. A "closed-string topological field theory" is a functor, respecting the monoidal structure, from $\mathbf{Cob}_c$ to $\mathbf{Vect}$.

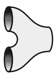
Such theories are specified by very little data. First we specify the vector space associated to a single circle. This is called the "Hilbert space" of the theory. Then we need to give the morphisms of vector spaces associated to a few basic morphisms. Then the full structure of the theory can be derived by sewing these basic morphisms together.

Let X be a Calabi–Yau threefold. That is, X is a quasi-projective complex algebraic variety of dimension three with trivial canonical class. We assume the covering space of X does not have an elliptic curve factor. Usually X will be smooth, and thus we often refer to it as a manifold. One of the richest applications of topological field theories is to such varieties. Given such a manifold, X, there are two associated theories—the A-model and the B-model.

The central object of study in this chapter is the B-model, which is a good deal easier than the A-model and which we define first. The B-model is completely algebraic in nature, although we will assume for purposes of presentation that we always work over $\mathbb{C}$. The data is:

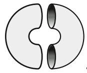
- The Hilbert space associated to a circle is

$$\mathscr{H}_\circ = \bigoplus_{p,q=0}^{3} H^q(X, \wedge^p T), \tag{1}$$

  where T is the tangent sheaf of X. Obviously we have a bigraded structure (p, q) here. The single grading given by $p + q$ is instrinsic to the topological field theory, as we discuss below. In the case that $b_1(X) = 0$, the topological field theory structure factorizes into even and odd $p + q$. We can then consistently restrict attention to the subspace where $p = q$.
- The left cap  is mapped by the functor to a map $k \to \mathscr{H}_\circ$. The image of 1 is $1 \in H^0(X, \mathcal{O}_X)$.
- The right cap  yields a map $\mathscr{H}_\circ \to k$ which is only nonzero on the degree (3, 3) part of $\mathscr{H}_\circ$. If $[A_{ijk}] \in H^3(X, \wedge^3 T)$, then the resulting value is $\int_X \Omega \wedge \bar{\Omega}^{ijk} A_{ijk}$, where Ω is a choice of nonzero holomorphic 3-form and indices are lowered and raised using the Kähler metric.
-  gives a product $\mathscr{H}_\circ \otimes \mathscr{H}_\circ \to \mathscr{H}_\circ$, which is the wedge product (extended to exterior powers of the tangent sheaf) on $\mathscr{H}_\circ$. Note that this wedge product is commutative only up to a sign. This is a *spin theory*, the sense of Sect. 2.1.6 of [4].

This information is enough to completely determine the topological field theory. For example, it follows that:

-  gives a nondegenerate pairing $\mathscr{H}_\circ \otimes \mathscr{H}_\circ \to k$.

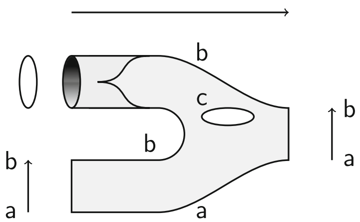
-  yields a map $1 \mapsto \sum_i \psi_i \otimes \psi_i^*$, where $\{\psi_1, \psi_2, \ldots\}$ is a basis for $\mathscr{H}_\circ$, and the ψ_i^*'s are dual with respect to the above pairing.
- The torus, which is the composition , is the map $k \to k$ given by multiplying by the Euler characteristic $\chi(X)$.
- Because of the grading, a Riemann surface of genus $\neq 1$ gives 0.

The A-model depends on the *symplectic* geometry of X. A complexified Kähler form $B + iJ \in H^2(X, \mathbb{C}/\mathbb{Z})$ is part of the basic data on which the model depends. Here J is the usual Kähler form, while $B \in H^2(X, \mathbb{R}/\mathbb{Z})$ represents the "B-field" which is ubiquitous in string theory. The Hilbert space of closed strings is given by the De Rham cohomology of X. The complexity of the A-model comes from the fact that the product $\mathscr{H}_\circ \otimes \mathscr{H}_\circ \to \mathscr{H}_\circ$ depends on "instanton corrections" coming from rational curves (see, e.g., [5]).

2.2 Open–Closed Strings

One obtains a much richer structure if one allows for open strings as well as closed strings. That is, the worldsheet may have a boundary. Thus, we need a new category $\mathbf{Cob}_{oc}$ whose objects are disjoint unions of oriented circles and line segments. The morphisms are cobordisms consisting of manifolds possibly with boundaries.

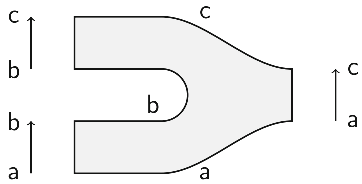
The boundaries allow us to decorate the category further. Each segment of the boundary of a cobordism and each end of a line segment object should be labeled by a "boundary condition." Such boundary conditions are called "D-branes." We consider the set of D-branes as part of the information in $\mathbf{Cob}_{oc}$. A morphism in $\mathbf{Cob}_{oc}$ may look something like



where the letters represent D-branes.

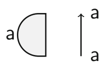
Definition 2. An "open–closed-string topological field theory" specifies a particular set of D-branes and gives a functor, respecting the monoidal structure, from $\mathbf{Cob}_{oc}$ to $\mathbf{Vect}$.

While this definition allows for an arbitrary set of D-branes, the A-model and B-model endow this set with a particular structure. It is key to note that such a field theory immediately gives the set of D-branes themselves a categorical structure, with D-branes as objects. The set of morphisms $\mathrm{Hom}(\mathsf{a}, \mathsf{b})$ is given by the Hilbert space associated with the line interval going from D-brane a to D-brane b. A composition of morphisms is given by



(2)

and the identity in $\mathrm{Hom}(\mathsf{a}, \mathsf{a})$ is the image of 1 in the map $k \to \mathrm{Hom}(\mathsf{a}, \mathsf{a})$ induced by the cobordism from nothing to the line interval:



(3)

Furthermore, D-branes can "bind together" to form other D-branes which gives this category a triangulated structure, but we will not use this structure in this chapter. In the case of the B-model on a Calabi–Yau threefold X, it is widely believed that the D-brane category is the *bounded derived category of coherent sheaves on* X.[1]

Similarly, the D-brane category for the A-model is generally taken to be the (derived) Fukaya category [9]. Objects in this category are (certain) Lagrangian 3-cycles in X. Fortunately, for the purposes of this chapter, we need to know little further about this category.

Definition 3. Two Calabi–Yau threefolds, X and Y, are said to be a "mirror pair" if there is a natural isomorphism between the open–closed topological field theory functors associated with the A-model on Y and the B-model on X. This implies "homological mirror symmetry" which is an equivalence between the D-brane categories.

[1]The basic idea of a proof was proposed in [6] and further studied in [4, 7, 8]. The current physics proofs only say that D-brane category has the derived category as a full subcategory.

2.3 Topological Conformal Field Theory

One may obtain a further richer theory which is "almost topological" by retaining a little more information than just the diffeomorphism class of a cobordism. Fixing two objects M and N, let $\mathscr{M}(M,N)$ be the moduli space of Riemann surfaces giving a cobordism from M to N. The category dg$\mathbf{Cob}_{oc}$ is defined to have the same objects as $\mathbf{Cob}_{oc}$, but now the morphisms will be geometric cochain complexes on $\mathscr{M}(M,N)$. This gives dg$\mathbf{Cob}_{oc}$ the structure of a dg-category. Let dg**Vect** be the dg-category of cochain complexes of vector spaces. A "topological conformal field theory" (TCFT)[2] is then a dg-functor from dg$\mathbf{Cob}_{oc}$, together with a set of D-brane labels, to dg**Vect**. We refer to [11] for a full description.

Note that a Hilbert space is now a complex of vector spaces rather than a single vector space. The original topological field theory Hilbert space can be recovered from this simply by taking the cohomology groups of these complexes. The homological grading naturally gives a grading to the Hilbert spaces involved. These gradings were seen above in the A and B-models.

The dg-category structure now extends to the D-brane category too. The TCFT associated a line interval from a to b with a chain complex of vector spaces. The homological grading on this complex can be formally extended to the D-branes themselves by defining $\mathrm{Hom}(\mathsf{a}[i],\mathsf{b}[j])$ in the category to be the complex $\mathrm{Hom}(\mathsf{a},\mathsf{b})$ shifted left $j-i$ places. This agrees with the usual translation functor on the derived category.

In the case of the B-model on a smooth Calabi–Yau threefold, this dg structure arises naturally from Dolbeault cohomology on vector bundles. That is, if two D-branes are locally free sheaves $\mathscr{E}$ and $\mathscr{F}$, the complex of morphisms between them is given by

$$0 \longrightarrow \Gamma(\mathscr{A}^{0,0}\otimes\mathscr{E}^\vee\otimes\mathscr{F}) \xrightarrow{\bar\partial} \Gamma(\mathscr{A}^{0,1}\otimes\mathscr{E}^\vee\otimes\mathscr{F}) \xrightarrow{\bar\partial} \ldots, \tag{4}$$

where $\mathscr{A}^{0,q}$ is the sheaf of C^∞ $(0,q)$-forms and Γ is the global section functor. This can then be extended in the usual way when the D-branes are complexes of locally free sheaves.

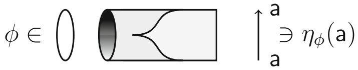
2.4 Hochschild Cohomology

We saw above that a collection of D-brane labels and a topological field theory give the data to construct a D-brane category. In the TCFT case the converse is also true.

---

[2] In physics language this is phrased as coupling a topological field theory to topological gravity. The term TCFT actually means something quite different in the physics literature [10].

That is, given a D-brane dg-category, one can completely reconstruct the TCFT. This was proven by Costello [11], but we can give a quick idea of how this works here. First define η_ϕ mapping a D-brane object a to a morphism $\eta_\phi(\mathsf{a}) \in \mathrm{Hom}(\mathsf{a}, \mathsf{a})$ as the image of ϕ in the Hilbert space of closed strings via the diagram:



(5)

Now note the equivalence of the morphisms:

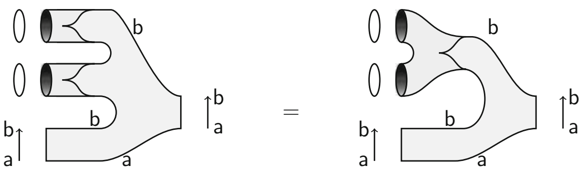
(6)

Let id be the identity functor from the D-brane category to itself. Rewriting (6) in algebraic form gives a commutative diagram that states precisely that η_ϕ *is a natural transformation from* id *to* id. Given that the D-brane category is a dg-category, the work of [12] shows that this can be viewed as Hochschild cohomology. To be precise, for a D-brane category $\mathcal{B}$, the Hochschild cohomology is written

$$HH^i(\mathcal{B}) = \mathrm{Nat}(\mathrm{id}_{\mathcal{B}}, \mathrm{id}_{\mathcal{B}}[i]). \tag{7}$$

In this way one identifies the Hochschild cohomology of the D-brane category with the Hilbert space of closed string states (complete with its homological grading) [11].

Furthermore we can recover the "pants diagram" combining two closed strings. This comes from the natural product rule on Hochschild cohomology coming from combining natural transformations and can be put in diagrammatic form:



(8)

We close this section by noting that all the constructions above for the B-model can be stated in purely algebraic terms. This is what allows us to attack the B-model by using the tools of commutative algebra and what makes the B-model much more amenable to study than the A-model.

3 Toric Geometry and Phases

A versatile context in which to apply the tools of commutative algebra is to consider compact Calabi–Yau threefolds which are complete intersections in toric varieties. We also need to consider noncompact Calabi–Yau varieties (of dimension possibly greater than 3) which are toric varieties themselves. We begin with the latter.

Let N be a lattice of rank d. Let $\mathscr{P}$ be a convex polytope in $N \otimes \mathbb{R}$ such that the vertices of the convex hull lie in N. Furthermore, we demand that $\mathscr{P}$ lies in a hyperplane of $N \otimes \mathbb{R}$. Let $\mathscr{A}$ denote the set of points $\mathscr{P} \cap N$ and let n denote the number of elements of $\mathscr{A}$.

The coordinates of the points of $\mathscr{A}$ form a $d \times n$ matrix defining a map $A : \mathbb{Z}^{\oplus n} \to N$ which we assume is surjective. Form an exact sequence

$$0 \longrightarrow L \longrightarrow \mathbb{Z}^{\oplus n} \xrightarrow{A} N \longrightarrow 0, \tag{9}$$

where L is the "lattice of relations" of rank $r = n - d$. Dual to this we have

$$0 \longrightarrow M \longrightarrow \mathbb{Z}^{\oplus n} \xrightarrow{\Phi} D \longrightarrow 0, \tag{10}$$

where Φ is the $r \times n$ matrix of "charges" of the points in $\mathscr{A}$.

Let

$$S = \mathbb{C}[x_1, \ldots, x_n]. \tag{11}$$

The matrix Φ gives an r-fold multigrading to this ring. In other words, we have a $(\mathbb{C}^*)^r$ torus action:

$$x_i \mapsto \lambda_1^{\Phi_{1i}} \lambda_2^{\Phi_{2i}} \ldots \lambda_r^{\Phi_{ri}} x_i, \tag{12}$$

where $\lambda_j \in \mathbb{C}^*$. Let S_0 be the $(\mathbb{C}^*)^r$-invariant subalgebra of S. The algebra S then decomposes into a sum of S_0-modules labeled by their r-fold grading:

$$S = \bigoplus_{\boldsymbol{\alpha} \in D} S_{\boldsymbol{\alpha}}, \tag{13}$$

where $D \cong \mathbb{Z}^{\oplus r}$ from (10). As usual we denote a shift in grading by parentheses, i.e., $S(\boldsymbol{\alpha})_{\boldsymbol{\beta}} = S_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$.

Consider a simplicial decomposition of the point set $\mathscr{A}$ which is *regular* in the sense of [13]. This simplicial decomposition may or may not include points in the interior of the convex hull of $\mathscr{A}$. We refer to a choice of simplicial decomposition as a "phase."

To each phase we associate the "Cox ideal" defined in [14] as follows.

Definition 1. Let $\Sigma = \{\sigma_1, \sigma_2, \ldots\}$ denote the set of simplices of maximum dimension. If σ is a simplex, we say $i \in \sigma$ if the ith element of $\mathscr{A}$ is a vertex of σ. Then

$$B_\Sigma = \left(\prod_{i \notin \sigma_1} x_i, \prod_{i \notin \sigma_2} x_i, \dots \right). \tag{14}$$

Clearly B_Σ is a square-free monomial ideal in S.

Definition 2. Let $V(B_\Sigma)$ denote the subvariety of $\mathbb{C}^n$ given by B_Σ. Then

$$Z_\Sigma = \frac{\mathbb{C}^n - V(B_\Sigma)}{(\mathbb{C}^*)^r}. \tag{15}$$

For examples of this construction we refer to Sect. 5.1 of [15]. Cox [14] shows that there is a correspondence between finitely generated graded S-modules and coherent sheaves on a smooth Z_Σ which follows the usual correspondence between sheaves and projective varieties as in Chap. II.5 of [16]. If U is an S-module, we denote $\widetilde{U}$ as the corresponding sheaf. $\widetilde{U}$ is zero as a sheaf if and only if U is killed by some power of B_Σ. This yields

Proposition 1. *Assume Z_Σ is a smooth toric variety. Then*

$$\mathbf{D}^b(Z_\Sigma) = \frac{\mathbf{D}^b(\mathrm{gr}{-}S)}{T_\Sigma}, \tag{16}$$

where $\mathbf{D}^b(\mathrm{gr}{-}S)$ is the bounded derived category of finitely generated multigraded S-modules and T_Σ is the full triangulated subcategory generated by modules killed by a power of B_Σ. This quotient of triangulated categories is the "Verdier quotient" (see, e.g., [17]).

Actually we can extend this proposition to the case where Z_Σ is not smooth. A graded S-module corresponds to a sheaf on the quotient *stack* $\mathbb{C}^n/(\mathbb{C}^*)^r$. That is, $\mathbf{D}^b(\mathrm{gr}{-}S)$ is the derived category of coherent sheaves on this stack. Quotienting by T_Σ in (16) is equivalent to removing the pointset $V(B_\Sigma)$. Thus we see a direct correspondence between (15) and (16) and the above proposition remains true in the singular case so long as we view sheaves in this quotient stack sense. Actually, this is very natural from the physics perspective. One may construct the topological B-model as a symplectic quotient in terms of the "gauged linear σ-model" [18]. D-branes can be described directly in this context [19]. It has also been argued elsewhere [20] that stacks are the correct language for D-branes.

We now have the following statement from physics [5] which we assume, for purposes of this chapter, to be true:

Physics Proposition 1. *The B-model on X does not depend on the Kähler form of X.*

This has an immediate consequence. The Kähler form is associated to the moment map of the symplectic quotient of the gauged linear σ-model, which, when varied, can change the triangulation Σ. Thus, assuming the above proposition, we have

Proposition 2. $\mathbf{D}^b(Z_\Sigma)$ *does not depend on the chosen triangulation* Σ*, i.e., it is independent of the phase.*

This is quite easy to verify directly when r (the rank of torus action) is one, as done in [19, 21]. The combinatorics become more intricate for $r > 1$.

Let I_Σ be the Alexander dual of the monomial ideal B_Σ. I_Σ is then the Stanley–Reisner ideal of the triangulation Σ [22]. Suppose

$$I_\Sigma = \langle m_1, m_2, \ldots \rangle, \tag{17}$$

for monomials $m_1, m_2, \ldots$. Then we have a primary decomposition

$$B_\Sigma = m_1^\vee \cap m_2^\vee \cap \ldots, \tag{18}$$

where, if $m_j = x_\alpha x_\beta x_\gamma \ldots$, then $m_j^\vee = \langle x_\alpha, x_\beta, x_\gamma, \ldots \rangle$.

It follows that $(S/m_j^\vee)(\boldsymbol{\alpha})$ is annihilated by B_Σ, where $(\boldsymbol{\alpha})$ is any shift in multidegree. In fact these modules form the building blocks of T_Σ as proven in [23]:

Proposition 3. T_Σ *is the smallest triangulated full subcategory of* $\mathbf{D}^b(\text{gr}-S)$ *containing the objects* $(S/m_j^\vee)(\boldsymbol{\alpha})$*. That is,* T_Σ *is generated by iteratively applying mapping cones between objects of the form* $(S/m_j^\vee)(\boldsymbol{\alpha})[n]$ *for any* $\boldsymbol{\alpha}$ *and* n*.*

3.1 Tilting Collections

The derived category of an arbitrary algebraic variety can be difficult to describe in concrete terms. However, in some cases, the description can be simplified due to the existence of tilting objects.

A *tilting sheaf* $\mathscr{T}$ on Z_Σ satisfies:

1. $\text{Ext}^i_{Z_\Sigma}(\mathscr{T}, \mathscr{T}) = 0$ for all $i < 0$.
2. $A = \text{Hom}_{Z_\Sigma}(\mathscr{T}, \mathscr{T})$ has finite global dimension.[3]
3. The direct summands of $\mathscr{T}$ generate $\mathbf{D}^b(Z_\Sigma)$.

In this case the functors

$$\begin{aligned} \mathbf{R}\,\text{Hom}_{Z_\Sigma}(\mathscr{T}, -) &: \mathbf{D}^b(Z_\Sigma) \to \mathbf{D}^b(\text{mod-}A) \\ - \overset{\mathbf{L}}{\otimes}_A \mathscr{T} &: \mathbf{D}^b(\text{mod-}A) \to \mathbf{D}^b(Z_\Sigma) \end{aligned} \tag{19}$$

are mutual inverses, where mod-A is the category of finitely generated right A-modules.

[3] In the context of this chapter, this condition is automatically satisfied.

In practice one can frequently, if not always, construct such a tilting sheaf as a sum of sheaves of the form $\tilde{S}(\boldsymbol{\alpha})$ as follows. For each $m_j^\vee$ in (18), we can consider the free resolution of the S-module $S/m_j^\vee$:

$$\dots \longrightarrow \bigoplus_i S(\mathbf{q}_{i2}) \longrightarrow \bigoplus_i S(\mathbf{q}_{i1}) \longrightarrow S \longrightarrow \frac{S}{m_j^\vee} \longrightarrow 0. \tag{20}$$

Sheafifying, this gives an exact sequence

$$\dots \longrightarrow \bigoplus_i \tilde{S}(\mathbf{q}_{i2}) \longrightarrow \bigoplus_i \tilde{S}(\mathbf{q}_{i1}) \longrightarrow \tilde{S} \longrightarrow 0, \tag{21}$$

relating $\tilde{S}$ to other objects in $\mathbf{D}^b(Z_\Sigma)$ composed of $\tilde{S}(\boldsymbol{\alpha})$'s for various multidegrees $\boldsymbol{\alpha}$. By considering such relations for all $m_j^\vee$'s, one can try to find a minimal generating set $\{\tilde{S}(\boldsymbol{\alpha}_1), \tilde{S}(\boldsymbol{\alpha}_2), \dots, \tilde{S}(\boldsymbol{\alpha}_k)\}$ from which all others may be built. This set thus generates $\mathbf{D}^b(Z_\Sigma)$. Then, if all higher Ext's between these sheaves vanish,

$$\mathscr{T} = \tilde{S}(\boldsymbol{\alpha}_1) \oplus \tilde{S}(\boldsymbol{\alpha}_2) \oplus \dots \oplus \tilde{S}(\boldsymbol{\alpha}_k) \tag{22}$$

is a tilting object. Obviously one may shift all the $\boldsymbol{\alpha}_i$'s by some common multidegree, and $\mathscr{T}$ will remain a tilting sheaf.

We will refer to $\{\tilde{S}(\boldsymbol{\alpha}_1), \tilde{S}(\boldsymbol{\alpha}_2), \dots, \tilde{S}(\boldsymbol{\alpha}_k)\}$ as a *tilting collection*. It is useful to impose one further condition:

Definition 6. A Σ-tilting collection of sheaves $\{\tilde{S}(\boldsymbol{\alpha}_1), \tilde{S}(\boldsymbol{\alpha}_2), \dots, \tilde{S}(\boldsymbol{\alpha}_k)\}$ is a collection of sheaves forming a tilting object (22) and such that (using (16))

$$\operatorname{Hom}_{\frac{\mathbf{D}^b(\mathrm{gr}-S)}{T_\Sigma}}(\tilde{S}(\boldsymbol{\alpha}_i), \tilde{S}(\boldsymbol{\alpha}_j)) \cong \operatorname{Hom}_{\mathbf{D}^b(\mathrm{gr}-S)}(S(\boldsymbol{\alpha}_i), S(\boldsymbol{\alpha}_j)), \tag{23}$$

for all i, j.

If $\mathscr{T}$ given by (22) corresponds to a Σ-tilting object then let $\mathbf{D}^b(\mathrm{gr}-S)_{\mathscr{T}}$ be the full triangulated subcategory of $\mathbf{D}^b(\mathrm{gr}-S)$ generated by $\{S(\boldsymbol{\alpha}_1), S(\boldsymbol{\alpha}_2), \dots, S(\boldsymbol{\alpha}_k)\}$. It immediately follows that

Proposition 4. *There is an equivalence of categories*

$$\frac{\mathbf{D}^b(\mathrm{gr}-S)}{T_\Sigma} \cong \mathbf{D}^b(\mathrm{gr}-S)_{\mathscr{T}}. \tag{24}$$

This is easy to see given that there is an equivalence between morphisms on both sides for objects in the tilting collection, and furthermore all objects may be built from the tilting objects.

The conditions on a Σ-tilting collection can be expressed conveniently in terms of local cohomology. Let $H_B^i(S)$ be the local cohomology groups of the monomial ideal B. These groups carry the same multigrading structure as S, and we denote the graded parts $H^i_B(S)_{\boldsymbol{\alpha}}$ accordingly. We then have the following copied from [24]:

Proposition 5. *For $i > 0$ there are isomorphisms*

$$H^i(Z_\Sigma, \tilde{S}(\boldsymbol{\alpha})) \cong H^{i+1}_B(S)_{\boldsymbol{\alpha}} \tag{25}$$

and an exact sequence

$$0 \longrightarrow S_{\boldsymbol{\alpha}} \longrightarrow H^0(Z_\Sigma, \tilde{S}(\boldsymbol{\alpha})) \longrightarrow H^1_B(S)_{\boldsymbol{\alpha}} \longrightarrow 0. \tag{26}$$

This gives

Proposition 6. *If $\{\tilde{S}(\boldsymbol{\alpha}_1), \tilde{S}(\boldsymbol{\alpha}_2), \ldots, \tilde{S}(\boldsymbol{\alpha}_k)\}$ forms a Σ-tilting collection then*

$$H^i_B(S)_{\boldsymbol{\alpha}_i - \boldsymbol{\alpha}_j} = 0 \tag{27}$$

for all $i \geq 0$ and all i, j.

In the case that $r = 1$ and we have a single grading, the process of finding a tilting sheaf is very straightforward [19, 21, 25]. Here, one obtains simply a range

$$\mathscr{T} = \tilde{S}(a) \oplus \tilde{S}(a+1) \oplus \ldots \oplus \tilde{S}(a+k-1) \tag{28}$$

for any choice of a.

Σ-tilting collections become very useful if they are simultaneously Σ_1-tilting and Σ_2-tilting for two phases Σ_1 and Σ_2. Then one can use it to explicitly map objects in the D-brane category between different phases. That is, we have an explicit equivalence between two phases given by Proposition 4:

$$\mathbf{D}^b(Z_{\Sigma_1}) \longrightarrow \mathbf{D}^b(\text{gr}{-}S)_{\mathscr{T}} \longrightarrow \mathbf{D}^b(Z_{\Sigma_2}). \tag{29}$$

In the case of $r = 1$, this idea has been explored in [19, 21, 25]. The combinatorics of finding simultaneous tilting objects for many phases when $r = 2$ was discussed in [26].

3.2 Complete Intersections

Because of Serre duality it is impossible to find a tilting collection on a compact Calabi–Yau variety. That said, we may still make use of the above technology by considering embedding the variety into a toric variety. Divide the homogeneous coordinates of S into two sets by relabeling $x_1, \ldots, x_n$ as $p_1, \ldots, p_s, z_1, \ldots, z_{n-s}$. Let $S' = \mathbb{C}[z_1, \ldots, z_{n-s}]$. Assume there is a $(\mathbb{C}^*)^r$-invariant polynomial, called a *superpotential*, $\mathcal{W} \in S$ that can be written

$$\mathcal{W} = p_1 f_1(z_1, z_2, \ldots) + p_2 f_2(z_1, z_2, \ldots) + \ldots + p_s f_s(z_1, z_2, \ldots), \tag{30}$$

where $f_1, f_2, \ldots$ forms a *regular sequence* in S'.

Let $X_\Sigma \subset Z_\Sigma$ be the critical point set of $\mathcal{W}$. If the functions f_j are sufficiently generic, the matrix of derivatives of these functions will have maximum rank. This maximal rank condition will imply that, for suitable Σ, X_Σ is a *smooth* complete intersection in Z_Σ. We are interested in the case where X_Σ is a compact Calabi–Yau threefold.

Now assume that the triangulation Σ is such that all the points corresponding to the p_j's are vertices of every simplex. That is, $B_\Sigma \cap \langle p_j \rangle = 0$ for $j = 1, \ldots, s$, which implies B_Σ may be viewed as an ideal of S'.

Define the ring

$$A = \frac{S'}{\langle f_1, f_2, \ldots, f_s \rangle}. \tag{31}$$

The category of coherent sheaves on X_Σ is then the Serre categorical quotient

$$\frac{\text{gr}-A}{M_\Sigma}, \tag{32}$$

where M_Σ is the abelian subcategory of all graded A-modules killed by some power of B_Σ. Thus, in analogy to (16), we have

$$\mathbf{D}^b(X_\Sigma) = \frac{\mathbf{D}^b(\text{gr}-A)}{T_\Sigma}. \tag{33}$$

Now let us introduce the notion of the *category of matrix factorizations*. S has an r-fold grading from the toric data. In addition, we add one further grading which we call the R-grading. This grading lives in $2\mathbb{Z}$, i.e., it is always an even number. For the superpotential (30) we may choose the variables p_j to have R-degree 2 and the z_k's to have degree 0.

Define the category $\text{DGr}S(\mathcal{W})$ of matrix factorizations of $\mathcal{W}$ as follows. An object is a pair

$$\bar{P} = \Big(P_1 \underset{u_0}{\overset{u_1}{\rightleftarrows}} P_0 \Big), \tag{34}$$

where P_0 and P_1 are two finite rank graded free S-modules. The two maps satisfy the matrix factorization condition

$$u_0 u_1 = u_1 u_0 = \mathcal{W}.\text{id}, \tag{35}$$

where both u_0 and u_1 have degree 0 with respect to the toric gradings (i.e., they preserve the toric grading). u_0 is a map of degree 2 with respect to the R-grading while u_1 has degree 0. Morphisms are defined in the obvious way up to homotopy. We refer to [23, 27] for more details. The category $\text{DGr}S(\mathcal{W})$ is a triangulated category with a shift functor

$$\bar{P}[1] = \Big(P_0 \underset{u_1}{\overset{u_0}{\rightleftarrows}} P_1\{2\} \Big), \tag{36}$$

where {} denotes a shift in the R-grading. Thus

$$\bar{P}[2] = \bar{P}\{2\}, \tag{37}$$

and the R-symmetry grading is identified with the homological grading (and extended from $2\mathbb{Z}$-valued to $\mathbb{Z}$-valued). That is, there is no difference between $[m]$ and $\{m\}$ in this category, although we will sometimes use both notations for clarity.

We now have

Proposition 7. *There is an equivalence of triangulated categories:*

$$DGrS(\mathcal{W}) \cong \mathbf{D}^b(\text{gr}-A). \tag{38}$$

This result follows very explicitly from the way that Macaulay 2 performs the computation of Ext groups in the category of graded A-modules as explained in detail in [28]. This algorithm was based on observations in [29, 30] as follows. An A-module typically has an infinite free resolution. In the case that $s = 1$ (i.e., a hypersurface), the resolution is 2-periodic. This resolution can then be reinterpreted as a system of maps between S'-modules, where the product of two maps forms a matrix factorization as above. To extend this to the case $s > 1$, and in order to correctly keep track of Ext data, one introduces new variables $p_1, \ldots, p_s$ extending the ring S' to S. We refer to [23] for more details.

It might help to view $\text{DGr}S(\mathcal{W})$ as the D-brane category on the stack $\mathscr{X}$, where $\mathscr{X}$ is the critical point set of $\mathcal{W}$ on the quotient $\mathbb{C}^n/(\mathbb{C}^*)^r$. Anyway, we have an equivalence of categories

$$\mathbf{D}^b(X_\Sigma) \cong \frac{\text{DGr}S(\mathcal{W})}{T_\Sigma}. \tag{39}$$

This gives us an analogy between D-branes on X_Σ and D-branes on Z_Σ. In the case of Z_Σ we began with D-branes on the quotient stack $\mathbb{C}^n/(\mathbb{C}^*)^r$ in the form $\mathbf{D}^b(\text{gr}-S)$ and then removed the pointset $V(B_\Sigma)$ by quotienting the category by T_Σ. Now in the case of X_Σ the original stack D-brane category is $\text{DGr}S(\mathcal{W})$, and we again remove the pointset $V(B_\Sigma)$ by quotienting the category by T_Σ.

The fact that the D-brane category is of this form has a direct physics derivation using the gauged linear σ-model in [19]. That is,

Physics Proposition 2. *The D-brane category for the B-model on the critical point set of the superpotential $\mathcal{W}$ is*

$$\frac{\text{DGr}S(\mathcal{W})}{T_\Sigma}. \tag{40}$$

Most importantly, combining this with Physics Proposition 1 gives us the following notion:

Physics Proposition 3. *The category*

$$\frac{\mathrm{DGr}S(\mathcal{W})}{T_\Sigma} \tag{41}$$

does not depend on the choice of triangulation Σ.

We refer to [21] and the recent work [31] for more rigorous statements using GIT language.

Note that the quotient in the above statement needs some interpretation. Let M be an S-module that is annihilated by $\mathcal{W}$. Then M is an $S/\langle\mathcal{W}\rangle$-module. A free resolution of M as an $S/\langle\mathcal{W}\rangle$-module results in a matrix factorization and thus maps M into $\mathrm{DGr}S(\mathcal{W})$. One should therefore consider T_Σ in (40) and (41) to be generated by S-modules annihilated both by a power of B_Σ *and* by $\mathcal{W}$. Note that in very simple examples, $\mathcal{W} \in B_\Sigma$ anyway, but this need not be the case generally.

As an example let us consider the following. Let the toric geometry be given by a homogeneous coordinate ring $S = \mathbb{C}[p, t, x_0, x_1, x_2, x_3, x_4]$ with degrees

	p	t	x_0	x_1	x_2	x_3	x_4
Q_1	-4	1	1	1	1	0	0
Q_2	0	-2	0	0	0	1	1
R	2	0	0	0	0	0	0

(42)

and superpotential

$$\mathcal{W} = p(x_0^4 + x_1^4 + x_2^4 + t^4x_3^8 + t^4x_4^8). \tag{43}$$

There is a triangulation of the pointset that has p as a vertex of every simplex and t is ignored completely. This gives a monomial ideal $I_\Sigma = \langle x_0x_1x_2x_3x_4, t\rangle$ and

$$B_\Sigma = \langle x_0, x_1, x_2, x_3, x_4\rangle\langle t\rangle. \tag{44}$$

Z_Σ then corresponds to the total space of the canonical line bundle over the weighted projective space $\mathbb{P}^4_{\{22211\}}$. One calls this the "orbifold" phase.

Now $S/\langle t\rangle$ is annihilated by B_Σ, but it is not annihilated by $\mathcal{W}$. Thus $S/\langle t\rangle$ is not naturally associated with any object in $\mathrm{DGr}S(\mathcal{W})$. However, consider $S/\langle p, t\rangle$, which is annihilated by B_Σ and $\mathcal{W}$. We may construct the corresponding matrix factorization as follows. First note that $S/\langle p\rangle$ is annihilated by $\mathcal{W}$ and so gives a matrix factorization m_p. But now it is easy to show that $S/\langle p, t\rangle$ is given by the mapping cone of

$$\mathsf{m}_p(-1, 2) \xrightarrow{\ t\ } \mathsf{m}_p, \tag{45}$$

which can be shown to be

$$\begin{array}{c} S(4,0)[-2] \\ \oplus \\ S(-1,2) \end{array} \; \underset{\begin{pmatrix} f & -t \\ 0 & p \end{pmatrix}}{\overset{\begin{pmatrix} p & t \\ 0 & f \end{pmatrix}}{\rightleftarrows}} \; \begin{array}{c} S \\ \oplus \\ S(3,2) \end{array}, \tag{46}$$

for $f = x_0^4 + x_1^4 + x_2^4 + t^4 x_3^8 + t^4 x_4^8$. Indeed, *any* mapping cone corresponding to multiplication by t will give something in T_Σ.

It would be nice to assert that we can perform this quotient by using a tilting collection as in Sect. 3.1. Let $\mathscr{T}$ be a tilting object for a triangulation Σ. Then the quotient is performed by considering matrix factorizations involving only summands of our tilting object. Indeed, if $\mathrm{DGr}S(\mathcal{W})_{\mathscr{T}}$ is the category of matrix factorizations involving only summands of $\mathscr{T}$ then we expect an equivalence

$$\mathbf{D}^b(X_\Sigma) \longrightarrow \mathrm{DGr}S(\mathcal{W})_{\mathscr{T}}, \tag{47}$$

in analogy with (29). This is expected from linear σ-model arguments in [19] and proven explicitly for $r = 1$ in [21]. In this chapter we will only use this assertion in the $r = 1$ case.

3.3 Landau–Ginzburg Theories

A particularly simple phase in which to work is the so-called Landau–Ginzburg theory. This is where the effective target space is a fat point. The simplest way to obtain such a theory is when the convex hull of the pointset $\mathscr{A}$ is a simplex. In this case there is a trivial triangulation—a single simplex consisting of the convex hull itself. In some sense this is the "opposite" phase of a smooth Calabi–Yau manifold, which corresponds to a maximal triangulation. In terms of the Kähler form, the smooth Calabi–Yau is a "large radius limit" while the Landau–Ginzburg theory is a "small radius limit."

Suppose a point corresponding to the homogeneous coordinate x_j is *not* included in the triangulation Σ. It follows that x_j is an element of the Stanley–Reisner ideal I_Σ and $B_\Sigma \subset \langle x_j \rangle$. This means that the mapping cone of a map between any two objects in $\mathrm{DGr}S(\mathcal{W})$, given by multiplication by x_j, lies in T_Σ. Performing the triangulated quotient by T_Σ means that x_j becomes a unit. We will thus simply impose $x_j = 1$. This will reduce the effective multigrading to that under which x_j is neutral.

Thus, for the Landau–Ginzburg theory, we set all the homogeneous coordinates equal to one, except for the ones corresponding to vertices of the simplex. The $(\mathbb{C}^*)^r$-action of the original toric action is reduced to a finite group G as a result of setting all these coordinates equal to one. This means that we are really considering a Landau–Ginzburg *orbifold*. The D-brane category on this Landau–Ginzburg theory is therefore of G-equivariant matrix factorizations of $\mathcal{W}$ [32, 33].

To illustrate this, we choose the canonical example of the quintic threefold $X \in \mathbb{P}^4$, where $S = \mathbb{C}[p, z_0, z_1, z_2, z_3, z_4]$ with a single grading of degrees $(-5, 1, 1, 1, 1, 1)$, and

$$\mathcal{W} = p(z_0^5 + z_1^5 + z_2^5 + z_3^5 + z_4^5) = pf(z_i). \tag{48}$$

This has two phases, the Calabi–Yau phase Σ_1 with $B_{\Sigma_1} = \langle z_0, z_1, z_2, z_3, z_4 \rangle$ and the Landau–Ginzburg phase Σ_0 with $B_{\Sigma_0} = \langle p \rangle$.

There are two matrix factorizations of particular note. First consider the S-module

$$\mathsf{w} = \frac{S}{\langle z_0, z_1, z_2, z_3, z_4 \rangle}. \tag{49}$$

Since w is annihilated by $\mathcal{W}$, it may also be viewed as an R-module, where $R = S/\langle \mathcal{W} \rangle$. We now compute a minimal free R-module resolution of w:

$$\longrightarrow \begin{matrix} R(-5) \\ \oplus \\ R(-3)\{-2\}^{\oplus 10} \\ \oplus \\ R(-1)\{-4\}^{\oplus 5} \end{matrix} \longrightarrow \begin{matrix} R(-4)^{\oplus 5} \\ \oplus \\ R(-2)\{-2\}^{\oplus 10} \\ \oplus \\ R\{-4\} \end{matrix} \longrightarrow \begin{matrix} R(-3)^{\oplus 10} \\ \oplus \\ R(-1)\{-2\}^{\oplus 5} \end{matrix} \longrightarrow$$

$$\begin{matrix} R(-2)^{\oplus 10} \\ \oplus \\ R\{-2\} \end{matrix} \xrightarrow{\begin{pmatrix} -x_1 & 0 & & px_0^4 \\ x_0 & -x_2 & & px_1^4 \\ 0 & x_1 & \cdots & px_2^4 \\ 0 & 0 & & px_3^4 \\ 0 & 0 & & px_4^4 \end{pmatrix}} R(-1)^{\oplus 5} \xrightarrow{(\, x_0 \; x_1 \; \ldots \; x_4 \,)} R \longrightarrow \mathsf{w}. \tag{50}$$

To write this infinite resolution as a matrix factorization, we replace R with S and "roll it up" following [23, 28]. w then corresponds to a 16×16 matrix factorization:

$$\begin{matrix} S(-1)^{\oplus 5} \\ \oplus \\ S(-3)[2]^{\oplus 10} \\ \oplus \\ S(-5)[4] \end{matrix} \;\overset{\longrightarrow}{\longleftarrow}\; \begin{matrix} S \\ \oplus \\ S(-2)[2]^{\oplus 10} \\ \oplus \\ S(-4)[4]^{\oplus 5}. \end{matrix} \tag{51}$$

Hoping context makes usage clear, we will use w to denote this matrix factorization.

The other object of note is given by $\mathsf{s} = S/\langle f \rangle$. This corresponds to the obvious matrix factorization (also denoted by s):

$$S(-5) \underset{p}{\overset{f}{\rightleftarrows}} S. \tag{52}$$

Note that $S/\langle p \rangle$ then corresponds to a similar matrix factorization given by $\mathsf{s}(5)[-1]$.

In the Calabi–Yau phase T_{Σ_1} is generated by w and in the Landau–Ginzburg phase T_{Σ_0} is generated by s. A tilting object can be chosen as

$$S(-4) \oplus S(-3) \oplus S(-2) \oplus S(-1) \oplus S. \tag{53}$$

We refer to this range, $-4 \leq m \leq 0$, as the tilting "window" following [19]. We may take any matrix factorization and convert it to a matrix factorization involving only summands $S(-4), \ldots, S$ of this tilting object. In the Landau–Ginzburg phase we iteratively take mapping cones with shifts of s to eliminate $S(m)$ for $m \leq -5$ or $m > 0$. As stated above, this is equivalent to setting $p = 1$. This gives an identification

$$S(m)[2] \cong S(m+5) \tag{54}$$

for any $m \in \mathbb{Z}$. Our double grading thus collapses to a single grading. To this effect, we denote $S(Q)[R]$ by $S\langle 5R + 2Q \rangle$. Thus, the D-brane category is simply the category of matrix factorizations of f with this new grading. This is a feature of all Landau–Ginzburg phases—we have a category of matrix factorizations with some single grading.

In the Calabi–Yau phase we iteratively take mapping cones with shifts of w to eliminate $S(m)$ for $m \leq -5$ or $m > 0$. In terms of matrix factorizations, this phase is not as simple as the Landau–Ginzburg phase.

As an example consider the structure sheaf Ø$_X$ as an object in $\mathbf{D}^b(X)$. This corresponds to the A-module A itself. The equivalence in Proposition 7 tells us that we can find an associated matrix factorization by viewing it as the $S/\langle \mathcal{W} \rangle$-module given by $\operatorname{coker}(f)$. This is none other than s given in (52). We may get this into the tilting window by considering the following triangle defining u

$$\begin{array}{ccc} \mathsf{w}[-4] & \xrightarrow{\;f\;} & \mathsf{s} \\ & {\scriptstyle [1]}\nwarrow \quad \swarrow & \\ & \mathsf{u} & \end{array} \tag{55}$$

where f contains the identity map $S(-5) \to S(-5)$ to cancel these summands in u. Thus u involves only summands from the tilting collection. Obviously u also represents $\mathcal{O}_X$ since it differs from s by $\mathsf{w}[-4]$, which is trivial in the Calabi–Yau phase.

So u represents the image of the D-brane $\mathcal{O}_X$ in the Landau–Ginzburg phase. But in the Landau–Ginzburg phase the matrix factorization s is trivial, so we may use the triangle (55) once again to identify u with $\mathsf{w}[-3]$. That is, the structure sheaf $\mathcal{O}_X$ corresponds in the Landau–Ginzburg phase to the matrix factorization

$$\begin{array}{ccc} S\langle -2\rangle^{\oplus 5} & & S \\ \oplus & & \oplus \\ S\langle 4\rangle^{\oplus 10} & \underset{\longleftarrow}{\longrightarrow} & S\langle 6\rangle^{\oplus 10} \\ \oplus & & \oplus \\ S\langle 10\rangle & & S\langle 12\rangle^{\oplus 5}, \end{array} \tag{56}$$

with maps coming from (50) with $p = 1$.

As another example, let us consider a particular line on the quintic (48). Let this rational curve be defined by the ideal

$$I = \langle x_0 + x_1, x_2 + x_3, x_4\rangle. \tag{57}$$

Matrix factorizations associated to this line were first studied in [34]. The sheaf supported on this curve is associated to the A-module $M = A/I$. This gives a matrix factorization

$$\begin{array}{ccc} S(-1)^{\oplus 3} & & S \\ \oplus & \underset{\longleftarrow}{\longrightarrow} & \oplus \\ S(-3)[2] & & S(-2)[2]^{\oplus 3} \end{array} \tag{58}$$

with Macaulay 2 yielding maps

$$\begin{bmatrix} x_0+x_1 & x_2+x_3 & x_4 & 0 \\ p(x_2^3x_3-px_2^4-x_2^2x_3^2+x_2x_3^3-x_3^4) & p(x_0^4-x_0^3x_1+x_0^2x_1^2-x_0x_1^3+x_1^4) & 0 & x_4 \\ -px_4^4 & 0 & p(x_0^4-x_0^3x_1+x_0^2x_1^2-x_0x_1^3+x_1^4) & -x_2-x_3 \\ 0 & -px_4^4 & p(x_2^4-x_2^3x_3+x_2^2x_3^2-x_2x_3^3+x_3^4) & x_0+x_1 \end{bmatrix} \tag{59}$$

and

$$\begin{bmatrix} p(x_0^4-x_0^3x_1+x_0^2x_1^2-x_0x_1^3+x_1^4) & -x_2-x_3 & -x_4 & 0 \\ p(x_2^4-x_2^3x_3+x_2^2x_3^2-x_2x_3^3+x_3^4) & x_0+x_1 & 0 & -x_4 \\ px_4^4 & 0 & x_0+x_1 & x_2+x_3 \\ 0 & px_4^4 & p(x_2^3x_3-px_2^4-x_2^2x_3^2+x_2x_3^3-x_3^4) & p(x_0^4-x_0^3x_1+x_0^2x_1^2-x_0x_1^3+x_1^4) \end{bmatrix} \tag{60}$$

Note that (58) contains only summands in the tilting collection, and so no further action is required to get it in the right form appropriate for the Landau–Ginzburg phase. It was shown in [35] that this is always the case (for a suitable choice of tilting collection) for projectively normal rational curves.

Now let us compute some dimensions of Ext groups between this twisted cubic and its degree shifts by computing in the Landau–Ginzburg phase. This is conveniently done using Macaulay 2.

```
i1 : kk = ZZ/31469
i2 : B = kk[x_0..x_4]
i3 : W = x_0^5+x_1^5+x_2^5+x_3^5+x_4^5
i4 : A = B/(W)
i5 : M = coker matrix {{x_0+x_1,x_2+x_3,x_4}}
```

Now we use the internal Macaulay routine described in [28] to compute the S-module $\mathrm{Ext}^*_A(M,M)$:[4]

```
i6 : ext = Ext(M,M)

o7 = cokernel {0, 0}   | x_4 x_2+x_3 x_0+x_1 0   0       0       X_1x_3^4
              {-1, -1} | 0   0       0       x_4 x_2+x_3 x_0+x_1 0          ...
              {-1, 3}  | 0   0       0       0   0       0       0
              {-2, 2}  | 0   0       0       0   0       0       0


o7 : kk[X , x , x , x , x , x ]-module, quotient of (kk[X , x , x , x , x ,
                                                              4
                                                          x ])
         1   0   1   2   3   4                             1   0   1   2   3   4
```

We have suppressed part of the output. The first column of the output represents the bi-degrees of the generators of this module. The first degree is the homological degree discussed in the previous section, and the second degree is the original degree associated to our graded ring B.

Next we need to pass to the quotient category $\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)$ by setting $X_1 = -1$. This collapses to a single grading as described above. The following code sets `pr` equal to the map whose cokernel defines $\mathrm{Ext}^*_A(M,M)$ above, and we define our rings S and a singly graded `B2`.

```
i7 : pr = presentation ext
i8 : S = ring target pr
i9 : B2 = kk[x_0..x_4,Degrees=>{5:2}]
```

[4]The Macaulay 2 variable X_1 is our $-p$. Note that Macaulay 2 views complexes in terms of homology rather than cohomology and so X_1 has R charge -2.

```
i10: toB = map(B2, S, {-1,x_0,x_1,x_2,x_3,x_4},
                      DegreeMap => ( i -> {-5*i#0+2*i#1}))
```

The last line above is the heart of our algorithm. It defines a ring map which sets $X_1 = -1$ and defines how we map degrees. It is now simple to compute the Ext's in the quotient category $\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)$ by constructing the tensor product:

```
i11 : extq = prune coker(toB ** pr)

o11 = cokernel {0} | x_4 x_2+x_3 x_0+x_1 0   0       0
  x_3^4 x_1^4 0      0      |
                {3} | 0   0       0       x_4 x_2+x_3 x_0+x_1 0
                  0      x_3^4 x_1^4 |
                                    2
o11 : B2-module, quotient of B2
```

Finally we may compute the dimensions of spaces of morphisms in the D-brane category by computing the dimensions of the above module at specific degrees. This, of course, is the Hilbert function:

```
i12 : apply(20, i -> hilbertFunction(i, extq))

o12 = {1, 0, 2, 1, 3, 2, 4, 3, 3, 4, 2, 3, 1, 2, 0, 1, 0, 0, 0,
       0}

o12 : List
```

The above list represents the dimensions of $\mathrm{Hom}_{\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)}(M, M\langle i\rangle)$ for

$$\mathrm{Ext}^k(M, M(r)) = \mathrm{Hom}_{\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)}(M, M\langle 5k + 2r\rangle). \tag{61}$$

Note that Serre duality implies $\mathrm{Hom}_{\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)}(M, M\langle i\rangle) \cong \mathrm{Hom}_{\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)}(M, M\langle 15 - i\rangle)$ which is consistent with the above output. For open strings beginning and ending on the same untwisted 2-brane M, we immediately see

$$\begin{aligned} \mathrm{Ext}^0(M, M) &= \mathbb{C}, & \mathrm{Ext}^1(M, M) &= \mathbb{C}^2 \\ \mathrm{Ext}^2(M, M) &= \mathbb{C}^2, & \mathrm{Ext}^3(M, M) &= \mathbb{C}. \end{aligned} \tag{62}$$

This shows that our twisted cubic curve has normal bundle $\mathcal{O}(-3) \oplus \mathcal{O}(1)$.

We could also compute open string Hilbert spaces between M and twists of M. For example, $\mathrm{Ext}^1(M, M(1)) = \mathbb{C}^3$.

3.4 Hochschild Cohomology

The above computation in Macaulay also reveals the closed to open string map $\eta_\phi(\mathsf{a})$ from (5) in the case that a is the D-brane given by the line on the quintic in question. The closed string Hilbert space of Landau–Ginzburg theories has long been understood [36]. For the quintic it is given by the quotient ring

$$H_c = \frac{\mathbb{C}[z_0, z_1, z_2, z_3, z_4]}{\left\langle \frac{\partial f}{\partial z_0}, \ldots, \frac{\partial f}{\partial z_4} \right\rangle}. \tag{63}$$

This was rederived in terms of Hochschild cohomology in [37, 38]. In our case of the quintic, we are looking at a $\mathbb{Z}_5$-orbifold of the Landau–Ginzburg theory. This means that we should restrict attention to the subring of H_c invariant under $\mathbb{Z}_5$. In addition, when orbifolding, one needs to worry about additional contributions to the Hilbert space coming from "twisted sectors." This does not occur in the B-model for the quintic.

The Hilbert space of closed strings in the quintic is therefore generated by monomials in $\mathbb{C}[z_0, z_1, z_2, z_3, z_4]$ of degree $5d$, for some integer d, where each variable appears with degree ≤ 3. As above, with the collapsing to a single degree, this monomial corresponds to an element of Hochschild cohomology of degree $2d$.

The closed to open string map $\eta_\phi(\mathsf{a})$ therefore maps a $\mathbb{Z}_5$-invariant polynomial into a pair of matrices representing an endomorphism of the matrix factorization corresponding to a. These endomorphisms must commute with all other matrices representing elements of the open string Hilbert space according to (6).

But output `o11` above tells us exactly the structure of $\mathrm{Hom}(\mathsf{a}, \mathsf{a}[i])$ for any i. The two rows tell us we have two generators. The first one is obviously the identity element in $\mathrm{Hom}(\mathsf{a}, \mathsf{a})$, and thus we denote it by $\mathbf{1}$.

If ϕ is a degree $5d$ monomial then $\eta_\phi(\mathsf{a})$ is an element of $\mathrm{Hom}_{\mathbf{D}^{\mathrm{gr}}_{\mathrm{Sg}}(A)}(M, M\langle 10d \rangle))$. If $d = 0$, we have the trivial case corresponding to the identity. If $d = 1$, we have two elements $x_1^2 x_3^3 \mathbf{1}$ and $x_1^3 x_3^2 \mathbf{1}$. Clearly then

$$\begin{aligned} \eta_1(\mathsf{a}) &= \mathbf{1} \\ \eta_{x_1^2 x_3^3}(\mathsf{a}) &= x_1^2 x_3^3 \mathbf{1} \\ \eta_{x_1^3 x_3^2}(\mathsf{a}) &= x_1^3 x_3^2 \mathbf{1}. \end{aligned} \tag{64}$$

For $d > 1$ the map is zero by degree considerations. The content of `o11` thus fully determines $\eta_\phi(\mathsf{a})$. Methods of commutative algebra can thus be utilized to compute the closed to open string map for the full untwisted sector for any model with a Landau–Ginzburg orbifold phase.

4 Monodromy

The B-model data is dependent on the complex structure of X. Mirror to this statement is the fact that the A-model data depends on a complexified Kähler form $B + iJ \in H^2(Y, \mathbb{C})$. This statement suggests an interesting question about monodromy. Suppose we consider a Lagrangian 3-cycle on Y. This defines a class in $H_3(Y, \mathbb{Z})$. Now go around a loop in the moduli space of complex structures of Y. This can have a nontrivial monodromy action on $H_3(Y, \mathbb{Z})$, and thus it must act on the set of A-brane objects. We can think of $H_3(Y, \mathbb{Z})$ as the Grothendieck group for the A-brane category. The corresponding Grothendieck group $K(X)$ for B-branes on a Calabi–Yau threefold[5] is

$$K(X) = H^0(X, \mathbb{Z}) \oplus H^2(X, \mathbb{Z}) \oplus H^4(X, \mathbb{Z}) \oplus H^6(X, \mathbb{Z}). \tag{65}$$

For mirror symmetry to work, there must be some action of monodromy in the moduli space of complexified Kähler forms on the derived category and thus $K(X)$.

Part of this monodromy can be explained nicely in terms of the B-field. We state this very briefly here and refer to [8] for a more complete description. The B-field describes a 2-form flat gerbe connection on X for worldsheets with no boundary. When boundaries are added, the gerbe connection restricts to a line bundle connection on the boundary in some sense. This is exactly part of the bundle data on D-branes. This intimate relationship between the B-field and the D-brane bundle implies that a transformation $B \to B + e$, for some $e \in H^2(X, \mathbb{Z})$, must be equivalent to a change in the bundle curvature $F \to F + e$.

Thus, monodromy associated to $B + iJ \to (B + e) + iJ$ maps any sheaf $\mathscr{F}$ to the twisted sheaf $\mathscr{F}(D_e)$ where D_e is the divisor class dual to e. Such a twist by D_e is an obvious automorphism of the derived category of X.

This monodromy preserves the decomposition (65). What about the other monodromies? One may add further data, called a “stability condition,” to the topological field theory. This is easier to explain in terms of the A-model. An object in the Fukaya category is a Lagrangian 3-manifold in Y (with a line bundle with a flat connection over it). This object is “stable” if it can be represented by a *special* Lagrangian (see, e.g., [39]). As one moves in the moduli space of complex structures, some objects in the Fukaya category will become unstable, and some previously unstable objects may become stable. The monodromy action on the Fukaya category will map the original stable set of objects to the new stable set upon going around a loop in the moduli space. Accordingly, there must be some similar story for the derived category where stability depends on the complexified Kähler form [6, 40–42].

The monodromy $B \to B + e$ can be viewed as monodromy around the large radius limit. Consider the toric Calabi–Yau Z_Σ with its r-fold multigraded homogeneous coordinate ring S. Let $(\mathbf{e}_i)$ denote a shift by one in the ith grading. Now consider:

Definition 7. $m_\Sigma(\mathbf{e}_i)$ is the automorphism of the derived category $\mathbf{D}^b(Z_\Sigma)$ induced by the action $S(\mathbf{q}) \to S(\mathbf{q} + \mathbf{e}_i)$.

[5] Let us assume the Calabi–Yau threefold is simply connected and the cohomology is torsion-free.

That it is an automorphism follows from proposition 1, since it is obviously an automorphism of $\mathbf{D}^b(\text{gr}-S)$, and it is an automorphism of T_Σ by the definition of T_Σ. In the Calabi–Yau phase it is easy to show that this monodromy indeed corresponds to $B \to B + e$. It is argued from a physics point of view in [19] that this is the correct automorphism to associate to *any* phase.

4.1 K-Theory

We will restrict analysis of monodromy to the K-theory associated to D-branes. Let the K-theory class of $S(\mathbf{q})$ be represented by

$$K(S(\mathbf{q})) = s_1^{q_1} s_2^{q_2} \dots s_r^{q^r}. \tag{66}$$

Then the K-theory of $\mathbf{D}^b(\text{gr}-S)$ is obviously the additive group of Laurent polynomials $\mathbb{Z}[s_1, s_1^{-1}, \dots, s_r, s_r^{-1}]$. The action of the automorphism $S(\mathbf{q}) \to S(\mathbf{q} + \mathbf{e}_i)$ clearly corresponds to multiplication by s_i in this Laurent polynomial ring.

The K-theory of T_Σ is invariant under such automorphisms and so must correspond to an ideal. We can compute it as follows:

Proposition 8. *Let* $B_\Sigma = m_1^\vee \cap m_2^\vee \cap \dots$ *and* $m_j^\vee = \langle x_\alpha, x_\beta, x_\gamma, \dots \rangle$ *as in (18). Define*

$$f_j = (1 - t^{q_\alpha})(1 - t^{q_\beta})(1 - t^{q_\gamma}) \dots, \tag{67}$$

where q_α *is the multidegree of* x_α*, etc. Then*

$$K(T_\Sigma) = \langle f_1, f_2, \dots \rangle. \tag{68}$$

This result follows from Proposition 3. The f_j's are the K-theory classes of $m_j^\vee$ given by Koszul resolutions of $m_j^\vee$.

It follows that the K-theory for D-branes in a phase given by Σ is

$$K(Z_\Sigma) = \frac{\mathbb{Z}[s_1, s_1^{-1}, \dots, s_r, s_r^{-1}]}{K(T_\Sigma)}. \tag{69}$$

We will call this the "monodromy ring" for phase Σ. In a smooth Calabi–Yau phase it corresponds to the toric Chow ring.

Let us consider the K-theory and associated monodromy ring in the compact case. Let M be an $S/\langle \mathcal{W} \rangle$-module. It typically has an infinite resolution in terms of free $S/\langle \mathcal{W} \rangle$-modules. Thus we associate to any object a *power series* which expresses the associated element of K-theory. It is convenient to include an extra variable to express the R-grading.

Let P denote the ring of formal power series

$$\mathbb{Z}[[s_1, s_1^{-1}, s_2, s_2^{-1}, \dots, s_r, s_r^{-1}, \sigma, \sigma^{-1}]] \tag{70}$$

and define a map on free $S/\langle\mathcal{W}\rangle$-modules

$$k(S/\langle\mathcal{W}\rangle(\mathbf{v})\{m\}) = s_1^{v_1} s_2^{v_3} \dots s_r^{v_r} \sigma^{-m}. \tag{71}$$

By writing free $S/\langle\mathcal{W}\rangle$-module resolutions this extends to a map

$$k : \mathbf{D}^b(\text{gr-}S/\langle\mathcal{W}\rangle) \to P. \tag{72}$$

Ultimately the resolution of any object is periodic with period 2, and the product of two consecutive maps, lifted to an S-module map, is of homological degree (and hence R-degree since the two are identified) 2. It follows that for any object a, we have

$$\begin{aligned} k(\mathsf{a}) &= f(s_1, s_2, \dots, s_r, \sigma)(1 + \sigma^2 + \sigma^4 + \dots) + g(s_1, s_2, \dots, s_r, \sigma) \\ &= \frac{f(s_1, s_2, \dots, s_r, \sigma)}{1 - \sigma^2} + g(s_1, s_2, \dots, s_r, \sigma), \end{aligned} \tag{73}$$

where f and g are finite polynomials. Clearly, it is the polynomial f that expresses the K-theory information for $\mathrm{DGr}S(\mathcal{W})$. Unfortunately this approach to K-theory is not easy since the quotient by the K-theory of T_Σ in (40) is awkward. Instead, we will translate the monodromy statements about X_Σ back to the noncompact case Z_Σ as we describe in an example in Sect. 4.3.

4.2 The GKZ System

The monodromy of integral 3-cycle D-branes of the A-model around loops in the moduli space of complex structures is encoded in the Picard–Fuchs differential equations. These in turn can be described torically in terms of the GKZ system [43–45]. Let us simplify the discussion a little by assuming that X_Σ is a hypersurface in the toric variety. Also we will shift our indexing so that $i = 0, \dots, n-1$; $j = 0, \dots, d-1$. The index $i = 0$ will correspond to the unique point in $\mathscr{A}$ properly in the interior of the convex hull of $\mathscr{A}$, and $j = 0$ corresponds to a row of 1's in the matrix A imposing the hyperplane condition. We write the partial differential equations in terms of variables $a_0, \dots, a_{n-1}$. Define the operators

$$\begin{aligned} Z_j &= \sum_{i=0}^{n-1} \alpha_{ji} a_i \frac{\partial}{\partial a_i} - \beta_j \\ \Box_{\mathbf{u}} &= \prod_{u_j>0} \left(\frac{\partial}{\partial a_j}\right)^{u_j} - \prod_{u_j<0} \left(\frac{\partial}{\partial a_j}\right)^{-u_j}, \end{aligned} \tag{74}$$

where α_{ji} are the entries in the matrix A and $\mathbf{u}$ is any vector with coordinates $(u_0, u_1, \dots, u_{n-1})$ in the row space of the matrix Φ. We also have

$$\beta_i = \begin{cases} (-1, 0, 0, \ldots) & \text{if } i = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{75}$$

A period, ϖ, is then a solution of

$$Z_j \varpi = \Box_{\mathbf{u}} \varpi = 0 \tag{76}$$

for all j and $\mathbf{u}$. The equations $Z_j \varpi = 0$ can be used to replace the n variables a_i with r variables ξ_i. This is nothing other than going from the homogeneous coordinate ring $\mathbb{C}[a_1, \ldots, a_n]$ to a set of affine coordinates $(\xi_1, \ldots, \xi_r)$ on the toric variety associated to the *secondary fan*. A choice of cone Σ in the secondary fan corresponds to a phase and a choice of coordinates $(\xi_1, \ldots, \xi_r)$ [46].

The remaining differential equations $\Box_{\mathbf{u}} \varpi = 0$ written in these affine coordinates exhibit monodromy around the origin. This, according to mirror symmetry, is exactly the same as the monodromy m_Σ of Definition 7.

That this story works in the large radius Calabi–Yau phase has been understood for some time. See, for example, [47]. An important point, which complicates the analysis, is that the GKZ system does not give the full information on monodromy. We will see that we need a systematic way of discarding "extra" solutions. The K-theory picture of the monodromy ring above gives an interesting way of doing this, which we now explore.

4.3 A Calabi–Yau and Landau–Ginzburg Example

Rather than exploring the full generalities of monodromy, it is easier to demonstrate the general idea with an example. The quintic threefold and its associated Landau–Ginzburg phase have been studied extensively, and the monodromy is well understood (see, e.g., [8]). Here we discuss a slightly more complicated example where the concept of the monodromy ring is more powerful.

Let the matrices A and Φ be given by (written as $A^T|\Phi$):

						Q_1	Q_2	Q_3	
$p = x_0$	0	0	0	0	1	−5	0	0	0
x_1	1	0	0	0	1	1	1	0	0
x_2	0	1	0	0	1	1	0	0	1
x_3	0	0	1	0	1	1	0	0	1
x_4	0	0	0	1	1	1	0	0	1
x_5	−5	−3	−3	−3	1	0	0	1	2
x_6	−3	−2	−2	−2	1	0	1	−2	0
x_7	−1	−1	−1	−1	1	1	−2	1	0
x_8	−2	−1	−1	−1	1	0	0	0	−5

(77)

We have labeled each row by the associated homogeneous coordinate. A particular simplicial decomposition of the corresponding pointset yields the canonical line bundle of the weighted projective space $\mathbb{P}^4_{\{5,3,3,3,1\}}$.

The point in $\mathscr{A}$ corresponding to x_8 is in the interior of a codimension one face of the convex hull. This means that it should be completely irrelevant for our purposes, and we will ignore it [46, 48]. As such we also ignore the last column of Φ. The first three columns are associated to charges giving the multigrading (Q_1, Q_2, Q_3). In this example we have $r = 3$, and thus the dimension of the moduli space of complexified Kähler forms is 3.

To obtain a smooth hypersurface we may set

$$\mathcal{W} = p(x_1^3 x_6 x_7^2 + x_2^5 + x_3^5 + x_4^5 + x_5^{15} x_6^{10} x_7^5). \tag{78}$$

There are 12 possible triangulations of $\mathscr{A}$ (always ignoring x_8), but we will concentrate on only two of them. One of the triangulations corresponds to the Calabi–Yau phase. The corresponding Stanley–Reisner ideal is

$$I_{\mathrm{CY}} = \langle x_1x_5, x_1x_6, x_5x_7, x_2x_3x_4x_6, x_2x_3x_4x_7 \rangle. \tag{79}$$

The corresponding ideal we quotient by in (69) is

$$\begin{aligned} K(T_{\mathrm{CY}}) = \langle &(1-s_1s_2)(1-s_3), (1-s_1s_2)(1-s_2s_3^{-2}), (1-s_3)(1-s_1s_2^{-2}s_3), \\ &(1-s_1)^3(1-s_2s_3^{-2}), (1-s_1)^3(1-s_1s_2^{-2}s_3) \rangle. \end{aligned} \tag{80}$$

The other is the Landau–Ginzburg phase with

$$\begin{aligned} I_{\mathrm{LG}} &= \langle p, x_6, x_7 \rangle \\ K(T_{\mathrm{LG}}) &= \langle 1-s_1^{-5}, 1-s_2s_3^{-2}, 1-s_1s_2^{-2}s_3 \rangle. \end{aligned} \tag{81}$$

In both of these phases we may choose a tilting collection consisting of 15 objects given by the union of $\{S(-n,0,0), S(-n,0,-1), S(-n,-1,0)\}$ for $n = 0, \ldots, 4$. This implies that the collection of monomials given by the union of $\{s_1^{-n}, s_1^{-n}s_3^{-1}, s_1^{-n}s_2^{-1}\}$ for $n = 0, \ldots, 4$ spans $K(Z_\Sigma)$ as a vector space. Let us denote this collection of monomials $\mathscr{T}$.

First consider the noncompact geometries Z_Σ. Given any object in $\mathbf{D}^b(Z_\Sigma)$ we can therefore find its class in K-theory by:

1. Find a free resolution of the object and thus express its K-theory class as a Laurent polynomial in $\mathbb{Z}[s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}]$.
2. Reduce this polynomial mod $K(T_\Sigma)$ so that all its terms lie in $\mathscr{T}$.

This would be most easily done if there were some monomial ordering so that this second step was reduction to the *normal form*. Sadly, there is *no* monomial ordering that can do this.[6]

Moving to the compact examples X_Σ we have a subtlety in that we do not know precisely how to construct $K(T_\Sigma)$. To evade this issue we work in the monodromy ring of Z_Σ. Let

$$i : X_{\mathrm{CY}} \to Z_{\mathrm{CY}} \tag{82}$$

be the inclusion map for a Calabi–Yau phase. The key observation is that the monodromy action given in Definition 7 is the same on X_Σ as on Z_Σ. Thus, if we begin in the image of i_*, then we should expect to remain in the image of i_*.

4.4 Monodromy of the Structure Sheaf

We begin in the Calabi–Yau phase, where $K(T_{\mathrm{CY}})$ is given by (80). Consider the structure sheaf $\mathcal{O}_X$ (i.e., $i_*\mathcal{O}_X$). This has a free resolution

$$0 \longrightarrow \tilde{S}(-5,0,0) \xrightarrow{f} \tilde{S} \longrightarrow \mathcal{O}_X \longrightarrow 0. \tag{83}$$

This implies that $\mathcal{O}_X$ has a K-theory class $1 - s_1^{-5}$ as an object of the K-theory of Z_{CY}.

The quotient ring $\mathbb{Q}[s_1, s_1^{-1}, \dots]/K(T_{\mathrm{CY}})$ is a vector space of dimension 15, corresponding to the fact that the rank of the K-theory for Z_Σ is 15. But $h^{1,1}(X_{\mathrm{CY}}) = 3$, and so the rank of the K-theory is only 8. Consider the ideal in $\mathbb{Q}[s_1, s_1^{-1}, \dots]/K(T_{\mathrm{CY}})$ generated by $1 - s_1^{-5}$. This is exactly the orbit of the structure sheaf $\mathcal{O}_X$ under the action of the monodromy. Since monodromy is tensoring by $\mathcal{O}_X(D)$ for various divisors D, this will span the Chow ring of X_{CY}. In this example, over the rationals, the Chow ring gives the full even-dimensional cohomology and thus the full K-theory.

We need the ideal generated by the image of the structure sheaf in the monodromy ring of Z_Σ. There is an exact sequence [49]

$$0 \longrightarrow \frac{R}{(I:a)} \xrightarrow{a} \frac{R}{I} \longrightarrow \frac{R}{I+(a)} \longrightarrow 0, \tag{84}$$

describing the image of a in R/I. This motivates the definition:

[6]This is proven by showing that it is incompatible with any weight order.

Definition 8. The *reduced* monodromy ring is the quotient

$$\frac{\mathbb{Q}[s_1, s_1^{-1}, \dots, s_r, s_r^{-1}]}{(K(T_\Sigma) : \mathsf{O}_X)} \tag{85}$$

where we have an ideal quotient in the denominator and O_X is the K-theory class of the structure sheaf.

This reduced monodromy ring is expected to describe the monodromy of X_Σ. In our case, using (80), the ideal quotient $(K(T_\Sigma) : \mathsf{O}_X)$ is easily computed using Macaulay 2, but the result is rather messy. The important fact is that the degree of this quotient is 8, exactly as expected.

It is more interesting to find the reduced monodromy ring in the Landau–Ginzburg phase. It is easy to see that

$$\frac{\mathbb{Q}[s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}]}{K(T_{\mathrm{LG}})} \cong \frac{\mathbb{Q}[s]}{\langle s^{15} - 1\rangle}. \tag{86}$$

This reflects the fact that it is a $\mathbb{Z}_{15}$-orbifold of a Landau–Ginzburg theory. The monodromy corresponds to the $\mathbb{Z}_{15}$ "quantum symmetry" acting on the moduli space. Again, we clearly get a 15-dimensional vector space for the K-theory of Z_{LG}.

Now in order to find the reduced monodromy ring for the Landau–Ginzburg phase, we need to find the K-theory charge of the structure sheaf. We know how to do that using the tilting collection above. This again becomes an exercise in computer commutative algebra. We have an equivalence

$$\begin{aligned} 1 - s_1^{-5} \cong{}& -s_1^{-4}s_2^{-1} + 3s_1^{-3}s_2^{-1} - s_1^{-3}s_3^{-1} - 3s_1^{-2}s_2^{-1} + 3s_1^{-2}s_3^{-1} + s_1^{-1}s_2^{-1} \\ &- 3s_1^{-1}s_3^{-1} - 3s_1^{-4} + s_3^{-1} + 4s_1^{-3} - 4s_1^{-2} + 3s_1^{-1} \pmod{K(T_{\mathrm{CY}})}. \end{aligned} \tag{87}$$

The polynomial on the right is purely in terms of monomials with degree in the tilting collection. This polynomial, which we denote ξ, may therefore be used as the K-theory class of $\mathcal{O}_X$ in the Landau–Ginzburg phase.

Finally, one can show with a little more computer algebra

$$\frac{\mathbb{Q}[s_1, s_1^{-1}, s_2, s_2^{-1}, s_3, s_3^{-1}]}{(K(T_{\mathrm{LG}}) : \xi)} \cong \frac{\mathbb{Q}[s]}{\langle s^8 - s^7 + s^5 - s^4 + s^3 - s + 1\rangle}. \tag{88}$$

This eight-dimensional space is the K-theory of the Landau–Ginzburg theory.

The companion matrix of $s^8 - s^7 + s^5 - s^4 + s^3 - s + 1$ has eigenvalues

$$\alpha, \alpha^2, \alpha^4, \alpha^7, \alpha^8, \alpha^{11}, \alpha^{13}, \alpha^{14}, \tag{89}$$

where $\alpha = \exp(2\pi i/15)$.

The GKZ system has 15 linearly independent solutions. The above monodromy around the Landau–Ginzburg point in the moduli space tells us which of the 8 solutions are pertinent for X_{LG}. The older method for determining this would have been to use intersection theory to establish the correct 8 solutions at large radius limit and then analytically continue these solutions to the Landau–Ginzburg phase. This method using the monodromy ring is much less arduous.

Acknowledgements I thank R. Plesser for many useful discussions. This work was partially supported by NSF grant DMS–0905923. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

1. Grayson, D.R., Stillman, M.E.: Macaulay 2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/
2. Atiyah, M.: Topological quantum field theories. Inst. Hautes Études Sci. Publ. Math. **68**, 175–186 (1989)
3. Lurie, J.: On the classification of topological field theories. In: Current Developments in Mathematics, 2008, pp. 129–280. International Press, Somerville (2009). arXiv:0905.0465.
4. Douglas, M.R., Gross, M. (eds.): Dirichlet branes and mirror symmetry. Clay Mathematical Monographs, vol. 4. American Mathematical Society, New York, Clay Mathematics Institute, Cambridge (2009)
5. Witten, E.: Mirror manifolds and topological field theory. In: Yau, S.-T. (ed.) Essays on Mirror Manifolds. International Press, Somerville (1992). hep-th/9112056
6. Douglas, M.R.: D-Branes, categories and $N=1$ supersymmetry. J. Math. Phys. **42**, 2818–2843 (2001). hep-th/0011017
7. Aspinwall, P.S., Lawrence, A.E.: Derived categories and zero-brane stability. J. High Energy Phys. **08**, 004 (2001). hep-th/0104147
8. Aspinwall, P.S.: D-Branes on Calabi–Yau manifolds. In: Maldacena, J.M. (ed.) Progress in String Theory. TASI 2003 Lecture Notes, pp. 1–152. World Scientific, Singapore (2005). hep-th/0403166
9. Fukaya, K., Oh, Y.-G., Ohta, H., Ono, K.: Lagrangian Intersection Floer Theory—Anomaly and Obstruction. (2000). http://www.math.kyoto-u.ac.jp/~fukaya/fukaya.html
10. Dijkgraaf, R., Verlinde, H.L., Verlinde, E.P.: Notes on topological string theory and 2-D quantum gravity. In Green, M., et al. (eds.) String Theory and Quantum Gravity: Proceedings. World Scientific, Singapore (1991). Proceedings of Trieste Spring School: String Theory and Quantum Gravity, Trieste, Italy, Apr 23–1 May (1990)
11. Costello, K.: Topological conformal field theories and Calabi–Yau categories. Adv. Math. **210**, 165–214 (2007). arXiv:math/0412149
12. Toën, B.: The homotopy theory of dg-categories and derived Morita theory. Invent. Math. **167**, 615–667 (2007)
13. Oda, T., Park, H.S.: Linear Gale transforms and Gelfand–Kapranov–Zelevinskij decompositions. Tôhoku Math. J. **43**, 375–399 (1991)
14. Cox, D.A.: The homogeneous coordinate ring of a toric variety. J. Algebraic Geom. **4**, 17–50 (1995). alg-geom/9210008
15. Cox, D.A., Little, J.B., Schenck, H.K.: Toric varities. Graduate Studies in Mathematics, vol. 124. American Mathematical Society, Providence (2010)

16. Hartshorne, R.: Algebraic geometry. Graduate Texts in Mathematics, vol. 52. Springer, New York (1977)
17. Bondal, A., Orlov, D.: Semiorthogonal decomposition for algebraic varieties. alg-geom/9506012
18. Witten, E.: Phases of $N = 2$ theories in two dimensions. Nucl. Phys. **B403**, 159–222 (1993). hep-th/9301042
19. Herbst, M., Hori, K., Page, D.: Phases of $N = 2$ theories in $1 + 1$ dimensions with boundary. arXiv:0803.2045
20. Pantev, T., Sharpe, E.: String compactifications on Calabi–Yau stacks. Nucl. Phys. **B733**, 233–296 (2006). hep-th/0502044
21. Segal, E.: Equivalence between GIT quotients of Landau–Ginzburg B-models. Comm. Math. Phys. **304**, 411–432 (2011)
22. Miller, E., Sturmfels, B.: Combinatorial Commutative Algebra. Springer, New York (2005)
23. Aspinwall, P.S.: Topological D-branes and commutative algebra. Commun. Num. Theory. Phys. **3**, 445–474 (2009). hep-th/0703279
24. Eisenbud, D., Mustaţă, M., Stillman, M.: Cohomology on toric varieties and local cohomology with monomial supports. J. Symbolic Comput. **29**, 583–600 (2000). arXiv:math/0001159
25. Van den Bergh, M.: Non-commutative crepant resolutions. In: The Legacy of Niels Henrik Abel: The Abel Bicentennial, Oslo 2002, pp. 749–770. Springer, Berlin (2004). arXiv:math/0211064
26. Aspinwall, P.S.: D-Branes on toric Calabi–Yau varieties. arXiv:0806.2612
27. Orlov, D.: Derived categories of coherent sheaves and triangulated categories of singularities. In: Algebra, Arithmetic, and Geometry: in Honor of Yu.I. Manin. Vol. II. Progress in Mathematics, vol. 270. pp. 503–531. Birkhäuser, Boston (2009). math.AG/ 0506347
28. Avramov, L.L., Grayson, D.R.: Resolutions and cohomology over complete intersections. In: Eisenbud, D., et al. (eds.) Computations in Algebraic Geometry with Macaulay 2, Algorithms and Computations in Mathematics, vol. 8, pp. 131–178. Springer, New York (2001)
29. Gulliksen, T.H.: A change of ring theorem with applications to Poincare series and intersection multiplicity. Math. Scand. **34**, 167–183 (1974)
30. Eisenbud, D.: Homological algebra on a complete intersection, with an application to group representations. Trans. Amer. Math. Soc. **260**, 35–64 (1980)
31. Ballard, M., Favero, D., Katzarkov, L.: Variation of geometric invariant theory quotients and derived categories. arXiv:1203.6643
32. Kapustin, A., Li, Y.: D-branes in Landau–Ginzburg models and algebraic geometry. J. High Energy Phys. **12** 005 (2003). hep-th/0210296
33. Hori, K., Walcher, J.: D-branes from matrix factorizations. Compt. Rendus Phys. **5**, 1061–1070 (2004). hep-th/0409204
34. Ashok, S.K., Dell'Aquila, E., Diaconescu, D.-E., Florea, B.: Obstructed D-branes in Landau–Ginzburg orbifolds. Adv. Theor. Math. Phys. **8**, 427–472 (2004). hep-th/0404167
35. Aspinwall, P.S.: The Landau–Ginzburg to Calabi–Yau dictionary for D-branes, J. Math. Phys. **48**, 082304 (2007). arXiv:hep-th/0610209
36. Vafa, C., Warner, N.: Catastrophes and the classification of conformal theories. Phys. Lett. **218B**, 51–58 (1989)
37. Kapustin, A., Rozansky, L.: On the relation between open and closed topological strings. Commun. Math. Phys. **252**, 393–414 (2004). arXiv:hep-th/0405232
38. Dyckerhoff, T.: Compact generators in categories of matrix factorizations. Duke Math. J. **159**, 223–274 (2011). arXiv:0904.4713
39. Joyce, D.: Lectures on Calabi–Yau and special Lagrangian geometry. In: Global Theory of Minimal Surfaces. Clay Mathematical Proceedings, vol. 2. American Mathematical Society, Providence (2005). math.DG/0108088
40. Douglas, M.R., Fiol, B., Römelsberger, C.: Stability and BPS branes. J. High Energy Phys. 0509, 006 (2005). hep-th/0002037
41. Aspinwall, P.S., Douglas, M.R.: D-brane stability and monodromy. J. High Energy Phys. **05**, 031 (2002), hep-th/0110071

42. Bridgeland, T.: Stability conditions on triangulated categories. Ann. Math. (2) **166**, 317–345 (2007). math.AG/0212237
43. Gel'fand, I.M., Zelevinskiĭ A.V., Kapranov, M.M.: Hypergeometric functions and toral (toric) manifolds. Func. Anal. Appl. **23**, 94–106 (1989). (from Russian in Funk. Anal. Pril. **23**)
44. Gelfand, I.M., Kapranov, M.M., Zelevinski, A.V.: Discriminants, Resultants and Multidimensional Determinants. Birkhäuser, Basel (1994)
45. Batyrev, V.V.: Variations of the mixed hodge structure of affine hypersurfaces in algebraic tori. Duke Math. J. **69**, 349–409 (1993)
46. Aspinwall, P.S., Greene, B.R., Morrison, D.R.: The monomial-divisor mirror map. Internat. Math. Res. Notices 319–33 (1993). alg-geom/9309007
47. Hosono, S., Lian, B.H., Yau, S.-T.: GKZ-generalized hypergeometric systems in mirror symmetry of Calabi-Yau hypersurfaces. Comm. Math. Phys. **182**, 535–577 (1996). alg-geom/9511001
48. Batyrev, V.V.: Dual polyhedra and mirror symmetry for Calabi–Yau hypersurfaces in toric varieties. J. Alg. Geom. **3**, 493–535 (1994)
49. Eisenbud, D.: Commutative algebra with a view towards algebraic geometry. Graduate Texts in Mathematics, vol. 150. Springer, New York (2004)

Measuring Singularities with Frobenius: The Basics

Angélica Benito, Eleonore Faber, and Karen E. Smith

1 Introduction

Consider a polynomial f over some field k, vanishing at some point x in k^n. By definition, f is smooth at x (or the hypersurface defined by f is smooth at x) if and only if some partial derivative $\frac{\partial f}{\partial x_i}$ is nonzero there. Otherwise, f is singular at x. But how singular? Can we quantify the singularity of f at x?

The multiplicity is perhaps the most naive measurement of singularities. Because f is singular at x if all the first-order partial derivatives of f vanish there, it is natural to say that f is even more singular if also all the second-order partials vanish, and so forth. The *order,* or *multiplicity,* of the singularity at x is the largest d such that for all differential operators ∂ of order less than d, ∂f vanishes at x. Choosing coordinates so that x is the origin, it is easy to see that the multiplicity is simply the degree of the lowest degree term of f.

The multiplicity is an important first step in measuring singularities, but it is too crude to give a good measurement of singularities. For example, the polynomials xy and $y^2 - x^3$ both define singularities of multiplicity two, though the former is clearly less singular than the latter. Indeed, xy defines a simple normal crossing divisor, whereas the singularity of the cuspidal curve defined by $y^2 - x^3$ is quite complicated, and that of, for example, $y^2 - x^{17}$ is even more so (See Fig. 1).

This chapter describes the first steps toward understanding a much more subtle measure of singularities which arises naturally in three different contexts—analytic, algebro-geometric, and finally, algebraic. Miraculously, all three approaches lead

---

A. Benito • K.E. Smith (✉)
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA
e-mail: abenitos@umich.edu; kesmith@umich.edu

E. Faber
Department of computer and mathematical science, University of Toronto at Scarborough, Toronto, Ont. M1A 1C4, Canada
e-mail: efaber@math.toronto.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_3,

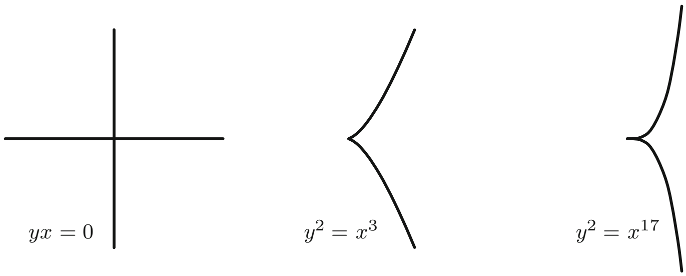


Fig. 1 Curves with multiplicity 2 at the origin

to essentially the same measurement of singularities: the log canonical threshold (in characteristic zero) and the closely related F-pure threshold (in characteristic p). The log canonical threshold, or complex singularity exponent, can be defined analytically (via integration) or algebro-geometrically (via resolution of singularities). As such, it is defined only for polynomials over $\mathbb{C}$ or other characteristic zero fields. The F-pure threshold, whose name we shorten to F-threshold here, by contrast, is defined only in prime characteristic. Its definition makes use of the Frobenius, or pth power map. Remarkably, these two completely different ways of quantifying singularities turn out to be intimately related. As we will describe, if we fix a polynomial with integer coefficients, the F-threshold of its "reduction mod p" approaches its log canonical threshold as p goes to infinity.

Both the log canonical threshold and the F-threshold can be interpreted as critical numbers for the behavior of certain associated ideals, called the multiplier ideals in the characteristic zero setting and the test ideals in the characteristic p world. Both naturally give rise to higher order analogs, called "jumping numbers." We will also introduce these refinements.

We present only the first steps in understanding these invariants, with an emphasis on the prime characteristic setting. Attempting only to demystify the concepts in the simplest cases, we make no effort to discuss the most general case or to describe the many interesting connections with deep ideas in analysis, topology, algebraic geometry, number theory, and commutative algebra. The reader who is inspired to dig deeper will find plenty of more sophisticated survey articles and a plethora of connections to many ideas throughout mathematics, including the Bernstein–Sato polynomial [34], Varchenko's work on mixed Hodge structures on the vanishing cycle [62], the Hodge spectrum [58], the Igusa zeta function [33], motivic integration and jet schemes [40], Lelong numbers [12], Tian's invariant for studying Kähler–Einstein metrics [45], various vanishing theorems for cohomology [37, Chap. 9], birational rigidity [10], Shokurov's approach to termination of flips [52], Hasse invariants [2], the monodromy action on the Milnor fiber, and Frobenius splitting and tight closure.

There are several surveys which are both more sophisticated and pay more attention to history. In particular, the classic survey by Kollár in [34, Sects. 8–10] contains a deeper discussion of the characteristic zero theory, as do the more recent lectures of Budur [9], mainly from the algebro-geometrical perspective. For a more analytic discussion, papers of Demailly are worth looking at, such as the article [11]. Likewise, for the full characteristic p story, Schwede and Tucker's survey of test ideals [48] is very nice. The survey [41] contains a modern account of both the characteristic p and characteristic zero theory.

2 Characteristic Zero: Log Canonical Threshold and Multiplier Ideals

In this section we work with polynomials over the complex numbers $\mathbb{C}$. Let $\mathbb{C}^N \xrightarrow{f} \mathbb{C}$ be a polynomial (or analytic) function, vanishing at a point x.

2.1 Analytic Approach

Approaching singularities from an analytic point of view, we consider how fast the (almost everywhere defined) function

$$\begin{aligned} \mathbb{C}^N = \mathbb{R}^{2N} &\longrightarrow \mathbb{R} \\ z &\longmapsto \frac{1}{|f(z)|} \end{aligned}$$

"blows up" at a point x in the zero set of f. We attempt to measure this singularity via integration. For example, is this function square integrable in a neighborhood of x? The integral

$$\int \frac{1}{|f|^2}$$

never converges in any small ball around x, but we can dampen the rate at which $\frac{1}{|f|}$ blows up by raising to a small positive power λ. Indeed, for sufficiently small positive real numbers λ, depending on f, the integral

$$\int_{B_\varepsilon(x)} \frac{1}{|f|^{2\lambda}}$$

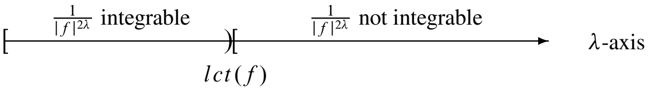
is finite, where $B_\varepsilon(x)$ denotes a ball of sufficiently small radius around x. As we vary the parameter λ from very small positive values to larger ones, there is a critical value at which the function $\frac{1}{|f|^\lambda}$ suddenly fails to be L^2 locally at x. This is the log canonical threshold or complex singularity exponent of f at x. That is:

Definition 1. The *complex singularity exponent* of f (at x) is defined as

$$lct_x(f) := \sup\left\{\lambda \in \mathbb{R}_+ \;\middle|\; \text{there exists a neighborhood } B \text{ of } x \text{ such that } \int_B \frac{1}{|f|^{2\lambda}} < \infty\right\}.$$

When the point x is understood, we denote the complex singularity exponent simply by $lct(f)$.

The following figure depicts the λ-axis: for small values of λ the function $z \mapsto \frac{1}{|f(z)|}$ belongs to L^2 locally in a neighborhood of a point x; for larger λ it does not. It is not clear whether or not the function is integrable *at* the complex singularity exponent; we will see shortly that it is.



This numerical invariant is more commonly known as the *log canonical threshold,* but it is natural to use the original analytic name when approaching it from the analytic point of view. See Remark 11.

Example 2. If f is smooth at x, then its complex singularity exponent at x is 1. Indeed, in this case the polynomial f can be taken to be part of a system of local coordinates for $\mathbb{C}^n$ at x. It is then easy to compute that the integral

$$\int_{B(x)} \frac{1}{|f|^{2\lambda}}$$

always converges on any bounded ball $B(x)$ for any positive $\lambda < 1$. Indeed, this computation is a special case of Example 3.

Example 3. Let $f = z_1^{a_1} \cdots z_N^{a_N}$ be a monomial in $\mathbb{C}[z_1, \ldots, z_N]$, which defines a singularity at the origin in $\mathbb{C}^N$. Let us compute its complex singularity exponent. By definition, we need to integrate

$$\frac{1}{|z_1^{a_1} \cdots z_N^{a_N}|^{2\lambda}}$$

over a ball around the origin. To do so, we use polar coordinates. We have each $|z_i| = r_i$ and each $dz_i \wedge d\bar{z}_i = r_i dr_i \wedge d\vartheta_i$. Hence we see that

$$\int \frac{1}{|z_1|^{2a_1\lambda} \cdots |z_N|^{2a_N\lambda}}$$

converges in a neighborhood B of the origin if and only if

$$\int_B \frac{r_1 \cdots r_N}{r_1^{2a_1\lambda} \cdots r_N^{2a_N\lambda}} = \int_B \frac{1}{r_1^{2a_1\lambda-1} \cdots r_N^{2a_N\lambda-1}}$$

converges. By Fubini's theorem, this integral converges if and only if $1-2a_i\lambda > -1$ for all i, that is, $\lambda < \frac{1}{a_i}$ for all i. Thus

$$lct(z_1^{a_1} \cdots z_N^{a_N}) = \min_i \left\{\frac{1}{a_i}\right\}.$$

If f has "worse" singularities, the function $\frac{1}{|f|}$ will blow up faster, and the complex singularity exponent will typically be smaller. In particular, the complex singularity exponent is always less than or equal to the complex singularity exponent of a smooth point or one.[1] Although it is not obvious, the complex singularity exponent is always a positive rational number. We prove this in the next subsection using Hironaka's theorem on resolution of singularities.

2.2 Computing Complex Singularity Exponent by Monomializing

Hironaka's beautiful theorem on resolution of singularities allows us to reduce the computation of the integral in the definition of the complex singularity exponent for any polynomial (or analytic) function to the monomial case. Let us recall Hironaka's theorem (cf. [27]).

Theorem 4. *Every polynomial (or analytic) function on $\mathbb{C}^N$ has a monomialization. That is, there exists a proper birational morphism $X \xrightarrow{\pi} \mathbb{C}^N$ from a smooth variety X such that both*

$$f \circ \pi \quad \text{and} \quad \text{Jac}_{\mathbb{C}}(\pi)$$

are monomials (up to unit) in local coordinates locally at each point of X.

Since X is a smooth complex variety, it has a natural structure of a complex manifold. Saying that π is a morphism of algebraic varieties means simply that it is defined locally in coordinates by polynomial (hence analytic) functions; therefore, π is also a holomorphic mapping of complex manifolds. The word "proper" in this context can be understood in the usual analytic sense: the preimage of a compact set

[1]One caveat: the singularity exponent behaves somewhat differently over $\mathbb{R}$ due to the possibility that a polynomial's zeros are hidden over $\mathbb{R}$, so that $\frac{1}{|f|}$ may fail to blow up as expected or even at all!

is compact[2]. The fact that π is birational (meaning it has an inverse on some dense open set) is not relevant at the moment, beyond the fact that the dimension of X is necessarily N.

The condition that both $f \circ \pi$ and $\mathrm{Jac}_{\mathbb{C}}(\pi)$ are monomials (up to unit) locally at a point $y \in X$ means that we can find local coordinates $z_1, \ldots, z_N$ at y, such that both

$$f \circ \pi = u z_1^{a_1} \cdots z_N^{a_N} \tag{1}$$

and the holomorphic Jacobian[3]

$$\mathrm{Jac}_{\mathbb{C}}(\pi) = v z_1^{k_1} \cdots z_N^{k_N} \tag{2}$$

where u and v are some regular (or analytic) functions defined in a neighborhood of y but not vanishing at y.

The properness of the map π guarantees that the integral

$$\int \frac{1}{|f|^{2\lambda}}$$

converges in a neighborhood of the point x if and only if the integral

$$\int \frac{\mathrm{Jac}_{\mathbb{R}}(\pi)}{|f \circ \pi|^{2\lambda}}$$

converges in a neighborhood of $\pi^{-1}(x)$, where $\mathrm{Jac}_{\mathbb{R}}(\pi)$ is the (real) Jacobian of the map π considered as a smooth map of real $2N$-dimensional manifolds. Recalling that

$$\mathrm{Jac}_{\mathbb{R}}(\pi) = |\mathrm{Jac}_{\mathbb{C}}(\pi)|^2,$$

(see [17, pp. 17–18]) and using that $\pi^{-1}(x)$ is compact, Hironaka's theorem reduces the convergence of this integral to a computation with monomials in each of finitely many charts U covering X:

$$\int_{\pi^{-1}(B(x))\cap U} \frac{|z_1^{k_1} \cdots z_N^{k_N}|^2}{|z_1^{a_1} \cdots z_N^{a_N}|^{2\lambda}}.$$

Doing an analogous computation to that in Example 3, we can conclude that the integral is finite if and only if in each chart we have

$$k_i - \lambda a_i > -1 \tag{3}$$

---

[2] Alternatively, it can be taken in the usual algebraic sense as defined in Hartshorne, [22, Ch. 2, Sect. 4].

[3] If we write π in local holomorphic coordinates as $(z_1, \ldots, z_N) \mapsto (f_1, \ldots, f_N)$, then $\mathrm{Jac}_{\mathbb{C}}(\pi)$ is the holomorphic function obtained as the determinant of the $N \times N$ matrix $\left(\frac{\partial f_i}{\partial z_j}\right)$.

for all i, or equivalently,

$$\lambda < \frac{k_i + 1}{a_i} \tag{4}$$

for all i. Hence

$$lct_x(f) = \min_{\substack{i \\ \text{all charts}}} \left\{ \frac{k_i + 1}{a_i} \right\}. \tag{5}$$

In particular, remembering that the map was proper so that only finitely many charts are at issue here, we have:

Corollary 5. *The complex singularity exponent of a complex polynomial at any point is a rational number.*

2.3 Algebro-Geometric Approach

In the world of algebraic geometry, we might attempt to measure the singularities of f by trying to measure the complexity of a resolution of its singularities. Hironaka's theorem can be stated as follows:

Theorem 6. *Fix a polynomial (or analytic) function f on $\mathbb{C}^N$. There exists a proper birational morphism $X \xrightarrow{\pi} \mathbb{C}^N$ from a smooth variety X such that the pullback of f defines a divisor F_π whose support has simple normal crossings and which is also in normal crossings with the exceptional divisor (the locus of points on X at which π fails to be an isomorphism). Furthermore, the morphism π can be assumed to be an isomorphism outside the singular set of f.*

The proper birational morphism π is usually called a *log resolution* of f in this context. The support of the divisor defined by the pullback of f is simply the zero set of $f \circ \pi$. The condition that it has normal crossings means that it is a union of smooth hypersurfaces meeting transversely. In more algebraic language, a divisor with normal crossing support is one whose equation can be written as a monomial in local coordinates at each point of X. Thus Theorem 6 is really just a restatement of Theorem 4.[4]

Hironaka actually proved more: such a log resolution can be constructed by a sequence of blowings up at smooth centers. We might consider the polynomial f to be "more singular" if the number of blowings up required to resolve f, and their relative complicatedness, is great. However, because there is no canonical way to resolve singularities, we need a way to compare across different resolutions. This is done with the canonical divisor.

---

[4]As stated here, Theorem 6 is actually a tiny bit stronger, since the condition that we have *simple* normal crossings rules out self-crossings. The difference is immaterial to our discussion.

2.4 The Canonical Divisor of a Map

Fix a proper birational morphism $X \xrightarrow{\pi} Y$ between smooth varieties. The holomorphic Jacobian (determinant) $\mathrm{Jac}_{\mathbb{C}}(\pi)$ can be viewed as a regular function locally on charts of X. Its zero set (counting multiplicity) is the *canonical divisor* of π (or *relative canonical divisor* of X over Y), denoted by K_π. Because the Jacobian matrix is invertible at $x \in X$ if and only if π is an isomorphism there, the canonical divisor of π is supported precisely on the exceptional set E, which by definition consists of the points in X at which π is not an isomorphism. In particular, since it is locally the zero set of this Jacobian determinant, the exceptional set E is always a codimension one subvariety of X. Moreover, this exceptional set is more naturally considered as a *divisor:* we label each of the components of E by the order of vanishing of the Jacobian along it. This is the canonical divisor K_π. That is,

$$K_\pi = \mathrm{div}(\mathrm{Jac}_{\mathbb{C}}(\pi)) = \sum k_i E_i,$$

where the sum ranges through all of the components E_i of the exceptional set E and where k_i is the order of vanishing[5] of $\mathrm{Jac}_{\mathbb{C}}(\pi)$ along E_i. Thus we can view the canonical divisor K_π as a precise "difference" between birationally equivalent varieties X and Y.

To measure the singularities of a polynomial f, consider a log resolution $X \xrightarrow{\pi} \mathbb{C}^N$. The polynomial f defines a simple crossing divisor F_π on X, namely the zero set (with multiplicities) of the regular function $f \circ \pi$,

$$F_\pi = \mathrm{div}(f \circ \pi) = \sum a_i D_i$$

where the D_i range through all irreducible divisors on X and the a_i are the orders of vanishing[6] of $f \circ \pi$ along each. If we denote the divisor of f in $\mathbb{C}^N$ by F, then F_π is simply $\pi^* F$. There are two types of divisors in the support of F_π: the birational transforms $\widetilde{F}_i$ of the components of F, and exceptional divisors E_i. All are smooth. Note that locally in charts, both types of divisors—the E_i and the $\widetilde{F}_i$—are defined by some local coordinates z_i on X.

Using this language, we examine our computation for the convergence of the integral

$$\int_{B(x)} \frac{1}{|f|^{2\lambda}}.$$

[5] These are the same k_i appearing in expressions (2) as we range over all charts of X. Note that there are typically many more than N components E_i despite the fact that in the expression (2) we were only seeing at most N of them at a time in each chart.

[6] Of course, the order of vanishing is zero along any irreducible divisor not in the support of F, so the sum is finite. Again, these are the same a_i as in formula (1); there are typically many divisors in the support of $\pi^* F$ although in formula (1) we see at most N in each chart.

The condition (4) that $k_i - \lambda a_i > -1$ is equivalent to the condition that all coefficients of the $\mathbb{R}$-divisor

$$K_\pi - \lambda F_\pi$$

are greater than -1. Put differently, the integral $\int \frac{1}{|f|^{2\lambda}}$ converges in a neighborhood of x if and only if the "round-up" divisor[7]

$$\lceil K_\pi - \lambda F_\pi \rceil$$

is effective. [Strictly speaking, since we are computing the complex singularity exponent at a particular point x, we should throw away any components of $K_\pi - \lambda F_\pi$ whose image on $\mathbb{C}^N$ does not contain x; that is, we should consider a log resolution of singularities only in a sufficiently small neighborhood of x].

Again we arrive at the following formula for the complex singularity exponent of f at x:

Corollary 7. *Let $\pi : X \longrightarrow \mathbb{C}^N$ be a log resolution of the polynomial f. If we write*

$$K_\pi = \sum k_i D_i \qquad \text{and} \qquad \operatorname{div}(f \circ \pi) = \sum a_i D_i, \tag{6}$$

where the D_i range through all irreducible divisors on X, then the complex singularity exponent or the log canonical threshold of f at point x is the minimum, taken over all indices i such that $x \in \pi(D_i)$, of the rational numbers

$$\frac{k_i + 1}{a_i}.$$

The complex singularity exponent is better known in algebraic geometry as the log canonical threshold.

Remark 8. The condition that $\lceil K_\pi - \lambda F_\pi \rceil$ is effective is independent of the choice of log resolution. This follows from our characterization of the convergence of the integral but can also be shown directly using the tools of algebraic geometry (see [35]). Although we did not motivate the study of $K_\pi - \lambda F_\pi$ in purely algebro-geometric terms, the $\mathbb{R}$- divisors $K_\pi - \lambda F$ turn out to be quite natural in birational algebraic geometry, without reference to the integrals. See, for example, [34]. In any case, our discussion shows that the definition of log canonical threshold can be restated as follows:

Definition 9. The *log canonical threshold* of a polynomial f is defined as

$$lct_x(f) := \sup \left\{ \lambda \in \mathbb{R}_+ \;\middle|\; \lceil K_\pi - \lambda F_\pi \rceil \ \text{ is effective} \right\},$$

[7]Given a divisor D with real coefficients, we define the roundup $\lceil D \rceil$ as the integral divisor obtained by rounding up all coefficients of prime divisors to the nearest integer. In the same way, $\lfloor D \rfloor$ is obtained by rounding down.

where $X \xrightarrow{\pi} \mathbb{C}^N$ is any log resolution of f (in a neighborhood of x), K_π is its relative canonical divisor, and F_π is the divisor on X defined by $f \circ \pi$. We can also define the global log canonical threshold by taking π to be a resolution at all points, not just in a neighborhood of x.

Remark 10. Note that loosely speaking, the more complicated the resolution, the more likely λ will have to be small in order to make $K_\pi - \lambda F_\pi$ close to effective. This essentially measures the complexity of the pullback of f to the log resolution. The presence of the K_π term accounts for the added multiplicity that would have been present in *any* resolution because of the nature of blowup $\mathbb{C}^N$ to get X, thus "standardizing" across different resolutions.

It is also clear from this point of view that $\lceil K_\pi - \lambda F_\pi \rceil$ is always effective for very small (positive) λ and that as we enlarge λ it stays effective until we suddenly hit the log canonical threshold of f, at which point at least one coefficient is exactly negative one.

Remark 11. The name *log canonical* comes from birational geometry. A pair (Y, D) consisting of a $\mathbb{Q}$-divisor on a smooth variety Y is said to be log canonical if, for any proper birational morphism $\pi : X \longrightarrow Y$ with X smooth (or equivalently, any fixed log resolution), the divisor $K_\pi - \pi^* D$ has all coefficients ≥ -1. This condition is independent of the choice of π (see [35]). Thus the log canonical threshold of f at x is the supremum, over positive $\lambda \in \mathbb{R}$ such that $(\mathbb{C}^n, \lambda \mathrm{div}(f))$ is log canonical in a neighborhood of x.

Example 12. The log canonical threshold of any complex polynomial f is bounded above by one. Indeed, suppose for simplicity that f is irreducible, defining a hypersurface D with isolated singularity at x. Let $\pi : X \longrightarrow \mathbb{C}^N$ be a log resolution of f. We have

$$K_\pi = \sum k_i E_i,$$

where all the E_i are exceptional, and

$$\mathrm{div}(f \circ \pi) = \sum a_i E_i + \widetilde{D},$$

where the E_i are exceptional and $\widetilde{D}$ is the birational transform of D on X. Then the log canonical threshold is the minimum value of

$$\min_i \left\{ \frac{k_i + 1}{a_i}, 1 \right\} \tag{7}$$

as we range through the exceptional divisors of π. More generally, the argument adapts immediately to show that if f factors into irreducibles as $f = f_1^{b_1} f_2^{b_2} \cdots f_t^{b_t}$, then the log canonical threshold is bounded above by the minimal value of $\frac{1}{b_i}$.

2.5 Computations of Log Canonical Thresholds

The canonical divisor of a morphism plays a starring role in birational geometry, and in particular, as we have seen, in the computation of the log canonical threshold. Before computing some more examples, we isolate two helpful properties of K_π.

Fact 13. Let $X \xrightarrow{\pi} Y$ be the blowup along a smooth subvariety of codimension c in the smooth variety Y. Then the relative canonical divisor is

$$K_\pi = (c-1)E,$$

where E denotes the exceptional divisor of the blowup.

Fact 14. Consider a sequence of proper birational morphisms $X_3 \xrightarrow{\pi} X_2 \xrightarrow{\nu} X_1$, where all the X_i are smooth. Then,

$$K_{\nu\circ\pi} = \pi^* K_\nu + K_\pi.$$

The proof of both these facts is easy exercises in local coordinates and left to the reader.

Example 15 (A cuspidal singularity). We compute the log canonical threshold of the cuspidal curve D given by $f = x^2 - y^3$ in $\mathbb{C}^2$ (at the origin, its unique singular point). The curve is easily resolved (i.e., the polynomial f is easily monomialized) by a sequence of three point blowups at points: $X_3 \xrightarrow{\psi} X_2 \xrightarrow{\nu} X_1 \xrightarrow{\phi} \mathbb{C}^2$, whose composition we denote by π and which create exceptional divisors E_1, E_2, and E_3, respectively.[8] [Here ϕ is the blowup at the origin, ν is the blowup of the unique intersection point with the birational transform of D with E_1, and ψ is the blowup of the unique intersection point of the birational transform of D on X_2 with E_2]. There are four relevant divisors on X_3 to consider: the three exceptional divisors E_1, E_2, and E_3, and the birational transform of D on X_3. Using the two facts above, it is easy to compute that

$$K_\pi = E_1 + 2E_2 + 4E_3$$

and

$$F = \operatorname{div}(f \circ \pi) = D + 2E_1 + 3E_2 + 6E_3$$

Hence $lct(f) = \frac{5}{6}$.

[8]In a slight, but very helpful, abuse of terminology, we use the same symbol to denote an irreducible divisor and its birational transform on any model.

Example 16. As an exercise, the reader can compute that for $f = x^m - y^n$ with $\gcd(m,n) = 1$, then $lct(f) = \frac{1}{m} + \frac{1}{n}$. The resolution is constructed as in Example 15 but may require a few more blowups to resolve.

Example 17. Let f be a homogenous polynomial of degree d in N variables, with an isolated singularity at the origin. Then $lct(f) = \frac{N}{d}$, if $d \geq N$ and 1 otherwise. Indeed, one readily checks that blowup the origin, we obtain a log resolution, $X \xrightarrow{\pi} \mathbb{C}^N$, with one exceptional component E. Using Fact 13, we compute that

$$K_\pi = (N-1)E.$$

Also, the divisor D defined by f pulls back tovspace*6pt

$$F = dE + D,$$

where again, D denotes also the birational transform of D on X. Thus

$$lct(f) = \min\left\{\frac{(N-1)+1}{d}, \frac{0+1}{1}\right\} = \min\left\{\frac{N}{d}, 1\right\}.$$

Remark 18. The log canonical threshold describes the singularity but it does not characterize it. For example, the previous example gives examples of numerous non-isomorphic non-smooth points whose log canonical threshold is one, the same as a smooth point.

Remark 19. In general it is hard to compute the log canonical threshold, but there are algorithms to compute it in special cases such as the monomial case [32], the toric case [3], or the case of two variables [61]. In all these cases, the reason the log canonical threshold can be computed is that a resolution of singularities can be explicitly understood.

2.6 Multiplier Ideals and Jumping Numbers

Our definition of log canonical threshold leads naturally to a family of richer invariants called the *multiplier ideals* of f, which are ideals in the polynomial ring indexed by the positive real numbers.

Again, multiplier ideals can be defined analytically or algebro-geometrically.

Definition 20 (Analytic Definition, cf. [11]). Fix $f \in \mathbb{C}[x_1, \ldots, x_n]$. For each $\lambda \in \mathbb{R}_+$, define the *multiplier ideal* of f as

$$\mathcal{J}(f^{\lambda}) = \Big\{ h \in \mathbb{C}[x_1, \ldots, x_n] \;\Big|\; \text{there exists a neighborhood } B \text{ of } x \text{ such that} \\ \int_B \frac{|h|^2}{|f|^{2\lambda}} < \infty \Big\} .$$

Thus the multiplier ideals consist of functions that can be used as "multipliers" to make the integral converge. It is easy to check that this set $\mathcal{J}(f^{\lambda})$ is in fact an ideal of the ring $\mathbb{C}[x_1, \ldots, x_n]$.

Equivalently, we define the multiplier ideal in an algebro-geometric context using a log resolution.

Definition 21 (Algebro-Geometric Definition). Fix $f \in \mathbb{C}[x_1, \ldots, x_n]$. For each $\lambda \in \mathbb{R}_+$, define the *multiplier ideal* of f as

$$\mathcal{J}(f^{\lambda}) = \pi_* \mathcal{O}_X(\lceil K_\pi - \lambda F_\pi \rceil),$$

where $X \xrightarrow{\pi} \mathbb{C}^n$ is a log resolution of f, K_π is its relative canonical divisor, and F_π is the divisor on X determined by $f \circ \pi$. These are the polynomials whose pull-backs to X have vanishing no worse than that of the divisor $K_\pi - \lambda F_\pi$.

Recall that if D is a divisor on a smooth variety X, the notation $\mathcal{O}_X(D)$ denotes the sheaf of rational functions g on X such that $\operatorname{div}(g) + D$ is effective. Thus, in concrete terms, the multiplier ideal is

$$\pi_* \mathcal{O}_X(\lceil K_\pi - \lambda F \rceil) = \{h \in \mathbb{C}[x_1, \ldots, x_n] \mid \operatorname{div}(h \circ \pi) + \lceil K_\pi - \lambda F \rceil \geq 0\}.$$

It is straightforward to check that the argument we gave for translating the analytic definition of the log canonical threshold into algebraic geometry can be used to see that these two definitions of multiplier ideals are equivalent. Moreover, this also shows that Definition 21 is independent of the choice of resolution. For a direct algebro-geometric proof, see [37, Chap. 9].

Remark 22. Because K_π is an integral divisor, we have $\lceil K_\pi - \lambda F_\pi \rceil = K_\pi - \lfloor \lambda F_\pi \rfloor$. We caution the reader delving deeper into the subject, however, that when the notion of multiplier ideals and log canonical thresholds are generalized to divisors on singular ambient spaces, there are situations in which K_π is a nonintegral $\mathbb{Q}$-divisor. In this case, we cannot assume that $\lceil K_\pi - \lambda F_\pi \rceil = K_\pi - \lfloor \lambda F_\pi \rfloor$.

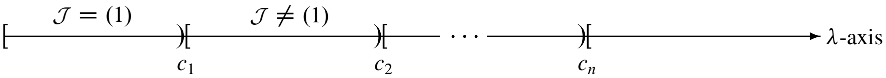
Proposition 23. *Fix a polynomial f and view its multiplier ideals $\mathcal{J}(f^{\lambda})$ as a family of ideals varying with λ. Then the following properties hold:*

1. *For $\lambda \in \mathbb{R}_+$ sufficiently small, $\mathcal{J}(f^{\lambda})$ is the unit ideal.*
2. *If $\lambda > \lambda'$, then $\mathcal{J}(f^{\lambda'}) \supset \mathcal{J}(f^{\lambda})$.*
3. *The log canonical threshold of f is*

$$lct(f) = \sup\{\lambda \mid \mathcal{J}(f^{\lambda}) = (1)\}.$$

4. *For each fixed value of* λ, *we have* $\mathcal{J}(f^{\lambda}) = \mathcal{J}(f^{\lambda+\varepsilon})$ *for small enough positive* ε. *(How small is small enough depends on* λ*).*
5. *There exist certain* $\lambda \in \mathbb{R}_{+}$ *such that* $\mathcal{J}(f^{\lambda-\varepsilon}) \supsetneq \mathcal{J}(f^{\lambda})$ *for all positive values of* ε.

All of these properties are easy to verify, thinking of what happens with the rounding up of the divisor $K_\pi - \lambda F$ as λ changes. As we imagine starting with a very small λ and increasing it, the properties above can be summarized by the following diagram:



$$(1) \supsetneq \mathcal{J}(f^{c_1}) \supsetneq \mathcal{J}(f^{c_2}) \supsetneq \cdots \supsetneq \mathcal{J}(f^{c_n}) \supsetneq \ldots$$

There are certain critical exponents c_i for which the multiplier ideal "jumps."

The critical numbers λ described in (5) and denoted by c_i in the diagram give a sequence of numerical invariants refining the log canonical threshold, which is the smallest of these. Formally:

Definition 24. The jumping numbers of $f \in \mathbb{C}[x_1, \ldots, x_n]$ are the positive real numbers c such that $\mathcal{J}(f^c) \subsetneq \mathcal{J}(f^{c-\varepsilon})$ for all positive ϵ.

Proposition 25. *The jumping numbers of* $f \in \mathbb{C}[x_1, \ldots, x_n]$ *are discrete and rational.*

Proof. Let $\pi : X \to \mathbb{C}^N$ be a log resolution of f, with $K_\pi = \sum k_i D_i$ and $F_\pi = \sum a_i D_i$, as before. It is easy to see that the critical values of $\lceil K_\pi - \lambda F_\pi \rceil$ occur only when $k_i - \lambda a_i \in \mathbb{N}$. That is, the jumping numbers are a subset of the numbers $\{\frac{k_i+m}{a_i}\}_{m\in\mathbb{N}}$. In particular, they are discrete and rational. □

Although an infinite sequence, the jumping numbers are actually determined by finitely many.

Theorem 26 (See, e.g., [37, Theorem 9.6.21]). *Fix a polynomial* f *as above. Then,*

$$\mathcal{J}(f^{1+\lambda}) = (f)\mathcal{J}(f^{\lambda}),$$

for $\lambda \geq 0$. *In particular, a positive real number* c *is a jumping number if and only if* $c + 1$ *is a jumping number.*

Thus the jumping numbers are periodic and completely determined by the finite set of jumping numbers less than or equal to 1.

Remark 27. It is quite subtle to determine which "candidate jumping numbers" of the form $\frac{k_i+1}{a_i}$ are actual jumping numbers. When f is a polynomial in 2 variables, there is some very pretty geometry behind understanding this question. See [56, 61].

Exercise 28. Show that 1 is a jumping number of every polynomial. [Hint: if f is irreducible, the jumping number 1 is contributed by the birational transform of $\mathrm{div}(f)$ on the log resolution].

Exercise 29. Show that the jumping numbers of a smooth f are the natural numbers $1, 2, 3, \ldots$

Exercise 30. Using the resolution described in Example 15, show that the multiplier ideals of $f = x^2 - y^3$ are as follows:

1. $\mathcal{J}(f^\lambda)$ is trivial for values of λ less than $\frac{5}{6}$.
2. $\mathcal{J}(f^\lambda) = \mathfrak{m} = (x, y)$ for $\frac{5}{6} \leq \lambda < 1$.
3. $\mathcal{J}(f^\lambda) = (f)$ for $1 \leq \lambda < \frac{11}{6}$.

Using Theorem 26, describe the multiplier ideal of $x^2 - y^3$ for any value of λ.

Exercise 31 (Harder). Show that the multiplier ideals of $f = x^2 - y^5$ are as follows:

(1) $\mathcal{J}(f^\lambda) = R$ is for values of λ less than $\frac{7}{10}$.
(2) $\mathcal{J}(f^\lambda) = \mathfrak{m} = (x, y)$ for $\frac{7}{10} \leq \lambda < \frac{9}{10}$.
(3) $\mathcal{J}(f^\lambda) = (x, y^2)$ for $\frac{9}{10} \leq \lambda < 1$.
(4) $\mathcal{J}(f^\lambda) = (f)$ for $1 \leq \lambda < \frac{17}{10}$.

Remark 32. The jumping numbers turn out to be related to many other well-studied invariants. For example, it is shown in [14] that the jumping numbers of f in the interval $(0, 1]$ are always (negatives of) roots of the so-called Bernstein–Sato polynomial b_f of f. The jumping numbers can also be viewed in terms of the Hodge spectrum arising from the monodromy action on the cohomology of the Milnor fiber of f [8].

Multiplier ideals have many additional properties, in addition to many deep applications which we don't even begin to describe. Lazarsfeld's book [37] gives an idea of some of these.

3 Positive Characteristic: The Frobenius Map and F-Thresholds

A natural question arises: What about positive characteristic?

Fix a polynomial f over a perfect field k. We wish to measure the singularity of f at some point where it vanishes. For concreteness and with no essential loss of generality, say, the field is $\mathbb{F}_p$ and the point is the origin, so that $f \in \mathfrak{m} = (x_1, \ldots, x_n) \subset \mathbb{F}_p[x_1, \ldots, x_n]$. How can we try to define an analog of log canonical threshold? In characteristic zero, we used real analysis to control the growth of the function $\frac{1}{|f|^\lambda}$ as we approached the singular points of f. But in characteristic p, can we even talk about taking fractional powers of f? Remarkably, the Frobenius map gives us a tool for raising polynomials to non-integer powers, and for considering their behavior near $\mathfrak{m}$.

3.1 The Frobenius Map

Let R be any ring of characteristic p, with no non-zero nilpotent elements.

Definition 1. The *Frobenius map* F is the ring homomorphism

$$\begin{aligned} R &\xrightarrow{F} R \\ r &\longmapsto r^p. \end{aligned}$$

The image is the subring R^p of pth powers of R, which is of course isomorphic to R via F (provided R has no nontrivial nilpotents, so that F is injective).

Nothing like this is true in characteristic zero. The point is that in characteristic p, the Frobenius map respects addition $[(r+s)^p = r^p + s^p$ for all $r, s \in R]$, because the binomial coefficients $\binom{p}{j}$ are congruent to 0 modulo p for every $1 \leq j \leq p-1$.

By iterating the Frobenius map we get an infinite chain of subrings of R:

$$R \supset R^p \supset R^{p^2} \supset R^{p^3} \supset \cdots, \tag{8}$$

each isomorphic to R. Alternatively, we can imagine adjoining pth roots: inside a fixed algebraic closure of the fraction field of R, for example, each element of R has a *unique* pth root. Now the ring inclusion $R^p \subset R$ is equivalent to the ring inclusion $R \subset R^{\frac{1}{p}}$; the Frobenius map gives an isomorphism between these two chains of rings. Iterating we have an increasing, but essentially equivalent chain of rings:

$$R \subset R^{\frac{1}{p}} \subset R^{\frac{1}{p^2}} \subset \ldots.$$

Viewing these $R^{\frac{1}{p^e}}$ as R-modules, it turns out that a remarkable wealth of information about singularities is revealed by their R-module structure as $e \to \infty$. Let us consider an example.

Example 2. Let $R = \mathbb{F}_p[x]$. The subring of pth powers is the ring $\mathbb{F}_p[x^p]$ of polynomials in x^p, and similarly the overring of pth roots is $\mathbb{F}_p[x^{\frac{1}{p}}]$. Given any polynomial $g(x) \in \mathbb{F}_p[x]$, there is a unique way to write

$$g(x) = g_0(x^p) \cdot 1 + g_1(x^p) \cdot x + \cdots + g_{p-1}(x^p) \cdot x^{p-1},$$

where each $g_j(x^p) \in R^p$. In fancier language, $\mathbb{F}_p[x]$ is a free $\mathbb{F}_p[x^p]$ module on the basis $\{1, x, x^2, \ldots, x^{p-1}\}$. That is, there is an $\mathbb{F}_p[x^p]$-module homomorphism

$$\mathbb{F}_p[x] \cong R^p \oplus R^p \cdot x \oplus \cdots \oplus R^p \cdot x^{p-1}.$$

Iterating, we see that $\mathbb{F}_p[x]$ is free over $\mathbb{F}_p[x^{p^e}]$ on the basis $\{1, x, x^2, \ldots, x^{p^e-1}\}$. Equivalently, each $\mathbb{F}_p[x^{\frac{1}{p^e}}]$ is a free R-module on the basis $\{1, x^{\frac{1}{p^e}}, x^{\frac{2}{p^e}}, \ldots, x^{\frac{p^e-1}{p^e}}\}$.

Example 3. Similarly, if R is the polynomial ring $\mathbb{F}_p[x_1, \ldots, x_n]$, then R is a free module over $R^{p^e} = \mathbb{F}_p[x_1^{p^e}, \ldots, x_n^{p^e}]$ with basis

$$\{x_1^{a_1} \cdots x_n^{a_n}\}_{0 \leq a_j \leq p^e-1,\ 1 \leq j \leq n}.$$

The freeness of the polynomial ring over its subring of pth powers is no accident, but rather reflects the fact that the corresponding affine variety is smooth. The Frobenius map can be used to detect singularities quite generally.

Theorem 4 (Kunz [36, Theorem 2.1]). *Let R be a ring of prime characteristic p without nilpotent elements. Then R is regular if and only if the Frobenius map is flat.*

Let us put Kunz's theorem more concretely in the case we care about—the case where the Frobenius map $R \xrightarrow{F} R$ is finite, that is, when R is *finitely generated* as an R^p-module.[9] In this case, Kunz's theorem says that R is regular if and only if R is a locally free R^p-module, or, equivalently, if and only if $R^{1/p}$ is locally free as an R-module.

This leads to the natural question: if R is *not* regular, can we use Frobenius to measure its singularities? The answer is a resounding YES. This is the topic of a large and active body of research in "F-singularities" which classifies singularities according to the structure of the chain of R-modules:

$$R \subset R^{\frac{1}{p}} \subset R^{\frac{1}{p^2}} \subset \ldots.$$

The F-threshold, which we now discuss, is only the beginning of a long and beautiful story.

3.2 F-Threshold

Now we fix a polynomial f in the ring $R = \mathbb{F}_p[x_1, \ldots, x_n]$. For a rational number of the form $c = \frac{a}{p^e}$, we can consider the fractional power $f^c = f^{\frac{a}{p^e}}$ as an element of the overring $R^{1/p^e} = \mathbb{F}_p[x_1^{1/p^e}, \ldots, x_n^{1/p^e}]$. This allows us to "take fractional

[9]All rings in which we are interested in this survey satisfy this condition. It is easy to check, for example, that if R is finitely generated over a perfect field, or a localization of such, then the Frobenius map is finite. Similarly, so do rings of power series over perfect fields.

powers" of polynomials, analogously to what the analysis allowed us to do in Sect. 1, at least if we restrict ourselves to fractional powers whose denominators are powers of p.

In the analytic setting, we tried to measure how badly the function $\frac{1}{|f|^\lambda}$ "blows up" at the singular point using integrability—this led to the complex singularity exponent or log canonical threshold. In this characteristic p world, we can not integrate, nor does even absolute value make sense. Amazingly, however, the most naive possible way to talk about the function $\frac{1}{f^c}$ "blowup" *does* lead to a sensible invariant, which turns out to be very closely related to the complex singularity index. Indeed, we can agree that $\frac{1}{f^c}$ certainly *does not* blow up at any point where the denominator does not vanish.

Recall that each R-module M can be interpreted as a coherent sheaf on the affine scheme Spec$R$, in which case each element $s \in M$ is interpreted as a section of this coherent sheaf. Grothendieck defined the "value" of a section s at the point $\mathfrak{P} \in \operatorname{Spec} R$ to be the image of s under the natural map from M to $M \otimes L$, where L is the residue field at $\mathfrak{P}$ [22, Chap 2.5]. In particular, the "function" f^c (when c is a rational number of the form $\frac{a}{p^e}$) is an element of the R-module R^{1/p^e} (for some e), and as such its "value" at the point $\mathfrak{m}$ is zero if and only if $f^c \in \mathfrak{m}R^{1/p^e}$.

So, given that integration does not make sense, we can at least look at values of c for which $\frac{1}{f^c}$ "does not blow up at all," and take the supremum over all such c. This extremely naive attempt to mimic the analytic definition then leads to the following definition.

Definition 5. The *F-threshold* of $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ at the maximal ideal $\mathfrak{m} = (x_1, \ldots, x_n)$ is defined as

$$FT_{\mathfrak{m}}(f) = \sup\Big\{c = \frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \ \Big|\ f^c \notin \mathfrak{m}R^{1/p^e}\Big\}.$$

Amazingly, this appear to be the "right" thing to do! Although we have stated the definition for polynomials over $\mathbb{F}_p$, any perfect field k, or indeed, any field k of characteristic p such that $[k : k^p]$ is finite works just as well.

Let us check that this definition is independent of how we write c. First note that viewing f^c as an element of the free R-module R^{1/p^e}, we can write it uniquely as an R-linear combination of the basis elements

$$\{x_1^{\frac{a_1}{p^e}} \cdots x_n^{\frac{a_n}{p^e}}\}_{0 \le a_j \le p^e - 1},$$

which we abbreviate $\{\mathbf{x}^{A/p^e}\}$.

$$f^c = f^{a/p^e} = \sum r_A \mathbf{x}^{A/p^e}, \tag{9}$$

for some uniquely determined $r_A \in R$. So, an equivalent formulation of the F-threshold of $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ at the maximal ideal $\mathfrak{m} = (x_1, \ldots, x_n)$ is

$$FT_{\mathfrak{m}}(f) = \sup\left\{c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid f^c \text{ has some coefficient } r_A \notin \mathfrak{m}\right\},$$

where the coefficients r_A are as in (9).

It is now easy to see that this supremum is independent of the way we write c. That is, if we instead had written $c = \frac{ap}{p^{e+1}}$ and viewed f^c as an element in the larger ring $R^{\frac{1}{p^{e+1}}}$, then when expressed f^c uniquely as an R-linear combination of the basis elements $\mathbf{x}^{A'/p^{e+1}}$ for $R^{\frac{1}{p^{e+1}}}$, the coefficients $r_{A'}$ that appear are the same elements of R as appearing in expression (9).

By Nakayama's Lemma, $f^c \notin \mathfrak{m}R^{1/p^e}$ if and only if f^c is part of a minimal generating set for the R-module R^{1/p^e} (after localizing at $\mathfrak{m}$). So equivalently:

Definition 6. The *F-threshold* of $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ at $\mathfrak{m}$ is

$$FT_{\mathfrak{m}}(f) = \sup\left\{c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid f^c \text{ is part of a free basis for } R_{\mathfrak{m}}^{\frac{1}{p^e}} \text{ over } R_{\mathfrak{m}}\right\}.$$

Example 7. The F-threshold of any polynomial is always bounded above by one. Indeed, let f be any polynomial in $\mathfrak{m}$. Since $f^1 = f \cdot 1 \in \mathfrak{m}R^{\frac{1}{p^e}}$ for all e, we see that f^1 is never part of a basis for R^{1/p^e} over R. We must raise f to numbers *less than one* to get a basis element. Thus $FT_{\mathfrak{m}}(f) \leq 1$ always.

Example 8. Assume that f is non-singular at $\mathfrak{m}$. Then f is part of a local system of regular parameters at $\mathfrak{m}$, that is, one of the minimal generators for the ideal $\mathfrak{m}$. Changing coordinates, we can assume $f = x_1$. For any p^e, note that

$$f^{\frac{p^e-1}{p^e}} = x_1^{\frac{p^e-1}{p^e}}$$

is part of a free basis for $R_{\mathfrak{m}}^{\frac{1}{p^e}}$ over $R_{\mathfrak{m}}$. This shows that the F-threshold is greater or equal than $\frac{p^e-1}{p^e}$ for all $e \geq 1$. In other words, taking the limit we see that the F-threshold at a smooth point is bounded below by one. Combining with the previous example, we conclude that the F-threshold of a smooth point is exactly one.

Example 9. Take $f = xy \in \mathbb{F}_p[x, y]$, then $(xy)^{\frac{p^e-1}{p^e}}$ is also part of a basis of the free R-module R^{1/p^e}. The same argument applies here to show that $FT_{\mathfrak{m}}(xy) = 1$. In general, $FT_{\mathfrak{m}}(x_1 \cdots x_\ell) = 1$, which is to say, the F-threshold of a simple normal crossings divisor is always one. In particular, this shows that $FT_{\mathfrak{m}}(f) = 1$ does not imply that f is non-singular.

Example 10. Consider $f = x^m \in \mathbb{F}_p[x]$. Then, $(x^m)^{\frac{a}{p^e}}$ is part of a free basis if and only if $\frac{ma}{p^e} \leq \frac{p^e-1}{p^e}$. So $f^{\frac{a}{p^e}}$ is part of a free basis if and only if $\frac{a}{p^e} \leq \frac{1}{m} - \frac{1}{mp^e}$. Taking the limit, this leads to $FT_{\mathfrak{m}}(x^m) = \frac{1}{m}$. In general,

$$FT_{\mathfrak{m}}(x_1^{a_1} \cdots x_\ell^{a_\ell}) = \min\left\{\frac{1}{a_i}\right\}.$$

Examples 7 through 10 indicate that the F-threshold has many of the same features as the log canonical threshold. Is the F-threshold capturing exactly "the same" measurement of singularities as the log canonical threshold? If a polynomial has integer coefficients, do we get the same value of the F-threshold modulo p for all p? Of course, the "same" polynomial can be more singular in some characteristics, so we expect not. But does the F-threshold for "large p" perhaps agree with the log canonical threshold? The following example is typical:

Example 11. Consider the polynomial $f = x^2 + y^3$, which we can view as a polynomial over any of the fields $\mathbb{F}_p$ (or $\mathbb{C}$). Its F-threshold depends on the characteristic:

$$FT_{\mathfrak{m}}(f) = \begin{cases} 1/2 & \text{if } p = 2, \\ 2/3 & \text{if } p = 3, \\ 5/6 & \text{if } p \equiv 1 \mod 6, \\ 5/6 - \frac{1}{6p} & \text{if } p \equiv 5 \mod 6. \end{cases}$$

Viewing f as a polynomial over $\mathbb{C}$, we computed in Example 15 that its log canonical threshold is $\frac{5}{6}$. Interestingly, we see that as $p \longrightarrow \infty$, the F-thresholds approach the log canonical threshold. Also, there are some (in fact, infinitely many) characteristics where the F-threshold agrees with the log canonical threshold. On the other hand, there are other characteristics where the polynomial is "more singular" than expected, as reflected by a smaller F-threshold. For example, in characteristics 2 and 3, of course, we expect "worse" singularities, and indeed, we see the F-threshold is smaller in these cases. But also the F-threshold detects a worse singularity for this curve in characteristics congruent to 5 mod 6, reflecting subtle number theoretic issues in that case (see [43, Question 3.9]).

Example 12. Daniel Hernández has computed many examples of F-thresholds in his Ph.D. thesis (cf. [23]), including any "diagonal" hypersurfaces $x_1^{a_1} + \cdots + x_n^{a_n}$ (see [25]). There is also an algorithm to compute the F-threshold of any binomial as well; see [26, 51]. See also [43].

Example 13. Let $f \in \mathbb{Z}[x, y, z]$ be homogeneous of degree 3 with an isolated singularity. In particular, f defines an elliptic curve in $\mathbb{P}^2$ over $\mathbb{Z}$. The F-threshold has been computed in this case by Bhatt [1]:

$$FT(f) = \begin{cases} 1 & \text{if } E \in \mathbb{P}^2_{\mathbb{F}_p} \text{ is } \textit{ordinary}, \\ 1 - \dfrac{1}{p} & \text{if } E \in \mathbb{P}^2_{\mathbb{F}_p} \text{ is } \textit{supersingular}. \end{cases}$$

Again we see that "more singular" polynomials have smaller F-thresholds. As in Example 11, there are infinitely many p for which the log canonical threshold and the F-threshold agree, and infinitely many p for which they do not, by a result of Elkies [13].

3.3 Comparison of F-Threshold and Multiplicity

To compare the F-threshold with the multiplicity, we rephrase the definition still one more time. Let us first recall a well-known notation: for an ideal I in a ring R of characteristic p, let $I^{[p^e]}$ denote the ideal of R generated by the p^e-th powers of the elements of I. That is, $I^{[p^e]}$ is the expansion of I under Frobenius $R \to R$ sending $r \mapsto r^{p^e}$.

Definition 14. The F-threshold of $f \in k[x_1, \ldots, x_n]_{\mathfrak{m}} = R$ is

$$FT_{\mathfrak{m}}(f) = \sup\left\{c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid f^a \notin \mathfrak{m}^{[p^e]}\right\},$$
$$= \inf\left\{c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid f^a \in \mathfrak{m}^{[p^e]}\right\}.$$

This is patently the same as Definition 5, simply by raising to the p^eth power.

On the other hand, the multiplicity of f at $\mathfrak{m}$ is defined as the largest n such that $f \in \mathfrak{m}^n$. It is trivial to check that this is equivalent to

$$\text{mult}_{\mathfrak{m}}(f) = sup\left\{\frac{t}{a} \in \mathbb{Q} \mid f^a \in \mathfrak{m}^t\right\}.$$

That is, the formula that computes the F-threshold is similar to formula that computes the *reciprocal* of the multiplicity, but with "Frobenius powers" replacing ordinary powers:

$$\frac{1}{\text{mult}_{\mathfrak{m}}(f)} = inf\left\{\frac{a}{t} \in \mathbb{Q} \mid f^a \in \mathfrak{m}^t\right\}.$$

It is also not hard to check in all these cases that the infimum (supremum) is in fact a limit.

This similarity allows us to easily prove the following comparison between multiplicity and F-threshold:

Proposition 15. *For $f \in \mathfrak{m} \subset k[x_, \ldots, x_N]$,*

$$\frac{N}{mult_{\mathfrak{m}}(f)} \geq FT_{\mathfrak{m}}(f) \geq \frac{1}{mult_{\mathfrak{m}}(f)}.$$

Proof. Since $\mathfrak{m}$ is generated by N elements, we have the inclusions

$$\mathfrak{m}^{Np^e} \subset \mathfrak{m}^{[p^e]} \subset \mathfrak{m}^{p^e}$$

for all e. So we also obviously have inclusions of sets

$$\left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{Np^e}\right\} \subset \left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{[p^e]}\right\}$$
$$\subset \left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{p^e}\right\}.$$

Taking the infimum, we have

$$\inf\left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{Np^e}\right\} \geq \inf\left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{[p^e]}\right\}$$
$$\geq \inf\left\{\frac{a}{p^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{p^e}\right\}.$$

Since the infimum on the left can be interpreted as N times $\inf\left\{\frac{a}{Np^e} \in \mathbb{Z}\Big[\frac{1}{p}\Big] \mid f^a \in \mathfrak{m}^{Np^e}\right\}$, the result is proved. □

Remark 16. A deeper inequality is proven in [60, Prop 4.5]: $mult(f) \geq \frac{N^N}{(FT(f))^N}$. This in turn, is a "characteristic p analog" of a corresponding statement in characteristic zero about log canonical threshold [10, Theorem 1].

3.4 Computing F-Thresholds

To get a feeling how to compute F-thresholds, we begin the computation of Example 11, relegating the details to [23]. First recall that $\mathbb{F}_p[x, y]$ is a free module over $\mathbb{F}_p[x^{p^e}, y^{p^e}]$ with basis $\{x^{a_1}y^{a_2}\}_{0\leq a_1,a_2\leq p^e-1}$. By definition,

$$FT(f) = \sup\left\{\frac{a}{p^e} \mid f^a = \sum r_A^{p^e} x^{a_1}y^{a_2} \text{has some coefficient } r_A^{p^e} \notin (x^{p^e}, y^{p^e})\right\}.$$

For each a, we expand using the binomial theorem to get

$$(x^2 + y^3)^a = \sum_{i=0}^{a} \binom{a}{i} x^{2i} y^{3(a-i)}.$$

Note that none of the terms in this expression can cancel, since each has a unique bi-degree in x and y, unless its corresponding binomial coefficient is zero. So, thinking over $\mathbb{F}[x^{p^e}, y^{p^e}]$, we see that $FT(f) \geq \frac{a}{p^e}$ if and only if there is an index i such that

$$2i < p^e, \quad 3(a-i) < p^e, \quad \text{and} \quad \binom{a}{i} \not\equiv 0 \pmod{p}. \qquad (*)$$

Note that if $2a < p^e$, then the index $i = a$ fulfills the three conditions in $(*)$. This, in particular, implies that $FT(f) \geq \frac{1}{2}$, independent of p. If the characteristic is 2, it is easy to see immediately that $FT(f) = \frac{1}{2}$.

On the other hand, if $\frac{a}{p^e} \geq \frac{5}{6}$, then condition $(*)$ is never satisfied so that $FT(f) \leq \frac{5}{6}$, independent of p. Indeed, in this case, either $2i \geq p^e$ or $3(a-i) \geq p^e$. For otherwise, we have both

$$i \leq \frac{p^e - 1}{2} \qquad \text{and} \qquad (a - i) \leq \frac{p^e - 1}{3},$$

in which case, adding them, we have that $a \leq \frac{p^e-1}{2} + \frac{p^e-1}{3}$. This implies that $\frac{a}{p^e} \leq \frac{5}{6} - \frac{5}{6p^e}$, a contradiction.

For the exact computation of the F-threshold, it remains to analyze the binomial coefficients in the critical terms in which both $2i < p^e$ and $3(a - i) < p^e$. These are the terms indexed by i satisfying $a - \frac{p^e}{3} < i < \frac{p^e}{2}$. For this, it is crucial to understand the behavior of binomial coefficients modulo p. One of the main tools is the following theorem:

Theorem 17 (Lucas [38]). *Fix nonnegative integers $m \geq n \in \mathbb{N}$ and a prime number p. Write m and n in their base p expansions: $m = \sum_{j=0}^{r} m_j p^j$ and $n = \sum_{j=0}^{r} n_j p^j$. Then, modulo p,*

$$\binom{m}{n} \equiv \binom{m_0}{n_0}\binom{m_1}{n_1}\cdots\binom{m_r}{n_r},$$

where we interpret $\binom{a}{b}$ as zero if $a < b$. In particular, $\binom{m}{n}$ is nonzero mod p if and only if $m_j \geq n_j$ for all $j = 1, \ldots, r$.

Thus, to compute the F-threshold of $x^2 + y^3$, it is helpful to write a in its base p expansion, and then try to understand, using Lucas's theorem, whether there exist values of i in the critical range for which $\binom{a}{i}$ is not zero. For example, if $p \equiv 1 \mod 6$, then $5p^e = 5 \mod 6$, so the number $a = \frac{5p^e-5}{6}$ is an integer. With this choice of a, it is not hard to check that the conditions $(*)$ are satisfied for the index $i = \frac{p^e-1}{2}$. This shows that when $p \equiv 1 \mod 6$, then the F-threshold of $x^2 + y^3$ is at least $\frac{a}{p^e} = \frac{5}{6} - \frac{5}{p^e}$ for all e. It follows that for $p \equiv 1 \mod 6$, the F-threshold of $x^2 + y^3$ is exactly $\frac{5}{6}$. The details, as well as the computation for other p, are carried out in [43, Example 4.3] or [23, Example 8.2].

3.5 Comparison of F-Thresholds and Log Canonical Thresholds

Once F-thresholds and log canonical thresholds are defined, it is natural to compare them when it is possible. This is the case when $f \in \mathbb{Z}[x_1, \ldots, x_n]$. We can view f as a polynomial over $\mathbb{C}$ and compute its log canonical threshold. After reduction modulo p, we can calculate the F-threshold of $f \mod p$ in $\mathbb{F}_p[x_1, \ldots, x_n]$.

Question. For which values of p is $lct(f) = FT(f \mod p)$? What happens when $p \gg 0$?

The following theorem provides a partial answer:

Theorem 18. *Fix $f \in \mathbb{Z}[x_1, \ldots, x_n]$. Then,*

- $FT(f \mod p) \leq lct(f)$ *for all $p \gg 0$ prime.*
- $\lim_{p\to\infty} FT(f \mod p) = lct(f)$.

The proof of this theorem is the culmination mainly of the work of the Japanese school of tight closure, who generalized the theory of tight closure to the case of pairs. The first important step was the work of Hara and Watanabe in [20] Theorem 3.3, with Theorem 18 essentially following from Theorem 6.8 in the paper [21] of Hara and Yoshida; the proof there in turn generalizes the proofs given in [19] and [55] in the nonrelative case to pairs.

Open Problem. Are there infinitely many primes p for which $FT(f \mod p) = lct(f)$?

A positive answer to this question would settle a long-standing conjecture in F-singularities: every log canonical pair (X, D) where X is smooth (over $\mathbb{C}$) is of "F-pure type." Daniel Hernández shows that this is the case for a "very general" polynomial f in $\mathbb{C}[x_1, \ldots, x_n]$ [24]. Versions of this question have been around since the early eighties, for example, as early as Rich Fedder's thesis in 1983 [16], when similarities between Hochster and Robert's notion of "F-purity" and rational singularities began to emerge. Watanabe pointed out that the log canonicity ought to correspond to F-purity, circulating a preprint in the late eighties proving that "F-pure implies log canonical" for rings. He did not publish this result for several years, when together with Hara, they introduced a notion of F-purity for pairs and proved the corresponding result for a pair (X, D). The field of "F-singularities of pairs," including the F-pure threshold, developed rapidly in Japan, with numerous papers of Watanabe, Hara, Takagi, Yoshida, and others filling in the theory. The question in the form stated here may have first appeared in print in [43].

3.6 Test Ideals and F-Thresholds

We wish to construct a family of ideals of a polynomial $f \in \mathbb{F}_p[x_1, \ldots, x_n] = R$, say, $\tau(f^c)$, called test ideals, which are analogous to the multiplier ideals.

We first restate the definition of F-threshold yet one more time, in a way that will make the definition of the test ideals very natural.

Definition 19. The F-threshold of $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ at $\mathfrak{m}$ is

$$FT_{\mathfrak{m}}(f) = \sup\left\{c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid \exists \varphi \in \mathrm{Hom}_R(R^{1/p^e}, R),\ \varphi(f^c) \notin \mathfrak{m}\right\}.$$

To see that this is equivalent to the previous definitions, we apply the following simple lemma in the case where $J = \mathfrak{m}$:

Lemma 20. *Consider $f \in R = \mathbb{F}_p[x_1, \ldots, x_n]$ and $c = \frac{a}{p^e}$. Fix any ideal $J \subset R$. Then $f^c \notin JR^{\frac{1}{p^e}}$ if and only if there is an R-linear map $R^{\frac{1}{p^e}} \xrightarrow{\phi} R$ sending f^c to an element not in J.*

Proof. Suppose that $f^c \notin JR^{\frac{1}{p^e}}$. Then in writing f^c uniquely in some basis for $R^{\frac{1}{p^e}}$ as in expression (9), there is some coefficient $r_A \notin J$. If we let ϕ be the projection onto this direct summand, we have a map ϕ satisfying the required conditions. Conversely, if $f^c \in JR^{\frac{1}{p^e}}$, then the R-linearity of ϕ forces $\phi(f^c) \in J$ for any $\phi \in \mathrm{Hom}_R(R^{1/p^e}, R)$. So every such ϕ must send f^c to an element in J. □

With this definition of F-threshold in mind, it is quite natural to define test ideals, at least for certain c. First note that for each $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ and each $c \in \mathbb{Z}[\frac{1}{p}]$, we have a natural R-module map:

$$\mathrm{Hom}(R^{1/p^e}, R) \longrightarrow R$$
$$\phi \longmapsto \phi(f^c),$$

where e is chosen so that $c = \frac{a}{p^e}$ for some natural number a. The test ideal is the image of this map.

Definition 21. Let $f \in \mathbb{F}_p[x_1, \ldots, x_n]$ and $c = \frac{a}{p^e} \in \mathbb{Z}[\frac{1}{p}]$. The *test ideal* $\tau(f^c)$ is the ideal

$$\tau(f^c) = \mathrm{im}[\mathrm{Hom}(R^{1/p^e}, R) \longrightarrow R, \text{ defined by evaluation at } f^c].$$

In practical terms, if we write $f^c = f^{a/p^e}$ uniquely as in Expression (9), then $\tau(f^c)$ is generated by the coefficients r_A which appear in this expression. Note that this is independent of the way we write c. Indeed, if we instead think of c as $\frac{ap}{p^{e+1}}$, then the expression for f^c becomes

$$f^c = \sum_A r_A x^{A/p^e} = \sum r_A x^{pA/p^{e+1}}$$

which is a valid expression for f^c as an element of the free R-module $R^{1/p^{e+1}}$ since the monomials $x^{pA/p^{e+1}}$ are still part of free basis for $R^{1/p^{e+1}}$.

Remark 22. The test ideal $\tau(f^{a/p^e})$ is the *smallest* ideal J such that $f^{a/p^e} \in JR^{1/p^e}$. Indeed, Lemma 20 can be reinterpreted as saying that if there is some ideal J of R such that $f^{a/p^e} \in JR^{1/p^e}$, then $\tau(f^{a/p^e}) \subset J$.

Although we have not yet defined test ideals with arbitrary real exponents c, let us pause and see whether, at least for $c \in \mathbb{Z}[1/p]$, we have a collection of ideals $\{\tau(f^c)\}$ satisfying the desired basic properties analogous to Proposition 23. That is, we want:

1. $\tau(f^c)$ is the unit ideal for sufficiently small positive c.
2. If $c > c'$, then $\tau(f^{c'}) \supseteq \tau(f^c)$;
3. The F-threshold of f is

$$FT_{\mathfrak{m}}(f) = \sup\{c \mid \tau(f^c) = (1)\}.$$

4. If $\varepsilon > 0$ small enough, then $\tau(f^c) = \tau(f^{c+\varepsilon})$.
5. There exist certain c such that $\tau(f^{c-\varepsilon}) \supsetneq \tau(f^c)$ for all positive ε.

The first property is easy. Indeed, for fixed a, consider $f^{a/p^e} \in R^{1/p^e}$ as e gets very large. If $f^{a/p^e} \in \mathfrak{m}R^{1/p^e}$, then $f^a \in m^{[p^e]} \subset m^{p^e}$ in R. But this cannot be the case for all e. The third property follows immediately from the second, which is also quite straightforward. Note that since the test ideal $\tau(f^c)$ is independent of the way we represent c as a quotient of two integers whose denominator is a power of p, we may assume that $c = \frac{a}{p^e}$ and $c' = \frac{b}{p^e}$ have a common denominator. Then:

Lemma 23. *For* $\frac{a}{p^e} \geq \frac{b}{p^e}$*, we have*

$$\tau(f^{a/p^e}) \subseteq \tau(f^{b/p^e}).$$

Proof. Suppose that $s \in \tau(f^{a/p^e})$. This means there is an R-linear map $\psi : R^{1/p^e} \longrightarrow R$ so that $\psi(f^{a/p^e}) = s$. Precomposing this with the R-linear map $R^{1/p^e} \xrightarrow{\mu} R^{1/p^e}$ given by multiplication by $f^{(a-b)/p^e}$, we have an R-linear map

$$\begin{array}{ccccc} R^{1/p^e} & \xrightarrow{\quad\mu\quad} & R^{1/p^e} & \xrightarrow{\quad\psi\quad} & R \\ f^{b/p^e} & \longmapsto & f^{(a-b)/p^e} f^{b/p^e} & \longmapsto & \varphi(f^{a/p^e}) = s, \end{array}$$

showing that $s \in \tau(f^{b/p^e})$ as well. □

The fourth property is also quite simple to prove.

Lemma 24. *Fix* $c = \frac{a}{p^e}$*. Then, for all* n *sufficiently large (how large could depend on* c*),*

$$\tau(f^c) = \tau(f^{c+\frac{1}{p^n}}).$$

Proof. Fix $c = \frac{a}{p^e}$ and let $s \in \tau(f^c)$. This means that there exists an R-linear map

$$R^{1/p^e} \xrightarrow{\phi} R \qquad s.t. \quad f^{a/p^e} \mapsto s.$$

Take n sufficiently large so that $f^{1/p^{n-e}}$ is part of a free basis for $R^{1/p^{n-e}}$. Projection onto the submodule spanned by $f^{1/p^{n-e}}$ is an R-linear map $R^{1/p^{n-e}} \to R$ sending $f^{1/p^{n-e}}$ to 1. Taking the p^eth roots, we have an R^{1/p^e}-linear map

$$R^{1/p^n} \xrightarrow{\psi} R^{1/p^e} \quad s.t. \quad f^{1/p^n} \mapsto 1.$$

In particular $\psi(f^{\frac{a}{p^e}+\frac{1}{p^n}}) = f^{\frac{a}{p^e}}$. Composing, we get an R-linear map

$$\begin{array}{ccccc} R^{1/p^n} & \xrightarrow{\psi} & R^{1/p^e} & \xrightarrow{\phi} & R \\ f^{a/p^e+1/p^n} & \longmapsto & f^{a/p^e} & \longmapsto & \varphi(f^{a/p^e}) = s. \end{array}$$

This shows that $s \in \tau(f^{c+\frac{1}{p^n}})$, as desired. ☐

The fifth property, however, is *not true* if we restrict attention to c that are rational numbers whose denominator is a power of p. To get property (5), we need to define test ideals also for arbitrary positive real numbers c; if we can do so in such a way that the first four properties are satisfied for all real numbers, then the completeness property of the real numbers will automatically grant (5).

Lemma 23 encourages us to define $\tau(f^c)$ for any positive real c by approximating c by a sequence of numbers $\{c_n\}_{n\in\mathbb{N}} \in \mathbb{Z}[\frac{1}{p}]$ converging to c *from above* and taking advantage of the Noetherian property of the ring. That is, we take any monotone decreasing sequence of numbers in $\mathbb{Z}[\frac{1}{p}]$

$$c_1 > c_2 > c_3 \dots$$

converging to c. There is a corresponding increasing sequence of test ideals:

$$\tau(f^{c_1}) \subseteq \tau(f^{c_2}) \subseteq \tau(f^{c_3}) \cdots .$$

Because R is Noetherian, this chain of ideals must stabilize. Since any other strictly decreasing sequence converging to c is cofinal with this one (meaning that if $\{c_i'\}$ is some other sequence, then for all i, there exists j such that $c_i > c_j'$ and vice versa), it is easy to check that the stable ideal is independent of the choice of approximating sequence. So we have the following definition:

Definition 25. For any $c \in \mathbb{R}_+$ the *test ideal* is defined as

$$\tau(f^c) := \bigcup_{n \geq 0} \tau(f^{c_n}),$$

where $\{c_n\}_{n \in \mathbb{N}}$ is any decreasing sequence of rational numbers in $\mathbb{Z}[1/p]$ approaching c. In particular,

$$\tau(f^c) := \bigcup_{n \geq 0} \tau(f^{\lceil cp^n \rceil / p^n}),$$

or equivalently as $\tau(f^{\lceil cp^n \rceil / p^n})$ for $n \gg 0$.

There is one slight ambiguity to address: If the real number c *happens* to be a rational number whose denominator is a power of p, then we have already defined $\tau(f^c)$ in Definition 21. Do the two definitions produce the same ideal in this case? That is, we need to check that, if we had instead approximated c by a sequence $\{c_n\}_{n \in \mathbb{N}} \in \mathbb{Z}[1/p]$ converging to a/p^e from above, then the ideals $\tau(f^{c_n})$ stabilize to $\tau(f^{a/p^e})$. But this is essentially the content of Lemma 24. So test ideals are well-defined for any positive real number.

As before with multiplier ideals, the following follows easily from the definition:

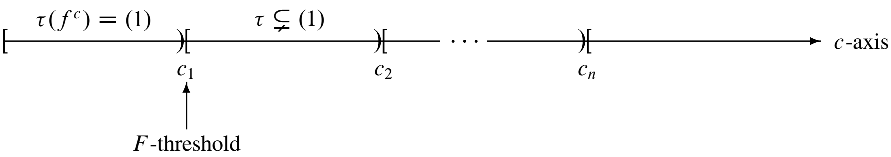
Proposition 26. *Fix a polynomial f and view its test ideals $\tau(f^c)$ as a family of ideals varying with a positive real parameter c. Then the following properties hold:*

1. *For $c \in \mathbb{R}_+$ sufficiently small, $\tau(f^c)$ is the unit ideal.*
2. *If $c > c'$, then $\tau(f^c) \subseteq \tau(f^{c'})$.*
3. *The F-threshold of f is*

$$FT(f) = \sup\{c \mid \tau(f^c) = (1)\}.$$

4. *For each fixed c, we have $\tau(f^c) = \tau(f^{c+\varepsilon})$ for sufficiently small positive ε (how small is small enough depends on c).*
5. *There exist certain $c \in \mathbb{R}_+$ such that $\tau(f^{c-\varepsilon}) \supsetneq \tau(f^c)$ for all positive ε.*

The proposition is summarized by the following diagram of the c-axis, which shows intervals where the test ideal remains constant:



This leads naturally to the F-jumping numbers.

Definition 27. The *F-jumping numbers* of f are the real numbers $c_i \in \mathbb{R}_+$ for which $\tau(f^{c_i}) \neq \tau(f^{c_i - \varepsilon})$, for every $\varepsilon > 0$.

It is not hard to see that test ideals and F-jumping numbers enjoy many of the same properties as do the multiplier ideals and jumping numbers defined in characteristic zero. First we have an analog of the Briançon–Skoda theorem.[10]

Proposition 28 ([4, Proposition 2.25]). *Let f be a polynomial in $\mathbb{F}_p[x_1, \dots, x_n]$. Then for every $c \in \mathbb{R}_+$, we have*

$$\tau(f^{c+1}) = (f) \cdot \tau(f^c).$$

In particular, a positive real number c is an F-jumping number if and only if $c + 1$ is an F-jumping number.

Proof. Without loss of generality, we may replace both c and $c + 1$ by rational numbers in $\mathbb{Z}[\frac{1}{p}]$ approximating each from above. Thus we may assume $c = a/p^e$ for some $a, e \in \mathbb{N}$.

For any R-linear map $R^{1/p^e} \xrightarrow{\varphi} R$, it is clear that

$$\varphi(f^{(a/p^e)+1}) = f\varphi(f^{(a/p^e)}),$$

since $f \in R$. It immediately follows that $\tau(f^{c+1}) = (f) \cdot \tau(f^c)$. □

Like the jumping numbers in characteristic zero, the F-jumping numbers are discrete and rational. Interestingly, the proof of the characteristic zero statement follows trivially from the (algebro-geometric) definition of multiplier ideals, while the characteristic p proof took some time to find.

Theorem 29 ([4, Theorem 3.1]). *The F-jumping numbers of a polynomial are discrete and rational.*

The proof takes advantage of an additional symmetry the F-jumping numbers enjoy for which there is no analog in characteristic zero.

Lemma 30 ([4, Proposition 3.4]). *Let f be a polynomial in $\mathbb{F}_p[x_1, \dots, x_n]$. If c is a jumping number for f, then also pc is a jumping number for f.*

Proof. Let c be a jumping number, that is, suppose $\tau(f^c) \subsetneq \tau(f^{c-\varepsilon})$ for all $\epsilon > 0$. For any $a/p^e < c \leq b/p^e$ where a, b are positive integers, we thus have

$$\tau(f^{\frac{b}{p^e}}) \subset \tau(f^c) \subsetneq \tau(f^{\frac{a}{p^e}}),$$

and the first inclusion is an equality when $\frac{b}{p^e}$ is sufficiently close to c. For such close $\frac{b}{p^e}$, write

$$f^{b/p^e} = \sum_B r_B \mathbf{x}^{B/p^e}$$

[10]Starting in [37], the name of this theorem, which belongs to a collection of inter-related theorems comparing powers of an ideal to its integral closure, has been sometimes shortened to "Skoda's theorem." We follow here the tradition in commutative algebra to include Briançon's name.

so that $\tau(f^{b/p^e})$ is generated by the coefficients r_B. Raising to the power p gives that

$$f^{b/p^{e-1}} = \sum_B r_B^p \mathbf{x}^{B/p^{e-1}},$$

which means, by Lemma 20, that every R-linear map $R^{1/p^{e-1}} \xrightarrow{\phi} R$ sends $f^{b/p^{e-1}}$ to something in $\langle r_B^p \rangle$. In other words,

$$\tau(f^{b/p^{e-1}}) \subseteq \langle r_B^p \rangle. \tag{10}$$

In contrast, since $\tau(f^{b/p^e}) \subsetneq \tau(f^{a/p^e})$, we know that $f^{a/p^e} \notin \langle r_B \rangle R^{1/p^e}$. Raising to pth powers again, we have that $f^{a/p^{e-1}} \notin \langle r_B^p \rangle R^{1/p^{e-1}}$. But then (10) forces $f^{a/p^{e-1}} \notin \tau(f^{\frac{b}{p^{e-1}}}) R^{1/p^{e-1}}$. By Lemma 20, it follows that there is an R-module homomorphism $R^{1/p^e} \xrightarrow{\phi} R$ such that $\phi(f^{a/p^{e-1}})$ is not in $\tau(f^{\frac{b}{p^{e-1}}})$. But this exactly means that there is an element of $\tau(f^{a/p^{e-1}})$ that is not in $\tau(f^{\frac{b}{p^{e-1}}})$. Letting $a/p^e, b/p^e$ go to c, we get that pc is a jumping number as well as c. □

Proof of Theorem 29. To prove discreteness, we fix an f of degree d, and $c = \frac{a}{p^e}$. We claim that $\tau(f^c)$ is generated by elements of degree smaller or equal to $\lfloor cd \rfloor$. Indeed, $\tau(f^c)$ is generated by the coefficients r_A appearing in $f^c = \sum r_A \mathbf{x}^{a/p^e} \in R^{1/p^e}$. Since f^c has degree dc, then $r_A \in R$ has degree $\le \lfloor cd \rfloor$. This proves the claim.

Now assume that the F-jumping numbers of f, say, $\alpha_1 < \alpha_2 < \cdots$, were clustering to some α. Without loss of generality, each of the α_i can be assumed in $\mathbb{Z}[\frac{1}{p}]$. By definition of F-jumping number,

$$\tau(f^{\alpha_1}) \supsetneq \tau(f^{\alpha_2}) \supsetneq \cdots. \tag{11}$$

The previous claim ensures that each $\tau(f^{\alpha_i})$ is generated in degree $\le \lfloor d\alpha_i \rfloor \le \lfloor d\alpha \rfloor = D$. Now intersect each of these test ideals with the finite-dimensional vector space $V \subseteq \mathbb{F}_p[x_1, \ldots, x_n]$ consisting of polynomials of degree $\le D$. The sequence

$$\tau(f^{\alpha_1}) \cap V \supsetneq \tau(f^{\alpha_2}) \cap V \supsetneq \cdots$$

stabilizes, since V is finite dimensional. Hence (11) also stabilizes, which is a contradiction to clustering of the F-jumping numbers.

To prove the rationality of F-jumping numbers, let $c \in \mathbb{R}$ be an F-jumping number. Then for all $e \in \mathbb{N}$, the real numbers $p^e c$ are also F-jumping numbers. One can write $p^e c = \lfloor p^e c \rfloor + \{p^e c\}$, where the fractional part $\{p^e c\}$ is also an F-jumping number by Lemma 30. By the discreteness of the F-jumping numbers it follows that $\{p^e c\} = \{p^{e'} c\}$ for some e and e' in $\mathbb{N}$, and hence $p^e c - p^{e'} c = m \in \mathbb{Z}$. Thus $c = \frac{m}{p^e - p^{e'}}$ is rational. □

Remark 31. Test ideals and multiplier ideals can be defined not just for one polynomial but for any ideal $\mathfrak{a}$ in any polynomial ring, and even for sheaves of ideals on (certain) singular ambient schemes. While not much more complicated that what we have introduced here, we refer the interested reader to the literature for this generalization. Many properties of multiplier ideals (see [37, Chap. 9]) can be directly proven, or adapted, to test ideals. For example, the fact that the F-jumping numbers are discrete and rational holds more generally—essentially for all ideals in normal $\mathbb{Q}$-Gorenstein ambient schemes [6]. In addition to the few properties discussed here, other properties of multiplier ideals that have analogs for test ideals include the "restriction theorem," the "subadditivity theorem," the "summation theorem" [21, 59], and the behavior of test ideals under finite morphisms [49]. Interestingly, some of the more difficult properties to prove in characteristic zero turn out to be extraordinarily simple in characteristic p. For example, in characteristic zero, the proof of the Briançon–Skoda theorem (for ideals that are not necessarily principal) uses the "local vanishing theorem" (see [37]). Although this vanishing theorem fails in characteristic p, the characteristic p analog of the Briançon–Skoda theorem (i.e., Theorem 28 for non-principal ideals) is nonetheless true and in fact quite simple to prove immediately from the definition. On the other hand, some properties of multiplier ideals that follow immediately from the definition in terms of resolution of singularities turn out to be *false* for test ideals. For example, while multiplier ideals are easily seen to be integrally closed, test ideals are not. In fact, *every* ideal in a polynomial ring is the test ideal $\tau(\mathfrak{a}^{\lambda})$ for some ideal $\mathfrak{a}$ and some positive $\lambda \in \mathbb{R}$, as shown in [42].

3.7 An Interpretation of F-Thresholds and Test Ideals Using Differential Operators

Our definition of F-threshold can be viewed as a measure of singularities using differential operators. The point is that a differential operator on a ring R of characteristic p is precisely the same as a R^{p^e}-linear map.

Differential operators can be defined quite generally. Let A be any base ring and R a commutative A-algebra. Grothendieck defined the ring of A-linear differential operators of R using a purely algebraic approach (see [18]), which in the case where $A = k$ is a field and R is a polynomial ring over k results in the "usual" differential operators.

Definition 32. Let R be a commutative A-algebra. The ring $D_A(R)$ of A-linear differential operators is the subring of the (noncommutative) ring $\mathrm{End}_A(R)$ obtained as the union of the A-submodules of differential operators $D^n_A(R)$ of order less than or equal to n, where $D^n_A(R)$ is defined inductively as follows: First the zeroth order operators $D^0_A(R)$ are the elements $r \in R$ interpreted as the A-linear endomorphisms "multiplication by r"; that is, $r : R \longrightarrow R$ sends $x \mapsto rx$. Then, for $n > 0$,

$$D^n_A(R) := \{\partial \in \mathrm{End}_A(R) \mid [r, \partial] \in D^{n-1}_A(R) \ \text{ for all } \ r \in R\},$$

where $[r, \partial]$ is the usual Lie bracket of operators, that is, $[r, \partial] = r \circ \partial - \partial \circ r$.

Example 33. The elements of $D^1_A(R)$ consist of the endomorphisms of the form $r + d$, where d is an A-linear derivation of R and $r \in R$.

Example 34. If k has characteristic zero and R is the polynomial ring $k[x_1, \ldots, x_n]$, then $D_k(R)$ is the Weil algebra $k[x_1, \ldots, x_n, \partial_1, \ldots, \partial_n]$ where each ∂_i denotes the derivation $\frac{\partial}{\partial x_i}$. This is the noncommutative subalgebra of $\mathrm{End}_k(R)$ generated by the $\frac{\partial}{\partial x_i}$ and the multiplication by x_j.

Example 35. In characteristic p, the differential operators on $k[x_1, \ldots, x_n]$ are essentially "the same" as in Example 34 as k-vector spaces, but not as rings. For example, if k has characteristic p, the operator

$$\left(\frac{\partial}{\partial x_i}\right)^p = \underbrace{\left(\frac{\partial}{\partial x_i} \circ \cdots \circ \frac{\partial}{\partial x_i}\right)}_{p \text{ times}}$$

obtained by composing the first-order operator $\frac{\partial}{\partial x_i}$ with itself p-times is the *zero operator.* Nonetheless, there is a differential operator

$$\frac{1}{p!}\frac{\partial^p}{\partial x_i^p}$$

sending x_i^p to 1, which is *not* the composition of lower order operators but which essentially has the same effect as the corresponding composition in characteristic zero. In particular, a k-basis for $D_k(k[x_1, \ldots, x_n])$, where k has characteristic p, is

$$\left\{ x_1^{j_1} \cdots x_n^{j_n}, \ \frac{1}{i_1!}\frac{\partial^{i_1}}{\partial x_1^{i_1}} \circ \frac{1}{i_2!}\frac{\partial^{i_2}}{\partial x_2^{i_2}} \circ \cdots \circ \frac{1}{i_n!}\frac{\partial^{i_n}}{\partial x_n^{i_n}} \right\}$$

as j_ℓ and i_ℓ range over all nonnegative integers. In characteristic p, $D_k(k[x_1, \ldots, x_n])$ is not finitely generated.

The following alternate interpretation of differential operators in characteristic p ties into the definition of F-threshold:

Proposition 36. *Let R be any ring of prime characteristic p such that the Frobenius map is finite (main case for us: $R = \mathbb{F}_p[x_1, \ldots, x_n]$). Then an $\mathbb{F}_p$-linear map $R \xrightarrow{d} R$ is a differential operator if and only if it is linear over some subring of p^eth powers. That is,*

$$D_{\mathbb{F}_p}(R) = \bigcup_{e \in \mathbb{N}} \mathrm{End}_{R^{p^e}}(R).$$

Proof. This is not a difficult fact to prove. It is first due to [63], or see [57] for a detailed proof in this generality. □

This gives us an alternate filtration of $D_{\mathbb{F}_p}(R)$ by "Frobenius order:" $D_{\mathbb{F}_p}(R)$ is the union of the chain

$$\mathrm{End}_R(R) = R \subset \mathrm{End}_{R^p}(R) \subset \mathrm{End}_{R^{p^2}}(R) \subset \mathrm{End}_{R^{p^3}}(R) \subset \dots$$

Alternatively, by taking p^eth roots, we can interpret the ring of differential operators as the union

$$\mathrm{End}_R(R) = R \subset \mathrm{End}_R(R^{\frac{1}{p}}) \subset \mathrm{End}_R(R^{\frac{1}{p^2}}) \subset \mathrm{End}_R(R^{\frac{1}{p^3}}) \subset \dots$$

Using this filtration, we can give an alternative definition of test ideals and F-threshold in terms of differential operators.

Definition 37. The *F-threshold* of $f \in \mathbb{F}_p[x_1, \dots, x_n]$ at the maximal ideal $\mathfrak{m}$ is defined as

$$FT_{\mathfrak{m}}(f) = \sup \left\{ c = \frac{a}{p^e} \in \mathbb{Z}\left[\frac{1}{p}\right] \mid \exists\, \partial \in \mathrm{End}_R(R^{1/p^e}) \right.$$
$$\left. \text{such that } \partial(f^{a/p^e}) \notin \mathfrak{m}R^{1/p^e} \right\}.$$

Note that this interprets the F-threshold as very much like the multiplicity: it is defined as the maximal (Frobenius) order of a differential operator which, when applied to (a power of) f, we get a nonvanishing function. However, here, the operators are filtered using Frobenius.

Similarly, the test ideal is the image of f^{a/p^e} over all differential operators of R^{1/p^e} which have image in R.

Remark 38 (Historical Remarks and Further Work). The F-threshold was first defined by Shunsuke Takagi and Kei-ichi Watanabe in [60], who called it the F-pure threshold. The definition looked quite different, since they defined it using ideas from Hochster and Huneke's tight closure theory [28]. Expanding on this idea, Hara and Yoshida [21] introduced the test ideals (under the name "generalized test ideals" in reference to the original test ideal of Hochster and Huneke, which was not defined for pairs), and soon later using this train of thought, F-jumping numbers were introduced in [43], where they are called F-thresholds. The definition of F-pure threshold, test ideals, and the higher F-jumping numbers we presented here is essentially from [4], where it is also proven that this point of view is equivalent (in regular rings) to the previously defined concepts. This point of view removes explicit mention of tight closure, focusing instead on R^{p^e}-linear maps (or differential operators).

We have presented the definition and only the simplest properties of F-threshold and test ideals, and only in the simplest possible case: of one polynomial in a polynomial ring or what amounts to a hypersurface in a smooth ambient scheme. Nor have we included any substantial applications. We urge the reader to investigate the survey [48] or others previously mentioned. In particular, test ideals are defined not only for individual polynomials but for any ideal, and the ambient ring need not be regular (e.g., the F-jumping numbers are discrete and rational in greater generality [6]). There are a great number of beautiful applications to developing tools from birational geometry in characteristic p using test ideals, such as Schwede's "centers of F-purity" and F-adjunction [46, 47]; see the papers of Schwede, Tucker, Takagi, Hara, Watanabe, Yoshida, Zhang, Blickle, and others listed in the bibliographies of [48].

4 Unifying the Prime Characteristic and Zero Characteristic Approaches

We defined the log canonical threshold for complex polynomials using integration and the F-threshold for characteristic p polynomials using differential operators. However, as we have seen, in characteristic zero, our approach was equivalent to a natural approach to measuring singularities in birational geometry. Since birational geometry makes sense over any field, might we also be able to define the F-threshold directly in this world as well?

This approach does not work as well as we would hope in characteristic p. Two immediate problems come to mind. First, resolution of singularities is not known in characteristic p. It turns out that this is not a very serious problem. Second, and more fatally, some of the vanishing theorems for cohomology that make multiplier ideals such a useful tool in characteristic zero actually *fail* in prime characteristic.

The lack of Hironaka's theorem in characteristic p can be circumvented as follows. We look at *all proper birational maps* $X \xrightarrow{\pi} \mathbb{C}^N$ with X normal. If X is a normal variety, the needed machinery of divisors goes through as in the smooth case, because the singular locus of a normal variety is of codimension two or higher. To define the order of vanishing of a function along an irreducible divisor D, we restrict to any (sufficiently small) smooth open set meeting the divisor. Thus, the relative canonical divisor K_π can be defined for a map $X \xrightarrow{\pi} \mathbb{C}^N$, for any *normal* X, as can the divisor $F = \text{div}(f \circ \pi)$. That is, if $X \xrightarrow{\pi} \mathbb{C}^N$ with X normal, we can define K_π and F as the divisor of the Jacobian determinant of $\pi_{|U}$ and the divisor of $f \circ \pi_{|U}$, respectively, where $U \subset X$ is the smooth locus of X (or any smooth subset of X whose complement is codimension two or more).

So we can attempt to define the log canonical threshold in arbitrary characteristic as follows:

Definition 1. The *log canonical threshold* of a polynomial $f \in k[x_1, \ldots, x_n]$ is defined as

$$lct(f) := \sup\left\{\lambda \in \mathbb{R}_+ \;\middle|\; \lceil K_\pi - \lambda F \rceil \text{ is effective}\right\},$$

as we range over all proper birational morphisms $X \xrightarrow{\pi} \mathbb{A}^n_k$ with X normal.

Remark 2. It is not hard to show that, in characteristic zero, this definition produces the same value as Definition 9. See [37, Thm 9.2.18].

Similarly, for any divisor on a normal variety X, the sheaves $\mathcal{O}_X(D)$ are defined.[11] So we can also attempt to define the multiplier ideal in characteristic p similarly by considering *all* proper birational models.

Definition 3. Let f be a polynomial in n variables and λ a positive real number. The multiplier ideal is

$$\mathcal{J}(f^\lambda) = \bigcap_{X \xrightarrow{\pi} \mathbb{A}^n_k} \pi_*\mathcal{O}_X(\lceil K_\pi - \lambda F \rceil) = \{h \in \mathbb{C}[x_1, \dots x_n] \mid \mathrm{div}(h) + \lceil K_\pi - \lambda F \rceil \geq 0\}$$

as we range over normal varieties X, mapping properly and birationally to $\mathbb{A}^n_k$ via π, where K_π is the relative canonical divisor and F is the divisor $\mathrm{div}(f \circ \pi)$ on X. Equivalently, this amounts to

$$\mathcal{J}(f^\lambda) = \{h \in k[x_1, \dots, x_n] \mid \mathrm{ord}_E(h) \geq \lambda \mathrm{ord}_E(f \circ \pi) - \mathrm{ord}_E(\mathrm{Jac}_{\mathbb{C}}(\pi))\},$$

where we range over all irreducible divisors E lying on a normal X mapping properly and birationally to $\mathbb{A}^n_k$, say, $X \xrightarrow{\pi} \mathbb{A}^n_k$.

Again, in characteristic 0, this produces the same definition as before. Does the multiplier ideal in characteristic p (produced by Definition 3) have the same good properties as in characteristic zero? The answer is NO. The problem occurs with the behavior of multiplier ideals in prime characteristic under wildly ramified maps: they simply do not have the properties we expect of multiplier ideals based on their behavior in characteristic zero (see [48, Example 6.33] or [48, Example 7.12]). The test ideals have better properties in char p than the multiplier ideals. They accomplish much of what multiplier ideals do in characteristic zero. The survey [48] gives an excellent introduction to this topic.

One reason the multiplier ideals fail to be useful in prime characteristic is that certain vanishing theorems fail that contribute to the magical properties of multiplier ideals over $\mathbb{C}$. For example, a very useful statement is "local vanishing": If $\pi : X \to \mathbb{A}^N_k$ is a log resolution of a complex polynomial f, then

$$R^i\pi_*\mathcal{O}_X(K_\pi - \lfloor cF \rfloor) = 0$$

for all $i > 0$ (cf. [37, Theorem 9.4.1]). For example, local vanishing is needed to prove the Briançon–Skoda theorem for non-principal ideals in characteristic zero.

[11] Although unlike the smooth case, they need not be invertible sheaves in general.

Unfortunately, this vanishing theorem is false in characteristic p. Fortunately, the Briançon–Skoda theorem for test ideals can be proven quite simply in characteristic p, using Frobenius instead of vanishing theorems.

On the other hand, for "large p," it is true that the multiplier ideals "reduce mod p" to the test ideals.

4.1 Idea of Reduction Modulo p

Fix a polynomial $f \in \mathbb{Q}[x_1, \dots, x_n]$ or (by clearing denominators) in $\mathbb{Z}[x_1, \dots, x_n]$. Fix a log resolution of f over $\mathbb{Q}$, say, given by $X_\mathbb{Q} \xrightarrow{\pi} \mathbb{A}^n_\mathbb{Q}$. We can "thicken" $X_\mathbb{Q}$ to a scheme $X_\mathbb{Z}$ over $\mathbb{Z}$ and so get a family of maps over $\operatorname{Spec}\mathbb{Z}$ described by the following diagram:

$$\begin{array}{ccccc}
X_\mathbb{Q} & & X_\mathbb{Z} & \hookleftarrow & X_{\mathbb{F}_p} \\
\downarrow \pi & & \downarrow & & \downarrow \\
\mathbb{A}^n_\mathbb{Q} & \hookrightarrow & \mathbb{A}^n_\mathbb{Z} & \hookleftarrow & \mathbb{A}^n_{\mathbb{F}_p} \\
\downarrow & & \downarrow & & \downarrow \\
\operatorname{Spec}\mathbb{Q} & \hookrightarrow & \operatorname{Spec}\mathbb{Z} & \hookleftarrow & \operatorname{Spec}(\mathbb{F}_p)
\end{array}$$

where the right-hand side gives a fiber over a closed point p in $\operatorname{Spec}\mathbb{Z}$ and the left-hand side shows the generic fiber. Because the generic fiber is a log resolution of f, it follows that for an open set of closed fibers, we also have a log resolution of f. That is, we can assume $X_{\mathbb{F}_p} \longrightarrow \mathbb{A}^n_{\mathbb{F}_p}$ is a log resolution of f for $p \gg 0$.

The multiplier ideal $\mathcal{J}(\mathbb{A}^n_\mathbb{Q}, f^c) \subset \mathbb{Q}[x_1, \dots, x_n]$ can be viewed as an ideal in $\mathbb{Z}[x_1, \dots, x_n]$ by clearing denominators if necessary; abusing notation, we denote the ideal in $\mathbb{Z}[x_, \dots, x_n]$ and $\mathbb{Q}[x_1, \dots, x_n]$ the same way. So we can reduce modulo p and obtain an analog of multiplier ideals in positive characteristic given by

$$\mathcal{J}(\mathbb{A}^n_{\mathbb{F}_p}, f^c) = \mathcal{J}(\mathbb{A}^n_\mathbb{Q}, f^c) \otimes \mathbb{F}_p.$$

These turn out to be the test ideals for $p \gg 0$!

Theorem 4 ([55, Theorem 3.1], [19], [21, Theorem 6.8]).

- *For all $p \gg 0$ and for all c,*

$$\tau(\mathbb{F}_p[x_1, \dots, x_n], f^c) \subseteq \mathcal{J}(\mathbb{Q}[x_1, \dots, x_n], f^c) \otimes \mathbb{F}_p.$$

- *Fix c, for all $p \gg 0$,*

$$\tau(\mathbb{F}_p[x_1, \ldots, x_n], f^c) = \mathcal{J}(\mathbb{Q}[x_1, \ldots, x_n], f^c) \otimes \mathbb{F}_p.$$

(How large is large enough for p depends on c.)

As we have explained, however, the test ideals are probably the "right" objects to use in each particular characteristic p.

Recently, Blickle, Schwede, and Tucker have found an interesting way to unify test ideals and multiplier ideals [7]. The idea is to look at a broader class of proper maps, not just birational ones. Recall that a surjective morphism of varieties $X \xrightarrow{\pi} Y$ is an *alteration* if it is proper and generically finite. We say an alteration is separable if the corresponding extension of function fields $k(Y) \subset k(X)$ is separable. Note that such π always factors as $X \xrightarrow{\phi} \tilde{Y} \xrightarrow{\nu} Y$ where ϕ is proper birational and ν is finite.

Consider a separable alteration $X \xrightarrow{\pi} \mathbb{A}^n$, with X normal. Denote by F_π the divisor on X defined by $f \circ \pi$ and by K_π the divisor on X defined by the Jacobian.[12] As before in our computation of the multiplier ideal, the idea is to push down the sheaf of ideals $\mathcal{O}_X(\lceil K_\pi - \lambda F_\pi \rceil)$ to $\mathcal{O}_X$. However, this will only produce a subsheaf of $\pi_* \mathcal{O}_X$, which is not $\mathcal{O}_{\mathbb{A}^n}$ but rather some normal finite extension. Let us denote its global sections by S, which is a normal finite extension of the polynomial ring R. To produce an ideal in R, we can use the *trace map.*

4.2 Trace

Let $R \subset S$ be a finite extension of normal domains, with corresponding fraction field extension $K \subset L$. The field trace is a K-linear map

$$L \to K$$

sending each $\ell \in L$ to the trace of the K-linear map $L \to L$ given by multiplication by ℓ. Because S is integral over the normal ring R, it is easy to check that this restricts to an R-linear map

$$S \xrightarrow{tr} R.$$

In particular, every ideal of S is sent, under the trace map, to an ideal in R. Using this, we can give a uniform definition of the multiplier ideal and test ideal. For any separable alteration $X \xrightarrow{\pi} \mathbb{A}^n$ with X normal, denote by tr_π the trace of the ring extension map $R \hookrightarrow S = \mathcal{O}_X(X)$. Then we have:

[12] By which we mean the unique divisor on X which agrees with these divisors on the smooth locus of X. This is possible since X is normal; see the beginning paragraphs of Sect. 4.

Theorem 5 ([7]). *Fix a polynomial $f \in k[x_1, \dots, x_n]$ where k is an arbitrary field, and let c be any positive real number. Define*

$$J := \bigcap_{\pi} tr_{\pi}(\pi_* \mathcal{O}_X(\lceil K_\pi - cF_\pi \rceil)).$$

where π varies over all possible normal varieties X mapping properly and generically separably to $\mathbb{A}^n$. Then,

$$J = \begin{cases} \mathcal{J}(f^c) & \text{if the characteristic is } 0, \\ \tau(f_p^c) & \text{if the characteristic is } p > 0. \end{cases}$$

Note that each $\pi_* \mathcal{O}_X(\lceil K_\pi - c\pi^* D \rceil)$ is an ideal in $\pi_* \mathcal{O}_X$, whose global sections form some finite extension S of the polynomial ring $R = k[x_1, \dots, x_n]$. So its image under the trace map is an ideal in R. The theorem says that if we intersect all such ideals of R, we get the test/multiplier ideal of f^λ.

In fact, Blickle, Schwede, and Tucker prove even more: The intersection stabilizes. So there is *one* alteration $X \xrightarrow{\pi} \mathbb{A}^n$ for which $\tau(f^c) = tr_\pi(\pi_* \mathcal{O}_X(\lceil K_\pi - c\pi^* D \rceil))$. In fact, it has been shown that, fixing f, there is one alteration which computes all test ideals $\tau(f^\lambda)$, for any λ [50]. Of course, this is already known for multiplier ideals: it suffices to take one log resolution of X to compute the multiplier ideal. Indeed, in characteristic zero, one need not take any finite covers at all. Interestingly, in characteristic p, it is precisely the finite covers that matter most.

It is worth remarking that many of the features of multiplier ideals—including the important local vanishing theorem—can be shown in characteristic p "up to finite cover" [7, Theorem 5.5]. This can be viewed as a generalization of Hochster and Huneke's original theorem on the Cohen-Macaulayness of the absolute integral closure of a domain of characteristic p [29]. See also [54].

Acknowledgements The authors are grateful to Daniel Hernández, Jeffrey Lagarias, Kai Rajala, and Kevin Tucker for many useful comments. In addition, two referees made excellent suggestions and spotted many misstatements. These notes began as a lecture notes from a minicourse by the third author at a workshop at Luminy entitled "Multiplier Ideals in Commutative Algebra and Singularity Theory" organized by Michel Granger, Daniel Naie, and Anne Pichon in January 2011; the authors are grateful to the organizers of that conference. They were then polished at the workshop "Computational F-singularities" organized by Karl Schwede, Kevin Tucker, and Wenliang Zhang in Ann Arbor Michigan in May 2012.

The first author is partially supported by MTM2009-07291, by the Fellowship Fundación Ramón Areces para Estudios Postdoctorales 2011/12, and by Spanish Goverment in the frame of Postdoctoral mobility abroad Fellowship (code: EX-2010-0128). The second author has been supported by a For Women in Science award 2011 of L'Oréal Austria, the Austrian commission for UNESCO and the Austrian Academy of Sciences and by the Austrian Science Fund (FWF) in frame of projects J-3326N25 and P21461. The third author is partially supported by NSF grant DMS-1001764.

References

1. Bhatt, B.: The F-pure threshold of an elliptic curve, available at http://www-personal.umich.edu/~bhattb/math/cyfthreshold.pdf
2. Bhatt, B., Singh, A.: F-thresholds of Calabi-Yau hypersurfaces. In preparation.
3. Blickle, M.: Multiplier ideals and modules on toric varieties. Math. Z. **248**(1), 113–121 (2004)
4. Blickle, M., Mustaţă, M., Smith, K.E.: Discreteness and rationality of F-thresholds. Michigan Math. J. **57**, 43–61 (2008)
5. Blickle, M., Mustaţă, M., Smith, K.E.: F-thresholds of hypersurfaces. Trans. Amer. Math. Soc. **361**, 6549–6565 (2009)
6. Blickle, M., Schwede, K., Takagi, S., Zhang, W.: Discreteness and rationality of F-jumping numbers on singular varieties. Math. Ann. **347**(4), 917–949 (2010)
7. Blickle, M., Schwede, K., Tucker, K.: F-singularities via alterations. `arXiv:1107.3807v3 [math.AG]`.
8. Budur, N.: On Hodge spectrum and multiplier ideals. Math. Ann. **327**(2), 257–270 (2003)
9. Budur, N.: Multiplier ideals, Milnor fibers, and other singularity invariants. Notes from a 5-lectures course at CIRM, Luminy, January 2011. Available at http://www.nd.edu/~nbudur/LNLuminy.pdf
10. de Fernex, T., Ein, L., Mustata, M.: Bounds for log canonical thresholds with applications to birational rigidity. Math. Res. Lett. **10**, 219–236 (2003)
11. Demailly, J.P.: L^2 vanishing theorems for positive line bundles and adjunction theory. Transcendental Methods in Algebraic Geometry (Cetraro, 1994), pp. 1–97. Lecture Notes in Mathematics, vol. 1646. Springer, Berlin (1996)
12. Demailly, J.P., Hiep, P.H.: A sharp lower bound for the log canonical threshold. Arxiv 1201-4086 preprint
13. Elkies, N.: Supersingular primes for elliptic curves over real number fields. Compositio **72**, 165–172 (1989)
14. Ein, L., Lazarsfeld, R., Smith, K.E., Varolin, D.: Jumping coefficients of multiplier ideals. Duke Math. J. **123**, 469–506 (2004)
15. Esnault, H., Viehweg, E.: Lectures on vanishing theorems. DMV Seminar, vol. 20. Birkhäuser, Basel (1992). vi+164 pp. ISBN: 3-7643-2822-3
16. Fedder, R.: F-purity and rational singularity. Trans. Amer. Math. Soc. **278**, 461–480 (1983)
17. Griffiths, P., Harris, J.: Principles of algebraic geometry. Pure and Applied Mathematics. Wiley, New York (1978). xii+813 pp. ISBN: 0-471-32792-1
18. Grothendieck, A.: Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas (Quatrième Partie). Inst. Hautes Études Sci. Publ. Math. **32**, 361 (1967)
19. Hara, N.: Geometric interpretation of tight closure and test ideals. Trans. Amer. Math. Soc. **353**(5), 1885–1906 (2001)
20. Hara, N., Watanabe, K.I.: F-regular and F-pure rings vs. log terminal and log canonical singularities. J. Algebraic Geom. **11**(2), 363–392 (2002)
21. Hara, N., Yoshida, K.I.: A generalization of tight closure and multiplier ideals. Trans. Amer. Math. Soc. **355**(8), 3143–3174 (2003)
22. Hartshorne, R.: Algebraic geometry. Graduate Texts in Mathematics, vol. 52. Springer, New York (1977). xvi+496 pp. ISBN: 0-387-90244-9
23. Hernández, D.: F-purity of hypersurfaces. Ph.D. thesis, University of Michigan (2011). Available at http://deepblue.lib.umich.edu/handle/2027.42/86491
24. Hernández, D.: F-purity versus log canonicity for polynomials. Available at `arXiv:1112.2423v1`
25. Hernández, D.: F-invariants of diagonal hypersurfaces. Available at `arXiv:1112.2425v1`
26. Hernández, D.: F-pure thresholds of binomial hypersurfaces. Proceedings of the AMS, Available at `arXiv:1112.2427v1`

27. Hironaka, H.: Resolution of singularities of an algebraic variety over a field of characteristic zero I-II. Ann. of Math. (2) **79**, 109–326 (1964)
28. Hochster, M., Huneke, C.: Tight closure, invariant theory, and the Briançon-Skoda Theorem. J. Amer. Math. Soc. **3**, 31–116 (1990)
29. Hochster, M., Huneke, C.: Infinite integral extensions and big Cohen–Macaulay algebras. Ann. Math. **135**(1), 53–89 (1992)
30. Hochster, M., Huneke, C.: Comparison of symbolic and ordinary powers of ideals. Invent. Math. **147**, 349–369 (2002)
31. Hochster, M., Roberts, J.L.: Rings of invariants of reductive groups acting on regular rings are Cohen–Macaulay. Adv. Math. **13**, 115–175 (1974)
32. Howald, J.A.: Multiplier ideals of monomial ideals. Trans. Amer. Math. Soc. **353**, 2665–2671 (2001)
33. Igusa, J.-I.: An introduction to the theory of local zeta functions. AMS/IP Studies in Advanced Mathematics, vol. 14. American Mathematical Society, Providence, International Press, Cambridge (2000)
34. Kollár, J.: Singularities of pairs. Algebraic geometry–Santa Cruz 1995. In: Proceedings of Symposium Pure Mathematics, vol. 62, Part 1, pp. 221–287. American Mathematical Society, Providence, RI (1997)
35. Kollár, J., Mori, S.: Birational geometry of algebraic varieties. Cambridge Tracts in Mathematics, vol. 134. Cambridge University Press, Cambridge (1998). viii+254 pp.
36. Kunz, E.: Characterizations of regular local rings for characteristic p. Amer. J. Math. **91** 772–784 (1969)
37. Lazarsfeld, R.: Positivity in algebraic geometry II. Ergeb. Math. Grenzgeb. (3), vol. 48. Springer, Berlin (2004)
38. Lucas, E.: Theorie des fonctions numeriques simplement periodiques. Amer. J. Math. **1**(3), 197–240 (1878)
39. Mehta, V.B., Ramanathan, A.: Frobenius splitting and cohomology vanishing for Schubert varieties. Ann. Math. (2) **122**(1), 27–40 (1985)
40. Mustata, M.: Singularities of pairs via jet schemes. J. Amer. Math. Soc. **15**, 599–615 (2002)
41. Mustaţă, M.: IMPANGA lecture notes on log canonical thresholds. Proceedings of the IMPANGA 2010 Summer school. To appear
42. Mustaţă, M., Yoshida, K.I.: Test ideals vs. Multiplier ideals. Nagoya Math. J. **193**, 111–128 (2009)
43. Mustaţă, M., Takagi, S., Watanabe, K.I.: F-thresholds and Bernstein–Sato polynomials. European Congress Math. vol. 341–364. Eur. Math. Soc., Zürich (2005)
44. Nadel, A.M., Multiplier ideal sheaves and existence of Kähler–Einstein metrics of positive scalar curvature. Proc. Nat. Acad. Sci. U.S.A. **86**(19), 7299–7300 (1989)
45. On Tian's invariant and log canonical thresholds. Appendix to Cheltsov and C. Shramov article "Log canonical thresholds of smooth Fano threefolds". Uspekhi Mat. Nauk **63**(5(383)), 73–180 (2008)
46. Schwede, K.: F-adjunction Algebra Number Theory **3**(8), 907–950 (2009)
47. Schwede, K.: Centers of F-purity. Math. Z. **265**(3), 687–714 (2010)
48. Schwede, K., Tucker, K.: A survey of test ideals. In: Francisco, C.L., Klingler, C., Sather-Wagstaff, S.M., Vassilev, J.C. (eds.) Progress in Commutative Algebra 2 Closures, Finiteness and Factorization, pp. 39–100. De Gruyter, Berlin, New York (2012)
49. Schwede, K., Tucker, K.: On the behavior of test ideals under finite morphisms. Available at `arXiv: 1003.4333`
50. Schwede, K., Tucker, K., Zhang, W.: Test ideals via a single alteration and discreteness and rationality of F-jumping numbers. Math. Res. Lett. **19**(01), 191–197 (2012)
51. Shibuta, T., Takagi, S.: Log canonical threshold of binomial ideals. Manuscripta Math. **130**(1), 45–61 (2009)
52. Shokurov, V.V., Letters of a bi-rationalist. V. Minimal log discrepancies and termination of log flips. (Russian) Tr. Mat. Inst. Steklova **246**, (2004). Algebr. Geom. Metody, Svyazi i Prilozh., 328–351; translation in Proceedings of Steklov Institute of Mathematics, 3 (246), 315–336 (2004)

53. Smith, K.E.: F-rational rings have rational singularities. Amer. J. Math. **119**(1), 159–180 (1997)
54. Smith, K.E.: Vanishing, singularities and effective bounds via prime characteristic local algebra. Algebraic geometry–Santa Cruz 1995. In: Proceedings of Symposium Pure Mathematics, vol. 62, Part 1, pp. 289–325. American Mathematical Society, Providence (1997)
55. Smith, K.E.: The multiplier ideal is a universal test ideal. Special issue in honor of Robin Hartshorne. Comm. Algebra **28**(12), 5915–5929 (2000)
56. Smith, K.E., Thompson, H.M.: Irrelevant exceptional divisors for curves on a smooth surface. Algebra, geometry and their interactions.Contemporary Mathematics, vol. 448, pp. 245–254. Amer. Math. Soc., Providence (2007)
57. Smith, K.E., Van den Bergh, M.: Simplicity of rings of differential operators in prime characteristic. Proc. London Math. Soc. (3) **75**(1), 32–62 (1997)
58. Stenbrink, J.H.M.: The spectrum of hypersurface singularities. In: Théorie de Hodge, Luminy, vol. 1987. Astérisque 179–180 (1989), 163–184
59. Takagi, S.: An interpretation of test ideals via tight closure, J. Alg. Geom. **13**, 393–415 (2004)
60. Takagi, S., Watanabe, K.I.: On F-pure thresholds. J. Algebra **282**(1), 278–297 (2004)
61. Tucker, K.: Jumping numbers on algebraic surfaces with rational singularities. Trans. Amer. Math. Soc. **362**(6), 3223–3241 (2010)
62. Varchenko, Asymptotic Hodge structures in the vanishing cohomology. Math. USSR Izv. **18**, 469–512 (1982)
63. Yekutieli, A.: An explicit construction of the Grothendieck residue complex. With an appendix by Pramathanath Sastry. Astérisque **208**, 127 (1992)

Three Flavors of Extremal Betti Tables

Christine Berkesch, Daniel Erman, and Manoj Kummini

1 Introduction

Classification problems can be often discretized by replacing a collection of complicated objects by numerical invariants. For instance, if we are interested in modules over a local or graded ring, then we can study their Hilbert polynomials, Betti numbers, Bass numbers, and more. Describing the behavior of these invariants becomes a proxy for understanding the modules; identifying the extremal behavior of an invariant provides structural limitations.

The conjectures of Boij and Söderberg [6], proven by Eisenbud and Schreyer [11], link the extremal properties of invariants of free resolutions over the graded polynomial ring $S = \Bbbk[x_1, \ldots, x_n]$ with the Herzog–Huneke–Srinivasan Multiplicity Conjectures. Here $\Bbbk$ is any field, S has the standard $\mathbb{Z}$-grading, and we study the graded Betti tables of S-modules. The Boij–Söderberg conjectures state that the extremal rays of the cone of Betti tables are given by Betti tables of Cohen–Macaulay modules with pure resolutions. There exist two excellent introductions to Boij–Söderberg theory [12, 15].

C. Berkesch
Department of Mathematics, Duke University, Box 90320,
Durham, NC 27708, USA
e-mail: cberkesc@math.duke.edu

D. Erman (✉)
Department of Mathematics, University of Michigan,
Ann Arbor, MI, 48109, USA
e-mail: derman@umich.edu

M. Kummini
Chennai Mathematical Institute, Siruseri, Tamilnadu, 603103, India
e-mail: mkummini@cmi.ac.in

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_4,

In this chapter, we explore the notion of an extremal Betti table in three different contexts: in the original setting of a standard graded polynomial ring, over a regular local ring, and over a finely graded polynomial ring.

Previous work has considered the extremal behavior of free resolutions, in a manner unconnected to Boij–Söderberg theory. Each graded Betti number of the Eliahou–Kervaire resolution of a lex-segment ideal is known to be maximal among cyclic modules with the same Hilbert function [5, 19, 22]. Also, [1] studies the Betti numbers of modules with extremal homological dimensions, complexity, or curvature. Though we will not discuss these types of results further, the interested reader might consider [2, 20, 23].

Throughout this chapter, S will denote a standard graded polynomial ring, R will denote a regular local ring, and T will denote a finely graded polynomial ring. For a graded S-module M, we define the **graded Betti numbers** $\beta_{i,j}(M) := \dim_{\Bbbk} \operatorname{Tor}_i^S(M,\Bbbk)_j$. Betti numbers also have a more concrete interpretation: if $\mathbf{F} = [\mathbf{F}_0 \leftarrow \mathbf{F}_1 \leftarrow \cdots \leftarrow \mathbf{F}_n \leftarrow 0]$ is a minimal graded free resolution of M, then $\beta_{i,j}(M)$ is the number of minimal generators of $\mathbf{F}_i$ of degree j. The **graded Betti table** of M, denoted $\beta(M)$, is the vector with coordinates $\beta_{i,j}M$ in the vector space $\mathbb{V} = \bigoplus_{i=0}^n \bigoplus_{j\in\mathbb{Z}} \mathbb{Q}$.

For local and multigraded rings, there are analogous definitions. For a regular local ring R with residue field $\Bbbk$, we define the **(local) Betti numbers** of an R-module as $\beta_i^R(M) = \dim_{\Bbbk} \operatorname{Tor}_i^R(M,\Bbbk)$. Over a $\mathbb{Z}^m$-graded polynomial ring T, we define the **multigraded Betti numbers** of a T-module M as $\beta_{i,\alpha}^T(M) = \dim_{\Bbbk} \operatorname{Tor}_i^T(M,\Bbbk)_\alpha$, where $\alpha \in \mathbb{Z}^m$. We denote the respective Betti tables by $\beta^R(M)$ and $\beta^T(M)$.

To streamline the exposition, we focus on modules of finite length. With minor adjustments, most results we discuss can be extended to the case of finitely generated modules. See [7, 12, 15] for the standard graded case and [4] for the local case.

Let M be a graded S-module (or an R-module or a multigraded T-module) of finite length. We say that $\beta(M)$ is **extremal** if, for any decomposition of the form

$$\beta(M) = \beta(M') + \beta(M'')$$

with M', M'' graded S-modules (or R-modules or multigraded T-modules, respectively), we have that $\beta(M')$ is a scalar multiple of $\beta(M)$. Extremal Betti tables correspond to extremal rays of the cone of Betti tables of finite length. In the case of S, this is the cone

$$B_{\mathbb{Q}}^{\text{fin}}(S) := \mathbb{Q}_{\geq 0} \cdot \{\beta(M) \mid M \text{ is a graded } S\text{-module of finite length}\} \subseteq \mathbb{V}.$$

Boij and Söderberg observed that for graded S-modules, there is a natural sufficient condition for extremality.

Claim 1.1. *For a graded S-module M of finite length, if M has a pure resolution, then $\beta(M)$ is extremal.*

Here we say that M has a **pure resolution** if, for each i, $\beta_{i,j}(M) \neq 0$ for at most one j. After proving the claim, Boij and Söderberg conjectured that this condition is not only sufficient but also necessary. In fact, after imposing some obvious degree restrictions on the Betti table, they conjecture the existence of pure resolutions of Cohen–Macaulay modules of essentially any combinatorial type. This was later proven by [13] in characteristic 0 and by [11] in a characteristic-free manner; see Theorem 3.

In Sect. 3, we first quickly review why Claim 1.1 provides a sufficient condition for extremality. The remainder of the section is an expository overview of Eisenbud and Schreyer's construction of modules with pure resolutions.

We then turn our attention to the case of a regular local ring, as considered in [4]. In contrast with the graded case, there is no obvious analogue of Claim 1.1. In retrospect this is inevitable, as there are no modules of finite length whose Betti tables are extremal.

In the final section, we move in the opposite direction, refining the grading to a finely graded polynomial ring T. One possibility for understanding extremal Betti tables in the multigraded setting is to seek out multigraded lifts of pure resolutions from the standard $\mathbb{Z}$-graded setting. This approach is taken in [14], which considers the linear space of such multigraded Betti tables. Moreover, in the case of $\mathbb{k}[x, y]$ with $\mathbb{Z}^2$-grading, [8] constructs the entire cone of bigraded Betti tables spanned by such lifted pure resolutions.

Not all extremal Betti tables arise in this way in the multigraded setting, and we provide a sufficient condition for a bigraded Betti table to be extremal, which demonstrates this fact. The extra rigidity induced by the bigrading seems to greatly complicate the picture. We use this condition to show the existence of a zoo of extremal Betti tables.

2 Preliminaries

Given a ring R (or a scheme X) and a complex $\mathbf{F}$ of R-modules (or $\mathcal{O}_X$-modules) with differential $\partial_i : \mathbf{F}_i \longrightarrow \mathbf{F}_{i-1}$, we denote the homology modules of $\mathbf{F}$ by $\mathrm{H}_i(\mathbf{F}) = (\ker \partial_i)/(\mathrm{im} \partial_{i+1})$. The derived category of R-modules (or of $\mathcal{O}_X$-modules) is the category consisting complexes of R-modules (or $\mathcal{O}_X$-modules) modulo the equivalence relation generated by quasi-isomorphisms. We may represent any object in the derived category by a genuine complex of modules.

For a projection of the form $\pi_1 \colon X \times \mathbb{P}^m \longrightarrow X$ of schemes, there are well-defined higher direct image functors $R^i\pi_{1*}$ that take a sheaf on $X \times \mathbb{P}^m$ (or a complex of sheaves on $X \times \mathbb{P}^m$) to a sheaf on X (or a complex of sheaves on X). Further, if we are willing to work with the derived category, then there is a single functor $R\pi_{1*}$ that combines all of these higher direct image functors: the functor $R\pi_{1*}$ takes a sheaf $\mathcal{F}$ on $X \times \mathbb{P}^m$ (or a complex $\mathbf{F}$ of sheaves) and returns an object in the derived category of $\mathcal{O}_X$-modules. The functor $R\pi_{1*}$ combines the higher direct image functors in the sense that if $\mathbf{G}$ is any complex that represents $R\pi_{1*}\mathbf{F}$, then $\mathrm{H}_i(\mathbf{G}) \cong R^{-i}\pi_{1*}\mathbf{F}$ for all i. In the special case where $X = \mathrm{Spec}(A)$, we will view each $R^i\pi_{1*}\mathcal{F}$ as an A-module (instead of writing $\Gamma(X, R^i\pi_{1*}\mathcal{F})$) and similarly for $R\pi_{1*}$. If $\mathcal{F}$ is an $\mathcal{O}_{X\times\mathbb{P}^m}$-module, then $R^i\pi_{1*}\mathcal{F} = 0$ for all $i < 0$. Since computing $R\pi_{1*}\mathcal{F}$ depends only on the quasi-isomorphism class of $\mathcal{F}$, the same fact holds for any (locally free) resolution $\mathbf{F}$ of an $\mathcal{O}_{X\times\mathbb{P}^m}$-module.

Let π_2 be the second projection $X \times \mathbb{P}^m \longrightarrow \mathbb{P}^m$. Given a sheaf $\mathcal{G}$ on X and a sheaf $\mathcal{L}$ on $\mathbb{P}^m$, we set

$$\mathcal{G} \boxtimes \mathcal{L} := \pi_1^*\mathcal{G} \boxtimes \pi_2^*\mathcal{L}.$$

If $\mathcal{L} = \mathcal{O}_{\mathbb{P}^m}(-e)$ is a line bundle on $\mathbb{P}^m$, then by way of the projection formula [17, III, Ex. 8.3], computing $R\pi_{1*}(\mathcal{G} \boxtimes \mathcal{L})$ is straightforward, and we will use this computation repeatedly. There are three cases, depending on the value of e:

1. If $-e \geq 0$, then the only nonzero cohomology of $\mathcal{O}_{\mathbb{P}^m}(-e)$ is $H^0(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(-e))$, and we have that $R\pi_{1*}(\mathcal{G} \boxtimes \mathcal{O}_{\mathbb{P}^m}(-e))$ is the complex consisting of the sheaf $\mathcal{G} \otimes H^0(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(-e))$ in homological degree 0.
2. If $-1 \geq -e \geq -m$, then $\mathcal{O}_{\mathbb{P}^m}(-e)$ has no cohomology, so $R\pi_{1*}(\mathcal{G} \boxtimes \mathcal{O}_{\mathbb{P}^m}(-e)) = 0$.
3. If $-m-1 \geq -e$, then the only nonzero cohomology of $\mathcal{O}_{\mathbb{P}^m}(-e)$ is $H^m(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(-e))$, and we have that $R\pi_{1*}(\mathcal{G} \boxtimes \mathcal{O}_{\mathbb{P}^m}(-e))$ is the complex consisting of sheaf $\mathcal{G} \otimes H^m(\mathbb{P}^m, \mathcal{O}_{\mathbb{P}^m}(-e))$ in homological degree $-m$.

3 Extremal Betti Tables in the Graded Case

In this section, we first prove Claim 1.1, providing a sufficient condition for extremality in the graded case. We then focus on the Eisenbud–Schreyer construction of pure resolutions.

We assume throughout this section that $\Bbbk$ is an infinite field. By [10, Lemma 9.6], this assumption will not affect questions related to cones of Betti tables. A strictly increasing sequence of integers $d = (d_0 < d_1 < \cdots < d_n) \in \mathbb{Z}^{n+1}$ is called a **degree sequence** of S. We say a free resolution $\mathbf{F}$ is **pure of type** d if it has the form

$$\mathbf{F}: \quad S(-d_0)^{\beta_0} \longleftarrow S(-d_1)^{\beta_1} \longleftarrow \cdots \longleftarrow S(-d_n)^{\beta_n} \longleftarrow 0.$$

Proof of Claim 1.1. Our argument follows [6, Sect. 2.1], which extends a computation of Herzog and Kühl [18]; see also [12, Proposition 2.1].

Let M be a finite length module with a pure resolution

$$0 \leftarrow M \leftarrow S(-d_0)^{\beta_{0,d_0}} \leftarrow S(-d_1)^{\beta_{1,d_1}} \leftarrow \cdots \leftarrow S(-d_n)^{\beta_{n,d_n}} \leftarrow 0.$$

Suppose that $\beta(M) = \beta(M') + \beta(M'')$. Since M has finite length, it follows that M' would also have to be a finite length module (the Hilbert series is determined by the Betti table and is additive). Thus, by the Auslander–Buchsbaum theorem, the projective dimension of M' is n. It then follows from the decomposition of $\beta(M)$ that M' admits a pure resolution of type $(d_0 < d_1 < \cdots < d_n)$. Thus, if the Betti table of a pure resolution is unique up to scalar multiple, then $\beta(M')$ will be a scalar multiple of $\beta(M)$.

To prove that the β_{i,d_i} are determined (up to scalar multiple), we consider the Herzog–Kühl equations for M from [18]. Since M has finite length, the following n equations must vanish:

$$\begin{cases} \sum_{i=0}^{n}(-1)^i \beta_{i,d_i} & = 0; \\ \sum_{i=0}^{n}(-1)^i d_i \beta_{i,d_i} & = 0; \\ \vdots & \vdots \\ \sum_{i=0}^{n}(-1)^i d_i^{n-1} \beta_{i,d_i} & = 0. \end{cases} \tag{1}$$

Thinking of this as a system of n linear equations in the $(n+1)$-unknowns β_{i,d_i}, the solutions are given by the kernel of the matrix

$$\begin{pmatrix} 1 & -1 & \dots & (-1)^n \\ d_0 & -d_1 & \dots & (-1)^n d_n \\ \vdots & & \ddots & \vdots \\ d_0^{n-1} & -d_1^{n-1} & \dots & (-1)^n d_n^{n-1} \end{pmatrix}.$$

This is a rank n matrix; in fact, the $n \times n$ minor given by the first n columns is nonzero. To see this, rescale every other column by -1 to obtain an $n \times n$ Vandermonde matrix for $(d_0, \dots, d_{n-1})$. Since the d_i are strictly increasing, this Vandermonde determinant is nonzero. It thus follows that the kernel of this matrix has rank 1, so the β_{i,d_i} are uniquely determined, up to scalar multiple. $\square$

Remark 2. Using Cramer's rule and the formula for Vandermonde determinants, any solution $(\beta_{0,d_0}, \beta_{1,d_1}, \dots, \beta_{n,d_n})$ to the system (1) is a scalar multiple of

$$\left(\frac{1}{\prod_{j \neq 0} |d_0 - d_j|}, \frac{1}{\prod_{j \neq 1} |d_1 - d_j|}, \dots, \frac{1}{\prod_{j \neq n} |d_n - d_j|} \right).$$

We now show that any degree sequence of S is realized by a pure resolution. The first two constructions of pure resolutions are due to Eisenbud, Fløystad, and Weyman [13]. Their constructions are based on representation theory and Schur functors, and they thus require that $\Bbbk$ has characteristic 0. See [15, Sect. 3] for an expository treatment of those constructions. The first characteristic-free construction is due to Eisenbud and Schreyer [11]. Their construction, which relies on a spectral sequence or, equivalently, on the Kempf–Lascoux–Weyman Geometric Technique, was later generalized in [3].

The complex **K**

	Spec(S)	$\times \mathbb{P}^3$
$\mathbf{K}_0$	S^1	$\boxtimes \mathcal{O}_{\mathbb{P}^3}$
$\mathbf{K}_1$	$S(-1)^6$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-1)$
$\mathbf{K}_2$	$S(-2)^{15}$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-2)$
$\mathbf{K}_3$	$S(-3)^{20}$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-3)$
$\mathbf{K}_4$	$S(-4)^{15}$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-4)$
$\mathbf{K}_5$	$S(-5)^6$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-5)$
$\mathbf{K}_6$	$S(-6)^1$	$\boxtimes \mathcal{O}_{\mathbb{P}^3}(-6)$

$\xrightarrow{R\pi_{1*}}$

The complex **F**

	Spec(S)
$\mathbf{F}_0$	$S^1 \otimes H^0(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3})$
-	-
-	-
-	-
$\mathbf{F}_1$	$S(-4)^{15} \otimes H^3(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3}(-4))$
$\mathbf{F}_2$	$S(-5)^6 \otimes H^3(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3}(-5))$
$\mathbf{F}_3$	$S(-6)^1 \otimes H^3(\mathbb{P}^3, \mathcal{O}_{\mathbb{P}^3}(-6))$

Fig. 1 To construct a pure resolution **F** of type $(0, 4, 5, 6)$ on Spec(S'), we begin with a Koszul complex **K** on Spec$(S') \times \mathbb{P}^3$ and then use a pushforward construction to collapse three of the terms. A term $\mathbf{K}_i$ gets collapsed if the second factor is a line bundle on $\mathbb{P}^3$ with no cohomology

Theorem 3 ([11, Theorem 5.1]). *For any degree sequence $d = (d_0 < d_1 < \cdots < d_n)$, there exists a finite length graded S-module whose minimal free resolution is pure of type d.*

Of course, it suffices to prove the theorem in the case where $d_0 = 0$, as we can obtain a pure resolution of type $(d_0 < \cdots < d_n)$ by tensoring a pure resolution of type $(0 < d_1 - d_0 < \cdots < d_n - d_0)$ with $S(-d_0)$. When Boij and Söderberg conjectured the existence of pure resolutions, there were very few known examples. One family of examples that was known came from the Eagon–Northcott complex, the Buchsbaum–Rim complex, and other related complexes [9]. Lascoux had shown that these complexes could be constructed by applying a pushforward construction to a Koszul complex [21]. This pushforward construction has the effect of collapsing strands of the Koszul complex, and Eisenbud and Schreyer realized that (with the appropriate setup) this collapsing effect could be iterated. This became the key to their construction of pure resolutions.[1]

Before presenting Eisenbud and Schreyer's general construction for a pure resolution, we review the original collapsing technique in the following lemma. This produces a pure resolution of type $(0, q + 1, \ldots, q + n)$, which is the Eagon–Northcott complex for an $n \times (q + 1)$ matrix of linear forms over $\Bbbk[x_1, \ldots, x_{n+q}]$. The proof of this lemma contains all of the technical features required for the general case. An example is provided in Fig. 1.

Lemma 4. *Let q be a positive integer and let $S' := \Bbbk[x_1, \ldots, x_{n+q}]$. Let $f_1, \ldots, f_{n+q}$ be generic bilinear forms on* Spec$(S') \times \mathbb{P}^q$ *and let* **K** *be the Koszul complex of locally free sheaves on* Spec$(S') \times \mathbb{P}^q$ *given by the f_i. Then $R\pi_{1*}(\mathbf{K})$ is represented by a pure resolution* **F** *of type* $(0, q + 1, q + 2, \ldots, q + n)$ *that resolves a Cohen–Macaulay S'-module of codimension n.*

[1]The idea that Eisenbud and Schreyer's construction of pure resolutions is a higher-dimensional analogue of the Eagon–Northcott and Buchsbaum–Rim complexes is developed explicitly in [3, Sect. 10].

Proof. Since $\Bbbk$ is infinite, we may assume that the f_i form a regular sequence, and hence they define a q-dimensional subscheme $Z \subseteq \mathbb{A}^{n+q} \times \mathbb{P}^q$. The Koszul complex $\mathbf{K}$ is thus a resolution of $\mathcal{O}_Z$. The support of $\pi_{1*}\mathcal{O}_Z$ has dimension at most q and therefore has codimension at least n. In fact, we will later see that the S'-module $\pi_{1*}\mathcal{O}_Z$ is a Cohen–Macaulay of codimension n.

For $0 \leq i \leq n+q$, the $\mathbb{P}^q$-degree of the generators of $\mathbf{K}_i$ is i. By taking the direct images under the map $\pi_1 : \operatorname{Spec}(S') \times \mathbb{P}^q \longrightarrow \operatorname{Spec}(S')$, we will collapse the terms $\mathbf{K}_1, \mathbf{K}_2, \ldots, \mathbf{K}_q$, resulting in the desired pure resolution.

Our first goal is to show that $R^\ell \pi_{1*}\mathbf{K} \neq 0$ if and only if $\ell = 0$. We do this in two steps. As noted in Sect. 2, since $\mathbf{K}$ is a resolution of $\mathcal{O}_Z$, it follows that $R^\ell \pi_{1*}\mathbf{K} \neq 0$ only if $\ell \geq 0$.

By computing $R\pi_{1*}(\mathbf{K})$ in a second way, we will now show that $R^\ell \pi_{1*}\mathbf{K} \neq 0$ only if $\ell \leq 0$. Note that $\mathbf{K}_i = S'^{\binom{n+q-1}{i}}(-i) \boxtimes \mathcal{O}_{\mathbb{P}^q}(-i)$. For each i, let $C^{-i,\bullet}$ be the Čech resolution of $\mathbf{K}_i$ with respect to the standard Čech cover $\{\operatorname{Spec}(S') \times U_0, \ldots \operatorname{Spec}(S') \times U_q\}$ of $\operatorname{Spec}(S') \times \mathbb{P}^q$. Since the construction of Čech resolutions is functorial, we obtain a double complex $C^{\bullet,\bullet}$ consisting of π_{1*}-acyclic sheaves on $\operatorname{Spec}(S') \times \mathbb{P}^q$, which has the form:

$$
\begin{array}{llcll}
C^{\bullet,\bullet}: & & \vdots & & \vdots \\
& & \uparrow & & \uparrow \\
0 \longleftarrow & S' \boxtimes \left(\bigoplus_{k,k'=0}^{q} \mathcal{O}|_{U_k \cap U_{k'}}\right) & \longleftarrow & S'(-1)^{n+q} \boxtimes \left(\bigoplus_{k,k'=0}^{q} \mathcal{O}(-1)|_{U_k \cap U_{k'}}\right) & \longleftarrow \ldots \\
& \uparrow & & \uparrow & \\
0 \longleftarrow & S' \boxtimes \left(\bigoplus_{k=0}^{q} \mathcal{O}|_{U_k}\right) & \longleftarrow & S'(-1)^{n+q} \boxtimes \left(\bigoplus_{k=0}^{q} \mathcal{O}(-1)|_{U_k}\right) & \longleftarrow \ldots \\
& \uparrow & & \uparrow & \\
& 0 & & 0 &
\end{array}
$$

We may now compute $R\pi_{1*}\mathbf{K}$ by applying π_{1*} to this double complex $C^{\bullet,\bullet}$ and running the vertical spectral sequence for the resulting double complex of S'-modules. After taking vertical homology of $C^{\bullet,\bullet}$, we obtain the ${}_vE_1^{\bullet,\bullet}$-page with differential $\partial_1^{\bullet,\bullet}$:

$$
\begin{array}{llcll}
{}_vE_1^{\bullet,\bullet}: & & \vdots & & \vdots \\
0 \longleftarrow & S' \otimes H^1(\mathbb{P}^q, \mathcal{O}) & \xleftarrow{\partial_1^{-1,1}} & S'(-1)^{n+q} \otimes H^1(\mathbb{P}^q, \mathcal{O}(-1)) & \longleftarrow \ldots \\
0 \longleftarrow & S' \otimes H^0(\mathbb{P}^q, \mathcal{O}) & \xleftarrow{\partial_1^{-1,0}} & S'(-1)^{n+q} \otimes H^0(\mathbb{P}^q, \mathcal{O}(-1)) & \longleftarrow \ldots \\
& 0 & & 0 &
\end{array}
$$

The general entry on the ${}_vE_1$-page is given by

$${}_vE_1^{-i,j} = S'(-i)^{\binom{n+q}{i}} \otimes H^j(\mathbb{P}^q, \mathcal{O}(-i)).$$

Since $H^j(\mathbb{P}^q, \mathcal{O}(-i)) = 0$ unless $j = 0$ or q, most of these entries of ${}_vE_1$ are equal to 0. In fact, ${}_vE_1$ has a single nonzero entry on row 0, with the only remaining nonzero entries appearing on row q, as shown below:

$$\begin{array}{lccccccccccccc}
{}_vE_1^{\bullet,\bullet}: & 0 & \longleftarrow & 0 & \longleftarrow & 0 & \leftarrow & \ldots & \leftarrow & S'(-q-1)^{\binom{n+q}{q+1}} \otimes H^q(\mathcal{O}(-q-1)) & \xleftarrow{\partial_1^{-q-2,q}} & \ldots \\
 & 0 & \longleftarrow & 0 & \longleftarrow & 0 & \leftarrow & \ldots & \longleftarrow & 0 & \longleftarrow & \ldots \\
 & & & \vdots & & \vdots & & & & \vdots & & \\
 & 0 & \longleftarrow & 0 & \longleftarrow & 0 & \leftarrow & \ldots & \longleftarrow & 0 & \longleftarrow & \ldots \\
 & 0 & \leftarrow & S' \otimes H^0(\mathcal{O}) & \leftarrow & 0 & \leftarrow & \ldots & \longleftarrow & 0 & \longleftarrow & \ldots
\end{array}$$

Since all of the terms of the ${}_vE_1$ page lie in total degree $-i + j \leq 0$, we see that $R^\ell \pi_{1*}\mathbf{K} \neq 0$ only if $\ell \leq 0$, as claimed.

Note that after passing the ${}_vE_1$-page, the only other differential exiting or entering a nonzero term will occur on ${}_vE_{q+1}$, from position $(-i, j) = (-q-1, q)$ to $(-i, j) = (0, 0)$. Since this spectral sequence satisfies ${}_vE_1^{-i,j} \Rightarrow R^{-i+j}\pi_{1*}\mathbf{K}$ and our previous computation shows that $R^\ell \pi_{1*}\mathbf{K} \neq 0$ if only if $\ell = 0$, only positions $(0, 0)$ and $(-q-1, q)$ may contain nonzero entries on the ${}_vE_2$-page. In addition, since $R^{-1}\pi_{1*}\mathbf{K} = 0$, we see that ${}_vE_\infty^{-q-1,q} = 0$. Hence the differential $\partial_{q+1}^{-q-1,q}$ must be injective:

$$\begin{array}{lcccccc}
{}_vE_{q+1}^{\bullet,\bullet}: & 0 & 0 & 0 & \ldots & \operatorname{coker}\partial_1^{-q-2,q} & 0 \\
 & & \vdots & \vdots & & \vdots & \\
 & 0 & S' \otimes H^0(\mathcal{O}) & 0 & \ldots & 0 & \ldots
\end{array}$$

$$\partial_{q+1}^{-q-1,q} : \operatorname{coker}\partial_1^{-q-2,q} \hookrightarrow S' \otimes H^0(\mathcal{O})$$

The differential $\partial_{q+1}^{-q-1,q}$ lifts to a map ϕ of the free modules on the ${}_vE_1$ page:

$$\begin{array}{ccc} S' \otimes H^0(\mathcal{O}) & \overset{\phi}{\dashleftarrow} & S'(-q-1)^{\binom{n+q}{q+1}} \otimes H^q(\mathcal{O}(-q-1)) \\ \uparrow = & & \downarrow \\ S' \otimes H^0(\mathcal{O}) & \xleftarrow{\partial_{q+1}^{-q-1,q}} & \operatorname{coker}\partial_1^{-q-2,q}. \end{array}$$

We thus conclude that $R^0\pi_{1*}\mathbf{K} = \pi_{1*}\mathcal{O}_Z$ is represented by a minimal complex of the form

$$S' \xleftarrow{\phi} S'(-q-1)^{\binom{n+q}{q+1}} \otimes H^q(\mathcal{O}(-q-1)) \longleftarrow S'(-q-2)^{\binom{n+q}{q+2}} \otimes H^q(\mathcal{O}(-q-2)) \longleftarrow \cdots .$$

Notice this is a pure complex of type $(0, q+1, \ldots, q+n)$. Since it is acyclic, it is actually a resolution of the S'-module $\pi_{1*}\mathcal{O}_Z$. Hence this module has projective dimension n, and since we noted initially that it has codimension at least n, it follows that $\pi_{1*}\mathcal{O}_Z$ is a Cohen–Macaulay module of codimension n. □

The following proposition, due to Eisenbud and Schreyer, provides a more general framework than Lemma 4 for collapsing terms from a resolution. The proof is nearly identical. See Fig. 2 for an illustration of this result.

Proposition 5 ([11, Proposition 5.3]). *Let $\mathcal{F}$ be a sheaf on $X \times \mathbb{P}^m$ that has a resolution $\mathbf{G}$ arising from $\mathcal{O}_X$-modules $\mathcal{G}_i$, such that*

$$\mathbf{G}_i = \mathcal{G}_i \boxtimes \mathcal{O}(-e_i) \text{ for } 0 \le i \le N$$

and $e_0 < \cdots < e_N$. If this sequence contains the subsequence $(e_{k+1}, \ldots, e_{k+m}) = (1, 2, \ldots, m)$ for some $k \ge -1$, then

$$R^\ell\pi_{1*}\mathcal{F} \cong R^\ell\pi_{1*}\mathbf{G} = 0 \quad \text{for } \ell \ne 0,$$

and $\pi_{1*}\mathcal{F}$ has a resolution $\mathbf{G}'$, where

$$\mathbf{G}'_i = \begin{cases} \mathcal{G}_i \otimes H^0(\mathbb{P}^m, \mathcal{O}(-e_i)) & \text{for } 0 \le i \le k, \\ \mathcal{G}_{i+m} \otimes H^m(\mathbb{P}^m, \mathcal{O}(-e_{i+m})) & \text{for } k+1 \le i \le N-m. \end{cases}$$

Proof. We proceed in a manner similar to the proof of Lemma 4. Our first goal is to show in two steps that $R^\ell p_*\pi_{1*}\mathbf{G} \ne 0$ if and only if $\ell = 0$. First, since $\mathbf{G}$ is a resolution of $\mathcal{F}$, it follows that $R^\ell p_*\pi_{1*}\mathbf{K} \ne 0$ only if $\ell \ge 0$.

We now compute $R\pi_{1*}(\mathbf{G})$ in a second way to show that $R^\ell\pi_{1*}\mathbf{G} \ne 0$ only if $\ell \le 0$. For each i, let $C^{-i,\bullet}$ be the Čech resolution of $\mathbf{G}_i$ with respect to the

Fig. 2 Proposition 5 uses a pushforward and the vanishing cohomology of line bundles on $\mathbb{P}^m$ to collapse terms from a free resolution. The above illustrates the proposition when $m = k = 2$ and $N = 8$

The complex $\mathbf{G}$

	X	$\times \mathbb{P}^2$
$\mathbf{G}_0$	$\mathcal{G}_0$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_0)$
$\mathbf{G}_1$	$\mathcal{G}_1$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_1)$
$\mathbf{G}_2$	$\mathcal{G}_2$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_2)$
$\mathbf{G}_3$	$\mathcal{G}_3$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-1)$
$\mathbf{G}_4$	$\mathcal{G}_4$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-2)$
$\mathbf{G}_5$	$\mathcal{G}_5$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_5)$
$\mathbf{G}_6$	$\mathcal{G}_6$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_6)$
$\mathbf{G}_7$	$\mathcal{G}_7$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_7)$
$\mathbf{G}_8$	$\mathcal{G}_8$	$\boxtimes \mathcal{O}_{\mathbb{P}^2}(-e_8)$

$\xrightarrow{Rp_* R\pi_{1*}}$

The complex $\mathbf{G}'$

	X
$\mathbf{G}'_0$	$\mathcal{G}_0 \otimes H^0(\mathcal{O}_{\mathbb{P}^2}(-e_0))$
$\mathbf{G}'_1$	$\mathcal{G}_1 \otimes H^0(\mathcal{O}_{\mathbb{P}^2}(-e_1))$
$\mathbf{G}'_2$	$\mathcal{G}_2 \otimes H^0(\mathcal{O}_{\mathbb{P}^2}(-e_2))$
-	-
-	-
$\mathbf{G}'_3$	$\mathcal{G}_3 \otimes H^2(\mathcal{O}_{\mathbb{P}^2}(-e_5))$
$\mathbf{G}'_4$	$\mathcal{G}_4 \otimes H^2(\mathcal{O}_{\mathbb{P}^2}(-e_6))$
$\mathbf{G}'_5$	$\mathcal{G}_5 \otimes H^2(\mathcal{O}_{\mathbb{P}^2}(-e_7))$
$\mathbf{G}'_6$	$\mathcal{G}_6 \otimes H^2(\mathcal{O}_{\mathbb{P}^2}(-e_8))$

standard Čech cover $\{X \times U_0, \ldots, X \times U_m\}$ of $X \times \mathbb{P}^m$. Since the construction of Čech resolutions is functorial, we obtain a double complex $C^{\bullet,\bullet}$ consisting of π_{1*}-acyclic sheaves on $X \times \mathbb{P}^m$. To compute $R\pi_{1*}\mathbf{G}$, we apply π_{1*} to the double complex $C^{\bullet,\bullet}$ and run the vertical spectral sequence for the resulting double complex of $\mathcal{O}_X$-modules. This yields an ${}_vE_1$-page with general entry

$$ {}_vE_1^{-i,j} = \mathcal{G}_i \otimes H^j(\mathbb{P}^m, \mathcal{O}(-e_i)). $$

Since $H^j(\mathbb{P}^m, \mathcal{O}(-e_i)) = 0$ unless $j = 0$ or m, most of these entries are equal to 0. In fact, the resulting ${}_vE_1$-page consists of a strand of nonzero entries in row 0, followed by all zeroes in columns $k+1, \ldots, k+m$, followed by a strand of nonzero entries in row m:

$$
\begin{array}{ccccccccccccc}
{}_vE_1^{\bullet,\bullet}: & & \underline{-i=k} & & \underline{k+1} & & & & \underline{k+m} & & \underline{k+m+1} & & \\
\cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots & \longleftarrow & 0 & \longleftarrow & \mathcal{G}_{k+m+1} \otimes H^m(\mathcal{O}(-e_{k+m+1})) & \xleftarrow{\partial_1^{-k-m-2,m}} & \cdots \\
\cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots \\
 & & \vdots & & \vdots & & & & \vdots & & \vdots & & \\
\cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots \\
\cdots & \longleftarrow & \mathcal{G}_k \otimes H^0(\mathcal{O}(-e_k)) & \longleftarrow & 0 & \longleftarrow & \cdots & \longleftarrow & 0 & \longleftarrow & 0 & \longleftarrow & \cdots
\end{array}
$$

Since all of the nonzero terms of this ${}_vE_1$-page lie in total cohomological degree $-i + j \leq 0$, we see that $R^\ell \pi_{1*}\mathbf{G} \neq 0$ only if $\ell \leq 0$, as desired.

We have now nearly constructed our complex $\mathbf{G}'$. The nonzero entries on the ${}_vE_1$-page are precisely the terms we use in $\mathbf{G}'$, and as its differential, we will use ∂_1 everywhere except for the map $\mathbf{G}'_k \longleftarrow \mathbf{G}'_{k+1}$:

$$ \mathbf{G}'_0 \xleftarrow{\partial_1^{-1,0}} \mathbf{G}'_1 \longleftarrow \cdots \xleftarrow{\partial_1^{-k,0}} \mathbf{G}'_k \xleftarrow{???} \mathbf{G}'_{k+1} \xleftarrow{\partial_1^{-k-m-2,m}} \mathbf{G}'_{k+2} \longleftarrow \cdots \xleftarrow{\partial_1^{-N,m}} \mathbf{G}'_{N-m} \longleftarrow 0. $$

To complete the construction of $\mathbf{G}'$ and to check its exactness, we note that after the ${}_vE_2$-page, the only other differential exiting or entering a nonzero term will occur on the ${}_vE_{m+1}$-page, from position $(-i, j) = (-k-m-1, m)$ to $(-i, j) = (-k, 0)$. Since we have ${}_vE_1^{-i,j} \Rightarrow R^{-i+j}\pi_{1*}\mathbf{G}$ and our previous computation shows that $R^\ell\pi_{1*}\mathbf{G} \neq 0$ if and only if $\ell = 0$, only positions $(-k, 0)$ and $(-k-m-1, m)$ may contain nonzero entries on the ${}_vE_2$-page. In particular, although we have not yet fully constructed the differential for $\mathbf{G}'$, we already see that our complex is exact in every position except possibly at $\mathbf{G}'_k$ or $\mathbf{G}'_{k+1}$.

We now examine the differential $\partial_{m+1}^{-k-m-1,m}$ on ${}_vE_{m+1}$. This differential must be an isomorphism when $k > 0$, as otherwise $R^{k+1}\pi_{1*}\mathbf{G}$ and $R^k\pi_{1*}\mathbf{G}$ would be nonzero. When $k = 0$, it must be injective for the same reason:

$$
\begin{array}{lcccccc}
{}_vE_{m+1}^{\bullet,\bullet}: & 0 & 0 & 0 & \dots & \operatorname{coker}\partial_1^{-k-m-2,m} & 0 \\
& & \vdots & \vdots & & \vdots & \\
& 0 & \ker\partial_1^{-k,1} & 0 & \dots & 0 & \dots
\end{array}
\qquad \operatorname{coker}\partial_1^{-k-m-2,m} \xrightarrow{\partial_{m+1}^{-k-m-1,m}} \ker\partial_1^{-k,1}
$$

This differential $\partial_{m+1}^{-k-m-1,m}$ lifts to a map $\phi\colon \mathbf{G}'_{k+1} \longrightarrow \mathbf{G}'_k$,

$$
\begin{array}{ccc}
\mathbf{G}'_k = \mathcal{G}_k \otimes H^0(\mathcal{O}(-e_k)) & \xleftarrow{\ \phi\ } & \mathbf{G}'_{k+1} = \mathcal{G}_k \otimes H^m(\mathcal{O}(-e_m)) \\
\uparrow & & \downarrow \\
\ker\partial_1^{-k,1} & \xleftarrow{\ \partial_{m+1}^{-k-m-1,m}\ } & \operatorname{coker}\partial_1^{-k-m-2,m}
\end{array}
$$

completing our construction of $\mathbf{G}'$:

$$
\mathbf{G}'_0 \xleftarrow{\partial_1^{-1,0}} \mathbf{G}'_1 \longleftarrow \dots \xleftarrow{\partial_1^{-k,0}} \mathbf{G}'_k \xleftarrow{\ \phi\ } \mathbf{G}'_{k+1} \xleftarrow{\partial_1^{-k-m-2,m}} \mathbf{G}'_{k+2}
$$

$$
\longleftarrow \dots \xleftarrow{\partial_1^{-N,m}} \mathbf{G}'_{N-m} \longleftarrow 0.
$$

It follows that $\mathbf{G}'$ is exact at $\mathbf{G}'_k$ and at $\mathbf{G}'_{k+1}$. Since $\mathbf{G}'$ is acyclic, it follows that it is a resolution $\pi_{1*}\mathcal{F}$, as desired. □

Proposition 5 provides a tool to construct a pure free resolution with a prescribed degree sequence. We illustrate this by explaining how to construct a pure resolution of type $d = (0, 3, 5, 6)$ over $S = \Bbbk[x_1, x_2, x_3]$ (see Fig. 3). Since the highest degree term has degree 6, we define the ring $S' := \Bbbk[y_1, \dots, y_6]$ and consider a Koszul complex involving 6 multilinear forms. The gaps in the degree sequence d tell us how to choose the projective spaces we use to collapse the various terms. For instance, this degree sequence has two gaps: the gap between 0 and 3 consisting of the integers $\{1, 2\}$ and the gaps between 3 and 5 consisting of $\{4\}$. To collapse degrees 1 and 2, we will use a copy of $\mathbb{P}^2$; to collapse degree 4, we will use a copy of $\mathbb{P}^1$.

We thus define a Koszul complex $\mathbf{K}$ involving 6 multidegree $(1, 1, 1)$-forms on $\mathrm{Spec}(S') \times \mathbb{P}^2 \times \mathbb{P}^1$, and we set $\mathbf{G} := \mathbf{K} \otimes_{\mathcal{O}_{\mathrm{Spec}(S') \times \mathbb{P}^2 \times \mathbb{P}^1}} (\mathcal{O}_{\mathrm{Spec}\, S'} \boxtimes \mathcal{O}_{\mathbb{P}^2} \boxtimes \mathcal{O}_{\mathbb{P}^1}(3))$. This twist of the Koszul complex is engineered so that we are able collapse the proper terms, as shown in Fig. 3. Put another way, we have attached a line bundle with vanishing cohomology to each of the terms in $\mathbf{G}$ that we want to collapse. By applying Proposition 5 twice to $\mathbf{G}$, we obtain a pure resolution of type $(0, 3, 5, 6)$ on $\mathrm{Spec}\, S'$ that resolves a Cohen–Macaulay module of codimension 3. Finally, we mod out by 3 generic linear forms to obtain a pure resolution $\mathbf{F}$ of type $(0, 3, 5, 6)$ on $\mathrm{Spec}(S)$ that resolves a module of finite length:

$$\mathbf{F} = \left[S^4 \leftarrow S(-3)^{20} \leftarrow S(-5)^{36} \leftarrow S(-6)^{20} \leftarrow 0 \right].$$

Proof of Theorem 3. Without loss of generality, we may assume that $d_0 = 0$. We define $S' = \Bbbk[y_1, \dots, y_{d_n}]$. It suffices to construct a Cohen–Macaulay S'-module of codimension n with a pure resolution of type d, as we may then mod out by generic linear forms to obtain a pure resolution of a finite length S-module.

We define an auxiliary space $\mathbb{P}$ which is a product of projective spaces corresponding to the gaps in the degree sequence $d = (d_0 < d_1 < \cdots < d_n)$. To record these gaps, set

$$m_i := d_i - d_{i-1} - 1 \quad \text{for } 1 \le i \le n.$$

Set $\mathbb{P} := \mathbb{P}^{m_1} \times \cdots \times \mathbb{P}^{m_n}$, which has dimension $d_n - n$. Choose d_n generic multilinear forms of multidegree $(1, 1, \dots, 1)$. Since $\Bbbk$ is an infinite field, these forms have no common zeroes in $\mathbb{P}$.[2] Let $\mathbf{K}$ denote the Koszul complex on these multilinear forms, and define

$$\mathbf{G} := \mathbf{K} \otimes (\mathcal{O}_{\mathrm{Spec}(S')} \boxtimes \mathcal{O}_{\mathbb{P}^{m_1}} \boxtimes \mathcal{O}_{\mathbb{P}^{m_2}}(-d_1) \boxtimes \cdots \boxtimes \mathcal{O}_{\mathbb{P}^{m_n}}(-d_{n-1})).$$

Note that $\mathbf{G}$ is an exact complex with

$$\mathbf{G}_i = S'(-i)^{\binom{d_n}{i}} \boxtimes \mathcal{O}_{\mathbb{P}^{m_1}}(-i) \boxtimes \mathcal{O}_{\mathbb{P}^{m_2}}(-d_1 - i) \boxtimes \cdots \boxtimes \mathcal{O}_{\mathbb{P}^{m_n}}(-d_{n-1} - i).$$

By repeatedly applying Proposition 5 (the order in which we pushforward does not matter), all terms from $\mathbf{G}$ will eventually be collapsed away with exception of

[2] In fact, this is also true over a finite field by [11, Proposition 5.2].

Fig. 3 We iterate Proposition 5 to build a pure resolution **F** of type (0, 3, 5, 6) over S'. Modding out by linear forms yields a resolution over S

The original complex $\mathbf{G}$

	Spec(S')	$\times\ \mathbb{P}^2$	$\times \mathbb{P}^1$
$\mathbf{G}_0$	$(S')^1$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^2}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(3)$
$\mathbf{G}_1$	$S'(-1)^6$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^2}(-1)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(2)$
$\mathbf{G}_2$	$S'(-2)^{15}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-2)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(1)$
$\mathbf{G}_3$	$S'(-3)^{20}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-3)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}$
$\mathbf{G}_4$	$S'(-4)^{15}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-4)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(-1)$
$\mathbf{G}_5$	$S'(-5)^6$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-5)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(-2)$
$\mathbf{G}_6$	$S'(-6)^1$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-6)$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^1}(-3)$

$$R(\pi_1 \times \pi_2)_* \downarrow$$

The complex $\mathbf{G}'$after one projection

	Spec(S')	$\times\ \mathbb{P}^1$
$\mathbf{G}'_0$	$(S')^1$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^2} \otimes H^0(\mathcal{O}_{\mathbb{P}^1}(3))$
$\mathbf{G}'_1$	$S'(-1)^6$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^2}(-1) \otimes H^0(\mathcal{O}_{\mathbb{P}^1}(2))$
$\mathbf{G}'_2$	$S'(-2)^{15}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-2) \otimes H^0(\mathcal{O}_{\mathbb{P}^1}(1))$
$\mathbf{G}'_3$	$S'(-3)^{20}$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-3) \otimes H^0(\mathcal{O}_{\mathbb{P}^1})$
-	-	-
$\mathbf{G}'_4$	$S'(-5)^6$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-5) \otimes H^1(\mathcal{O}_{\mathbb{P}^1}(-2))$
$\mathbf{G}'_5$	$S'(-6)^1$	$\boxtimes\, \mathcal{O}_{\mathbb{P}^3}(-6) \otimes H^1(\mathcal{O}_{\mathbb{P}^1}(-3))$

$$R\pi_{2*} R\pi_{1*} \downarrow$$

The pure resolution $\mathbf{F}$

	Spec(S')
$\mathbf{F}_0$	$(S')^1 \otimes H^0(\mathcal{O}_{\mathbb{P}^2}) \otimes H^0(\mathcal{O}_{\mathbb{P}^1}(3))$
-	-
-	-
$\mathbf{F}_1$	$S'(-3)^{20} \otimes H^2(\mathcal{O}_{\mathbb{P}^3}(-3)) \otimes H^0(\mathcal{O}_{\mathbb{P}^1})$
-	-
$\mathbf{F}_2$	$S'(-5)^6 \otimes H^2(\mathcal{O}_{\mathbb{P}^3}(-5)) \otimes H^1(\mathcal{O}_{\mathbb{P}^1}(-2))$
$\mathbf{F}_3$	$S'(-6)^1 \otimes H^2(\mathcal{O}_{\mathbb{P}^3}(-6)) \otimes H^1(\mathcal{O}_{\mathbb{P}^1}(-3))$

$\mathbf{G}_{d_i}$ for $0 \leq i \leq n$. More precisely, when we push away from $\mathbb{P}^{m_i}$, Proposition 5 implies that we will collapse away the terms that originally corresponded to $\mathbf{G}_{d_i+1}, \ldots \mathbf{G}_{d_{i+1}-1}$.

This process produces a pure resolution $\mathbf{F}$ of graded S'-modules, where

$$\mathbf{F}_k = S'(-d_k)^{\binom{d_n}{k}} \otimes \bigotimes_{i=1}^{k-1} H^0(\mathbb{P}^{m_i}, \mathcal{O}(-d_{i-1}-k)) \otimes \bigotimes_{i=k}^{n} H^{m_i}(\mathbb{P}^{m_i}, \mathcal{O}(-d_{i-1}-k)).$$

Since $\mathbf{G}$ resolves a module of codimension d_n and the fibers of the projection $p : X \times \mathbb{P} \longrightarrow X$ have dimension $d_n - n$, it follows that the cokernel of $\mathbf{F}$ has support of codimension at least n. However, since $\mathbf{F}$ is a resolution of projective dimension n, we conclude that the cokernel of $\mathbf{F}$ is a Cohen–Macaulay S'-module of codimension n, as desired. □

If one works with the base scheme Proj(S) instead of Spec(S), then there is a slightly different argument which eliminates the need to pass to the intermediate ring S', but this requires different steps to check exactness. This was Eisenbud and Schreyer's original approach in [11, Sect. 5].

Remark 6. There is a useful shorthand for reverse-engineering the Eisenbud–Schreyer construction of a pure resolution. For instance, to construct a pure resolution of type $(0, 3, 5, 6)$, begin by considering the table on the left, where we have marked with an asterisk the degrees that we need to collapse. We may then use a copy of $\mathbb{P}^2$ to collapse the first two asterisks and a copy of $\mathbb{P}^1$ to collapse the last asterisk. To do so, line up integers as in the middle table so that the vanishing cohomology degrees of $\mathbb{P}^2$ and $\mathbb{P}^1$ align with the asterisks. Now fill in the remaining entries of the table linearly.

Spec(S')	
0	
-1	$*$
-2	$*$
-3	
-4	$*$
-5	
-6	

$\longrightarrow$

Spec(S')	$\mathbb{P}^2$	$\mathbb{P}^1$
0		
-1	-1	
-2	-2	
-3		
-4		-1
-5		
-6		

$\longrightarrow$

Spec(S')	$\mathbb{P}^2$	$\mathbb{P}^1$
0	0	3
-1	-1	2
-2	-2	1
-3	-3	0
-4	-4	-1
-5	-5	-2
-6	-6	-3

The last table tells us that we should build a Koszul complex of six $(1, 1, 1)$-forms on $\text{Spec}(S') \times \mathbb{P}^2 \times \mathbb{P}^1$ and then twist by the degrees we see in the top row: $\mathcal{O}(0, 0, 3)$. Note that this is precisely the construction from Fig. 3.

Although the proof of Theorem 3 is constructive, it does not provide an efficient technique for understanding the differentials of the resulting pure resolution **F**. To obtain explicit formulas for the differentials from the proof, we would have to carry a description of the differential through the spectral sequence.

A more efficient approach to understanding the differentials of these Eisenbud–Schreyer pure resolutions is given in [3, Sect. 4]. That article constructs a generic version of the Eisenbud–Schreyer pure resolution, referred to as a **balanced tensor complex**, which is defined over a polynomial ring in many more variables. The differentials for the tensor complex can be expressed in terms of explicit multilinear constructions (e.g., (co)multiplication maps on symmetric and exterior products). Since the Eisenbud–Schreyer pure resolutions are obtained as specializations of balanced tensor complexes [3, Theorem 10.2], this construction provides closed formulas for the various differentials in the Eisenbud–Schreyer pure resolutions.

Example 7. There is a Macaulay2 package `TensorComplexes` that can be used to compute the Eisenbud–Schreyer pure resolutions explicitly [16]. With $\mathbb{k} = \mathbb{F}_{101}$, the following code computes the first differential for a pure resolution of type $(0, 1, 3, 5)$:

```
i1 : loadPackage "TensorComplexes";
i2 : FF = pureResES({0,1,3,5},ZZ/101);
i3 : betti FF
```

```
             0  1  2 3
o3 = total: 8 15 10 3
         0: 8 15  . .
         1: .  . 10 .
         2: .  .  . 3

i4 : FF.dd_1
o4 = | x_0 0   0   0   x_1 0   0    0   x_2 0   0   0   0   0   0   |
     | 0   x_0 0   0   0   x_1 0    0   0   x_2 0   0   0   0   0   |
     | 0   0   x_0 0   0   0   x_1  0   0   0   x_2 0   0   0   0   |
     | 0   0   0   x_0 0   0   0    x_1 0   0   0   x_2 0   0   0   |
     | 0   0   0   x_2 x_0 0   -x_2 0   x_1 x_2 0   0   0   0   0   |
     | 0   0   0   0   0   x_0 0    0   0   x_1 0   0   x_2 0   0   |
     | 0   0   0   0   0   0   x_0  0   0   0   x_1 0   0   x_2 0   |
     | 0   0   0   0   0   0   0    x_0 0   0   0   x_1 0   0   x_2 |
```

4 Extremal Betti Tables in the Local Case

Let M be a finitely generated module over a regular local ring R of dimension n. From the minimal free resolution of M,

$$0 \leftarrow M \leftarrow R^{\beta_0^R(M)} \leftarrow R^{\beta_1^R(M)} \leftarrow \cdots \leftarrow R^{\beta_n^R(M)} \leftarrow 0,$$

we obtain the **(local) Betti table** $\beta^R(M) = (\beta_0^R(M), \ldots, \beta_n^R(M))$. Here we again restrict our attention to the case when M is of finite length.

As in the graded case, we would like to find modules M of finite length where $\beta^R(M)$ is extremal. However, unlike the graded case, there are no natural candidates for such vectors. It turns out that this is because no local Betti table is extremal.

Claim 4.1. *If* $\dim(R) > 1$*, then there does not exist any R-module M of finite length whose Betti table is extremal.*

Example 2. Let $R = \Bbbk[[x, y]]$. Then the local Betti table of the residue field $\Bbbk$ is $\beta^R(\Bbbk) = (1, 2, 1)$. If $M = R/\langle x^2, xy, y^2\rangle$ and $N = \mathrm{Hom}(M, \Bbbk)$, then we have the decomposition

$$(1, 2, 1) = \beta^R(\Bbbk) = \tfrac{1}{3}\beta^R(M) + \tfrac{1}{3}\beta^R(N) = \tfrac{1}{2}(1, 3, 2) + \tfrac{1}{2}(2, 3, 1).$$

To understand how this comes to pass, we now assume that $n > 1$ and view each $\beta^R(M) \in \mathbb{Q}^{n+1}$. An extremal local Betti table corresponds to a ray of the cone

$$B_{\mathbb{Q}}^{\mathrm{fin}}(R) := \mathbb{Q}_{\geq 0} \cdot \left\{\beta^R(M) \mid M \text{ is an } R\text{-module of finite length}\right\} \subset \mathbb{Q}^{n+1}.$$

Theorem 3 ([4, Theorem 1.1]). *If R is an n-dimensional regular local ring with $n > 1$, then $B_{\mathbb{Q}}^{fin}(R)$ is an open cone that has no extremal rays. More precisely,*

$$B_{\mathbb{Q}}^{fin}(R) = \mathbb{Q}_{>0} \cdot \{\rho_0, \rho_1, \dots, \rho_{n-1}\},$$

where $\rho_i = e_i + e_{i+1}$ *is the sum of the* i*th and* $(i+1)$*st standard basis vectors of* $\mathbb{Q}^{n+1}$.

The story for finitely generated modules is similar; see [4, Sect. 4]. Of course, if $\dim(R) = 1$, then ρ_0 is an extremal ray, as it spans the entire cone.

Proof of Theorem 3. For brevity, set $C := \mathbb{Q}_{>0} \cdot \{\rho_0, \rho_1, \dots, \rho_{n-1}\}$. Clearly the cone C lies in the linear subspace of $\mathbb{Q}^{n+1}$ defined by $\sum_{k=0}^{n}(-1)^k \beta_k^R = 0$. Inside this subspace, an elementary computation confirms that C equals the open cone defined by the inequalities:

$$0 < \sum_{k=i}^{n}(-1)^{i-k}\beta_k^R \quad \text{for } 1 \leq i \leq n.$$

When applied to the Betti numbers of a module M, the above sum is a partial Euler characteristic (computed from the back of the resolution) that computes the rank of the kth syzygy module of M. In particular, each such linear functional is strictly nonnegative when evaluated on the Betti table of a finite length module, and hence we have $B_{\mathbb{Q}}^{\text{fin}}(R) \subseteq C$.

The reverse containment $C \subseteq B_{\mathbb{Q}}^{\text{fin}}(R)$ requires a limiting argument. We show that for each i, there is a sequence of pairs of positive scalars and modules $\{(\lambda_{i,j}, M_{i,j})\}_{j=1}^{\infty}$ such that

$$\rho_i = \lim_{j \longrightarrow \infty} \lambda_{i,j} \beta^R(M_{i,j}).$$

The key fact used in the construction of these R-modules is that there exist local ring analogues to the S-modules with pure resolutions constructed in Theorem 3. Thus, given a degree sequence $d \in \mathbb{Z}^{n+1}$, we may construct an R-module $M(d)$ whose total Betti numbers are computed (up to scalar multiple) by the Herzog–Kühl equations. The precise existence statement for the R-module $M(d)$ is given in Lemma 4 below.

As noted in Remark 2, the Betti table of $M(d)$ is, up to scalar multiple, given by

$$\begin{aligned}\mathfrak{b}(d) &= \prod_{\ell \neq i}|d_\ell - d_i| \left(\frac{1}{\prod_{\ell \neq 0}|d_\ell - d_0|}, \frac{1}{\prod_{\ell \neq 1}|d_\ell - d_1|}, \dots, \frac{1}{\prod_{\ell \neq n}|d_\ell - d_n|} \right) \\ &= \left(\frac{\prod_{\ell \neq i}|d_\ell - d_i|}{\prod_{\ell \neq 0}|d_\ell - d_0|}, \frac{\prod_{\ell \neq i}|d_\ell - d_i|}{\prod_{\ell \neq 1}|d_\ell - d_1|}, \dots, \frac{\prod_{\ell \neq i}|d_\ell - d_i|}{\prod_{\ell \neq n}|d_\ell - d_n|} \right) \in \mathbb{Q}^{n+1}.\end{aligned}$$

Note that $\mathfrak{b}(d)_i = 1$. By carefully choosing degree sequences $d^{i,j}$, we will realize ρ_i as the desired limit using $M_{i,j} = M(d^{i,j})$. To make this choice, set $d^{i,j} := (0, j, 2j, \dots, ij, ij+1, (i+1)j+1, \dots, (n-1)j+1)$, so that

$$d_k^{i,j} = \begin{cases} kj & \text{if } k \le i, \\ (k-1)j + 1 & \text{if } k > i. \end{cases}$$

For these degree sequences, the Herzog–Kühl equations imply that, as $i \longrightarrow \infty$, the ith and $(i+1)$st Betti numbers go to infinity more quickly than the other Betti numbers do. Of course, this limit does not make sense for graded Betti numbers. In the local case, where the Betti numbers are ungraded, we may consider such limits.

We thus set $M_{i,j} := M(d^{i,j})$ and $\lambda_{i,j} := \frac{1}{\beta_i^R(M(d^{i,j}))}$. This yields

$$\lambda_{i,j}\beta^R(M_{i,j}) = \mathfrak{b}(d^{i,j}),$$

since they are equal up to scalar multiple and the ith entry in both vectors is equal to 1.

We now claim that $\lim_{j\longrightarrow\infty} \mathfrak{b}(d^{i,j}) = \rho_i$. By construction, the limit equals 1 in the ith position. Also, each element $\mathfrak{b}(d^{i,j})$ lies in the linear subspace given by $\sum_{k=0}^{n}(-1)^k\beta_k = 0$. Thus it suffices to show that $\lim_{j\longrightarrow\infty} \mathfrak{b}(d^{i,j})_k = 0$ for $k \neq i, i+1$, which we directly compute:

$$\lim_{j\longrightarrow\infty} \mathfrak{b}(d^{i,j})_k = \lim_{j\longrightarrow\infty} \frac{\prod_{\ell\neq j} |d_\ell^{i,j} - d_i^{i,j}|}{\prod_{\ell\neq k} |d_\ell^{i,j} - d_k^{i,j}|} = \lim_{j\longrightarrow\infty} \frac{O(j^{n-1})}{O(j^n)} = 0.$$

Thus $B_{\mathbb{Q}}^{\text{fin}}(R)$ contains points that are arbitrarily close to each ρ_i. Since C equals the interior of the closed cone spanned by the ρ_i, we have shown that $B_{\mathbb{Q}}^{\text{fin}}(R)$ contains C, as desired. □

The following lemma is proven in [4, Proposition 2.1].

Lemma 4. *Let R be an n-dimensional regular local ring, and let $d = (d_0, \ldots, d_n)$ be a degree sequence. If N is the cokernel of the pure resolution of type d constructed in Theorem 3, then there exists a finite length R-module $M(d)$ where $\beta_i^R(M(d)) = \beta_{i,d_i}(N)$.*

5 Extremal Betti Tables in the Multigraded Case

Whereas in the previous section, we considered regular local rings, we now move in the opposite direction by refining the grading on the polynomial ring. As we will see, this greatly increases the complexity of the situation. The results discussed in this section stem from our original work plus extended discussions with Eisenbud and Schreyer.

We restrict attention to the simplest example of a finely graded polynomial ring, namely $T := \Bbbk[x, y]$ with the bigrading $\deg(x) = (1, 0)$ and $\deg(y) = (0, 1)$. We seek T-modules M of finite length such that $\beta^T(M)$ is extremal.

Over S, extremal was synonymous with having a pure resolution, but over T this is not the case. In fact, there cannot exist a finite length module M with a resolution of $\mathbf{F}$ where each $\mathbf{F}_i$ is generated in a single bidegree. This is because T is finely graded, so the cokernel of any map $T^a(-\mu_1, -\mu_2) \longleftarrow T^b(-\lambda_1, -\lambda_2)$ has codimension at most 1.

There are, however, other natural candidates for extremal Betti tables. For instance, in the standard $\mathbb{Z}$-graded case, every pure resolution over $\Bbbk[x, y]$ can be a realized by taking the resolution of a quotient of monomial ideals [6, Remark 3.2]. Since each of these modules is naturally bigraded, we might expect that these provide extremal Betti tables in the bigraded sense as well as in the graded sense. While this is quite often the case (see Example 4), there are many other extremal bigraded Betti tables as well.

To describe a sufficient condition for extremality, we introduce the notion of the **matching graph** $\Gamma(M)$ of a bigraded T-module of finite length. By imposing rather weak conditions on matching graphs, we produce a wide array of bigraded T-modules with extremal Betti tables. This illustrates the additional complexity that arises from refined gradings.

Claim 5.1. *Let M be a bigraded T-module of finite length. If its matching graph $\Gamma(M)$ is $(1,1)$-valent and connected, then $\beta^T(M)$ is extremal.*

For a bigraded T-module M of finite length, let $\mathbf{F}$ be the bigraded minimal free resolution of M. The **matching graph** of M is a graph whose vertices have weights in $\mathbb{Z}$ and whose edges are of two types: x-edges and y-edges. The vertices correspond to the degrees of the generators of the $\mathbf{F}_i$; to a vertex $\alpha \in \mathbb{Z}^2$, we assign the weight $\beta^T_{0,\alpha}(M) + \beta^T_{1,\alpha}(M) + \beta^T_{2,\alpha}(M)$. We then include an x-edge (or y-edge, respectively) between any two vertices with the same x-degree (or y-degree).

If a vertex of $\Gamma(M)$ meets precisely a of the x-edges and precisely b of the y-edges, then we say that this vertex has **valency** (a, b). If all of the vertices of $\Gamma(M)$ have valency (a, b), then we say that $\Gamma(M)$ is an (a, b)-valent graph. In addition, we say that $\Gamma(M)$ is **connected** if the underlying graph (i.e., the graph on the same vertices whose edges are the union of the x-edges and y-edges of $\Gamma(M)$) is connected.

Example 2. Let $M = T/\langle x^2, xy, y^2\rangle$. The minimal free resolution of M has the form

$$T^1 \longleftarrow \begin{matrix} T^1(-2,0) \\ \oplus \\ T^1(-1,-1) \\ \oplus \\ T^1(0,-2) \end{matrix} \longleftarrow \begin{matrix} T^1(-2,-1) \\ \oplus \\ T^1(-1,-2) \end{matrix} \longleftarrow 0.$$

Using the natural embedding of the matching graph $\Gamma(M)$ in the first orthant, $\Gamma(M)$ has x-edges as shown in the figure on the left.

We omit the weights on the vertices, since all weights are 1. The graph $\Gamma(M)$ appears on the right, and is (1, 1)-valent and connected. Hence $\beta^T(M)$ is extremal by Claim 5.1.

Example 3. Let $M = \langle x, y\rangle/\langle x^2, xy^2, y^3\rangle$. Then $\Gamma(M)$ fails to be (1, 1)-valent. In fact, at each vertex of the form $(1, *)$, there are 3 x-edges. In this case, $\beta^T(M)$ equals $\beta^T(\langle x\rangle/\langle x^2, xy^2\rangle) + \beta^T(\langle y\rangle/\langle xy, y^3\rangle)$. These last two Betti tables are extremal by Claim 5.1.

Proof of Claim 5.1. For any $\lambda_1 \in \mathbb{Z}$, we can consider the subgraph of $\Gamma(M)$ obtained by restricting to the vertices of $\Gamma(M)$ whose degrees have the form $(\lambda_1, *)$. By definition of the x-edges, there will be an x-edge between any two vertices of this subgraph. Hence, by the (1, 1)-valency, we see that $\Gamma(M)$ has at most two vertices of the form $(\lambda_1, *)$.

In fact, for each $\lambda_1 \in \mathbb{Z}$, we claim that $\Gamma(M)$ has either zero or two vertices of the form $(\lambda_1, *)$. The bigraded Hilbert series of M is given by the rational function

$$H_M(s_1, s_2) = \frac{K_M(s_1, s_2)}{(1-s_1)(1-s_2)} := \frac{\sum_{i=0}^{2} \sum_{\lambda\in\mathbb{Z}^2} (-1)^i \beta^T_{i,\lambda}(M)\mathbf{s}^\lambda}{(1-s_1)(1-s_2)}.$$

Since M has finite length, $H_M(s_1, s_2)$ is actually a polynomial. This implies that the K-polynomial of M, K_M, is in $\langle 1-s_1\rangle \cap \langle 1-s_2\rangle \subseteq \mathbb{Z}[s_1, s_2]$. We thus have

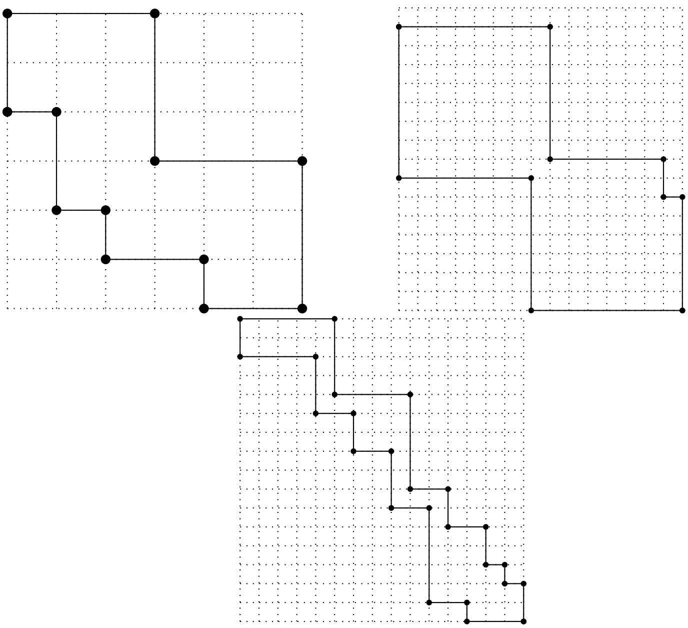
$$K_M(s_1, 1) = \sum_{\lambda_1\in\mathbb{Z}} \left(\sum_{\lambda_2\in\mathbb{Z}} \beta^T_{0,(\lambda_1,\lambda_2)} - \beta^T_{1,(\lambda_1,\lambda_2)} + \beta^T_{2,(\lambda_1,\lambda_2)} \right) s_1^{\lambda_1} = 0.$$

Thus, if $\Gamma(M)$ has a vertex of the form $(\lambda_1, *)$, then it has at least two such vertices of this form. Further, one of these vertices must correspond to a generator of $\mathbf{F}_1$ and the other must correspond to a generator of either $\mathbf{F}_0$ or $\mathbf{F}_2$, and the corresponding Betti numbers be equal. By alternately considering Betti numbers with the same x-degrees and Betti numbers with the same y-degrees, we may show that any two Betti numbers in the same connected component of $\Gamma(M)$ must have the same value. The connectedness of $\Gamma(M)$ implies that each nonzero Betti number of M has the same positive value.

Suppose now that $\beta^T(M) = a'\beta^T(M') + a''\beta^T(M'')$ for some bigraded modules M', M'' of finite length and some $a', a'' \in \mathbb{Q}_{>0}$. We start by considering a bidegree (λ_1, λ_2) where $\beta^T_{0,(\lambda_1,\lambda_2)}(M') = r$. Since $K_{M'}(s_1, 1) = 0$ and $\Gamma(M)$ is (1, 1)-valent, the argument above implies that there is a unique μ_2 such that $\beta^T_{1,(\lambda_1,\mu_2)}(M') \neq 0$, and hence this Betti number must also equal r. We then consider y-degrees, and a similar argument shows that there is a unique μ_1 such that either (but not both) $\beta^T_{0,(\mu_1,\mu_2)}(M') \neq 0$ or $\beta^T_{2,(\mu_1,\mu_2)}(M') \neq 0$. In either case, this Betti number must also equal r.

Continuing to alternate between x-degrees and y-degrees, we eventually form a subcycle of $\Gamma(M)$. However, since $\Gamma(M)$ is $(1, 1)$-valent and connected, this cycle must equal $\Gamma(M)$, so we have shown that $\beta^T(M')$ is simply r times $\beta^T(M)$. □

Example 4. Quotients of monomial ideals provide many examples of extremal bigraded Betti tables. For instance, let $M = I/J$, where $I = \langle x^4, xy^2, x^2y, y^4\rangle$ and $J = \langle x^6, x^3y^3, y^6\rangle$. Then $\Gamma(M)$ is the graph on the left (each vertex has weight 1), which is extremal by Claim 5.1.



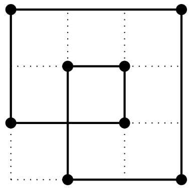
The graphs on the right above correspond to the matching graphs of other quotients of monomial ideals: the upper right graph is the matching graph of $\langle x^7, y^7\rangle/\langle x^{15}, x^{14}y^6, x^8y^8, y^{15}\rangle$. These types of examples may be very far from the pure resolutions we saw in Sect. 3. For instance, we can produce an extremal bigraded Betti table given by a resolution $\mathbf{F}$, where $\mathbf{F}_i$ has minimal generators in arbitrarily many different bidegrees.

Note that these examples are not pure with respect to the $\mathbb{Z}$-grading.

Claim 5.1 begs the question of which $(1, 1)$-valent, connected graphs can be realized as $\Gamma(M)$ for some M. If $\Gamma(M)$ is a $(1, 1)$-valent, connected graph that comes from a quotient of monomial ideals, then it decomposes as the union of two nonintersecting monotonic paths (from the upper left corner to the lower right corner). But the following example illustrates that not all extremal Betti tables arise in this way.

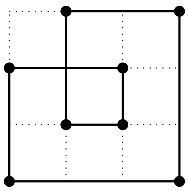
Example 5. The cokernel of the matrix below induces the following matching graph:

$$\begin{array}{c} \\ \binom{1}{0} \\ \binom{0}{1} \end{array} \begin{array}{c} \begin{array}{cccc} \binom{3}{0} & \binom{2}{1} & \binom{1}{2} & \binom{0}{3} \end{array} \\ \begin{pmatrix} x^2 & xy & y^2 & 0 \\ 0 & x^2 & xy & y^2 \end{pmatrix} \end{array}$$



However, not every $(1,1)$-valent connected graph arises as the matching graph of a module.

Example 6. Suppose that the graph on the left below is the matching graph of a module M of finite length. Then the free resolution of M has the form shown on the right.



$$\begin{array}{c} T \\ \oplus \\ T(-1,-1) \end{array} \xleftarrow{\phi} \begin{array}{c} T(-3,0) \\ \oplus \\ T(-2,-1) \\ \oplus \\ T(-1,-3) \\ \oplus \\ T(0,-2) \end{array} \xleftarrow{\psi} \begin{array}{c} T(-2,2) \\ \oplus \\ T(-3,-3) \end{array} \longleftarrow 0$$

In this case, the matrix ϕ would have the form:

$$\phi = \begin{array}{c} \\ \binom{0}{0} \\ \binom{1}{1} \end{array} \begin{array}{c} \begin{array}{cccc} \binom{3}{0} & \binom{2}{1} & \binom{1}{3} & \binom{0}{2} \end{array} \\ \begin{pmatrix} x^3 & a_1x^2y & a_2xy^3 & y^2 \\ 0 & -x & y^2 & 0 \end{pmatrix} \end{array}$$

for some scalars a_1, a_2. After performing an appropriate row operation and column operation, we can assume that $a_1 = a_2 = 0$. However, the kernel of the resulting matrix is generated by $T(-3,-2) \oplus T(-2,-3)$, providing the contradiction.

Though we know of no condition for determining which $(1,1)$-valent, connected graphs arise as $\Gamma(M)$ for some M, Claim 5.1 provides a zoo of extremal rays. If we restrict to Betti tables whose support is contained in the square with corners $(0,0)$ and $(3,3)$, Claim 5.1 produces 74 extremal rays. They are generated by the tables of quotients of monomial ideals, along with the table in Example 5 and its dual. We conclude with a conjecture.

Conjecture 5.7. *All extremal Betti tables of the cone of bigraded T-modules with finite length are generated by Betti tables of modules M that satisfy Claim 5.1.*

Acknowledgements David Eisenbud shaped all aspects of this chapter. He taught many of these topics to us firsthand, patiently explaining the details, and our goal in this chapter is largely to reproduce some of the insights we learned from him. Further, his support and enthusiasm were crucial to our subsequent research in related areas.

We would also like to thank Frank–Olaf Schreyer; many of the ideas in Sect. 5 were refined in conversation with him and David Eisenbud. We came to understand the details of the proof of Eisenbud and Schreyer during a Mathematical Research Communities workshop at Snowbird, Utah, organized by Mike Stillman and Hal Schenck; we thank them and the other participants for their help. We thank the National Science Foundation which supported this AMS MRC workshop through grant DMS-0751449. We thank Steven Sam for helpful ideas. We also thank the anonymous referee for suggestions that improved the chapter.

References

1. Avramov, L.L.: Modules with extremal resolutions. Math. Res. Lett. **3**(3), 319–328 (1996)
2. Avramov, L.L.: Infinite free resolutions. In: Six Lectures On Commutative Algebra (Bellaterra, 1996). Progress in Mathematics, vol. 166, pp. 1–118. Birkhäuser, Basel (1998)
3. Berkesch, C., Erman, D., Kummini, M., Sam, S.V: Tensor complexes: Multilinear free resolutions constructed from higher tensors. J. Eur. Math. Soc. (to appear). `arXiv:1101.4604`
4. Berkesch, C., Erman, D., Kummini, M., Sam, S.V: Shapes of free resolutions over a local ring. Math. Ann. 354 (2012), no. 3, 939–954. (to appear). arXiv:1105.2204
5. Bigatti, A.M.: Upper bounds for the Betti numbers of a given Hilbert function. Comm. Algebra **21**(7), 2317–2334 (1993)
6. Boij, M., Söderberg, J.: Graded Betti numbers of Cohen-Macaulay modules and the multiplicity conjecture. J. Lond. Math. Soc. (2) **78**(1), 85–106 (2008)
7. Boij, M., Söderberg, J.: Betti numbers of graded modules and the multiplicity conjecture in the Non-Cohen–Macaulay case. J. Algebra Number Theory **6**(3), 437–454 (2009)
8. Boij, M., Fløystad, G.: The cone of Betti diagrams of bigraded Artinian modules of codimension two. In: Combinatorial Aspects Of Commutative Algebra And Algebraic Geometry. Abel Symposium, vol. 6, pp. 1–16. Springer, Berlin (2011)
9. Buchsbaum, D.A., Eisenbud, D.: Remarks on ideals and resolutions. In: Symposia Mathematica, Vol. XI (Convegno di Algebra Commutativa, INDAM, Rome, 1971), pp. 193–204. Academic, London (1973)
10. Eisenbud, D., Erman, D.: Categorified duality in Boij–Söderberg theory and invariants of free complexes. `arXiv:1205.0449`
11. Eisenbud, D., Schreyer, F.-O.: Betti numbers of graded modules and cohomology of vector bundles. J. Amer. Math. Soc. **22**(3), 859–888 (2009)
12. Eisenbud, D., Schreyer, F.-O.: Betti numbers of syzygies and cohomology of coherent sheaves. In: Proceedings of the International Congress of Mathematicians. Hyderabad, India (2010)
13. Eisenbud, D., Fløystad, G., Weyman, J.: The existence of equivariant pure free resolutions. Ann. l'institut Fourier **61**(3), 905–926 (2011)
14. Fløystad, G.: The linear space of betti diagrams of multigraded artinian modules. Math. Res. Lett. **17**(5), 943–958 (2010)
15. Fløystad, G.: Boij-Söderberg theory: Introduction and survey. (2011). arXiv:1106.0381
16. Grayson, D.R., Stillman, M.E.: Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/Citing/
17. Hartshorne, R.: Algebraic Geometry. Graduate Texts in Mathematics, vol. 52. Springer, New York (1977)
18. Herzog, J., Kühl, M.: On the Betti numbers of finite pure and linear resolutions. Comm. Algebra **12**(13–14), 1627–1646 (1984)

19. Hulett, H.A.: Maximum Betti numbers of homogeneous ideals with a given Hilbert function. Comm. Algebra **21**(7), 2335–2350 (1993)
20. Iyengar, S., Pardue, K.: Maximal minimal resolutions. J. Reine Angew. Math. **512**, 27–48 (1999)
21. Lascoux, A.: Syzygies des variétés déterminantales. Adv. Math. **30**(3), 202–237 (1978)
22. Pardue, K.: Deformation classes of graded modules and maximal Betti numbers. Illinois J. Math. **40**(4), 564–585 (1996)
23. Peeva, I.: Graded Syzygies. Algebra and Applications, vol. 14. Springer, London (2011)

p^{-1}-Linear Maps in Algebra and Geometry

Manuel Blickle and Karl Schwede

1 Introduction

In this survey we study the basic properties of p^{-1}-linear morphisms between coherent sheaves on a scheme X over a perfect field of positive characteristic p. If $F\colon X \to X$ is the Frobenius morphism (i.e., the pth power map on the structure sheaf) we denote by F_* the restriction functor along F (cf. Sect. 2.2). A p^{-1}-linear map is then an $\mathcal{O}_X$-linear map $\varphi : F_*\mathscr{F} \to \mathscr{G}$ for two $\mathcal{O}_X$-modules $\mathscr{F}$ and $\mathscr{G}$. The name stems from the fact that if we view φ as a map on the underlying sheaves of Abelian groups, φ satisfies the condition $\varphi(r^p f) = r\varphi(f)$ for local sections $r \in \mathcal{O}_X$ and $f \in \mathscr{F}$. In particular, if r has a p^{th} root, then we may write this relation as $\varphi(rf) = r^{p^{-1}}\varphi(f)$.

As an example for a p^{-1}-linear map, we start with a splitting of the Frobenius map, i.e., an $\mathcal{O}_X$-linear map φ: $F_*\mathcal{O}_X \to \mathcal{O}_X$ such that the composition

$$\mathcal{O}_X \xrightarrow{F} F_*\mathcal{O}_X \xrightarrow{\varphi} \mathcal{O}_X$$

is equal to the identity. The mere existence of such a φ has strong implications for the local geometry of X (it is reduced, for example). Furthermore, it immediately implies a highly effective version of Serre vanishing: the higher cohomology of *any* ample line bundle vanishes. In the light of such strong implications, it is somewhat surprising that there are varieties of interest that are Frobenius split.

M. Blickle
Institut für Mathematik, Johannes Gutenberg-Universität Mainz,
55099 Mainz, Germany
e-mail: blicklem@uni-mainz.de

K. Schwede (✉)
Department of Mathematics,
The Pennsylvania State University, University Park, PA, 16802, USA
e-mail: schwede@math.psu.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_5,

For example, regular affine varieties, projective spaces, normal toric varieties, and most prominently flag and Schubert varieties are Frobenius split. And it was precisely for the latter varieties where the above vanishing yields a simple proof of Kempf's vanishing theorem [66]; see also [24]. Frobenius split varieties have been extensively studied [11], and in Sect. 5 we give a detailed account of their theory, explaining some of the more delicate vanishing and extension results, and discussing criteria to decide if a given variety is Frobenius split.

In Sect. 6 we show how some of the results and techniques for Frobenius splittings can be extended to more general contexts (where the variety is not F-split) to derive similar conclusions (vanishing and extension results). For example, a systematic use of certain p^{-1}-linear maps can replace Kodaira and Kawamata–Viehweg vanishing theorems [47, 94] in some applications (see Sect. 6.1). These techniques rely on an explicit connection between p^{-e}-linear maps $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathcal{O}_X)$ and $\mathbb{Q}$-divisors Δ such that $\mathcal{O}_X((p^e - 1)(K_X + \Delta)) \cong \mathscr{L}^{-1}$ which is explained in detail in Sect. 4. Indeed, this correspondence between p^{-e}-maps and $\mathbb{Q}$-divisors pervades much of this chapter. This correspondence also provides us with valuable geometric intuition in working with p^{-e}-linear maps.

In Sect. 7 we state a number of general results on the behavior of p^{-e}-linear maps under certain functorial operations, such as pullback along closed immersions, localization, pushforward along a birational map, and finally pullback along a finite map. In all these cases, viewing p^{-e}-linear maps as $\mathbb{Q}$-divisors and performing operations on divisors is the guiding principle.

A second key example of a p^{-1}-linear map is the classical Cartier operator $C\colon F_*\omega_X \to \omega_X$ introduced in [13]. There are various guises in which this operator on the dualizing sheaf appears, but most generally one may view it as the *trace of Frobenius* under the duality for finite morphisms (see Sect. 3.1). The Cartier operator has been extensively studied in connection to residues of differentials in positive characteristic and plays a crucial role in Deligne and Illusie's [16] algebraic proof of Kodaira vanishing.

In the final two Sects. 8 and 9 we describe the category of Cartier modules introduced in [6]. This category consists of coherent $\mathcal{O}_X$-modules $\mathscr{F}$ equipped with a p^{-e}-linear endomorphism, i.e., a $\mathcal{O}_X$-linear map $F^e_*\mathscr{F} \to \mathscr{F}$. We show that the Abelian category of Cartier modules satisfies some remarkable properties. Most importantly, Cartier modules have finite length up to nilpotence.[1] Furthermore, Cartier modules are related to a number of other categories which have been extensively used in the study of local cohomology in positive characteristic. Hence the finiteness results about Cartier modules imply and generalize previous finiteness results about local cohomology, see Sect. 9.1, where we indicate how results of Hartshorne–Speiser [37], Lyubeznik [59], and Enescu and Hochster [18] can be derived easily.

In the final section we explain a certain degree-reducing property of p^e-linear maps and show how this property yields a completely elementary approach to the

[1] A coherent Cartier module $\mathscr{F}$ if nilpotent is some power of the structural map is zero.

above mentioned finiteness result. In the last section we finally close the gap to the theory of tight closure [38, 41, 42], which im- and explicitly heavily relied on p^{-e}-linear maps since its beginnings, in showing how the test ideals of Hara and Yoshida [29] are obtained from certain generalizations of Cartier modules. We include as another demonstration of the utility of this viewpoint a quick proof of the discreteness of the jumping numbers of the test ideal.

The target audience for this chapter is a researcher or student who is familiar with commutative algebra and algebraic geometry and who wishes to learn how to use p^{-1}-linear maps in a wide variety of contexts. We do not assume the reader has one particular background (i.e., representation theory/Frobenius splitting, tight closure theory, $\mathcal{D}$-modules, or higher dimensional complex algebraic geometry). Because we view this chapter as a place where material can be learned, at the end of each section, there are many exercises. The more difficult exercises are decorated with a *. The exercises are a fundamental part of this document.

2 Preliminaries on Frobenius

In this section we introduce our conventions on notation—in particular with regards to the Frobenius morphism.

2.1 Prerequisites and Notation

We assume that the reader is familiar with the basics of commutative algebra and algebraic geometry, all of which is covered in the standard reference works [34] and [64]. Beyond this, a familiarity with Grothendieck duality, [14, 32], will be particularly helpful. Explicitly, Serre vanishing, canonical modules, dualizing and Serre duality, and the connection between divisors and line bundles will appear frequently (see also [12]). The notion of $\mathbb{Q}$-divisors will be used extensively (see [50] or [55, 56]). The process of reflexification of sheaves on normal varieties and its relation to Weil divisors will be recalled in Appendix A for the convenience of the reader; also see [36] where the same theory is worked out in substantially greater generality.

Throughout this chapter all rings and schemes are assumed to be of finite type over a perfect field k of characteristic $p > 0$, or they are a localization or completion of such at a prime. This implies that our schemes are excellent and possess canonical modules and dualizing complexes [32, 64]. We further assume that all schemes are separated.

2.2 Frobenius and Pushforward

We begin by reviewing the most basic notation (since it varies wildly in the literature).

The key structure in algebra and geometry over a field of positive characteristic $p > 0$ is the *(absolute) Frobenius* endomorphism. For a ring R this is just the pth power ring endomorphism

$$F = F_R \colon R \to R$$

given by sending $r \in R$ to r^p.

Since the Frobenius is canonical it induces a morphism for any scheme X over a field k of characteristic $p > 0$, also called the Frobenius endomorphism and also denoted by

$$F = F_X \colon X \to X.$$

Supposing that k is perfect and X is a k-variety (or a scheme according to our convention) then F_X is a finite map[2] by Exercise 2. Note that F_X is in general not a morphism of k-schemes—however this point can be rectified by changing the k-structure on the first copy of X, if desired. We denote by F^e the e-fold self composition of Frobenius.

Even in the affine situation $X = \operatorname{Spec} R$ we use geometric notation and denote the Frobenius on R by $F \colon R \to F_* R$ to remind us that it is not R-linear. This has the added benefit that we now can distinguish the source and target of $F = F_R$.

Given an ideal $I = \langle f_1, \dots, f_m \rangle \subseteq R$, we define its *$p^e$th Frobenius power* to be $I^{[p^e]} = \langle f_1^{p^e}, \dots, f_m^{p^e} \rangle$. This is independent of the choice of generators f_j (see Exercise 3). The formation of $I^{[p^e]}$ commutes with localization, and so for any ideal sheaf $\mathscr{I} \subseteq \mathcal{O}_X$, we can define $\mathscr{I}^{[p^e]}$ in the obvious way.

Note $F^e_* \mathcal{O}_X$ is isomorphic to $\mathcal{O}_X$ as a sheaf of rings—but as $\mathcal{O}_X$-modules they are distinct: namely, $\mathcal{O}_X$ acts on $F^e_* \mathcal{O}_X$ via p^eth powers. More generally, for any $\mathcal{O}_X$-module $\mathscr{M}$, one observes that $F^e_* \mathscr{M}$ is isomorphic to $\mathscr{M}$ as a sheaf of Abelian groups, but the $\mathcal{O}_X$-module structure is given by $r.m = r^{p^e} m$ for a local section $r \in \mathcal{O}_X$ and $m \in F^e_* \mathscr{M}$. Of course, $F^e_* \mathscr{M}$ also has an $F^e_* \mathcal{O}_X$-module structure, which coincides with $\mathscr{M}$'s original $\mathcal{O}_X$-module structure. We also use the notation $F^e_* M$ in the affine case $X = \operatorname{Spec} R$ to denote an R-module with the twisted (restriction of scalars) Frobenius structure.

One immediately verifies that $F_* \widetilde{M}$ coincides with $\widetilde{F_* M}$ as $\mathcal{O}_X$-modules, where $\widetilde{M}$ denotes the $\mathcal{O}_X$-module associated to the R-module M. However, we caution the reader that the same identification does not hold in the graded case with respect to Proj; see, for example, [78, Lemma 5.6] and Exercise 7.

Notation 2.2.1. Given an element $m \in M$, we will sometimes use $F^e_* m$ to denote the corresponding element of $F^e_* M$. Likewise, for sheaves of $\mathcal{O}_X$-modules $\mathscr{M}$ on X.

[2] An abstract scheme with a finite Frobenius is called *F-finite*.

2.3 Frobenius Pullback and the Projection Formula

Let X be a scheme over a perfect field k of characteristic $p > 0$, and let $\mathscr{F}$ be a coherent sheaf and $\mathscr{L}$ a line bundle on X. Since the Frobenius is an isomorphism on the underlying topological space, the pullback $F^{e*}\mathscr{F}$ (as an $\mathcal{O}_X$-module) can be identified with $\mathscr{F} \otimes_{\mathcal{O}_X} F^e_*\mathcal{O}_X$ as an $F^e_*\mathcal{O}_X$-module, again using that $F^e_*\mathcal{O}_X$ is isomorphic with $\mathcal{O}_X$ as sheaves of rings. If the line bundle $\mathscr{L}$ is given by the datum of a local trivialization and transition functions, then the line bundle $F^{e*}\mathscr{L}$ is given by the p^eth powers of the transition functions in that datum for $\mathscr{L}$. This shows that

$$(F^e)^*\mathscr{L} \cong \mathscr{L}^{p^e}, \tag{1}$$

i.e., the pullback along the Frobenius of a line bundle just raises that line bundle to the p^eth tensor power. Combining this observation with the projection formula [34, Chap. II, Exercise 5.1(d)] we obtain

$$(F^e_*\mathscr{F}) \otimes_{\mathcal{O}_X} \mathscr{L} \cong F^e_*(\mathscr{F} \otimes_{\mathcal{O}_X} F^{e*}\mathscr{L}) \cong F^e_*(\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{L}^{p^e}). \tag{2}$$

This basic equality is used frequently throughout the theory and will be referred to as the *projection formula*.

2.4 Exercises

Exercise 1. Set $X = \operatorname{Spec} k[x_1, \ldots, x_n]$ for some perfect field k. Show that $F^e_*\mathcal{O}_X$ is a free $\mathcal{O}_X$-module with basis $F^e_* x_1^{\lambda_1} \cdots x_n^{\lambda_n}$ where $0 \leq \lambda_i \leq p^e - 1$. Show that the same result also holds for power series $\operatorname{Spec} k[\![x_1, \ldots, x_n]\!]$.

Exercise 2. Suppose that k is a perfect field and that X is scheme of (essentially) finite type over k. Prove that the Frobenius map on X is a finite map.

Exercise 3. Suppose that $I \subseteq R$ is an ideal in a ring R of characteristic $p > 0$. Show that $I^{[p^e]}$ can be identified with $\operatorname{Image}(F^{e*}I \to F^{e*}R) \subseteq F^{e*}R \cong R$ where the last isomorphism is the canonical one identifying R with F^e_*R sending r to F^e_*r. Conclude that $I^{[p^e]}$, the Frobenius power of I, is independent of the choice of generators of I.

Exercise 4. Suppose that X is a smooth d-dimensional variety and $\mathscr{L}$ is a vector bundle of rank m on X. Prove that $F_*\mathscr{L}$ is also a vector bundle and find its rank.

Hint: Complete, use Cohen structure theorem [65, Theorem 28.3], and use Exercise 1.

Exercise 5. Suppose that $\mathscr{E}$ is a locally free sheaf of finite rank on X. Is $\mathscr{E}^{\otimes p^e}$ isomorphic to $(F^e)^*\mathscr{E}$?

Exercise 6. Suppose that R is (essentially) of finite type over a perfect field.

(a) If $W \subseteq R$ is any multiplicatively closed set, then show that $W^{-1}(F^e_* R) \cong F^e_*(W^{-1}R)$. Here the first $F^e_* R$ means as an R-module, and the second is as an $W^{-1}R$-module.
(b) If $\mathfrak{m} \subseteq R$ is a maximal ideal, prove that $F^e_* \widehat{R} \cong \widehat{F^e_* R}$ where $\widehat{\_}$ denotes completion along $\mathfrak{m}$. Again, the first F^e_* is the Frobenius for $\hat{R}$, and the second is that of R-modules.

Exercise 7. Suppose that X is a projective variety with ample line bundle $\mathscr{L}$, and suppose that $\mathscr{F}$ is a coherent sheaf on X. Set $S = \bigoplus_{i \in \mathbb{Z}} H^0(X, \mathscr{L}^i)$ to be the section ring with respect to $\mathscr{L}$, and set $M = \bigoplus_{i \in \mathbb{Z}} H^0(X, \mathscr{F} \otimes \mathscr{L}^i)$ to be the saturated graded S-module corresponding to $\mathscr{F}$. Verify that $F^e_* S$ is a $(\frac{1}{p^e} \cdot \mathbb{Z})$-graded ring,[3] the natural map $S \to F^e_* S$ is graded, and $F^e_* M$ is a graded $F^e_* S$-module. Of course, $F^e_* M$ is also a graded S-module.

Show that $F^e_* M$ is not in general isomorphic to $\bigoplus_{i \in \mathbb{Z}} H^0(X, (F^e_* \mathscr{F}) \otimes \mathscr{L}^i)$. Instead, prove that $\bigoplus_{i \in \mathbb{Z}} H^0(X, (F^e_* \mathscr{F}) \otimes \mathscr{L}^i)$ is isomorphic to a (graded) direct summand of $F^e_* M$, the summand whose terms have integral gradings.

Exercise 8. A ring R (or scheme X) such that the Frobenius map $F : R \to F_* R$ is a finite map is called *F-finite*. Essentially all rings considered in this chapter are F-finite, but not all rings are. Find an example of a field which is not F-finite.

If X is a smooth variety, then we have already seen that $F_* \mathcal{O}_X$ is a locally free (in other words flat) $\mathcal{O}_X$-module. In this exercise, you will prove the converse. First we introduce a definition.

Definition 2.4.1. Suppose that $(R, \mathfrak{m})$ is a local ring. A sequence of elements $f_1, \dots, f_n \in \mathfrak{m} \subseteq R$ is called *Lech-independent* if for any $a_1, \dots, a_n \in R$ such that $a_1 f_1 + \dots + a_n f_n = 0$, then each $a_i \in \langle f_1, \dots, f_n \rangle$.

Now we come to the exercise.

Exercise* 2.9 (Kunz's regularity criterion [53]). Suppose that $(R, \mathfrak{m})$ is a local ring. We will show that if $F_* R$ is flat, then R is regular. We need some lemmas due to Lech [57].

(i) [57, Lemma 3]. If $f_1, \dots, f_n$ are Lech-independent elements and $f_1 \in gR$ for some $g \in R$, then $g, f_2, \dots, f_n$ is also Lech-independent. Furthermore, $\langle f_2, \dots, f_n \rangle : g \subseteq \langle f_1, \dots, f_n \rangle$.
(ii) [57, Lemma 4]. If $f_1, \dots, f_n$ are Lech-independent, $\sqrt{\langle f_1, \dots, f_n \rangle} = \mathfrak{m}$, and $f_1 = gh$. Then

$$l_R(R/\langle f_1, \dots, f_n \rangle) = l_R(R/\langle g, f_2, \dots, f_n \rangle) + l_R(R/\langle h, f_2, \dots, f_n \rangle).$$

[3] Here $\frac{1}{p^e} \cdot \mathbb{Z}$ is the subgroup of $\mathbb{Q}$ generated by $\frac{1}{p^e}$.

(iii) Now we return to the proof of the theorem of Kunz. Show that $\mathfrak{m}^{[p^e]}/(\mathfrak{m}^{[p^e]})^2$ is a free $R/\mathfrak{m}^{[p^e]}$-module. Conclude that if $\mathfrak{m} = \langle x_1, \dots, x_n \rangle$ is generated by a minimal set of generators, then $x_1^{p^e}, \dots, x_n^{p^e}$ is Lech-independent.

(iv) Use the previous parts of the exercise to conclude that $l_R(R/\mathfrak{m}^{[p^e]}) = p^{ne}$.

(v) Reduce to the case that R is complete and write $R = S/\mathfrak{a} = k[[x_1, \dots, x_n]]/\mathfrak{a}$ using the Cohen structure theorem [65, Theorem 28.3] where $k = R/\mathfrak{m}$. Then notice that $l_S(S/\mathfrak{m}_S^{[p^e]}) = p^{ne}$ for all $e \geq 0$. Complete the proof of Kunz' regularity criterion. The proof given in the exercise was Kunz' original proof.

Remark 2.4.2. A simpler proof of Kunz' result using the Buchsbaum–Eisenbud acyclicity criterion can be found on page 12 of [42]. Alberto Fernandez Boix pointed out to us that another short proof can be found in [63, Theorem 4.4.2].

3 p^{-e}-Linear Maps: Definition and Examples

In this section we introduce p^{-e}-linear maps and give a number of examples which will be discussed in more detail throughout the rest of this chapter.

Definition 3.0.3 (p^{-e}-linear map). Suppose that X is a scheme and $\mathscr{M}$ and $\mathscr{N}$ are $\mathcal{O}_X$-modules. A *p^{-e}-linear map* is an additive map $\varphi : \mathscr{M} \to \mathscr{N}$ such that

$$\varphi(r^{p^e} m) = r\varphi(m) \tag{3}$$

for all local sections $r \in \mathcal{O}_X$ and $m \in \mathscr{M}$.

Equivalently, we may specify a *p^{-e}-linear map* by the data of an $\mathcal{O}_X$-linear map

$$\varphi\colon F^e_* \mathscr{M} \to \mathscr{N}\,.$$

We will *frequently and freely* switch between these two points of view, depending on the context.

If R is a ring, then a p^{-e}-linear map between R-modules M and N is simply an additive map between them satisfying the rule from (3). If k is a perfect field, then p^{-e}-linearity for an additive map $\varphi : k \to k$ just means $\varphi(\lambda x) = \lambda^{1/p^e}\varphi(x)$ for all $x, \lambda \in k$. In particular, such a map is completely determined by where it sends any nonzero element.

Example 3.0.4. Consider $R = k[x]$. Then $F_* R$ is a free module with basis

$$\{F_*1, F_*x, F_*x^2, \dots, F_*x^{p-1}\},$$

(see Exercise 1). Therefore any p^{-1}-linear map from $k[x]$ to any other R-module N is simply a choice of where to send these basis elements.

Example 3.0.5. Consider $R = k[x_1, \dots, x_n]$ then as we saw in Exercise 1, $F_* R$ is a free R-module with basis $F_* x_1^{\lambda_1} \cdots x_n^{\lambda_n}$ for $0 \le \lambda_i \le p-1$. A map $F_* R \to R$ is uniquely determined by where it sends the elements of this basis.

Consider the R-linear map $\Phi : F_* R \to R$ which sends $F_* x_1^{p-1} \cdots x_n^{p-1}$ to 1 and all other basis elements to zero. In other words

$$\Phi\left(F_*(x_1^{\lambda_1} \cdots x_n^{\lambda_n})\right) = \begin{cases} x_1^{\frac{\lambda_1-(p-1)}{p}} \cdots x_n^{\frac{\lambda_n-(p-1)}{p}}, & \text{if all } \frac{\lambda_i-(p-1)}{p} \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}$$

For each tuple $\lambda = (\lambda_1, \dots, \lambda_n) \in \{0, 1, \dots, p-1\}^n$, consider the map $\varphi^\lambda : F_* R \to R$ defined by the rule $\varphi_\lambda(F_* \_\_) = \Phi(F_*(x_1^{p-1-\lambda_1} \cdots x_n^{p-1-\lambda_n} \cdot \_\_))$. It is easy to see that φ^λ sends $\mathbf{x}^\lambda$ to 1 and all other basis monomials to zero.

Because we can thus obtain all of the projections this way, it follows that the map $F_* R \to \operatorname{Hom}_R(F_* R, R)$ which sends $F_* z$ to the map $\varphi_z(F_* \_\_) = \Phi(F_*(z \cdot \_\_))$ is surjective as a map of $F_* R$-modules. On the other hand, it is clearly injective as well, and so it is an isomorphism. In other words, we just showed that $\operatorname{Hom}_R(F_* R, R)$ is a free $F_* R$–module generated by Φ. In other words, Φ generates $\operatorname{Hom}_R(F_* R, R)$ as an $F_* R$-module.

The most pervasive type of p^{-1}-linear maps are maps $\varphi : R \to R$. Of course, for fixed e, the set of p^{-e}-linear maps $\{\varphi : R \to R \mid \varphi \text{ is } p^{-e}\text{-linear}\}$ form a group under addition. However, as we vary e, we have a multiplication of these maps as well. Indeed, suppose that $\varphi : R \to R$ is p^{-e}-linear and $\psi : R \to R$ is p^{-d}-linear. Then both $\varphi \circ \psi$ and $\psi \circ \varphi$ are p^{-e-d}-linear. However, they need not be equal as the following example shows. It follows that

$$\bigoplus_{e \ge 0} \{\varphi : R \to R \mid \varphi \text{ is } p^{-e}\text{-linear}\}$$

forms a noncommutative graded ring. This graded ring will be studied more in Sect. 9.3.

Example 3.0.6. Suppose that $R = \mathbb{F}_p[x]$. We will describe two p^{-1}-linear maps, φ, ψ presented as in Exercise 3.0.4.

- $\varphi : R \to R$ satisfies $\varphi(x^{p-1}) = 1$ and $\varphi(x^i) = 0$ for $0 \le i < p-1$.
- $\psi : R \to R$ satisfies $\psi(1) = 1$ and $\psi(x^i) = 0$ for $0 < i \le p-1$.

Then $\psi \circ \varphi$ and $\varphi \circ \psi$ are p^{-2}-linear maps. However, notice that

$$\varphi(\psi(x^{p-1})) = \varphi(0) = 0$$

but that

$$\psi(\varphi(x^{p-1})) = \psi(1) = 1.$$

In particular, $\psi \circ \varphi \ne \varphi \circ \psi$.

An important class of examples of p^{-1}-linear maps are the splittings of Frobenius.

Example 3.0.7 ([66, 71]). Let X be a scheme. A *Frobenius splitting* is any p^{-1}-linear map $\varphi : \mathcal{O}_X \to \mathcal{O}_X$ that sends 1 to 1. Equivalently, it is an $\mathcal{O}_X$-linear map $\varphi : F_* \mathcal{O}_X \to \mathcal{O}_X$ that sends $F_* 1$ to 1. This, in particular, implies that the composition

$$\mathcal{O}_X \xrightarrow{F} F_* \mathcal{O}_X \xrightarrow{\varphi} \mathcal{O}_X \tag{4}$$

is an isomorphism; hence φ "splits" the Frobenius.

If X has a Frobenius splitting, then it satisfies many remarkable properties as we shall discuss in detail in Sect. 5. Let us just mention two of them to taste.

If X is a scheme that has some Frobenius splitting $\varphi: F_* \mathcal{O}_X \to \mathcal{O}_X$ (we call such X Frobenius split), then X is reduced: Indeed, if $x \in \Gamma(U, \mathcal{O}_X)$ is such that $0 = x^{p^e} = F^e(x)$ for some $e \geq 0$, then $0 = \varphi^e(F^e(x)) = x$, simply by that fact that $\varphi \circ F = \mathrm{id}$. This is a simple but important local property of Frobenius split varieties.

A similarly fundamental global result is the following *vanishing theorem*: Suppose that $\mathscr{L}$ is a line bundle and that $H^i(X, \mathscr{L}^p) = 0$ for some $i > 0$ (e.g., $H^i(X, \mathscr{L}^{p^e}) = 0$ holds for $e \gg 0$ for ample $\mathscr{L}$ by Serre vanishing), then $H^i(X, \mathscr{L}) = 0$ as well since we have the following isomorphism obtained by tensoring (4) by $\mathscr{L}$, using the projection formula and taking cohomology

$$H^i(X, \mathscr{L}) \xrightarrow{F} H^i(X, F_* \mathscr{L}^p) = 0 \xrightarrow{\varphi} H^i(X, \mathscr{L}).$$

If $e > 1$, rinse and repeat. We will study vanishing theorems for Frobenius split varieties in much greater detail in Theorem 5.2.4.

3.1 The Cartier Isomorphism

We now come to the most important example of a p^{-1}-linear map, coming from the Cartier operator. Suppose that X is a smooth variety over a perfect field k of characteristic $p > 0$. Consider the de Rham complex, $\Omega^{\bullet}_X$. This is not a complex of $\mathcal{O}_X$-modules (the differentials are not $\mathcal{O}_X$-linear). However, the complex

$$F_* \Omega^{\bullet}_X$$

is a complex of $\mathcal{O}_X$-modules (notice that $d(x^p) = 0$). We now state the Cartier isomorphism. We take this presentation from [11, 13, 19, 44].

Definition-Proposition 3.1.1. *There is a natural isomorphism (of $\mathcal{O}_X$-modules):*

$$C^{-1} : \Omega^i_X \to \mathbf{h}^i(F_* \Omega^{\bullet}_X).$$

Remark 3.1.2. It might strike the reader as odd that we put an inverse on C. This is because the isomorphism in the other direction is called the *Cartier operator* and represented by C. It is just more convenient for us to define C^{-1} than it is to define C.

We will not use the details of this isomorphism later in this chapter. However, the map T we obtain from it in Sect. 3.2 will be indispensable.

Let us explain how to construct this isomorphism C^{-1}. We follow [19, 9.13] and [44]. We begin with C^{-1} in the case that $i = 1$. We work locally on X (which we assume is affine), and we define C^{-1} by its action on $dx \in \Omega^i_X$, $x \in \mathcal{O}_X$:

$$C^{-1}(dx) := F_* x^{p-1} dx,$$

or rather, the image of $F_* x^{p-1} dx$ in cohomology. In order for this to make sense, we observe that $d(x^{p-1}dx) = 0$, in other words, that $C^{-1}(dx)$ is in the kernel of d. We now must show that C^{-1} is additive.

Now $C^{-1}(d(x) + d(y)) = C^{-1}(d(x + y)) = F_*(x + y)^{p-1} d(x + y)$, we need to compare this to $F_* x^{p-1} dx + F_* y^{p-1} dy = C^{-1}(dx) + C^{-1}(dy)$. Write $f = \frac{1}{p}\left((x + y)^p - x^p - y^p\right)$. Here the $\frac{1}{p}$ is a formal operation that simply cancels ps from the binomial coefficients. Then

$$\begin{aligned} df &= d\left(\sum_{i=1}^{p-1} \gamma_i x^i y^{p-i}\right) \\ &= \left(\sum_{i=1}^{p-1} \gamma_i i x^{i-1} y^{p-i}\right) dx + \left(\sum_{i=1}^{p-1} \gamma_i (p-i) x^i y^{p-i-1}\right) dy \\ &= \left(\sum_{i=1}^{p-1} \gamma_i i x^{(p-1)-(p-i)} y^{p-i}\right) dx + \left(\sum_{i=1}^{p-1} \gamma_i (p-i) x^i y^{(p-1)-i}\right) dy \end{aligned}$$

where $\gamma_i = \frac{1}{p}\binom{p}{i} = \frac{(p-1)(p-2)\cdots 1}{i!(p-i)!} = \frac{1}{i}\binom{p-1}{p-i} = \frac{1}{p-i}\binom{p-1}{i}$. Thus,

$$df = (x + y)^{p-1}(dx + dy) - x^{p-1}dx - y^{p-1}dy.$$

Therefore, $x^{p-1}dx + y^{p-1}dy$ and $(x+y)^{p-1}d(x+y)$ are the same in $\Omega^1_X / d\left(\Omega^0_X\right)$. This proves that C^{-1} is additive. Finally, we extend by p-linearity to obtain that

$$C^{-1}(f dx) = F_* f^p x^{p-1} dx.$$

We should also show that C^{-1} is an isomorphism. We only show that this initial C^{-1} is injective—in a special case. Set $X = \operatorname{Spec} \mathbb{F}_p[x, y]$ (see, e.g., [19, Theorem 9.14] for how to reduce to the polynomial ring case in general).

Suppose that $C^{-1}(f dx + g dy) = 0$. Let $h \in \mathcal{O}_X$ be such that we have $f^p x^{p-1} dx + g^p y^{p-1} dy = dh = \frac{\partial h}{\partial x} dx + \frac{\partial h}{\partial y} dy$. Therefore, if $f = \sum \lambda_{i,j} y^i x^j$, we see that

$$\sum \lambda_{i,j} y^{ip} x^{jp+p-1} = f^p x^{p-1} = \frac{\partial h}{\partial x}.$$

However, this is ridiculous unless $f dx + g dy = 0$. If you take a derivative of some non-zero polynomial in x with respect to x, no output can ever have x^{jp+p-1} in it. This completes the proof of injectivity of $C^{-1} : \Omega^1_X \to \mathbf{h}^1(F_* \Omega^{\bullet}_X)$ in the case that $X = \operatorname{Spec} \mathbb{F}_p[x, y]$. The general case is similar.

The surjectivity of C^{-1} is more involved. See, for example, [19, Theorem 9.14(d)] or [11, Theorem 1.3.4] or do Exercise *3.1.

At this point, we have only defined

$$\Omega^1_X \to \mathbf{h}^1(F_* \Omega^{\bullet}_X).$$

We define $C^{-1} : \Omega^i_X \to \mathbf{h}^i(F_* \Omega^{\bullet}_X)$ for $i > 1$ using wedge powers of C^{-1} for $i = 1$. We make this definition for any X.

Example 3.1.3 (Cartier isomorphism $\mathbb{A}^2$). Returning again to $X = \mathbb{A}^2 = \mathbb{F}_p[x, y]$, we explicitly compute $C^{-1} : \Omega^2_X \to \mathbf{h}^2(F_* \Omega^{\bullet}_X)$ at the top cohomology.

By definition

$$\begin{aligned} C^{-1}(f dx dy) = C^{-1}(f dx \wedge dy) &:= F_*\big(f^p (x^{p-1} dx) \wedge (y^{p-1} dy)\big) \\ &= F_* f^p x^{p-1} y^{p-1} dx dy \end{aligned}$$

or rather its image in cohomology. Again, this map is an isomorphism (Exercise 2).

3.2 Grothendieck Trace of Frobenius

Suppose that X is a smooth n-dimensional variety over a perfect field k of characteristic $p > 0$. Then coming from the Cartier isomorphism, Theorem 3.1.1, we have an exact sequence

$$F_* \Omega^{n-1}_{X/k} \xrightarrow{d} F_* \Omega^n_{X/k} \xrightarrow{T} \Omega^n_{X/k} \to 0.$$

The surjective map $T : F_* \Omega^n_{X/k} =: F_* \omega_X \to \omega_X := \Omega^n_{X/k}$ is often called the *trace map* or *Cartier map/operator*.

This map can be constructed in other ways. With X as above, again set $\omega_X = \Omega^n_{X/k}$. Then ω_X is a *dualizing/canonical* module in the sense of [34, Chap. III, Sect. 7] or more generally, [32, Chap. V].

For any finite dominant map $\pi : Y \to X$ with Y and X smooth, it is a fact (black boxed for now [32, Chap. V, Proposition 2.4], [50, Proposition 5.68]) that $\pi_* \omega_Y \cong \mathscr{H}om_{\mathcal{O}_Y}(\pi_* \mathcal{O}_Y, \omega_X)$ as a $\pi_* \mathcal{O}_Y$-module. This is described in greater generality on the next pages (see the diagram (7)). Note that this *completely determines* ω_Y as well, since π is finite, and so the data of a coherent $\pi_* \mathcal{O}_Y$-module on X is equivalent to the data of a coherent $\mathcal{O}_Y$-module on Y. Now, we also have the following map:

$$\pi_* \omega_Y \cong \mathscr{H}om_{\mathcal{O}_X}(\pi_* \mathcal{O}_Y, \omega_X) \xrightarrow{\text{eval @ 1}} \omega_X. \tag{5}$$

This is the map which sends a section $\varphi \in \omega_Y \cong \Gamma(U, \mathscr{H}om_{\mathcal{O}_Y}(\pi_* \mathcal{O}_Y, \omega_X))$ to the element $\varphi(1) \in \Gamma(U, \omega_X)$.

Now we specialize to the case that $Y = X$ and $\pi = F$ the Frobenius map.

Theorem 3.2.1. *The map described in (5) is the map T described above (up to choice of isomorphism).*

Sketch of Proof. We only show this for $X = \operatorname{Spec} \mathbb{F}_p[x, y] = \mathbb{A}^2$. By considering Example 3.1.3, we see that the map T sends

$$F_* f^p x^{p-1} y^{p-1} dxdy \mapsto f dxdy$$

and everything not of that form to zero.

So we then consider

$$\begin{array}{ccccc} \mathscr{H}om_{\mathcal{O}_X}(F_* \mathcal{O}_X, \omega_X) & \xleftrightarrow{\simeq} & F_* \omega_X & \xleftrightarrow{\simeq} & F_* \mathcal{O}_X. \\ & & F_* dxdy & \longleftrightarrow & F_* 1 \end{array}$$

Now, we identify the $\Phi \in \operatorname{Hom}_{\mathcal{O}_X}(F_* \mathcal{O}_X, \omega_X)$ which generates $\mathscr{H}om_{\mathcal{O}_X}(F_* \mathcal{O}_X, \omega_X)$ as an $F_* \mathcal{O}_X$-module just as in Example 3.0.5. Since $\omega_X = \mathcal{O}_X \cdot (dxdy) \cong \mathcal{O}_X$, we notice that Φ sends $F_* f^p x^{p-1} y^{p-1} \mapsto f dxdy$ and Φ sends things not of this form to zero.

Choosing then $\varphi(F_* \_\_) = \Phi(F_* c \cdot \_\_) \in \mathscr{H}om_{\mathcal{O}_X}(F_* \mathcal{O}_X, \omega_X)$, we see that the evaluation-at-1 map (5) sends φ to $\Phi(F^e_* c)$. Making the identification

$$(F_* \mathcal{O}_X) \cdot (F_* dxdy) = F_* \omega_X \cong \mathscr{H}om_{\mathcal{O}_X}(F_* \mathcal{O}_X, \omega_X) = (F_* \mathcal{O}_X) \cdot \Phi,$$

we immediately observe that T and the evaluation-at-1 map (5) coincide.

The general case for $X \neq \mathbb{A}^2$ is similar but slightly more technical to write down. Both the map T and the evaluation-at-1 map can be shown to be a local generator of the same $\mathscr{H}$om-sheaf. Thus they coincide up to multiplication by a unit of $\Gamma(X, \mathcal{O}_X)$. □

3.3 The Trace Map for Singular Varieties

Suppose that X is a normal variety with $U \subseteq X$ the regular locus. Consider the map $T : F^e_* \omega_U \to \omega_U$ as described above. This is an element of $\operatorname{Hom}_U(F^e_* \omega_U, \omega_U)$. However, there is an isomorphism $\operatorname{Hom}_{\mathcal{O}_U}(F^e_* \omega_U, \omega_U) \cong \operatorname{Hom}_{\mathcal{O}_X}(F^e_* \omega_X, \omega_X)$ since $X \setminus U$ is a codimension 2 subset of X and X is normal (see Appendix A). Therefore we obtain the following proposition.

Proposition 3.3.1. *Given any normal variety X, there is a trace map $T : F^e_* \omega_X \to \omega_X$ which agrees with and is completely determined by the map T described in terms of the Cartier isomorphism on the regular locus $U \subseteq X$.*

Even for non-normal schemes, we can do something similar if we are willing to work in the derived category. Suppose that X and Y are schemes of finite type over a field k with a map $f : X \to Y$. Then there is the extraordinary inverse image functor $f^!$ from $D^+_{\text{coh}}(Y)$ to $D^+_{\text{coh}}(X)$ (bounded below complexes of $\mathcal{O}_Y$-modules the extraordinary inverse image, respectively $\mathcal{O}_X$-modules, with coherent cohomology). For a precise definition of $f^!$, please see [32]. Its abstract existence can nowadays be shown quite formally from general principles, cf. [58]. Its key property is that it is right adjoint to $\mathbf{R}f_*$ in the case that f is proper (see Exercise *3.5). We will define $f^!$ in two cases which will suffice for our purposes.

Finite: If f is finite (e.g., Frobenius or a closed immersion), then $\mathscr{F} \in D^b_{\text{coh}}(X)$ we have an isomorphism of $f_*\mathcal{O}_X$-complexes

$$f_* f^! \mathscr{F} = \mathbf{R}\mathscr{H}\text{om}^\bullet_{\mathcal{O}_Y}(f_*\mathcal{O}_X, \mathscr{F}) \tag{6}$$

where $\mathbf{R}\mathscr{H}\text{om}^\bullet_{\mathcal{O}_Y}(f_*\mathcal{O}_X, \mathscr{F})$ is the complex obtained by taking an injective resolution of $\mathscr{F}$ and applying $\mathscr{H}\text{om}^\bullet_{\mathcal{O}_Y}(f_*\mathcal{O}_X, \_\_)$. Note that this completely describes $f^!$ since f is finite so that f_* is harmless.

Smooth: If f is smooth of relative dimension n, for any $\mathscr{F} \in D^b_{\text{coh}}(X)$, we have an isomorphism

$$f^! \mathscr{F} = (\mathbf{L}f^* \mathscr{F}) \otimes (\wedge^n \Omega^1_{Y/X})[n].$$

If $f : X \to \operatorname{Spec} k$ is itself the structural map, then we define the *dualizing complex of* X, denoted $\omega^\bullet_X$ as follows. View $k \in D^b_{\text{coh}}(\operatorname{Spec} k)$ as the complex which is trivial except in degree zero where it is k. Then we define $\omega^\bullet_X := f^! k$ to be the dualizing complex on X.

Consider the following diagram:

$$\begin{array}{ccc} X & \xrightarrow{F^e} & X \\ {\scriptstyle f}\downarrow & & \downarrow{\scriptstyle f} \\ \operatorname{Spec} k & \xrightarrow[F^e]{} & \operatorname{Spec} k \end{array} \tag{7}$$

where the top row is the absolute e-iterated Frobenius on X and the bottom row is the e-iterated Frobenius on k. Notice that the bottom row is an isomorphism (although not the identity), and so $(F^e)^! k \cong k$. The fact that $(f \circ g)^! = f^! \circ g^!$ then implies that $\omega^\bullet_X$ is independent of the choice of Frobenius-variant of the k-structure on X. In particular, we see that

$$\omega^\bullet_X \cong f^! k \cong (f \circ F^e)^! k \cong (F^e \circ f)^! k \cong (F^e)^! \omega^\bullet_X. \tag{8}$$

Now we will apply the *duality functor* $\mathbf{R}\mathscr{H}\mathrm{om}^{\bullet}_{\mathcal{O}_X}(\_\_, \omega_X^{\bullet})$ to the Frobenius map $\mathcal{O}_X \to F^e_* \mathcal{O}_X$. This operation yields

$$\omega_X^{\bullet} \cong \mathbf{R}\mathscr{H}\mathrm{om}^{\bullet}_{\mathcal{O}_X}(\mathcal{O}_X, \omega_X^{\bullet}) \leftarrow \mathbf{R}\mathscr{H}\mathrm{om}^{\bullet}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \omega_X^{\bullet}) \cong F^e_*(F^e)^! \omega_X^{\bullet} \cong F^e_* \omega_X^{\bullet}$$

where the isomorphisms are in the derived category and the final two isomorphisms are due to Eqs. (6) and (8), respectively. Taking cohomology of this map of complexes gives us maps

$$\mathbf{h}^i \omega_X^{\bullet} \leftarrow F^e_* \mathbf{h}^i \omega_X^{\bullet} \cong \mathbf{h}^i F^e_* \omega_X^{\bullet} \tag{9}$$

for each integer $i \in \mathbb{Z}$.

For any equidimensional scheme X of finite type over k with dualizing complex $\omega_X^{\bullet} := f^! k$, we define $\omega_X = \mathbf{h}^{-\dim X}(\omega_X^{\bullet})$ and call it the *canonical module of* X. It follows that (9) induces a map $F^e_* \omega_X \to \omega_X$. As expected, we then have

Proposition 3.3.2. *The map $F^e_* \omega_X \to \omega_X$ coincides with the map T defined previously on the regular locus of X.*

3.4 Exercises

Exercise* 3.1. Suppose that k is a perfect field and that $X = \operatorname{Spec} k[x, y] = \mathbb{A}^2$, prove that $C^{-1} : \Omega^1_X \to \mathbf{h}^1(F_* \Omega^{\bullet}_X)$ is surjective.

Hint: First prove the result for $\mathbb{A}^1 = \operatorname{Spec} \mathbb{F}_p[x]$. Now consider $\sum_j y^j(\alpha_j + \beta_j x^{b_j} dy) = \alpha \in \Omega^1_X$ such that $d\alpha = 0$ where $\alpha_j \in \Omega^1_{\mathbb{A}^1}$ and $\beta_j \in \mathbb{F}_p[x]$. Deduce that $y^{j+1}\alpha_{j+1} + y^j \beta_j dy \in d\Omega^0_X$ if $j+1$ is not divisible by p. Use this to rewrite α and then use the result for $\mathbb{A}^1$.

This method can be used to do the general proof by induction (see [11, Theorem 1.3.4]).

Exercise 2. Suppose that k is a perfect field and that $X = \operatorname{Spec} k[x, y] = \mathbb{A}^2$, prove that $C^{-1} : \Omega^2_X \to \mathbf{h}^2(F_* \Omega^{\bullet}_X)$ is an isomorphism.

Exercise 3. Suppose that R is a regular local ring. We have seen that $F_* R$ is a flat R-module by Exercise* 2.9. Consider the evaluation-at-1 map

$$\begin{aligned} \operatorname{Hom}_R(F_* R, R) &\xrightarrow{\ e\ } R \\ \varphi &\longmapsto \varphi(F_* 1). \end{aligned}$$

Fix an isomorphism $\gamma : F_* R \to \operatorname{Hom}_R(F_* R, R)$ and consider the composition

$$e \circ \gamma : F_* R \to R.$$

Prove that $(e \circ \gamma)$ generates $\operatorname{Hom}_R(F_* R, R)$ as an $F_* R$-module.

Exercise 4. A variety X is called *Cohen–Macaulay* if $\omega_X^{\bullet} \cong \omega_X[\dim X]$ is a complex with cohomology only in degree equal to $-\dim X$. Suppose that H is a Cartier divisor on a Cohen–Macaulay scheme X. Prove that H is also Cohen–Macaulay. Conversely, suppose that H is Cohen–Macaulay, prove that X is Cohen–Macaulay in a neighborhood of H.

Hint: Apply the *duality functor* to the short exact sequence $0 \to \mathcal{O}_X(-H) \to \mathcal{O}_X \to \mathcal{O}_H \to 0$ and observe that $\mathbf{R}\mathscr{H}om^{\bullet}_{\mathcal{O}_X}(\mathcal{O}_H, \omega_X^{\bullet}) \cong \omega_H^{\bullet}$ by (6). For the converse statement, use Nakayama's lemma.

Exercise* 3.5 (Grothendieck duality). (*For those who wish to learn some homological algebra*) Grothendieck duality says the following:

Theorem. *If $f : X \to Y$ is a proper map of schemes of finite type over a field k, then we have an isomorphism in $D^b_{\text{coh}}(Y)$*

$$\mathbf{R}\mathscr{H}om^{\bullet}_{\mathcal{O}_Y}(\mathbf{R}f_*\mathscr{F}, \mathscr{G}) \cong \mathbf{R}f_*\mathbf{R}\mathscr{H}om^{\bullet}_{\mathcal{O}_X}(\mathscr{F}, f^!\mathscr{G})$$

for $\mathscr{F} \in D^b_{\text{coh}}(X)$ and $\mathscr{G} \in D^b_{\text{coh}}(Y)$.

Set $Y = \operatorname{Spec} k$ and learn enough about the symbols above to deduce the variant of Serre duality found in Hartshorne [34, Chap. III, Sect. 7].

4 Connections with Divisors

In this section we explain why maps $\varphi \in \operatorname{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)$ contain roughly the same information as a $\mathbb{Q}$-divisor Δ such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier (i.e., such that there exists an integer n such that $n\Delta$ is integral and $nK_X + n\Delta$ is Cartier). These ideas go back at least to the original papers on Frobenius splittings [66, 71]. The difference between this section and those original papers is that we normalize our divisors with respect to iterates of Frobenius and thus obtain $\mathbb{Q}$-divisors.[4] The statements in this section are somewhat technical. Therefore, the reader may wish to skim this section for the main idea and refer back to the numbered bijections as needed throughout the remainder of this chapter.

Fix X to be a smooth variety of finite type over a perfect field. Consider an element $\varphi \in \operatorname{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)$. We claim that

$$\operatorname{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X) \cong F^e_*\mathcal{O}_X((1-p^e)K_X). \tag{10}$$

Let us prove this claim. By applying the projection formula as in Eq. (2), taking global sections and using the fact that $\mathscr{H}om_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X(K_X)) \cong F^e_*\mathcal{O}_X(K_X)$, we have

[4] Formal sums of codimension 1 subvarieties with rational coefficients.

$$\begin{aligned}\mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X) &\cong \mathrm{Hom}_{\mathcal{O}_X}((F^e_*\mathcal{O}_X) \otimes \mathcal{O}_X(K_X), \mathcal{O}_X(K_X)) \\ &\cong \mathrm{Hom}_{\mathcal{O}_X}(F^e_*(\mathcal{O}_X(p^e K_X)), \mathcal{O}_X(K_X)) \\ &\cong \mathrm{Hom}_{F^e_*\mathcal{O}_X}(F^e_*(\mathcal{O}_X(p^e K_X)), \mathscr{H}om_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X(K_X))) \\ &\cong F^e_*\mathrm{Hom}_{\mathcal{O}_X}((\mathcal{O}_X(p^e K_X)), \mathcal{O}_X(K_X)) \\ &\cong F^e_*\mathcal{O}_X((1-p^e)K_X).\end{aligned} \tag{11}$$

See (8), [50, Proposition 5.68] or [32]. Alternately, it follows from Grothendieck duality for the finite map $F : X \to X$ (see Exercise 1).

Therefore, any nonzero map $\varphi : F^e_*\mathcal{O}_X \to \mathcal{O}_X$ induces a nonzero global section of $F^e_*\mathcal{O}_X((1-p^e)K_X)$. By using the fact that F^e_* does not change the underlying structure of sheaves of Abelian groups, we see that there is a bijective correspondence:

$$\left\{\begin{array}{c}\text{nonzero elements}\\ \varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)\end{array}\right\} \longleftrightarrow \left\{\begin{array}{c}\text{nonzero elements}\\ z \in \Gamma\big(X, \mathcal{O}_X((1-p^e)K_X)\big)\end{array}\right\}.$$

Note every nonzero global section of $\mathcal{O}_X((1-p^e)K_X)$ induces an effective Weil divisor $0 \leq D \sim (1-p^e)K_X$; see Theorem A.2.6.

We notice also that two nonzero elements $z_1, z_2 \in \Gamma\big(X, \mathcal{O}_X((1-p^e)K_X)\big)$ induce the same divisor if and only if there exists a *unit* $u \in \Gamma\big(U, \mathcal{O}_X\big)$ such that $uz_1 = z_2$. Therefore, we have the following bijection:

$$\left\{\begin{array}{c}\text{nonzero } \varphi \in\\ \mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)\end{array}\right\} \Big/ \left(\begin{array}{c}\text{multiplication}\\ \text{by units in}\\ \Gamma(X, F^e_*\mathcal{O}_X)\end{array}\right) \longleftrightarrow \left\{\begin{array}{c}\text{effective divisors}\\ \text{linearly equivalent}\\ \text{to } (1-p^e)K_X\end{array}\right\}. \tag{12}$$

Now suppose that X is normal but not necessarily smooth. Of course, the previous argument works fine on $U = X_{\mathrm{reg}} \subseteq X$. However, Weil divisors are determined off a set of codimension 2. Likewise $\Gamma\big(X, \mathcal{O}_X((1-p^e)K_X)\big) = \Gamma\big(U, \mathcal{O}_X((1-p^e)K_X)\big)$ since $X \setminus U$ has codimension ≥ 2 cf. [36, Proposition 2.9]. In particular, we see that

(12) holds on normal varieties.

We continue now to work with normal X. Given an effective Weil divisor $D = D_\varphi \sim (1-p^e)K_X$ corresponding to φ, set $\Delta = \Delta_\varphi = \frac{1}{p^e-1}D_\varphi$. This is an effective $\mathbb{Q}$-divisor. Notice that

$$K_X + \Delta = K_X + \frac{1}{p^e-1}D \sim K_X + \frac{1}{p^e-1}(1-p^e)K_X = K_X - K_X = 0.$$

In particular, we obtain a bijective correspondence:

$$\left\{\begin{matrix}\text{nonzero } \varphi \in \\ \mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)\end{matrix}\right\} \Big/ \begin{pmatrix}\text{multiplication} \\ \text{by units in} \\ \Gamma(X, F^e_*\mathcal{O}_X)\end{pmatrix} \longleftrightarrow \left\{\begin{matrix}\mathbb{Q}\text{-divisors} \\ \Delta \geq 0 \text{ such that} \\ (p^e-1)(K_X+\Delta) \text{ is} \\ \text{an integral Weil} \\ \text{divisor linearly} \\ \text{equivalent to } 0\end{matrix}\right\}. \tag{13}$$

At this point, it is natural to ask why should one divide by $p^e - 1$. This division is a normalizing factor as described below.

Suppose that $\varphi : F^e_*\mathcal{O}_X \to \mathcal{O}_X$ is an $\mathcal{O}_X$-linear map. We apply the functor F^e_* and obtain $F^e_*\varphi : F^{2e}_*\mathcal{O}_X \xrightarrow{F^e_*\mathcal{O}_X} F^e_*\mathcal{O}_X$. Composing this with φ we obtain $\varphi \circ (F^e_*\varphi) : F^{2e}_*\mathcal{O}_X \to \mathcal{O}_X$. We use φ^2 to denote this map (note if we view φ as an honest p^{-e} linear map, then this is really just φ composed with itself). More generally, for each $n \geq 1$, we obtain maps

$$\varphi^n : F^{ne}_*\mathcal{O}_X \to \mathcal{O}_X \tag{14}$$

in the same way.

Lemma 4.0.1 ([73, Theorem 3.11(e)]). *Suppose that X is a normal variety. Then the map $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)$ induces the same $\mathbb{Q}$-divisor Δ via (13) as does the map*

$$\varphi^n \in \mathrm{Hom}_{\mathcal{O}_X}(F^{ne}_*\mathcal{O}_X, \mathcal{O}_X)$$

for any $n \geq 1$.

Proof. The divisor section correspondence is determined in codimension 1, and so we may assume that $X = \mathrm{Spec}\, R$ where $(R, \mathfrak{m})$ is a DVR with $\mathfrak{m} = \langle r \rangle$. We will simply verify the claim in the Lemma for $n = 2$ and leave the general case to the reader Exercise 3. Now, since R is regular (and so Gorenstein) and local, $K_X \sim 0$. Thus as

$$\mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X) \cong \Gamma(X, F^e_*\mathcal{O}_X((1-p^e)K_X)) \cong F^e_*R;$$

we fix $\Phi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_*\mathcal{O}_X, \mathcal{O}_X)$ corresponding to F^e_*1. In other words we pick Φ such that $D_\Phi = 0$ (note that the Sect. 1 doesn't vanish anywhere).

It is an exercise left to the reader that $\Phi^2 = \Phi \circ (F^e_*\Phi)$ generates $\mathrm{Hom}_R(F^{2e}_*R, R)$ as an F^{2e}_*-module (Exercise 2). This is the key point though!

Now consider $\varphi(F^e_* \_\_) = \Phi(F^e_*ur^a \cdot \_\_)$ for some unit $u \in R$ and integer $a \geq 0$. It follows immediately that $D_\varphi = a\mathrm{div}(r)$ and so $\Delta_\varphi = \frac{a}{p^e-1}\mathrm{div}(r)$.

Now we consider φ^2. We observe that

$$\begin{aligned}\varphi^2(F^{2e}_* \_\_) &= \Phi(F^e_*ur^a\Phi(F^e_*ur^a \cdot \_\_)) = \Phi(F^e_*\Phi(F^e_*u^{p^e+1}r^{a(p^e+1)} \cdot \_\_)) \\ &= \Phi^2(F^{2e}_*u^{p^e+1}r^{a(p^e+1)} \cdot \_\_).\end{aligned}$$

Thus $D_{\varphi^2} = a(p^e+1)\mathrm{div}(r)$ and so that $\Delta_{\varphi^2} = \frac{a(p^e+1)}{p^{2e}-1}\mathrm{div}(r) = \frac{a}{p^e-1}\mathrm{div}(r) = \Delta_\varphi$ as desired. □

Therefore, we obtain a bijection:

$$\left\{ \begin{array}{c} \text{nonzero } \varphi \in \\ \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X) \\ \text{as } e \geq 0 \text{ varies} \end{array} \right\} \Big/ \left(\begin{array}{c} \text{relation generated} \\ \text{by multiplication} \\ \text{by units in} \\ \Gamma(X, F^e_* \mathcal{O}_X) \text{ and by} \\ \text{composition in (14)} \end{array} \right) \longleftrightarrow \left\{ \begin{array}{c} \mathbb{Q}\text{-divisors} \\ \Delta \geq 0 \text{ such that} \\ n(K_X + \Delta) \sim 0 \\ \text{for some } n > 0 \\ \text{with } p \text{ not} \\ \text{dividing } n. \end{array} \right\} \tag{15}$$

Here we notice that $(p^e-1)(K_X+\Delta) \sim 0$ for some $e > 0$ is equivalent to requiring that $n(K_X + \Delta) \sim 0$ for some $n > 0$ which is not divisible by p (Exercise 5).

Example 4.0.2. Consider $X = \mathbb{A}^n = \mathrm{Spec}\, k[x_1, \ldots, x_n] = \mathrm{Spec}\, R$ over a perfect field k. Consider $\Phi : F^e_* R \to R$ defined by the following action on monomials:

$$\Phi\left(F^e_*(x_1^{\lambda_1} \cdots x_n^{\lambda_n})\right) = \begin{cases} x_1^{\frac{\lambda_1-(p^e-1)}{p^e}} \cdots x_n^{\frac{\lambda_n-(p^e-1)}{p^e}}, & \text{if all } \frac{\lambda_i-(p^e-1)}{p^e} \in \mathbb{Z} \\ 0, & \text{otherwise} \end{cases}$$

In other words, Φ sends $F^e_*(x_1^{p^e-1} \cdots x_n^{p^e-1})$ to 1 and all other elements of the obvious basis of $F^e_* R$ over R to zero. We already saw in Exercise 3.0.5 that $\Phi : F^e_* R \to R$ generates $\mathrm{Hom}_R(F^e_* R, R)$ as an $F^e_* R$-module (at least when $e = 1$, but the general case is no different).

But then it immediately follows that the divisor D_Φ is the zero divisor. In particular, D_Φ corresponds to the element in $\mathrm{Hom}_R(F^e_* R, R) \cong F^e_* R$ that doesn't vanish anywhere.

4.1 A Generalization with Line Bundles

Previously we considered nonzero maps $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X)$. In this section, we generalize this to maps $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathcal{O}_X)$ where $\mathscr{L}$ is a line bundle on X. This generality actually simplifies some of the statements considered in the previous section. Indeed, just as in (11), it is easy to see that for a smooth variety X

$$\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathcal{O}_X) \cong F^e_* \mathscr{L}^{-1}((1-p^e)K_X).$$

Just as before, this extends to normal varieties as well. Therefore, for any line bundle on a normal variety X, we have a bijection of sets.

$$\left\{ \begin{array}{c} \text{nonzero } \varphi \in \\ \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathcal{O}_X) \end{array} \right\} \Big/ \left(\begin{array}{c} \text{multiplication} \\ \text{by units in} \\ \Gamma(X, F^e_* \mathcal{O}_X) \end{array} \right) \longleftrightarrow \left\{ \begin{array}{c} \text{effective Weil} \\ \text{divisors } D \text{ such} \\ \text{that } \mathcal{O}_X(D) \cong \\ \mathscr{L}^{-1}((1-p^e)K_X) \end{array} \right\}. \tag{16}$$

Thus, just as before, $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathcal{O}_X)$ induces $\mathbb{Q}$-divisors $\Delta_\varphi = \frac{1}{p^e-1} D$ such that $\mathcal{O}_X((p^e-1)(K_X + \Delta)) \cong \mathscr{L}^{-1}$.

Definition 4.1.1. Given $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$, we use Δ_φ to denote the $\mathbb{Q}$-*divisor associated to* φ as above.

Finally, consider the data of a line bundle $\mathscr{L}$ and an $\mathcal{O}_X$-linear map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$. We will compose φ with itself in the following way. We tensor φ with $\mathscr{L}$ and use the projection formula to obtain

$$F^e_*(\mathscr{L}^{p^e+1}) \to \mathscr{L}.$$

Pushing forward by F^e_* we obtain

$$F^{2e}_*(\mathscr{L}^{p^e+1}) \to F^e_* \mathscr{L}.$$

Composing with φ again we obtain a map

$$F^{2e}_*(\mathscr{L}^{p^e+1}) \to \mathcal{O}_X$$

which we denote by φ^2. Continuing in this way, we obtain maps

$$\varphi^n : F^{ne}_*(\mathscr{L}^{p^{(n-1)e}+\cdots+p^e+1}) \to \mathcal{O}_X \tag{17}$$

for all $n \geq 1$.

It is then straightforward to verify that

Lemma 4.1.2. *The* $\mathbb{Q}$*-divisor* Δ_φ *induced by* $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ *is equal to the* $\mathbb{Q}$*-divisor* Δ_{φ^n} *induced by* $\varphi^n : F^{ne}_*(\mathscr{L}^{p^{(n-1)e}+\cdots+p^e+1}) \to \mathcal{O}_X$.

Proof. Left as an exercise to the reader Exercise 7. □

In other words, forming the $\mathbb{Q}$-divisor $\Delta = \frac{1}{p^e-1} D$ normalizes the divisor with respect to self composition just as in the case that $\mathscr{L} = \mathcal{O}_X$.

Given two line bundles $\mathscr{L}, \mathscr{M}$, we declare maps $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ and $\psi : F^e_* \mathscr{M} \to \mathcal{O}_X$ equivalent if there exists a commutative diagram:

$$\begin{array}{ccc} F^e_* \mathscr{L} & \xrightarrow{\alpha} & F^e_* \mathscr{M} \\ \varphi \downarrow & & \downarrow \psi \\ \mathcal{O}_X & \xrightarrow[\mathrm{id}]{} & \mathcal{O}_X \end{array}$$

where α is an isomorphism. We also declare φ and φ^n to be equivalent. These relations generate an equivalence relation $\sim$ on pairs $(\mathscr{L}, \varphi : F^e_* \mathscr{L} \to \mathcal{O}_X)$.

Theorem 4.1.3. *For a normal variety X over a field of characteristic $p > 0$, there is a bijection of sets*

$$\left\{ \begin{array}{c} \textit{Line bundles } \mathscr{L} \textit{ and} \\ \textit{maps } \varphi : F^e_* \mathscr{L} \to \mathcal{O}_X \\ \textit{modulo equivalence} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \textit{Effective } \mathbb{Q}\textit{-divisors } \Delta \\ \textit{such that} \\ \mathcal{O}_X((p^e - 1)(K_X + \Delta)) \cong \mathscr{L}^{-1} \end{array} \right\}.$$

Proof. Left to the reader Exercise* 4.8. ☐

We compute a final example.

Example 4.1.4. Set $X = \mathbb{P}^n_k$ and consider the line bundle $\mathscr{L} = \mathcal{O}_X((1-p)K_X) = \mathcal{O}_X((n+1)(p-1))$. Then

$$\begin{aligned} & \mathscr{H}\mathrm{om}_{\mathcal{O}_X}(F_* \mathscr{L}, \mathcal{O}_X) \\ = \ & \mathscr{H}\mathrm{om}_{\mathcal{O}_X}(F_* \mathcal{O}_X((1-p)K_X), \mathcal{O}_X) \\ = \ & \mathscr{H}\mathrm{om}_{\mathcal{O}_X}(F_* \mathcal{O}_X(K_X), \mathcal{O}_X(K_X)) \\ = \ & F_* \mathscr{H}\mathrm{om}_{\mathcal{O}_X}(\mathcal{O}_X(K_X), \mathcal{O}_X(K_X)) \\ = \ & F_* \mathcal{O}_X. \end{aligned}$$

In particular, there is only one nonzero element $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F_* \mathscr{L}, \mathcal{O}_X)$ up to scaling by elements of k. In particular, it follows that $D_\varphi = \Delta_\varphi = 0$ since a nonzero global section of $F_* \mathcal{O}_X$ doesn't vanish anywhere. On the affine charts, this element is easily seen to coincide with the map described in Example 4.0.2 (at least for $e = 1$).

On the other hand, there is an obvious map

$$\psi : F_* \mathcal{O}_X \to \mathcal{O}_X$$

defined by the rule $\psi(F_* y) = y^{1/p}$, for $y \in \Gamma(U, \mathcal{O}_X)$, if $y^{1/p} \in \Gamma(U, \mathcal{O}_X)$ and $\psi(F_* y) = 0$ otherwise.

It is an exercise left to the reader that $D_\psi = (p-1)F$ where F is the union of the various coordinate axes in $\mathbb{P}^n$. For example, if $n = 2$ and $X = \mathrm{Proj}\, k[x, y, z]$, then $F = V(xyz)$.

4.2 Exercises

Exercise 1 (Grothendieck duality for a finite map). Suppose that $R \subseteq S$ is a finite inclusion of Cohen–Macaulay local rings and M is an S-module. Grothendieck duality for this inclusion says that there is an isomorphism of S-modules:

$$\mathrm{Hom}_R(M, \omega_R) \cong \mathrm{Hom}_S(M, \omega_S).$$

Here ω_R and ω_S are canonical modules for R and S, respectively. Verify that this is an easy consequence of the formula $\mathrm{Hom}_R(S, \omega_R) = \omega_S$, a formula which was given to you (8).

Exercise 2. Suppose that R is a ring and S is an R-algebra such that $\mathrm{Hom}_R(S, R) \cong S$ as S-modules. Suppose that M is any S-module and prove that the natural map:

$$\mathrm{Hom}_S(M, S) \times \mathrm{Hom}_R(S, R) \to \mathrm{Hom}_R(M, R)$$

defined by $(\psi, \varphi) \mapsto \varphi \circ \psi$ is surjective.

In particular, every map in $\mathrm{Hom}_R(M, R)$ can be factored through a map in $\mathrm{Hom}_R(S, R)$. A solution can be found in [54, Appendix F].

Exercise 3. Prove the general case of Lemma 4.0.1.

Exercise 4. Suppose that $R \subseteq S$ is a finite extension of Gorenstein local rings. Prove that $\mathrm{Hom}_R(S, R)$ is a rank-1 free S-module. Conclude that if R is Gorenstein and local, $\mathrm{Hom}_R(F^e_* R, R)$ is isomorphic to $F^e_* R$ as an $F^e_* R$-module.

Hint: Since R is Gorenstein and local (semi-local is good enough), $\omega_R \cong R$.

Exercise 5. Suppose we are given an integer $n > 0$ such that p does not divide n, prove that $n | (p^e - 1)$ for some integer $e > 0$. Conclude that $(p^e - 1)(K_X + \Delta) \sim 0$ for some $e > 0$ if and only if $n(K_X + \Delta) \sim 0$ for some $n > 0$ which is not divisible by p.

Exercise 6. Compute D_ψ and Δ_ψ where ψ is as in Example 4.1.4.

Exercise 7. Prove Lemma 4.1.2. See [73, Theorem 3.11(e)].

Exercise* 4.8. Prove Theorem 4.1.3.

Exercise 9. Suppose that X is a smooth (or Gorenstein) variety and $T : F_* \omega_X \to \omega_X$ is the trace map as described in Sect. 3.3. By twisting by $-K_X$ and reflexifying, we obtain a map $\Phi : F_* \mathcal{O}_X((1-p)K_X) \to \mathcal{O}_X$. Prove that Φ corresponds to the zero divisor by (16).

Exercise 10. A normal variety X is called $\mathbb{Q}$-*Gorenstein* if $\mathcal{O}_X(nK_X)$ is a line bundle for some $n > 0$ (in other words, nK_X is Cartier). Note that we do not require $\mathbb{Q}$-Gorenstein varieties to be Cohen–Macaulay. In this case, the *index of* K_X is the smallest $n > 0$ such that nK_X is a Cartier divisor.

Suppose that X is $\mathbb{Q}$-Gorenstein with index not divisible by p. Suppose that $R = \mathcal{O}_{X,x}$ is the stalk of R at some point $x \in X$. Prove that we have an isomorphism of R-modules, $F^e_* R \cong \mathrm{Hom}_R(F^e_* R, R)$, for all sufficiently divisible e.

Exercise 11. Suppose that R is a normal domain and that $\varphi : F^e_* R \to R$ is an R-linear map corresponding to a divisor Δ_φ as in Definition 4.1.1. Fix a nonzero $g \in R$. Form a new map

$$\psi(F^e_* \_\_) = \varphi(F^e_*(g \cdot \_\_)).$$

Prove that $\Delta_\psi = \Delta_\varphi + \frac{1}{p^e - 1}\mathrm{div}(g)$.

Exercise* 4.12. Suppose that $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ and $\psi : F^f_* \mathscr{M} \to \mathcal{O}_X$ are two $\mathcal{O}_X$-linear maps. Form the twisted composition $\varphi \circ (F^e_* \psi')$ as follows. Twist ψ by $\mathscr{L}$ to get $\psi' : F^f_* (\mathscr{M} \otimes \mathscr{L}^{p^f}) \to \mathscr{L}$. Now pushforward by F^e_* and compose with φ and obtain:

$$\varphi \circ (F^e_* \psi') : F^{f+e}_* (\mathscr{M} \otimes \mathscr{L}^{p^f}) \xrightarrow{F^e_* \psi'} F^e_* \mathscr{L} \xrightarrow{\varphi} \mathcal{O}_X.$$

Find a relation between Δ_φ, Δ_ψ, and $\Delta_{\varphi \circ (F^e_* \psi')}$ where the Δ are $\mathbb{Q}$-divisors defined as in Definition 4.1.1. For a solution, see the proof of [78, Lemma 4.9(i)].

Exercise* 4.13 (Noneffective divisors). Fix a line bundle $\mathscr{L}$ on a variety X. There is a bijection between nonzero elements of $\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathscr{K}(X))$ and (*not necessarily effective*) Weil divisors D such that $\mathcal{O}_X(D) \cong \mathscr{L}^{-1}((1-p^e)K_X)$.

Indeed, suppose that $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathscr{L}, \mathscr{K}(X))$. Then, working locally if needed, for some sufficiently large Cartier divisor $E \geq 0$, we have that $\varphi(F^e_* \mathscr{L}((1-p^e)E)) \subseteq \mathcal{O}_X$. Set $\psi : F^e_* \mathscr{L}((1-p^e)E) \to \mathcal{O}_X$ to be the restriction map. Then ψ induces a divisor $D_\psi > 0$. Set $D_\varphi = D_\psi + (1-p^e)E$, and prove that D_φ is independent of the choice of E.

5 Frobenius Splittings

In this section we give a brief introduction to Frobenius splittings. A more complete treatment can be found in [11, Chap. 1].

Suppose that X is a scheme over a perfect field of characteristic $p > 0$.

Definition 5.0.1. We say that X is *Frobenius split* (or *F-split*) if the map

$$\mathcal{O}_X \to F_* \mathcal{O}_X$$

splits as a map of $\mathcal{O}_X$-modules. In this case the splitting map $\varphi : F_* \mathcal{O}_X \to \mathcal{O}_X$ is called a *Frobenius splitting*. Of course, there may be multiple different Frobenius splittings $\varphi \in \mathrm{Hom}_{\mathcal{O}_X}(F_* \mathcal{O}_X, \mathcal{O}_X)$.

Likewise, we say that a ring R is *Frobenius split* (or *F-pure*) if the map

$$R \to F_* R$$

splits as a map of R-modules.

A scheme X is said to be *F-pure* (or *locally F-split*) if every point $x \in X$ has a neighborhood which is F-split.

Remark 5.0.2. Frobenius split varieties were formally introduced in [66] (also see [71]), although very closely related concepts were studied in [24,37,40,70]. Indeed, Frobenius split affine varieties (i.e., rings) had been heavily studied by Hochster and his students in the 1970s and 1980s cf. [20].

We shall see below that every regular variety is F-pure Proposition 5.1.2, but not every regular variety is F-split Lemma 5.2.2.

Lemma 5.0.3. *A variety X is Frobenius split if and only if either of the following equivalent conditions hold.*

(a) The e-iterated Frobenius $\mathcal{O}_X \to F^e_* \mathcal{O}_X$ splits for some e.
(b) The e-iterated Frobenius $\mathcal{O}_X \to F^e_* \mathcal{O}_X$ splits for all e.

Proof. This is left as an exercise to the reader (see Exercise 1). □

Suppose that X is a variety, we will look for Frobenius splittings inside $\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X)$. Indeed, notice that for any $c \in \Gamma(X, \mathcal{O}_X)$, we have a map $\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X) \to \Gamma(X, \mathcal{O}_X)$ defined by evaluation at c, in other words, $\varphi \mapsto \varphi(F^e_* c)$. Now we observe that

Lemma 5.0.4. *A variety X is Frobenius split if and only if the evaluation-at-1 map*

$$\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X) \to \Gamma(X, \mathcal{O}_X)$$

is surjective.

Proof. Left as an exercise to the reader in Exercise 2. □

Finally, we observe that a normal X is Frobenius split if and only if the regular locus of X is Frobenius split.

Lemma 5.0.5. *Suppose that X is normal and $U \subseteq X$ is the regular locus. Then X is Frobenius split if and only if U is Frobenius split.*

Proof. The natural restriction map $\mathrm{Hom}_{\mathcal{O}_X}(F^e_* \mathcal{O}_X, \mathcal{O}_X) \to \mathrm{Hom}_{\mathcal{O}_U}(F^e_* \mathcal{O}_U, \mathcal{O}_U)$ is an isomorphism since $X \setminus U$ has codimension ≥ 2 and the $\mathscr{H}om$ sheaves are reflexive. See Appendix A and [11, Lemma 1.1.7] for additional discussion. □

5.1 *Local Properties of Frobenius Split Varieties*

The easiest property to prove about Frobenius split varieties is that they are reduced.

Lemma 5.1.1. *Suppose that a scheme X is F-pure, then X is reduced.*

Proof. Without loss of generality we may assume that $X = \operatorname{Spec} R$ is affine and Frobenius split. Suppose that $x \in R$ satisfies $x^n = 0$. Then $x^{p^e} = 0$ for some $e > 0$ (where p is the characteristic of R). Therefore $x = x\varphi(F^e_* 1) = \varphi(F^e_* x^{p^e}) = \varphi(F^e_* 0) = 0$. □

First we identify some Frobenius split varieties.

Proposition 5.1.2 (Regular affine varieties are Frobenius split). *Suppose that $X = \operatorname{Spec} R$ is a regular affine variety. Then X is Frobenius split.*

Proof. We prove the result for $R_{\mathfrak{m}} = \mathcal{O}_{X,x}$, the stalk of X at a closed point $x \in X$. The global affine case is Exercise 6. Let $\hat{R}$ denote the completion of $R_{\mathfrak{m}}$ at the maximal ideal $\mathfrak{m}$. Now consider the evaluation-at-1 map $\Phi : \mathrm{Hom}_{R_{\mathfrak{m}}}(F^e_* R_{\mathfrak{m}}, R_{\mathfrak{m}}) \to R_{\mathfrak{m}}$. Tensoring with $\hat{R}$ gives us a map

$$\hat{\Phi} : \mathrm{Hom}_{\hat{R}}(F^e_* \hat{R}, \hat{R}) \cong \mathrm{Hom}_{R_{\mathfrak{m}}}(F^e_* R_{\mathfrak{m}}, R_{\mathfrak{m}}) \otimes_{R_{\mathfrak{m}}} \hat{R} \to R_{\mathfrak{m}} \otimes_{R_{\mathfrak{m}}} \hat{R} \cong \hat{R}.$$

Here we have used Exercise 6. Note that by the Cohen-structure theorem, [65, Theorem 28.3], we have that $\hat{R} = k[\![x_1, \dots, x_n]\!]$. It follows then from the argument of Exercise 1 that $F^e_* \hat{R}$ is free as an $\hat{R}$-module and in particular, that there is a splitting of $\hat{R} \to F^e_* \hat{R}$. In particular, $\hat{\Phi}$ is surjective. But therefore Φ is surjective as well since tensoring with $\hat{R}$ is faithfully flat. Thus, by Lemma 5.0.4, we are done. □

Of course, not all Frobenius split varieties are regular.

Lemma 5.1.3 (Simple normal crossings are F-split). *The ring*

$$R = k[x_1, \dots, x_n]/\langle x_1 \cdot x_2 \cdots x_n \rangle = S/J$$

is Frobenius split.

Proof. Observe we have an “obvious” Frobenius splitting $\varphi : F^e_* k[x_1, \dots, x_n] \to k[x_1, \dots, x_n]$ coming from Exercise 1, which sends the basis element corresponding to $F^e_* 1$ to 1 and sends all the other basis elements $x_1^{\lambda_1} \dots x_n^{\lambda_n}$ to 0. We want to consider what this map does to the ideal $\langle x_1 \cdot x_2 \cdots x_n \rangle = J$. Consider any monomial in $\mathbf{x}^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in \langle x_1 \cdot x_2 \cdots x_n \rangle = J$. Then $\varphi(F^e_* \mathbf{x}^\alpha) \neq 0$ if and only if $p^e | \alpha_i$ for each i. In particular, this means that $\varphi(F^e_* \mathbf{x}^\alpha) = x_1^{\beta_1} \cdots x_n^{\beta_n}$ with each $\beta_i \geq 1$. Therefore, $\varphi(F^e_* \mathbf{x}^\alpha) \in J$. Since every element of J is a sum of such monomials, we have that $\varphi(F^e_* J) \subseteq J$.

But now consider the commutative diagram:

$$\begin{array}{ccc} F^e_* J & \xrightarrow{\varphi|_J} & J \\ \big\downarrow & & \big\downarrow \\ F^e_* R & \xrightarrow{\varphi} & R \\ \big\downarrow & & \big\downarrow \\ F^e_*(R/J) & \xrightarrow[\varphi/J]{} & R/J \end{array} \quad . \tag{18}$$

Since φ sends 1 to 1, so does φ/J. □

In the next section, we will introduce a highly effective tool, based upon similar analysis, which can be used to test whether an affine variety is Frobenius split—Fedder's criterion.

Definition 5.1.4. Suppose that $\varphi : F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is a Frobenius splitting, then an ideal sheaf $J \subseteq \mathcal{O}_X$ is called *compatibly (φ-)split* if $\varphi(F^e_* J) \subseteq J$. If the subscheme $Y = V(J) \subseteq X$, then we also say that Y is *compatibly (φ-)split*.

Note that in Lemma 5.1.3, we showed that the coordinate hyperplanes were compatibly split with the obvious Frobenius splitting on $X = \operatorname{Spec} k[x_1, \ldots, x_n]$. Indeed, consider the following proposition:

Proposition 5.1.5 (Properties of compatibly split varieties). *Suppose that $\varphi : F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is a Frobenius splitting. Then:*

(a) If $J \subseteq \mathcal{O}_X$ is compatibly φ-split, then $V(J)$ is Frobenius split as well. In particular, J is a radical ideal.
(b) If $J \subseteq \mathcal{O}_X$ is compatibly φ-split, then $\varphi(F^e_* J) = J$ (instead of just contained in J).
(c) If $I, J \subseteq \mathcal{O}_X$ is compatibly φ-split, then so are $I + J$ and $I \cap J$.
(d) If Q is a minimal prime over J, then Q is also compatibly φ-split.
(e) If $I \subseteq \mathcal{O}_X$ is compatibly φ-split, then so is $I : K$ for any ideal sheaf $K \subseteq \mathcal{O}_X$.
(f) A prime ideal sheaf Q is compatibly φ-split if and only if $Q \cdot \mathcal{O}_{X,Q}$ is compatibly φ_Q split where φ_Q is the map induced on the stalk $\varphi_Q : F^e_* \mathcal{O}_{X,Q} \to \mathcal{O}_{X,Q}$.

Proof. This is left as an exercise to the reader in Exercise 9. □

Remark 5.1.6. Suppose that $\varphi : F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is a Frobenius splitting. It is easy to see that a sort of converse to Proposition 5.1.5(a) holds. In particular, suppose there is a commutative diagram

$$\begin{array}{ccc} F^e_* \mathcal{O}_X & \xrightarrow{\ \varphi\ } & \mathcal{O}_X \\ \downarrow & & \downarrow \\ F^e_* (\mathcal{O}_X / J) & \xrightarrow[\ \varphi/J\ ]{} & \mathcal{O}_X / J \end{array}$$

then J is φ-compatibly split (simply take the kernel of the vertical arrows).

One important point about Frobenius splittings is that compatibly split subvarieties intersect normally. In particular

Corollary 5.1.7. *If $\varphi : F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is a Frobenius splitting, if I and J are compatibly φ-split, then $I + J$ is a radical ideal.*

Proof. Combine properties (a) and (c) from Proposition 5.1.5. □

Also see Exercise 3.

5.2 Global Properties of Frobenius Split Varieties

Now we turn to projective (or more generally complete) Frobenius split varieties. First we introduce another definition.

Definition 5.2.1. Suppose that D is an effective Weil divisor on a normal variety X. Then we say that X is *e-Frobenius split relative to D* if the composition

$$\mathcal{O}_X \to F^e_* \mathcal{O}_X \hookrightarrow F^e_*(\mathcal{O}_X(D))$$

is split.

Notice that if X is e-Frobenius split relative to D, then X is Frobenius split. We mentioned earlier that regular affine varieties are Frobenius split Proposition 5.1.2, but not every smooth projective variety is Frobenius split. We prove that now.

Lemma 5.2.2. *If X is proper, Frobenius split and normal, then $H^0(X, \mathcal{O}_X(-nK_X)) \neq 0$ for some $n > 0$. In particular X is not of general type. Even more, if X is e-Frobenius split relative to an ample divisor A, then $-K_X$ is big.*

Proof. The fact that X is Frobenius split implies that there is a nonzero element $\varphi \in \mathrm{Hom}_X(F^e_*\mathcal{O}_X, \mathcal{O}_X) \cong H^0(X, F^e_*\mathcal{O}_X((1-p^e)K_X))$ by Sect. 4. In particular, $H^0(X, F^e_*\mathcal{O}_X((1-p^e)K_X)) \neq 0$. But $F^e_*\mathcal{O}_X((1-p^e)K_X)$ is isomorphic to $\mathcal{O}_X((1-p^e)K_X)$ as an Abelian group, and so the result follows for $n = (p^e - 1)$.

For the second statement, we notice that we have a section $\varphi \in \mathrm{Hom}_X(F^e_*\mathcal{O}_X(D), \mathcal{O}_X) \cong H^0(X, F^e_*\mathcal{O}_X((1-p^e)K_X - A))$, and so there is an effective divisor $H \sim (1-p^e)K_X - A$, and thus $(1-p^e)K_X \sim A + H =$ "ample + effective," and so K_X is big.[5] □

Our next goal is to prove vanishing theorems for Frobenius split varieties. First, however, we need the following lemma.

Lemma 5.2.3. *If X is e-Frobenius split relative to D, then for any integer $n > 0$, X is ne-Frobenius split relative to $(p^{(n-1)e} + \cdots + p^e + 1)D$.*

Proof. Suppose that $\mathcal{O}_X \to F^e_*\mathcal{O}_X \to F^e_*\mathcal{O}_X(D) \xrightarrow{\varphi} \mathcal{O}_X$ is the Frobenius splitting. By tensoring this with D, taking the reflexification of the sheaves, and applying the functor F^e_*, we obtain a splitting

$$F^e_*(\mathcal{O}_X(D)) \to F^e_*(\mathcal{O}_X(p^eD)) \to F^{2e}_*(\mathcal{O}_X(D + p^eD)) \xrightarrow{F^e_*\varphi(D)} F^e_*(\mathcal{O}_X(D)).$$

But now composing with Frobenius and φ on the left and right sides respectively, we obtain our desired splitting

[5] On a projective variety X, you can take the definition of *big* to be a divisor which has a multiple which is linearly equivalent to an ample divisor plus an effective divisor [55, Corollary 2.2.7].

$$\begin{aligned}\mathcal{O}_X &\to F^e_*(\mathcal{O}_X(D)) \to F^e_*(\mathcal{O}_X(p^e D)) \to F^{2e}_*(\mathcal{O}_X(D + p^e D)) \\ &\xrightarrow{F^e_*\varphi(D)} F^e_*(\mathcal{O}_X(D)) \xrightarrow{\varphi} \mathcal{O}_X.\end{aligned}$$

Continuing in this way yields the desired result. □

Theorem 5.2.4 (Vanishing theorems for Frobenius split varieties). *Suppose that X is a projective Frobenius split variety. Then:*

(a) $H^i(X, \mathscr{L}) = 0$ for any ample line bundle $\mathscr{L}$ and any $i > 0$.
(b) $H^i(X, \mathscr{L} \otimes \omega_X) = 0$ for any ample line bundle $\mathscr{L}$ and any $i > 0$.
(c) If X is normal and e-Frobenius split relative to an ample Cartier divisor D, then $H^i(X, \mathscr{L}) = 0$ for any nef line bundle $\mathscr{L}$ and any $i > 0$.
(d) If X is normal and e-Frobenius split relative to an ample Cartier divisor D such that $X \setminus D$ is regular, then $H^i(X, \mathscr{L} \otimes \omega_X) = 0$ for any big and nef line bundle $\mathscr{L}$ and any $i > 0$.

Proof. For (a), notice that we have a splitting of $\mathscr{L} \cong \mathcal{O}_X \otimes \mathscr{L} \to (F^e_*\mathcal{O}_X) \otimes \mathscr{L} \cong F^e_*\mathscr{L}^{p^e}$. Thus $H^i(X, \mathscr{L}) \hookrightarrow H^i(X, F^e_*\mathscr{L}^{p^e})$ injects. On the other hand, $H^i(X, F^e_*\mathscr{L}^{p^e}) \cong H^i(X, \mathscr{L}^{p^e})$ as Abelian groups, and the latter vanishes for $i > 0$ and $e \gg 0$ by Serre vanishing.

For (b), notice that an application of $\mathscr{H}\mathrm{om}_{\mathcal{O}_X}(\underline{\quad}, \omega_X)$ to the splitting $\mathcal{O}_X \to F^e_*\mathcal{O}_X \to \mathcal{O}_X$ induces a splitting:

$$\omega_X \xleftarrow{T} F^e_*\omega_X \hookleftarrow \omega_X.$$

Twisting by $\mathscr{L}$ and applying the projection formula gives us

$$\omega_X \otimes \mathscr{L} \xleftarrow{T} F^e_*(\omega_X \otimes \mathscr{L}^{p^e}) \hookleftarrow \omega_X \otimes \mathscr{L}.$$

Taking cohomology for $i > 0$ we obtain maps

$$H^i(X, \omega_X \otimes \mathscr{L}) \xleftarrow{T} H^i(X, F^e_*(\omega_X \otimes \mathscr{L}^{p^e})) \hookleftarrow H^i(X, \omega_X \otimes \mathscr{L})$$

whose composition is an isomorphism. But the middle term vanishes by Serre vanishing since we may take $e \gg 0$.

For (c), we first notice that by using Lemma 5.2.3, we may assume that D is as ample as we wish (at the expense of increasing e). Thus, using the same strategy as in (a), it is sufficient to prove that $H^i(X, \mathcal{O}_X(D) \otimes \mathscr{L}^{p^e}) = 0$ for all $i > 0$. But this follows from Fujita's vanishing theorem [21].

Part (d) is left as a somewhat involved exercise to the reader Exercise* 5.10. □

Finally, we notice that sections on Frobenius split subvarieties often extend to sections on the ambient spaces.

Theorem 5.2.5. *Suppose that $Y \subseteq X$ is compatibly Frobenius split. Then the natural maps*

$$H^0(X, \mathscr{L}) \to H^0(Y, \mathscr{L}|_Y)$$

are surjective for any ample line bundle $\mathscr{L}$.

Proof. By composition of the Frobenius splitting with itself, we have the following diagram for any $e > 0$.

$$\begin{array}{ccccc} H^0(X, F^e_*(\mathscr{L}^{p^e})) & \xrightarrow{\beta} & H^0(X, F^e_*(\mathscr{L}^{p^e}|Y)) & \longrightarrow & H^1(X, F^e_*(I_Y \otimes \mathscr{L}^{p^e})) = 0 \\ \downarrow & & \downarrow & & \\ H^0(X, \mathscr{L}) & \xrightarrow[\alpha]{} & H^0(X, \mathscr{L}|_Y). & & \end{array}$$

Note, we have the top-right vanishing by Serre vanishing which implies that β is surjective. The vertical maps are surjective because they are obtained from twisting the Frobenius splitting $F^e_*\mathcal{O}_X \to \mathcal{O}_X$ by $\mathscr{L}$. The diagram then implies that α is surjective, this completes the proof. □

5.3 Tools for Proving Proper Varieties Are Frobenius Split

There are two common tools for proving that proper varieties are Frobenius split. The first involves a study of the singularities of sections of $H^0(X, \mathcal{O}_X((1-p^e)K_X))$. The second is a general fact that images of Frobenius split varieties often remain Frobenius split. In many applications, these tools are combined.

Theorem 5.3.1 ([66], [11, Sect. 1.3]). *Suppose X is a proper normal d-dimensional variety of finite type over an algebraically closed field of characteristic $p > 0$. Further suppose that there is an effective divisor D, linearly equivalent to $(1 - p^e)K_X$ for some e, that satisfies the following condition:*

- *There exists a smooth point $x \in X$ and divisors $D_1, D_2, \ldots, D_d$ intersecting in a simple normal crossings divisor at $x \in X$ such that $D = (p^e - 1)D_1 + \cdots + (p^e - 1)D_d + G$ for some effective divisor G not passing through $x \in X$.*

Then X is Frobenius split by a map $\varphi : F^e_*\mathcal{O}_X \to \mathcal{O}_X$ which corresponds to D as in (12).

Proof. There are two main ideas in this proof.

(a) D corresponds to some map, $\varphi : F^e_*\mathcal{O}_X \to \mathcal{O}_X$ by (12). Thus $\varphi(F^e_*1) = \lambda \in H^0(X, \mathcal{O}_X) = k$ is a constant. If we can show that $\lambda \neq 0$, then by rescaling φ, we are done.

(b) The value of $\varphi(F^e_*1)$ can be detected at any point. In particular, we can try to compute it at the stalk of $x \in X$.

For simplicity, we denote the stalk at x by $R := \mathcal{O}_{X,x}$ and we use $\mathfrak{m}$ to denote the maximal ideal.

Fix φ corresponding to D as in (12) and consider $\varphi_x : F^e_* R \to R$. Suppose that

$$D|_{\operatorname{Spec} R} = V(f_1^{p^e-1} \cdots f_d^{p^e-1}) = V(f)$$

where the f_i are the local equations for D_i near x.

Set $\widehat{R}$ to be the completion of $R = \mathcal{O}_{X,x}$. We know that φ corresponds to D, so it can be factored as

$$F^e_* \mathcal{O}_X((1-p^e)K_X - D) \hookrightarrow F^e_* \mathcal{O}_X((1-p^e)K_X) \to \mathcal{O}_X.$$

Taking the completion of this factorization, we obtain

$$F^e_* \widehat{R} \xrightarrow{\cdot (F^e_* f)} F^e_* \widehat{R} \xrightarrow{\psi} \widehat{R}, \qquad \widehat{\varphi} : F^e_* \widehat{R} \to \widehat{R}.$$

By construction, ψ, viewed as an element of $M = \operatorname{Hom}(F^e_* \widehat{R}, \widehat{R})$, generates M as an $F^e_* \widehat{R}$-module (use Exercise 9).

On the other hand, $\widehat{R} = k[\![f_1, \ldots, f_d]\!]$, and so the map $\Psi : F^e_* \widehat{R} \to \widehat{R}$ which sends $f = f_1^{p^e-1} \cdots f_d^{p^e-1}$ to 1 and the other basis monomials $\{f_1^{a_1} \cdots f_d^{a_d} \neq f \mid 0 \leq a_i \leq p^e - 1\}$ to zero also generates M as an $F^e_* \widehat{R}$-module by Example 4.0.2.

It follows that $\psi(F^e_* \_\_) = \Psi(F^e_*(c \cdot \_\_))$ for some invertible element $c \in \widehat{R}$. But notice that c is invertible, so it has a non-zero constant term $c_0 \in k$ where $c = c_0 + c'$, $c' \in \langle f_1, \ldots, f_d \rangle_{\widehat{R}}$. Thus

$$\begin{aligned} \lambda &= \varphi_x(F^e_* 1) \\ &= \widehat{\varphi}(F^e_* 1) \\ &= \psi(F^e_* f) \\ &= \Psi(F^e_*(c \cdot f)) \\ &= \Psi(F^e_*(c_0 \cdot f)) + \Psi(F^e_*(c' \cdot f)) \\ &= c_0^{1/p^e} + \Psi(F^e_*(c' \cdot f)). \end{aligned}$$

But $\Psi(F^e_*(c' \cdot f)) \in \langle f_1, \ldots, f_d \rangle_{\widehat{R}}$ by our choice of Ψ (note that $c' \cdot f \in \langle f_1^{p^e}, \ldots, f_d^{p^e} \rangle$). Since $c_0^{1/p^e} + \Psi(F^e_*(c' \cdot f)) = \lambda \in k$ is a constant, we see that $\Psi(F^e_*(c' \cdot f)) = 0$. Thus, $\lambda = c_0^{1/p^e} \neq 0$ as desired. □

Remark 5.3.2. A more general, simpler, and more conceptual version of the above result is described in Exercise 12 in the next section. We lack the language to describe it here however.

Now we study the behavior of Frobenius splittings under maps between varieties. We will study some complementary constructions later in Sect. 7.

Theorem 5.3.3 ([40, 66]). *Suppose that $\pi : Y \to X$ is a map of varieties such that $\mathcal{O}_X \to \pi_* \mathcal{O}_Y$ splits as a map of $\mathcal{O}_X$-modules (e.g., if $\pi_* \mathcal{O}_Y = \mathcal{O}_X$). Then if Y is Frobenius split, so is X.*

Before proving the theorem, we point out just how common the condition that $\mathcal{O}_X \to \pi_* \mathcal{O}_Y$ splits is. Indeed, if $\pi : Y \to X$ is a proper surjective map between normal varieties with connected fibers, then $\pi_* \mathcal{O}_Y = \mathcal{O}_X$. Alternately, if $\pi : Y \to X$ is proper, dominant, generically finite, Y and X are normal, and p does not divide $[\mathscr{K}(Y) : \mathscr{K}(X)] = n$, then the normalized field trace $\frac{1}{n}\mathrm{Tr} : \mathscr{K}(Y) \to \mathscr{K}(X)$ restricts to a map $\pi_* \mathcal{O}_Y \to \mathcal{O}_X$ which sends 1 to 1.

Proof of Theorem 5.3.3. Set $\varphi : F^e_* \mathcal{O}_Y \to \mathcal{O}_Y$ to be the Frobenius splitting of Y, and fix $\alpha : \pi_* \mathcal{O}_Y \to \mathcal{O}_X$ to be the splitting of $i : \mathcal{O}_X \to \pi_* \mathcal{O}_Y$. Pushing down φ we obtain

$$(\pi_* \varphi) : \pi_* F^e_* \mathcal{O}_Y \to \pi_* \mathcal{O}_Y .$$

Now, we simply form the composition:

$$F^e_* \mathcal{O}_X \xhookrightarrow{F^e_* i} F^e_* \pi_* \mathcal{O}_Y = \pi_* F^e_* \mathcal{O}_Y \xrightarrow{\pi_* \varphi} \pi_* \mathcal{O}_Y \xrightarrow{\alpha} \mathcal{O}_X .$$

By chasing through the composition, we see that $F^e_* 1$ is sent to 1 and that X is F-split. ☐

5.4 Exercises

Exercise 1. Prove Lemma 5.0.3.
Hint: Compose Frobenius and Frobenius splittings by using the *functor* F^e_*.

Exercise 2. Prove Lemma 5.0.4.

Exercise 3. A domain R containing a field of characteristic $p > 0$ is said to be *weakly normal* if any $r \in K(R)$ satisfying $r^p \in R$ also satisfies $r \in R$ as well (see [96, Lemma 3] and [95, Sect. 3]). Show that any F-pure/split R is weakly normal. You can find a solution in [11, Proposition 1.2.5], cf. [40, Proposition 5.31].

Exercise 4. Suppose that X is Frobenius split relative to a Cartier divisor D such that $X \setminus D$ is Cohen–Macaulay. Prove that X is Cohen–Macaulay.

Hint: Working locally we may assume that $X = \operatorname{Spec} R$ and $D = V(f)$. Fix a maximal ideal $\mathfrak{m} \in \operatorname{Spec} R$ and consider the composition $H^i_{\mathfrak{m}}(R) \to H^i_{\mathfrak{m}}(F^e_* R) \xrightarrow{\cdot (F^e_* f)} H^i_{\mathfrak{m}}(F^e_* R)$ recalling that a variety can be proven to be Cohen–Macaulay by examining its local cohomology modules as in [34, Chap. III, Exercises 3.3 and 3.4].

Exercise 5. Suppose that $F^e_* R \cong R \oplus M$ as R-modules where M is some arbitrary R-module. Prove that R is Frobenius split. More generally, prove the same result if there is any surjective map $F^e_* R \to R$.

Exercise 6. Suppose that $X = \operatorname{Spec} R$ is an affine variety and suppose that for every maximal ideal $\mathfrak{m} \in \operatorname{Spec} R$, we have that $R_{\mathfrak{m}}$ is F-split. Prove that X is F-split.

Hint: The given splittings definitely do not glue. However, consider the evaluation-at-1 map $\operatorname{Hom}_R(F_* R, R) \to R$.

Exercise 7 (Toric varieties). Suppose that X is a normal toric variety. Consider the map $\Psi : F_* \mathcal{O}_X \to \mathcal{O}_X$ defined as follows. We define

$$\Psi(F_* \mathbf{x}^\lambda) = \begin{cases} \mathbf{x}^{\lambda/p} & \text{if } \lambda/p \text{ has integer entries} \\ 0 & \text{otherwise} \end{cases}$$

acting on each affine toric chart (where $\mathbf{x}^\lambda$ is a monomial). Show that this induces a Frobenius splitting on X which compatibly splits all the torus invariant divisors. What is the Δ_Ψ (as defined as in (13))?

Exercise 8 (Affine section rings). Suppose that X is a projective algebraic variety with ample line bundle $\mathscr{A}$. Consider

$$S := \bigoplus_{i \in \mathbb{Z}} H^0(X, \mathscr{A}^i),$$

the section right with respect to $\mathscr{A}$. Prove that X is Frobenius split if and only if S is Frobenius split. For additional discussion of related topics, see [89].

Exercise 9. Prove Proposition 5.1.5.

Hint: For part (a), use a diagram similar to the one in Lemma 5.1.3. For solutions, see [11, Chap. 1].

Exercise* 5.10. Prove Theorem 5.2.4(d).

Hint: This is somewhat involved. There exists a Cartier divisor B such that $\mathscr{L}^n(-B)$ is ample for all $n \gg 0$ since $\mathscr{L}$ is big and nef. For some $m \gg 0$, we also know that $mD + B$ is still ample. First show that X is r-Frobenius split relative to $mD + B$ for some integer $r \gg 0$ (this is hard). Then notice we have a composition

$$\omega_X \otimes \mathscr{L} \xleftarrow{T} F^r_*(\omega_X \otimes \mathscr{L}) \hookleftarrow F^r_*(\omega_X(-B) \otimes \mathscr{L}) \hookleftarrow F^r_*(\omega_X(-mD-B) \otimes \mathscr{L}) \leftarrow \omega_X$$

which is an isomorphism (we type this with the arrows going backwards to suggest that this arises by duality). Now, by composing the map $\omega_X \leftarrow F^r_*(\omega_X(-B) \otimes \mathscr{L})$ with itself as in (17), we can obtain the desired vanishing. For a solution, see [78, Theorem 6.8].

6 Frobenius Non-splittings

Our goal in this section is to develop a theory for p^{-1}-linear maps generalizing the theory of Frobenius split varieties demonstrated in the previous section. First we start with a definition.

Definition 6.0.1. Suppose that we are given a line bundle $\mathscr{L}$ on a variety X. Consider an $\mathcal{O}_X$-linear map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$. We say that an ideal J is *φ-compatible* if we have that $\varphi(F^e_*(J \cdot \mathscr{L})) \subseteq J$. If $Y = V(J) \subseteq X$, then we say that Y is *φ-compatible* if J is.

For example, if $\mathscr{L} = \mathcal{O}_X$ and φ is a Frobenius splitting, then any φ-compatibly split ideal is φ-compatible. We also have a slight variation on this definition.

Definition 6.0.2. Given Δ corresponding to $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ as in (16). A subvariety $Y \subseteq X$ is called an *F-pure center of* (X, Δ) if Y is φ-compatible and φ_η is surjective where η is the generic point of Y.

Lemma 6.0.3. *If $J \subseteq \mathcal{O}_X$ is an ideal sheaf, then J is $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ compatible if and only if φ induces a map $\varphi_Y : F^e_*(\mathscr{L}|_Y) \to \mathcal{O}_Y$.*

Proof. Left as an exercise to the reader Exercise 1. □

We explore compatibility after composing maps as in (17).

Lemma 6.0.4. *Suppose that $J \subseteq \mathcal{O}_X$ is $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$-compatible. Then J is*

$$\varphi^n : F^{ne}_* \mathscr{L}^{p^{e(n-1)} + \cdots + 1} \to \mathcal{O}_X$$

compatible for all $n > 0$. Conversely, suppose that φ is surjective. If J is φ^n-compatible, then J is also φ-compatible.

Proof. The statement is local, so we may as well only check this at the stalks $\mathcal{O}_{X,x}$ and in particular assume that $\mathscr{L} \cong \mathcal{O}_{X,x}$. The first statement is obvious and will be left to the reader. For the second statement, we sketch the idea of the proof.

The first step is to show that any $J \subseteq \mathcal{O}_{X,x}$ which is $\varphi : F^e_* \mathcal{O}_{X,x} \twoheadrightarrow \mathcal{O}_{X,x}$-compatible is also radical (see Exercise *6.3). One can then show it is sufficient to verify the statement at the minimal primes of J. In particular, we can assume that J is the maximal ideal of $\mathcal{O}_{X,x}$ by localizing.

Now then, suppose that J is φ^n compatible but not φ-compatible. Then $\varphi(F^e_* J) = \mathcal{O}_{X,x}$ (since otherwise, it would be in the maximal ideal, which coincides with J). But then it is easy to see that $\varphi^2(F^{2e}_* J) = \varphi(F^e_* \varphi(F^e_* J)) = \mathcal{O}_{X,x}$ as well. Continuing in this way, we obtain a contradiction. □

We also generalize the notion of F-pure to non-Frobenius splittings and to pairs.

Definition 6.0.5. Suppose that X is a normal variety and that Δ is a $\mathbb{Q}$-divisor such that

$$K_X + \Delta \text{ is a } \mathbb{Q}\text{-Cartier divisor with index not divisible by } p. \qquad (\dagger)$$

We say that (X, Δ) is *sharply F-pure* if the map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$, corresponding to Δ as in (15), is surjective as a map of $\mathcal{O}_X$-modules.

If we do not satisfy (†), then we say that (X, Δ) is *sharply F-pure* if for every point $x \in X$, there exists a neighborhood U of $x \in X$ and a divisor Δ_U on U such that $\Delta_U \geq \Delta|_U$ and such that (U, Δ_U) is sharply F-pure in the above sense.

It is an exercise below, Exercise 5, that the definition of sharply F-pure above and the definition given in Definition 5.0.1 coincide.

6.1 Global Considerations

In this section, we briefly demonstrate that some of the *global* methods from the Frobenius splitting section can still bear fruit, even if the actual vanishing theorems do not hold.

Our first goal is to consider a generalization of a proof due to Keeler [49] (also independently obtained by N. Hara [unpublished]). Related results were first proven by [88] and also [26]. Before doing that, we recall a Definition and a Lemma.

Definition 6.1.1 (Castelnuovo–Mumford regularity [55, Sect. 1.8]). Suppose that $\mathscr{F}$ is a coherent sheaf on a projective variety X and that $\mathscr{A}$ is a globally generated ample divisor on X. Then $\mathscr{F}$ is called 0-*regular (with respect to $\mathscr{A}$)* if $H^i(X, \mathscr{F} \otimes \mathscr{A}^{-i}) = 0$ for all $i > 0$.

Lemma 6.1.2 (Mumford's theorem [55, Theorem 1.8.5]). *If $\mathscr{F}$ is 0-regular with respect to a globally generated ample line bundle $\mathscr{A}$, then $\mathscr{F}$ is globally generated.*

Now we are in a position to prove that certain sheaves are globally generated.

Theorem 6.1.3 ([49, 75]). *Suppose that $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ is a* surjective *$\mathcal{O}_X$-linear map and $\mathscr{L}$ is a line bundle. Additionally suppose that $\mathscr{A}$ is a globally generated ample line bundle and that $\mathscr{M}$ is any other line bundle such that $\mathscr{L} \otimes \mathscr{M}^{p^e-1}$ is ample (e.g., if $\mathscr{L}$ is itself ample, then we may take $\mathscr{M} = \mathcal{O}_X$). In this case, the line bundle*

$$\mathscr{M} \otimes \mathscr{A}^{\dim X}$$

is globally generated.

Proof. Choose $n \gg 0$. Then we have a surjective map:

$$\varphi^n : F^{ne}_* \mathscr{L}^{p^{(n-1)e} + \cdots + p^e + 1} \to \mathcal{O}_X$$

from (17). Twisting by $\mathscr{M} \otimes \mathscr{A}^{\dim X}$, we obtain a surjective map:

$$F^{ne}_* (\mathscr{L}^{p^{(n-1)e} + \cdots + p^e + 1} \otimes \mathscr{M}^{p^{ne}} \otimes \mathscr{A}^{p^{ne} \dim X}) \to \mathscr{M} \otimes \mathscr{A}^{\dim X}.$$

It is sufficient to show that the left side is globally generated as an $\mathcal{O}_X$-module since then the right side is a quotient of a globally generated module and thus globally generated itself. Note that it is definitely *not* sufficient to show that the left side is globally generated as an $F^{ne}_* \mathcal{O}_X$-module. We will proceed by proving that the left side is 0-regular as an $\mathcal{O}_X$-module. Note

$$\mathscr{L}^{p^{(n-1)e}+\cdots+p^e+1} \otimes \mathscr{M}^{p^{ne}} = (\mathscr{L} \otimes \mathscr{M}^{p^e-1})^{p^{(n-1)e}+\cdots+p^e+1} \otimes \mathscr{M}.$$

But now we have

$$\begin{aligned} & H^i\Big(X, F^{ne}_*\big((\mathscr{L} \otimes \mathscr{M}^{p^e-1})^{p^{(n-1)e}+\cdots+p^e+1} \otimes \mathscr{M} \otimes \mathscr{A}^{p^{ne}\dim X}\big) \otimes \mathscr{A}^{-i}\Big) \\ = {} & H^i\Big(X, F^{ne}_*\big((\mathscr{L} \otimes \mathscr{M}^{p^e-1})^{p^{(n-1)e}+\cdots+p^e+1} \otimes \mathscr{M} \otimes \mathscr{A}^{p^{ne}(\dim X - i)}\big)\Big) \\ = {} & H^i\Big(X, F^{ne}_*\big((\mathscr{L} \otimes \mathscr{M}^{p^e-1} \otimes \mathscr{A}^{(\dim X - i)(p^e-1)})^{p^{(n-1)e}+\cdots+p^e+1} \\ & \otimes (\mathscr{M} \otimes \mathscr{A}^{\dim X - i})\big)\Big). \end{aligned}$$

We already have the vanishing for $i > \dim X$. Now the F^{ne}_* does not effect the vanishing or nonvanishing of the cohomology since it does not change the underlying sheaf of Abelian groups. Therefore, the above cohomology groups vanish by Serre vanishing, since $\mathscr{L} \otimes \mathscr{M}^{p^e-1} \otimes \mathscr{A}^{(\dim X - i)(p^e-1)}$ is ample and each of the *finitely many* $\mathscr{M} \otimes \mathscr{A}^{\dim X - i}$ are coherent sheaves. □

Example 6.1.4. If X is smooth (or even F-pure), then there is always a *surjective* map $F^e_* \mathcal{O}_X((1-p^e)K_X) \to \mathcal{O}_X$. It follows that if M is a divisor such that $M - K_X$ is ample, and $\mathscr{A}$ is any globally generated ample line bundle, then $\mathcal{O}_X(M) \otimes \mathscr{A}^{\dim X}$ is globally generated.

Remark 6.1.5. It is worth pointing out that not only is $\mathscr{M} \otimes \mathscr{A}^{\dim X}$ globally generated, one even has that it is globally generated by the image of the map

$$H^0\Big(X, F^{ne}_*\big(\mathscr{L}^{p^{(n-1)e}+\cdots+p^e+1} \otimes \mathscr{M}^{p^{ne}} \otimes \mathscr{A}^{p^{ne}\dim X}\big)\Big) \to H^0\big(X, \mathscr{M} \otimes \mathscr{A}^{\dim X}\big).$$

This special sub-vector space of global sections also behaves well with respect to restriction to compatible subvarieties as we shall see shortly.

Similar arguments to those in the proof Theorem 6.1.3 also yield the following result.

Proposition 6.1.6 ([26, 49, 88]). *If X is any F-pure variety, $\mathscr{A}$ is a globally generated ample line bundle, and $\mathscr{M}$ is any other ample line bundle, then*

$$\omega_X \otimes \mathscr{A}^{\dim X} \otimes \mathscr{M}$$

is globally generated.

Proof. The proof is left to the reader in Exercise 7. □

Finally, we also remark that compatible ideals also play a special role with regard to lifting of sections.

Theorem 6.1.7. *Suppose that $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ is an $\mathcal{O}_X$-linear map and that $J \subseteq \mathcal{O}_X$ is φ-compatible. Set $Y = V(J)$ and set $\varphi_Y : F^e_*(\mathscr{L}|_Y) \to \mathcal{O}_Y$ to be the map φ restricted to Y as in Lemma 6.0.3. Suppose that $\mathscr{H}$ is a line bundle on X such that $\mathscr{H}^{p^e-1} \otimes \mathscr{L}$ is ample and also such that the map induced by φ_Y*

$$H^0\big(Y, F^{ne}_*((\mathscr{L}^{p^{(n-1)e}+\dots+p^e+1} \otimes \mathscr{H}^{p^{ne}})|_Y)\big) \xrightarrow{\gamma} H^0(Y, \mathscr{H}|_Y) \tag{19}$$

is nonzero for some $n \gg 0$. Then $H^0(X, \mathscr{H}) \neq 0$ as well. Even more, the sections in the image of γ all extend to sections on $H^0(X, \mathscr{H})$.

Before starting the proof, let us note some conditions under which the map γ is nonzero. For example, if $\mathscr{L}|_Y = \mathcal{O}_Y$ and φ_Y is a Frobenius splitting, then γ is in fact surjective (e.g., if Y is a point and φ_Y is nonzero). Alternately, if φ_Y is surjective and also $\mathscr{H}|_Y = \mathscr{A}^{\dim Y} \otimes \mathscr{M}$ where $\mathscr{A}$ is a globally generated ample line bundle on Y and $\mathscr{M}^{p^e-1} \otimes \mathscr{L}|_Y$ is ample on Y, then we can apply Theorem 6.1.3. In the case that Y is a curve, see Exercise 2.

Proof. We fix $n \gg 0$; for simplicity of notation set $\eta = p^{(n-1)e} + \dots + p^e + 1$ and consider the following diagram:

$$\begin{array}{ccccc}
H^0\big(X, F^{ne}_*(\mathscr{L}^\eta \otimes \mathscr{H}^{p^{ne}})\big) & \longrightarrow & H^0\big(Y, F^{ne}_*((\mathscr{L}^\eta \otimes \mathscr{H}^{p^{ne}})|_Y)\big) & \longrightarrow & H^1\big(X, F^{ne}_*(J \otimes \mathscr{L}^\eta \otimes \mathscr{H}^{p^{ne}})\big) \\
\downarrow{\scriptstyle\varphi} & & \downarrow{\scriptstyle\varphi_Y} & & \downarrow \\
H^0(X, \mathscr{H}) & \longrightarrow & H^0(Y, \mathscr{H}|_Y) & \longrightarrow & H^1(X, J \otimes \mathscr{H}).
\end{array}$$

However, note that

$$H^1\big(X, F^{ne}_*(J \otimes \mathscr{L}^\eta \otimes \mathscr{H}^{p^{ne}})\big) = H^1\big(X, F^{ne}_*(J \otimes \mathscr{H} \otimes (\mathscr{L} \otimes \mathscr{H}^{p^e-1})^\eta\big) = 0$$

by Serre vanishing since the F^{ne}_* does not effect the underlying sheaf of Abelian groups. ☐

6.2 Fedder's Lemma

We now delve into the *local* theory of p^{-e}-linear maps and in particular state *Fedder's lemma*. This is a particularly effective tool for explicitly writing down these maps and also for identifying which of them are surjective.

Suppose that $S = k[x_1, \dots, x_n]$ and $R = S/I$ for some ideal $I \subseteq R$. The point is that if $R = S/I$, then maps $\bar{\varphi} : F^e_* R \to R$ come from maps $\varphi : F^e_* S \to S$, which Fedder's lemma *precisely* identifies. Set $\Phi_S : F^e_* S \to S$ to be the map which generates $\mathrm{Hom}_S(F^e_* S, S)$ as an $F^e_* S$-module as identified in Example 4.0.2. Recall that Φ_S sends the monomial $F^e_*(x_1^{p^e-1} \dots x_n^{p^e-1})$ to 1 and all other basis monomials to zero.

Lemma 6.2.1 (Fedder's Lemma [20, Lemma 1.6]). *With $S \supseteq I$, R and Φ_S as above, then*

$$\left\{ \begin{array}{c} \textit{Maps } \varphi \in \mathrm{Hom}_S(F^e_* S, S) \\ \textit{compatible with } I \end{array} \right\} = \left\{ \begin{array}{c} \varphi \mid \varphi(F^e_* \_\_) = \Phi_S(F^e_*(z \cdot \_\_)), \\ \textit{for some } z \in I^{[p^e]} : I \end{array} \right\}.$$

More generally, there is an isomorphism of S-modules

$$\mathrm{Hom}_R(F^e_* R, R) \longleftrightarrow \frac{\left(F^e_*(I^{[p^e]} : I)\right) \cdot \Phi_S}{\left(F^e_* I^{[p^e]}\right) \cdot \Phi_S}$$

induced by restricting $\psi \in (F^e_*(I^{[p^e]} : I)) \cdot \Phi_S \subseteq \mathrm{Hom}_S(F^e_* S, S)$ to $R = S/I$ as in Lemma 6.0.3.

Finally, for any point $\mathfrak{q} \in V(I) \subseteq \mathrm{Spec}\, S$, there exists a map $\varphi \in \mathrm{Hom}_R(F^e_* R, R)$ which is surjective at $\mathfrak{q}/I \in \mathrm{Spec}\, R$ if and only if $I^{[p^e]} : I \not\subseteq \mathfrak{q}^{[p^e]}$. In other words, R is F-pure in a neighborhood of $\mathfrak{q}$ if and only if $I^{[p]} : I \not\subseteq \mathfrak{q}^{[p]}$.

Proof. There are a lot of statements here. First we notice that any map of the form $\varphi(F^e_* \_\_) = \Phi_S(F^e_*(z \cdot \_\_))$ for some $z \in I^{[p^e]} : I$ is clearly compatible with I since

$$\Phi_S(F^e_*(z \cdot I)) \subseteq \Phi_S(F^e_*(I^{[p^e]} : I) \cdot I) \subseteq \Phi_S(F^e_* I^{[p^e]}) = I \cdot \Phi_S(F^e_* S) = I.$$

This gives us the containment $\supseteq$ in the first equality. For the other containment, we first prove the following claim.

Claim 6.2.2. *For ideals $I, J \subseteq S$ we have*

$$\Phi_S(F^e_* J) \subseteq I$$

if and only if $J \subseteq I^{[p^e]}$.

Proof of Claim. Certainly the *if* direction is obvious, so suppose then that $\Phi_S(F^e_* J) \subseteq I$. This implies that $\varphi(F^e_* J) \subseteq I$ for every $\varphi \in \mathrm{Hom}_S(F^e_* S, S)$ since Φ_S generates that set as an $F^e_* S$-module. But $F^e_* S$ is a free S-module of rank p^{en}, so we see that

$$F^e_* J \subseteq \underbrace{I \oplus \cdots \oplus I}_{p^{ne}-\text{times}}$$

since we could take the φ as the various projections. Now, $I \oplus \cdots \oplus I = I \cdot (F^e_* S) = F^e_* I^{[p^e]}$. This proves the claim. □

Now we return to the proof of Fedder's lemma. We observe that if $\varphi(F^e_* \_\_) = \Phi_S(F^e_*(z \cdot \_\_))$ is I-compatible, then $z \cdot I \subseteq I^{[p^e]}$ by the claim, which proves that $z \in I^{[p^e]} : I$, and so the equality is proven.

Now we come to the bijection. We certainly have a natural map

$$\Lambda : (F^e_*(I^{[p^e]} : I)) \cdot \Phi_S \to \operatorname{Hom}_R(F^e_* R, R)$$

induced by sending $F^e_* z$ first to $(F^e_* z) \cdot \Phi_S(\_\_) = \Phi_S((F^e_* z) \cdot \_\_)$ and then second, inducing a map in $\operatorname{Hom}_R(F^e_* R, R)$ as in Lemma 6.0.3. The kernel of Λ is $(F^e_* I^{[p^e]}) \cdot \Phi_S$ by the claim, and so we only need to show that this map is surjective.

Given $\varphi \in \operatorname{Hom}_R(F^e_* R, R) = \operatorname{Hom}_S(F^e_* R, R)$, consider the following diagram of S-linear maps where the horizontal maps are the canonical surjections:

$$\begin{array}{ccc} F^e_* S & \twoheadrightarrow & F^e_*(R/I) \\ \exists\psi \downarrow & & \downarrow \varphi \\ S & \twoheadrightarrow & (R/I). \end{array}$$

Because $F^e_* S$ is a free (and so projective) S-module, the dotted map ψ exists. By construction, ψ is compatible with I. By the earlier parts of the theorem, ψ corresponds to a $z \in I^{[p^e]} : I$ which restricts to φ, completing the proof of the bijection.

The last part of the theorem is left as an exercise to the reader. □

Remark 6.2.3 (Regular local rings are fine). The proof given above goes through without change if one assumes that S is a regular local[6] ring *instead* of assuming that S is a polynomial ring.

One of the most important corollaries of this is the following.

Corollary 6.2.4. *Given $f \in k[x_1, \ldots, x_n] = S$, then $S/\langle f \rangle$ is F-split in a neighborhood of the origin if and only if $f^{p-1} \notin \langle x_1^p, \ldots, x_n^p \rangle$.*

Proof. Note that $S/\langle f \rangle = R$ is F-split if and only if there exists a surjective map $\varphi \in \operatorname{Hom}_R(F^e_* R, R)$ by Exercise 5. The result then follows from Fedder's lemma since $\langle f^p \rangle : \langle f \rangle = \langle f^{p-1} \rangle$. □

We now apply Fedder's lemma in a number of examples of hypersurface singularities:

Example 6.2.5. We consider S to be a polynomial ring in the following examples.

Node: Consider the ring $S = k[x, y]$ and $R = k[x, y]/\langle xy \rangle$. Then R is F-split near the origin since

$$(xy)^{p-1} = x^{p-1} y^{p-1} \notin \langle x^p, y^p \rangle.$$

[6] Or even semilocal.

Cusp: Consider the ring $S = k[x, y]$ and $R = k[x, y]/\langle x^3 - y^2\rangle$. Then we claim that R is *not* F-split near the origin since (for odd primes). To see this observe that for some constant c

$$(x^3 - y^2)^{p-1} = x^{3(p-1)} + \cdots + cx^{3(p-1)/2}y^{p-1} + \cdots + x^{2(p-1)} \in \langle x^p, y^p\rangle.$$

The computation for $p = 2$ is similar (or follows from the work below).

Pinch point: Consider the ring $S = k[x, y, z]$ and $R = k[x, y, z]/\langle xy^2 - z^2\rangle$. If $p \neq 2$, this is F-split near the origin since

$$\begin{aligned} &(xy^2 - z^2)^{p-1} \\ &= x^{p-1}y^{2(p-1)} + \cdots + \tbinom{p-1}{(p-1)/2}(-1)^{(p-1)/2}x^{(p-1)/2}y^{p-1}z^{p-1} + \cdots + z^{(p-1)/2} \\ &\notin \langle x^p, y^p, z^p\rangle, \end{aligned}$$

noting that p does not divide $\binom{p-1}{(p-1)/2}$.

Characteristic 2: If $R = k[x_1, \ldots, x_n]/\langle f\rangle$ and $\operatorname{char} k = 2$, then R is F-split near the origin if and only if $f \notin \langle x_1^2, \ldots, x_n^2\rangle$. In particular, it is immediate that the cusp and the pinch point are also not F-split near the origin in characteristic 2.

Characteristic 3: Just like characteristic 2, if $R = k[x_1, \ldots, x_n]/\langle f\rangle$ and $\operatorname{char} k = 3$, then R is F-split near the origin if and only if $f^2 \notin \langle x_1^3, \ldots, x_n^3\rangle$.

Finally, we point out that complete intersection singularities are nearly as easy to compute as hypersurfaces.

Proposition 6.2.6. *Suppose that $f_1, \ldots, f_m \subseteq \langle x_1, \ldots, x_n\rangle \subseteq k[x_1, \ldots, x_n] = S$ is a regular sequence.*[7] *Set $I = \langle f_1, \ldots, f_m\rangle$. Then*

$$(I^{[p^e]} : I) = \langle f_1^{p^e-1} \cdots f_n^{p^e-1}\rangle + I^{[p^e]}$$

In particular, S/I is F-split near the origin $\mathfrak{m} = \langle x_1, \ldots, x_n\rangle$ if and only if the product

$$f_1^{p^e-1} \cdots f_n^{p^e-1} \notin \mathfrak{m}^{[p^e]}$$

for some $e > 0$.

Proof. The containment $\supseteq$ is trivial. The converse direction is left as Exercise 15. □

6.3 Exercises

Exercise 1. Prove Lemma 6.0.3.

[7] This means that f_i is not a zero divisor in $S/\langle f_1, \ldots, f_{i-1}\rangle$ for all $i > 0$.

Exercise 2. Suppose that C is a smooth curve and that $\mathscr{L}$ is a line bundle of degree ≥ 2. Prove that the image of the map

$$H^0\big(X, F^e_*(\omega_X \otimes \mathscr{L}^{p^e})\big) \to H^0(X, \omega_X \otimes \mathscr{L})$$

globally generates $\omega_X \otimes \mathscr{L}$ for any $e \gg 0$.

Hint: Mimic the proof in [34, Chap. IV, Proposition 3.1]. For a solution, see [75, Theorem 3.3].

Exercise* 6.3. Consider a map $\varphi : F^e_*\mathscr{L} \to \mathcal{O}_X$ for some line bundle $\mathscr{L}$ and $e > 0$. Formulate analogs of the properties from Proposition 5.1.5 and Corollary 5.1.7 for such a map (and φ-compatible ideals / subvarieties). Which of these properties hold for all φ? Which hold for surjective φ? Prove those that do and give counterexamples to those that do not. Some of the answers can be found in [73, 74].

Exercise* 6.4. Suppose that X is a Frobenius split normal variety. Suppose that X embeds into $\mathbb{P}^n$ as a closed subvariety. Prove that X is compatibly F-split by a Frobenius splitting of $\mathbb{P}^n$ if and only if the embedding $X \subseteq \mathbb{P}^n$ is projectively normal, cf. [34, Chap. II, Exercise 5.14].

Hint: Projective normality can be detected by the difference between the affine cone and the section ring as in Exercise 8. Develop then a "graded variant" of Fedder's lemma that will allow you to prove the result.

Exercise 5. We can define X to be F-pure if $(X, 0)$ is sharply F-pure in the sense of Definition 6.0.5. Show that this coincides with the definition of F-pure given in Definition 5.0.1.

Exercise 6. Suppose that $\mathscr{L}$ is an ample line bundle on a smooth variety X. Prove that $H^0(X, F^e_*(\omega_X \otimes \mathscr{L}^{mp^e})) \to H^0(X, \omega_X \otimes \mathscr{L}^m)$ is surjective for all $m \gg 0$. For one solution, see [75, Lemma 3.1].

Exercise 7. Use the method of Theorem 6.1.3 to prove Proposition 6.1.6.

Hint: Dualize a local splitting $\mathcal{O}_U \to F_*\mathcal{O}_U \to \mathcal{O}_U$ to obtain a surjective map $T : F_*\omega_U \to \omega_U$. Use T instead of φ in the proof of Theorem 6.1.3.

Exercise 8. Consider $\overline{\mathbb{F}_5}[x, y, z] = S$ and $f = x^4 + y^4 + z^4$. Consider the map $\Phi_S : F_*S \to S$ which sends $F_*x^4y^4z^4$ to 1 and sends all the other monomials $x^iy^jz^k$ to 0 for $0 \leq i, j, k \leq 4$ as in Example 4.0.2. Consider the map $\varphi : F_*S \to S$ defined by

$$\varphi(F_*\_\_) = \Phi_S(F_*(f^4 \cdot \_\_)).$$

(a) Prove that $\langle f \rangle$ is φ-compatible and let $\overline{\varphi} : F_*R \to R$ be the induced map on $R = S/\langle f \rangle$ as in Lemma 6.0.3.

(b)* Set $\mathfrak{m} = \langle x, y, z \rangle \in S$. Fix $a, b, c \in \mathbb{F}_{5^2} \setminus \mathbb{F}_5$. Show that $J = \mathfrak{m}^2 + \langle ax + by + cz \rangle$ is φ^2-compatible. However, show that J is not φ-compatible.

Exercise 9. With Φ_S as in Sect. 6.2, fix $f \in S$ and consider the map φ defined by the rule $\varphi(F^e_* \_\_) = \Phi_S(F^e_*(f \cdot \_\_))$. Show that φ is compatible with an ideal $J \subseteq S$ if and only if $f \in J^{[p^e]} : J$.

Exercise 10. Complete the proof of Fedder's lemma by proving the following. For any point $\mathfrak{q} \in V(I) \subseteq \operatorname{Spec} S$, there exists a map $\varphi \in \operatorname{Hom}_R(F^e_* R, R)$ which is surjective at $\mathfrak{q}/I \in \operatorname{Spec} R$ if and only if $I^{[p^e]} : I \not\subseteq \mathfrak{q}^{[p^e]}$. In other words, R is F-pure in a neighborhood of $\mathfrak{q}$ if and only if $I^{[p^e]} : I \not\subseteq \mathfrak{q}^{[p^e]}$.

Hint: Note that a map $\varphi_{\mathfrak{q}} : F^e_*(R_{\mathfrak{q}}) \to R_{\mathfrak{q}}$ is surjective if and only if $\operatorname{Image}(\varphi_{\mathfrak{q}}) \not\subseteq \mathfrak{q} R_{\mathfrak{q}}$.

Exercise 11. Suppose that $X = \operatorname{Spec} R$ is a regular ring and $\Delta = \frac{1}{p^e-1}\operatorname{div}_X(f)$ is a $\mathbb{Q}$-divisor on X. Show that (X, Δ) is sharply F-pure near a point $\mathfrak{m} \in \operatorname{Spec} R = X$ if and only if $f^{p^e-1} \notin \mathfrak{m}^{[p^e]}$.

Hint: Use Fedder's lemma in the form of Remark 6.2.3.

Exercise 12. Suppose that X is a proper variety and that $\varphi : F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is a map that is compatible with $\mathfrak{m}$, the ideal of a closed point $x \in X$. Further suppose that (X, Δ_φ) is sharply F-pure in a neighborhood of $\mathfrak{m}$. Prove that $0 \neq \varphi(F^e_* 1) \in k$, and so in particular X is F-split. This generalizes Theorem 5.3.1 by the following argument.

Given a $D = (p^e-1)D_1 + \cdots + (p^e-1)D_d + G$ and φ as in Theorem 5.3.1, set $\Delta = \frac{1}{p^e-1}D$. Observe that $\mathfrak{m}_x$, the maximal ideal of x, is φ-compatible since each D_i is φ-compatible, cf. Lemma 5.1.3. Use Exercise 11 to conclude that φ is surjective in a neighborhood of $x \in X$.

Obtain a new proof of Theorem 5.3.1 by combining the above.

Exercise 13. Suppose that $X = \operatorname{Spec} k[\![x, y]\!]$ where k has characteristic 7 and that $\Delta = \frac{1}{2}\operatorname{div}_X(y^2 - x^3) + \frac{1}{3}\operatorname{div}_X(x) + \frac{1}{2}\operatorname{div}_X(y)$. Prove that (X, Δ) is sharply F-pure at the origin $\mathfrak{m}$ and also that if φ corresponds to Δ, then $\mathfrak{m}$ is φ-compatible.

Now suppose that Y is a smooth projective variety with a $\mathbb{Q}$-divisor $\Theta \geq 0$ such that

- $(p-1)(K_Y + \Theta) \sim 0$.
- (Y, Θ) has a point $y \in Y$ analytically isomorphic to (X, Δ) above.

Show that Y is Frobenius split using Exercise 12.

Exercise 14. Suppose that R is an integral domain with normalization R^{N} in $K(R)$, the field of fractions of R. In this exercise, we will prove that every map $\varphi : F^e_* R \to R$ induces an R^{N}-linear map $\varphi^{\mathrm{N}} : F^e_* R^{\mathrm{N}} \to R^{\mathrm{N}}$ which is compatible with the conductor ideal $\mathfrak{c} := \operatorname{Ann}_R(R^{\mathrm{N}})$, an ideal in both R and R^{N}. We do this in two steps.

(a) Prove that φ is compatible with $\mathfrak{c}$ (when viewed as an ideal in R).
(b) Notice that φ induces a map $\varphi_0 : F^e_* K(R) \to K(R)$ by localization. Prove that $\varphi_0(R^{\mathrm{N}}) \subseteq R^{\mathrm{N}}$ which proves that we can take $\varphi^{\mathrm{N}} = \varphi_0|_{R^{\mathrm{N}}}$.

Hint: Recall that $x \in K(R)$ is integral over R if there exists a nonzero $c \in R$ such that $cx^n \in R$ for all $n \gg 0$ (see [43, Exercise 2.26]).

Exercise 15. Prove Proposition 6.2.6.

Hint: A very easy proof (pointed out to us by Alberto Fernandez Boix), follows from [31, Corollary 1]. Alternately, the $\supseteq$ containment is easy. For the reverse proceed by induction on the number of f_i. Notice that $\mathrm{Hom}_{S/I}(F^e_* S/I, S/I)$ is a free $F^e_* S/I$-module of rank 1. Thus, a generator of that module corresponds to an element $h \in I^{[p^e]} : I$.

For a generalization to Gorenstein rings (instead of just complete intersections), see [73, Corollary 7.5].

Exercise 16 (Macaulay2 Fedder's criterion). The following Macaulay2 code, written by Mordechai Katzman and available at

`http://katzman.staff.shef.ac.uk/FSplitting/`

can be quite useful.

```
frobeniusPower=method();

frobeniusPower(Ideal,ZZ) := (I,e) ->(
R:=ring I;
p:=char R;
local u;
local answer;
G:=first entries gens I;
if (#G==0) then answer=ideal(0_R) else answer=ideal(apply (G,
u->u^(p^e)));
answer
);
```

This takes an ideal I and raises it to the p^eth Frobenius power, $I \mapsto I^{[p^e]}$. Using this as a starting place, implement within Macaulay2 a method which determines whether a given ring is F-pure near the origin. Check your method against the following examples:

(a) $R = k[x, y, z]/\langle xy, xz, yz\rangle$ in whatever characteristics you feel like.
(b) $R = k[w, x, y, z]/\langle xy, z^2 + wx^2, yz\rangle$ in characteristic 2 and 3.
(c) $R = k[x, y, z]/\langle x^3 + y^3 + z^3\rangle$ in characteristics 7, 11, and 13.
(d) $R = k[x, y, z]/\langle x^2 + y^3 + z^5\rangle$ in characteristics 2, 3, 5, 7, and 11.

Exercise* 6.17. Use Fedder's criterion to determine for which $p > 0$, the ring $k[x, y, z]/\langle x^3 + y^3 + z^3\rangle$ is F-pure near the origin. For some related computations, see [87, Chap. V, Sect. 4].

Exercise* 6.18. If $(R, \mathfrak{m})$ is a regular local ring and $0 \neq f \in \mathfrak{m}$, then the *F-pure threshold* $c_{\mathfrak{m}}(f)$ of $f \in k[x_1, \dots, x_n]$, at the origin $\mathfrak{m} = \langle x_1, \dots, x_n\rangle$, is defined as follows:

$$\lim_{e \to \infty} \frac{\max\{l \mid f^l \notin \mathfrak{m}^{[p^e]}\}}{p^e}.$$

Prove that this limit exists in general and then show that $c_{\mathfrak{m}}(x^3 - y^2) = \frac{5}{6}$ if $p = 7$. See [69] for solutions, cf. [91].

7 Change of Variety

In this section, we describe how p^{-e}-linear maps change under common change of variety operations.

7.1 Closed Subschemes

We have already studied the behavior of p^{-e}-linear maps for subschemes extensively. Indeed, suppose that $\varphi : F^e_* R \to R$ is an R-linear map which is compatible with an ideal $I \subseteq R$. Then we have an induced map $\varphi_{R/I} : F^e_*(R/I) \to (R/I)$. It is natural to ask what the divisor associated to R/I is.

Lemma 7.1.1. *Suppose that R is a normal Gorenstein local ring and that $D = V(f)$ is a normal Cartier divisor on $X = \operatorname{Spec} R$. Fix $\Phi : F^e_* R \to R$ to be map generating $\operatorname{Hom}_R(F^e_* R, R)$ as an $F^e_* R$-module as in Exercise 4. Set $\varphi(F^e_* \_\_) = \Phi(F^e_*(f^{p^e-1} \cdot \_\_))$. Then φ is compatible with D, and furthermore, φ_D generates $\operatorname{Hom}_{R/\langle f \rangle}(F^e_*(R/\langle f \rangle), R/\langle f \rangle)$ as an $F^e_* R/\langle f \rangle$-module. It follows that the $\mathbb{Q}$-divisor Δ on D associated to φ_D, as in (13) is the zero divisor.*

Proof. See Exercise* 7.2. ☐

However, things are not always nearly so nice. In particular the divisor associated to φ_D need not always be zero.

Example 7.1.2. Consider $S = k[x, y, z]$ with $p = \operatorname{char} k \neq 2$, set $R := k[x, y, z]/\langle xy - z^2 \rangle$, and fix $D = V(\langle x, z \rangle)$. Set $\Phi_S \in \operatorname{Hom}_S(F^e_* S, S)$ to be the $F^e_* S$-module generator as in Exercise 4.0.2. We notice that by Fedder's lemma, Lemma 6.2.1, that $\Psi(F^e_* \_\_) = \Phi(F^e_*((xy - z^2)^{p^e-1} \cdot \_\_))$ induces the generator of $\operatorname{Hom}_R(F^e_* R, R)$ by restriction. Notice that $\mathcal{O}_X(-2nD) = \langle x^n \rangle$ and consider the map

$$\varphi(F^e_* \_\_) = \Psi(F^e_*(x^{\frac{p^e-1}{2}} \cdot \_\_)) = \Phi(F^e_*(x^{\frac{p^e-1}{2}}(xy - z^2)^{p^e-1} \cdot \_\_)).$$

If we set $X = \operatorname{Spec} R$, then the induced map $\varphi_X \in \operatorname{Hom}_R(F^e_* R, R)$ corresponds to the divisor $(p^e - 1)D$.

However, it is easy to see that φ_X also is compatible with D. Thus we obtain φ_D. To compute the divisor associated to D, we need only read off the term containing $x^{p^e-1} z^{p^e-1}$ in

$$(x^{\frac{p^e-1}{2}})(xy-z^2)^{p^e-1} = x^{\frac{3(p^e-1)}{2}} y^{p^e-1} + \cdots + \binom{p^e-1}{\frac{p^e-1}{2}} x^{p^e-1} y^{\frac{p^e-1}{2}} z^{p^e-1}$$
$$+ \cdots + z^{2(p^e-1)}.$$

Again, the reason this works is because the map $\Phi_S(F^e_*(x^{p^e-1}z^{p^e-1} \cdot \_\_))$ induces the generator on $\mathrm{Hom}_{\mathcal{O}_D}(F^e_*\mathcal{O}_D, \mathcal{O}_D)$. But $\binom{p^e-1}{\frac{p^e-1}{2}} \neq 0 \bmod p$, and so if $\Phi_D : F^e_*k[y] \to k[y]$ is the map generating $\mathrm{Hom}_{\mathcal{O}_D}(F^e_*\mathcal{O}_D, \mathcal{O}_D)$, then φ_D (which is just φ_X restricted to D) is defined by the rule

$$\varphi_D(F^e_*\_\_) = \Phi_D(F^e_* y^{\frac{p^e-1}{2}} \cdot \_\_)$$

at least up to multiplication by an element of k. Thus, in the terminology of (13),

$$\Delta_{\varphi_D} = \frac{1}{p^e-1}\mathrm{div}(y^{\frac{p^e-1}{2}}) = \frac{1}{2}\mathrm{div}(y).$$

In particular, in contrast with Lemma 7.1.1, $\Delta_{\varphi_D} \neq 0$.

Theorem 7.1.3 (F-adjunction). *If X is a normal variety, $\Delta \geq 0$ is a $\mathbb{Q}$-divisor on X such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier with index not divisible by p. Suppose that Y is an F-pure center (see Definition 6.0.2) of (X, Δ) and that φ corresponds to Δ as in (15). Then there exists a canonically determined $\mathbb{Q}$-divisor $\Delta_Y \geq 0$ such that:*

(a) $(K_Y + \Delta)|_Y \sim_{\mathbb{Q}} K_Y + \Delta_Y$.
(b) (X, Δ) *is sharply F-pure near Y if and only if* (Y, Δ_Y) *is sharply F-pure.*

Proof. Set φ_Y to be the restriction of φ to Y as in Lemma 6.0.3. Set Δ_Y to be the $\mathbb{Q}$-divisor associated to φ_Y as in (15). The first result then follows easily. The second follows since φ is surjective near Y if and only if φ_Y is surjective. □

Remark 7.1.4. The previous result should be compared with subadjunction and inversion of adjunction in birational geometry. See for example [25, 46, 48] and [50, Chap. 5, Sect. 4].

7.2 Birational Maps

Suppose that X is a normal variety, $\mathscr{L}$ is a line bundle on X, and $\varphi : F^e_*\mathscr{L} \to \mathcal{O}_X$ is an $\mathcal{O}_X$-linear map corresponding to the $\mathbb{Q}$-divisor Δ as in (15). Suppose $\pi : \widetilde{X} \to X$ is a birational map with $\widetilde{X}$ normal. Fix $K_{\widetilde{X}}$ and K_X which agree wherever π is an isomorphism. We can write

$$K_{\widetilde{X}} + \Delta_{\widetilde{X}} = \pi^*(K_X + \Delta)$$

where now $\Delta_{\widetilde{X}}$ is uniquely determined. Notice that $\Delta_{\widetilde{X}}$ need not be effective. The main result of this section is the following:

Lemma 7.2.1. *The map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ induces a map $\widetilde{\varphi} : F^e_*(\pi^* \mathscr{L}) \to \mathscr{K}(\widetilde{X})$ where $\mathscr{K}(\widetilde{X})$ is the fraction field sheaf of $\widetilde{X}$ (which we can also identify with the fraction field on X since π is birational). Furthermore, $\widetilde{\varphi}$ agrees with φ wherever π is an isomorphism.*

Even more, using the fact that maps to the fraction field correspond to possibly non-effective divisors via Exercise 4.13, we have that $\Delta_{\widetilde{\varphi}} = \Delta_{\widetilde{X}}$.*

Proof. We construct $\widetilde{\varphi}$ as follows. We note that $\mathscr{L} = \mathcal{O}_X((1-p^e)(K_X + \Delta))$ by (15), and so after fixing K_X, we obtain an embedding of $\mathscr{L} \subseteq \mathscr{K}(X)$. In particular, for each affine open set U, we have an embedding $\Gamma(U, \mathscr{L}) \subseteq K(X)$. But then we also obtain for each affine open set $V \subseteq \widetilde{X}$, an embedding $\Gamma(V, \pi^* \mathscr{L}) \subseteq K(\widetilde{X}) = K(X)$.

Now, by taking the map $F^e_* \mathscr{L} \to \mathcal{O}_X$ at the generic point η of X, we obtain $\varphi_\eta : F^e_* K(X) \to K(X)$ (note that our embedding of $\mathscr{L} \subseteq \mathscr{K}(X)$ fixes the isomorphism $\mathscr{L}_\eta \cong K(X)$). But we identify η with the generic point $\widetilde{\eta}$ of $\widetilde{X}$ (since they have isomorphic neighborhoods), and so we have a map $\varphi_{\widetilde{\eta}} : F^e_* K(\widetilde{X}) \to K(\widetilde{X})$. By restricting $\varphi_{\widetilde{\eta}}$ to $\Gamma(V, \pi^* \mathscr{L})$ for each open set V, we obtain a map $\widetilde{\varphi} : F^e_* \pi^* \mathscr{L} \to \mathscr{K}(\widetilde{X})$.

By construction, $\widetilde{\varphi}$ agrees with φ wherever π is an isomorphism. For the statement $\Delta_{\widetilde{\varphi}} = \Delta_{\widetilde{X}}$ we proceed as follows. We notice that $\Delta_{\widetilde{\varphi}}$ and $\Delta_{\widetilde{X}}$ already agree wherever π is an isomorphism so that $\Delta_{\widetilde{\varphi}} - \Delta_{\widetilde{X}}$ is π-exceptional. Furthermore, by the construction done in Exercise* 4.13, $\mathcal{O}_{\widetilde{X}}((1-p^e)(K_{\widetilde{X}} + \Delta_{\widetilde{\varphi}})) \cong \pi^* \mathscr{L} \cong \pi^* \mathcal{O}_X((1-p^e)(K_X + \Delta))$. Thus, $\Delta_{\widetilde{\varphi}} \sim_{\mathbb{Q}} \Delta_{\widetilde{X}}$, and so

$$\Delta_{\widetilde{\varphi}} - \Delta_{\widetilde{X}} \sim_{\mathbb{Q}} 0$$

is π-exceptional. Therefore $\Delta_{\widetilde{\varphi}} = \Delta_{\widetilde{X}}$ as desired, cf. [50]. ☐

We now come to the definition of log canonical singularities (in arbitrary characteristic).

Definition 7.2.2. Suppose that X is a normal variety and that Δ is a $\mathbb{Q}$-divisor such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier. Then we say that (X, Δ) is *log canonical* if the following condition holds. For every proper birational map $\pi : \widetilde{X} \to X$ with $\widetilde{X}$ normal, when we write

$$\sum a_i E_i = K_{\widetilde{X}} - \pi^*(K_X + \Delta),$$

each a_i is ≥ -1.

Theorem 7.2.3 ([28, Main Theorem]). *If (X, Δ) is sharply F-pure, then (X, Δ) is log canonical.*

Proof. The statement is local on X, and so we may assume that $\mathscr{L} = \mathcal{O}_X$ and that $X = \operatorname{Spec} R$ is affine. We only prove the case where the index of $K_X + \Delta$ is not divisible by p. To reduce to this case, use Exercise 6. Set $\varphi : F^e_* R \to R$ to be a map corresponding to Δ. Thus there exists an element $c \in R = \Gamma(X, \mathscr{L})$ such that $\varphi(F^e_* c) = 1$ since (X, Δ) is sharply F-pure.

Set $\pi : \widetilde{X} \to X$ a proper birational map with $\widetilde{X}$ normal and write $\sum a_i E_i = K_{\widetilde{X}} - \pi^*(K_X + \Delta)$. Suppose that some $a_i < -1$ (with corresponding fixed E_i). Then in particular $a_i \leq 0$. Set η_i to be the generic point of E_i. It follows that $-a_i$, the E_i-coefficient of $\Delta_{\widetilde{\varphi}}$, is positive, and so we have a factorization

$$F^e_* \mathcal{O}_{\widetilde{X},\eta_i} \subseteq F^e_* \mathcal{O}_{\widetilde{X},\eta_i}((1-p^e)a_i E_i) \xrightarrow{\widetilde{\varphi}} \mathcal{O}_{\widetilde{X},\eta_i},$$

where $\widetilde{\varphi}$ is as in Lemma 7.2.1. But now it is easy to see that if $a_i < -1$, then $(1-p^e)a_i \geq p^e$ so that we have the factorization

$$F^e_* \mathcal{O}_{\widetilde{X},\eta_i} \subseteq F^e_* \mathcal{O}_{\widetilde{X},\eta_i}(p^e E_i) \xrightarrow{\widetilde{\varphi}|_{F^e_* \mathcal{O}_{\widetilde{X},\eta_i}(p^e E_i)}} \mathcal{O}_{\widetilde{X},\eta_i}$$

which sends $F^e_* c \in F^e_* R \subseteq F^e_* \mathcal{O}_{\widetilde{X},\eta_i}$ to 1. But that is impossible since if $d \in \mathcal{O}_{\widetilde{X},\eta_i}$ is the local parameter for E_i, then $\widetilde{\varphi}$ sends $F^e_*(c/d^{p^e}) \in F^e_* \mathcal{O}_{\widetilde{X},\eta_i}(p^e E_i)$ to $1/d \notin \mathcal{O}_{\widetilde{X},\eta_i}$. □

7.3 Finite Maps

Finally, suppose that $\pi : Y \to X$ is a finite surjective map of normal varieties. Then there is an inclusion $\mathcal{O}_X \subseteq \pi_* \mathcal{O}_Y$. Given a line bundle $\mathscr{L}$ on X and a map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$, it is natural to ask when φ can be extended to a map $F^e_* \pi_*(\pi^* \mathscr{L}) \to \pi_* \mathcal{O}_Y$. Since π is finite, the π_* is harmless, and so we can ask when φ can be extended to a map $\varphi_Y : F^e_* \pi^* \mathscr{L} \to \mathcal{O}_Y$.

The local version of this statement is as follows. Suppose that $R \subseteq S$ is a finite extension of semi-local normal rings and suppose that $\varphi : F^e_* R \to R$ is a finite map. Then when does there exist a commutative diagram as follows?

$$\begin{array}{ccc} F^e_* S & \xrightarrow{\varphi_S} & S \\ \uparrow & & \uparrow \\ F^e_* R & \xrightarrow[\varphi]{} & R \end{array}$$

It is easy to see that the answer is not always.

Example 7.3.1. Consider $k[x^2] \subseteq k[x]$ with $p = \operatorname{char} k \neq 2$. Consider the map $\varphi : F_* k[x^2] \to k[x^2]$ which sends $F_* x^{2(p-1)}$ to 1 and other monomials $F_* x^{2i}$, for $0 \leq i < p-1$ to zero. Note $\Delta_\varphi = 0$.

Suppose this map extended to a map $\psi : F_* k[x] \to k[x]$. Then $\varphi(F_* x^{2(p-1)}) = 1$, and so since φ and ψ are the same on $k[x^2]$, we have

$$1 = \psi(F_* x^{2(p-1)}) = \psi(F_* x^p x^{p-2}) = x\psi(F_* x^{p-2})$$

which implies that x is a unit. But that is a contradiction.

On the other hand, consider the map $\alpha : F_* k[x^2] \to k[x^2]$ which sends $F_* x^{2(p-1)/2} = F_* x^{p-1}$ to 1 and the other monomials $F_* x^{2i}$ to 0 for $0 \leq i \leq p-1$ to zero. Note $\Delta_\alpha = \frac{1}{2}\mathrm{div}(x^2)$.

We will show that α extends to a map $\beta : F_* k[x] \to k[x]$. It is in fact easy to show that α extends to a map on the fraction field $\beta : F_* k(x) \to k(x)$ (see Exercise 7). Therefore, it is enough to show that $\beta(F_* x^j) \in k[x]$ for each $0 \leq j \leq p-1$. Fix such a j. If j is even, then there is nothing to do since $\beta(F_* x^j) = \alpha(F_* x^j) \in k[x^2] \subseteq k[x]$. Therefore, we may suppose that j is odd. But then $j+p$ is even and $p \leq j+p \leq 2(p-1)$. Thus

$$\beta(F_* x^j) = \frac{1}{x}\beta(F_* x^{j+p}) = \frac{1}{x}\alpha(F_* x^{j+p}) = \frac{1}{x} \cdot 0 = 0 \in k[x].$$

This proves that β exists and is well defined.

Theorem 7.3.2 ([80]). *Fix $\pi : Y \to X$ as above. Fix a nonzero map $\varphi : F^e_* \mathscr{L} \to \mathcal{O}_X$ as above. If π is inseparable then φ never extends to φ_Y. If π is separable, then there exists a map $\varphi_Y : F^e_* \pi^* \mathscr{L} \to \mathcal{O}_Y$ extending φ if and only if Δ_φ is bigger than or equal to the ramification divisor of $\pi : Y \to X$.*

Proof. We won't prove this, but we will sketch the main steps and leave the details as an exercise. We first work in the separable case.

Step 1: The statement is local on X, and so we may suppose that $X = \operatorname{Spec} R$, $Y = \operatorname{Spec} S$, and $\mathscr{L} = \mathcal{O}_X$. In fact, we may even assume that R is a DVR and that S is a Dedekind domain.

Step 2: There is a map φ_S and a commutative diagram

$$\begin{array}{ccc} F^e_* S & \xrightarrow{\varphi_S} & S \\ \uparrow & & \uparrow \\ F^e_* R & \xrightarrow[\varphi]{} & R \end{array}$$

if and only if there exists a map φ_S and a commutative diagram

$$\begin{array}{ccc} F^e_* S & \xrightarrow{\varphi_S} & S \\ {\scriptstyle F^e_*\mathrm{Tr}}\downarrow & & \downarrow{\scriptstyle \mathrm{Tr}} \\ F^e_* R & \xrightarrow[\varphi]{} & R \end{array}$$

where $\mathrm{Tr} : S \to R$ is simply the restriction of the field trace $\mathrm{Tr} : K(S) \to K(R)$ to S.

Step 3: $\mathrm{Hom}_R(S, R)$ is isomorphic to S as an S-module. The map $\mathrm{Tr} : S \to R$ is a section of this and so corresponds to a divisor D on $\mathrm{Spec}\, R$. This divisor is the ramification divisor Ram_π of $\pi : \mathrm{Spec}\, S \to \mathrm{Spec}\, R$.

Step 4: Supposing φ_S exists, compute the divisor corresponding to $\mathrm{Tr} \circ \varphi_S = \varphi \circ (F^e_* \mathrm{Tr})$. This gives one direction of the if and only if. Working with the fraction fields, as in Exercise* 4.13, yields the other direction.

For the inseparable case, it turns out that the only map that can extend is the zero map (see Exercise 9). □

7.4 Exercises

Exercise 1. In the setting of Lemma 7.1.1, prove that the divisor D is F-pure (as a variety) if and only if φ is surjective.

Exercise* 7.2. Prove Lemma 7.1.1.

Hint: Consider the map $\langle \varphi \rangle_{F^e_* R} \to \mathrm{Hom}_{R/\langle f \rangle}(F^e_* R/\langle f \rangle, R/\langle f \rangle)$, and prove it is surjective at the codimension 1 points of $R/\langle f \rangle$. For a solution, see [73, Proposition 7.2].

Exercise 7.3 (The F-different).** Suppose that X is a normal variety and D is an effective normal Weil divisor such that $K_X + D$ is $\mathbb{Q}$-Cartier with index not divisible by p. Thus there exists a map $\varphi_D : F^e_* \mathscr{L} \to \mathcal{O}_X$ as in (15) corresponding to D for any e such that $(p^e - 1)(K_X + D)$ is Cartier. It is easy to see that this map is compatible with D, and so it induces a map

$$\varphi_D : F^e_* \mathscr{L}|_D \to \mathcal{O}_D.$$

This map corresponds to a $\mathbb{Q}$-divisor Δ_D on D, again by (15), which is called the *F-different*. Verify all the statements made above.

It is an open question whether or not the F-different always coincides with the different, as described in [51, Chap. 17] or [86, 10.6]. Prove that it either does or does not and write a paper about it, and then tell the authors of this survey paper what you found (this is why the problem gets $**$). For more discussion see [73, Remark 7.6].

Exercise* 7.4. Consider the family of cones over elliptic curves:

$$X = \operatorname{Spec} k[x, y, z, t]/\langle y^2 - x(x-1)(x-t)\rangle \to \mathbb{A}^1 = \operatorname{Spec} k[t].$$

Set $\Phi \in \operatorname{Hom}_X(F_*\mathcal{O}_X, \mathcal{O}_X)$ to be the map generating $\operatorname{Hom}_X(F_*\mathcal{O}_X, \mathcal{O}_X)$ as an $F_*\mathcal{O}_X$-module. Show that Φ is compatible with the ideal $J = \langle x, y, z\rangle$. Consider $\Phi_J = \Phi/J$, the map obtained by restricting Φ to $V(J) \cong \mathbb{A}^1$. Show that Δ_{Φ_J} is supported exactly at those points whose fibers correspond to supersingular elliptic curves.

Exercise 5. Using the notation of Lemma 7.2.1, suppose that Δ_φ is the effective divisor associated to φ. Show that there is a map

$$\varphi' : F^e_*\big((\pi^*\mathscr{L})(\lceil K_{\widetilde{X}} - \pi^*(K_X + \Delta_\varphi)\rceil)\big) \to \mathcal{O}_{\widetilde{X}}(\lceil K_{\widetilde{X}} - \pi^*(K_X + \Delta_\varphi)\rceil)$$

that agrees with φ wherever π is an isomorphism.

*H*int: It is sufficient to show that there is a map $\varphi'' : F^e_*\big((\pi^*\mathscr{L})(\lceil K_{\widetilde{X}} - \pi^*(K_X + \Delta_\varphi)\rceil - p^e\lceil K_{\widetilde{X}} - \pi^*(K_X + \Delta_\varphi)\rceil)\big) \to \mathcal{O}_{\widetilde{X}}$. Now, use the roundings to your advantage and the fact that $\pi^*\mathscr{L} = \mathcal{O}_{\widetilde{X}}(\pi^*(1-p^e)(K_X + \Delta_\varphi))$.

Exercise 6. Suppose that (X, Δ) is sharply F-pure. Prove that for every point $x \in X$ there exists a divisor Δ_U on a neighborhood U of x such that $\Delta_U \geq \Delta|_U$, such that (U, Δ_U) is sharply F-pure, *and* such that $K_U + \Delta_U$ has index not divisible by p. Conclude that Theorem 7.2.3 holds in full generality. For a solution, see [78, Theorem 4.3(ii)].

Exercise 7. Suppose that $R \subseteq S$ is an extension of integral domains with induced *separable* extension of fraction fields $K(R) \subseteq K(S)$. Fix $\varphi : F^e_* R \to R$ to be an R-linear map. Prove that there is always a map $\psi : F^e_* K(S) \to K(S)$ such that $\psi|_R = \varphi$.

Hint: First form $\varphi_\eta : F^e_* K(R) \to K(R)$ by localization. Then tensor this map with $K(S)$ and use the fact that $K(R) \subseteq K(S)$ is separable (unlike $K(R) \subseteq F^e_* K(R) \cong (K(R))^{1/p^e}$).

Exercise* 7.8. Prove the separable case of Theorem 7.3.2 by filling in the details of the Steps 1–4. Step 3 is somewhat involved; see for example [15, 68, 72]. On the other hand, see [80] for a complete proof.

Exercise 9. Prove the inseparable case of Theorem 7.3.2 as follows. First suppose that $K \subseteq L$ is a purely inseparable extension of fields. Suppose that $\varphi : F_*K \to K$ is a K-linear map that extends to an L-linear map $\varphi_L : F_*L \to L$. Prove that $\varphi = 0$.

Use the above to prove that now if $L \supseteq K$ is any inseparable map, the only map $\varphi : F_*K \to K$ that extends to $\varphi_L : F_*L \to L$ is the zero map. For a complete solution, see [80, Proposition 5.2].

8 Cartier Modules

Perhaps the most natural example of a p^{-e}-linear map is the trace of the Frobenius $F_*\omega_X \to \omega_X$ on the canonical sheaf of a normal variety as discussed in detail in Sect. 3.2. In generalizing one is led to consider the category consisting of (coherent) $\mathcal{O}_X$-modules $\mathscr{F}$ equipped with a p^{-e}-linear map $\kappa\colon F^e_*\mathscr{F} \to \mathscr{F}$. We will outline here the resulting theory in a slightly more general setting than considered in [6].

Definition 8.0.1. If $\mathscr{L}$ is a line bundle on X, then a $(\mathscr{L}, p^e)$–*Cartier module* is a coherent $\mathcal{O}_X$-module $\mathscr{F}$ equipped with an $\mathcal{O}_X$-linear map

$$\kappa\colon F^e_*(\mathscr{F} \otimes \mathscr{L}) \to \mathscr{F}\,.$$

(or equivalently, equipped with a p^{-e} linear map $\mathscr{F} \otimes \mathscr{L} \to \mathscr{F}$). If $\mathscr{L} \cong \mathcal{O}_X$, we call these objects mostly just Cartier modules.

Remark 8.0.2. Cartier modules as originally defined in the work of [6] were always defined with $\mathscr{L} \cong \mathcal{O}_X$. The addition of the $\mathscr{L}$ adds little to the complication of the basic theory (which generally reduces to the local case where $\mathscr{L}$ is trivialized). Although admittedly, it does add some notational complications. However, this generalization does show up naturally. Regardless, little will be lost if the reader always assumes that $\mathscr{L} = \mathcal{O}_X$.

A morphism of $(\mathscr{L}, p^e)$–Cartier modules $(\mathscr{F}, \kappa_{\mathscr{F}})$ and $(\mathscr{G}, \kappa_{\mathscr{G}})$ is an $\mathcal{O}_X$-linear map $\varphi\colon \mathscr{F} \to \mathscr{G}$ such that the diagram

$$\begin{array}{ccc} F^e_*(\mathscr{F} \otimes \mathscr{L}) & \xrightarrow{\kappa_{\mathscr{F}}} & \mathscr{F} \\ \Big\downarrow{\scriptstyle F^e_*(\varphi\otimes \mathrm{id})} & & \Big\downarrow{\scriptstyle \varphi} \\ F^e_*(\mathscr{G} \otimes \mathscr{L}) & \xrightarrow[\kappa_{\mathscr{G}}]{} & \mathscr{G} \end{array}$$

commutes. If $(\mathscr{F}, \kappa)$ is a $(\mathscr{L}, p^e)$–Cartier module, then we can apply F^e_* to $\kappa \otimes \mathscr{L}$ to obtain—using the projection formula—a map

$$\kappa^2\colon F^{2e}_*(\mathscr{F} \otimes \mathscr{L} \otimes \mathscr{L}^{p^e}) \cong F^e_*(F^e_*(\mathscr{F} \otimes \mathscr{L}) \otimes \mathscr{L}) \xrightarrow{F^e_*(\kappa\otimes\mathscr{L})} F^e_*(\mathscr{F} \otimes \mathscr{L}) \xrightarrow{\kappa} \mathscr{F}$$

which equips $\mathscr{F}$ with the structure of a $(\mathscr{L}^{1+p^e}, p^{2e})$–Cartier module. Iterating this construction in the obvious way (similar to (17)) we obtain morphisms

$$\kappa^e\colon F^{ne}_*\left(\mathscr{F} \otimes \mathscr{L}^{1+p^e+p^{2e}+\cdots+p^{(n-1)e}}\right) \to \mathscr{F}$$

for all $n \geq 1$, making $\mathscr{F}$ into a $(\mathscr{L}^{\frac{p^{ne}-1}{p^e-1}}, p^{ne})$-Cartier module.

Proposition 8.0.3. *The category of (coherent) $(\mathscr{L}, p^e)$-Cartier modules is an Abelian category. The kernel and cokernel of the underlying quasi-coherent sheaves carry an obvious Cartier module structure and are the kernel and cokernel in the category of Cartier modules.*

Proof. This is easy to verify since $\_\_ \otimes \mathscr{L}$ as well as $F^e_* \_\_$ are exact functors. Alternatively, we may view $(\mathscr{L}, p^e)$-Cartier modules as the right module category over a certain (noncommutative) sheaf of rings, see Exercise 4, which immediately implies that the category is Abelian. □

Compared to a Frobenius splitting, which is nothing but a Cartier module structure on the coherent sheaf $\mathcal{O}_X$, the advantages of working in this larger category of Cartier modules are manifold. For one, there are a number of natural examples of Cartier modules, most prominently the canonical sheaf ω_X together with the trace of Frobenius as Cartier module structure. Furthermore one has in this category methods to construct new Cartier modules by functorial operations. Most notably there is the notion of a pushforward for proper maps (in the case that $\mathscr{L} \cong \mathcal{O}_X$), localization and étale pullback, and even an extraordinary pullback $f^!$ can be defined [5, 6]. We conclude this section by illustrating some of these concepts in special cases. First however, we state some examples.

Example 8.0.4 (Examples of Cartier modules).

(a) The canonical sheaf ω_X is a Cartier module with structural map $\kappa\colon F_*\omega_X \to \omega_X$ given by the trace map. More generally, if $\omega_X^\bullet$ is the dualizing complex of X, then the trace of Frobenius is a map (in the derived category) $F_*\omega_X^\bullet \to \omega_X^\bullet$. This induces for each i the structure of a Cartier module on the cohomology $\mathbf{h}^i\omega_X^\bullet$.
(b) Suppose that D is a Cartier divisor on X, then the map

$$F^e_*(\omega_X(p^e D)) \xrightarrow{\mathrm{Tr}} \omega_X(D)$$

equips $\omega_X(D)$ with the structure of an $\mathcal{O}_X((p^e-1)D)$-Cartier module.
(c) Suppose that D is an effective integral divisor on X, then the composition

$$F_*\omega_X(D) \hookrightarrow F_*\omega_X(pD) \xrightarrow{\mathrm{Tr}} \omega_X(D)$$

equips $\omega_X(D)$ with the structure of a Cartier module as well.
(d) Suppose that $\pi : Y \to X$ is a proper map of varieties. Then $R^i\pi_*\omega_Y$ is a Cartier module for any $i \geq 0$. This is because $F_*R^i\pi_*\omega_Y = R^i\pi_*F_*\omega_Y$.
(e) Set $X = \mathbb{A}^2$ and let $\pi : Y \to X$ be the blowup at the origin with exceptional divisor E. Thus we have $\mathrm{Tr}_Y : F_*\omega_Y \to \omega_Y$ as the trace on Y. Now, $\omega_Y \cong \mathcal{O}_Y(E)$. Thus, by twisting by $-E$,we have a $\mathcal{O}_Y((1-p)E)$-Cartier module structure on $\mathcal{O}_Y$. Namely, a map $\mathrm{Tr} : F_*(\mathcal{O}_Y((1-p)E)) \to \mathcal{O}_Y$.

Since localization at any multiplicative set commutes with pushforward along the Frobenius (see Exercise 6 and Exercise 3) we observe that localization preserves the Cartier module structure.

Lemma 8.0.5. *Let $S \subseteq R$ be a multiplicative system and $\mathscr{F}$ a $(\mathscr{L}, p^e)$–Cartier module on $X = \operatorname{Spec} R$. Then the map*

$$F^e_{S^{-1}R*}(S^{-1}\mathscr{F} \otimes_{S^{-1}R} S^{-1}\mathscr{L}) \cong S^{-1}F^e_*(\mathscr{F} \otimes_R \mathscr{L}) \xrightarrow{S^{-1}\kappa_{\mathscr{F}}} S^{-1}\mathscr{F}$$

is a $(S^{-1}\mathscr{L}, p^e)$–Cartier module structure on $S^{-1}\mathscr{F}$.

In particular, if $j\colon U \subseteq X = \operatorname{Spec} R$ is the inclusion of a basic open subset $U = \operatorname{Spec} R_f$ for some $f \in R$, then the pullback j^* induces a functor from $\mathscr{L}$–Cartier modules on X to $j^*\mathscr{L}$–Cartier modules on U. Using a Čech-complex construction, this globalizes to an arbitrary open immersion $U \subseteq X$. Even more generally this holds for any essentially étale[8] morphism $j\colon U \to X$ [5].

Proposition 8.0.6. *Let $j\colon U \to X$ be essentially étale, and let $\mathscr{F}$ be a $\mathscr{L}$-Cartier module on X. Then the pullback $j^*\mathscr{F}$ carries a natural functorial structure of a $j^*\mathscr{L}$-Cartier module on U. The structural map is given by*

$$F_{U*}(j^*\mathscr{F} \otimes j^*\mathscr{L}) \cong F_{U*}j^*(\mathscr{F} \otimes \mathscr{L}) \cong j^*F_{X*}(\mathscr{F} \otimes \mathscr{L}) \xrightarrow{j^*\kappa} j^*\mathscr{F}\,.$$

Proof. The key point is the fact that for an essentially étale morphism $j\colon U \to X$ the diagram

$$\begin{array}{ccc} U & \xrightarrow{j} & X \\ {\scriptstyle F_Y}\downarrow & & \downarrow{\scriptstyle F_X} \\ U & \xrightarrow{j} & X \end{array}$$

is Cartesian and that the base change morphism $j^*F_{X*} \cong F_{U_*}j^*$ is an isomorphism since j is flat (see [41]). This justifies the definition of the Cartier structure on $j^*\mathscr{F}$. □

For a closed immersion $i\colon Y \to X$, the pullback i^* does not give a functor on Cartier modules. The reason is precisely that the above diagram is not Cartesian in this case. However, there is an exotic restriction functor one can define. For concreteness, let $X = \operatorname{Spec} R$ be affine and let $Y = \operatorname{Spec} R/I$ for some ideal $I \subseteq R$. Then, for an R-module M, the R/I submodule $i^\flat(M) := \operatorname{Hom}_R(R/I, M) = \{m \in M \mid Im = 0\}$ is just the I-torsion submodule $M[I] \subseteq M$. Note that $F_*(M[I]) \subseteq F_*(M[I^{[p]}]) = (F_*M)[I]$ which shows that $F_*(M[I])$, is contained

[8] Essentially étale means essentially of finite type and formally étale, i.e., a morphism that can be factored as a localization followed by a finite type étale morphism.

in the I-torsion submodule $(F_*M)[I]$ of F_*M. Hence, if $\kappa\colon F_*M \to M$ is a Cartier module structure on M, then we have that this restricts to a map

$$\kappa\colon F_*(M[I]) \to M[I]$$

giving $M[I]$ a natural Cartier module structure. The same construction works globally and more generally for $(\mathscr{L}, p^e)$-Cartier modules:

Proposition 8.0.7. *Let $i\colon Y \hookrightarrow X$ be a closed immersion given by a sheaf of ideals I of $\mathcal{O}_X$, and let $\mathscr{F}$ be a $(\mathscr{L}, p^e)$-Cartier module on X. Then the $\mathcal{O}_Y$-module (via action on the first argument) $i^\flat(\mathscr{F}) = \mathrm{Hom}_{\mathcal{O}_X}(i_*\mathcal{O}_Y, \mathscr{F}) = \mathscr{F}[I]$ carries a natural functorial structure of a $(\mathscr{L}|_Y, p^e)$-Cartier module on Y. The structural map is given by*

$$\begin{aligned} F^e_*(\mathscr{F}[I] \otimes_{\mathcal{O}_Y} \mathscr{L}|_Y) &\subseteq F^e_*((\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{L})[I^{[p^e]}]) \\ &= (F^e_*(\mathscr{F} \otimes_{\mathcal{O}_X} \mathscr{L}))[I] \xrightarrow{\kappa_{\mathscr{F}}} \mathscr{F}[I] = i^\flat \mathscr{F}\,. \end{aligned}$$

Finally, let us consider a proper morphism of varieties $\pi\colon Y \to X$. Since the Frobenius commutes with any morphism one has a natural isomorphism of functors $F^e_{X*} \circ \pi_* \cong \pi_* \circ F^e_{Y*}$ which implies that the pushforward induces a functor on Cartier modules as well.

Proposition 8.0.8. *Let $\pi\colon Y \to X$ be a proper morphism and $\kappa\colon F^e_*\mathscr{F} \to \mathscr{F}$ a Cartier module on Y. Then the map*

$$F^e_*(\pi_*(\mathscr{F})) \cong \pi_*(F^e_*\mathscr{F}) \xrightarrow{\pi_*(\kappa)} \pi_*\mathscr{F}$$

is a Cartier module structure on $\pi_*\mathscr{F}$. The same construction also holds for the higher derived images $R^i\pi_*\mathscr{F}$.

Note, however, that if $\mathscr{F}$ is a $(\mathscr{L}, p^e)$–Cartier module, there is no obvious way to equip $\pi_*\mathscr{F}$ with such a structure unless $\mathscr{L}$ is of the form $\pi^*\mathscr{L}'$ for some invertible sheaf on X. In this case, using the projection formula, one obtains

$$F^e_*(\pi_*\mathscr{F} \otimes \mathscr{L}') \cong F^e_*(\pi_*(\mathscr{F} \otimes \pi^*\mathscr{L}') \cong \pi_*(F^e_*(\mathscr{F} \otimes \mathscr{L}) \xrightarrow{\pi_*(\kappa)} \pi_*\mathscr{F}$$

as a Cartier structure on $\pi_*\mathscr{F}$.

Example 8.0.9. Let $\kappa\colon F^e_*\mathscr{F} \to \mathscr{F}$ be a Cartier module, then the pushforward along the Frobenius (which is an affine map) equips $F_*\mathscr{F}$ with the Cartier module structure

$$F^e_*\kappa\colon F^e_*(F^e_*(\mathscr{F})) \to F^e_*\mathscr{F}$$

making $\kappa\colon F^e_*\mathscr{F} \to \mathscr{F}$ into a map of Cartier modules.

8.1 Finiteness Results for Cartier Modules

In this section we state, and outline the proofs of two key structural results which make the category of Cartier modules interesting. But first we introduce the basic concept of nilpotence of a Cartier module and recall some elementary constructions, starting with the following simple lemma whose verification we leave to the reader in Exercise 1.

Lemma 8.1.1. *Let* $\kappa\colon F^e_*(\mathscr{F} \otimes \mathscr{L}) \to \mathscr{F}$ *be a Cartier module. Then the images* $\mathscr{F}_n \kappa^n(F^{ne}_*(\mathscr{F} \otimes \mathscr{L}^{1+p^e+\dots+p^{(n-1)e}}) \subseteq \mathscr{F}$ *are Cartier submodules of* $\mathscr{F}$ *and satisfy the properties:*

(a) $\mathscr{F}_n \supseteq \mathscr{F}_{n+1}$.
(b) $\kappa(F^e_*(\mathscr{F}_n \otimes \mathscr{L})) = \mathscr{F}_{n+1}$.
(c) If $S \subseteq \mathcal{O}_X$ *is a multiplicative set, then* $S^{-1}\mathscr{F}_n = (S^{-1}\mathscr{F})_n$.
(d) The sequence of closed subsets $Y_n := \operatorname{Supp} \mathscr{F}_n/\mathscr{F}_{n+1}$ *is descending.*

An important notion in the theory of Cartier modules, and in particular, for its applications to finiteness results for local cohomology for local rings, is the notion of nilpotence.

Definition 8.1.2. Let $\mathscr{F}$ be a coherent Cartier module on X. We say that $\mathscr{F}$ is *nilpotent* if for some $n \geq 0$ the nth power κ^n of the structural map κ is zero.

Some basic properties of this notion are collected in the following lemma.

Lemma 8.1.3. *Let* $\kappa\colon F^e_*(\mathscr{F} \otimes \mathscr{L}) \to \mathscr{F}$ *be a Cartier module. Denote by* $\mathscr{F}^n \subseteq \mathscr{F}$ *the Cartier submodule of* $\mathscr{F}$ *consisting of all local sections* s *such that* $\kappa^n(F^{ne}_*(\mathcal{O}_C \cdot s \otimes \mathscr{L}^{1+p^e+\dots+p^{(n-1)e}}) = 0$. *Then:*

(a) $\mathscr{F}^n \subseteq \mathscr{F}^{n+1}$ *for all* $n \geq 0$.
(b) $\kappa(F^e_*(\mathscr{F}^{n+1} \otimes \mathscr{L})) \subseteq \mathscr{F}^n$.
(c) If $S \subseteq \mathcal{O}_X$ *is a multiplicative set, then* $S^{-1}\mathscr{F}^n = (S^{-1}\mathscr{F})^n$.
(d) If $\mathscr{F}$ *is coherent, then the ascending sequence stabilizes and the stable member* $\mathscr{F}_{\text{nil}} = \bigcup_n \mathscr{F}^n$ *is the maximal nilpotent Cartier submodule of* $\mathscr{F}$.

Nilpotent Cartier modules form a Serre subcategory of all coherent Cartier modules, i.e., they form an Abelian subcategory which is closed under extension. The only non-trivial part here is the non-closedness under extensions (see Exercise 5).

The first structural result for Cartier modules we will show is that the descending sequence of iterated images stabilizes. This result was first proved in [22, Lemma 13.1]. In fact, this result is essentially Matlis dual to a famous result of Hartshorne and Speiser [37, Proposition 1.11] and generalized by Lyubeznik [59], also cf. [3, 82, 84].

Proposition 8.1.4. *Let* $(\mathscr{F}, \kappa)$ *be a coherent* $(\mathscr{L}, p^e)$*–Cartier module. Then the descending sequence of images*

$$\mathscr{F}_n := \kappa^n(F_*^{ne}(\mathscr{F} \otimes \mathscr{L}^{1+p^e+\cdots+p^{(n-1)e}}) \subseteq \mathscr{F}$$

stabilizes. In particular, the stable image $\sigma(\mathscr{F}) \subseteq \mathscr{F}$ is the largest $(\mathscr{L}, p^e)$-Cartier submodule with the property that the structural map κ is surjective.

Proof. To show the stabilization of a sequence of subsheaves on a Noetherian scheme X can be done on an affine open cover. Choosing the open sets of the cover sufficiently small we may assume that $\mathscr{L}$ is trivial. Hence we may assume that $X = \operatorname{Spec} R$ and M is a finitely generated R module equipped with a p^{-e}-linear map $\kappa\colon M \to M$. And we have to show that the descending sequence of Cartier submodules of M

$$M \supseteq \kappa(M) \supseteq \kappa^2(M) \supseteq \cdots$$

stabilizes. The sets

$$Y_n := \operatorname{Supp}(\kappa^n(M)/\kappa(\kappa^n(M)))$$

form a descending sequence of closed subsets of X, by Lemma 8.1.1. Since X is Noetherian, the descending sequence must stabilize. After truncating we may assume that $Y = Y_n = Y_{n+1}$ for all n. We have to show that Y is empty. Assuming otherwise, let $\mathfrak{p}$ be the generic point of a component of Y. Localizing at $\mathfrak{p}$ we may assume that R is local with maximal ideal $\mathfrak{p}$ and that $Y = \{\mathfrak{p}\} = \operatorname{Supp}(\kappa^n(M)/\kappa(\kappa^n(M)))$ for all n. In particular, for $e = 0$, we obtain that there is an integer k such that $\mathfrak{p}^k M \subseteq \kappa(M)$. Hence, for any $x \in \mathfrak{p}^k$

$$x^2 M \subseteq x\mathfrak{p}^k M \subseteq x\kappa(M) = \kappa(x^{p^e} M) \subseteq \kappa(x^2 M)$$

and iterating we get $x^2 M \subseteq \kappa^n(M)$ for all n. Hence $\mathfrak{p}^k(b-1) \subseteq \kappa^n(M)$ for all e where b is the number of generators of $\mathfrak{p}^k$. Hence the original chain stabilizes if and only if the chain $\kappa^n(M)/\mathfrak{p}^{k(b-1)}M$ does. But the latter is a chain in the finite length module $M/\mathfrak{p}^{k(b-1)}M$. □

A characterization of this stable image is as follows. $\sigma(\mathscr{F}) \subseteq \mathscr{F}$ is the smallest Cartier submodule of $\mathscr{F}$ such that on the quotient $\mathscr{F}/\sigma(\mathscr{F})$ some power of the structural map is zero. If this property is satisfied for some Cartier submodule $\mathscr{N} \subseteq \mathscr{F}$, then it is also satisfied for its image. The minimality now implies that for $\sigma(\mathscr{F})$ the structural map

$$F_*^e(\sigma(\mathscr{F}) \otimes \mathscr{L}) \to \sigma(\mathscr{F})$$

is surjective. The Cartier modules with surjective structural map play an important role in the theory. For example, one can see immediately (Exercise 2) that for such Cartier module $\kappa\colon F_*^e(\mathscr{F} \otimes \mathscr{L}) \twoheadrightarrow \mathscr{F}$ its annihilator $\operatorname{Ann}\mathscr{F}$ is a sheaf of radical ideals, i.e., $\mathscr{F}$ has reduced support. This may be viewed as a generalization of the reducedness of Frobenius split varieties alluded to earlier. It is also a key ingredient in the following Kashiwara-type equivalence which will be used repeatedly below (the easy but rewarding proof is left to the reader as Exercise 10, see also [6, Proposition 2.6 and Sect. 3.3]):

Proposition 8.1.5. *Let $\mathscr{F}$ be a coherent Cartier module on X with surjective structural map $\kappa_{\mathscr{F}}$ (i.e., $\sigma(\mathscr{F}) = \mathscr{F}$). Then $I = \operatorname{Ann}_{\mathcal{O}_X}\mathscr{F}$ is a sheaf of radical ideals, and hence $\mathscr{F} = \mathscr{F}[I] = i^\flat(\mathscr{F})$ is a Cartier module on $Y = \operatorname{Supp}\mathscr{F}$, the closed reduced subset of X given by I.*

More precisely, if $i\colon Y \to X$ denotes a closed immersion, then the functors $i^\flat$ and i_* induce a (inclusion preserving) bijection between

$$\left\{\begin{matrix}\textit{coherent } \mathscr{L}\textit{–Cartier modules on } X\\ \textit{with surjective structural map}\\ \textit{and } \operatorname{Supp}\mathscr{F} \subseteq Y\end{matrix}\right\} \longleftrightarrow \left\{\begin{matrix}\textit{coherent } \mathscr{L}|_Y\textit{-Cartier modules on } Y\\ \textit{with surjective structural map}\end{matrix}\right\}.$$

The most important structural result for Cartier modules is the following theorem which asserts that for a coherent Cartier module $\mathscr{F}$, the lattice of Cartier submodules with surjective structural map satisfies the ascending and descending chain conditions.

Theorem 8.1.6. *Let X be a scheme and $\kappa\colon F^e_*(\mathscr{F} \otimes \mathscr{L}) \to \mathscr{F}$ a coherent Cartier module. Then any chain of Cartier submodules*

$$\cdots \mathscr{F}_i \supseteq \mathscr{F}_{i+1} \supseteq \mathscr{F}_{i+2} \supseteq \cdots,$$

each of whose structural map $\kappa_{\mathscr{F}_i}$ is surjective, is eventually constant (in both directions).

Proof. The ascending chain stabilizes simply because the underlying $\mathcal{O}_X$-module is coherent and our schemes are Noetherian. So it remains to show the descending chain condition. One way to prove this result is to show that there is a unique smallest Cartier submodule $\tau(\mathscr{F}) \subseteq \mathscr{F}$ which agrees with $\sigma(\mathscr{F})$ on each generic point of X, i.e., $\tau(\mathscr{F})_\eta = \sigma(\mathscr{F})_\eta$ for each η the generic point of an irreducible component of X. This is a generalization of the notion of a *test ideal* which will be discussed in some detail in Sect. 9.3.

Assuming the existence of $\tau(\mathscr{F})$ for now, the proof can be outlined as follows: We show that the chain

$$\mathscr{F}_0 \supseteq \mathscr{F}_1 \supseteq \mathscr{F}_2 \supseteq \cdots$$

stabilizes by induction on $\dim X$, the case $\dim X = 0$ being clear. Since a chain stabilizes if it stabilizes after restriction to each of the finitely many irreducible components of X, we may assume that X is irreducible. Since X is Noetherian, the descending sequence of supports $\operatorname{Supp}\mathscr{F}_i$ stabilizes. After truncating we may assume that $Y = \operatorname{Supp}\mathscr{F}_i$ for all i. Since the structural map of each $\mathscr{F}_i$ is surjective, we have by Proposition 8.1.5 that $\mathscr{F}_i$ is annihilated by the ideal sheaf defining the reduced structure of Y. Hence we may view the $\mathscr{F}_i$ as $(\mathscr{L}|_Y, p^e)$ Cartier modules on Y. If $\dim Y < \dim X$ then we are done by induction. So let us assume otherwise that $\dim X = \dim Y$. Further truncating the sequence $\mathscr{F}_i$ we may assume that all $\mathscr{F}_i$'s have the same generic rank. Now, by definition, $\tau(\mathscr{F}_0)$ is contained in $\mathscr{F}_i$ for all i (in fact $\tau(\mathscr{F}_0) = \tau(\mathscr{F}_i)$) such that it is enough to show the stabilization of the sequence

$$\mathscr{F}_0/\tau(\mathscr{F}_0) \supseteq \mathscr{F}_1/\tau(\mathscr{F}_0) \supseteq \mathscr{F}_2/\tau(\mathscr{F}_0) \supseteq \cdots .$$

But since $\tau(\mathscr{F}_0)$ generically agrees with each $\mathscr{F}_i$ this is a sequence of Cartier modules $\mathscr{F}_i/\tau(\mathscr{F}_0)$ whose entries have strictly smaller support than X. As above, we are done by induction. □

A corollary of the proof is the following result.

Proposition 8.1.7. *Let $\mathscr{F}$ be a coherent Cartier module on X with surjective structural map. Then the set*

$$\{\operatorname{supp} \mathscr{F}/\mathscr{G} \mid \mathscr{G} \subseteq \mathscr{F} \text{ a Cartier submodule}\}$$

is a finite set of reduced subschemes that is closed under finite unions and taking irreducible components.

Proof. We only prove the finiteness and leave the rest as an exercise Exercise* 8.6. We proceed by induction on $\dim X$. By Proposition 8.1.5 we may view $\mathscr{F}$ as a Cartier module on $\operatorname{supp} \mathscr{F}$; hence we may assume that $\operatorname{supp} \mathscr{F} = X$. Since X is Noetherian it has only finitely many irreducible components, so we may assume that X itself is irreducible. If $\operatorname{supp} \mathscr{F}/\mathscr{G} \neq X$, then $\mathscr{F}$ and $\mathscr{G}$ agree on the generic point of X. Hence, the test module $\tau(\mathscr{F}) \subseteq \mathscr{G}$. Therefore

$$\operatorname{supp} \mathscr{F}/\mathscr{G} \subseteq \operatorname{supp} \mathscr{F}/\tau(\mathscr{F}) =: Y,$$

and Y is a proper closed subset of X. Again using Proposition 8.1.5 we can apply the induction hypothesis to the Cartier module $\mathscr{F}/\tau(\mathscr{F})$ on Y whose dimension is strictly less than $\dim X$. □

This yields the following corollary which was obtained in [52] and also independently obtained by the second author in [73]. In the case that $X = \operatorname{Spec} R$ and R is local, proofs of this fact were first obtained in [83] and [18].

Corollary 8.1.8. *Let X be Frobenius split, then $\mathcal{O}_X$ has only finitely many ideals which are compatible with the splitting.*

Proof. If $\varphi: F^e_* \mathcal{O}_X \to \mathcal{O}_X$ is the splitting of Frobenius, note that φ is surjective. The Cartier submodules of $\mathcal{O}_X$ are just the ideals which are φ-compatible. Since $\operatorname{Ann}(\mathcal{O}_X/I) = I$, there is a one-to-one correspondence between the set of φ-compatible ideals, and the set $\operatorname{supp} \mathcal{O}_X/I$ for I a Cartier submodule of $\mathcal{O}_X$. The latter set is finite by the preceding proposition. □

8.2 Cartier Crystals

The finiteness results for Cartier modules of the preceding section receive a more natural formulation if one deals with the notion of nilpotence in a more systematic manner. This is done by localizing the category of coherent Cartier modules at its Serre subcategory[9] of nilpotent Cartier modules. That is, we invert morphisms which are *nil-isomorphisms*, i.e., maps of Cartier modules $\varphi\colon \mathscr{F} \to \mathscr{G}$ whose kernel and cokernel are nilpotent. For the formal definition, see [23, 67], but roughly speaking the localization is defined as follows:

Definition 8.2.1. Let X be a scheme. The *category of $\mathscr{L}$–Cartier crystals* has as objects the coherent Cartier modules on X. A morphism $\varphi\colon \mathscr{F} \to \mathscr{G}$ of Cartier crystals is an equivalence class (left fraction) of diagrams of morphisms of the underlying Cartier modules

$$\varphi\colon \mathscr{F} \leftarrow \mathscr{F}' \xrightarrow{\varphi'} \mathscr{G}$$

where $\mathscr{F}'$ is some Cartier module and $\mathscr{F} \leftarrow \mathscr{F}'$ is a nil-isomorphism. More precisely

$$\mathrm{Hom}_{Crys}(\mathscr{F}, \mathscr{G}) = \mathrm{colim}_{\mathscr{F}' \to \mathscr{F}} \mathrm{Hom}_{Cart}(\mathscr{F}', \mathscr{F})$$

where $\mathscr{F}' \to \mathscr{G}$ ranges over all nil-isomorphisms.

It follows from general principles that the category of Cartier crystals on X is again Abelian. Using this point of view the preceding result can be phrased (and extended) as follows (see [6, Theorem 4.17 and Corollary 4.7]):

Theorem 8.2.2. *Let X be a scheme:*

(a) Each Cartier crystal $\mathscr{F}$ has finite length in the category of Cartier crystals.
(b) Hom-sets in the category of Cartier crystals are finite sets (finite dimensional $\mathbb{F}_{p^e}$ vector spaces).
(c) Each Cartier crystal $\mathscr{F}$ has only finitely many Cartier sub-crystals.

Proof. The first statement follows from Theorem 8.1.6 by noting that $\mathscr{F}$ and $\sigma(\mathscr{F})$ are isomorphic as Cartier crystals (i.e., nil-isomorphic as Cartier modules). The second statement is shown in [6, Theorem 4.17] (but see Exercise 11 for an idea why such a statement may hold), and the last one follows formally from the other two. □

In [5] the category of Cartier crystals (for $\mathscr{L} \cong \mathcal{O}_X$) is thoroughly studied on an arbitrary Noetherian scheme such that $F : X \to X$ is finite. In particular it is shown that half of Grothendieck's six operations, namely, $f^!$, Rf_* and an exotic tensor product, can be defined on a suitable derived category of Cartier crystals. In particular the construction of the functors $f^!$ and Rf_* is rather subtle and bears some

[9] That is, a full Abelian subcategory which is closed under extensions; see [5].

interesting insights. This greatly extends the examples of the pullback for open and closed immersions and the proper pushforward that was discussed in the preceding section.

If $f : Y \to X$ is a proper morphism, then $R^i f_*$ induces a functor on (coherent) Cartier modules, which can be shown to preserve nilpotence. Hence it descends to a functor on Cartier crystals. However, if f is not proper, then already $f_*\mathscr{F}$ of a coherent sheaf is no longer coherent. It is a crucial observation in [5] that if $\mathscr{F}$ is a coherent Cartier crystal on Y, then $R^i f_*\mathscr{F}$ is a *locally nil-coherent* Cartier crystal on X. Nil-coherent for a Cartier module $\mathscr{F}$ means that $\mathscr{F}$ has a coherent Cartier submodule $\mathscr{E} \subseteq \mathscr{F}$ such that the quotient $\mathscr{F}/\mathscr{E}$ is locally nilpotent, i.e., is the union of nilpotent Cartier submodules. This implies the following result:

Theorem 8.2.3. *For an arbitrary finite type morphism $f : Y \to X$, the usual pushforward functor Rf_* on quasi-coherent sheaves induces an exact functor*

$$Rf_* \colon D^b_{\text{crys}}(\text{QCrys}(Y)) \to D^b_{\text{crys}}(\text{QCrys}(X)),$$

where $D^b_{\text{crys}}(\text{QCrys}(\_\_))$ denotes the bounded derived category of quasi-coherent Cartier crystals whose cohomology is locally nil-coherent.

The proof of this result, though not difficult, is somewhat subtle, so we won't attempt it here but instead refer to [5]. However the basic idea is already present in Exercise* 8.7.

The situation with the functor $f^!$ is similar but more subtle. As we have already seen in Sect. 3.3, on quasi-coherent sheaves, the construction of the functor $f^!$ is generally quite involved. Already in the finite case, in particular for a closed immersion $Y \subseteq X$ with X not smooth, one sees that $f^!$ does not have bounded cohomological dimension, hence does not preserve the bounded derived category. However, in [5], it is shown quite generally that $f^!$ preserves local nilpotence, and hence induces a functor on quasi-coherent Cartier crystals. The induced functor on Cartier crystals preserves boundedness up to local nilpotence.

Theorem 8.2.4. *If $f : Y \to X$ is essentially of finite type, then the twisted inverse image functor $f^!$ on quasi-coherent sheaves induces an exact functor*

$$f^! \colon D^b_{\text{crys}}(\text{QCrys}(X)) \to D^b_{\text{crys}}(\text{QCrys}(Y))$$

of bounded cohomological dimension.

Besides a number of obvious compatibilities between these functors which are induced from the corresponding ones of the underlying quasi-coherent sheaves, there are two adjointness statements which are important in the theory.

Proposition 8.2.5. *(a) Let $f : Y \to X$ be a proper morphism. Then as functors on categories $D^b_{\text{crys}}(\text{QCrys}(\_\_))$, the functor Rf_* is naturally left adjoint to $f^!$.*

(b) If $j : Y \to X$ is an open immersion, then j_* is naturally right adjoint to $j^! = j^*$.

For an open immersion $j: U \hookrightarrow X$ and a closed complement $i: Z \hookrightarrow X$ the above adjunction yields natural isomorphisms $i_* i^! \to \mathrm{id}$ and $\mathrm{id} \to j_* j^*$. This yields the following technically important result regarding their combination:

Theorem 8.2.6. *In $D^b_{\mathrm{crys}}(\mathrm{QCrys}(X))$, there is a natural exact triangle*

$$i_* i^! \to \mathrm{id} \to Rj_* j^* \xrightarrow{+1} .$$

This in turn yields a very general form of the Kashiwara equivalence that was alluded to in Proposition 8.1.5.

Theorem 8.2.7. *Let $i: Y \to X$ be a closed immersion. Then $i^!$ and i_* define natural isomorphisms*

$$D^b_{\mathrm{crys}}(\mathrm{QCrys}(Y)) \underset{i^!}{\overset{i_*}{\rightleftarrows}} D^b_{\mathrm{crys},Y}(\mathrm{QCrys}(X))$$

where the right hand category consists of bounded complexes of quasi-coherent Cartier crystals on X whose cohomology is coherent and supported in Y.

8.3 Arithmetic Aspects of p^{-e}-Linear Maps

We conclude with a brief discussion of connections between Cartier crystals and more arithmetic constructions. What follows is much less explicit than previous sections of this chapter, so if the terms used are not familiar to you, we suggest the reader use this as a place to jump off for further reading.

The finite length result for Cartier crystals in Theorem 8.2.2 suggests—in analogy with the Riemann–Hilbert correspondence for D-modules (i.e., modules of the ring of differential operators) on smooth complex manifolds—a connection of Cartier crystals with a category of constructible sheaves. Indeed, in [22], Gabber introduces a family of t-structures on the derived category of bounded complexes of constructible $\mathbb{F}_p$-vector spaces on the étale site of X. He shows that for the middle perversity the heart of this t-structure (i.e., the perverse sheaves with respect to this t-structure) forms an Abelian category which also is Noetherian and Artinian. The connection between Cartier crystals and constructible $\mathbb{F}_p$-vector spaces is a combination of [5, 10] and yields an equivalence of derived categories:

$$D^b_{\mathrm{crys}}(\mathrm{QCrys}(X)) \xrightarrow{\cong} D^b_c(X_{et}, \mathbb{F}_p)$$

where the right hand side is the category of constructible sheaves of $\mathbb{F}_p$-vector spaces on X_{et}. This correspondence is a two-step procedure: First is a Grothendieck–

Serre duality between Cartier crystals (coherent $\mathcal{O}_X$-modules with a *right* action of Frobenius) with the category of τ-crystals (coherent $\mathcal{O}_X$-modules with a *left* Frobenius action) of [10] and was largely motivated by our desire to understand the precise connection of the theory in [10] with the work of Emerton and Kisin [17] and Lyubeznik [59]. This Grothendieck–Serre duality is the main result of [5]. The step from τ-crystals to constructible sheaves is just by taking Frobenius fix-points, i.e., the Artin–Schreier sequence [10].

The first author's PhD student Tobias Schedlmeier has shown in his upcoming thesis that the equivalence is given directly by the functor $\mathrm{Sol}(\_\_) := \mathbf{R}\mathscr{H}om_{\mathrm{crys}}(\_\_, \omega_X^\bullet)$ and proved that the image of the Abelian subcategory of Cartier crystals under Sol is precisely Gabbers category of perverse sheaves $\mathrm{Perv}(X_{et}, \mathbb{F}_p)$ for the middle perversity.

8.4 Exercises

Exercise 1. Prove Lemma 8.1.1.

Exercise 2. Show that the annihilator of any coherent Cartier module $\mathscr{F}$ on X with surjective structural map is a sheaf of radical ideals, *i.e.,* its support is reduced.

Exercise 3. Let R be a ring and $S \subseteq R$ a multiplicative set. Then for any module M, show that $S^{-1}(F_R)_* M \cong (F_{S^{-1}R})_* S^{-1} M$.
Hint: Localize with respect to the multiplicative set S^p is the as with respect to S. This generalizes Exercise 6.

Exercise 4. Let X be a scheme and $\mathscr{L}$ a line bundle. We define a sheaf of rings $\mathcal{O}_X^{\mathscr{L}}[F^e]$ as

$$\mathcal{O}_X \oplus (\mathscr{L} \cdot F^e) \oplus (\mathscr{L}^{1+p^e} \cdot F^{2e}) \oplus (\mathscr{L}^{1+p^e+p^{2e}} \cdot F^{3e}) \oplus \cdots$$

where F^{ne} are formal symbols, and the multiplication of homogeneous elements lF^{ne} and $l'F^{ne'}$ is defined as $lF^{ne} l' F^{n'e} = l(l')^{p^{n'e}} F^{(n+n')e}$.

(a) Show that this defines the structure of a sheaf of rings on $\mathcal{O}_X^{\mathscr{L}}[F^e]$.
(b) Show that the category of $(\mathscr{L}, p^e)$–Cartier modules is equivalent to the category of (sheaves of) right $\mathcal{O}_X^{\mathscr{L}}[F^e]$-modules.

Hint: Do the case of $\mathscr{L} \cong \mathcal{O}_X$ first, and then attempt the general case.

Exercise 5. If $0 \to \mathscr{F}' \to \mathscr{F} \to \mathscr{F}'' \to 0$ is an exact sequence of coherent Cartier modules, show that $\mathscr{F}'$ and $\mathscr{F}''$ are nilpotent (of order $\leq e, e'$) if and only if $\mathscr{F}$ is nilpotent (of order $\leq e + e'$).

Exercise* 8.6. Let $\mathscr{F}$ be a quasi-coherent Cartier module with surjective structural map. Show that the collection

$$\{\mathrm{supp}(\mathscr{F}/\mathscr{G}) \,|\, \mathscr{G} \subseteq \mathscr{F} \text{ a Cartier submodule}\}$$

is a collection of reduced subschemes that is closed under finite unions and taking irreducible components.

Exercise* 8.7. Let $X = \operatorname{Spec} R$ be an affine scheme and $U = \operatorname{Spec} R_f$ a basic open subset with $f \in R$, and denote the open inclusion $U \subseteq X$ by j. Let $\mathscr{F}$ be a coherent Cartier module on U. Show that $j_*\mathscr{F}$ has a coherent Cartier submodule F such that the quotient $j_*\mathscr{F}/F$ is locally nilpotent, i.e., the union of its nilpotent Cartier submodules.

Exercise 8. Let $\mathscr{F}$ be a coherent Cartier module on X. The *test submodule* $\tau(\mathscr{F})$ is defined as the smallest Cartier submodule $\mathscr{G} \subseteq \mathscr{F}$ which agrees with $\sigma(\mathscr{F})$ for each generic point of X. Show that Theorem 8.1.6 implies the existence and uniqueness of $\tau(\mathscr{F})$.

Exercise 9. Suppose that R is a ring and (M, φ) is a Cartier module on M. Suppose further that $R \to S$ is a finite ring homomorphism. Prove that $\operatorname{Hom}_R(S, M)$ has the structure of a Cartier module induced by φ and by the Frobenius map $S \to F_*S$.

Exercise 10. Prove Proposition 8.1.5.

Exercise 11. Let R be a regular F-finite ring with dualizing sheaf ω_R with its standard Cartier structure $T : F_*\omega_R \to \omega_R$ (see Sect. 3.2). Show that the homomorphisms of Cartier modules $\operatorname{Hom}_{\mathrm{Cart}}(\omega_R, \omega_R) = R^F = \mathbb{F}_p$ are just the Frobenius fixed points of the action of F on R. In particular, this Hom-set is finite.

Exercise* 8.12. Suppose that $R = k[x_1, \ldots, x_4]_{\langle x_1,\ldots,x_4\rangle}$ and that $\varphi : F^e_* R \to R$ is a Frobenius splitting. In Corollary 8.1.8, it was shown that there are at most finitely many φ-compatible ideals.

Prove that there at most $\binom{4}{d}$ prime ideals Q which are compatibly split by φ such that $\dim(R/Q) = d$.

Hint: Prove it for $d = 0$ first (very easy), then $d = 1$ (use the fact that compatibly split subvarieties must intersect normally, Corollary 5.1.7, but we only have 4 "directions" in $\operatorname{Spec} R$, which is just the origin in $\mathbb{A}^4$). For $d = 2, 3$, simply consider all possibilities exhaustively (keeping in mind the normal intersections). For a complete proof for any $\mathbb{A}^n$ (not just $n = 4$), see [79].

Exercise 13. Suppose that $(\mathscr{F}, \kappa)$ is an $(\mathscr{L}, p^e)$-Cartier module on a projective variety X such that the structural map $\kappa : F^e_*(\mathscr{L} \otimes \mathscr{F}) \to \mathscr{F}$ is surjective. Further suppose that $\mathscr{A}$ is a globally generated ample line bundle and that $\mathscr{N}$ is another line bundle such that $\mathscr{N}^{p^e-1} \otimes \mathscr{L}$ is ample. Prove that

$$\mathscr{F} \otimes \mathscr{A}^{\dim X} \otimes \mathscr{N}$$

is a globally generated sheaf.

Hint: Use the same strategy as in Theorem 6.1.3.

9 Applications to Local Cohomology and Test Ideals

In this section we discuss in detail the relation of the theory of Cartier modules to other theories of modules with a Frobenius action, with an emphasis on applications to local cohomology. Then we discuss a simple but interesting degree-reducing property of Cartier linear maps, which allows an elementary treatment of the theory of Cartier modules in the case that X is of finite type over a perfect field. We use this approach to study the test ideals and show the discreteness of their jumping numbers.

9.1 Cartier Modules and Local Cohomology

The category of Cartier modules, besides enjoying some extraordinary finiteness conditions, is useful due to its connection to other categories which are studied, in particular in connection with local cohomology. Besides the connection to constructible p-torsion sheaves that we hinted at above, we show the relation to two further categories which are particularly important in the study of the local cohomology of rings in positive characteristic. Our goal is to explain the following diagram of categories and to derive a number of finiteness results for local cohomology from the above finiteness result for Cartier modules.

$$\left\{\begin{matrix}\text{cofinite } R\text{-modules}\\ \text{with left Frobenius action}\\ (R \text{ complete local ring})\end{matrix}\right\} \leftrightarrow \left\{\begin{matrix}\text{coherent Cartier modules on } X\\ (X \text{ Noetherian and } F\text{-finite})\end{matrix}\right\} \to \left\{\begin{matrix}\text{Lyubeznik's } F\text{-modules over } R\\ (R \text{ regular, Noetherian ring})\end{matrix}\right\}.$$

The parenthetical parts indicate in what generality the categories are defined, and the arrows are defined when both assumption holds, for example, the first double arrow holds for complete local and F-finite rings. The left double arrow is an equivalence of categories given by Matlis duality $\mathrm{Hom}_R(\_\_, E_{R/\mathfrak{m}})$ where $E_{R/\mathfrak{m}}$ is an injective hull of the perfect residue field of R. The right arrow is a functor which gives an equivalence after inverting Cartier modules at nil-isomorphisms, i.e., it induces an equivalence of categories from Cartier crystals to F-finite modules. Lyubeznik's F-finite modules and this equivalence will be explained in detail below.

Let us begin with Matlis duality. Let $(R, \mathfrak{m})$ be complete and local and denote by $E = E_R$ an injective hull of the prefect residue field of R. Since R is F-finite one has that $F^e_* F^! E_R := \mathrm{Hom}_R(F_* R, E_R) \cong E_{F_* R}$ which we identify with E_R since R and $F_* R$ are isomorphic as rings. We fix hence an isomorphism $F^! E \cong E$. If we denote by $(\_\_)^\vee = \mathrm{Hom}_R(\_\_, E_R)$ the Matlis duality functor, we have the following lemma whose proof we leave as Exercise 2.

Lemma 9.1.1. *For* $(R, \mathfrak{m})$ *local and* F*-finite there is a (functorial) isomorphism* $F_*(\_\_)^\vee \cong (F_*\_\_)^\vee$.

This immediately implies the first of the equivalences above, also cf. [85].

Proposition 9.1.2. *Let $(R, \mathfrak{m})$ be complete, local, and F-finite. Then Matlis duality induces an equivalence between the categories of*

$$\left\{ \begin{array}{c} \text{co-finite } R\text{-modules} \\ \text{with left Frobenius action} \end{array} \right\} \leftrightarrow \left\{ \begin{array}{c} \text{finitely generated } R\text{-modules} \\ \text{with right Frobenius action} \end{array} \right\}.$$

Of course the R-modules with right Frobenius action are just the coherent Cartier modules on $X = \operatorname{Spec} R$. The equivalence preserves nilpotence.

Proof. A left action of Frobenius on M is an R-linear map $\varphi: M \to F_*M$. Applying Matlis duality and the preceding lemma, this yields a map

$$F_*(M^\vee) \cong (F_*M)^\vee \xrightarrow{\varphi^\vee} M^\vee$$

which is the desired Cartier structure (=right Frobenius action) on the dual $M^\vee$. The same construction works in the opposite direction, and one immediately checks that this induces an equivalence of categories. □

With this result we can translate the finiteness theorems for Cartier modules obtained above to the setting of cofinite R-modules with a left Frobenius action. In particular the results hold for local cohomology modules $H^i_\mathfrak{m}(R)$ with support in the maximal ideal $\mathfrak{m}$ of R.

Theorem 9.1.3. *Let N be a cofinite R module equipped with a p-linear map $F: N \to N$ (i.e., F is additive and $F(rm) = r^p F(m)$).*

(a) The ascending chain of submodules $\ker F \subseteq \ker F^2 \subseteq \ker F^3 \subseteq \cdots$ stabilizes ([37, Proposition 1.1]).
(b) Any chain $\cdots \subseteq N_i \subseteq N_{i+1} \subseteq N_{i+2} \subseteq \cdots$ of submodules $N_i \subseteq N$ which are stable under F (i.e., $F(N_i) \subseteq N_i$) has eventually F-nilpotent quotients ([59, Theorem 4.7]).
(c) N has up to nilpotent action of F, only finitely many F-stable submodules. Concretely, there are only finitely many F stable submodules N' for which the action of F on the quotient N/N' is injective.

Proof. These are just the Matlis dual statements of Proposition 8.1.4, Theorem 8.1.6, and Theorem 8.2.2 part (c). □

An immediate consequence of these observations is the following result originally obtained by Enescu and Hochster [18]; see [62] for a recent extension showing that F-split alone is sufficient in the assumptions below.

Proposition 9.1.4. *If R is quasi-Gorenstein (i.e., $H^d_\mathfrak{m}(R) \cong E_R$) and F-split, then the top local cohomology module $H^d_\mathfrak{m}(R)$ with its left action of the Frobenius has only finitely many F-stable submodules.*

Proof. The existence of a splitting $\varphi \colon R \to S$ implies that the Cartier module (R, φ) has only finitely many Cartier submodules. Hence, by the above duality result, its dual $(r^\vee = H^d_{\mathfrak{m}}(R), \varphi^\vee)$ has only finitely many submodules stable under the action of $\varphi^\vee$. But $H^d_{\mathfrak{m}}(R)$ also has a natural Frobenius action F_H induced by the Frobenius on R by functoriality of $H^d_{\mathfrak{m}}(\_\_)$. One can show (Exercise 7) that there is a $r \in R$ such that $\varphi^\vee = rF_H$. Hence, all submodules which are stable under F_H are also stable under $\varphi^\vee$, but of the latter, there are only finitely many as just argued. □

The connection of Cartier modules with Lyubeznik's F-finite modules also relies on a certain commutation of functors which we recall first. Lyubeznik's theory [59] is phrased for a regular ring R, and even though there is an extension to schemes by Emerton and Kisin [17], we will stick to this setting and assume from now on that $X = \operatorname{Spec} R$, with R regular (and such that the Frobenius morphism $F : R \to R$ is finite).

Lemma 9.1.5. *Let $f \colon Y \to X$ be a finite flat morphism and M a $\mathcal{O}_X$ module, then there is a functorial isomorphism*

$$f^! \mathcal{O}_X \otimes_{\mathcal{O}_Y} f^* M \cong f^! M \,,$$

where $f^!(\_\_) = \operatorname{Hom}_{\mathcal{O}_X}(f_* \mathcal{O}_X, \_\_)$.

Proof. See Exercise 4. □

Applying this to the case of the Frobenius on the regular scheme X and $M = \omega_X$ the dualizing sheaf (which is invertible!), we obtain an isomorphism

$$F^! \mathcal{O}_X \cong F^! \omega_X \otimes F^* \omega_X^{-1} \,.$$

Further using that the adjoint of the map $F_* \omega_X \to \omega_X$ coming from the Cartier isomorphism in Theorem 3.1.1, is an isomorphism $\omega_X \to F^! \omega_X$, we obtain

$$F^! M \otimes \omega_X^{-1} \cong F^* M \otimes F^! \omega_X \otimes \omega_X^{-1} \otimes F^* \omega \cong F^*(M \otimes \omega_X^{-1})$$

which allows us to describe the functor from Cartier modules to Lyubeznik's F-finite modules. Starting with a Cartier module M with structural map $\kappa \colon F_* M \to M$ we first consider its adjoint $\kappa' \colon M \to F^! M$ and tensor it with ω_X^{-1} to obtain

$$\gamma \colon M \otimes \omega_X^{-1} \xrightarrow{\kappa' \otimes \mathrm{id}} F^!(M) \otimes \omega_X^{-1} \cong F^*(M \otimes \omega_X^{-1})$$

where the final isomorphism is the one derived above. Let us pause for a moment to recall the definition of Lyubeznik's F-finite modules, which we phrase in a way convenient for our purpose:

Definition 9.1.6. Let R be regular. Given a finitely generated R-module N together with a map $\gamma \colon N \to F^* N$, then an *$F$-finite module* is the limit $\mathcal{N}$ of the directed system

$$N \xrightarrow{\gamma} F^* N \xrightarrow{F^* \gamma} F^{2*} N \xrightarrow{F^{2*} \gamma} F^{3*} N \to \cdots$$

together with the induced map $\vartheta\colon \mathscr{N} \xrightarrow{\cong} F^*\mathscr{N}$ which is immediately verified to be an isomorphism.

Phrased differently, an *F-finite module* is a (not necessarily finitely generated) R-module $\mathscr{N}$ together with an isomorphism $\vartheta\colon \mathscr{N} \xrightarrow{\cong} F^*\mathscr{N}$ which arises in the above described manner from a *finitely generated* R-module N.

It is shown in [59] that F-finite modules are an Abelian category which is closed under extensions, that local cohomology modules $H^i_I(R)$ are F-finite modules, and that F-finite modules enjoy a number of important finiteness results. For example, they have only finitely many associated primes, and all Bass numbers are finite.

From this definition it is immediate how to connect the Cartier modules with F-finite modules. The F-finite module attached to a Cartier module M is just the limit of

$$M \otimes \omega_X^{-1} \to F^*(M \otimes \omega_X^{-1}) \to F^{2*}(M \otimes \omega_X^{-1}) \to \cdots .$$

One obtains the following Proposition [6].

Proposition 9.1.7. *For a regular ring R, the just described construction assigning to a coherent Cartier module M on R an F-finite module is an essentially surjective functor*

$$\{\textit{coherent Cartier modules}\} \to \{F\textit{-finite modules}\}$$

which sends nilpotent Cartier modules to zero. The induced functor

$$\{\textit{coherent Cartier}\ \text{crystals}\} \xrightarrow{\cong} \{F\textit{-finite modules}\}$$

is an equivalence of categories.

Proof. All statements are shown in [6], but with the above preparations none of them is particularly difficult. □

Hence we obtain as an immediate consequence of Theorem 8.2.2 the following finiteness result for F-finite modules, which partially extends one of the main results of [59]:

Theorem 9.1.8. *Let R be regular and F-finite, then*

(a) F-finite modules over R-have finite length.
(b) The Hom-*sets in the category of F-finite modules are finite.*
(c) An F-finite module has only finitely many F-finite submodules.

Part (a) of the theorem has been proven for R regular and of finite type over a regular local ring in [59] and for arbitrary F-finite schemes X in [5]. The latter results also are shown for regular rings (part (b) even without the F-finiteness assumption) in [39]. Finally, let us state the aforementioned finiteness result for local cohomology modules.

Theorem 9.1.9. *Let M be an F-finite module, $I \subseteq R$ an ideal in a regular ring, then $H_I^j(M)$ is an F-finite R-module and hence has only finitely many associated primes.*

Proof. We only have to show that $H_I^j(M)$ is an F-finite module. The crucial step is to show that for $f \in R$ we have that the localization M_f is also F-finite (cf. Exercise* 8.7). Once this is established, the Čech-complex finishes the proof. □

9.2 Contracting Property of p^{-e}-Linear Maps

In this section we point out a simple fact about p^{-e}-linear map which has a number of interesting consequences. In particular we give an elementary proof of the finite length result for Cartier modules. The idea goes back at least to a paper of Anderson [2] and says that a p^{-e}-linear endomorphism reduces the degree in a graded context. For this we consider $X = \operatorname{Spec} S$ with $S = k[x_1, \ldots, x_n]$ a polynomial ring over a perfect field. Then we consider the filtration of S given by the finite-dimensional vector spaces

$$S_d := k\langle x_1^{i_1} \cdots x_n^{i_n} \mid 0 \leq i_j \leq d \text{ for } j = 1, \ldots, n\rangle .$$

Hence S_d is the k–subspace of S freely generated by the monomials with degree $\leq d$ *in each variable*. One immediately verifies that

$$S_{-\infty} := 0, \quad S_0 = k, \quad S_d S_{d'} \subseteq S_{d+d'}, \text{ and } \quad S_d + S_d' \subseteq S_{\max d+d'} .$$

For each choice of a set of generators $m_1, \ldots, m_k$ of an S module M we define the induced filtration on M given by

$$M_{-\infty} := 0 \text{ and } M_d = S_d\langle m_1, \ldots, m_k\rangle.$$

For $m \in M$ we write $\delta(m) = d$ if and only if $m \in M_d \setminus M_{d-1}$ and call $\delta = \delta_M$ a *gauge* for M. One should think of the gauge δ as a substitute for a degree on M, and the contracting property of p^{-e}-linear maps on M is measured in terms of the gauge δ. Spelling out the definition we see that $\delta(m) \leq d$ if m can be written as a S-linear combination of the m_i such that all coefficients are in S_d. S itself has a gauge, induced by the generator 1. We summarize the immediate properties of a gauge (the proof is left to the reader in Exercise 3):

Lemma 9.2.1. *Let M be finitely generated over $S = k[x_1, \ldots, x_n]$, and δ a gauge corresponding to some generators $m_1, \ldots, m_k$ of M. Then*

(a) $\delta(m) = -\infty$ if and only if $m = 0$.
(b) Each M_d is finite dimensional over k (since S_d is).
(c) $\bigcup_d M_d = M$ (since the m_i generate M).

(d) $\delta(m + m') \leq \max\{\delta(m), \delta(m')\}$.
(e) $\delta_M(fm) \leq \delta_S(f) + \delta_M(m)$.

Proposition 9.2.2 ([2], Proposition 3). *Let M be a finitely generated S-module, and $\delta = \delta_M$ a gauge corresponding to some generators $m_1, \ldots, m_k$ of M, and let $\varphi : M \to M$ be a p^{-e}-linear map. Then there is a constant K such that for all $m \in M$:*

$$\delta(\varphi(m)) \leq \frac{\delta(m)}{p^e} + \frac{K}{p^e}$$

Furthermore, for all $n \geq 0$, we have

$$\delta(\varphi^n(m)) \leq \frac{\delta(m)}{p^{ne}} + \frac{K}{p^e - 1}.$$

Proof. By definition, we may write $m = \sum_{l=1}^{k} f_l m_l$ with $\delta_S(f_l) \leq \delta(m)$. For each l write uniquely $f_l = \sum_{\mathbf{x}^{\mathbf{i}} \in S_{\delta(m)}} r_{l,i}^{p^e} \mathbf{x}^{\mathbf{i}}$ with $\mathbf{x}^{\mathbf{i}} = x_1^{i_1} \cdots x_n^{i_n}$. Then Exercise 1 shows that $\delta_S(r_{l,i}) \leq \lfloor \delta(m)/p^e \rfloor$. Writing this out

$$\varphi(m) = \sum_{l=1}^{k} \sum_{x^i \in S_{\delta(m)}} r_{l,i} \varphi(\mathbf{x}^i m_l),$$

we consequently obtain

$$\delta(\varphi(m)) \leq \max_{l,i}\{\delta_S(r_{l,i}) + \delta(\varphi(x^i m_l))\} \leq \lfloor \frac{\delta(m)}{p^e} \rfloor + \frac{K}{p^e}.$$

Taking for $K = p^e \cdot \max_{l,i}\{\delta(x^i m_l)\}$ we obtain the claimed inequality.

The final inequality follows by applying the first inequality iteratively and then to use the geometric series (exercise!). □

This proposition has an important consequence about the generators of the images of a submodule under a p^{-e}-linear map.

Lemma 9.2.3. *Let $\varphi \colon M \to M$ be a p^{-e}-linear map on the finitely generated S-module M with gauge δ and bound K as in Proposition 9.2.2. Suppose that the S-submodule $N \subseteq M$ is generated by elements with gauge $\leq d$. Then $\varphi^n(N) \subseteq M$ is generated by elements of gauge at most $d/p^{ne} + K/(p^e - 1) + 1$.*

Proof. If N is generated by $n_1, \ldots, n_t$, then $\varphi^n(N)$ is generated by $\varphi(x^i n_j)$ where $0 \leq i_1, \ldots, i_n \leq p^{ne} - 1$ and $j = 1, \ldots, t$. Now, if each $\delta(n_j) \leq d$, then

$$\delta(\varphi(x^i n_j)) \leq \frac{\delta(x^i n_j)}{p^{ne}} + \frac{K}{p^e - 1} \leq \frac{(p^{ne} - 1) + d}{p^{ne}} + \frac{K}{p^e - 1} \leq 1 + \frac{d}{p^{ne}} + \frac{K}{p^e - 1}.$$

□

Corollary 9.2.4. *Let $\varphi\colon M \to M$ be a Cartier module with gauge δ and bound K (as in Proposition 9.2.2). Then every Cartier submodule $N \subseteq M$ with surjective structural map $\varphi\colon N \twoheadrightarrow N$ is generated by elements in the finite dimensional k-vector space $M_{\frac{K}{p^e-1}+1}$ (independently of N).*

Proof. If N has surjective structural maps, then for each n, we have $\varphi^n(N) = N$. Since N is finitely generated, it is generated by elements of some gauge $\leq d$. By the above lemma, we have hence for all n that N is generated by elements of gauge $\leq d/p^{ne} + K/(p^e-1) + 1$. But for n big enough the first term is irrelevant (less than 1), and the result follows. □

Corollary 9.2.5. *In a coherent Cartier module M there are no infinite proper chains of Cartier submodules N_i, each with surjective structural map.*

Proof. Each N_i has generators in the finite dimensional k-vector space $M_{\frac{K}{p^e-1}+1}$; hence there cannot be any infinite proper chains. □

As we have alluded to (in Exercise 8) before, the fact that there are no infinite chains of Cartier submodules with surjective structural maps implies the existence of the test module $\tau(M)$. By definition of being the smallest Cartier submodule of M which generically agrees with $\sigma(M)$ it is clear that $\tau(M)$ has surjective structural map (since the image under the structural map would again be of that type). The intersection of two Cartier submodules agreeing generically with $\sigma(M)$ clearly also has this property. Now, the existence of $\tau(M)$ follows from the stabilization of any chain of submodules generically agreeing with $\sigma(M)$ and with surjective structural maps, which we just showed.

In the next section, Sect. 9.3, we will show this approach to Cartier modules via gauges also gives an elementary proof of the discreteness of jumping numbers for test ideals.

We conclude this section with pointing out that our restriction to the polynomial ring $S = k[x_1, \ldots, x_n]$ is not very restrictive after all. In the case of an arbitrary scheme X we may reduce to the affine case by considering an affine cover. Then any finite type k-algebra $R = S/I$ is the quotient of a polynomial ring. Then we can use the Kashiwara-equivalence Proposition 8.1.5 to reduce to the case of the polynomial ring itself.

Remark 9.2.6 (Historical discussion). The major source of inspiration to explore the contracting property of p^{-e}-linear maps in [4] came from a paper of Anderson [2] where he uses this property to study L-functions $\mod p$ on varieties over $\mathbb{F}_p$. The key observation there is that if $\varphi\colon M \to M$ is a p^{-e}-linear map of R-modules (say R of finite type over $\mathbb{F}_p$) then there is a finite dimensional $\mathbb{F}_p$-subspace into which every element of M is eventually contracted by iterated application of φ. This allows him, inspired by Tate's work [93], to develop a trace calculus for these operators. This is then used to show the rationality of L-functions $\mod p$ attached to a finitely generated R-module M with a left action of Frobenius F on M. In fact, he shows that if R is the polynomial ring and M is projective, this L-function is equal to the characteristic polynomial (or its inverse) of the action of the dual of F on $M^\vee$. This dual is a Cartier linear endomorphism of $M^\vee$, and the characteristic polynomial is defined via the important contracting property of Cartier linear maps.

9.3 Algebras of Maps and the Test Ideal

Suppose that $X = \operatorname{Spec} R$ is an affine variety (for simplicity). Previously we considered finitely generated R-modules M and p^{-e}-linear maps $\varphi : M \to M$. Unless $M = \omega_R$ (or is obtained functorially from $T : \omega_Y \to \omega_Y$ from some other variety Y), there probably is no *natural* choice of φ.

The obvious solution is to choose all possible φ; see [4, 76] and cf. [61] for a dual formulation. For any *finitely generated* module M, we set $\operatorname{End}_e(M)$ to be the set of p^{-e}-linear maps from M to M. In other words, $\operatorname{End}_e(M)$ is just $\operatorname{Hom}_R(F^e_* M, M)$. Of course $\operatorname{End}_e(M)$ has an R-module structure via both the source and target R-module structures. Notice that if $\varphi \in \operatorname{End}_e(M)$ and $\psi \in \operatorname{End}_d(M)$, then we can form the composition $\psi \circ \varphi \in \operatorname{End}_{e+d}(M)$. Thus $\operatorname{End}_*(M) = \oplus_{e \geq 0} \operatorname{End}_e(M)$ forms a noncommutative graded ring. Unfortunately, the ring $\operatorname{End}_0(M)$ is often too big, and so we set $\mathscr{C}_0^M$ to denote the image of R inside $\operatorname{End}_0(M)$ via the natural map that sends $r \in R$ to the multiplication by r map on M.

Definition 9.3.1 (Cartier algebras). An *(abstract) Cartier algebra* over R[10] is an $\mathbb{N}$-graded ring $\mathscr{C} = \bigoplus_{e \geq 0} \mathscr{C}_e$ satisfying the rule $r \cdot \varphi_e = \varphi_e \cdot r^{p^e}$ for all $\varphi_e \in \mathscr{C}_e$ and $r \in R$ and furthermore such that $\mathscr{C}_0 \cong R/I$ for some ideal I.

Example 9.3.2. Suppose that M is a finitely generated R-module. The *total Cartier algebra on M*, denoted $\mathscr{C}^M$, is the following graded subring of $\operatorname{End}_*(M)$:

$$\mathscr{C}^M := \mathscr{C}_0^M \oplus \left(\bigoplus_{e>0} \operatorname{End}_e(M) \right) = \bigoplus_{e \geq 0} \mathscr{C}_e^M.$$

It is obviously a Cartier algebra.

A *Cartier subalgebra (on M)* is any graded subring $\mathscr{C} \subseteq \mathscr{C}^M$ such that $[\mathscr{C}]_0 = \mathscr{C}_0^M$.

With the above definitions, if $\mathscr{C}$ is an (abstract) Cartier algebra, and M is any left-$\mathscr{C}$-module, then there is a natural map $\mathscr{C} \to \mathscr{C}^M$, the image of which is a Cartier subalgebra on M. Conversely, note that any Cartier subalgebra $\mathscr{C} \subseteq \mathscr{C}^M$ acts on M by the application of functions. In particular, M is also a $\mathscr{C}$-module.

---

[10] It is important to note that while we call it an *algebra*, it is not generally an R-algebra because R is not central.

Remark 9.3.3. Most commonly, we will consider $\mathscr{C}^R$, in which case $\mathscr{C}_0^R = \mathrm{Hom}_R(R, R) = \mathrm{End}_0(R)$ automatically.

Now suppose that $\mathscr{C}$ is a Cartier algebra and that M is a left $\mathscr{C}$-module (or that $\mathscr{C}$ is a Cartier-submodule on M), we use $\mathscr{C}_+$ to denote $\oplus_{e>0}\mathscr{C}_e$. It is easy to see that $\mathscr{C}_+$ is a 2-sided ideal. For any $\mathscr{C}$-submodule $N \subseteq M$, we define

$$\mathscr{C}_+ N := \langle \varphi(x) \,|\, x \in N, \varphi \in \mathscr{C}_e \text{ for some } e > 0\rangle_R \subseteq N$$

to be the submodule generated by all $\varphi(x)$ for homogeneous $\varphi \in \mathscr{C}_+$ and $n \in N$. We set

$$(\mathscr{C}_+)^n N := \underbrace{\mathscr{C}_+(\mathscr{C}_+(\cdots \mathscr{C}_+}_{n\text{-times}}(N))) \subseteq N.$$

A crucial step from dealing with an algebra of Cartier linear operators as opposed to a single one is to establish the right notion of nilpotence. With following definition the theory develops in surprising analogy to the single operator case dealt with above.

Definition 9.3.4. We say that N is *$\mathscr{C}$-nilpotent* if $(\mathscr{C}_+)^n N = 0$ for some $n > 0$.

It is obvious we have a chain of inequalities:

$$N \supseteq \mathscr{C}_+ N \supseteq (\mathscr{C}_+)^2 N \supseteq \cdots \supseteq (\mathscr{C}_+)^i N \supseteq (\mathscr{C}_+)^{i+1} N \supseteq \cdots. \tag{20}$$

The following remarkable theorem about this chain generalizes Proposition 8.1.4.

Theorem 9.3.5 ([4, Proposition 2.14]). *Suppose that M is a finitely generated R-module that is also a left $\mathscr{C}$-module for some Cartier algebra $\mathscr{C}$. Then $(\mathscr{C}_+)^n M = (\mathscr{C}_+)^{n+1} M$ for all $n \gg 0$. In other words, the chain of submodules in (20) eventually stabilizes.*

Proof. The proof is similar to that of Proposition 8.1.4 and left to the reader in Exercise* 9.6. □

As an immediate corollary we obtain

Corollary 9.3.6 ([4, Corollary 2.14]). *Let M be a finitely generated R-module that is also an $\mathscr{C}$-module for some Cartier algebra $\mathscr{C}$. Then there is a unique $\mathscr{C}$-submodule $\sigma(M) \subseteq M$ such that*

(a) The quotient $M/\sigma(M)$ is nilpotent.
(b) $\mathscr{C}_+\sigma(M) = \sigma(M)$, and so $\sigma(M)$ does not have nilpotent quotients.

Proof. Set $\sigma(M) = (\mathscr{C}_+)^n M$ for $n \gg 0$, then verify the statements in Exercise 8. □

Suppose that M is a finitely generated R-module and a left $\mathscr{C}$-module. We can now define a notion of the test ideal on M.

Definition 9.3.7. Suppose that M and $\mathscr{C}$ are as above. Then we define the *test submodule* $\tau(M, \mathscr{C})$ to be the unique smallest submodule N of M which:

(a) Is a $\mathscr{C}$-module
(b) Satisfies $(\sigma(M))_\eta = N_\eta$ for every minimal prime of R[11]

if it exists.

The existence of $\tau(M, \mathscr{C})$ is known in many important cases, but not in all generality. It is known to exist if R is of finite type over a field (or a localization of such), or if $\mathscr{C}$ is generated by a single operator; see [4, Theorem 4.13, Corollary 3.18]. It is also known to exist if $M = R$ by the same argument as Proposition 9.3.10.

For the rest of the section, we consider $\mathscr{C}^R$, the total Cartier algebra on R, and subalgebras of it. Indeed, a common way to construct a Cartier algebra is as follows.

Definition 9.3.8. Suppose that R is a normal domain with $X = \operatorname{Spec} R$. Suppose further that $\Delta \geq 0$ is an effective $\mathbb{Q}$-divisor, $\mathfrak{a} \subseteq R$ is a nonzero ideal, and $t \geq 0$ is a real number. Then we define the following Cartier subalgebra of $\mathscr{C}^R$. For each $e \geq 0$ first identify $\operatorname{Hom}_R(F^e_* R, R)$ with $\mathscr{C}^R_e$ and fix $\mathscr{C}^\Delta_e$ to be the subset

$$\operatorname{Hom}_R(F^e_* R(\lceil (p^e-1)\Delta \rceil), R) \subseteq \operatorname{Hom}_R(F^e_* R, R) = \mathscr{C}^R_e.$$

Here $R(\lceil (p^e-1)\Delta \rceil) = \Gamma(X, \mathcal{O}_X(\lceil (p^e-1)\Delta \rceil))$.

It follows that

$$\mathscr{C}^\Delta := \bigoplus_{e \geq 0} \mathscr{C}^\Delta_e$$

is a Cartier subalgebra of $\mathscr{C}^R$ (the details will be left as Exercise 9).

Furthermore, we can form $\mathscr{C}^{\Delta, \mathfrak{a}^t}_e := \mathscr{C}^\Delta_e \cdot \mathfrak{a}^{\lceil t(p^e-1) \rceil}$ (where multiplication on the right is pre-composition; in other words, $\mathscr{C}^{\Delta, \mathfrak{a}^t}_e$ is identified with $\operatorname{Hom}_R(F^e_* R(\lceil (p^e - 1)\Delta \rceil), R) \cdot (F^e_* \mathfrak{a}^{\lceil t(p^e-1) \rceil})$). Again the direct sum

$$\mathscr{C}^{\Delta, \mathfrak{a}^t} := \bigoplus_{e \geq 0} \mathscr{C}^{\Delta, \mathfrak{a}^t}_e$$

is a Cartier subalgebra of $\mathscr{C}^R$; see Exercise 9.

With these definitions, we can now define the test ideal $\tau(R; \Delta, \mathfrak{a}^t) := \tau(R, \mathscr{C}^{\Delta, \mathfrak{a}^t})$ [4, 76].

Remark 9.3.9. Test ideals (with $\Delta = 0$ and $\mathfrak{a} = R$) were originally introduced by Hochster and Huneke in their theory of tight closure [41]. In fact, what we call the test ideal is often called the *big test ideal* [38] and is denoted by $\widetilde{\tau}$ or τ_b. This object

[11] This definition differs slightly from the original one given in [4] where one requires equality for every minimal prime of $\sigma(M)$ instead of R. Though this yields different results in general, in light of the Kashiwara equivalence Proposition 8.1.5, the respective theories imply each other.

though is better behaved with respect to geometric operations (such as localization Exercise 11). It is conjectured that $\widetilde{\tau}$ and τ coincide in general [60, 61].

Even with $\Delta \neq 0$ and $\mathfrak{a} \neq R$, this definition we gave is not the original one. For $\mathfrak{a} \neq R$, $\tau(R; \mathfrak{a}^t)$ was originally defined in [29] (and $\widetilde{\tau}(R; \mathfrak{a}^t)$ was studied in [30]). For $\Delta \neq 0$, $\tau(R; \Delta)$ was introduced in [90].

Proposition 9.3.10. *Suppose R is a normal domain. The test ideal $\tau(R, \mathscr{C}^{\Delta,\mathfrak{a}^t}) = \tau(R; \Delta, \mathfrak{a}^t)$ exists.*

Proof. The main point is the following lemma, which is a generalization of a result of Hochster and Huneke.

Lemma 9.3.11 ([41, Sect. 6], [76, Lemma 3.21]). *There exists an element $0 \neq c \in R$ such that for every $0 \neq d \in R$, there exists $e > 0$ such that $c \in \mathscr{C}_e^{\Delta,\mathfrak{a}^t}(dR)$.*

Now choose c as in the lemma, and it follows that $c \in I$ for any nonzero $\mathscr{C}^{\Delta,\mathfrak{a}^t}$-submodule $I \subseteq R$. However,

$$\sum_{e \geq 0} \mathscr{C}_e^{\Delta,\mathfrak{a}^t}(Rc)$$

is evidently the smallest $\mathscr{C}^{\Delta,\mathfrak{a}^t}$ submodule containing c. □

One of the aspects of the test ideal which has attracted the most interest over the past few years is how the test ideal $\tau(R; \Delta, \mathfrak{a}^t)$ changes as t varies. First we mention the following lemma which serves as a baseline for how the test ideal behaves.

Lemma 9.3.12 ([69, Remark 2.12], [7, Proposition 2.14], [9, Lemma 3.23]). *With notation as above, for every real number $t \geq 0$, there exists an $\varepsilon > 0$ such that*

$$\tau(R; \Delta, \mathfrak{a}^t) = \tau(R; \Delta, \mathfrak{a}^s)$$

for every $s \in [t, t + \varepsilon]$.

Proof. The containment $\supseteq$ is obvious. A substantial hint is given in Exercise* 9.12. □

Because of this, we make the following definition:

Definition 9.3.13 (F-jumping numbers). Suppose that $(R, \Delta, \mathfrak{a}^t)$ are as above. Then a number $t > 0$ is called an *F-jumping number* if

$$\tau(R; \Delta, \mathfrak{a}^t) \neq \tau(R; \Delta, \mathfrak{a}^{t-\varepsilon})$$

for all $1 \gg \varepsilon > 0$.

Based on the above lemma and a connection between test ideals and multiplier ideals [29, 90] it is natural to expect that the set of jumping numbers for the test ideal is discrete. In the case that X is smooth this was shown to be the case in [7, 8]. The singular case was obtained in [9]; see also [1, 27, 45, 77, 81, 92]. We will outline here an elementary proof based on the contracting property of p^{-e} linear maps that was investigated in the preceding section. In order to be able to handle not only a single p^{-e}-linear map but a whole Cartier algebra, we need to generalize the results on gauge bounds obtained above slightly. To keep things simple we will consider a Cartier algebra of the type

$$\mathscr{C} = \bigoplus R \cdot \varphi^n \mathfrak{a}^{\lceil t(p^{ne}-1)\rceil}$$

where $\mathfrak{a}$ is an ideal in R, $t \geq 0$ is a real number, and φ is a single p^{-e}-linear operator on R. This is essentially the case $\mathscr{C} = \mathscr{C}^{\Delta,\mathfrak{a}^t}$ for $(p^e-1)(K_R+\Delta)$ is a Cartier divisor (i.e., the pair (R,Δ) is $\mathbb{Q}$-Cartier with index not divisible by p). We first state a generalization of Lemma 9.2.3 to this context.

Lemma 9.3.14. *Let $\mathscr{C}$ be the Cartier subalgebra of $\mathscr{C}^R$ generated by φ, a p^{-e}-linear map on $R = k[x_1,\dots,x_n]/I$. Let M be a coherent $\mathscr{C}$ module and suppose that for all $m \in M$ and $n > 0$ one has*

$$\delta(\varphi^n(m)) \leq \frac{\delta(m)}{p^{ne}} + \frac{K}{p^e-1}$$

for some bound $K \geq 0$ as in Proposition 9.2.2. Then, if $\mathfrak{a} \subseteq R$ is an ideal generated by element of gauge $\leq d$, and $N \subseteq M$ is a R-submodule generated by elements of gauge $\leq D$, then $(\mathscr{C}_+^{\mathfrak{a}^t})^n(N) \subseteq N$ is generated by elements of gauge $\leq \frac{D}{p^{ne}} + \frac{K}{p^e-1} + td + 1$.

Proof. Note that R has a set of generators over $R^{p^{ne-1}}$ each of gauge $\leq p^{ne}-1$ (the images of the relevant monomials of $k[x_1,\dots,x_n]$ in R will do fine). Next, it is easy to check that $\mathfrak{a}^{\lceil t(p^{ne}-1)\rceil}$ is generated by element with gauge $\leq tdp^{ne}+1$. Hence, as in the proof of Lemma 9.2.3, the ideal $(\mathscr{C}_+^{\mathfrak{a}^t})^n$ is generated as a left R-modules by elements ψ of the form $\psi = \varphi^l \cdot b \cdot a$ where $l \geq n$ and b (resp. a) is one of the just described generators of R over $R^{p^{ln}}$ (resp. of $\mathfrak{a}^{\lceil t(p^{ne}-1)\rceil}$). Ranging over all such ψ and a set of R-generators m of N we see that $(\mathscr{C}_+^{\mathfrak{a}^t})^n(N)$ is generated by elements of the form $\psi(m)$. Now we just compute

$$\begin{aligned}\delta(\psi(m)) = \delta(\varphi^l \cdot b \cdot a \cdot m) &\leq \frac{\delta(bam)}{p^{le}} + \frac{K}{p^e-1} \leq \frac{\delta(m)}{p^{le}} \\ &\quad + \frac{(p^{le}-1)+(tdp^{le}+1)}{p^{le}} + \frac{K}{p^e-1} \\ &\leq \frac{\delta(m)}{p^{ne}-1} + \frac{K}{p^e-1} + td + 1.\end{aligned}$$

This shows the claim. ◻

Corollary 9.3.15. *With notation as in the lemma, let N be an R-submodule of M such that $\mathscr{C}_+^{\mathfrak{a}^t}(N) = N$. Then N is generated by elements of gauge $\leq \frac{K}{p^e-1} + td + 1$.*

For $T \geq 0$ there no infinite chains of R-submodules N of M for which $\mathscr{C}_+^{\mathfrak{a}^t}(N) = N$ for some $t < T$.

Proof. Clearly, $\mathscr{C}_+^{\mathfrak{a}^t}(N) = N$ implies that $(\mathscr{C}_+^{\mathfrak{a}^t})^n(N) = N$ for all n and hence the first claim follows from the preceding lemma. The second claim follows from the first one since each such N is generated by elements in the finite dimensional vector space $M_{\leq \frac{K}{p^e-1}+Td+1}$; hence there cannot be infinite chains. □

Now, the discreteness of the jumping numbers for the test ideal is an immediate consequence.

Theorem 9.3.16. *For $(R, \Delta, \mathfrak{a}^t)$ as above, the F-jumping numbers form a discrete subset of $\mathbb{Q}$.*

Proof. In the case that $(p^e - 1)(K_X + \Delta)$ is Cartier, the Cartier algebra $\mathscr{C}^\Delta$ is of the form considered above. Since each test ideal $\tau(R, \Delta, \mathfrak{a}^t)$ has the properties $\tau(R, \Delta, \mathfrak{a}^t) \supseteq \tau(R, \Delta, \mathfrak{a}^{t'})$ for $t' \geq t$ and $\mathscr{C}_+^{\Delta,\mathfrak{a}^t}\tau(R, \Delta, \mathfrak{a}^t) = \tau(R, \Delta, \mathfrak{a}^t)$, the preceding corollary shows that there is only finitely for t below a fixed bound T. Hence the jumping numbers must be discrete. The general case is similar or can be reduced to this case by using the methods of Sect. 7.3; see [80, 81]. □

9.4 Exercises

Exercise 1. Let $f \in S = k[x_1, \ldots, x_n]$ with k perfect and with gauge δ corresponding to the generator 1. Show that if $\delta(f) \leq d$ and writing uniquely

$$f = \sum_{x^i \in S_{p^e-1}} s_i^{p^e} x^i,$$

one has $\delta(s_i) \leq \lfloor d/p^e \rfloor$. (Here we used multi-exponent notation x^i as shorthand for $x_1^{i_1} \cdots x_n^{i_n}$.)

Exercise 2. Use the duality for finite morphisms to prove Lemma 9.1.1.

Exercise 3. Prove Lemma 9.2.1.

Exercise 4. Let $R \hookrightarrow S$ be a module-finite and flat ring extension. Show that the natural map

$$\mathrm{Hom}_R(S, R) \otimes_R M \xrightarrow{\varphi \otimes n \mapsto (r \mapsto \varphi(r)n)} \mathrm{Hom}_R(S, M)$$

is an isomorphism. Derive from this the statement of Lemma 9.1.5.

Exercise 5. Consider the example of a Cartier structure κ on the polynomial ring $k[x]$ given by sending $1 \mapsto x^t$ and $x, x^2, \ldots, x^{p-1} \mapsto 0$. Show that $\delta(\kappa(f)) \leq \delta(f)/p + t$ where δ is the gauge on $k[x]$ induced by the generator $1 \in k[x]$.

Exercise* 9.6. Prove Theorem 9.3.5 by using the same strategy as in Proposition 8.1.4.

Exercise 7. Let $(R, \mathfrak{m})$ be complete local of dimension d, and denote by $F\colon H^d_{\mathfrak{m}}(R) \to H^d_{\mathfrak{m}}(R)$ the natural Frobenius action. Show that any left action φ on $H^d_{\mathfrak{m}}(R)$ of the Frobenius is of the form $\varphi = r \cdot F$ for some $r \in R$.

Exercise 8. Prove Corollary 9.3.6.

Exercise 9. With notation as in Definition 9.3.8, show that $\mathscr{C}^{\Delta}$ and $\mathscr{C}^{\Delta,\mathfrak{a}^t}$ are Cartier subalgebras of $\mathscr{C}^R$. For a proof, see [76, Remark 3.10].

Exercise 10. Suppose that R is a normal local domain and that $\Delta \geq 0$ is a $\mathbb{Q}$-divisor on $X = \operatorname{Spec} R$ such that $K_X + \Delta$ is $\mathbb{Q}$-Cartier with index not divisible by $p > 0$. Prove that $\mathscr{C}^{\Delta}$ is a finitely generated ring over $\mathscr{C}^{\Delta}_0 = R$.

Hint: Show that $\operatorname{Hom}_R(F^e_* R(\lceil (p^e - 1)\Delta \rceil), R) \cong F^e_* R$ for some $e > 0$, and then use Exercise 2. For additional discussion see [76, Sect. 4].

Exercise 11. Suppose that R is a normal domain, $W \subseteq R$ is a multiplicative system, $\Delta \geq 0$ is a $\mathbb{Q}$-divisor on $X = \operatorname{Spec} R$, $\mathfrak{a} \subseteq R$ is a nonzero ideal, and $t \geq 0$ is a real number. Set $U = \operatorname{Spec}(W^{-1}R) \subseteq \operatorname{Spec} R = X$. Prove that

$$W^{-1}\tau(R; \Delta, \mathfrak{a}^t) = \tau(W^{-1}R; \Delta|_U, (W^{-1}\mathfrak{a})^t).$$

Exercise* 9.12. Prove Lemma 9.3.12.

Hint: Use the description of $\tau(R; \Delta, \mathfrak{a}^t)$ from the proof of Proposition 9.3.10. Also use the fact that R is Noetherian to see that the sum from Proposition 9.3.10 is a finite sum ($e = 0$ to m). Now notice that if c works in that sum, then so does bc where $0 \neq b \in \mathfrak{a}$. Set $\varepsilon = \frac{1}{p^m}$.

Exercise* 9.13. Suppose that R is a normal ring and that $X = \operatorname{Spec} R$. Consider the anticanonical ring

$$K := \bigoplus_{n \geq 0} \mathcal{O}_X(-nK_X).$$

Set $K_F := \bigoplus_{e \geq 0} \mathcal{O}_X((1 - p^e)K_X)$ to be the summand of K made up of terms of degree $p^e - 1$ for some $e \geq 0$. This is not a subring of K. However, define a noncommutative multiplication on K_F as follows. If $\alpha \in \mathcal{O}_X((1 - p^e)K_X)$ and $\beta \in \mathcal{O}_X((1 - p^d)K_X)$, then define $\alpha \star \beta = \alpha^{p^d}\beta \in \mathcal{O}_X(((1 - p^e)p^d + p^e)K_X) = \mathcal{O}_X((1 - p^{e+d})K_X)$.

With this ring operation, prove that K_F is isomorphic to $\mathscr{C}^R$.

Appendix A: Reflexification of Sheaves and Weil Divisors

In this section, we briefly recall basic properties of reflexive sheaves and Weil divisors on normal varieties. This material is all "well known," but there isn't a good source for it in the literature (we note that it is certainly assumed in [50]). We note that substantial generalizations of all this material (and complete proofs) can be found in [36]. As before, all schemes are of finite type over a field (or localizations or completions of such schemes). We assume the reader is familiar with the basic notion of depth and S_n (Serre's nth condition) and the connections with local cohomology / cohomology with support. See, for example, [34, Chap. III, Exercises in Sect. 3], [12] or [33].

A.1. Reflexive Sheaves

Given a coherent sheaf $\mathscr{F}$ on any scheme X, there is the following (dualizing) operation: $\mathscr{F}^\vee = \mathscr{H}om_{\mathcal{O}_X}(\mathscr{F}, \mathcal{O}_X)$. Furthermore, there is a natural map from $\mathscr{F}$ to the double-dual, $\mathscr{F} \to (\mathscr{F}^\vee)^\vee$.

Definition A.1.1. If this map is an isomorphism, we say that $\mathscr{F}$ *is reflexive* (or more specifically that it is $\mathcal{O}_X$-reflexive).

Note that if a sheaf is reflexive, it is also coherent (by definition). If $X = \operatorname{Spec} R$ and M is a coherent R-module, we say that M is reflexive if the corresponding sheaf is reflexive (equivalently, if $M \to \operatorname{Hom}_R(\operatorname{Hom}_R(M, R), R)$ is an isomorphism).

Notice first that any locally free sheaf is reflexive. But there are other reflexive sheaves as well. If one is careful, one can check that $\langle x, z \rangle \subseteq k[x, y, z]/(xy - z^2)$ corresponds to a reflexive ideal sheaf after taking Spec (Exercise A.1). There are a few basic facts about reflexive sheaves that should be mentioned. We now limit ourselves to varieties (i.e., integral schemes) which makes dealing with torsion much easier. One can do analogues of the following in more general situations (say for reduced schemes), but the statements become much more involved.

Lemma A.1.2. *Suppose that X is a variety and suppose that $\mathscr{F}$ is a coherent sheaf on X. Then $\mathscr{F}^\vee$ is torsion-free. (That is, if $U \subset X$ is open and $0 \neq r \in \mathcal{O}_X(U)$ and $0 \neq z \in \mathscr{F}^\vee(U)$, then $rz \neq 0$). In particular, a reflexive sheaf is torsion-free.*

Note that a torsion-free sheaf is necessarily S_1 (any nonzero element makes up a rather short regular sequence).

Lemma A.1.3. *Suppose that X is a variety and that $\mathscr{F}$ is a torsion-free coherent sheaf. Then the natural map $\alpha : \mathscr{F} \to \mathscr{F}^{\vee\vee}$ is injective.*

Lemma A.1.4 ([35, Proposition 1.1]). *A coherent sheaf $\mathscr{F}$ on a quasi-projective variety X is reflexive if and only if it can be included in an exact sequence*

$$0 \to \mathscr{F} \to \mathscr{E} \to \mathscr{G} \to 0$$

where $\mathscr{E}$ is locally free and $\mathscr{G}$ is torsion-free.

We note that the $\mathcal{O}_X$-dual of any coherent sheaf is always reflexive.

Theorem A.1.5. *If $\mathscr{F}$ is a coherent sheaf on a variety X, then $\mathscr{F}^\vee$ is reflexive. More generally, if $\mathscr{F}$ is coherent and $\mathscr{G}$ is reflexive, then $\mathscr{H}om_{\mathcal{O}_X}(\mathscr{F}, \mathscr{G})$ is reflexive.*

We now come to a very useful criterion for checking whether a sheaf is reflexive.

Theorem A.1.6 ([36, Theorem 1.9]). *Suppose that X is a normal (not necessarily quasi-projective) variety and that $\mathscr{F}$ is a coherent sheaf on X such that* $\mathrm{Supp}(\mathscr{F}) = X$. *Then $\mathscr{F}$ is* S_2 *if and only if $\mathscr{F}$ is reflexive.*

The key reason why the previous criterion is so useful is the Hartog's phenomenon associated with S_2 sheaves.

Corollary A.1.7. *Let X be an integral, normal (not necessarily quasi-projective) variety and suppose that $\mathscr{F}$ is a reflexive sheaf on X (defined as above). Let $Y \subset X$ be a closed subset of codimension ≥ 2 and set $U = X \setminus Y$. Then if $i : U \to X$ is the natural inclusion, then the natural map $\mathscr{F} \to i_* \mathscr{F}|_U$ is an isomorphism.*

Corollary A.1.8. *Suppose that $\mathscr{F}$ is a reflexive sheaf on $U \subseteq X$ (where X is as above) such that $X - U$ is codimension two. Let us denote by $i : U \to X$ the inclusion. Then $i_* \mathscr{F}$ is a reflexive sheaf on X.*

A.2. Divisors

Let X be a normal variety of finite type over a field. By a *Weil divisor* on X, we mean a formal sum of integral codimension 1 subschemes (prime divisors). Recall that a divisor D is called *effective* if the coefficients of D are nonnegative. Just like in the regular case, each prime divisor D corresponds to some discrete valuation v_D of the fraction field of X (although the reverse direction is not true).

Definition A.2.1. Choose $f \in \mathscr{K}(X)$, $f \neq 0$. We define the *principal divisor* $\mathrm{div}(f)$ as in the regular case: $\mathrm{div}(f) = \Sigma_i v_{D_i}(f) D_i$. Likewise, we say that two Weil divisors D_1 and D_2 are *linearly equivalent*, if $D_1 - D_2$ is principal.

Definition A.2.2. Given a divisor D, we define $\mathcal{O}_X(D)$ be the sheaf associated to the following rule:

$$\Gamma(V, \mathcal{O}_X(D)) = \{f \in \mathscr{K}(X) \mid \mathrm{div}(f)|_V + D|_V \geq 0\}.$$

A divisor D is called *Cartier* if $\mathcal{O}_X(D)$ is an invertible sheaf. It is called $\mathbb{Q}$-*Cartier* if nD is Cartier for some $n > 0$.

Note that D is effective if and only if $\mathcal{O}_X(D) \supseteq \mathcal{O}_X$.

Proposition A.2.3. *Suppose that D is a prime divisor, then $\mathcal{O}_X(-D) = \mathscr{I}_D$, the ideal sheaf defining D. Furthermore, if D is any divisor, then $\mathcal{O}_X(D)$ is reflexive.*

Proof. We first show the equality. The object defined above is clearly a sheaf. We will prove the equality of the sheaves in the setting where U is affine. Then $\Gamma(U, \mathcal{O}_X(D))$ is just the functions in $\mathcal{O}_X$ which vanish to order at least 1 along D, in other words the ideal of D.

We now want to show that this sheaf is reflexive (or equivalently, that it is S_2). First notice that clearly if U is the regular locus of X, then $\Gamma(V \cap U, \mathcal{O}_X(D)) \cong \Gamma(V, \mathcal{O}_X(D))$ for any open set V. This is because $V \cap U = U \setminus \{\text{non-regular locus}\}$, the non-regular locus, is codimension 2, and the sections of $\mathcal{O}_X(D)$ obviously do not change when removing a codimension 2 subset. This implies that the natural map $\mathcal{O}_X(D) \to i_* \mathcal{O}_X(D)|_U$ is an isomorphism, but then we notice that $\mathcal{O}_X(D)|_U$ is reflexive (since it is invertible), and thus, by Corollary A.1.8, $\mathcal{O}_X(D)$ is also reflexive. □

We now list some basic properties of rank-1 reflexive sheaves which completely link their behavior to divisors.

Proposition A.2.4. *Suppose that X is a normal variety. Then*

(a) If X is regular, then every reflexive rank-1 sheaf $\mathscr{F}$ on X is invertible [35, Proposition 1.9].
(b) Every rank one reflexive sheaf $\mathscr{F}$ on a normal scheme X embeds as a subsheaf of $\mathscr{K}(X)$.
(c) Any reflexive rank 1 subsheaf of $\mathscr{K}(X)$ is $\mathcal{O}_X(D)$ for some (uniquely determined) divisor D.

Proof. Left to the reader in Exercise A.4. □

The addition operations for divisors translates into the tensor of the associated sheaves, up to reflexification.

Proposition A.2.5. *Suppose that X is a normal variety and D and E are divisors on X. Then*

(a) If E is Cartier, then $\mathcal{O}_X(D) \otimes \mathcal{O}_X(E) \cong \mathcal{O}_X(D + E)$.
(b) In general, $\mathcal{O}_X(D + E) \cong (\mathcal{O}_X(D) \otimes \mathcal{O}_X(E))^{\vee\vee}$.
(c) $\mathcal{O}_X(-D) = \mathscr{H}om_{\mathcal{O}_X}(\mathcal{O}_X(D), \mathcal{O}_X) = \mathcal{O}_X(-D)^{\vee}$.

Proof. Left to the reader; see Exercise A.5 □

Finally, we mention a result relating sections and linearly equivalent divisors, which will be a key part of this chapter.

Theorem A.2.6. *Suppose that X is a normal variety and D is a Weil divisor on X. Then there is a bijection between the following two sets:*

$$\left\{ \begin{matrix} \text{Effective divisors } E \\ \text{linearly equivalent to } D \end{matrix} \right\} \longleftrightarrow \left\{ \begin{matrix} \text{Nonzero sections } \gamma \in H^0(X, \mathcal{O}_X(D)) \\ \text{modulo equivalence} \end{matrix} \right\}$$

where we define γ and γ' in $H^0(X, \mathcal{O}_X(D))$ to be equivalent if there exists a unit $u \in H^0(X, \mathcal{O}_X)$ such that $u\gamma = \gamma'$.

Proof. Set $\mathscr{M} = \mathcal{O}_X(D)$. The choice γ induces an embedding $i_\gamma : \mathscr{M} \hookrightarrow \mathscr{K}(X)$ which sends γ to 1. Thus γ induces a divisor via Proposition A.2.4. It follows from the same argument that γ and γ' induce the same divisor if and only if i_γ and $i_{\gamma'}$ have the same image in $\mathscr{K}(X)$. But this happens if and only if γ and γ' are unit multiplies of one another. □

A.3. Exercises

Exercise A.1. Show that $\langle x, z\rangle \in k[x, y, z]/\langle xy - z^2\rangle$ corresponds to a reflexive ideal sheaf after taking Spec.

Exercise A.2. Which of the following $k[x, y] = R$-modules are reflexive? If a module is not reflexive, compute its double dual $M^{\vee\vee}$.

(a) The ideal $\langle x\rangle$
(b) The ideal $\langle x, y\rangle$
(c) The module $R/\langle x, y\rangle$
(d) The module $R/\langle x\rangle$
(e) The ideal $\langle x^2, xy\rangle = \langle x, y\rangle^2 \cap \langle y\rangle$

Exercise A.3. Suppose that $\pi : Y \to X$ is a finite dominant map of normal varieties and $\mathscr{F}$ is a coherent sheaf on Y. Then $\mathscr{F}$ is reflexive on Y if and only if $\pi_*\mathscr{F}$ is reflexive on X.

Hint: Use the fact that you can check whether a sheaf is reflexive by checking whether it is S_2. Then use the criterion for checking depth via local cohomology.

Exercise A.4. Prove Proposition A.2.4.

Exercise A.5. Prove Proposition A.2.5.

Acknowledgements The authors are deeply indebted to Alberto Fernandez Boix, Lance Miller, Claudiu Raicu, Kevin Tucker, Wenliang Zhang, and the referee for innumerable valuable comments on previous drafts of this chapter.

The first author was partially supported by a Heisenberg Fellowship and the SFB/TRR45. The second author was partially supported by the NSF grant DMS #1064485 and a Sloan Research Fellowship.

References

1. Àlvarez Montaner, J., Boix, A., Zarzuela, S.: Frobenius and Cartier algebras of Stanley-Reisner rings. J. Algebra **358**(15), 162–177 (2012)
2. Anderson, G.W.: An elementary approach to L-functions mod p. J. Number Theory **80**(2), 291–303 (2000)

3. Blickle, M.: Minimal γ–sheaves. Algebra Number Theory **2**(3), 347–368 (2008)
4. Blickle, M.: Test ideals via algebras of p^{-e}-liner maps. J. Algebraic Geom. **22**, 49–83 (2013)
5. Blickle, M., Böckle, G.: Cartier crystals. manuscript in preparation, started 2006
6. Blickle, M., Böckle, G.: Cartier modules: finiteness results. J. Reine Angew. Math. **661**, 85–123 (2011). 2863904
7. Blickle, M., Mustaţă, M., Smith, K.: Discreteness and rationality of F-thresholds. Michigan Math. J. **57**, 43–61 (2008)
8. Blickle, M., Mustaţă, M., Smith, K.E.: F-thresholds of hypersurfaces. Trans. Amer. Math. Soc. **361**(12), 6549–6565 (2009) 2538604 (2011a:13006)
9. Blickle, M., Schwede, K., Takagi, S., Zhang, W.: Discreteness and rationality of F-jumping numbers on singular varieties. Math. Ann. **347**(4), 917–949 (2010) 2658149
10. Böckle, G., Pink, R.: Cohomological Theory of crystals over function fields. Tracts in Mathematics. European Mathematical Society, Zürich (2009)
11. Brion, M., Kumar, S.: Frobenius splitting methods in geometry and representation theory. Progress in Mathematics, vol. 231. Birkhäuser, Boston (2005). MR2107324 (2005k:14104)
12. Bruns, W., Herzog, J.: Cohen-Macaulay rings. Cambridge Studies in Advanced Mathematics, vol. 39. Cambridge University Press, Cambridge, 1993. MR1251956 (95h:13020)
13. Cartier, P.: Une nouvelle opération sur les formes différentielles. C. R. Acad. Sci. Paris **244**, 426–428 (1957) 0084497 (18,870b)
14. Conrad, B.: Grothendieck duality and base change. Lecture Notes in Mathematics, vol. 1750. Springer, Berlin (2000). MR1804902 (2002d:14025)
15. de Smit, B.: The different and differentials of local fields with imperfect residue fields. Proceedings of Edinburgh Mathematical Society (2) **40**(2), 353–365 (1997). MR1454030 (98d:11148)
16. Deligne, P., Illusie, L.: Relvements modulo p^2 et dcomposition du complexe de de rham. Inventiones Mathematica **89**, 247–270 (1987)
17. Emerton, M., Kisin, M.: Riemann–Hilbert correspondence for unit $\mathcal{F}$-crystals. Astérisque **293**, vi+257 pp (2004)
18. Enescu, F., Hochster, M.: The Frobenius structure of local cohomology. Algebra Number Theory **2**(7), 721–754 (2008). MR2460693 (2009i:13009)
19. Esnault, H., Viehweg, E.: Lectures on vanishing theorems. DMV Seminar, vol. 20. Birkhäuser, Basel (1992). MR1193913 (94a:14017)
20. Fedder, R.: F-purity and rational singularity. Trans. Amer. Math. Soc. **278**(2), 461–480 (1983) MR701505 (84h:13031)
21. Fujita, T.: Vanishing theorems for semipositive line bundles. Algebraic geometry (Tokyo/Kyoto, 1982), Lecture Notes in Mathematics, vol. 1016, pp. 519–528. Springer, Berlin (1983). 726440 (85g:14023)
22. Gabber, O.: Notes on some t-structures. Geometric Aspects of Dwork Theory, Vol. I, II, pp. 711–734. Walter de Gruyter GmbH & Co. KG, Berlin (2004)
23. Gabriel, P.: Des catégories abéliennes. Bull. Soc. Math. France **90**, 323–448 (1962). 0232821 (38 #1144)
24. Haboush, W.J.: A short proof of the Kempf vanishing theorem. Invent. Math. **56**(2), 109–112 (1980) 558862 (81g:14008)
25. Hacon, C.D.: On log canonical inversion of adjunction. arXiv:1202.0491, PEMS, volume in honour of V. Shokurov. To appear
26. Hara, N.: A characteristic p analog of multiplier ideals and applications. Comm. Algebra **33**(10), 3375–3388 (2005). MR2175438 (2006f:13006)
27. Hara, N.: F-pure thresholds and F-jumping exponents in dimension two. Math. Res. Lett. **13**(5–6), 747–760 (2006). With an appendix by Paul Monsky. MR2280772
28. Hara, N., Watanabe, K.-I.: F-regular and F-pure rings vs. log terminal and log canonical singularities. J. Algebraic Geom. **11**(2), 363–392 (2002). MR1874118 (2002k:13009)
29. Hara, N., Yoshida, K.-I.: A generalization of tight closure and multiplier ideals. Trans. Amer. Math. Soc. **355**(8), 3143–3174 (2003) (electronic). MR1974679 (2004i:13003)

30. Hara, N., Takagi, S.: On a generalization of test ideals. Nagoya Math. J. **175**, 59–74 (2004). MR2085311 (2005g:13009)
31. Hartshorne, R.: A property of A-sequences. Bull. Soc. Math. France **94**, 61–65 (1966) 0209279 (35 #181)
32. Hartshorne, R.: Residues and duality. Lecture notes of a seminar on the work of A. Grothendieck, given at Harvard 1963/64. With an appendix by P. Deligne. Lecture Notes in Mathematics, vol. 20, Springer, Berlin (1966) MR0222093 (36 #5145)
33. Hartshorne, R.: Local cohomology. A seminar given by A. Grothendieck, Harvard University, Fall, vol. 1961. Springer, Berlin (1967). MR0224620 (37 #219)
34. Hartshorne, R.: Algebraic geometry, Graduate Texts in Mathematics, vol. 52. Springer, New York (1977). MR0463157 (57 #3116)
35. Hartshorne, R.: Stable reflexive sheaves. Math. Ann. **254**(2), 121–176, (1980). 597077 (82b:14011)
36. Hartshorne, R.: Generalized divisors on Gorenstein schemes. Proceedings of Conference on Algebraic Geometry and Ring Theory in Honor of Michael Artin, Part III (Antwerp, 1992), vol. 8, pp. 287–339 (1994). MR1291023 (95k:14008)
37. Hartshorne, R., Speiser, R.: Local cohomological dimension in characteristic p. Ann. Math. (2) **105**(1), 45–79 (1977) MR0441962 (56 #353)
38. Hochster, M.: Foundations of tight closure theory. Lecture Notes from a Course Taught on the University of Michigan Fall 2007 (2007)
39. Hochster, M.: Some finiteness properties of Lyubeznik's $\mathscr{F}$-modules. Algebra, Geometry and Their Interactions. Contemporary Mathematics, vol. 448, pp. 119–127. American Mathematical Society, Providence (2007) 2389238 (2009c:13010)
40. Hochster, M., Roberts, J.L.: The purity of the Frobenius and local cohomology. Adv. Math. **21**(2), 117–172 (1976) MR0417172 (54 #5230)
41. Hochster, M., Huneke, C.: Tight closure, invariant theory, and the Briançon-Skoda theorem. J. Amer. Math. Soc. **3**(1), 31–116 (1990) MR1017784 (91g:13010)
42. Huneke, C.: Tight closure and its applications. CBMS Regional Conference Series in Mathematics, vol. 88. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1996, With an appendix by Melvin Hochster. MR1377268 (96m:13001)
43. Huneke, C., Swanson, I.: Integral closure of ideals, rings, and modules. In: London Mathematical Society Lecture Note Series, vol. 336. Cambridge University Press, Cambridge (2006) MR2266432 (2008m:13013)
44. Katz, N.M.: Nilpotent connections and the monodromy theorem: Applications of a result of Turrittin, Inst. Hautes Études Sci. Publ. Math. (1970), (39), 175–232. 0291177 (45 #271)
45. Katzman, M., Lyubeznik, G., Zhang, W.: On the discreteness and rationality of F-jumping coefficients. J. Algebra **322**(9), 3238–3247 (2009). 2567418 (2011c:13005)
46. Kawakita, M.: Inversion of adjunction on log canonicity. Invent. Math. **167**(1), 129–133 (2007) MR2264806 (2008a:14025)
47. Kawamata, Y.: A generalization of Kodaira-Ramanujam's vanishing theorem. Math. Ann. **261**(1), 43–46 (1982). MR675204 (84i:14022)
48. Kawamata, Y.: Subadjunction of log canonical divisors. II. Amer. J. Math. **120**(5), 893–899 (1998). MR1646046 (2000d:14020)
49. Keeler, D.S.: Fujita's conjecture and Frobenius amplitude. Amer. J. Math. **130**(5), 1327–1336 (2008) 2450210 (2009i:14006)
50. Kollár, J., Mori, S.: Birational geometry of algebraic varieties. Cambridge Tracts in Mathematics, vol. 134, Cambridge University Press, Cambridge (1998), With the collaboration of C.H. Clemens and A. Corti. Translated from the 1998 Japanese original. MR1658959 (2000b:14018)
51. Kollár, J., and 14 coauthors: Flips and abundance for algebraic threefolds. Société Mathématique de France, Paris, 1992. Papers from the Second Summer Seminar on Algebraic Geometry held at the University of Utah, Salt Lake City, Utah, August 1991, Astérisque No. 211 (1992). MR1225842 (94f:14013)
52. Kumar, S., Mehta, V.B.: Finiteness of the number of compatibly split subvarieties. Int. Math. Res. Not. IMRN **2009**(19), 3595–3597 (2009). 2539185 (2010j:13012)

53. Kunz, E.: Characterizations of regular local rings for characteristic p. Amer. J. Math. **91**, 772–784 (1969). MR0252389 (40 #5609)
54. Kunz, E.: Kähler differentials. Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig (1986). MR864975 (88e:14025)
55. Lazarsfeld, R.: Positivity in algebraic geometry. I. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 48. Springer, Berlin (2004). Classical setting: line bundles and linear series. MR2095471 (2005k:14001a)
56. Lazarsfeld, R.: Positivity in algebraic geometry. II. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 49. Springer, Berlin (2004). Positivity for vector bundles, and multiplier ideals. MR2095472 (2005k:14001b)
57. Lech, C.: Inequalities related to certain couples of local rings. Acta Math. **112**, 69–89 (1964). 0161876 (28 #5080)
58. Lipman, J.: Notes on derived functors and Grothendieck duality, Foundations of Grothendieck duality and diagrams of schemes, Lecture Notes in Math., Springer Berlin (2009)
59. Lyubeznik, G.: F-modules: applications to local cohomology and D-modules in characteristic $p > 0$. J. Reine Angew. Math. **491**, 65–130 (1997). MR1476089 (99c:13005)
60. Lyubeznik, G., Smith, K.E.: Strong and weak F-regularity are equivalent for graded rings. Amer. J. Math. **121**(6), 1279–1290 (1999) MR1719806 (2000m:13006)
61. Lyubeznik, G., Smith, K.E.: On the commutation of the test ideal with localization and completion. Trans. Amer. Math. Soc. **353**(8), 3149–3180 (2001). (electronic). MR1828602 (2002f:13010)
62. Ma, L.: Finiteness property of local cohomology for F-pure local rings. arXiv:1204.1539 (2012).
63. Majadas, J., Rodicio, A.G.: Smoothness, regularity and complete intersection. In: London Mathematical Society Lecture Note Series, vol. 373. Cambridge University Press, Cambridge (2010). 2640631 (2011m:13028)
64. Matsumura, H.: Commutative algebra. In: Mathematics Lecture Note Series, vol. 56, 2nd edn. Benjamin/Cummings Publishing, Reading (1980). MR575344 (82i:13003)
65. Matsumura, H.: Commutative ring theory. Cambridge Studies in Advanced Mathematics, vol. 8, 2nd edn. Cambridge University Press, Cambridge (1989). Translated from the Japanese by M. Reid. MR1011461 (90i:13001)
66. Mehta, V.B., Ramanathan, A.: Frobenius splitting and cohomology vanishing for Schubert varieties. Ann. Math. (2) **122**(1), 27–40 (1985). MR799251 (86k:14038)
67. Miyachi, J.-I.: Localization of triangulated categories and derived categories. J. Algebra **141**(2), 463–483 (1991) 1125707 (93b:18016)
68. Moriya, M.: Theorie der Derivationen und Körperdifferenten. Math. J. Okayama Univ. **2**, 111–148 (1953) MR0054643 (14,952f)
69. Mustaţă, M., Takagi, S., Watanabe, K.-I.: F-thresholds and Bernstein-Sato polynomials. European Congress of Mathematics, European Mathematical Society, Zürich, pp. 341–364 (2005). MR2185754 (2007b:13010)
70. Peskine, C., Szpiro, L.: Dimension projective finie et cohomologie locale. Applications à la démonstration de conjectures de M. Auslander, H. Bass et A. Grothendieck, Inst. Hautes Études Sci. Publ. Math. **42**, 47–119 (1973). MR0374130 (51 #10330)
71. Ramanan, S., Ramanathan, A.: Projective normality of flag varieties and Schubert varieties. Invent. Math. **79**(2), 217–224 (1985). MR778124 (86j:14051)
72. Scheja, G., Storch, U.: Über Spurfunktionen bei vollständigen Durchschnitten. J. Reine Angew. Math. **278/279**, 174–190 (1975). MR0393056 (52 #13867)
73. Schwede, K.: F-adjunction. Algebra Number Theory **3**(8), 907–950 (2009)
74. Schwede, K.: Centers of F-purity. Math. Z. **265**(3), 687–714 (2010). 2644316
75. Schwede, K.: A canonical linear system associated to adjoint divisors in characteristic $p > 0$. J. Reine Angew. Math. arXiv:1107.3833

76. Schwede, K.: Test ideals in non-$\mathbb{Q}$-Gorenstein rings. Trans. Amer. Math. Soc. **363**(11), 5925–5941 (2011) 2817415 (2012c:13011)
77. Schwede, K., Takagi, S.: Rational singularities associated to pairs. Michigan Math. J. **57**, 625–658 (2008)
78. Schwede, K., Smith, K.E.: Globally F-regular and log Fano varieties. Adv. Math. **224**(3), 863–894 (2010) 2628797 (2011e:14076)
79. Schwede, K., Tucker, K.: On the number of compatibly Frobenius split subvarieties, prime F-ideals, and log canonical centers. Ann. Inst. Fourier (Grenoble) **60**(5), 1515–1531 (2010). 2766221 (2012d:13007)
80. Schwede, K., Tucker, K.: On the behavior of test ideals under finite morphisms. arXiv:1003.4333, J. Algebraic Geom. To appear
81. Schwede, K., Tucker, K., Zhang, W.: Test ideals via a single alteration and discreteness and rationality of f-jumping numbers. Math. Res. Lett. **19**(01), 191–197 (2012)
82. Sharp, R.Y.: Tight closure test exponents for certain parameter ideals. Michigan Math. J. **54**(2), 307–317 (2006) 2252761 (2007e:13008)
83. Sharp, R.Y.: Graded annihilators of modules over the Frobenius skew polynomial ring, and tight closure. Trans. Amer. Math. Soc. **359**(9), 4237–4258 (2007) (electronic). MR2309183 (2008b:13006)
84. Sharp, R.Y.: On the Hartshorne–Speiser–Lyubeznik theorem about Artinian modules with a Frobenius action. Proc. Amer. Math. Soc. **135**(3), 665–670 (2007) (electronic). 2262861 (2007f:13008)
85. Sharp, R.Y., Yoshino, Y.: Right and left modules over the Frobenius skew polynomial ring in the F-finite case. Math. Proc. Cambridge Philos. Soc. **150**(3), 419–438 (2011). 2784768 (2012e:13009)
86. Shokurov, V.V.: Three-dimensional log perestroikas. Izv. Ross. Akad. Nauk Ser. Mat. **56**(1), 105–203 (1992). MR1162635 (93j:14012)
87. Silverman, J.H.: The arithmetic of elliptic curves. In: Graduate Texts in Mathematics, vol. 106, 2nd edn. Springer, Dordrecht (2009). 2514094 (2010i:11005)
88. Smith, K.E.: Fujita's freeness conjecture in terms of local cohomology. J. Algebraic Geom. **6**(3), 417–429 (1997). MR1487221 (98m:14002)
89. Smith, K.E.: Globally F-regular varieties: applications to vanishing theorems for quotients of Fano varieties. Michigan Math. J. **48**, 553–572 (2000). Dedicated to William Fulton on the occasion of his 60th birthday. MR1786505 (2001k:13007)
90. Takagi, S.: An interpretation of multiplier ideals via tight closure. J. Algebraic Geom. **13**(2), 393–415 (2004). MR2047704 (2005c:13002)
91. Takagi, S., Watanabe, K.-I.: On F-pure thresholds. J. Algebra **282**(1), 278–297 (2004). MR2097584 (2006a:13010)
92. Takagi, S., Takahashi, R.: D-modules over rings with finite F-representation type. Math. Res. Lett. **15**(3), 563–581 (2008). MR2407232 (2009e:13003)
93. Tate, J.: Residues of differentials on curves. Ann. Sci. École Norm. Sup. (4) **1**, 149–159 (1968). 0227171 (37 #2756)
94. Viehweg, E.: Vanishing theorems. J. Reine Angew. Math. **335**, 1–8 (1982). MR667459 (83m:14011)
95. Vitulli, M.A.: Weak normality and seminormality. Commutative Algebra—Noetherian and non-Noetherian Perspectives, pp. 441–480. Springer, New York (2011). 2762521 (2012d:13040)
96. Yanagihara, H.: On an intrinsic definition of weakly normal rings. Kobe J. Math. **2**(1), 89–98 (1985). MR811809 (87d:13007)

Castelnuovo–Mumford Regularity of Annihilators, Ext and Tor Modules

Markus Brodmann, Cao Huy Linh, and Maria-Helena Seiler

1 Introduction

In his opening address to the *Workshop on Castelnuovo–Mumford Regularity and Applications* at the Max Planck Institute for Mathematics in the Sciences at Leipzig in June 2007, we learned from Professor Eberhard Zeidler, former Director of that Institute, that physicists have a high esteem for algebraic geometry, because it provides so many invariants. Among these invariants Castelnuovo–Mumford regularity is particularly interesting. For example, mathematical physics are interested in degrees of defining equations of characteristic varieties of D-modules, a subject which is closely related to Castelnuovo–Mumford regularity. So, in the PhD thesis [2] of Michael Bächtold, we find the result that the Hilbert function (with respect to an appropriate filtration) of a D-module W over a standard Weyl algebra A bounds from above the degrees of polynomials which are needed to cut out set theoretically the characteristic variety of W. This is true, because the Hilbert function h_M of a graded module M which is generated over the polynomial ring $R = K[x_1, \ldots, x_r]$ by finitely many elements of degree 0 bounds from above the Castelnuovo–Mumford regularity $\text{reg}(\text{Ann}_R(M))$ of the annihilator $\text{Ann}_R(M)$ of M. This was worked out in the MSc thesis [21] of the third author. Let us also mention here the recently finished PhD thesis [5] of Roberto Boldini, which is devoted to a different aspect of characteristic varieties of D-modules.

M. Brodmann (✉) • M.-H. Seiler
Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190,
CH-8057 Zürich, Switzerland
e-mail: brodmann@math.uzh.ch; mariahelena_seiler@hotmail.com

C.H. Linh
Department of Mathematics, College of Education, Hue University,
32 Le Loi, Hue City, Viet Nam
e-mail: huylinh2002@yahoo.com

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_6,

Later it turned out that the ideas used in [21] may be combined with some earlier bounding results of [8] to get a number of a priori bounds for the Castelnuovo–Mumford regularity of Ext- and Tor-modules, for example, bounds which hold over arbitrary Noetherian homogeneous rings with local Artinian base ring and for arbitrary finitely generated graded modules over them. We do not insist that one has to use exclusively the results of [8]. Indeed, instead one also could use, for example, results of Chardin–Fall–Nagel [13] to end up with similar bounds.

Our original question asks whether a certain finite collection of invariants of a finitely generated graded module M over a homogeneous Noetherian ring R bounds the Castelnuovo–Mumford regularity of M. This leads to ask for a priori bounds, hence for bounds which apply in a most general setting. Here, being bounded in terms of certain invariants usually is more interesting than the size of the bound. On the other hand, one also can ask for *specific bounds*, for example, bounds which apply only for a specified class of graded R-modules, but which in turn are smaller (and possibly sharp). Already at its beginning, the investigation of Castelnuovo–Mumford regularity shows an interplay of these two aspects (see, e.g., [3, 4, 6, 10, 12, 16, 20]). In the this chapter, clearly the first aspect plays a dominant role. Nevertheless, in the last section, we shall give a bound on the Castelnuovo–Mumford regularity of certain specified Tor-modules which extends earlier bounding results of Eisenbud–Huneke–Ulrich [15] and Caviglia [11].

In Sect. 2 of this chapter, we present some preliminaries, and we give an extension to graded modules of Mumford's basic bounding result for graded ideals in a polynomial ring [20] in terms of Hilbert polynomials—an extension which to some extend may be viewed as folklore. It says that over a Noetherian homogeneous (e.g., standard graded) ring R with local Artinian base ring R_0, the Castelnuovo–Mumford regularity of a finitely generated graded R-module M is bounded in terms of the length of R_0, the degree vector of a homogeneous system of generators of M, the Hilbert polynomial p_M, and the postulation number $p(M)$, of M (see Proposition 5). In view of our first goal, which is to bound the Castelnuovo–Mumford regularity of the annihilator $\mathrm{Ann}_R(M)$ of a finitely generated graded R-module M in terms of the Hilbert function h_M of M, we clearly have to use this result.

In Sect. 3, we give a few preliminaries on filtered modules over filtered rings, especially on D-modules, and introduce in more detail the original question asked by Bächtold on the degrees of equations cutting out set theoretically the characteristic variety of such modules. As this chapter has an expository touch, we allow ourselves to include here a short introduction to characteristic varieties of modules over appropriately filtered K-algebras, especially over Weyl algebras. Readers familiar with the subject therefore might jump what is said in Reminder 1 to Remark 5. For readers who aim to learn more about the subject, we recommend to consult [18, 19], or [14]. After this expository introduction, we tie the link to the Castelnuovo–Mumford regularity of annihilators of graded modules and prove the requested bounding results on their Castelnuovo–Mumford regularity in terms of Hilbert functions (see Theorem 14 and Corollaries 15 and 16). We shall do this by first proving that the Castelnuovo–Mumford regularity of the annihilator $\mathrm{Ann}_R(M)$

of a finitely generated graded R-module M is bounded in terms of invariants of R, the initial degree and the Castelnuovo–Mumford regularity of M (see Theorem 10 and Corollaries 11–13). Then we apply Proposition 5 to get the requested bound for $\mathrm{reg}(\mathrm{Ann}_R(M))$ in terms of the Hilbert function h_M of M.

In Sect. 4 we give an a priori bound for the Castelnuovo–Mumford regularity of the modules $\mathrm{Ext}^i_R(M, N)$ in terms of i, of invariants of R, the initial degrees, the Castelnuovo–Mumford regularities, and the number of generators of M and N, where M and N are finitely generated graded modules over a Noetherian homogeneous ring R with local Artinian base ring R_0 (see Theorem 4 and Corollary 5). As an application we get a simply shaped bound for the Castelnuovo–Mumford regularity of the deficiency modules $K^i(M)$ in terms of i, invariants of R, the initial degree, the Castelnuovo–Mumford regularity and the number of generators of M (see Corollaries 7 and 8).

In Sect. 5—under the same hypothesis as in Sect. 4—we first give a bound for the Castelnuovo–Mumford regularity of the tensor product $M \otimes_R N$ in terms of the invariants mentioned above (see Proposition 3). We then deduce a corresponding a priori bound for the Castelnuovo–Mumford regularity of the modules $\mathrm{Tor}_i^R(M, N)$ (see Theorem 4, Corollary 5, and Remark 6). Then, we leave the field of a priori bounds and establish—over arbitrary rings R as above—an upper bound on the Castelnuovo–Mumford regularity of the modules $\mathrm{Tor}_k^R(M, N)$, provided that at least one of the two modules M or N has finite projective dimension and that $\mathrm{Tor}_i^R(M, N)$ is if dimension ≤ 1 for all $i > 0$ (see Proposition 8). As an application we prove a bounding result for the Castelnuovo–Mumford regularity of the modules $\mathrm{Tor}_k^R(M, N)$ which holds under the hypotheses that $\mathrm{Tor}_1^R(M, N)$ is of dimension ≤ 1 and the singular locus of the scheme $\mathrm{Proj}(R)$ is finite. This will extend the previously mentioned results of Eisenbud–Huneke–Ulrich and Caviglia (see Theorem 10 and Corollary 11).

2 Some Preliminaries

In this section, we fix a few notations and recall some basic facts which we shall use throughout this chapter. For the reader's convenience we also present and prove a result of folklore type which extends Mumford's basic regularity bound [20]. As a basic reference for this section we use [7].

Notation 1. Let $\mathbb{N}_0$ denote the set of nonnegative integers and let $\mathbb{N}$ denote the set of positive integers.

Throughout let $R = \bigoplus_{n \in \mathbb{N}_0} R_n$ be a Noetherian homogeneous ring with Artinian local base ring $(R_0, \mathfrak{m}_0)$ and irrelevant ideal $R_+ := \bigoplus_{n \in \mathbb{N}} R_n$. Observe in particular that there are finitely many elements $l_1, l_2, \ldots, l_r \in R_1$ such that $R = R_0[l_1, l_2, \ldots, l_r]$, $R_+ = \langle l_1, l_2, \ldots, l_2 \rangle$ and $\mathfrak{m} := \mathfrak{m}_0 \oplus R_+$ is the unique homogeneous maximal ideal of R.

Next we recall a few basic facts on local cohomology of graded R-modules and Castelnuovo–Mumford regularity.

Reminder 2. If $T = \bigoplus_{n\in\mathbb{Z}} T_n$ is a graded R-module we define the *beginning* (or the *initial degree*) and the *end* of T, respectively, by

$$\operatorname{beg}(T) := \inf\{n \in \mathbb{Z} \mid T_n \neq 0\}, \quad \operatorname{end}(T) := \sup\{n \in \mathbb{Z} \mid T_n \neq 0\}.$$

Moreover, the *generating degree* of the graded R-module T is defined by

$$\operatorname{gendeg}(T) := \inf\{n \in \mathbb{Z} \mid T = \sum_{m\leq n} RT_m\}.$$

We always use the convention that $\inf(\bullet)$ and $\sup(\bullet)$ are formed in $\mathbb{Z} \cup \{-\infty, +\infty\}$ with $\inf(\varnothing) := \infty$ and $\sup(\varnothing) := -\infty$. Obviously, we have

$$T \neq 0 \Rightarrow \operatorname{beg}(T) \leq \operatorname{gendeg}(T) \leq \operatorname{end}(T).$$

If the R-module T is finitely generated, we have $\operatorname{gendeg}(T) \leq \infty$.

For each nonnegative integer $i \in \mathbb{N}_0$ and each graded R-module $M = \bigoplus_{n\in\mathbb{Z}} M_n$ let $H^i_{R_+}(M)$ denote the ith local cohomology module of M with respect to the irrelevant ideal R_+ of R. The R-modules $H^i_{R_+}(M) = \bigoplus_{n\in\mathbb{Z}} H^i_{R_+}(M)_n$ carry a natural grading, the graded R-modules $H^i_{R_+}(M)$ are Artinian, and so their graded parts $H^i_{R_+}(M)_n$ are R_0-modules of finite length in all degrees $n \in \mathbb{Z}$ and vanish for all $n \gg 0$. Moreover, if $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1)$ denotes the minimal number of generators of the R_0 module R_1, we have $H^i_{R_+}(M) = 0$ for all $i > r$.

Let M be a finitely generated graded R-module and let $k \in \mathbb{N}_0$. The *(Castelnuovo–Mumford) regularity of M at and above level k* is defined by

$$\operatorname{reg}^k(M) := \sup\{\operatorname{end}(H^i_{R_+}(M)) + i \mid i \geq k\}.$$

Observe that $\operatorname{reg}^k(M) < \infty$. The *(Castelnuovo–Mumford) regularity of M* at all is defined as the Castelnuovo–Mumford regularity of M at and above level 0, thus by

$$\operatorname{reg}(M) := \operatorname{reg}^0(M) = \sup\{\operatorname{end}(H^i_{R_+}(M)) + i \mid i \in \mathbb{N}_0\}.$$

We always have the inequality

$$\operatorname{gendeg}(M) \leq \operatorname{reg}(M).$$

We constantly use without further mention the behavior of regularities in short exact sequences of finitely generated graded R-modules and the fact that regularities are not affected if one considers M as a graded S-module by means of a surjective homomorphism of homogeneous Noetherian rings $\phi : S \twoheadrightarrow R$.

For simplicity, we also introduce the *width* of the finitely generated graded R-module M. e.g. the span between the regularity and the initial degree of M:

$$\mathrm{w}(M) := \max\{0, \mathrm{reg}(M) - \mathrm{beg}(M) + 1\}.$$

Observe that $\mathrm{w}(M) > 0$ if and only if $M \neq 0$, whereas $\mathrm{w}(M) = 0$ means that $M = 0$. When $R = K[x_1, x_2, \ldots, x_r]$ is a polynomial ring over a field K, the width of M is precisely the number of rows in the Betti-diagram of M, so

$$\mathrm{w}(M) = \sup\{\mathrm{end}(\mathrm{Tor}^i_R(R/R_+, M)) - \mathrm{beg}(M) - i \mid i \in \mathbb{N}_0\}.$$

We recall a few basic facts on Hilbert polynomials of graded R-modules.

Reminder 3. Let $M = \bigoplus_{n\in\mathbb{Z}} M_n$ be a finitely generated graded R-module. We denote the *Hilbert polynomial* of M by p_M so that

$$\mathrm{length}_{R_0}(M_n) = p_M(n) \text{ for all } n \gg 0.$$

We also introduce the *postulation number* of M that is the invariant

$$p(M) := \sup\{n \in \mathbb{Z} \mid \mathrm{length}_{R_0}(M_n) \neq p_M(n)\} \quad \in \mathbb{Z} \cup \{-\infty\}.$$

Hilbert polynomials behave additively in short exact sequences of finitely generated graded R-modules. Moreover, the Hilbert polynomial and the postulation number of a finitely generated graded R-module are not affected if one considers M as a graded S-module by means of a surjective homomorphism $S \twoheadrightarrow R$ of Noetherian homogeneous R_0-algebras.

For each $i \in \mathbb{N}_0$ and all $n \in \mathbb{Z}$ we may consider the nonnegative integer

$$h^i_M(n) := \mathrm{length}_{R_0}(H^i_{R_+}(M)_n),$$

which vanishes for all $n \gg 0$ and for all $i > \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1)$. *Serre's formula* yields (see [7, 17.1.6])

$$p_M(n) = \mathrm{length}_{R_0}(M_n) - \sum_{i\in\mathbb{N}_0}(-1)^i h^i_M(n) \text{ for all } n \in \mathbb{Z}.$$

One obvious consequence of this formula is the estimate

$$\mathrm{reg}(M) \leq \max\{\mathrm{reg}^1(M), p(M) + 1\}.$$

Next, we quote the following auxiliary result, which will play a crucial role in our later arguments

Lemma 4. *Assume that R is a Cohen–Macaulay ring of dimension $r > 0$ and multiplicity e. Let $f : W \longrightarrow V$ be a homomorphism of finitely generated graded R-modules. If $V \neq 0$ is generated by μ homogeneous elements and*

$$\alpha := \min\{\text{beg}(V), \text{reg}(V) - \text{reg}(R)\},$$

then we have

$$\text{reg}(\text{Im}(f)) \leq \left[\max\{\text{gendeg}(W), \text{reg}(V) + 1\} + e(\mu + 1) - \alpha\right]^{2^{r-1}} + \alpha.$$

Proof. This is nothing else than Corollary 6.2 of [8]. □

Finally, we give the announced extension of Mumford's regularity bound. It says that the regularity of a finitely generated graded R-module M is bounded in terms of the length of the base ring R_0, the Hilbert polynomial p_M, the postulation number $p(M)$ and the degrees $a_1, a_2, \ldots, a_\mu$ of generators of M.

Proposition 5. *Let $p \in \mathbb{Q}[x]$ be a polynomial, let $\mu \in \mathbb{N}$, and let $\boldsymbol{a} := (a_1, a_2, \ldots, a_\mu) \in \mathbb{Z}^\mu$ with $a_1 \leq a_2, \leq \cdots \leq a_\mu$. Then, there is a function*

$$F_{p,\boldsymbol{a}} : \mathbb{N}^2 \times \mathbb{Z} \longrightarrow \mathbb{Z}$$

such that whenever $\lambda :=$ length(R_0), $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1)$ and M, is a finitely generated graded R-module such that $p_M = p$, $p(M) \leq \pi$, and $M = \sum_{i=1}^{\mu} Rm_i$ with $m_i \in M_{a_i}$ for $i = 1, 2, \ldots, \mu$, we have

$$\text{reg}(M) \leq F_{p,\boldsymbol{a}}(\lambda, r, \pi).$$

Proof. Let $r, \mu \in \mathbb{N}$, let $\boldsymbol{a} := (a_1, a_2, \ldots, a_\mu) \in \mathbb{Z}^\mu$ with $a_1 \leq a_2 \leq \cdots \leq a_\mu$, and let $p \in \mathbb{Q}[x]$. According to Theorem 17.2.7 of [7], there is a function

$$G_{r,p,\boldsymbol{a}} : \mathbb{N} \longrightarrow \mathbb{Z}$$

such that whenever $S = R_0[x_1, x_2, \ldots, x_r]$ is a polynomial with Artinian local base ring R_0 ring with length$(R_0) = \lambda$ and $N \subset \bigoplus_{i=1}^{\mu} S(-a_i) =: U$ is a graded submodule such that the graded S-module $M := U/N$ satisfies $p_M = p$, we have

$$\text{reg}^2(N) \leq G_{r,p,\boldsymbol{a}}(\lambda).$$

In view of the short exact sequence of graded S-modules

$$0 \longrightarrow N \longrightarrow U \longrightarrow M \longrightarrow 0,$$

as $\text{reg}(U) = a_\mu$ and by Reminder 3 we get

$$\text{reg}(M) \leq \max\{a_\mu, G_{r,p,\boldsymbol{a}}(\lambda), p(M) + 1\}.$$

Define

$$F_{p,\boldsymbol{a}}(\lambda, r, \pi) := \max\{a_\mu, G_{r,p,\boldsymbol{a}}(\lambda).\pi + 1\}.$$

If R and M satisfy the requirements of our proposition, then there is a surjective homomorphism of homogeneous Noetherian R_0-algebras $\phi : S \twoheadrightarrow R$ and we may consider M as an S-module by means of ϕ. In particular, we may write $M = U/N$ for some graded submodule $N \subset U$. As $\operatorname{reg}(M)$, p_M and $p(M)$, are not affected if we consider M as an S-module we get the requested inequality. □

3 Characteristic Varieties of *D*-Modules and the Regularity of Annihilators

As mentioned in the introduction, this chapter grew out of a problem concerning characteristic varieties of D-modules. In this section, we aim to introduce this problem in more detail and present its solution, which bases on a bound for the regularity of the annihilator of a finitely generated graded module over a polynomial ring over a field. We first recall a few elementary facts on Weyl algebras and D-modules. Our suggested reference for this is [14], although we partly use our own terminology. We start in a slightly more general setting, for which we recommend the references [18] and [19].

Reminder 1. Let K be a field and let A be a unital associative K-algebra which carries a *filtration* $A_\bullet = (A_i)_{i\in\mathbb{N}_0}$ so that each A_i is a K-subspace of A such that

$$A_i \subseteq A_{i+1} \text{ for all } i \in \mathbb{N}_0, \quad 1 \in A_0, \quad A = \bigcup_{i\in\mathbb{N}_0} A_i \quad \text{and}$$

$$A_i A_j \subseteq A_{i+j} \text{ for all } i, j \in \mathbb{N}_0,$$

where by definition $A_i A_j := \sum_{(f,g)\in A_i\times A_j} Kfg$. To simplify notation, we set $A_i = 0$ for all $i < 0$. The *associated graded* ring of A with respect to the filtration $A_\bullet$ is defined as the graded K-algebra

$$\operatorname{Gr}(A) = \operatorname{Gr}_{A_\bullet}(A) = \bigoplus_{i\in\mathbb{N}_0} A_i/A_{i-1},$$

with multiplication induced by $(f + A_{i-1})(g + A_{j-1}) := fg + A_{i+j-1}$ for all $i, j \in \mathbb{N}_0$, all $f \in A_i$ and all $g \in A_j$. The filtration $A_\bullet$ is said to be *commutative* if

$$fg - gf \in A_{i+j-1} \text{ for all } i, j \in \mathbb{N}_0 \text{ and for all } f \in A_i \text{ and all } g \in A_j.$$

In this situation, the associated graded ring $\operatorname{Gr}(A)$ is commutative. The filtration $A_\bullet$ is said to be *very good* if is commutative and

$$A_0 = K, \quad \dim_K(A_1) < \infty, \quad \text{ and } A_i = A_1 A_{i-1} \text{ for all } i \in \mathbb{N}.$$

Clearly in this situation, the associated graded ring is a commutative homogeneous Noetherian K-algebra. If $A_\bullet$ is a very good filtration of A, we say that $(A, A_\bullet)$—or briefly A—is a *very well-filtered* K-algebra.

Let W be left A-module which carries an $A_\bullet$*-filtration* $W_\bullet = (W_i)_{i\in\mathbb{Z}}$ so that each W_i is a K-subspace of W and moreover

$$W_i \subseteq W_{i+1} \text{ for all } i \in \mathbb{Z}, \quad W = \bigcup_{i\in\mathbb{Z}} W_i \quad \text{and}$$

$$A_i W_j \subseteq W_{i+j} \text{ for all } (i, j) \in \mathbb{N}_0 \times \mathbb{Z},$$

where by definition $A_i W_j := \sum_{(f,w)\in A_i\times W_j} Kfw$. The *associated graded module* of W with respect to the filtration $W_\bullet$ is the graded $\mathrm{Gr}(A)$-module

$$\mathrm{Gr}(W) = \mathrm{Gr}_{W_\bullet}(W) := \bigoplus_{j\in\mathbb{Z}} W_j / W_{j-1},$$

with scalar multiplication induced by $(f + A_{i-1})(w + W_{j-1}) := fw + W_{i+j-1}$ for all $(i, j) \in \mathbb{N}_0 \times \mathbb{Z}$, all $f \in A_i$ and all $w \in W_j$.

We say that two $A_\bullet$-filtrations $W_\bullet^{(1)}$, $W_\bullet^{(2)}$ are *equivalent* if there is some $r \in \mathbb{N}_0$ such that

$$W_{i-r}^{(1)} \subseteq W_i^{(2)} \subseteq W_{i+r}^{(1)} \text{ for all } i \in \mathbb{Z}.$$

Note that in this situation for all $i \in \mathbb{N}$ and all $f \in A_i$ we have the implication

$$fW_j^{(1)} \subseteq W_{j+i-1}^{(1)} \text{ for all } j \in \mathbb{Z} \quad \Rightarrow \quad f^{2r+1}W_j^{(2)} \subseteq W_{j+(2r+1)i-1}^{(2)} \text{ for all } j \in \mathbb{Z}.$$

So, if the filtration $A_\bullet$ is commutative, we can say:

If $W_\bullet^{(1)}$ is equivalent to $W_\bullet^{(2)}$, then $\sqrt{\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}_{W_\bullet^{(1)}}(W))} = \sqrt{\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}_{W_\bullet^{(2)}}(W))}$.

Remark and Definition 2. Let $V \subseteq W$ be a K-subspace such that $AV = W$. Then $A_\bullet V := (A_j V)_{j\in\mathbb{Z}}$ defines an $A_\bullet$-filtration on W, which we call the $A_\bullet$*-filtration induced by* V. If $V^{(1)}, V^{(2)} \subseteq W$ are two K-subspaces of finite dimension such that $W = AV^{(k)}$ for $k = 1, 2$, the induced filtrations $A_\bullet V^{(1)}$ and $A_\bullet V^{(2)}$ are equivalent so that by Reminder 1 we have $\sqrt{\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}_{A_\bullet V^{(1)}}(W))} = \sqrt{\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}_{A_\bullet V^{(2)}}(W))}$.

Assume now, that the filtration $A_\bullet$ of A is commutative and that the left A-module W is finitely generated. Then, there is a finite-dimensional K-subspace $V \subseteq W$ such that $W = AV$. According to our previous observation, the closed subset

$$\mathbb{V}(W) = \mathbb{V}_{A_\bullet}(W) := \mathrm{Spec}\left(\mathrm{Gr}(A)/\left(\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}_{A_\bullet V}(W))\right)\right) \subseteq \mathrm{Spec}(\mathrm{Gr}(A))$$

does not depend on our choice of V and hence is determined by the filtration $A_\bullet$ and the module W. It is called the *characteristic variety* of the finitely generated left A-module W with respect to the commutative filtration $A_\bullet$ of A.

Remark and Definition 3. Let W be a left A-module equipped with an $A_\bullet$-filtration $W_\bullet$. We say that the $A_\bullet$-filtration $W_\bullet$ is *very good*, if

$$W_j = 0 \text{ for all } j < 0, \quad \dim_K(W_0) < \infty \quad \text{and } W_j = A_j W_0 \text{ for all } j \in \mathbb{N}.$$

Thus, the very good $A_\bullet$-filtrations of W are precisely the filtrations $A_\bullet V$ induced by a K-subspace $V \subseteq W$ of finite dimension. So, W admits a very good filtration if and only if it is finitely generated, and then all good filtrations are equivalent. If $W_\bullet$ is a good filtration of W, we say that $(W, W_\bullet)$—or briefly W—is *very well-filtered* (with respect to the filtration $A_\bullet$).

Assume that $W = (W, W_\bullet)$ is very well-filtered with respect to $A_\bullet$. Then, the associated graded module $\mathrm{Gr}(W) = \mathrm{Gr}_{W_\bullet}(W)$ of W with respect to $W_\bullet$ is generated by finitely many homogeneous elements of degree 0. In particular one may define the *Hilbert function* $h_W = h_{(W,W_\bullet)}$ of W with respect to $W_\bullet$ as the Hilbert function of the graded $\mathrm{Gr}(A)$-module $\mathrm{Gr}(W) = \mathrm{Gr}_{W_\bullet}(W)$, hence

$$h_W(j) = h_{(W,W_\bullet)}(j) = h_{\mathrm{Gr}(W)}(j) = \dim_K(W_j / W_{j-1}) \text{ for all } j \in \mathbb{Z}.$$

Example 4. Let K be a field of characteristic 0, let $n \in \mathbb{N}$, and let $E_n(K) := \mathrm{End}_K(K[X_1, \ldots, X_n])$ denote the endomorphism ring of the polynomial ring $K[X_1, \ldots, X_n]$. For all $i \in \{1, \ldots, n\}$ we identify X_i with the K-endomorphism on $K[X_1, \ldots, X_n]$ given by multiplication with X_i, and we write D_i for the partial derivative with respect to X_i on $K[X_1, \ldots, X_n]$. Then, the nth *Weyl algebra* over K is defined as the subring

$$A_n(K) := K\langle X_1, \ldots, X_n, D_1, \ldots, D_n\rangle \subseteq E_n(K)$$

of $E_n(K)$ generated by the multiplication endomorphisms X_i and the partial derivatives D_i. The ring $A_n(K)$ is a unital associative central K-algebra and its elements are called *partial differential operators* on $K[X_1, \ldots, X_n]$. The elements

$$X^\nu D^\mu := X_1^{\nu_1} \ldots X_n^{\nu_n} D_1^{\mu_1} \ldots D_n^{\mu_n} \in A_n(K), \text{ with } \nu := (\nu_1, \ldots, \nu_n),$$
$$\mu = (\mu_1, \ldots, \mu_n) \in \mathbb{N}_0^n$$

are called *elementary partial differential operators*. One has the *Heisenberg relations*

$$[X_i, X_j] = 0, \quad [D_i, D_j] = 0, \quad [D_i, X_j] = \delta_{ij} \text{ for all } i, j \in \{1, \ldots, n\},$$

where $[\bullet, \bullet]$ denotes the commutator operation and δ_{ij} denotes the Kronecker symbol.

It follows from the Heisenberg relations that the elementary differential operators form a K-vector space basis of $A_n(K)$. Therefore, each element $f \in A_n(K)$ may be written as $f = \sum_{\nu,\mu \in \mathbb{N}_0^n} a_{\nu\mu} X^\nu D^\mu$ with uniquely determined coefficients $a_{\nu\mu} \in K$

which vanish for all but finitely many pairs (ν, μ). So, if $f \neq 0$, we may define the *degree* of f by $\deg(f) := \max\{|\nu| + |\mu| \mid (\nu, \mu) \in \mathbb{N}_0^n : a_{\nu\mu} \neq 0\}$. In addition, we set $\deg(0) := -\infty$. Now, one gets a filtration $A_\bullet$ on $A = A_n(K)$ given by $A_i := \{f \in A \mid \deg(f) \leq i\}$, the so-called *degree filtration*—a commutative very good filtration on A. More precisely, if $x_1, \ldots, x_{2n}$ are indeterminates, one has an isomorphism of graded K algebras:

$$K[x_1, \ldots, x_{2n}] \xrightarrow{\cong} \mathrm{Gr}_{A_\bullet}(A), \quad x_i \mapsto X_i + A_0, \quad x_{i+n} \mapsto D_i + A_0$$
$$\text{for all } i \in \{1, \ldots, n\}.$$

We assume from now on, that the Weyl algebra $A = A_n(K)$ is always endowed with its degree filtration.

Remark 5. The finitely generated left A-modules are called *D-modules* over A. For each D-module W over A, the characteristic variety of W is a closed subset of an affine $2n$-space over K:

$$\mathbb{V}(W) \subset \mathrm{Spec}(\mathrm{Gr}(A)) = \mathrm{Spec}(K[x_1, \ldots, x_{2n}]) = \mathbb{A}_K^{2n}.$$

We endow W with a very good filtration $W_\bullet$ so that its associated graded module $\mathrm{Gr}(W) = \mathrm{Gr}_\bullet(W)$ is generated by finitely many homogeneous elements of degree 0. The very well-filtered D-module $W = (W, W_\bullet)$ has a Hilbert function $h_W = h_{(W,W_\bullet)}$.

We now formulate in its original form the problem concerning the degrees of homogeneous polynomials which set theoretically cut out the characteristic variety of a D-module, posed to us by Bächtold.

Problem 6. Let $W = (W, W_\bullet)$ be a very well-filtered D-module. Do n and the Hilbert function $h_W = h_{(W,W_\bullet)}$ bound from above the degree of homogeneous polynomials in $K[x_1, \ldots, x_{2n}]$ which are needed to cut out the set $\mathbb{V}(W)$ from $\mathbb{A}_K^{2n}$?

Remark 7. By the definition of characteristic variety, the bound we are asking for in Problem 6 is on its turn bounded from above by $\mathrm{gendeg}\big(\mathrm{Ann}_{\mathrm{Gr}(A)}(\mathrm{Gr}(W))\big)$. So, it suffices to bound from above this latter invariant in terms of the Hilbert function $h_{(W,W_\bullet)} = h_{\mathrm{Gr}(W)}$. This is what we are heading for, and this is also what finally was stated in Lemma 7.41 of [2]. This Lemma was used there, to prove a certain uniformity result, which says that, over a C^∞-manifold M, the "global characteristic generically agrees with the point-wise characteristic" (see Theorem 7.39 of [2]).

Remark 8. According to Remark 7, the problem posed in Problem 6 is solved if, for a polynomial ring $R = K[x_1, \ldots, x_r]$ over a field K and a graded R-module M which is generated by finitely many homogeneous elements of degree 0, the generating degree $\mathrm{gendeg}(\mathrm{Ann}_R(M))$ of the annihilator of M is bounded in terms of r and the Hilbert function h_M of M. As $\mathrm{gendeg}(\mathrm{Ann}_R(M)) \leq \mathrm{reg}(\mathrm{Ann}_R(M))$, it suffices indeed to show that the regularity of the annihilator of M is bounded in terms of r and the Hilbert function of M.

Remark 9. Let the notation and hypotheses as in Remark 8 and assume that M is generated by μ homogeneous elements of degree 0. We aim to find an upper bound on $\text{reg}(\text{Ann}_R(M))$ which depends only on r and h_M. In fact, the Hilbert function is a somehow enigmatic object, as it is not clear (e.g., from the computational point of view) what it means "to know" a function $h : \mathbb{Z} \to \mathbb{N}_0$. Such arithmetic functions may encode an uncountable variety of information, and thus are not accessible for finitistic considerations. We therefore prefer to replace the function h_M by finitistic invariants (which are known in a "philosophical sense" if h_M is). We thus aim to bound $\text{reg}(\text{Ann}_R(M))$ in terms of r, $\mu = h_M(0)$, the Hilbert polynomial p_M of M, and the postulation number $p(M)$ of M.

We shall do this in a more general context. Hence, from now on, let $R = \bigoplus_{n\in\mathbb{N}_0} R_n$ be as in Notation 1, that is, a Noetherian homogeneous ring with Artinian local base ring $(R_0, \mathfrak{m}_0)$ and let $M = \bigoplus_{n\in\mathbb{Z}} M_n$ be a finitely generated graded R-module.

We begin with the following bounding result for the regularity $\text{reg}(\text{Ann}_R(M))$ of the annihilator $\text{Ann}_R(M)$ of the graded R-module M.

Theorem 10. *Let $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0R_1) > 0$, set $\lambda := \text{length}(R_0)$, $\rho := \text{reg}(R)$. If $M \neq 0$ is generated by μ homogeneous elements,*

$$\beta := \text{reg}(M) + \text{gendeg}(M) - 2\text{beg}(M) \quad \textit{and}\ \alpha := \text{beg}(M) - \text{gendeg}(M),$$

then we have

$$\text{reg}(\text{Ann}_R(M)) \leq \max\{\rho, \left[\beta + \lambda(\mu^2+1)+1\right]^{2^{r-1}} + \alpha + 1\}.$$

Proof. By our definition of the number r there is a surjective homomorphism of homogeneous R_0-algebras $\phi : S = R_0[x_1, x_2, \ldots, x_r] \twoheadrightarrow R$, where S is a standard graded polynomial ring over R_0. Clearly, the invariants $\text{gendeg}(M)$, $\text{beg}(M)$, and μ are not affected, if we consider M as a graded S-module by means of ϕ. In addition, the invariants $\text{reg}(M)$ and $\text{reg}(\text{Ann}_R(M))$ are not affected if we consider M and $\mathfrak{a} := (\text{Ann}_R(M))$ as graded S-modules by means of ϕ.

We now set $\mathfrak{b} := \text{Ann}_S(M) = \phi^{-1}(\mathfrak{a})$ so that we have an isomorphism of graded S-modules $R/\mathfrak{a} \cong S/\mathfrak{b}$ and hence a short exact sequence of graded S-modules

$$0 \longrightarrow \mathfrak{a} \longrightarrow R \longrightarrow S/\mathfrak{b} \longrightarrow 0.$$

Consequently we have $\text{reg}(\mathfrak{a}) \leq \max\{\text{reg}(R), \text{reg}(S/\mathfrak{b}) + 1\}$. So, it suffices to show that

$$\text{reg}(S/\mathfrak{b}) \leq \left[\beta + \lambda(\mu^2+1)+1\right]^{2^{r-1}} + \alpha.$$

Observe that we have an exact sequence of R-modules

$$0 \longrightarrow \mathfrak{b} \longrightarrow S \xrightarrow{\epsilon} \mathrm{Hom}_S(M, M), \quad (x \mapsto \epsilon(x) := x\mathrm{Id}_M)$$

and an epimorphism of graded S-modules

$$\pi : \bigoplus_{i=1}^{\mu} S(-a_i) \twoheadrightarrow M, \quad \mathrm{beg}(M) = a_1 \leq a_2 \leq \cdots \leq a_\mu = \mathrm{gendeg}(M).$$

In particular we obtain an induced monomorphism of graded S-modules

$$0 \longrightarrow \mathrm{Hom}_S(M, M) \xrightarrow{g:=\mathrm{Hom}_S(\pi, M)} \mathrm{Hom}_S\left(\bigoplus_{i=1}^{\mu} S(-a_i), M\right) = \bigoplus_{i=1}^{\mu} M(a_i) =: V.$$

So we get a composition map

$$S \xrightarrow{f:=g\circ\epsilon} V, \quad \text{with} \quad \mathrm{Im}(f) = \mathrm{Im}(\epsilon) \cong S/\mathfrak{b}.$$

Now, observe that S is a Cohen–Macaulay ring of dimension r with $\mathrm{gendeg}(S) = \mathrm{reg}(S) = 0$ and with multiplicity λ. Moreover the S-module V is generated by μ^2 homogeneous elements. Furthermore, we have

$$\mathrm{beg}(V) = \mathrm{beg}(M) - a_\mu = \mathrm{beg}(M) - \mathrm{gendeg}(M) = \alpha$$

and

$$\mathrm{reg}(V) = \mathrm{reg}(M) - a_1 = \mathrm{reg}(M) - \mathrm{beg}(M) \geq 0 = \mathrm{gendeg}(S).$$

In particular we have $\min\{\mathrm{beg}(V), \mathrm{reg}(V) - \mathrm{reg}(S)\} = \mathrm{beg}(V) = \alpha$ and $\mathrm{reg}(V) \geq 0$. So, if we apply Lemma 4 to the above homomorphism $f : S \longrightarrow V$ and observe that $\mathrm{Im}(f) \cong S/\mathfrak{b}$, we obtain indeed

$$\begin{aligned}\mathrm{reg}(S/\mathfrak{b}) &\leq \left[\mathrm{reg}(M) - \mathrm{beg}(M) + 1 + \lambda(\mu^2+1) - \mathrm{beg}(M) + \mathrm{gendeg}(M)\right]^{2^{r-1}} + \alpha \\ &= \left[\beta + \lambda(\mu^2 + 1) + 1\right]^{2^{r-1}} + \alpha.\end{aligned}$$

□

As an immediate consequence we now get the an upper bound for the regularity of the annihilator of M in terms of the two invariants $\rho := \mathrm{reg}(R), \lambda := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1)$ of the ring R and the three invariants $\mathrm{reg}(M), \mathrm{beg}(M), \mu := \dim_{R_0/\mathfrak{m}_0}(M/(\mathfrak{m}_0 R + R_1 R)M)$ of the module M of Theorem 10.

Corollary 11. *Let R, M, r, λ, ρ, and μ be as in Theorem 10. Then it holds*

$$\mathrm{reg}(\mathrm{Ann}_R(M)) \leq \max\{\rho, \left[2\big(\mathrm{reg}(M) - \mathrm{beg}(M)\big) + \lambda(\mu^2 + 1) + 1\right]^{2^{r-1}} + 1\}.$$

Proof. This is clear by Theorem 10 as $\text{gendeg}(M) \leq \text{reg}(M)$ and $\alpha = \text{beg}(M) - \text{gendeg}(M) \leq 0$. □

This bound becomes particularly simple if R is a polynomial ring.

Corollary 12. *Let $r \in \mathbb{N}$ and assume that $R = R_0[x_1, x_2, \ldots, x_r]$ is a standard graded polynomial ring. If $M \neq 0$ is generated by μ homogeneous elements, then we have*

$$\text{reg}(\text{Ann}_R(M)) \leq \left[2\big(\text{reg}(M) - \text{beg}(M)\big) + \text{length}(R_0)(\mu^2 + 1) + 1\right]^{2^{r-1}} + 1.$$

Proof. This follows immediately from Corollary 11 as $\text{reg}(R) = 0$. □

The following special case covers the situation of primary interest.

Corollary 13. *Let $R = K[x_1, x_2, \ldots, x_r]$ be a polynomial ring over the field K. If $M \neq 0$ is a graded R-module which is generated by μ homogeneous elements of degree* 0*, then we have*

$$\text{reg}(\text{Ann}_R(M)) \leq \left[2\text{reg}(M) + \mu^2 + 2\right]^{2^{r-1}} + 1.$$

To answer affirmatively our original question on characteristic varieties of D-modules, we can use the following result, in which the function $F_{p,\boldsymbol{a}}$ is as in Proposition 5.

Theorem 14. *Let r, λ, and ρ be as in Theorem 10, let $\mu \in \mathbb{N}$, and let $\boldsymbol{a} := (a_1, a_2, \ldots, a_\mu) \in \mathbb{Z}^\mu$ with $a_1 \leq a_2 \leq \cdots \leq a_\mu$. If $M = \sum_{i=1}^{\mu} Rm_i$ is a finitely generated graded R-module with $m_i \in M_{a_i}$ for $i = 1, 2, \ldots .\mu$, then we have*

$$\text{reg}(\text{Ann}_R(M)) \leq \max\{\rho, \left[F_{p_M,\boldsymbol{a}}(\lambda, r, p(M)) + a_\mu - 2a_1 + \lambda(\mu^2 + 1) + 1\right]^{2^{r-1}} + a_1 - a_\mu + 1\}.$$

In particular, we have

$$\text{reg}(\text{Ann}_R(M)) \leq \max\{\rho, \left[2\big(F_{p_M,\boldsymbol{a}}(\lambda, r, p(M)) - a_1\big) + \lambda(\mu^2 + 1) + 1\right]^{2^{r-1}} + 1\}.$$

Proof. This is immediate by Theorem 10, respectively, Corollary 11 and Proposition 5 as $\text{beg}(M) = a_1$ and $\text{gendeg}(M) = a_\mu$. □

Now, we have reached the goal set out in Remark 9 by the special case of the previous bound in which $\boldsymbol{0} \in \mathbb{Z}^\mu$.

Corollary 15. *Let K be a field and R be a Noetherian homogeneous K-algebra set $r := \dim_K(R_1)$ and $\rho := \text{reg}(R)$. If $M \neq 0$ is a graded R-module which is generated by μ homogeneous elements of degree* 0*, then we have*

$$\text{reg}(\text{Ann}_R(M)) \leq \max\{\rho, \left[2F_{p_M,\boldsymbol{0}}(1, r, p(M)) + \mu^2 + 2\right]^{2^{r-1}} + 1\}.$$

Finally, we also recover the bound we suggested to look for in Remark 9.

Corollary 16. *Let $R = K[x_1, x_2, \ldots, x_r]$ be a polynomial ring over the field K and let $M \neq 0$ be a graded R-module which is generated by μ homogeneous elements of degree 0. Then the regularity of the annihilator of M is bounded in terms of the number r of indeterminates, the Hilbert polynomial p_M of M, the postulation number $p(M)$ of M, and the number μ of generators of M. More precisely, we have*

$$\mathrm{reg}(\mathrm{Ann}_R(M)) \leq \left[2F_{p_M,\mathbf{0}}(1, r, p(M)) + \mu^2 + 2)\right]^{2^{r-1}} + 1.$$

4 A Regularity Bound for Ext-Modules

The aim of this section is to give an upper bound on the Castelnuovo–Mumford regularity of the modules $\mathrm{Ext}^i_R(M, N)$ in terms of the number r of linear forms, which are needed to generate R as an R_0-algebra, and the regularities and initial degrees of the modules M and N. We begin with the case $i = 0$ and give a bound on the regularity of the graded R-module $\mathrm{Hom}_R(M, N)$.

Lemma 1. *Let $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, $\lambda := \mathrm{length}(R_0)$ and let*

$$0 \longrightarrow U \xrightarrow{f} V \xrightarrow{g} W$$

be an exact sequence of finitely generated graded R-modules. If $W \neq 0$ is generated by μ homogeneous elements, then we have

$$\mathrm{reg}(U) \leq \max\{\mathrm{reg}(V), \left[\max\{\mathrm{gendeg}(V), \mathrm{reg}(W) + 1\} + \lambda(\mu + 1) - \mathrm{beg}(W)\right]^{2^{r-1}} + \mathrm{beg}(W) + 1\}$$

Proof. The short exact sequence of graded R-modules

$$0 \longrightarrow U \xrightarrow{f} V \longrightarrow \mathrm{Im}(g) \longrightarrow 0$$

gives $\mathrm{reg}(U) \leq \max\{\mathrm{reg}(V), \mathrm{reg}(\mathrm{Im}(g)) + 1\}$. Hence, it suffices to show that

$$\mathrm{reg}(\mathrm{Im}(g)) \leq \left[\max\{\mathrm{gendeg}(V), \mathrm{reg}(W)+1\}+\lambda(\mu+1)-\mathrm{beg}(W)\right]^{2^{r-1}} + \mathrm{beg}(W).$$

According to our definition of the number r there is a surjective homomorphism of homogeneous R_0-algebras $\phi : S = R_0[x_1, x_2, \ldots, x_r] \twoheadrightarrow R$, in which S is a standard graded polynomial ring over R_0. None of the numerical invariants occurring in the requested inequality are affected if we consider U, V and W as graded S-modules by means of ϕ. Thus, we may replace R by S and hence assume that $R = R_0[x_1, x_2, \ldots, x_r]$ is a polynomial ring. In particular R then is

Cohen–Macaulay of dimension r of multiplicity λ and satisfies $\operatorname{reg}(R) = 0$. Now, we get the requested inequality if we apply Lemma 4 to the homomorphism $V \xrightarrow{g} W$. □

Lemma 2. *Let* $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, $\lambda := \operatorname{length}(R_0)$; *let* M *and* N *be two non-zero finitely generated graded* R*-modules and suppose that there is an exact sequence of graded* R*-modules*

$$\bigoplus_{j=1}^{\sigma} R(-b_j) \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \longrightarrow M \longrightarrow 0$$

with integers $b_1 \le b_2 \le \cdots \le b_\sigma$ *and* $a_1 \le a_2 \le \cdots \le a_\mu$. *If* $N \ne 0$ *is generated by* ν *homogeneous elements,*

$$\beta := \max\{\operatorname{gendeg}(N) - a_1, \operatorname{reg}(N) - b_1 + 1\}, \quad \textit{and } \gamma := \operatorname{beg}(N) - b_\sigma,$$

then we have

$$\operatorname{reg}(\operatorname{Hom}_R(M,N)) \le \max\{\operatorname{reg}(N) - a_1, \left[\beta + \lambda(\sigma\nu + 1) - \gamma\right]^{2^{r-1}} + \gamma + 1\}.$$

Proof. Apply Lemma 1 to the induced exact sequence

$$0 \longrightarrow \operatorname{Hom}_R(M,N) \longrightarrow \bigoplus_{i=1}^{\mu} N(a_i) \longrightarrow \bigoplus_{j=1}^{\sigma} N(b_j)$$

and observe that

$$\operatorname{gendeg}\left(\bigoplus_{i=1}^{\mu} N(a_i)\right) = \operatorname{gendeg}(N) - a_1, \quad \operatorname{reg}\left(\bigoplus_{i+1}^{\mu} N(a_i)\right) = \operatorname{reg}(N) - a_1,$$

$$\operatorname{beg}\left(\bigoplus_{j=1}^{\sigma} N(b_j)\right) = \operatorname{beg}(N) - b_\sigma, \quad \operatorname{reg}\left(\bigoplus_{j=1}^{\sigma} N(b_j)\right) = \operatorname{reg}(N) - b_1$$

and that $\bigoplus_{j=1}^{\sigma} N(b_j)$ is generated by $\sigma\nu$ homogeneous elements. □

Proposition 3. *Let* $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, $\lambda = \operatorname{length}(R_0)$. *If* M *and* N *are two nonzero graded* R*-modules which are generated, respectively, by* μ *and* ν *homogeneous elements, then we have*

$$\operatorname{reg}(\operatorname{Hom}_R(M,N)) \le \left[\operatorname{w}(M) + \operatorname{w}(N) - 1 + \left(\binom{\operatorname{w}(M)+r}{r-1}\lambda\mu\nu + 1\right)\lambda\right]^{2^{r-1}} + \operatorname{beg}(N) - \operatorname{beg}(M).$$

Proof. Again, there is a surjective homomorphism $\phi : S = R_0[x_1, x_2, \dots, x_r] \twoheadrightarrow R$ of homogeneous R_0-algebras. Observe in particular that the graded S-modules $\mathrm{Hom}_S(M, N)$ and $\mathrm{Hom}_R(M, N)$ are isomorphic. So, the numerical invariants occurring in our statement are not affected if we consider M and N as graded S-modules by means of ϕ. Hence, we may once more assume that $R = R_0[x_1, x_2, \dots, x_r]$ is a polynomial ring. Now, let

$$\bigoplus_{j=1}^{\sigma} R(-b_j) \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \longrightarrow M \longrightarrow 0$$

with $b_1 \le b_2 \le \cdots \le b_\sigma$ and $a_1 \le a_2 \le \cdots \le a_\mu$ be a minimal free presentation of M. Then, as

$$\mathrm{gendeg}(N) \le \mathrm{reg}(N), \text{ and } \mathrm{reg}(M) + 1 \ge b_\sigma \ge b_1 \ge a_1 + 1 = \mathrm{beg}(M) + 1,$$

we get the following inequalities:

$$\begin{aligned}
\beta &:= \max\{\mathrm{gendeg}(N) - a_1, \mathrm{reg}(N) - b_1 + 1\} \le \mathrm{reg}(N) - \mathrm{beg}(M),\\
\gamma &:= \mathrm{beg}(N) - b_\sigma \le \mathrm{beg}(N) - \mathrm{beg}(M) - 1,\\
-\gamma &\le \mathrm{reg}(M) - \mathrm{beg}(N) + 1.
\end{aligned}$$

Moreover, by the minimality of our presentation, we have

$$\sigma \le \mathrm{length}_{R_0}\left(\Big(\bigoplus_{i=1}^{\mu} R(-a_i) \Big)_{\le \mathrm{reg}(M)+1} \right) \le \lambda\mu \binom{\mathrm{w}(M)+r}{r-1}.$$

Thus, we may conclude by Lemma 2. □

Now, we are ready to prove the main result of this section.

Theorem 4. *Let $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, let $\lambda := \mathrm{length}(R_0)$, $\rho = \mathrm{reg}(R)$, and let M and N be two nonzero graded R-modules which are generated, respectively, by μ and ν homogeneous elements. Then, for each $i \in \mathbb{N}_0$, we have*

$$\begin{aligned}
&\mathrm{reg}(\mathrm{Ext}^i_R(M, N)) \le\\
&\Big[\mathrm{w}(M) + \mathrm{w}(N) + i\rho - 1 + \Big(\lambda^{i+1}\mu\nu\binom{\mathrm{w}(M)+r+i\rho}{r-1} \prod_{j=1}^{i} \Big(\binom{\mathrm{w}(M)+j\rho+r}{r-1} + 1\Big)\lambda\Big]^{2^{r-1}}\\
&+ \mathrm{beg}(N) - \mathrm{beg}(M) - i.
\end{aligned}$$

Proof. The case $i = 0$ is clear by Proposition 3. To treat the cases with $i > 0$ we choose a short exact sequence of graded R-modules

$$0 \longrightarrow M' \longrightarrow \bigoplus_{k=1}^{\mu} R(-a_k) \xrightarrow{\pi} M \longrightarrow 0, \quad \operatorname{beg}(M) = a_1 \leq a_2 \leq \cdots \leq a_\mu$$
$$= \operatorname{gendeg}(M),$$

in which the epimorphism π is minimal. If $M' = 0$, the module M is free, and hence our claim is obvious. So, let $M' \neq 0$ and consider the induced exact sequence of graded R-modules

$$0 \longrightarrow \operatorname{Hom}_R(M, N) \longrightarrow \bigoplus_{k=1}^{\mu} N(a_k) \xrightarrow{f} \operatorname{Hom}_R(M', N) \longrightarrow \operatorname{Ext}^1_R(M, N) \longrightarrow 0$$

and the induced isomorphisms of graded R-modules

$$\operatorname{Ext}^i_R(M, N) \cong \operatorname{Ext}^{i-1}_R(M', N) \text{ for all } i > 1.$$

We first aim to prove our statement in the case $i = 1$. From the above four term exact sequence, we get the estimates

$$\operatorname{reg}(\operatorname{Ext}^1_R(M, N)) \leq \max\{\operatorname{reg}(\operatorname{Im}(f)) - 1, \operatorname{reg}(\operatorname{Hom}_R(M', N))\}$$

and

$$\operatorname{reg}(\operatorname{Im}(f)) \leq \max\{\operatorname{reg}(\operatorname{Hom}_R(M, N)) - 1, \operatorname{reg}(\bigoplus_{k=1}^{\mu} N(a_k))\}.$$

Our next aim is to make explicit the second estimate. According to Proposition 3, we have

$$\operatorname{reg}(\operatorname{Hom}_R(M, N)) - 1$$
$$\leq \left[\mathrm{w}(M) + \mathrm{w}(N) - 1 + \left(\binom{\mathrm{w}(M)+r}{r-1}\lambda\mu\nu + 1\right)\right]^{2^{r-1}} + +\operatorname{beg}(N) - \operatorname{beg}(M) - 1.$$

Moreover, the term $\operatorname{reg}\left(\bigoplus_{k=1}^{\mu} N(a_k)\right) = \operatorname{reg}(N) - a_1 = \operatorname{reg}(N) - \operatorname{beg}(M) = \mathrm{w}(N) + \operatorname{beg}(N) - \operatorname{beg}(M) \leq \mathrm{w}(N) + \mathrm{w}(M) + \operatorname{beg}(N) - \operatorname{beg}(M) - 1$ cannot exceed the right-hand side of the above inequality, so that we get the following explicit estimate:

$$\operatorname{reg}(\operatorname{Im}(f)) \leq \left[\mathrm{w}(M)+\mathrm{w}(N)-1+\left(\binom{\mathrm{w}(M)+r}{r-1}\lambda\mu\nu+1\right)\lambda\right]^{2^{r-1}} +\operatorname{beg}(N)-\operatorname{beg}(M)-1.$$

Our next aim is to bound the invariant $\operatorname{reg}(\operatorname{Hom}_R(M', N))$. By our initial minimal short exact sequence, we have $\operatorname{reg}(M') \leq \operatorname{reg}(M)+\rho+1$ and $\operatorname{beg}(M') \geq \operatorname{beg}(M)+1$, so that we obtain

$$\mathrm{w}(M') \leq \mathrm{w}(M) + \rho, \quad -\operatorname{beg}(M') \leq -\operatorname{beg}(M) - 1.$$

Let μ' denote the minimal number of homogeneous generators of M'. As gendeg(M') $\leq \operatorname{reg}(M') \leq \operatorname{reg}(M) + \rho + 1$, we have

$$\mu' \leq \operatorname{length}\big((\bigoplus_{k=1}^{\mu} R(-a_k))_{\leq \operatorname{reg}(M)+\rho+1}\big) \leq \tbinom{\mathrm{w}(M)+\rho+r}{r-1}\lambda\mu.$$

Using these estimates and applying Proposition 3, we obtain

$$\begin{aligned}
&\operatorname{reg}(\operatorname{Hom}_R(M', N)) \\
&\leq \big[\mathrm{w}(M) + \rho + \mathrm{w}(N) + \rho - 1 + \big(\tbinom{\mathrm{w}(M)+\rho+r}{r-1}^2 \lambda^2\mu\nu + 1\big)\lambda\big]^{2^{r-1}} \\
&\quad + \operatorname{beg}(N) - \operatorname{beg}(M) - 1.
\end{aligned}$$

Observe, that this term exceeds our previous upper bound for reg(Im(f)). So on use of our very first inequality, we end up with

$$\begin{aligned}
&\operatorname{reg}(\operatorname{Ext}^1_R(M, N)) \\
&\leq \big[\mathrm{w}(M) + \mathrm{w}(N) + \rho - 1 + \big(\tbinom{\mathrm{w}(M)+\rho+r}{r-1}^2 \lambda^2\mu\nu + 1\big)\lambda\big]^{2^{r-1}} \\
&\quad + \operatorname{beg}(N) - \operatorname{beg}(M) - 1.
\end{aligned}$$

This proves our claim if $i = 1$.

For $i > 1$ we now may proceed by induction on use of the previously observed isomorphisms of Ext-modules and keeping in mind the above inequalities $\mathrm{w}(M') \leq \mathrm{w}(M) + \rho$, $-\operatorname{beg}(M') \leq -\operatorname{beg}(M) - 1$, and $\mu' \leq \tbinom{\mathrm{w}(M)+\rho+r}{r-1}\lambda\mu$. □

In case R is a polynomial ring, this bound becomes simpler in appearance.

Corollary 5. *Assume that M and N are two non-zero graded modules generated by μ respectively, ν homogeneous elements over the polynomial ring $R = R_0[x_1, x_2, \ldots, x_r]$ with $\lambda :=$ length(R_0). Then, for each $i \in \mathbb{N}_0$, we have*

$$\operatorname{reg}(\operatorname{Ext}^i_R(M, N)) \leq$$

$$\big[\mathrm{w}(M) + \mathrm{w}(N) - 1 + \big(\tbinom{\mathrm{w}(M)+r}{r-1}^{i+1} \lambda^{i+1}\mu\nu + 1\big)\lambda\big]^{2^{r-1}} + \operatorname{beg}(N) - \operatorname{beg}(M) - i.$$

Proof. This is clear by Theorem 4 as $\operatorname{reg}(R) = 0$. □

In [17], Hoa and Hyry did give upper bounds for the Castelnuovo–Mumford regularity of deficiency modules of graded ideals in polynomial rings over a field. In [9], Brodmann, Jahangiri, and Linh took up this idea and gave upper bounds for the Castelnuovo–Mumford regularity of deficiency modules of finitely generated graded modules over a standard graded Noetherian ring R with local Artinian base ring $(R_0, \mathfrak{m}_0)$. We aim to take up this theme again.

Remark and Notation 6. We set $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1)$. Then there is a surjective homomorphism of graded R_0-algebras $R' := R'_0[x_1, x_2, \ldots, x_r] \twoheadrightarrow R$ where $(R'_0, \mathfrak{m}'_0)$ is an Artinian Gorenstein ring. Let M be a finitely generated graded R-module, which we also consider as R'-module by means of the above homomorphism. Then, for each $i \in \mathbb{N}_0$, the *i th deficiency module of M* is given by

$$K^i(M) = \mathrm{Ext}_{R'}^{r-i}(M, R'(-r)).$$

We write $\lambda := \mathrm{length}(R_0)$ and λ' for the minimum length of all local Artinian Gorenstein rings R'_0 such that R_0 is a homomorphic image of R'_0. We may write R_0 as a homomorphic image of a complete regular local ring S_0 of dimension $\mathrm{e} := \mathrm{edim}(R_0) = \mathrm{length}_{R_0}(\mathfrak{m}/\mathfrak{m}^2)$. Let $a_1, a_2, \ldots, a_\mathrm{e}$ be a regular system of parameters of S_0. Then $(a_j)^\lambda$ is mapped to 0 under the canonical map $S_0 \twoheadrightarrow R_0$ for all $j \in \{1, 2, \ldots, \mathrm{e}\}$. Therefore R_0 is a homomorphic image of the Artinian Gorenstein ring

$$R'_0 := S_0/\langle (a_1)^\lambda, (a_2)^\lambda, \ldots, (a_\mathrm{e})^\lambda \rangle$$

But this means that we have

$$\lambda \leq \lambda' \leq 1 + \lambda \mathrm{edim}(R_0).$$

Now, as an application of Theorem 4 and with the above notations, we get the following bounding result on the regularity of deficiency modules. Observe in particular that the estimates given in statements (a) and (c) allow to bound the regularity of the ith deficiency module of a finitely generated graded R-module M only in terms of i, the initial degree of M, the regularity of M and invariants of R.

Corollary 7. *Let r and ρ be as in Theorem 4 and let M be a nonzero graded R-module which is generated by μ homogeneous elements. Let $t := \mathrm{reg}^2(M)$, let $p_M(n)$ denote the Hilbert polynomial of M, and let λ' be defined as in Remark and Notation 6. Then the following statements hold:*

(a) $\mathrm{reg}(K^0(M)) \leq -\mathrm{beg}(M)$.
(b) $\mathrm{reg}(K^1(M)) \leq \max\{0, 1 + t - \mathrm{beg}(M)\} + (d-1)p_M(t) - t$.
(c) For all $i \in \mathbb{N}$ we have

$$\mathrm{reg}(K^i(M)) \leq i - \mathrm{beg}(M) +$$

$$[\mathrm{w}(M) + (r-i-1)\rho - 1 + (\lambda')^{i+1}\mu\binom{\mathrm{w}(M)+r+(r-i)\rho}{r-1}\prod_{j=1}^{r-i}(\binom{\mathrm{w}(M)+j\rho+r}{r-1}+1)]^{2^{r-1}}.$$

Proof. (a) and (b) were proved in [9, Theorem 4.2]. (c) is implied directly by Theorem 4 as in the notations of Remark and Notation 6 we may replace R by $R' := R'_0[x_1, x_2, \ldots, x_r]$, where R'_0 is an Artinian Gorenstein ring of length λ'. □

In case R is a polynomial ring over a field, statement (c) of the above result takes a particularly simple form.

Corollary 8. *Let K be a field and let M be a finitely generated graded module over the polynomial ring $K[x_1, x_2 \dots, x_r]$. Let $i \in \mathbb{N}$. Then it holds*

$$\operatorname{reg}(K^i(M)) \leq i - \operatorname{beg}(M) + [\mathrm{w}(M) - 1 + \mu\binom{\mathrm{w}(M)+r}{r-1}^{r-i+1} + 1)]^{2^{r-1}}.$$

Proof. Observe that in our situation we have $\rho = 0$ and $\lambda' = 1$. □

5 Regularity Bounds for Tor-Modules

For two finitely generated graded R-modules M and N, the modules $\operatorname{Tor}_i^R(M, N)$ are finitely generated and carry a natural grading for all $i \in \mathbb{N}_0$. The aim of this section is to give an upper bound for the Castelnuovo–Mumford regularity of the modules $\operatorname{Tor}_i^R(M, N)$ in terms of the same bounding invariants as in Sect. 4. As in Sect. 4 we begin with the case $i = 0$ and give a regularity bound for the graded R-module $M \otimes_R N$.

Lemma 1. *Let $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, $\lambda = \operatorname{length}(R_0)$ and let*

$$U \xrightarrow{f} V \xrightarrow{g} W \longrightarrow 0$$

be an exact sequence of finitely generated graded R-modules. If $V \neq 0$ is generated by μ homogeneous elements, then we have

$$\operatorname{reg}(W) \leq \max\{\operatorname{reg}(V), \big[\max\{\operatorname{gendeg}(U), \operatorname{reg}(V) + 1\} + \lambda(\mu + 1) - \operatorname{beg}(V)\big]^{2^{r-1}} + \operatorname{beg}(V) - 1\}.$$

Proof. In view of the short exact sequence of graded R-modules

$$0 \longrightarrow \operatorname{Im}(f) \longrightarrow V \xrightarrow{g} W \longrightarrow 0,$$

we have $\operatorname{reg}(W) \leq \max\{\operatorname{reg}^1(\operatorname{Im}(f)) - 1, \operatorname{reg}(V)\}$. So it suffices to show that

$$\operatorname{reg}(\operatorname{Im}(f)) \leq \big[\max\{\operatorname{gendeg}(U), \operatorname{reg}(V) + 1\} + \lambda(\mu + 1) - \operatorname{beg}(V)\big]^{2^{r-1}} + \operatorname{beg}(V).$$

As in the proof of Lemma 1, we may assume that $R = R_0[x_1, x_2, \dots, x_r]$ is a polynomial ring, so that R is CM of dimension r of multiplicity λ and satisfies $\operatorname{reg}(R) = 0$. Now, we get the requested inequality if we apply once more Lemma 4 to the homomorphism $U \xrightarrow{f} V$. □

Lemma 2. *Let* $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, $\lambda := \operatorname{length}(R_0)$, *let* M *and* N *be two non-zero finitely generated graded* R*-modules, and suppose that there is an exact sequence of graded* R*-modules*

$$\bigoplus_{j=1}^{\sigma} R(-b_j) \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \longrightarrow M \longrightarrow 0$$

with integers $b_1 \leq b_2 \leq \cdots \leq b_\sigma$ *and* $a_1 \leq a_2 \leq \cdots \leq a_\mu$. *Suppose in addition that* N *is generated by* ν *homogeneous elements and set*

$$\delta := \max\{\operatorname{gendeg}(N) + b_\sigma, \operatorname{reg}(N) + a_\mu + 1\}, \quad \varepsilon := \operatorname{beg}(N) + a_1.$$

Then we have

$$\operatorname{reg}(M \otimes_R N) \leq \max\{\operatorname{reg}(N) + a_\mu, \left[\delta + \lambda(\mu\nu + 1) - \varepsilon\right]^{2^{r-1}} + \varepsilon - 1\}.$$

Proof. Apply Lemma 1 to the induced exact sequence

$$\bigoplus_{j=1}^{\sigma} N(-bj) \longrightarrow \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$$

and observe that

$$\operatorname{gendeg}\Big(\bigoplus_{j=1}^{\sigma} N(-b_j)\Big) = \operatorname{gendeg}(N) + b_\sigma, \quad \operatorname{reg}\Big(\bigoplus_{i=1}^{\mu} N(-a_i)\Big) = \operatorname{reg}(N) + a_\mu,$$

$$\operatorname{beg}\Big(\bigoplus_{i=1}^{\mu} N(-a_i)\Big) = \operatorname{beg}(N) + a_1$$

and that $\bigoplus_{i=1}^{\mu} N(-a_i)$ is generated by $\mu\nu$ homogeneous elements. This gives the requested bound. □

Proposition 3. *Let* $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 0$, *let* $\lambda = \operatorname{length}(R_0)$. *If* M *and* N *are two non-zero graded* R*-modules which are generated respectively by* μ *and* ν *homogeneous elements, then we have*

$$\operatorname{reg}(M \otimes_R N) \leq \left[\operatorname{w}(M) + \operatorname{w}(N) + \lambda(\mu\nu + 1) - 1\right]^{2^{r-1}} + \operatorname{beg}(M) + \operatorname{beg}(N) - 1.$$

Proof. Again, there is a surjective homomorphism $\phi : S = R_0[x_1, x_2, \ldots, x_r] \twoheadrightarrow R$ of homogeneous R_0-algebras and the graded S-modules $M \otimes_S N$ and $M \otimes_R N$ are isomorphic. So none of the numerical invariants which occur in our statement is affected if we consider M and N as graded S-modules by means of ϕ. Therefore we may again assume that $R = R_0[x_1, x_2, \dots, x_r]$ is a polynomial ring and chose a minimal graded free presentation

$$\bigoplus_{j=1}^{\sigma} R(-b_j) \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \longrightarrow M \longrightarrow 0$$

of M with $b_1 \leq b_2 \leq \cdots \leq b_\sigma$ and $a_1 \leq a_2 \cdots \leq a_\mu$. Then, as

$$a_1 = \operatorname{beg}(M), \quad a_\mu \leq \operatorname{reg}(M), \quad \operatorname{gendeg}(N) \leq \operatorname{reg}(N), \quad b_\sigma \leq \operatorname{reg}(M) + 1$$

we get the relations

$$\operatorname{reg}(N) + a_\mu \leq \operatorname{reg}(M) + \operatorname{reg}(N),$$
$$\delta := \max\{\operatorname{gendeg}(N) + b_\sigma, \operatorname{reg}(N) + a_\mu + 1\} \leq \operatorname{reg}(M) + \operatorname{reg}(N) + 1,$$
$$\varepsilon := \operatorname{beg}(N) + a_1 = \operatorname{beg}(M) + \operatorname{beg}(N).$$

Now, it follows by Lemma 2 that

$$\operatorname{reg}(M \otimes_R N) \leq$$
$$\leq \max\{\operatorname{reg}(M)+\operatorname{reg}(N), \left[\operatorname{w}(M)+\operatorname{w}(N)+\lambda(\mu\nu+1)-1\right]^{2^{r-1}}+\operatorname{beg}(M)+\operatorname{beg}(N)-1\}.$$

As

$$\operatorname{reg}(M) + \operatorname{reg}(N) = [\operatorname{w}(M) + \operatorname{w}(N) - 1] + \operatorname{beg}(M) + \operatorname{beg}(N) - 1 \leq$$
$$\leq \left[\operatorname{w}(M) + \operatorname{w}(N) + \lambda(\mu\nu + 1) - 1\right]^{2^{r-1}} + \operatorname{beg}(M) + \operatorname{beg}(N) - 1,$$

we finally get our claim. □

As an application we get the following estimate for the regularity of Tor-modules, which is not symmetric in the two occurring modules. So, to get out the best of it, one should apply the result after eventually exchanging M and N such that $\operatorname{w}(M) \leq \operatorname{w}(N)$. Observe also that the case $r = 1$ is omitted in this result.

Theorem 4. *Let $r := \dim_{R_0/\mathfrak{m}_0}(R_1/\mathfrak{m}_0 R_1) > 1$, let $\lambda := \operatorname{length}(R_0)$, let $\rho = \operatorname{reg}(R)$ and let M and N be two non-zero finitely generated graded R modules which are generated respectively by μ and ν homogeneous elements. Then, for all $i \in \mathbb{N}_0$ we have*

$$\operatorname{reg}(\operatorname{Tor}_i^R(M, N)) \leq$$
$$\left[\operatorname{w}(M) + \operatorname{w}(N) + i\rho - 1 + \left(\lambda^i \mu\nu \prod_{j=1}^{i} \left(\binom{\operatorname{w}(M)+r+j\rho}{r-1} + 1\right)\lambda\right]^{2^{r-1}}$$
$$+ \operatorname{beg}(N) + \operatorname{reg}(M) + i\rho.$$

Proof. We proceed by induction on i. The case $i = 0$ is clear by Proposition 3. We first treat the case $i = 1$. Consider a graded short exact sequence

$$0 \longrightarrow M' \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \xrightarrow{\pi} M \longrightarrow 0.$$

As in the proof of Theorem 4 we see that

$$\operatorname{reg}(M') \leq \operatorname{reg}(M) + \rho + 1, \quad \operatorname{w}(M') \leq \operatorname{w}(M) + \rho$$

and that the minimal number μ' of homogeneous generators of M' satisfies

$$\mu' \leq \binom{\operatorname{w}(M)+\rho+r}{r-1}\lambda\mu.$$

By Proposition 3 and as $\operatorname{beg}(M') \leq \operatorname{reg}(M')$, we thus have

$$\operatorname{reg}(M' \otimes_R N) \leq$$

$$\left[\operatorname{w}(M) + \operatorname{w}(N) + \rho + \lambda(\binom{\operatorname{w}(M)+\rho+r}{r-1}\lambda\mu\nu + 1) - 1\right]^{2^{r-1}} + \operatorname{reg}(M) + \operatorname{beg}(N) + \rho.$$

Next, look at the induced exact sequence

$$0 \longrightarrow \operatorname{Tor}_1^R(M, N) \longrightarrow M' \otimes_R N \xrightarrow{f} \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$$

and the two resulting short exact sequences

$$0 \longrightarrow \operatorname{Im}(f) \longrightarrow \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0,$$

$$0 \longrightarrow \operatorname{Tor}_1^R(M, N) \longrightarrow M' \otimes_R N \longrightarrow \operatorname{Im}(f) \longrightarrow 0.$$

If follows (see [7] Exercise 15.2.15) that

$$\operatorname{reg}(\operatorname{Im}(f)) \leq \max\{\operatorname{reg}(\bigoplus_{i=1}^{\mu} N(-a_i)), \operatorname{reg}(M \otimes_R N) + 1\},$$

$$\operatorname{reg}(\operatorname{Tor}_1^R(M, N)) \leq \max\{\operatorname{reg}(M' \otimes_R N), \operatorname{reg}(\operatorname{Im}(f)) + 1\}.$$

But $\operatorname{reg}(\bigoplus_{i=1}^{\mu} N(-a_i)) + 1 = \operatorname{gendeg}(M) + \operatorname{reg}(N) + 1 \leq \operatorname{reg}(M) + \operatorname{beg}(N) + \operatorname{w}(N)$ as well as $\operatorname{reg}(M \otimes_R N) + 2$ cannot exceed the previously given upper bound for $\operatorname{reg}(M' \otimes_R N)$ (see also Proposition 3 and observe that $r > 1$). Therefore we end up with the estimate

$$\operatorname{reg}(\operatorname{Tor}_1^R(M, N)) \leq$$

$$\left[\operatorname{w}(M) + \operatorname{w}(N) + \rho + \lambda(\binom{\operatorname{w}(M)+\rho+r}{r-1}\lambda\mu\nu + 1) - 1\right]^{2^{r-1}} + \operatorname{reg}(M) + \operatorname{beg}(N) + \rho.$$

This proves the case $i = 1$. Now, assume that $i > 1$. Then the isomorphism of graded R-modules

$$\mathrm{Tor}_i^R(M, N) \cong \mathrm{Tor}_{i-1}^R(M', N)$$

allow to proceed by induction as in the proof of Theorem 4. □

In case R is a polynomial ring, the upper bound of the previous theorem takes a simpler form. It follows because in this case $\rho = 0$.

Corollary 5. *Let $r > 1$, let R_0 be a local Artinian ring of length λ, and let M and N be two non-zero graded modules generated, respectively, by μ and ν homogeneous elements over the polynomial ring $R = R_0[x_1, \dots, x_r]$. Then we have*

$$\mathrm{reg}(\mathrm{Tor}_i^R(M, N)) \leq \left[\mathrm{w}(M) + \mathrm{w}(N) - 1\left(\lambda^i \mu\nu \binom{\mathrm{w}(M)+r}{r-1}^i + 1\right)\lambda\right]^{2^{r-1}} + \mathrm{beg}(N) + \mathrm{reg}(M).$$

Remark 6. As already observed above, the case $r = 1$ is not included in the previous two bounding results. But a look at the proof of Theorem 4 shows that for $r = 1$ we have the estimate

$$\mathrm{reg}(\mathrm{Tor}_i^R(M, N)) \leq \mathrm{w}(M) + \mathrm{reg}(M) + \mathrm{reg}(N) + 2i\rho + (\lambda^i \mu\nu + 1)\lambda + 1.$$

Up to now, the bounding results of this section where of *a priori type*, for example, valid without any further conditions on the Noetherian homogeneous ring R and the finitely generated graded R-modules M and N. We now follow the direction pointed out by earlier work of Caviglia and Eisenbud–Huneke–Ulrich and give a bound for the regularity of the modules $\mathrm{Tor}_k^R(M, N)$ under the additional condition that one of the modules M or N has finite projective dimension and that the modules $\mathrm{Tor}_i^R(M, N)$ are of dimension ≤ 1 for all $i \in \mathbb{N}$. We end up by generalizing the corresponding results of the mentioned authors (proved by them in case R is a polynomial ring over a field) to the case of homogeneous Noetherian rings R with Artinian base ring R_0 such that the singular locus of $\mathrm{Proj}(R)$ is a finite set (see Theorem 10).

Lemma 7. *Let $\rho = \mathrm{reg}(R)$. If M and N are finitely generated graded R-modules such that $p := \mathrm{pdim}_R(M) < \infty$ and $\dim_R(\mathrm{Tor}_i^R(M, N)) \leq 1$ for all $i > 0$, then it holds*

$$\mathrm{reg}(M \otimes_R N) \leq \mathrm{reg}(M) + \mathrm{reg}(N) + p\rho.$$

Proof. We proceed by induction on p. If $p = 0$, the graded R-module M is free, and our claim is obvious. So, let $p > 0$ and consider a graded short exact sequence

$$0 \longrightarrow M' \longrightarrow \bigoplus_{i=1}^{\mu} R(-a_i) \xrightarrow{\pi} M \longrightarrow 0$$

in which the homomorphism π is minimal. Look at the induced exact sequence

$$0 \longrightarrow \mathrm{Tor}_1^R(M,N) \longrightarrow M' \otimes_R N \stackrel{f}{\longrightarrow} \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$$

and the two resulting short exact sequences

$$0 \longrightarrow \mathrm{Im}(f) \longrightarrow \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$$

$$0 \longrightarrow \mathrm{Tor}_1^R(M,N) \longrightarrow M' \otimes_R N \longrightarrow \mathrm{Im}(f) \longrightarrow 0.$$

The first of these two sequences implies

$$\mathrm{reg}(M \otimes_R N) \leq \max\{\mathrm{reg}^1(\mathrm{Im}(f)) - 1, \mathrm{reg}(\bigoplus_{i=1}^{\mu} N(-a_i))\},$$

whereas the second of these sequences implies

$$\mathrm{reg}^1(\mathrm{Im}(f)) \leq \max\{\mathrm{reg}^2(\mathrm{Tor}_1^R(M,N)) - 1, \mathrm{reg}(M' \otimes_R N)\}.$$

As $\dim_R(\mathrm{Tor}_1^R(M,N)) \leq 1$, we have $\mathrm{reg}^2(\mathrm{Tor}_1^R(M,N)) = -\infty$. Hence, we obtain

$$\mathrm{reg}^1(\mathrm{Im}(f)) \leq \mathrm{reg}(M' \otimes_R N))$$

Therefore, we deduce that

$$\mathrm{reg}(M \otimes_R N) \leq \max\{\mathrm{reg}(M' \otimes_R N) - 1, \mathrm{reg}(\bigoplus_{i=1}^{\mu} N(-a_i))\}.$$

As $\mathrm{pdim}_R(M') = p - 1$ and $\mathrm{Tor}_j^R(M', N) \cong \mathrm{Tor}_{j+1}^R(M,N)$ for all $j \in \mathbb{N}$, the inductive hypothesis implies that $\mathrm{reg}(M' \otimes_R N) \leq \mathrm{reg}(M') + \mathrm{reg}(N) + (p-1)\rho$. Our initial graded short exact sequence yields that $\mathrm{reg}(M') \leq \mathrm{reg}(M) + \mathrm{reg}(R) + 1$. Moreover, $\mathrm{reg}(\bigoplus_{i=1}^{\mu} N(-a_i)) = a_\mu + \mathrm{reg}(N) \leq \mathrm{reg}(M) + \mathrm{reg}(N)$. Therefore

$$\begin{aligned}
\mathrm{reg}(M \otimes_R N) &\leq \max\{\mathrm{reg}(M' \otimes_R N) - 1, \mathrm{reg}(\bigoplus_{i=1}^{\mu} N(-a_i))\} \\
&\leq \max\{\mathrm{reg}(M') + \mathrm{reg}(N) + (p-1)\rho - 1, \mathrm{reg}(\bigoplus_{i=1}^{\mu} N(-a_i))\} \\
&\leq \mathrm{reg}(M) + \rho + 1 + \mathrm{reg}(N) + (p-1)\rho - 1 \\
&= \mathrm{reg}(M) + \mathrm{reg}(N) + p\rho.
\end{aligned}$$

□

Proposition 8. *Let* $\rho = \operatorname{reg}(R)$*. If* M *and* N *are finitely generated graded* R*-modules such that* $p := \operatorname{pdim}_R(M) < \infty$ *and* $\dim_R(\operatorname{Tor}_i^R(M,N)) \leq 1$ *for all* $i > 0$*. Then it holds*

$$\operatorname{reg}(\operatorname{Tor}_k^R(M,N)) \leq \operatorname{reg}(M) + \operatorname{reg}(N) + (k+1)p\rho + k \text{ for all } k \in \mathbb{N}_0.$$

Proof. The case $k = 0$ is clear by Lemma 7. To treat the cases with $k > 0$, we choose a short exact sequence of graded R-modules

$$0 \longrightarrow M' \longrightarrow \oplus_{i=1}^{\mu} R(-a_i) \xrightarrow{\pi} M \longrightarrow 0$$

in which π is minimal, such that $\operatorname{beg}(M) = a_1 \leq a_2 \leq \cdots \leq a_\mu = \operatorname{gendeg}(M)$. We proceed by induction on p. If $p = 0$ the module M is free and hence our claim is obvious. So, let $p > 0$ and consider the induced exact sequence of graded R-modules

$$0 \longrightarrow \operatorname{Tor}_1^R(M,N) \longrightarrow M' \otimes_R N \xrightarrow{f} \oplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$$

and the induced isomorphisms of graded R-modules

$$\operatorname{Tor}_k^R(M,N) \cong \operatorname{Tor}_{k-1}^R(M',N) \text{ for } k > 1.$$

As $\operatorname{pdim}_R(M') = p - 1$, these isomorphisms and the inductive hypotheses imply that

$$\operatorname{reg}(\operatorname{Tor}_k^R(M,N,)) \leq \operatorname{reg}(M') + \operatorname{reg}(N) + k(p-1)\rho + k - 1 \text{ for all } k > 1.$$

Our initial short exact sequence yields that $\operatorname{reg}(M') \leq \operatorname{reg}(M) + \rho + 1$. From this our claim follows for all $k > 1$. It thus remains to treat the case $k = 1$. The above exact sequence, induces two short exact sequences:

(1) $0 \longrightarrow \operatorname{Im}(f) \longrightarrow \bigoplus_{i=1}^{\mu} N(-a_i) \longrightarrow M \otimes_R N \longrightarrow 0$
(2) $0 \longrightarrow \operatorname{Tor}_1^R(M,N) \longrightarrow M' \otimes_R N \longrightarrow \operatorname{Im}(f) \longrightarrow 0$

As $\operatorname{reg}(\bigoplus_{i=1}^{\mu} N(-a_i)) = \operatorname{reg}(N) + \mu \leq \operatorname{reg}(M) + \operatorname{reg}(N)$, sequence (1) implies that

$$\operatorname{end}(H_{R_+}^0(\operatorname{Im}(f)) \leq \operatorname{reg}(M) + \operatorname{reg}(N).$$

As $\operatorname{pdim}_R(M') = p - 1$ and $\operatorname{reg}(M') \leq \operatorname{reg}(M) + \rho + 1$, it follows by Lemma 7 that

$$\operatorname{reg}(M' \otimes_R N) \leq \operatorname{reg}(M) + \operatorname{reg}(N) + p\rho + 1.$$

As $\dim(\operatorname{Tor}_1^R(M,N)) \leq 1$, sequence (2) yields an epimorphism of graded R-modules $H_{R_+}^1(M' \otimes_R N) \twoheadrightarrow H_{R_+}^1(\operatorname{Im}(f))$. Therefore

$$\operatorname{end}(H_{R_+}^1(\operatorname{Im}(f)) \leq \operatorname{reg}(M' \otimes_R N) - 1 \leq \operatorname{reg}(M) + \operatorname{reg}(N) + p\rho.$$

But now by sequence (2), we get

$$\operatorname{end}(H^1_{R_+}(\operatorname{Tor}^R_1(M,N)) \le \max\{\operatorname{end}(H^0_{R_+}(\operatorname{Im}(f))), \operatorname{end}(H^1_{R_+}(M' \otimes_R N))\} \le$$

$$\le \max\{\operatorname{reg}(M)+\operatorname{reg}(N), \operatorname{reg}(H^1(M' \otimes_R N))-1\} \le \operatorname{reg}(M)+\operatorname{reg}(N)+p\rho.$$

Another use of sequence (2) yields that

$$\operatorname{end}(H^0_{R_+}(\operatorname{Tor}^R_1(M,N))) \le \operatorname{reg}(M' \otimes_R N) \le \operatorname{reg}(M)+\operatorname{reg}(N)+p\rho+1.$$

As $\dim_R(\operatorname{Tor}^R_1(M,N)) \le 1$, it follows that

$$\operatorname{reg}(\operatorname{Tor}^R_1(M,N)) \le \operatorname{reg}(M)+\operatorname{reg}(N)+p\rho+1$$

and this proves our claim. □

Lemma 9. *Let $i \in \mathbb{N}$, $d \in \mathbb{N}_0$ and assume that the local ring $R_\mathfrak{p}$ is regular for all graded primes $\mathfrak{p} \subset R$ with $\dim(R/\mathfrak{p}) > d$. Let M and N be finitely generated graded R-modules such that $\dim_R(\operatorname{Tor}^R_i(M,N)) \le d$. Then it holds*

$$\dim_R(\operatorname{Tor}^R_j(M,N)) \le d \textit{ for all } j \ge i.$$

Proof. Let $\mathfrak{p} \subset R$ be a graded prime with $\dim(R/\mathfrak{p}) > d$. As $\dim_R(\operatorname{Tor}^R_i(M,N)) \le d$, it follows

$$\operatorname{Tor}^{R_\mathfrak{p}}_i(M_\mathfrak{p}, N_\mathfrak{p}) \cong \operatorname{Tor}^R_i(M,N)_\mathfrak{p} = 0.$$

The regular local ring $R_\mathfrak{p}$ contains the field $R_0/\mathfrak{m}_0$ and hence is unramified. So, by Auslander's Rigidity Theorem (see [1] Corollary 2.2), we have

$$\operatorname{Tor}^R_j(M,N)_\mathfrak{p} \cong \operatorname{Tor}^{R_\mathfrak{p}}_j(M_\mathfrak{p}, N_\mathfrak{p}) = 0 \text{ for all } j \ge i.$$

Therefore $\mathfrak{p} \notin \operatorname{Supp}_R(\operatorname{Tor}^R_j(M,N))$ for all $j \ge i$ and all graded primes $\mathfrak{p} \subset R$ with $\dim(R/\mathfrak{p}) > d$. As the R-modules $\operatorname{Tor}^R_j(M,N)$ are graded, our claim follows. □

Theorem 10. *Let $\rho := \operatorname{reg}(R)$ and assume that the local ring $R_\mathfrak{p}$ is regular for all graded primes $\mathfrak{p} \subset R$ with $\dim(R/\mathfrak{p}) \ge 2$. Let M and N be finitely generated graded R-modules such that $p = \operatorname{pdim}_R(M) < \infty$ and $\dim_R(\operatorname{Tor}^R_1(M,N)) \le 1$. Then it holds*

$$\operatorname{reg}(\operatorname{Tor}^R_k(M,N)) \le \operatorname{reg}(M)+\operatorname{reg}(N)+(k+1)p\rho+k \textit{ for all } k \in \mathbb{N}_0.$$

Proof. If we apply Lemma 9 with $d=1$ and $i=1$, we obtain that $\dim_R(\operatorname{Tor}^R_i(M,N)) \le 1$ for all $i > 0$. Now, our claim follows by Proposition 8. □

Corollary 11. *Let $r > 0$ and let $R = K[x_1, x_2, \ldots, x_r]$ be a polynomial ring over the field K. Let M and N be finitely generated graded R-modules such that* $\dim_R(\operatorname{Tor}_1^R(M, N)) \leq 1$. *Then it holds*

$$\operatorname{reg}(\operatorname{Tor}_k^R(M, N)) \leq \operatorname{reg}(M) + \operatorname{reg}(N) + k \textit{ for all } k \in \{0, 1, \ldots, r\}.$$

Proof. This is clear from Theorem 10 as $\operatorname{reg}(R) = 0$ and R is a regular ring. □

Remark 12. Corollary 11 has been proved by Eisenbud–Huneke–Ulrich (see[15] Corollary 3.1). The special case with $k = 0$ has been proved by Caviglia [11].
The conclusion of Theorem 10 need not hold if $\dim_R(\operatorname{Tor}_1^R(M, N)) > 1$, even in the special case where $R = K[x_1, x_2, \ldots, x_r]$ is a polynomial ring over the field K and for $k = 0$. Indeed Caviglia has constructed in this situation an example with $\dim_R(\operatorname{Tor}_1^R(M, N)) = 2$ and $\operatorname{reg}(M \otimes_R N) > \operatorname{reg}(M) + \operatorname{reg}(N)$.

Finally, we aim to conclude this section with slightly more geometric formulations of Theorem 10 and Corollary 11. To do so, we write

$$\operatorname{Sing}(X) := \{x \in X \mid \mathcal{O}_{X,x} \text{ is not regular}\}$$

for the *singular locus* of the Noetherian scheme X. If $\mathcal{H}$ is a coherent sheaf of $\mathcal{O}_X$-modules, we write

$$\operatorname{Sing}(\mathcal{H}) := \{x \in X \mid \mathcal{H}_x \text{ is not free over } \mathcal{O}_{X,x}\}$$

for the set of all points $x \in X$ at which the stalk $\mathcal{H}_x$ of $\mathcal{H}$ in x is not free.

Corollary 13. *Let $\rho := \operatorname{reg}(R)$, and set $X := \operatorname{Proj}(R)$. Let M and N be finitely generated graded R-modules such that $p = \operatorname{pdim}_R(M) < \infty$. Let $\mathcal{F} := \widetilde{M}$ and $\mathcal{G} := \widetilde{N}$ be the coherent sheaves of $\mathcal{O}_X$-modules induced, respectively, by M and N. Assume that the sets* $\operatorname{Sing}(X)$ *and* $\operatorname{Sing}(\mathcal{F}) \cap \operatorname{Sing}(\mathcal{G})$ *are finite. Then it holds*

$$\operatorname{reg}(\operatorname{Tor}_k^R(M, N)) \leq \operatorname{reg}(M) + \operatorname{reg}(N) + (k + 1)p\rho + k \textit{ for all } k \in \mathbb{N}_0.$$

Proof. The finiteness of the singular locus of X implies that $R_{\mathfrak{p}}$ is a regular local ring for all graded primes $\mathfrak{p} \subset R$ with $\dim(R/\mathfrak{p}) \geq 2$. Our hypothesis on the stalks of $\mathcal{F}$ and $\mathcal{G}$ imply that at least one of the two finitely generated $R_{\mathfrak{p}}$-modules $M_{\mathfrak{p}}$ or $N_{\mathfrak{p}}$ is free for each graded prime $\mathfrak{p} \subset R$ with $\dim(R/\mathfrak{p}) \geq 2$. Therefore $\operatorname{Tor}_1^R(M, N)_{\mathfrak{p}} \cong \operatorname{Tor}_1^{R_{\mathfrak{p}}}(M_{\mathfrak{p}}, N_{\mathfrak{p}}) = 0$ for all such $\mathfrak{p}$—and hence $\dim_R(\operatorname{Tor}_1^R(M, N)) \leq 1$. Now, we get our claim by Theorem 10. □

To formulate Corollary 11 in geometric terms, we recall a few notions from sheaf cohomology.

Reminder 14. (See Chap. 20 of [7] for example.) Let $X := \operatorname{Proj}(R)$ and let $\mathcal{H}$ be a coherent sheaf of $\mathcal{O}_X$-modules. Then, the *(Castelnuovo–Mumford) regularity* of $\mathcal{H}$ is defined as

$$\operatorname{reg}(\mathcal{H}) := \inf\{r \in \mathbb{Z} \mid H^i(X, \mathcal{H}(r-i)) = 0 \text{ for all } i > 0\},$$

where $H^i(X, \mathcal{H}(n))$ denotes the ith *sheaf cohomology* group of (X with coefficients in) the *nth twist* $\mathcal{H}(n) := \mathcal{H} \otimes_{\mathcal{O}_X} \mathcal{O}(n)$ of $\mathcal{H}$.
The *total group of sections* of $\mathcal{H}$ is defined by

$$\Gamma_*(\mathcal{H}) := \bigoplus_{n \in \mathbb{Z}} H^0(X, \mathcal{H}(n))$$

and carries a natural structure of graded R-module. Moreover, the sheaf $\widetilde{\Gamma_*(\mathcal{H})}$ of $\mathcal{O}_X$-modules induced by the graded R-module $\Gamma_*(\mathcal{H})$ coincides with $\mathcal{H}$. Finally, the R-module $\Gamma_*(\mathcal{H})$ is finitely generated and only if the set $\operatorname{Ass}_X(\mathcal{H})$ contains no closed points of X—and if this is the case, we have

$$\operatorname{reg}(\mathcal{H}) = \operatorname{reg}(\Gamma_*(\mathcal{H})).$$

Corollary 15. *Let $r \in \mathbb{N}$, let K be a field, and let $\mathcal{F}$ and $\mathcal{G}$ be two coherent sheaves of $\mathcal{O}_{\mathbb{P}^r_K}$-modules such that the set* $\operatorname{Ass}_{\mathbb{P}^r_K}(\mathcal{F}) \cup \operatorname{Ass}_{\mathbb{P}^r_K}(\mathcal{G})$ *contains no closed points and the set* $\operatorname{Sing}(\mathcal{F}) \cap \operatorname{Sing}(\mathcal{G})$ *is finite. Then it holds*

$$\operatorname{reg}\big(\operatorname{Tor}_k^{\Gamma_*(\mathcal{O}_{\mathbb{P}^r_K})}(\Gamma_*(\mathcal{F}), \Gamma_*(\mathcal{G}))\big) \le \operatorname{reg}(\mathcal{F}) + \operatorname{reg}(\mathcal{G}) + k \text{ for all } k \in \{0, 1, \ldots, r+1\}.$$

Proof. Consider the polynomial ring $R := K[x_0, x_1, \ldots, x_r]$ and write $\mathbb{P}^r_K = \operatorname{Proj}(R)$. As $r > 0$ we have $\Gamma_*(\mathcal{O}_{\mathbb{P}^r_K}) = R$. According to Reminder 14, the graded R-modules $\Gamma_*(\mathcal{F})$ and $\Gamma_*(\mathcal{G})$ are finitely generated and induce, respectively, the coherent sheaves $\mathcal{F}$ and $\mathcal{G}$. Now, we get our claim by Corollary 13 and Reminder 14. □

Acknowledgements The first author expresses his gratitude toward David Eisenbud for the many things he could learn either from David's published work, David's conference talks, or from a number of inspiring personal discussions with David, for example, at the Forschungsinstitut Oberwolfach, at Brandeis University, on a mountain walk in les Plans sur Bex, on a Winter excursion to Arosa, and at several other occasions—discussions which were of great influence on the first author's work. The second author was partially supported by NAFOSTED (Vietnam).

References

1. Auslander, M.: Modules over unramified local rings. Illinois J. Math. **5**, 631–647 (1961)
2. Bächtold, M.: Fold-type solution singularities and characteristic varieties of non-linear PDEs. Ph.D. Dissertation, Institute of Mathematics, University of Zürich (2009)
3. Bayer, D., Stillman, M.: A criterion for detecting m-regularity. Invent. Math. **87**, 1–11 (1987)
4. Bayer, D., Mumford, M.: What can be computed in algebraic geometry? In: Eisenbud, D., Robbiano, L. (eds.) Computational Algebraic Geometry and Commutative Algebra. Proceedings of Cortona (1991), pp. 1–48. Cambridge University Press, Cambridge (1993)

5. Boldini, R.: Finiteness of leading monomial ideals and critical cones of characteristic varieties. Ph.D. Dissertation, Institute of Mathematics, University of Zürich (2012)
6. Brodmann, M.: Castelnuovo–Mumford regularity and degrees of generators of graded submodules. Illinois J. Math. **47**(3), 749–767 (2003)
7. Brodmann, M., Sharp, R.Y.: Local cohomology—an algebraic introduction with geometric applications. In: Cambridge Studies in Advanced Mathematics, vol. 60. Cambridge University Press, Cambridge (1998)
8. Brodmann, M., Götsch, T.: Bounds for the Castelnuovo–Mumford regularity. J. Commutative Algebra **1**(2), 197–225 (2009)
9. Brodmann, M., Jahangiri, M., Linh, C.H.: Castelnuovo–Mumford regularity of deficiency modules. J. Algebra **322**, 12816–12838 (2009)
10. Castelnuovo, G.: Sui multipli di una serie lineare di gruppi di punti appartente ad una curva algebrica. Rend. circ. Math. Palermo **7**, 89–110 (1893)
11. Caviglia, G.: Bounds on the Castelnuovo–Mumford regularity of tensor products. Proc. AMS **135**, 1449–1957 (2007)
12. Chardin, M.: Some results and questions on Castelnuovo–Mumford regularity. In: Syzygies and Hilbert Functions. Lecture Notes in Pure and Applied Mathematics, vol. 254, pp. 1–40. Chapman & Hall/CRC, Boca Raton (2007)
13. Chardin, M., Fall, A.L., Nagel, U.: Bounds for the Castelnuovo–Mumford regularity of modules. Math. Z. **258**, 69–80 (2008)
14. Coutinho, S.C.: A primer of algebraic D-modules. In: London Mathematical Society Student Texts, vol. 33. Cambridge University Press, Cambridge (1995)
15. Eisenbud, D., Huneke, C., Ulrich, B.: The regularity of Tor and graded Betti Numbers. Amer. J. Math. **128**(3), 573–605 (2006)
16. Gruson, L., Lazarsfeld, R., Peskine, C.: On a theorem of Castelnuovo and the equations defining space curves. Inventiones Math. **72**, 491–506 (1983)
17. Hoa, L.T., Hyry, E.: Castelnuovo–Mumford regularity of canonical and deficiency modules. J. Algebra **305**(2), 877–900 (2006)
18. Li, H., van Oystaeyen, F.: Zariski filtrations. In: K-Monographs in Mathematics, vol. 2. Kluwer, Dordrecht (1996)
19. McConnell, J.C., Robson, J.C.: Noncommutative noetherian rings. In: AMS Graduate Studies in Mathematics, vol. 30. American Mathematical Society, Providence (2001)
20. Mumford, D.: Lectures on curves on an algebraic surface. In: Annals of Mathematics Studies, vol. 59. Princeton University Press, Princeton (1966)
21. Seiler, M.-H.: Castelnuovo–Mumford regularity of annihilators. Diploma Thesis, Institute of Mathematics, University of Zürich, Zürich (2010)

Selections from the Letter-Place Panoply

David A. Buchsbaum

This paper is dedicated to David Eisenbud, mentee, mentor and, above all, dear friend.

1 Introduction

There is a fairly extensive literature on letter-place algebras, but mostly for the edification of those working in algebraic combinatorics (see, for instance, [15, 16, 18, 20]). I don't think that letter-place algebras have had much play yet in commutative and homological algebra, so I thought I'd talk about them here and perhaps arouse a bit of interest in that subject.

I was introduced to the letter-place notion by Gian–Carlo Rota, with whom I had the pleasure of collaborating for almost a decade. He attributed it to a physicist, I believe to Feynman. The fact that letter-place techniques, including place polarizations, could simplify a good deal of the work Akin and I had been doing earlier on resolving Weyl modules [1, 2], appealed to both of us, and we decided to use them to push ahead to find the projective resolutions of those representations. Much of this is spelled out in great detail in [7], and in a fairly long article written with Rota (posthumously) [13], we focused on the resolutions themselves.

In Sect. 2, we will see some examples to show how representation theory first reared its head for Eisenbud and me in our joint work on generalized Koszul complexes [8, 10]. From there we discuss Lascoux's use of classical representation theory to describe the terms of the resolutions of determinantal ideals and of Schur modules [21]. This led to the development of characteristic-free representation theory of the general linear group, to the general definition of Schur and Weyl modules (which are in a precise sense dual to each other; this will be explained

D.A. Buchsbaum (✉)
Brandeis University, Hue City, Viet Nam
e-mail: buchsbau@brandeis.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_7,

in Sect. 5.8), and an attempt by Akin, Weyman, and myself to replicate the Lascoux results for determinantal ideals in characteristic-free form [4,5]. An interesting early development along those lines was the discovery of the role $\mathbb{Z}$-forms play in these results, and this led Akin and me to study the resolutions of Weyl modules in a serious systematic way [2].

In Sect. 3, we will give a precise definition of Schur and Weyl modules, so that the terms used above will make sense; these are, in fact, the major objects of study in the rest of this article.

In Sect. 4 we will consider resolutions of two-rowed Weyl modules associated to skew-shapes, and see the letter-place techniques in play. In particular, we'll see how use of letter-place enables us to define a splitting homotopy for these resolutions [11].

In Sect. 5, we'll take the bull by the horns and give the general definition of letter-place algebras, indicate the proof of the basis theorem for them, and discuss the "straight basis" theorem for Weyl modules (based on Taylor's work [23]). Limits of space and typography will make it impossible to describe here in detail the terms of the resolutions of Weyl modules in general. However, the book [7], has all of that spelled out. Space constraints also constrain us to omit all detailed proofs of basis theorems although an indication of the proof will be given where possible. However, in Sect. 6 we give outlines of proofs of some of the more complex results.

2 Some Background

Many years ago, a family of complexes was introduced (see [8]) which were generalizations of the usual Koszul complex. They were introduced and studied for a number of reasons: we wanted to generalize the usual Koszul complex for purposes of grade sensitivity and generalized multiplicity [8]; we wanted to apply them to the Grothendieck Lifting Problem [9]; we wanted to apply them to the problem of resolving the ideals generated by the minors of a generic matrix, that is, determinantal ideals [4].

Further work with these complexes led to a "trimming down" of the terms, and this introduced certain representations, called "hooks", into the picture.

From this hint of representation theory arising in resolutions, we move to Lascoux's use of classical representation theory to describe the terms of resolutions of determinantal ideals. From there, we're led to the development of the classical theory in a characteristic-free context, and to the emulation of the Lascoux resolutions in that context. Problems then arose with certain integral representations, called $\mathbb{Z}$-forms, whose general study led to the problem of resolving fairly arbitrary Weyl and Schur modules.

2.1 Generalized Koszul Complexes

The following family of complexes was introduced in [8].

Let $F = R^m$ and $G = R^n$ $(m \geq n)$, and take a map $f : F \to G$. For each integer k, with $1 \leq k \leq n$, we associate a complex related to the map $\Lambda^k f : \Lambda^k F \to \Lambda^k G$ (we'll denote it by $\mathbf{C}(k; f)$) as follows:

$$0 \to C^k_{m-n+1} \to \cdots \to C^k_q \to \cdots \to \sum_{s_i \geq 1} \Lambda^{n-k+s_0} G^* \otimes \Lambda^{s_1} G^* \otimes \Lambda^{n+|s|} F \to$$

$$\sum_{s \geq 1} \Lambda^{n-k+s} G^* \otimes \Lambda^{n+s} F \to \Lambda^k F \to \Lambda^k G,$$

where

$$C^k_q = \sum_{s_i \geq 1} \Lambda^{n-k+s_0} G^* \otimes \Lambda^{s_1} G^* \otimes \cdots \otimes \Lambda^{s_{q-2}} G^* \otimes \Lambda^{n+|s|} F, \quad q \geq 2,$$

$|s| = \sum s_i$, and the maps (except for $\Lambda^k f : \Lambda^k F \to \Lambda^k G$) are the bar complex maps associated to the action of the algebra ΛG^* on ΛF.

At about the same time these complexes were defined, another—and much more efficient—complex was developed by Eagon and Northcott [17] which was associated to the map $\Lambda^n f$. This raised the following question: how is this complex related to $\mathbf{C}(n; f)$?

A quick look at the map

$$\sum_{s \geq 1} \Lambda^{n-k+s} G^* \otimes \Lambda^{n+s} F \to \Lambda^k F$$

tells us that its image is the same as that of the map restricted to just the one summand:

$$\Lambda^{n-k+1} G^* \otimes \Lambda^{n+1} F \to \Lambda^k F.$$

The reason for throwing in all the extra summands is that the bar construction involves multiplication in the algebra ΛG^*, and the extra terms are there to "catch" terms as they come flying in from

$$\sum_{s_i \geq 1} \Lambda^{n-k+s_0} G^* \otimes \Lambda^{s_1} G^* \otimes \Lambda^{n+|s|} F.$$

In short, if we could replace all the summands here by

$$\operatorname{Ker}(\Lambda^{n-k+1} G^* \otimes \Lambda^1 G^* \otimes \Lambda^{n+2} F \to \Lambda^{n-k+2} G^* \otimes \Lambda^{n+2} F),$$

(call it, for the moment, $K_{(1^{n-k+1},2)}G^*$) we could start slimming down our complex so that it starts out looking like

$$K_{(1^{n-k+1},2)}G^* \otimes \Lambda^{n+2}F \to \Lambda^{n-k+1}G^* \otimes \Lambda^{n+1}F \to \Lambda^k F \to \Lambda^k G.$$

2.2 Hooks

Here is where the first hint of representation theory appears, for the modules $K_{(1^{n-k+1},2)}G^*$ are representations known as "hooks". To make sense of all this, we'll take a short detour.

We're all familiar with the classical family of Koszul-type complexes (one for each q):

$$0 \to \Lambda^q F \to S_1 \otimes \Lambda^{q-1}F \to \cdots \to S_{q-l} \otimes \Lambda^l F \to \cdots \to S_{q-1} \otimes \Lambda^1 F \to S_q \to 0$$

where F is a given free R-module, and S_j stands for the symmetric power, $S_j F$. We'll call this complex $\Lambda^q(F)$.

If we take the "dual" of this complex, replacing the symmetric powers by divided powers (and omitting the asterisk), we obtain the complex:

$$0 \to D_q \to D_{q-1}(F) \otimes \Lambda^1 F \to \cdots \to D_{q-l} \otimes \Lambda^l F$$
$$\to \cdots \to D_1 \otimes \Lambda^{q-1}F \to \Lambda^q F \to 0.$$

We'll call this complex $\mathbf{D}_q(F)$.

It's important to notice that while the boundary map in $\Lambda^q(F)$ entails diagonalization in the exterior algebra and multiplication in the symmetric algebra of F, the boundary in $\mathbf{D}_q(F)$ is given by diagonalization in the divided power algebra and multiplication in the exterior algebra of F. Also, except for $q = 0$, both complexes are exact.

Now all the modules involved here are representations of $GL(F)$, the general linear group of F (or, to be very concrete, the group of invertible $n \times n$ matrices over R, where n is the rank of F), and all the maps are equivariant. So the cycles (which are the same as the boundaries) of these complexes are also representations of $GL(F)$.

Definition 1. We define the Weyl and Schur hooks as follows:

(a) The kernel of the map $D_p F \otimes \Lambda^l F \to D_{p-1}F \otimes \Lambda^{l+1}F$
is denoted by $\mathbf{K_{(1^l,p+1)}F}$; this is the **Weyl hook**.

(b) The kernel of the map $S_p F \otimes \Lambda^l F \to S_{p+1}F \otimes \Lambda^{l-1}F$
is denoted by $\mathbf{L_{(1,1^{p-1})})F}$; this is the **Schur hook**.

When $p = 1$, we have our hook, $K_{(1^l,2)}$ above, and we see that when $p = 0$, we have $K_{(1^l,1)} = \Lambda^l$.

These observations led Eisenbud and me [10] to construct another family of complexes which were associated to the maps $L_{(k,1^q)}(f) : L_{(k,1^q)}R^m \to L_{(k,1^q)}R^n$ induced on these hooks from the map f. In particular, for $q = 0$, we had complexes associated to $\Lambda^k f$ for all $1 \leq k \leq n$ (which we will denote by $\mathbf{T}(k; f)$), and for $k = n$, this was just the Eagon–Northcott complex mentioned above. As was the case of the Eagon–Northcott complex, the ones in [10] were much slimmer than the corresponding complexes constructed earlier in [8].[1] In [6], the two families of complexes were shown to be homotopically equivalent.

2.3 Determinantal Ideals

One of the main motivations for constructing these families of complexes was to try to find resolutions of the ideal generated by the minors of a (generic) matrix corresponding to a map $f : F \to G$. We've already noted that the families we constructed gave resolutions of a class of modules all of which have the ideal of maximal minors of the given map as support, and for certain values of the parameters, actually provided the resolution of the ideal of maximal minors itself. But it was still an open problem to find a resolution of the ideal generated by minors of any given order.

While it was already apparent way back in the 1960s that the modules that would comprise such a resolution were representations of the product of general linear groups, $GL(F) \times GL(G)$, it wasn't until Lascoux [21] tackled the problem in characteristic zero that it became clear just which representation modules they were, namely, the direct sum of tensor products of certain $GL(F)$-Schur modules with $GL(G)$-Schur modules (see Sect. 3 for definitions of these modules). And he not only defined those resolutions, he defined the resolutions of certain classes of Schur and Weyl modules as well, but always in characteristic zero.[2] This naturally led to the question whether it was possible to do in the characteristic-free case what he had accomplished in the classical case.

The upshot is that it was possible to define the various representation modules over an arbitrary commutative ring that arose in the Lascoux results (see [1, 2, 4, 5]), as well as an even larger class that was necessary for homological purposes.[3] But when we tried to replicate Lascoux's resolutions, we ran into a few snags, most of which revolved around the issue of $\mathbb{Z}$-forms (see below). Akin, Weyman, and I

[1] This explains the notation $\mathbf{C}(k; f)$ and $\mathbf{T}(k; f)$: the **C** stands for "corpulent", while the **T** stands for "thin".

[2] It should be added that while Lascoux indicated what the boundary maps in these resolutions might be, it took a bit of time before they were explicitly described (for the determinantal ideals), and it's still an open problem to describe the boundary maps of the Lascoux resolutions of the Schur and Weyl modules.

[3] Previous work on characteristic-free representation theory had been done in [14, 24], but the categories of modules were too small for our purposes.

did succeed in describing the terms and maps of a characteristic-free resolution of the ideal of submaximal minors [4], but we could go no further. Not too surprising when one considers the fact that several years later, Hashimoto [19] proved the non-existence of a universal characteristic-free resolution of the ideal of minors of lower degree.

2.4 $\mathbb{Z}$-Forms

Strange $\mathbb{Z}$-forms arose even in the case of submaximal minors, but it was possible to handle these. But before going any further, we'll define what we mean by $\mathbb{Z}$-forms.

Let F be a free abelian group of rank m, and let $\overline{F}$ be its extension to the rationals, that is, $\overline{F} = \mathbb{Q} \otimes_{\mathbb{Z}} F$. We know that D_2F and S_2F are both $GL(F)$-representations, where $GL(F)$ means the general linear group over the integers, $\mathbb{Z}$. Furthermore, $D_2\overline{F} = \mathbb{Q} \otimes_{\mathbb{Z}} D_2F$, $S_2\overline{F} = \mathbb{Q} \otimes_{\mathbb{Z}} S_2F$, and these are $GL(\overline{F})$-representations. We know that there is a $GL(F)$-equivariant map from D_2F to S_2F, namely, the composition

$$D_2F \xrightarrow{\Delta} F \otimes_{\mathbb{Z}} F \xrightarrow{\mathrm{m}} S_2F,$$

where Δ is the diagonal map and m is the usual multiplication map. However, this is not an isomorphism of the two integral representations. Nevertheless, the corresponding map over the rationals is an isomorphism. We therefore say that D_2F and S_2F are $\mathbb{Z}$-forms of the same rational representation (in this case, $S_2\overline{F}$). So, we are led to make the following definition.

Definition 2. Let F be a free abelian group. Two $GL(F)$-representations are $\mathbb{Z}$-**forms of the same representation** if, when tensored with the rationals, $\mathbb{Q}$, they are isomorphic $GL(\overline{F})$-representations.

We've indicated that in the construction of the resolution of the ideal of submaximal minors, certain of the terms in the Lascoux resolution had to be replaced by their $\mathbb{Z}$-forms in order to get an integral complex which was acyclic. These representations that arose were $\mathbb{Z}$-forms of certain hooks, and they came about in the following way.

We know that for all $l > 0$ the complexes

$$0 \to \Lambda^l F \to \Lambda^{l-1} F \otimes_{\mathbb{Z}} F \to \cdots \to \Lambda^{l-t} F \otimes_{\mathbb{Z}} S_t F \to \cdots \to F \otimes_{\mathbb{Z}} S_{l-1} F \to S_l F \to 0$$

are exact, where the boundary map is given by diagonalizing the exterior powers and multiplying the symmetric powers. But what happens if we replace the symmetric powers by divided powers, that is, if we consider the complex

$$0 \to \Lambda^l F \to \Lambda^{l-1} F \otimes_{\mathbb{Z}} F \to \cdots \to \Lambda^{l-t} F \otimes_{\mathbb{Z}} D_t F \to \cdots \to F \otimes_{\mathbb{Z}} D_{l-1} F \to D_l F \to 0$$

where we still diagonalize the exterior powers and multiply, this time, into the divided powers (something that we generally avoid doing)? As the reader may strongly suspect, this complex is no longer exact; the surprising thing, however, is that, counting from the left, it is exact up to the middle of the complex, that is, from $t = 0$ to $t = [\frac{l-1}{2}]$ where $[x]$ indicates the integral part of x ([4], Proposition 2.22). As a result, the cycles of this complex are $\mathbb{Z}$-forms of the corresponding cycles of the complex above, involving the symmetric powers, and these, as we know, are just the hooks.

Another, simpler, way to construct non-isomorphic $\mathbb{Z}$-forms is the following: Consider the short exact sequence

$$(\dagger) \quad 0 \to D_{k+2} \to D_{k+1} \otimes_{\mathbb{Z}} D_1 \to K_{(k+1,1)} \to 0$$

where $K_{(k+1,1)}$ is the Weyl hook, as defined earlier. (We are leaving out the module F, as that is understood throughout.)

If we take an integer, t, and multiply D_{k+2} by t, we get an induced exact sequence and a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & D_{k+2} & \to & D_{k+1} \otimes_{\mathbb{Z}} D_1 & \to & K_{(k+1,1)} & \to & 0 \\ & & \downarrow t & & \downarrow & & \downarrow & & \\ 0 & \to & D_{k+2} & \to & E(t; k+1, 1) & \to & K_{(k+1,1)} & \to & 0, \end{array}$$

where $E(t; k+1, 1)$ stands for the cofiber product of D_{k+2} and $D_{k+1} \otimes_{\mathbb{Z}} D_1$. Each of these modules is a $\mathbb{Z}$-form of $D_{k+1} \otimes_{\mathbb{Z}} D_1$, but for t_1 and t_2, two such are isomorphic if and only if $t_1 \equiv t_2 \bmod k+2$ (see[3]). This says that $\mathrm{Ext}^1_A(K_{(k+1,1)}, D_{k+2}) \cong \mathbb{Z}/(k+2)$, where A is the Schur algebra of appropriate degree. We should explain that the Schur algebra is the universal enveloping algebra of $GL(F)$, that is, the $GL(F)$-(polynomial) representations of degree n are modules over the Schur algebra of degree n (see [1] or [7] for a complete definition of this algebra).

The above observations should give an idea of why Akin and I turned to the study of resolutions of Weyl and Schur modules.

3 Weyl and Schur Modules

We've already had a very small dose of representation theory of the general linear group in our discussion of the hook shapes. In this section, we will deal more comprehensively with representations of that group, over an arbitrary commutative ring, R. To start, we have to talk about shape matrices and tableaux.

3.1 Shape Matrices and Tableaux

In the classical theory, the fundamental shapes that are the basis of definition of Schur and Weyl modules, are the Ferrers diagrams corresponding to "partitions", and the closely related "skew-partitions". For our purposes, we will have to consider a slightly larger class of shapes, corresponding to the so-called "almost skew-partitions".[4]

Definition 1. A **shape matrix** is an infinite integral matrix $A = \left(a_{ij}\right)$ of finite support, with all the a_{ij} equal to 0 or 1. (To say it has finite support is to say that $a_{ij} \neq 0$ for only a finite number of indices i and j.) The **last row (column)** of the shape matrix A is the last row (column) in which a non-zero term appears. Such a matrix is said to be **row-convex (column-convex)** if, in each row (column), there are no zeroes lying between ones. *(All the shapes that we consider will be row- and column-convex.)* The shape matrix $B = \left(b_{ij}\right)$ is a **subshape** of A (written $B \subseteq A$) if $b_{ij} \leq a_{ij}$ for every i and j. The shape matrix, A, is said to correspond to a **partition** if for all i, j, $a_{ij} = 0$ implies $a_{i+1j} = 0$ and $a_{ij+1} \neq 0$ implies $a_{ij} \neq 0$. It is said to correspond to a **skew-partition** or **skew-shape** if $A = B - C$, where B and C correspond to partitions, and $C \subseteq B$. It is said to be a **bar shape** if its only non-zero entries are in its last row, and it is row-convex. Finally, it is said to correspond to an **almost skew-partition** or **almost skew-shape** if $A = B - C$, where B corresponds to a skew-shape, and C is a bar subshape matrix of B the index of whose last row coincides with that of B, and whose first nonzero entry in that row occurs in the same place as the first nonzero entry of B.

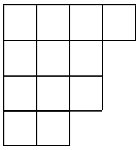
- Notice that, unless a given partition shape matrix is the zero matrix, $a_{11} = 1$.

We now illustrate each of these types of shape matrices. The typical partition shape looks like this:

$$(P) \qquad \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0\cdots \\ 1 & 1 & 1 & 0 & 0 & 0\cdots \\ 1 & 1 & 1 & 0 & 0 & 0\cdots \\ 1 & 1 & 0 & 0 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0\cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\cdots \end{pmatrix},$$

[4] We should add that the class of shapes that is currently being studied is far broader than this. However, to study all of these would require quite a bit more of combinatorics than we propose to talk about here.

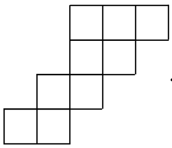
and is often represented by the Ferrers diagram:



The typical skew-shape looks like this

$$(S) \quad \begin{pmatrix} 0 & 0 & 1 & 1 & 1 & 0\cdots \\ 0 & 0 & 1 & 1 & 0 & 0\cdots \\ 0 & 1 & 1 & 0 & 0 & 0\cdots \\ 1 & 1 & 0 & 0 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0\cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\ldots \end{pmatrix},$$

and is often represented by the Ferrers diagram:



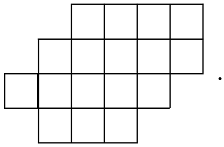
A bar shape looks like:

$$(B) \quad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0\cdots \\ 0 & 1 & 1 & 1 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0\cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\ldots \end{pmatrix},$$

and an almost skew-shape looks like

$$(AS) \quad \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0\cdots \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0\cdots \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0\cdots \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0\cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\ldots \end{pmatrix},$$

and would be represented by the Ferrers diagram:



Note that the diagram ignores the fact that the left-most column of the shape matrix consists completely of zeroes. Also notice that the empty diagram corresponds to the zero matrix.

The shapes illustrated above are those that we will have most to do with, but clearly we can associate to any shape matrix, $A = (a_{ij})$, a Ferrers diagram, or simply diagram: we just set up a grid equal to the effective size of the matrix (say, $s \times t$), and throw away the boxes whose entries are equal to zero. That is, the (i, j)th box lies in the diagram if and only if $a_{ij} = 1$. If A is the shape matrix, we will sometimes denote by (A) its corresponding diagram.

The partition shape has a uniquely associated partition, namely, the sequence $\lambda = (\lambda_1, \ldots, \lambda_s, \ldots)$ where λ_i is the non-negative integer equal to the number of ones in row i. Clearly, the sequence is decreasing: $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda s \geq \cdots$. We say the **length** of λ is s if s is the smallest nonnegative integer such that $\lambda_{s+t} = 0$ for every positive integer t.

The skew-shape has two partitions uniquely associated to it, namely, $\lambda = (\lambda_1, \ldots, \lambda_s, \ldots)$ and $\mu = (\mu_1, \ldots, \mu_t, \ldots)$, with $\mu_i \leq \lambda_i$ for all i, and such that the length of μ is strictly less than that of λ. One then thinks of the shape as the result of removing from the shape of λ the subshape corresponding to μ. In fact, the notation most often used for a skew-shape is λ/μ. We say the **length** of λ/μ is the length of λ. If one removes the condition that the length of μ be strictly less than that of λ, then we have other pairs of partitions (λ', μ') that will yield the same diagram. In that case, we still use the same notation, λ'/μ' (but the length of λ'/μ' stays equal to the length of the previous λ).

Finally, we see that the almost skew-shape would be a skew-shape but for its last row, which, rather than projecting beyond (or flush with) the penultimate row, doesn't make it out that far to the left. In short, it would be a skew-partition but for that inadequacy in the last row.

In our examples above, the partition λ associated to (P) is $(4, 3, 3, 2)$; the pair of partitions associated to (S) are $\lambda = (5, 4, 3, 2)$ and $\mu = (2, 2, 1)$. (For convenience we have eliminated the zeroes to the right in our notation.) For (AS), we might take the skew-partition to be $(7, 7, 6, 5)/(3, 2, 1)$ with a bar having the entries $(1, 1)$ in the fourth row, or we might take $(7, 7, 6, 5)/(3, 2, 1, 1)$, with a bar having entries $(0, 1)$ in the fourth row.

Another way to denote an almost skew-shape, which closely parallels the notation for a skew-shape, is to first define an **almost partition** to be a sequence $(\mu_1, \ldots, \mu_n)$ such that $\mu_1 \geq \cdots \geq \mu_{n-1}$ and $0 \leq \mu_n \leq \mu_1$. We then can denote (not necessarily uniquely) an almost skew-shape by λ/μ, where λ is a partition of length n, and μ is an almost partition having exactly n terms and satisfying $\mu_i \leq \lambda_i$ for all i. We can go one step further and say that an almost partition, μ, is **of type**

$n-(i+1)$ if i is the largest integer less than n such that $\mu_n \leq \mu_i$, and we say that the type of the almost skew-shape λ/μ is equal to the type of μ. Clearly, the type is independent of the choice of λ and μ used to describe the almost skew-shape. The choice of the pair (λ, μ) can be made canonical if, in the case of type zero, we choose $\mu_n = 0$, while for type greater than zero, we choose $\mu_{n-1} = 0$. We define the **length** of the almost skew-shape to be the length of the canonical partition, λ.

With this terminology, we see that an almost skew-shape of type 0 is a skew-shape and that for almost skew-shapes of length n, we can have types $0, 1, \ldots, n-2$. In particular, almost skew-shapes of length 2 are necessarily skew-shapes; for length 3, there are only skew-shapes and almost skew-shapes of type 1 and so on.

We spoke of shape matrices as infinite matrices in order not to have to specify last row or column when we talked about subshapes. However, we see that a shape matrix whose last row is row s and whose last column is column t, can be thought of as an $s \times t$-matrix; when we draw diagrams or shapes, we will generally avoid the dots that we were forced to put into the illustrations above. If we have two shape matrices A and B, with $B \subseteq A$, we will assume that they're both $s \times t$-matrices, simply by augmenting where necessary by zeroes.

Remark 2. Three immediate observations should be made here:

(1) If A is a partition (or skew-partition) matrix, then its transpose, $\widetilde{A}$ is also a partition (or skew-partition) matrix.
(2) If A is a partition matrix with associated partition λ, then we denote the partition associated to $\widetilde{A}$ by $\widetilde{\lambda}$.
(3) We see that if A is a skew-partition matrix with associated partitions λ and μ, then the matrix $\widetilde{A}$ has associated partitions $\widetilde{\lambda}$ and $\widetilde{\mu}$.

Definition 3. We introduce some standard terminology for shapes, and partitions in particular:

(1) The **weight** of a shape matrix $A = (a_{ij})$ is $\sum a_{ij}$, and is denoted by $|A|$.
(2) If $\lambda = (\lambda_1, \ldots, \lambda_n)$ and $\lambda' = (\lambda'_1, \ldots, \lambda'_m)$ are two partitions, we say $\lambda \geq \lambda'$ if either $\lambda = \lambda'$ or if for some i, $\lambda_j = \lambda'_j$ for all $j < i$, and $\lambda_i > \lambda'_i$.

And now we turn our attention to tableaux.

Definition 4. Let A be a shape matrix and S a set. A **tableau,** T**, of shape** A **with values in** S is a filling-in of the diagram (A) by elements of S. We denote the tableau by the ordered pair $T = ((A); \tau)$, where τ is the filling-in of (A) by S.

We could have said that τ is a function from (A) to S, but for the fact that we haven't given a formal enough definition of "diagram" to do this. But if one regards the diagram as a collection of cells, then τ would be a function with domain (A). In most cases, we will simply refer to the tableau as T, with the set S an understood ordered basis of a finitely generated free module.

Occasionally we will use the term **row tableau**; this is simply a tableau the diagram of whose shape consists of one row.

Assume now that our set S is totally ordered. We have the following definitions of "standardness". (Later we'll give a more general definition, but these will suffice for the next section.)

Definition 5. We say that a tableau (of any shape) is **Weyl-row-standard** if in each row it is weakly increasing; we say it is **Weyl-column-standard** if in each column it is strictly increasing. We say it is **Weyl-standard** if it is both Weyl-row- and Weyl-column-standard.

Definition 6. We say that a tableau (of any shape) is **Schur-row-standard** if in each row it is strictly increasing; we say it is **Schur-column-standard** if in each column it is weakly increasing. We say it is **Schur-standard** if it is both Schur-row- and Schur-column-standard.

Later it will be convenient to have a quasi-order on tableaux with values in a totally ordered set. Suppose, again, that S is a totally ordered set, say $S = \{s_1, \ldots, s_n\}$ with $s_1 < \cdots < s_n$, and suppose T is a tableau with values in S. Define T_{ij} to be the number of elements in $\{s_1 \ldots, s_i\}$ that appear in at least one of the first j rows of the diagram of T. Now suppose that T' is another tableau.

Definition 7. We say that $T' \leq T$ if $T'_{ij} \geq T_{ij}$ for every i and j. We say that $T' < T$ if $T' \leq T$ and for some i, j we have $T'_{ij} > T_{ij}$.

To see that this is a quasi-order and not an order, consider our set S with three elements: $S = \{s_1, s_2, s_3\}$, with $s_1 < s_2 < s_3$, and consider the diagram corresponding to the partition $\lambda = (4, 2, 1)$. Then the two tableaux

$$T = \begin{array}{|c|c|c|c|} \hline s_1 & s_2 & s_2 & s_3 \\ \hline s_2 & s_3 \\ \cline{1-2} s_3 \\ \cline{1-1} \end{array}$$

and

$$T' = \begin{array}{|c|c|c|c|} \hline s_2 & s_3 & s_2 & s_1 \\ \hline s_3 & s_2 \\ \cline{1-2} s_3 \\ \cline{1-1} \end{array}$$

are such that $T \leq T'$ and $T' \leq T$, but T and T' are clearly not equal. However, the "$\leq$" relation is both reflexive and transitive, as can easily be checked.

3.2 *Associating Weyl and Schur Modules to Shape Matrices*

To each finite free module, F, over a commutative ring, R, and each shape matrix we will associate two maps, a Weyl map and a Schur map, whose images will be called the Weyl and Schur modules of that shape. To do that, we first look at some auxiliary ideas.

If $\mathbf{a} = (a_1, \ldots, a_l)$ is a sequence of non-negative integers, let $\alpha = a_1 + \cdots + a_l$. We define the maps $\delta'_{\mathbf{a}} : D_\alpha F \to D_{a_1} F \otimes \cdots \otimes D_{a_l} F$ and $\delta''_{\mathbf{a}} : \Lambda^\alpha F \to \Lambda^{a_1} F \otimes \cdots \otimes \Lambda^{a_l} F$ to be the diagonalization maps of the indicated divided and exterior powers of F into the indicated tensor products. We define the maps $\mu'_{\mathbf{a}} : \Lambda^{a_1} F \otimes \cdots \otimes \Lambda^{a_l} F \to \Lambda^\alpha F$ and $\mu''_{\mathbf{a}} : S_{a_1} F \otimes \cdots \otimes S_{a_l} F \to S_\alpha F$ to be the multiplication maps from the indicated tensor products of exterior and symmetric powers to the indicated exterior and symmetric powers.

Definition 8 (Weyl and Schur maps). Let F be a free module over the commutative ring, R. For the $s \times t$ shape matrix $A = \left(a_{ij}\right)$, set $\mathbf{r}_i = (a_{i1}, \ldots, a_{it})$, $\mathbf{c}_j = \left(a_{1j}, \ldots, a_{sj}\right)$, $\rho_i = \sum_{j=1}^{t} a_{ij}$, $\gamma_j = \sum_{i=1}^{s} a_{ij}$. The **Weyl map associated to** A, ω_A, is the map

$$\omega_A : D_{\rho_1} F \otimes \cdots \otimes D_{\rho_s} F \to \Lambda^{\gamma_1} F \otimes \cdots \otimes \Lambda^{\gamma_t} F$$

defined as the composition

$$\omega_A = \left(\mu'_{\mathbf{c}_1} \otimes \cdots \otimes \mu'_{\mathbf{c}_t}\right) \theta_W \left(\delta'_{\mathbf{r}_1} \otimes \cdots \otimes \delta'_{\mathbf{r}_s}\right)$$

where, since all the a_{ij} are equal to zero or one, we have identified $D_{a_{ij}} F$ with $\Lambda^{a_{ij}} F$ for all i, j; the map θ_W is the isomorphism comprising all of these identifications together with rearrangement of the factors. "Pictorially" what we have is the following:

$$D_{\rho_1} F \otimes \cdots \otimes D_{\rho_s} F \to \left\{ \begin{array}{c} D_{a_{11}} F \otimes \cdots \otimes D_{a_{1t}} F \\ \otimes \\ \vdots \\ \otimes \\ D_{a_{s1}} F \otimes \cdots \otimes D_{a_{st}} F \end{array} \right\} \xrightarrow{\theta_W}$$

$$\left\{ \begin{array}{c} \Lambda^{a_{11}} F \otimes \cdots \otimes \Lambda^{a_{1t}} F \\ \otimes \\ \vdots \\ \otimes \\ \Lambda^{a_{s1}} F \otimes \cdots \otimes \Lambda^{a_{st}} F \end{array} \right\} \to \Lambda^{\gamma_1} F \otimes \cdots \otimes \Lambda^{\gamma_t} F.$$

The **Schur map associated to** A, σ_A, is the map

$$\sigma_A : \Lambda^{\rho_1} F \otimes \cdots \otimes \Lambda^{\rho_s} F \to S_{\gamma_1} F \otimes \cdots \otimes S_{\gamma_t} F$$

defined as the composition

$$\sigma_A = \left(\mu''_{\mathbf{c}_1} \otimes \cdots \otimes \mu''_{\mathbf{c}_t}\right) \theta_S \left(\delta''_{\mathbf{r}_1} \otimes \cdots \otimes \delta''_{\mathbf{r}_s}\right)$$

where, since all the a_{ij} are zero or one, we have identified $\Lambda^{a_{ij}} F$ with $S_{a_{ij}} F$ for all i, j; the map θ_S is the isomorphism comprising all of these identifications together with rearrangement of the factors. We can view the definition of the Schur map "pictorially" in the same way we did the Weyl map.

Definition 9 (Weyl and Schur modules). Let F be a free R-module, and A a shape matrix. We define the **Weyl module of F associated to** A, denoted $K_A F$, to be the image of ω_A. We define the **Schur module of F associated to** A, denoted $L_A F$, to be the image of σ_A.

Remark 10. The following observations are easy to check and very useful:

(1) If A is a nonzero shape matrix with its initial column consisting only of zeroes and B is the shape matrix with that initial column removed, it is clear that the Weyl and Schur maps associated to A and B are the same. Hence, we will generally assume that our shape matrices have at least one entry in the first column equal to one.
(2) If A is a shape matrix and B is the shape matrix obtained from A by a permutation of its rows (columns), then the associated Weyl and Schur modules of these matrices are isomorphic.

With the definition of Weyl and Schur modules to hand, a natural question to consider is whether these modules are free over the ground ring and, if so, how can we describe a basis. For example, if we take a one-rowed partition λ, what are the Weyl and Schur modules associated to it? The Weyl module is the image of the map $D_\lambda F \overset{\omega_\lambda}{\to} \underbrace{\Lambda^1 F \otimes \cdots \otimes \Lambda^1 F}_{\lambda}$, where ω_λ is the diagonalization map. Clearly, then, the image is isomorphic to $D_\lambda F$ itself, that is, $K_\lambda F = D_\lambda F$. In a similar way we can show that $L_\lambda F = \Lambda^\lambda F$. In both of these cases, the modules are clearly free R-modules, and we have a very concrete description of bases for them. Not only do we have explicit descriptions of their bases, we even have a description in terms of tableaux. In the case of $D_\lambda F$, a basis is parametrizable by the set of all one-rowed tableaux: $\boxed{x_{i_1}}\boxed{x_{i_2}}\boxed{\cdots}\boxed{x_{i_\lambda}}$ where $\{x_1, \ldots, x_m\}$ is an ordered basis of the free module, F, and $i_1 \leq \cdots \leq i_\lambda$. In the case of $\Lambda^\lambda F$, we have that a basis is parametrizable by all one-rowed tableaux: $\boxed{x_{i_1}}\boxed{x_{i_2}}\boxed{\cdots}\boxed{x_{i_\lambda}}$ where $i_1 < \cdots < i_\lambda$.

Later, we will give an outline of a proof that if λ/μ is an almost skew-partition, then $K_{\lambda/\mu}$ and $L_{\lambda/\mu}$ are free, and bases for them can be parametrized by certain sets of tableaux of shape λ/μ satisfying combinatorial conditions.

4 Two-Rowed Modules

Before getting into that, we will look at the very particular case of two-rowed shapes and there introduce heuristically the idea of letter-place for a small number of places. This will enable us to then describe the terms of the resolutions of two-rowed Weyl (and Schur) modules and prove they are truly resolutions by means of an explicit homotopy.

4.1 Illustration of Letter-Place for Two Places

If we take an element $w \otimes w' \in D_p \otimes D_q$, we know that w is in the first factor, and w' is in the second. If $p = q$, we might still want to indicate that these elements are in the first and second factors, but just how would we explicitly denote this fact? The idea of letter-place is to introduce the notion of "place" to indicate that an element (denoted by "letters") in the tensor product is in either place 1 or place 2. So in the "letter-place algebra", $w \otimes w' \in D_p \otimes D_q$ would be written as $(w|1^{(p)})(w'|2^{(q)})$ to indicate that it is the tensor product of an element of degree p in the first factor, and one of degree q in the second. This is then collected in double tableau form as

$$\begin{pmatrix} w & \big| & 1^{(p)} \\ w' & \big| & 2^{(q)} \end{pmatrix}.$$

If we further agree that the symbol $(v|1^{(p)}2^{(q)})$ means $\sum v(p) \otimes v(q) \in D_p \otimes D_q$, where v is an element of degree $p + q$ and the sum represents the diagonalization of v in $D_p \otimes D_q$, then we can also talk about the double tableau

$$\begin{pmatrix} w & \big| & 1^{(p)}2^{(k)} \\ w' & \big| & 2^{(q-k)} \end{pmatrix},$$

which means $\sum w(p) \otimes w(k)w'$. Ordering the basis elements of the underlying free module, we can now talk about "standard" and "double standard" double tableaux, where we're here using "standard" to mean Weyl-standard, since we're talking about tensor products of divided powers.[5] A major result on letter-place algebra is that the set of double standard double tableaux form a basis for $D_p \otimes D_q$ [7].

All of the above discussion tacitly assumed that the places were "positive". When we discuss this in Sect. 5, we will see that we can have positive and negative places, as well as positive and negative letters. But in this section, all letters and places will

[5] In Sect. 5, our more general approach will make it unnecessary to keep talking about different kinds of standardness.

be considered positive. This is reflected in the notation, $a^{(p)}$ for letters, and $1^{(p)}$ for places; we're essentially working in the context of divided powers.

To illustrate the basis theorem, suppose $p < q$, and we have the element $a^{(p)} \otimes b^{(q)} \in D_p \otimes D_q$. Then, although $\begin{pmatrix} a^{(p)} & \big| & 1^{(p)} \\ b^{(q)} & \big| & 2^{(q)} \end{pmatrix}$ is a basis element of $D_p \otimes D_q$, it isn't a double standard tableau (even assuming $a < b$ and $1 < 2$) since $p < q$.

SLIGHT DIGRESSION: Although we will give a general definition of "standard tableau" in the next section, let me give a rough idea of how it applies here. We may think of the tableau we've written $\begin{pmatrix} a^{(p)} & \big| & 1^{(p)} \\ b^{(q)} & \big| & 2^{(q)} \end{pmatrix}$ as a shortcut for writing a strung out p times in the first row, b strung out q times in the second row, and the same for the $1^{(p)}$ and $2^{(q)}$. Now "standard" would mean that the first rows are no shorter than the second (which is already false if we assume that $p < q$), and that each column of the array is strictly increasing (which is the case if we make the assumption that $a < b$ and $1 < 2$, and don't worry about the fact that the top rows are too short to make sense of the inequality beyond the pth term). If, however, we were to assume in addition that $p \geq q$, we would have a double standard tableau.

To write $\begin{pmatrix} a^{(p)} & \big| & 1^{(p)} \\ b^{(q)} & \big| & 2^{(q)} \end{pmatrix}$ as a linear combination of standard tableaux, we clearly must have[6]

$$\begin{pmatrix} a^{(p)} & \big| & 1^{(p)} \\ b^{(q)} & \big| & 2^{(q)} \end{pmatrix} = \sum_{l=0}^{p} c_l \begin{pmatrix} a^{(p)} b^{(q-p+l)} & \big| & 1^{(p)} 2^{(q-p+l)} \\ b^{(p-l)} & \big| & 2^{(p-l)} \end{pmatrix},$$

and we want to determine the coefficients c_l. Rewriting the above, we get

$$a^{(p)} \otimes b^{(q)} = \sum_{l=0}^{p} c_l \sum_{k=0}^{p} \binom{q-k}{p-l} a^{(p-k)} b^{(k)} \otimes a^{(k)} b^{(q-k)};$$

we want the c_l to be such that

$$\sum_{l=0}^{p} c_l \binom{q-k}{p-l} = \begin{cases} 1 \text{ for } k = 0 \\ 0 \text{ otherwise} \end{cases}.$$

Clearly, if we set $c_l = \binom{p-q}{l}$, then for $k = 0$, the sum above is

$$\sum_{l=0}^{p} \binom{p-q}{l} \binom{q}{p-l} = \binom{p}{p} = 1,$$

[6] In the following, we are using the shortcut of not stringing out the letters or numbers that occur with exponents. If you go through the "stringing out" procedure, you will see that the tableaux on the right of the equation below are standard.

while for $k > 0$, we get

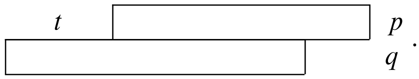
$$\sum_{l=0}^{p} \binom{p-q}{l}\binom{q-k}{p-l} = \binom{p-k}{p} = 0$$

as we wanted.[7]

4.2 Examples of Place Polarization Maps

To illustrate how certain maps can be thought of as "place polarization" maps, take skew-shape:

(A)



To be more precise, we take the skew-shape represented by λ/μ, where $\lambda = (\lambda_1, \lambda_2), \mu = (\mu_1, \mu_2), \lambda_1 - \mu_1 = p, \lambda_2 - \mu_2 = q$, and $\mu_1 - \mu_2 = t$.

It has been defined (see above) as the image of $D_p \otimes D_q$ under the Weyl map. In Sect. 4, we will see that it is also the cokernel of the "box map" (usually designated, unimaginatively, as $\Box_{\lambda/\mu}$) which is the map

$$\Box_{\lambda/\mu} : \sum_{k>t} D_{p+k} \otimes D_{q-k} \to D_p \otimes D_q$$

which sends an element

$$x \otimes y \in D_{p+k} \otimes D_{q-k} \text{ to } \sum x_p \otimes x_k^{'} y,$$

where $\sum x_p \otimes x_k^{'}$ is the component of the diagonal of x in $D_p \otimes D_k$.

LETTER-PLACE PERSPECTIVE: Again, we will wait until Sect. 4 to make all of the following precise, but in the context of letter-place algebra, there is the notion of "place polarization" which "replaces" one positive place by another, say it replaces the occurrence of the positive place 1 by the place 2. This replacement is written (in this case) as ∂_{21}. In general, if we replace a positive place r by a positive place s, we would write this operation as ∂_{sr}. If, moreover, we want to replace a number, say k, of occurrences in the place r by the place s, we would write this operation as $\partial_{sr}^{(k)}$. In this notation, we see that the box map is the direct sum of the place polarization maps, $\partial_{21}^{(k)}$, where $k > t$. To illustrate, we take a double standard

[7] To clear up any misunderstanding about our binomial coefficients, let's define, for $l \geq 0$, $\binom{X}{l}$ to be $\frac{X(X-1)\cdots(X-l+1)}{l!}$. This allows us to substitute negative integers for X.

tableau in $D_{p+k} \otimes D_{q-k}$, let's say

$$\left(\begin{array}{c|c} w & 1^{(p+k)}2^{(l)} \\ w' & 2^{(q-k-l)} \end{array}\right) \in D_{p+k} \otimes D_{q-k},$$

and $\partial_{21}^{(k)}$ would send this to

$$\left(\begin{array}{c|c} w & 1^{(p)}2^{(k)}2^{(l)} \\ w' & 2^{(q-k-l)} \end{array}\right) = \binom{k+l}{k} \left(\begin{array}{c|c} w & 1^{(p)}2^{(k+l)} \\ w' & 2^{(q-k-l)} \end{array}\right) \in D_p \otimes D_q.$$

To explain the mysterious binomial coefficient that comes into play here, take, for example, the case when $w = a^{(p+k+l)}$ and $w' = b^{(q-l-k)}$. That would give us as a starting element,

$$\left(\begin{array}{c|c} a^{(p+k+l)} & 1^{(p+k)}2^{(l)} \\ b^{(q-k-l)} & 2^{(q-k-l)} \end{array}\right) = a^{(p+k)} \otimes a^{(l)}b^{(q-k-l)} \in D_{p+k} \otimes D_{q-k},$$

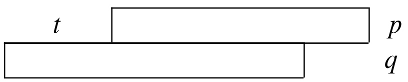
and as the image:

$$\binom{k+l}{k} \left(\begin{array}{c|c} a^{(p+k+l)} & 1^{(p)}2^{(k+l)} \\ b^{(q-k-l)} & 2^{(q-k-l)} \end{array}\right) = \binom{k+l}{k} a^{(p)} \otimes a^{(k+l)}b^{(q-k-l)} \in D_p \otimes D_q.$$

4.3 The Two-Rowed Resolution

With these notations (and assertions) now introduced, we can describe the resolution of our skew-shape, (A), described in Sect. 4.2. We will also describe a contracting homotopy for the nonnegative part of the resolution and a basis for the syzygies.

Recall that the Weyl module associated to the skew-shape

(A) 

is the image of $D_p \otimes_R D_q$ under the Weyl map. The "box map" referred to at the very beginning of Sect. 4.2, and denoted by $\square_{\lambda/\mu}$, was seen to be the sum of place polarizations,

$$\sum_{k>t} \partial_{2,1}^{(k)} : \sum_{k>t} D_{p+k} \otimes_R D_{q-k} \to D_p \otimes_R D_q.$$

If we let $Z_{2,1}$ stand for the generator of a divided power algebra in one free generator, we can let $Z_{2,1}^{(k)}$ act on $D_{p+k} \otimes_R D_{q-k}$ and carry it to $D_p \otimes_R D_q$. (in short, we are letting this formal generator act as the place polarization). Thus, we may take the (t^+)-graded strand of degree q of the normalized bar complex of this algebra acting on $\sum D_{p+k} \otimes_R D_{q-k}$ (where the degree of the second factor determines the grading) to get a complex over the Weyl module:

$$\cdots \to \sum_{k_i>0} Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l+1)} x \otimes_R (D_{t+p+|k|} \otimes_R D_{q-t-|k|}) \to$$
$$\sum_{k_i>0} Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l)} x \otimes_R (D_{t+p+|k|} \otimes_R D_{q-t-|k|}) \to \cdots$$
$$\to \sum_{k>0} Z_{2,1}^{(t+k)} x \otimes_R (D_{t+p+k} \otimes_R D_{q-t-k}) \to D_p \otimes_R D_q \to 0,$$

where the symbol "x" is a "separator variable" to replace the usual bar symbol used in the bar construction (see [11, 12] for a full explanation of this notation) and $|k|$ stands for the sum of the indices k_i. Here, the boundary operator is called ∂_x (or, what is the same thing, it is obtained by polarizing the variable x to the element 1). This, then, describes a left complex over the Weyl module in terms of bar complexes and letter-place algebra. We also know from the fact that the Weyl module is the cokernel of the box map and that the zero-dimensional homology of this complex is the Weyl module itself.

Now the question is: how do we show that this complex is an exact left complex over the Weyl module? In other words, that it is in fact a resolution. One way, is to produce a splitting contracting homotopy, which is what we will do here. Another way is to use our fundamental exact sequences and a mapping cone argument; we refer the reader to [3] for this approach.

Definition 1. With our complex given as above, define the homotopy as follows:

$$s_0 : D_p \otimes_R D_q \to \sum_{k>0} Z_{2,1}^{(t+k)} x \otimes_R D_{t+p+k} \otimes_R D_{q-t-k}$$

sends the double standard tableau $\left(\begin{matrix} w \\ w' \end{matrix} \middle| \begin{matrix} 1^{(p)}2^{(k)} \\ 2^{(q-k)} \end{matrix}\right)$ to zero if $k \le t$, and to $Z_{2,1}^{(k)} x \otimes \left(\begin{matrix} w \\ w' \end{matrix} \middle| \begin{matrix} 1^{(p+k)} \\ 2^{(q-k)} \end{matrix}\right)$ if $k > t$. For higher dimensions ($l > 0$),

$$s_l : \textstyle\sum_{k_i>0} Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l)} x \otimes_R D_{t+p+|k|} \otimes_R D_{q-t-|k|} \to$$
$$\textstyle\sum_{k_i>0} Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l+1)} x \otimes_R D_{t+p+|k|} \otimes_R D_{q-t-|k|}$$

is defined by sending $Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l)} x \otimes \left(\begin{matrix} w \\ w' \end{matrix} \middle| \begin{matrix} 1^{(t+p+|k|)}2^{(m)} \\ 2^{(q-t-|k|-m)} \end{matrix}\right)$ to zero if $m = 0$, and to $Z_{2,1}^{(t+k_1)} x Z_{2,1}^{(k_2)} x \cdots x Z_{2,1}^{(k_l)} x Z_{2,1}^{(m)} x \otimes \left(\begin{matrix} w \\ w' \end{matrix} \middle| \begin{matrix} 1^{(t+p+|k|+m)} \\ 2^{(q-t-|k|-m)} \end{matrix}\right)$ if $m > 0$.

The proofs of the following statements are in [11].

Proposition 2. *The collection of maps* $\{s_l\}_{l\geq 0}$ *provides a splitting contracting homotopy for the complex above.*

Theorem 3. *The complex above is a projective resolution of the Weyl module associated to the shape A, over the Schur algebra of appropriate weight.*[8]

5 The Letter-Place Panoply

We have talked about tensor products of divided powers, exterior powers, and symmetric powers. As we've strongly asserted, the **letter-place algebra** is an effective tool for dealing with these kinds of tensor products. In this section we will define this algebra in (almost) complete generality, and develop some important combinatorial properties of it. Our treatment will be a little less general than that given in [22]; the interested reader may go to that reference to see how multi-signed alphabets are treated in a uniform and general way. We'll deal with the divided powers, case in some detail, and then quickly treat the cases of exterior powers and symmetric powers. Most of the proofs will be found in Sect. 6.

5.1 Positive Places and the Divided Power Algebra

Usually we are given a fixed number, say n, of terms in the tensor product: $D_{k_1}(F)\otimes \cdots \otimes D_{k_n}(F)$, where F is a free module. As we said in the previous section, we intuitively look at such a product and know which is the first factor, the second, and so on. The idea behind the letter-place approach is to clearly designate the places that the terms in the product are actually in. As an example of what we mean, suppose that $x \in D_{k_i}(F)$, and we want to write the element $\underbrace{1 \otimes \cdots \otimes x}_{i} \otimes \cdots \otimes 1$ in the tensor product above. The letter-place algebra will allow us to write this element as $(x|i^{(k_i)})$. How this will help besides just shortening the amount we have to type and the space it takes to type it will become evident as we develop and use this approach.

Just as with the symmetric and exterior algebras, we have that $D(F \oplus G) = D(F) \otimes D(G)$; it is, after all, the graded dual of the symmetric algebra. So, if we take $D(F \otimes R^n) \cong D(\underbrace{F \oplus \cdots \oplus F}_{n})$, we see that $D(F \otimes R^n)$ is equal to

[8] **Personal note:** While the definition of this homotopy may look complicated, I actually "discovered" it while swimming laps in a local lake, after promising Rota we could come up with one using these letter-place techniques. Attempts to define a homotopy using the methods Akin and I had employed earlier were woefully unsuccessful.

$\underbrace{D(F) \otimes \cdots \otimes D(F)}_{n \text{ times}}$. This is natural with respect to the action of $GL(F)$, but clearly not with respect to the action of $GL(R^n)$. In fact, we moved to the notation R^n rather than G to indicate that we have made a choice of basis in our free module, G. We can, though, use G in our preliminary discussion and, assuming that the rank of this free module is n, still see that "in some way," $D(F \otimes G) \cong \underbrace{D(F) \otimes \cdots \otimes D(F)}_{n}$. Now we want to introduce convenient notation to exhibit this isomorphism, as well as to get to the letter-place conventions.

To this end, let us suppose that G has the (ordered) basis, $\{y_1, \ldots, y_n\}$ with $y_1 < \cdots < y_n$, and for any $x \in D_1(F)$, let us denote by $(x|y_i)$ the element $x \otimes y_i$, and by $(x^{(k)}|y_i^{(k)})$ the element corresponding in $\underbrace{D(F) \otimes \cdots \otimes D(F)}_{n}$ to $(x|y_i)^{(k)}$, that is, to the element in that n-fold tensor product of $D(F)$ having $x^{(k)}$ in the ith factor.

- The picture to keep in mind is $(x|y_i)$ is the element $\underbrace{1 \otimes \cdots \otimes x}_{i} \otimes \cdots \otimes 1$. Now

$$k!(x \otimes y_i)^{(k)} = (x \otimes y_i)^k = (\underbrace{1 \otimes \cdots \otimes x}_{i} \otimes \cdots \otimes 1)^k =$$

$$\underbrace{1 \otimes \cdots \otimes x^k}_{i} \otimes \cdots \otimes 1 = k!\,(\underbrace{1 \otimes \cdots \otimes x^{(k)}}_{i} \otimes \cdots \otimes 1),$$

so we see that the above definition of $(x^{(k)}|y_i^{(k)})$ makes sense.

Finally, if $l = l_1 + \cdots + l_n$, and $x \in D_l(F)$, we set

$$(x|y_1^{(l_1)} \cdots y_n^{(l_n)}) = \sum (x(l_1)|y_1^{(l_1)}) \cdots (x(l_n)|y_n^{(l_n)}),$$

where $\sum x(l_1) \otimes \cdots \otimes x(l_n)$ indicates the image of the diagonal map into $D_{l_1}(F) \otimes \cdots \otimes D_{l_n}(F)$ applied to our element x.

Remark 1. The identities and conventions that we adopt for our discussion are those that are clearly valid if one works over the ring of integers (as is the case illustrated above, where we have cancelled $k!$ because there is no torsion over the integers). We will continue to do this in our treatment of the letter-place algebra and all other structures that are transportable from $\mathbb{Z}$ to arbitrary commutative base rings.

A simple illustration, just to fix our ideas, is this:
Suppose x_1, x_2, and x_3 are in $D_1(F)$, and consider the element

$$(x_1 x_2^{(2)} x_3 | y_1^{(2)} y_2 y_3) \in D_2(F) \otimes D_1(F) \otimes D_1(F).$$

Then this element is equal to

$$(x_1x_2|y_1^{(2)})(x_2|y_2)(x_3|y_3) + (x_1x_2|y_1^{(2)})(x_3|y_2)(x_2|y_3)+$$
$$(x_1x_3|y_1^{(2)})(x_2|y_2)(x_2|y_3) + (x_2^{(2)}|y_1^{(2)})(x_1|y_2)(x_3|y_3)+$$
$$(x_2^{(2)}|y_1^{(2)})(x_3|y_2)(x_1|y_3) + (x_2x_3|y_1^{(2)})(x_1|y_2)(x_2|y_3)+$$
$$(x_2x_3|y_1^{(2)})(x_2|y_2)(x_1|y_3).$$

(♠) We agree to set the symbol $(w|y_1^{(a_1)} \cdots y_n^{(a_n)})$ equal to zero if the degree of w is not equal to $\sum_{i=1}^{n} a_i$. The element w is supposed to be a homogeneous element of $D(F)$.

As we saw in the previous section, the letter-place notation we've been using above lends itself very naturally to writing tableaux. That is, suppose we wanted to write the product of the above element, $(x_1x_2^{(2)}x_3|y_1^{(2)}y_2y_3)$ with, say, $(x_3^{(2)}x_1|y_1y_2y_3)$. As we saw above, each of these terms is a sum of a number of addends, so that the notation we have for each of these terms is already of considerable convenience. But now, instead of using juxtaposition to denote the product of these two terms, let us use "double tableau" notation, that is, let us write

$$\left(\begin{array}{l|lll} x_1x_2^{(2)}x_3 & y_1^{(2)} & y_2 & y_3 \\ x_3^{(2)}x_1 & y_1 & y_2 & y_3 \end{array}\right)$$

for this product.

Suppose that we choose an ordered basis for F, say $\{x_1, \dots, x_m\}$ with $x_1 < \cdots < x_m$, and let us say that the elements x_i above are among these basis elements. Then the double tableau above does not change value if we write it as:

$$(DT)\ \left(\begin{array}{l|lll} x_1x_2^{(2)}x_3 & y_1^{(2)} & y_2 & y_3 \\ x_1x_3^{(2)} & y_1 & y_2 & y_3 \end{array}\right).$$

We point this out to indicate that we may always assume that our tableaux are given in such a way that in each row, the elements are increasing. The terminology for this is that the tableaux are **row-standard** (a notion that we've already encountered previously). We could agree to write out the rows of the tableau repeating letters instead of using divided powers. This helps to talk about the columns of a tableau; for instance, the tableau above has two rows and four columns (the number refers to the arrays in the letters as well as the places).

Usually, we call the basis of F the **letters**, while the basis of G is called the **places**. A **basic word of degree** k is simply a basis element of $D_k(F)$, while a **word of degree** k is a linear combination of basic words of degree k. Usually we will write a word as w, and we will write a general double tableau as

$$(G)\qquad \left(\begin{array}{c|cccc} w_1 & 1^{(a_{11})} & 2^{(a_{21})} & 3^{(a_{31})} & \dots \\ w_2 & 1^{(a_{12})} & 2^{(a_{22})} & 3^{(a_{32})} & \dots \\ \dots & \dots & \dots & \dots & \dots \\ w_n & 1^{(a_{1n})} & 2^{(a_{2n})} & 3^{(a_{3n})} & \dots \end{array}\right)$$

where $\alpha_i = (a_{1i} + a_{2i} + a_{3i} + \cdots) \geq \alpha_j$ for $1 \leq i < j \leq n$, and we have written i for y_i. We will continue to write i for y_i as long as there is no danger of confusion. Also, in most cases, the words w_i will be basic words, in which case (since they are basis elements of $D_k(F)$), they are increasing. Because of our convention (♠) above, we see that we may assume that the degree of the element w_i is equal to α_i. Note that our tableau is an element of $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$ when, for each $j = 1, \dots, n, \sum_l a_{jl} = k_j$.

We will call a double tableau **standard** if the words w_i are basic, the lengths of the rows are decreasing (from the top), it is row-standard, and also column-standard in the sense that when we have used repeat notation instead of divided powers, the columns are strictly increasing from top to bottom. Our double tableau (DT) above is **not** a standard double tableau; if we replace the element x_1 in the second row by x_2, however, it will be standard.

Clearly there is a set of double tableaux that form a basis for $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$, namely:

$$(W)\qquad \left(\begin{array}{c|c} w_1 & 1^{(k_1)} \\ w_2 & 2^{(k_2)} \\ \dots & \dots \\ w_n & n^{(k_n)} \end{array}\right)$$

where the w_i run through the basis elements of $D_{k_i}(F)$. But these tableaux are not in general standard. Even if it were the case that $k_1 \geq \cdots \geq k_n$, so that the "place" side of the tableau were standard, the "word" side of the tableau would in general not be so. And if we had to reorder the rows so that they were decreasing in length, we would upset standardness in the column of places.

What we do have is the following theorem:

Theorem 2. *The set of standard double tableaux having the ith place counted k_i times is a basis for $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$.*

The proof breaks up into two parts: the double tableaux of type (G), with $\sum_l a_{jl} = k_j$, generate, and the number of such tableaux is equal to the number of tableaux of type (W) above, for fixed $k_1, \dots, k_n$. The first part is given in Sect. 6.1, and the second part in Sect. 6.2.

5.2 *Negative Places and the Exterior Algebra*

Now we sketch the letter-place approach to the tensor product of a fixed number of copies of ΛF for a fixed free module, F. As in the previous discussion, we use the fact that $\Lambda(F \otimes R^n) \cong \Lambda(\underbrace{F \oplus \cdots \oplus F}_{n})$, which is, in turn, isomorphic to $\underbrace{\Lambda F \otimes \cdots \otimes \Lambda F}_{n}$. There are two natural ways to proceed with this discussion from a letter-place point of view: we could make the letters of F be positive and the places of R^n negative, or vice versa. We will deal with the first case, and indicate the necessary changes if we reverse sign.

Take the basis of R^n to be $\{\mathbf{1}, \ldots, \mathbf{n}\}$, but this time we will treat them as "negative" places (in fact, we have written them in bold face to distinguish them from the "positive" places of the previous subsection). To make the meaning of this clearer (if not altogether clear), we can think of the bases of our free modules as "alphabets" from which we make "words" by stringing them together (as we have been doing). But we can also think of the letters of our alphabet as being "signed", that is, either positive or negative. In the preceding discussion of tensor products of divided powers, all of our letters and places were positive, so that we can assign the number 0 to all of them (to indicate that they're positive). However, in this case, we want to consider the basis elements of F as positive, while those of R^n as negative. So, we assign the value 0 to the basis elements of F, and we assign the value 1 to the basis elements $1, \ldots, n$ to indicate that they are negative. In general, if you have signed alphabets $\mathcal{A}$ and $\mathcal{B}$ which are the bases of A and B, respectively, then the element $a \otimes b \in A \otimes B$ is assigned the value $|a \otimes b| = |a| + |b| \mod 2$, where $|x|$ stands for the sign of x. Of course, we will write the element $a \otimes b$ as $(a|b)$ when we adopt the "letter-place" language as we did in the foregoing subsection.

As before, then, we write the element $(x|\mathbf{i})$ to stand for the element $x \otimes \mathbf{i} \in \Lambda(F \otimes R^n)$, where x is a basis element of F. We think of this, under the identifications made above, as the element $\underbrace{1 \otimes \cdots \otimes x}_{i} \otimes 1 \otimes \cdots \otimes 1 \in \underbrace{\Lambda F \otimes \cdots \otimes \Lambda F}_{n}$. Since x has sign 0 and $\mathbf{i}$ has sign 1, the sign of $(x|\mathbf{i})$ is $0 + 1 = 1$. From the identifications we have made, we see that $(x|\mathbf{i})(y|\mathbf{i}) = -(y|\mathbf{i})(x|\mathbf{i})$. This, and the commutativity of multiplication in the case of divided powers, is consistent with the sign convention:

$$(a_1|b)(a_2|b) = (-1)^{|(a_1|b)||(a_2|b)|}(a_2|b)(a_1|b).$$

Our object is to work toward the same sort of double tableau notation for this tensor product that we had earlier. But before it was possible to take a positive place, i, say, and consider the element $i^{(2)}$ as in $(xy|i^{(2)})$. In this case, since a place $\mathbf{i}$ is negative, we see that $\mathbf{i}^{(2)} = 0$, so we have to define what we mean by the element $(w|p_1 \wedge \cdots \wedge p_k)$ where w is an element (word) of a basis of $D_k F$, and $p_1, \ldots, p_k$ are distinct basis elements of R^n (so that $p_1 \wedge \cdots \wedge p_k$ is plus or minus a basis element of $\Lambda^k R^n$).

Suppose that $w = a_1^{(k_1)} \cdots a_l^{(k_l)}$, let $k = k_1 + \cdots + k_l$ and let $b_1, \ldots, b_k$ be the sequence $\underbrace{a_1, \ldots, a_1}_{k_1}, \ldots, \underbrace{a_l, \ldots, a_l}_{k_l}$. Let $\mathfrak{S}_{k_1,\ldots,k_l}$ denote the Young subgroup of the symmetric group $\mathfrak{S}_k$ consisting of those permutations that permute the first k_1 elements of $1, \ldots, k$ among themselves, the next k_2 elements among themselves, and so on. (This is a subgroup isomorphic to $\mathfrak{S}_{k_1} \times \cdots \times \mathfrak{S}_{k_l}$ consisting of $k_1! \cdots k_l!$ elements.) Then we define

$$(FI) \ \ (w|p_1 \wedge \cdots \wedge p_k) = \sum_{\sigma} (b_{\sigma(1)}|p_1) \cdots (b_{\sigma(k)}|p_k),$$

where σ runs through representatives of distinct cosets of $\mathfrak{S}_k / \mathfrak{S}_{k_1,\ldots,k_l}$.

In the summation above we have written the product in our exterior algebras as simple juxtaposition instead of using wedges. We do this to conserve a uniform notation for multiplication in the letter-place algebra, in which (as we will see later) letters and places may sometimes be positive and sometimes negative.

Three simple examples will make this clear:

- Consider $(a^{(2)}|p_1 \wedge p_2)$. We have

$$(a^{(2)}|p_1 \wedge p_2) = (a|p_1)(a|p_2).$$

- Consider $(a^{(2)}b^{(3)}|p_1 \wedge \cdots \wedge p_5)$. We have

$$\begin{aligned}(a^{(2)}b^{(3)}|p_1 \wedge \cdots \wedge p_5) = {} & (a|p_1)(a|p_2)(b|p_3)(b|p_4)(b|p_5) \\ & +(a|p_1)(b|p_2)(a|p_3)(b|p_4)(b|p_5) \\ & +(a|p_1)(b|p_2)(b|p_3)(a|p_4)(b|p_5) \\ & +(a|p_1)(b|p_2)(b|p_3)(b|p_4)(a|p_5) \\ & +(b|p_1)(a|p_2)(a|p_3)(b|p_4)(b|p_5) \\ & +(b|p_1)(a|p_2)(b|p_3)(a|p_4)(b|p_5) \\ & +(b|p_1)(a|p_2)(b|p_3)(b|p_4)(a|p_5) \\ & +(b|p_1)(b|p_2)(a|p_3)(a|p_4)(b|p_5) \\ & +(b|p_1)(b|p_2)(a|p_3)(b|p_4)(a|p_5) \\ & +(b|p_1)(b|p_2)(b|p_3)(a|p_4)(a|p_5),\end{aligned}$$

  in other words, the $10 = \frac{5!}{2!3!}$ terms that correspond to the ten distinct cosets of $\mathfrak{S}_5 / \mathfrak{S}_{2,3}$.

  We can consider double tableaux as we did earlier, but now the left side consists of words in the positive alphabet which is the basis of F, and the left side consists of words in the negative alphabet of places, the basis $\{\mathbf{1}, \ldots, \mathbf{n}\}$ of R^n.

To see that our usual basis elements of $\Lambda^{k_1}F \otimes \cdots \otimes \Lambda^{k_n}F$ can be expressed as double tableaux, consider the following example:

- The element $x_2 \wedge x_3 \wedge x_5 \otimes x_1 \wedge x_3 \otimes x_2 \wedge x_4 \otimes x_3 \wedge x_5 \wedge x_6 \in \Lambda^3F \otimes \Lambda^2F \otimes \Lambda^2F \otimes \Lambda^3F$ can be expressed as the double tableau:

$$\left(\begin{array}{c|lll} x_2^{(2)} & \mathbf{1} & \mathbf{3} & \\ x_3^{(3)} & \mathbf{1} & \mathbf{2} & \mathbf{4} \\ x_5^{(2)} & \mathbf{1} & \mathbf{4} & \\ x_1 & \mathbf{2} & & \\ x_4 & \mathbf{3} & & \\ x_6 & \mathbf{4} & & \end{array}\right).$$

On the right-hand side of the tableau we have omitted the wedge, and simply spread the basis elements out along the row. We used the divided power notation on the left hand of the column to simplify writing. Really, the top row of the tableau above should look like

$$(x_2\ x_2|\mathbf{1}\ \mathbf{3}).$$

In our situation, we see that if we interchange rows of the tableau, we must take sign into account. For example,

$$\left(\begin{array}{c|lll} x_2^{(2)} & \mathbf{1} & \mathbf{3} & \\ x_3^{(3)} & \mathbf{1} & \mathbf{2} & \mathbf{4} \\ x_5^{(2)} & \mathbf{1} & \mathbf{4} & \\ x_1 & \mathbf{2} & & \\ x_4 & \mathbf{3} & & \\ x_6 & \mathbf{4} & & \end{array}\right) = -\left(\begin{array}{c|lll} x_2^{(2)} & \mathbf{1} & \mathbf{3} & \\ x_3^{(3)} & \mathbf{1} & \mathbf{2} & \mathbf{4} \\ x_5^{(2)} & \mathbf{1} & \mathbf{4} & \\ x_1 & \mathbf{2} & & \\ x_6 & \mathbf{4} & & \\ x_4 & \mathbf{3} & & \end{array}\right).$$

As in the previous case, we now have to define what we mean by a double standard tableau. We will call a double tableau **standard** if it is standard in the old sense on the left-hand side of the vertical column, but on the right-hand side, has the property that it is strictly increasing in the rows and weakly increasing in the columns. Notice that this definition implies that the shape of the tableau is that of a partition.

In the same way the double standard tableaux generate the tensor product of divided powers, these double standard tableaux generate the tensor product of exterior powers. We have the following theorem:

Theorem 3. *The set of standard double tableaux having the* **i***th place counted k_i times is a basis for $\Lambda^{k_1}F \otimes \cdots \otimes \Lambda^{k_n}F$.*

The sketch of the proof Theorem 2 given in Sect. 6, suitably (and easily) modified, gives a proof of this result.

It should be fairly clear that the discussion above could just as well have been carried out if we had assumed that the alphabet for F were signed negatively, and that for the places signed positively. In that case, we would have simply written the basis elements of F in boldface, and those for R^n in ordinary typeface. There are one or two differences that we would have to remark in this case. One is that we would set $(\mathbf{x}_1|i)(\mathbf{x}_2|i) = (\mathbf{x}_1 \wedge \mathbf{x}_2|i^{(2)})$. Another is that we would modify the fundamental identity (FI) earlier in this subsection as follows. If **w** were the word $\mathbf{w} = \mathbf{x}_1 \wedge \cdots \wedge \mathbf{x}_k$ and we had $p = p_1^{(k_1)} \cdots p_l^{(k_l)}$ with $k = k_1 + \cdots + k_l$, then setting $\{q_1, \ldots, q_k\}$ equal to the sequence $\{\underbrace{p_1, \ldots, p_1}_{k_1}, \ldots, \underbrace{p_l, \ldots, p_l}_{k_l}\}$, we define

$$(FI)' \quad (\mathbf{w}|p) = \sum_{\sigma} (\mathbf{x}_1|q_{\sigma(1)}) \cdots (\mathbf{x}_k|q_{\sigma(k)}),$$

where σ ranges over representatives of the cosets of the appropriate Young subgroup.

For "standardness" of double tableaux, we would have strictly increasing rows in the letters, weakly increasing rows in the places; weakly increasing columns in the letters, strictly increasing columns in the places. The proof that these double standard tableaux form a basis is indicated in Sect. 6.

There is one last canonical algebra to consider, namely the tensor product of a fixed number of copies of the symmetric algebra of F : $\underbrace{S(F) \otimes \cdots \otimes S(F)}_{n}$. In this case, we consider the basis elements of both F and R^n negative. We'll skip the discussion here and move on to the next subsection, whose generality will include all of the cases above.

5.3 *Almost Full Generality*

We now put these various pieces together and consider what happens when we have "letter alphabets" and "place alphabets" that contain both positive and negative elements. To use more descriptive notation, we'll let $\mathcal{L}$ and $\mathcal{P}$ stand for the letter and place alphabets, respectively. Further, we'll suppose that $\mathcal{L} = \mathcal{L}^+ \uplus \mathcal{L}^-$ and $\mathcal{P} = \mathcal{P}^+ \uplus \mathcal{P}^-$, where the plus and minus superscripts indicate the signs of the elements of these alphabets. If we now let L^+, L^-, P^+, P^- stand for the free modules generated by these alphabets (or bases), we may consider what is called the **letter-place superalgebra**:

$$\mathcal{S}(\mathcal{L}|\mathcal{P}) = \Lambda(L^+ \otimes P^-) \otimes \Lambda(L^- \otimes P^+) \otimes D(L^+ \otimes P^+) \otimes S(L^- \otimes P^-).$$

The individual factors of the tensor product above have been described in detail; the product of two terms from different components of the product is simply the tensor product of these terms, while the product $(l_1|p)(l_2|p) = (-1)^{|(l_1|p)||(l_2|p)|}(l_2|p)(l_1|p)$.

5.4 Place Polarization Maps and Capelli Identities

In Sect. 3 we defined the Weyl and Schur maps, which entailed a good deal of diagonalization, identification, and multiplication from a tensor product of divided (exterior) powers to a tensor product of exterior (symmetric) powers. We now know that these tensor products of various powers can be expressed in letter-place terms, and we may ask if these complicated maps may be viewed in a different way (hopefully, a simpler way) using the letter-place approach. The answer, as was no doubt anticipated, is yes, and the method will be that of *place polarizations*.

We will consider two types of maps, both of which are called place polarizations: those from "positive places to positive places" and those from "positive places to negative places."

Definition 4. Let $q \in \mathcal{P}^+, s \in \mathcal{P}, s \neq q$, and let $(l|p)$ be a basis element of $\mathcal{S}(\mathcal{L}|\mathcal{P})$. Define the **place polarization**, $\partial_{s,q}$, to be the unique derivation on $\mathcal{S}(\mathcal{L}|\mathcal{P})$ defined by

$$\partial_{s,q}(l|p) = \delta_{q,p} \cdot (l|s),$$

where $\delta_{q,p}$ is the Kronecker delta.

When we say that this map is a derivation on $\mathcal{S}(\mathcal{L}|\mathcal{P})$, we mean that it has the property

$$\partial_{s,q}\{(l_1|p_1)(l_2|p_2)\} = \{\partial_{s,q}(l_1|p_1)\}(l_2|p_2) + (-1)^{|s||p_1|}(l_1|p_1)\partial_{s,q}(l_2|p_2).$$

A straightforward calculation shows that if s is a negative place, then $\partial^2_{s,q} = 0$. On the other hand, we can see easily that if s is positive, $\partial^2_{s,q}\{(l_1|q)(l_2|q)\} = 2\{(l_1|s)(l_2|s)\}$, so that for q and s positive places, it makes sense to talk about the higher divided powers of the place polarizations, $\partial_{s,q}$, namely, $\partial^{(k)}_{s,q}$. In the case of the divided square just discussed, for instance, we see that the equation may be interpreted as $\partial^{(2)}_{s,q}(l_1 l_2|q^{(2)}) = (l_1 l_2|s^{(2)})$. In general, then, we have

$$\partial^{(k)}_{s,q}(w|q^{(m)}) = (w|q^{(m-k)}s^{(k)}),$$

where q and s are positive places.

One fundamental identity, which it is easy to prove, is the following:

Fact 5. Let p, q, r be places with q and p positive, and consider the place polarizations $\partial_{r,q}$, $\partial_{q,p}$, and $\partial_{r,p}$. Then

$$\partial_{r,p} = \partial_{r,q}\partial_{q,p} - \partial_{q,p}\partial_{r,q}.$$

In short, $\partial_{r,p}$ is the commutator of $\partial_{q,p}$ and $\partial_{r,q}$.

We'll first look at **positive-to-positive** place polarizations.

Assume that our places p, q, and r are all positive. Then as we know, we can form the divided powers of all of the place polarizations involving these places, and ask if there are identities associated to these that generalize the basic identity proved above.

Proposition 6 (Capelli Identities). *Let p, q, r be places with p, q, and r all positive, and consider the place polarizations $\partial_{r,q}$, $\partial_{q,p}$ and $\partial_{r,p}$. Then*

$$(Cap) \quad \partial_{r,q}^{(a)}\partial_{q,p}^{(b)} = \sum_{k\geq 0} \partial_{q,p}^{(b-k)}\partial_{r,q}^{(a-k)}\partial_{r,p}^{(k)};$$

$$(Cap') \quad \partial_{q,p}^{(b)}\partial_{r,q}^{(a)} = \sum_{k\geq 0}(-1)^k \partial_{r,q}^{(a-k)}\partial_{q,p}^{(b-k)}\partial_{r,p}^{(k)}.$$

For full details, see [7].

We next turn to **positive-to-negative** place polarizations.

In this case, we consider what happens if r is a negative place, with both p and q still positive. If we look at the above proof and keep in mind that higher divided powers of positive-to-negative polarizations are zero, the identities (Cap) and (Cap') make sense only when $a = 1$, and our proof in the case $a = 1$ above, is still valid. Hence we have the following proposition:

Proposition 7 (Capelli Identities). *Let p, q, r be places with p, q positive, and r negative. Consider the place polarizations $\partial_{r,q}$, $\partial_{q,p}$ and $\partial_{r,p}$. Then*

$$(Cap)_{+} \partial_{r,q}\partial_{q,p}^{(b)} = \partial_{q,p}^{(b)}\partial_{r,q} + \partial_{q,p}^{(b-1)}\partial_{r,p},$$

$$(Cap')_{-} \partial_{q,p}^{(b)}\partial_{r,q} = \partial_{r,q}\partial_{q,p}^{(b)} - \partial_{q,p}^{(b-1)}\partial_{r,p}.$$

5.5 Return to Weyl and Schur Maps

Recall the set-up for the definitions of the Weyl and Schur maps. We let F be a finite free module over the commutative ring, R. For the $n \times m$ shape matrix $A = (a_{ij})$, set $p_i = \sum_{j=1}^{m} a_{ij}, \gamma_j = \sum_{i=1}^{n} a_{ij}$. The *Weyl map associated to* A, ω_A, is a map

$$\omega_A : D_{p_1}F \otimes \cdots \otimes D_{p_n}F \to \Lambda^{\gamma_1}F \otimes \cdots \otimes \Lambda^{\gamma_m}F$$

that we defined using many diagonalizations, identifications, and multiplications.

Similarly, we defined the Schur map

$$\sigma_A : \Lambda^{p_1}F \otimes \cdots \otimes \Lambda^{p_n}F \to S_{\gamma_1}F \otimes \cdots \otimes S_{\gamma_m}F.$$

We now maintain that these maps can be described using place polarizations; in particular, positive-to-negative place polarizations.

For the Weyl map, we are going to consider the basis, $\mathcal{L}^+$, of F as a positive letter alphabet (in the letter-place language), and our place alphabet $\mathcal{P} = \mathcal{P}^+ \uplus \mathcal{P}^-$, where $\mathcal{P}^+ = \{1, \dots, n\}$ and $\mathcal{P}^- = \{\mathbf{1}, \dots, \mathbf{m}\}$. For the Schur map, we are going to regard the basis of F as a negatively signed letter alphabet, $\mathcal{L}^-$, and our place alphabet the same as the above.

We next observe that $\mathcal{S}(\mathcal{L}^+|\mathcal{P}) = D(F \otimes R^n) \otimes \Lambda(F \otimes R^m)$, which contains the subalgebras $\mathcal{S}(\mathcal{L}^+|\mathcal{P}^+) = D(F \otimes R^n)$ and $\mathcal{S}(\mathcal{L}^+|\mathcal{P}^-) = \Lambda(F \otimes R^m)$. Our discussion of the letter-place algebra tells us that $D(F \otimes R^n) = \underbrace{DF) \otimes \cdots \otimes DF}_{n}$ while $\Lambda(F \otimes R^m) = \underbrace{\Lambda F \otimes \cdots \otimes \Lambda F}_{m}$. A similar discussion applies to the algebra $\mathcal{S}(\mathcal{L}^-|\mathcal{P}) = \Lambda(F \otimes R^n) \otimes S(F \otimes R^m)$.

What we will show is that our Weyl (or Schur) maps are compositions of place polarizations that take us from our desired domain to our desired target through $\mathcal{S}(\mathcal{L}^+|\mathcal{P})$ (or $\mathcal{S}(\mathcal{L}^-|\mathcal{P})$).

Although we can carry out this project for arbitrary shapes, we'll restrict ourselves to almost skew-shapes. Recall that an almost skew-shape can be represented as λ/μ where λ is a partition and μ is an almost partition. In order to conform to the notation used to describe the shape matrix, A, above, we'll assume that our partition λ has length n, and that $\lambda_1 - \mu_n = m$ if μ is a partition, and that $\lambda_1 - \mu_{n-1} = m$ if μ is not a partition. A quicker way to say this is that $\lambda_1 - \min(\mu_n, \mu_{n-1}) = m$. As we've noted before, we may as well set $\min(\mu_n, \mu_{n-1}) = 0$.

Using this notation for our shapes, we see that the numbers p_i and γ_j above become:

$$p_i = \lambda_i - \mu_i; \gamma_j = \tilde{\lambda}_j - \tilde{\mu}_j,$$

for $i = 1, \dots, n$ and $j = 1, \dots, m$, where the tilde denotes the transpose shape matrices of λ and μ.

For each $i = 1, \dots, n$, let

$$\Delta_i = \partial_{\lambda_i}, i \cdots \partial_{\mu i} + 1, i.$$

(Recall that we are assuming that $\min(\mu_n, \mu_{n-1}) = 0$, so that $m = \lambda_1$.) Now we set

$$\Delta_{\lambda/\mu} = \Delta_n \cdots \Delta_1.$$

We see that each Δ_i is a composition of positive-to-negative place polarizations from the positive place, i, to the negative places $\mu_{\mathbf{i}} + 1$ to $\lambda_{\mathbf{i}}$. Hence the map $\Delta_{\lambda/\mu}$ is a composition of such place polarizations from $1, \dots, n$ to $\mathbf{1}, \dots, \mathbf{m}$. We see, therefore, that the image of $\Delta_{\lambda/\mu}$ is contained in that part of $\mathcal{S}(\mathcal{L}^+|\mathcal{P})$ which contains no positive places, namely, in $\Lambda(F \otimes R^m)$ or, what is the same thing, it is a map

$$\Delta_{\lambda/\mu} : \underbrace{DF \otimes \cdots \otimes DF}_{n} \to \underbrace{\Lambda F \otimes \cdots \otimes \Lambda F}_{m}.$$

If we restrict it to $D_{p_1}F \otimes \cdots \otimes D_{p_n}F$, it is immediate to see that we end in $\Lambda^{\gamma_1}F \otimes \cdots \otimes \Lambda^{\gamma_m}F$. It's laborious but straightforward to prove that this last map is the same as the Weyl map ω_A for $A = \lambda/\mu$; we will sketch a procedure for carrying out such an argument.

We know that a basis for $D_{p_1}F \otimes \cdots \otimes D_{p_n}F$ consists of double tableaux

$$(W) \qquad \left(\begin{array}{c|c} w_1 & 1^{(p_1)} \\ w_2 & 2^{(p_2)} \\ \cdots & \cdots \\ w_n & n^{(p_n)} \end{array}\right).$$

The result of applying $\Delta_{\lambda/\mu}$ to such a tableau yields the tableau

$$\left(\begin{array}{c|c} w_1 & \boldsymbol{\mu_1} + \mathbf{1} \cdots \boldsymbol{\lambda_1} \\ w_2 & \boldsymbol{\mu_2} + \mathbf{1} \cdots \lambda_2 \\ \cdots & \cdots \quad \cdots\cdots \\ w_n & \boldsymbol{\mu_n} + \mathbf{1} \cdots \boldsymbol{\lambda_n} \end{array}\right).$$

If one now reads this tableau as the element one obtains by diagonalizing w_i over the negative places $\boldsymbol{\mu_i} + \mathbf{i}, \ldots, \boldsymbol{\lambda_i}$ and multiplying, one sees that this is precisely the definition of the map $\omega_{\lambda/\mu}$.

The discussion of the Schur map is identical to this one, with the proviso that we now consider the letters to be negative. However, we are still going from positive places to negative ones, in exactly the same way, so that while the domain and range of the Weyl and Schur maps are different, the expression of them as composites of place polarizations is identical.

5.6 Some Kernel Elements of Weyl and Schur Maps

In this section, we will define some maps from the sum of tensor products of divided powers (exterior powers) to the domain of the Weyl (Schur) map and show that the images are in the kernel of the Weyl (Schur) map. These maps are what were called in [5] the "box map"; here we will see that they are expressible in terms of positive-to-positive place polarizations.

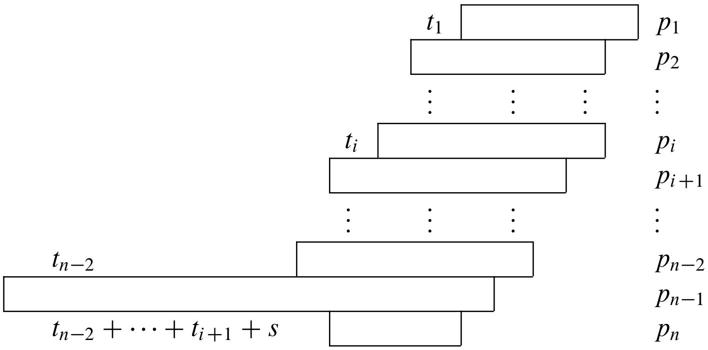
Consider our almost skew-shape $\lambda/\mu : \lambda = (\lambda_1, \ldots, \lambda_n),\ \mu = (\mu_1, \ldots, \mu_n)$. Remember that the shape is of type $\tau = n-(i+1)$ if i is the largest integer different from n such that $\mu_n \leq \mu_i$. Thus, $\tau = 0$ means that λ/μ is a skew-shape; $\tau > 0$ means that the bottom row of the diagram of λ/μ is indented on the left from the penultimate row.

We will introduce some more notation that we will use uniformly when we discuss these almost skew-shapes.

Notation (Almost Skew-Shapes)

We will set $t_i = \mu_i - \mu_{i+1}$ for $i = 1, \dots, n-1$. If $\tau = 0$, this means that $\mu_n \leq \mu_{n-1}$ and $t_{n-1} = \mu_{n-1} - \mu_n = \mu_{n-1}$. If $\tau > 0$, this means that $\mu_{n-1} - \mu_n = -\mu_n < 0$; moreover there is an $i = n - 1 - \tau$ such that $\mu_{i+1} < \mu_n \leq \mu_i$, and we set $s = \mu_n - \mu_{i+1}$. Finally, we denote our shape λ/μ by the notation $(p_1, \dots, p_n; t_1, \dots, t_{n-1})$.

With this notation, we see that the diagram of an almost skew-shape of type $\tau = n - (i + 1) > 0$ looks like this:



with $0 < s \leq t_i$. Of course, $t_{n-2} + \cdots + t_{i+1} + s = \mu_n - \mu_{n-1} = -t_{n-1} > 0$.

We will now restrict ourselves to the Weyl case until the end of this subsection, where we indicate how the results apply to the Schur case as well.

Assume that our shape $(p_1, \dots, p_n; t_1, \dots, t_{n-1})$ is a skew-shape, that is, assume that $t_{n-1} \geq 0$.

For each $i = 1, \dots, n-1$ and for each $k_i > 0$, we consider the module $D_{p_1} \otimes \cdots \otimes D_{p_i+t_i+k_i} \otimes D_{p_{i+1}-t_i-k_i} \otimes \cdots \otimes D_{p_n}$ and the (positive-to-positive) place polarization

$$\partial_{i+1,i}^{(t_i+k_i)} : D_{p_1} \otimes \cdots \otimes D_{p_i+t_i+k_i} \otimes D_{p_{i+1}-t_i-k_i} \otimes \cdots \otimes D_{p_n} \to D_{p_1} \otimes \cdots \otimes D_{p_n}.$$

Here, and from now on in most cases, we omit the underlying free module, F, from our notation.

Define $\square_{\lambda/\mu,i}$ to be the map

$$\square_{\lambda/\mu,i} : \sum_{k_i>0} D_{p_1} \otimes \cdots \otimes D_{p_i+t_i+k_i} \otimes D_{p_{i+1}-t_i-k_i} \otimes \cdots \otimes D_{p_n} \to D_{p_1} \otimes \cdots \otimes D_{p_n},$$

which, on each summand, is equal to $\partial_{i+1,i}^{(t_i+k_i)}$. Now define

$$\mathrm{Rel}(\lambda/\mu) = \sum_i \sum_{k_i} D_{p_1} \otimes \cdots \otimes D_{p_i+t_i+k_i} \otimes D_{p_{i+1}-t_i-k_i} \otimes \cdots \otimes D_{p_n},$$

where the sum is taken over $i = 1, \ldots, n-1$ and all positive k_i. And now define

$$\Box_{\lambda/\mu} : \mathrm{Rel}(\lambda/\mu) \to D_{p_1} \otimes \cdots \otimes D_{p_n}$$

to be the map which, for each i, is the map $\Box_{\lambda/\mu,i}$.

In short, $\Box_{\lambda/\mu}$ is the sum of many, many place polarizations.

We will often write $\mathrm{Rel}(p_1, \ldots, p_n; t_1, \ldots, t_{n-1})$ for $\mathrm{Rel}(\lambda/\mu)$ when we want to make the data for the shape more explicit. The reason for this elaborate notation is that we will eventually show that the image of the map $\Box_{\lambda/\mu}$ is the kernel of the Weyl map $\Delta_{\lambda/\mu} = \omega_{\lambda/\mu}$.

For an almost skew-shape of type $\tau > 0$, the kernel of the Weyl map will be given by relations of the kind above, plus τ additional kinds of terms. It's evident from the definition of the map $\Box_{\lambda/\mu}$ above that the relations on the Weyl map for a skew-shape involve shuffling between consecutive pairs of rows of the shape. The additional terms that we must consider for the almost skew-shape of type $\tau > 0$ involve shuffling between the last row and those rows beyond which it doesn't protrude (to the left), as well as the lowest row beyond which it does protrude. In our diagram of the almost skew-shape of type $\tau > 0$, this means that we have to shuffle the last row with the rows from $n-1$ up through the ith. This makes $n-(i+1) = \tau$ rows, and hence τ kinds of terms to describe these shuffles.

We now formally describe these additional terms. For $j = i+1, \ldots, n-2$, define

$$\triangle_{\lambda/\mu,j} : \sum_{k_j=1}^{t_j} D_{p_1} \otimes \cdots \otimes D_{p_j+k_j} \otimes D_{p_{j+1}} \otimes \cdots \otimes D_{p_{n-1}} \otimes D_{p_n-k_j} \to D_{p_1} \otimes \cdots \otimes D_{p_n}$$

to be the map which on each component is the place polarization $\partial_{n,j}^{(k_j)}$, and for $i = n-(\tau+1)$, define

$$\triangle_{\lambda/\mu,i} : \sum_{k=1}^{s} D_{p_1} \otimes \cdots \otimes D_{p_i+t_i-s+k} \otimes D_{p_{i+1}} \otimes \cdots \otimes D_{p_{n-1}} \otimes D_{p_n-(t_i-s)-k} \to D_{p_1} \otimes \cdots \otimes D_{p_n}$$

to be, again, the map which on each component is the place polarization $\partial_{n,i}^{(t_i-s+k)}$.

We next define, for an almost skew-shape, λ/μ of type $\tau > 0$, the overall relations $\mathrm{Rel}(\lambda/\mu) = \mathrm{Rel}(p_1, \ldots, p_n; t_1, \ldots, t_{n-1})$ by

$$\mathrm{Rel}(p_1, \ldots, p_n; t_1, \ldots, t_{n-2}, 0) \quad \bigoplus$$

$$\sum_{j=i+1}^{n-2}\sum_{k_j=1}^{t_j} D_{p_1}\otimes\cdots\otimes D_{p_j+k_j}\otimes D_{p_{j+1}}\otimes\cdots\otimes D_{p_{n-1}}\otimes D_{p_n-k_j} \quad \bigoplus$$

$$\sum_{k=1}^{s} D_{p_1}\otimes\cdots\otimes D_{p_i+t_i-s+k}\otimes D_{p_{i+1}}\otimes\cdots\otimes D_{p_{n-1}}\otimes D_{p_n-(t_i-s)-k},$$

and the map

$$\square_{\lambda/\mu} : \mathrm{Rel}(p_1,\ldots,p_n;t_1,\ldots,t_{n-1}) \to D_{p_1}\otimes\cdots\otimes D_{p_n}$$

in the by now obvious way.

The thrust of this subsection is the statement of the following essential result.

Theorem 8. *Let λ/μ be any almost skew-shape. Then the composition*

$$\mathrm{Rel}(\lambda/\mu) \xrightarrow{\square_{\lambda/\mu}} D_{p_1}\otimes\cdots\otimes D_{p_n} \xrightarrow{\Delta_{\lambda/\mu}} \Lambda^{\gamma_1}\otimes\cdots\otimes\Lambda^{\gamma_m}$$

is zero. That is, the image of $\square_{\lambda/\mu}$ is contained in the kernel of the Weyl map.

Again we defer to the limits of space and simply refer the reader to [7]. I'll just say that the proof depends heavily on the Capelli identity (Cap) involving positive-to-negative polarizations.

Corollary 9. *Let us define $\bar{K}_{\lambda/\mu}$ to be the cokernel of $\square_{\lambda/\mu}$. Then the identity map on $D_{p_1}\otimes\cdots\otimes D_{p_n}$ induces a map $\theta_{\lambda/\mu} : \bar{K}_{\lambda/\mu} \to K_{\lambda/\mu}$.*

Proof. This follows immediately from the result above. □

All of the above discussion carries over to the Schur map and Schur modules, simply by replacing divided powers by exterior powers and exterior powers by symmetric powers. Or, if one wishes, one can simply replace the positive letter alphabet by its negative counterpart. All the maps that we define are in terms of the place alphabets, and these haven't changed.

5.7 Tableaux, Straightening, and the Straight Basis Theorem

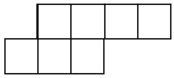
The last theorem is a step toward giving us a presentation of our Weyl (Schur) modules: since the image of $\Delta_{\lambda/\mu}$ is the Weyl module, $K_{\lambda/\mu}$, it suggests that perhaps the sequence

$$\mathrm{Rel}(\lambda/\mu) \to D_{p_1}\otimes\cdots\otimes D_{p_n} \to K_{\lambda/\mu} \to 0$$

is exact. At least we know it's a complex. In this subsection, we will state a basis theorem for our Weyl (Schur) modules, from which the exactness of the above sequence will follow. At this point, a certain amount of combinatorics will enter the picture.

5.7.1 Tableaux for Weyl and Schur Modules

The Weyl module corresponding to the shape,



is the image of $D_4 \otimes D_3$ under the map $\Delta_2\Delta_1$, where

$$\Delta_1 = \partial_{\mathbf{5},1}\partial_{\mathbf{4},1}\partial_{\mathbf{3},1}\partial_{\mathbf{2},1};$$

$$\Delta_2 = \partial_{\mathbf{3},2}\partial_{\mathbf{2},2}\partial_{\mathbf{1},2}.$$

Suppose $\{x_i\}$ is a basis for our free module, F (unspecified rank at this point), and suppose we take the basis element of $D_4 \otimes D_3 : x_2^{(2)}x_3x_4 \otimes x_1x_2x_4$. In our double tableau notation for $D_4 \otimes D_3$, this would be written

$$\left(\begin{array}{c|l} x_2x_2x_3x_4 & 1^{(4)} \\ x_1x_2x_4 & 2^{(3)} \end{array}\right),$$

and its image under $\Delta_{\lambda/\mu}$ would be

$$\left(\begin{array}{r|cccc} x_2x_2x_3x_4 & \mathbf{2} & \mathbf{3} & \mathbf{4} & \mathbf{5} \\ x_1x_2x_4 & \mathbf{1} & \mathbf{2} & \mathbf{3} & \end{array}\right).$$

What we will do is write this element as

$$\begin{array}{ccccc} & \boxed{x_2} & \boxed{x_2} & \boxed{x_3} & \boxed{x_4} \\ \boxed{x_1} & \boxed{x_2} & \boxed{x_4} & & \end{array},$$

namely, as a tableau. This may cause some initial confusion as the element we are representing by this tableau is in reality a sum of basis elements in $\Lambda^1 \otimes \Lambda^2 \otimes \Lambda^2 \otimes \Lambda^1 \otimes \Lambda^1$ rather than simply a filling of a diagram. To be more meticulous, we should really introduce some term such as Weyl tableau to indicate that it is more than just a filled diagram. However, it will be clear from the context of our discussions, when we are using the term "tableau" in this extended sense, and when we are using it in the strictly combinatorial or typographic sense. This notation is not only the standard one used for these modules, but it is also extremely efficient.

All of the above carries over *mutatis mutandis* for Schur modules: the divided powers are replaced by exterior powers, and the exterior powers are replaced by symmetric powers. In addition, the positive letters are replaced by negative letters.

The next definitions of various kinds of standardness and straightness of tableaux apply to tableaux of positive or negative letters; we will therefore introduce a notation that will apply to both cases simultaneously.

Notation (Signed Inequalities)

If $\mathcal{A}$ is a multi-signed alphabet, we say that $a <+ b$ if $a < b$ or a and b are positive, and $a = b$. We say that $a <- b$ if $a < b$ or a and b are negative, and $a = b$.

In less formal language, $a <+ b$ means, for example, that if a and b are both positive, then $a \leq b$. Otherwise, $a < b$.

With this notation we proceed with some definitions.

Definition 10. We say that a tableau (of any shape) is **row-standard** if in each row it is $<+$ -increasing; we say it is **column-standard** if in each column it is $<-$ -increasing. We say it is **standard** if it is both row- and column-standard.

Definition 11. In a row-standard tableau, two elements a_{ik}, a_{jk} (with $i < j$) in the same column are said to form (or be) an **inversion** if they violate column-standardness. That is, if a_{ik} is not $<- a_{jk}$. The inversion is said to be **unflippable** if there is an element in the tableau, b_{ik-1}, immediately to the left of a_{ik}, such that $a_{jk} <- b_{ik-1}$. Otherwise the inversion is called **flippable**. The row-standard tableau is said to be **straight** if every inversion is unflippable.

Clearly, since a standard tableau has no inversions, a standard tableau is necessarily straight. We do have a partial converse.

Proposition 12. *If λ/μ is a skew-shape and T is a row-standard tableau of that shape, then T is straight if and only if it is standard.*

Proof. This is pretty straightforward. □

From now on, we will focus our attention on Weyl modules; the appropriate changes for Schur modules will mostly be left to the reader. In the Weyl case, $<+$ -increasing means weakly increasing, while $<-$ -increasing means strictly increasing.

5.7.2 Straightening Tableaux

What we want to do now is show that our Weyl modules are generated by the straight tableaux. There is a procedure that is called "straightening", that is used to do this. Before we get into the general details, we give two examples of straightening.

Example 13. Let's take as our first example the tableau we've just looked at, namely, $\begin{array}{|c|c|c|c|c|} \cline{2-5} \multicolumn{1}{c|}{} & x_2 & x_2 & x_3 & x_4 \\ \hline x_1 & x_2 & x_4 & \multicolumn{2}{c}{} \\ \cline{1-3} \end{array}$. It fails to be standard because of an inversion in the second column. This element is the image of $x_2^{(2)} x_3 x_4 \otimes x_1 x_2 x_4 \in D_4 \otimes D_3$. However, let us look at $x_1 x_2^{(3)} x_3 x_4 \otimes x_4 \in D_6 \otimes D_1 \subset \text{Rel}((5,3)/(1,0))$, and at its image under $\square_{(5,3)/(1,0)}$ which, in this instance, means applying the map $\partial_{2,1}^{(2)}$. We get

$$x_2^{(2)}x_3x_4 \otimes x_1x_2x_4 + x_2^{(3)}x_4 \otimes x_1x_3x_4 + 2x_2^{(3)}x_3 \otimes x_1x_4^{(2)} +$$
$$x_1x_2x_3x_4 \otimes x_2^{(2)}x_4 + x_1x_2^{(2)}x_4 \otimes x_2x_3x_4 +$$
$$2x_1x_2^{(2)}x_3 \otimes x_2x_4^{(2)} + 2x_1x_2^{(3)} \otimes x_3x_4^{(2)}.$$

Notice that the first summand is the term we started with. If we apply the Weyl map to this sum, we get zero by the theorem of the last section. This means that our original tableau is equal to the negative of the sum of the tableaux we obtain by writing the images of all the remaining summands as tableaux. One sees easily that all of these tableaux are standard, so that our original tableau is a linear combination of standard tableaux.

Example 14. Another example to consider is the following tableau corresponding to an almost skew-shape of type 1: $\begin{array}{cccc} & x_3 & x_3 & x_3 \\ x_2 & x_4 & x_4 & \\ & x_4 & & \end{array}$. Here we have an inversion that is flippable (in the second column, second and third rows). In this case, we swoop up the third with the second row, and consider the element $x_3^{(3)} \otimes x_2x_4^{(3)} \otimes 1 \in D_3 \otimes D_4 \otimes D_0 \subset \mathrm{Rel}((4,3,2)/1,0,1))$. This maps (under $\partial_{3,2}$) to the sum: $x_3^{(3)} \otimes x_2x_4^{(2)} + x_3^{(3)} \otimes x_4^{(3)} \otimes x_2$, one of whose summands is our original term. If we arrange these in tableaux, we have our original tableau plus another, namely, $\begin{array}{cccc} & x_3 & x_3 & x_3 \\ x_4 & x_4 & x_4 & \\ & x_2 & & \end{array}$, which isn't straight because of a flippable inversion in the second column, first and third row. We will see soon that this tableau is "better" than the one we started with, but we can go one step further to actually express our original tableau as a sum of straight ones. For now we can swoop up the third and first rows together, getting $-x_2x_3^{(3)} \otimes x_4^{(3)} \otimes 1 \in D_4 \otimes D_3 \otimes D_0 \subset \mathrm{Rel}((4,3,2)/(1,0,1))$, and the image of this element under $\partial_{3,1}$ is $-x_3^{(3)} \otimes x_4^{(3)} \otimes x_2 - x_2x_3^{(2)} \otimes x_4^{(3)} \otimes x_3$. This shows that our original tableau is equal to the straight tableau: $\begin{array}{cccc} & x_2 & x_3 & x_3 \\ x_4 & x_4 & x_4 & \\ & x_3 & & \end{array}$.

The main task is to formalize the argument underlying the procedure illustrated above in order to prove that the straight tableaux generate. We will have to omit the details, and refer the reader to the book, [7], where these can be found. But we should at least say what we mean by one tableau being "better" than another.

Definition 15. Let T be a tableau. We define the **column word of** T, denoted u_T, to be the word we obtain by writing down the elements of the columns of T starting from the bottom of the left-most column of T, working up that column, returning to the bottom of the next column, etc. We define the **modified column word of** T, denoted w_T, to be the word obtained from T reading the columns in decreasing order. Finally, we define the **reversed column word of** T, denoted w'_T, to be the word obtained from T reading the columns in increasing order.

Just to be sure that there is no confusion about this definition, let us look at the tableau, $\begin{array}{cccc} & \boxed{x_3} & \boxed{x_3} & \boxed{x_3} \\ \boxed{x_4} & \boxed{x_4} & \boxed{x_4} & \\ & \boxed{x_2} & & \end{array}$. The column word is $u_T = x_4x_2x_4x_3x_4x_3x_3$; the modified column word is $w_T = x_4x_4x_3x_2x_4x_3x_3$, and the reversed column word is $w'_T = x_4x_2x_3x_4x_3x_4x_4$.

Notice that the tableau above appeared in Example VI.7.5, but there it was being treated as an element of $K_{\lambda/\mu}$. Here, however, we are regarding the tableau purely combinatorially.

Definition 16. Given two tableaux, T and T', corresponding to the same diagram, we say that $T' < T$ if $u_T < u_{T'}$ in the lexicographic ordering of words.

In Examples 13 and 14 above, the tableaux that were produced via straightening were all less than the original tableaux we started with.

These definitions provide the tools to show that the straight tableaux generate our Weyl module for an almost skew-shape.

5.7.3 Taylor-Made Tableaux, or a Straight-Filling Algorithm

To prove that the set of straight tableaux is linearly independent, there is a very clever algorithm, which we will call the **Taylor algorithm**, that produces straight tableaux of a given shape from certain reverse column words. The algorithm was developed by B. Taylor, and, as with most of these constructions involving straight tableaux, applies to the larger class of row-convex shapes. We, however, will describe this algorithm just for our almost skew-shapes.

Algorithm (Taylor Algorithm)

Let D be the diagram of an almost skew-shape with columns $c_1, \ldots, c_m$, and let $w = x_{11} \wedge \cdots \wedge x_{1\gamma_1} \otimes \cdots \otimes x_{m1} \wedge \cdots \wedge x_{m\gamma_m}$ be a basis element of $\Lambda^{\gamma_1} \otimes \cdots \otimes \Lambda^{\gamma_m}$. Arrange $x_{11}, \ldots, x_{1\gamma_1}$ in increasing order in column c_1. Next, place x_{21} in the first box of c_2 which either has no neighbor to its immediate left or has such a neighbor whose entry is less than or equal to x_{21}. If there is no such box in c_2, the output of the algorithm is "no straight filling." If there is such a box, fill it with x_{21}. Assuming that we have placed $x_{21}, \ldots, x_{2\,j}$, we place $x_{2\,j+1}$ in the first empty box of c_2 which either has no neighbor to its immediate left, or has such a neighbor whose entry is less than or equal to $x_{2\,j+1}$. Again, if there is no such box, our output is "no straight filling"; if there is, we fill it with $x_{2\ j+1}$.[9] We continue in this way with the remaining columns, obtaining an output of "no straight filling," or a filling which we shall call $T(w)$.

Let's look at an example or two. As our shape, we will take one we have used earlier, namely, $(4,3,2)/(1,0,1)$. This has three rows and four columns, with $\gamma_1 = 1, \gamma_2 = 3, \gamma_3 = 2, \gamma_4 = 1$. The word $w_1 = x_4 \otimes x_1 \wedge x_2 \wedge x_4 \otimes x_4 \wedge x_5 \otimes x_6$ produces the filling $\begin{array}{|c|c|c|c|} \cline{2-4} \multicolumn{1}{c|}{} & x_1 & x_4 & x_6 \\ \hline x_4 & x_4 & x_5 & \multicolumn{1}{c}{} \\ \cline{1-3} \multicolumn{1}{c|}{} & x_2 & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-2} \end{array}$, while the word $w_2 = x_4 \otimes x_1 \wedge x_2 \wedge x_5 \otimes x_4 \wedge x_6 \otimes x_4$ produces the filling $\begin{array}{|c|c|c|c|} \cline{2-4} \multicolumn{1}{c|}{} & x_1 & x_4 & x_4 \\ \hline x_4 & x_5 & x_6 & \multicolumn{1}{c}{} \\ \cline{1-3} \multicolumn{1}{c|}{} & x_2 & \multicolumn{1}{c}{} & \multicolumn{1}{c}{} \\ \cline{2-2} \end{array}$. On the other hand, words such as $x_4 \otimes x_1 \wedge x_2 \wedge x_4 \otimes x_5 \wedge x_6 \otimes x_4$ or $x_5 \otimes x_1 \wedge x_2 \wedge x_4 \otimes x_4 \wedge x_6 \otimes x_4$ produce "no straight filling."

We point out a few things about this algorithm. First of all, if we start with a straight tableau, T, with reverse column word w' (written as a basis element of our tensor product of exterior powers), then $T(w')$ will clearly have the reverse column word w', if $T(w')$ exists. But from our second fact above, $T(w')$ clearly does exist and equals T. We therefore see that if two straight tableaux have the same reverse column word (and hence the same modified column word), then they are equal. Hence the straight tableaux are precisely those whose reverse column words produce a successful outcome of the straight-filling algorithm.

This algorithm makes it possible to prove the linear independence of the straight tableaux in a manner almost identical to the classical proof for skew-shapes (for which straight tableaux are standard). It was with the idea of "straight tableaux" that Taylor was able to extend that proof to the almost skew-shapes.

We now have the straight basis theorem for almost skew-shapes.

Theorem 17. *Let F be a free R-module, and λ/μ an almost skew-shape. The following statements are true:*

(1) The Weyl module, $K_{\lambda/\mu}(F)$, is a free R-module with basis consisting of the straight tableaux in a basis of F.
(2) The map $\theta_{\lambda/\mu} : \bar{K}_{\lambda/\mu}(F) \to K_{\lambda/\mu}(F)$ is an isomorphism.
(3) The functor $K_{\lambda/\mu}(F)$, considered as a functor of F, is universally free.

We should add that from the proof of this theorem, we also obtain a proof of the fact that the $\square_{\lambda/\mu}$ map provides a presentation of the Weyl (Schur) module corresponding to an almost skew-shape.

Since the proofs (and examples) have dealt almost exclusively with Weyl modules, I should add that with slight (and evident) modifications for the case of Schur modules, the same theorems can be proven.

---

[9] We have put a larger space between the indices 2 and j and 2 and $j+1$ to avoid possible confusion of juxtaposition with multiplication.

5.8 Duality

As anticipated in our discussion of the special case of hooks, there is a duality between some Weyl and Schur modules. It is well-known, [5], that if A is the shape matrix of a skew-shape, λ/μ, and $\widetilde{A}$ is its transpose (again the shape matrix of a skew-shape), then $K_A(F^*) \cong (L_{\widetilde{A}}(F))^*$. The proof depends in part on the fact that for such shapes, the modules in question are universally free. Now, if A is the shape matrix of an almost skew-shape of positive type, its transpose is no longer of the same kind. Therefore, if we want an isomorphism like the one stated, we would at least have to saturate the class of almost skew-shapes with respect to transposition, and develop all of the preceding material for that larger class of shapes. Since all of the shapes in that class would be row-convex, and the straightening techniques and algorithms used in this chapter apply to row-convex shapes, one could probably arrive at such a duality statement.

6 Further Discussion of the Proofs

In this section, we'll try to furnish an idea of the proof of the theorems stated in Sect. 5.1 on letter-place methods. More precisely, Sects. 6.1 and 6.2 indicate the proof of Theorem 2 which involves straightening considerations (Sect. 6.1) as well as a combinatorial procedure known as the Robinson–Schensted–Knuth correspondence (Sect. 6.2).

6.1 Theorem 2, Part 1: The Double Standard Tableaux Generate

To show the generation by double standard tableaux, we can assume that they're row-standard all the time, since we can always row-standardize them without changing their values. We can also assume that the lengths of the rows are decreasing, since we can always arrange that without changing their values. So we're talking about double tableaux of the form

$$(S) \qquad \left(\begin{array}{c|cccc} w_1 & 1^{(a_{11})} & 2^{(a_{21})} & 3^{(a_{31})} & \cdots \\ w_2 & 1^{(a_{12})} & 2^{(a_{22})} & 3^{(a_{32})} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ w_n & 1^{(a_{1n})} & 2^{(a_{2n})} & 3^{(a_{3n})} & \cdots \end{array}\right)$$

where $\alpha_i = (a_{1i} + a_{2i} + a_{3i} + \cdots) \geq \alpha_j$ for $1 \leq i < j \leq n$, and the words, when written out, are increasing. Our tableaux of the form (W):

$$(W) \qquad \left(\begin{array}{c|c} w_1 & 1^{(k_1)} \\ w_2 & 2^{(k_2)} \\ \cdots & \cdots \\ w_n & n^{(k_n)} \end{array} \right)$$

are, by rearranging the order of the rows if necessary, of type (S) above. Since they form a basis for $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$, it certainly suffices to prove that any double tableau of type (S) is a linear combination of standard double tableaux.

The procedure used is to put a quasi-order on these double tableaux and show that we can express a double tableau that is not standard in the places (or letters) as a linear combination of others which are "lower" in the quasi-order. Since the set of tableaux is finite, this will show that we can express any double tableau as a linear combination of such that are standard in the places. We then do the same for the letters. As this process cannot go on forever, each double tableau must eventually be a linear combination of standard double tableaux; that is, the standard ones generate.

At the cost of being ultra pedantic, we shall denote a double tableau, T, by a triple: $T = (\lambda; L, P)$, to denote the diagram, the "letter", and the "place" parts of the tableau. We stress that the diagram is always to be that of a partition in this case. In Definition 3, we described a well-known partial order on partitions: we said that $\lambda' > \lambda$ if the first row of λ' (from the top) which differs in length from the corresponding row of λ, is longer than that of λ. We also defined in Definition 7, a quasi-order on single tableaux. Using these orderings, we now define a quasi-ordering of our double tableaux.

Definition 1. Given tableaux, $T = (\lambda; L, P)$ and $T' = (\lambda'; L', P')$, we define $(\lambda'; L', P') \leq (\lambda; L, P)$ if $\lambda' \geq \lambda$, $L'_{ij} \geq L_{ij}$ for all i, j, and $P'_{ij} \geq P_{ij}$ for all i, j. We then say that $(\lambda'; L', P') < (\lambda; L, P)$ if $(\lambda'; L', P') \leq (\lambda; L, P)$, and either $\lambda' > \lambda$ or $\lambda' = \lambda$ and either $L'_{ij} > L_{ij}$ for some i, j, or $P'_{ij} > P_{ij}$ for some i, j.

To give an example, and also to see how the general "straightening" proceeds, consider an elementary, yet prototypical situation:

$$T = \left(\begin{array}{c|cc} w_1 & 1^{(a_1)} & 2^{(b_1)} \\ w_2 & 1^{(2)} & 2^{(b_2)} \end{array} \right),$$

where w_1 and w_2 are words of degrees $a_1 + b_1$ and $2 + b_2$, respectively. (We're now dropping the cumbersome designation of a tableau as a triple.)

Our problem in the "place" section of the tableau is that we have 1 in the bottom row, so that no matter what the value of a_1, that section isn't standard.

Recall that

$$T = \sum w_1(a_1) w_2(2) \otimes w_1(b_1) w_2(b_2)$$

where the letters in parentheses indicate the degrees of the w_i under appropriate diagonalization. (We will continue to use this notation to indicate that we have diagonalized our terms in the indicated degrees.)

Consider now:

$$T' = \left(\begin{array}{l|ll} \sum w_1 w_2(2) & 1^{(a_1+2)} \, 2^{(b_1)} & \\ \sum w_2(b_2) & 2^{(b_2)} & \end{array} \right),$$

$$T'' = \left(\begin{array}{l|ll} w_1 & 1^{(a_1+1)} & 2^{(b_1-1)} \\ w_2 & 1 & 2^{(b_2+1)} \end{array} \right),$$

$$T''' = \left(\begin{array}{l|ll} w_1 & 1^{(a_1+2)} & 2^{(b_1-2)} \\ w_2 & 2^{(b_2+2)} & \end{array} \right).$$

Then it is straightforward to check that

$$T = T' - (b_2+1)T'' - \binom{b_2+2}{2} T'''.$$

- It is extremely important to observe at this point that the diagram for T' has changed shape, but it is bigger than that of our original T. The shapes of T'' and T''' are the same as our original, the letter part of the tableau hasn't changed (i.e., $L'' = L''' = L$), but we have both P'' and P''' which are different from P. In both cases, though, we have decreased these tableaux (in the quasi-order). Notice that if we were to "straighten" our letter tableau similarly, the changes would either lead to a bigger diagram, in which case we go down in the order, or the places would remain unchanged, and the letter tableaux that intervened would be lower in the quasi-order.
- When checking this type of calculation, it's most often convenient to assume that the words, w_i, are simply divided powers. That is, we might assume that $w_1 = x^{(a_1+b_1)}$ and that $w_2 = y^{(2+b_2)}$. This makes keeping track of the diagonalizations much easier.

The point of this exercise is to show that when we get rid of an instance of non-standardness, the tableaux that emerge in the process are all less, in the quasi-order we introduced, than the tableau we started with. In the case of T', this is due to the fact that the diagram of T' is properly larger than that of T; in the other two cases, the "letter" side of the tableau hasn't changed, but the "place" side has properly decreased. Of course, the tableau, T'', is manifestly nonstandard (the others may or may not be, depending on the relative values of a_1, b_1, and b_2).

To see how a greater number of places affect the straightening procedure, we'll look at one more example. To this end, consider

$$T = \left(\begin{matrix} w_1 \\ w_2 \end{matrix} \middle| \begin{matrix} 1^{(a_1)} & 2^{(b_1)} & 3^{(c_1)} \\ 1 & 2^{(b_2)} & 3^{(c_2)} \end{matrix} \right).$$

Then if we set

$$T' = \left(\begin{matrix} \sum w_1 w_2(1) \\ \sum w_2(b_2 + c_2) \end{matrix} \middle| \begin{matrix} 1^{(a_1+1)} & 2^{(b_1)} & 3^{(c_1)} \\ 2^{(b_2)} & 3^{(c_2)} & \end{matrix} \right),$$

$$T'' = \left(\begin{matrix} w_1 \\ w_2 \end{matrix} \middle| \begin{matrix} 1^{(a_1+1)} & 2^{(b_1-1)} & 3^{(c_1)} \\ 2^{(b_2+1)} & 3^{(c_2)} & \end{matrix} \right),$$

and

$$T''' = \left(\begin{matrix} w_1 \\ w_2 \end{matrix} \middle| \begin{matrix} 1^{(a_1+1)} & 2^{(b_1)} & 3^{(c_1-1)} \\ 2^{(b_2)} & 3^{(c_2+1)} & \end{matrix} \right),$$

we have

$$T = T' - (b_2 + 1)T'' - (c_2 + 1)T'''.$$

Again we see that we've eliminated the offending 1 in the bottom row, and we have "replaced" T by the tableaux T', T'' and T''' which are strictly lower in our quasi-order than T.

As these two examples show, it is enough to work either on the letter or place side of the tableau to push us further toward standardness. These examples are prototypical in the sense that it's enough to work on straightening two adjacent rows, for if we have a tableau that is not standard, this is because it is not strictly increasing in the columns (since we start with row-standardness). But then a violation of strict increase must also occur in two adjacent rows. This is the reason that we can focus on two-rowed double tableaux such as

$$T = \left(\begin{matrix} w_1 \\ w_2 \end{matrix} \middle| \begin{matrix} z_{11} & z_{12} & z_{13} & \cdots & z_{1k} \\ z_{21} & z_{22} & z_{23} & \cdots & z_{2l} \end{matrix} \right)$$

where we write z_{ij} to represent the integers from 1 to n, with possible repeats. Suppose that the first violation of strict increase occurs in the ith column, that is, $z_{11} < z_{21}, \ldots, z_{1i-1} < z_{2i-1}$, but $z_{1i} \geq z_{2i}$. (The ith column could be the first, as it was in the above examples.) Next, suppose that $z_{2i} = \cdots = z_{2t} < z_{2t+1}$ (possibly $t = i$), and consider the sum of tableaux:

$$T' = \sum \left(\begin{matrix} w_1 w_2(t) \\ w_2(l-t) \end{matrix} \middle| \begin{matrix} z_{11} & \cdots z_{1i-1} & z_{21} \cdots z_{2t} & z_{1i} \cdots z_{1k} \\ z_{2t+1} & \cdots & z_{2l} & \end{matrix} \right).$$

Now, for each subset J of $U = \{z_{21}, \dots, z_{2t}, z_{1i}, \dots, z_{1k}\}$, of order $0 < s < t$, let J' be the complement of J in U, let $Z_{J'}$ be the row tableau of that set, and let Z_J be the corresponding row tableau of J. We let I stand for an arbitrary subset of U of order precisely t, other than the subset $\{z_{21}, \dots, z_{2t}\}$ itself, I' its complementary subset, and $Z_{I'}, Z_I$ as for J. Set

$$T'_J = \quad \sum \left(\begin{matrix} w_1 w_2(t-s) \\ w_2(l-t+s) \end{matrix} \middle| \begin{matrix} z_{11} & \cdots & z_{1i-1} & Z_{J'} \\ Z_J & z_{2\,t+1} & \cdots & z_{2l} \end{matrix} \right),$$

and let c_J equal plus or minus the product of binomial coefficients that would be appropriate due to the multiplication in the divided power algebra, of Z_J with $z_{2\,t+1} \cdots z_{2l}$ (see the foregoing examples to get a more concrete picture of this description). Finally, let

$$T''_I = \quad \sum \left(\begin{matrix} w_1 \\ w_2 \end{matrix} \middle| \begin{matrix} z_{11} & \cdots & z_{1i-1} & Z_{I'} \\ Z_I & z_{2\,t+1} & \cdots & z_{2l} \end{matrix} \right),$$

and let c_I equal plus or minus the corresponding product of binomial coefficients. Then it is tedious, but straightforward, to prove that

$$T = \pm T' + \sum_J c_J T'_J + \sum_I c_I T''_I.$$

The important thing to observe in this equation is that the diagrams of T' and T'_J are all bigger than that of T, and so all these terms are strictly lower in the quasi-order. The terms T''_I all have the letter half of the tableau unchanged from that of T. However, since our sets I must have t elements, and none can be the set $\{z_{21}, \dots, z_{2t}\}$ itself, at least one of these elements must remain in the top row, and one of the elements $z_{1u} > z_{2i}$ must come down into the bottom row. Thus, the resulting tableaux in this case are all less in the quasi-order than the original.

One may ask, when dealing with the tableaux other than the T''_I what we do about keeping these tableaux within the class we're considering, namely, the lengths of the rows decreasing. In order to keep to this recipe, we simply push the top row of the two up as far as it has to go, and the bottom one down to where it has to go to make it into a legitimate shape. But this still makes the shape lower in our quasi order than the original.

6.2 *Theorem 2 Part 2: Linear Independence of Double Standard Tableaux*

Here we introduce the Robinson–Schensted–Knuth correspondence (which will be written R-S-K in the future) to set up a one-to-one count between the double standard tableaux and the usual basis elements of $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$.

Let us agree to write $x^{(k)}$ as $\underbrace{x \cdots x}_{k}$, when there is no danger of confusion. That is, we write $\underbrace{x \cdots x}_{k}$ to mean the row tableau consisting of k copies of x in a row. Then a basis element of $D_k(F)$ is the same as a non-decreasing sequence of elements: $x_{i_1} \leq x_{i_2} \leq \cdots \leq x_{i_k}$, and a basis element of $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$ is a long string of such sequences. For instance, the basis element $x_1^{(2)} x_3 \otimes x_2^{(3)} x_3 \in D_3 \otimes D_4$ corresponds to the long sequence $\beta = x_1 x_1 x_3 x_2 x_2 x_2 x_3$. Now R-S-K sets up a correspondence between such sequences and pairs of tableaux. Before we describe (loosely) this correspondence in general, let us look at the example at hand.

To β we want to associate two tableaux, $L(\beta)$ and $P(\beta)$; first we will describe how we get $L(\beta)$. Since $x_1 x_1 x_3$ is increasing, we put these elements in a row tableau: $\begin{array}{|c|c|c|}\hline x_1 & x_1 & x_3 \\ \hline\end{array}$. But now we hit up against x_2 (the next term in our sequence), and if we were to put that in as the next element of the row, we would spoil row-standardness. So, we "bump" the x_3 from the first row and replace it by x_2 while moving x_3 to the second row of a now two-rowed tableau: $\begin{array}{|c|c|c|}\hline x_1 & x_1 & x_2 \\ \hline x_3 \\ \cline{1-1}\end{array}$. And now we see that the remaining three terms of the sequence, β, can all be placed in the first row without necessitating any bumping, so the tableau we end up with is $L(\beta) = \begin{array}{|c|c|c|c|c|c|}\hline x_1 & x_1 & x_2 & x_2 & x_2 & x_3 \\ \hline x_3 \\ \cline{1-1}\end{array}$. (Notice that this tableau is standard, an end result guaranteed by the nature of the bumping process.)

To assign the next tableau, $P(\beta)$, we will make a slight modification of the usual R-S-K, and make use of the fact that we are looking very distinctly at $D_3 \otimes D_4$, namely, a twofold tensor product of divided powers of designated degrees. This means that we have only two places to deal with, namely, 1 and 2, and we want $P(\beta)$ to be of the same shape as $L(\beta)$, standard, and filled with the places 1 and 2. Now the first part of the construction of our tableau involved using the first three terms of the sequence, β, so that the first three boxes of the first row should be filled in with the place, 1. The second row was next produced by bumping, so we put the place 2 in that box of the second row. The next three entries in the first row were produced by inserting entries from the second factor, so we fill them in with 2, and the resulting tableau we get is $P(\beta) = \begin{array}{|c|c|c|c|c|c|}\hline 1 & 1 & 1 & 2 & 2 & 2 \\ \hline 2 \\ \cline{1-1}\end{array}$.

On the other hand, given the pair of tableaux, $L(\beta)$ and $P(\beta)$, we can reconstruct the sequence β. We look at the tableau, $P(\beta)$, and peel off the **highest place from the highest row first**, with its corresponding letters in $L(\beta)$. This tells us that in our second factor, we had $x_2 x_2 x_3$, and we still are left with the pair of tableaux: $\begin{array}{|c|c|c|}\hline x_1 & x_1 & x_2 \\ \hline x_3 \\ \cline{1-1}\end{array}$ and $\begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 2 \\ \cline{1-1}\end{array}$. This means that we still have a term from the second factor, and it was obtained by bumping from a single-rowed tableau with three entries. The only way that bumping could have happened was for x_3 to have been in the upper row previously, and to have been bumped by x_2 (for if an element of the first row had been bumped by x_1, that element would have been x_2, and we would have had

x_2 in the second row). Therefore, the full second factor must be $x_2x_2x_2x_3$, and our first factor is what's left, namely, $x_1x_1x_3$.

To explain the bumping procedure that leads to the construction of the first tableau, let's consider the general situation of an ordered set, $S = \{s_1, s_2, \ldots, s_t\}$, and a sequence $\beta = s_{i_1}s_{i_2}\ldots s_{i_k}$ formed from these elements. To the element s_{i_1} we attach the tableau consisting of one box, and the one entry, s_{i_1}. If $s_{i_2} \geq s_{i_1}$, then we attach the one-rowed tableau consisting of s_{i_1} and s_{i_2}; if $s_{i_2} < s_{i_1}$, we attach the two-rowed tableau having s_{i_2} in the top row, and s_{i_1} in the second row. Suppose we have attached by this procedure a standard tableau, λ, to the first l elements of the sequence. We then take the element $s_{i_{l+1}}$ and try to add it to λ. If it exceeds (in the weak sense) every element in the top row of λ, then we stick it onto the end of that row. If it doesn't, then we look for the first element of the first row that is strictly greater than it, and replace it by $s_{i_{l+1}}$. This leaves us with the exiled element of the first row, and we take it and look at the second row. If it fits, we add it on to the second row, otherwise bump as before and continue in this way. By this procedure, we arrive at the tableau, $L(\beta)$, associated to the sequence, β.

What if, one may ask, the first two rows of λ were of the same length and the bumped element exceeded all the elements of the second row? Wouldn't sticking it onto the second row produce an inadmissible shape? But we have assumed that the tableau, λ, is standard, and so we see that this situation cannot arise.

To associate a "place" tableau to the sequence β, we have to be given a bit more information along with the sequence (as when we saw in our example that we were dealing with two factors—hence two places—and subsequences of fixed length). So let us assume that our integer $k = k_1 + \cdots + k_n$ and that our sequence β has the property that the first k_1 elements are increasing, the next k_2 are increasing, and so on. We want to assign to our β a standard place tableau having n places. Notice that since the first k_1 elements are increasing, they all fit into a single-row tableau with k_1 boxes. We record this by attaching a single-rowed tableau with k_1 boxes all filled in with the place 1. We then run through the next k_2 elements, keeping track of all the new boxes created in the $L(\beta)$ construction by labeling them with the place 2. Notice that if elements from the second strand bump elements from the first row, the elements they bump grow in size, so that these elements all go into the second row (they bump no more). Also, once an element from this second strand gets placed in the first row, all the others do. This means that the place tableau so far associated has 1's in the first row, and 2's in the first and second rows. Obviously the number of 2's in the second row cannot exceed the number of 1's in the first row (namely, k_1). We then proceed with the third strand placing 3's, and so on.

Rather than spend more words on this description, a not too trivial example may be in order. For simplicity, we will use numbers for letters as well as places; we believe that this should cause no confusion.

Example 2. Consider the element $223446 \otimes 12335 \otimes 3446$ in $D_6 \otimes D_5 \otimes D_4$. Following the recipes above, one ends with the following pair of tableaux (the first one the "letter" tableau, the second the "place" tableau filled with only three places):

1	2	2	3	3	3	4	4	6
2	3	4	4	5				
6								

and

1	1	1	1	1	1	3	3	3
2	2	2	2	2				
3								

.

The original "sequence" (we write it in quotes, since we've taken the liberty to use the tensor product symbol to mark off where the substrands are to be seen) can be reconstructed from this pair of tableaux as we indicated in the first example. By removing all the entries labeled 3 in the first row, we see that the third factor has to end with 446. But then we see that the 6 must have been bumped into the third row by the first term of the third factor. If 6 had been bumped from the second row, it would have had to be the 5 that did it, but if the 5 had been the bumper, it would have already fit nicely onto the first row. So it must have been the 3 that bumped, which means that the third factor was 3446, and that 3 bumped into the two-rowed tableau having 122335 in the top row and 23446 in the second. The corresponding place tableau now has only the places 1 and 2 with all the entries labeled 1 in the top row and all those labeled 2 in the second. Now the previous tableau had six entries in the top row and only four in the second. The only way we could have arrived at the current stage is if we had bumped a 5 into the tableau having 122346 in the top row, 2344 in the second, and so forth. In this way we reconstruct our original sequence of departure.

This procedure establishes a one-to-one correspondence between the usual basis elements of $D_{k_1}(F) \otimes \cdots \otimes D_{k_n}(F)$ and the double standard tableaux having k_i places i for $i = 1, \ldots, n$. This, then finishes our discussion of the linear independence of the double standard tableaux, and hence of Theorem 2.

References

1. Akin, K., Buchsbaum, D.: Characteristic-free representation theory of the general linear group. Adv. Math. **58**, 149–200 (1985)
2. Akin, K., Buchsbaum, D.A.: Characteristic-free representation theory of the general linear group. II. Homological considerations. Adv. Math. **72**(2), 171–210 (1988)
3. Akin, K., Buchsbaum, D.: Representations, resolutions and intertwining numbers. Communications in Algebra, pp. 1–19. MSRI Publications, Springer, New York (1989)
4. Akin, K., Buchsbaum, D.A., Weyman, J.: Resolutions of determinantal ideals: the submaximal minors. Adv. Math. **39**, 1–10 (1981)

5. Akin, K., Buchsbaum, D.A., Weyman, J.: Schur functors and Schur complexes. Adv. Math. **44**, 207–278 (1982)
6. Boffi, G., Buchsbaum, D.A.: Homotopy equivalence of two families of complexes. Trans. Am. Math. Soc. **356**, 3077–3107 (2004)
7. Boffi, G., Buchsbaum, D.A.: Threading Homology Through Algebra: Selected Patterns. Oxford University Press, Oxford (2006)
8. Buchsbaum, D.A.: A generalized Koszul complex, I. Trans. Amer. Math Soc. **111**, 183–196 (1964)
9. Buchsbaum, D., Eisenbud, D.: Some structure theorems for finite free resolutions. Adv. Math. **12**, 84–139 (1974)
10. Buchsbaum, D.A., Eisenbud, D.: Generic free resolutions and a family of generically perfect ideals. Adv. Math. **18**, 245–301 (1975)
11. Buchsbaum, D.A., Rota, G.-C.: Projective resolutions of Weyl modules. Proc. Nat. Acad. Sci. USA **90**, 2448–2450 (1993)
12. Buchsbaum, D.A., Rota, G.-C.: A new construction in homological algebra. Proc. Nat. Acad. Sci. USA **91**, 4115–4119 (1994)
13. Buchsbaum, D.A., Rota, G.-C.: Approaches to resolution of Weyl modules. Adv. App. Math. **27**, 82–191 (2001)
14. Carter, R.W., Lusztig, G.: On the modular representations of the general linear and symmetric groups. Math. Z. **136**, 193–242 (1974)
15. Desarmenien, J., Kung, J., Rota, G.-C.: Invariant theory, Young tableaux and combinatorics. Adv. Math. **27**, 63–92 (1978)
16. Doubilet, P., Rota, G.-C., Stein, J.A.: On the foundations of combinatorial theory. Stud. Appl. Math. **53**, 185–216 (1976)
17. Eagon, J., Northcott, D.G.: Ideals defined by matrices and a certain complex associated with them. Proc. Roy. Soc. London Ser. A **269**, 188–204 (1962)
18. Grosshans, F.D., Rota, G.-C., Stein, J.A.: Invariant Theory and Superalgebras. American Mathematical Society, Providence (1987)
19. Hashimoto, M.: Determinantal ideals without minimal free resolutions. Nagoya Math. J. **118**, 203–216 (1990)
20. Huang, R.Q., Zhang, J.J.: Standard basis theorem for quantum linear groups. Adv. Math. **102**, 202–229 (1993)
21. Lascoux, A.: Syzygies des variétés déterminantales. Adv. Math. **30**, 202–237 (1978)
22. Taylor, B.D.: Generalized straightening laws for products of determinants. Ph.D. Thesis, MIT (1997)
23. Taylor, B.D.: A straightening algorithm for row-convex tableaux. J. Alg. **236**, 155–191 (2001)
24. Towber, J.: Two new functors from modules to algebras. J. Alg. **47**, 80–109 (1977)

Koszul Algebras and Regularity

Aldo Conca, Emanuela De Negri, and Maria Evelina Rossi

1 Introduction

This is a chapter on commutative Koszul algebras and Castelnuovo–Mumford regularities. Koszul algebras, originally introduced by Priddy [48], are graded K-algebras R whose residue field K has a linear free resolution as an R-module. Here linear means that the nonzero entries of the matrices describing the maps in the resolution have degree 1. For example, over the symmetric algebra $S = \mathrm{Sym}_K(V)$ of a finite dimensional K-vector space V, the residue field K, is resolved by the Koszul complex which is linear. Similarly, for the exterior algebra $\bigwedge_K V$ the residue field K is resolved by the Cartan complex which is also linear. In this chapter we deal mainly with standard graded commutative K-algebras, that is, quotient rings of the polynomial ring S by homogeneous ideals. The (absolute) Castelnuovo–Mumford regularity $\mathrm{reg}_S(M)$ is, after Krull dimension and multiplicity, perhaps the most important invariant of a finitely generated graded S-module M, as it controls the vanishing of both syzygies and the local cohomology modules of M. By definition, $\mathrm{reg}_S(M)$ is the least integer r such that the ith syzygy module of M is generated in degrees $\leq r + i$ for every i. By local duality, $\mathrm{reg}_S(M)$ can be characterized also as the least number r such that the local cohomology module $H^i_{\mathbf{m}_S}(M)$ vanishes in degrees $> r - i$ for every i. Analogously when $R = S/I$ is a standard graded K-algebra and M is a finitely generated graded R-module one can define the relative Castelnuovo–Mumford regularity as the least integer r such that the ith syzygy module over R of M is generated in degrees $\leq r + i$ for every i. The main difference between the relative and the absolute regularity is that over R most of the resolutions are infinite, that is, there are infinitely many syzygy modules, and hence it is not at all clear whether $\mathrm{reg}_R(M)$ is finite. Avramov, Eisenbud and

A. Conca (✉) • E. De Negri • M.E. Rossi
Dipartimento di Matematica, Università degli Studi di Genova, Genova, Italy
e-mail: conca@dima.unige.it; denegri@dima.unige.it; rossim@dima.unige.it

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_8,

Peeva gave in [5, 6] a beautiful characterization of the Koszul property in terms of the relative regularity: R is Koszul iff $\operatorname{reg}_R(M)$ is finite for every M iff $\operatorname{reg}_R(K)$ is finite.

From certain point of views, Koszul algebras behave homologically as polynomial rings. For instance $\operatorname{reg}_R(M)$ can be characterized in terms of regularity of truncated submodules (see Proposition 8). On the other hand, "bad" homological behaviors may occur over Koszul algebras. For instance, modules might have irrational Poincaŕe series over Koszul algebras. Furthermore, Koszul algebras appear quite frequently among the rings that are classically studied in commutative algebra, algebraic geometry and combinatorial commutative algebra. This mixture of similarities and differences with the polynomial ring and their frequent appearance in classical constructions are some of the reasons that make Koszul algebras fascinating, studied and beloved by commutative algebraists and algebraic geometers. In few words, a homological life is worth living in a Koszul algebra. Of course there are other reasons for the popularity of Koszul algebras in the commutative and noncommutative setting, as, for instance, Koszul duality, a phenomenon that generalizes the duality between the symmetric and the exterior algebra (see [13, 14, 50]).

The structure of this chapter is the following. Section 2 contains the characterization, due to Avramov, Eisenbud and Peeva, of Koszul algebras in terms of the finiteness of the regularity of modules (see Theorem 7). It contains also the definition of G-quadratic and LG-quadratic algebras and some fundamental questions concerning the relationships between these notions and the syzygies of Koszul algebras (see Questions 12 and 14).

In Sect. 3 we present three elementary but powerful methods for proving that an algebra is Koszul: the existence of a Gröbner basis of quadrics, the transfer of Koszulness to quotient rings and Koszul filtrations. To illustrate these methods we apply them to Veronese algebras and Veronese modules. We prove that Veronese subalgebras of Koszul algebras are Koszul and that high-enough Veronese subalgebras of any algebra are Koszul. These and related results were proved originally in [3, 11, 12, 25, 32].

Section 4 is devoted to two very strong versions of Koszulness: universally Koszul [21] and absolutely Koszul [43]. An algebra R is universally Koszul if for every ideal $I \subset R$ generated by elements of degree 1 one has $\operatorname{reg}_R(I) = 1$. Given a graded R-module M and $i \in \mathbf{Z}$ one defines $M_{\langle i \rangle}$ as the submodule of M generated by the homogeneous component M_i of degree i of M. The R-module M is componentwise linear if $\operatorname{reg}_R(M_{\langle i \rangle}) = i$ for every i with $M_i \neq 0$. The K-algebra R is absolutely Koszul if any finitely generated graded R-module M has a componentwise linear ith syzygy module for some $i \geq 0$. Two major achievements are the complete characterization of the Cohen–Macaulay domains that are universally Koszul (see [21] or Theorem 4) and the description of two classes of absolutely Koszul algebras (see [43] or Theorem 10). We also present some questions related to these notions, in particular Questions 13 and 14.

In Sect. 5 we discuss some problems regarding the regularity of modules over Koszul algebras. Some are of computational nature, for instance Question 12, and others are suggested by the analogy with the polynomial ring, for example, Question 9. This section contains also some original results, in particular Proposition 5 and Theorem 11, motivating the questions presented.

Finally Sect. 6 contains a discussion on local variants of the notion of Koszul algebra and the definition of Koszul modules. A local ring $(R, \mathbf{m}, K)$ is called a Koszul ring if the associated graded ring $\mathrm{gr}_{\mathbf{m}}(R)$ is Koszul as a graded K-algebra. The ring R is called Fröberg if its Poincaré series equals to $H_R(-z)^{-1}$, where $H_R(z)$ denotes the Hilbert series of R. Any Koszul ring is Fröberg. The converse holds in the graded setting and is unknown in the local case (see Question 5). Large classes of local rings of almost minimal multiplicity are Koszul. In [41] and [43] a characterization of Koszulness of graded algebras is obtained in terms of the finiteness of the linear defect of the residue field (see Proposition 12). It is an open problem whether the same characterization holds in the local case too (see Question 13).

2 Generalities

Let K be a field and R be a (commutative) standard graded K-algebra, that is, a K-algebra with a decomposition $R = \oplus_{i \in \mathbf{N}} R_i$ (as an Abelian group) such that $R_0 = K$, R_1 is a finite dimensional K-vector space and $R_i R_j = R_{i+j}$ for every $i, j \in \mathbf{N}$. Let S be the symmetric algebra over K of R_1. One has an induced surjection

$$S = \mathrm{Sym}_K(R_1) \to R \tag{1}$$

of standard graded K-algebras. We call Eq. (1) the canonical presentation of R. Hence R is isomorphic (as a standard graded K-algebra) to S/I where I is the kernel of Eq. (1). In particular, I is homogeneous and does not contain elements of degree 1. We say that I defines R. Choosing a K-basis of R_1 the symmetric algebra S gets identified with the polynomial ring $K[x_1, \ldots, x_n]$, with $n = \dim_K R_1$, equipped with its standard graded structure (i.e., $\deg x_i = 1$ for every i). Denote by $\mathbf{m}_R$ the maximal homogeneous ideal of R. We may consider K as a graded R-module via the identification $K = R/\mathbf{m}_R$.

Assumption. With the exception of the last section, K-algebras are always assumed to be standard graded, modules and ideals are graded and finitely generated, and module homomorphisms have degree 0.

For an R-module $M = \oplus_{i \in \mathbf{Z}} M_i$ we denote by $\mathrm{HF}(M, i)$ the Hilbert function of M at i, that is, $\mathrm{HF}(M, i) = \dim_K M_i$ and by $H_M(z) = \sum \dim_K M_i z^i \in \mathbf{Q}[|z|][z^{-1}]$ the associated Hilbert series.

Recall that a minimal graded free resolution of M as an R-module, is a complex of free R-modules

$$\mathbf{F} : \cdots \to F_{i+1} \xrightarrow{\phi_{i+1}} F_i \xrightarrow{\phi_i} F_{i-1} \to \cdots \to F_1 \xrightarrow{\phi_1} F_0 \to 0$$

such that $H_i(\mathbf{F}) = 0$ for $i > 0$ and $H_0(\mathbf{F}) = M$, Image $\phi_{i+1} \subseteq \mathbf{m}_R F_i$ for every i. Such a resolution exists and it is unique up to an isomorphism of complexes, that is why we usually talk of "the" minimal free (graded) resolution of M. By definition, the ith Betti number $\beta_i^R(M)$ of M as an R-module is the rank of F_i. Each F_i is a direct sum of shifted copies of R. The (i,j)th graded Betti number $\beta_{ij}^R(M)$ of M is the number of copies of $R(-j)$ that appear in F_i. By construction one has $\beta_i^R(M) = \dim_K \operatorname{Tor}_i^R(M,K)$ and $\beta_{ij}^R(M) = \dim_K \operatorname{Tor}_i^R(M,K)_j$. The Poincaré series of M is defined as

$$P_M^R(z) = \sum_i \beta_i^R(M) z^i \in \mathbf{Q}[|z|],$$

and its bigraded version is

$$P_M^R(s,z) = \sum_{i,j} \beta_{i,j}^R(M) z^i s^j \in \mathbf{Q}[s][|z|].$$

We set

$$t_i^R(M) = \sup\{j : \beta_{ij}^R(M) \neq 0\}$$

where, by convention, $t_i^R(M) = -\infty$ if $F_i = 0$. By definition, $t_0^R(M)$ is the largest degree of a minimal generator of M. Two important invariants that measure the "growth" of the resolution of M as an R-module are the projective dimension

$$\operatorname{pd}_R(M) = \sup\{i : F_i \neq 0\} = \sup\{i : \beta_{ij}^R(M) \neq 0 \text{ for some } j\}$$

and the Castelnuovo–Mumford regularity

$$\operatorname{reg}_R(M) = \sup\{j - i : \beta_{ij}^R(M) \neq 0\} = \sup\{t_i^R(M) - i : i \in \mathbf{N}\}.$$

We may as well consider M as a module over the polynomial ring S via Eq. (1). The regularity $\operatorname{reg}_S(M)$ of M as an S-module has also a cohomological interpretation via local duality (see , e.g. [15, 31]). Denoting by $H^i_{\mathbf{m}_S}(M)$ the ith local cohomology module with support on the maximal ideal of S one has

$$\operatorname{reg}_S(M) = \max\{j + i : H^i_{\mathbf{m}_S}(M)_j \neq 0\}.$$

Since $H^i_{\mathbf{m}_R}(M) = H^i_{\mathbf{m}_S}(M)$ for every i, nothing changes if on right-hand side of the formula above we replace S with R. So $\operatorname{reg}_S(M)$ is in some sense the "absolute" Castelnuovo–Mumford regularity. Both $\operatorname{pd}_R(M)$ and $\operatorname{reg}_R(M)$ can be infinite.

Example 1. Let $R = K[x]/(x^3)$ and $M = K$. Then $F_{2i} = R(-3i)$ and $F_{2i+1} = R(-3i-1)$ so that $\operatorname{pd}_R(M) = \infty$ and $\operatorname{reg}_R(M) = \infty$.

Note that, in general, $\operatorname{reg}_R(M)$ is finite if $\operatorname{pd}_R M$ is finite, but, as we will see, not the other way round.

In the study of minimal free resolutions over R, the resolution $\mathbf{K}_R$ of the residue field K as an R-module plays a prominent role. This is because $\operatorname{Tor}_*^R(M,K) =$

$H_*(M \otimes \mathbf{K}_R)$ and hence $\beta_{ij}^R(M) = \dim_K H_i(M \otimes \mathbf{K}_R)_j$. A very important role is played also by the Koszul complex $K(\mathbf{m}_R)$ on a minimal system of generators of the maximal ideal $\mathbf{m}_R$ of R.

When is $\mathrm{pd}_R(M)$ finite for every M? The answer is given by one of the most classical results in commutative algebra: the Auslander–Buchsbaum–Serre Theorem. We present here the graded variant of it that can be seen as a strong version of the Hilbert syzygy theorem.

Theorem 2. *The following conditions are equivalent:*

(1) $\mathrm{pd}_R M$ *is finite for every* R*-module* M.
(2) $\mathrm{pd}_R K$ *is finite.*
(3) R *is regular, that is,* R *is a polynomial ring.*

When the conditions hold, then for every M*, one has* $\mathrm{pd}_R M \leq \mathrm{pd}_R K = \dim R$*, and the Koszul complex* $K(\mathbf{m}_R)$ *resolves* K *as an* R*-module, that is,* $\mathbf{K}_R \cong K(\mathbf{m}_R)$.

Remark 3. The Koszul complex $K(\mathbf{m}_R)$ has three important features:

(1) It is finite.
(2) It has an algebra structure. Indeed it is a DG-algebra and this has important consequences such as the algebra structure on the Koszul cycles and Koszul homology. See [4] for the definition (and much more) on DG-algebras.
(3) The matrices describing its differentials have nonzero entries only of degree 1.

When R is not a polynomial ring $\mathbf{K}_R$ does not satisfy condition (1) in Remark 3. Can $\mathbf{K}_R$ nevertheless satisfy (2) or (3) of Remark 3?

For (2) the answer is yes: $\mathbf{K}_R$ has always a DG-algebra structure. Indeed a theorem, proved independently by Gulliksen and Schoeller (see [4, 6.3.5]), asserts that $\mathbf{K}_R$ is obtained by the so-called Tate construction. This procedure starts from $K(\mathbf{m}_R)$ and builds $\mathbf{K}_R$ by "adjoining variables to kill homology" while preserving the DG-algebra structure (see [4, 6.3.5]).

Algebras R such that $\mathbf{K}_R$ satisfies condition (3) in Remark 3 in above are called Koszul:

Definition 4. The K-algebra R is Koszul if the matrices describing the differentials of $\mathbf{K}_R$ have nonzero entries only of degree 1, that is, $\mathrm{reg}_R(K) = 0$ or, equivalently, $\beta_{ij}^R(K) = 0$ whenever $i \neq j$.

Koszul algebras were originally introduced by Priddy [48] in his study of homological properties of graded (noncommutative) algebras arising from algebraic topology, leaving the commutative case "for the interested reader". In the recent volume [50] Polishchuk and Positselski present various surprising aspects of Koszulness. We collect below a list of important facts about Koszul commutative algebras. We always refer to the canonical presentation Eq. (1) of R. First we introduce a definition.

Definition 5. We say that R is G-quadratic if its defining ideal I has a Gröbner basis of quadrics with respect to some coordinate system of S_1 and some term order τ on S.

Remark 6. (1) If R is Koszul, then I is generated by quadrics (i.e., homogeneous polynomials of degree 2). Indeed, the condition $\beta^R_{2j}(K) = 0$ for every $j \neq 2$ is equivalent to the fact that I is defined by quadrics. But there are algebras defined by quadrics that are not Koszul. For example, $R = K[x, y, z, t]/I$ with $I = (x^2, y^2, z^2, t^2, xy + zt)$ has $\beta^R_{34}(K) = 5$.

(2) If I is generated by monomials of degree 2 with respect to some coordinate system of S_1, then a simple filtration argument that we reproduce in Sect. 3, (see Theorem 15) shows that R is Koszul in a very strong sense.

(3) If I is generated by a regular sequence of quadrics, then R is Koszul. This follows from a result of Tate [59] asserting that if R is a complete intersection, then $\mathbf{K}_R$ is obtained by $K(\mathbf{m}_R)$ by adding polynomial variables in homological degree 2 to kill $H_1(K(\mathbf{m}_R))$.

(4) If R is G-quadratic, then R is Koszul. This follows from (2) and from the standard deformation argument showing that $\beta^R_{ij}(K) \leq \beta^A_{ij}(K)$ with $A = S/\mathrm{in}_\tau(I)$.

(5) On the other hand there are Koszul algebras that are not G-quadratic. One notes that an ideal defining a G-quadratic algebra must contain quadrics of “low” rank. For instance, if R is Artinian and G-quadratic then its defining ideal must contain the square of a linear form. But most Artinian complete intersection of quadrics do not contain the square of a linear form. For example, $I = (x^2 + yz, y^2 + xz, z^2 + xy) \subset \mathbf{C}[x, y, z]$ is an Artinian complete intersection not containing the square of a linear form. Hence I defines a Koszul and not G-quadratic algebra. See [32] for a general result in this direction.

(6) The Poincaré series $P^R_K(z)$ of K as an R-module can be irrational, see [2]. However, for a Koszul algebra R, one has

$$P^R_K(z)H_R(-z) = 1, \tag{2}$$

and hence $P^R_K(z)$ is rational. Indeed the equality Eq. (2) turns out to be equivalent to the Koszul property of R, [37, 1]. A necessary (but not sufficient) numerical condition for R to be Koszul is that the formal power series $1/H_R(-z)$ has non-negative coefficients (indeed positive unless R is a polynomial ring). Another numerical condition is the following: expand $1/H_R(-z)$ as

$$\frac{\Pi_{h\in 2\mathbf{N}+1}(1+z^h)^{e_h}}{\Pi_{h\in 2\mathbf{N}+2}(1-z^h)^{e_h}}$$

with $e_h \in \mathbf{Z}$ (see [4, 7.1.1]). The numbers e_h are the “expected” deviations. If R is Koszul then $e_h \geq 0$ for every h, (indeed $e_h > 0$ for every h unless R is a complete intersection). For example, if $H(z) = 1+4z+5z^2$, then the coefficient of z^6 in $1/H(-z)$ is negative and the third expected deviation is 0. So for two reasons an algebra with Hilbert series $H(z)$, as the one in (1), cannot be Koszul.

The following characterization of the Koszul property in terms of regularity is formally similar to the Auslander–Buchsbaum–Serre Theorem 2.

Theorem 7 (Avramov–Eisenbud–Peeva). *The following conditions are equivalent:*

(1) $\operatorname{reg}_R(M)$ *is finite for every* R*-module* M.
(2) $\operatorname{reg}_R(K)$ *is finite.*
(3) R *is Koszul.*

Avramov and Eisenbud proved in [5] that every module has finite regularity over a Koszul algebra. Avramov and Peeva showed in [6] that if $\operatorname{reg}_R(K)$ is finite then it must be 0. Indeed they proved a more general result for graded algebras that are not necessarily standard.

If M is an R-module generated by elements of a given degree, say d, we say that it has a linear resolution over R if $\operatorname{reg}_R(M) = d$. For $q \in \mathbf{Z}$ we set $M_{\langle q\rangle}$ to be the submodule of M generated by M_q and set $M_{\geq q} = \oplus_{i\geq q} M_i$. The module M is said to be componentwise linear over R if $M_{\langle q\rangle}$ has a linear resolution for every q. The (absolute) regularity of a module can be characterized as follows:

$$\begin{aligned}\operatorname{reg}_S(M) &= \min\{q \in \mathbf{Z} : M_{\geq q} \text{ has a linear resolution}\}\\ &= \min\{q \geq t_0^S(M) : M_{\langle q\rangle} \text{ has a linear resolution}\}\end{aligned}$$

One of the motivations of Avramov and Eisenbud in [5] was to establish a similar characterization for the relative regularity over a Koszul algebra. They proved:

Proposition 8. *Let* R *be a Koszul algebra and* M *be an* R*-module. Then:*

$$\operatorname{reg}_R(M) \leq \operatorname{reg}_S(M)$$

and

$$\begin{aligned}\operatorname{reg}_R(M) &= \min\{q \in \mathbf{Z} : M_{\geq q} \textit{ has a linear } R\text{–resolution}\}\\ &= \min\{q \geq t_0^R(M) : M_{\langle q\rangle} \textit{ has a linear } R - \text{resolution}\}.\end{aligned}$$

Another invariant that measures the growth of the degrees of the syzygies of a module is the slope:

$$\operatorname{slope}_R(M) = \sup\left\{\frac{t_i^R(M) - t_0^R(M)}{i} : i > 0\right\}.$$

A useful feature of the slope is that it is finite (no matter if R is Koszul or not). Indeed with respect to the canonical presentation Eq. (1), one has

$$\operatorname{slope}_R(M) \leq \max\{\operatorname{slope}_S(R), \operatorname{slope}_S(M)\}$$

(see [8, 1.2]), and the right-hand side is finite since S is a polynomial ring. Backelin defined in [10] the (Backelin) rate of R to be

$$\mathrm{Rate}(R) = \mathrm{slope}_R(\mathbf{m}_R)$$

as a measure of the failure of the Koszul property. By the very definition, one has $\mathrm{Rate}(R) \geq 1$ and R is Koszul if and only if $\mathrm{Rate}(R) = 1$.

We close the section with a technical lemma:

Lemma 9. *(1) Let* $0 \to M_1 \to M_2 \to M_3 \to 0$ *be a short exact sequence of* R*-modules. Then one has*

$$\begin{aligned} \mathrm{reg}_R(M_1) &\leq \max\{\mathrm{reg}_R(M_2), \mathrm{reg}_R(M_3) + 1\}, \\ \mathrm{reg}_R(M_2) &\leq \max\{\mathrm{reg}_R(M_1), \mathrm{reg}_R(M_3)\}, \\ \mathrm{reg}_R(M_3) &\leq \max\{\mathrm{reg}_R(M_1) - 1, \mathrm{reg}_R(M_2)\}. \end{aligned}$$

(2) Let

$$\mathbf{M} : \cdots \to M_i \to \cdots \to M_2 \to M_1 \to M_0 \to 0$$

be a complex of R*-modules. Set* $H_i = H_i(\mathbf{M})$*. Then for every* $i \geq 0$ *one has*

$$t_i^R(H_0) \leq \max\{a_i, b_i\}$$

where $a_i = \max\{t_j^R(M_{i-j}) : j = 0, \dots, i\}$ *and* $b_i = \max\{t_j^R(H_{i-j-1}) : j = 0, \dots, i-2\}$.

Moreover one has

$$\mathrm{reg}_R(H_0) \leq \max\{a, b\}$$

where $a = \sup\{\mathrm{reg}_R(M_j) - j : j \geq 0\}$ *and* $b = \sup\{\mathrm{reg}_R(H_j) - (j+1) : j \geq 1\}$.

Proof. (1) follows immediately by considering the long exact sequence obtained by applying $\mathrm{Tor}(K, -)$. For (2) one breaks the complex into short exact sequences and proves by induction on i the inequality for $t_i^R(H_0)$. Then one deduces the second inequality by translating the first into a statement about regularities. □

We collect below some problems about the Koszul property and the existence of Gröbner bases of quadrics. Let us recall the following.

Definition 10. A K-algebra R is LG-quadratic if there exists a G-quadratic algebra A and a regular sequence of linear forms $y_1, \dots, y_c$ such that $R \simeq A/(y_1, \dots, y_c)$.

We have the following implications:

$$\text{G-quadratic} \Rightarrow \text{LG-quadratic} \Rightarrow \text{Koszul} \Rightarrow \text{quadratic} \tag{3}$$

As discussed in Remark 6 the third implication in Eq. (3) is strict. The following remark, due to Caviglia, in connection with Remark 6(5) shows that also the first implication in Eq. (3) is strict.

Remark 11. Any complete intersection R of quadrics is LG-quadratic.
Say $R = K[x_1, \dots, x_n]/(q_1, \dots, q_m)$ then set

$$A = R[y_1, \dots, y_m]/(y_1^2 + q_1, \dots, y_m^2 + q_m)$$

and note that A is G-quadratic for obvious reasons and $y_1, \dots, y_m$ is a regular sequence in A by codimension considerations.

But we do not know an example of a Koszul algebra that is not LG-quadratic. So we ask:

Question 12. Is any Koszul algebra LG-quadratic?

Our feeling is that the answer should be negative. But how can we exclude that a Koszul algebra is LG-quadratic? One can look at the h-vector (i.e., the numerator of the Hilbert series) since it is invariant under Gröbner deformation and modifications as the one involved in the definition of LG-quadratic. Alternatively one can look at syzygies over the polynomial ring because they can only grow under such operations. These observations lead to a new question:

Question 13. Is the h-vector of any Koszul algebra R the h-vector of an algebra defined by quadratic monomials? And, if yes, does there exist an algebra A with quadratic monomial relations, h-vector equal to that of R and satisfying $\beta_{ij}^S(R) \leq \beta_{ij}^T(A)$ for every i and j? Here S and T denote the polynomial rings canonically projecting onto R and A.

A negative answer to Question 13 would imply a negative answer to Question 12. Note that any h-vector of an algebra defined by quadratic monomials is also the h-vector of an algebra defined by square-free quadratic monomials (by using the polarization process). The simplicial complexes associated to square-free quadratic monomial ideals are called flag. There has been a lot of activity concerning combinatorial properties and characterizations of h-vectors and f-vectors of flag simplicial complexes, see [28] for recent results and for a survey of what is known and conjectured. Here we just mention that Frohmader has proved in [39] a conjecture of Kalai asserting that the f-vectors of flag simplicial complexes are f-vectors of balanced simplicial complexes.

Regarding the inequality for Betti numbers in Question 13, LG-quadratic algebras R satisfy the following restrictions:

1. $t_i^S(R) \leq 2i$
2. $t_i^S(R) < 2i$ if $t_{i-1}^S(R) < 2(i-1)$
3. $t_i^S(R) < 2i$ if $i > \dim S - \dim R$
4. $\beta_i^S(R) \leq \binom{\beta_1^S(R)}{i}$

deduced from the deformation to the (non-minimal) Taylor resolution of quadratic monomial ideals (see for instance [47, 4.3.2]). As shown in [8] the same restrictions are satisfied by any Koszul algebra, with the exception of possibly (4). So we ask:

Question 14. Let R be a Koszul algebra quotient of the polynomial ring S. Is it true that $\beta_i^S(R) \leq \binom{\beta_1^S(R)}{i}$?

It can be very difficult to decide whether a given Koszul algebra is G-quadratic. In the 1990s, Peeva and Sturmfels asked whether the coordinate ring

$$PV = K[x^3, x^2y, x^2z, xy^2, xz^2, y^3, y^2z, yz^2, z^3]$$

of the pinched Veronese is Koszul. For about a decade this was a benchmark example for testing new techniques for proving Koszulness. In 2009 Caviglia [18] gave the first proof of the Koszulness of PV. Recently a new one has been presented in [19] that applies also to a larger family of rings including all the general projections to $\mathbf{P}^8$ of the Veronese surface in $\mathbf{P}^9$. The problem remains to decide whether:

Question 15. Is PV G-quadratic?

The answer is negative if one considers the toric coordinates only (as it can be checked by computing the associated Gröbner fan using CaTS [1]), but unknown in general. There are plenty of quadratic monomial ideals defining algebras with the Hilbert function of PV and larger Betti numbers.

The algebra PV is generated by all monomials in $n = 3$ variables of degree $d = 3$ that are supported on at most $s = 2$ variables. By varying the indices n, d, s one gets a family of pinched Veronese algebras $PV(n, d, s)$, and it is natural to ask:

Question 16. For which values of n, d, s is $PV(n, d, s)$ quadratic or Koszul?

Not all of them are quadratic, for instance, $PV(4, 5, 2)$ is not. Questions as Question 16 are very common in the literature: in a family of algebras one asks which ones are quadratic or Koszul or if quadratic and Koszul are equivalent properties for the algebras in the family. For example, in [26, 6.10] the authors ask:

Question 17. Let R be a quadratic Gorenstein algebra with Hilbert series $1 + nz + nz^2 + z^3$. Is R Koszul?

For $n = 3$ the answer is obvious as R must be a complete intersection of quadrics and for $n = 4$ the answer is positive by [26, 6.15]. See Theorem 12 for results concerning this family of algebras.

3 How to Prove that an Algebra is Koszul?

To prove that an algebra is Koszul is usually a difficult task. There are examples, due to Roos, showing that a sort of Murphy's law (anything that can possibly go wrong, does) holds in this context. Indeed there exists a family of quadratic algebras $R(a)$ depending on an integer $a > 1$ such that the Hilbert series of $R(a)$ is $1 + 6z + 8z^2$ for every a. Moreover K has a linear resolution for a steps and a nonlinear syzygy in homological position $a + 1$ (see [51]). So there is no statement of the kind: if R is an algebra with Hilbert series H then there is a number N depending on H, such that if the resolution of K over R is linear for N steps, it will be linear forever.

The goal of this section is to present some techniques to prove that an algebra is Koszul (without pretending they are the most powerful or interesting). For the sake of illustration we will apply these techniques to discuss the Koszul properties of Veronese algebras and modules. The material we present is taken from various sources (see [3, 8, 10–12, 16, 24–27, 32, 42, 58]).

3.1 Gröbner Basis of Quadrics

The simplest way to prove that an algebra is Koszul is to show that it is G-quadratic. A weak point of this prospective is that Gröbner bases refer to a system of coordinates and a term order. As said earlier, not all the Koszul algebras are G-quadratic. On the other hand many of the classical constructions in commutative algebra and algebraic geometry lead to algebras that have a privileged, say, natural, system of coordinates. For instance, the coordinate ring of the Grassmannian comes equipped with the Plücker coordinates. Toric varieties come with their toric coordinates. So one looks for a Gröbner basis of quadrics with respect to the natural system of coordinates. It turns out that many of the classical algebras (Grassmannian, Veronese, Segre, etc..) do have Gröbner bases of quadrics in the natural system of coordinates. Here we treat in details the Veronese case:

Theorem 1. *Let $S = K[x_1, \dots, x_n]$ and $c \in \mathbf{N}$. Then the Veronese subring $S^{(c)} = \oplus_{j \in \mathbf{N}} S_{jc}$ is defined by a Gröbner basis of quadrics.*

Proof. For $j \in \mathbf{N}$ denote by M_j the set of monomials of degree j of S. Consider $T_c = \mathrm{Sym}_K(S_c) = K[t_m : m \in M_c]$ and the surjective map $\Phi : T_c \to S^{(c)}$ of K-algebras with $\Phi(t_m) = m$ for every $m \in M_c$. For every monomial m we set $\max(m) = \max\{i : x_i | m\}$ and $\min(m) = \min\{i : x_i | m\}$. Furthermore for monomials $m_1, m_2 \in M_c$ we set $m_1 \prec m_2$ if $\max(m_1) \leq \min(m_2)$. Clearly $\prec$ is a transitive (but not reflexive) relation. We say that $m_1, m_2 \in M_c$ are incomparable if $m_1 \not\prec m_2$ and $m_2 \not\prec m_1$ and that are comparable otherwise. For a pair of incomparable elements $m_1, m_2 \in M_c$, let $m_3, m_4 \in M_c$ be the uniquely determined elements in M_c such that $m_1 m_2 = m_3 m_4$ and $m_3 \prec m_4$. Set

$$F(m_1, m_2) = t_{m_1} t_{m_2} - t_{m_3} t_{m_4}.$$

By construction $F(m_1, m_2) \in \mathrm{Ker}\Phi$ and we claim that the set of the $F(m_1, m_2)$'s is a Gröbner basis of $\mathrm{Ker}\Phi$ with respect to any term order τ of T_c such that $\mathrm{in}_\tau(F(m_1, m_2)) = t_{m_1} t_{m_2}$. Such a term order exists: order the $t_m's$ totally as follows:

$$t_u \geq t_v \text{ iff } u \geq v \text{ lexicographically}$$

and then consider the degree reverse lexicographic term order associated to that total order. Such a term order has the required property as it is easy to see. It remains to prove that the $F(m_1, m_2)$'s form a Gröbner basis of KerΦ. Set

$$U = (t_{m_1} t_{m_2} : m_1, m_2 \in M_c \text{ are incomparable})$$

By construction we have $U \subset \text{in}_\tau(\text{Ker}\Phi)$ and we have to prove equality. We do it by checking that the two associated quotients have the same Hilbert function. The inequality $\text{HF}(T_c/\text{Ker}\Phi, i) \leq \text{HF}(T_c/U, i)$ follows from the inclusion of the ideals. For the other note that

$$\text{HF}(T_c/\text{in}_\tau(\text{Ker}\Phi), i) = \text{HF}(T_c/\text{Ker}\Phi, i) = \text{HF}(S^{(c)}, i) = \#M_{ic}$$

The key observations are:

1. A monomial in the t's, say $t_{m_1} \cdots t_{m_i}$, is not in U if (after a permutation) $m_1 \prec m_2 \prec \cdots \prec m_i$.
2. Every monomial $m \in M_{ic}$ has a uniquely determined decomposition $m = m_1 \cdots m_i$ with $m_1 \prec m_2 \prec \cdots \prec m_i$.

   This implies that

$$\text{HF}(T_c/U, i) \leq \#M_{ic},$$

proving the desired assertion. □

3.2 Transfer of Koszulness

Let A be a K-algebra A and $B = A/I$ a quotient of it. Assume one of the two algebras is Koszul. What do we need to know about the relationship between A and B to conclude that the other algebra is Koszul too? Here is an answer:

Theorem 2. *Let A be a K-algebra and B be a quotient of A.*

(1) If $\text{reg}_A(B) \leq 1$ and A is Koszul, then B is Koszul.
(2) If $\text{reg}_A(B)$ is finite and B is Koszul, then A is Koszul.

The theorem is a corollary of the following:

Proposition 3. *Let A be a K-algebra and B a quotient algebra of A. Let M be a B-module. Then:*

(1) $\text{reg}_A(M) \leq \text{reg}_B(M) + \text{reg}_A(B)$.
(2) *If* $\text{reg}_A(B) \leq 1$ *then* $\text{reg}_B(M) \leq \text{reg}_A(M)$.

Proof. One applies Lemma 9(2) to the minimal free resolution **F** of M as a B-module and one has:

$$\text{reg}_A(M) \leq \sup\{\text{reg}_A(F_j) - j : j \geq 0\}.$$

Since $\operatorname{reg}_A(F_j) = \operatorname{reg}_A(B) + t_j^B(M)$, we can conclude that (1) holds.

For (2) it is enough to prove that the inequality

$$t_i^B(M) - i \le \max\{t_j^A(M) - j : j = 0, \dots, i\}$$

holds for every i. We argue by induction on i; the case $i = 0$ is obvious because $t_0^A(M) = t_0^B(M)$. Assume $i > 0$ and take a minimal presentation of M as a B-module

$$0 \to N \to F \to M \to 0$$

where F is B-free. Since $t_i^B(M) = t_{i-1}^B(N)$, by induction we have:

$$t_i^B(M) - i = t_{i-1}^B(N) - i \le \max\{t_j^A(N) - j - 1 : j = 0, \dots, i-1\}$$

Since $t_j^A(N) \le \max\{t_j^A(F), t_{j+1}^A(M)\}$ and $t_j^A(F) = t_j^A(B) + t_0^A(M) \le j + 1 + t_0^A(M)$ we may conclude that the desired inequality holds. □

Proof of Theorem 2. (1) Applying Proposition 3(1) with M equal to K one has that $\operatorname{reg}_A(K) \le \operatorname{reg}_A(B)$ which is finite by assumption. It follows then from Theorem 7 that A is Koszul. For (2) one applies Proposition 3(2) with $M = K$, and one gets $\operatorname{reg}_B(K) \le \operatorname{reg}_A(K)$ which is 0 by assumption; hence $\operatorname{reg}_B(K) = 0$ as required. □

Lemma 4. *Let R be Koszul algebra and M be an R-module. Then*

$$\operatorname{reg}_R(\mathbf{m}_R M) \le \operatorname{reg}_R(M) + 1.$$

In particular, $\operatorname{reg}_R(\mathbf{m}_R^u) = u$, *(unless $\mathbf{m}_R^u = 0$) that is, $\mathbf{m}_R^u$ has a linear resolution for every $u \in \mathbf{N}$.*

Proof. Apply Lemma 9 to the short exact sequence

$$0 \to \mathbf{m}_R M \to M \to M/\mathbf{m}_R M \to 0$$

and use the fact that $M/\mathbf{m}_R M$ is a direct sum of copies of K shifted at most by $-t_0^R(M)$. □

We apply now Theorem 2 to prove that the Veronese subrings of a Koszul algebra are Koszul.

Let $c \in \mathbf{N}$ and $R^{(c)} = \oplus_{j \in \mathbf{Z}} R_{jc}$ be the cth Veronese subalgebra of R. Similarly one defines $M^{(c)}$ for every R-module M. The formation of the cth Veronese submodule is an exact functor from the category of R-modules to the category of graded $R^{(c)}$-modules (recall that, by convention, modules are graded and maps are homogeneous of degree 0). For $u = 0, \dots, c-1$ consider the Veronese submodules $V_u = \oplus_{j \in \mathbf{Z}} R_{jc+u}$. Note that V_u is an $R^{(c)}$-module generated in degree 0 and that for $a \in \mathbf{Z}$ one has

$$R(-a)^{(c)} = V_u(-\lceil a/c \rceil)$$

where $u = 0$ if $a \equiv 0 \bmod(c)$ and $u = c - r$ if $a \equiv r \bmod(c)$ and $0 < r < c$.

Theorem 5. *Let R be Koszul. Then $R^{(c)}$ is Koszul and* $\operatorname{reg}_{R^{(c)}}(V_u) = 0$ *for every* $u = 0, \dots, c-1$.

Proof. Set $A = R^{(c)}$. First we prove that $\operatorname{reg}_A(V_u) = 0$ for every $u = 0, \dots, c-1$. To this end, we prove by induction on i that $t_i^A(V_u) \le i$ for every i. The case $i = 0$ is obvious. So assume $i > 0$. Let $M = \mathbf{m}_R^u(u)$. By Lemma 4 and by construction we have $\operatorname{reg}_R(M) = 0$ and $M^{(c)} = V_u$. Consider the minimal free resolution $\mathbf{F}$ of M over R and apply the functor $-^{(c)}$. We get a complex $\mathbf{G} = \mathbf{F}^{(c)}$ of A-modules such that $H_0(\mathbf{G}) = V_u$, $H_j(\mathbf{G}) = 0$ for $j > 0$ and $G_j = F_j^{(c)}$ is a direct sum of copies of $R(-j)^{(c)}$. Applying Lemma 9 we get $t_i^A(V_u) \le \max\{t_{i-j}^A(G_j) : j = 0, \dots, i\}$. Since G_0 is A-free we have $t_i^A(G_0) = -\infty$. For $j > 0$ we have $R(-j)^{(c)} = V_w(-\lceil j/c \rceil)$ for some number w with $0 \le w < c$. Hence, by induction, $t_{i-j}^A(G_j) \le i - j + \lceil j/c \rceil \le i$. Summing up,

$$t_i^A(V_u) \le \max\{i - j + \lceil j/c \rceil : j = 1, \dots, i\} = i.$$

In order to prove that A is Koszul we consider the minimal free resolution $\mathbf{F}$ of K over R and apply $-^{(c)}$. We get a complex $\mathbf{G} = \mathbf{F}^{(c)}$ of A-modules such that $H_0(\mathbf{G}) = K$, $H_j(\mathbf{G}) = 0$ for $j > 0$ and $G_j = F_j^{(c)}$ is a direct sum of copies of $V_u(-\lceil j/c \rceil)$. Hence $\operatorname{reg}_A(\mathbf{G}_j) = \lceil j/c \rceil$ and applying Lemma 9 we obtain

$$\operatorname{reg}_A(K) \le \sup\{\lceil j/c \rceil - j : j \ge 0\} = 0.$$

□

We also have:

Theorem 6. *Let R be a K-algebra, then the Veronese subalgebra $R^{(c)}$ is Koszul for $c \gg 0$. More precisely, if $R = A/I$ with A Koszul, then $R^{(c)}$ is Koszul for every* $c \ge \sup\{t_i^A(R)/(1+i) : i \ge 0\}$.

Proof. Let $\mathbf{F}$ be the minimal free resolution of R as an A-module. Set $B = A^{(c)}$ and note that B is Koszul because of Theorem 5. Then $\mathbf{G} = \mathbf{F}^{(c)}$ is a complex of B-modules such that $H_0(\mathbf{G}) = R^{(c)}$, $H_j(\mathbf{G}) = 0$ for $j > 0$. Furthermore $G_i = F_i^{(c)}$ is a direct sum of shifted copies of the Veronese submodules V_u. Using Theorem 5 we get the bound $\operatorname{reg}_B(G_i) \le \lceil t_i^A(R)/c \rceil$. Applying Lemma 9 we get

$$\operatorname{reg}_B(R^{(c)}) \le \sup\{\lceil t_i^A(R)/c \rceil - i : i \ge 0\}.$$

Hence, for $c \ge \sup\{t_i^A(R)/(1+i) : i \ge 0\}$ one has $\operatorname{reg}_B(R^{(c)}) \le 1$ and we conclude from Theorem 2 that $R^{(c)}$ is Koszul. □

Remark 7. (1) Note that the number $\sup\{t_i^A(R)/(1+i) : i \ge 0\}$ in Theorem 6 is finite. For instance it is less than or equal to $(\operatorname{reg}_A(R) + 1)/2$ which is finite because $\operatorname{reg}_A(R)$ is finite. Note however that $\sup\{t_i^A(R)/(1+i) : i \ge 0\}$ can be much smaller than $(\operatorname{reg}_A(R) + 1)/2$; for instance if $R = A/I$ with I generated by a regular sequence of r elements of degree d, then $t_i^A(R) = id$ so that $\operatorname{reg}_A(R) = r(d-1)$ while $\sup\{t_i^A(R)/(1+i) : i \geq 0\} = dr/(r+1)$.

(2) In particular, if we take the canonical presentation $R = S/I$ Eq. (1), then we have that $R^{(c)}$ is Koszul if $c \geq \sup\{t_i^S(R)/(1+i) : i \geq 0\}$. In [32, 2] it is proved that if $c \geq (\operatorname{reg}_S(R)+1)/2$, then $R^{(c)}$ is even G-quadratic. See [57] for other interesting results in this direction.

(3) Backelin proved in [10] that $R^{(c)}$ is Koszul if $c \geq \operatorname{Rate}(R)$.

(4) The proof of Theorem 6 shows also that $\operatorname{reg}_{A^{(c)}}(R^{(c)}) = 0$ if $c \geq \operatorname{slope}_A(R)$.

3.3 Filtrations

Another tool for proving that an algebra is Koszul is a "divide and conquer" strategy that can be formulated in various technical forms, depending on the goal one has in mind. We choose the following:

Definition 8. A Koszul filtration of a K-algebra R is a set $\mathscr{F}$ of ideals of R such that:

(1) Every ideal $I \in \mathscr{F}$ is generated by elements of degree 1.
(2) The zero ideal 0 and the maximal ideal $\mathbf{m}_R$ are in $\mathscr{F}$.
(3) For every $I \in \mathscr{F}$, $I \neq 0$, there exists $J \in \mathscr{F}$ such that $J \subset I$, I/J is cyclic and $\operatorname{Ann}(I/J) = J : I \in \mathscr{F}$.

By the very definition a Koszul filtration must contain a complete flag of R_1, that is, an increasing sequence $I_0 = 0 \subset I_1 \subset \cdots \subset I_{n-1} \subset I_n = \mathbf{m}_R$ such that I_i is minimally generated by i elements of degree 1. The case where $\mathscr{F}$ consists of just a single flag deserves a name:

Definition 9. A Gröbner flag for R is a Koszul filtration that consists of a single complete flag of R_1. In other words, $\mathscr{F} = \{I_0 = 0 \subset I_1 \subset \cdots \subset I_{n-1} \subset I_n = \mathbf{m}_R\}$ with $I_{i-1} : I_i \in \mathscr{F}$ for every i.

One has:

Lemma 10. *Let $\mathscr{F}$ be a Koszul filtration for R. Then one has:*

(1) $\operatorname{reg}_R(R/I) = 0$ *and R/I is Koszul for every $I \in \mathscr{F}$.*
(2) *R is Koszul.*
(3) *If $\mathscr{F}$ is a Gröbner flag, then R is G-quadratic.*

Proof. (1) and (2): One easily proves by induction on i and on the number of generators of I that $t_i^R(R/I) \leq i$ for every i and $I \in \mathscr{F}$. This implies that R is Koszul (take $I = \mathbf{m}_R$) and that $\operatorname{reg}_R(R/I) = 0$, hence R/I is Koszul by Theorem 2.

(3) We just sketch the argument: let $x_1, \ldots, x_n$ be a basis for the flag, that is, $\mathscr{F} = \{I_0 = 0 \subset I_1 \subset \cdots \subset I_{n-1} \subset I_n = \mathbf{m}_R\}$ and $I_i = (x_1, \ldots, x_i)$ for every i.

For every i there exists $j_i \geq i$ such that $(x_1, \ldots, x_i) : x_{i+1} = (x_1, \ldots, x_{j_i})$. For every $i < h \leq j_i$ the assertion $x_h x_{i+1} \in (x_1, \ldots, x_i)$ is turned into a quadratic equation in the defining ideal of R. The claim is that these quadratic equations form a Gröbner basis with respect to a term order that selects $x_h x_{i+1}$ as leading monomial. To prove the claim one shows that the identified monomials define an algebra, call it A, whose Hilbert function equals that of R. This is done by showing that the numbers $j_1, \ldots, j_n$ associated to the flag of R determine the Hilbert function of R and then by showing that also A has a Gröbner flag with associated numbers $j_1, \ldots, j_n$. □

There are Koszul algebras without Koszul filtrations and G-quadratic algebras without Gröbner flags see the examples given in [26, pp. 100 and 101]. Families of algebras having Koszul filtrations or Gröbner flags are described in [26]. For instance, it is proved that the coordinate ring of a set of at most $2n$ points in $\mathbf{P}^n$ in general linear position has a Gröbner flag, and that the general Gorenstein Artinian algebra with socle in degree 3 has a Koszul filtration. The results for points in [26] generalize results of [27,44] and are generalized in [49]. Filtrations of more general type are used in [24] to control the Backelin rate of coordinate rings of sets of points in the projective space.

The following notion is very natural for algebras with privileged coordinate systems (e.g. in the toric case).

Definition 11. An algebra R is said to be strongly Koszul if there exists a basis X of R_1 such that for every $Y \subset X$ and for every $x \in X \setminus Y$ there exists $Z \subseteq X$ such that $(Y) : x = (Z)$.

Our definition of strongly Koszul is slightly different than the one given in [42]. In [42] it is assumed that the basis X of R_1 is totally ordered, and in the definition one adds the requirement that x is larger than every element in Y.

Remark 12. If R is strongly Koszul with respect to a basis X of R_1, then the set $\{(Y) : Y \subseteq X\}$ is obviously a Koszul filtration.

We have:

Theorem 13. *Let $R = S/I$ with $S = K[x_1, \ldots, x_n]$ and $I \subset S$ an ideal generated by monomials of degrees $\leq d$. Then $R^{(c)}$ is strongly Koszul for every $c \geq d - 1$.*

Proof. In the proof we use the following basic facts:

Fact (1): If $m_1, \ldots, m_t, m$ are monomials of S, then $(m_1, \ldots, m_t) :_S m$ is generated by the monomials $m_i / \gcd(m_i, m)$ for $i = 1, \ldots, t$.

Fact (2): If T is an algebra and $A = T^{(c)}$, then for every ideal $I \subset A$ and $f \in A$ one has $IT \cap A = I$ and $(IT :_T f) \cap A = I :_A f$.

The first is an elementary and well-know, property of monomials; the second holds true because A is a direct summand of T.

Let $A = R^{(c)}$. Let X be the set of the residue classes in R of the monomials of degree c that are not in I. Clearly X is a basis of A_1. Let $Y \subset X$ and $z \in X \setminus Y$, say $Y = \{\bar{m}_1, \dots, \bar{m}_v\}$ and $z = \bar{m}$. We have to compute $(Y) :_A z$. To this end let us consider $J = (I + (m_1, \dots, m_v)) :_S m$ and note that $J = I + H$ with H a monomial ideal generated in degrees $\leq c$. Then $(Y) :_A z = (\bar{m} : m \in H \setminus I$ is a monomial of degree $c)$. □

Let us single out two interesting special cases:

Theorem 14. *Let $S = K[x_1, \dots, x_n]$. Then $S^{(c)}$ is strongly Koszul for every $c \in \mathbf{N}$.*

Theorem 15. *Let $S = K[x_1, \dots, x_n]$ and let $I \subset S$ be an ideal generated by monomials of degree 2. Then S/I is strongly Koszul.*

Given a Koszul filtration $\mathscr{F}$ for an algebra R we may also look at modules having a filtration compatible with $\mathscr{F}$. This leads us to the following:

Definition 16. Let R be an algebra with a Koszul filtration $\mathscr{F}$. Let M be an R-module. We say that M has linear quotients with respect to $\mathscr{F}$ if M is minimally generated by elements $m_1, \dots, m_v$ such that $\langle m_1, \dots, m_{i-1} \rangle :_R m_i \in \mathscr{F}$ for $i = 1, \dots, v$.

One easily deduces:

Lemma 17. *Let R be an algebra with a Koszul filtration $\mathscr{F}$ and M an R-module with linear quotients with respect to $\mathscr{F}$. Then $\operatorname{reg}_R(M) = t_0^R(M)$.*

As an example we have:

Proposition 18. *Let $S = K[x_1, \dots, x_n]$ and I be a monomial ideal generated in degree $\leq d$. Consider $R = S/I$ and the Veronese ring $R^{(c)}$ equipped with the Koszul filtration described in the proof of Theorem 13. For every $u = 0, \dots, c-1$ the Veronese module $V_u = \oplus_j R_{u+jc}$ has linear quotients with respect to $\mathscr{F}$.*

The proof is easy, again, based on Fact (1) in the proof of Theorem 13. In particular, this gives another proof of the fact that the Veronese modules V_u have a linear $R^{(c)}$-resolution.

The results and the proofs presented for Veronese rings and Veronese modules have their analogous in the multigraded setting (see [25]). For later applications we mention explicitly one case.

Let $S = K[x_1, \dots, x_n, y_1, \dots, y_m]$ with $\mathbf{Z}^2$-grading induced by the assignment $\deg(x_i) = (1, 0)$ and $\deg(y_i) = (0, 1)$. For every $c = (c_1, c_2)$ we look at the diagonal subalgebra $S_\Delta = \oplus_{a \in \Delta} S_a$ where $\Delta = \{ic : i \in \mathbf{Z}\}$. The algebra S_Δ is nothing but the Segre product of the c_1th Veronese ring of $K[x_1, \dots, x_n]$ and the c_2th Veronese ring of $K[y_1, \dots, y_m]$. We have:

Proposition 19. *For every $(a, b) \in \mathbf{Z}^2$ the S_Δ-submodule of S generated by $S_{(a,b)}$ has a linear resolution.*

4 Absolutely and Universally

We have discussed in the previous sections some notions, such as being G-quadratic, strongly Koszul, having a Koszul filtration or a Gröbner flag that imply Koszulness. In this section we discuss two variants of the Koszul property: universally Koszul and absolutely Koszul.

4.1 Universally Koszul

When looking for a Koszul filtration, among the many families of ideals of linear forms one can take the set of all ideals of linear forms. This leads to the following definition:

Definition 1. Let R be a K-algebra and set

$$\mathscr{L}(R) = \{I \subset R : I \text{ ideal generated by elements of degree } 1\}.$$

We say that R is universally Koszul if the following equivalent conditions hold:

(1) $\mathscr{L}(R)$ is a Koszul filtration of R.
(2) $\operatorname{reg}_R(R/I) = 0$ for every $I \in \mathscr{L}(R)$.
(3) For every $I \in \mathscr{L}(R)$ and $x \in R_1 \setminus I$, one has $I : x \in \mathscr{L}(R)$.

That the three conditions are indeed equivalent is easy to see (see [21, 1.4]). In [21, 2.4] it is proved that:

Theorem 2. *Let $S = K[x_1, \dots, x_n]$ and $m \in \mathbf{N}$. If $m \le n/2$, then a generic space of quadrics of codimension m in the vector space of quadrics defines a universally Koszul algebra.*

One should compare the result above with the following:

Theorem 3. *Let $S = K[x_1, \dots, x_n]$ and $m \in \mathbf{N}$.*

(1) *A generic space of quadrics of codimension m defines a Koszul algebra if $m \le n^2/4$.*
(2) *A generic space of quadrics of codimension m defines an algebra with a Gröbner flag if $m \le n - 1$.*

For (1) see [27, 3.4], for (2) [20, 10]. Fröberg and Löfwall proved in [38] that, apart from spaces of quadrics of codimension $\le n^2/4$, the only generic spaces of quadrics defining Koszul algebras are the complete intersections. Returning to universally Koszul algebras, in [21] it is also proved that:

Theorem 4. *Let R be a Cohen–Macaulay domain K-algebra with K algebraically closed of characteristic* 0. *Then R is universally Koszul if and only if R is a polynomial extension of one of the following algebras:*

(1) *The coordinate ring of a quadric hypersurface.*
(2) *The coordinate ring of a rational normal curve, that is,* $K[x,y]^{(c)}$ *for some* c.
(3) *The coordinate ring of a rational normal scroll of type* (c,c)*, that is, the Segre product of* $K[x,y]^{(c)}$ *with* $K[s,t]$.
(4) *The coordinate ring of the Veronese surface in* $\mathbf{P}^5$*, that is,* $K[x,y,z]^{(2)}$.

4.2 Absolutely Koszul

Let $\mathbf{F}_M^R$ be the minimal free resolution of a graded module M over R. One defines a $\mathbf{m}_R$-filtration on $\mathbf{F}_M^R$ whose associated graded complex $\mathrm{lin}(\mathbf{F}_M^R)$ has, in the graded case, a very elementary description. The complex $\mathrm{lin}(\mathbf{F}_M^R)$ is obtained from $\mathbf{F}_M$ by replacing with 0 all entries of degree > 1 in the matrices representing the homomorphisms. In the local case the definition of $\mathrm{lin}(\mathbf{F}_M^R)$ is more complicated (see Sect. 6 for details). One defines

$$\mathrm{ld}_R(M) = \sup\{i : H_i(\mathrm{lin}(\mathbf{F}_M^R)) \neq 0\}. \tag{4}$$

Denote by $\Omega_i^R(M)$ the ith syzygy module of a module M over R. It is proved in Römer PhD thesis and also in [43] that:

Proposition 5. *Assume R is Koszul. Then:*

(1) *M is componentwise linear iff* $\mathrm{ld}_R(M) = 0$.
(2) $\mathrm{ld}_R(M) = \inf\{i : \Omega_i(M)$ *is componentwise linear*$\}$.
(3) *If* $\Omega_i^R(M)$ *is componentwise linear then* $\Omega_{i+1}^R(M)$ *is componentwise linear.*

Iyengar and Römer introduced in [43] the following notion:

Definition 6. A K-algebra R is said to be absolutely Koszul if $\mathrm{ld}_R(M)$ is finite for every module M.

It is shown in [41] that:

Proposition 7. *If* $\mathrm{ld}_R(M)$ *is finite, then* $\mathrm{reg}_R(M)$ *is finite as well. Furthermore the Poincaré series* $P_M(z)$ *of* M *is rational and its "denominator" only depends on* R.

One obtains the following characterization of the Koszul property:

Corollary 8. *Let R be a K-algebra. Then R is Koszul if and only if* $\mathrm{ld}_R(K)$ *is finite. In particular, if R is absolutely Koszul then R is Koszul.*

On the other hand there are Koszul algebras that are not absolutely Koszul.

Example 9. The algebra

$$R = K[x_1,x_2,x_3,y_1,y_2,y_3]/(x_1,x_2,x_3)^2 + (y_1,y_2,y_3)^2$$

is Koszul but not absolutely Koszul because there are R-modules with non-rational Poincaré series. This and other examples of "bad" Koszul algebras are discussed by Roos in [52].

One also has [41, 5.10].

Theorem 10. *Let $R = S/I$ with $S = K[x_1, \dots, x_n]$. Then R is absolutely Koszul if either R is a complete intersection of quadrics or* $\operatorname{reg}_S(R) = 1$.

There is however an important difference between the two cases [41, 6.2, 6.7]:

Remark 11. When $\operatorname{reg}_S(R) = 1$ one has $\operatorname{ld}_R(M) \le 2\dim R$ for every M and even $\operatorname{ld}_R(M) \le \dim R$ is furthermore R is Cohen–Macaulay. But when R if a complete intersection of quadrics of codimension > 1 (or more generally when R is Gorenstein of with socle in degree > 1) one has $\sup_M \operatorname{ld}_R(M) = \infty$.

Another important contribution is the following:

Theorem 12. *Let R be a Gorenstein Artinian algebra with Hilbert function* $1 + nz + nz^2 + z^3$ *and* $n > 2$. *Then:*

(1) *If there exist $x, y \in R_1$ such that* $0 : x = (y)$ *and* $0 : y = (x)$ *(an exact pair of zero-divisors in the terminology of [40]), then R has a Koszul filtration and it is absolutely Koszul.*
(2) *If R is generic then it has an exact pair of zero-divisors.*

See [26, 2.13,6.3] for the statement on Koszul filtration in (1) and for (2) and see [40, 3.3] for the absolutely Koszulness.

What are the relationships between the properties discussed in this and the earlier sections? Here are some questions:

Question 13. (1) Strongly Koszul $\Rightarrow$ G-quadratic?
(2) Universally Koszul $\Rightarrow$ G-quadratic?
(3) Universally Koszul $\Rightarrow$ absolutely Koszul?

Question 13 (1) is mentioned in [42, p. 166] in the toric setting. Another interesting question is:

Question 14. What is the behavior of absolutely Koszul algebras under standard algebra operations (e.g. forming Veronese subalgebras or Segre and fiber products)?

The same question for universally Koszul algebras is discussed in [21] and for strongly Koszul in [42]. Note however that in [42] the authors deal mainly with toric algebras and their toric coordinates. Universally Koszul algebras with monomial relations have been classified in [22]. We may ask:

Question 15. Is it possible to classify absolutely Koszul algebras defined by monomials?

5 Regularity and Koszulness

We list in this section some facts and some questions that we like concerning Koszul algebras and regularity. We observe the following.

Remark 1. Regularity over the polynomial ring S behaves quite well with respect to products of ideals and modules:

(1) $\operatorname{reg}_S(I^u M)$ is a linear function in u for large u (see [29, 45, 60]).
(2) $\operatorname{reg}_S(IM) \leq \operatorname{reg}_S(M) + \operatorname{reg}_S(I)$ (does not hold in general but it) holds provided $\dim S/I \leq 1$, [23].
(3) More generally,

$$\operatorname{reg}_S(\operatorname{Tor}_i^S(N, M)) \leq \operatorname{reg}_S(M) + \operatorname{reg}_S(N) + i$$

provided the Krull dimension of $\operatorname{Tor}_1^S(N, M)$ is ≤ 1, [17, 33].
(4) $\operatorname{reg}_S(I_1 \cdots I_d) = d$ for ideals I_i generated in degree 1, [23]

where M, N are S-modules and I, I_i are ideals of S.

What happens if we replace in Remark 1 the polynomial ring S with a Koszul algebra R and consider regularity over R? Trung and Wang proved in [60] that $\operatorname{reg}_S(I^u M)$ is asymptotically a linear function in u when I is an ideal of R and M is a R-module. If R is Koszul, $\operatorname{reg}_R(I^u M) \leq \operatorname{reg}_S(I^u M)$, and hence $\operatorname{reg}_R(I^u M)$ is bounded above by a linear function in u.

Question 2. Let R be a Koszul algebra $I \subset R$ an ideal and M an R-module. Is $\operatorname{reg}_R(I^u M)$ a linear function in u for large u?

The following examples show that statements (2) and (3) in Remark 1 do not hold over Koszul algebras.

Example 3. Let $R = \mathbf{Q}[x, y, z, t]/(x^2+y^2, z^2+t^2)$. With $I = (x, z)$ and $J = (y, t)$ one has $\operatorname{reg}_R(I) = 1$, $\operatorname{reg}_R(J) = 1$ because x, z and y, t are regular sequences on R, $\dim R/I = 0$ and $\operatorname{reg}_R(IJ) = 3$.

Example 4. Let $R = K[x, y]/(x^2 + y^2)$. Let $M = R/(x)$ and $N = R/(y)$ and note that $\operatorname{reg}_R(M) = 0$, $\operatorname{reg}_R(N) = 0$ because x and y are non-zero divisors in R while $\operatorname{Tor}_1^R(M, N) = H_1(x, y, R) = K(-2)$.

Nevertheless statements (2), (3) of Remark 1 might hold for special type of ideals/modules over special type of Koszul algebras. For example, one has:

Proposition 5. *Let R be a Cohen–Macaulay K-algebra with $\operatorname{reg}_S(R) = 1$, let I be an ideal generated in degree 1 such that $\dim R/I \leq 1$ and M an R-module. Then $\operatorname{reg}_R(IM) \leq \operatorname{reg}_R(M) + 1$. In particular, $\operatorname{reg}_R(I) = 1$.*

Proof. We may assume K is infinite. The short exact sequence

$$0 \to IM \to M \to M/IM \to 0$$

implies that $\text{reg}_R(IM) \leq \max\{\text{reg}_R(M), \text{reg}_R(M/IM)+1\}$. It is therefore enough to prove that $\text{reg}_R(M/IM) \leq \text{reg}_R(M)$. Then let $J \subset I$ be an ideal generated by a maximal regular sequence of elements of degree 1 and set $A = R/J$. Since $\text{reg}_R(A) = 0$ and since M/IM is an A-module, by virtue of Proposition 3, we have $\text{reg}_R(M/IM) \leq \text{reg}_A(M/IM)$. By construction, A is Cohen–Macaulay of dimension ≤ 1 and has regularity 1 over the polynomial ring projecting onto it. So, by Lemma 14 we have $\text{reg}_A(M/IM) = \max\{t_0^A(M/IM), t_1^A(M/IM) - 1\}$. Summing up, since $t_0^A(M/IM) = t_0^R(M)$, it is enough to prove $t_1^A(M/IM) \leq \text{reg}_R(M) + 1$. Now we look at

$$0 \to IM/JM \to M/JM \to M/IM \to 0$$

that gives $t_1^A(M/IM) \leq \max\{t_1^A(M/JM), t_0^A(IM/JM)\}$. Being $t_0^A(IM/JM) \leq t_0^R(M) + 1 \leq \text{reg}_R(M) + 1$, it remains to prove that $t_1^A(M/JM) \leq \text{reg}_R(M) + 1$, and this follows from the right exactness of the tensor product. □

The following example shows that the assumption $\dim R/I \leq 1$ in Proposition 5 is essential.

Example 6. The algebra $R = K[x, y, z, t]/(xy, yz, zt)$ is Cohen–Macaulay of dimension 2 and $\text{reg}_S(R) = 1$. The ideal $I = (y - z)$ has $\text{reg}_R(I) = 2$ and $\dim R/I = 2$.

Example 3 shows also that statement (4) of Remark 1 does not hold over a Koszul algebra even if we assume that each I_i is an ideal of regularity 1 and of finite projective dimension. Statement (4) of Remark 1 might be true if one assumes that the ideals I_i belongs to a Koszul filtration. We give a couple of examples in this direction:

Proposition 7. *Let $S = K[x_1, \ldots, x_n]$, $R = S/I$ with I generated by monomials of degree 2. Let $X = \{\bar{x}_1, \ldots, \bar{x}_n\}$ and $\mathscr{F} = \{(Y) : Y \subset X\}$. Let $I_1, \ldots, I_d \in \mathscr{F}$. Then $\text{reg}_R(I_1 \cdots I_d) = d$ unless $I_1 \cdots I_d = 0$.*

Proof. First we observe the following. Let $m_1, \ldots, m_t$ be monomials of degree d and $J = (m_1, \ldots, m_t)$. Assume that they have linear quotients (in S), that is, $(m_1, \ldots, m_{i-1}) :_S m_i$ is generated by variables for every i. Fact (1) in the proof of Theorem 13 implies that JR has linear quotients with respect to the Koszul filtration $\mathscr{F}$ of R. By Lemma 17 we have that $\text{reg}_R(JR) = d$ (unless $JR = 0$). Now the desired result follows because products of ideals of variables have linear quotients in S by [23, 5.4]. □

Example 4.3 in [23] shows that the inequality $\text{reg}_R(IM) \leq \text{reg}_R(M) + \text{reg}_R(I)$ does not even hold over a K-algebra R with a Koszul filtration $\mathscr{F}$, $I \in \mathscr{F}$ and M an R-module with linear quotient with respect to $\mathscr{F}$. The following are natural questions:

Question 8. Let R be an algebra with a Koszul filtration $\mathscr{F}$. Is it true that $\text{reg}_R(I_1 \cdots I_d) = d$ for every $I_1, \ldots, I_d \in \mathscr{F}$ whenever the product is non-zero?

In view of the analogy with statement (4) of Remark 1 the following special case deserves attention:

Question 9. Let R be a universally Koszul algebra. Is it true that $\operatorname{reg}_R(I_1 \cdots I_d) = d$ for every $I_1, \dots, I_d$ ideals of R generated in degree 1 (whenever the product is non-zero)?

Remark 10. In a universally Koszul algebra a product of elements of degree 1 has a linear annihilator. This can be easily shown by induction on the number of factors. Hence, the answer to Question 9 is positive if each I_i is principal.

We are able to answer Question 8 in the following cases:

Theorem 11. *Products of ideals of linear forms have linear resolutions over the following rings:*

(1) *R is Cohen–Macaulay with* $\dim R \le 1$ *and* $\operatorname{reg}_S(R) = 1$.
(2) $R = K[x, y, z]/(q)$ *with* $\deg q = 2$.
(3) $R = K[x, y]^{(c)}$ *with* $c \in \mathbf{N}_{>0}$.
(4) $R = K[x, y, z]^{(2)}$.
(5) $R = K[x, y] * K[s, t]$ *($*$ denotes the Segre product).*

Proof. The rings in the list are Cohen–Macaulay with $\operatorname{reg}_S(R) = 1$. Let $I_1, \dots, I_d$ be ideals generated by linear forms. We prove by induction on d that $\operatorname{reg}_R(I_1 \cdots I_d) = d$. The case $d = 1$ follows because the rings in the list are universally Koszul. If for one of the I_i we have $\dim R/I_i \le 1$ then we may use Proposition 5 and conclude by induction. Hence we may assume $\dim R/I_i \ge 2$ for every i. For the ring (1) and (3) (which is a 2-dimensional domain) we are done. In the case (2), the only case left is when the I_i are principal. But then we may conclude by virtue of Remark 10. In cases (3) and (4) we have that $\dim R/I_i = 2$ for each i, that is, $\operatorname{height} I_i = 1$. If one of the I_i is principal, then we are done by induction (because the R is a domain). Denote by A either $K[x, y, z]$ in case (3) or $K[x, y, s, t]$ in case (4). Since R is a direct summand of A we have $IA \cap R = I$ for every ideal I of R. It follows that $\operatorname{height}(I_i A) = 1$ for every i and hence there exist non-units $f_i \in A$ such that $I_i A \subset (f_i)$. In case (3) we have that each f_i must have degree 1 in A and $I_i A = (f_1) J_1$ with J_1 an ideal generated by linear forms of A. Hence $I_1 I_2 = (f_1 f_2) H$ where $H = J_1 J_2$ is an ideal generated by linear forms of R. Hence we are done because one of the factor is principal. In case (4) we have that each f_i is either a linear form in x, y or a linear form in s, t. If one of the f_i's is a linear form in x, y and another one is a linear form in s, t we can proceed as in the case (3). So we are left with the case that every f_i is a linear form in, say, x and y and $I_i = (f_i) J_i$ with J_i generated by linear forms in z, t. Since none of the I_i is principal we have that $J_i = (z, t)$ for every i. Hence $I_1 \cdots I_d$ is generated by $(\prod_{i=1}^{d} f_i)(z, t)^d$ and it isomorphic to the R-submodule of A generated by its component of degree $(0, d)$. That such a module has a linear resolution over R follows from Proposition 19. □

We state now a very basic question of computational nature.

Question 12. Let R be a Koszul algebra and M an R-module. How does one compute $\operatorname{reg}_R(M)$? Can one do it algorithmically?

Few comments concerning Question 12. We assume to be able to compute syzygies over R and so to be able to compute the first steps of the resolution of a R-module M. Let $S \to R$ the canonical presentation Eq. (1) of R. We know that $\operatorname{reg}_R(M) \leq \operatorname{reg}_S(M)$ and $\operatorname{reg}_S(M)$ can be computed algorithmically because $\operatorname{pd}_S(M)$ is finite. A special but already interesting case of Question 12 is:

Question 13. Let R be a Koszul algebra and M an R-module generated in degree 0, with $M_1 \neq 0$ and $M_i = 0$ for $i > 1$. Can one decide algorithmically whether $\operatorname{reg}_R(M) = 0$ or $\operatorname{reg}_R(M) = 1$?

Set

$$r_R(M) = \min\{i \in \mathbf{N} : t_i^R(M) - i = \operatorname{reg}_R(M)\}.$$

So $r_R(M)$ is the first homological position where the regularity of M is attained. If one knows $r_R(M)$ or a upper bound $r \geq r_R(M)$ for it, then one can compute $\operatorname{reg}_R(M)$ by computing the first r steps of the resolution of M. Note that

$$r_R(M) \leq \operatorname{ld}_R(M)$$

because $\operatorname{reg}_R(N) = t_0^R(N)$ if N is componentwise linear. One has:

Lemma 14. *Let R be a K-algebra with $\operatorname{reg}_S(R) = 1$. Then $r_R(M) \leq 2\dim R$ for every M, that is, the regularity of any R-module is attained within the first $2\dim R$ steps of the resolution. If furthermore R is Cohen–Macaulay, $r_R(M) \leq \dim R - \operatorname{depth} M$.*

The first assertion follows from Remark 11; the second is proved by a simple induction on $\operatorname{depth} M$.

Note that the ith syzygy module of M cannot have a free summand if $i > \dim R$ by [30, 0.1] and so

$$t_{j+1}^R(M) > t_j^R(M) \text{ if } j > \dim R.$$

Unfortunately there is no hope to get a bound for $r_R(M)$ just in terms of invariants of R for general Koszul algebras. The argument of [41, 6.7] that shows that if R is a Gorenstein algebra with socle degree > 1 then $\sup \operatorname{ld}_R(M) = \infty$ shows also that $\sup_M r_R(M) = \infty$. For instance, over $R = K[x, y]/(x^2, y^2)$ let M_n be the dual of the nth syzygy module $\Omega_n^R(K)$ shifted by n. One has that M_n is generated be in degree 0, $\operatorname{reg}_R(M_n) = 1$ and $r_R(M_n) = n$. On the other hand, the number of generators of M_n is n. So we ask:

Question 15. Let R be a Koszul algebra. Can one bound $r_R(M)$ in terms of invariants of R and "computable" invariants of M such as its Hilbert series or its Betti numbers over S?

The questions above make sense also over special families of Koszul rings. For instance, there has been a lot of activity to understand resolutions of modules over short rings, that is, rings with $\mathbf{m}_R^3 = 0$ or $\mathbf{m}_R^4 = 0$, both in the graded and local case (see [7, 9, 40]). It would be very interesting to answer Questions 12, 13, and 15 for short rings.

6 Local Variants

This section is concerned with "Koszul-like" behaviors of local rings and their modules.

Assumption: From now on, when not explicitly said, R is assumed to be a local or graded ring with maximal ideal $\mathbf{m}$ and residue field $K = R/\mathbf{m}$. Moreover all modules and ideals are finitely generated, and homogeneous in the graded case.

We define the associated graded ring to R with respect to the $\mathbf{m}$-adic filtration

$$G = \mathrm{gr}_{\mathbf{m}}(R) = \oplus_{i \geq 0} \mathbf{m}^i / \mathbf{m}^{i+1}.$$

The Hilbert series and the Poincaré series of R are

$$H_R(z) = H_G(z) = \sum_{i \geq 0} \dim(\mathbf{m}^i / \mathbf{m}^{i+1}) z^i \quad \text{and} \quad P_R(z) = \sum_{i \geq 0} \dim \mathrm{Tor}_i^R(K, K) z^i.$$

6.1 Koszul Rings

Following Fröberg [35] we extend the definition of Koszul ring to the local case as follows:

Definition 1. The ring R is Koszul if its associated graded ring G is a Koszul algebra (in the graded sense), that is, R is Koszul if K has a linear resolution as a G-module.

As it is said in Remark 6 (6) in the graded setting the Koszul property holds equivalent to the following relation between the Poincaré series of K and the Hilbert series of R:

$$P_K^R(z) H_R(-z) = 1. \tag{5}$$

The following definition is due to Fitzgerald [34]:

Definition 2. The ring R is Fröberg if the relation Eq. (5) is verified.

We want to explain why every Koszul ring is Fröberg. To this end we need to introduce few definitions.

Let A be a regular local ring with maximal ideal $\mathbf{m}_A$ and let I be an ideal of A such that $I \subseteq \mathbf{m}_A^2$. Set $R = A/I$. Then $G \simeq S/I^*$ where S is the polynomial ring and I^* is the homogeneous ideal generated by the initial forms f^* of the elements $f \in I$.

Definition 3. (1) A subset $\{f_1, \dots, f_t\}$ of I is a standard basis of I if $I^* = (f_1^*, \dots, f_t^*)$;
(2) The ideal I is d-isomultiple if I^* is generated in degree d.

If $\{f_1, \dots, f_t\}$ is a standard basis of I, then $I = (f_1, \dots, f_t)$. See [53] for more details on d-isomultiple ideals. Notice that by Remark 6 (1) we have

$$R \quad \text{Koszul} \quad \Longrightarrow \quad I \text{ is 2-isomultiple.}$$

Obviously the converse does not hold true because a quadratic K-algebra is not necessarily Koszul.

We now explore the connection between Koszul and Fröberg rings. By definition $H_R(z) = H_G(z)$, and

$$P_K^R(z) \le P_K^G(z)$$

(see, e.g. [36, 4]). Conditions are known under which $\beta_i^R(K) = \beta_i^G(K)$, for instance, this happens if

$$t_i^G(K) = \max\{j : \beta_{ij}^G(K) \ne 0\} \le \min\{j : \beta_{i+1j}^G(K) \ne 0\} \text{ for every } i; \quad (6)$$

see [36, 4]. This is the case if K has a linear resolution as a G-module. Hence,

Proposition 4. *If R is Koszul, then R is Fröberg.*

Proof. By definition, $H_R(z) = H_G(z)$. If R is a Koszul ring, then G is Koszul, in particular $P_K^G(z)H_G(-z) = 1$. The result follows because the graded resolution of K as a G-module is linear and hence Eq. (6) and therefore $P_K^R(z) = P_K^G(z)$. □

Since in the graded case R is Fröberg iff it is Koszul, it is natural to ask the following question.

Question 5. Is a Fröberg (local) ring Koszul?

We give a positive answer to this question for a special class of rings. If f is a non-zero element of R, denote by $v(f) = v$ the valuation of f, that is the largest integer such that $f \in \mathbf{m}^v$.

Proposition 6. *Let I be an ideal generated by a regular sequence in a regular local ring A. The following facts are equivalent:*

(1) *A/I is Koszul.*
(2) *A/I is Fröberg.*
(3) *I is 2-isomultiple.*

Proof. By Proposition 4 we know (1) implies (2). We prove that (2) implies (3). Let $I = (f_1, \dots, f_r)$ with $v(f_i) = v_i \ge 2$. By [59] we have $P_K^{A/I}(z) = (1+z)^n/(1-z^2)^r$. Since A/I is Fröberg, one has that $H_{A/I}(z) = (1-z^2)^r/(1-z)^n$, in particular the multiplicity of A/I is 2^r. From [53, 1.8], it follows that $v_i = 2$ for every $i = 1, \dots, r$ and $f_1^*, \dots, f_r^*$ is a regular sequence in G. Hence $I^* = (f_1^*, \dots, f_r^*)$, so I is 2-isomultiple. If we assume (3), then G is a graded quadratic complete intersection, hence $P_K^G(z)H_G(-z) = 1$ and since G is graded this implies that G is Koszul. □

Next example is interesting to better understand what happens in case the regular sequence is not 2-isomultiple.

Example 7. Consider $I_s = (x^2 - y^s, xy) \subset A = K[[x, y]]$ where s is an integer ≥ 2. Then, as we have seen before, $P_K^{A/I}(z) = (1+z)^2/(1-z^2)^2$ and it does not depend on s. On the contrary the Hilbert series depends on s, precisely $H_{A/I}(z) = 1 + 2z + \sum_{i=2}^{s} z^i$. It follows that A/I is Koszul (hence Fröberg) if and only if $s = 2$ if and only if I is 2-isomultiple. In fact if $s > 2$, then $I_s^* = (x^2, xy, y^{s+1})$ is not quadratic.

In the following we denote by $e(M)$ the multiplicity (or degree) of an R-module M and by $\mu(M)$ its minimal number of generators. Let R be a Cohen–Macaulay ring. Abhyankar proved that $e(R) \geq h + 1$ and $h = \mu(\mathbf{m}) - \dim R$ is the so-called embedding codimension. If equality holds R is said to be of minimal multiplicity.

Proposition 8. *Let R be a Cohen–Macaulay ring of multiplicity e and Cohen–Macaulay type τ. If one of the following conditions holds:*

(1) $e = h + 1$
(2) $e = h + 2$ *and* $\tau < h$

then R is a Koszul ring.

Proof. In both cases the associated graded ring is Cohen–Macaulay and quadratic (see [53, 3.3, 3.10]). We may assume that the residue field is infinite; hence there exist $x_1^*, \dots, x_d^*$ filter regular sequence in G and it is enough to prove that $G/(x_1^*, \dots, x_d^*) \simeq \mathrm{gr}_{\mathbf{m}/(x_1,\dots,x_d)}(R/(x_1, \dots, x_d))$ is Koszul (see, e.g. [43, 2.13]). Hence the problem is reduced to an Artinian quadratic K-algebra with $\mu(\mathbf{m}) = h > 1$ and $\dim_K \mathbf{m}^2 \leq 1$, and the result follows (see [34] or [20]). □

Remark 9. (1) There are Cohen–Macaulay rings R with $e = h + 2$ and $\tau = h$ whose associated graded ring G is not quadratic, hence not Koszul. For example this is the case if $R = k[[t^5, t^6, t^{13}, t^{14}]]$, where $e = h + 2 = 3 + 2 = 5$ and $\tau = 3$.

(2) Let R be Artinian of multiplicity $e = h + 3$. Then R is stretched if its Hilbert function is $1 + hz + z^2 + z^3$, and short if its Hilbert function is $1 + hz + 2z^2$ (for details see [54]). For example, if R is Gorenstein, then R is stretched. Sally classified, up to analytic isomorphism, the Artinian local rings which are stretched in terms of the multiplicity and the Cohen–Macaulay type. As a consequence one verifies that if R is stretched of multiplicity $\geq h + 3$, then I^* is never quadratic, hence R is never Koszul. If R is short, then R is Koszul if and only if G is quadratic. In fact, by a result in Backelin's PhD thesis (see also [20]), if $\dim_K G_2 = 2$ and G is quadratic, then G is Koszul, so R is Koszul.

6.2 Koszul Modules and Linear Defect

Koszul modules have been introduced in [41]. Let us recall the definition.

Consider $(\mathbf{F}_M^R, \delta)$ a minimal free resolution of M as an R-module. The property $\delta(\mathbf{F}_M^R) \subseteq \mathbf{m}\mathbf{F}_M^R$ (the minimality) allows us to form for every $j \geq 0$ a complex

$$\mathrm{lin}_j(\mathbf{F}_M^R): \quad 0 \to \frac{F_j}{\mathbf{m}F_j} \to \cdots \to \frac{\mathbf{m}^{j-i}F_i}{\mathbf{m}^{j-i+1}F_i} \to \cdots \to \frac{\mathbf{m}^j F_0}{\mathbf{m}^{j+1}F_0} \to 0$$

of K-vector spaces. Denoting $\mathrm{lin}(\mathbf{F}_M^R) = \oplus_{j\geq 0}\mathrm{lin}_j(\mathbf{F}_M^R)$, one has that $\mathrm{lin}(\mathbf{F}_M^R)$ is a complex of free graded modules over $G = \mathrm{gr}_{\mathbf{m}}(R)$ whose ith free module is

$$\oplus_j \mathbf{m}^{j-i}F_i/\mathbf{m}^{j-i+1}F_i = \mathrm{gr}_{\mathbf{m}}(F_i)(-i) = G(-i) \otimes_K F_i/\mathbf{m}F_i.$$

By construction the differentials can be described by matrices of linear forms.

Accordingly with the definition given by Herzog and Iyengar in [41]:

Definition 10. M is a Koszul module if $H_i(\mathrm{lin}_j(\mathbf{F}_M^R)) = 0$ for every $i > 0$ and $j \geq 0$, that is, $H_i(\mathrm{lin}(\mathbf{F}_M^R)) = 0$ for every $i > 0$.

Remark 11. Notice that, if R is graded, the K-algebras G and R are naturally isomorphic. In particular $\mathrm{lin}(\mathbf{F}_M^R)$ coincides with the complex already defined in Sect. 4.2. This is why $\mathrm{lin}(\mathbf{F}_M^R)$ is called the linear part of $\mathbf{F}_M^R$.

As in the graded case (see Eq. (4)), one defines the linear defect of M over R:

$$\mathrm{ld}_R(M) = \sup\{i : H_i(\mathrm{lin}(\mathbf{F}_M^R)) \neq 0\}. \tag{7}$$

The linear defect gives a measure of how far is $\mathrm{lin}(\mathbf{F}_M^R)$ from being a resolution of $\mathrm{gr}_{\mathbf{m}}(M) = \oplus_{j\geq 0}\mathbf{m}^j M/\mathbf{m}^{j+1}M$. By the uniqueness of minimal free resolution, up to isomorphism of complexes, one has that $\mathrm{ld}_R(M)$ does not depend on $\mathbf{F}_M^R$, but only on the module M. When $\mathrm{ld}_R(M) < \infty$, we say that in the minimal free resolution the linear part predominates.

Koszul modules have appeared previously in the literature under the name "modules with linear resolution" in [56] and "weakly Koszul" in [46].

By definition R is Koszul as an R-module because it is R-free. But, accordingly with Definition 1, R is a Koszul ring if and only if K is a Koszul R-module. We have:

R is a Koszul ring $\iff$ K is a Koszul R-module $\iff$ K is a Koszul G-module.

If R is a graded K-algebra, Corollary 8 in particular says that R is a Koszul ring if K has finite linear defect or equivalently K is a Koszul module. By [41, 1.13] and [43, 3.4] one gets the following result, that is, the analogous of Theorem 7, with the regularity replaced by the linear defect.

Proposition 12. *Let R be a graded K-algebra. The following facts are equivalent:*

(1) *R is Koszul.*
(2) $\mathrm{ld}_R(K) = 0$.
(3) $\mathrm{ld}_R(K) < \infty$.
(4) *There exists a Koszul Cohen–Macaulay R-module M with $\mu(M) = e(M)$.*
(5) *Every Cohen–Macaulay R-module M with $\mu(M) = e(M)$ is Koszul.*

In [43] the modules verifying $\mu(M) = e(M)$ are named modules of minimal degree. When R itself is Cohen–Macaulay, the maximal Cohen–Macaulay modules of minimal degree are precisely the so-called Ulrich modules. Cohen–Macaulay modules of minimal degree exist over any local ring, for example, the residue field is one.

The following question appears in [41, 1.14].

Question 13. Let R be a local ring. If $\mathrm{ld}_R(K) < \infty$, then is $\mathrm{ld}_R(K) = 0$?

To answer Question 13 one has to compare $\mathrm{lin}(\mathbf{F}_K^R)$ and $\mathrm{lin}(\mathbf{F}_K^G)$. From a minimal free resolution of K as a G-module we can build up a free resolution (not necessarily minimal) of K as an R-module. In some cases the process for getting the minimal free resolution is under control via special cancellations (see [55, 3.1]), but in general it is a difficult task.

We may define absolutely Koszul local rings exactly as in the graded case. A positive answer to Question 13 would give a positive answer to the following:

Question 14. Let R be an absolutely Koszul local ring. Is R Koszul?

Acknowledgements We thank Rasoul Ahangari, Lucho Avramov, Giulio Caviglia, Ralf Fröberg, Jüergen Herzog, Srikanth Iyengar, Liana Şega, Bart Snapp, Rekha Thomas and Matteo Varbaro for useful discussion concerning the material presented in this chapter.

References

1. Anders, J.: CaTS, a software system for toric state polytopes. Available at http://www.soopadoopa.dk/anders/cats/cats.html
2. Anick, D.: A counterexample to a conjecture of Serre. Ann. Math. **115**, 1–33 (1982)
3. Aramova, A., Bărcănescu, Ş., Herzog, J.: On the rate of relative Veronese submodules. Rev. Roumaine Math. Pures Appl. **40**, 243–251 (1995)
4. Avramov, L.L.: Infinite free resolutions. Six lectures on commutative algebra (Bellaterra, 1996). In: Progress in Mathematics, vol. 166, pp. 1–118. Birkhäuser, Basel (1998)
5. Avramov, L.L., Eisenbud, D.: Regularity of modules over a Koszul algebra. J. Algebra **153**, 85–90 (1992)
6. Avramov, L.L., Peeva, I.: Finite regularity and Koszul algebras. Amer. J. Math. **123**, 275–281 (2001)
7. Avramov, L.L., Iyengar, S., Şega, L.: Free resolutions over short local rings. Proc. London Math. Soc. **78**, 459–476 (2008)

8. Avramov, L.L., Conca, A., Iyengar, S.: Free resolutions over commutative Koszul algebras. Math. Res. Lett. **17**, 197–210 (2010)
9. Avramov, L.L., Iyengar, S., Şega, L.: Short Koszul modules. J. Commutative Algebra **2**, 249–278 (2010)
10. Backelin, J.: On the rates of growth of the homologies of Veronese subrings. Algebra, algebraic topology, and their interactions (Stockholm, 1983). In: Lecture Notes in Mathematics, vol. 1183, pp. 79–100. Springer, Berlin (1986)
11. Backelin, J., Fröberg, R.: Koszul algebras, Veronese subrings, and rings with linear resolutions. Rev. Roumaine Math. Pures Appl. **30**, 85–97 (1985)
12. Bărcănescu, S., Manolache, N.: Betti numbers of Segre–Veronese singularities. Rev. Roumaine Math. Pures Appl. **26**, 549–565 (1981)
13. Beilinson, A., Ginsburg, V., Schechtman, V.: Koszul duality. J. Geom. Phys. **5**, 317–350 (1988)
14. Beilinson, A., Ginsburg, V.A., Soergel, W.: Koszul duality patterns in representation theory. J. Amer. Math. Soc. **9**, 473–527 (1996)
15. Bruns, W., Herzog, J.: Cohen–Macaulay rings. In: Revised edition, Cambridge Studies in Advanced Mathematics, vol. 39. University Press, Cambridge (1998)
16. Bruns, W., Conca, A., Römer, T.: Koszul homology and syzygies of Veronese subalgebras. Math. Ann. **351**, 761–779 (2011)
17. Caviglia, G.: Bounds on the Castelnuovo–Mumford regularity of tensor products. Proc. Amer. Math. Soc. **135**, 1949–1957 (2007)
18. Caviglia, G.: The pinched Veronese is Koszul. J. Algebr. Comb. **30**, 539–548 (2009)
19. Caviglia, G., Conca, A.: Koszul property of projections of the Veronese cubic surface. Adv. Math. arXiv:1203.1949, preprint (2012)
20. Conca, A.: Gröbner bases for spaces of quadrics of low codimension. Adv. in Appl. Math. **24**, 111–124 (2000)
21. Conca, A.: Universally Koszul Algebras. Math. Ann. **317**, 329–346 (2000)
22. Conca, A.: Universally Koszul algebras defined by monomials. Rend. Sem. Mat. Univ. Padova **107**, 95–99 (2002)
23. Conca, A., Herzog, J.: Castelnuovo-Mumford regularity of products of ideals. Collect. Math. **54**, 137–152 (2003)
24. Conca, A., De Negri, E., Rossi, M.E.: On the rate of points in projective spaces. Israel J. Math. **124**, 253–265 (2001)
25. Conca, A., Herzog, J., Trung, N.V., Valla, G.: Diagonal subalgebras of bigraded algebras and embeddings of blow-ups of projective spaces. Amer. J. Math. **119**, 859–901 (1997)
26. Conca, A., Rossi, M.E., Valla, G.: Gröbner flags and Gorenstein algebras. Compositio Math. **129**, 95–121 (2001)
27. Conca, A., Trung, N.V., Valla, G.: Koszul property for points in projectives spaces. Math. Scand. **89**, 201–216 (2001)
28. Constantinescu, A., Varbaro, M.: On the h-vectors of Cohen-Macaulay flag complexes. Math. Scand. available at arXiv:1004.0170v2, To appear
29. Cutkosky, D., Herzog, J., Trung, N.V.: Asymptotic behavior of the Castelnuovo–Mumford regularity. Compositio Math. **118**, 243–261 (1999)
30. Eisenbud, D.: Homological algebra on a complete intersection, with an application to group representations. Trans. Amer. Math. Soc. **260**, 35–64 (1980)
31. Eisenbud, D., Goto, S.: Linear free resolutions and minimal multiplicity. J. Algebra **88**, 89–133 (1984)
32. Eisenbud, D., Reeves, A., Totaro, B.: Initial ideals, Veronese subrings, and rates of algebras. Adv. Math. **109**, 168–187 (1994)
33. Eisenbud, D., Huneke, C., Ulrich, B.: The regularity of Tor and graded Betti numbers. Amer. J. Math. **128**, 573–605 (2006)
34. Fitzgerald, R.W.: Local Artinian rings and the Fröberg relation. Rocky Mountain J. Math. **26**, 1351–1369 (1996)
35. Fröberg, R.: Determination of a class of Poincaré Series. Mat. Scand. **37**, 29–39 (1975)

36. Fröberg, R.: Connections between a local ring and its associated graded ring. J. Algebra **111**, 300–305 (1987)
37. Fröberg, R.: Koszul algebras. Advances in Commutative Ring Theory, Proceedings of Fez Conference 1997. In: Lectures Notes in Pure and Applied Mathematics, vol. 205, Marcel Dekker, New York (1999)
38. Fröberg, R., Löfwall, C.: Koszul homology and Lie algebras with application to generic forms and points. Homology Homotopy Appl. **4**, 227–258 (2002)
39. Frohmader, A.: Face vectors of flag complexes. Israel J. Math. **164**, 153–164 (2008)
40. Henriques, I., Şega, L.: Free resolutions over short Gorenstein local rings. Math. Zeit. **267**, 645–663 (2011)
41. Herzog, J., Iyengar, S.: Koszul modules. J. Pure Appl. Alg. **201**, 154–188 (2005)
42. Herzog, J., Hibi, T., Restuccia, G.: Strongly Koszul algebras. Math. Scand. **86**, 161–178 (2000)
43. Iyengar, S., Römer, T.: Linearity defects of modules over commutative rings. J. Algebra **322**, 3212–3237 (2009)
44. Kempf, G.: Syzygies for points in projective space. J. Algebra **145**, 219–223 (1992)
45. Kodiyalam, V.: Asymptotic behavior of Castelnuovo-Mumford regularity. Proc. Amer. Math. Soc. **128**, 407–411 (2000)
46. Martinez-Villa, R., Zacharia, D.: Approximations with modules having linear resolutions. J. Algebra **266**, 671–697 (2003)
47. Miller, E., Sturmfels, B.: Combinatorial commutative algebra. In: Graduate Texts in Mathematics, vol. 227. Springer, New York (2004)
48. Priddy, S.B.: Koszul resolutions. Trans. Amer. Math. Soc. **152**, 39–60 (1970)
49. Polishchuk, A.: On the Koszul property of the homogeneous coordinate ring of a curve. J. Algebra **178**, 122–135 (1995)
50. Polishchuk, A., Positselski, L.: Quadratic algebras. University Lecture Series, vol. 37. Americna Mathematical Society, Providence (2005)
51. Roos, J.E.: Commutative non Koszul algebras having a linear resolution of arbitrarily high order, Applications to torsion in loop space homology. C. R. Acad. Sci. **316**, 1123–1128 (1993)
52. Roos, J.E.: Good and bad Koszul algebras and their Hochschild. J. Pure Appl. Alg. **201**, 295–327 (2005)
53. Rossi, M.E., Valla, G.: Multiplicity and t-isomultiple ideals. Nagoya Math. J. **110**, 81–111 (1988)
54. Rossi, M.E., Valla, G.: Cohen–Macaulay local rings of embedding dimension $e + d - 3$. Proc. London Math. Soc. **80**, 107–126 (2000)
55. Rossi, M.E., Sharifan, L.: Consecutive cancellations in Betti numbers of local rings. Proc. Amer. Math. Soc. **138**, 61–73 (2009)
56. Şega, L.: Homological properties of powers of the maximal ideal of a local ring. J. Algebra **241**, 827–848 (2001)
57. Shibuta, T.: Gröbner bases of contraction ideals. J. Algebr. Comb. **36**(1), 1–19 (2012) arXiv:1010.5768
58. Sturmfels, B.: Gröbner bases and convex polytopes. In: University Lecture Series, **8**. American Mathematical Society, Providence (1996)
59. Tate, J.: Homology of local and Noetherian rings. Illinois J. Math. **1**, 14–27 (1957)
60. Trung, N., Wang, H.: On the asymptotic linearity of Castelnuovo–Mumford regularity. J. Pure Appl. Alg. **201**, 42–48 (2005)

Powers of Ideals: Betti Numbers, Cohomology and Regularity

Marc Chardin

1 Introduction

The aim of this chapter is to provide a coherent approach to a collection of advances over the last decade concerning homological invariants of powers of a graded ideal. At this level of generality, the common denominator of all results is the use of the Rees algebra of the ideal. Most, if not all, of the results are derived from the existence, when the ring is Noetherian, of a graded free resolution with finite summands for the Rees ring, seen as a (properly graded) module over a polynomial extension of the ring.

Besides few improvements or generalizations, most of the results in this text are not new. Three results that make a progress are Propositions 1 and 7, Theorem 2. Many authors contributed to advances on the understanding of homological properties of powers of ideals, and we are conscious that we certainly omit part of them (an example: the work [31] of Trung on the stabilization of the regularity index of the Hilbert function, see also [20] for earlier advances). This is both because we only cover part of this field and due to the fact that we very likely overlooked some of the contributions on these questions. It may very well reflect a personal way of looking at these questions, and I apologize in advance for these non-intentional omissions.

In the first section after this introduction, we recall some early advances that are connected to the more recent results we will focus on. We then present results on Betti numbers of powers and their eventual behaviour. The main idea, exploited in different ways in [12, 33] or [1], is that the Rees algebra provides structured free resolutions of all the powers in a single object. These are not minimal, but the defect of minimality corresponds to a finitely generated module, hence also stabilizes.

M. Chardin (✉)
Institut de Mathématiques de Jussieu, CNRS & UPMC, 4 place Jussieu,
F-75005, Paris, France
e-mail: chardin@math.jussieu.fr

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_9,

Precise results on powers out of this remark are given in the quoted articles, and we just reproduce two of them to give a flavour of what can be done.

We then recall, in Sect. 3, the definition of Castelnuovo–Mumford regularity of a graded module over a (finitely generated) standard graded algebra. One key feature of this notion is that it can be defined in two ways: via Tor modules or via local cohomology with respect to the positive part. It was noticed, very likely first by Jouanolou, that the two definitions agree without assuming neither the base ring to be Noetherian nor the module to be finitely generated. This makes the proof of a linear bound for the powers completely straightforward in the case of an ideal generated in a single degree (for that application, it in fact suffices to notice that local cohomology is invariant under arbitrary base change, for a determined module), and a general linear bound on the regularity of powers follows.

Section 4 considers the case of $\mathfrak{m}$-primary ideals for standard graded algebras over a field. In that case, pretty precise results on the behaviour were given by Eisenbud and Ulrich and by D. Berlekamp. We prove a result that completes the one of Eisenbud and Ulrich and covers partly the work of Berlekamp for ideals that need not be generated in a single degree. The more precise results for the single-degree case are explained in Sect. 6.

The next section explains the geometric description of the constant that appears in the linear function that equals the regularity of a high enough power. This very nice description, for ideals generated in a single degree, is due to Eisenbud and Harris. They were able to prove it for $\mathfrak{m}$-primary ideals, and it was then extended by Tai Hà to arbitrary ideals, for a slight variation of the notion of regularity (the a^*-invariant). We here present the general result we obtain that provides information on the eventual behaviour of local cohomology of powers, which implies the result of Hà as well as the natural conjecture he had on the regularity of powers. It also provides further applications, and we mention one about the regularity of powers of ideal sheaves. In the last section, we give some further results on the regularity of powers of ideal sheaves (saturated powers, in the commutative algebra terminology) and symbolic powers that were obtained by several authors, including Cutkosky, Ein, Lazarsfeld, Niu and myself. These only cover a small part of the recent advances on symbolic powers; in particular we do not report about the behaviour of ideals of fat points.

In Sect. 6 we first provide a slight improvement of a result of Eisenbud and Ulrich on the behaviour of the regularity of powers of $\mathfrak{m}$-primary ideals that are generated in a single degree. The description in terms of regularities of graded strands of the Rees ring is a little more precise and general, and gives a coherent and pretty complete view of what happens. An extension to ideals that are not generated in a single degree can be obtained via similar techniques. However, this is more technical to state and prove and uses regularity of modules that need not be finitely generated as invariants. We preferred to include the proof of a generalization to ideals generated in a single degree that need not be $\mathfrak{m}$-primary. For this purpose, we introduce approximation complexes, in order to characterize Rees algebras of regularity 0. We could have only quoted the ingredient that we needed to prove Lemma 5 from the work of Herzog, Simis and Vasconcelos, or proved it under a more restrictive hypothesis, but it seemed to us that Proposition 4 can be of general use and was not stated in this simple form, even though none of its claims is new.

2 Some Early Advances on the Regularity of Powers of Ideals

Homological invariants of powers of ideals have been studied for quite some time. The stabilization of the associated primes attached to powers and the stabilization of the depth of powers, by Markus Brodmann, are among the first examples of this kind of result (see [3, 4, 25, 28] for these results and connected ones; [24] for a book collecting several of these; [5] for some geometric consequences; and [8] for a few more recent results).

Some years later, Vijay Kodiyalam proved in [22] the asymptotic stability of several homological invariants: Betti numbers, Bass numbers, length and minimal number of generators of some Tor and Ext modules attached to powers of an ideal. All these invariants for I^n (or more generally for I^nM, with M finitely generated) are eventually a polynomial in n. These results essentially follow from a version of the Hilbert-Serre theorem when properly applied to graded modules over the Rees algebra $\oplus_n I^n$ such as $\oplus_n \operatorname{Tor}_i^A(MI^n, k)$ when $I \subset A$ is an ideal of the local Noetherian ring A with residue field k, and M is a finitely generated A-module.

The first result on the regularity of powers of a graded ideal is due to Irena Swanson in [30]: for any homogeneous ideal I, there exists k such that $\operatorname{reg}(I^n) \leq kn$ for all n. This result easily follows from the existence of an integer ℓ and primary decompositions $I^n = \mathfrak{q}_{n,1} \cap \cdots \cap \mathfrak{q}_{n,s_n}$ with $\sqrt{\mathfrak{q}_{n,i}}^{\ell n} \subseteq \mathfrak{q}_{n,i}$, for all i. Indeed, this in particular shows that $(I^n)_\mu = ((I^n) : \mathfrak{m}^\infty)_\mu$ for $\mu \geq \ell n$ and the result then follows by induction on the number of variables, or on the dimension of A/I.

In small dimension, Chandler and Geramita, Gimigliano and Pitteloud [9, 17] discovered independently that $\operatorname{reg}(I^n) \leq n \operatorname{reg}(I)$ whenever $\dim(A/I) \leq 1$. In fact the statements are slightly more precise and were complemented by a collection of results on the regularity of modules of the form $\operatorname{Tor}_i^A(M, N)$ obtained by several authors. These are detailed in Sects. 5–7, of my survey [6] and I will not return to these here.

3 The Eventual Linearity

We first recall some definitions and notations. A standard graded algebra is a $\mathbf{Z}$-graded algebra generated by its elements of degree 1 over its component of degree 0, A_0. The initial degree of a $\mathbf{Z}$-graded module M is $\operatorname{indeg}(M) := \inf\{\mu \mid M_\mu \neq 0\}$, and its end is $\operatorname{end}(M) := \sup\{\mu \mid M_\mu \neq 0\}$. We recall that $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$. For a graded ideal I, $I_{\leq\mu}$ is the set of elements in I of degree at most μ and $(I_{\leq\mu})$ the A-ideal they generate.

The most significant simple result on the regularity of powers of graded ideals is the following one, due independently to Vijay Kodiyalam [23] and to Cutkosky, Herzog and Trung [12]. Let us state it in a slightly generalized form due to Trung and Wang in [32]:

Theorem 1. *Let A be a standard graded Noetherian algebra. If $I \neq 0$ is a graded ideal and $M \neq 0$ a finitely generated graded A-module, there exists n_1 and b such that*

$$\operatorname{reg}(I^n M) = nd + b, \quad \forall n \geq n_1,$$

with

$$d := \min\left\{\mu \mid \exists m \geq 1,\ (I_{\leq \mu}) I^{m-1} M = I^m M\right\}.$$

Notice that the number d is bounded above by the maximal degree of a generator of I.

The first statements were for the case where A is a polynomial ring over a field. The value of d was given by Kodiyalam in that case. On the other hand, Cutkosky, Herzog and Trung deduced this result from the asymptotic linearity of the highest degree of an element in the ith syzygy module, for any i.

The result on the linearity of the highest degree of a ith syzygy admits several refinements and generalizations. The simplest to state is as follows. It was obtained by Gwyneth Whieldon in [33] for a standard graded polynomial ring over a field and independently by Amir Bagheri, Tai Hà and myself in more generality; in the statement, G denotes an abelian group:

Theorem 2. *Let A be a G-graded Noetherian algebra over $S := A_0$. Let $I \neq 0$ be a graded ideal generated in a single degree $d \in G$ and $M \neq 0$ be a finitely generated graded A-module. Then, for every i, there exists a finite set $\Delta_i \subseteq G$ such that:*

(1) *For all $n \in \mathbf{N}$, $\operatorname{Tor}_i^S(MI^n, A)_\eta = 0$ if $\eta \notin \Delta_i + nd$.*
(2) *There exists a subset $\Delta_i' \subset \Delta_i$ such that $\operatorname{Tor}_i^S(MI^n, A)_\eta \neq 0$ for $n \gg 0$ and $\eta \in \Delta_i' + nd$, and $\operatorname{Tor}_i^S(MI^n, A)_\eta = 0$ for $n \gg 0$ and $\eta \notin \Delta_i' + nd$.*
(3) *Let $A \to k$ be a ring homomorphism to a field k. Then for any δ and any j, the function*

$$\dim_k \operatorname{Tor}_i^S(MI^n, k)_{\delta + nd}$$

is polynomial in n for $n \gg 0$.

More general and precise statements are given in [1], that considers the G-graded Betti numbers of $I_1^{n_1} \cdots I_s^{n_s} M$ for a collection of ideals which need not be generated in a single degree.

To provide a simple corollary of these general statements, let us mention that they imply the following:

Theorem 3. *Let A be a positively graded Noetherian algebra over $S := A_0$. Let $I \neq 0$ be a graded ideal generated in degrees $d_1, \ldots, d_s$, and $M \neq 0$ a finitely generated graded A-module. Then, for any i, there exists $e_i, f_i \in \{d_1, \ldots, d_s\}$ and*

$a_i, b_i \in \mathbf{Z}$, or $(a_i, b_i) = (+\infty, -\infty)$, such that for $n \gg 0$

$$\mathrm{indeg}(\mathrm{Tor}_i^A(MI^n, S)) = nf_i + a_i$$

and

$$\mathrm{end}(\mathrm{Tor}_i^A(MI^n, S)) = ne_i + b_i.$$

The statement for the initial degree is possibly new, and the second one is a slight refinement of the initial result of Cutkosky, Herzog and Trung: it ensures that the number e_i is one of the degrees of generators of I.

The proof of these results in [1] relies on the following remark. If I is generated by forms $f_1, \dots, f_r$ of degrees $d_1, \dots, d_r$, the Rees algebra $\oplus_n I^n$ is a $G \times \mathbf{Z}$-graded quotient of $A[T_1, \dots, T_r]$, obtained by mapping T_i to f_i, setting $\deg(T_i) := (d_i, 1)$ and $\deg(a) := (\deg(a), 0)$ for $a \in A$.

Taking a free $G \times \mathbf{Z}$-graded resolution $F_\bullet$ of the Rees algebra, one notices that, as $\deg(a) \in G \times 0$ for $a \in A$, $\oplus_n \mathrm{Tor}_i^A(I^n, S) = H_i(F_\bullet \otimes_A S)$ is a finitely generated $G \times \mathbf{Z}$-graded $S' := S[T_1, \dots, T_r]$-module. One then studies the structure of the subset $\mathrm{Supp}_{G\times\mathbf{Z}}(N) := \{\gamma \in G \times \mathbf{Z} \mid N_\gamma \neq 0\}$ for a finitely generated $G \times \mathbf{Z}$-graded S'-module N to provide a description of $\mathrm{Supp}_{G\times\mathbf{Z}}(H_i(F_\bullet \otimes_A S)) \cap G \times \{n\}$ for $n \gg 0$.

Let us now recall the two equivalent definitions of Castelnuovo–Mumford regularity. It is worth noticing that the equivalence of definitions is valid not only for algebras over a field. Sometimes it is also useful to not have to assume finite generation. This general version is probably due to Jouanolou, and is proved in [8, 2.4].

Theorem 4. *Let A be a finitely generated standard graded algebra over a commutative ring $S := A_0$ and M be a graded A-module. Set*

$$a^i(M) := \sup\{\mu \mid H^i_{A_+}(M)_\mu \neq 0\} \text{ and } b_i(M) := \sup\{\mu \mid \mathrm{Tor}_i^A(M, S)_\mu \neq 0\}.$$

Then $\mathrm{reg}(M) := \max_i\{a^i(M) + i\} = \max_j\{b_j(M) - j\} \in \mathbf{Z} \cup \{\pm\infty\}$ *is called the Castelnuovo–Mumford regularity of M.*

An important fact about local cohomology, which comes from its computation using a Čech complex, is that, for a fixed module, it commutes with arbitrary base change. If this base change does not affect the grading, this will transfer to regularity. Let us give a very useful specific example. Let A be a finitely generated standard graded algebra over $S := A_0$ and $I = (f_1, \dots, f_r)$ be a graded A-ideal generated in a single degree $d := \deg f_i$, and set $S' := S[T_1, \dots, T_r]$. The graded map $A' := A[T_1, \dots, T_r] \to \mathcal{R}_I := \oplus_{n\geq 0} I^n(nd)$ with $\deg(T_i) = 0$ makes the Rees algebra a quotient of the standard graded algebra A' over $S' = A'_0$, and a (not finitely generated) graded A-module. Notice that we do use here that I is generated in a single degree, a case that is much better understood than the general case. (This hypothesis could be weakened without much changes to the case where I has a reduction generated in a single degree). Considering $\mathcal{R}_I$ as a A-module or as a A'-module does not affect its regularity, and the following result follows.

Theorem 5. *With the above notations,*

$$\sup_n\{\text{reg}(I^n) - nd\} = \text{reg}(\mathcal{R}_I).$$

Furthermore, $\text{reg}(\mathcal{R}_I) \in \mathbf{Z}_{\geq 0}$ *if S is Noetherian.*

Notice that the finiteness of $\text{reg}(\mathcal{R}_I)$ comes from the Noetherianity of A', using the definition in terms of Tor. The inequality $\text{reg}(I^n) \leq nd + \text{reg}(\mathcal{R}_I)$ was proved by Tim Römer in [29, 5.3] using a notion of bigeneric initial ideal. This work of Römer also contains several corollaries and connected results, as well as estimates of the power were linearity occurs in terms of the bigeneric initial ideal.

A linear bound for the regularity of powers can easily be deduced. Let

$$d_1 := \min\{\mu \mid (I_\mu) \text{ and } I \text{ coincide off } V(A_+)\},$$

and notice that it is bounded above by the highest degree of an element in any set of generators of I. It follows from the definition of d_1 that I^n and K^n coincide off $V(A_+)$, and using the definition of regularity in terms of local cohomology and the inclusion $K^n \subseteq I^n$ one has

$$\text{reg}(I^n) \leq \text{reg}(K^n) \leq nd_1 + \text{reg}(\mathcal{R}_K).$$

We will see below that, when I is supported in $V(A_+)$, the number d_1 coincides with d as defined in Theorem 1.

4 The $\mathfrak{m}$-Primary Case

Let $A := k[x_1, \ldots, x_m]$ be a standard graded algebra over a field k, $\mathfrak{m} := (x_1, \ldots, x_m)$, M be a finitely generated graded A-module and I be a graded ideal such that M/IM is $\mathfrak{m}$-primary. Let $d := \min\{\mu \mid \exists p,\ (I_{\leq\mu})I^pM = I^{p+1}M\}$ and $J := (I_{\leq d})$. By [23] in the case that A is regular and [32] in general, there exists $b \geq \text{indeg}(M)$ such that

$$\text{reg}(M/I^nM) = nd + b - 1, \quad \forall t \gg 0.$$

Set $n_0 := \min\{n \mid \mathfrak{m}^d I^n M \subseteq JM\}$.

Proposition 1. *With notations as above:*

(i) $d = \min\{\mu \mid M/(I_{\leq\mu})M$ *is* $\mathfrak{m}$*-primary*$\}$.

(ii) *The function* $f(n) := \text{reg}(M/I^nM) - nd = \text{end}(M/I^nM) - nd$ *is weakly decreasing for* $n \geq \min\{1, n_0\}$.

(iii) One has

$$n_0 \leq \left\lceil \frac{\operatorname{reg}(M/JM) - \operatorname{indeg}(M) - d + 1}{d+1} \right\rceil$$

and $\operatorname{reg}(M/JM) \leq \operatorname{reg}(M) + (d-1)\dim M$.

Proof. Let $d_0 := \min\{\mu \mid M/(I_{\leq\mu})M$ is $\mathfrak{m}$-primary$\}$, $K := (I_{\leq d_0})$ and $t_0 := \min\{t > 0 \mid \mathfrak{m}^{d_0} I^t M \subseteq KM\}$.

Notice first that $d_0 \leq d$ as M/JM is $\mathfrak{m}$-primary.

As $(I^t M \cap KM)/I^t KM = \operatorname{Tor}_1^A(M/I^t M, A/K)$, and the latter module is a subquotient of $F \otimes_A M/I^t M$ if $F \to A$ presents A/K, it follows that $\operatorname{reg}((I^t M \cap KM)/I^t KM) \leq \operatorname{reg}(M/I^t M) + d_0$ and $\operatorname{reg}(M/\mathfrak{m}^{d_0} I^t M) \leq \operatorname{reg}(M/I^t M) + d_0$ for $t \geq 1$. Hence, if $t \geq t_0$,

$$\begin{aligned}
\operatorname{reg}(M/I^{t+1}M) &\leq \operatorname{reg}(M/I^t KM) \\
&\leq \max\{\operatorname{reg}(M/I^t M) + d_0, \operatorname{reg}(M/KM)\} \\
&\leq \max\{\operatorname{reg}(M/I^t M) + d_0, \operatorname{reg}(M/\mathfrak{m}^{d_0} I^t M)\} \\
&\leq \operatorname{reg}(M/I^t M) + d_0,
\end{aligned}$$

which implies that $d \leq d_0$; hence $d_0 = d$, $J = K$ and $t_0 = n_0$. This shows (i) and (ii).

For (iii) notice that $I \subseteq J + \mathfrak{m}^{d+1}$; hence $\mathfrak{m}^d I^t \subseteq J + \mathfrak{m}^{t(d+1)+d}$. Now $\mathfrak{m}^{t(d+1)+d} M \subseteq M_{\geq \operatorname{indeg}(M)+t(d+1)+d}$ and $M_{\geq\nu} \subseteq (JM)_{\geq\nu}$ if (and only if) $\nu > \operatorname{reg}(M/JM)$. Finally notice that $\operatorname{reg}(M/JM) \leq \operatorname{reg}(M) + (d-1)\dim M$ (e.g. by [6, 1.5.2 (i)]). ☐

For ideals generated in a single degree, point (ii) is Proposition 1.1 in [14], and the argument above shows that it also follows from [16, 7.5].

In the case where A is a polynomial ring and $M = A$, it shows the following result that is of interest in the unequal degree case:

Corollary 2. *Let I be a homogeneous ideal of a standard graded polynomial ring $A := k[X_1, \ldots, X_m]$ over a field k. Let*

$$d := \min\{\mu \,|(I_{\leq\mu}) \text{ is } \mathfrak{m}\text{-primary}\}$$

and $n_0 := \left\lceil \frac{(m-1)(d-1)}{d+1} \right\rceil \leq m-1$.

Then $\operatorname{reg}(I^n) - dn$ *is a nonnegative weakly decreasing function of n, for $n \geq \min\{1, n_0\}$.*

In Example 2.3 of [15], $m = 4$ and $d = 5$ and the bound for n_0 is sharp.

Several interesting results on the behaviour of the regularity of powers of $\mathfrak{m}$-primary ideals were obtained by David Berlekamp. In particular, results closely related to the above proposition figures in his article [2]. For instance, he proves the following [2, 2.8]:

$\operatorname{reg}(I^n) - dn$ is a nonnegative weakly decreasing function of n for $n > \frac{\operatorname{reg}(J)}{d'}$, where d' is the highest degree of a minimal generator of I. The bound above $n \geq \frac{\operatorname{reg}(J)-d}{d+1}$ could not be directly compared to this bound.

5 The Constant

The first important result towards a description of the number b such that $\operatorname{reg}(I^n) = nd + b$ for $n \gg 0$ is due to Eisenbud and Harris (who also quote early advances together with Huneke). They provide a geometric description in the equal degree $\mathfrak{m}$-primary case that we will recall below. Then Tai Hà proved in [18] an extension for the equal degree case, without the $\mathfrak{m}$-primary assumption, for a very closely related invariant, $a^*(I^n) := \max_i\{a^i(I^n)\}$. More recently, I provided a general description of the number $\lim_{n\to\infty}(a^i(I^n) - nd) \in \mathbf{Z} \cup \{-\infty\}$ that in particular implies the natural conjecture given by Hà in view of his results. Although the result in [7] treats the unequal degree case, the geometric description is much less satisfying, except perhaps when (I_d) is a reduction of I.

We will only discuss here the equal degree case and for simplicity assume that the base ring is a field and further not involve a module M as in the previous section.

Let A be a positively graded Noetherian algebra over a field k and I be a graded A-ideal generated by r forms $f_1, \ldots, f_r$ of degree d. These forms define a rational map :

$$\phi : \mathcal{S} := \operatorname{Proj}(A) \cdots\to \mathbf{P}^{r-1}.$$

This rational map is a morphism when I is $\mathfrak{m}$-primary.

The closure Γ of the graph of ϕ is the irreducible subscheme of $\mathcal{S} \times \mathbf{P}^{r-1}$ defined by the Rees algebra. More precisely, let $A' := A[T_1, \ldots, T_r]$, $B := k[T_1, \ldots, T_r]$, set $\operatorname{bideg}(T_i) := (0, 1)$ and $\operatorname{bideg}(a) := (\deg(a), 0)$ for $a \in A$. The natural bigraded morphism of bigraded k-algebras

$$A' \xrightarrow{\psi} \mathcal{R}_I := \oplus_{t\geq 0} I(d)^t = \oplus_{t\geq 0} I^t(dt),$$

sending T_i to f_i, is onto and correspond to the embedding $\Gamma \subset \mathcal{S} \times \mathbf{P}^{r-1}$.

Notice that taking the projectivization of $\mathcal{R}_I$ with respect to the simple grading coming from A ($\deg(T_i) = 0$) gives an irreducible scheme $\widetilde{\Gamma} \subset \mathcal{S} \times \mathbf{A}^r$ whose projectivization is Γ. When A is standard graded, there is a natural embedding $\mathcal{S} \subseteq \mathbf{P}^s$ with $s + 1 := \dim_k(A_1)$; hence

$$\Gamma \subset \mathbf{P}^s \times \mathbf{P}^{r-1} = \mathbf{P}^s_{\mathbf{P}^{r-1}} \quad \text{and} \quad \widetilde{\Gamma} \subset \mathbf{P}^s \times \mathbf{A}^r = \mathbf{P}^s_{\mathbf{A}^r}.$$

We have already seen that $\max_n\{\operatorname{reg}(I^n) - nd\} = \operatorname{reg}(\mathcal{R}_I) = \operatorname{reg}(\widetilde{\Gamma})$.

The regularity of a subscheme $X \subset \mathbf{P}^n_Y$, $Y = \mathrm{Spec}(C_0)$ is defined as the regularity of the unique graded quotient Q of $C := C_0[X_0, \dots, X_n]$ with $X = \mathrm{Proj}(Q)$ and $H^0_{C_+}(Q) = 0$. Q is the quotient of C by the graded C-ideal $I_X := \oplus_\mu H^0(\mathbf{P}^n_Y, \mathcal{I}_X(\mu))$, where $\mathcal{I}_X$ is the sheaf of ideals defining X as a subscheme of $\mathbf{P}^n_Y$. Equivalently, $\mathrm{reg}(X) = \max_{i \geq 1}\{a^i(Q) + i\}$ for any Q, quotient of C such that $X = \mathrm{Proj}(Q)$.

We define the geometric regularity as $\mathrm{greg}(X) := \max_{i \geq 2}\{a^i(Q) + i\}$, with Q as above. Alternatively, one has $\mathrm{greg}(X) := \mathrm{reg}(\oplus_\mu H^0(\mathbf{P}^n_Y, \mathcal{O}_X(\mu)))$. Notice also that $\mathrm{reg}(X) = \mathrm{greg}(\mathbb{X})$ with $\mathbb{X} \subset \mathbf{P}^{n+1}_Y$ defined by $I_X C[X_{n+1}]$. The name geometric regularity has the (weak) following justification : if X' is an isomorphic projection of X to a linear subspace of $\mathbf{P}^n_Y$, then $\mathrm{greg}(X') = \mathrm{greg}(X)$ while $\mathrm{reg}(X') \geq \mathrm{reg}(X)$, and this inequality can be strict. However, this notion depends very much on the embedding of X. Regularity is based on the vanishing of graded pieces of cohomology (*i.e.* of specific twists in sheaf cohomology), thus will typically change after a Veronese embedding.

Local cohomology commutes with localization, in particular with localization on the base (degree zero elements). It follows that (geometric) regularity is a notion that is local on the base. Indeed, if M is a graded module over $C_0[x_1, \dots, x_n]$, then $a^i(M) = \sup_{\mathfrak{p} \in Y}\{a^i(M \otimes_{C_0} (C_0)_{\mathfrak{p}})\}$, and this supremum is a maximum if $a^i(M) < \infty$. Also one may take the supremum among maximal ideals of Y. Alternatively, one can similarly remark that the formation of Tor modules commutes with localization on the base and uses the second definition of regularity.

From this fact, one easily derives a notion of regularity for a subscheme of $\mathbf{P}^s_Y$ for any scheme Y (see [7] for a detailed treatment). With this definition, considering $\Gamma \subset \mathbf{P}^s_{\mathbf{P}^{r-1}}$ and $\widetilde{\Gamma} \subset \mathbf{P}^s_{\mathbf{A}^r}$, one has :

Theorem 1. *In the above situation;*

(1) $\max_n\{\mathrm{reg}(I^n) - nd\} = \mathrm{reg}(\widetilde{\Gamma})$.
(2) $\mathrm{reg}(I^n) - nd = \mathrm{reg}(\Gamma)$*, for* $n \gg 0$.
(3) $\lim_{n \to +\infty}(\mathrm{reg}(I^n)^{sat} - nd) = \mathrm{greg}(\Gamma)$.

The regularity of Γ is defined as the maximal regularity of the stalks over points in $\mathbf{P}^{r-1}$, while in the work of Eisenbud and Harris the maximal is taken over the fibers. In fact, according to [7, Sect. 6], both coincide as long as all non-empty fibers have a common Hilbert polynomial up to constant term (a weakening of the flatness condition that requires this polynomial to be constant). This condition holds when all fibers are finite, as in the case considered by Eisenbud and Harris.

Let us point out that by [27, 2.9], the fibers are all finite if and only if $\mathrm{Proj}(A/I^n)$ is defined scheme-theoretically by equations of degree $< nd$ for some n.

The theorem follows from the more precise result below that shows such an equality for each of the invariants $a^i(I^n)$ individually.

Theorem 2. *In the above situation, let* $B' := B/\mathrm{ann}_B(\ker(\psi))$. *Then,*

$$\lim_{t \to +\infty}(a^i(I^t) - td) = \max_{\mathfrak{q} \in \mathrm{Proj}(B')}\{a^i(\mathcal{R}_I \otimes_{B'} B'_{\mathfrak{q}})\}.$$

This result has another interesting consequence : if the projection of Γ to its image $\mathrm{Proj}(B') \subseteq \mathbf{P}^{r-1}$ has a fiber of positive dimension, then $\lim_{t\to+\infty}(a^i(I^t) - td) \geq -i$ for some $i \geq 2$, showing that the regularity of the powers of the ideal sheaf associated to I is eventually a linear function with leading coefficient d.

6 The Stabilization Index

From the result of Kodiyalam and Cutkosky, Herzog and Trung, a natural question arises : from which point on the regularity of I^n is a linear function of n ?

In this section we address this question. Alternative arguments are given to derive a result of Eisenbud and Ulrich, that actually leads to some improvements. We then present results of Berlekamp and provide an extension from the $\mathfrak{m}$-primary case to a more general situation. We first introduce some notation.

Let $I = (f_1,\dots,f_r)$ with $\deg f_i = d$ for $1 \leq i \leq r$, $B := k[T_1,\dots,T_r]$, $\mathfrak{n} := (T_1,\dots,T_r)$ and $A' := A[T_1,\dots,T_m]$.

We set $\mathrm{bideg}(T_i) := (0,1)$ and $\mathrm{bideg}(a) := (\deg(a),0)$ for $a \in A$. The natural onto map $A' \to \mathcal{R}_I$ makes $M\mathcal{R}_I$ a bigraded A'-module.

Let $N := \oplus_n (M/I^nM)(nd)$ with the A'-module structure obtained from the one of A-module by setting $T^\alpha m := f^\alpha m$. In other words, N is the cokernel of the map $\oplus_n I^nM(nd) \to \oplus_n M(nd)$. If M has positive depth, $N \subset H^1_{\mathfrak{m}}(M\mathcal{R}_I)$ and equality holds if M has depth at least 2. In particular, $N_{\mu,*}$ is a finitely generated graded B-module for any μ. Notice that if A is standard graded, $N_{\mu,n} = 0$ implies $N_{\mu+1,n} = 0$ for any $\mu \geq \mathrm{end}(M/\mathfrak{m}M) - nd$.

The following lemma will be very useful to compute the stabilization index. Its proof is a little bit involved at this level of generality. However, if for instance depth $M \geq 2$, its proof simplifies quite a lot.

Lemma 1. *One has an exact sequence*

$$0 \to \oplus_{n<0} M(nd) \to H^1_{\mathfrak{n}}(M\mathcal{R}_I) \to H^0_{\mathfrak{n}}(N) \to 0,$$

and for any $i \neq 0$,

$$H^i_{\mathfrak{n}}(N) = H^{i+1}_{\mathfrak{n}}(M\mathcal{R}_I) = H^{i+1}_{\mathfrak{m}+\mathfrak{n}}(M\mathcal{R}_I).$$

Furthermore, $H^i_{\mathfrak{n}}(N) = H^i_{\mathfrak{n}}(H^1_{\mathfrak{m}}(M\mathcal{R}_I))$ *if* $H^0_{\mathfrak{m}}(M) = 0$.

Proof. We may replace A by $A/\mathrm{ann}_A(M)$ to assume that I is $\mathfrak{m}$-primary. Let $\mathcal{R}^e_I := \oplus_{n\in\mathbf{Z}} I^n(nd)$. Consider the two spectral sequences arising from the double complex $\mathcal{C}^\bullet_{\mathfrak{n}}\mathcal{C}^\bullet_{\mathfrak{m}} M\mathcal{R}^e_I$ abutting to $H^\bullet_{\mathfrak{n}+\mathfrak{m}}(M\mathcal{R}^e_I)$. As I is $\mathfrak{m}$-primary, after inverting any X_i, the ideal I is the unit ideal, and $(M\mathcal{R}^e_I)_{X_i}$ is isomorphic to $M_{X_i}[t^{-1},t]$ where t is the class of T_j for j such that f_j is a unit in A_{X_i}, and $\mathfrak{n}$ coincides with (t) after this localization. It follows that $H^i_{\mathfrak{n}}(\mathcal{C}^j_{\mathfrak{m}} M\mathcal{R}^e_I) = 0$ for $j > 0$; hence

$$H^i_{\mathfrak{n}}(M\mathcal{R}^e_I) = H^i_{\mathfrak{n}+\mathfrak{m}}(M\mathcal{R}^e_I).$$

On the other hand, $H^j_{\mathfrak{m}}(M\mathcal{R}^e_I) = \oplus_{n\in\mathbf{Z}} H^j_{\mathfrak{m}}(M)(nd)$ for $j \geq 2$, and there is a natural graded exact sequence :

$$0\to H^0_{\mathfrak{m}}(M\mathcal{R}^e_I)\to \oplus_{n\in\mathbf{Z}} H^0_{\mathfrak{m}}(M)(nd)\to N\to H^1_{\mathfrak{m}}(M\mathcal{R}^e_I)\to \oplus_{n\in\mathbf{Z}} H^1_{\mathfrak{m}}(M)(nd)\to 0.$$

It follows that $H^i_{\mathfrak{n}}(H^j_{\mathfrak{m}}(M\mathcal{R}^e_I)) = 0$ for $j \geq 2$, and there is a long exact sequence:

$$0\to H^1_{\mathfrak{n}}(H^0_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^0_{\mathfrak{n}}(N)\to H^0_{\mathfrak{n}}(H^1_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^2_{\mathfrak{n}}(H^0_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^1_{\mathfrak{n}}(N)$$
$$\to H^1_{\mathfrak{n}}(H^1_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^3_{\mathfrak{n}}(H^0_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^2_{\mathfrak{n}}(N)\to\cdots$$

The second spectral sequence has second terms $H^i_{\mathfrak{n}}H^j_{\mathfrak{m}}(M\mathcal{R}^e_I)$ and abutment $E^{i,j}_\infty$ equal to the cokernel of $H^{i-2}_{\mathfrak{n}}(H^1_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^i_{\mathfrak{n}}(H^0_{\mathfrak{m}}(M\mathcal{R}^e_I))$ if $j = 0$ and equal to the kernel of $H^i_{\mathfrak{n}}(H^1_{\mathfrak{m}}(M\mathcal{R}^e_I))\to H^{i+2}_{\mathfrak{n}}(H^0_{\mathfrak{m}}(M\mathcal{R}^e_I))$ if $j = 1$.

It in particular provides the identification $H^i_{\mathfrak{n}}(N) = H^{i+1}_{\mathfrak{m}+\mathfrak{n}}(M\mathcal{R}^e_I)$ for $i \geq 0$.

Now, the exact sequence

$$0\to M\mathcal{R}_I\to M\mathcal{R}^e_I\to \oplus_{n<0} M(nd)\to 0$$

shows that $H^i_{\mathfrak{n}}(\mathcal{R}_I) = H^i_{\mathfrak{n}}(\mathcal{R}^e_I)$ for $i \neq 1$ (recall that $H^0_{\mathfrak{n}}(M\mathcal{R}^e_I) = 0$) and provides the graded exact sequence

$$0\to \oplus_{n<0} M(nd)\to H^1_{\mathfrak{n}}(M\mathcal{R}_I)\to H^1_{\mathfrak{n}}(M\mathcal{R}^e_I)\to 0. \qquad \square$$

The following result is a slight improvement of the result of Eisenbud and Ulrich in [15] on the stabilization of the regularity of powers of an ideal generated in a single degree.

Theorem 2. *Let A be a standard graded algebra over a field, $\mathfrak{m} := A_+$, M be a finitely generated graded A-module and I be a graded ideal generated in degree d such that M/IM is $\mathfrak{m}$-primary. Then there exists $b \geq \mathrm{indeg}(M)$ such that*

$$\mathrm{reg}(M/I^nM) = nd + b - 1, \quad \forall n \geq \mathrm{reg}((M\mathcal{R}_I)_{b,*}) = \mathrm{end}(H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{b,*}) + 1.$$

More precisely,

(i) $H^i_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,*} = H^i_{\mathfrak{n}}((M\mathcal{R}_I)_{\mu,*}) = 0$ *for* $\mu \geq b$ *and* $i \neq 1$.

(ii) $r(\mu) := \mathrm{reg}((M\mathcal{R}_I)_{\mu,*}) < \infty$ *is a weakly decreasing function of* μ, *for* $\mu \geq \max\{b, \mathrm{end}(M/\mathfrak{m}M) - d\}$.

(iii) The values $\mu_1 > \mu_2 > \cdots > \mu_l = b$ *of* $\mathrm{reg}(M/I^nM) - nd + 1$ *for* $n > 0$ *satisfy*

$$\mathrm{reg}(M/I^nM) = nd + \mu_i - 1 \quad \Leftrightarrow r(\mu_i) \leq n < r(\mu_{i+1})$$

if one sets $r(\mu_{l+1}) := +\infty$.

Proof. Let b be defined by $\mathrm{reg}(M/I^nM) = nd + b - 1$ for $n \gg 0$.

For (i), notice that if $\mu \geq b$, $N_{\mu,n} = 0$ for $n \gg 0$; hence $N_{\mu,*} = H^0_{\mathfrak{n}}(N_{\mu,*})$ and $H^i_{\mathfrak{n}}(N_{\mu,*}) = 0$ for $i > 0$; the result then follows from Lemma 1. This implies that $\mathrm{reg}((M\mathcal{R}_I)_{\mu,*}) = \mathrm{end}(H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,*}) + 1$ for $\mu \geq b$.

Furthermore it shows that $H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,n} = 0 \Rightarrow H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,n+1} = 0$ if $\mu \geq b$.

By Proposition 1 (ii), $\mathrm{reg}(M/I^nM) - nd$ is a weakly decreasing function of n whose eventual value is $b-1$. By Lemma 1 $H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,n} = H^0_{\mathfrak{n}}(N)_{\mu,n}$, for $n \geq 0$; hence $H^1_{\mathfrak{n}}(M\mathcal{R}_I)_{\mu,n} = N_{\mu,n}$ if $\mu \geq b$ and $n \geq 0$, which proves that

$$\max\{n \mid \mathrm{reg}(M/I^nM) - nd = \mu_i - 1\} = \mathrm{end}(N_{\mu_i,*}), \forall i < l.$$

Now $N_{\mu,n} = 0$ implies $N_{\mu+1,n} = 0$ if M has no minimal generator in degree $\mu + nd + 1$, which implies (ii). □

Notice that if $b \geq \mathrm{end}(M/\mathfrak{m}M) - d$, as is for instance the case when M is cyclic or generated in a single degree, then (iii) takes the nicer following form, using (ii) :

(iii) For $n \geq 1$ and $\mu \geq b$,

$$\mathrm{reg}(M/I^nM) = nd + \mu - 1 \quad \Leftrightarrow \quad \begin{cases} r(\mu) \leq n < r(\mu - 1) & \text{if } \mu > b, \\ n \geq r(b) \quad \text{if } \mu = b. \end{cases}$$

In the case of unequal degree, Berlekamp obtained an explicit bound for monomial ideals. An $\mathfrak{m}$-primary monomial ideal I in $k[X_1, \dots, X_m]$ is minimally generated (after permuting the variables, if needed) by monomials as follows:

$$I = (X_1^d, \dots, X_l^d, X_{l+1}^{d_1}, \dots, X_m^{d_{m-l}}, h_1, \dots, h_r),$$

where the h_i's are not powers of variables and $d > d_1 \geq d_2 \geq \cdots \geq d_{m-l}$. Notice that d is the minimal degree μ such that $(I_{\leq\mu})$ is $\mathfrak{m}$-primary; hence $\mathrm{reg}(I^n) = nd + b$ for some $b \geq 0$ and $n \gg 0$. Then,

Theorem 3. [2, 3.1]. *The value of* $\mathrm{reg}(I^n) - nd$ *is a non negative constant for* $n > (m-1)\max\{1, l(d-1) - 1\}$.

For proving stabilization results for non-$\mathfrak{m}$-primary ideals, it is important to be able to control the regularity of the Rees algebra. To this end we will use approximation complexes. Let us briefly recall their construction.

If $I = (f_1, \dots, f_r)$ is a R-ideal for some commutative ring R, the Koszul complexes $K^f_\bullet := K_\bullet(f_1, \dots, f_r; R[T_1, \dots, T_r])$ (with $f_i \in R \subset R[T_1, \dots, T_r]$) and $K^T_\bullet := K_\bullet(T_1, \dots, T_r; R[T_1, \dots, T_r])$ have same modules as components, and their differentials satisfy $\delta^T \circ \delta^f + \delta^f \circ \delta^T = 0$, proving that the i-cycles Z^f_i of $K^f_\bullet$ are mapped to Z^f_{i-1} by δ^T (this map has degree one). The map δ^T induces a map in homology from H^f_i to H^f_{i-1}. The complexes $\mathcal{Z} := (\mathcal{Z}^f_\bullet, \delta^T)$ and $\mathcal{M} := (\mathcal{H}^f_\bullet, \delta^T)$, with $\mathcal{Z}_i := Z_i(-i)$ and $\mathcal{H}_i := H_i(-i)$, are two of the three approximation complexes constructed by Herzog, Simis and Vasconcelos. They depend on the choice of generators, but their homology does not. One has $H_0(\mathcal{Z}) = \mathrm{Sym}_R(I)$ and $H_0(\mathcal{M}) = \mathrm{Sym}_R(I/I^2)$. The following result, that follows from [19] and [21], demonstrates their importance.

Proposition 4. *Let R be a Noetherian local ring and I an R-ideal. The following are equivalent :*

(i) *The $\mathcal{Z}$-complex resolves $\mathcal{R}_I$.*
(ii) *The $\mathcal{M}$-complex resolves $\mathcal{R}_I/I\mathcal{R}_I$.*
(iii) *The $\mathcal{M}$-complex has only 0th homology.*
(iv) $\mathrm{reg}(\mathcal{R}_I) = 0$.
(v) $\mathrm{reg}(\mathcal{R}_I/I\mathcal{R}_I) = 0$.
If further R has infinite residue field, these are also equivalent to :
(vi) *I is generated by a d-sequence.*

Recall that $I\mathcal{R}_I = (\mathcal{R}_I)_+(-1) \subset \mathcal{R}_I$ and $\mathcal{R}_I/I\mathcal{R}_I = R/I \oplus_R I/I^2 \oplus_R \cdots$.

Proof. Replacing R by $R(U)$, properties (i) to (v) are unchanged. Hence we may assume that the residue field is infinite.

By [21, 4.1], (iv) and (v) are equivalent. By [19, 12.9], (i), (iii) and (vi) are equivalent. The implication (ii)⇒(iii) is trivial, and the reverse one holds since (i) is implied by (iii) and shows that $\mathrm{Sym}_R(I/I^2) = \mathcal{R}_I/I\mathcal{R}_I$ by [19, 3.1]. Finally (v) and (vi) are equivalent by [19, 12.7, 12.8 and 12.10]. □

Let $A := k[X_1, \ldots, X_n]$, $\mathfrak{m} := (X_1, \ldots, X_n)$ and $I = (f_1, \ldots, f_m)$ be an ideal with $\deg f_i = d$ for all i. Set $B := k[T_1, \ldots, T_r]$, $\mathfrak{n}$ and $A' := A[T_1, \ldots, T_r]$.

We will denote by $\mathcal{Z}'_\bullet(f; A)$ the augmented $\mathcal{Z}$-complex, with augmentation map the epimorphism $\mathcal{Z}'_0(f; A) = A' \to \mathcal{R}_I$.

Lemma 5. *Assume that $I_\mathfrak{p} \subset A_\mathfrak{p}$ satisfies one of the equivalent conditions of Proposition 4 for all graded prime ideals $\mathfrak{p} \not\supseteq A_+$. Then*

$$H^i_\mathfrak{n}(\mathcal{C}^j_\mathfrak{m}(\mathcal{R}_I))_{*,t} = 0$$

for $j > 0$ and $t > -i$.

Proof. It suffices to show that, for any j, $H^i_\mathfrak{n}((\mathcal{R}_I) \otimes_A A_{X_j})_{*,t} = 0$ for $t > -i$. Our assumptions imply that the augmented $\mathcal{Z}$-complex $\mathcal{Z}'_\bullet(f; A_{X_j})$ is exact. Denote by (–) shifts in the degrees in the T_i's, then $\mathcal{Z}_i = Z_i(f; A) \otimes_A A'(-i)$, where $Z_i(f; A)$ is the A-module of ith cycles in the Koszul complex $K(f; A)$. The double complex $\mathcal{C}^\bullet_\mathfrak{n}(\mathcal{Z}'_\bullet)$ gives rise to a spectral sequence abutting to zero with first terms:

$$_1E^p_q = \begin{cases} H^p_\mathfrak{n}((\mathcal{R}_I) \otimes_A A_{X_j}) & \text{if } q = -1, \\ Z_q(f; A_{X_j}) \otimes_{A_{X_j}} H^r_\mathfrak{n}(A')(-q) & \text{if } p = r \text{ and } q \geq 0, \\ 0 & \text{else.} \end{cases}$$

It follows that $H^r_{\mathfrak{n}}(A')_{*,\mu-i} = 0$ implies that $H^{r-i}_{\mathfrak{n}}(\mathcal{R}_I)_{*,\mu} = 0$. As $H^r_{\mathfrak{n}}(A')_{>-r} = 0$, the conclusion follows. □

Corollary 6. *Assume that $I_{\mathfrak{p}} \subset A_{\mathfrak{p}}$ satisfies one of the equivalent conditions of Proposition 4 for all graded prime ideals $\mathfrak{p} \not\supseteq A_+$. Then for $t > 0$,*

$$H^i_{\mathfrak{m}+\mathfrak{n}}(\mathcal{R}_I)_{\mu,t} = H^i_{\mathfrak{n}}(\mathcal{R}_I)_{\mu,t}.$$

Proof. The double complex $\mathcal{C}^\bullet_{\mathfrak{n}}\mathcal{C}^\bullet_{\mathfrak{m}}\mathcal{R}_I$ gives rise to two spectral sequences abutting to $H^\bullet_{\mathfrak{m}+\mathfrak{n}}(\mathcal{R}_I)$. One of them has first terms:

$${}_1E^{p,q} = H^p_{\mathfrak{n}}(\mathcal{C}^q_{\mathfrak{m}}(\mathcal{R}_I)).$$

For $t > 0$, $({}_1E^{p,q})_{*,t} = 0$ for $q > 0$ by Lemma 5, and the conclusion follows since $\mathcal{C}^0_{\mathfrak{m}}(\mathcal{R}_I) = \mathcal{R}_I$. □

The double complex $\mathcal{C}^\bullet_{\mathfrak{n}}\mathcal{C}^\bullet_{\mathfrak{m}}\mathcal{R}_I$ gives rise to another spectral sequence, whose first terms are ${}_1E^{p,q} = \mathcal{C}^p_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I))$, and ${}_2E^{p,q} = H^p_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I))$. Taking graded components, one has

$$({}_2E^{p,q})_{\mu,t} = H^p_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*})_{\mu,t}.$$

Hence,

$$a^p_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}) = \mathrm{end}(({}_2E^{p,q})_{\mu,*}).$$

Let μ be such that $H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}$ is supported in $V(\mathfrak{n})$ for all q. Then $({}_2E^{p,q})_{\mu,*} = 0$ for $p > 0$. Therefore $({}_2E^{p,q})_{\mu,*} = ({}_\infty E^{p,q})_{\mu,*}$ and

$$\mathrm{reg}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}) = a^0_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}) = \mathrm{end}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}) = \mathrm{end}(H^q_{\mathfrak{m}+\mathfrak{n}}(\mathcal{R}_I)_{\mu,*}).$$

The following result offers a generalization (in terms of the a^*-invariant) of the stabilization result for $\mathfrak{m}$-primary ideals; the hypothesis on I is for instance fulfilled when the projective scheme $\mathrm{Proj}(A/I)$ is locally a complete intersection.

Proposition 7. *Let I be a graded A-ideal generated in a single degree d, with $a^*(I^t) = dt + b$ for $t \gg 0$. Assume that $I_{\mathfrak{p}} \subset A_{\mathfrak{p}}$ satisfies one of the equivalent conditions of Proposition 4 for all graded prime ideals $\mathfrak{p} \not\supseteq A_+$. Then,*

$$a^*_{\mathfrak{m}}(I^t) \leq dt + b, \quad \forall t > \max_{\mu > b} a^*_{\mathfrak{n}}((\mathcal{R}_I)_{\mu,*}).$$

Proof. Let $\mu > b$. Then $H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}$ is supported in $V(\mathfrak{n})$ for all q. Hence $H^p_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,*}) = 0$ for all $p > 0$, and for $t > 0$,

$$H^q_{\mathfrak{m}}(I^t)_{\mu+dt} = H^0_{\mathfrak{n}}(H^q_{\mathfrak{m}}(\mathcal{R}_I)_{\mu,t}) = H^q_{\mathfrak{n}}(\mathcal{R}_I)_{\mu,t}$$

by Lemma 5. Now $H^q_{\mathfrak{n}}(\mathcal{R}_I)_{\mu,t} = (H^q_{\mathfrak{n}}(\mathcal{R}_I)_{\mu,*})_{\mu,t}$ vanishes for $t > a^*_{\mathfrak{n}}((\mathcal{R}_I)_{\mu,*})$. The result follows. □

Let us finally recall that Conca provided in [10] examples of ideals such that $\text{reg}(I^n)$ is a linear function of n for $n = 1, \ldots, -p, p$ arbitrary, without being a linear function of n for all $n \geq 1$.

7 Powers of Ideal Sheaves and Symbolic Powers

Let A be a standard graded algebra over a field, $\mathfrak{m} := A_+$, and I be a graded ideal generated in degree at most d.

We denote by $I^{sat} := \cup_s I : \mathfrak{m}^s$ the saturation of an ideal I with respect to $\mathfrak{m}$. This ideal is equal to the sections over the projective scheme defined by A of the sheaf of ideals defined by I.

Recall that $\text{reg}(A/I^{sat}) \leq \text{reg}(A/I)$ as both quotients share the same local cohomology modules of positive index and $H^0_{\mathfrak{m}}(A/I^{sat}) = 0$.

Cutkosky, Ein and Lazarsfeld proved in [13] that the limit

$$s(I) := \lim_{t \to \infty} \text{reg}((I^t)^{sat})/t$$

exists and is equal to the inverse of a Seshadri constant. It need not be in Q, see [11].

Using the existence of c such that $\text{reg}(MI^t) \leq dt + c$ for all t when I is generated in degree at most d and M is finitely generated, one can easily derive the existence of this limit.

For proving this we introduce some notation : $r_p := \text{reg}((I^p)^{sat})$ and

$$d_p := \min\{\mu \mid (I^p)^{sat} = ((I^p)^{sat}_{\leq \mu})^{sat}\},$$

the so-called degree of generation by global section for the sheaf associated to I^p.

First, one has $d_{p+q} \leq d_p + d_q$, proving that $s := \lim_{p \to \infty}(d_p/p)$ exists.

For any p there exists c_p such that $\text{reg}(((I^p)^{sat}_{\leq d_p})^t I^q) \leq td_p + c_p$ for all $t \geq 1$ and $0 \leq q < p$.

One has $(((I^p)^{sat}_{\leq d_p})^t I^q)^{sat} = (I^{pt+q})^{sat}$, and the inequalities

$$d_{pt+q} \leq r_{pt+q} \leq td_p + c_p$$

show that $\lim_{p \to \infty}(r_p/p) = s$ and that $d_p \geq ps$ for all p.

The same argument applies to any graded ideal J such that $\dim(A/J) \leq 1$. Setting $r^J_p := \text{reg}(I^p :_A J^\infty) \leq \text{reg}(I^p)$ and defining d^J_p similarly as above,

$$d^J_p := \min\{\mu \mid ((I^p : J^\infty)_{\leq \mu}) : J^\infty = I^p : J^\infty\},$$

the limits of r^J_p/p and d^J_p/p exist and are equal. For example, if $\text{Proj}(A/I)$ is a scheme with isolated non-locally complete intersection points, then $\lim_{p \to \infty} \text{reg}(I^{(p)}/p)$ exists, where $I^{(p)}$ denotes the pth symbolic power of I.

Similar arguments apply in a much more general setting : A, a Noetherian positively graded algebra; I, a graded A-ideal; and J a graded A-ideal of cohomological dimension (relative to A_+) at most 1 (see [7, 6.12]).

On the other hand, when A/J has dimension 2, it may be that $\mathrm{reg}(I : J^\infty) > \mathrm{reg}(I)$ for J, an embedded prime of I. This shows that the above argument is not directly applicable for symbolic powers in general.

In [27], Niu, using the definition of $s(I)$ as (the inverse of) a Seshadri constant, gave a geometric characterization of the condition $s(I) < d$. Recall that I is generated in degree at most d, and denote by Γ the closure of the graph of the rational map defined by generators of the vector space I_d. One has $\Gamma \subset \mathrm{Proj}(A) \times \mathbf{P}^t$ for some t, and we denote by π the projection of Γ to its image in $\mathbf{P}^t$.

With this definition, $s(I) < d$ if and only if the morphism π is finite. More precisely, π is finite if and only if $\mathrm{Proj}(A/I^t)$ is defined by equations of degree $< dt$ for some t, and if not, $\mathrm{reg}((I^t)^{sat}) - td$ is a non-negative constant for $t \gg 0$.

This follows from Theorem 2 and our arguments above. The article of Niu contains several other interesting results on the regularity of powers of ideal sheaves, in particular concerning the regularity of powers in terms of the regularity of the ideal sheaf in the case where it defines a locally complete intersection scheme.

In another work of Niu [26], it is proved that, if k is of characteristic 0, then there exists e such that

$$0 \leq \mathrm{reg}((I^t)^{sat}) - ts(I) \leq e, \quad \forall t.$$

This is a strong refinement of the existence of the limit defining $s(I)$ proved by Cutkosky, Ein and Lazarsfeld. It relies on vanishing theorems (Fujita's result). A more self-contained proof would be interesting.

References

1. Bagheri, A., Chardin, M., Ha, H.T.: Asymptotic linearity of multigraded betti numbers. arXiv:1110.0383 (2012)
2. Berlekamp, D.: Regularity defect stabilization of powers of an ideal. Math. Res. Lett. **19**(1), 109–119 (2012)
3. Brodmann, M.: Asymptotic stability of Ass(M/InM). Proc. Amer. Math. Soc. **74**(1), 16–18 (1979)
4. Brodmann, M.: The asymptotic nature of the analytic spread. Math. Proc. Cambridge Philos. Soc. **86**(1), 35–39 (1979)
5. Brodmann, M.: Asymptotic depth and connectedness in projective schemes. Proc. Amer. Math. Soc. **108**(3), 573–581 (1990)
6. Chardin, M.: Some results and questions on Castelnuovo-Mumford regularity. Syzygies and Hilbert Functions, vol. 140. Lecture Notes in Pure Applied Mathematics, vol. 254. Chapman & Hall/CRC, Boca Raton (2007)
7. Chardin, M.: Powers of ideals and the cohomology of stalks and fibers of morphisms. Algebra Number Theory. To appear
8. Chardin, M., Jouanolou, J.-P., Rahimi, A.: The eventual stability of depth, associated primes and cohomology of a graded module. J. Commut. Algebra. To appear

9. Chandler, K.A.: Regularity of the powers of an ideal. Comm. Algebra **25**(12), 3773–3776 (1997)
10. Conca, A.: Regularity jumps for powers of ideals. Commutative Algebra with a Focus on Geometric and Homological Aspects. In: Lecture Notes in Pure Applied Mathematics, vol. 244. Chapman & Hall/CRC, Boca Raton (2006)
11. Cutkosky, D.: Irrational asymptotic behaviour of Castelnuovo-Mumford regularity. J. Reine Angew. Math. **522**, 93–103 (2000)
12. Cutkosky, D., Herzog, J., Trung, N.V.: Asymptotic behaviour of the Castelnuovo-Mumford regularity. Compositio Math. **118**(3), 243–261 (1999)
13. Cutkosky, D., Ein, L., Lazarsfeld, R.: Positivity and complexity of ideal sheaves. Math. Ann. **321**(2), 213–234 (2001)
14. Eisenbud, D., Harris, J.: Powers of ideals and fibers of morphisms. Math. Res. Lett. **17**(2), 267–273 (2010)
15. Eisenbud, D., Ulrich, B.: Notes on regularity stabilization. Proc. Amer. Math. Soc. **140**(4), 1221–1232 (2012)
16. Eisenbud, D., Huneke, C., Ulrich, B.: The regularity of Tor and graded Betti numbers. Amer. J. Math. **128**(3), 573–605 (2006)
17. Geramita, A.V., Gimigliano, A., Pitteloud, Y.: Graded Betti numbers of some embedded rational n-folds. Math. Ann. **301**(2), 363–380 (1995)
18. Hà, H.T.: Asymptotic linearity of regularity and a*-invariant of powers of ideals. Math. Res. Lett. **18**(01), 1–9 (2011)
19. Herzog, J., Simis, A., Vasconcelos, W.V.: Koszul homology and blowing-up rings. In: Commutative algebra (Trento, 1981). In: Lecture Notes in Pure and Applied Mathematics, vol. 84. pp. 79–169. Dekker, New York (1983)
20. Hoa, L.T., Hyry, E.: On local cohomology and Hilbert function of powers of ideals. Manuscripta Math. **112**(1), 77–92 (2003)
21. Johnson, M., Ulrich, B.: Artin–Nagata properties and Cohen-Macaulay associated graded rings. Compositio Math. **108**, 7–29 (1996)
22. Kodiyalam, V.: Homological invariants of powers of an ideal. Proc. Amer. Math. Soc. **118**(3), 757–764 (1993)
23. Kodiyalam, V.: Asymptotic behaviour of Castelnuovo-Mumford regularity. Proc. Amer. Math. Soc. **128**(2), 407–411 (2000)
24. McAdam, S.: Asymptotic prime divisors. In: Lecture Notes in Mathematics, vol. 1023. Springer, Berlin (1983)
25. McAdam, S., Eakin, P.: The asymptotic ass. J. Algebra **61**, 71–81 (1979)
26. Niu, W.: A vanishing theorem and asymptotic regularity of powers of ideal sheaves. arXiv:1009.0314
27. Niu, W.: Some results on asymptotic regularity of ideal sheaves. arXiv:1106.2585
28. Ratliff, L.J. Jr.: On prime divisors of I^n, n large. Michigan Math. J. **23**, 337–352 (1976)
29. Römer, T.: Homological properties of bigraded algebras. Illinois J. Math. **45**(4), 1361–1376 (2001)
30. Swanson, I.: Powers of ideals. Primary decompositions, Artin–Rees lemma and regularity. Math. Ann. **307**(2), 299–313 (1997)
31. Trung, T.N.: Regularity index of Hilbert functions of powers of ideals. Proc. Amer. Math. Soc. **137**(7), 2169–2174 (2009)
32. Trung, N.V., Wang, H.: On the asymptotic linearity of Castelnuovo–Mumford regularity. J. Pure Appl. Algebra **201**(1–3), 42–48 (2005)
33. Whieldon, G.: Stabilization of Betti Tables. arXiv:1106.2355

Some Homological Properties of Modules over a Complete Intersection, with Applications

Hailong Dao

1 Introduction

The purpose of this note is to survey some classical theory and recent developments in homological algebra over complete intersection rings. A tremendous amount of work has been done on this topic during the last fifty years or so, and it would not be possible for us to summarize even a sizable part of it. We shall focus on properties of modules (rather than following the modern and powerful trend of using complexes and derived categories) and emphasize the somewhat unexpected applications of such properties. While such narrow focus obviously reflects the author's own bias and ignorance, it will hopefully make the paper friendlier to researchers with less experience in this area. In addition, the connections to problems that are not homological in appearance reveal numerous interesting yet simple-looking open questions about modules over complete intersections that we shall try to highlight.

Throughout this note let $(R, \mathfrak{m}, k)$ be a commutative local Noetherian ring. We say that R is a *complete intersection* if there exists a regular local ring T and a regular sequence $f_1, \cdots, f_c$ in T such that $R \cong T/(f_1, \cdots, f_c)T$. We say that R is an *abstract complete intersection* if the $\mathfrak{m}$-adic completion $\hat{R}$ is a complete intersection. These two definitions are very recently proved to be different in a stunning preprint by Heitmann–Jorgensen [45]. However, for virtually all the local rings that arise in nature (i.e., those encountered by people outside of commutative algebra), these two notions coincide. Furthermore, in many situations of interest, one can reduce a statement about an abstract complete intersection to the same statement over a complete intersection via completion.

It is very well known that regular local rings have finite global dimension; thus every finitely generated module has a finite resolution by projective (even better,

---

H. Dao (✉)
Department of Mathematics, University of Kansas, Lawrence, KS 66045-7523, USA
e-mail: hdao@math.ku.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_10,

free) modules. Non-regular rings no longer posses this property. However, the free resolutions of modules over a complete intersection still enjoy remarkable finiteness properties, to be made precise later in this note; see Sect. 6. Such properties allow us to better control the behavior of the Ext and Tor functors. In particular, the vanishing of certain Ext or Tor modules often has much stronger consequences over complete intersections than other classes of singularities.

Why should we care about the vanishing of Ext or Tor then? Apart from intrinsic interests from homological algebra, one important reason is that such vanishing often comes up when one is trying to understand when the tensor product or Hom of two modules is nice (e.g., free). As an example let us discuss a famous result by Auslander–Buchsbaum, that is, every regular local ring R is a unique factorization domain (UFD). This classical result is quite fundamental in commutative algebra and algebraic geometry. An equivalent statement is that every reflexive R-module M of rank one is free. However, any such module satisfies the condition that $\mathrm{Hom}_R(M, M) \cong R$ since R is normal. When R is of dimension at least 3, this would imply that $\mathrm{Ext}^1_R(M, M) = 0$ (see Lemma 2.3.2). Thus, the fact that regular local rings are UFDs is a consequence of the following homological statement, which appeared in [46]:

Theorem 1.0.1 (Jothilingham). *Let R be a regular local ring and M be a finitely generated R-module. Then $\mathrm{Ext}^1_R(M, M) = 0$ if and only if M is free.*

Note that the above theorem works without any assumption on the rank or reflexivity of M; so it is quite a bit more general than the statement about unique factorization.

In this survey we shall discuss how such a statement and similar ones may still hold, with proper modifications, for non-regular rings. A natural class of rings for such results turns out to be complete intersections, due to the good homological properties they enjoy. Obviously, some new technical tools are needed, and we try to indicate the most essential ones.

Although most of the results in this note are known, a few appear for the first time, to the best of our knowledge. In addition, a number of old results have been given more simple or streamlined proofs, and some are strengthened considerably. We also discuss a rather large number of open questions.

Throughout the note we shall follow a rather informal approach, with intuition given priority over detailed proofs. When full proofs are given, it is usually because they are concerned with enhanced or modified versions of the original results (see for instance Lemma 2.3.2, Proposition 3.1.2, and Theorem 5.3.8). The payoff of such minor modifications can be quite satisfactory as they allow us to present sleek proofs or improved versions of several well-known results, such as the Danilov–Lipman theorem on discrete divisor class groups (7.1.1) or certain generalizations of the Grothendieck–Lefschetz theorem (7.2.5).

We now briefly describe the content of the paper. In Sect. 1 we recall the basic notations and some important preliminary results. Section 2 introduces the notion of Tor rigidity. This notion is crucial for many of the applications that follow, and we record some consequences for later use, the most important being Proposition 3.1.2 which is enhanced slightly from the original version.

Section 3 treats modules over regular local rings. Since in this situation, any finitely generated module is Tor rigid, the proofs can be simplified considerably. Thus this section serves as both a motivation and a guideline to what we try to do later with complete intersections.

Section 4 concerns with the hypersurface case. We give a quick proof of the fact that all resolutions are eventually periodic of period at most two and use that fact to define Hochster's pairing on finitely generated modules. This numerical function captures the Tor-rigidity property for modules over hypersurfaces, a fact we explain in Proposition 5.2.2. We discuss many results about vanishing of $\theta^R(-,-)$ as well as one of the most intriguing open questions, Conjecture 5.3.5.

Section 6 focuses on complete intersections. The main theme here is that the total modules of Ext (respectively Tor) have the Noetherian (respectively artinian) property over certain rings of operators. We discuss two useful exploitations of such properties: the theory of support varieties by Avramov–Buchweitz and a generalization of Hochster's pairing to higher codimensions.

The next two sections deal with applications. In Sect. 7 we explain how some classical results on class groups and Picard groups can be viewed as consequences of certain homological properties of modules. In particular we discuss a statement which can be viewed as a far-reaching generalization of the Grothendieck–Lefschetz theorem in certain cases (Theorem 7.2.5).

In Sect. 8, we turn our attention to several other applications: intersection of closed subschemes, splitting of vector bundles, and noncommutative crepant resolutions.

In Sect. 9 we collect and discuss the many open questions that arise from the recent developments. Most of them are motivated by results and problems outside of commutative algebra.

2 Notations and Preliminary Results

2.1 Generalities on Rings and Modules

Throughout this note $(R, \mathfrak{m}, k)$ will always be a commutative Noetherian local ring. Recall that a maximal Cohen–Macaulay R-module M is a finitely generated module satisfying depth $M = \dim R$ where depth M can be characterized as the infimum of the set of integers i such that $\mathrm{Ext}^i_R(k, M) \neq 0$. Let mod R and $\mathrm{MCM}(R)$ be the categories of finitely generated and finitely generated maximal Cohen–Macaulay R-modules, respectively. For an R-module M we let $\mathrm{pd}_R M$ and $\mathrm{id}_R M$ denote the projective and injective dimensions of M, respectively. Recall that R is said to be *regular* if $\mathrm{pd}_R k \leq \infty$.

The ring R is called *Cohen–Macaulay* if $R \in \mathrm{MCM}(R)$. It is called *Gorenstein* if $\mathrm{id}_R R < \infty$. We say that R is a *complete intersection* if there exists a regular local ring T and a regular sequence $f_1, \cdots, f_c$ in T such that $R \cong T/(f_1, \cdots, f_c)T$.

We say that R is an *abstract complete intersection* if the $\mathfrak{m}$-adic completion $\hat{R}$ is a complete intersection. The presentation of a ring as a complete intersection is typically not unique. However, if we assume that the elements $f_1, \cdots, f_c$ are in the square of the maximal ideal of T, then c is uniquely determined by R and it is called the *codimension* of R. A hypersurface is a complete intersection of codimension one.

The following of hierarchy of singularity type of R is very well known:

$$\text{regular} \Rightarrow \text{complete intersection} \Rightarrow \text{Gorenstein} \Rightarrow \text{Cohen–Macaulay}.$$

For $M \in \operatorname{mod} R$, let M^* denote the R-dual $\operatorname{Hom}_R(M, R)$. M is said to be *torsion-free* (respectively, *reflexive*) if the natural map $M \to M^{**}$ is injective (respectively, bijective). Let ΩM denote the first module of syzygy of M and $\Omega^n M$ the nth syzygy for each $n > 0$.

Given an element $x \in M$, the *order ideal* of x is by definition the ideal of R generated by $\{f(x) | f \in M^*\}$ and is denoted by $O_M(x)$. Clearly $\mathfrak{p} \in \operatorname{Spec} R$ does not contain $O_M(x)$ if and only if the $R_\mathfrak{p}$ submodule of $M_\mathfrak{p}$ generated by x splits off as a free summand. For other unexplained terminologies we refer to [50].

2.2 The Auslander–Bridger Transpose

We now recall an important construction known as the *Auslander–Bridger transpose*. Let $F_1 \to F_0 \to M \to 0$ be a minimal free resolution of M. The transpose of M, denoted by $\mathsf{Tr} M$, is defined as the cokernel of the dual map $F_0^* \to F_1^*$. It is not hard to show that the isomorphism class of $\mathsf{Tr} M$ does not depend on our choice of the minimal resolution. The crucial property here is the following exact sequence for each integer $n \geq 0$ (see [34, 44, 46]):

$$\operatorname{Tor}_2^R(\mathsf{Tr}\Omega^n M, N) \to \operatorname{Ext}_R^n(M, R) \otimes_R N \to \operatorname{Ext}_R^n(M, N) \to \operatorname{Tor}_1^R(\mathsf{Tr}\Omega^n M, N) \to 0.$$

2.3 Serre's Condition (S_n)

The Serre's conditions on a module are concerned with somewhat subtle conditions on depths over *all localizations* of the modules. While looking a bit cumbersome when one first encounters them, these conditions turn out to be rather natural for many results. One reason is that depth does not behave very well when one localizes, so in many cases these conditions can be considered as necessary nondegenerate conditions. Let us now recall the definition.

For a nonnegative integer n, M is said to satisfy (S_n) if

$$\operatorname{depth}_{R_\mathfrak{p}} M_\mathfrak{p} \geq \min\{n, \dim(R_\mathfrak{p})\} \ \forall \mathfrak{p} \in \operatorname{Spec}(R).$$

When R is Gorenstein, M is (S_1) if and only if it is torsion-free and (S_2) if and only if it is reflexive.

Remark 2.3.1. There are several definitions of Serre's condition for modules in the literature. For a detailed discussion, see [21]. These inconsistencies may lead to subtle problems; see, for example, [42].

The following technical but useful result gives an intimate connection between Serre's condition on the module of homomorphisms and vanishing of extension modules. We state it in a slightly more general version than the reference.

Lemma 2.3.2 (Dao [25, Lemma 2.3]). *Let R be a local ring, M, N be finitely generated R-modules, and $n > 1$ be an integer such that R satisfies Serre's condition* (S_{n+1}). *Consider the two conditions:*

(1) $\operatorname{Hom}(M, N)$ *satisfies Serre's condition* (S_{n+1}).
(2) $\operatorname{Ext}^i_R(M, N) = 0$ *for* $1 \le i \le n-1$.

If M is locally free in codimension n and N satisfies (S_n)*, then (1) implies (2). If N satisfies* (S_{n+1})*, then (2) implies (1).*

Proof. The first claim is obvious if $\dim R \le n$, as then M is free by assumptions. By localizing at the primes on the punctured spectrum of R and using induction on dimension, we can assume that all the modules $\operatorname{Ext}^i_R(M, N)$, $1 \le i \le n-1$ have finite length. Take a free resolution $P_\bullet$ of M, and look at the first n terms of $\operatorname{Hom}(P_\bullet, N)$. As all the cohomology of this complex are Ext modules, the claim now follows from the Acyclicity Lemma (see [13], Exercise 1.4.23).

For the second claim, one again takes a free resolution of P of M and looks at $\operatorname{Hom}(P_\bullet, N)$. The vanishing of the Ext modules gives the long exact sequence:

$$0 \to \operatorname{Hom}_R(M, N) \to N^{b_0} \to \cdots \to N^{b_{n-1}} \to B \to 0.$$

Counting depth shows that $\operatorname{depth} \operatorname{Hom}_R(M, N)_\mathfrak{p} \ge \min\{n+1, \operatorname{depth}(N_\mathfrak{p})\}$ for any $\mathfrak{p} \in \operatorname{Spec}(R)$, which is what we want. □

We record here a slightly more general version of Lemma 2.3.2 for the cases when N is allowed to be a torsion module. This will be used later in the proof of the Danilov–Lipman Theorem 7.1.2. The proof requires very little modification, and we shall omit it.

Lemma 2.3.3. *Let R be a local ring, M, N be finitely generated R-modules, and $n > 1$ be an integer. Consider the two conditions:*

(1) $\operatorname{depth} \operatorname{Hom}(M, N) \ge n+1$
(2) $\operatorname{Ext}^i_R(M, N) = 0$ *for* $1 \le i \le n-1$

Assume that $M_\mathfrak{p}$ is a free $R_\mathfrak{p}$-module for any $\mathfrak{p} \in \operatorname{Spec} R$ *such that* $\operatorname{depth} N_\mathfrak{p} \le n-1$. *Then (1) implies (2). If* $\operatorname{depth} N \ge n+1$*, then (2) implies (1).*

2.4 Partial Euler Characteristics and Intersection Multiplicities

Let T be a regular local ring.

$$\chi_i(M,N) = \sum_{j \geq i} (-1)^{j-i} \ell(\mathrm{Tor}_j^R(M,N)).$$

When $i = 0$ we simply write $\chi^R(M,N)$ or $\chi(M,N)$. The χ_i^R for $i > 0$ are called the *partial Euler characteristic*, and χ^R is known as the *Serre's intersection multiplicity*. Serre [58] introduced $\chi^R(M,N)$ as a homological definition of intersection multiplicity for modules over a regular local ring and showed that it satisfied many of the properties in the sense of intersection theory:

Theorem 2.4.1 (Serre). *Let T be a regular local ring that is equicharacteristic or unramified. Then for any pair of T-modules M, N such that $\ell(M \otimes_T N) < \infty$, we have:*

(1) $\dim(M) + \dim(N) \leq \dim(T)$.
(2) (Vanishing) If $\dim(M) + \dim(N) < \dim(T)$*, then* $\chi^T(M,N) = 0$.
(3) (Nonnegativity) It is always true that $\chi^T(M,N) \geq 0$.
(4) (Positivity) If $\dim(M) + \dim(N) = \dim(T)$*, then* $\chi^T(M,N) > 0$.

In fact, (1), (2), and (3) are known for all regular local rings due to the work of Serre, Roberts, Gillet–Soule, and Gabber. See [57] for a thorough discussion of these results and related open questions.

Concerning the partial Euler characteristics, we have the following important result which tells us exactly when they vanish [37, 48].

Theorem 2.4.2 (Lichtenbaum, Hochster). *Consider finitely generated modules M, N over a regular local ring T that is equicharacteristic or unramified and an integer i such that $\ell(\mathrm{Tor}_j^T(M,N)) < \infty$ for $j \geq i$. Then $\chi_i^T(M,N) \geq 0$ and it is 0 if and only if $\mathrm{Tor}_j^T(M,N) = 0$, for all $j \geq i$.*

Conjecturally this is true for all regular local rings or if the modules are of finite projective dimension. See [28] for some recent developments.

2.5 Change of Rings Sequences

In studying modules over complete intersection, it is crucial to track their behavior as the codimension increases. In other words, we need to know how the behavior of modules over $T/(f)$ and T differs for a regular element f in T. The following change-of-rings exact sequence is important for many applications.

Let T be a regular local ring. Let $R = T/f$ where f is a nonzero divisor on T, and let M, N be R-modules. Then we have the long exact sequence of Tors (see, e.g., [39]):

$$\begin{aligned}&\ldots \to \operatorname{Tor}_n^R(M,N) \to \operatorname{Tor}_{n+1}^T(M,N) \to \operatorname{Tor}_{n+1}^R(M,N)\\ &\to \operatorname{Tor}_{n-1}^R(M,N) \to \operatorname{Tor}_n^T(M,N) \to \operatorname{Tor}_n^R(M,N)\\ &\to \ldots\\ &\to \operatorname{Tor}_0^R(M,N) \to \operatorname{Tor}_1^T(M,N) \to \operatorname{Tor}_1^R(M,N) \to 0.\end{aligned}$$

There is a similar sequence for Ext modules:

$$\begin{aligned}0 \to \operatorname{Ext}_R^1(M,N) \to \operatorname{Ext}_T^1(M,N) \to \operatorname{Ext}_R^0(M,N) \to\\ \cdots \to \operatorname{Ext}_R^n(M,N) \to \operatorname{Ext}_T^n(M,N) \to \operatorname{Ext}_R^{n-1}(M,N)\\ \to \operatorname{Ext}_R^{n+1}(M,N) \to \operatorname{Ext}_T^{n+1}(M,N) \to \operatorname{Ext}_R^n(M,N) \to \ldots\end{aligned}$$

2.6 Chow and Grothendieck Groups

Suppose X is a Noetherian scheme. Let $\mathfrak{Coh}(X)$ denote the category of coherent sheaves on X, and $\mathfrak{Vect}(X)$ the subcategory of vector bundles on X. By $G(X), \operatorname{Pic}(X), \operatorname{CH}^i(X), \operatorname{Cl}(X)$ we shall denote the Grothendieck group of coherent sheaves on X, the Picard group of invertible sheaves on X, the Chow group of codimension i irreducible, the closed subschemes of X, and the class group of X, respectively. When $X = \operatorname{Spec} R$ we shall write $G(R), \operatorname{Pic}(R), \operatorname{CH}^i(R), \operatorname{Cl}(R)$. Let $\overline{G}(R) := G(R)/\mathbb{Z}[R]$ be the reduced Grothendieck group and $\overline{G}(R)_{\mathbb{Q}} := \overline{G}(R) \otimes_{\mathbb{Z}} \mathbb{Q}$ be the reduced Grothendieck group of R with rational coefficients. Let $\operatorname{Spec}^o R$ denote the punctured spectrum of R.

In the rest of this section we assume that R is a local ring such that depth $R \geq 2$. For $i = 0, 1$, there are maps $c_i : G(R) \to \operatorname{CH}^i(R)$. These maps admit a very elementary definition which we now recall. Suppose M is an R-module, and choose any prime filtration $\mathcal{F}$ of M. Then one can take $c_i([M]) = \sum [R/p]$, where p runs over all prime ideals such that R/p appears in $\mathcal{F}$ and height$(p) = i$; note that a prime can occur multiple times in the sum (for a proof that this is welldefined, see the main Theorem of [19]). When R is a normal algebra essentially of finite type over a field and N is locally free (i.e., a vector bundle) on $\operatorname{Spec}^o R$, c_1 agrees with the first Chern class of N, as defined in [31, Chap. 3], but we shall not need that fact.

One has the following diagram of maps of abelian groups:

$$\begin{array}{ccc} & & \operatorname{Pic}(\operatorname{Spec}^o R)\\ & & \big\downarrow p\\ G(R) & \xrightarrow{\;c_1\;} & \operatorname{CH}^1(R).\end{array}$$

Here p is induced by the well-known map between Cartier and Weil divisors (see Chap. 2 of [31]).

Note that we do not indicate any map between $\mathrm{Pic}(\mathrm{Spec}^o\ R)$ and $G(R)$. However, the diagram "commutes" in a weak sense: if $\mathcal{E}$ represents an element in $\mathrm{Pic}(X)$ and $I = \Gamma_X(\mathcal{E})$ then $p([\mathcal{E}]) = c_1([I])$ in $\mathrm{CH}^1(R)$.

Obviously, $c_1([R]) = 0$, so c_1 induces a map $q : \overline{G}(R) \to \mathrm{CH}^1(R)$. In particular, if M is a module such that $[M] = 0$ in $\overline{G}(R)_{\mathbb{Q}}$ then $c_1([M])$ is torsion in $\mathrm{CH}^1(R)$.

3 Tor Rigidity and Some Consequences

3.1 Tor Rigidity

In this section we discuss the notion of Tor rigidity, a condition which plays an essential role in many of the following results. First, the definition.

Definition 3.1.1. A pair of finitely generated R-modules (M, N) is called Tor rigid if for any integer $i \geq 0$, $\mathrm{Tor}_i^R(M, N) = 0$ implies $\mathrm{Tor}_j^R(M, N) = 0$ for all $j \geq i$. Moreover, M is called Tor rigid if for all $N \in \mathrm{mod}\ R$, the pair (M, N) is Tor rigid.

It was Auslander who first recognized the powerful consequences of these properties. In fact, he made Tor rigidity a central point of his 1962 ICM address on modules over (unramified) regular local rings [2]. In Auslander's and consequent work, Tor rigidity has been typically considered for a module and not a pair. However, for many recent applications the rigidity of pairs has proved to be not only more flexible, but necessary; see, for example, the proof of Theorem 7.2.5.

In the commutative algebra literature, Tor rigidity is often just called rigidity. However, rigidity has a different meaning related to deformation theory. Namely, a module is called *rigid* if $\mathrm{Ext}_R^1(M, M) = 0$. The name comes from the fact that such modules have no first-order deformations. To avoid confusion we shall try not to use that terminology and just write $\mathrm{Ext}_R^1(M, M) = 0$ when needed. In fact, it is rather amusing that the two notions of rigidity have the following connection.

Proposition 3.1.2 (Jothilingham). *Let R be a local ring and M, N be finitely generated R-modules such that the pair $(\mathrm{Tr}\Omega M, N)$ is* Tor *rigid (the first module is the transpose of the first syzygy of M; see Sect. 2.2). If* $\mathrm{Ext}_R^1(M, N) = 0$ *then* $M^* \otimes_R N \cong \mathrm{Hom}_R(M, N)$ *via the canonical map, and* $\mathrm{Tor}_i^R(M^*, N) = 0$ *for all* $i > 0$. *In particular, if the pair $(\mathrm{Tr}\Omega M, M)$ is* Tor *rigid then* $\mathrm{Ext}_R^1(M, M) = 0$ *if and only if M is free.*

The above result was first proved in [46], when it was assumed that R is regular, so Torrigidity is guaranteed; see the main theorem and the discussion of the last proposition. For a more modern presentation, see [44].

Proof. The key point of the proof is the short exact sequence in Sect. 2.2:

$$\mathrm{Tor}_2^R(\mathrm{Tr}\Omega^n M, N) \to \mathrm{Ext}_R^n(M, R) \otimes_R N \to \mathrm{Ext}_R^n(M, N) \to \mathrm{Tor}_1^R(\mathrm{Tr}\Omega^n M, N) \to 0.$$

Suppose $\mathrm{Ext}_R^1(M, N) = 0$. Then the above exact sequence shows that $\mathrm{Tor}_1^R(\mathrm{Tr}\Omega M, N) = 0$. Because of Tor rigidity, we have $\mathrm{Tor}_i^R(\mathrm{Tr}\Omega M, N) = 0$ for all $i > 0$. In particular, the exact sequence forces $\mathrm{Ext}_R^1(M, R) = 0$. Thus the R-dual of a minimal resolution of M is exact at the F_1^* spot, so $\mathrm{Tr}M \cong \Omega\mathrm{Tr}\Omega M$. Thus, $\mathrm{Tor}_i^R(\mathrm{Tr}M, N) = 0$ for all $i > 0$, and applying the above exact sequence with $n = 0$ yields an isomorphism $M^* \otimes_R N \cong \mathrm{Hom}_R(M, N)$. Of course, we also know that $M^* \cong \Omega^2\mathrm{Tr}M$ and the vanishing of $\mathrm{Tor}_i^R(M^*, N)$ follows. For the last statement, it is well known that the canonical map $M^* \otimes_R M \to \mathrm{Hom}_R(M, M)$ is an isomorphism if and only if M is free. □

The Tor-rigidity property is sometimes used in tandem with the so-called "depth formula" which we recall next.

Theorem 3.1.3. *Let R be a local complete intersection and $M, N \in \mathrm{mod}\ R$. If $\mathrm{Tor}_i^R(M, N) = 0$, for all $i > 0$ then*

$$\mathrm{depth}\ M + \mathrm{depth}\ N = \mathrm{depth}\ R + \mathrm{depth}\ M \otimes_R N$$

The above result was first proved by Auslander for regular local rings in [1, 1.2]. It was given in this form by Huneke and Wiegand in [39, 2.5]. There have been generalizations of this formula to much broader contexts, see for example [20].

Proof. We first prove it for the case when R is regular. Since $\mathrm{Tor}_i^R(M, N)$ vanishes for all i, the tensor product of the minimal resolutions of M and N becomes a minimal resolution of $M \otimes_R N$. Thus we have

$$\mathrm{pd}_R M \otimes_R N = \mathrm{pd}_R M + \mathrm{pd}_R N.$$

The depth formula follows from Auslander–Buchsbaum formula. The general case is handled by induction on the codimension of R. It suffices to assume that $R = T/(f)$ and that the depth formula holds for modules over T. One can use the change-of-rings exact sequence of Tor in Sect. 2.5 to deduce that $\mathrm{Tor}_i^T(M, N) = 0$ for all $i > 1$. Let $E = \Omega^T(M)$; then by induction hypotheses we have

$$\mathrm{depth}\ E + \mathrm{depth}\ N = \mathrm{depth}\ T + \mathrm{depth}\ E \otimes_T N.$$

However, since $\mathrm{depth}\ E = \mathrm{depth}\ M + 1$,

$$\mathrm{depth}\ M + \mathrm{depth}\ N = \mathrm{depth}\ R + \mathrm{depth}\ E \otimes_T N.$$

It just remains to prove that $\mathrm{depth}\ E \otimes_T N = \mathrm{depth}\ M \otimes_R N$. We leave this as an exercise. For details, see [39, Theorem 2.5]. □

It would not be a complete discussion of Tor rigidity without mentioning the following famous conjecture, usually attributed to Auslander:

Conjecture 3.1.4. *Let R be a Noetherian local ring and $M \in \operatorname{mod} R$ such that $\operatorname{pd}_R M < \infty$. Then M is* Tor *rigid.*

This conjecture was disproved by an ingenious construction by Heitmann [35]. The module he constructed has projective dimension two. However, the following modified version is still open:

Question 3.1.5. Let R be a Noetherian local ring and $M, N \in \operatorname{mod} R$ such that $\operatorname{pd}_R M, \operatorname{pd}_R N < \infty$. Is the pair (M, N) Tor rigid?

4 Warm-Up: Modules over Regular Local Rings

In this section we quickly review Tor rigidity over regular local rings and its various consequences. We hope they will provide some insights and motivations for the following sections.

4.1 Tor Rigidity and the UFD Property

We begin with a famous result.

Theorem 4.1.1 (Auslander–Lichtenbaum). *Let R be a regular local ring. Any finitely generated module is* Tor *rigid.*

This is a deep theorem and we shall only sketch a proof in the equicharacteristic case (i.e., when R contains a field). There is no harm in assuming R is complete, and now the Cohen structure theorem tells us that R is isomorphic to the power series ring $k[[x_1, \cdots, x_d]]$. We now use an ingenious trick devised by Serre [58]. First we form the so-called completed tensor product $R \hat{\otimes}_k R$, which will just be the power series rings in $2d$ variables $S = k[[x_1, \cdots, x_d, y_1, \cdots, y_d]]$. Letting $L = M \hat{\otimes}_k N$, one has isomorphisms:

$$\operatorname{Tor}_i^R(M, N) \cong \operatorname{Tor}_i^S(L, S/(x_1 - y_1, \cdots, x_d - y_d)S).$$

Since the elements $x_i - y_i$, $1 \leq i \leq d$ form a regular sequence on S, the rigidity property of $\operatorname{Tor}_i^R(M, N)$ follows from the well-known rigidity of Koszul homologies.

When R is of mixed characteristic, the proof is much harder and uses in an essential way the non-negativity of the partial Euler characteristic functions. We refer to [1, 48] for details.

Corollary 4.1.2. *Let R be a regular local ring and M be a finitely generated R-module. Then* $\operatorname{Ext}^1_R(M, M) = 0$ *if and only if M is free.*

Proof. This is a direct consequence of Proposition 3.1.2 and Theorem 4.1.1. □

Corollary 4.1.3. *A regular local ring R is a UFD.*

Proof. Recalled that elements of the class group of R can be identified with isomorphism classes of reflexive ideals of rank one satisfying the following rule:

$$[\operatorname{Hom}_R(I, J)] = [J] - [I]$$

where I, J are reflexive ideals and $[I]$ denotes the class of I in $\operatorname{Cl}(R)$. In particular, it implies that $\operatorname{Hom}_R(I, I) \cong R$. We proceed by induction on $d = \dim R$. If $d \leq 2$ then reflexive modules are automatically maximal Cohen–Macaulay, thus are free since they also have finite projective dimension. Suppose $d > 2$. By localizing at prime ideals in $\operatorname{Spec}^o(R)$ we can assume by induction that I is locally free on $\operatorname{Spec}^o(R)$. We now invoke Lemma 2.3.2 to show that $\operatorname{Ext}^1_R(I, I) = 0$; thus I is free which is what we want to prove. □

We note that the proof of the last corollary gives a much more general result in [40]:

Corollary 4.1.4 (Huneke–Wiegand). *Let R be a regular local ring and M be a finitely generated reflexive R-module. Then* $\operatorname{Hom}_R(M, M)$ *is satisfying Serre's condition* (S_3) *if and only if M is free.*

The assumption on $\operatorname{Hom}_R(M, M)$ appears to be quite weak, since the module of homomorphisms is always reflexive, which is equivalent to (S_2) over Gorenstein rings.

4.2 Torsion-Freeness of Tensor Products and a Flatness Criterion

Next we discuss another direction, namely, that tensor products can rarely have high-depth over a regular local ring. The key point is that such high depth forces certain Tor modules to vanish, and then one can apply Tor rigidity and the depth formula (Theorem 3.1.3) to derive strong properties of the modules involved.

Proposition 4.2.1 (Auslander). *Let R be a regular local ring and $M, N \in \operatorname{mod} R$. If $M \otimes_R N$ is torsion-free then* $\operatorname{Tor}^R_i(M, N) = 0$ *for $i > 0$, and*

$$\operatorname{depth} M + \operatorname{depth} N = \operatorname{depth} R + \operatorname{depth} M \otimes_R N.$$

Proof. Let $t(M)$ denote the torsion submodule of M. By tensoring the exact sequence $0 \to t(M) \to M \to M/t(M) = M' \to 0$ with N, we get that $M' \otimes_R N$ is torsion-free. Suppose we have already proven the assertions for M', N. Then it follows that $\mathrm{Tor}_1^R(M', N) = 0$ and thus $t(M) \otimes_R N = 0$ which forces $t(M) = 0$. Thus we can assume that M is torsion-free to begin with. As such we can embed M into a free module F to get

$$0 \to M \to F \to M_1 \to 0.$$

Tensoring with N shows that $\mathrm{Tor}_1^R(M_1, N)$ embeds into $M \otimes_R N$. Since the former is torsion and the latter is torsion-free, we get that the former is 0. Tor rigidity and the depth formula (Theorem 3.1.3) now complete the proof. □

The above seemingly technical result implies an interesting criterion for flatness over R.

Corollary 4.2.2 (Auslander). *Let R be a regular local ring and $M \in \mathrm{mod}\, R$. If $M^{\otimes_n}$ is torsion-free for some $n \geq \dim R$, then M is free.*

Proof. Let $M_i = M^{\otimes_i}$. Proposition 4.2.1 shows that $\mathrm{depth}\, M + \mathrm{depth}\, M_{n-1} = \mathrm{depth}\, R + \mathrm{depth}\, M_n$. Furthermore, the formula holds true when we localize at any prime in $\mathrm{Spec}(R)$ since vanishing of Tor still holds. Since torsion-free is equivalent to (S_1), it follows that M_{n-1} is (S_1). One can use induction to show that M_{n-i} is (S_1) and $\mathrm{depth}\, M + \mathrm{depth}\, M_{n-i} = \mathrm{depth}\, R + \mathrm{depth}\, M_{n-i+1}$ for $1 \leq i \leq \dim R$. Set $i = d = \dim R$ we get that

$$\mathrm{depth}\, M_n = \mathrm{depth}\, M_{n-d} - d(\mathrm{depth}\, R - \mathrm{depth}\, M).$$

If $\mathrm{depth}\, R - \mathrm{depth}\, M > 0$ then the left-hand side will be nonpositive, a contradiction. Thus, $\mathrm{depth}\, M = \mathrm{depth}\, R$ and M is free. □

Recently, Avramov and Iyengar have generalized this flatness criterion to modules essentially of finite type over smooth k-algebras for any field k; see [9].

4.3 An Equivalent Condition for Vanishing of Tor and Intersection Multiplicities

We end this section with an interesting result noticed by Auslander. It says that, roughly speaking, the vanishing of all $\mathrm{Tor}_i^R(M, N)$, $i > 0$ over a regular local ring only depends on the local depths of the arguments M, N. We state it in a slightly different form from Auslander's original formulation ([2, Theorem 2]) to emphasize this point.

Theorem 4.3.1 (Auslander). *Let R be a regular local ring and $M, N \in \mathrm{mod}\, R$. The following are equivalent:*

1. $\mathrm{Tor}_i^R(M, N) = 0$ *for* $i > 0$.
2. $\mathrm{depth}\, M_\mathfrak{p} + \mathrm{depth}\, N_\mathfrak{p} \geq \mathrm{depth}\, R_\mathfrak{p}$ *for all* $\mathfrak{p} \in \mathrm{Spec}\, R$ *(the depth of the zero module is ∞ by convention).*

Proof. Assume (1). Then the depth formula (Theorem 3.1.3) gives us (2). Now assume (2) and also assume that (1) does not hold. Let s be the largest integer such that $\mathrm{Tor}_i^R(M,N) \neq 0$; it is finite since R is regular. Since we assume (1) does not hold, s is positive. Let $\mathfrak{p}$ be a minimal prime of $\mathrm{Tor}_s^R(M,N)$. Replacing R by $R_{\mathfrak{p}}$ we may assume $\mathrm{Tor}_s^R(M,N)$ has finite length. We can also assume that depth $R \geq 1$, as otherwise there is nothing to prove.

Let $F_\bullet, G_\bullet$ be minimal free resolutions of M and N, respectively. Consider the complex $H_\bullet = F_\bullet \otimes_R G_\bullet$. Let B, C be the modules of boundaries and cycles at the sth spot of $H_\bullet$, respectively. We have a short exact sequence:

$$0 \to B \to C \to \mathrm{Tor}_s^R(M,N) \to 0$$

which shows that depth $B = 1$. The complex $H_{i>s}$ is acyclic by our choice of s and is a minimal resolution of B. Thus the Auslander–Buchsbaum formula tells us that $s +$ depth R equals to the length of $H_\bullet$ which is $\mathrm{pd}_R M + \mathrm{pd}_R N$. It follows that

$$\operatorname{depth} M + \operatorname{depth} N = \operatorname{depth} R - s < \operatorname{depth} R,$$

a contradiction. □

The above result is quite relevant to intersection theory as follows. For $M, N \in \operatorname{mod} R$ such that $\ell(M \otimes_R N) < \infty$, recall the definition of intersection multiplicity:

$$\chi^R(M,N) = \sum_{i \geq 0} (-1)^i \ell(\mathrm{Tor}_j^R(M,N)).$$

When $M = R/I, N = R/J$ defining closed subschemes of Spec R, it is of considerable interest to know when $\chi^R(R/I, R/J)$ is just equal to $\ell(R/I \otimes_R R/J)$ (the "naive" definition of intersection multiplicity). The following result, a nice consequence of Theorem 4.3.1, tells us exactly when it happens. It can be viewed as a vast generalization of the Bezout theorem for curves.

Corollary 4.3.2 (Serre [58]). *Let R be a regular local ring and $M, N \in \operatorname{mod} R$ such that $M \otimes_R N$ has finite length and* $\dim M + \dim N = \dim R$. *Consider the following:*

1. *M, N are Cohen–Macaulay modules.*
2. $\mathrm{Tor}_i^R(M,N) = 0$ *for* $i > 0$.
3. $\chi^R(M,N) = \ell(M \otimes_R N)$.

We have (1) $\Leftrightarrow$ (2) $\Rightarrow$ (3). *If R is unramified or equicharacteristic then they are all equivalent.*

Proof. Assume (1). As the intersection of Supp$M$ and Supp$N$ consists of the maximal ideal, we only need to check the depth condition there to prove (2). But since M, N are Cohen–Macaulay, we have

$$\operatorname{depth} M + \operatorname{depth} N = \dim M + \dim N = \dim R = \operatorname{depth} R,$$

so (2) follows by Theorem 4.3.1. Now assume (2). Then by Theorem 4.3.1 again we have depth M + depth $N \geq$ depth $R = \dim R = \dim M + \dim N$. Since depth is bounded above by dimension, one must have depth $M = \dim M$ and depth $N = \dim N$, so M, N are Cohen–Macaulay.

That (3) follows from (2) is obvious. For the last assertion, assuming (3) forces

$$\chi_1^R(M, N) = \sum_{i \geq 1} (-1)^{i-1} \ell(\mathrm{Tor}_i^R(M, N) = 0.$$

Since we know (2.4.2) that $\chi_1^R = 0$ if and only if all the Tor modules vanish when R is unramified or equicharacteristic, the conclusion follows. □

5 The Hypersurface Case and Hochster's Theta Pairing

In this section we focus on the hypersurface case; in other words, R is a complete intersection of codimension one. Through this section we will assume $R = T/fT$ where T is a regular local ring. Set $d = \dim R$.

5.1 Periodic Resolutions

The most important fact about homological algebra over R from our point of view is the following result, which was noticed by quite a few people but perhaps first appeared explicitly in [29]:

Lemma 5.1.1. *Let R be a local hypersurface and $M \in \mathrm{MCM}(R)$. Then the minimal resolution of M is periodic of period (at most) two. In other words, $M \cong \Omega^2 M$.*

Proof. Suppose $R = T/fT$ where T is regular. The Auslander–Buchsbaum formula tells us that $\mathrm{pd}_T M = \operatorname{depth} T - \operatorname{depth} M = 1$. Thus over T, M has the following minimal resolution:

$$0 \to T^n \to T^n \to M \to 0$$

(the free modules have same rank since M is torsion as a T-module). Tensoring with $R = T/fT$ gives an exact sequence:

$$0 \to \mathrm{Tor}_1^T(M, R) \to R^n \to R^n \to M \to 0.$$

But one can easily compute $\mathrm{Tor}_1^T(M, R)$ using the resolution of R as T module:

$$0 \to T \to T \to R \to 0$$

and see that it is isomorphic to M. The map between the free R-modules is still minimal, so we are done. □

Remark 5.1.2. Observe that the map between the free T-modules in the resolution of M is represented by a $n \times n$ matrix A with entries in T. If B is the corresponding matrix for $\Omega_R M$, then it is not hard to see from basic homological algebra and the periodicity that $AB = BA = fI_n$. This is the origin of Eisenbud's theory of matrix factorization [29], a topic which has become very active recently due to connections to a number of areas in algebraic geometry and even mathematical physics.

The above result says that although hypersurfaces do not have finite global dimension, the resolutions of their modules are still finite in a rather specific way, as all the information about the resolution is captured after finitely many steps. Next we discuss a very effective way to exploit this eventual periodicity of minimal resolutions over hypersurfaces.

5.2 Hochster's Theta Pairing and Tor Rigidity

Let $M, N \in \operatorname{mod} R$ such that $\ell(\operatorname{Tor}_i^R(M,N)) < \infty$ for $i \geq d$; here $\ell(-)$ denotes length. The pairing $\theta^R(M,N)$, introduced by Hochster [36], is defined as follows:

$$\theta^R(M,N) := \ell(\operatorname{Tor}_{2e+2}^R(M,N)) - \ell(\operatorname{Tor}_{2e+1}^R(M,N))$$

where e is any integer such that $2e \geq \dim R$. By Lemma 5.1.1 this function is well defined. The theta pairing satisfies the following properties:

Proposition 5.2.1 (Hochster [36]).

(1) If $M \otimes_R N$ has finite length, then

$$\theta^R(M,N) = \chi^T(M,N) := \sum_{i \geq 0} (-1)^i \ell(\operatorname{Tor}_i^T(M,N)).$$

Here χ^T is the Serre's intersection multiplicity. In particular, if $\dim M + \dim N \leq \dim R = \dim T - 1$, *then* $\theta^R(M,N) = 0$ *(note that vanishing for χ^T is proved for all regular local rings; see [57, 13.1])*

(2) $\theta^R(M,N)$ is bi-additive on short exact sequences, assuming it is defined on all pairs. In particular, if M is locally of finite projective dimension on $\operatorname{Spec}^o(R)$, *then the rule* $[N] \mapsto \theta^R(M,N)$ *induces maps* $\overline{G}(R) \to \mathbb{Z}$ *and* $\overline{G}(R)_{\mathbb{Q}} \to \mathbb{Q}$.

A crucial property for this section is that vanishing of Hochster's pairing implies Tor rigidity for a big class of hypersurfaces including all the ones that contain a field.

Proposition 5.2.2 (Dao [22]). *Let $R = T/fT$ be a local hypersurface such that T is an unramified regular local ring. Let $M, N \in \operatorname{mod} R$ such that $\ell(\operatorname{Tor}_i^R(M,N)) < \infty$ for $i \geq d$. If $\theta^R(M,N) = 0$, then the pair (M,N) is* Tor *rigid.*

Proof. First, we prove the statement assuming that $\ell(\operatorname{Tor}_j^R(M,N)) < \infty$ for all $j > i$. We truncate the change-of-rings long exact sequence for Tor (Sect. 2.5) as follows (note that all $\operatorname{Tor}^T(M,N)$ vanish after $d+1$ spots):

$$\begin{aligned}
&0 \to \operatorname{Tor}_{2e+2}^R(M,N) \\
&\to \operatorname{Tor}_{2e}^R(M,N) \to \operatorname{Tor}_{2e+1}^T(M,N) \to \operatorname{Tor}_{2e+1}^R(M,N) \\
&\to \dots \\
&\to \operatorname{Tor}_i^R(M,N) \to \operatorname{Tor}_{i+1}^T(M,N) \to \operatorname{Tor}_{i+1}^R(M,N) \to C \to 0.
\end{aligned}$$

It is easy to see that all the modules in this sequence have finite lengths. Therefore we can take the alternating sum of the lengths and get

$$l(C) + \chi_{i+1}^T(M,N) = (-1)^{2e+2-i}\theta^R(M,N) + l(\operatorname{Tor}_i^R(M,N)) = 0.$$

This equation and Theorem 2.4.2 forces $C = 0$ and $\operatorname{Tor}_j^T(M,N) = 0$ for all $j \geq i+1$.

Now we prove the general case by induction on $d = \dim R$. If d=0 then all modules involved have finite lengths, so we are done. Assume $d > 0$. By localizing at primes in $\operatorname{Spec}^o R$ and applying the induction hypothesis, it is clear that $\ell(\operatorname{Tor}_j^R(M,N)) < \infty$ for all $j > i$; thus the proof is complete. □

Note that almost by definition, the vanishing of $\theta^R(M,N) = 0$ implies rigidity of Tor *asymptotically*. Namely, for $i > d$, we would have $\ell(\operatorname{Tor}_i^R(M,N)) = \ell(\operatorname{Tor}_{i+1}^R(M,N))$, so they vanish together. The content of the proposition which is crucial for many applications is that rigidity holds even if we only know that $\operatorname{Tor}_i^R(M,N) = 0$ for small i (in fact, $i = 1$ for most applications).

5.3 *Vanishing of $\theta^R(-,-)$*

A particular situation when $\theta^R(-,-)$ becomes very useful is when R has *isolated singularity*, namely, that $R_{\mathfrak{p}}$ is regular for any $\mathfrak{p} \in \operatorname{Spec}^o(R)$. In this case the higher Tor modules always have finite length, so the Hochster's formula gives a pairing:

$$\theta^R(M,N) : \overline{G}(R) \times \overline{G}(R) \to \mathbb{Z}.$$

We now briefly discuss Conjecture 5.3.5, starting with small dimensions. We start with a lemma which exploits the vanishing condition of Serre's intersection multiplicity.

Lemma 5.3.1 (Huneke–Wiegand [39]). *Let R be a hypersurface with isolated singularity. Let $M, N \in \operatorname{mod} R$ such that $M \otimes_R N$ has finite length and* $\dim M + \dim N \leq \dim R$. *Then* $\theta^R(M, N) = 0$.

Proof. Let $R = T/fT$ with T a regular local ring. Then $M \otimes_T N$ has finite length, and $\dim M + \dim N < \dim T$. By the vanishing property of the Serre's intersection multiplicity, we know that $\chi^T(M, N) = 0$, so what we want to show follows from Proposition 5.2.1, part (1). □

The next result illustrates a often used technique to prove vanishing of $\theta^R(-,-)$: one can try to move the support of the modules involved (inside the Chow group of R) so that one gets to the situation of Lemma 5.3.1.

Theorem 5.3.2. *Assume $d > 1$, and M, N are finitely generated R-modules. Then $\theta^R(M, N) = 0$ if* $\dim M \leq 1$.

Proof. Since any module has a filtration by prime cyclic modules, we may assume that $M = R/P$ and $N = R/Q$ for some $P, Q \in \operatorname{Spec} R$. If $\dim R/P = 0$, so $P = m$, then $[R/P] = 0$ in $\overline{G}(R)_{\mathbb{Q}}$, and θ vanishes. Also, we may assume $Q \neq 0$. We now consider two cases:

Case 1: Q is not contained in P. Then $l(R/(P + Q)) < \infty$ because $\dim R/P = 1$, and since $\dim R/P + \dim R/Q \leq \dim R$, we have $\theta(R/P, R/Q) = 0$ by Proposition 5.2.1.

Case 2: So now we only need to consider the case $0 \neq Q \subset P$. We claim that there is cycle $\alpha = \sum l_i[R/Q_i] \in \mathrm{CH}^*(R)_{\mathbb{Q}}$ such that $\alpha = [R/Q]$ and $Q_i \not\subseteq P$. Consider the element $[R_P/Q] \in \mathrm{CH}^*(R_P)_{\mathbb{Q}}$. Since R_P is regular, $[R_P/Q] = 0$. Therefore, formally, we have a collection of primes q_i and elements f_i and integers n, n_i such that $n[R_P/Q] = \sum \operatorname{div}(R_P/q_i, f_i)$. Now in $\mathrm{CH}^*(R)_{\mathbb{Q}}$ we will have $\sum \operatorname{div}(R/q_i, f_i) = n[R/Q] + \sum n_i[R/Q_i]$, with $Q_i \not\subseteq P$, which proves our claim. The fact that $[R/Q] = \sum_i l_i[R/Q_i]$ in $\mathrm{CH}^*(R)_{\mathbb{Q}}$ means that in $G(R)_{\mathbb{Q}}$, $[R/Q] = \sum_i l_i[R/Q_i]$ + terms of lower dimension. By Case 1, $\theta^R(R/P, R/Q_i) = 0$, and we may conclude our proof by induction on $\dim R/Q$. □

Using the real "moving lemma" from intersection theory one can prove the following vast generalization of Lemma 5.3.1.

Theorem 5.3.3. *Suppose that R is an excellent local hypersurface containing a field, and has only isolated singularity. Then $\theta^R(R/P, R/Q) = 0$ if* $\dim R/P + \dim R/Q \leq \dim R$.

Proof. The strategy is clear, as indicated above. We need to move P and Q inside the Chow group of $\operatorname{Spec} R$ so that their support intersects only at the maximal ideal, and then apply Lemma 5.3.1. For details, see [22, Theorem 3.5]. □

Theorem 5.3.4. *Suppose that R is a local hypersurface with isolated singularity of even dimension. If* $\dim R \leq 2$ *or* $\dim R = 4$ *and R is excellent and contains a field, then $\theta^R(M, N) = 0$ for all pairs (M, N).*

Proof. When $d = 0$, $\overline{G}(R)$ is torsion; thus the statement is true by Proposition 5.2.1. It suffices to assume that M, N are cyclic prime modules; let's say $M = R/P, N = R/Q$. Then by the previous theorems, we only need to worry if both of them have dimensions at least 2. If $\dim R = 2$, they must both be R (note that since R is normal and local, it is a domain); thus θ^R certainly vanishes. If $\dim R = 4$ and R contains a field we can apply Theorem 5.3.3 and assume $\dim R/P + \dim R/Q \geq 5$. Then one of the primes, say P, is height 1 (if the minimal height is 0, the assertion is trivial). Thus we will be done if $\mathrm{CH}^1(R) = 0$. But by the Grothendieck–Lefschetz theorem, the Picard group of $X = \mathrm{Spec}(R) - \{m\}$ is 0. Since X is regular, the Picard group of X is the same as $\mathrm{CH}^1(X) = \mathrm{CH}^1(R)$. □

We now state an intriguing open question, first proposed in [22].

Conjecture 5.3.5. *Let $R = T/fT$ be a hypersurface of even dimension with isolated singularity. Then $\theta^R(M, N) = 0$ for any pair of modules $M, N \in \mathrm{mod}\, R$.*

There are several sources of supporting evidence for this conjecture, the small dimensional cases, a conjecture by Hartshorne that the Chow groups of a smooth hypersurface X vanish for codimension below half the dimension of X (see [22, Sect. 3] for a discussion). But perhaps the strongest supporting evidence is that it has already been proven in important cases, using several different approaches.

Theorem 5.3.6. *Conjecture 5.3.5 is true in the following cases:*

(1) (Moore–Piepmeyer–Spiroff–Walker [52]) $R = k[x_0, \cdots, x_d]_{\mathfrak{m}}/(f)$, where f is a homogenous polynomial defining a smooth hypersurface in $\mathbf{P}^d_k$ over a perfect field k and $\mathfrak{m} = (x_0, \cdots, x_d)$.
(2) (Polishchuk–Vaintrob [54], Buchweitz–Van Straten [15]) $R = k[[x_0, \cdots, x_d]]/(f)$, a hypersurface isolated singularity with k a field of characteristic 0.

We will comment briefly about the proofs, which are quite sophisticated and beyond the scope of this survey. The key point in all of them is to compare $\theta^R(-,-)$ with some maps from $G(R)$ to certain cohomology theories (étale cohomology, topological K-theory, Hochschild cohomology) and use what we already know about such theories. There is perhaps a lot more to be understood about these connections.

Before moving on we give a lemma which is frequently used to prove vanishing of $\theta^R(-,-)$.

Lemma 5.3.7. *Suppose R is a local Noetherian ring. A module of finite length is zero in $\overline{G}(R)_{\mathbb{Q}}$.*

Proof. Since any module of finite length is a multiple of $[k]$, it is enough to prove the claim for one such module. If $\dim R = 0$ then there is nothing to prove as $\overline{G}(R)_{\mathbb{Q}} = 0$. If $\dim R > 0$ then pick a prime $\mathfrak{p}$ such that $\dim R/\mathfrak{p} = 1$ and a non-unit $x \notin \mathfrak{p}$. The exact sequence $0 \to R/\mathfrak{p} \to R/\mathfrak{p} \to R/(\mathfrak{p}, x) \to 0$ shows that $[R/(\mathfrak{p}, x)] = 0$, which is all we need. □

Next we give a strong supporting evidence for Conjecture 5.3.5. This will also be used to prove certain generalizations of the Grothendieck–Lefschetz theorem in Sect. 7. Note that we do not assume isolated singularity in the following, which improves on [24, Theorem 4.1].

Theorem 5.3.8. *Let R be a hypersurface of even dimension. Let $M \in \operatorname{mod} R$ such that M is locally free on the punctured spectrum* $\operatorname{Spec}^o R$. *Then* $\theta^R(M, M^*) = 0$.

Proof. Assume M is locally free on $\operatorname{Spec}^o R$. Let $K = \Omega M$. We want to prove that $\theta^R(M, M^*) = \theta^R(K, K^*)$. Dualizing the short exact sequence $0 \to K \to F \to M \to 0$ we get an exact sequence:

$$0 \to M^* \to F^* \to K^* \to \operatorname{Ext}^1_R(M, R) \to 0.$$

So $[M^*] + [K^*] = [\operatorname{Ext}^1_R(M, R)]$ in $\overline{G}(R)$. Note that $\operatorname{Ext}^1_R(M, R)$ has finite length because M is locally free on the punctured spectrum of R. By Lemma 5.3.7 any module of finite length is equal to 0 in $\overline{G}(R)_{\mathbb{Q}}$. So $[M^*] = -[K^*]$ and $\theta^R(M^*, -) = -\theta^R(K^*, -)$. Since $\theta^R(M, -) = -\theta^R(K, -)$, we have $\theta^R(M, M^*) = \theta^R(K, K^*)$.

Repeating the equality above we get $\theta^R(M, M^*) = \theta^R(L, L^*)$ when $L = \Omega^n M$ for any $n > 0$. But for $n \gg 0$ L is an MCM R-module, so $\theta^R(L, L^*) = 0$ by Proposition 5.3.9 . ☐

The following result is essentially due to Buchweitz, although it was originally stated in the language of stable cohomology. For a detailed explanation, see [24] (note that the proof of Proposition 4.3 there works verbatim in our situation).

Proposition 5.3.9 (Buchweitz [14], 10.3.3). *Let R be a local hypersurface with* $d = \dim R$. *Then for any two MCM R-modules M, N such that M is locally free on* $\operatorname{Spec}^o R$, *we have*

$$\theta^R(M, N) = (-1)^{d-1}\theta^R(M^*, N^*).$$

6 Asymptotic Behavior of Ext and Tor Over Complete Intersections

In this section we move on to complete intersections. Here resolutions of modules are still finite in a rather strong sense, which we shall now discuss. Let us assume throughout this section that $R \cong T/(f_1, \cdots, f_c)T$ with the f_is a T-regular sequence in $\mathfrak{m}_T^2$. In this situation, the codimension of R, $\operatorname{codim}(R)$, is equal to c.

6.1 Cohomology Operators and Support Varieties

The key point now is that the presentation $R = T/(f_1, \cdots, f_c)T = T/(\underline{f})$ gives rise to cohomology operators under which the total module of extensions $\mathrm{Ext}^*_R(M,N) := \oplus_{i\geq 0}\mathrm{Ext}^i_R(M,N)$ becomes a *finitely generated* module over the ring of said operators [33].

We sketch the construction of these operators. Start with a minimal resolution of M over R:

$$\cdots \xrightarrow{\delta_{i+1}} F_i \xrightarrow{\delta_i} F_{i-1} \xrightarrow{\delta_{i-1}} \cdots \xrightarrow{\delta_1} F_0 \to M \to 0.$$

One can lift this resolution to T to get

$$\cdots \xrightarrow{\tilde{\delta}_{i+1}} \tilde{F}_i \xrightarrow{\tilde{\delta}_i} \tilde{F}_{i-1} \xrightarrow{\tilde{\delta}_{i-1}} \cdots \xrightarrow{\tilde{\delta}_1} \tilde{F}_0.$$

Of course, this liftings will not form a complex anymore. However, we know that $\tilde{\delta}_{i-1}\tilde{\delta}_i(\tilde{F}_i) \subseteq (\underline{f})\tilde{F}_{i-2}$. Now since $\underline{f}$ is a regular sequence in T, $(\underline{f})/(\underline{f})^2$ is a free $R = T/(\underline{f})$-module of rank c. Thus, we write:

$$\tilde{\delta}_{i-1}\tilde{\delta}_i = \sum_{j=1}^{c} f_j \tilde{\theta}_j$$

where the $\tilde{\theta}_j : \tilde{F}_i \to \tilde{F}_{i-2}$ are T-linear maps. Set $\theta_j = \tilde{\theta}_j \otimes_T R$, which will be a map from F_i to F_{i-2}. These induce R-linear maps:

$$\chi_j : \mathrm{Ext}^i_R(M,N) \to \mathrm{Ext}^{i+2}_R(M,N)$$

for $1 \leq j \leq c$ and $i \geq 0$. These operators can be shown to be well behaved (commuting up to homotopies, functorial, cf. [29, 33]). There are subtle issues with the commutativity with Yoneda products which are fully resolved in [7]. The upshot is that they turn $\mathrm{Ext}^*_R(M,N)$ into a module over the (commutative) *ring of cohomology operators* $\mathcal{S} := R[\chi_1, \ldots, \chi_c]$.

Note that the χ_i have cohomological degree 2. Furthermore, for every $M \in \mathrm{mod}\, R$, there is a graded ring homomorphism:

$$\mathcal{S} \xrightarrow{\varphi_M} \mathrm{Ext}^*_R(M,M).$$

We now recall the definition of support varieties for modules over local complete intersections $(R, \mathfrak{m}, k)$ of codimension c: for details, see [8] and [5].

Let M, N be finitely generated R-modules. The *support variety* of the pair (M,N) is defined as

$$V_R(M,N) := \mathrm{Supp}_{\mathcal{S}\otimes_R k}(\mathrm{Ext}^*_R(M,N) \otimes_R k) \subset \mathbf{P}^{c-1}_k.$$

The support variety of M is defined as $V_R(M) := V_R(M,M)$.

Recall that M has *complexity* s, written as $\mathrm{cx}_R(M) = s$, provided that s is the least nonnegative integer for which there exists a real number γ such that $b_n^R(M) \leq \gamma \cdot n^{s-1}$ for all $n \gg 0$ [5, 3.1]. One can define the complexity of a pair of modules $\mathrm{cx}_R(M, N)$ similarly by using the minimal number of generators of $\mathrm{Ext}_R^n(M, N)$ instead of $b_n^R(M)$. Note that $\mathrm{cx}_R(M, N) = 0$ if and only if $\mathrm{Ext}_R^i(M, N) = 0$ for $i \gg 0$.

Theorem 6.1.1 (Avramov–Buchweitz [8, Theorem I]).

1. $V_R(M, N) = V_R(M) \cap V_R(N) = V_R(N, M)$
2. $V_R(M) = V_R(M, k)$
3. $\mathrm{cx}_R(M, N) = \dim V_R(M, N) + 1$ *(the dimension of the empty set is* -1 *by convention)*

Corollary 6.1.2 (Avramov–Buchweitz [8, Theorem 4.2]). *Let* R *be a local complete intersection and* $M \in \mathrm{mod}\, R$. *If* $\mathrm{Ext}_R^{2i}(M, M) = 0$ *for some* i *then* $\mathrm{pd}_R M < \infty$.

Proof. The key point here is the graded ring homomorphism

$$\mathcal{S} \xrightarrow{\varphi_M} \mathrm{Ext}_R^*(M, M).$$

As each χ_i has degree 2 and $\mathrm{Ext}_R^{2i}(M, M) = 0$, it follows that $\mathrm{Ext}_R^*(M, M)$ is $(\chi_1, \ldots, \chi_c)$-torsion. This remains true after tensoring with k, so the support variety $V_R(M, M)$ is empty. By Theorem 6.1.1 it follows that $\mathrm{cx}_R M = 0$; that is M, has finite projective dimension. □

When $i = 1$ one has a much more specific result, whose proof we omit.

Theorem 6.1.3 (Auslander–Ding–Solberg [4]). *Let* T *be a complete local ring and* $R = T/(f_1, \cdots, f_c)T$ *be a quotient of* T *by a regular sequence. Let* $M \in \mathrm{mod}\, R$. *If* $\mathrm{Ext}_R^2(M, M) = 0$ *for some* i *then* M *lifts to* T*; namely there is a module* $N \in \mathrm{mod}(T)$ *such that* $M \cong N/(f_1, \cdots, f_c)N$ *and the* f_i*s form a regular sequence on* N.

6.2 Length of Tor and a Generalized Version of Hochster's Pairing

In this section we study a more general version of Hochster's theta pairing. This needs a bit of preparatory work. The key point is that over a complete intersection, if all the higher Tor of two modules have finite length, such lengths display a quasi-polynomial behavior. This allows us to define a numerical value using those lengths. We now describe the process in more detail.

First, the construction in Sect. 6.1 yields operators on the total Tor module:

$$\mathcal{T}_R(M, N) := \oplus_{i \geq 0} \mathrm{Tor}_i^R(M, N),$$

as well (these are known as Eisenbud operators as they were constructed explicitly in [29]). Of course, these operators, which we shall also call $\chi_1, \dots, \chi_c$, now have degree (-2), and we have the following dual version of the key point in Sect. 6.1:

Theorem 6.2.1 (Gulliksen [33]). *Suppose that $M \otimes_R N$ is artinian. Then $\mathcal{T}(M, N)$ is an artinian module over the ring of operators $R[\chi_1, \dots, \chi_c]$.*

This result has a serious restriction, namely, that $M \otimes_R N$ has finite length. As we have seen in the previous section, the most interesting results live naturally in the setting where we only know that $\operatorname{Tor}_i^R(M, N)$ have finite length for $i \gg 0$ (this is always the case if R has an isolated singularity). Thus we need to first establish the following, which should be compared to similar results in [6, 11]

Lemma 6.2.2. *Suppose there exists an integer j such that $\operatorname{Tor}_i^R(M, N)$ has finite length for every $i \geq j$. Then $\mathcal{T}_R(M, N)_{i \geq j}$ is an artinian module over the ring of Eisenbud operators $R[\chi_1, \dots, \chi_c]$.*

Proof. See [23, Lemma 3.2] or [53, Appendix 1]. □

The significance of this result is that we now know that the lengths of $\operatorname{Tor}_i^R(M, N)$ must have quasi-polynomial behavior. This fact allows us to make the following definition which generalizes Hochster's pairing to complete intersections.

Definition 6.2.3. Suppose R is a local complete intersection of codimension c. Let $M, N \in \operatorname{mod} R$ such that $\operatorname{Tor}_i^R(M, N)$ has finite length for every $i \geq j$. We define

$$\eta^R(M, N) := \lim_{n \to \infty} \frac{\sum_j^n (-1)^i \ell(\operatorname{Tor}_i^R(M, N))}{n^c}.$$

Remark 6.2.4. The definition does not depend on the value of j we start with since $c \geq 1$. There is another way to define $\eta^R(M, N)$ as follows. By virtue of Lemma 6.2.2 there exist polynomials $P_{odd}(x)$ and $P_{even}(x)$ of degrees at most $c-1$ such that for $i \gg 0$ there are equalities:

$$P_{odd}(i) = \ell(\operatorname{Tor}_i^R(M, N))$$

for i odd and

$$P_{even}(i) = \ell(\operatorname{Tor}_i^R(M, N))$$

for i even. Let a, b be the coefficients of x^{c-1} in P_{even} and P_{odd}, respectively. Then

$$a - b = 2^c \eta^R(M, N).$$

Thus, one can use $a - b$ as an alternative definition for $\eta^R(M, N)$. However, the original definition seems to make the proof of additivity on short exact sequences (see below) easier, and this is a crucial property that allows us to make η^R into a pairing on the Grothendieck group.

We obtain the following properties of $\eta^R(M, N)$.

Theorem 6.2.5 (Dao [23]). *Suppose $R \cong T/(f_1, \cdots, f_c)T$ with the f_is a T-regular sequence in $\mathfrak{m}_T^2$ (T regular local). Let $M, N \in \operatorname{mod} R$ such that $\operatorname{Tor}_i^R(M, N)$ has finite length for all $i \gg 0$. Then the following statements hold:*

(1) *$\eta^R(M, N)$ is finite and rational.*
(2) *(Biadditivity) Let $0 \to M_1 \to M_2 \to M_3 \to 0$ and $0 \to N_1 \to N_2 \to N_3 \to 0$ be exact sequences such that $\eta^R(M, N_i)$ and $\eta^R(M_i, N)$ can be defined for all i. Then,*

$$\eta^R(M_2, N) = \eta^R(M_1, N) + \eta^R(M_3, N)$$

and

$$\eta^R(M, N_2) = \eta^R(M, N_1) + \eta^R(M, N_3).$$

(3) *(Change of rings) Suppose that $c > 1$. Let $R' = T/(f_1, \ldots, f_{c-1})$. Note that we have $\ell(\operatorname{Tor}_i^{R'}(M, N)) < \infty$ for $i \gg 0$. Then,*

$$\eta^R(M, N) = \frac{1}{2c}\eta^{R'}(M, N).$$

The above properties afford us some mild generalizations of the connections between $\theta^R(M, N)$ and vanishing of Tor as well as intersection multiplicity.

Corollary 6.2.6. *Let R, M, N be as in Theorem 6.2.5. Then we have:*

(1) *Suppose that $\eta^R(M, N) = 0$ and T is equicharacteristic or unramified. Then if $\operatorname{Tor}_i^R(M, N) = 0$ for $j \leq i \leq j + c - 1$ with some $j > 0$ then $\operatorname{Tor}_i^R(M, N) = 0$ for all $i \geq j$.*
(2) *If $M \otimes_R N$ has finite length and $\dim M + \dim N < \dim R + c$ then $\eta^R(M, N) = 0$. The converse is true if T is equicharacteristic or unramified.*

Proof. These are direct consequences of Theorem 6.2.5 and Propositions 5.2.1 and 5.2.2 (Theorem 6.2.5 allows us to reduce everything to the hypersurface case). □

One can push these results a bit further with additional assumptions, see [16, 17, 23]. One can also define a similar version of $\eta^R(M, N)$ using Ext modules, see [18].

7 Applications: Class Groups and Picard Groups

In this section we discuss the connection between Tor-rigidity property and some well-known results and open problems regarding class groups and Picard groups of varieties. Even though such connections have sometimes been used explicitly in certain proofs, they seem to have escaped wide attention. We hope what follows will stimulate further research in this direction.

7.1 The Lipman–Danilov theorem

First, we recast a well-known proof by Lipman on class groups in terms of Tor rigidity. The following theorem is the key result in Lipman's paper [49, Theorem 1] (it was stated there slightly more generally in the non-local setting).

Theorem 7.1.1 (Lipman). *Let R be a local ring and t be a regular element in the maximal ideal of R. Let $M \in \operatorname{mod} R$ such that t is also M-regular. Let $S = R/tR$. Assume that $\operatorname{Hom}_R(M, S)$ is a free S-module. Also assume that $(M/tM)_{\mathfrak{p}}$ is a free $S_{\mathfrak{p}}$-module for any $\mathfrak{p} \in \operatorname{Spec} S$ such that $\operatorname{depth} S_{\mathfrak{p}} \leq 2$. Then $\operatorname{Hom}_R(M, R)$ is a free R-module.*

Proof. Consider any prime $\mathfrak{p} \in \operatorname{Spec} S$ such that $\operatorname{depth} S_{\mathfrak{p}} \leq 2$. Let $\mathfrak{q}$ be the preimage of $\mathfrak{p}$ in $\operatorname{Spec} R$. Then $M_{\mathfrak{q}}$ is a free $R_{\mathfrak{q}}$-module. We may assume that $\operatorname{depth} S \geq 3$; otherwise M is a free R-module automatically. Applying Lemma 2.3.3 with $S = N$ and $n = 2$, we obtain $\operatorname{Ext}^1_R(N, S) = 0$. But $\operatorname{pd}_R S = 1$; thus S is obviously Tor-rigid as an R-module. Proposition 3.1.2 implies that $\operatorname{Hom}_R(M, R) \otimes_R S \cong \operatorname{Hom}_R(M, S)$. So $\operatorname{Hom}_R(M, R) \otimes_R S$ is a free S-module, and Nakayama's Lemma forces $\operatorname{Hom}_R(M, R)$ to be free over R. □

As explained in [49], the above theorem allows us to prove a famous result by Danilov on discrete divisor class groups. Recall that a normal domain S is said to have *discrete divisor class group* (abbreviated DCG) if the map between class groups $\operatorname{Cl}(S) \to \operatorname{Cl}(S[[t]])$ is bijective.

Corollary 7.1.2. *Let S be a noetherian normal domain. If $S_{\mathfrak{p}}$ has DCG for all $\mathfrak{p} \in \operatorname{Spec} S$ such that $\operatorname{depth} S_{\mathfrak{p}} \leq 2$ then S has DCG.*

Let $R = S[[t]]$. The point is that the natural map $\phi : \operatorname{Cl}(S) \to \operatorname{Cl}(R)$ is obviously injective. There is another map $\psi : \operatorname{Cl}(R) \to \operatorname{Cl}(S)$ induced by the map from $\operatorname{mod} R$ to $\operatorname{mod} S$ given by $M \mapsto M/tM$. It is not hard to check that $\psi\phi = \mathrm{id}$. But the injectivity of ψ is guaranteed by Theorem 7.1.1 (or rather the nonlocal version of it as in [49, Theorem 1]).

Remark 7.1.3. We note that the proof above allows one to generalize Lipman's Theorem 7.1.1 to the situation when S is only assumed to have projective dimension one over R. This was pursued in [55].

7.2 The Grothendieck–Lefschetz theorem

Next we discuss some relationships between Tor rigidity and a strong version of the famous Grothendieck–Lefschetz theorem, which we now recall.

Theorem 7.2.1 (Grothendieck–Lefschetz). *Let R be an abstract complete intersection of dimension at least 4 and $\operatorname{Spec}^o R$ be the punctured spectrum of R. Then $\operatorname{Pic}(\operatorname{Spec}^o R) = 0$.*

To see the connection to what we have discussed, let us rephrase the above result purely in the language of commutative algebra. Let $X = \operatorname{Spec}^o R$ and $\mathcal{L}$ denote an element in $\operatorname{Pic}(X)$. Let $M = \Gamma_X(\mathcal{L})$ generated by the global sections of the pushforward $i_*\mathcal{L}$ to $\operatorname{Spec} R$. Then M is a reflexive module over R which is locally free of rank 1 on X. Furthermore

$$\operatorname{Hom}_R(M, M) \cong \Gamma_X(\mathcal{H}om_{\mathcal{O}_X}(\mathcal{L}, \mathcal{L})) \cong \Gamma_X(\mathcal{O}_X) = R.$$

Thus, the Grothendieck–Lefschetz theorem can be viewed as a special case of the following:

Conjecture 7.2.2. *Let R be an abstract complete intersection of dimension at least* 4*. Suppose M is a reflexive module in* mod R *which is locally free on* $\operatorname{Spec}^o R$*. If* depth $\operatorname{Hom}_R(M, M) \geq 4$ *then M is free.*

The above assertion follows rather easily if we make an additional assumption that depth $M \geq 3$.

Proposition 7.2.3. *Let R be an abstract complete intersection of dimension at least* 4*. Suppose $M \in$* mod R *is locally free on* $\operatorname{Spec}^o R$ *and* depth $M \geq 3$*. If* depth $\operatorname{Hom}_R(M, M) \geq 4$ *then M is free.*

Proof. The issues are not affected by completion, so we may assume that R is a quotient of a complete regular local ring T by a regular sequence. By Lemma 2.3.2 we know that $\operatorname{Ext}^1_R(M, M) = \operatorname{Ext}^2_R(M, M) = 0$. It follows from Theorem 6.1.3 that M lifts to a T-module M'. Since for any $N \in \operatorname{mod} R$ there is an isomorphism $\operatorname{Tor}_i^T(M', N) \cong \operatorname{Tor}_i^R(M, N)$ and M' is Tor rigid (as a T-module) it follows that M is Tor rigid. Now we can use Proposition 3.1.2 to conclude that M is free. □

If we only assume M to be reflexive then one can only deduce that $\operatorname{Ext}^1_R(M, M) = 0$ by Lemma 2.3.2. Of course, by the virtue of Proposition 3.1.2, we can conclude that M is free if it is known to be Tor rigid to begin with.

Corollary 7.2.4. *Let R be a local ring of depth at least* 3 *satisfying Serre's condition* (S_2)*. Let M be a reflexive R-module which is locally free on* $\operatorname{Spec}^o R$*. Suppose that the pair* $(\operatorname{Tr}\Omega M, M)$ *is Tor rigid. Then* depth $\operatorname{Hom}_R(M, M) \geq 3$ *if and only if M is free.*

By the above discussion, the following can be viewed as a vast generalization of the Grothendieck–Lefschetz theorem in the hypersurface case.

Theorem 7.2.5. *Let R be a formal hypersurface such that $\hat{R} \cong T/(f)$ where T is an equicharacteristic or unramified regular local ring and $f \in T$ is a regular element. Assume that $d =$* dim R *is even and greater than* 3*. Let M be a reflexive R-module which is locally free on* $\operatorname{Spec}^o R$*. If* depth $\operatorname{Hom}_R(M, M) \geq 3$ *then M is free.*

Proof. Using Lemma 2.3.3 we get $\operatorname{Ext}^1_R(M, M) = 0$. By Proposition 3.1.2 it is enough to show that the pair of modules $(\operatorname{Tr}\Omega M, M)$ is Tor rigid, so it suffices to show that $\theta^R(\operatorname{Tr}\Omega M, M) = 0$ due to Proposition 5.2.2. By definition of the transpose one has the following complex:

$$0 \to \mathsf{Tr} M \to F_1^* \to \mathsf{Tr}\Omega M \to 0$$

whose only nonzero homology in the middle is isomorphic to $\mathrm{Ext}^1_R(M, R)$, which has finite length. By Lemma 5.3.7 it follows that $[\mathsf{Tr} M] = -[\mathsf{Tr}\Omega M]$ in $\overline{G}(R)_{\mathbb{Q}}$. Thus $\theta^R(\mathsf{Tr}\Omega M, M) = -\theta^R(\mathsf{Tr} M, M) = -\theta(M^*, M) = 0$ (the last two equalities are consequences of Proposition 5.2.2 and Theorem 5.3.8). □

7.3 Gabber's Conjecture

The Grothendieck–Lefschetz theorem says that a complete intersection in dimension 4 is parafactorial. In dimension three, the statement is no longer true, but one can instead consider a very interesting conjecture, made by Gabber in [32]:

Conjecture 7.3.1 (Gabber). *Let R be a local complete intersection of dimension* 3. *Then* $\mathrm{Pic}(\mathrm{Spec}^o\ R)$ *is torsion-free.*

The above conjecture is equivalent to the statement that the local flat cohomology group $H^2_{\{m\}}(\mathrm{Spec}(R), \mu_n) = 0$ when R is a local complete intersection of dimension 3, and they are both implied by the following (see [32]):

Conjecture 7.3.2. *Let R be a strictly Henselian local complete intersection of dimension at least* 4. *Then the cohomological Brauer group of* $\mathrm{Spec}^o\ R$ *vanishes.*

In fact, the characteristic 0 case of Conjecture 7.3.1 follows from Grothendieck's techniques on local Leftschetz theorems (cf. [12,56]), and the positive characteristic case can be found in [27] (it is probably known to experts, though we can not find an exact reference. It was claimed in [32] that Conjecture 7.3.2 is known in positive characteristic). The proof in [27] actually uses the Tor rigidity of the Frobenius in an essential way. We now describe briefly how one can prove such a statement.

Suppose that R is a local ring of characteristic $p > 0$. Then the map $F : R \to R$ that takes r to r^p is a ring homomorphism, famously known as the Frobenius map. Under this map, the target R now has a new structure of an R-module, which we denote by ${}^f R$.

Now, let us assume that R is a complete intersection of characteristic $p > 0$. The difficult part is to show that $\mathrm{Pic}(\mathrm{Spec}^o\ R)$ has no p-torsion elements (as the ones whose order is relatively prime to p can be ruled out by identical proof to the characteristic 0 case). But we can easily see that the map

$$\phi_F : \mathrm{Pic}(\mathrm{Spec}^o\ R) \to \mathrm{Pic}(\mathrm{Spec}^o\ {}^f R) \cong \mathrm{Pic}(\mathrm{Spec}^o\ R)$$

induced by tensoring with ${}^f R$ is actually a self-map on the abelian group $\mathrm{Pic}(\mathrm{Spec}^o\ R)$ given by multiplying with $p^{\dim R}$. Thus, it is enough to show that ϕ_F is injective. This can be achieved (at least when F is a finite morphism) by adapting the argument in Theorem 7.1.1 together with the following (cf. [10]):

Theorem 7.3.3 (Dutta, Avramov–Miller). *Let R be a local complete intersection. Then ${}^f R$ is a* Tor*-rigid R-module.*

Thus, Gabber's conjecture only remains in the mixed characteristic case. We shall see next that the hypersurface situation can also be proved, even in the mixed characteristic case. In fact, one can prove the following more general result (see Sects. 2.6 and 7.2 for notations and discussions of how such a statement implies what we want):

Theorem 7.3.4. *Let R be local hypersurface of dimension 3. Let N be a finitely generated reflexive R-module which is locally free on* $\mathrm{Spec}^o\, R$. *Furthermore, assume that the first local Chern class of N is torsion in* $\mathrm{CH}^1(R)$. *Then* $\mathrm{Hom}_R(N, N)$ *is a maximal Cohen–Macaulay R-module if and only if N is free.*

Proof. We sketch the main ideas. As in the proof of Theorem 7.2.5, one first proves that $\theta^R(\mathsf{Tr}\Omega N, N) = 0$. By taking direct sum of copies of N, we can assume that $c_1(N) = 0$. Then a special version of the Bourbaki sequence as in [43, Theorem 1.4] produces a short exact sequence:

$$0 \to F \to N \to I \to 0$$

where I is an ideal of height two, that is, $\dim R/I = 1$. A similar argument as in Theorem 5.3.2 shows that indeed $\theta^R(\mathsf{Tr}\Omega N, N) = 0$.

In the next part of the proof, we utilize the theory of maximal Cohen–Macaulay approximation developed in [3]. Given N, we can fit it in a short exact sequence:

$$0 \to G \to M \to N \to 0$$

where $M \in \mathrm{MCM}(R)$ and G has finite projective dimension. Since N is reflexive and $\dim R = 3$, we see that G is actually free. As in the proof of Theorem 7.2.5 and Proposition 3.1.2, we know that $\mathrm{Ext}^1_R(N, N) = \mathrm{Tor}^R_1(\mathsf{Tr}\Omega N, N) = 0$. Using the exact sequence above, we get $\mathrm{Tor}^R_1(\mathsf{Tr}\Omega N, M) = 0$. We also know that $\theta^R(\mathsf{Tr}\Omega N, M) = \theta^R(\mathsf{Tr}\Omega N, N) = 0$. However, since M is maximal Cohen–Macaulay,

$$\theta^R(\mathsf{Tr}\Omega N, M) = \ell(\mathrm{Tor}^R_2(\mathsf{Tr}\Omega N, M)) - \ell(\mathrm{Tor}^R_1(\mathsf{Tr}\Omega N, M)).$$

Thus, $\mathrm{Tor}^R_i(\mathsf{Tr}\Omega N, M) = 0$ for all $i \gg 0$. It follows from Theorem 6.1.1 that either $\mathrm{pd}_R M < \infty$ or $\mathrm{pd}_R \mathsf{Tr}\Omega N < \infty$. In the former case, we deduce that $\mathrm{pd} N < \infty$ and in the latter, $\mathrm{pd}_R(\Omega N)^* < \infty$. But if $\mathrm{pd} N < \infty$ and it is not free then $\mathrm{pd}_R N = 1$. Then $\mathrm{Ext}^1_R(N, N)$ cannot be 0 by Nakayama's Lemma, a contradiction. On the other hand, if $\mathrm{pd}_R(\Omega N)^* < \infty$ then $(\Omega N)^*$ must be free since it is maximal Cohen–Macaulay. Therefore ΩN is free; in other words $\mathrm{pd}_R N < \infty$, and we already saw that this implies N must be free. □

The above theorem gives the following characterization of the UFD property for dimension three hypersurfaces.

Corollary 7.3.5. *Let R be a local hypersurface with isolated singularity and* dim R=3. *The following are equivalent:*

(1) $\theta^R(M, N) = 0$ *for all* $M, N \in \mathrm{mod}(R)$.
(2) The class group $\mathrm{Cl}(R)$ *is torsion.*
(3) R *is a UFD (equivalently,* $\mathrm{Cl}(R) = 0$*).*

8 Applications: Intersection of Subvarieties, Splitting of Vector Bundles and Non-commutative Crepant Resolutions

In this section we discuss further applications which are somewhat more geometric in nature. Thus we begin with a brief discussion of how some questions on projective varieties can be reduced to local algebra.

8.1 Local Rings of Cones of Projective Varieties

Let A be a finitely generated standard graded ring over a field k and $R = A_{\mathfrak{m}}$ where $\mathfrak{m}$ is the irrelevant ideal. Let $X = \mathrm{Proj}\ A$. Then $\dim X + 1 = \dim A = \dim R$. When k is perfect, X is smooth if and only if R has isolated singularity.

Let $\mathcal{F}$ be a quasi-coherent sheaf on X and

$$\Gamma_*(\mathcal{F}) = \oplus_{i \in \mathbb{Z}} \Gamma(X, \mathcal{F}(i)).$$

Let M be a graded A-module and $\widetilde{M}$ be the corresponding sheaf on X. Given any coherent sheaf $\mathcal{F}$ on X one can find a finitely generated module M such that $\widetilde{M} \cong \mathcal{F}$. In general such a module is not unique, but we have the short exact sequence:

$$0 \to H^0_{\mathfrak{m}}(M) \to M \to \Gamma_*(\mathcal{F}) \to H^1_{\mathfrak{m}}(M) \to 0.$$

and for $i > 0$, isomorphisms

$$H^{i+1}_{\mathfrak{m}}(M) \cong \oplus_{i \in \mathbb{Z}} H^i(X, \mathcal{F}(i))$$

where $H^i_{\mathfrak{m}}(M)$ denotes the ith local cohomology supported at $\mathfrak{m}$ of M. When A is normal of dimension at least 2 and F is a vector bundle on X, M can be chosen to be a reflexive module over A (which is locally free on Spec $A - \{\mathfrak{m}\}$). In this situation, $\mathcal{F} \cong \oplus_{i \in S} \mathcal{O}_X(i)$ for some set of indexes S (such $\mathcal{F}$ are sometimes called *dissocié*) if and only if M is a free A-module. One can localize at $\mathfrak{m}$, and all the information such as reflexivity, local freeness on the punctured spectrum, and the local cohomology modules are preserved between M and $M_{\mathfrak{m}}$. Thus, certain statements about vector bundles over the projective variety X can be deduced from their analogues on the punctured spectrum $\mathrm{Spec}^o\ R$.

8.2 Intersections of Subvarieties

A well-known fact of projective geometry is that in projective space any two lines must intersect. A much more general fact is the following

Theorem 8.2.1. *Let U, V be subschemes of the projective space $\mathbf{P}_k^n$ for some field k. If $\dim U + \dim V \geq n$, then $U \cap V$ is nonempty.*

For convenience we isolate this property, which we do not know a standard reference for.

Definition 8.2.2. A Noetherian scheme X is called decent if for any subschemes U, V such that $\dim U + \dim V \geq \dim X$, $U \cap V$ is non-empty.

Proposition 8.2.3. *Let $R = T/(f)$ be a local hypersurface with isolated singularity such that T is an equicharacteristic or unramified regular local ring. Assume that $\theta^R(M, N) = 0$ for all $M, N \in \operatorname{mod} R$. Then $X = \operatorname{Spec}^o R$ is decent in the sense of Definition 8.2.2.*

Proof. Let $I, J \subsetneq R$ be the ideals defining U, V, respectively. Suppose that $U \cap V = \emptyset$. Then it follows that $R/I \otimes_R R/J$ has finite length. But as $\theta^R(R/I, R/J) = 0$, we have that $\chi^T(R/I, R/J) = 0$; hence $\dim R/I + \dim R/J < \dim T$ (cf. Theorem 2.4.1 and Proposition 5.2.1) or $\dim U + 1 + \dim V + 1 < \dim X + 2$, a contradiction. □

Thus, we have the following consequences of Theorems 5.3.4 and 5.3.6.

Theorem 8.2.4. *For the following situations $\operatorname{Spec}^o R$ is decent in the sense of Definition 8.2.2.*

(1) R is an excellent (formal) hypersurface with isolated singularity, containing a field and $\dim R = 4$.
(2) $R = k[x_0, \cdots, x_d]_{\mathfrak{m}}/(f)$, where f is a homogenous polynomial defining a smooth hypersurface in $\mathbf{P}_k^d$ over a perfect field k and $\mathfrak{m} = (x_0, \cdots, x_d)$.
(3) $R = k[[x_0, \cdots, x_d]]/(f)$, a hypersurface-isolated singularity with k a field of characteristic 0.

One can say quite a bit about the decency property of projective varieties since a lot is known about intersection theory on such varieties. For example, we note below that smooth projective complete intersections are decent. The proof is essentially in [24, Theorem 4.10].

Theorem 8.2.5. *Let X be a smooth complete intersection in $\mathbf{P}_k^n$ for some algebraically closed field k. Then X is decent in the sense of Definition 8.2.2.*

Proof. We are going to use l-adic cohomology (for basic properties and notations, we refer to [51] or [59]). Let l be a prime number such that $l \neq char(k)$. There is a class map:

$$cl : \mathrm{CH}^r(X) \to H^{2r}(X, \mathbb{Q}_l(r))$$

This map gives a graded rings homomorphism $\mathrm{CH}^*(X) \to \oplus H^{2r}(X, \mathbb{Q}_l(r))$ (with the intersection product on the left-hand side and the cup product on the right-hand side, see [51], VI, 10.7 and 10.8). Let $a = \dim U$ and $b = \dim V$, and we may assume $a \geq b$. Suppose $a + b = \dim X = n$ (if $a + b > n$, we can always choose some subvariety of smaller dimension inside U or V such that equality occurs). Then $2a \geq n$, but n is odd, so $2a > n$. Let $h \in \mathrm{CH}^1(X)$ represent the hyperplane section. By the weak Lefschetz theorem (see, for example, [59], 7.7, page 112) and the fact that $2(n-a) < n$, we have:

$$H^{2(n-a)}(X, \mathbb{Q}_l(n-a)) \cong H^{2(n-a)}(\mathbf{P}_k^n, \mathbb{Q}_l(n-a)).$$

The latter is generated by a power of the class of the hyperplane section. Thus $cl(U) = cl(h)^{n-a}$ in $H^{2(n-a)}(X, \mathbb{Q}_l(n-a))$. We then have

$$cl(U.V) = cl(U) \cup cl(V) = cl(h)^{n-a} \cup cl(V) = cl(h^{n-a}.V).$$

The last term is equal to the degree of $h^{n-a}.V \in \mathrm{CH}^n(X)$, so it is nonzero. But the first term has to be 0 by assumption. This contradiction proves the theorem. □

8.3 Splitting of Vector Bundles

In this section we discuss the problem of splitting of vector bundles on schemes. In general this is a non-trivial question. For example, it is not known whether there are indecomposable vector bundles of rank two on projective spaces in dimension at least 6. As discussed in Sect. 8.1, there is an intimate connection between such questions on a projective variety and the punctured spectrum of a local ring. Thus we shall focus on the latter. We begin with an algebraic generalization of a result by Faltings, who proved that any vector bundle on $\mathbf{P}^n$ that is globally generated by at most n sections must be a direct sum of line bundles [30]. See also [38].

Theorem 8.3.1 (After Faltings). *Suppose R is a local ring satisfying* (S_2) *and* $\mathrm{Spec}^o\, R$ *is decent in the sense of Definition 8.2.2. Suppose $M \in \mathrm{mod}\, R$ is locally free of constant rank on* $\mathrm{Spec}^o\, R$ *and* $\operatorname{depth} M \geq 1$. *If M has less than* $\dim R$ *generators then M is free.*

Proof. If $\dim R \leq 1$ there is no content, so we assume that $\dim R \geq 2$. As R is (S_2), depth R is at least 2 also. We use induction on r, the number of generators of M. If $r = 1$ the conclusion follows immediately, so we assume $r > 1$.

We may also assume that M has no free summand. Consider a short exact sequence:

$$0 \to N \xrightarrow{\alpha} F \xrightarrow{\beta} M \to 0$$

with $N \cong \Omega M$. Pick a standard basis for F and e a basis element. Let $x = \beta(e)$ and $y = \alpha^*(e^*)$ where e^* is the basis element of F^* corresponding to e. We can assume that the order ideals $I = O_M(x)$ and $J = O_{N^*}(y)$ (see Sect. 2.1) are proper; otherwise M would have a free summand. Let $a = \operatorname{rk} M$ and $b = \operatorname{rk} N$. Clearly $a + b = \operatorname{rk} F = r$. Locally on $\operatorname{Spec}^o R$, I and J can be generated by at most a and b generators, respectively.

As M is locally free on $\operatorname{Spec}^o R$, we must have $\operatorname{Supp} R/I \cap \operatorname{Supp} R/J = \{\mathfrak{m}\}$. In other words, let U, V be the subschemes of $\operatorname{Spec}^o R$ defined by I, J, respectively, then $U \cap V = \emptyset$. If both I, J are not $\mathfrak{m}$-primary, then by localizing at some minimal primes, we see that their heights are at most a and b, respectively. Thus

$$\dim R/I + \dim R/J \leq 2 \dim R - a - b = 2 \dim R - r > \dim R$$

or

$$\dim U + \dim V \geq \dim \operatorname{Spec}^o R$$

which contradicts the assumption that $\operatorname{Spec}^o R$ is decent. However, if I is $\mathfrak{m}$-primary, then it follows that the quotient M/Rx is locally free on $\operatorname{Spec}^o R$ and has $r-1$ generators. Furthermore the exact sequence

$$0 \to Rx \to M \to M/Rx \to 0$$

shows that depth $M/Rx \geq 1$, thus by induction it is free, and so is M. If J is $\mathfrak{m}$-primary then similarly N^* has a free summand generated by y. Dualizing again we see that $x = 0$, a contradiction.

□

Corollary 8.3.2. *Let R be an regular local ring or a local hypersurface as in Theorem 8.2.4 with* $\dim R \geq 2$. *Let $X = \operatorname{Spec}^o R$. Then a vector bundle over X that is globally generated by at most* $\dim X$ *sections must be trivial.*

The next result shows how one can detect triviality of vector bundles from vanishing of certain local cohomology modules.

Corollary 8.3.3. *Let R be a abstract hypersurface such that $\hat{R} \cong T/(f)$ where T is an equicharacteristic or unramified regular local ring and $f \in T$ is a regular element. Assume that $d = \dim R$ is even and greater than 3. Let M be a reflexive R-module which is locally free on $\operatorname{Spec}^o R$. If $H^2_{\mathfrak{m}}(M^* \otimes_R M) = 0$ then M is free.*

Proof. One has the canonical map:

$$M^* \otimes_R M \to \operatorname{Hom}_R(M, M).$$

By assumption the kernel and cokernel have finite lengths. From that one can easily see that $H^2_{\mathfrak{m}}(M^* \otimes_R M) = H^2_{\mathfrak{m}}(\operatorname{Hom}_R(M, M))$. It follows that depth $\operatorname{Hom}_R(M, M) \geq 3$, so one can apply Theorem 7.2.5. □

One can now deduce the following from the discussion at the beginning of this section.

Corollary 8.3.4. *Let X be a projective hypersurface over a field such that* $\dim X$ *is odd and at least* 3*. Let $\mathcal{F}$ be a vector bundle on X such that*

$$\oplus_{i \in \mathbb{Z}} H^1(X, \mathcal{F} \otimes \mathcal{F}^*(i)) = 0.$$

Then $\mathcal{F}$ is a direct sum of line bundles.

8.4 Non-commutative Crepant Resolutions

Recently, the study of $\mathrm{Hom}_R(M, N)$ over Gorenstein rings has taken on a renewed significance, due to the following concept.

Definition 8.4.1 (Van den Bergh [61]). Let R be a Gorenstein domain. Suppose that there exists a reflexive module N satisfying:

(1) $A = \mathrm{Hom}_R(N, N)$ is a maximal Cohen–Macaulay R-module.
(2) A has finite global dimension equal to $d = \dim R$.

Then A is called a noncommutative crepant resolution (henceforth NCCR) of R.

This concept was suggested by Van den Bergh to give a conceptual proof of the three-dimensional case of a famous conjecture by Bondal–Orlov that two birational Calabi–Yau varieties have equivalent derived categories. In dimension three one can reduce to the case of two such varieties X, X' related by a "flop." In such situations one can prove that each derived category is equivalent to a third category, the derived category of a non-commutative algebra which has exactly the property of the endomorphism ring A described in the definition above. See [61, 62] for details and [47] for a very nice survey on this rapidly developing topic.

The existence of such NCCR is a subtle question. For example, it was shown by Stafford and Van den Bergh [60] that the existence of NCCRs over a Gorenstein affine k-algebra R where k is an algebraically closed field of characteristic 0 forces Spec R to have only rational singularities. This was recently extended to non-Gorenstein setting in certain cases [26]. However, one can immediately derive, from what we know so far, many necessary conditions. The following is a sample of such results, more can be said if one makes additional assumptions on the module M (see Sect. 7).

Corollary 8.4.2. *Let R be a local hypersurface with an isolated singularity. In the following situations there are no NCCRs over* Spec R*:*

(1) $\dim R = 3$ *and R is a UFD.*
(2) $\dim R$ *is even and at least* 4*, and $R \cong T/(f)$ where T is an equicharacteristic or unramified regular local ring.*

Proof. By Theorems 7.3.4 and 7.2.5 such a module giving an NCCR would have to be free. However, if M is free, then its endomorphism ring would be Morita equivalent to R, which does not have finite global dimension. □

9 Open Questions

9.1 Some Open Questions

In this section we describe many open questions that are actually quite natural in view of what has been discussed so far. As is evident from the previous sections, they are mostly motivated by results or questions from outside commutative algebra. We will be slightly provocative and call the questions we are more confident about "conjectures"; the rest will be stated as mere "questions." We already mentioned three of them, Question 3.1.5 and Conjectures 5.3.5 and 7.2.2. Note that Conjecture 7.2.2 can be viewed as a rather ambitious generalization of the Grothendieck–Lefschetz theorem.

Next we discuss open questions related to Gabber's Conjecture 7.3.1. Obviously, we would like to know if the stronger version for hypersurfaces, Theorem 7.3.4, can be proved for complete intersections.

Conjecture 9.1.1. *Let R be local complete intersection of dimension 3. Let N be a reflexive R-module which is locally free of constant rank on* $\operatorname{Spec}^o R$*. Furthermore, assume that $[N] = 0$ in $\overline{G}(R)_{\mathbb{Q}}$, the reduced Grothendieck group of R with rational coefficients. Then* $\operatorname{Hom}_R(N, N)$ *is a maximal Cohen–Macaulay R-module if and only if N is free.*

In view of the proof of the Theorem 7.3.4 and the known Tor-rigidity results for regular and hypersurface rings, we feel it is reasonable to make the following.

Conjecture 9.1.2. *Let R be local complete intersection (of arbitrary dimension). Let M, N be R-modules such that M is locally free of constant rank on* $\operatorname{Spec}^o R$ *and $[N] = 0$ in $\overline{G}(R)_{\mathbb{Q}}$. Then (M, N) is* Tor *rigid, namely, that for any $i > 0$,* $\operatorname{Tor}_i^R(M, N) = 0$ *forces* $\operatorname{Tor}_j^R(M, N) = 0$ *for $j \geq i$.*

By virtue of Proposition 5.2.2, the above statement is true for hypersurfaces defined over an equicharacteristic or unramified regular local ring. An interesting consequence that is worth pointing out is when R is artinian.

Conjecture 9.1.3. *Let R be an artinian local complete intersection. Let $M, N \in$* mod R*. The following are equivalent:*

(1) $\operatorname{Tor}_i^R(M, N) = 0$ *for some $i > 0$*
(2) $\operatorname{Ext}_R^i(M^*, N) = 0$ *for some $i > 0$*
(3) $\operatorname{Ext}_R^i(M, N) = 0$ *for some $i > 0$.*
(4) $\operatorname{Tor}_i^R(M^*, N) = 0$ *for some $i > 0$*
(5) $\operatorname{Tor}_i^R(M, N) = 0$ *for all $i > 0$.*

(6) $\operatorname{Ext}^i_R(M^*, N) = 0$ *for all* $i > 0$
(7) $\operatorname{Ext}^i_R(M, N) = 0$ *for all* $i > 0$.
(8) $\operatorname{Tor}_i^R(M^*, N) = 0$ *for all* $i > 0$
(9) $V_R(M) \cap V_R(N) = \emptyset$

Assuming Conjecture 9.1.2 we have that (1) $\Leftrightarrow$ (5) and (4) $\Leftrightarrow$ (8). The equivalence of $(5), (6), (7), (8), and (9)$ holds unconditionally by Theorem 6.1.1. It would be enough now to show (3) $\Rightarrow$ (8) and (2) $\Rightarrow$ (5).

Recall the exact sequence we used in the proof of Proposition 3.1.2:

$$\operatorname{Tor}_2^R(\mathsf{Tr}\Omega^n M, N) \to \operatorname{Ext}^n_R(M, R) \otimes_R N \to \operatorname{Ext}^n_R(M, N) \to \operatorname{Tor}_1^R(\mathsf{Tr}\Omega^n M, N) \to 0$$

In our situation $\operatorname{Ext}^n_R(M, R) = 0$ as R is artinian and Gorenstein, so the sequence degenerates to an isomorphism: $\operatorname{Ext}^n_R(M, N) \cong \operatorname{Tor}_1^R(\mathsf{Tr}\Omega^n M, N)$ for all $n > 0$.

Since Tor rigidity holds by Conjecture 9.1.2 and $M^* \cong \Omega^{n+2}\mathsf{Tr}\Omega^n M$, we have that (3) $\Rightarrow$ (8). The implication (2) $\Rightarrow$ (5) holds by symmetry.

In fact, even for artinian Gorenstein rings we do not know of any module with no non-trivial self-extensions. Thus we ask the following question

Question 9.1.4. Let R be an artinian, Gorenstein local ring and $M \in \operatorname{mod} R$. If $\operatorname{Ext}^1_R(M, M) = 0$, is M free?

We expect the answer to the above question to be negative. However, it looks a rather difficult problem to construct a counterexample. An affirmative answer to this question is easily seen to be equivalent to an affirmative answer to the following:

Question 9.1.5. Let R be a Gorenstein local ring of dimension d and $M \in \operatorname{MCM}(R)$. If $\operatorname{Ext}^i_R(M, M) = 0$ for $1 \le i \le d+1$, is M free?

Obviously this is a strengthened version of the famous Auslander–Reiten conjecture for the commutative Gorenstein case. For some results in this direction, see [41].

In view of Theorem 8.2.5 one can make the following:

Conjecture 9.1.6. *Let* R *be a local complete intersection with isolated singularity. Then* $\operatorname{Spec}^o R$ *is decent in the sense of Definition 8.2.2.*

Unfortunately the generalized version of Hochster's theta pairing introduced in Sect. 6.2 cannot be used to approach this conjecture. It does suggest that η^R should vanish when R has isolated singularity by Corollary 6.2.6.

Conjecture 9.1.7 ([53, Conjecture 2.4]). *Let* R *be a local complete intersection of codimension at least two and isolated singularity. Then* $\eta^R(M, N) = 0$ *for any* $M, N \in \operatorname{mod} R$.

It is easy to see that the above statement holds when $\dim R = 0$ or $\dim R = 1$ and R is a domain, as $\overline{G}(R)_{\mathbb{Q}} = 0$ in such situations. A very recent result by Moore–Piepmeyer–Spiroff–Walker in [53, Theorem 4.5] shows that this is true in the standard graded case.

Acknowledgements The author would like to thank Jesse Burke and Olgur Celikbas for some useful comments on the materials. He is also grateful to the anonymous referee whose careful reading and suggestions significantly improve the manuscript. The author is partially supported by NSF grants DMS 0834050 and DMS 1104017.

References

1. Auslander, M.: Modules over unramified regular local rings. Ill. J. Math. **5**, 631–647 (1961)
2. Auslander, M.: Modules over unramified regular local rings. Proc. Internat. Congr. Mathematicians (Stockholm, 1962). Inst. Mittag-Leffler, Djursholm, pp. 230–233 (1963)
3. Auslander, M., Buchweitz, R.-O.: The homological theory of maximal Cohen–Macaulay approximations. Mémoires de la Société Mathématique de France Sr. 2, **38**, 5–37 (1989)
4. Auslander, M., Ding, S., Solberg, Ø.: Liftings and weak liftings of modules. J. Algebra **156**, 273–317 (1993)
5. Avramov, L.L.: Modules of finite virtual projective dimension. Invent. Math. **96**, 71–101 (1989)
6. Avramov, L.L.: Infinite free resolutions. Six Lectures in Commutative Algebra (Bellaterra, 1996). Progress in Mathematics, vol. 166, pp. 1–118. Birkhausër, Boston (1998)
7. Avramov, L.L., Sun, L.-C.: Cohomology operators defined by a deformation. J. Algebra **204**, 684–710 (1998)
8. Avramov, L.L., Buchweitz, R.-O.: Support varieties and cohomology over complete intersections. Invent. Math. **142**, 285–318 (2000)
9. Avramov, L.L., Iyengar, S.: Detecting flatness over smooth bases. J. Algebraic Geom. **22**, 35–47 (2013)
10. Avramov, L.L., Miller, C.: Frobenius powers of complete intersections. Math. Res. Lett. **8**, 225–232 (2001)
11. Avramov, L.L., Gasharov, V., Peeva, I.: Complete intersection dimension. Publ. Math. I.H.E.S. **86**, 67–114 (1997)
12. Badescu, L.: A remark on the Grothendieck–Lefschetz theorem about the Picard group. Nagoya Math. J. **71**, 169–179 (1978)
13. Bruns, W., Herzog, J.: Cohen–Macaulay Rings. Cambridge University Press, Cambridge (1996)
14. Buchweitz, R.-O.: Maximal Cohen–Macaulay modules and Tate cohomology over Gorenstein rings. Preprint, University Hannover (1986)
15. Buchweitz, R.-O., van Straten, D.: An Index Theorem for Modules on a Hypersurface Singularity. Moscow M. Journal **12** (2012). arXiv:1103.5574
16. Celikbas, O.: Vanishing of Tor over complete intersections. J. Commutative Algebra **3**, 168–206 (2011)
17. Celikbas, O., Dao, H.: Necessary conditions for the depth formula over Cohen–Macaulay local rings. J. Pure Appl. Algebra. To appear
18. Celikbas, O., Dao, H.: Asymptotic Behavior of Ext functors for modules of finite complete intersection dimension. Math. Z. **269**, 1005–1020 (2011)
19. Chan, C.-Y.J.: Filtrations of modules, the Chow group, and the Grothendieck group. J. Algebra **219**, 330–344 (1999)
20. Choi, S., Iyengar, S.: On a depth formula for modules over local rings. Comm. Algebra **29**, 3135–3143 (2001)
21. Dao, H.: What is Serre's condition (S_n) for sheaves? Mathoverflow question. http://mathoverflow.net/questions/22228
22. Dao, H.: Decency and rigidity over hypersurfaces. Tran. Amer. Math. Soc. To appear.
23. Dao, H.: Asymptotic behavior of Tor over complete intersections and applications. Preprint (2007). arXiv: 0710.5818

24. Dao, H.: Some observations on local and projective hypersurfaces. Math. Res. Lett. **15**, 207–219 (2008)
25. Dao, H.: Remarks on non-commutative crepant resolutions of complete intersections. Adv. Math. **224**, 1021–1030 (2010)
26. Dao, H., Iyama, O., Takahashi, R., Vial, C.: Non-commutative resolutions and Grothendieck group. Preprint. arXiv:1205.4486
27. Dao, H., Li, J., Miller, C.: On (non)rigidity of the Frobenius over Gorenstein rings. Algebra Number Theory **4–8**, 1039–1053 (2010)
28. Dutta, S.: On negativity of higher Euler characteristics. Amer. J. Math. **126**, 1341–1354 (2004)
29. Eisenbud, D.: Homological algebra on a complete intersection, with an application to group representations. Tran. Amer. Math. Soc. **260**, 35–64 (1980)
30. Faltings, G.: Ein Kriterium für vollständige Durchschnitte. Invent. Math. **62**(3), 393–401 (1981)
31. Fulton, W.: Intersection Theory. Springer, Berlin (1998)
32. Gabber, O.: On purity for the Brauer group. Arithmetic Algebraic Geometry. Oberwolfach Report No. 34, 1975–1977 (2004)
33. Gulliksen, T.H.: A change of ring theorem with applications to Poincaré series and intersection multiplicity. Math. Scan **34**, 167–183 (1974)
34. Hartshorne, R.: Coherent Functors. Adv. Math. **140**, 44–94 (1998).
35. Heitmann, R.: A counterexample to the rigidity conjecture for rings. Bull. Am. Math. Soc. 29, 94–97 (1993)
36. Hochster, M.: The dimension of an intersection in an ambient hypersurface. In: Proceedings of the First Midwest Algebraic Geometry Seminar (Chicago Circle, 1980), Lecture Notes in Mathematics, vol. 862, pp. 93–106. Springer, Berlin (1981)
37. Hochster, M.: Euler characteristics over unramified regular local rings. Ill. J. Math. **28**, 281–288 (1984)
38. Horrocks, G.: Vector bundles on the punctured spectrum of a local ring. Proc. London Math. Soc. **14**, 689–713 (1964)
39. Huneke, C., Wiegand, R.: Tensor products of modules and the rigidity of Tor. Math. Ann. **299**, 449–476 (1994)
40. Huneke, C., Wiegand, R.: Tensor products of modules, rigidity and local cohomology. Math. Scan. **81**, 161–183 (1997)
41. Huneke, C., Leuschke, G.: On a conjecture of Auslander and Reiten. J. Algebra **275**, 781–790 (2004)
42. Huneke, C., Wiegand, R.: Correction to: "Tensor products of modules and the rigidity of Tor". Math. Ann. **338**, 291–293 (2007)
43. Huneke, C., Wiegand, R., Jorgensen, D.: Vanishing theorems for complete intersections. J. Algebra **238**, 684–702 (2001)
44. Jorgensen, D.: Finite projective dimension and the vanishing of Ext(M,M). Comm. Alg. **36**(12), 4461–4471 (2008)
45. Jorgensen, D., Heitmann, R.: Are complete intersections complete intersections? J. Algebra **371**, 276–299 (2012)
46. Jothilingam, P.: A note on grade. Nagoya Math. J. **59**, 149–152 (1975)
47. Leuschke, G.: Non-commutative crepant resolutions: scenes from categorical geometry. Progress in commutative algebra, **1**, 293–361. de Gruyter, Berlin, (2012), available at http://www.leuschke.org/research/scenes-from-categorical-geometry/
48. Lichtenbaum, S.: On the vanishing of Tor in regular local rings. Illinois J. Math. **10**, 220–226 (1966)
49. Lipman, J.: Ring with discrete divisor class group: Theorem of Danilov–Samuel. Amer. J. Math. **101**, 203–211 (1979)
50. Matsumura, H.: Commutative ring theory. In: Cambridge Studies in Advanced Mathematics, vol. 8. Cambridge University Press, Cambridge (1986)
51. Milne, J.S.: Etale Cohomology. Princeton University Press, Princeton (1980)

52. Moore, F., Piepmeyer, G., Spiroff, S., Walker, M.: Hochster's theta invariant and the Hodge-Riemann bilinear relations. Adv. Math. **226**, 1692–1714 (2011)
53. Moore, F., Piepmeyer, G., Spiroff, S., Walker, M.: The vanishing of a higher codimension analogue of Hochster's theta invariant. Adv. Math. **226**(2), 1692–1714 (2011)
54. Polishchuk, A., Vaintrob, A.: Chern characters and Hirzebruch–Riemann–Roch formula for matrix factorizations. Duke Math. J. **161**(10), 1863–1926 (2012)
55. Pruyn, S.: On the divisor class group. Master's Thesis, University of Kansas, Lawrence (2011)
56. Robbiano, L.: Some properties of complete intersections in "good" projective varieties. Nagoya Math. J. **61**, 103–111 (1976)
57. Roberts, P.: Multiplicities and Chern classes in Local Algebra. Cambridge University Press, Cambridge (1998)
58. Serre, J.P.: Algèbre locale. multiplicités. In: Lecture Note in Mathematics, vol. 11. Springer, Berlin (1965)
59. Shafarevich, I.R. (ed.): Algebraic Geometry II. Cohomology of algebraic varieties. Algebraic surfaces. In: Encyclopaedia of Mathematical Sciences, vol. 35. Springer, Berlin (1995)
60. Stafford, J.T., Van den Bergh, M.: Noncommutative resolutions and rational singularities. Mich. Math. J. **57**, 659–674 (2008)
61. Van den Bergh, M.: Non-commutative crepant resolutions. In: The Legacy of Niels Henrik Abel, pp. 749–770. Springer, Berlin (2004)
62. Van den Bergh, M.: Three-dimensional flops and noncommutative rings. Duke Math. J. **122**(3), 423–455 (2004)

Powers of Square-Free Monomial Ideals and Combinatorics

Christopher A. Francisco, Huy Tài Hà, and Jeffrey Mermin

1 Introduction

Powers of ideals are instrumental objects in commutative algebra. In addition, square-free monomial ideals are intimately connected to combinatorics. In this chapter, we survey work on secant, symbolic, and ordinary powers of square-free monomial ideals and their combinatorial consequences in (hyper)graph theory and linear integer programming.

There are two well-studied basic correspondences between square-free monomial ideals and combinatorics. Each arises from the identification of square-free monomials with sets of vertices of either a simplicial complex or a hypergraph. The Stanley–Reisner correspondence associates to the nonfaces of a simplicial complex Δ the generators of a square-free monomial ideal, and vice-versa. This framework leads to many important results relating (mostly homological) ideal-theoretic properties of the ideal to properties of the simplicial complex; see [4, Chap. 5] and [25, Sects. 61–64].

The edge and cover ideal constructions identify the minimal generators of a square-free monomial ideal with the edges (covers) of a simple hypergraph. The edge ideal correspondence is more naïvely obvious but less natural than the Stanley–Reisner correspondence, because the existence of a monomial in this ideal does not translate easily to its presence as an edge of the (hyper)graph. Nevertheless, this correspondence has proven effective at understanding properties

C.A. Francisco • J. Mermin (✉)
Department of Mathematics, Oklahoma State University, 401 Mathematical Sciences, Stillwater, OK 74078, USA
e-mail: chris@math.okstate.edu; mermin@math.okstate.edu

H.T. Hà
Department of Mathematics, Tulane University, 6823 St. Charles Ave., New Orleans, LA 70118, USA
e-mail: tha@tulane.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_11,

of (hyper)graphs via algebra. We focus on powers of square-free monomial ideals when they are viewed as edge (or cover) ideals of hypergraphs. To the best of our knowledge, there has been little systematic study of the powers of square-free ideals from the Stanley–Reisner perspective.

The general theme of this chapter is the relationship between symbolic and ordinary powers of ideals. This topic has been investigated extensively in the literature (cf. [2, 8, 17, 20]). Research along these lines has revealed rich and deep interactions between the two types of powers of ideals, and often their equality leads to interesting algebraic and geometric consequences (cf. [15, 22, 29–31]). We shall see that examining symbolic and ordinary powers of square-free monomial ideals also leads to exciting and important combinatorial applications.

The chapter is organized as follows. In the next section, we collect notation and terminology. In Sect. 3, we survey algebraic techniques for detecting important invariants and properties of (hyper)graphs. We consider three problems:

1. Computing the chromatic number of a hypergraph
2. Detecting the existence of odd cycles and odd holes in a graph
3. Finding algebraic characterizations of bipartite and perfect graphs

We begin by describing two methods for determining the chromatic number of a hypergraph via an ideal-membership problem, one using secant ideals, and the other involving powers of the cover ideal. Additionally, we illustrate how the associated primes of the square of the cover ideal of a graph detect its odd induced cycles.

The results in Sect. 3 lead naturally to the investigation of associated primes of higher powers of the cover ideal. This is the subject of Sect. 4. We explain how to interpret the associated primes of the sth power of the cover ideal of a hypergraph in terms of coloring properties of its sth expansion hypergraph. Specializing to the case of graphs yields two algebraic characterizations of perfect graphs that are independent of the Strong Perfect Graph Theorem.

Section 5 is devoted to the study of when a square-free monomial ideal has the property that its symbolic and ordinary powers are equal. Our focus is the connection between this property and the Conforti–Cornuéjols conjecture in linear integer programming. We state the conjecture in its original form and discuss an algebraic reformulation. This provides an algebraic approach for tackling this long-standing conjecture.

2 Preliminaries

We begin by defining the central combinatorial object of the chapter.

Definition 2.1. A *hypergraph* is a pair $G = (V, E)$ where V is a set, called the *vertices* of G, and E is a subset of 2^V, called the *edges* of G. A hypergraph is *simple* if no edge contains another; we allow the edges of a simple hypergraph to contain only one vertex (i.e., isolated *loops*). Simple hypergraphs have also been studied under other names, including *clutters* and *Sperner systems*. All hypergraphs in this chapter will be simple.

A *graph* is a hypergraph in which every edge has cardinality exactly two. We specialize to graphs to examine special classes, such as cycles and perfect graphs.

If W is a subset of V, the *induced subhypergraph* of G on W is the pair (W, E_W) where $E_W = E \cap 2^W$ is the set of edges of G containing only vertices in W.

Notation 2.2. Throughout the chapter, let $V = \{x_1, \ldots, x_n\}$ be a set of vertices. Set $S = K[V] = K[x_1, \ldots, x_n]$, where K is a field. We will abuse notation by identifying the square-free monomial $x_{i_1} \ldots x_{i_s}$ with the set $\{x_{i_1}, \ldots, x_{i_s}\}$ of vertices. If the monomial m corresponds to an edge of G in this way, we will denote the edge by m as well.

Definition 2.3. The *edge ideal* of a hypergraph $G = (V, E)$ is

$$I(G) = (m : m \in E) \subset S.$$

On the other hand, given a square-free monomial ideal $I \subset S$, we let $G(I) = (V, \text{gens}(I))$ be the hypergraph associated to I, where $\text{gens}(I)$ is the unique set of minimal monomial generators of I.

Definition 2.4. A *vertex cover* for a hypergraph G is a set of vertices w such that every edge hits some vertex of w, i.e., $w \cap e \neq \varnothing$ for all edges e of G.

Observe that, if w is a vertex cover, then appending a variable to w results in another vertex cover. In particular, abusing language slightly, the vertex covers form an ideal of S.

Definition 2.5. The *cover ideal* of a hypergraph G is

$$J(G) = (w : w \text{ is a vertex cover of } G).$$

In practice, we compute cover ideals by taking advantage of duality.

Definition 2.6. Given a square-free monomial ideal $I \subset S$, the *Alexander dual* of I is

$$I^\vee = \bigcap_{m \in \text{gens}(I)} \mathfrak{p}_m,$$

where $\mathfrak{p}_m = (x_i : x_i \in m)$ is the prime ideal generated by the variables of m.

Observe that if $I = I(G)$ is a square-free monomial ideal, its Alexander dual $I^\vee$ is also square-free. We shall denote by G^* the hypergraph corresponding to $I^\vee$, and call G^* the *dual hypergraph* of G. That is, $I^\vee = I(G^*)$. The edge ideal and cover ideal of a hypergraph are related by the following result.

Proposition 2.7. *The edge ideal and cover ideal of a hypergraph are dual to each other: $J(G) = I(G)^\vee = I(G^*)$ (and $I(G) = J(G)^\vee$). Moreover, minimal generators of $J(G)$ correspond to minimal vertex covers of G, covers such that no proper subset is also a cover.*

Proof. Suppose w is a cover. Then for every edge e, $w \cap e \neq \varnothing$, so $w \in \mathfrak{p}_e$. Conversely, suppose $w \in I(G)^\vee$. Then, given any edge e, we have $w \in \mathfrak{p}_e$, i.e., $w \cap e \neq \varnothing$. In particular, w is a cover. □

We shall also need generalized Alexander duality for arbitrary monomial ideals. We follow Miller and Sturmfels's book [21], which is a good reference for this topic. Let $\mathbf{a}$ and $\mathbf{b}$ be vectors in $\mathbb{N}^n$ such that $b_i \le a_i$ for each i. As in [21, Definition 5.20], we define the vector $\mathbf{a} \setminus \mathbf{b}$ to be the vector whose ith entry is given by

$$a_i \setminus b_i = \begin{cases} a_i + 1 - b_i & \text{if } b_i \ge 1 \\ 0 & \text{if } b_i = 0. \end{cases}$$

Definition 2.8. Let $\mathbf{a} \in \mathbb{N}^n$, and let I be a monomial ideal such that all the minimal generators of I divide $\mathbf{x}^{\mathbf{a}}$. The *Alexander dual* of I *with respect to* $\mathbf{a}$ is the ideal

$$I^{[\mathbf{a}]} = \bigcap_{\mathbf{x}^{\mathbf{b}} \in \mathrm{gens}(I)} (x_1^{a_1 \setminus b_1}, \ldots, x_n^{a_n \setminus b_n}).$$

For square-free monomial ideals, one obtains the usual Alexander dual by taking $\mathbf{a}$ equal to $\mathbf{1}$, the vector with all entries 1, in Definition 2.8.

By Definition 2.6, Alexander duality identifies the minimal generators of a square-free ideal with the primes associated to its dual. The analogy for generalized Alexander duality identifies the minimal generators of a monomial ideal with the *irreducible components* of its dual.

Definition 2.9. A monomial ideal I is *irreducible* if it has the form $I = (x_1^{e_1}, \ldots, x_n^{e_n})$ for $e_i \in \mathbb{Z}_{>0} \cup \{\infty\}$. (We use the convention that $x_i^\infty = 0$.) Observe that the irreducible ideal I is $\mathfrak{p}$-primary, where $\mathfrak{p} = (x_i : e_i \neq \infty)$.

Definition 2.10. Let I be a monomial ideal. An *irreducible decomposition* of I is an irredundant decomposition

$$I = \bigcap Q_j$$

with the Q_j irreducible ideals. We call these Q_j *irreducible components* of I. By Corollary 2.12 below, there is no choice of decomposition, so the irreducible components are an invariant of the ideal.

Proposition 2.11. *Let I be a monomial ideal, and* $\mathbf{a}$ *be a vector with entries large enough that all the minimal generators of I divide* $\mathbf{x}^{\mathbf{a}}$. *Then* $(I^{[\mathbf{a}]})^{[\mathbf{a}]} = I$.

Corollary 2.12. *Every monomial ideal has a unique irreducible decomposition.*

A recurring idea in our paper is the difference between the powers and symbolic powers of square-free ideals. We recall the definition of the symbolic power.

For a square-free monomial ideal I, the *sth symbolic power* of I is

$$I^{(s)} = \bigcap_{\mathfrak{p} \in \mathrm{Ass}(S/I)} \mathfrak{p}^s.$$

(This definition works because square-free monomial ideals are the intersection of prime ideals. For general ideals (even general monomial ideals) the definition is more complicated.) In general we have $I^s \subseteq I^{(s)}$, but the precise nature of the relationship between the symbolic and ordinary powers of an ideal is a very active area of research.

In commutative algebra, symbolic and ordinary powers of an ideal are encoded in the symbolic Rees algebra and the ordinary Rees algebra. More specifically, for any ideal $I \subseteq S = K[x_1, \ldots, x_n]$, the *Rees algebra* and the *symbolic Rees algebra* of I are

$$\mathcal{R}(I) = \bigoplus_{q \geq 0} I^q t^q \subseteq S[t] \quad \text{and} \quad \mathcal{R}_s(I) = \bigoplus_{q \geq 0} I^{(q)} t^q \subseteq S[t].$$

The symbolic Rees algebra is closely related to the Rees algebra, but often is richer and more subtle to understand. For instance, while the Rees algebra of a homogeneous ideal is always Noetherian and finitely generated, the symbolic Rees algebra is not necessarily Noetherian. In fact, non-Noetherian symbolic Rees algebras were used to provide counterexamples to Hilbert's Fourteenth Problem (cf. [24, 26]).

3 Chromatic Number and Odd Cycles in Graphs

In this section, we examine how to detect simple graph-theoretic properties of a hypergraph G from (powers of) its edge and cover ideals. Since the results in this section involving chromatic number are the same for graphs as for hypergraphs, modulo some essentially content-free extra notation, we encourage novice readers to ignore the hypergraph case and think of G as a graph.

Definition 3.1. Let k be a positive integer. A *k-coloring* of G is an assignment of colors $c_1, \ldots, c_k$ to the vertices of G in such a way that every edge of cardinality at least 2 contains vertices with different colors. We say that G is *k-colorable* if a k-coloring of G exists, and that the *chromatic number* $\chi(G)$ of G is the least k such that G is k-colorable.

Remark 3.2. Since loops do not contain two vertices, they cannot contain two vertices of different colors. Thus the definition above considers only edges with cardinality at least two. Furthermore, since the presence or absence of loops has no effect on the chromatic number of the graph, we will assume throughout this section that all edges have cardinality at least two.

Remark 3.3. For hypergraphs, some texts instead define a coloring of G to be an assignment of colors to the vertices such that no edge contains two vertices of the same color. However, this is equivalent to a coloring of the one-skeleton of G, so the definition above allows us to address a broader class of problems.

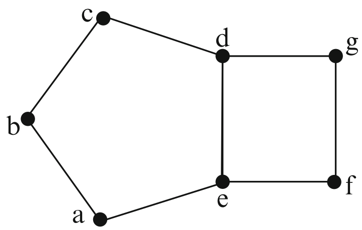


Fig. 1 The graph G in the running example

Running Example 3.4. Let G be the graph obtained by gluing a pentagon to a square along one edge, shown in Fig. 1. The edge ideal of G is $I(G) = (ab, bc, cd, de, ae, ef, fg, dg)$. The chromatic number of G is 3: for example, we may color vertices a, c, and g red, vertices b, d, and f yellow, and vertex e blue.

The chromatic number of G can be determined from the solutions to either of two different ideal-membership problems.

Observe that a graph fails to be k-colorable if and only if every assignment of colors to its vertices yields at least one single-colored edge. Thus, it suffices to test every color-assignment simultaneously. To that end, let $Y_1, \dots, Y_k$ be distinct copies of the vertices: $Y_i = \{y_{i,1}, \dots, y_{i,n}\}$. We think of Y_i as the ith color and the vertices of Y_i as being colored with this color. Now let $I(Y_i)$ be the edge ideal $I = I(G)$, but in the variables Y_i instead of V. Now an assignment of colors to G corresponds to a choice, for each vertex x_j, of a colored vertex $y_{i,j}$; or, equivalently, a monomial of the form $y_{i_1,1}y_{i_2,2}\dots y_{i_n,n}$. This monomial is a coloring if and only if it is not contained in the monomial ideal $\widetilde{I} = I(Y_1) + \cdots + I(Y_k)$. In particular, G is k-colorable if and only if the sum of all such monomials is not contained in $\widetilde{I}$.

We need some more notation to make the preceding discussion into a clean statement. Let $\mathbf{m} = x_1 \dots x_n$, let $T_k = K[Y_1, \dots, Y_k]$, and let $\phi_k : S \to T_k$ be the homomorphism sending x_i to $y_{1,i} + \cdots + y_{k,i}$. Then $\phi_k(\mathbf{m})$ is the sum of all color-assignments, and we have shown the following:

Lemma 3.5. With notation as above, G is k-colorable if and only if $\phi_k(\mathbf{m}) \notin \widetilde{I}$.

We recall the definition of the kth secant ideal. Secant varieties are common in algebraic geometry, including in many recent papers of Catalisano, Geramita, and Gimigliano (e.g., [5]), and, as Sturmfels and Sullivant note in [28], are playing an important role in algebraic statistics.

Definition 3.6. Let $I \subset S$ be any ideal, and continue to use all the notation above. Put $T = K[V, Y_1, \dots, Y_k]$ and regard S and T_k as subrings of T. Then the *kth secant power* of I is

$$I^{\{k\}} = S \cap \left(\widetilde{I} + (\{x_i - \phi_k(x_i)\})\right).$$

Lemma 3.5 becomes the following theorem of Sturmfels and Sullivant [28]:

Theorem 3.7. *G is k-colorable if and only if $\mathbf{m} \notin I(G)^{\{k\}}$. In particular,*

$$\chi(G) = \min\{k \mid \mathbf{m} \notin I(G)^{\{k\}}\}.$$

Running Example 3.8. Let G and I be as in Example 3.4. Then $I^{\{1\}} = I$ and $I^{\{2\}} = (abcde)$ both contain the monomial $abcdefg$. However, $I^{\{3\}} = 0$. Thus G is 3-colorable but not 2-colorable.

Alternatively, we can characterize chromatic number by looking directly at the powers of the cover ideal.

Observe that, given a k-coloring of G, the set of vertices which are not colored with any one fixed color forms a vertex cover of G. In particular, a k-coloring yields k different vertex covers, with each vertex missing from exactly one. That is, if we denote these vertex covers $w_1, \dots, w_k$, we have $w_1 \dots w_k = \mathbf{m}^{k-1}$. In particular, we have the following result of Francisco, Hà, and Van Tuyl. [11].

Theorem 3.9. *G is k-colorable if and only if $\mathbf{m}^{k-1} \in J(G)^k$. In particular,*

$$\chi(G) = \min\{k \mid \mathbf{m}^{k-1} \in J(G)^k\}.$$

Proof. Let $J = J(G)$. Given a k-coloring, let w_i be the set of vertices assigned a color other than i. Then $\mathbf{m}^{k-1} = w_1 \dots w_k \in J^k$. Conversely, if $\mathbf{m}^{k-1} \in J^k$, we may write $\mathbf{m}^{k-1} = w_1 \dots w_k$ with each w_i a square-free monomial in J. Assigning the color i to the complement of w_i yields a k-coloring: indeed, we have $\prod \frac{\mathbf{m}}{w_i} = \frac{\mathbf{m}^k}{\mathbf{m}^{k-1}} = \mathbf{m}$, so the $\frac{\mathbf{m}}{w_i}$ partition V. □

Running Example 3.10. In Example 3.4, let $\mathbf{m} = abcdefg$. The cover ideal $J(G)$ is $(abdf, acdf, bdef, aceg, bceg, bdeg)$. Because J does not contain $\mathbf{m}^0 = 1$, G is not 1-colorable. All 21 generators of J^2 are divisible by the square of a variable, so G is not 2-colorable. Thus $\mathbf{m} \notin J^2$, so J is not 2-colorable. However, J^3 contains $\mathbf{m}^2$, so G is 3-colorable.

Remark 3.11. One can adapt the proof of Theorem 3.9 to determine the b-fold chromatic number of a graph, the minimum number of colors required when each vertex is assigned b colors, and adjacent vertices must have disjoint color sets. See [11, Theorem 3.6].

Remark 3.12. The ideal membership problems in Theorems 3.7 and 3.9 are for monomial ideals, and so they are computationally simple. On the other hand, computing the chromatic number is an **NP**-complete problem. The bottleneck in the algebraic algorithms derived from Theorems 3.7 and 3.9 is the computation of the secant ideal $I(G)^{\{k\}}$ or the cover ideal $J(G)$ given G; these problems are both **NP**-complete.

It is naturally interesting to investigate the following problem.

Problem 3.13. Find algebraic algorithms to compute the chromatic number $\chi(G)$ based on algebraic invariants and properties of the edge ideal $I(G)$.

For the rest of this section, we shall restrict our attention to the case when G is a graph (i.e., not a hypergraph), and consider the problem of identifying odd cycles and odd holes in G. As before, let $I = I(G)$ and $J = J(G)$.

Recall that a *bipartite graph* is a two-colorable graph, or, equivalently, a graph with no odd circuits. This yields two corollaries to Theorem 3.9:

Corollary 3.14. *G is a bipartite graph if and only if* $\mathbf{m} \in J^2$.

Corollary 3.15. *If G is a graph, then G contains an odd circuit if and only if* $\mathbf{m} \notin J^2$.

It is natural to ask if we can locate the offending odd circuits. In fact, we can identify the *odd induced cycles* from the associated primes of J^2.

Definition 3.16. Let $C = (x_{i_1}, \dots, x_{i_s}, x_{i_1})$ be a circuit in G. We say that C is an *induced cycle* if the induced subgraph of G on $W = \{x_{i_1}, \dots, x_{i_s}\}$ has no edges except those connecting consecutive vertices of C. Equivalently, C is an induced cycle if it has no chords.

Running Example 3.17. G has induced cycles $abcde$ and $defg$. The circuit $abcdgfe$ isn't an induced cycle, since it has the chord de.

Simis and Ulrich prove that the odd induced cycles are the generators of the second secant ideal of I [27].

Theorem 3.18. *Let G be a graph with edge ideal I. Then a square-free monomial m is a generator of $I^{\{2\}}$ if and only if G_m is an odd induced cycle.*

Sketch of proof. If G_m is an odd induced cycle, then G_m and hence G are not 2-colorable. On the other hand, if $m \in I^{\{2\}}$, then G_m is not 2-colorable and so has an odd induced cycle. □

Now suppose that G is a cycle on $(2\ell - 1)$ vertices, so without loss of generality $I = (x_1x_2, x_2x_3, \dots, x_{2\ell-1}x_1)$. Then the generators of J include the $(2\ell-1)$ vertex covers $w_i = x_i x_{i+2} x_{i+4} \dots x_{i+2\ell-2}$ obtained by starting anywhere in the cycle and taking every second vertex until we wrap around to an adjacent vertex. (Here we have taken the subscripts mod $(2\ell - 1)$ for notational sanity.) All other generators have higher degree. In particular, the generators of J all have degree at least ℓ, so the generators of J^2 have degree at least 2ℓ. Thus $\mathbf{m} \notin J^2$, since $\deg(\mathbf{m}) = 2\ell - 1$. However, we have $\mathbf{m}x_i = w_i w_{i+1} \in J^2$ for all x_i. Thus $\mathbf{m}$ is in the socle of S/J^2, and in particular this socle is nonempty, so $\mathfrak{p}_{\mathbf{m}} = (x_1, \dots, x_{2\ell-1})$ is associated to J^2. In fact, it is a moderately difficult computation to find an irredundant primary decomposition:

Proposition 3.19. *Let G be the odd cycle on $x_1, \dots, x_{2\ell-1}$. Then*

$$J^2 = \left[\bigcap_{i=1}^{2\ell-1} (x_i, x_{i+1})^2 \right] \cap (x_1^2, \dots, x_{2\ell-1}^2).$$

Remark 3.20. Proposition 3.19 picks out the difference between J^2 and the symbolic square $J^{(2)}$ when G is an odd cycle. The product of the variables $\mathbf{m}$ appears in $\mathfrak{p}^2$ for all $\mathfrak{p} \in \mathrm{Ass}(S/J)$, but is missing from J^2. (Combinatorially, this corresponds to $\mathbf{m}$ being a double cover of G that cannot be partitioned into two single covers.) Thus $\mathbf{m} \in J^{(2)} \setminus J^2$.

Remark 3.21. We can attempt a similar analysis on an even cycle, but we find only two smallest vertex covers, $w_{\mathrm{odd}} = x_1 \dots x_{2\ell-1}$ and $w_{\mathrm{even}} = x_2 \dots x_{2\ell}$. Then $\mathbf{m} = w_{\mathrm{odd}} w_{\mathrm{even}} \in J^2$ is not a socle element. In this case Theorem 3.22 will tell us that J^2 has primary decomposition $\bigcap (x_i, x_{i+1})^2$, i.e., $J^{(2)} = J^2$.

In fact, Francisco, Hà, and Van Tuyl show that, for an arbitrary graph G, the odd cycles can be read off from the associated primes of J^2 [9]. Given a set $W \subset V$, put $\mathfrak{p}_W^{\langle 2\rangle} = (x_i^2 : x_i \in W)$. Then we have:

Theorem 3.22. *Let G be a graph. Then J^2 has irredundant primary decomposition*

$$J^2 = \left[\bigcap_{e \in E(G)} \mathfrak{p}_e^2 \right] \cap \left[\bigcap_{G_W \text{ is an induced odd cycle}} \mathfrak{p}_W^{\langle 2\rangle} \right].$$

Corollary 3.23. *Let G be a graph. Then we have*

$$\mathrm{Ass}(S/J^2) = \{\mathfrak{p}_e : e \in E(G)\} \cup \{\mathfrak{p}_W : G_W \text{ is an induced odd cycle}\}.$$

Corollary 3.23 and Theorem 3.18 are also connected via work of Sturmfels and Sullivant [28], who show that generalized Alexander duality connects the secant powers of an ideal with the powers of its dual.

Running Example 3.24. We have $\mathrm{Ass}(S/J^2) = E(G) \cup \{(a,b,c,d,e)\}$. The prime (a,b,c,d,e) appears here because $abcde$ is an odd induced cycle of G. The even induced cycle $defg$ does not appear in $\mathrm{Ass}(S/J^2)$, nor does the odd circuit $abcdgfe$, which is not induced. Furthermore, per Theorem 3.18, $I^{\{2\}}$ is generated by the odd cycle $abcde$.

Theorem 3.22 and Corollary 3.23 tell us that the odd cycles of a graph G exactly describe the difference between the symbolic square and ordinary square of its cover ideal $J(G)$. It is natural to ask about hypergraph-theoretic interpretations of the differences between higher symbolic and ordinary powers of $J(G)$, and of the differences between these powers for the edge ideal $I(G)$. The answer to the former question involves *critical hypergraphs*, discussed in Sect. 4. The latter question is closely related to a problem in combinatorial optimization theory. We describe this relationship in Sect. 5.

The importance of detecting odd induced cycles in a graph is apparent in the Strong Perfect Graph Theorem, proven by Chudnovsky, Robertson, Seymour, and Thomas in [6] after the conjecture had been open for over 40 years. A graph G is *perfect* if for each induced subgraph H of G, the chromatic number $\chi(H)$ equals the clique number $\omega(H)$, where $\omega(H)$ is the number of vertices in the largest clique (i.e., complete subgraph) appearing in H. Perfect graphs are an especially important class of graphs, and they have a relatively simple characterization. Call any odd cycle of at least five vertices an *odd hole*, and define an *odd antihole* to be the complement of an odd hole.

Theorem 3.25 (Strong Perfect Graph Theorem). *A graph is perfect if and only if it contains no odd holes or odd antiholes.*

Let G be a graph with complementary graph G^c (i.e., G^c has the same vertex set as G but the complementary set of edges). Let $J(G)$ be the cover ideal of G and $J(G^c)$ be the cover ideal of G^c. Using the Strong Perfect Graph Theorem along with Corollary 3.23, we conclude that a graph G is perfect if and only if neither $S/J(G)^2$ nor $S/J(G^c)^2$ has an associated prime of height larger than three. It is clear from the induced pentagon that the graph from Running Example 3.4 is imperfect; this is apparent algebraically from the fact that (a, b, c, d, e) is associated to $R/J(G)^2$.

4 Associated Primes and Perfect Graphs

Theorem 3.22 and Corollary 3.23 exhibit a strong interplay between coloring properties of a graph and associated primes of the square of its cover ideal. In this section, we explore the connection between coloring properties of hypergraphs in general and associated primes of higher powers of their cover ideals. We also specialize back to graphs and give algebraic characterizations of perfect graphs.

Definition 4.1. A *critically d-chromatic hypergraph* is a hypergraph G with $\chi(G) = d$ whose proper induced subgraphs all have smaller chromatic number; G is also called a *critical hypergraph*.

The connection between critical hypergraphs and associated primes begins with a theorem of Sturmfels and Sullivant on graphs that generalizes naturally to hypergraphs.

Theorem 4.2. *Let G be a hypergraph with edge ideal I. Then the square-free minimal generators of $I^{\{s\}}$ are the monomials W such that G_W is critically $(s+1)$-chromatic.*

Higher powers of the cover ideal $J = J(G)$ of a hypergraph have more complicated structure than the square. It is known that the primes associated to S/J^2 persist as associated primes of all S/J^s for $s \geq 2$ [11, Corollary 4.7]. As one might expect from the case of J^2, if H is a critically $(d+1)$-chromatic induced subhypergraph of G, then $\mathfrak{p}_H \in \operatorname{Ass}(S/J^d)$ but $\mathfrak{p}_H \notin \operatorname{Ass}(S/J^e)$ for any $e < d$. However, the following example from [11] illustrates that other associated primes may arise as well.

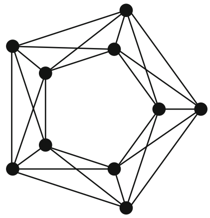


Fig. 2 The second expansion graph of a 5-cycle

Example 4.3. Let G be the graph with vertices $\{x_1, \ldots, x_6\}$ and edges

$$x_1x_2, x_2x_3, x_3x_4, x_4x_5, x_5x_1, x_3x_6, x_4x_6, x_5x_6,$$

where we have abused notation by writing edges as monomials. Thus G is a five-cycle on $\{x_1, \ldots, x_5\}$ with an extra vertex x_6 joined to x_3, x_4, and x_5. Let J be the cover ideal of G. The maximal ideal $\mathfrak{m} = (x_1, \ldots, x_6)$ is associated to S/J^3 but to neither S/J nor S/J^2. However, G is not a critically 4-chromatic graph; instead, $\chi(G) = 3$.

Consequently, the critical induced subhypergraphs of a hypergraph G may not detect all associated primes of S/J^s. Fortunately, there is a related hypergraph whose critical induced subhypergraphs do yield a complete list of associated primes. We define the expansion of a hypergraph, the crucial tool.

Definition 4.4. Let G be a hypergraph with vertices $V = \{x_1, \ldots, x_n\}$ and edges E, and let s be a positive integer. We create a new hypergraph G^s, called *the sth expansion of* G, as follows. We create vertex sets $V_1 = \{x_{1,1}, \ldots, x_{n,1}\}$, ..., $V_s = \{x_{1,s}, \ldots, x_{n,s}\}$. (We think of these vertex sets as having distinct flavors. In the literature, the different flavors $x_{i,j}$ of a vertex x_i are sometimes referred to as its *shadows*.) The edges of G^s consist of all edges $x_{i,j}x_{i,k}$ connecting all differently flavored versions of the same vertex, and all edges arising from possible assignments of flavors to the vertices in an edge of G.

We refer to the map sending all flavors $x_{i,j}$ of a vertex x_i back to x_i as *depolarization*, by analogy with the algebraic process of polarization.

Example 4.5. Consider a five-cycle G with vertices $x_1, \ldots, x_5$. Then G^2 has vertex set $\{x_{1,1}, x_{1,2}, \ldots, x_{5,1}, x_{5,2}\}$. Its edge set consists of edges $x_{1,1}x_{1,2}, \ldots, x_{5,1}x_{5,2}$ as well as all edges $x_{i,j}x_{i+1,j'}$, where $1 \leq j \leq j' \leq 2$, and the first index is taken modulo 5. Thus, for example, the edge x_1x_2 of G yields the four edges $x_{1,1}x_{2,1}$, $x_{1,1}x_{2,2}$, $x_{1,2}x_{2,1}$, and $x_{1,2}x_{2,2}$ in G^2 (Fig. 2).

Our goal is to understand the minimal monomial generators of the generalized Alexander dual $(J(G)^s)^{[\mathbf{s}]}$, where $\mathbf{s}$ is the vector $(s, \ldots, s)$, one entry for each vertex of G. Under generalized Alexander duality, these correspond to the ideals in an irredundant irreducible decomposition of $J(G)^s$, yielding the associated primes of $S/J(G)^s$.

By generalized Alexander duality, Theorem 4.2 identifies the square-free minimal monomial generators of $(J(G)^s)^{[\mathbf{s}]}$. Understanding the remaining monomial generators requires the following theorem [11, Theorem 4.4]. For a set of vertices T, write $\mathbf{m}_T$ to denote the product of the corresponding variables.

Theorem 4.6. *Let G be a hypergraph with cover ideal $J = J(G)$, and let s be a positive integer. Then*

$$(J^s)^{[\mathbf{s}]} = (\overline{\mathbf{m}_T} \mid \chi(G_T^s) > s)$$

where $\overline{\mathbf{m}_T}$ is the depolarization of $\mathbf{m}_T$.

The proof relies on a (hyper)graph-theoretic characterization of the generators of $I(G^s)^{\{s\}}$ from Theorem 4.2. One then needs to prove that $(J^s)^{[\mathbf{s}]}$ is the depolarization of $I(G^s)^{\{s\}}$, which requires some effort; see [11].

Using Theorem 4.6, we can identify all associated primes of $S/J(G)^s$ in terms of the expansion graph of G.

Corollary 4.7. *Let G be a hypergraph with cover ideal $J = J(G)$. Then $P = (x_{i_1}, \ldots, x_{i_r}) \in \operatorname{Ass}(S/J^s)$ if and only if there is a subset T of the vertices of G^s such that G_T^s is critically $(s+1)$-chromatic, and T contains at least one flavor of each variable in P but no flavors of other variables.*

We outline the rough idea of the proof. If $P \in \operatorname{Ass}(S/J^s)$, then $(x_{i_1}^{e_{i_1}}, \ldots, x_{i_r}^{e_{i_r}})$ is an irreducible component of J^s, for some $e_{i_j} > 0$. This yields a corresponding minimal generator of $(J^s)^{[\mathbf{s}]}$, which gives a subset W of the vertices of G^s such that G_W^s is critically $(s+1)$-chromatic, and W depolarizes to $x_{i_1}^{e_{i_1}} \ldots x_{i_r}^{e_{i_r}}$. Conversely, given a critically $(s+1)$-chromatic expansion hypergraph G_T^s, we get a minimal generator of $(J^s)^{[\mathbf{s}]}$ of the form $x_{i_1}^{e_{i_1}} \cdots x_{i_r}^{e_{i_r}}$, where $1 \leq e_{i_j} \leq s$ for all i_j. Duality produces an irreducible component of J^s with radical P.

Corollary 4.7 explains why $\mathfrak{m} \in \operatorname{Ass}(S/J^3)$ in Example 4.3. Let T be the set of vertices

$$T = \{x_{1,1}, x_{2,1}, x_{2,2}, x_{3,1}, x_{4,1}, x_{5,1}, x_{6,1}\},$$

a subset of the vertices of G^3. Then G_T^3 is critically 4-chromatic.

As a consequence of this work, after specializing to graphs, we get two algebraic characterizations of perfect graphs that are independent of the Strong Perfect Graph Theorem. First, we define a property that few ideals satisfy (see, e.g., [18]).

Definition 4.8. An ideal $I \subset S$ has the *saturated chain property for associated primes* if given any associated prime P of S/I that is not minimal, there exists an associated prime $Q \subsetneq P$ with $\text{height}(Q) = \text{height}(P) - 1$.

We can now characterize perfect graphs algebraically in two different ways [11, Theorem 5.9]. The key point is that for perfect graphs, the associated primes of powers of the cover ideal correspond exactly to the cliques in the graph.

Theorem 4.9. *Let G be a simple graph with cover ideal J. Then the following are equivalent:*

(1) *G is perfect.*
(2) *For all s with* $1 \leq s < \chi(G)$*,* $P = (x_{i_1}, \ldots, x_{i_r}) \in \operatorname{Ass}(R/J^s)$ *if and only if the induced graph on* $\{x_{i_1}, \ldots, x_{i_r}\}$ *is a clique of size* $1 < r \leq s+1$ *in* G*.*
(3) *For all* $s \geq 1$*,* J^s *has the saturated chain property for associated primes.*

Proof. We sketch (1) implies (2) to give an idea of how expansion is used. Suppose G is a perfect graph. A standard result in graph theory shows that G^s is also perfect. Let $P \in \operatorname{Ass}(S/J^s)$, so P corresponds to some subset T of the vertices of G^s such that G^s_T is critically $(s+1)$-chromatic. Because G^s is perfect, the clique number of G^s_T is also $s+1$, meaning there exists a subset T' of T such that $G^s_{T'}$ is a clique with $s+1$ vertices. Thus $G^s_{T'}$ is also a critically $(s+1)$-chromatic graph contained inside G^s_T, forcing $T = T'$. Hence G^s_T is a clique, and the support of the depolarization of $\bar{\mathbf{m}}_T$ is a clique with at most $s+1$ vertices. Therefore G_P is a clique. □

Remark 4.10. If J is the cover ideal of a perfect graph, its powers satisfy a condition stronger than that of Definition 4.8. If $P \in \operatorname{Ass}(S/J^s)$, and Q is any monomial prime of height at least two contained in P, then $Q \in \operatorname{Ass}(S/J^s)$. This follows from the fact that P corresponds to a clique in the graph.

Theorem 4.9 provides information about two classical issues surrounding associated primes of powers of ideals. Brodmann proved that for any ideal J, the set of associated primes of S/J^s stabilizes [3]. However, there are few good bounds in the literature for the power at which this stabilization occurs. When J is the cover ideal of a perfect graph, Theorem 4.9 demonstrates that stabilization occurs at $\chi(G) - 1$. Moreover, though in general associated primes may disappear and reappear as the power on J increases (see, e.g., [1,14] and also [23, Example 4.18]), when J is the cover ideal of a perfect graph, we have $\operatorname{Ass}(S/J^s) \subseteq \operatorname{Ass}(S/J^{s+1})$ for all $s \geq 1$. In this case, we say that J has the *persistence property for associated primes*, or simply the *persistence property*. Morey and Villarreal give an alternate proof of the persistence property for cover ideals of perfect graphs in [23, Example 4.21].

While there are examples of arbitrary monomial ideals for which persistence fails, we know of no such examples of *square-free* monomial ideals. Francisco, Hà, and Van Tuyl (see [9, 10]) have asked:

Question 4.11. Suppose J is a square-free monomial ideal. Is $\operatorname{Ass}(S/J^s) \subseteq \operatorname{Ass}(S/J^{s+1})$ for all $s \geq 1$?

While Question 4.11 has a positive answer when J is the cover ideal of a perfect graph, little is known for cover ideals of imperfect graphs. Francisco, Hà, and Van Tuyl answer Question 4.11 affirmatively for odd holes and odd antiholes in [10], but we are not aware of any other imperfect graphs whose cover ideals are known to have this persistence property. One possible approach is to exploit the machinery of expansion again. Let G be a graph, and let x_i be a vertex of G. Form the expansion of G at $\{x_i\}$ by replacing x_i with two vertices $x_{i,1}$ and $x_{i,2}$, joining them with an edge. For each edge $\{v, x_i\}$ of G, create edges $\{v, x_{i,1}\}$ and $\{v, x_{i,2}\}$. If W is any subset of the vertices of G, form $G[W]$ by expanding all the vertices of W. Francisco, Hà, and Van Tuyl conjecture:

Conjecture 4.12. *Let G be a graph that is critically s-chromatic. Then there exists a subset W of the vertices of G such that $G[W]$ is critically $(s+1)$-chromatic.*

In [10], Francisco, Hà, and Van Tuyl prove that if Conjecture 4.12 is true for all $s \geq 1$, then all cover ideals of graphs have the persistence property. One can also state a hypergraph version of Conjecture 4.12; if true, it would imply persistence of associated primes for all square-free monomial ideals.

Finally, in [23], Morey and Villarreal prove persistence for edge ideals I of any graphs containing a leaf (a vertex of degree 1). Their proof passes to the associated graded ring, and the vital step is identifying a regular element of the associated graded ring in I/I^2. Morey and Villarreal remark that attempts to prove persistence results for more general square-free monomial ideals lead naturally to questions related to the Conforti–Cornuéjols conjecture, discussed in the following section.

5 Equality of Symbolic and Ordinary Powers and Linear Programming

We have seen in the last section that comparing symbolic and ordinary powers of the cover ideal of a hypergraph allows us to study structures and coloring properties of the hypergraph. In this section, we address the question of when symbolic and ordinary powers of a square-free monomial ideal are the same and explore an algebraic approach to a long-standing conjecture in linear integer programming, the Conforti–Cornuéjols conjecture. In what follows, we state the Conforti–Cornuéjols conjecture in its original form, describe how to translate the conjecture into algebraic language, and discuss its algebraic reformulation and related problems.

The Conforti–Cornuéjols conjecture states the equivalence between the packing and the max-flow-min-cut (MFMC) properties for *clutters* which, as noted before, are essentially simple hypergraphs.

As before, $G = (V, E)$ denotes a hypergraph with n vertices $V = \{x_1, \ldots, x_n\}$ and m edges $E = \{e_1, \ldots, e_m\}$. Let A be the *incidence matrix* of G, i.e., the (i, j)-entry of A is 1 if the vertex x_i belongs to the edge e_j and 0 otherwise. For a nonnegative integral vector $\mathbf{c} \in \mathbb{Z}^n_{\geq 0}$, consider the following dual linear programming system:

$$\max\left\{\langle \mathbf{1}, \mathbf{y}\rangle \mid \mathbf{y} \in \mathbb{R}^m_{\geq 0}, A\mathbf{y} \leq \mathbf{c}\right\} = \min\left\{\langle \mathbf{c}, \mathbf{z}\rangle \mid \mathbf{z} \in \mathbb{R}^n_{\geq 0}, A^{\mathrm{T}}\mathbf{z} \geq \mathbf{1}\right\}. \quad (1)$$

Definition 5.1. Let G be a simple hypergraph.

(1) The hypergraph G is said to *pack* if the dual system (1) has integral optimal solutions $\mathbf{y}$ and $\mathbf{z}$ when $\mathbf{c} = \mathbf{1}$.

(2) The hypergraph G is said to have the *packing property* if the dual system (1) has integral optimal solutions $\mathbf{y}$ and $\mathbf{z}$ for all vectors $\mathbf{c}$ with components equal to 0, 1, and $+\infty$.
(3) The hypergraph G is said to have the *MFMC property* or to be *Mengerian* if the dual system (1) has integral optimal solutions $\mathbf{y}$ and $\mathbf{z}$ for all nonnegative integral vectors $\mathbf{c} \in \mathbb{Z}^n_{\geq 0}$.

Remark 5.2. In Definition 5.1, setting an entry of $\mathbf{c}$ to $+\infty$ means that this entry is sufficiently large, so the corresponding inequality in the system $A\mathbf{y} \leq \mathbf{c}$ can be omitted. It is clear that if G satisfies the MFMC property, then it has the packing property.

The following conjecture was stated in [7, Conjecture 1.6] with a reward prize of \$5,000 for the solution.

Conjecture 5.3 (*Conforti–Cornuéjols*). *A hypergraph has the packing property if and only if it has the MFMC property.*

As we have remarked, the main point of Conjecture 5.3 is to show that if a hypergraph has the packing property then it also has the MFMC property.

The packing property can be understood via more familiar concepts in (hyper) graph theory, namely, *vertex covers* (also referred to as *transversals*), which we recall from Sect. 1, and *matchings*.

Definition 5.4. A *matching* (or *independent set*) of a hypergraph G is a set of pairwise disjoint edges.

Let $\alpha_0(G)$ and $\beta_1(G)$ denote the minimum cardinality of a vertex cover and the maximum cardinality of a matching in G, respectively. We have $\alpha_0(G) \geq \beta_1(G)$ since every edge in any matching must hit at least one vertex from every cover.

The hypergraph G is said to be *König* if $\alpha_0(G) = \beta_1(G)$. Observe that giving a vertex cover and a matching of equal size for G can be viewed as giving integral solutions to the dual system (1) when $\mathbf{c} = \mathbf{1}$. Thus, G is König if and only if G packs.

There are two operations commonly used on a hypergraph G to produce new, related hypergraphs on smaller vertex sets. Let $x \in V$ be a vertex in G. The *deletion* $G \setminus x$ is formed by removing x from the vertex set and deleting any edge in G that contains x. The *contraction* G/x is obtained by removing x from the vertex set and removing x from any edge of G that contains x. Any hypergraph obtained from G by a sequence of deletions and contractions is called a *minor* of G. Observe that the deletion and contraction of a vertex x in G has the same effect as setting the corresponding component in $\mathbf{c}$ to $+\infty$ and 0, respectively, in the dual system (1). Hence,

G satisfies the packing property if and only if G and all of its minors are König.

Example 5.5. Let G be a 5-cycle. Then G itself is not König ($\alpha_0(G) = 3$ and $\beta_1(G) = 2$). Thus, G is does not satisfy the packing property.

Example 5.6. Any bipartite graph is König. Therefore, if G is a bipartite graph, then (since all its minors are also bipartite) G satisfies the packing property.

We shall now explore how Conjecture 5.3 can be understood via commutative algebra, and, more specifically, via algebraic properties of edge ideals.

As noted in Sect. 2, symbolic Rees algebras are more complicated than the ordinary Rees algebras, and could be non-Noetherian. Fortunately, in our situation, the symbolic Rees algebra of a square-free monomial ideal is always Noetherian and finitely generated (cf. [15, Theorem 3.2]). Moreover, the symbolic Rees algebra of the edge ideal of a hypergraph G can also be viewed as the *vertex cover algebra* of the dual hypergraph G^*.

Definition 5.7. Let $G = (V, E)$ be a simple hypergraph over the vertex set $V = \{x_1, \ldots, x_n\}$.

(1) We call a nonnegative integral vector $\mathbf{c} = (c_1, \ldots, c_n)$ a *k-cover* of G if $\sum_{x_i \in e} c_i \geq k$ for any edge e in G.
(2) The *vertex cover algebra* of G, denoted by $\mathcal{A}(G)$, is defined to be

$$\mathcal{A}(G) = \bigoplus_{k \geq 0} \mathcal{A}_k(G),$$

where $\mathcal{A}_k(G)$ is the k-vector space generated by all monomials $x_1^{c_1} \ldots x_n^{c_n} t^k$ such that $(c_1, \ldots, c_n) \in \mathbb{Z}^n_{\geq 0}$ is a k-cover of G.

Lemma 5.8. Let G be a simple hypergraph with edge ideal $I = I(G)$, and let G^* be its dual hypergraph. Then

$$\mathcal{R}_s(I) = \mathcal{A}(G^*).$$

We are now ready to give an algebraic interpretation of the MFMC property.

Lemma 5.9. Let $G = (V, E)$ be a simple hypergraph with n vertices and m edges. Let A be its incidence matrix. For a nonnegative integral vector $\mathbf{c} \in \mathbb{Z}^n_{\geq 0}$, define

$$\sigma(\mathbf{c}) = \max\{\langle \mathbf{1}, \mathbf{y} \rangle \mid \mathbf{y} \in \mathbb{Z}^m_{\geq 0}, A\mathbf{y} \leq \mathbf{c}\} \text{ and}$$
$$\gamma(\mathbf{c}) = \min\{\langle \mathbf{c}, \mathbf{z} \rangle \mid \mathbf{z} \in \mathbb{Z}^n_{\geq 0}, A^{\mathrm{T}}\mathbf{z} \geq \mathbf{1}\}.$$

Then

(1) $\mathbf{c}$ is a k-cover of G^* if and only if $k \leq \gamma(\mathbf{c})$.
(2) $\mathbf{c}$ can be written as a sum of k vertex covers of G^* if and only if $k \leq \sigma(\mathbf{c})$.

Proof. By definition, a nonnegative integral vector $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{Z}^n_{\geq 0}$ is a k-cover of G^* if and only if

$$k \leq \min \left\{ \sum_{x_i \in e} c_i \mid e \text{ is any edge of } G^* \right\}. \tag{2}$$

Let $\mathbf{z}$ be the (0, 1)-vector representing e. Observe that e is an edge of G^* if and only if e is a minimal vertex cover of G, and this is the case if and only if $A^{\mathrm{T}}\mathbf{z} \geq \mathbf{1}$. Therefore, the condition in (2) can be translated to

$$\begin{aligned} k &\leq \min\{\langle \mathbf{c}, \mathbf{z}\rangle \mid \mathbf{z} \in \{0,1\}^n, A^{\mathrm{T}}\mathbf{z} \geq \mathbf{1}\} \\ &= \min\{\langle \mathbf{c}, \mathbf{z}\rangle \mid \mathbf{z} \in \mathbb{Z}^n_{\geq 0}, A^{\mathrm{T}}\mathbf{z} \geq \mathbf{1}\} = \gamma(\mathbf{c}). \end{aligned}$$

To prove (2), let $\mathbf{a}_1, \ldots, \mathbf{a}_m$ be representing vectors of the edges in G (i.e., the columns of the incidence matrix A of G). By Proposition 2.7, $\mathbf{a}_1, \ldots, \mathbf{a}_m$ represent the minimal vertex cover of the dual hypergraph G^*. One can show that a nonnegative integral vector $\mathbf{c} \in \mathbb{Z}^n$ can be written as the sum of k vertex covers (not necessarily minimal) of G^* if and only if there exist integers $y_1, \ldots, y_m \geq 0$ such that $k = y_1 + \cdots + y_m$ and $y_1\mathbf{a}_1 + \cdots + y_m\mathbf{a}_m \leq \mathbf{c}$. Let $\mathbf{y} = (y_1, \ldots, y_m)$. Then

$$\langle \mathbf{1}, \mathbf{y}\rangle = y_1 + \cdots + y_m \text{ and } A\mathbf{y} = y_1\mathbf{a}_1 + \cdots + y_m\mathbf{a}_m.$$

Thus,

$$\sigma(\mathbf{c}) = \max\{k \mid \mathbf{c} \text{ can be written as a sum of } k \text{ vertex covers of } G^*\}. \qquad \square$$

Theorem 5.10. *Let G be a simple hypergraph with dual hypergraph G^*. Then the dual linear programming system (1) has integral optimal solutions $\mathbf{y}$ and $\mathbf{z}$ for all nonnegative integral vectors $\mathbf{c}$ if and only if $\mathcal{R}_s(I(G)) = \mathcal{A}(G^*)$ is a standard graded algebra or, equivalently, if and only if $I(G)^{(q)} = I(G)^q$ for all $q \geq 0$.*

Proof. Given integral optimal solutions $\mathbf{y}$ and $\mathbf{z}$ of the dual system (1) for a nonnegative integral vector $\mathbf{c}$, we get

$$\sigma(\mathbf{c}) = \gamma(\mathbf{c}).$$

The conclusion then follows from Lemmas 5.8 and 5.9. □

The following result (see [16, Corollary 1.6] and [12, Corollary 3.14]) gives an algebraic approach to Conjecture 5.3.

Theorem 5.11. *Let G be a simple hypergraph with edge ideal $I = I(G)$. The following conditions are equivalent:*

(1) G satisfies the MFMC property.
(2) $I^{(q)} = I^q$ for all $q \geq 0$.
(3) The associated graded ring $\mathrm{gr}_I := \bigoplus_{q\geq 0} I^q/I^{q+1}$ is reduced.
(4) I is normally torsion-free*, i.e., all powers of I have the same associated primes.*

Proof. The equivalence between (1) and (2) is the content of Theorem 5.10. The equivalences of (2), (3), and (4) are well-known results in commutative algebra (cf. [19]). □

The Conforti–Cornuéjols conjecture now can be restated as follows.

Conjecture 5.12. *Let G be a simple hypergraph with edge ideal $I = I(G)$. If G has packing property then the associated graded ring* gr_I *is reduced. Equivalently, if G and all its minors are König, then the associated graded ring* gr_I *is reduced.*

It remains to give an algebraic characterization for the packing property. To achieve this, we shall need to interpret minors and the König property. Observe that the deletion $G \setminus x$ at a vertex $x \in \mathcal{X}$ has the effect of setting $x = 0$ in $I(G)$ (or equivalently, of passing to the ideal $(I(G), x)/(x)$ in the quotient ring $S/(x)$), and the contraction G/x has the effect of setting $x = 1$ in $I(G)$ (or equivalently, of passing to the ideal $I(G)_x$ in the localization S_x). Thus, we call an ideal I' a *minor* of a square-free monomial ideal I if I' can be obtained from I by a sequence of taking quotients and localizations at the variables. Observe further that $\alpha_0(G) = \mathrm{ht} I(G)$, and if we let m-grade I denote the maximum length of a regular sequence of monomials in I then $\beta_1(G) =$ m-grade $I(G)$. Hence, a simple hypergraph with edge ideal I is König if $\mathrm{ht} I =$ m-grade I. This leads us to a complete algebraic reformulation of the Conforti–Cornuéjols conjecture:

Conjecture 5.13. *Let I be a square-free monomial ideal such that I and all of its minors satisfy the property that their heights are the same as their m-grades. Then* gr_I *is reduced; or equivalently, I is normally torsion-free.*

The algebraic consequence of the conclusion of Conjecture 5.13 (and equivalently, Conjecture 5.3) is the equality $I^{(q)} = I^q$ for all $q \geq 0$ or, equivalently, the normally torsion-freeness of I. If one is to consider the equality $I^{(q)} = I^q$, then it is natural to look for an integer l such that $I^{(q)} = I^q$ for $0 \leq q \leq l$ implies $I^{(q)} = I^q$ for all $q \geq 0$, or to examine square-free monomial ideals with the property that $I^{(q)} = I^q$ for all $q \geq q_0$. On the other hand, if one is to investigate the normally torsion-freeness then it is natural to study properties of minimally not normally torsion-free ideals. The following problem is naturally connected to Conjectures 5.3 and 5.13, and part of it has been the subject of work in commutative algebra (cf. [13]).

Problem 5.14. Let I be a square-free monomial ideal in $S = K[x_1, \ldots, x_n]$.

(1) Find the least integer l (may depend on I) such that if $I^{(q)} = I^q$ for $0 \leq q \leq l$ then $I^{(q)} = I^q$ for all $q \geq 0$.
(2) Suppose that there exists a positive integer q_0 such that $I^{(q)} = I^q$ for all $q \geq q_0$. Study algebraic and combinatorial properties of I.
(3) Suppose I is minimally not normally torsion-free (i.e., I is not normally torsion-free, but all its minors are). Find the least power q such that $\mathrm{Ass}(S/I^q) \neq \mathrm{Ass}(S/I)$.

Acknowledgements This work was partially supported by grants from the Simons Foundation (#199124 to Francisco and #202115 to Mermin). Hà is partially supported by NSA grant H98230-11-1-0165.

References

1. Bandari, S., Herzog, J., Hibi, T.: Monomial ideals whose depth function has any given number of strict local maxima. Preprint, 2012. `arXiv:1205.1348`
2. Bocci, C., Harbourne, B.: Comparing powers and symbolic powers of ideals. J. Algebraic Geom. **19**(3), 399–417 (2010)
3. Brodmann, M.: Asymptotic stability of Ass(M/I^nM). Proc. Amer. Math. Soc. **74**, 16–18 (1979)
4. Bruns, W., Herzog, J.: Cohen-Macaulay rings. Cambridge Studies in Advanced Mathematics, vol. 39. Cambridge University Press, Cambridge (1993)
5. Catalisano, M.V., Geramita, A.V., Gimigliano, A.: On the ideals of secant varieties to certain rational varieties. J. Algebra **319**(5), 1913–1931 (2008)
6. Chudnovsky, M., Robertson, N., Seymour, P., Thomas, R.: The strong perfect graph theorem. Ann. Math. (2) **164**, 51–229 (2006)
7. Cornuéjols, G.: Combinatorial optimization: packing and covering. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 74. SIAM, Philadelphia (2001)
8. Ein, L., Lazarsfeld, R., Smith, K.E.: Uniform bounds and symbolic powers on smooth varieties. Invent. Math. **144**(2), 241–252 (2001)
9. Francisco, C.A., Hà, H.T., Van Tuyl, A.: Associated primes of monomial ideals and odd holes in graphs. J. Algebraic Combin. **32**(2), 287–301 (2010)
10. Francisco, C.A., Hà, H.T., Van Tuyl, A.: A conjecture on critical graphs and connections to the persistence of associated primes. Discrete Math. **310**(15–16), 2176–2182 (2010)
11. Francisco, C.A., Hà, H.T., Van Tuyl, A.: Colorings of hypergraphs, perfect graphs, and associated primes of powers of monomial ideals. J. Algebra **331**, 224–242 (2011)
12. Gitler, I., Valencia, C.E., Villarreal, R.: A note on Rees algebras and the MFMC property. Beiträge Algebra Geom. **48**(1), 141–150 (2007)
13. Hà, H.T., Morey, S.: Embedded associated primes of powers of square-free monomial ideals. J. Pure Appl. Algebra **214**(4), 301–308 (2010)
14. Herzog, J., Hibi, T.: The depth of powers of an ideal. J. Algebra **291**, 534–550 (2005)
15. Herzog, J., Hibi, T., Trung, N.V.: Symbolic powers of monomial ideals and vertex cover algebras. Adv. Math. **210**(1), 304–322 (2007)
16. Herzog, J., Hibi, T., Trung, N.V., Zheng, X.: Standard graded vertex cover algebras, cycles and leaves. Trans. Amer. Math. Soc. **360**(12), 6231–6249 (2008)
17. Hochster, M., Huneke, C.: Comparison of symbolic and ordinary powers of ideals. Invent. Math. **147**(2), 349–369 (2002)
18. Hosten, S., Thomas, R.: The associated primes of initial ideals of lattice ideals. Math. Res. Lett. **6**(1), 83–97 (1999)
19. Huneke, C., Simis, A., Vasconcelos, W.: Reduced normal cones are domains. Invariant theory (Denton, TX, 1986), pp. 95–101. Contemporary Mathematics, vol. 88. American Mathematical Society, Providence (1989)
20. Huneke, C., Katz, D., Validashti, J.: Uniform equivalence of symbolic and adic topologies. Illinois J. Math. **53**(1), 325–338 (2009)
21. Miller, E., Sturmfels, B.: Combinatorial commutative algebra. Graduate Texts in Mathematics, vol. 227. Springer, New York (2004)
22. Minh, N.C., Trung, N.V.: Cohen-Macaulayness of monomial ideals and symbolic powers of Stanley–Reisner ideals. Adv. Math. **226**(2), 1285–1306 (2011)
23. Morey, S., Villarreal, R.H.: Edge ideals: algebraic and combinatorial properties. Progress in Commutative Algebra: Ring Theory, Homology, and Decomposition, pp. 85–126. de Gruyter, Berlin (2012)
24. Nagata, M.: On the 14th problem of Hilbert. Amer. J. Math. **81**, 766–772 (1959)
25. Peeva, I.: Graded syzygies. Algebra and Applications, vol. 14. Springer, London (2011)
26. Roberts, P.: A prime ideal in a polynomial ring whose symbolic blow-up is not Noetherian. Proc. Amer. Math. Soc. **94**, 589–592 (1985)

27. Simis, A., Ulrich, B.: On the ideal of an embedded join. J. Algebra **226**(1), 1–14 (2000)
28. Sturmfels, B., Sullivant, S.: Combinatorial secant varieties. Pure Appl. Math. Q. **2**(3), part 1, 867–891 (2006)
29. Terai, N., Trung, N.V.: Cohen–Macaulayness of large powers of Stanley–Reisner ideals. Adv. Math. **229**(2), 711–730 (2012)
30. Trung, N.V., Tuan, T.M.: Equality of ordinary and symbolic powers of Stanley-Reisner ideals. J. Algebra **328**, 77–93 (2011)
31. Varbaro, M.: Symbolic powers and matroids. Proc. Amer. Math. Soc. **139**(7), 2357–2366 (2011)

A Brief History of Order Ideals

E. Graham Evans and Phillip Griffith

The notion of *order ideal* is no doubt implicit in Serre's 1958 paper [41] on free summands of projective modules. A formal definition is given in Bass' fundamental article [4] on K-theory and stable algebra. However, any algebraist contemplating the question, "On what locus of prime ideals in Spec R does an element e in a module E generate a free summand?", has in fact encountered the concept of an order ideal. In the account on order ideals and their applications that follows, it is our intent to elaborate on four basic theorems—as we see them—that give insight into the height properties of these ideals. We do this both from a historical view and a view of their utility.

Definition 1. Let R be a Noetherian ring, let E be an R-module, and let $e \in E$. Then the order ideal of e in E is defined by the equation

$$O_E(e) = \{f(e) \in R | f : E \to R \text{ is a homomorphism}\}.$$

Defining $E^* = \text{Hom}(E, R)$, the "R-dual of E," we may express $O_E(e)$ as the image of the R-homomorphism $e : E^* \to R$ in which $e(f) = f(e)$. In fact the latter version is the point of view taken in algebraic geometry in describing the concept where $\text{Spec}(R/O_E(e))$ is referred to as the "zeros of the section e" (see Hartshorne [29, p. 431]). An investigation into the properties and behavior of order ideals in commutative algebra began in earnest in the 1970s and early 1980s. Most of the attention during this period was devoted to questions concerning upper and lower bounds for the height (or grade) of order ideals relative to rank or depth of the module in question. The usual starting point for such a discussion is to consider

E.G. Evans • P. Griffith (✉)
Department of Mathematics, University of Illinois at Urbana Champaign,
Urbana, IL 61801, USA
e-mail: pgriffit@illinois.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_12,

how the notion plays out in the case of free modules. To this end, we let $E = R^n$ and let $e = \langle r_1, \ldots, r_n \rangle \in E$. The order ideal in this case is easily computed to be

$$O_E(e) = (r_1, \ldots, r_n),$$

that is, $O_E(e)$ is simply the ideal generated by the coordinates of e relative to a (any) free basis for E. When the ideal $(r_1, \ldots, r_n)$ is a proper ideal, the Krull altitude theorem gives that its height is bounded above by n, that is, $\operatorname{ht} O_E(e) \leq n$. The first important order ideal theorem is a generalization of this observation and is the central focus of the 1976 article by Eisenbud and Evans [16]. One should consult Bruns [6] for a "characteristic-free" argument. In Sect. 1 we discuss the Eisenbud–Evans–Bruns result along with a recent improvement in the paper of Eisenbud–Huneke–Ulrich [17]. Going back to our free R-module example $E = R^n$ we also see that $e = \langle r_1, \ldots, r_n \rangle$ generates a free R-summand of E exactly when $O_E(e) = (r_1, \ldots r_n) = R$. In particular, our opening question, "On what locus of prime ideals in Spec R is e a free generator of E?", can now be framed from this viewpoint. The question simply put is: "For what prime ideals $\mathfrak{p}$ in Spec R is it true that $\mathfrak{p} \not\supseteq O_E(e)$?." Thus, one is left to compute the "size" of the Zariski open set Spec $R - V(O_E(e))$.

The above line of thought led the authors to establish a second key result on behavior of order ideals in the setting of local algebra and kth syzygy modules having finite projective dimension—at least when the local ring $(R, \mathfrak{m})$ contains a field. Specifically, it is shown in [22, Theorem 3.14]: if e is a minimal generator in E, i.e., $e \in E - \mathfrak{m}E$, and if pd $E < \infty$, then $\operatorname{ht} O_E(e) \geq k$. Two other principles that govern the height behavior of order ideals are also described in Sect. 1. One of these results is implicit in the work of Bruns [5] which guarantees that a kth syzygy of finite projective dimension and having rank $> k$ necessarily has a minimal generator e with $\operatorname{ht} O_E(e) > k$.

Also useful is the authors' result [20] for torsion-free modules over a local domain R which states: some minimal generator e of a torsion-free R-module E must have $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$, when E is not free. The residue field is required to be algebraically closed. In [18, Theorem 2.1] Eisenbud–Huneke–Ulrich remove the assumption of "algebraically closed" residue field by making a base change via a local homomorphism to a localized polynomial extension. In Sect. 2 we get an example of a module with a large set of minimal generators having order ideal the maximal ideal.

In Sect. 2 we take up several important applications of the main theorems outlined in Sect. 1. We devote much of our attention here to various aspects of our syzygy theorem which was first proved in [19]. This result provides lower bounds for the ranks of non-free syzygies having finite projective dimension. We sketch the relationship of this theorem with the *improved new intersection theorem*. In the more restrictive setting of regular local rings in Sect. 3 we explore a point of view that has its inception in the article of Eisenbud–Huneke–Ulrich [17] where the notion of *perpendicular element* is successfully utilized. Our main result here makes use of the Serre intersection theorem and shows that consecutive syzygies influence one another with respect to heights of order ideals and rank—in the setting of locally free modules on the punctured spectrum. In Sect. 4 we outline more recent advances in the special context of *mixed characteristic*. Several of these results fit under the general heading of *comparison theorems* where one encounters a homomorphism from the syzygy module being studied to one where *good* properties are known. Often, these homomorphisms arise when one restricts ones attention to a hypersurface. The manuscript is organized as follows:

1. Bounding heights of order ideals: the main theorems.
2. The syzygy theorem for modules of finite projective dimension and applications.
3. The Serre intersection theorem and order ideals of consecutive syzygy modules.
4. The state of mixed characteristic.
5. Acknowledgement.

For unexplained references and notation we refer the reader to our monograph [22] or Bruns–Herzog's excellent book [8]. One may consult the Hochster–Huneke article [34, Sect. 10] to see how some of this material may be approached from the point of view of *tight closure*.

1 Bounding Heights of Order Ideals: The Main Theorems

Our discussion in the introduction provides evidence that the rank of a module should play a role in determining an upper bound for the height of an ideal. Moreover, such an inequality might be viewed as a natural generalization of the Krull Altitude Theorem. The "first" order ideal theorem stated next makes this connection precise and was first established by Eisenbud–Evans [16] in case the ring contains a field. Their argument relied on exterior algebra and a comparison theorem now known as the "canonical element theorem" (see [7, pp. 358–361]). In 1981 Bruns [6] gave a more elementary and characteristic-free proof. Bruns' proof makes use of adjoining formal polynomial variables to the ring—enough to accommodate a generating set—and then reduces the problem in a clever way to the case of a free module over a polynomial ring.

Theorem 1 ([16, 1976], [6, 1981]). *Let R be a Noetherian ring, let E be an R-module, and let I be a proper ideal of R. If $e \in IE$, then* $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$.

The notion of "rank" in the statement of Theorem 1 is referred to as "big rank" (see [7, p. 360]). Nonetheless, the most important case is that of an integral domain especially since the general case reduces to this situation. In their 2004 article [17] Eisenbud–Huneke–Ulrich establish yet a more general statement which takes into account the behavior of order ideals after base change. In particular, they extend the above inequality to include elements in the integral closure of IE (see [17, Theorem 3.1 and Corollary 3.4]). In addition they determine invariants that give other upper bounds besides just rank.

In our work (discussed here in Sect. 2) related to lower bounds for syzygy modules having finite projective dimension, we encounter heights of order ideals in the following context.

Let $(R, \mathfrak{m})$ be a Noetherian local ring and let $\phi : R^n \to R^m$ be an R-homomorphism of free modules such that the associated matrix (r_{ij}) has all of its entries in the maximal ideal of R. The column vectors of ϕ represent the images of the respective standard basis vectors of R^n. Taken together these image vectors generate in a minimal way the image E of ϕ. The ideal generated by the coordinates of the jth column vector e^j of ϕ is contained in the order ideal $O_E(e^j)$, $1 \le j \le n$. The second order ideal theorem (Theorem 4) will provide a sharp lower bound for heights of order ideals of minimal generators of modules E having finite projective dimension. However, Theorem 1 can be used to establish a lower bound for the rank of ϕ in terms of heights of order ideals as follows.

Corollary 2. *Let $\phi : R^n \to R^m$ be an R-homomorphism and let $E = \operatorname{im}\phi \subseteq R^m$. Let $v \in \mathfrak{m}E$ and let I be the ideal generated by the coordinates of $\phi(v) \in R^m$. One has the inequalities*

$$\operatorname{ht} I \le \operatorname{ht} O_E(\phi(v)) \le \operatorname{rank} E = \operatorname{rank}\phi.$$

The above observation can be used to illustrate the difference in behavior of heights of order ideals when one compares order ideals of minimal generators with elements in $\mathfrak{m}E$. Let R be the 3-dimensional power series ring $R = k[[x, y, z]]$ in which k is a field. The 3×2 matrix

$$\phi = \begin{bmatrix} x & z \\ y & x \\ z & y \end{bmatrix}$$

defines an R-homomorphism $R^2 \to R^3$ of rank 2. The image of the basis vector

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ is } \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in E = \operatorname{im}\phi$$

has coordinate ideal (x, y, z) of height $3 > 2$ while any image of the vectors

$$u \begin{bmatrix} x \\ y \\ z \end{bmatrix} + v \begin{bmatrix} z \\ x \\ y \end{bmatrix}$$

with $u, v \in \mathfrak{m}$ will have order ideal of height ≤ 2 as predicted by Theorem 1. Further, one can plainly see that adding more rows of indeterminates to the matrix, e.g., making an "$n \times 2$" example, only makes the disparity of heights even greater.

Thus minimal generators in the local case generally do not obey the inequality of Theorem 1.

In order to facilitate our discussion surrounding heights of order ideals of minimal generators in local algebra we need to bring standard facts about (minimal) free resolutions into the picture.

Definition 3. We say that an R-module E is a kth syzygy provided E fits into an exact sequence

$$0 \to E \to R^{n_{k-1}} \to \cdots \to R^{n_1} \to R^{n_0} \to M \to 0 \qquad (*)$$

If E does fit into an exact sequence as in $(*)$ then we may extend the coordinate maps $f_1, \ldots, f_{n_{k-1}}$ (if necessary) to a generating set $f_1, \ldots, f_{n_{k-1}}, \ldots, f_q$ of $E^* = \mathrm{Hom}(E, R)$. In this way we obtain an embedding

$$0 \to E \to R^q \to W \to 0$$

in which the dual sequence with respect to R

$$0 \to W^* \to (R^q)^* \to E^* \to 0$$

is exact, i.e., $\mathrm{Ext}^1_R(W, R) = 0$. In this situation the order ideal of an element in E is exactly equal to the ideal generated by its coordinate representation in R^q. We may now repeat this process for the cokernel W of $E \to R^q$, provided W is "torsion free." In fact, if we begin with a kth syzygy E of finite projective dimension we are guaranteed that this process will continue until E is realized as a kth syzygy in a free resolution which has an exact dual (see [22, p. 49] for more details). Thus, under the assumption that E is a kth syzygy satisfying $\mathrm{pd}\, E < \infty$ we may assume that our original exact sequence $(*)$ has the property that its dual sequence

$$0 \leftarrow E^* \leftarrow (R^{n_{k-1}})^* \leftarrow \cdots \leftarrow (R^{n_1})^* \leftarrow (R^{n_0})^* \leftarrow M^* \leftarrow 0$$

is exact, and in particular represents the first k terms in a free resolution of E^*. One may properly view the preceding construction as a "universal" representation of the module E as a kth syzygy.

For ease of presentation we make the following underlying assumption on the ambient local ring, denoted by R. While this assumption is not necessary in order to state and discuss our second theorem on order ideals (e.g., see [7, section 9.5] or [34, section 10]), the condition makes our discussion flow more easily.

One can characterize kth syzygy modules of finite projective dimension as follows. We assume $\mathrm{pd}\, E < \infty$:

Throughout the remainder of this section when discussing kth syzygy modules E of finite projective dimension we assume that the ring R satisfies the Serre condition (S_k). (For definition see [22, p. 3].)

1. E is a first syzygy $\iff$ E satisfies the Serre condition (S_1) $\iff$ E can be embedded as a submodule of a free R-module (i.e., E is "torsion free").
2. E is a second syzygy $\iff$ E satisfies (S_2) $\iff$ E is a reflexive R-module.
3. E is a kth syzygy for $k \geq 3$ $\iff$ E satisfies (S_k) and $\operatorname{Ext}^j(E^*, R) = 0$ for $1 \leq j \leq k-2$.

We can rephrase the above conditions to be a kth syzygy in terms of the open subscheme $X = \operatorname{Spec} R - \{\mathfrak{m}\}$. Thinking of E as defining a coherent sheaf on $X = \operatorname{Spec} R - \{\mathfrak{m}\}$ and making the assumption $\dim R = d+1$ one has that the $\operatorname{Ext}^j(E^*, R)$ are dual to the sheaf cohomology $H^{d-j}(X, E^*)$. Thus one can rephrase the fact E is a kth syzygy module in terms of the cohomology of E^*. In the situation where E is a locally free sheaf on X, one can use the duality between $H^j(X, E)$ and $H^{d-j}(X, E^*)$ to phrase the condition in terms of E alone. Next we describe the fundamental connection between kth syzygies and order ideals.

Theorem 4 (Evans–Griffith [22, Theorem 3.14]). *Let $(R, \mathfrak{m})$ be a local ring that contains a field, and let E be a kth syzygy module without free summands. If $e \in E - \mathfrak{m}E$ then* $\operatorname{ht} O_E(e) \geq k$.

Proof. (Sketch) Our starting point is to let

$$0 \to E \to R^{n_{k-1}} \to \cdots \to R^{n_1} \to R^{n_0} \to M \to 0$$

be a universal presentation of E as a kth syzygy module and extend this sequence to

$$0 \to R^{n_d} \to R^{n_{d-1}} \to \cdots \to R^{n_k} \to R^{n_{k-1}} \to \cdots \to R^{n_0} \to M \to 0$$

a free resolution of M. We may assume that the maps $\phi_d, \ldots, \phi_k$ in the resolution have their entries in the maximal ideal $\mathfrak{m}$. Note if E were to have a nonzero free summand then the map $R^{n_k} \to R^{n_{k-1}}$ would have some entries consisting of units. Further we may assume that a basis is chosen for R^{n_k} so that the "first" basis vector v_1 is carried to e under the map $R^{n_k} \to R^{n_{k-1}}$. By way of contradiction we suppose that $\operatorname{ht} I < k$ where $I = O_E(e)$.

By [22, Theorem 1.11] there is a maximal Cohen–Macaulay R-module C over R/I so that $\operatorname{pd} C = \operatorname{ht} I < k$ and such that $\operatorname{Tor}_k^R(C, M) = 0$. However, after tensoring the above free resolution with C, we obtain a complex

$$0 \to C^{n_d} \to \cdots \to C^{n_k} \to C^{n_{k-1}} \to \cdots \to C^{n_0}$$

which must be exact at C^{n_k}. But elements of the form $v_1 \otimes c$ are sent to $IC = 0$; thus the $v_1 \otimes c$ must necessarily be in the image of $C^{n_{k+1}}$. However this implies that $C = \mathfrak{m}C$ since the entries of ϕ_k are in $\mathfrak{m}$. This statement contradicts the fact that C is a maximal Cohen–Macaulay module. ☐

A simpler version of the above argument where the "intersecting" module C is replaced by a cyclic module gives the following result (see [23, pp. 486, 487] for details):

Theorem 4* **(Evans–Griffith [23, Theorem A]).** *Let R be a local ring and let E be a kth syzygy module of finite projective dimension. Let $I \neq 0$ be an ideal such that* $\operatorname{pd} R/I < k$ *and let $e \in E$. Then*

(i) The order ideal $O_E(e)$ cannot equal I.
(ii) If $e \in E - \mathfrak{m}E$ then $O_E(e)$ is not contained in I.

The conclusions reached in Theorem 4* have their clearest implication in the situation where R/I is Cohen–Macaulay of projective dimension $l < k$. In this context we may restate them to say, if $Y \subseteq X$ is a Cohen–Macaulay subvariety of codimension l, then no section of a kth syzygy has zeros which are ideal-theoretically Y, and further no section which is a minimal generator has zeros completely inside Y.

We should also mention other proofs of Theorem 4 (or slight variations). We recorded a characteristic p proof of Huneke in our 1987 article [25, p. 220], and this argument was repeated in Hochster–Huneke [34, section 10]. Their treatment applies their extensive theory of "tight closure." Bruns–Herzog also present a slight variation in Sect. 9.5 of their book [8] (see also Bruns [7]).

Our original argument for Theorem 4 (see [19, Proposition 1.6]) only proved a special case of the result—the case in which the kth syzygy module E is locally free on the subscheme $X = \operatorname{Spec} R - \{\mathfrak{m}\}$, i.e., the case where E represents a vector bundle on X. For our proof of the syzygy theorem (see Sect. 2) this was the only case needed since we had reduced the problem on rank to this setting. The outgrowth of this particular point of view spawned an "intersection theorem" that is a natural generalization of the "new intersection theorem" due to Peskine–Szpiro [39] and Roberts [40]. Hochster [33] referred to the new result as the "improved new intersection theorem" (we denote the name by "INIT"). The theorem can be stated:

INIT (Hochster [33] and [22, Theorem 1.13]). *Let $(A, \mathfrak{m})$ denote a local ring that contains a field and let $F_\bullet$ be a nontrivial free complex that satisfies:*

(i) The positive homology of $F_\bullet$ is supported only at $\{\mathfrak{m}\}$.
(ii) The homology module $H_0(F_\bullet)$ contains a minimal generator w supported at $\{\mathfrak{m}\}$.

Then $\operatorname{length}(F_\bullet) \geq \dim A$.

One applies INIT to the special case of Theorem 4 described above in the following way. Let $e \in E - \mathfrak{m}E$, let $I = O_E(e)$ and let $F_\bullet \to E$ be a minimal free resolution of E. We set $A = R/I$ and apply INIT to the free A-complex $F\bullet/IF\bullet = A \otimes F_\bullet$. Since E is locally free on $\operatorname{Spec} R - \{m\}$ the hypothesis of INIT is easily verified where $w = e + IE$. Thus one concludes

$$\operatorname{pd} E = \operatorname{length} A \otimes F_\bullet \geq \dim A,$$

and from here (using the Auslander–Buchsbaum formula [1]

$$\operatorname{depth} E + \operatorname{pd} E = \operatorname{depth} R$$

and the harmless assumption that R is complete) one gets the inequality $\operatorname{ht} I \geq \operatorname{depth} E$. However, $\operatorname{depth} E \geq k$ since E is locally free on $\operatorname{Spec} R - \{\mathfrak{m}\}$. Articles of Hochster [33], Dutta [13], and Ogoma [37] show that INIT is equivalent to the canonical element conjecture for general characteristic. For a thorough development and understanding of the relationship between the various "homological theorems/conjectures," one should consult Hochster's excellent articles [32, 33].

The third order ideal theorem from our perspective arises in a natural way out of a theorem of Bruns [5] (see also [22, Theorem 3.11]). We begin with a kth syzygy of finite projective dimension and consider a short exact sequence

$$0 \to R \to E \to E' \to 0 \qquad (***)$$

in which $1 \in R$ is sent to $e \in E - \mathfrak{m}E$. So $\operatorname{pd} E' < \infty$ as well. A basic question here is: "When is the module E' also a kth syzygy?" The answer is in fact quite easy. The module E' is a kth syzygy precisely when $(***)$ becomes split exact locally in codimension $\leq k$. Moreover, this property is clearly equivalent to $\operatorname{ht} O_E(e) > k$. Using "basic element" theory developed by Eisenbud–Evans [15], Bruns [5] proved: if E is a kth syzygy module of finite projective dimension and if $\operatorname{rank} E = k+s$ then E contains a free submodule F of rank s such that E/F is also a kth syzygy. It was this circle of ideas that led Bruns to contemplate the "syzygy theorem" (first noted by Hackman [28] in his Ph.D. thesis at University of Stockholm, 1969) discussed in Sect. 2. The third order ideal theorem is simply a recasting of Bruns' result along the lines of [22, Theorem 3.11].

Theorem 4 (Bruns [5]). *Let R be a local ring and let E be a kth syzygy module of finite projective dimension. If* $\operatorname{rank} E > k$ *then E must have a minimal generator e such that* $\operatorname{ht} O_E(e) > k$.

That one cannot improve on this result is the essence of our syzygy theorem in Sect. 2.

Now that we have established a condition under which we can locate minimal generators with larger than expected order ideals, one might ask if there are any restrictions that would indicate an upper bound should exist. As examples show one cannot expect too much in the way of a positive answer. For example in our monograph [22, Appendix], we describe a construction due to Horrocks and Mumford that shows the existence of a rank two-second syzygy E for R regular local of dimension 5 such that E is locally free on $\operatorname{Spec} R - \{\mathfrak{m}\}$. In addition $\operatorname{Ext}^1_R(E, R) \neq 0$. Thus, a nontrivial extension

$$0 \to R \to M \to E \to 0$$

gives rise to an element $m \in M - \mathfrak{m}M$ (the image of $1 \in R$) such that $O_M(m)$ is $\mathfrak{m}$ primary, i.e.,

$$\operatorname{ht} O_M(m) = \dim R = 5 > 3 = \operatorname{rank} M.$$

In Sect. 2 we give a different method for producing minimal generators with order ideal the maximal ideal. However, our fourth order ideal theorem guarantees there are (many) minimal generators that have a more stable upper bound on the height of their order ideals. The organization of our proof of the next result requires the residue field be algebraically closed, although one eventually can see that a finite residue field extension of the local ring in question is all that is needed.

Theorem 5 (Evans–Griffith [20]). *Let* $(R, \mathfrak{m}, k)$ *be a local domain which is universally catenary and such that the residue field* k *is algebraically closed. Let* E *be a nonfree torsion free* R*-module. Then some minimal generator* $e \in E$ *must have the property* $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$.

The theorem was stated in our 1985 monograph [22, Theorem 2.12] in contrapositive form, that is, if each $e \in E - \mathfrak{m}E$ has $\operatorname{ht} O_E(e) > \operatorname{rank} E$, then E must be free. The key to our argument was a lemma [22, Lemma 2.13] which showed there is a homogenous ideal J in $k[X_1, \ldots, X_t]$ such that to each maximal ideal $(X_1 - a_1, \ldots, X_t - a_t)$ in $V(J) - \{(X_1, \ldots, X_t)\}$ corresponds a minimal generator $e = a_1e_1 + \cdots + a_te_t$ in E with $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$. (Here $e_1, \ldots, e_t$ is a prescribed minimal generating set for E.)

Perhaps the first question that comes to mind after going over our argument is: "Does one really need to make the residue field algebraically closed?" To see that the answer is "no" let's begin with a local domain R and nonfree, torsion-free R-module E. To obtain the desired minimal generator e with order ideal height bounded above by $\operatorname{rank} E$ we simply base change E to $R' \otimes E$ where $R \to R'$ represents the faithfully flat local homomorphism (integral as well) which allows one to algebraically close the residue field. From the above lemma we capture a minimal generator e of $R' \otimes E$ where $e = \sum^t a_i \otimes e_i$ and $E = \sum^t Re_i$. Thus we really only need to get our hands on the elements $a_1, \ldots, a_t$ in R' in order to secure the desired minimal generator—and this step can be accomplished via a finite faithfully flat residue field extension R'' of our original local ring R. The extension $R \to R''$ is even finite etale if separability is not an issue (see [21]).

We wish to point out here that since the writing of our monograph (now 26 years ago), Eisenbud–Huneke–Ulrich addressed the issue of Theorem 5 in their 2004 article [18, see Theorem 2.1]. Their result also employs a faithfully flat base change. The trade-off in their construction is that while not adjoining algebraic elements to the residue field, they adjoin polynomial variables to the local ring and then localize. Their theorem works quite well in getting bounds on heights for determinantal ideals.

A curious fact about kth syzygies results from applying the fourth order ideal theorem to the following situation. Let R be a local ring and E a kth syzygy of finite projective dimension and let $e_1, \ldots, e_{k-1}$ be any collection of $k - 1$ minimal generators that are linearly independent modulo $\mathfrak{m}E$, i.e., they are k-linearly independent where k denotes the residue field. Let $F = \sum_{i=1}^{k-1} Re_i$. It is easy to see that the submodule F is "$\mathfrak{m}$-pure" in E, i.e., $F \cap \mathfrak{m}E = \mathfrak{m}F$. We assume Theorem 4 holds, e.g., say R contains a field. Therefore, we have $\operatorname{ht} O_F(e) \geq \operatorname{ht} O_E(e) \geq k$

for each $e \in F - \mathfrak{m}F$. Of course, if $\operatorname{rank} F = k - 1$ then F is free and all is well. So what about the case $\operatorname{rank} F < k - 1$? With a little fuss we may pass to the situation where R is a local domain having algebraically closed residue field (we lose finite projective dimension, but this is not required to invoke Theorem 5). This step produces a contradiction to the fact F is not free. Thus we observe that for equicharacteristic local rings and kth syzygy modules E that k-linear independence implies R-linear independence as long as the collection of elements has less than k-members. Hochster–Huneke [34, Corollary 10.10] first noticed this fact about kth syzygies and present an elegantargument based on induction.

The observation above concerning the two notions of linear independence in kth syzygies deserves a note of caution. To get our point across we let $(R, \mathfrak{m})$ be a regular local ring of dimension ≥ 4 and let $I = (a_1, a_2, a_3, a_4)$ be a 4-generated ideal of height $= 2$, e.g., $I = (x, y) \cap (w, z)$ where x, y, w, z are part of a regular system of parameters. Let E be the module defined by the short exact sequence:

$$0 \to R \to R^4 \to E \to 0$$

where 1 is sent to $\langle a_1, a_2, a_3, a_4 \rangle$. The sequence does not split for all prime ideals of height ≤ 2. Therefore one sees that E is a torsion-free module that is not a second syzygy. Let e_1, e_2, e_3 denote 3 minimal generators that are linearly independent modulo $\mathfrak{m}E$ and consider the equation

$$\lambda_1 e_1 + \lambda_2 e_2 + \lambda_3 e_3 = 0.$$

After performing an automorphism of R^4 we may assume that e_1, e_2, e_3 are images of the first 3 standard basis vectors in R^4. Note $v = \langle a_1, a_2, a_3, a_4 \rangle$ will be sent to $v' = \langle b_1, b_2, b_3, b_4 \rangle$ where (b_1, b_2, b_3, b_4) has height 2 and is 4-generated, i.e., the order ideals will be the same! The upshot here is we may reduce our consideration for the above equation in E to the following one:

$$\lambda_1 \langle 1, 0, 0, 0 \rangle + \lambda_2 \langle 0, 1, 0, 0 \rangle + \lambda_3 \langle 0, 0, 1, 0 \rangle = \lambda_4 \langle b_1, b_2, b_3, b_4 \rangle$$

in R^4. The only possible solution is $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$. Thus E has the linear independence property exhibited by a fourth syzygy but in fact is not even a second syzygy.

We summarize our discussion on order ideals. For this purpose let $(R, \mathfrak{m})$ be a local ring and let E be a finitely generated kth syzygy module over R. (When Theorem 4 is invoked the ring R should contain a field.) We assume that $\operatorname{pd} E < \infty$ but that E is not free. Let us examine properties of the order ideal $O_E(e)$, $e \in E$. If $e \in \mathfrak{m}E$ then the first order ideal theorem gives $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$. Moreover, even if $e \in E - \mathfrak{m}E$ one can expect some such e with $\operatorname{ht} O_E(e) \leq \operatorname{rank} E$—perhaps after a local faithfully flat (finite) ring extension. Theorem 4 guarantees a lower-bound $\operatorname{ht} O_E(e) \geq k$ for all minimal generators $e \in E - \mathfrak{m}E$—and Theorem 4 provides the existence of a minimal generator e for which the strict inequality $\operatorname{ht} O_E(e) > k$ holds should $\operatorname{rank} E > k$. In each of the inequalities noted above, the syzygy index k and the rank of E are the key invariants for bounding heights of order ideals from below and above, respectively. Moreover, these inequalities suggest that the inequality $k \leq \operatorname{rank} E$ should hold as well. In the next section (Sect. 2) we argue that this inequality is indeed true provided the local ring R contains a field. The inequality $k \leq \operatorname{rank} E$ is known as the "syzygy theorem"—for modules of finite projective dimension—and of course should not be confused with the famous "Hilbert syzygy theorem" on existence of finite free resolutions.

2 The Syzygy Theorem and Applications

Having digested the four basic order ideal theorems in Sect. 1, the reader will no doubt have guessed that the so-called syzygy theorem is but a mere corollary of Theorem 4 together with Theorem 5, at least in the case of a local domain with algebraically closed residue field. In fact the small technical difficulty concerning the assumptions of "domain" and "algebraically closed" residue field can themselves be easily circumvented by appealing to the version of Theorem 5 developed by Eisenbud–Huneke–Ulrich [18, Theorem 2.1]. We describe yet another elementary approach below—but first we state the theorem.

Syzygy Theorem. *Let $(R, \mathfrak{m})$ be a local ring that contains a field and let E be kth syzygy module having finite projective dimension. If E is not a free module, then* $\operatorname{rank} E \geq k$.

Proof (Sketch). We give a simple argument based on Theorems 4 and 5 of Sect. 1. By way of contradiction let's suppose that k is the smallest integer for which the claim is not true. Since the case of $k = 2$ and $\operatorname{rank} E = 1$ can be treated in way similar to Kaplansky's argument (see [36, Theorem 20.3, p. 163]) we must have $k > 2$ and $\operatorname{rank} E < k$. Let $e \in E - \mathfrak{m}E$. By Theorem 4 one has $\operatorname{ht} O_E(e) \geq k$ and $E/Re = E'$ is necessarily a $(k-1)$st syzygy of finite projective dimension having rank $< k - 1$. We have achieved a contradiction at this point. Thus, it must be that $\operatorname{rank} E \geq k$. □

Our original proof [19] relied on induction in much the same way as the one above. However, we also assumed that $\dim R$ was as small as possible for a counterexample to occur. This assumption allowed us to have the property that E was locally free on $\operatorname{Spec} R - \{\mathfrak{m}\}$ and that the coset $e + IE$ was supported only at $\{\mathfrak{m}\}$, where $I = O_E(e)$. In this way we could invoke a version of INIT (see Sect. 1) in order to apply a more limited form of Theorem 4. In 2002 Heitmann [30] established the direct summand theorem for all regular local rings of dimension ≤ 3. Putting this fact together with Hochster's result [33, section 2] allows one to conclude that INIT holds for all free complexes of length two that satisfy the hypothesis of INIT. In [14, Corollary 3.5] it was shown that the above observations could be used to prove that all $(n-2)$nd syzygies of finite projective dimension satisfied the syzygy theorem for $\dim R = n$. The next theorem is a slight improvement of [14, Corollary 3.5].

Theorem 1. *Let* $(R, \mathfrak{m})$ *be a local ring and let* E *be a* k*th syzygy module of projective dimension* ≤ 2*. If* E *is not free then* $\operatorname{rank} E \geq k$.

Proof. One immediately reduces the argument to the situation E is locally free on $\operatorname{Spec} R - \{\mathfrak{m}\}$. Then one may conclude the proof by making application of INIT that follows from Heitmann's theorem [30]. □

To see that the syzygy theorem is sharp we return to Theorem 4 and state an equivalent form.

Theorem 2 (Bruns [5]). *Let* $(R, \mathfrak{m})$ *be a local ring and let* E *be a* k*th syzygy module of finite projective dimension. If* $\operatorname{rank} E = k + s$ *then there is a free submodule* F *of* E *such that* $\operatorname{rank} F = s$ *and* E/F *is a* k*th syzygy of rank* k.

In our monograph [22, p. 54, 55] we illustrated how to apply the above theorem to construct free resolutions over a Cohen–Macaulay local ring of dimension n in which the kth syzygy had rank exactly k for $0 < k < n$. Moreover, we applied the same technique [22, Corollary 3.13] to further illustrate how one may start with any finite free resolution

$$0 \to F_n \to F_{n-1} \to \cdots \to F_2 \to F_1 \to F_0 \to M \to 0$$

and perturb the first few terms in order to obtain a free resolution $F'_\bullet$ in which all terms and differentials are identical until F'_2 (summand of F_2) and $F'_1 = R^3$, $F_0 = R$. Thus any pathology exhibited by $F_\bullet$ in the "back" $n - 2$ terms is also exhibited by $F'_\bullet$. In addition the first syzygy of $F'_\bullet$ is a 3-generated ideal. This phenomena was discovered by Bruns [5]. One may properly conclude that the pathology that takes place in all finite free resolutions can be observed when one restricts the focus to free resolutions of three-generated ideals. The following three observations are examples of this philosophy. We let R denote the local ring under consideration:

(a) R is a local domain $\iff$ each principal ideal has finite projective dimension.
(b) R is a local UFD $\iff$ each two-generated ideal has finite projective dimension (see [22, Theorem 4.3]).
(c) R is a regular local ring $\iff$ each 3-generated ideal has finite projective dimension.

In contemplating the implication "$\Leftarrow$" in part c) one notes that the hypothesis localizes, so one may use part b) to see that R is a local UFD—and from there it is not difficult to see that R is regular in codimension ≤ 2. Thus all second syzygies are free locally in codimension ≤ 2. A slightly more general version of Theorem 4 (see [22, Corollary 2.6]) allows one to reduce the whole question to whether second syzygy modules of rank two have finite projective dimension. The final step is accomplished by noting that second syzygy modules of rank two are always first syzygies for 3-generated ideals (see [22, Corollary 3.13 (proof)]).

There is one case of 3-generated ideals in which a "nice" conclusion is reached and is the subject of our next theorem.

Theorem 3. *Let* $(R, \mathfrak{m})$ *be a local ring that satisfies* (S_3) *and contains a field. If* I *is a* 3*-generated ideal of* R *that is unmixed and of finite projective dimension then* R/I *is Cohen–Macaulay and* $\operatorname{pd} I = \operatorname{ht} I - 1$.

Proof. (case $\operatorname{ht} I = 2$) One has the short exact sequence

$$0 \to E \to R^3 \to I \to 0$$

in which $\operatorname{pd} E < \infty$ and $\operatorname{rank} E = 2$. The "unmixed" hypothesis allows one to see that E satisfies the Serre condition (S_3) and is therefore a third syzygy of rank two and has finite projective dimension. Thus the syzygy theorem applies and E must be free. □

The above theorem is most interesting in case I is a height two unmixed ideal (e.g., a 3-generated prime ideal of height two) in a regular local ring. Here we see that all such ideals behave as in the generic case of the ideal generated by the 2×2 minors of a 2×3 matrix of variables (see [22, Corollary 4.6]).

Remark 4. One can use the Bruns' modification of finite free resolutions described above to get more exotic examples of modules with lots of elements with order ideals having height larger than the rank. Let $(R, \mathfrak{m}, k)$ be a regular local ring of dimension at least 4. Let M be any module with $\dim \operatorname{Soc}(M) = d$ and let $F_\bullet$ be its free resolution. Apply Bruns' theorem to $F_\bullet$ to get a resolution of R modulo a three generated ideal, $I = (a, b, c)$, with the same last $\dim R - 2$ terms as $F_\bullet$. Then R/I has the same socle as M. In fact they have the same zeroth local cohomology module. Let $m_1, \ldots, m_d$ be representatives of a k-basis of $\operatorname{Soc}(R/I)$. We define E to be R^{d+3} modulo the element $v = \langle m_1, \ldots, m_d, a, b, c \rangle$. Then if x is in the maximal ideal we have that $x m_i = ra + sb + tc$. This relation gives an element of E^* sending the ith generator of E to x and the other of the first d generators of E to zero and sending the last three generators to r, s, and t, respectively. Let F be the submodule of E generated by the first d generators. Then

(i) $\operatorname{pd} E = 1$.
(ii) E is a first syzygy but not a second syzygy.
(iii) $\mu(E) = d + 3$.
(iv) the rank of E is $d + 2$.
(v) F is a free submodule of E of rank d.
(vi) If f is in $F - \mathfrak{m}F$ then the order ideal of f is $\mathfrak{m}$.

Since the dimension of R can be arbitrarily large, the difference between the rank of E and the height of the order ideals can be as large as you want. We suspect that there is no example of a module of rank $d + 1$ having a free submodule F of rank d with the order ideals equal to the maximal ideal for all f in $F - \mathfrak{m}F$.

The syzygy theorem also proved useful in generalizing results of Serre [42] in which he sought conditions for a prime ideal of codimension two to be generated by two elements. The connection to Serre's work was noted by Simon [45] in which she observed the following result (also see [22, Theorem 4.7, Corollary 4.8]).

Theorem 5. *Let R be a regular local ring containing a field and let $\mathfrak{p}$ be a prime ideal such that $R/\mathfrak{p}$ has cyclic canonical module. Then $\mathfrak{p}$ is generated by a 2-sequence (e.g., this occurs if $R/\mathfrak{p}$ is a UFD).*

Proof. The canonical module in this case is $\omega_{R/\mathfrak{p}} = \mathrm{Ext}^2(R/\mathfrak{p}, R) \cong \mathrm{Ext}^1(\mathfrak{p}, R)$. Thus, there is a class

$$[0 \to R \to E \to \mathfrak{p} \to 0]$$

which generates $\mathrm{Ext}^1(\mathfrak{p}, R)$. It follows easily that E is a rank two-third syzygy (i.e., $\mathrm{Ext}^1(E^*, R) = 0$). By the syzygy theorem E must be free from which the conclusion follows. □

A kth syzygy module in certain instances can possess special properties that force its rank to be even larger than predicted by the syzygy theorem. An example of such behavior is the subject of our next result.

Theorem 6 (Griffith–Seceleanu [27]). *Let $(R, \mathfrak{m})$ be a regular local ring and let E be a kth syzygy module such that $\mathrm{pd}_R\, E \geq 2$. Suppose E contains a free submodule F for which $x(E/F) = 0$, where x is a nonzero element in $\mathfrak{m}$. If $T = E/F$ has finite projective dimension over $S = R/xR$ then* rank $E \geq 2k - 1$.

Proof. One argues that $0 \to F \to E \to T \to 0$ gives rise to the 4-term exact sequence modulo (x) where T is a kth syzygy over S of finite projective dimension.

$$\begin{array}{ccccccccccc} 0 & \longrightarrow & T & \xrightarrow{\partial} & \overline{F} & & \longrightarrow & & \overline{E} & \longrightarrow & T \longrightarrow 0 \\ & & & & & \searrow & & \nearrow & & & \\ & & & & & & K & & & & \end{array}$$

Moreover, K must be a $(k-1)$st syzygy of finite projective dimension over S. Neither K nor T can be free S-modules since $\mathrm{pd}_R\, E \geq 2$. By the syzygy theorem one has rank E = rank $\overline{E}$ and rank $\overline{E} \geq (k-1) + k \geq 2k - 1$. □

The above situation always occurs if $x(E/F) = 0$ and $x \in \mathfrak{m} - \mathfrak{m}^2$ when R is regular local. In this special circumstance the result follows in a trivial way from Shamash's article [44] (see also Avramov [3, 3.3.5]). Moreover, in some cases, the map ∂ provides splitting over the ring R/xR which is equivalent to the property $T = E/F$ "weakly" lifts to R, i.e., $T \oplus \mathrm{Syz}_1^R(T)$ lifts to R. Indeed this is precisely the situation for the residue field k of a local ring of positive depth. Namely, $\mathfrak{m}/x\mathfrak{m}^2 \cong k \oplus \mathfrak{m}/(x)$ for each regular hyperplane "x." One can conclude from the decomposition that the kth syzygy module of the residue field k is—modulo the hyperplane $x = 0$—isomorphic to the direct sum of the kth and $(k-1)$st syzygy modules for k over $\overline{R} = R/xR$. From here a simple induction argument gives the standard formula

$$\operatorname{rank}\operatorname{Syz}_k(k) = \binom{n-2}{k-1} + \binom{n-2}{k-2} = \binom{n-1}{k-1}$$

where $\dim R = n$ and R is regular local.

More generally, examples of syzygies of modules of finite length and finite projective dimension appear to exhibit Betti numbers and syzygy ranks having similar characteristics to those of the residue field. For this reason Buchsbaum–Eisenbud [9] and Horrocks (see [24]) independently formulated the following conjecture: If M represents an R-module of finite length and finite projective dimension then the ith Betti number of $M \geq \binom{n}{i}$ where $n = \dim R$. Various special cases of the Buchsbaum–Eisenbud–Horrocks conjecture have been proven, e.g., Herzog–Kuhl [31] for the case of "pure graded resolutions" and Evans–Griffith [24] in the context of monomial ideals (see Charambolus [11] and Charambolus–Evans [12] for additional insight into behavior of the resolutions).

Aside from obvious interpretations of the syzygy theorem into the language of coherent sheaves in algebraic geometry as we observed in Sect. 1, there have been more recent subtle applications in the articles of Pareschi–Popa [38] and Lazarsfeld–Popa [35] in developing a suitable theory of generic vanishing indices involving kth syzygy sheaves (see [38, section 3] for a definition).

3 Serre Intersection Theorem and Order Ideals of Consecutive Syzygy Modules

In the current section we explore a circle of ideas that is at the heart of the Eisenbud–Huneke–Ulrich article [17, section 2]. In that article the authors are most interested in extending Theorem 1 to order ideals of minimal generators—at least to determine when it is possible to do so. Their main result [17, Theorem 3.1] accomplishes this task in a quite satisfactory manner. The key point in their development is to follow what transpires after base changing to local domains. Here we consider how the upper bounds of heights of order ideals of minimal generators of a kth syzygy module E influence order ideals of the dual of its first syzygy.

We set some notation to be used throughout this section. Let $(R, \mathfrak{m})$ be a regular local ring and let E be a kth syzygy module for which $k \geq 2$, so E is necessarily a reflexive R-module. We consider a short exact sequence

$$0 \to Z \to F \to E \to 0 \qquad (\dagger)$$

where F is a free R-module and $Z \subseteq \mathfrak{m}F$. Since most of the homological problems we encounter that involve syzygy modules come down to analyzing the case where

E is assumed to be locally free on Spec $R - \{\mathfrak{m}\}$, we make this assumption here, that is, the short exact sequence (†) is locally split on Spec $R - \{\mathfrak{m}\}$. Before getting to the main result we state an elementary lemma.

Lemma 1. *If F is a free R-module and if $\pi : F \to R$ represents a free generator of F^* then π is necessarily surjective.*

Proof. Expand the set $\{\pi\}$ to a free basis for F^* and consider a dual basis. □

The notion of "perpendicular element" (language used in [17, section 2]) is at the heart of the arguments presented below. Our assumption of "algebraically closed residue field" is purely technical and is a result of an application of Theorem 5 (Sect. 1).

Theorem 2. *Let $(R, \mathfrak{m})$ be a regular local ring having algebraically closed residue field. Let E be a kth syzygy module over R with $k \geq 2$. Let $0 \to Z \to F \to E \to 0$ be a minimal free presentation of E and assume the presentation is locally split on* Spec $R - \{\mathfrak{m}\}$. *Then*

(a) For $e \in E - \mathfrak{m}E$ there is a $z^* \in Z^*$ such that

$$\operatorname{ht} O_E(e) + \operatorname{ht} O_Z(z^*) \geq \dim R.$$

(b) If $\operatorname{ht} O_E(e) \leq h$ *for each minimal generator e of E then*

$$\operatorname{rank} Z \geq \dim R - h.$$

(c) If equality $\operatorname{ht} O_E(e) = h$ *holds for all $e \in E - \mathfrak{m}E$ then*

$$\operatorname{rank} E + \operatorname{rank} Z \geq \dim R,$$

i.e., this statement is equivalent to the "kth Betti number is $\geq$ $\dim R$."

Proof. Consider the 4-term induced exact sequence

$$\begin{array}{ccccccccc} 0 \longrightarrow & E^* & \longrightarrow & F^* & \longrightarrow & Z^* & \longrightarrow & \operatorname{Ext}^1(E,R) & \longrightarrow 0. \\ & & & & \searrow \quad \nearrow & & & & \\ & & & & W & & & & \end{array}$$

Then $\operatorname{Supp}\operatorname{Ext}^1(E, R) \subseteq \{\mathfrak{m}\}$ implies that W and Z^* are locally equal on Spec $R - \{\mathfrak{m}\}$. Moreover, there are natural isomorphisms $W^* = Z^{**} \cong Z$. Therefore, if $w \in W$, then $O_W(w) = O_{Z^*}(w)$.

For $e \in E - \mathfrak{m}E$ one also has the commutative triangle

$$\begin{array}{ccc} E^* & \longrightarrow & F^* \\ {\scriptstyle e}\downarrow & \swarrow {\scriptstyle \pi} & \\ R & & \end{array}$$

where π is a free generator of F^*. We note the map π corresponds to $v \in F - \mathfrak{m}F$, where $v \mapsto e$, under the natural identification $F = F^{**}$. Therefore π is a surjection. The image of the map e above is simply $I = O_E(e)$. Thus π induces a surjection

$$W \to R/I.$$

Let $w \in W - \mathfrak{m}W$ such that $w \mapsto 1 + I$ under the above induced map and let $J = O_W(w) = O_{Z^*}(w)$. Note $w + JW$ is supported at $\{\mathfrak{m}\}$ which implies its image after base change:

$$R/J \otimes W \to R/J \otimes R/I = R/(I + J)$$

is also supported at $\{\mathfrak{m}\}$. Therefore length$(R/(I+J)) < \infty$. By Serre's intersection theorem [43] we have

$$\operatorname{ht} I + \operatorname{ht} J \geq \dim R.$$

Thus, part a) is proved. For part b) we observe that $w \in W - \mathfrak{m}W$ such that $w \mapsto 1 + I \in R/I$ must satisfy $\operatorname{ht} O_W(w) \geq \dim R - h$. We argue that all minimal generators of W must satisfy this inequality. For $w \in W - \mathfrak{m}W$ choose $\pi \in F^* - \mathfrak{m}F^*$ such that $\pi \mapsto w$ under $F^* \to W$. Since π is necessarily surjective one has a commutative triangle

$$\begin{array}{ccc} E^* & \longrightarrow & F^* \\ {\scriptstyle e = \pi|_{E^*}} \downarrow & \swarrow {\scriptstyle \pi} & \\ R & & \end{array} .$$

(Each map $E^* \to R$ is given by an associated $e \in E$.)

Hence, $\operatorname{ht} O_W(w) \geq \dim R - h$. By Theorem 5 (Sect. 1) we have

$$\operatorname{rank} W = \operatorname{rank} Z^* = \operatorname{rank} Z \geq \dim R - h.$$

This inequality proves part c) in view of Theorem 5 since $h \leq \operatorname{rank} E$ is necessary. □

Corollary 3. *If* $\operatorname{rank} E \leq k$ *then* $\operatorname{rank} Z \geq \dim R - k$.

The above corollary may be interpreted to say: if E has minimal rank for a kth syzygy module—under hypothesis of the theorem—then its first syzygy must compensate with sufficiently larger rank so that part c) holds, i.e., $\operatorname{rank} Z \geq \dim R - \operatorname{rank} E$.

In mixed characteristic the "syzygy theorem" remains a conjecture—although it is known that the conclusion can be off by at most "1", i.e., $\operatorname{rank} E \geq k - 1$ (see our discussion in Sect. 4). The above corollary does show for $k < \frac{1}{2}(\dim R + 1)$ it is not possible that successive syzygy modules can be counter examples.

Corollary 4. *Suppose R has mixed characteristic and* $k < \frac{1}{2}(\dim R + 1)$ *where* $\operatorname{rank} E = k - 1$ *and E not free. Then* $\operatorname{rank} Z \geq k + 1$.

In summary, one should take away the following general rule (under our hypothesis on R and E): Every "small" order ideal $O_E(e)$ with $e \in E - \mathfrak{m}E$ produces a "large" order ideal in Z^*. The converse holds when $\mathrm{Ext}^1(E, R) = 0$, i.e., when E is universal with respect to Z (see Sect. 1).

4 The State of Mixed Characteristic

For our discussion of what is known in mixed characteristic "p" we restrict our attention to the setting in which R is a standard graded ring

$$R = R_0 \oplus R_1 \oplus R_2 \oplus \cdots ,$$

where R_0 is a discrete valuation ring, or R is local and catenary (e.g., R is complete). In all cases we assume that the base ring R contains p as a regular element and satisfies (S_k). Let E be a kth syzygy of finite projective dimension in either of the two situations just mentioned and let $e \in E$ be a minimal generator for E (e is required to be homogenous in the graded case). The difficulty with establishing Theorem 4 for the order ideal $O_E(e)$ in this context is that INIT may not hold for the factor ring $R/O_E(e)$. However, there is one obvious situation when Theorem 4 is valid.

Theorem 1. *Let R, E, and $e \in E$ be as above with E nonfree.*

(a) If p is in some prime $\mathfrak{p}$ that contains $O_E(e)$ such that $\mathrm{ht}\,\mathfrak{p} = \mathrm{ht}\,O_E(e)$ *then* $\mathrm{ht}\,O_E(e) \geq k$.
(b) The element p is in the radical ideal $\sqrt{O_E(e)}$ if and only if e is a free generator for $E[p^{-1}]$ over $R[p^{-1}]$.

In case there is at least one minimal generator e such that $\mathrm{ht}\,O_E(e) \geq k$ we say that the kth syzygy E satisfies the *weak order ideal property*, that is, E satisfies a weaker version of Theorem 4. In their article [27, Proposition 3.1] Griffith–Seceleanu argue the weak order ideal property is sufficient for proving the syzygy theorem.

Theorem 2 (Griffith–Seceleanu [27, 3.1]). *If a nonfree kth syzygy E (as above) satisfies the weak order ideal property then* $\mathrm{rank}\,E \geq k$.

Thus for questions concerning $\mathrm{rank}\,E$ it suffices to know the weak order ideal property holds. Note, by Theorem 4 on order ideals, if $\mathrm{rank}\,E \geq k$, then the weak order ideal property must hold for E. One sees this claim by viewing E as a $(k-1)$st syzygy having rank $> k-1$. Thus by Theorem 4 there is a minimal generator $e \in E$ such that E/Re is a $(k-1)$st syzygy, i.e., $\mathrm{ht}\,O_E(e) \geq k$. For graded modules the following result provides a satisfactory criterion for when the weak order ideal property must hold. The argument outlined here is essentially identical to the one given in our article [26, Theorem 6].

Proposition 3. *Let R be a standard graded ring*

$$R = R_0 \oplus R_1 \oplus R_2 \oplus \cdots$$

where R_0 is a DVR having uniformizing parameter p such that p is regular on R. If

$$E = E_0 \oplus E_1 \oplus \cdots$$

is a graded p-regular R-module such that $E[p^{-1}]$ is $R[p^{-1}]$-projective, then any homogenous element $e \in E_0 - pE_0$ satisfies $p \in \sqrt{O_E(e)}$.

Proof. Let K denote the fraction field of R_0, i.e., $K = R_0[p^{-1}]$. We may assume E_0 is the first nonzero graded piece (otherwise we "twist" the grading on E until this is so). The key point here is that $E[p^{-1}]$ remains a graded $R[p^{-1}]$-module where

$$R[p^{-1}] = K \oplus R_1[p^{-1}] \oplus R_2[p^{-1}] \oplus \cdots$$

is graded over a field K. Since graded projective modules in such a case are known to be free one can see that any nonzero element in $E_0[p^{-1}]$ must be part of a free basis. Thus $p \in \sqrt{O_E(e)}$ for any such minimal generator. □

Corollary 4. *The syzygy theorem holds for graded kth syzygies of finite projective dimension over standard graded rings (as above).*

Proof. Should there be a counterexample to the syzygy theorem in this setting, i.e., a graded kth syzygy module

$$E = E_0 \oplus E_1 \oplus \cdots$$

of finite projective dimension such that E is nonfree and rank $E < k$, then $E[p^{-1}]$ must be locally free over $R[p^{-1}]$ since we are back in the equicharacteristic situation when p^{-1} exists. From Proposition 3 we see that E must have the weak order ideal property which contradicts the conclusion of Theorem 2. □

Unfortunately the above line of reasoning does not hold in the local setting—even when R is regular local. Two examples are given in [27, section 7]. We provide a repeat discussion of them here.

Example 5. Let $R = V[[x, y]]$ where V is a DVR and p generates the maximal ideal of V. Let

$$I = (x^2 + y^2 + p^2, px, py).$$

It is easy to check $p^3 \in \mathfrak{m}I$; note

$$p^3 = p(x^2 + y^2 + p^2) - px(x + y) - py(y - x).$$

However, one can argue $p^2 \notin I$. Observe $I[p^{-1}] = R[p^{-1}]$. Since grade $I = 2$ the R-homomorphisms $I \to R$ are just given by multiplication by elements of R.

Thus, if $a \in I$ then $O_I(a) = (a)$. Consequently no minimal generator e of I has $p \in \sqrt{O_I(e)}$. A similar analysis works for a ramified regular local ring in which $p \in \mathfrak{m}^2$ and $E = \mathfrak{m}$.

In the remainder we emphasize the regular local case since such rings seem to present most of the challenges one meets when studying behavior of order ideals. Thus we let $(R, \mathfrak{m})$ be a complete regular local ring of mixed characteristic p. The R-module E will always denote a kth syzygy R-module. Similar to our analysis of the order ideal situation that occurs in Proposition 3 above, one easily comes to the realization that it is sufficient to study kth syzygy modules E for which $E[p^{-1}]$ is $R[p^{-1}]$-projective. In fact it suffices to study syzygies of free resolutions of modules M where $p^s M = 0$—and bringing into play a result of Auslander–Bridger [2], [22, Corollary 5.3], we may even assume that $\mathrm{pd}_S\, M < \infty$ where $S = R/p^s R$ (see [27, sections 2 and 3]). We summarize these facts in the following statement.

Observation 6. *Let R and E be as above and assume E is not free. We may assume that E has the properties:*

(i) $E = \mathrm{Syz}_k(M)$ *where M is an $R/p^s R$ module for $s > 0$ and .* $\mathrm{pd}_S\, M < \infty$.
(ii) *E is locally free on* $\mathrm{Spec}\, R - \{\mathfrak{m}\}$.
(iii) $\mathrm{ht}\, O_E(e) \geq k - 1$ *for $e \in E - \mathfrak{m}E$.*
(iv) $\mathrm{Rank}\, E \geq k - 1$.

Property (iii) holds because INIT holds modulo p and (iv) is a consequence of (iii).

We may add a positive observation to the above list. Namely, Theorem 4 and consequently the syzygy theorem holds for $\mathrm{pd}_R\, E \leq 2$. (See Theorem 4*.)

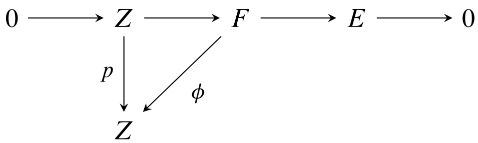
In particular, these results hold when E is a kth syzygy for $k \geq \dim R - 2$.

We can gain a certain level of information by considering a finite ring morphism $R \to S$, especially in case there is a "natural" map $S \otimes_R E \to E'$ where E' is a kth syzygy of finite projective dimension. Such a map will prove fruitful in the circumstance $1 \otimes e \mapsto e' \in E' - \mathfrak{m}E'$ and $\mathrm{ht}_S\, O_E(e') \geq k$.

We follow the presentation in [27] and address two such situations. The setup for the first one goes as follows: Consider a minimal free R-presentation

$$[\epsilon] : 0 \to Z \to F \to E \to 0$$

and assume $p[\epsilon] = 0$ in $\mathrm{Ext}^1(E, Z)$. Note the hypotheses $E = \mathrm{Syz}_k(M)$ where $pM = 0$ would be sufficient to force $p[\epsilon] = 0$. We obtain the following commutative triangle:



which induces a map $E/pE \to Z/pZ$. Moreover, some minimal generator of E/pE, say $e + pE$, will map to a minimal generator $z + pZ$ of Z/pZ provided $\phi(F) \nsubseteq \mathfrak{m}Z$. In [27, section 4] this situation is analyzed and the following theorem is proven.

Theorem 7. *Let R be an unramified regular local ring of mixed characteristic p, and let M be an R-module such that $p\mathrm{Ext}^{k+1}(M,-) = 0$ for some $k > 0$. Then the weak order ideal property holds for each jth syzygy of M where $j \geq k$; and consequently the syzygy theorem holds in the same range as well.*

The point of the argument in getting $e + pE \mapsto z + pZ$ for $z \in Z - \mathfrak{m}Z$ is that $O_{\overline{Z}}(\overline{z}) \subseteq O_{\overline{E}}(\overline{e})$, where $^{-}$ means modulo p. So $\mathrm{ht}\, O_{\overline{E}}(\overline{e}) \geq \mathrm{ht}\, O_{\overline{Z}}(\overline{z}) \geq k$. Moreover, it is an easy argument to also check $\mathrm{ht}\, O_E(e) \geq \mathrm{ht}\, O_{\overline{E}}(\overline{e})$ (see [26, Corollary 3]).

A rather immediate corollary follows:

Theorem 8. *The syzygy theorem holds for syzygies of $R/\mathfrak{q}$ when $\mathfrak{q} \in \mathrm{Spec}\, R$ and R is unramified.*

Proof. If $p \in \mathfrak{q}$ then the conclusion is a consequence of Theorem 7. When $p \notin \mathfrak{q}$ then the entire free resolution restricts intact to the hypersurface ring R/pR where our questions on $\mathrm{ht}\, O_E(e)$ and $\mathrm{rank}\, E$ have affirmative answers. □

The second comparison begins with E being the kth syzygy in a free R-resolution $F_\bullet \to M$ where $p^s M = 0$ and $\mathrm{pd}_S\, M < \infty$ and where $S = R/p^s R$ (see Observation 6 (i)). Restricting the complex $F_\bullet \to M$ to $R/p^s R$ one gets an exact complex up to the degree 1-term. Here we have the 4-term exact sequence

$$0 \to M \to \overline{Z_1} \to \overline{F_0} \to M \to 0.$$

Thus, the truncated complex $(\overline{F}_\bullet)_{i\geq 1}$ becomes a free S-resolution for $\overline{Z_1}$ where

$$0 \to M \to \overline{Z_1} \to K_1 \to 0$$

is exact; $K_1 = \mathrm{Syz}_1^S(M)$. Denoting the S-syzygies of M by $K_\bullet$ and computing simultaneous free resolutions of M, $\overline{Z_1}$, and K_1 (use mapping cone) we achieve a short exact sequence of syzygy modules over S:

$$0 \to K_{k-1} \to L_{k-1} \oplus \overline{E} \to K_k \to 0,$$

where L_{k-1} is S-free and where $E = Z_k = k$th syzygy for M and $\overline{E} = E/p^s E$. If the induced map $\overline{E} \to K_k$ has an image in $K_k - \mathfrak{m}K_k$ then $\overline{E}$ and E will have the weak order ideal property. If this is not the case then the induced map $L_{k-1} \to K_k$ is necessarily surjective, and we obtain the induced commutative diagram

$$\begin{array}{ccccccccc}
0 & \longrightarrow & K_{k+1} & \longrightarrow & L_{k-1} & \longrightarrow & K_k & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & & & \\
0 & \longrightarrow & K_{k-1} & \longrightarrow & L_{k-1} \oplus \overline{E} & \longrightarrow & K_k & \longrightarrow & 0. \\
 & & \downarrow & & \downarrow & & & & \\
 & & E & & E & & & &
\end{array}$$

The induced diagram yields a short exact sequence

$$0 \to K_{k+1} \to K_{k-1} \to \overline{E} \to 0$$

which gives the rank inequality

$$\operatorname{rank} K_{k-1} \geq \operatorname{rank} K_{k+1} + \operatorname{rank} E \geq (k+1) + (k-1) = 2k$$

since the syzygy theorem holds for M over S (INIT holds in this case). As a corollary we obtain the following statement

Corollary 9 (Notation as above). *If* $\operatorname{rank}_R K_{k-1} < 2k$ *then E must satisfy the weak order ideal property (because E necessarily has an order ideal of height $\geq k$ over R).*

We remark that Theorem 4 shows there are S-modules M of finite projective dimension such that the $(k-1)$st syzygy K_{k-1} for M can satisfy $\operatorname{rank} K_{k-1} = k - 1 < 2k$ for $1 < k < \dim S$.

If K_{k-1} in the above commutative diagram has no free S-summands then an analysis of the diagram yields that the induced S-map $L_{k-1} \to K_k$ is necessarily minimal, i.e., has kernel contained in $\mathfrak{m}L_{k-1}$. (If the map were not minimal then the syzygy module K_{k+1} would share a free S-summand with L_{k-1} which would then show K_{k-1} also shares a free S-summand with L_{k-1} as well.) Thus the top two rows of the diagram yield the equalities

$$\beta_k = \operatorname{rank} K_{k+1} + \operatorname{rank} K_k = \operatorname{rank} L_{k-1}$$

$$\beta_{k-1} = \operatorname{rank} K_{k-1} + \operatorname{rank} K_k = \operatorname{rank} L_{k-1} + \operatorname{rank} E$$

where β_i represents the ith Betti number for M as an S-module. Thus, one has the equality $\beta_{k-1} - \beta_k = \operatorname{rank} E$.

Corollary 10 (Notation as above). *If the $(k-1)$st S-syzygy module for M has no nonzero free S-summands and if $\beta_{k-1} - \beta_k \neq k - 1$ then E necessarily satisfies the weak order ideal property as an R-module. In this case* $\operatorname{rank} E \geq k$.

The above corollaries show that, if we know sufficient information about the free S-resolution for an S-module M having $\operatorname{pd}_S M < \infty$, then we can often determine specialized information about the ranks of the R-syzygy modules for M. In particular when R is merely a local ring and p^s is replaced by any regular element in R, while other properties of M and E remain the same, one observes the general facts:

(a) One always has rank $E \leq \beta_{k-1}$, and equality holds if M weakly lifts to R (see discussion following Theorem 6).
(b) If $E \to K_k$ does not provide a proper comparison map, then rank $E \geq \beta_{k-1} - \beta_k$, and equality holds if K_{k-1} has no nonzero free S-summands.

To further illustrate how the above corollary might play out in a specific situation let K be an (S_k) module of finite projective dimension over the hypersurface S. We assume that K has no nontrivial free S-summands. Let $G_\bullet \to M$ represent the minimal free S-resolution that is universal for representing K as a kth syzygy over S, that is, $G_\bullet$ is formed by splicing together a minimal free resolution of K together with a minimal universal pushforward of K (see Sect. 1). We note that the $(k-1)$st syzygy module K' in $G_\bullet$ has no nonzero free summands since the universal pushforward is dual exact. We let E be the kth R-syzygy module for M and consider the possibility that rank $E = k - 1$. For this equality to hold we see from Corollary 10 that $\beta_{k-1} - \beta_k = k - 1$, and this statement translates into

$$\mu(K) = \mu(K^*) + (k-1) \qquad (*)$$

where $\mu(-)$ denotes the size of a minimal generating set. One would expect that a constraint such as (*) is rarely satisfied, and thus "most" kth syzygies E that arise in this fashion will satisfy at least the weak order ideal property and so rank $E \geq k$. We invite the reader to consider examples of torsion-free modules K of projective dimension one to see the variance of $\mu(K)$ versus $\mu(K^*)$.

Acknowledgements We first met David Eisenbud in the late 1960s at University of Chicago. David and the first author were graduate students and the second author was a post-doctoral fellow. Even at this young age David was a catalyst in seminars and a central force for ideas at afternoon tea. It was patently clear that he would become a leader in mathematical research and the greater mathematical community. To each of the authors he was an inspiration—especially in their early work. We conclude by recounting memorable conversations between David Eisenbud, Mike Stillman, and the first author. The essence of these conversations served as motivation for our joint work on syzygies.

In the fall of 1976 the first author was talking with Mike Stillman about computing syzygies of a collection of homogeneous polynomials generating an ideal and realized that if they bounded the degree they were looking at they had a simple set of linear equations over a field. They could solve those and build a complex and check the Buchsbaum–Eisenbud criteria for exactness [10]. This was possible to do by hand as the complexes were quite small and the entries were not too complicated. They noticed that if a column in the kth matrix map had entries in the ideal generated by fewer than k of the variables which might as well be $x_1, \ldots, x_{k-1}$, then that entry was a linear combination of entries of lower degrees. The first author mentioned this to David and he said one should be able to prove this by computing $\mathrm{Tor}_k(R/I, R/(x_1, \ldots, x_{k-1}))$. It is zero since R modulo the x's has projective dimension $k-1$ but would be nonzero if the syzygy were a minimal generator. Of course it isn't too hard to see the same proof would show that the order ideal of a minimal generator of a kth syzygy couldn't be in the annihilator of a finitely generated module

of projective dimension less than k (see Theorem 4*, Sect. 1). Thus a Cohen–Macaulay module over R modulo the order ideal of a minimal generator of a kth syzygy would be useful. Alas it took the authors some time to realize that. Happily Mike went to Harvard and met David Bayer. They learned about Grobner bases and created the first version of Macaulay. The authors then used Macaulay to compute lots of resolutions. Eventually they understood how to use the above ideas to formulate a version of the improved new intersection theorem that applied to order ideals of minimal generators which led to the proof of their syzygy theorem.

References

1. Auslander, M., Buschbaum, D.: Codimension and multiplicity. Ann. Math. **68**, 625–657 (1958)
2. Auslander, M., Bridger, M.: Stable module theory. In: Memoirs of the American Mathematical Society, vol. 94. American Mathematical Society, Providence (1969)
3. Avravov, L.: Infinite free resolutions. Six lectures on commutative algebra. Progr. Math. **166**, 1–118 (1998)
4. Bass, H.: K-Theory and stable algebra. Publications Mathématiques, No. 22, IHES, Paris (1964)
5. Bruns, W.: "Jede" endliche freie Auflösung is freie Auflösung eines von drei elementen erzeugten Ideals. J. Algebra **39**, 429–439 (1976)
6. Bruns, W.: The Eisenbud–Evans principal ideal theorem and determinantal ideals. Proc. Amer. Math. Soc. **83**, 19–24 (1981)
7. Bruns, W.: The Evans–Griffith Syzygy theorem and Bass numbers. Proc. Amer. Math. Soc. **115**(4), 939–946 (1992)
8. Bruns, W., Herzog, J.: Cohen–Macaulay rings. In: Cambridge Studies in Advanced Mathematics, vol. 39. Cambridge University Press, Cambridge (1993)
9. Buchsbaum, D., Eisenbud, D.: Lifting modules and a theorem on finite free resolutions. In: Ring Theory, pp. 63–64. Academic, New York (1972)
10. Buchsbaum, D., Eisenbud, D.: What makes a complex exact? J. Algebra **13**, 259–268 (1973)
11. Charalambous, H.: Betti numbers of multigraded modules. J. Algebra **137**(2), 491–500 (1991)
12. Charalambous, H., Evans, E.G.: A deformation theory approach to Betti numbers of finite length modules. J. Algebra **143**(1), 246–251 (1991)
13. Dutta, S.P.: On the canonical element conjecture. Trans. Amer. Math. Soc. **299**, 803–811 (1987)
14. Dutta, S.P., Griffith, P.: Intersection multiplicities, the canonical element conjecture and the syzygy problem. Michigan Math. J. **57**, 227–247 (2008)
15. Eisenbud, D., Evans, E.G.: Generating modules efficiently: theorems from algebraic K-theory. J. Algebra **27**, 278–305 (1973)
16. Eisenbud, D., Evans, E.G.: A generalized principal ideal theorem. Nagoya Math. J. **62**, 41–53 (1976)
17. Eisenbud, D., Huneke, C., Ulrich, B.: Order ideals and a generalized Krull height theorem. Math. Ann. **330**(3), 417–439 (2004)
18. Eisenbud, D., Huneke, C., Ulrich, B.: Heights of ideals of minors. Amer. Math. J. **126**(2), 417–438 (2004)
19. Evans, E.G., Griffith, P.: The syzygy problem. Ann. Math. (2) **114**, 323–333 (1981)
20. Evans, E.G., Griffith, P.: Order ideals of minimal generators. Proc. Amer. Math. Soc. **86**, 375–378 (1982)
21. Evans, E.G., Griffith, P.: The syzygy problem: A new proof and historical perspective. Commutative algebra (Durham 1981). In: London Mathematical Society Lecture Note Series, vol. 72. Cambridge University Press, Cambridge (1982)
22. Evans, E.G., Griffith, P.: Syzygies. In: London Mathematical Society Lecture Note Series, vol. 106. Cambridge University Press, Cambridge (1985)

23. Evans, E.G., Griffith, P.: Syzygies: the codimension of zeros of a nonzero section. Algebraic Geometry, Bowdin, pp. 485–489 (1985). Proceedings of Symposium in Pure Mathematics, vol. 46, part 2, American Mathematical Society, Providence, RI (1987)
24. Evans, E.G., Griffith, P.: Binomial behavior of Betti numbers for modules of finite length. Pacific Math. J. **133**(2), 267–276 (1988)
25. Evans, E.G., Griffith, P.: Order ideals. Commutative Algebra, Berkeley, CA, pp. 213–225 (1987). Mathematical Sciences Research Institute Publications, vol. 15. Springer, New York (1989)
26. Evans, E.G., Griffith, P.: A graded syzygy theorem in mixed characteristic. Math. Res. Lett. **8**, 605–611 (2001)
27. Griffith, P., Seceleanu, A.: Syzygy theoremsvia comparison of order ideals on a hypersurface. J. Applied Alg. **216**(2), 468–479 (2012)
28. Hackman, P.: Exterior powers of homology. Ph.D. Thesis, University of Stockholm, Sweden (1969)
29. Hartshorne, R.: Algebraic Geometry. Springer, New York (1977)
30. Heitmann, R.: The direct summand conjecture in dimension three. Ann. Math. (2) **156**, 695–712 (2002)
31. Herzog, J., Kühl, M.: On the Betti numbers of finite pure and linear resolutions. Commun. Algebra **12**(13–14), 1627–1646 (1984)
32. Hochster, M.: Topics in the homological theory of modules over commutative rings. In: CBMS Regional Conference Series in Applied Mathematics, vol. 24. American Mathematical Society, Providence, RI (1975)
33. Hochster, M.: Canonical elements in local cohomology modules and the direct summand conjecture. J. Algebra **84**, 503–553 (1983)
34. Hochster, M., Huneke, C.: Tight Closure, Invariant Theory, and the Briancon–Skoda Theorem. J. Amer. Math. Soc. **3**(4), 31–116 (1990)
35. Lazarsfeld, R., Popa, M.: Derivative complex, BGG correspondence, and numerical inequalities for compact Kähler manifolds. Ivent. Math. **182**(3), 605–633 (2010)
36. Matsumura, H.: Commutative ring theory. In: Cambridge Studies in Advanced Mathematics, vol. 8. Cambridge University Press, Cambridge (1986)
37. Ogoma, T.: A note on the syzygy problem. Comm. Algebra **17**(8), 2061–2066 (1989)
38. Pareschi, G., Popa, M.: Strong generic vanishing and a higher-dimensional Castelnuovo-de Franchis inequality. Duke J. **150**(2), 269–285 (2009)
39. Peskine, C., Szpiro, L.: Dimension projective finie et cohomologie locale. Inst. Hautes Études Sci. Publ. Math. **42**, 49–119 (1973)
40. Roberts, P.: Le théorèm dintersection. C. R. Acad. Sci. Paris Sér. I Math. **304**, 177–180 (1987)
41. Serre, J.-P.: Modules Projectifs et Espaces Fibrés à Fibre Vectorielle. Seminaire P. Dubreil, Expose 23, Paris (1957–58)
42. Serre, J.-P.: Sur les modules projectifs. Seminaire Dubriel-Pisot 2, Expose 13, Paris (1960–61)
43. Serre, J.-P.: Algebra Locale-Multiplicites. In: Lecture Notes in Mathematics, vol. 11. Springer, Berlin (1965)
44. Shamash, J.: The Poincaré series of a local ring. J. Algebra **12**, 453–470 (1969)
45. Simon, A.-M.: A corollary to the Evans-Griffith syzygy theorem. Rend. Sem. Mat. Univ. Padova **94**, 71–77 (1995)

Moduli of Abelian Varieties, Vinberg θ-Groups, and Free Resolutions

Laurent Gruson, Steven V. Sam, and Jerzy Weyman

Introduction

The contents of these notes sit at the crossroads of representation theory, algebraic geometry, and commutative algebra, so we will explain each of these perspectives on our work before getting into the details.

From representation theory, we are considering the problem of classifying orbits in Vinberg θ-representations (G, U) [56]. From the point of view of (geometric) invariant theory these are the representations that are the simplest. One naturally gets a representation from a $\mathbf{Z}$-grading of a Kac–Moody algebra $\mathfrak{g}$, and so one naturally gets a trichotomy of these representations according to the structure of $\mathfrak{g}$: finite type, affine type, and wild type. The θ-representations come from finite and affine type. The θ-representations of finite type have finitely many orbits and include many cases of classical interest, such as determinantal varieties. The study of the geometry and algebra of these orbits is undertaken in work in progress by Kraśkiewicz–Weyman (starting with [38]). In the affine case, there are typically infinitely many orbits, but the θ-representation (G, U) has the property that the ring of semi-invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring and its unstable locus (nullcone) has finitely many orbits. This class of representations includes the adjoint representations of semisimple Lie algebras and share many features in common with them.

L. Gruson
Université de Versailles Saint-Quentin-en-Yvelines, Versailles, France
e-mail: laurent.gruson@math.uvsq.fr

S.V. Sam
University of California, Berkeley, CA, USA
e-mail: svs@math.berkeley.edu

J. Weyman (✉)
Northeastern University, Boston, MA, USA
e-mail: j.weyman@neu.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_13,

From the point of view of algebraic geometry, we are giving geometric constructions of (torsors of) Abelian varieties from the data of a G-orbit in U. The GIT quotient $U/\!\!/G$ is isomorphic to a quotient space $\mathfrak{h}/W$ where $W \subset \mathbf{GL}(\mathfrak{h})$ is a complex reflection group. In most of the cases that we consider, the quotients $\mathfrak{h}/W$ were considered previously and contain an open subset isomorphic to a moduli space of curves with a special kind of marked data. However, the constructions for Abelian varieties that we give seem to be new. One of the nice features of our work is that it is possible to make a lot of it explicit and turn it into input for the computer algebra system Macaulay2. For example, in the case of $\bigwedge^3 V$ (V is a vector space of dimension 9), we give in Sect. 5.3 a detailed explanation of how to calculate the ideals of the degeneracy loci under study (which include $(3,3)$-polarized Abelian surface torsors) starting with a $\mathbf{GL}(V)$-orbit.

From commutative algebra, this circle of ideas illustrates the power of a systematic use of perfect resolutions. As we will soon explain, the main tool that we employ is the Eagon–Northcott generic perfection theorem: using the minimal free resolutions of Kraśkiewicz–Weyman (many of which are perfect resolutions), we can construct certain global sheafy complexes which specialize to locally free resolutions of some varieties of interest, such as the Abelian variety torsors mentioned above.

In particular, for θ-representations of affine type, the GIT quotient $U/\!\!/G$ is a weighted projective space, and hence rational, and it is of interest to know if the orbit space U/G has a modular interpretation. The problem of studying the nilpotent orbits requires a different kind of approach and will not be discussed in these notes. For some of the examples of θ-representations, it is easy to find such interpretations using constructions from standard linear algebra, such as determinants. However, many of them do not seem to have any obvious constructions associated with them. We take up a systematic approach to dealing with these representations which we now outline (see Construction 3 for details). The main idea is to use information of the orbits in representations of finite type to bootstrap to the affine type case.

1. Using the Borel–Weil construction, one may realize U as the space of sections of a homogeneous bundle $\mathcal{U}$ on a homogeneous space G/P. This can usually be done in many different ways. For any point $x \in G/P$, the stabilizer of x is a subgroup in G which is conjugate to P, and for our choices of P, the action of P on the fiber $\mathcal{U}(x)$ will have finitely many orbits (and one can restrict the action to a certain reductive group $G' \subset P$ without affecting the orbit structure).
2. These orbits can be glued together to get "global orbit closures" in the total space of $\mathcal{U}$. Any vector $v \in U$ is then a section of $\mathcal{U}$ and each of these global orbit closures gives "degeneracy loci" in G/P by considering when $v(G/P)$ intersects a given global orbit closure. Hence the first step to understanding the geometry of these degeneracy loci should be to understand the "local" geometry of orbit closures in representations with finitely many orbits.
3. Vinberg's theory allows one to completely classify the orbits in these finite type representations. In many cases, one can calculate the minimal free

resolutions of the coordinate rings of these orbit closures and various equivariant modules which are of interest. This program is taken up in the work of Kraśkiewicz–Weyman.

4. For generic sections, the free resolutions can be turned into locally free resolutions for the degeneracy loci in question via the Eagon–Northcott generic perfection theorem. The power in this approach is that it allows one to obtain cohomological information about the varieties. This can sometimes be enough to determine the structure of the variety, such as showing it is an Abelian variety (which happens in many cases that we consider).

In the case of a representation with finitely many orbits, there is no interesting moduli over an algebraically closed field, but it is often interesting to consider **Z**-forms of the group and its representation and to study the arithmetic orbits. This has been done in a series of papers by Bhargava [6]. Under that perspective, a later goal would be to understand the arithmetic orbits for the representations that we consider. Some examples of representations which have positive-dimensional quotients have been worked out by Ho in her thesis [30].

For the contents of this chapter: Sects. 1–2 are introductory in nature and provide background on free resolutions and the requisite representation theory that will be used in the notes. Section 3 contains some geometric preliminaries on Abelian varieties and moduli spaces of vector bundles on curves, as well as a more precise description of the steps outlined above. The rest of the notes are devoted to studying examples of θ-representations of affine type.

Notation and Conventions

If R is a graded ring, then $R(d)$ is the R-module R with a grading shift: $R(d)_i = R_{d+i}$. We define a **graded local ring** to be a positively graded ring R whose degree 0 part is a field. In this case, we denote $\mathfrak{m} = \bigoplus_{i>0} R_i$. Given a free K-module E, we think of $\mathrm{Sym}(E)$ as a graded ring by $\mathrm{Sym}(E)_i = S^i E$. Furthermore, when we talk about modules over graded rings, we will implicitly assume that they are also graded.

If X is a scheme over a field K and E is a vector space, we let $\underline{E}$ denote the trivial vector bundle $E \otimes \mathcal{O}_X$. Also, given a vector bundle $\mathcal{E}$, we let $\det \mathcal{E}$ denote its top exterior power.

For the sections involving examples, we will work over an algebraically closed field of characteristic 0, which we will just denote by the complex numbers **C**. Ultimately, the goal is to relax this assumption to other characteristics, or non-algebraically closed fields, so the introductory sections are written in this more general context. We will also make some comments throughout about this point.

1 Free Resolutions

1.1 Basic Definitions

A lot of the foundational results in this section can be found in [10, Appendix].

Definition 1. Let R be a commutative ring and M be a finitely generated R-module. A complex of R-modules

$$\mathbf{F}_\bullet : \cdots \to \mathbf{F}_i \xrightarrow{d_i} \mathbf{F}_{i-1} \to \cdots \to \mathbf{F}_0$$

is a **projective resolution** of M if:

- Each $\mathbf{F}_i$ is a finitely generated projective R-module.
- $\mathrm{H}_i(\mathbf{F}_\bullet) = 0$ for $i > 0$ and $\mathrm{H}_0(\mathbf{F}_\bullet) = M$.

The **projective dimension** of M (denoted $\operatorname{pdim} M$) is the minimum length of any projective resolution of M. **Free resolutions** are projective resolutions where the $\mathbf{F}_i$ are free modules. If R is a (graded) local ring with maximal ideal $\mathfrak{m}$, then $\mathbf{F}_\bullet$ is **minimal** if:

- $d_i(\mathbf{F}_i) \subseteq \mathfrak{m}\mathbf{F}_{i-1}$ for all $i > 0$.

In the graded case, we will shift the gradings of the $\mathbf{F}_i$ to assume that the differentials are homogeneous of degree 0.

Definition 2. Let R be a Noetherian ring and let M be a finitely generated R-module. A sequence $(r_1, \ldots, r_n)$ of elements in R is a **regular sequence** on M if:

- r_1 is not a zero divisor or unit on M.
- r_i is not a zero divisor or unit on $M/(r_1, \ldots, r_{i-1})M$ for all $i > 1$.

For an ideal $I \subset R$, the **depth** of M (with respect to I) is the length of the longest regular sequence for M which is contained in I. It is denoted by $\operatorname{depth}_I M$. If R is local with maximal ideal $\mathfrak{m}$, we denote $\operatorname{depth} M = \operatorname{depth}_{\mathfrak{m}} M$. For an ideal $I \subset R$, the **grade** of I is the length of the longest regular sequence in I for R. M is **perfect of grade** g if $g = \operatorname{pdim} M = \operatorname{grade} \operatorname{Ann} M$. (In general, one has $\operatorname{pdim} M \geq \operatorname{grade} \operatorname{Ann} M$.)

Over a local Noetherian ring R, a finitely generated module M is **Cohen–Macaulay** if it is 0 or $\operatorname{depth} M = \dim M := \dim(R/\operatorname{Ann}(M))$. For a general Noetherian ring R, M is Cohen–Macaulay if the localization $M_\mathfrak{p}$ is Cohen–Macaulay over $(R_\mathfrak{p}, \mathfrak{p})$ for all prime ideals $\mathfrak{p}$ of R. A Noetherian ring is Cohen–Macaulay if it is a Cohen–Macaulay module over itself.

Theorem 3. *Let R be a Noetherian Cohen–Macaulay ring.*

1. For every ideal $I \subset R$, we have $\operatorname{grade} I = \operatorname{codim} I = \dim R - \dim(R/I)$.

2. *The polynomial ring $R[x]$ is Cohen–Macaulay.*
3. *If an R-module M is perfect, then it is Cohen–Macaulay.*

(The distinction between perfect and Cohen–Macaulay for a module over a Cohen–Macaulay ring is the property of having finite projective dimension.)

Theorem 4 (Auslander–Buchsbaum Formula). *Suppose R is a Noetherian (graded) local ring and that M is a finitely generated R-module with* $\operatorname{pdim} M < \infty$. *Then*

$$\operatorname{depth} M + \operatorname{pdim} M = \operatorname{depth} R.$$

Theorem 5. *Let R be a Noetherian (graded) local ring and M be a perfect R-module of grade g with minimal free resolution $\mathbf{F}_\bullet$. Then* $\operatorname{Hom}(\mathbf{F}_\bullet, R)$ *is a minimal free resolution of the perfect module* $M^\vee = \operatorname{Ext}^g_R(M, R)$*, and* $(M^\vee)^\vee \cong M$.

Definition 6. If $M = R/I$, for an ideal I, is perfect, then we write $\omega_{R/I} = M^\vee$ and call it the **canonical module** of R/I. If R/I is perfect and $\omega_{R/I} \cong R/I$ (ignoring grading if it is present), then we say that I is a **Gorenstein ideal**. This is equivalent to the last term in the minimal free resolution of R/I having rank 1.

Theorem 7 (Eagon–Northcott Generic Perfection). *Let R be a Noetherian ring and M a perfect R-module of grade g, and let $\mathbf{F}_\bullet$ be an R-linear free resolution of M of length g. Let S be a Noetherian R-algebra. If $M \otimes_R S \neq 0$ and* $\operatorname{grade}(M \otimes_R S) \geq g$*, then $M \otimes_R S$ is perfect of grade g and $\mathbf{F}_\bullet \otimes_R S$ is an S-linear free resolution of $M \otimes_R S$. If $M \otimes_R S = 0$, then $\mathbf{F}_\bullet \otimes_R S$ is exact.*

See [10, Theorem 3.5]. It is natural to ask what happens if the grade of $M \otimes_R S$ is some value less than g, and this can be answered by the Buchsbaum–Eisenbud acyclicity criterion [20, Theorem 20.9].

Remark 8. In particular, if R is Cohen–Macaulay (e.g., $R = K[x_1, \ldots, x_n]$), then we can replace grade in the above theorem with codimension. It is often much easier to calculate codimension. We will use it as follows. We first construct graded minimal free resolutions of perfect modules M over $A = K[x_1, \ldots, x_n]$. Then we specialize the variables x_i to elements of a Cohen–Macaulay K-algebra S in such a way that the codimension of M is preserved. Then the resulting specialized complex is still a resolution.

1.2 Examples

For this section, let K be a commutative ring and let E be a free module of rank N. In the following examples, we will construct some complexes that are functorial in E and compatible with change of rings. Two consequences of these properties is that the complex carries an action of the general linear group $\mathbf{GL}(E)$ and that the constructions make sense for vector bundles over an arbitrary scheme.

Example 9. Write $A = \mathrm{Sym}(E)$. We define a complex $\mathbf{F}_\bullet$ by setting $\mathbf{F}_i = \bigwedge^i E \otimes A(-i)$. The differential is defined by

$$\bigwedge^i E \otimes A(-i) \to \bigwedge^{i-1} E \otimes A(-i+1)$$

$$e_1 \wedge \cdots \wedge e_i \otimes f \mapsto \sum_{j=1}^{i} (-1)^j e_1 \wedge \cdots \widehat{e_j} \cdots \wedge e_i \otimes e_j f.$$

This is the **Koszul complex**. It is a resolution of $K = A/\mathfrak{m}$ and is functorial with respect to E. See [20, Chap. 17] for basic properties.

Example 10 (Buchsbaum–Eisenbud). We assume $N = 2n+1$ is odd. Set $A = \mathrm{Sym}(\bigwedge^2 E)$, which we can interpret as the coordinate ring of the space of all skew-symmetric matrices of size $2n+1$ with entries in K if we fix a basis $e_1, \dots, e_{2n+1}$ of E. Let Φ be the generic skew-symmetric matrix of size $2n+1$ whose (i, j) entry is $x_{ij} = e_i \wedge e_j \in A_1$. We construct a complex

$$\mathbf{F}_\bullet : 0 \to (\det E)^{\otimes 2} \otimes A(-2n-1) \to (\det E) \otimes E \otimes A(-n-1)$$
$$\to \bigwedge^{2n} E \otimes A(-n) \to A.$$

For $j = 1, \dots, 2n+1$, let $e'_j = e_1 \wedge \cdots \hat{e}_j \cdots \wedge e_{2n+1}$. We also define $\mathrm{Pf}(\hat{j})$ to be the Pfaffian of the submatrix of Φ obtained by deleting row and column j. Then we have

$$\bigwedge^{2n} E \otimes A(-n) \xrightarrow{d_1} A$$
$$e'_j \otimes f \mapsto \mathrm{Pf}(\hat{j}) f,$$

$$(\det E) \otimes E \otimes A(-n-1) \xrightarrow{d_2} \bigwedge^{2n} E \otimes A(-n)$$

$$(e_1 \wedge \cdots \wedge e_{2n+1}) \otimes e_j \otimes f \mapsto \sum_{i=1}^{2n+1} (-1)^i e'_i \otimes x_{ij} f,$$

$$(\det E)^{\otimes 2} \otimes A(-2n-1) \xrightarrow{d_3} (\det E) \otimes E \otimes A(-n-1)$$

$$(e_1 \wedge \cdots \wedge e_{2n+1})^2 \otimes f \mapsto (e_1 \wedge \cdots \wedge e_{2n+1}) \otimes \sum_{j=1}^{2n+1} (-1)^j \mathrm{Pf}(\hat{j}) e_j f.$$

This is the **Buchsbaum–Eisenbud complex**. It is a resolution of A/I where I is the ideal generated by the $2n \times 2n$ Pfaffians of Φ, and it is functorial with respect to E. Furthermore, A/I is a free K-module, and I is a Gorenstein ideal of codimension 3.

We can identify d_2 with the map Φ. Buchsbaum and Eisenbud showed that given a codimension 3 Gorenstein ideal I, there is an n such that its free resolution is a specialization of the above complex. See [11, Theorem 2.1] for more details.

Example 11 (Józefiak–Pragacz). We assume $N = 2n$ is even. Again we set $A = \mathrm{Sym}(\bigwedge^2 E)$ and let Φ be the generic skew-symmetric matrix. We will give the resolution for the ideal generated by the Pfaffians of size $2n-2$. We just give the functorial terms in the complex when K contains the field of rational numbers (the definitions of the Schur functors $\mathbf{S}$ are given in Sect. 2.1):

$$\begin{aligned}
\mathbf{F}_0 &= A \\
\mathbf{F}_1 &= \bigwedge^{2n-2} E \otimes A(-n+1) \\
\mathbf{F}_2 &= \mathbf{S}_{2,1^{2n-2}} E \otimes A(-n) \\
\mathbf{F}_3 &= (\det E) \otimes S^2 E \otimes A(-n-1) \oplus (\det E)^2 \otimes (S^2 E)^* \otimes A(-2n+1) \\
\mathbf{F}_4 &= (\det E) \otimes \mathbf{S}_{2,1^{2n-2}} E \otimes A(-2n) \\
\mathbf{F}_5 &= (\det E)^2 \otimes \bigwedge^2 E \otimes A(-2n-1) \\
\mathbf{F}_6 &= (\det E)^3 \otimes A(-3n).
\end{aligned}$$

When K is an arbitrary commutative ring, the functors must be defined differently, but the ranks of the modules remain the same. We refer the reader to [51] for the details.

Example 12 (Goto–Józefiak–Tachibana). We set $A = \mathrm{Sym}(S^2 E)$, which we can interpret as the coordinate ring of the space of symmetric matrices of size N. We let Φ be the generic symmetric matrix. We give the terms of the resolution of the ideal generated by the minors of size $N-1$:

$$\begin{aligned}
\mathbf{F}_0 &= A \\
\mathbf{F}_1 &= (\det E)^2 \otimes (S^2 E)^* \otimes A(-N+1) \\
\mathbf{F}_2 &= (\det E)^2 \otimes \ker(E \otimes E^* \xrightarrow{\text{eval}} K) \otimes A(-N) \\
\mathbf{F}_3 &= (\det E)^2 \otimes \bigwedge^2 E \otimes A(-N-1).
\end{aligned}$$

See [32, Sect. 3] for details. Over a field of characteristic 0, the term $\det E \otimes \ker(E \otimes E^* \xrightarrow{\text{eval}} K)$ can be replaced by the Schur functor $\mathbf{S}_{2,1^{n-2}} E$ (see Sect. 2.1 for the definition).

1.3 The Geometric Technique

The material in this section is not logically necessary for the rest of this chapter, but it is the main tool behind the work of Kraśkiewicz–Weyman, so we include it for completeness. For a reference, see [58, Chap. 5]. Note that we have changed notation.

Let K be a field, let X be a projective K-variety, and let V be a vector space. Suppose we are given a short exact sequence of locally free sheaves

$$0 \to \mathcal{S} \to \underline{V} \to \mathcal{T} \to 0.$$

We let $p_1 \colon \underline{V} \to V$ and $p_2 \colon \underline{V} \to X$ be the projection maps. Set $Y = p_1(\mathcal{S}) \subset V$ and $A = \mathcal{O}_V = \mathrm{Sym}(V^*)$. Note that Y is the affine cone over some projective variety in $\mathbf{P}(V)$. Also, let $\mathcal{E}$ be any vector bundle on X.

Theorem 13. *There is a minimal A-linear complex $\mathbf{F}_\bullet$ whose terms are*

$$\mathbf{F}_i = \bigoplus_{j \geq 0} \mathrm{H}^j(X; \bigwedge^{i+j}(\mathcal{T}^*) \otimes \mathcal{E}) \otimes A(-i-j).$$

Furthermore, $\mathrm{H}_i(\mathbf{F}_\bullet) = 0$ *for* $i > 0$ *and for* $i \leq 0$*, we have*

$$\mathrm{H}_i(\mathbf{F}_\bullet) = \mathrm{R}^{-i}(\mathcal{O}_{\mathcal{S}} \otimes_{\mathcal{O}_{\underline{V}}} p_2^*\mathcal{E}) = \mathrm{H}^{-i}(X; \mathrm{Sym}(\mathcal{S}^*) \otimes_{\mathcal{O}_X} \mathcal{E}).$$

In particular, if the higher direct images of $\mathcal{O}_{\mathcal{S}} \otimes p_2^*\mathcal{E}$ vanish, then $\mathbf{F}_\bullet$ is a free resolution of the pushforward. In the case that $\mathcal{E} = \mathcal{O}_X$, this pushforward is an A-algebra. The vanishing of the higher direct images is an intrinsic property of the variety Y.

The idea behind this theorem is to start with an affine cone variety Y and to find X and $\mathcal{S}$ that fit into the above framework. Then the theorem above gives a tool for calculating the minimal free resolution of Y.

Example 14 (Eagon–Northcott Complex). Let E and F be vector spaces of dimensions m and n and assume that $m \geq n$. We set $V = \mathrm{Hom}(E, F)$ and let $Y \subset V$ be the subvariety of linear maps of rank at most $n-1$.

This fits into the previous setup by taking $X = \mathbf{Gr}(n-1, F) \cong \mathbf{P}(F^*)$. This has a tautological exact sequence of vector bundles

$$0 \to \mathcal{R} \to \underline{F} \to \mathcal{O}(1) \to 0$$

where $\mathcal{R} = \{ (x, W) \in F \times X \mid x \in W \}$. Then we can take $\mathcal{S} = \mathcal{H}om(\underline{E}, \mathcal{R}) = \underline{E}^* \otimes \mathcal{R}$. We will see in Sect. 2.3 that the higher direct images of $\mathrm{Sym}(\mathcal{S}^*)$ vanish and so $\mathbf{F}_\bullet$ gives a minimal free resolution for $\mathcal{O}_Y$.

Since $\mathcal{T}^* = \underline{E} \otimes \mathcal{O}(-1)$, we have $\bigwedge^d \mathcal{T}^* = \bigwedge^d \underline{E} \otimes \mathcal{O}(-d)$. So we can calculate the terms of $\mathbf{F}_\bullet$ explicitly. For $i > 0$, we have

$$\mathbf{F}_i = \bigwedge^{n+i-1} E \otimes \det F^* \otimes (S^{i-1} F)^* \otimes A(-n-i+1).$$

The differentials can be calculated by noting that they will preserve the natural $\mathbf{GL}(E) \times \mathbf{GL}(F)$ action on V and Y and that equivariant maps of this form are unique up to a choice of scalar. In fact, this construction works with $K = \mathbf{Z}$, in which case we get uniqueness of scalars up to a choice of sign. A multilinear generalization of this complex, constructed using similar ideas, can be found in [5].

2 Representation Theory

2.1 Schur Functors

For the material in this section, see [58, Chap. 2]. What we call $\mathbf{S}_\lambda$ is denoted by $\mathbf{L}_{\lambda'}$ there.

Definition 1. A partition λ is a decreasing sequence of positive integers $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. We represent this as a Young diagram by drawing λ_i boxes left-justified in the ith row, starting from top to bottom. The dual partition λ' is obtained by letting λ'_i be the number of boxes in the ith column of λ. Given a box $b = (i, j) \in \lambda$, its **content** is $c(b) = j - i$, and its **hook length** is $h(b) = \lambda_i - i + \lambda'_j - j + 1$. If we have a sequence $(i, i, \dots, i)$ repeated j times, we abbreviate by the notation (i^j).

Example 2. Let $\lambda = (4, 3, 1)$. Then $\lambda' = (3, 2, 2, 1)$. The contents and hook lengths are given as follows:

c :

0	1	2	3
−1	0	1	
−2			

h :

6	4	3	1
4	2	1	
1			

Definition 3. Let R be a commutative ring and E a free R-module. Let λ be a partition with n parts and write $m = \lambda_1$. We use $S^n E$ to denote the nth symmetric power of E. The **Schur functor** $\mathbf{S}_\lambda(E)$ is the image of the map

$$\bigwedge^{\lambda'_1} E \otimes \cdots \otimes \bigwedge^{\lambda'_m} E \xrightarrow{\Delta} E^{\otimes \lambda'_1} \otimes \cdots \otimes E^{\otimes \lambda'_m} = E^{\otimes \lambda_1} \otimes \cdots \otimes E^{\otimes \lambda_n}$$
$$\xrightarrow{\mu} S^{\lambda_1} E \otimes \cdots \otimes S^{\lambda_n} E,$$

where the maps are defined as follows. First, Δ is the product of the comultiplication maps $\bigwedge^i E \to E^{\otimes i}$ given by $e_1 \wedge \cdots \wedge e_i \mapsto \sum_{w \in \Sigma_i} \mathrm{sgn}(w) e_{w(1)} \otimes \cdots \otimes e_{w(i)}$. The equals sign is interpreted as follows: pure tensors in $E^{\otimes \lambda'_1} \otimes \cdots \otimes E^{\otimes \lambda'_m}$ can be interpreted as filling the Young diagram of λ with vectors along the columns, which can be thought of as pure tensors in $E^{\otimes \lambda_1} \otimes \cdots \otimes E^{\otimes \lambda_n}$ by reading via rows. Finally, μ is the multiplication map $E^{\otimes i} \to S^i E$ given by $e_1 \otimes \cdots \otimes e_i \mapsto e_1 \cdots e_i$.

In particular, note that $\mathbf{S}_\lambda E = 0$ if the number of parts of λ exceeds rank E.

Example 4. Take $\lambda = (3, 2)$. Then the map is given by

$$\begin{aligned}(e_1 \wedge e_2) \otimes (e_3 \wedge e_4) \otimes e_5 &\mapsto \begin{array}{|c|c|c|}\hline e_1 & e_3 & e_5 \\ \hline e_2 & e_4 \\ \cline{1-2}\end{array} - \begin{array}{|c|c|c|}\hline e_2 & e_3 & e_5 \\ \hline e_1 & e_4 \\ \cline{1-2}\end{array} - \begin{array}{|c|c|c|}\hline e_1 & e_4 & e_5 \\ \hline e_2 & e_3 \\ \cline{1-2}\end{array} + \begin{array}{|c|c|c|}\hline e_2 & e_4 & e_5 \\ \hline e_1 & e_3 \\ \cline{1-2}\end{array} \\ &\mapsto (e_1 e_3 e_5 \otimes e_2 e_4) - (e_2 e_3 e_5 \otimes e_1 e_4) - (e_1 e_4 e_5 \otimes e_2 e_3) \\ &\quad + (e_2 e_4 e_5 \otimes e_1 e_3).\end{aligned}$$

Theorem 5. *The Schur functor* $\mathbf{S}_\lambda E$ *is a free* R*-module. If* $\operatorname{rank} E = n$*, then*

$$\operatorname{rank} \mathbf{S}_\lambda E = \prod_{b \in \lambda} \frac{n + c(b)}{h(b)}.$$

The construction of $\mathbf{S}_\lambda E$ is functorial with respect to E. This has two consequences: $\mathbf{S}_\lambda E$ is naturally a representation of $\mathbf{GL}(E)$, and we can also construct $\mathbf{S}_\lambda \mathcal{E}$ when $\mathcal{E}$ is a vector bundle.

If $\operatorname{rank} E = n$, then we have $\mathbf{S}_\lambda E \otimes \det E = \mathbf{S}_{(\lambda_1+1,\ldots,\lambda_n+1)} E$. Using this, it makes sense to define $\mathbf{S}_\lambda E$ when λ is any weakly decreasing sequence of integers. Furthermore, we have $\mathbf{S}_\lambda (E^*) = \mathbf{S}_{-\lambda_n,\ldots,-\lambda_1} E$ [58, Exercise 2.18] and over a field of characteristic 0, the isomorphism $\mathbf{S}_\lambda (E^*) = (\mathbf{S}_\lambda E)^*$ [58, Proposition 2.1.18].

2.2 Descriptions of Some Homogeneous Spaces

Let E be a vector space of rank N. We let $\mathbf{Fl}(E)$ be the flag variety of E. Its K-valued points are complete flags of subspaces $E_\bullet : E_1 \subset E_2 \subset \cdots E_N = E$ such that $\operatorname{rank} E_i = i$. The trivial bundle $\underline{E}$ contains a tautological flag of subbundles $\mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_N = \underline{E}$ where

$$\mathcal{R}_i = \{ (x, E_\bullet) \in E \times \mathbf{Fl}(E) \mid x \in E_i \}.$$

Given a subset $S \subset \{ 1, \ldots, N-1 \}$, we can also consider the partial flag varieties $\mathbf{Fl}(S; E)$ whose K-valued points only are partial flags whose ranks are the elements in S. Then $\underline{E}$ has a tautological partial flag of subbundles $\mathcal{R}_i$ $(i \in S)$.

Now assume that E is equipped with a symplectic form ω and set $n = N/2$. We say that a subspace $U \subset E$ is isotropic if $\omega(u, u') = 0$ for all $u, u' \in U$. Also, for any subspace U, we set $U^{\perp} = \{ x \in E \mid \omega(x,u) = 0 \text{ for all } u \in U \}$. Then $\operatorname{rank} U + \operatorname{rank} U^{\perp} = \operatorname{rank} E$ and U is isotropic if and only if $U \subseteq U^{\perp}$.

We define the symplectic flag variety to be the subvariety $\mathbf{Fl}_{\omega}(E)$ of the partial flag variety $\mathbf{Fl}(1,\dots,n;E)$ consisting of flags $E_1 \subset E_2 \subset \cdots \subset E_n$ such that each E_i is isotropic. We also let $\mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_n$ denote the restriction of the tautological bundles to $\mathbf{Fl}_{\omega}(E)$. We can also define the subbundles $\mathcal{R}_i^{\perp} = \{ (x, E_{\bullet}) \in E \times \mathbf{Fl}_{\omega}(E) \mid x \in E_i^{\perp} \}$. Note that we have

$$0 \subset \mathcal{R}_1 \subset \mathcal{R}_2 \subset \cdots \subset \mathcal{R}_n = \mathcal{R}_n^{\perp} \subset \mathcal{R}_{n-1}^{\perp} \subset \cdots \subset \mathcal{R}_1^{\perp} \subset \underline{E}.$$

Given a subset $S \subset \{ 1,\dots,n \}$, we can also define partial ω-isotropic flag varieties $\mathbf{Fl}_{\omega}(S;E)$. When $S = \{ i \}$ is a singleton, we also write $\mathbf{Gr}_{\omega}(i,E) = \mathbf{Fl}_{\omega}(\{ i \};E)$ and call it the ω-isotropic Grassmannian.

2.3 Borel–Weil–Bott Theorem

For the material in this section, see [58, Chap. 4].

Let $S = \{ i_1 < i_2 < \cdots < i_k \}$ be a subset of $\{ 1,\dots,N \}$ and consider the partial flag variety $\mathbf{Fl}(S;E)$. For each $j = 1,\dots,k+1$, let $\lambda^{(j)}$ be a weakly decreasing sequence of integers of length $i_j - i_{j-1}$ (set $i_0 = 0$ and $i_{k+1} = N$). Let λ be the sequence obtained by concatenating $\lambda^{(1)}, \lambda^{(2)}, \dots, \lambda^{(k+1)}$. Set

$$\mathcal{R}(\lambda) = \bigotimes_{j=1}^{k+1} \mathbf{S}_{\lambda^{(j)}}((\mathcal{R}_{i_j}/\mathcal{R}_{i_{j-1}})^*).$$

Theorem 6 (Borel–Weil). *If λ is a weakly decreasing sequence of integers, then we have a $\mathbf{GL}(E)$-equivariant isomorphism*

$$\mathrm{H}^0(\mathbf{Fl}(S;E);\mathcal{R}(\lambda)) = \mathbf{S}_{\lambda}(E^*),$$

and all higher cohomology of $\mathcal{R}(\lambda)$ vanishes.

Now we consider the analogue for the symplectic group. Let λ be a partition with at most n parts. We restrict the line bundle $\mathcal{R}(\lambda)$ on the flag variety $\mathbf{Fl}(E)$ to the symplectic flag variety $\mathbf{Fl}_{\omega}(E)$ and denote its space of sections by

$$\mathbf{S}_{[\lambda]}(E) = \mathrm{H}^0(\mathbf{Fl}_{\omega}(E);\mathcal{R}(\lambda)). \tag{2.7}$$

By considering a relative situation, this definition makes sense for any vector bundle E over a scheme X which is equipped with a nondegenerate skew-symmetric $\mathcal{O}_X$-bilinear form $\bigwedge^2 E^* \to \mathcal{O}_X$. (More generally, the form could take values in a line bundle, but we will not use this generalization.) For all tautological subbundles $\mathcal{R}$

on any symplectic flag variety, the restriction of the symplectic form to the quotient bundle $\mathcal{R}^{\perp}/\mathcal{R}$ is nondegenerate.

Now let $S = \{ i_1 < \cdots < i_k \}$ be a subset of $\{ 1, \dots, n \}$ and let $\lambda^{(j)}$ be as before, except now we assume that $\lambda^{(k+1)}$ only has nonnegative integers. We set

$$\mathcal{R}[\lambda] = \mathbf{S}_{[\lambda^{(k+1)}]}(\mathcal{R}_{i_k}^{\perp}/\mathcal{R}_{i_k}) \otimes \bigotimes_{j=1}^{k} \mathbf{S}_{\lambda^{(j)}}((\mathcal{R}_{i_j}/\mathcal{R}_{i_{j-1}})^*).$$

Theorem 7 (Borel–Weil). *If λ is a weakly decreasing sequence of nonnegative integers, then we have a $\mathbf{Sp}(E)$-equivariant isomorphism*

$$\mathrm{H}^0(\mathbf{Fl}_\omega(S; E); \mathcal{R}[\lambda]) = \mathbf{S}_{[\lambda]} E$$

and all higher cohomology of $\mathcal{R}[\lambda]$ vanishes.

Now we discuss what happens when λ is not a weakly decreasing sequence of integers. For this, we now need to assume that the characteristic of the ground field is 0. We first handle the partial flag varieties.

Define the vector $\rho = (N-1, N-2, \dots, 1, 0)$. Given a permutation $w \in \Sigma_N$, we define the dotted action of w on a sequence of integers α of length N by

$$w^\bullet(\alpha) = w(\alpha + \rho) - \rho.$$

We define the length of a permutation w to be $\ell(w) = \#\{ (i, j) \mid i < j,\ w(i) > w(j) \}$. Alternatively, let s_i be the simple transposition that swaps i and $i+1$. Then $s_1, \dots, s_{N-1}$ generate Σ_N, and we could also define $\ell(w)$ to be the minimal number k so that we can write $w = s_{i_1} \cdots s_{i_k}$.

Theorem 8 (Bott). *Assume that the field K has characteristic 0 and let λ be as before. Then exactly one of the following two cases occurs:*

- *There exists a non-identity $w \in \Sigma_N$ such that $w^\bullet(\lambda) = \lambda$. In this case, all cohomology of $\mathcal{R}(\lambda)$ vanishes.*
- *There is a unique $w \in \Sigma_N$ such that $w^\bullet(\lambda) = \mu$ is a decreasing sequence. In this case, we have a $\mathbf{GL}(E)$-equivariant isomorphism*

$$\mathrm{H}^{\ell(w)}(\mathbf{Fl}(S; E); \mathcal{R}(\lambda)) = \mathbf{S}_\mu(E^*)$$

and all other cohomology of $\mathcal{R}(\lambda)$ vanishes.

An example that uses the previous theorem is given in the proof of Theorem 3.

Now we consider Bott's theorem for the symplectic flag varieties $\mathbf{Fl}_\omega(S; E)$. Now we set $\rho = (N, N-1, \dots, 2, 1)$. We replace the symmetric group Σ_N with the group of signed permutations $W = \Sigma_N \ltimes (\mathbf{Z}/2)^N$, which we think of as $N \times N$ signed permutation matrices.

We consider the generators $s_1, \dots, s_N$ for W. The meaning of $s_1, \dots, s_{N-1}$ is the same as for the symmetric group Σ_N, and s_N is the diagonal matrix $\mathrm{diag}(1, \dots, 1, -1)$. Now given $w \in W$, we define $\ell(w)$ to be the minimal number k so that we can write $w = s_{i_1} \cdots s_{i_k}$.

Then the definition of the dotted action of W remains the same, and the analogue of Bott's theorem holds for the bundles $\mathcal{R}[\lambda]$ on $\mathbf{Fl}_\omega(S; E)$.

Theorem 9 (Bott). *Assume that the field K has characteristic 0 and let λ be as before. Then exactly one of the following two cases occurs:*

- *There exists a non-identity $w \in W$ such that $w^\bullet(\lambda) = \lambda$. In this case, all cohomology of $\mathcal{R}[\lambda]$ vanishes.*
- *There is a unique $w \in \Sigma_N$ such that $w^\bullet(\lambda) = \mu$ is a weakly decreasing sequence of nonnegative integers. In this case, we have an $\mathbf{Sp}(E)$-equivariant isomorphism*

$$\mathrm{H}^{\ell(w)}(\mathbf{Fl}_\omega(S; E); \mathcal{R}[\lambda]) = \mathbf{S}_{[\mu]}(E^*)$$

and all other cohomology of $\mathcal{R}[\lambda]$ vanishes.

The symplectic form gives $\mathbf{Sp}(E)$-equivariant isomorphisms $\mathbf{S}_{[\mu]}(E^*) = \mathbf{S}_{[\mu]}E$ for all μ.

2.4 Vinberg θ-Representations

For this section, we refer to [33, 34, 56, 57] for reference.

Let X_n be a Dynkin diagram and let $\mathfrak{g}$ be the corresponding simple Lie algebra. Let us distinguish a node $x \in X_n$. Let α_k be a corresponding simple root in the root system Φ corresponding to X_n. The choice of α_k determines a $\mathbf{Z}$-grading on Φ by letting the degree of a root β be equal to the coefficient of α_k when we write β as a linear combination of simple roots. On the level of Lie algebras, this corresponds to a $\mathbf{Z}$-grading

$$\mathfrak{g} = \bigoplus_{i \in \mathbf{Z}} \mathfrak{g}_i.$$

We define the group $G_0 := (G, G) \times \mathbf{C}^*$ where (G, G) is a connected semisimple group with the Dynkin diagram $X_n \setminus x$. A representation of type I is the representation of G_0 on $\mathfrak{g}_1$, and a representation of type II is what we get when we replace X_n with an affine Dynkin diagram. This notation follows [34, Proposition 3.1].

Remark 10. In the case of a type II representation (i.e., when X_n is an affine Dynkin diagram), each θ-representation $(G_0, \mathfrak{g}_1)$ has a Chevalley isomorphism: there is a subspace $\mathfrak{h} \subset \mathfrak{g}_1$ and a complex reflection group $W \subset \mathbf{GL}(\mathfrak{h})$ (defined as the normalizer of $\mathfrak{h}$ modulo the centralizer of $\mathfrak{h}$), which we call the **graded Weyl group**, such that the restriction map is an isomorphism

$$\mathrm{Sym}(\mathfrak{g}_1^*)^{(G,G)} \xrightarrow{\cong} \mathrm{Sym}(\mathfrak{h}^*)^W.$$

In particular, $\mathrm{Sym}(\mathfrak{g}_1^*)^{(G,G)}$ is always a polynomial ring. The complex reflection groups were classified by Shephard–Todd in [54], and we will refer to their numbering system when using names for these groups.

In [56], Vinberg gave a description of the G_0-orbits in the representations of type I in terms of conjugacy classes of nilpotent elements in $\mathfrak{g}$. Let $e \in \mathfrak{g}_1$ be a nilpotent element in $\mathfrak{g}$. Consider the irreducible components of the intersection of the conjugacy class of e in $\mathfrak{g}$

$$C(e) \cap \mathfrak{g}_1 = C_1(e) \cup \cdots \cup C_{n(e)}(e).$$

The sets $C_i(e)$ are clearly G_0-stable. Vinberg's result shows that these are precisely the G_0-orbits in $\mathfrak{g}_1$.

Theorem 11 (Vinberg). *The G_0-orbits of the action of G_0 on $\mathfrak{g}_1$ are the components $C_i(e)$, for all choices of the conjugacy classes $C(e)$ and all i, $1 \le i \le n(e)$.*

Theorem 11 makes a connection between the orbits in $\mathfrak{g}_1$ and the nilpotent orbits in $\mathfrak{g}$. The classification of nilpotent orbits in simple Lie algebras was obtained by Bala and Carter in the papers [1]. A good account of this theory is the book [16]. Here we recall that the nilpotent orbit of an element e in a simple Lie algebra $\mathfrak{g}$ is characterized by the smallest Levi subalgebra $\mathfrak{l}$ containing e. One must be careful because sometimes $\mathfrak{l}$ is equal to $\mathfrak{g}$. If the element e is a principal element in $\mathfrak{l}$, then this orbit is denoted by the Dynkin diagram of $\mathfrak{l}$ (but there might be different ways in which the root system $R(\mathfrak{l})$ sits as a subroot system of $R(\mathfrak{g})$).

There are, however, the non-principal nilpotent orbits that are not contained in a smaller reductive Lie algebra $\mathfrak{l}$. These are called *the distinguished nilpotent orbits* and are described in [16, Sect. Sect. 8.2–8.4]. They are characterized by their associated parabolic subgroups (as their Dynkin characteristics are even, [16, Sect. 8]). Let us remark that for Lie algebras of classical types, for type A_n the only distinguished nilpotent orbits are the principal ones, and for types B_n, C_n, D_n these are orbits corresponding to the partitions with different parts. For exceptional Lie algebras the distinguished orbits can be read off the tables in [16, Sect. 8.4].

Theorem 11 is not easy to use because it is not very explicit. In the next section we describe a more precise method from another of Vinberg's papers [57].

2.5 The Vinberg Method for Classifying Orbits

In this section we describe the second paper of Vinberg [57] in which he describes orbits of nilpotent elements in $\mathfrak{g}_1$. Similar to the Bala–Carter classification, the nilpotent elements in $\mathfrak{g}_1$ are described by means of some graded subalgebras of $\mathfrak{g}$. We need some preliminary notions.

All Lie algebras $\mathfrak{g}$ we will consider will be Lie algebras of some algebraic group G.

Let (X_n, α_k) be one of the representations from the previous section. As before, it defines the grading

$$\mathfrak{g} = \bigoplus_{i \in \mathbf{Z}} \mathfrak{g}_i$$

where $\mathfrak{g}_i$ is the span of the roots which, written as a combination of simple roots, have α_k with coefficient i. The component $\mathfrak{g}_0$ contains a Cartan subalgebra. G_0 denotes the connected closed subgroup of G corresponding to $\mathfrak{g}_0$.

In the sequel, $Z(x)$ denotes the centralizer of an element $x \in G$ and $Z_0(x) = Z(x) \cap G_0$, and $\mathfrak{z}, \mathfrak{z}_0$ denote the corresponding Lie algebras. Similarly, $N(x)$ denotes the normalizer of an element $x \in G$ and $N_0(x) = N(x) \cap G_0$.

We let $R(\mathfrak{g})$ denote the set of roots of $\mathfrak{g}$, and $\Pi(\mathfrak{g})$ denotes a set of simple roots.

Definition 12. A graded Lie subalgebra $\mathfrak{s} \subset \mathfrak{g}$ is **regular** if it is normalized by a maximal torus in $\mathfrak{g}_0$. A reductive graded Lie algebra $\mathfrak{s} \subset \mathfrak{g}$ is **complete** if it is not a proper graded Lie subalgebra of any regular reductive $\mathbf{Z}$-graded Lie algebra of the same rank.

Definition 13. A $\mathbf{Z}$-graded Lie algebra $\mathfrak{g}$ is **locally flat** if any of the following equivalent conditions is satisfied, for e a point in general position in $\mathfrak{g}_1$:

1. The subgroup $Z_0(e)$ is finite.
2. $\mathfrak{z}_0(e) = 0$.
3. $\dim \mathfrak{g}_0 = \dim \mathfrak{g}_1$.

Fix a nonzero nilpotent element $e \in \mathfrak{g}_a$, and choose some maximal torus H in $N_0(e)$. Its Lie algebra $\mathfrak{h}$ is **the accompanying torus** of the element e. We denote by φ the character of the torus H defined by the condition

$$[u, e] = \varphi(u)e$$

for $u \in \mathfrak{h}$. Consider the graded Lie subalgebra $\mathfrak{g}(\mathfrak{h}, \varphi)$ of $\mathfrak{g}$ defined by

$$\mathfrak{g}(\mathfrak{h}, \varphi) = \bigoplus_{i \in \mathbf{Z}} \mathfrak{g}(\mathfrak{h}, \varphi)_i$$

where

$$\mathfrak{g}(\mathfrak{h}, \varphi)_i = \{x \in \mathfrak{g}_{ia} \mid [u, x] = i\varphi(u)x \text{ for all } u \in H\}.$$

Definition 14. The **support** $\mathfrak{s}$ of the nilpotent element $e \in \mathfrak{g}_a$ is the commutator subalgebra of $\mathfrak{g}(\mathfrak{h}, \varphi)$ considered as a $\mathbf{Z}$-graded Lie algebra.

Clearly $e \in \mathfrak{s}_1$. We are ready to state the main theorem of [57].

Theorem 15 (Vinberg). *The supports of nilpotent elements of the space $\mathfrak{g}_i$ are exactly the complete regular locally flat semisimple $\mathbf{Z}$-graded subalgebras of the algebra $\mathfrak{g}$. The nilpotent element e can be recovered from the support subalgebra $\mathfrak{s}$ as the generic element in $\mathfrak{s}_1$.*

It follows from the theorem that the nilpotent element e is defined uniquely (up to conjugation by an element of G_0) by its support. This means it is enough to classify the regular semisimple $\mathbf{Z}$-graded subalgebras $\mathfrak{s}$ of $\mathfrak{g}$.

Let us choose a maximal torus $\mathfrak{t} \subset \mathfrak{g}_0$. The $\mathbf{Z}$-graded subalgebra $\mathfrak{s}$ is **standard** if it is normalized by $\mathfrak{t}$, i.e., if for all $i \in \mathbf{Z}$ we have

$$[\mathfrak{t}, \mathfrak{s}_i] \subset \mathfrak{s}_i .$$

Vinberg also proves that every $\mathbf{Z}$-graded subalgebra $\mathfrak{s}$ is conjugated to a standard subalgebra by an element of G_0. Moreover, he shows that if two standard $\mathbf{Z}$-graded subalgebras are conjugated by an element of G_0, then they are conjugated by an element of $N_0(\mathfrak{t})$. This gives combinatorial method for classifying regular semisimple $\mathbf{Z}$-graded subalgebras of $\mathfrak{g}$.

Let $\mathfrak{s}$ be a standard semisimple $\mathbf{Z}$-graded subalgebra of $\mathfrak{g}$. The grading on the subalgebra $\mathfrak{s}$ defines a degree map $\deg\colon R(\mathfrak{s}) \to \mathbf{Z}$. For a standard $\mathbf{Z}$-graded subalgebra $\mathfrak{s}$ we also get the map

$$f\colon R(\mathfrak{s}) \to R(\mathfrak{g}).$$

The map f has to be *additive*, i.e., it satisfies

$$f(\alpha+\beta) = f(\alpha) + f(\beta) \quad \forall \alpha, \beta \in R(\mathfrak{s}),$$
$$f(-\alpha) = -f(\alpha) \quad \forall \alpha \in R(\mathfrak{s}).$$

Moreover we have

Proposition 16. *The map f satisfies the following properties:*

(a)
$$\frac{(f(\alpha), f(\beta))}{(f(\alpha), f(\alpha))} = \frac{(\alpha, \beta)}{(\alpha, \alpha)} \quad \forall \alpha, \beta \in R(\mathfrak{s}).$$

(b) $f(\alpha) - f(\beta) \notin R(\mathfrak{g}) \quad \forall \alpha, \beta \in \Pi(\mathfrak{s})$.
(c) $\deg f(\alpha) = \deg \alpha, \quad \forall \alpha \in \Pi(\mathfrak{s})$.

Conversely, every map satisfying these conditions defines a standard regular $\mathbf{Z}$*-graded subalgebra* $\mathfrak{s}$ *of* $\mathfrak{g}$.

Remark 17. The subalgebra $\mathfrak{s}$ corresponding to the map f is complete if and only if there exists an element w in the Weyl group W of $\mathfrak{g}$ such that $wf(\Pi(\mathfrak{s})) \subset \Pi(\mathfrak{g})$ (see [57, p. 25]).

Theorem 15 means that in order to classify the nilpotent elements $e \in \mathfrak{g}_1$ we need to classify the possible maps f corresponding to its support, i.e., the corresponding complete regular $\mathbf{Z}$-graded subalgebra $\mathfrak{s}$. Since we are interested in the nilpotent elements $e \in \mathfrak{g}_1$, we need to classify the maps f for which $\deg(f(\alpha)) \in \{0, 1\}$ for every $\alpha \in \Pi(\mathfrak{s})$.

2.6 Example: $\bigwedge^3 \mathbf{C}^7$

Let $X_n = \mathrm{E}_7$, and $\alpha_k = \alpha_2$ in Bourbaki numbering.

The graded Lie algebra corresponding to the resulting grading is

$$\mathfrak{g}(E_7) = \mathfrak{g}_{-2} \oplus \mathfrak{g}_{-1} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2$$

with $G_0 = \mathbf{GL}_7(\mathbf{C})$, $\mathfrak{g}_0 = \mathfrak{gl}_7(\mathbf{C})$, $\mathfrak{g}_1 = \bigwedge^3 \mathbf{C}^7$, $\mathfrak{g}_2 = \bigwedge^6 \mathbf{C}^7$.

We choose a basis $\{e_1, \ldots, e_7\}$ of $\mathbf{C}^7$. The weight vectors in $\mathfrak{g}_1 = \bigwedge^3 \mathbf{C}^7$ are the vectors $e_i \wedge e_j \wedge e_k$ of weight $\varepsilon_i + \varepsilon_j + \varepsilon_k$ for $1 \le i < j < k \le 7$.

There is a natural bijection Σ of these weight vectors with the positive roots with label 1 on the node α_2. In order to describe it we identify the subroot system of $\mathfrak{g}_0$ with the root system of type A_6 with the simple roots $\beta_1, \ldots, \beta_6$ (corresponding to permutations $(1,2), (2,3), \ldots, (6,7)$, respectively) by identifying β_1 with α_1 and β_s with α_{s+1} for $2 \le s \le 6$. We denote this identification map by Λ. We set $\Sigma(e_5 \wedge e_6 \wedge e_7) = \alpha_2$ and extend it to a unique bijection satisfying $\Sigma(e_i \wedge e_j \wedge e_k + \beta_s) = \Sigma(e_i \wedge e_j \wedge e_k) + \Lambda(\beta_s)$ for every simple root β_s, for $1 \le s \le 6$.

The invariant scalar product on $\mathfrak{h}$ restricted to the roots from $\mathfrak{g}_1$ and transferred by Σ is

$$(e_i \wedge e_j \wedge e_k, e_p \wedge e_q \wedge e_r) = \#(\{i, j, k\} \cap \{p, q, r\}) - 1.$$

This can be checked by observing that both the invariant scalar product and the one defined by the formula above are invariant with respect to the Weyl group of $\mathfrak{g}_0$, i.e., the permutation group S_7, and checking the equality of one scalar product directly.

Next using the table from [55, p.248] we see that the general element of $\bigwedge^3 \mathbf{C}^7$ has the Bala–Carter label $\mathrm{A}_2 + 3\mathrm{A}_1$. Notice that the labeling of this orbit given on [16, p. 130] has weighted Dynkin diagram with labels 0 on all nodes except for α_k and label 2 on the node α_k (this also happens in the other cases except the case (E_6, α_3) where the weighted Dynkin diagram of the nilpotent orbit intersecting $\mathfrak{g}_1$ in the open orbit has labelings 1 at α_3 and α_5 and zero elsewhere). The support subalgebras of the smaller orbits in $\mathfrak{g}_1$ have Vinberg labels exactly matching the Bala–Carter labels that can be read off the Spaltenstein tables. In our analysis we checked also that the other types of support algebras given in [57] do not exist.

The combinatorial analysis is not difficult. We just give a few examples. Let us identify the support subalgebra of type $\mathrm{A}_2 + 3\mathrm{A}_1$. According to Proposition 16 we need to find all choices (up to the action of S_7) of five weight vectors in $\bigwedge^3 \mathbf{C}^7$ whose pairwise scalar products match those of the five simple roots of the root system $\mathrm{A}_2 + 3\mathrm{A}_1$. It is clear that for simple roots of the system A_2 we can choose $e_1 \wedge e_2 \wedge e_3$ and $e_4 \wedge e_5 \wedge e_6$. Then we need a choice of three subsets $[i, j, k]$ which have one element intersections with both $[1, 2, 3]$, $[4, 5, 6]$ and with each other. Up to the permutation group S_7 it is clear that there is only one choice $[1, 4, 7]$, $[2, 5, 7]$, $[3, 6, 7]$. One can check that the resulting orbit is open by calculating its dimension. This is an easy linear algebra exercise since the tangent space is obtained by hitting this element with all vectors in $\mathfrak{gl}_7(\mathbf{C})$.

Similarly, looking at the subsystems of type $3\mathrm{A}_1$ we need choices of three subsets $[i, j, k]$ such that every pair of subsets intersects in one element. This element can be common to all three subsets or not, which gives two possibilities $\{[1,2,3],[1,4,5],[1,6,7]\}$ and $\{[1,2,3],[1,4,5],[2,4,6]\}$ up to the S_7-action.

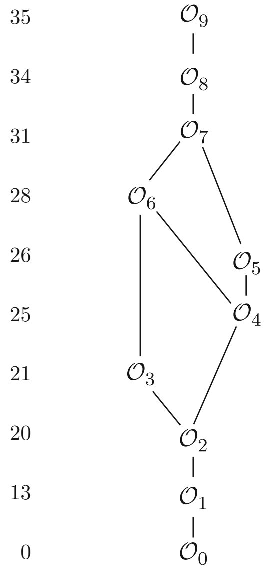
Finally, looking at the possibilities for $4\mathrm{A}_1$ we notice that the triple $[1,2,3]$, $[1,4,5],[1,6,7]$ can be only complemented in eight ways equivalent to $[2,4,6]$. The triple $[1,2,3],[1,4,5],[2,4,6]$ can be complemented in three ways equivalent to $[1,6,7]$. But these two choices of roots of $4\mathrm{A}_1$ are the same. We conclude that there is only one orbit of type $4\mathrm{A}_1$ in $\bigwedge^3 \mathbf{C}^7$.

A subtle point is checking that the obtained support algebra is complete. In the case of $\bigwedge^3 \mathbf{C}^7$ it also follows from the fact that all the representatives give different orbits. We omit this.

The result of the analysis is presented in the following table (writing $[i, j, k]$ for $e_i \wedge e_j \wedge e_k$).

Label	$\mathfrak{s}$	dim	Representative
0	0	0	0
1	A_1	13	[123]
2	$2\mathrm{A}_1$	20	[123] + [145]
3	$3\mathrm{A}_1$	21	[123] + [145] + [167]
4	$3\mathrm{A}_1$	25	[123] + [145] + [246]
5	A_2	26	[123] + [456]
6	$4\mathrm{A}_1$	28	[123] + [145] + [167] + [357]
7	$\mathrm{A}_2 + \mathrm{A}_1$	31	[123] + [456] + [147]
8	$\mathrm{A}_2 + 2\mathrm{A}_1$	34	[123] + [456] + [147] + [257]
9	$\mathrm{A}_2 + 3\mathrm{A}_1$	35	[123] + [456] + [147] + [257] + [367]

The containment diagram is



and it happens to be the same as the corresponding containment of the nilpotent orbits in $\mathfrak{g}_1$.

We can interpret the orbits in terms of skew-symmetric tensors as follows. Let $A = \mathrm{Sym}(\bigwedge^3 \mathbf{C}^{7*})$ be the coordinate ring of our representation. The orbit $\mathcal{O}_1$ is the orbit of the highest-weight vector, and its closure is the set of decomposable tensors, i.e., the tensors $t = v_1 \wedge v_2 \wedge v_3$ $(v_1, v_2, v_3 \in \mathbf{C}^7)$. There is an $\mathbf{SL}_7(\mathbf{C})$-invariant Δ of degree 7, the hyperdiscriminant. This is the projective dual variety of the decomposable tensors in $\mathbf{C}^{7*}$. In fact, we can canonically identify the orbits in $\mathbf{C}^7$ with those in its dual, so it makes sense to consider the projective dual $\overline{\mathcal{O}}^\vee$ of an orbit closure $\overline{\mathcal{O}}$ as a subset of $\mathbf{C}^7$ itself. By the **rank** of a tensor $t \in \bigwedge^3 \mathbf{C}^7$ we mean the subspace rank, i.e., the minimal number s such that there exists a subspace V of dimension s in $\mathbf{C}^7$ such that $t \in \bigwedge^3 V \subset \bigwedge^3 \mathbf{C}^7$. An orbit is **degenerate** if it consists of tensors of rank ≤ 6. Such orbits correspond to $\mathbf{GL}_6(\mathbf{C})$-orbits in $\bigwedge^3 \mathbf{C}^6$. By a 1-decomposable tensor we mean a tensor $t = v \wedge q$ with $v \in \mathbf{C}^7$, $q \in \bigwedge^2 \mathbf{C}^7$. Finally in the projective picture we interpret orbit closures as secant and tangential varieties of the orbit $\mathcal{O}_1$. We denote by $\sigma_k(\mathcal{O}_1)$ the kth secant variety of $\mathcal{O}_1$ and by $\tau(\mathcal{O}_1)$ the tangential variety of $\mathcal{O}_1$.

With this terminology, we have the following table describing the orbits.

Number	Projective picture	Tensor picture
0	0	0
1	Cone($\mathbf{Gr}(3,7)$)	Decomposable tensors
2		Tensors of rank ≤ 5
3	$\sigma_2(\overline{\mathcal{O}}_1)^\vee$	1-Decomposable tensors
4	$\tau(\overline{\mathcal{O}}_1)$	Degenerate and zero hyperdiscriminant for $\bigwedge^3 \mathbf{C}^6$
5	$\sigma_2(\overline{\mathcal{O}}_1)$	Tensors of rank ≤ 6
6	$J(\overline{\mathcal{O}}_1, \tau(\overline{\mathcal{O}}_1))$	Polarizations of hyperdiscriminant for $\bigwedge^3 \mathbf{C}^6$
7	$\sigma_3(\overline{\mathcal{O}}_1)$	Singular locus of $\overline{\mathcal{O}}_1$
8	$\overline{\mathcal{O}}_1^\vee$	Hyperdiscriminant is zero
9		Generic

Remark 18. The geometric description of these orbits is also considered in [31, Sect. 5] (he also considers $\bigwedge^3 \mathbf{C}^6$ and $\bigwedge^3 \mathbf{C}^8$). The classification of orbits of $\bigwedge^3 K^7$ is studied for many kinds of fields K (including algebraically closed fields of positive characteristic and finite fields) in [15]. In particular, the classification of orbits is independent of characteristic if the field is algebraically closed.

We will describe in detail the non-degenerate orbit closures in $\bigwedge^3 \mathbf{C}^7$. These are the orbits $\overline{\mathcal{O}_9}$, $\overline{\mathcal{O}_8}$, $\overline{\mathcal{O}_7}$, $\overline{\mathcal{O}_6}$, and $\overline{\mathcal{O}_3}$. The first of these is generic so there is not much to say. We also describe the generic degenerate orbit of tensors of rank ≤ 6.

We use the usual notation. Let $A = \mathrm{Sym}(\bigwedge^3 \mathbf{C}^{7*})$ and λ is notation for $\mathbf{S}_\lambda(\mathbf{C}^{7*})$. Also, let $x_1, \dots, x_7$ be a basis of $\mathbf{C}^{7*}$ dual to the basis $e_1, \dots, e_7$ of $\mathbf{C}^7$. We will also describe vector bundle desingularizations for these orbit closures. The bundles η and ξ correspond to $\mathcal{S}^*$ and $\mathcal{T}^*$ in the notation of Sect. 1.3.

- The hyperdiscriminant orbit $\mathcal{O}_8$.
  This is the hypersurface given by the tensors with vanishing hyperdiscriminant Δ. The orbit closure $\overline{\mathcal{O}}_8$ is characterized (set-theoretically) by the condition that the determinant of the multiplication map

$$\bigwedge^5 \mathbf{C}^{7*} \otimes A(-1) \to \bigwedge^2 \mathbf{C}^{7*} \otimes A$$

  given by multiplication is zero. The determinant of this matrix is equal to Δ^3.
- The codimension 4 orbit $\mathcal{O}_7$.
  This orbit closure is the singular locus of the hyperdiscriminant orbit $\overline{\mathcal{O}_8}$.
  We can find a desingularization by a vector bundle over $G/P = \mathbf{Fl}(2,6;\mathbf{C}^7)$. The bundle $\xi \subset \bigwedge^3 \underline{\mathbf{C}^{7*}}$ is induced from the largest P-submodule of $\bigwedge^3 \mathbf{C}^{7*}$ which does not contain $x_1 \wedge x_2 \wedge x_7$ and $x_2 \wedge x_5 \wedge x_7$. The bundle η has rank 17, so the dimension of the desingularization is $17 + 14 = 31$ as needed. One gets a very nice complex describing the resolution of $\mathbf{C}[\overline{\mathcal{O}}_7]$. The terms of the complex $\mathbf{F}(7)_\bullet$ are as follows

$$0 \to (6, 5^6) \to (5^2, 4^5) \to (4, 3^5, 2) \\ \to (3^4, 2^3) \to (0^7).$$

The orbit closure is normal and has rational singularities, and the complex $\mathbf{F}(7)_\bullet$ is pure.

- The codimension 7 orbit $\mathcal{O}_6$.
  We can find a desingularization by a vector bundle over $G/P = \mathbf{Fl}(1, 4; \mathbf{C}^7)$. The bundle $\xi \subset \bigwedge^3 \underline{\mathbf{C}^{7}}^*$ is induced from the P-submodule of $\bigwedge^3 \mathbf{C}^{7*}$ which does not contain $x_1 \wedge x_4 \wedge x_7$ and $x_2 \wedge x_3 \wedge x_4$. The bundle η has rank 13, so the dimension of the desingularization is $13 + 15 = 28$ as needed. The terms in the resulting complex $\mathbf{F}(6)_\bullet$ are

$$0 \to (7^6, 6) \to (7, 6^5, 5) \to (6^2, 5^4, 4) \to (5^3, 4^3, 3) \to (4^4, 3^2, 2) \to (3^5, 2, 1) \\ \to (2^6, 0) \to (0^7).$$

  The orbit closure is normal, with rational singularities, and the complex $\mathbf{F}(6)_\bullet$ is pure.
- The generic degenerate orbit closure $\overline{\mathcal{O}_5}$ of tensors of rank ≤ 6 (codimension 9). This orbit closure has a desingularization $Z(5)$ that lives on the Grassmannian $\mathbf{Gr}(1, \mathbf{C}^{7*})$. Denoting the tautological bundles $\mathcal{R}$, $\mathcal{Q}$ (rank $\mathcal{R} = 1$, rank $\mathcal{Q} = 6$), we have $\xi = \mathcal{R} \otimes \bigwedge^2 \mathcal{Q}$. It is normal and has rational singularities. Calculating the resolution is straightforward, as ξ is irreducible.
- The orbit $\mathcal{O}_3$ of 1-decomposable tensors (codimension 14).
  This orbit closure is the set of tensors $t \in \bigwedge^3 \mathbf{C}^7$ that can be expressed as $t = \ell \wedge \bar{t}$ where $\ell \in \mathbf{C}^7$, $\bar{t} \in \bigwedge^2 \mathbf{C}^7$. The desingularization $Z(3)$ lives on the Grassmannian $\mathbf{Gr}(6, \mathbf{C}^{7*})$. Denoting the tautological bundles as $\mathcal{R}$, $\mathcal{Q}$ (rank $\mathcal{R} = 6$, rank $\mathcal{Q} = 1$), we have $\xi = \bigwedge^3 \mathcal{R}$. It is normal and has rational singularities. Calculating the resolution is straightforward, as ξ is irreducible. The defining ideal is generated by the representation $(2^3, 1^3, 0)$ in degree 3.

Remark 19. Since we will need it later on, we ask, for a fixed $v \in \bigwedge^3 \mathbf{C}^7$, how many lines $\ell \subset \mathbf{C}^7$ there are such that the image of v in $\bigwedge^3(\mathbf{C}^7/\ell)$ is a pure tensor. This only depends on the orbit, so we can study specific representatives.

No such line exists for vectors in the generic orbit or codimension 1 orbit. A representative for the codimension 4 orbit is

$$e_1 \wedge e_2 \wedge e_3 + e_4 \wedge e_5 \wedge e_6 + e_1 \wedge e_4 \wedge e_7.$$

This is a pure tensor in $\bigwedge^3(\mathbf{C}^7/\ell)$ exactly for $\ell = \langle e_1 \rangle$ and $\ell = \langle e_4 \rangle$. A representative for the codimension 7 orbit is

$$e_1 \wedge e_2 \wedge e_3 + e_1 \wedge e_4 \wedge e_5 + e_1 \wedge e_6 \wedge e_7 + e_3 \wedge e_5 \wedge e_7.$$

This is a pure tensor in $\bigwedge^3(\mathbf{C}^7/\ell)$ exactly for $\ell = \langle e_1 \rangle$.

3 Some Geometry

3.1 Abelian Varieties

The following result is most likely well known, but we could not find it in the literature. The main points of the proof were communicated to us by Damiano Testa.

Theorem 1. *Let X be a g-dimensional geometrically connected projective non-singular variety over a field K of characteristic* 0*. If $\omega_X \cong \mathcal{O}_X$ and* $\dim \mathrm{H}^1(X; \mathcal{O}_X) = g$*, then X is a torsor over an Abelian variety (namely, its Albanese variety).*

There are two reasons that we emphasize the fact that X is only a torsor: in our applications of this theorem in the later sections, there may be no natural choice of base point (which is relevant when working over non-algebraically closed fields), and in later work we will be interested in working over families (in which case the existence of a section may be more subtle).

Proof. First, extend scalars to the algebraic closure $\overline{K}$ of K. By [35, Corollary 2], $X_{\overline{K}}$ is birationally equivalent to its Albanese variety $X_{0,\overline{K}}$, let $f \colon X_{\overline{K}} \to X_{0,\overline{K}}$ be a birational morphism. In fact, f can be (uniquely) extended to a morphism on all of $X_{\overline{K}}$ [7, Theorem 4.9.4]. One has an induced map $\mathrm{d}f$ on cotangent bundles. The determinant of $\mathrm{d}f$ is a map between the canonical bundles, which are trivial by assumption. The endomorphisms of the trivial bundle must be scalars since we assumed that $X_{\overline{K}}$ is projective, so $\det \mathrm{d}f$ is a scalar. This scalar is nonzero since f is generically an étale morphism. So in fact f is étale (and hence open). Also f is proper (and hence closed), so f is an étale covering, which implies that $X_{\overline{K}}$ is an Abelian variety. Furthermore, f must be an isomorphism since it is birational.

Hence, choosing any point $P \in X(\overline{K})$, we have a $\overline{K}$-isomorphism $X_{0,\overline{K}} \to X_{\overline{K}}$ via $x \mapsto P + x$. This map descends to a K-rational map $Y \to X$ where Y is a K-rational X_0-torsor given by the cocycle $\gamma \mapsto {}^{\gamma}P - P$, which gives the claim. □

Remark 2. If we drop the assumption that K be of characteristic 0, then this theorem already fails for $g = 2$. In particular, it is valid if the characteristic is different from 2 or 3, but in these small characteristics, there are new exotic examples, known as quasi-hyperelliptic surfaces which come from the Bombieri–Mumford classification of surfaces (the quasi-hyperelliptic surfaces have the property that their Picard varieties are non-reduced); see [8, p. 25, Table].

Given a smooth curve C of genus g, we let $\mathrm{Jac}(C)$ denote its Jacobian, which is an Abelian variety of dimension g (see [7, Chap. 11] for an analytic construction of $\mathrm{Jac}(C)$).

3.2 Moduli Space of Vector Bundles

The material in this section is provided for convenience and informative purposes, since later we will see some examples of the moduli spaces discussed in this section (see Example 6, Remark 6, and Sect. 8). We refer the reader to [4] for a more in-depth survey of the following.

Let C be a smooth curve of genus $g \geq 2$. There is a moduli space $SU_C(n,d)$ which parametrizes rank n semistable vector bundles of degree d on C. It has dimension $n^2(g-1)+1$. Let $SU_C(n,L)$ denote the moduli space of rank n semistable vector bundles on C with determinant equal to L. We write $SU_C(n)$ for $SU_C(n,\mathcal{O}_C)$. Then $SU_C(n)$ has dimension $(g-1)(n^2-1)$ and is Gorenstein and has rational singularities. The Picard group of $SU_C(n)$ is infinite cyclic. Let $\mathcal{L}$ be its ample generator, which we call the theta divisor. The canonical bundle is $\mathcal{L}^{-2n}$.

Now focus on $n=2$. If $g=2$, then $SU_C(2) \cong \mathbf{P}^3$, and for $g>2$, the singular locus of $SU_C(2)$ is bundles of the form $L \oplus L^{-1}$ and is naturally identified with the Kummer quotient of the Jacobian of degree $g-1$ line bundles on C. Also, $\mathrm{h}^0(SU_C(n);\mathcal{L}) = 2^g$ and the map given by $\mathcal{L}$ is an embedding if C is not hyperelliptic. Furthermore, the restriction of $\mathcal{L}$ is a $(2,\dots,2)$-polarization. See [9] and [25] for more details.

Finally, we state the Verlinde formula, which gives the dimension of the space of sections of powers of $\mathcal{L}$. For $n=2$, we have

$$\mathrm{h}^0(SU_C(2);\mathcal{L}^k) = \left(\frac{k+2}{2}\right)^{g-1} \sum_{j=1}^{k+1} (\sin\frac{\pi j}{k+2})^{-2g+2};$$

see [2, Sect. 5] for the case of general n.

3.3 Degeneracy Loci

The main idea of this chapter is in the following construction.

Construction 3. Start with a Vinberg representation (G,U) of affine type (see Sect. 2.4). Choose a parabolic subgroup P of G. We can realize U as the sections of a homogeneous bundle $\mathcal{U}$ over the homogeneous space G/P using Sect. 2.3. In each case that we consider, the fibers of $\mathcal{U}$ can naturally be interpreted as another Vinberg representation (G',U') of finite type (more specifically, G' will be the Levi subgroup of the stabilizer of the point where the fiber lives). We use information about the orbit closures in (G',U') and patch them together to get subvarieties $\mathcal{Y}$ of the total space of $\mathcal{U}$.

Given a section $v \in \mathrm{H}^0(X;\mathcal{U})$, we consider the subvarieties $v(G/P) \cap \mathcal{Y}$, and in particular, when the grade of the ideal sheaf does not change. In all of the orbit closures in U' of relevance, we give free resolutions for their coordinate rings and some related modules. Then this gives a locally free resolution of $v(X) \cap \mathcal{Y} \subset v(G/P) \cong G/P$ via Theorem 7, and this will allow us to read off properties of this variety. In particular, we can try to use this resolution to calculate the canonical sheaf of $v(G/P) \cap \mathcal{Y}$ and the cohomology of its structure sheaf.

In all cases, we will find a choice of P so that one of the degeneracy loci (or a variety closely related to it) is a torsor over an Abelian variety. We will also explore what happens when we vary the choice of P. In some cases, we are able to establish a direct link between these other degeneracy loci and the torsor via some classical geometric constructions (such as projective duality).

Lemma 4. *Let Y° be a G'-orbit in U' and let $\mathcal{Y}^\circ$ be the union of these G'-orbits over all fibers. There is an open subset $U_Y^{\text{gen}} \subset U$ such that for each $v \in U_Y^{\text{gen}}$, we have* $\operatorname{codim}(v(G/P) \cap \mathcal{Y}^\circ, v(G/P)) = \operatorname{codim}(\mathcal{Y}^\circ, \mathcal{U})$. *In particular, this intersection either has expected codimension or is empty. Furthermore, if the base field has characteristic* 0*, there is an open subset U_Y^{sm} such that $v(G/P) \cap \mathcal{Y}^\circ$ is smooth.*

Proof. Define $Z = \{(v, x) \in U \times G/P \mid v(x) \in Y^\circ\}$. Let $\pi_1 : Z \to U$ and $\pi_2 : Z \to G/P$ be the projection maps. We claim that π_2 is a fiber bundle with smooth fibers. Fix $x \in G/P$ and let $\mathcal{U}(x)$ be the fiber of $\mathcal{U}$ at x. Then the restriction map $\rho_x : \mathcal{U} \to \mathcal{U}(x)$ is G'-linear and it is surjective since $\mathcal{U}(x)$ is irreducible as a G'-module. Hence the map $\rho_x : \rho_x^{-1}(Y^\circ) \to Y^\circ$ is an affine bundle. But Y° is smooth, and $\rho_x^{-1}(Y^\circ) = \pi_2^{-1}(x)$, so the claim follows. In particular, Z is smooth and

$$\dim Z = \dim G/P + \dim U - \operatorname{codim}(Y^\circ, U').$$

Now consider the map π_1. If it is dominant, then the fibers over a nonempty open subset U_Y^{gen} have dimension $\dim G/P - \operatorname{codim}(Y^\circ, U')$ and hence have expected codimension. Otherwise, the fibers are empty over a nonempty open subset U_Y^{gen}. Given $v \in U$, we have an identification $\pi_1^{-1}(v) = v(G/P) \cap \text{U}$, so this proves the first claim.

The last statement follows from generic smoothness applied to π_1. □

We will define $U^{\text{gen}} \subset U$ to be the intersection of U_Y^{gen} over all orbits Y in U', and similarly we define U^{sm}.

Remark 5. The idea of studying degeneracy loci using perfect complexes rather than cohomology class formulas has been considered by the second author in [53] for the class of Schubert determinantal loci.

Example 6. Let V be a vector space of dimension $2n$ and let $q \in S^2V$ be a nondegenerate quadratic form. Denote the quotient $S^2V/\langle q\rangle$ by S_0^2V.

Consider the action of $\mathbf{SO}(V) \times \mathbf{C}^*$ on S_0^2V. Given a nondegenerate quadric $q' \in S_0^2\mathbf{C}^{2n}$, we can form the pencil $xq + yq'$. The determinant of the associated symmetric matrix gives us $2n$ points in $\mathbf{P}^1$ and hence a hyperelliptic curve C of genus $n-1$, and this process is reversible. This situation was considered by Weil.

Now consider the intersection of the quadrics defined by q and q' in $\mathbf{P}(V^*)$. Then the variety of $\mathbf{P}^{n-2}$'s in $q \cap q'$ is isomorphic to the Jacobian of C (after fixing a base point), and the variety of $\mathbf{P}^{n-3}$'s in $q \cap q'$ is isomorphic to the moduli space $SU_C(2, L)$ (see Sect. 3.2) where L is any line bundle of odd degree. See [17, Theorems 1, 2] for details.

These constructions can be interpreted as degeneracy loci as follows. Consider the orthogonal Grassmannian $\mathbf{OGr}(n-1, V^*)$, which is the subvariety of $\mathbf{Gr}(n-1, V^*)$ whose points are totally isotropic subspaces for q. The trivial bundle $\mathbf{Gr}(n-1, V^*) \times V^*$ has a tautological rank $n-1$ subbundle $\mathcal{R} = \{ (x, W) \mid x \in W \}$. Then we have $S_0^2 V = \mathrm{H}^0(\mathbf{OGr}(n-1, V^*); S^2\mathcal{R}^*)$, and q' gives a generic section whose zero locus is the variety of $\mathbf{P}^{n-2}$'s in $q \cap q'$. Similar comments apply to the variety of $\mathbf{P}^{n-3}$'s using $\mathbf{OGr}(n-2, V^*)$. Modular interpretations for the degeneracy loci for the other Grassmannians $\mathbf{OGr}(k, V^*)$ are given by Ramanan [52, Sect. 6, Theorem 3].

4 $\mathbf{C}^5 \otimes \bigwedge^2 \mathbf{C}^5$

In the rest of this chapter, we will work over the complex numbers $\mathbf{C}$. In fact, many results will hold over more general fields, but as we have not done a systematic investigation of the correct hypotheses, we will not make any attempt to be more general.

The analysis of the representation $\mathbf{C}^5 \otimes \bigwedge^2 \mathbf{C}^5$ is in fact easy to handle by more direct means, but we want to illustrate our approach. We also mention that Fisher has examined this case as well; see [22, 23].

Let A and B be vector spaces of dimension 5. The relevant data:

- $U = A \otimes \bigwedge^2 B$.
- $G = (\mathbf{GL}(A) \times \mathbf{GL}(B))/\{ (x, x^{-1}) \mid x \in \mathbf{C}^* \}$.
- $G/P = \mathbf{P}(A^*) = \mathbf{Gr}(1, A^*)$.
- $\mathcal{U} = \mathcal{R}^* \otimes \bigwedge^2 \underline{B} \cong \mathcal{O}(-1) \otimes \bigwedge^2 \underline{B}$.
- $U' = \bigwedge^2 \mathbf{C}^5$.
- $G' = \mathbf{GL}_5(\mathbf{C})$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 20, 30, and the graded Weyl group is Shephard–Todd group 16 [56, Sect. 9].

4.1 *Modules Over $\mathcal{O}_{U'}$*

We are only interested in the ideal of 4×4 Pfaffians of U'. This situation was explained in Example 10.

4.2 Geometric Data from a Section

The constructions in this section work over an arbitrary field.

We get the following locally free resolution over $\mathcal{O}_{\mathcal{U}} = \mathrm{Sym}(\bigwedge^2 \underline{B}^* \otimes \mathcal{O}_{\mathcal{U}}(1))$:

$$0 \to (\det \underline{B}^*)^{\otimes 2} \otimes \mathcal{O}_{\mathcal{U}}(-5) \to (\det \underline{B}^*) \otimes \underline{B}^* \otimes \mathcal{O}_{\mathcal{U}}(-3) \to \bigwedge^4 \underline{B}^* \otimes \mathcal{O}_{\mathcal{U}}(-2)$$
$$\to \mathcal{O}_{\mathcal{U}} \to \mathcal{O}_{\mathcal{C}} \to 0$$

where $\mathcal{C}$ has codimension 3 in the total space of $\mathcal{U}$. Its singular locus is the zero section of $\mathcal{U}$, and has codimension 10 in $\mathcal{U}$.

For $v \in U^{\text{gen}}$, $C_v = \mathcal{C} \cap v(\mathbf{P}(A^*))$ will have codimension 3 in $v(\mathbf{P}(A^*)) \cong \mathbf{P}(A^*)$. By generic perfection, we get a locally free resolution for $\mathcal{O}_C$:

$$0 \to (\det \underline{B}^*)^{\otimes 2} \otimes \mathcal{O}_{\mathbf{P}(A^*)}(-5) \to (\det \underline{B}^*) \otimes \underline{B}^* \otimes \mathcal{O}_{\mathbf{P}(A^*)}(-3)$$
$$\to \bigwedge^4 \underline{B}^* \otimes \mathcal{O}_{\mathbf{P}(A^*)}(-2) \to \mathcal{O}_{\mathbf{P}(A^*)} \to \mathcal{O}_C \to 0.$$

This gives enough information to see that $\omega_C = \mathcal{O}_C$, $\dim \mathrm{H}^0(C; \mathcal{O}_C) = 1$, and $\dim \mathrm{H}^1(C; \mathcal{O}_C) = 1$. In particular, C is a curve of genus 1. We can also deduce that C is projectively normal and embedded by a complete linear series.

Conversely, given a smooth curve C of genus 1 embedded in $\mathbf{P}(A^*)$ by a complete linear series, its homogeneous ideal I is generated by 5 quadrics and is a codimension 3 Gorenstein ideal. The Buchsbaum–Eisenbud classification of such ideals says that we can recover a section $v \in U$ which gives rise to C.

Theorem 1. *We have a bijection between G-orbits in U^{sm} and the set of pairs $(C, \mathcal{L})$ where C is a genus 1 curve and $\mathcal{L}$ is a degree 5 line bundle on C.*

4.3 Examples of Singular Quintic Curves

In this section, we give a few examples of degenerations of the smooth elliptic quintic C. Describing degenerations of the Abelian varieties in the later examples will require more effort and will appear in future work.

We pick homogeneous coordinates $z_1, \dots, z_5$ on $\mathbf{P}^4$.

Example 2. Here is one example of a section that gives a rational nodal curve:

$$\begin{pmatrix} 0 & z_5 & z_1 & z_2 & z_3 \\ -z_5 & 0 & z_2 & z_3 & z_4 \\ -z_1 & -z_2 & 0 & z_4 & z_5 \\ -z_2 & -z_3 & -z_4 & 0 & 0 \\ -z_3 & -z_4 & -z_5 & 0 & 0 \end{pmatrix}$$

It is given by the parametrization $[a:b] \mapsto [a^5+b^5 : ab^4 : a^2b^3 : a^3b^2 : a^4b]$, and its node is the point $[1:0:0:0:0]$. The stabilizer subgroup in G of this curve is the dihedral group of size 10 generated by the transformations $[a:b] \mapsto [b:a]$ and $[a:b] \mapsto [a:\zeta b]$ where ζ is a primitive 5th root of unity [23, proof of Lemma 2.3]. This is not an unstable orbit.

Furthermore, its secant variety is an irreducible quintic hypersurface.

Example 3. We can get a triangle consisting of two smooth rational quadrics and a line. Here is one example:

$$\begin{pmatrix} 0 & 0 & z_4 & z_3 & z_2 \\ 0 & 0 & 0 & z_2 & z_1 \\ -z_4 & 0 & 0 & 0 & -z_5 \\ -z_3 & -z_2 & 0 & 0 & -z_4 \\ -z_2 & -z_1 & z_5 & z_4 & 0 \end{pmatrix}$$

The quadrics are $[a:b] \mapsto [a^2 : ab : b^2 : 0 : 0]$ and $[a:b] \mapsto [0:0:a^2:ab:b^2]$, and the line is $[a:b] \mapsto [a:0:0:0:b]$. The secant variety is the union of $z_4 = 0$, $z_2 = 0$, and the cubic $z_1z_4^2 + z_2^2z_5 - z_1z_3z_5 = 0$.

Example 4. We can also get a union of 5 $\mathbf{P}^1$'s which are labeled with $i \in \mathbf{Z}/5$ such that $\mathbf{P}^1_i$ intersects $\mathbf{P}^1_j$ if and only if $j = i \pm 1$, and the intersection points are distinct. This is a **Néron pentagon**. All Néron pentagons form a single orbit since they are determined by their points of intersection. Here is one example of a section that gives a Néron pentagon:

$$\begin{pmatrix} 0 & z_1 & z_2 & 0 & 0 \\ -z_1 & 0 & 0 & z_3 & 0 \\ -z_2 & 0 & 0 & 0 & z_4 \\ 0 & -z_3 & 0 & 0 & z_5 \\ 0 & 0 & -z_4 & -z_5 & 0 \end{pmatrix}.$$

This is the set of points in $\mathbf{P}(A^*)$ with at least 3 coordinates equal to 0. Its stabilizer subgroup contains the normalizer of the diagonal subgroup in $\mathbf{SL}(A)$. So the orbit of Néron pentagons has codimension of at least 5.

Its secant variety is the hypersurface $z_1z_2z_3z_4z_5 = 0$.

Example 5. Here is a non-reduced example of a union of a rational normal cubic and a non-reduced line which intersect with multiplicity 2:

$$\begin{pmatrix} 0 & 0 & z_5 & z_3 & z_2 \\ 0 & 0 & 0 & z_2 & z_1 \\ -z_5 & 0 & 0 & z_4 & z_3 \\ -z_3 & -z_2 & -z_4 & 0 & 0 \\ -z_2 & -z_1 & -z_3 & 0 & 0 \end{pmatrix}.$$

The cubic is $[a : b] \mapsto [a^3 : a^2b : ab^2 : b^3 : 0]$ and the line is given by the ideal (z_1, z_2, z_3^2).

Its secant variety is the union of the hyperplane $z_5 = 0$ and the non-reduced quartic $(z_2^2 - z_1z_3)^2 = 0$.

Example 6. Here is a rational cuspidal cubic:

$$\begin{pmatrix} 0 & z_1 & z_4 & 0 & z_5 \\ -z_1 & 0 & 0 & z_5 & z_2 \\ -z_4 & 0 & 0 & z_2 & z_3 \\ 0 & -z_5 & -z_2 & 0 & z_4 \\ -z_5 & -z_2 & -z_3 & -z_4 & 0 \end{pmatrix}.$$

Its cusp point is $[0 : 0 : 1 : 0 : 0]$ and it has the parametrization

$$[s : t] \mapsto [\sqrt{-1}t^5 : s^3t^2 : s^5 : \sqrt{-1}s^2t^3 : st^4].$$

This vector lies in the unstable locus of the representation. The group of automorphisms of this curve that extend to automorphisms of $\mathbf{P}^4$ is generated by scaling t, so the orbit of this curve in $\mathbf{C}^5 \otimes \bigwedge^2 \mathbf{C}^5$ has codimension 2. In particular, it gives a generic point of the unstable locus.

Its secant variety is an irreducible quintic hypersurface.

4.4 Secant and Tangential Varieties

Here is a different approach:

- $G/P = \mathbf{Gr}(2, A^*)$.
- $\mathcal{U} = \mathcal{R}^* \otimes \bigwedge^2 \underline{B}$.
- $U' = \mathbf{C}^2 \otimes \bigwedge^2 \mathbf{C}^5$.
- $G' = (\mathbf{GL}_2(\mathbf{C}) \times \mathbf{GL}_5(\mathbf{C}))/\{ (x, x^{-1}) \mid x \in \mathbf{C}^* \}$.

The relevant G'-orbit closures in U' are of codimensions 2, 4, and 5. The singular locus and the non-normal locus of the codimension 2 orbit closure are both the codimension 4 orbit closure S'. Also, S' is smooth along the codimension 5 orbit closure T'. Furthermore, S' and T' can be identified as the secant and tangential varieties of the affine cone over the Segre variety $\mathbf{P}^1 \times \mathbf{Gr}(2, 5)$. Let $\mathcal{S}$ and $\mathcal{T}$ be the global versions of these varieties, and given a section $v \in U$, let S and T be the corresponding degeneracy loci.

Proposition 7. *S is the locus of planes W such that* $\deg(\mathbf{P}(W) \cap C) \geq 2$*, and T is the locus of planes W such that $\mathbf{P}(W)$ is tangent to some point of C.*

Proof. Pick $W \in \mathbf{Gr}(2, A^*)$. Then $W \in S \setminus T$ if and only if $v(S) \in (\mathcal{R}^* \otimes \bigwedge^2 \underline{B})(W)$ is a sum of the form $a_1 \otimes (b_1 \wedge c_1) + a_2 \otimes (b_2 \wedge c_2)$ where a_1 and a_2 are linearly independent. But we can also identify this fiber with $\mathrm{H}^0(\mathbf{P}(W); \mathcal{O}_{\mathbf{P}(A^*)}(1) \otimes \bigwedge^2 \underline{B})$ when we identify $\mathbf{P}(W)$ with a line in $\mathbf{P}(A^*)$. This means that $\mathbf{P}(W) \cap C$ consists of two points corresponding to the vanishing of a_1 and a_2. Since T is in the closure of $S \setminus T$, we finish via a limiting argument. □

In particular, T is a smooth curve. To calculate its genus, we can use a free resolution for T' (here we use $(\lambda; \mu)(-i)$ as shorthand for $\mathbf{S}_\lambda(\mathbf{C}^2) \otimes \mathbf{S}_\mu(\mathbf{C}^5) \otimes \mathrm{Sym}(U'^*)(-i)$):

$$\mathbf{F}_1 = (2,1;2,1,1,1,1)(-3) + (2,2;2,2,2,2,0)(-4)$$
$$\mathbf{F}_2 = (2,2;2,2,2,1,1)(-4) + (4,1;2,2,2,2,2)(-5) + (3,2;3,2,2,2,1)(-5)$$
$$\mathbf{F}_3 = (4,2;3,3,2,2,2)(-6) + (3,3;4,2,2,2,2)(-6) + (4,3;3,3,3,3,2)(-7)$$
$$\mathbf{F}_4 = (4,2;4,3,3,3,3)(-8) + (4,4;4,3,3,3,3)(-8)$$
$$\mathbf{F}_5 = (6,4;4,4,4,4,4)(-10).$$

So a locally free resolution for T is

$$\begin{aligned}
\mathbf{F}_1 &= \mathcal{R}(-1)^5 + \mathcal{O}(-2)^{15} \\
\mathbf{F}_2 &= \mathcal{O}(-2)^5 + S^3(\mathcal{R})(-1) + \mathcal{R}(-2)^{24} \\
\mathbf{F}_3 &= S^2(\mathcal{R})(-2)^{10} + \mathcal{O}(-3)^{15} + \mathcal{R}(-3)^5 \\
\mathbf{F}_4 &= S^2(\mathcal{R})(-2)^5 + \mathcal{O}(-4)^5 \\
\mathbf{F}_5 &= S^2(\mathcal{R})(-4),
\end{aligned}$$

and we see that T has genus 1.

Proposition 8. $C \cong T$.

Proof. Since C is smooth, we have a well-defined morphism $\psi\colon C \to T$ obtained by sending $x \in C$ to the tangent line at x. This is a finite morphism, and by Riemann–Hurwitz, the ramification divisor is 0. Hence, ψ is an isomorphism. □

4.5 Chow Forms

Yet another approach is as follows:

- $G/P = \mathbf{Gr}(3, A^*)$.
- $\mathcal{U} = \mathcal{R}^* \otimes \bigwedge^2 \underline{B}$.
- $U' = \mathbf{C}^3 \otimes \bigwedge^2 \mathbf{C}^5$.

- $G' = (\mathbf{GL}_3(\mathbf{C}) \times \mathbf{GL}_5(\mathbf{C}))/\{\,(x, x^{-1}) \mid x \in \mathbf{C}^*\,\}$.

The space U' contains a G'-invariant degree 15 hypersurface. The corresponding degeneracy locus X' is a section of $\mathcal{O}(5)$.

Proposition 9. *X' is the Chow form of C.*

Proof. To obtain the Chow form of C, let $\mathbf{Fl}(1, 3, A^*)$ be a partial flag variety with projections π_1, π_2 to $\mathbf{P}(A^*)$ and $\mathbf{Gr}(3, A^*)$. Then $\mathrm{R}\,\pi_{2*}\mathrm{L}\,\pi_1^*\mathcal{O}_C$ is quasi-isomorphic to a complex whose determinant is the Chow form. A locally free resolution of $\mathcal{O}_C$ in $\mathbf{P}(A^*)$ is

$$0 \to (\det B^*)^2(-5) \to \mathbf{S}_{2,1^4} B^*(-3) \to \bigwedge^4 B^*(-2) \to \mathcal{O}_{\mathbf{P}(A^*)}.$$

Applying $\mathrm{R}\,\pi_{2*}\pi_1^*$ to this complex gives a 2-term complex over $\mathcal{O}_{\mathbf{Gr}(3,A^*)}$:

$$0 \to (\det B^*)^2 \otimes S^2\mathcal{R}(-1) \to \mathcal{O}_{\mathbf{Gr}(3,A^*)} \oplus \mathbf{S}_{2,1^4} B^*(-1). \tag{4.10}$$

In this case, the determinant is just the actual determinant of a 6×6 matrix. This gives a section of $\mathcal{O}(5)$ which is the Chow form of C.

We claim that this map is a sheafy version of the following map. For $U' = \mathbf{C}^3 \otimes \bigwedge^2 \mathbf{C}^5$, we have $(\det \mathbf{C}^5)^{-2} \otimes \mathbf{S}_{3,1,1}(\mathbf{C}^3)^* \subset S^5(U'^*)$ with multiplicity 1, and also $(\det \mathbf{C}^5)^{-2} \otimes \mathbf{S}_{3,1,1}(\mathbf{C}^3)^* \subset S^2(U'^*) \otimes (\det \mathbf{C}^3)^* \otimes \mathbf{S}_{2,1^4}(\mathbf{C}^5)^*$ with multiplicity 1, so this gives a 6×6 matrix

$$(\det \mathbf{C}^5)^{-2} \otimes \mathbf{S}_{3,1,1}(\mathbf{C}^3)^* \otimes \mathrm{Sym}(U'^*) \to \mathrm{Sym}(U'^*)(5) \oplus [(\det \mathbf{C}^3)^* \otimes \mathbf{S}_{2,1^4}(\mathbf{C}^5)^* \otimes \mathrm{Sym}(U'^*)(2)]$$

whose determinant is the degree 15 invariant.

Taking sections of (4.10), we see that the process of constructing the Chow form of C is a map that takes a section v to an element in $[(\det B)^2 \otimes \mathbf{S}_{3,1,1} A] \oplus [\wedge^4 B \otimes S^2 A]$. This can be interpreted as $\mathbf{GL}(A) \times \mathbf{GL}(B)$-equivariant linear maps $S^5(A \otimes \bigwedge^2 B) \to (\det B)^2 \otimes \mathbf{S}_{3,1,1} A$ and $S^2(A \otimes \bigwedge^2 B) \to \wedge^4 B \otimes S^2 A$. But such maps are unique up to scalar (checked with LiE [39]), so the claim follows. □

4.6 Projective Duality

Here is another approach:

- $G/P = \mathbf{Gr}(4, A^*) = \mathbf{P}(A)$.
- $\mathcal{U} = \mathcal{R}^* \otimes \bigwedge^2 \underline{B}$.
- $U' = \mathbf{C}^4 \otimes \bigwedge^2 \mathbf{C}^5$.
- $G' = (\mathbf{GL}_4(\mathbf{C}) \times \mathbf{GL}_5(\mathbf{C}))/\{\,(x, x^{-1}) \mid x \in \mathbf{C}^*\,\}$.

The space U' contains a G'-invariant degree 40 hypersurface. The polynomial f is described as follows: let $a_1, \dots, a_{40}$ be a basis for the Lie algebra $\mathbf{C} \oplus \mathfrak{sl}_4 \oplus \mathfrak{sl}_5$. Then for $x \in \mathbf{C}^4 \otimes \bigwedge^2 \mathbf{C}^5$, we have $f(x) = \det(A_1 x \cdots A_{40} x)$.

The corresponding degeneracy locus is a degree 10 hypersurface C' in $\mathbf{P}(A)$. By the Katz–Kleiman formula [26, Sect. 2.3.C], the projective dual of an Abelian variety $X \subset \mathbf{P}^N$ is a hypersurface of degree $(\dim X + 1)(\deg X)$. In our case, for $X = C$ from the last section, we get a hypersurface of degree 10.

Proposition 10. *For generic $v \in V$, the projective dual of C is C'.*

Proof. Let f be the equation for the hyperdiscriminant of $\mathbf{C}^4 \otimes \bigwedge^2 \mathbf{C}^5$. An element $x \in \mathbf{C}^4 \otimes \bigwedge^2 \mathbf{C}^5$ can be thought of as a 5×5 skew-symmetric matrix of linear forms on $\mathbf{P}^3$. Generically, the 4×4 Pfaffians give a 0-scheme of degree 5, and $f(x) = 0$ if and only if this 0-scheme is non-reduced.

So for generic $v \in V$, we first write $V = \mathrm{H}^0(\mathbf{P}(A); \mathcal{Q} \otimes \bigwedge^2 B)$. Then as a 4×10 matrix (along the fibers), v has full rank 4 over each point in $\mathbf{P}(A)$. Now pick $H \in \mathbf{P}(A)$. Then H is a hyperplane in $\mathbf{P}(A^*)$ and we have a canonical identification $\mathrm{H}^0(H; \bigwedge^2 B(1)) = (\mathcal{Q} \otimes \bigwedge^2 B)(H)$ (where the notation (H) means "fiber at H"). So $H \cap C$ is identified with the 5 points mentioned above, and hence $H \in C^\vee$ if and only if $f(v(H)) = 0$, which proves our claim. $\square$

5 $\bigwedge^3 \mathbf{C}^9$

Let V be a vector space of dimension 9. The relevant data:

- $U = \bigwedge^3 V$.
- $G = \mathbf{GL}(V)$.
- $G/P = \mathbf{P}(V^*) = \mathbf{Gr}(1, V^*)$.
- $\mathcal{U} = \bigwedge^2 \mathcal{Q}^* \otimes \mathcal{R}^* \cong \bigwedge^2 \mathcal{Q}^* \otimes \mathcal{O}(1) \cong \Omega^2(3)$.
- $U' = \bigwedge^2 \mathbf{C}^8$.
- $G' = \mathbf{GL}_8(\mathbf{C})$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 12, 18, 24, 30, and the graded Weyl group W is Shephard–Todd group 32 [56, Sect. 9].

Remark 1. The invariants for W acting on its reflection representation $\mathfrak{h}$ were explicitly calculated by Maschke [41]. It is known that the GIT quotient $U /\!\!/ G \cong \mathfrak{h}/W$ contains an open set which is isomorphic to the moduli space of genus 2 curves C with a marked Weierstrass point (i.e., ramification point for the hyperelliptic map). This was shown by Burkhardt [12]. See also [19, Sect. 4.3] for a modern treatment and further discussion.

Remark 2. The orbits in $\bigwedge^3 \mathbf{C}^9$ were classified in [21], but the connection to geometric objects as treated here is not made.

5.1 Modules Over $\mathcal{O}_{U'}$

The orbits in U' are given by the vanishing of Pfaffians of various sizes. We are only interested in the vanishing locus of the 8×8 Pfaffian and the vanishing locus of the 6×6 Pfaffians. The latter is described in Example 11. We denote their global versions in $\mathcal{U}$ by $\mathcal{Y}$ and $\mathcal{X}$, respectively.

5.2 Geometric Data from a Section

The subvariety $\mathcal{Y}$ has the following resolutions over $\mathcal{O}_\mathcal{U} = \mathrm{Sym}(\bigwedge^2 \mathcal{Q} \otimes \mathcal{O}(-1))$:

$$0 \to (\det \mathcal{Q}) \otimes \mathcal{O}_\mathcal{U}(-4) \to \mathcal{O}_\mathcal{U} \to \mathcal{O}_\mathcal{Y} \to 0. \tag{5.3}$$

We can simplify this by noting that $\det \mathcal{Q} = \mathcal{O}(1)$.

From Example 11, the subvariety $\mathcal{X}$ has this resolution:

$$\begin{aligned} 0 \to \mathcal{O}_\mathcal{U}(-9) \to \bigwedge^2 \mathcal{Q} \otimes \mathcal{O}_\mathcal{U}(-7) \to \mathbf{S}_{2,1^6}\mathcal{Q} \otimes \mathcal{O}_\mathcal{U}(-7) \\ \to (S^2\mathcal{Q} \otimes \mathcal{O}_\mathcal{U}(-4)) \oplus ((S^2\mathcal{Q})^* \otimes \mathcal{O}_\mathcal{U}(-5)) \\ \to \mathbf{S}_{2,1^6}\mathcal{Q} \otimes \mathcal{O}_\mathcal{U}(-4) \to \bigwedge^6 \mathcal{Q} \otimes \mathcal{O}_\mathcal{U}(-3) \to \mathcal{O}_\mathcal{U} \to \mathcal{O}_\mathcal{X} \to 0. \end{aligned} \tag{5.4}$$

So for $v \in U^{\mathrm{gen}}$, we have that $Y = \mathcal{Y} \cap v(\mathbf{P}(V^*))$ and $X = \mathcal{X} \cap v(\mathbf{P}(V^*))$ will have codimensions 1 and 6 in $\mathbf{P}(V^*)$, respectively.

The self-duality of the resolution for $\mathcal{O}_X$ shows that $\omega_X = \mathcal{E}xt^6(\mathcal{O}_X, \mathcal{O}(-9)) = \mathcal{O}_X$.

Theorem 3. *For $v \in U^{\mathrm{gen}}$, we have $\mathrm{h}^i(X; \mathcal{O}_X) = \binom{2}{i}$. In particular, for $v \in U^{\mathrm{sm}}$, X is a torsor over an Abelian surface.*

(We will only do this calculation once. For the remaining examples, we leave it to the reader.)

Proof. First, we replace the sheaf of algebras $\mathcal{O}_\mathcal{U}$ with the structure sheaf of $\mathbf{P}(V^*)$ in (5.4). In the notation of Sect. 2.3, we have $\mathbf{P}(V^*) = \mathbf{Fl}(1; V^*)$ and $\mathbf{S}_\lambda \mathcal{Q} \otimes \mathcal{O}_{\mathbf{P}(V^*)}(d) = \mathcal{R}(\mu)$ where $\mu = (d, -\lambda_8, -\lambda_7, \dots, -\lambda_1)$ and $\rho = (8, 7, \dots, 1, 0)$. In particular, when we add ρ to the sequence μ for any term in homological degrees $\{2, 3, 4, 5\}$ of (5.4), all cohomology vanishes by Theorem 8 because there will always be a repeating term. For example, $\mathbf{S}_{2,1^6}\mathcal{Q} \otimes \mathcal{O}(-4)$ has vanishing cohomology because $\mu + \rho = (4, 7, 5, 4, 3, 2, 1, 0, -2)$.

Now consider the remaining terms. For $\bigwedge^6 \mathcal{Q} \otimes \mathcal{O}(-3)$, we have $\lambda = (1^6)$, so $\mu + \rho = (5, 7, 6, 4, 3, 2, 1, 0, -1)$. We can sort this using 2 consecutive swaps, and subtracting ρ again leaves us with a sequence of all -1. Hence, Theorem 8 says $\mathrm{H}^2(\mathbf{P}(V^*); \bigwedge^6 \mathcal{Q} \otimes \mathcal{O}(-3)) = \det V$. By similar considerations (or Serre duality), one can show that $\mathrm{h}^6(\bigwedge^2 \mathcal{Q} \otimes \mathcal{O}(-7)) = 1$. Finally, we already know that $\mathrm{h}^0(\mathcal{O}_{\mathbf{P}(V^*)}) = \mathrm{h}^8(\mathcal{O}(-9)) = 1$.

Now the result follows from a spectral sequence argument (or equivalently by splicing (5.4) into short exact sequences). The last statement follows from Theorem 1. □

From (5.3), we see that Y is a cubic hypersurface. In fact, it is the Coble cubic of X; see [3, 13] for more information on the Coble cubic.

Proposition 4. *The polarization on X induced by $\mathcal{O}_X(1)$ is indecomposable and of type* $(3, 3)$.

Proof. We have $\mathrm{h}^0(\mathcal{O}_X(1)) = 9$, so X is embedded via a complete linear series. So $K(\mathcal{O}(1)) = \mathbf{Z}/D \oplus \mathbf{Z}/D$ where $D = (3, 3)$ or $D = (1, 9)$. Since X is the singular locus of a cubic hypersurface, we conclude that it is the intersection of the quadrics (partial derivatives) that contain it. However, an Abelian surface in $\mathbf{P}^8$ with a $(1, 9)$-polarization cannot be the intersection of the quadrics containing it [28, Remark 3.2], so we conclude that the polarization is $(3, 3)$.

Furthermore, X is not the product of 2 elliptic curves as a polarized Abelian variety. To see this, first consider elliptic curves $E_1, E_2 \subset \mathbf{P}^2$ embedded as cubics. Then the product polarization is given by the Segre embedding $E_1 \times E_2 \subset \mathbf{P}^2 \times \mathbf{P}^2 \subset \mathbf{P}^8$. The quadratic equations vanishing on $\mathbf{P}^2 \times \mathbf{P}^2$ give $\bigwedge^2 \mathbf{C}^3 \otimes \bigwedge^2 \mathbf{C}^3$, and the quotient of $S^2(\mathbf{C}^3 \otimes \mathbf{C}^3)$ by this space is $S^2\mathbf{C}^3 \otimes S^2\mathbf{C}^3$, none of which vanish on $E_1 \times E_2$. Hence, $E_1 \times E_2$ is not the intersection of quadrics in $\mathbf{P}^8$. □

Remark 5. It could be the case that X is abstractly isomorphic to the product of two elliptic curves if we ignore the polarizations. See, for example, [50, Sect. 12].

Remark 6. Given a curve C of genus 2, the moduli space $SU_C(3)$ admits a degree 2 map to $\mathbf{P}^8$ which is branched along a degree 6 hypersurface. This hypersurface is projectively dual to the Coble cubic (see [43] and [46] for two different proofs of this, together with some more discussion of the hypersurfaces).

5.3 Macaulay2 Code

Here we give some Macaulay2 code for generating examples of sections. To do this, first consider the short exact sequence of vector bundles over $\mathbf{P}(V^*)$:

$$0 \to \bigwedge^2 \mathcal{Q}^* \otimes \mathcal{O}(1) \to \bigwedge^2 \underline{V} \otimes \mathcal{O}(1) \to \mathcal{Q}^* \otimes \mathcal{O}(2) \to 0.$$

Taking sections, we get the inclusion $\bigwedge^3 V \subset \bigwedge^2 V \otimes V$. Since this map is $\mathbf{GL}(V)$-equivariant, it must be the comultiplication map. Thus, given $v \in \bigwedge^3 V$, we can comultiply to get an element of $\bigwedge^2 V \otimes V$, which we may think of as a skew-symmetric 9×9 matrix φ_v of linear forms on $\mathbf{P}(V^*)$.

To restrict back to skew-symmetric matrix on $\mathcal{Q}^*$, we restrict our attention to an affine open set. Pick homogeneous coordinates $z_1, \dots, z_9$ on $\mathbf{P}(V^*)$ and consider the open set given by $z_9 = 1$. Then $z_1, \dots, z_8$ give a trivialization for $\mathcal{Q}$ over this open set, so the relevant data is the upper-left 8×8 submatrix φ'_v of φ_v. Now our degeneracy loci correspond to the usual Pfaffian loci of this submatrix. Assuming that there are no components contained in the hyperplane $z_9 = 0$, we get the homogeneous ideals of these degeneracy loci by saturating with respect to z_9.

The function `basicMat` below takes as input `s = (i,j,k)` and computes the matrix φ'_v for $v = z_i \wedge z_j \wedge z_k$. Then we take random coefficients and calculate the Pfaffian ideals.

```
P=101;
R = ZZ/P[z_1..z_9];
basicMat = s -> (
     ans := mutableMatrix(0*id_(R^8));
     ans_(s_0-1,s_1-1) = -z_(s_2);
     ans_(s_1-1,s_0-1) = z_(s_2);
     if not(member(9,s)) then (
          ans_(s_0-1,s_2-1) = z_(s_1);
          ans_(s_2-1,s_0-1) = -z_(s_1);
          ans_(s_1-1,s_2-1) = -z_(s_0);
          ans_(s_2-1,s_1-1) = z_(s_0);
          );
     matrix ans
     )

M=0;
for i in subsets(1..9,3) do M = M + random(ZZ/P)
                                   * basicMat(i);
I = saturate(pfaffians(8,M),ideal(z_9));
J = saturate(pfaffians(6,M),ideal(z_9));
K = saturate(pfaffians(4,M),ideal(z_9));
```

Then generically, I is the ideal of the cubic hypersurface, J is the homogeneous ideal of the Abelian surface, and K is the unit ideal.

Example 7. Here is an example calculation we can do with this. Define `M` by

```
M=0;
for i from 1 to 3 do
for j from 4 to 6 do
for k from 7 to 9 do M = M + random(ZZ/P)
                            * basicMat({i,j,k}).
```

This gives us a generic vector in $\mathbf{C}^3 \otimes \mathbf{C}^3 \otimes \mathbf{C}^3 \subset \bigwedge^3 V$, where the subspace corresponds to the choice of a decomposition $V = \mathbf{C}^3 \oplus \mathbf{C}^3 \oplus \mathbf{C}^3$. Using the command `primaryDecomposition J` we see that J now defines a reducible variety of degree 12 with 2 components each contained in $z_1 = z_2 = z_3 = 0$ and $z_4 = z_5 = z_6 = 0$. In fact, we know that J is supposed to have degree 18. There will be a third component of degree 6 inside $z_7 = z_8 = z_9 = 0$ which we do not see because of our choice of open affine (one can check this by modifying the definition of `basicMat` to use a different affine trivialization). In fact, if we intersect any two of the three ideals in the primary decomposition of J, we will get an ideal generated by 6 linear equations and a cubic, which is exactly a plane cubic curve.

6 $\bigwedge^4 \mathbf{C}^8$

Let V be a vector space of dimension 8. The relevant data:

- $U = \bigwedge^4 V$.
- $G = \mathbf{GL}(V)$.
- $G/P = \mathbf{P}(V^*) = \mathbf{Gr}(1, V^*)$.
- $\mathcal{U} = \bigwedge^3 \mathcal{Q}^* \otimes \mathcal{R}^* \cong \bigwedge^3 \mathcal{Q}^* \otimes \mathcal{O}(1) \cong \Omega^3(4)$.
- $U' = \bigwedge^3 \mathbf{C}^7$.
- $G' = \mathbf{GL}_7(\mathbf{C})$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 2, 6, 8, 10, 12, 14, 18, and the graded Weyl group W is the Weyl group of type E_7 [56, Sect. 9].

Remark 1. Letting $\mathfrak{h}$ be the 7-dimensional reflection representation of W, it is known that the GIT quotient $U /\!\!/ G \cong \mathfrak{h}/W$ has an open subset isomorphic to the moduli space of smooth plane quartics (i.e., non-hyperelliptic genus 3 curves) with a marked flex point (i.e., a point with tangency of order ≥ 3; there are 24 of them for a generic curve). See, for example, [18, Sect. IX.7, Remark 7] (but note that the 21 mentioned there should be 24) or [40, Proposition 1.11].

6.1 *Modules Over $\mathcal{O}_{U'}$*

We are interested in 3 orbits in U'. The first is a degree 7 hypersurface Y'. The singular locus of Y' is an orbit closure X' of codimension 4 in U', and the singular locus of X' is an orbit closure Z' of codimension 7 in U'. Let W be a 7-dimensional vector space with $U' = \bigwedge^3 W$. Set $A = \mathrm{Sym}(\bigwedge^3 W^*) = \mathcal{O}_{U'}$. The minimal free resolutions of these orbit closures are given as follows.

The minimal free resolution of $\mathcal{O}_{X'}$ is given by

$$0 \to (\det W^*)^5 \otimes W^*(-12) \to (\det W^*)^4 \otimes \bigwedge^2 W^*(-10) \to$$
$$(\det W^*)^2 \otimes \mathbf{S}_{2,1^5}(W^*)(-7) \to (\det W^*)^2 \otimes \bigwedge^4 W^*(-6) \to A \to \mathcal{O}_{X'} \to 0. \quad (6.2)$$

The minimal free resolution for $\mathcal{O}_{Z'}$ is given by

$$0 \to \mathbf{S}_{7^6,6}W^*(-16) \to \mathbf{S}_{7,6^5,5}W^*(-14) \to \mathbf{S}_{6^2,5^4,4}W^*(-12) \to \mathbf{S}_{5^3,4^3,3}W^*(-10)$$
$$\to \mathbf{S}_{4^4,3^2,2}W^*(-8) \to \mathbf{S}_{3^5,2,1}W^*(-6) \to \mathbf{S}_{2^6}W^*(-4) \to A \to \mathcal{O}_{Z'} \to 0. \quad (6.3)$$

We will also use the fact that the local ring of $\mathcal{O}_{X'}$ at the generic point of $\mathcal{O}_{Z'}$ is a Cohen–Macaulay ring of type 3, i.e., its canonical module is minimally generated by 3 elements. This calculation was done by Federico Galetto [24].

We set M to be a certain twist of the cokernel of the dual of the last differential in (6.2). In particular, it has a presentation of the form

$$(\det W^*)^2 \otimes \bigwedge^2 W(-4) \to \det W^* \otimes W(-2) \to M \to 0. \quad (6.4)$$

So up to a grading shift, M is the canonical module of $\mathcal{O}_{X'}$.

Lemma 2. *There is a G'-equivariant $\mathcal{O}_{X'}$-linear isomorphism $S^2M \cong I_{Z',X'}$.*

Proof. From (6.4), we get the presentation

$$(\det W^*)^3 \otimes W \otimes \bigwedge^2 W \otimes \mathcal{O}_{X'}(-6) \to (\det W^*)^2 \otimes S^2 W \otimes \mathcal{O}_{X'}(-4) \to S^2 M \to 0.$$

Also, the presentation of $I_{Z'}$ in $\mathcal{O}_{U'}$ is given by

$$(\det W^*)^3 \otimes \mathbf{S}_{2,1} W \otimes \mathcal{O}_{U'}(-6) \to (\det W^*)^2 \otimes S^2 W \otimes \mathcal{O}_{U'}(-4) \to I_{Z'} \to 0.$$

To get a presentation over $\mathcal{O}_{X'}$ we have to add in the relations $(\det W^*)^3 \otimes \bigwedge^3 W$ which come from the ideal generators of $I_{X'}$. Hence we see that the S^2M and $I_{Z',X'}$ have the same presentations. By equivariance, we get an isomorphism of these presentations up to a choice of a scalar, and this implies the desired isomorphism $S^2M \cong I_{Z',X'}$. □

We define $\mathcal{Y}$, $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{M}$ to be the global versions of Y', X', Z', and M.

6.2 Geometric Data from a Section

Now choose a section $v \in U^{\text{gen}}$, and set $Y = v(\mathbf{P}(V^*)) \cap \mathcal{Y}$, $X = v(\mathbf{P}(V^*)) \cap \mathcal{X}$, and $Z = v(\mathbf{P}(V^*)) \cap \mathcal{Z}$.

To get the locally free resolution of X over $\mathcal{O}_{\mathbf{P}^7}$, we replace W with $\mathcal{Q}^*$ and replace $(-i)$ with $\mathcal{O}(-i)$. After using $\det \mathcal{Q} = \mathcal{O}(1)$ to simplify, we get

$$0 \to \mathcal{Q} \otimes \mathcal{O}(-7) \to \bigwedge^2 \mathcal{Q} \otimes \mathcal{O}(-6) \to \mathbf{S}_{2,1^5}(\mathcal{Q}) \otimes \mathcal{O}(-5) \to \bigwedge^4 \mathcal{Q} \otimes \mathcal{O}(-4) \to \mathcal{O}_{\mathbf{P}^7} \to \mathcal{O}_X \to 0. \tag{6.6}$$

Furthermore, note that $\omega_X = \mathcal{M} \otimes \mathcal{O}_{v(\mathbf{P}(V^*))}$. From Lemma 2, we get an $\mathcal{O}_X$-linear map $\mu\colon S^2\omega_X \to \mathcal{O}_X$ and hence an $\mathcal{O}_X$-algebra structure on $\mathcal{O}_X \oplus \omega_X$. We define $\widetilde{X} = \operatorname{Spec}_{\mathcal{O}_X}(\mathcal{O}_X \oplus \omega_X)$. As in Theorem 3, we can check that $\omega_{\widetilde{X}} = \mathcal{O}_{\widetilde{X}}$ and that $\mathrm{h}^i(\widetilde{X}; \mathcal{O}_{\widetilde{X}}) = \mathrm{h}^i(X; \mathcal{O}_X \oplus \omega_X) = \binom{3}{i}$.

Define $U^{\mathrm{sm}+}$ to be the subset of U^{sm} where the Cohen–Macaulay type of X along Z remains 3.

Proposition 3. *If $v \in U^{\mathrm{sm}+}$, then $\widetilde{X}$ is smooth. In particular, $\widetilde{X}$ is a torsor over an Abelian threefold, and X is the Kummer variety of $\widetilde{X}$.*

Proof. Let $\mathcal{I}_Z$ denote the ideal sheaf of Z in $\mathcal{O}_X$. We check smoothness locally. Let P be a nonsingular point of X'. Then $\mathcal{I}_{Z,P} = \mathcal{O}_{X,P}$ and $\omega_{X,P} = \mathcal{O}_{X,P}$, and μ_P is just the multiplication map. Then $\mathcal{O}_{\widetilde{X},P} \cong \mathcal{O}_{X,P}[t]/(t^2-1)$ as a ring, so the points over P are nonsingular. Now let $P \in Z$ be a singular point. Set $R = \mathcal{O}_{X,P}$. Then $\mathfrak{m} = \mathcal{I}_{Z,P}$ is the maximal ideal of R. Write $\omega_R = \omega_{X,P}$ and $S = \mathcal{O}_{\widetilde{X},P}$. Then $S = R \oplus \omega_R$ as an R-module, and $\mathfrak{n} = \mathfrak{m} \oplus \omega_R$ is the unique maximal ideal of S: $S/\mathfrak{n} \cong R/\mathfrak{m}$, and any element not in $\mathfrak{n}$ is $(r, 0)$ for $r \in R \setminus \mathfrak{m}$, so is a unit. Since the multiplication map $S^2\omega_R \to \mathfrak{m}$ is surjective, we have that $\mathfrak{n}^2 = \mathfrak{m} \oplus \mathfrak{m} \cdot \omega_R$. Hence $\mathfrak{n}/\mathfrak{n}^2 \cong \omega_R/\mathfrak{m} \cdot \omega_R$. The dimension of this space over $R/\mathfrak{m}$ is the Cohen–Macaulay type of R, which is 3, and implies that S is a regular local ring. So we have shown that $\widetilde{X}$ is smooth.

Using Theorem 1, the above facts imply that $\widetilde{X}$ is a torsor over an Abelian threefold. Temporarily choose a point $P \in \pi^{-1}(Z)$ to be the origin of $\widetilde{X}$ so that it acquires a group structure. The map ι which swaps points in the same fiber of π gives an involution on $\widetilde{X}$. Then ι fixes P, and there is an induced linear map on the tangent space at P. Since ι has order 2, the induced linear map is the negation map. But this is the derivative of the inversion map on $\widetilde{X}$, so we conclude that ι is the inversion map on $\widetilde{X}$. So X is the Kummer variety $\widetilde{X}/\langle\iota\rangle$. □

Proposition 4. *$\mathcal{L} = \pi^*\mathcal{O}_X(1)$ defines an indecomposable $(2,2,2)$-polarization on $\widetilde{X}$.*

Proof. Note that

$$\mathrm{h}^0(\widetilde{X}; \mathcal{L}) = \mathrm{h}^0(X; \pi_*\mathcal{L}) = \mathrm{h}^0(X; (\mathcal{O}_X \oplus \omega_X) \otimes \mathcal{O}_X(1)) = 8,$$

so the map $\widetilde{X} \to \mathbf{P}(V^*)$ is given by a complete linear series. Hence $\mathcal{L} = \mathcal{L}(a,b,c)$ where $abc = 8$. By [42, Sect. 2, Corollary 4], each of a, b, c must be even, so we have $a = b = c = 2$. Furthermore, $(\widetilde{X}, \mathcal{L})$ is indecomposable since by [7, Theorem 4.8.2], if $(\widetilde{X}, \mathcal{L})$ is a product of s Abelian varieties, then the degree of the map $\widetilde{X} \to X \subset \mathbf{P}(V^*)$ is 2^s. □

In particular, this Abelian threefold is the Jacobian of some genus 3 curve C.

Proposition 5. *C is not hyperelliptic.*

Proof. Using the locally free resolution (6.6), the map $\mathrm{H}^0(\mathcal{O}_{\mathbf{P}(V^*)}(n)) \to \mathrm{H}^0(\mathcal{O}_X(n))$ is surjective for all $n \geq 0$. Hence, X is projectively normal, so C is not hyperelliptic by [36, Sect. 2.9.3]. □

Remark 6. The quartic hypersurface has an interpretation as the embedding of $SU_C(2)$ (see Sect. 3.2) via its theta divisor [45]. This is also known as the Coble quartic; see [3, 14] for more details.

6.3 Flag Variety

Here is another approach which was pointed out to us by Jack Thorne:

- $G/P = \mathbf{Fl}(1, 7, V^*)$
- $\mathcal{U} = \mathcal{R}_1^* \otimes \bigwedge^3 (\mathcal{R}_7/\mathcal{R}_1)^*$
- $U' = \bigwedge^3 \mathbf{C}^6$
- $G' = \mathbf{GL}_6(\mathbf{C})$

We are interested in the smallest orbit closure in U', which is the affine cone over $\mathbf{Gr}(3, 6)$. This has codimension 10. In characteristic 0, its minimal free resolution is as follows (the coordinate ring is $\mathrm{Sym}(\bigwedge^3 \mathbf{C}^6)$ and we only list the partitions):

$$\begin{aligned}
\mathbf{F}_1 &= (2, 1^4),\\
\mathbf{F}_2 &= (2^4, 1) + (3, 2, 1^4),\\
\mathbf{F}_3 &= (3^2, 2^2, 1^2) + (3^5) + (5, 2^5),\\
\mathbf{F}_4 &= (4^3, 2^3) + (4^2, 3^3, 1) + (5, 3^3, 2^2),\\
\mathbf{F}_5 &= (5, 4^2, 3^2, 2)^{\oplus 2},\\
\mathbf{F}_6 &= (5^2, 4^3, 2) + (5^3, 3^3) + (6, 4^3, 3^2),\\
\mathbf{F}_7 &= (7, 4^5) + (5^5, 2) + (6^2, 5^2, 4^2),\\
\mathbf{F}_8 &= (6^4, 5, 4) + (7, 6, 5^4),\\
\mathbf{F}_9 &= (7, 6^4, 5),\\
\mathbf{F}_{10} &= (7^6).
\end{aligned}$$

To see the ranks, we include the graded Betti table (this follows Macaulay2 notation, so a term d in row i and column j represents the free module $A(-i-j)^{\oplus d}$ in homological degree j).

	0	1	2	3	4	5	6	7	8	9	10
total:	1	35	140	301	735	1080	735	301	140	35	1
0:	1	.	.	.	.	.	.	.	.	.	.
1:	.	35	140	189	.	.	.	.	.	.	.
2:	.	.	.	112	735	1080	735	112	.	.	.
3:	.	.	.	.	.	.	.	189	140	35	.
4:	.	.	.	.	.	.	.	.	.	.	1

Remark 7. The resolution changes in characteristic 2:

	0	1	2	3	4	5	6	7	8	9	10
total:	1	35	141	302	735	1080	735	302	141	35	1
0:	1	.	.	.	.	.	.	.	.	.	.
1:	.	35	140	190	.	.	.	.	.	.	.
2:	.	.	1	112	735	1080	735	112	1	.	.
3:	.	.	.	.	.	.	.	190	140	35	.
4:	.	.	.	.	.	.	.	.	.	.	1

and in characteristic 3:

	0	1	2	3	4	5	6	7	8	9	10
total:	1	35	140	321	756	1082	756	321	140	35	1
0:	1	.	.	.	.	.	.	.	.	.	.
1:	.	35	140	189	20	1	.	.	.	.	.
2:	.	.	.	132	736	1080	736	132	.	.	.
3:	.	.	.	.	.	1	20	189	140	35	.
4:	.	.	.	.	.	.	.	.	.	.	1

In all other characteristics, the Betti numbers agree with characteristic 0.

The corresponding degeneracy locus X_2 is a torsor over an Abelian threefold: For each partition λ in the resolution for $\mathbf{Gr}(3,6)$, we get the sheaf $\mathcal{R}_1^{|\lambda|/3} \otimes \mathbf{S}_\lambda(\mathcal{R}_7/\mathcal{R}_1)$. Also, the canonical sheaf of $\mathbf{Fl}(1,7,V^*)$ is $(\det \mathcal{R}_7)^7 \otimes \mathcal{R}_1^7$, which is the last term of the resolution. This shows that $\omega_{X_2} = \mathcal{O}_{X_2}$. As in Theorem 3, we can use Borel–Weil–Bott to get that $\mathrm{h}^i(X_2; \mathcal{O}_{X_2}) = \binom{3}{i}$.

Proposition 8. $X_2 \cong \widetilde{X}$.

Proof. Let $\pi : \mathbf{Fl}(1,7,V^*) \to \mathbf{P}(V^*)$ be the projection map. We claim that $\pi(X_2) = X$ and that $\pi|_{X_2}$ is a finite morphism of degree 2. Note that $\pi^*\mathcal{Q} = V^*/\mathcal{R}_1$ and $\pi^*\mathcal{R} = \mathcal{R}_1$. For $x \in \mathbf{P}(V^*)$, pick $(x \subset H) \in \pi^{-1}(x)$. We have a surjection

$$(\bigwedge^3 \mathcal{Q}^* \otimes \mathcal{R}^*)(x) \to (\bigwedge^3 (\mathcal{R}_7/\mathcal{R}_1)^* \otimes \mathcal{R}_1^*)(x \subset H),$$

and so $(x \subset H) \in X_2$ if and only if $v(x) \in (V^*/x)^*$ is a pure tensor when mapped to $(H/x)^*$. By Remark 19, there exists such an H if and only if $x \in X$, so $\pi(X_2) = X$. Furthermore, there exists exactly 2 such H if $x \in X \setminus Z$ and there exists exactly 1 such H if $x \in Z$. So $\pi|_{X_2}$ is a finite morphism of degree 2. We conclude that $X_2 \cong \widetilde{X}$. □

6.4 Projective Duality

We can instead work with the following data:

- $G/P = \mathbf{Gr}(7, V^*) = \mathbf{P}(V)$.
- $\mathcal{U} = \bigwedge^4 \mathcal{R}^*$.
- $U' = \bigwedge^4 \mathbf{C}^7$.
- $G' = \mathbf{GL}_7(\mathbf{C})$.

The orbit classifications in $\bigwedge^4 \mathbf{C}^7$ and $\bigwedge^3 \mathbf{C}^7$ are the same, so we can proceed as in Sect. 6.2.

Under the identification $\mathbf{Gr}(7, V^*) = \mathbf{P}(V)$, the bundle $\bigwedge^4 \mathcal{R}^*$ becomes $\bigwedge^3 \mathcal{Q}^* \otimes \mathcal{O}(1)$. Hence, given a section $v \in U^{\mathrm{sm}+}$, we get a quartic hypersurface Y^d and a Kummer threefold X^d in $\mathbf{P}(V)$.

Proposition 9. *We have isomorphisms $X \cong X^d$ and $Y^d \cong Y$. Furthermore, Y and Y^d are projectively dual varieties.*

Proof. We first show that $X \cong X^d$. Consider the variety $X_2 \subset \mathbf{Fl}(1, 7, V^*)$ constructed in the previous section from v. Let $\pi_2 \colon \mathbf{Fl}(1, 7, V^*) \to \mathbf{Gr}(7, V^*)$ be the projection map. We claim that $\pi_2(X_2) = X^d$ and that π_2 is finite of degree 2. This will prove the claim.

Note that $\pi_2^* \mathcal{R} = \mathcal{R}_7$. For $H \in \mathbf{Gr}(7, V^*)$, pick $(x \subset H) \in \pi_2^{-1}(H)$. We have a surjection

$$(\bigwedge^4 \mathcal{R}^*)(H) \to (\bigwedge^3 (\mathcal{R}_7/\mathcal{R}_1)^* \otimes \mathcal{R}_1^*)(x \subset H)$$

given by comultiplication. So $(x \subset H) \in X_2$ if and only if $v(x) \in H^*$ is a pure tensor when mapped to $(H/x)^*$. After picking a volume form in $\bigwedge^7 \mathbf{C}^7$, we can identify $\bigwedge^4 \mathbf{C}^{7*}$ and $\bigwedge^3 \mathbf{C}^7$, in which case the rest of the argument is similar to the proof of Proposition 8.

So X and X^d are embedded by dual linear series. By [49, Theorem 3.1], the Coble quartics of X and X^d are isomorphic and projective dual to one another. □

6.5 Doing Calculations

We have written some Macaulay2 code for calculating the degeneracy loci in Sect. 6.2. However, it is messy so we do not include it here. We will just comment on how one can practically go about these calculations and explain the analogue of Example 7.

First, we need to find a way to calculate the ideals of the appropriate low-codimension orbits in $\bigwedge^3 \mathbf{C}^7$. An explicit construction of the equation f for the degree 7 hypersurface was given in [37, Remark 4.4]. Namely, he constructs two symmetric 7×7 matrices $\varphi(x)$ and $\varphi^*(x)$ whose entries are homogeneous polynomials of degrees 3 and 4, respectively, such that $\varphi(x)\varphi^*(x) = f(x)I_7$, i.e., the pair $(\varphi(x), \varphi^*(x))$ is a matrix factorization for f. Explicit summation formulas are given for the entries of these matrices.

From (6.2), the Jacobian ideal of f is reduced and gives the equations for the codimension 4 orbit. The ideal for the codimension 7 orbit is given by the entries of the matrix $\varphi^*(x)$. One possibility to check this is to use that the representation generating this ideal is $S^2\mathbf{C}^7$ (up to a power of determinant) (6.3) and that this representation has multiplicity 1 in $S^4(\bigwedge^3 \mathbf{C}^{7*})$.

In practical computations, the product of the matrices $\varphi(x)$ and $\varphi^*(x)$ can be calculated in a few seconds. A possible way to proceed is to evaluate these against a chosen section v (after choosing an open affine in $\mathbf{P}^7$ to work over) and then to take the product in order to calculate the equation of the codimension 1 degeneracy locus. If the section is generic, then the Jacobian ideal of this equation will be the same as evaluating the Jacobian ideal of f at v (note that calculating f and its Jacobian ideal before evaluating may take a long time).

Example 10. Suppose we choose a decomposition $V = \mathbf{C}^2 \oplus \mathbf{C}^2 \oplus \mathbf{C}^2 \oplus \mathbf{C}^2$, and we pick a generic vector in $\mathbf{C}^2 \otimes \mathbf{C}^2 \otimes \mathbf{C}^2 \otimes \mathbf{C}^2 \subset \bigwedge^4 V$. In this case, the codimension 4 degeneracy locus will have 4 irreducible components, all of degree 6. Taking any two of these components, the affine cone over their intersection is a bidegree $(2, 2)$ hypersurface in $\mathbf{C}[x, y] \otimes \mathbf{C}[s, t]$ whose multigraded Proj is a nonsingular curve of genus 1 (this agrees with the data obtained in [30, Sect. 3.1.2])

7 $\bigwedge_0^4 \mathbf{C}^8$

Let V be a vector space of dimension 8 equipped with a symplectic form $\omega \in \bigwedge^2 V^*$. The symplectic form gives an injective multiplication map $\bigwedge^2 V^* \to \bigwedge^4 V^*$ and we define $\bigwedge_0^4 V = \bigwedge_0^4 V^*$ to be the cokernel. The relevant data:

- $U = \bigwedge_0^4 V$.
- $G = \mathbf{Sp}(V)$.
- $G/P = \mathbf{P}(V^*) = \mathbf{Gr}(1, V^*)$.
- $\mathcal{U} = \bigwedge_0^3(\mathcal{R}^\perp/\mathcal{R}) \otimes \mathcal{R}^* \cong \bigwedge_0^3(\mathcal{R}^\perp/\mathcal{R}) \otimes \mathcal{O}(1)$.
- $U' = \bigwedge_0^3 \mathbf{C}^6$.
- $G' = \mathbf{Sp}_6(\mathbf{C})$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 2, 5, 6, 8, 9, 12, and the graded Weyl group W is the Weyl group of type E_6 [56, Sect. 9].

Remark 1. Letting $\mathfrak{h}$ be the 6-dimensional reflection representation of W, it is known that the GIT quotient $U/\!\!/G \cong \mathfrak{h}/W$ has an open subset isomorphic to the moduli space of smooth plane quartics (i.e., non-hyperelliptic genus 3 curves) with a marked hyperflex point (i.e., a point with tangency of order 4; the existence of such a point is a codimension 1 condition on the space of plane quartics). See [40, Proposition 1.15].

7.1 Modules Over $\mathcal{O}_{U'}$

Let W be a 6-dimensional vector space equipped with a symplectic form, so we can write $U' = \bigwedge_0^3 W$ and $G' = \mathbf{Sp}(W)$. Set $A = \mathrm{Sym}(\bigwedge_0^3 W^*) = \mathcal{O}_{U'}$.

There is a degree 4 G'-invariant hypersurface in U', which we denote by Y'. The other orbits X' and Z' have codimensions 4 and 7, respectively. The minimal free resolution of $\mathcal{O}_{X'}$ is

$$0 \to W^* \otimes A(-7) \to \bigwedge_0^2 W^* \otimes A(-6) \to S^2 W^* \otimes A(-4) \to \bigwedge_0^3 W^* \otimes A(-3) \\ \to A \to \mathcal{O}_{X'} \to 0, \tag{7.2}$$

which we can calculate by Macaulay2. Furthermore, the last matrix has corank 3 when specialized to a nonzero point in the highest weight orbit, so the localization of $\mathcal{O}_{X'}$ at such a point is a Cohen–Macaulay ring of type 3.

We set M to be a certain twist of the cokernel of the dual of the last differential of (7.2). In particular, it has a presentation of the form

$$\bigwedge_0^2 W(-2) \to W(-1) \to M \to 0. \tag{7.3}$$

The orbit closure Z' is the affine cone over the Lagrangian Grassmannian. The minimal free resolution for $\mathcal{O}_{Z'}$ is

$$0 \to A(-10) \to S^2 W^*(-8) \to \mathbf{S}_{[2,1]} W^*(-7) \to \mathbf{S}_{[2,1,1]} W^*(-6) \\ \to \mathbf{S}_{[2,1,1]} W^*(-4) \to \mathbf{S}_{[2,1]} W^*(-3) \to S^2 W^*(-2) \to A \to \mathcal{O}_{Z'} \to 0 \tag{7.4}$$

(recall that $\mathbf{S}_{[\lambda]}$ was defined in (2.7)).

Lemma 2. *There is a G'-equivariant $\mathcal{O}_{X'}$-linear isomorphism $S^2M \cong I_{Z',X'}$.*

Proof. The proof is the same as for Lemma 2. □

We define $\mathcal{Y}$, $\mathcal{X}$, $\mathcal{Z}$, and $\mathcal{M}$ to be the global versions of Y', X', Z', and M.

7.2 Geometric Data from a Section

Now choose a section $v \in U^{\text{gen}}$ and set $Y = v(\mathbf{P}(V^*)) \cap \mathcal{Y}$, $X = v(\mathbf{P}(V^*)) \cap \mathcal{X}$, and $Z = v(\mathbf{P}(V^*)) \cap \mathcal{Z}$.

To get the locally free resolution of X over $\mathcal{O}_{\mathbf{P}^7}$, we replace W with $\mathcal{R}^\perp/\mathcal{R}$ and replace $(-i)$ with $\mathcal{O}(-i)$. So the locally free resolution for $\mathcal{O}_X$ is

$$0 \to \mathcal{R}^\perp/\mathcal{R}(-7) \to \bigwedge_0^2(\mathcal{R}^\perp/\mathcal{R})(-6) \to S^2(\mathcal{R}^\perp/\mathcal{R})(-4) \to \bigwedge_0^3(\mathcal{R}^\perp/\mathcal{R})(-3)$$
$$\to \mathcal{O}_{\mathbf{P}^7} \to \mathcal{O}_X \to 0,$$

and the locally free resolution for ω_X is

$$0 \to \mathcal{O}_{\mathbf{P}^7}(-8) \to \bigwedge_0^3(\mathcal{R}^\perp/\mathcal{R})(-5) \to S^2(\mathcal{R}^\perp/\mathcal{R})(-4) \to \bigwedge_0^2(\mathcal{R}^\perp/\mathcal{R})(-2)$$
$$\to (\mathcal{R}^\perp/\mathcal{R})(-1) \to \omega_X \to 0,$$

Using Borel–Weil–Bott as in Theorem 3, we get $\mathrm{h}^0(\mathcal{O}_X) = 1$ and $\mathrm{h}^2(\mathcal{O}_X) = 3$, and all other cohomology vanishes. So by Serre duality, $\mathrm{h}^i(\mathcal{O}_X \oplus \omega_X) = \binom{3}{i}$.

From Lemma 2, we get a $\mathcal{O}_X$-linear multiplication map $\mu\colon S^2\omega_X \to \mathcal{O}_X$ and hence an $\mathcal{O}_X$-algebra structure on $\mathcal{O}_X \oplus \omega_X$. We define $\widetilde{X} = \mathrm{Spec}_{\mathcal{O}_X}(\mathcal{O}_X \oplus \omega_X)$ and let $\pi\colon \widetilde{X} \to X$ be the projection map.

We define $U^{\text{sm}+}$ to be the subset of U^{sm} where the Cohen–Macaulay type of X along Z remains 3.

Theorem 3. *If $v \in U^{\text{sm}+}$, then $\widetilde{X}$ is smooth. Furthermore, $\widetilde{X}$ is a torsor over an Abelian threefold, $\mathcal{L} = \pi^*\mathcal{O}_X(1)$ is an indecomposable $(2,2,2)$-polarization on $\widetilde{X}$, and $\widetilde{X}$ is not the Jacobian of a hyperelliptic curve.*

Proof. The proofs are analogous to the ones in Sect. 6.2. □

8 spin(16)

Let B be a vector space of dimension 16 equipped with a quadratic form $q \in S^2(B^*)$. Let $Z(q)$ be the zero locus of this quadratic form. We let $\mathbf{Spin}(B)$ be the simply connected double cover of $\mathbf{SO}(B)$. One construction is to realize it as a multiplicative subgroup of a Clifford algebra, and we set $\mathbf{GSpin}(B)$ to be the Clifford group generated by $\mathbf{Spin}(B)$ and the scalar matrices. There are two spin representations $\mathbf{spin}^{\pm}(B)$. The choice of spin representation will not affect our results, so we pick one and call it $\mathbf{spin}(B)$. Furthermore, $Z(q)$ possesses two "spinor bundles" which we call $\mathcal{S}^+$ and $\mathcal{S}^-$. These can be constructed as pushforwards of line bundles from the flag variety of $\mathbf{Spin}(B)$, or more direct geometric means (see [47] for details). There is a perfect pairing

$$\mathcal{S}^+ \otimes \mathcal{S}^- \to \mathcal{O}(-1)$$

[47, Theorem 2.8], and the sections of $\mathcal{S}^+(1)$ and $\mathcal{S}^-(1)$ give the two half-spin representations (which follows from their descriptions as pushforwards of line bundles).

The relevant data:

- $U = \mathbf{spin}(B)$.
- $G = \mathbf{GSpin}(B)$.
- $G/P = Z(q)$ (quadric hypersurface).
- $\mathcal{U} = \mathcal{S}^+(1)$.
- $U' = \mathbf{spin}^+(14)$.
- $G' = \mathbf{GSpin}_{14}(\mathbf{C})$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 2, 8, 12, 14, 18, 20, 24, 30, and the graded Weyl group is the Weyl group of type E_8 [56, Sect. 9].

Remark 1. A non-hyperelliptic genus 4 curve lies on a unique quadric in its canonical embedding, and the locus of curves C where this quadric is singular (i.e., C has a vanishing θ-characteristic) has codimension 1. This condition also implies that C has a unique degree 3 map to $\mathbf{P}^1$. If we further impose that this map has a point with non-simple ramification, the locus loses another dimension. Letting $\mathfrak{h}$ be the 8-dimensional reflection representation of W, the GIT quotient $U /\!\!/ G \cong \mathfrak{h}/W$ has an open subset which should be isomorphic to this moduli space. We could not find any mention of this in the existing literature.

To get the interpretation for $\mathbf{C}^7/W(E_7)$ (Remark 1), one first realizes $\mathbf{C}$ as the smooth locus of a cuspidal cubic. Then $(P_1, \dots, P_7)$ becomes 7 points in the plane, and blowing them up gives a del Pezzo surface. Its anticanonical divisor gives a map to $\mathbf{P}^2$ branched along the plane quartic mentioned in Remark 1.

When we do the same thing for 8 points, the corresponding del Pezzo surface is mapped to a quadric cone in $\mathbf{P}^3$ via twice its anticanonical divisor, and it is branched along a genus 4 curve. However, 8 points in the plane generically do not lie on a cuspidal cubic; this condition on the points corresponds to the ramification condition on the genus 4 curve mentioned above. This was brought to our attention by Igor Dolgachev.

The representation **spin**(14) has four orbits $Z(f)', Y', X', Z'$, of interest, which are of codimensions 1, 5, 10, and 14, respectively. All of them are Gorenstein and have rational singularities, but we have not yet calculated the minimal free resolutions for X' and Z' due to certain extension problems. The hypersurface $Z(f)$ has degree 8.

The results for this example are incomplete, so we will just state what we expect to be true. For a generic section $v \in U$, the above four orbits give degeneracy loci $Z(f), Y, X, Z$. What should happen is that there is a curve C of genus 4 such that X is the Kummer variety of Jac(C) and Z is its singular locus consisting of 256 points. Furthermore, we should have $Y = SU_C(2)$ (see Sect. 3.2 for the definition of $SU_C(2)$).

One can check that $Z(f)$ is a quartic hypersurface in $Z(q)$, and it should be an analogue of a Coble hypersurface. The curve C should be a non-hyperelliptic curve with vanishing theta characteristic: every non-hyperelliptic curve C can be written as a complete intersection of a quadric and cubic in its canonical embedding, and having a vanishing theta characteristic means that this quadric is singular. For comparison, the analogues of Coble hypersurfaces for genus 4 curves without a vanishing theta characteristic were studied in [48].

We remark that having the minimal free resolution of X' (and of an additional auxiliary module M') will allow one to prove that X is the Kummer variety of an Abelian fourfold, but there do not seem to be any cohomological characterizations of Jacobians amongst Abelian fourfolds.

9 $\mathbf{C}^4 \otimes$ spin(10)

We can write $\mathbf{C}^4 \otimes \mathbf{spin}(10) = \mathbf{spin}(6) \otimes \mathbf{spin}(10)$, so this can be considered as a subcase of **spin**(16) in the previous section.

Let A be a vector space of dimension 4 and B be a vector space of dimension 10 equipped with an orthogonal form $\omega \in S^2(B^*)$. The relevant data:

- $U = A \otimes \mathbf{spin}(B)$.
- $G = (\mathbf{GL}(A) \times \mathbf{GSpin}(B))/\{(x, x^{-1}) \mid x \in \mathbf{C}^*\}$.
- $G/P = \mathbf{P}(A) = \mathbf{Gr}(1, A)$.
- $\mathcal{U} = \mathcal{Q} \otimes \mathbf{spin}(\underline{B})$.
- $U' = \mathbf{C}^3 \otimes \mathbf{spin}(10)$.
- $G' = (\mathbf{GL}_3(\mathbf{C}) \times \mathbf{Spin}_{10}(\mathbf{C}))/\{(x, x^{-1}) \mid x \in \mathbf{C}^*\}$.

The ring of invariants $\mathrm{Sym}(U^*)^{(G,G)}$ is a polynomial ring with generators of degrees 8, 12, 20, 24, and the graded Weyl group is Shephard–Todd group 31 [56, Sect. 9].

9.1 Modules Over $\mathcal{O}_{U'}$

Let $A' = \mathbf{C}^3$ and let $G' = (\mathbf{GL}(A') \times \mathbf{Spin}(B))/\{(x, x^{-1}) \mid x \in \mathbf{C}^*\}$. Let $R = \mathrm{Sym}((A' \otimes \mathbf{spin}(B))^*)$.

There is a degree 12 G'-invariant hypersurface $X' \subset U'$, whose equation is described in [29, Sect. 3]. One can interpret the construction there as taking the determinant of a certain 3×3 symmetric matrix whose entries are quartic forms. More precisely, the representation $\mathbf{S}_{2,2}(\mathbf{spin}(B))$ contains a $\mathbf{Spin}(B)$-invariant, so we get polynomials P which span the representation $\mathbf{S}_{2,2}(A'^*) \subset S^4((A' \otimes \mathbf{spin}(B))^*)$. Then $\mathbf{S}_{2,2}(A'^*) \cong S^2(A') \otimes (\det A'^*)^2$, so we can interpret the space of P as the space of linear functions on symmetric 3×3 matrices of the form

$$\varphi\colon A' \to A'^* \otimes (\det A')^2, \tag{9.1}$$

and the equation for X' is the determinant of this matrix.

Proposition 1. *The 2×2 minors of φ define a radical ideal of codimension 3, and the corresponding variety is the singular locus Z' of X'.*

Proof. The maximal codimension of the submaximal minors of a symmetric matrix is 3. Since this ideal is G'-equivariant, it must cut out a union of orbit closures. Since they do not vanish on X' and there are no orbit closures of codimension 2, this ideal has codimension 3 and in particular defines a Cohen–Macaulay variety.

There are 2 orbits of codimension 3. We can check on representatives that the 2×2 minors vanish on only one of these orbits, and that the defined scheme is generically reduced on the other orbit. This gives that the ideal of minors is radical of codimension 3. To check the statement about the singular locus, it is enough to check the Jacobian matrix of the ideal of minors on orbit representatives. □

As a corollary, the resolution of $\mathcal{O}_{Z'}$ is given as follows (see Example 12):

$$0 \to A'^* \otimes (\det A'^*)^5 \otimes R(-16) \to \mathbf{S}_{2,1}(A'^*) \otimes (\det A'^*)^3 \otimes R(-12) \to$$
$$S^2(A'^*) \otimes (\det A'^*)^2 \otimes R(-8) \to R \to \mathcal{O}_{Z'} \to 0.$$

We set M to be a certain twist of the cokernel of φ, namely, we have the presentation

$$A' \otimes (\det A'^*)^3 \otimes R(-8) \to A'^* \otimes (\det A'^*) \otimes R(-4) \to M \to 0.$$

For $P \in X' \setminus Z'$, φ_P has corank 1, so $M_P \cong \mathcal{O}_{X',P}$. For P in the open orbit of Z', then φ_P has corank 2, so M_P is minimally generated by 2 elements.

Lemma 2. *There is a G'-equivariant $\mathcal{O}_{X'}$-linear isomorphism $S^2M \cong I_{Z',X'}$.*

Proof. The proof is similar to the proof of Lemma 2. □

The other codimension 3 orbit closure Z_2' can be described as maps $\mathbf{spin}(B)^* \to A'$ whose kernel contains a nonzero pure spinor. This variety fails to be normal, and the minimal free resolution of its normalization is

$$0 \to \det A' \otimes S^5 A' \otimes R(-8) \to \det A' \otimes S^3 A' \otimes B \otimes R(-6)$$
$$\to \det A' \otimes S^2 A' \otimes \mathbf{spin}(B)^* \otimes R(-5) \to R \oplus \det A' \otimes \mathbf{spin}(B)$$
$$\otimes R(-3) \to \widetilde{\mathcal{O}}_{Z_2'} \to 0.$$

We define $\mathcal{X}$, $\mathcal{Z}$, $\mathcal{Z}_2$, and $\mathcal{M}$ to be the global versions of X', Z', Z_2' and M.

9.2 Geometric Data from a Section

Now choose a section $v \in U^{\text{gen}}$ and set $X = v(\mathbf{P}(A)) \cap \mathcal{X}$ and $Z = v(\mathbf{P}(A)) \cap \mathcal{Z}$.

Lemma 3. $v(\mathbf{P}(A)) \cap \mathcal{Z}_2 = \varnothing$.

Proof. Since $\widetilde{\mathcal{O}}_{Z_2'}$ is a perfect module, we can specialize its resolution by replacing A' by $\mathcal{Q}$, so that we get a complex

$$0 \to S^5\mathcal{Q}^*(-1) \to S^3\mathcal{Q}^* \otimes \underline{B}(-1) \to S^2\mathcal{Q}^* \otimes \mathbf{spin}(\underline{B})^*(-1) \to \mathcal{O}_{\mathbf{P}^3} \oplus \mathbf{spin}(\underline{B})(-1).$$

If $v(\mathbf{P}(A)) \cap \mathcal{Z}_2$ is nonempty, then it consists of finitely many points. But taking sections above, we get an exact complex. Hence the intersection is empty. □

From Lemma 2, we get a $\mathcal{O}_X$-linear map $\mu\colon S^2\mathcal{M} \to \mathcal{O}_X$, which gives an $\mathcal{O}_X$-algebra structure on $\mathcal{O}_X \oplus \mathcal{M}$. We set $\mathcal{O}_{\widetilde{X}} = \operatorname{Spec}_{\mathcal{O}_X}(\mathcal{O}_X \oplus \mathcal{M})$ and let $\pi\colon \widetilde{X} \to X$ be the structure map.

Define $U^{\text{sm}+}$ to be the subset of U^{sm} where $\mathcal{M}$ is minimally generated by 2 elements along Z.

Theorem 4. *If $v \in U^{\text{sm}+}$, then $\widetilde{X}$ is an Abelian surface and $\mathcal{L} = \pi^*\mathcal{O}_X(1)$ is an indecomposable $(2,2)$-polarization.*

Proof. The proof is similar to the proofs in Sect. 6.2. □

From (9.1), we also get a symmetric matrix

$$\varphi_v\colon \mathcal{Q} \to \mathcal{Q}^* \otimes (\det \mathcal{Q})^2. \tag{9.6}$$

9.3 Quadratic Complexes

Here is another approach:

- $G/P = \mathbf{Gr}(2, A)$.

- $\mathcal{U} = \mathcal{Q}_2 \otimes \mathbf{spin}(\underline{B})$.
- $U' = \mathbf{C}^2 \otimes \mathbf{spin}(10)$.
- $G' = (\mathbf{GL}_2(\mathbf{C}) \times \mathbf{Spin}_{10}(\mathbf{C}))/\{\,(x, x^{-1}) \mid x \in \mathbf{C}^*\,\}$.

There is a G'-invariant quartic hypersurface in U'. Given $v \in U^{\mathrm{sm}}$, the corresponding degeneracy locus is a smooth section Q of $\mathcal{O}(2)$.

Proposition 5. *Q is the quadratic complex associated to the symmetric matrix* (9.6) *constructed in the previous section.*

Proof. There is a $\mathbf{GL}(A)$-equivariant isomorphism

$$k\colon \mathrm{H}^0(\mathbf{Gr}(2, A); (\det \mathcal{Q}_2)^2) \to \mathrm{H}^0(\mathbf{P}(A); S^2(\mathcal{Q}^*) \otimes (\det \mathcal{Q})^2)$$

constructed in [44, Sect. 8] which associates to a quadratic complex (i.e., section in the domain above) to a symmetric matrix $\mathcal{Q} \to \mathcal{Q}^* \otimes (\det \mathcal{Q})^2$ whose determinant is the equation of a Kummer surface. Since both sides are isomorphic to the irreducible $\mathbf{GL}(A)$-representation $\mathbf{S}_{2,2}A$, this map is uniquely determined up to scalar. In the previous section and in the construction above, we have constructed two $\mathbf{GL}(A)$-equivariant maps $S^4(A \otimes \mathbf{spin}(B)) \to \mathbf{S}_{2,2}A$. Since $\mathbf{S}_{2,2}A$ appears in $S^4(A \otimes \mathbf{spin}(B))$ with multiplicity 1, this map is also unique up to scalar. Hence, the diagram

$$\begin{array}{ccc} S^4(A \otimes \mathbf{spin}(B)) & & \\ \downarrow & \searrow & \\ \mathrm{H}^0(\mathbf{Gr}(2, A); (\det \mathcal{Q}_2)^2) & \xrightarrow{k} & \mathrm{H}^0(\mathbf{P}(A); S^2(\mathcal{Q}^*) \otimes (\det \mathcal{Q})^2) \end{array}$$

commutes (up to possibly nonzero scalar ambiguity), which proves our claim. □

9.4 Doing Calculations

Again, we have written code in Macaulay2 for calculating the quadratic complex in affine trivializations in $\mathbf{P}^3$, but it is too messy to include. We instead explain the concepts behind the calculation.

We will explicitly calculate the symmetric matrix $\mathcal{Q} \to \mathcal{Q}^* \otimes (\det \mathcal{Q})^2$ starting with $v \in A \otimes \mathbf{spin}(B)$. First, note that this symmetric matrix is a section of $S^2\mathcal{Q}^* \otimes (\det \mathcal{Q})^2 \subset S^2A^* \otimes (\det \mathcal{Q})^2$, and the space of sections of the latter is $S^2A^* \otimes \mathbf{S}_{2,2,2}A = S^2A^* \otimes S^2A^* \otimes (\det A)^2$.

First, we have $\mathbf{S}_{2,2}A^* \subset S^4(A^* \otimes \mathbf{spin}(B)^*)$. This is a 20-dimensional space of quartics which is described explicitly in [29, Sect. 3] by the polynomials

$P(x, y, z, w)$. More specifically, a basis is given by those x, y, z, w such that $x \leq y$, $x < z$, $y < w$, $z \leq w$, and $1 \leq x, y, z, w \leq 4$, i.e., such that $T = \begin{array}{|c|c|}\hline x & y \\ \hline z & w \\ \hline\end{array}$ is a semistandard Young tableau. Abbreviate this polynomial by P_T. Let $Q_T \in \mathbf{S}_{2,2}A$ be the dual basis vectors.

Then given v, we can produce the element $p_v = \sum_T P_T(v)Q_T \in \mathbf{S}_{2,2}A = \mathbf{S}_{2,2}A^* \otimes (\det A)^2$. The inclusion $\iota\colon \mathbf{S}_{2,2}A^* \to S^2A^* \otimes S^2A^*$ is defined by

$$\begin{array}{|c|c|}\hline x & y \\ \hline z & w \\ \hline\end{array} \mapsto 2(xy \otimes zw + zw \otimes xy) - (xz \otimes yw + xw \otimes yz + yw \otimes xz + yz \otimes xw).$$

So $\iota(p_v) \in S^2A^* \otimes S^2A^* \otimes (\det A)^2$. Identifying the latter $S^2A^* \otimes (\det A)^2$ as sections of $(\det \mathcal{Q})^2$, this can be interpreted as a 4×4 symmetric matrix whose entries are quadrics on $\mathbf{P}^3$. Let z_1, z_2, z_3, z_4 be homogeneous coordinates. Delete the last row and column of this matrix. This gives the symmetric matrix $\mathcal{Q} \to \mathcal{Q}^* \otimes (\det \mathcal{Q})^2$ over the affine open set $z_4 = 1$. Take the ideals defined by the $i \times i$ minors of this matrix for $i = 3, 2$ and saturate with respect to z_4 to get homogeneous ideals for the degeneracy loci.

Acknowledgements We thank Manjul Bhargava, Igor Dolgachev, Wei Ho, Steven Kleiman, Bjorn Poonen, and Jack Thorne for helpful discussions and Damiano Testa for explaining the proof of Theorem 1. We also thank Igor Dolgachev for pointing out numerous references related to this work. Finally, we thank an anonymous referee for making some suggestions.

We also thank Federico Galetto and Witold Kraśkiewicz for assistance with some computer calculations. The software LiE [39] and Macaulay2 [27] were helpful in our work.

Steven Sam was supported by an NDSEG fellowship while this work was done. Jerzy Weyman was partially supported by NSF grant DMS-0901185.

References

1. Bala, P., Carter, R.: Classes of unipotent elements in simple algebraic groups. I, II. Proc. Camb. Phil. Soc. **79**, 401–425 (1976); **80**, 1–18 (1976)
2. Beauville, A.: Vector bundles on curves and generalized theta functions: recent results and open problems. Current topics in complex algebraic geometry (Berkeley, CA, 1992/93), pp. 17–33. Mathematical Sciences Research Institute Publications, vol. 28. Cambridge University Press, Cambridge (1995)
3. Beauville, A.: The Coble hypersurfaces. C. R. Math. Acad. Sci. Paris **337**(3), 189–194 (2003). http://arxiv.org/abs/math/0306097v1
4. Beauville, A.: Vector bundles on curves and theta functions. Moduli Spaces and Arithmetic Geometry, pp. 145–156. Advanced Studies in Pure Mathematics, vol. 45. Mathematical Society, Japan, Tokyo (2006). http://arxiv.org/abs/math/0502179v2
5. Berkesch, C., Erman, D., Kummini, M., Sam, S.V.: Tensor complexes: multilinear free resolutions constructed from higher tensors. J. Eur. Math. Soc. (JEMS). To appear. http://arxiv.org/abs/1101.4604v4
6. Bhargava, M.: Higher composition laws. I–IV. Ann. Math. (2) **159**(1) 217–250 (2004); **159**(2), 865–886 (2004); **159**(3), 1329–1360 (2004); **167**(1), 53–94 (2008)

7. Birkenhake, C., Lange, H.: Complex Abelian Varieties, 2nd edn. Grundlehren der Mathematischen Wissenschaften, vol. 302. Springer, Berlin (2004)
8. Bombieri, E., Mumford, D.: Enriques' classification of surfaces in char. p. II. Complex analysis and algebraic geometry, pp. 23–42. Iwanami Shoten, Tokyo (1977)
9. Brivio, S., Verra, A.: The theta divisor of $SU_C(2, 2d)^s$ is very ample if C is not hyperelliptic. Duke Math. J. **82**(3), 503–552 (1996)
10. Bruns, W., Vetter, U.: Determinantal rings. In: Lecture Notes in Mathematics, vol. 1327. Springer, Berlin (1988)
11. Buchsbaum, D.A., Eisenbud, D.: Algebra structures for finite free resolutions, and some structure theorems for ideals of codimension 3. Amer. J. Math. **99**(3), 447–485 (1977)
12. Burkhardt, H.: Untersuchungen aus dem Gebiete der hyperelliptischen Modulfunctionen. Math. Ann. **41**(3), 313–343 (1892)
13. Coble, A.: Point sets and allied Cremona groups. III. Trans. Amer. Math. Soc. **18**, 331–372 (1917)
14. Coble, A.: Algebraic geometry and theta functions. In: American Mathematical Society Colloquium Publications, vol. 10 (1929). American Mathematical Society, Providence (1982)
15. Cohen, A.M., Helminck, A.G.: Trilinear alternating forms on a vector space of dimension 7. Comm. Algebra **16**(1), 1–25 (1988)
16. Collingwood, D.H., McGovern, W.M.: Nilpotent orbits in semisimple Lie algebras. In: Van Nostrand Reinhold Mathematics Series. Van Nostrand Reinhold, New York (1993)
17. Desale, U.V., Ramanan, S.: Classification of vector bundles of rank 2 on hyperelliptic curves. Invent. Math. **38**(2), 161–185 (1976/77)
18. Dolgachev, I., Ortland, D.: Point sets in projective spaces and theta functions. Astérisque **165** (1988)
19. Dolgachev, I., Lehavi, D.: On isogenous principally polarized abelian surfaces. Curves and Abelian Varieties. Contemporary in Mathematics, vol. 465, pp. 51–69 (2008). http://arxiv.org/abs/0710.1298v3
20. Eisenbud, D.: Commutative algebra with a view toward algebraic geometry. In: Graduate Texts in Mathematics, vol. 150. Springer, New York (1995)
21. Èlašvili, A.G., Vinberg, È.B.: A classification of the three-vectors of 9-dimensional space. Trudy Sem. Vektor. Tenzor. Anal. **18**, 197–233 (1978). Translated in Sel. Math. Sov. **7**(1) (1988)
22. Fisher, T.: Invariants for the elliptic normal quintic, preprint (2002). Available at http://www.dpmms.cam.ac.uk/taf1000/papers/invenq.html
23. Fisher, T.: Genus one curves defined by Pfaffians. preprint (2006). Available at http://www.dpmms.cam.ac.uk/taf1000/papers/genus1pf.html
24. Galetto, F.: work in progress.
25. van Geemen, B., Izadi, E.: The tangent space to the moduli space of vector bundles on a curve and the singular locus of the theta divisor of the Jacobian. J. Algebraic Geom. **10**(1), 133–177 (2001). http://arxiv.org/abs/math/9805037v1
26. Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V.: Discriminants, Resultants, and Multidimensional Determinants. Reprint of the 1994 edition, Modern Birkhäuser Classics. Birkhäuser, Boston (2008)
27. Grayson, D.R., Stillman, M.E.: Macaulay 2, a software system for research in algebraic geometry. Version 1.4 available at http://www.math.uiuc.edu/Macaulay2/
28. Gross, M., Popescu, S.: The moduli space of $(1, 11)$-polarized abelian surfaces is unirational. Compositio Math. **126**(1), 1–23 (2001). http://arxiv.org/abs/math/9902017v1
29. Gyoja, A.: Construction of invariants. Tsukuba J. Math. **14**(2), 437–457 (1999)
30. Ho, W.: Orbit parametrizations of curves. Ph.D. thesis, Princeton University, New Jersey (2009)
31. Holweck, F.: Singularities of duals of Grassmannians. J. Algebra **337**, 369–384 (2011). http://arxiv.org/abs/1206.1038v1
32. Józefiak, T.: Ideals generated by minors of a symmetric matrix. Comment. Math. Helv. **53**(4), 595–607 (1978)

33. Kac, V.G.: On the question of describing the orbit space of linear algebraic groups. Uspehi Mat. Nauk **30**(6(186)), 173–174 (1975)
34. Kac, V.G.: Some remarks on nilpotent orbits. J. Algebra **64**(1), 190–213 (1980)
35. Kawamata, Y.: Characterization of abelian varieties. Compositio Math. **43**(2), 253–276 (1981)
36. Khaled, A.: Projective normality and equations of Kummer varieties. J. Reine Angew. Math. **465**, 197–217 (1995)
37. Kimura, T.: Remark on some combinatorial construction of relative invariants. Tsukuba J. Math. **5**(1), 101–115 (1981)
38. Kraśkiewicz, W., Weyman, J.: Geometry of orbit closures for the representations associated to gradings of Lie algebras of types E_6, F_4 and G_2. preprint. http://arxiv.org/abs/1201.1102v2
39. van Leeuwen, M.A.A., Cohen, A.M., Lisser, B.: LiE, a package for Lie group computations, version 2.2.2. http://www-math.univ-poitiers.fr/maavl/LiE/
40. Looijenga, E.: Cohomology of $\mathcal{M}_3$ and $\mathcal{M}_3^1$. Mapping class groups and moduli spaces of Riemann surfaces (Göttingen, 1991/Seattle, WA, 1991), pp. 205–228. Contemporary Mathematics, vol. 150. American Mathematical Society, Providence, RI (1993)
41. Maschke, H.: Aufstellung des vollen Formensytems einer quaternären Gruppe von 51840 linearen substitutionen. Math. Ann. **33**, 317–344 (1889)
42. Mumford, D.: On the equations defining abelian varieties. I. Invent. Math. **1**, 287–354 (1966)
43. Minh, N.Q.: Vector bundles, dualities and classical geometry on a curve of genus two. Internat. J. Math. **18**(5), 535–558 (2007). http://arxiv.org/abs/math/0702724v1
44. Narasimhan, M.S., Ramanan, S.: Moduli of vector bundles on a compact Riemann surface. Ann. Math. (2) **89**, 14–51 (1969)
45. Narasimhan, M.S., Ramanan, S.: 2θ-linear systems on abelian varieties. Vector bundles on algebraic varieties (Bombay, 1984), pp. 415–427. Tata Institute of Fundamental Research Studies in Mathematics, vol. 11. Tata Institute of Fundamental Research, Bombay (1987)
46. Ortega, A,: On the moduli space of rank 3 vector bundles on a genus 2 curve and the Coble cubic. J. Algebraic Geom. **14**(2), 327–356 (2005)
47. Ottaviani, G.: Spinor bundles on quadrics. Trans. Amer. Math. Soc. **307**(1), 301–316 (1988)
48. Oxbury, W., Pauly, C.: Heisenberg invariant quartics and $\mathscr{SU}_C(2)$ for a curve of genus four. Math. Proc. Cambridge Philos. Soc. **125**(2), 295–319 (1999). http://arxiv.org/abs/alg-geom/9703026v1
49. Pauly, C.: Self-duality of Coble's quartic hypersurface and applications. Michigan Math. J. **50**(3), 551–574 (2002). http://arxiv.org/abs/math/0109218v1
50. Poonen, B.: Computational aspects of curves of genus at least 2. Algorithmic number theory (Talence, 1996), pp. 283–306. Lecture Notes in Computer Science, vol. 1122, Springer, Berlin (1996)
51. Pragacz, P.: Characteristic free resolution of $n-2$-order Pfaffians of $n \times n$ antisymmetric matrix. J. Algebra **78**(2), 386–396 (1982)
52. Ramanan, S.: Orthogonal and spin bundles over hyperelliptic curves. Proc. Indian Acad. Sci. Math. Sci. **90**(2), 151–166 (1981)
53. Sam, S.V.: Schubert complexes and degeneracy loci. J. Algebra **337**, 103–125 (2011). http://arxiv.org/abs/1006.5514v2
54. Shephard, G.C., Todd, J.A.: Finite unitary reflection groups. Canadian J. Math. **6**, 274–304 (1954)
55. Spaltenstein, N.: Classes unipotentes et sous-groupes de Borel. In: Lecture Notes in Mathematics, vol. 946. Springer, Berlin (1982)
56. Vinberg, È.B.: The Weyl group of a graded Lie algebra. Math. USSR-Izv. **10**(3), 463–495 (1976)
57. Vinberg, È.B.: Classification of homogeneous nilpotent elements of a semisimple graded Lie algebra. Selecta Mathematica Sovietica **6**(1), 15–35 (1987)
58. Weyman, J.: Cohomology of vector bundles and syzygies. In: Cambridge Tracts in Mathematics, vol. 149. Cambridge University Press, Cambridge (2003)

F-Purity, Frobenius Splitting, and Tight Closure

Melvin Hochster

1 Introduction

I became interested in the study of F-purity and F-splitting in the interval 1967–1973 while I was at the University of Minnesota. My colleague Jack Eagon and I did work on the properties of determinantal rings (discussed briefly in Sect. 2, Example (13)): see [13]. This led to work, joint with Joel Roberts [14], on proving that rings of invariants of linearly reductive groups acting on regular rings are Cohen–Macaulay, and ultimately to a further study of F-purity [15]. At the same time I became interested in the local homological conjectures. Irving Kaplansky sent me an early preprint of the joint thesis of Peskine and Szpiro, [28], which was a great source of inspiration for me. I became interested in a number of splitting questions [8, 11], in the technique of reduction to characteristic p and in the existence of big Cohen–Macaulay modules and algebras [9, 10, 12, 18, 21]. This led in turn to the development of tight closure theory [16, 17, 19, 20, 22] in joint work with Craig Huneke that began in the fall of 1986. I will return to these themes below.

2 Pure and Split Extensions

Throughout, R is a commutative, associative ring with 1. A homomorphism of R-modules $\alpha : N \to M$ is called *pure* if $W \otimes_R N \to W \otimes_R M$ is an injective map for every R-module W. Since we may take $W = R$, we have, in particular, that $N \to M$ must be injective. If N is a direct summand of M, i.e., if there is a splitting $\beta : M \to N$ such that $\beta \circ \alpha = \mathrm{id}_N$, then $N \to M$ is pure. If M/N is

M. Hochster (✉)
Department of Mathematics, University of Michigan, Ann Arbor, MI 48109–1043, USA
e-mail: hochster@umich.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_14,

finitely presented, then $N \to M$ is pure if and only if N is a direct summand of M. Thus, if R is Noetherian, $N \to M$ is pure if and only if it is a direct limit of split extensions $N \to M_0$, since M is the directed union of its submodules finitely generated over N (this is true even when R is not Noetherian, although the maps $M_0 \to M$ need not be injective in that case). For more details on purity, see [21], pp. 48–50.

A ring extension $R \to S$ is called *split* (respectively, *pure*) if $R \to S$ is split (respectively, pure) as a map of R-modules. When this holds, if $G_\bullet$ is any complex of R-modules, the maps $G_\bullet \to S \otimes_R G_\bullet$ are split (respectively, pure), and so are the induced maps of homology or cohomology between the two complexes.

If R has prime characteristic $p > 0$, R is *F-split* (respectively, F-*pure*) if the Frobenius endomorphism $F_R = F : R \to R$ is split (respectively, pure). If either condition holds, R is reduced. When R is reduced, the maps $F : R \to R$, $F(R) \hookrightarrow R$, and $R \hookrightarrow R^{1/p}$ are isomorphic maps.

Examples

1. If $R \to S$ is faithfully flat, it is pure.
2. Splitting and purity for ring homomorphisms are both preserved by composition.
3. Any map of fields $K \to L$ is split over K since 1 is part of a free basis for L.
4. If $R \to S$ is split, say by a map $\alpha : S \to R$, then $R[x_1, \dots, x_n] \to S[x_1, \dots, x_n]$ is split, and this is also true for the R-algebra map that sends $x_i \mapsto x_i^{m_i}$, $1 \le i \le n$. One may send the term $cx_1^{a_1} \cdots x_n^{a_n}$ for $c \in S$ to 0 unless for all i, a_i is divisible by m_i and to $\alpha(c)x_1^{a_1/m_1} \cdots x_n^{a_n/m_n}$ when $m_i | a_i$ for all i.
5. In particular, a polynomial ring over a field K is F-split. If α splits $F_K : K \to K$, one may construct a splitting β as follows: for each monomial μ in the x_j, $\beta(c\nu) = 0$ unless $\nu = \mu^p$ is a p th power, and then $\beta(c\mu^p) = F_K(c)\mu$.
6. The quotient of the polynomial ring $K[x_1, \dots, x_n]$ by an ideal I generated by square-free monomials is F-split. The map β described above induces a splitting.
7. Similarly, let G be a finite group of permutations of the variables $x_1, \dots, x_n$. The ring of invariants R^G is spanned over K by sums of orbits of monomials. Again, the map β described above induces a splitting. The ring R^G is normal but not necessarily Cohen–Macaulay.
8. If R is a normal domain of equal characteristic 0, every module-finite extension S of R is split. One can kill a minimal prime of S disjoint from $R - \{0\}$, so that both are domains. Let $\mathcal{K} \hookrightarrow \mathcal{L}$ be the induced map of fraction fields and $\mathrm{tr}_{\mathcal{L}/\mathcal{K}} : \mathcal{L} \to \mathcal{K}$ be field trace. Let $d = [\mathcal{L} : \mathcal{K}]$. Then the restriction of $\frac{1}{d}\,\mathrm{tr}_{\mathcal{L}/\mathcal{K}}$ to S is an R-module retraction $S \to R$, i.e., yields a splitting.
9. If R is regular of equal characteristic, then every module-finite extension of R is split. See [Ho2]. This is conjectured to be true in mixed characteristic, where

it is easy in dimension ≤ 2, known in dimension 3 [5], and an open question in dimension ≥ 4.

10. In particular, in characteristic $p > 0$, every regular ring is F-pure. Let S be a ring of characteristic p. If $I \subseteq S$, and $q = p^e$ is a power of p, then $I^{[q]}$ denotes the ideal $(s^q : s \in I)S$ generated by all q th powers of elements of I (it suffices to use q th powers of generators of I). The following result of Richard Fedder is called *Fedder's criterion* for F-purity: *in characteristic* $p > 0$*, if* (S, m) *is regular local, or else a polynomial ring over a field and its homogeneous maximal ideal, and* I *is a proper ideal of* S *(*I *is assumed to be homogeneous in the polynomial ring case), then* S/I *is F-pure if and only if* $I^{[p]} : I \nsubseteq m^{[p]}$. Cf. Theorem 1.12 in [2].
11. We can apply Fedder's criterion to understand what happens for the cubical cone $R = K[X, Y, Z]/(X^3 + Y^3 + Z^3)$ over a field K of characteristic $p > 0$, where $p \neq 3$. Let $S = K[X, Y, Z]$, the polynomial ring. Fedder's criterion asserts that R is F-pure if and only if $(X^3 + Y^3 + Z^3)^p :_S (X^3 + Y^3 + Z^3) \nsubseteq (X^p, Y^p, Z^p)$, i.e., if and only if $(X^3+Y^3+Z^3)^{p-1} \notin (X^p, Y^p, Z^p)$. When we expand the left-hand side, a typical term is $\binom{p-1}{i\ j\ k} X^{3i} Y^{3j} Z^{3k}$ where $i + j + k = p - 1$. The multinomial coefficient $\binom{p-1}{i\ j\ k} = (p-1)!/i!j!k!$ does not vanish. If p has the form $3h + 2$ then at least one of i, j, k is $\geq h + 1$, and when we multiply by 3 we get an exponent that is $\geq p$. Hence, Fedder's criterion shows that R is *not* F-pure when $p \equiv 2 \bmod 3$. When $p = 3h + 1$ there is a nonzero term that is a multiple of $x^{3h}y^{3h}z^{3h}$, where $i = j = k = h$, and so Fedder's criterion shows that R is F-pure if and only if $p \equiv 1 \bmod 3$.

The following three examples all use the fact that the rings considered are weakly F-regular (and, for that matter, strongly F-regular): see Sect. 7 and the results of [7] (for (12)) and [20] (for (13) and (14)). Moreover, all of the rings in these three examples split from *every* module-finite extension. See Sect. 7 and [20].

12. A normal K-subalgebra R of a polynomial ring $K[X_1, \ldots, X_n]$ such that R is generated over the field K by monomials in the variables $X_1, \ldots, X_n$ is F-split in characteristic p.
13. If K has prime characteristic $p > 0$, X is an $r \times s$ matrix of indeterminates over K, $1 \leq t \leq \min\{r, s\}$, and $I_t(X)$ denotes the ideal generated by the $t \times t$ minors of X, which is prime (cf. [13]) then $K[x_{ij}]/I_t(X)$ is F-split.
14. With the same notation as in (13), if S denotes the subring of $K[x_{ij}]$ generated by the $r \times r$ minors of X, which is the homogeneous coordinate ring of a Grassmann variety, then S is F-split.

When $R \subseteq S$ is pure, it is always true that for every ideal $I \subseteq R$, $IS \cap R = I$: this follows because $R \subseteq S$ remains injective after one applies $R/I \otimes_R \_\_$. When R is Noetherian, the converse is true under mild conditions on R: see [11].

3 Review of Local Cohomology

Several of the applications of F-splitting techniques that we discuss in the sequel make use of basic results about local cohomology. In this section we give a brief review of what we need. The reader may consult [4] for a detailed treatment. Let R be a Noetherian ring, I an ideal, and M be an R-module, which need not be finitely generated. Then we may take as a definition that

$$H_I^i(M) = \varinjlim{}_t \mathrm{Ext}_R^i(R/I^t,\, M).$$

The ideals I^t may be replaced by any decreasing sequence of ideals cofinal with the powers of I, and these modules depend only on $\mathrm{Rad}(I)$. If $f_1, \ldots, f_h$ generate an ideal with the same radical as I, these modules are also the cohomology of the complex

$$0 \to M \to \bigoplus_i M_{f_i} \to \bigoplus_{i_1<i_2} M_{f_{i_1} f_{i_2}} \to \cdots \to \bigoplus_{i_1<\cdots<i_t} M_{f_{i_1}\cdots f_{i_t}} \to \cdots$$
$$\to M_{f_1\cdots f_h} \to 0, \tag{$\dagger$}$$

which is the same as the tensor product of M with the total tensor product of all of the complexes $0 \to R \to R_{f_i} \to 0$. If we omit M and start the numbering with $\oplus M_i$ we have the Čech complex on $U = \mathrm{Spec}(R) - V(I)$ of the sheaf $M^{\sim}|_U$ with respect to the affine open cover given by the sets $D(f_i)$. If $I \subseteq R$, S is a Noetherian R-algebra, and M is an S-module, we may view $M = {}_RM$ as a module over R by restriction of scalars. In this case $H_I^i({}_RM) \cong H_{IS}^i({}_SM)$.

If R and M are $\mathbb{Z}$-graded and I is homogeneous we may choose the $f_1, \ldots, f_n$ to be homogeneous. Every term in the complex ($\dagger$) is $\mathbb{Z}$-graded, and the maps preserve the grading. Thus, we get a $\mathbb{Z}$-grading on the local cohomology modules that turns out to be independent of which homogeneous generators $f_1, \ldots, f_n$ we choose.

Also note that if M is an R-module and we denote by ${}^eM$ the R-module obtained by restricting scalars via the map $F^e : R \to R$ (so that for $u \in {}^eM$, the value of $r \cdot u$ is $r^{p^e}u$), then $H_I^i({}^eM) \cong {}^eH_I^i(M)$. To see why, denote by S the target copy of R when one applies F^e. Think of M as an S-module. Then ${}^eM$ is obtained from M by restriction of scalars, and $H_I^i({}_RM) \cong {}_RH_{IS}^i(M) \cong {}_RH_(^i I^{[p^e]}(M) = {}_RH_I^i(M)$, since $I^{[p^e]}$ and I have the same radical.

When R has prime characteristic $p > 0$, there is a natural action of the Frobenius endomorphism F of R on $H_I^i(R)$. One way to think of this is to think of the map $F : R \to R$ as a map $R \to S$, where $S = R$. Then $F : R \to S$ induces a map $H_I^i(R) \to H_I^i(S) \cong H_{IS}^i(S) \cong H_{I^{[p]}}^i(R) \cong H_I^i(R)$ since $I^{[p]}$ has the same radical as I, and this map $F : H_I^i(R) \to H_I^i(R)$ is the action of F that we want. It has the property that $F(ru) = r^pF(u)$ for all $r \in R$ and $u \in H_I^i(R)$. When R is $\mathbb{Z}$-graded and I is homogeneous, the action of F on $H_I^i(R)$ is such that if u is homogeneous of degree $d \in \mathbb{Z}$, then $F(u)$ has degree pd. Hence, $F^e(u)$ has degree

$p^e d$. If $F : R \to R$ splits or is pure, the action of F is injective. This is critically important in the sequel.

Note that every element of every $H_I^i(M)$ is killed by some power of I.

When M is Noetherian and $IM \neq M$, the first nonvanishing $H_I^i(M)$ occurs when $i = d$, the depth of M on I.

Now suppose that M is finitely generated and m is a maximal ideal of R. Then the modules $H_m^i(M)$ are Artinian modules, and since every element is killed by a power of m, they may be viewed as modules over R_m or even over its completion. If (R, m) is local and $M \neq 0$ is finitely generated, then $H_m^i(R)$ is nonzero when i is the depth of M on m and when $i = \dim(M)$. It vanishes if i is smaller than the depth of M or larger than $\dim(M)$. Hence, $M \neq 0$ is Cohen–Macaulay over (R, m) if and only if it has a unique nonvanishing local cohomology module $H_m^i(M)$, which occurs when i is the depth of M on m or, equivalently, the dimension of M.

If (R, m) is regular local (or Gorenstein) of Krull dimension n, then $E = H_m^n(R)$ is an injective hull for the residue class field $K = R/m$. In this case, we have local duality: if M is finitely generated, for all i, $H_m^i(M) \cong \mathrm{Ext}_R^{n-i}(M, R)^\vee$, where $\_^\vee$ denotes $\mathrm{Hom}_R(\_, E)$.

An important consequence of local duality is the following:

Lemma 1. *Let (R, m) be a Gorenstein local ring of Krull dimension n and let $M \neq 0$ be a finitely generated R-module of pure dimension d. Suppose that M_P is Cohen–Macaulay for every prime P of R in its support different from m. Then $H_m^i(M)$ has finite length for every $i < d = \dim(M)$. In particular, this holds when $M \neq 0$ is finitely generated and torsion-free over R/Q for some prime Q of R if M is Cohen–Macaulay when localized at any proper prime in its support.*

This follows from the fact that this local cohomology module $H_m^i(M)$ is the Matlis dual of $N = \mathrm{Ext}_R^{n-i}(M, R)$, and so it suffices to show that N has finite length for $i < d$. Since N is finitely generated, we need only show that N is not supported at any prime $P \neq m$ in the support of M. But $N_P \cong \mathrm{Ext}_{R_P}^{n-i}(M_P, R_P)$ which, by Matlis duality over R_P, will vanish if and only if $H_{PR_P}^{h-(n-i)}(M_P) = 0$, where $h = \dim(R_P)$ the height of P. Since M_P is a Cohen–Macaulay module over R_P of pure dimension $h - (n - d)$ (the height of its annihilator does not change when we localize at P, and that height is $n - d$), it has only one nonvanishing local cohomology module, namely $H_{PR_P}^{h-(n-d))}(M_P)$. Since $i < d$, $H_{PR_P}^{h-(n-i)}(M_P) = 0$, as required. □

We also note the following fact, which connects local cohomology with cohomology of sheaves on projective spaces.

Proposition 2. *Let K be a field and let R be a finitely generated $\mathbb{N}$-graded K-algebra of Krull dimension n such that $[R]_0 = K$ and R is generated by the vector space $[R]_1$ of forms of degree one. Let M be a finitely generated $\mathbb{Z}$-graded R-module, and let $\mathcal{M}$ denote the corresponding sheaf on $X = \mathrm{Proj}(R)$, so that if $f \in m$, the homogeneous maximal ideal of R, then $\Gamma(X_f, \mathcal{M}) = [M_f]_0$.*

Then for $i \geq 1$, $H^i(X, \mathcal{M}) \cong [H_m^{i+1}(M)]_0$. *More generally, for every* $t \in \mathbb{Z}$, $H^i(X, \mathcal{M}(t)) \cong [H_m^{i+1}(M)]_t$.

If, moreover, R *is a domain of positive dimension and* M *is a nonzero torsion-free* R*-module then the following conditions are equivalent:*

(1) M *is Cohen–Macaulay.*
(2) $H_m^i(M) = 0$, $0 \leq i < \dim(R)$.
(3) If $n \geq 2$, M *has depth at least two on* m *and for all* $t \in \mathbb{Z}$, $H^i(X, \mathcal{M}(t)) = 0$, $1 \leq i < \dim(X)$.

Proof. Let $f_1, \ldots, f_n$ be a homogeneous system of parameters for the $\mathbb{N}$-graded ring R, so that $I = (f_1, \ldots, f_n)R$ is primary to the homogeneous maximal ideal m. Then $H_m^\bullet(M) = H_I^\bullet(M)$ is the cohomology of the complex (†) displayed in the first paragraph of this section. If we drop the first term of this complex, shift the numbering by one, and take the degree 0 part, we get the Čech complex for computing the cohomology of the sheaf $\mathcal{M}$. This yields that $H^i(X, \mathcal{M}) \cong [H_I^{i+1}(M)]_0$ for $i \geq 1$. The final statement follows if one applies this fact to $M(t)$ (M with the grading shifted so that $[M(t)]_s = [M]_{s+t}$: the sheaf on X corresponding to $M(t)$ is $\mathcal{M}(t)$).

In the graded case, to check that M is Cohen–Macaulay of maximum dimension, it suffices to check that $\mathrm{depth}_m M = \dim(R)$ and the depth is the same as the smallest integer d such that $H_m^d(M) \neq 0$. Since $d \leq n$ in any case, we have that (2) is the equivalent to the Cohen–Macaulay property, while (3) is equivalent to (2) by the first part of the proposition. □

4 Proving that Rings Are Cohen–Macaulay

One of the motivations for studying F-pure and F-split rings is the following fact:

Theorem 1. *Let* R *be a domain that is finitely generated over a field* K *of characteristic* $p > 0$ *and that is generated by its forms of degree 1. Suppose that* R *has depth at least two on* m *(which holds, e.g., if* R *is normal), is Cohen–Macaulay when localized at a prime other than maximal ideal, and is F-pure. Let* $\mathrm{Proj}(R) = (X, \mathcal{O}_X)$. *Then* R *is Cohen–Macaulay if and only if* $H^i(X, \mathcal{O}_X) = 0$, $1 \leq i < \dim(X)$.

Proof. We may assume that $R \neq K$, since K is Cohen–Macaulay, and so $\dim(R) \geq 1$. We know that the depth is at least two, and so it suffices to show that $H_m^{i+1}(R) = 0$ for $1 \leq i < \dim(R) - 2$. Since R is Cohen–Macaulay when localized at any prime P except m, we know that $H_m^{i+1}(R)$ has finite length for all i in the specified range. Hence $[H_m^{i+1}(R)]_t = 0$ whenever $|t| \gg 0$. But the Frobenius endomorphism F and, hence, all of its iterates F^e act injectively on the local cohomology modules since R is F-pure. These modules are $\mathbb{Z}$-graded and $F^e : [H^{i+1}(R)]_t \to [H^{i+1}(R)]_{p^e t}$. The latter vanishes for $e \gg 0$ if $t \neq 0$, and this shows that $[H^{i+1}(R)]_t = 0$ for $1 \leq$

$i \le \dim(R) - 2$ if $t \neq 0$. But $[H^{i+1}(R)]_0 = H^i(X, \mathcal{O}_X) = 0$ for i in the specified range by hypothesis, and so $[H^{i+1}(R)]_t = 0$ for all t for $1 \le i \le \dim(R) - 2$, as required. □

The original proof of the following result, first established in [14], utilized a variant of this result. First note that when we say that an algebraic group G *acts rationally* on a K-vector space, we mean that the vector space is a directed union of finite-dimensional G-stable subspaces V such that the group action on V is given by a regular map $G \times V \to V$. For example, if G acts rationally on the vector space of one-forms in a polynomial ring S over K, the action extends uniquely to a rational action of G on S.

Theorem 2 (Hochster–Roberts). *Let G be a linearly reductive linear algebraic group over a field K acting rationally, by K-algebra automorphisms, on a Noetherian K-algebra S. Then $R = S^G$, the ring of invariants, is Cohen–Macaulay.*

This is very largely a theorem about equal characteristic 0, because there are very few linearly reductive groups in positive characteristic: there are finite groups of order invertible modulo p, products of $\mathrm{GL}(1, K)$ (called *algebraic tori*), and groups obtained from these by extension. In equal characteristic 0, one has the classical groups (cf. [33]) which have many interesting representations with rings of invariants that are of considerable importance in algebraic geometry. In addition to finite groups and algebraic tori, the semisimple groups (which include the special linear, special orthogonal, and symplectic groups) are linearly reductive.

The proof of the theorem uses the fact that if G is linearly reductive and acts on S as in the theorem, there is a canonical R-module retraction $S \to S^G = R$, called the *Reynolds operator*. But there are some rather subtle points in the argument. Although $R \to S$ is a split extension, this is not true when one passes to characteristic p—it is often false for every p.

For example, let X be a 2×3 matrix of indeterminates and let $A \in \mathrm{SL}(2, \mathbb{Q})$ act on the polynomial ring $\mathbb{Q}[X]$ in these indeterminates by mapping the entries of X to the entries of $A^{-1}X$. Let $\Delta_1, \Delta_2, \Delta_3$ denote the 2×2 minors of X. Then $S^G = \mathbb{Q}[\Delta_1, \Delta_2, \Delta_3]$ is the ring of invariants, and there is an R-module retraction $S \to R$. However, in characteristic $p > 0$, $(\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$ does not split over $(\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2, \Delta_3]$ for any prime $p > 0$. This means that if one restricts the canonical splitting $\mathbb{Q}[X] \to \mathbb{Q}[\Delta_1, \Delta_2, \Delta_3]$ to $\mathbb{Z}[X]$, it takes on values in such a way that every prime $p \in \mathbb{Z}$ is needed in the denominator in at least one of its values!

In fact, if

$$(\mathbb{Z}/pZ)[\Delta_1, \Delta_2, \Delta_3] \to (\mathbb{Z}/p\mathbb{Z})[X]$$

were split then, if we let $I = (\Delta_1, \Delta_2, \Delta_3)$, the map of local cohomology

$$H^3_I\big((\mathbb{Z}/p\mathbb{Z})[\Delta_1, \Delta_2, \Delta_3]\big) \to H^3_I\big((\mathbb{Z}/p\mathbb{Z})[X]\big)$$

would be injective. Since the former is not 0, this would imply that $H^3_I\big(\mathbb{Z}/p\mathbb{Z})]X]\big)$ $\neq 0$. But this local cohomology module is 0 by a result of Peskine and Szpiro [28] that we discuss in the next section.

In the original proof of the Hochster–Roberts theorem one uses induction on the dimension to reduce to the case of a supposed counterexample of minimum dimension. One can then pass to associated graded rings to get a counterexample in which G acts linearly on a polynomial ring S over a field. From the minimality, one can assume that R is Cohen–Macaulay except when localized at its homogeneous maximal ideal. One then makes use of reduction to characteristic p. Although one cannot preserve the splitting of $R \to S$ as one passes to characteristic $p > 0$, one can preserve finitely many consequences of the fact that one has a splitting. This is enough to imitate the argument in the characteristic p result stated at the beginning of this section, and thus one is able to show that for $t \neq 0$, the graded components of $[H^i_m(R)]_t$ for $i < \dim(R)$ vanish. One is left with the problem of studying the component in degree 0. Since it is easy to see that R is normal, what one needs to show is that with $\mathrm{Proj}(R) = (X, \mathcal{O}_X)$, one has that $H^i(X, \mathcal{O}_X) = 0$, $1 \leq i < \dim(X)$. Again, one uses reduction to characteristic p, but for this argument, one needs the fact that the Frobenius endomorphism is flat in a regular ring of characteristic $p > 0$. In retrospect, the argument given can be seen to be a precursor of tight closure theory, which is discussed in Sect. 7.

Kempf [23] gives a different treatment of the theorem. Boutot [1] showed that if R, S are affine algebras over a field of characteristic such that S rational singularities and $R \to S$ is split, then R has rational singularities. There is a brief treatment of rational singularities in [24], pp. 49–52. Boutot's argument uses a characterization of rational singularities in [24] that depends on the Grauert–Riemenschneider vanishing theorem [3].

Tight closure theory has been used to give substantial generalizations of the Hochster–Roberts theorem: see Sect. 7.

Here is another early application of Frobenius splitting ideas to proving that certain rings are Cohen–Macaulay. Let Σ be a finite simplicial complex with vertices $x_1, \dots, x_n$. This simply means that Σ is a set of subsets of $x_1, \dots, x_n$ closed under passing to subsets and containing each of the sets $\{x_i\}$. The elements σ of Σ are called *simplices*. The *dimension* of the simplex σ is one less than the number of vertices in σ, and the *dimension* of Σ is the largest dimension of any of its simplices. Let $e_1, \dots, e_n$ be the standard basis for $\mathbb{R}^n$. We can establish a bijection of the x_i with the e_i by letting x_i correspond to e_i, $1 \leq i \leq n$, and, hence, between the simplices of Σ and a set of subsets of $\{e_1, \dots, e_n\}$. The *geometric realization* $|\Sigma|$ of Σ is the topological subspace of $\mathbb{R}^n$ which is the union of the convex hulls of the subsets of $e_1, \dots, e_n$ corresponding to simplices in Σ. Note that Σ is a compact topological space. The *link* of $\sigma \in \Sigma$ is the simplicial complex consisting of all $\tau \in \Sigma$ disjoint from σ such that $\tau \cup \sigma \in \Sigma$. If $\sigma = \{x_i\}$, the union of all the simplices of Σ that contain x_i is a cone with vertex x_i over the link of $\{x_i\}$.

To a simplicial complex Σ one can associate the *Stanley–Reisner ring* or *face ring* over the field K, $K[x_1, \dots, x_n]/I_\Sigma$, where I_Σ is generated by all square-free monomials in the indeterminates $x_1, \ldots, x_n$ such that the set of variables that occurs is not a simplex in Σ. The following characterization of when $K[x_1, \ldots, x_n]/I_\Sigma$ is Cohen–Macaulay is given in [29]. Note that the *reduced simplicial cohomology* of Σ with coefficients in K agrees with the simplicial cohomology over K in positive degree (the simplicial cohomology is the same as, say, the singular cohomology of $|\Sigma|$ with coefficients in K). In degree 0, if $H^0(\Sigma;\, K)$ has dimension $r > 0$, the reduced simplicial cohomology $\widetilde{H}^0(\Sigma;\, K)$ has dimension $r - 1$, so that it vanishes when $|\Sigma|$ is connected.

Theorem 3 (G. Reisner). *Let K be a field, and let Σ be a finite simplicial complex. Then the Stanley–Reisner ring $K[x_1, \ldots, x_n]/I_\Sigma$, where, as above, $K[x_1, \ldots, x_n]$ is a polynomial ring in variables corresponding to the vertices of Σ, is Cohen–Macaulay if and only if the following two conditions hold:*

(1) The reduced simplicial cohomology $\widetilde{H}^i(\Sigma; K)$ of Σ with coefficients in K vanishes for $i < \dim(\Sigma)$.
(2) The reduced simplicial cohomology $\widetilde{H}^i(\Lambda; K)$ of every link Λ of every simplex of Σ vanishes for $i < \dim(\Lambda)$.

This characterization, combined with results of Macaulay on the Hilbert functions of graded Cohen–Macaulay rings, was used by Richard Stanley [32] to prove the upper bound conjecture for simplicial polytopes. Munkres [27] showed that Reisner's conditions actually constitute a purely topological condition on Σ.

Sketch of the Proof. The case where the field has characteristic 0 can be proved by reduction to characteristic p. The original proof in characteristic $p > 0$ used the fact that Stanley–Reisner rings are F-split. The condition on the links implies, by induction, that the Cohen–Macaulay property holds except possibly at the homogeneous maximal ideal. One can conclude that the local cohomology is of finite length except in the top dimension. There is a $\mathbb{Z}^n$-grading (or grading by monomials) on R/I_Σ, on m, and hence on the local cohomology modules $H^i_m(R/I_\Sigma)$. The action of Frobenius multiplies multi-degrees by p and is injective because of the F-split condition. It follows that any multi-graded component in which any of the n coordinates of the degree is nonzero must vanish. Therefore one can reduce the problem to the vanishing of the local cohomology modules in degree $(0, 0, \ldots, 0)$, and so one can use the degree $(0, 0, \ldots, 0)$ part of the complex displayed in (†) in the first paragraph of Sect. 3, with $M = R/I_\Sigma$ and the f_j are taken to be the images of the x_j, to calculate it. This turns out to be the same complex used to calculate the reduced simplicial cohomology of Σ. □

5 Some Results of Peskine and Szpiro

The joint work of Peskine and Szpiro in [28] had an enormous influence: they used techniques involving the application of the Frobenius endomorphism to prove several local homological conjectures due to M. Auslander and H. Bass in characteristic p, introducing conjectures of their own in the process. They also obtained many equal characteristic cases by reduction to characteristic $p > 0$. See also [10], where the existence of big Cohen–Macaulay modules is proved by reduction to characteristic $p > 0$ and then applied to settle the same conjectures in equal characteristic. Many of their results depend on the fact that Frobenius is flat relative to modules of finite projective dimension. (See also [6].) This means that if we write S for R viewed as an R-algebra via a power F^e (under composition) of the Frobenius endomorphism and M is an R-module of finite projective dimension, then $\mathrm{Tor}_i^R(M, S) = 0$ for all $i \geq 1$. This may be viewed as a generalization of the fact that S is R-flat when R is regular. In fact, the flatness of $F : R \to R$ is equivalent to the regularity of R: cf. [25].

Because of its remarkably simple proof via Frobenius techniques we want to discuss one further result of [28], which was applied in Sect. 4 to show that certain rings of invariants are not direct summands of (nor pure in) polynomial rings in characteristic p.

Theorem 4 (C. Peskine and L. Szpiro). *Let R be a regular domain of prime characteristic $p > 0$ and I an ideal of R such that R/I is Cohen–Macaulay. Let h denote the height of I. Then $H_I^j(R) = 0$ for $j > h$.*

Proof. The fact that $F^e : R \to R = S$ is flat implies that $S \otimes_R R/I = R/I^{[p^e]}$ is Cohen–Macaulay for all e. But then there is a unique nonvanishing $\mathrm{Ext}_R^j(R/I^{[p^e]}, R)$ for all e, occurring when $j = h$. Since the local cohomology may be obtained as the direct limit of these, it follows that $H_I^j(R) = 0$ except when $j = h$. □

6 Small Cohen–Macaulay Modules

It is known (cf. [10, 18, 21]) that over every equal characteristic local ring (R, m), there is a module (even an algebra) B such that $mB \neq B$ and every system of parameters for R is a regular sequence on B. B is called a *big Cohen–Macaulay* module (respectively, algebra) for R. This was first proved by reduction to characteristic p in [10], and all known proofs require reduction to characteristic p. This is an open question in mixed characteristic in dimension 4 and higher. (The dimension 3 case is settled using the results of [5] in [12].)

It has long been an open question whether, under mild conditions on a local ring (R, m) (e.g., if R is excellent), there exists a Cohen–Macaulay module that is finitely generated (hence, the use of the word "small") whose dimension is the same as $\dim(R)$. In this section we give an application of Frobenius splitting techniques to proving the existence of small Cohen–Macaulay modules in characteristic $p > 0$ in certain instances. The argument was first given by R. Hartshorne and later rediscovered independently first by C. Peskine and L. Szpiro and later by the author. (The argument is given, e.g., in [9].) For simplicity, we have not attempted to state the most general form of the result here. But the question remains open even for $\mathbb{N}$-graded affine algebras over a field of characteristic 0 in dimension 3, and it is an open question for local rings of affine algebras over a field of characteristic $p > 0$ in dimension 3.

Theorem 5 (Hartshorne). *Let R be a finitely generated $\mathbb{N}$-graded domain over a perfect field K of characteristic $p > 0$ with $[R]_0 = K$. Let M be a finitely generated $\mathbb{N}$-graded R-module that is torsion-free over R, and suppose that M_P is Cohen–Macaulay over R_P for every prime ideal P of R except possibly the homogeneous maximal ideal m. Then R has a graded finitely generated module N that has depth equal to the dimension of R.*

Sketch of the Proof. We may assume that R has positive dimension and is graded so that $[R]_i \neq 0$ for all $i \gg 0$, and then the same will be true for M. The fact that M_P is Cohen–Macaulay for $P \neq m$ implies, that the local cohomology modules $H^i_m(M)$ have finite length for $i < d = \dim(R)$ by the Lemma in Sect. 3 (R is a homomorphic image of a Gorenstein ring). Let $F^e : R \to R$, and consider M as module over the right-hand copy of R. Restriction of scalars produces a module ${}^eM$ over the left-hand copy of R as in the fourth paragraph of Sect. 3. The grading on M enables us to split ${}^eM$ into the direct sum of p^e nonzero R-modules N_j, $0 \leq j < p^e$, where

$$N_j = \bigoplus_{i \equiv j \bmod p^e} [M]_i.$$

Let B denote the sum of the lengths of the $H^i_m(M)$ for $i < \dim(R)$. We claim that for all e so large that $p^e > B$, at least one of the modules N_j is Cohen–Macaulay. For consider the sum of the lengths L_j of the local cohomology modules $H^i_m(N_j)$ for $i < \dim(R)$. All we need to show is that at least one L_j is 0. But the total of the L_j is the same as the sum of the lengths of the $H^i_m({}^eM)$ for $i < \dim(R)$, and, as noted in Sect. 3, $H^i_m({}^eM) = {}^eH^i_m(M)$, and, because K is perfect, this has the same length as $H^i_m(M)$. But then $\sum_{j=0}^{p^e-1} L_j = B$. Since $p^e > B$, at least one of the L_j must be zero. □

7 Tight Closure and Splinters

We give here the very brief introduction to tight closure theory, which has many interconnections with questions about F-splitting and F-purity.

Throughout this section, R is an excellent ring. In characteristic $p > 0$, u is defined to be in the *tight closure* of an ideal I of R if there is an element $c \in R$ not in any minimal prime such that $cu^{p^e} \in I^{[p^e]}$ for all $e \gg 0$. This holds if and only if it holds modulo every minimal prime of R. We focus primarily on the case where R is a domain. In that case, c is simply required to be nonzero. For the characteristic $p > 0$ theory see [16, 17, 19, 20, 22, 31].

Tight closure may also be defined in finitely generated $\mathbb{Q}$-algebras as follows: if R is such an algebra, $u \in R$, and $I \subseteq R$, we say that u is in the *tight closure* of J in R if there is a domain $R_0 \subseteq R$ finitely generated over $\mathbb{Z}$ such that $u \in R_0$, and an ideal $I \subseteq J \cap R_0$ such that the image of u is in the tight closure of IR_0/pR_0 in R_0/pR_0 for all but finitely many prime integers p. One can then extend the theory to all excellent Noetherian rings containing $\mathbb{Q}$ as follows: u is in the tight closure of J in R if there exists a finitely generated $\mathbb{Q}$-algebra A and ideal $I \subseteq A$, an element $t \in A$ in the tight closure of I, and a homomorphism $A \to R$ such that $t \mapsto u$ and I maps into J. This notion is called *equational tight closure* in [22].

There is also a tight closure theory for submodules of modules.

A ring such that every ideal is tightly closed is called *weakly F-regular*. If all localizations of R are weakly F-regular, R is called *F-regular*. It is not known whether weakly F-regular implies F-regular for excellent rings.

In the equicharacteristic case, one has the following for excellent rings:

1. Every ideal of a regular ring is tightly closed.
2. If $x_1, \dots, x_k$ is part of a system of parameters in a reduced equidimensional local ring and $rx_k \in (x_1, \dots, x_{k-1})$, then r is in the tight closure of $(x_1, \dots, x_{k-1})$.
3. If R is weakly F-regular, then R is Cohen–Macaulay.
4. If $R \to S$ is pure and S is weakly F-regular, then so is R.
5. If $R \subseteq S$ is an integral extension, $IS \cap R$ is contained in the tight closure of I.
6. If R is weakly F-regular, then R is normal.

These results imply that in the equicharacteristic case, every ring R pure in a regular ring is Cohen–Macaulay. This is a generalization of the Hochster–Roberts theorem discussed in Sect. 4. This is an open question in the mixed characteristic case.

We refer to a Noetherian ring that is a direct summand of every module-finite extension ring as a *splinter*. The results of [20] (see Corollary 5.23 and Theorem 5.25 on p. 630) coupled with the results of [11] imply that every weakly F-regular ring is a splinter and, hence, F-pure. In the Gorenstein case, in characteristic $p > 0$, the converse is true: splinters are weakly F-regular. This is also true in the $\mathbb{Q}$-Gorenstein case (cf. [30]). In general, it is known that in positive characteristic a splinter must be Cohen–Macaulay, but it is an open question whether splinters are weakly F-regular in general in the Cohen–Macaulay case.

A different point of view connecting splitting questions with tight closure in the characteristic $p > 0$ case is the following. Let S be a module-finite extension of a reduced ring R of characteristic $p > 0$. From the point of view of Yoneda Ext, the exact sequence

$$0 \to R \to S \to S/R \to 0$$

of finitely generated R-modules represents an element ϵ of $E = \mathrm{Ext}^1_R(S/R, R)$. If we compute E using a finite projective resolution $P_\bullet$ of S/R, then E may be viewed as a submodule of $Q = \mathrm{Hom}_R(P_1, R)/\mathrm{Im}\big(\mathrm{Hom}_R(P_0, R)\big)$. Theorem 5.17 of [HH5] yields:

Theorem 6. *With notation and hypotheses as in the paragraph just above, the element* $\epsilon \in \mathrm{Ext}^1_R(S/R, R)$ *represented by* $0 \to R \to S \to S/R \to 0$ *is in the tight closure of 0 when regarded as an element of* Q*. Hence, if* R *is weakly F-regular,* ϵ *is 0, and* $R \hookrightarrow S$ *splits.*

This yields a proof from a different perspective of the fact that weakly F-regular rings are splinters.

Tight closure is connected with Frobenius splitting in another way. Let R be Noetherian of characteristic $p > 0$. R is called *F-finite* if $F : R \to R$ is module-finite. F-finite rings are excellent (cf. [26]). An F-finite domain of characteristic p is called *strongly F-regular* if for every $c \neq 0$, the map $R \to R^{1/p^e}$ such that $1 \mapsto c^{1/p^e}$ splits for all sufficiently large e. See [16, 29]. It is easy to show that strongly F-regular rings are F-regular. In the F-finite Gorenstein case, weakly F-regular is equivalent to strongly F-regular. The converse is an open question in the general case.

The rings discussed in Examples (12), (13), and (14) of Sect. 2 are known to be strongly F-regular (see [7] for Example (12) and [20], Theorem 7.14, p. 651, for Examples (13) and (14): note that strong F-regularity follows from weak F-regularity in these cases because the rings are either Gorenstein or algebra retracts of F-regular Gorenstein rings) and so split from every module-finite extension. One can then deduce immediately that all of these rings are F-split.

Acknowledgments The author was partially supported by the National Science Foundation (DMS-9970702, DMS-0400633). Part of this manuscript is based on a talk presented at a conference entitled "Frobenius splitting in algebraic geometry, commutative algebra, and representation theory" held at the University of Michigan, Ann Arbor, May 17–22, 2010.

References

1. Boutot, J.-F.: Singularités rationelles et quotients par les groupes réductifs. Invent. Math. **88**, 65–68 (1987)
2. Fedder, R.: F-purity and rational singularity. Trans. Amer. Math. Soc. **278**, 461–480 (1983)
3. Grauert, H., Riemenschneider, O.: Verschwindungssätze für analytische Kohomologiegruppen auf komplexen Räumen. Inventiones Math. **11**, 263–292 (1970)
4. Grothendieck, A. (notes by Hartshorne, R.): Local cohomology. In: Lecture Notes in Mathematics, vol. 41. Springer, Heidelberg (1967)
5. Heitmann, R.: The direct summand conjecture in dimension three. Ann. Math. (2) **156**, 695–712 (2002)
6. Herzog, J.: Ringe der Characteristik p und Frobenius-funktoren. Math. Z. **140**, 67–78 (1974)
7. Hochster, M.: Rings of invariants of tori, Cohen–Macaulay rings generated by monomials, and polytopes. Ann. Math. **96**, 318–337 (1972)
8. Hochster, M.: Contracted ideals from integral extensions of regular rings. Nagoya Math. J. **51**, 25–43 (1973)
9. Hochster, M.: Big Cohen–Macaulay modules and algebras and embeddability in rings of Witt vectors. In: Proceedings of the Queen's University Commutative Algebra Conference (Kingston, Ontario, Canada, 1975). Queen's Papers in Pure and Applied Mathematics, vol. 42, pp. 106–195 (1975)

10. Hochster, M.: Topics in the homological theory of modules over commutative rings. In: Proceedings of Nebraska Regional C.B.M.S. Conference (Lincoln, Nebraska, 1974). American Mathematical Society, Providence, RI (1975)
11. Hochster, M.: Cyclic purity versus purity in excellent Noetherian rings. Trans. Amer. Math. Soc. **231**, 463–488 (1977)
12. Hochster, M.: Big Cohen–Macaulay algebras in dimension three via Heitmann's theorem. J. Algebra **254**, 395–408 (2002)
13. Hochster, M., Eagon, J.A.: Cohen–Macaulay rings, invariant theory, and the generic perfection of determinantal loci. J. Amer. Math. **93**, 1020–1058 (1971)
14. Hochster, M., Roberts, J.L.: Rings of invariants of reductive groups acting on regular rings are Cohen–Macaulay. Adv. Math. **13**, 115–175 (1974)
15. Hochster, M., Roberts, J.L.: The purity of the Frobenius and local cohomology. Adv. Math. **21**, 117–172 (1976)
16. Hochster, M., Huneke, C.: Tight closure and strong F-regularity. Mémoires de la Société Mathématique de France **38**, 119–133 (1989)
17. Hochster, M., Huneke, C.: Tight closure, invariant theory, and the Briançon–Skoda theorem. J. Amer. Math. Soc. **3**, 31–116 (1990)
18. Hochster, M., Huneke, C.: Infinite integral extensions and big Cohen–Macaulay algebras. Ann. Math. **35**, 53–89 (1992)
19. Hochster, M., Huneke, C.: F-regularity, test elements, and smooth base change. Trans. Amer. Math. Soc. **346**, 1–62 (1994)
20. Hochster, M., Huneke, C.: Tight closure of parameter ideals and splitting in module-finite extensions. J. Algebraic Geometry **3**, 599–672 (1994)
21. Hochster, M., Huneke, C.: Applications of the existence of big Cohen–Macaulay algebras. Adv. Math. **113**, 45–117 (1995)
22. Hochster, M., Huneke, C.: Tight closure in equal characteristic zero. Preprint. Available at http://www.math.lsa.umich.edu/~hochster/msr.html.
23. Kempf, G.: The Hochster–Roberts theorem of invariant theory. Michigan Math. J. **26**, 19–32 (1979)
24. Kempf, G., Knudsen, F., Mumford, D., Saint-Donat, B.: Toroidal Embeddings I. In: Springer-Verlag Lecture Notes in Mathematics, vol. 339. Springer, Berlin (1973)
25. Kunz, E.: Characterizations of regular local rings of characteristic p. Amer. J. Math. **91**, 772–784 (1969)
26. Kunz, E.: On Noetherian rings of characteristic p. Amer. J. Math. **98**, 999–1013 (1976)
27. Munkres, J.R.: Topological results in combinatorics. Michigan Math. J. **31**, 113–128 (1984)
28. Peskine, C., Szpiro, L.: Dimension projective finie et cohomologie locale. Publ. Math. I.H.E.S. (Paris) **42**, 323–395 (1973)
29. Reisner, G.A.: Cohen–Macaulay quotients of polynomial rings. Adv. Math. **21**, 30–49 (1976)
30. Singh, A.K.: $\mathbb{Q}$-Gorenstein splinter rings of characteristic p are F-regular. Math. Proc. Cambridge Philos. Soc. **27**, 201–205 (1999)
31. Smith, K.E.: Tight closure of parameter ideals. Invent. Math. **115**, 41–60 (1994)
32. Stanley, R.P.: The upper bound conjecture and Cohen–Macaulay rings. Stud. Appl. Math. **54**, 135–142 (1975)
33. Weyl, H.: The Classical Groups: Their Invariants and Representations. Princeton University Press, Princeton (1946)

Hilbert–Kunz Multiplicity and the F-Signature

Craig Huneke

1 Introduction

Throughout this chapter, $(R, \mathfrak{m}, k)$ will denote a Noetherian local ring of prime characteristic p with maximal ideal $\mathfrak{m}$ and residue field k. We let e be a varying non-negative integer, and let $q = p^e$. By $I^{[q]}$ we denote the ideal generated by x^q, $x \in I$. If M is a finite R-module, $M/I^{[q]}M$ has finite length. We will use $\lambda(-)$ to denote the length of an R-module. We assume knowledge of basic ideas in commutative algebra, including the usual Hilbert–Samuel multiplicity, Cohen–Macaulay, regular, and Gorenstein rings.

The basic question this chapter studies is how $\lambda(M/I^{[q]}M)$ behaves as a function on q, and how understanding this behavior leads to better understanding of the singularities of the ring R. In a seminal paper which appeared in 1969, [40], Ernst Kunz introduced the study of this function as a way to measure how close the ring R is to being regular.

The *Frobenius* homomorphism is the map $F : R \longrightarrow R$ given by $F(r) = r^p$. We say that R is *F-finite* if R is a finitely generated module over itself via the Frobenius homomorphism. It is not difficult to prove that if $(R, \mathfrak{m}, k)$ is a complete local Noetherian ring of characteristic p, or an affine ring over a field k of characteristic p, then R is F-finite if and only if $[k^{1/p} : k]$ is finite. When R is reduced we can identify the Frobenius map with the inclusion of R into $R^{1/p}$, the ring of pth roots of elements of R. If M is an R-module, we will usually write $M^{1/q}$ to denote what is more commonly denoted $F^e_*(M)$, where $q = p^e$, the module which is the same as M as abelian groups, but whose R-module structure is coming from restriction of scalars via e-iterates of the Frobenius map. This is an exact functor on the category of R-modules. Notice that $F^e_*(R)$ can be naturally identified with $R^{1/q}$.

C. Huneke (✉)
Department of Mathematics, University of Virginia, Charlottesville, VA 22902, USA
e-mail: huneke@virginia.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_15,

If the residue field k of R is perfect then the lengths of the R-modules $R^{1/q}/IR^{1/q}$ and $R/I^{[q]}$ are the same. If k is not perfect, but R is F-finite, then we can adjust by $[k^{1/q} : k]$. We define $\alpha(R) := \log_p([k^{1/p} : k])$, so that we can write $[k^{1/q} : k] = q^{\alpha(R)}$. With this notation, $\lambda_R(R^{1/q}/IR^{1/q}) = \lambda_R(R/I^{[q]})q^{\alpha(R)}$.

More broadly, the two numbers we will study, namely the Hilbert–Kunz multiplicity and the F-signature, are characteristic p invariants which give information about the singularities of R and lead to many interesting issues concerning how to use characteristic p methods to study singularities. There are four basic facts about characteristic p which make things work. Those facts are first, that $(r + s)^p = r^p + s^p$ for elements in a ring of characteristic p (i.e., the Frobenius is an endomorphism); second, that the map from $R \longrightarrow R^{1/p}$ is essentially the same map as that of $R^{1/q} \longrightarrow R^{1/qp}$ when R is reduced and $q = p^e$; third, that $\sum_i \frac{1}{p^i}$ converges (!); and lastly that the flatness of Frobenius characterizes regular rings. Virtually everything we prove comes down to these interrelated facts.

Throughout this chapter, whenever possible we have tried to give new (or at least not published) approaches to basic material. This is not done for the sake of whimsy, but to provide extra methods which may be helpful. Thus, the approach we take to proving the existence of the Hilbert–Kunz multiplicity and the F-signature, while following the general lines of the proofs of Paul Monsky [47] and Kevin Tucker [68], respectively, uses a lemma of Sankar Dutta [20] as a central point, which is not present in the usual proofs. When we present the proof of the existence of a second coefficient, we veer from the paper [38] to present another proof, based on the growth of the length of certain Tor modules, due to Moira McDermott and this author. In proving the theorem relating tight closure to the Hilbert–Kunz multiplicity, we use a lemma of Ian Aberbach [1] as a crucial point in the proof instead of presenting the original proof in [31]. We provide examples of Hilbert–Kunz multiplicities throughout this chapter, but often do not give details of the calculation.

We describe the contents of this chapter. In the second section we give some early results of Kunz on the relationship between regular local rings and the Hilbert–Kunz function. Kunz was ahead of his time in this regard, though characteristic p methods in commutative algebra were being using to study various homological conjectures at around the same time. In section three, we develop basic results and definitions needed to give our main existence theorems. The main technical tool we use is a lemma of Dutta [20] which gives information about the nature of prime filtrations of $R^{1/q}$. We prove that the Hilbert–Kunz multiplicity exists. Section four proves that for formally unmixed rings, the Hilbert–Kunz multiplicity is one if and only if R is regular. Here *formally unmixed* means that for all associated primes Q of the completion of a local ring R, $\dim \widehat{R}/Q = \dim R$. Section five provides the relationship between tight closure and Hilbert–Kunz multiplicity. In section six we prove that the F-signature exists and do some examples. Section seven proves the existence of a second coefficient in the Hilbert–Kunz function for normal rings. The final section takes up lower bounds on the Hilbert–Kunz multiplicity, introducing the volume estimates due to Watanabe and

Yoshida [74, 77], as well as the method of root adjunction of Aberbach and Enescu [6, 7] and recent improvements by Celikbas et al. [18]. We close with some results of Doug Hanes [29].

This survey does not present the considerable research dealing with the many remarkable and difficult calculations of Hilbert–Kunz multiplicity. For example, for work on plane cubics, see Pardue's thesis, [17, 48]. For plane curves in general see [66] and for general two-dimensional graded rings either [11] or [65]. For binomial hypersurfaces, see [19] or [69]. For flag varieties see [25]. The Hilbert–Kunz multiplicity of Rees algebras was the theme of [24]. Many other important examples or work are in [11–13, 19, 23, 24, 26, 46–53, 55, 56, 61, 65–68, 73–76]. We borrow freely from these papers for some of the examples presented in this chapter. We do not cover many new developments and calculations of the F-signature, for example, see [9, 10] and for toric rings see [61] and more recently [70]. See [22] for further extensions of Hilbert–Kunz multiplicity, and [71] for additional work. We also do not discuss the very interesting work being done on limiting value of Hilbert–Kunz multiplicities as p goes to infinity, for example, see [14, 26, 67]. For an excellent survey of other numerical invariants of singularities defined via Frobenius and their relationship to birational algebraic geometry and the theory of test ideals, see [4, 33, 45, 49, 50, 52, 58, 64, 75].

2 Early History

Ernst Kunz was a pioneer in this study, realizing that studying the colengths of Frobenius powers of $\mathfrak{m}$-primary ideals would be an interesting idea.

Theorem 1 (Kunz [40, Theorem 2.1, Proposition 3.2, Theorem 3.3]). *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic $p > 0$. For every $e \geq 0$ and $q = p^e$, $\lambda(R/\mathfrak{m}^{[q]}) \geq q^d$. Moreover, equality holds for some q if and only if R is regular, in which case equality holds for all q. If R is F-finite, then $R^{1/q}$ is a free module for some $q > 1$ if and only if R is regular.*

Proof. We may complete R and assume the residue field is algebraically closed to prove the first statement. We may also go modulo a minimal prime of R to assume that R is a complete local domain; this change will only potentially decrease $\lambda(R/\mathfrak{m}^{[q]})$. We claim that $R^{1/q}$ has rank q^d as an R-module in this case. Choose a coefficient field k and a minimal reduction $x_1, \ldots, x_d$ of the maximal ideal. Let A be the complete subring $k[[x_1, \ldots, x_d]]$ which is isomorphic with a formal power series. Note that $A^{1/q} \cong k[[x_1^{1/q}, \ldots, x_d^{1/q}]]$, which is a free A-module of rank q^d, whose basis is given by by arbitrary monomials of the form $x_1^{a_1/q} \cdots x_d^{a_d/q}$ where $0 \leq a_i \leq q - 1$. Since the rank of R over A and the rank of $R^{1/q}$ over $A^{1/q}$ are the same, it follows that the rank of $R^{1/q}$ over R is exactly q^d. (We note that if R is an F-finite complete domain but the residue field is not perfect, then essentially the same proof shows that the rank of $R^{1/q}$ is exactly $q^{(d+\alpha(R))}$.) Since $R^{1/q}$ is a finite R-module, $\mu_R(R^{1/q}) \geq q^d$, with equality if and only if $R^{1/q}$ is a free R-module. However, $\mu_R(R^{1/q}) = \lambda_R(R^{1/q}/\mathfrak{m}R^{1/q}) = \lambda_R(R/\mathfrak{m}^{[q]})$, which implies that $\lambda_R(R/\mathfrak{m}^{[q]}) \geq q^d$. Notice that equality occurs in this case if and only if $R^{1/q}$ is a free R-module.

If R is regular, then since $\mathfrak{m}$ is generated by a regular sequence, it easily follows that $\lambda(R/\mathfrak{m}^{[q]}) = q^d$. The second statement also easily is seen when R is regular and R is F-finite; one can complete and use the Cohen structure theorem to do the complete case and then descend using standard facts. It is the converse of both statements that is the most interesting part of the theorem.

Suppose that equality holds for some q, i.e., $\lambda(R/\mathfrak{m}^{[q]}) = q^d$. We can complete the ring and extend the residue field to be algebraically closed without changing this equality, so without loss of generality, R is F-finite and $\alpha(R) = 0$. Note that $\lambda(R/\mathfrak{m}^{[q^n]}) = q^{nd}$ for all $n \geq 1$, by a simple induction.

We claim that R is a domain; for if Q is a minimal prime of R of maximal dimension, then we have that $q^{nd} = \lambda(R/\mathfrak{m}^{[q^n]}) \geq \lambda(R/\mathfrak{m}^{[q^n]} + Q) \geq q^{nd}$. Hence we have equality throughout. But then $\lambda(R/\mathfrak{m}^{[q^n]}) = \lambda(R/\mathfrak{m}^{[q^n]} + Q)$ forces $\lambda((\mathfrak{m}^{[q^n]} + Q)/\mathfrak{m}^{[q^n]}) = 0$, so that $Q \subseteq \cap_n \mathfrak{m}^{[q^n]} = 0$. From the first part of this theorem, we then obtain that for all $n \geq 1$, R^{1/q^n} is a free R-module.

We next claim that R is Cohen–Macaulay. Let $x_1, \ldots, x_d$ be a system of parameters generating an ideal J. Then $\lambda(R/J^{[q^n]}) = \lambda(R^{1/q^n}/JR^{1/q^n}) = \lambda(R/J)q^{dn}$, since R^{1/q^n} is a free R-module of rank q^{dn}. By a formula of Lech [63, Theorem 11.2.10]: $\varinjlim \lambda(R/J^{[q^n]})/q^{dn} = e(J)$, the usual multiplicity of J. Hence the multiplicity of J is the colength of J. Since J is generated by a system of parameters, it follows that R is Cohen–Macaulay. (See [16, Theorem 4.6.10]).

Now choose a system of parameters as above, and fix n such that $\mathfrak{m}^{[q^n]} \subseteq J$, where J is the ideal generated by the parameters. Suppose that the projective dimension of k is infinite. We compute $\mathrm{Tor}_{d+1}(R/J, R/\mathfrak{m}^{[q^n]})$ in two ways. From the fact that J is generated by a regular sequence of length d, this Tor module is 0. On the other hand, we can take the free resolution of k and tensor with R^{1/q^n} and obtain an R^{1/q^n} minimal free resolution of $R^{1/q^n}/\mathfrak{m}R^{1/q^n}$. Identifying R^{1/q^n} with R, we see that a free resolution of $R/\mathfrak{m}^{[q^n]}$ is obtained by applying the Frobenius to the maps in the free resolution of k, which has the effect of raising all entries in matrices in the resolution (after fixing bases of the free modules) to the q^nth powers. Now tensoring with R/J, we see the homology at the $(d+1)$st stage is 0 if and only if the projective dimension of k is at most d, since the maps become 0 after tensoring with R/J. It follows that R is regular. □

Exercise 2. If $(R, \mathfrak{m}, k)$ is F-finite, and Q is a prime ideal, prove that $\alpha(R_Q) = \alpha(R)p^{\dim(R/Q)}$. (See [41, Proposition 2.3].)

Exercise 3. Let $(R, \mathfrak{m}, k)$ be a regular local ring of dimension d and prime characteristic p, and let I be an $\mathfrak{m}$-primary ideal. Prove that $\lambda(R/I^{[q]}) = q^d\lambda(R/I)$ so that in particular, $e_{HK}(I) = \lambda(R/I)$.

3 Basics

We begin with some estimates on the growth of the Hilbert–Kunz function, and some examples.

Lemma 1. *Let* $(R, \mathfrak{m}, k)$ *be a Noetherian local ring of dimension* d *and prime characteristic* $p > 0$. *We let* $e(I)$ *denote the multiplicity of the ideal* I. *Let* I *be an* $\mathfrak{m}$*-primary ideal. Then* ($q = p^e$),

$$e(I)/d! \le \liminf \lambda(R/I^{[q]})/q^d \le \limsup \lambda(R/I^{[q]})/q^d \le e(I).$$

Proof. We can make an extension of R to assume that the residue field is infinite without changing any of the relevant lengths. Let J be a minimal reduction of I, so that J is generated by a system of parameters. There are containments, $J^{[q]} \subseteq I^{[q]} \subseteq I^q$, which give inequalities on the lengths,

$$\lambda(R/J^{[q]}) \ge \lambda(R/I^{[q]}) \ge \lambda(R/I^q).$$

For large q, the right hand length is given by a polynomial in q of degree d with leading coefficient $e(I)/d!$. Dividing by q^d gives one inequality. For the other, we use a formula of Lech [63, Theorem 11.2.10]: $\varinjlim \lambda(R/J^{[q]})/q^d = e(J)$. Since J is a reduction of I, $e(J) = e(I)$. □

Corollary 2. *Let* $(R, \mathfrak{m}, k)$ *be a Noetherian local ring of dimension* 1 *and prime characteristic* $p > 0$ *Let* I *be an* $\mathfrak{m}$*-primary ideal. Then* $e(I) = \varinjlim \lambda(R/I^{[q]})/q^d$.

Proof. Set $d = 1$ in the above formula. □

Example 3. Although the one-dimensional case may seem very transparent, as the usual multiplicity equals the Hilbert–Kunz multiplicity, the actual Hilbert function is by no means obvious. Here is one example from [47]. Let k be a field of characteristic p congruent to 2 or 3 modulo 5. Set $R = k[[X, Y]]/(X^5 - Y^5)$. R is a one-dimensional local ring with maximal ideal $\mathfrak{m} = (x, y)$, and the multiplicity of R is 5. The difference $|\lambda(R/\mathfrak{m}^{[q]}) - 5q|$ is bounded by a constant. But it is not a constant in general. If we write the constant as d_e where $q = p^e$, then when e is even $d_e = -4$ while when e is odd, $d_e = -6$. For one-dimensional complete local rings, Monsky shows that the "constant" term is a periodic function. See [47] for details. See also [39] for work in the graded case.

Our goal of this section is to prove that $\varinjlim \lambda(R/I^{[q]})/q^d$ always exists. We call it the *Hilbert–Kunz multiplicity*. The history of how Monsky came to prove its existence is interesting. One might think that he was inspired by the paper of Kunz, but in fact, he did not know about it when he proved the existence. The situation was additionally complicated by the fact that Kunz had erroneously thought that the limit did not actually exist, and proposed a counterexample in his paper. This author asked Monsky how he came to think about it, and here is what he replied:

"Craig asked me how I was led into looking into Kunz's papers on the characterization of regular local rings in characteristic p (and defining and studying the Hilbert–Kunz multiplicity as a result). But that's not the order in which things occurred.

At Brandeis I was on the thesis committee of Al Cuoco, who was working in Iwasawa theory. He studied the growth of the p-part of the ideal class group as one moves up the levels in a tower of number fields, where the Galois group is a product of 2 copies of the p-adic integers. I extended his results to a product of s copies; this involved the study of modules over power series rings, with the base ring being the p-adics or $\mathbb{Z}/p\mathbb{Z}$. In particular I considered the following: let M be a finitely generated module over the power series ring in s variables over $\mathbb{Z}/p\mathbb{Z}$ and J be the ideal generated by the p^nth powers of the variables. How does the length of M/JM grow with n? I got an asymptotic formula for this growth, put it into a more general setting and wrote things up. In analogy with the Hilbert–Samuel terminology I intended to speak of the Hilbert–Frobenius function and the Hilbert–Frobenius multiplicity.

But when I showed my result to David Eisenbud he told me that it was wrong, and that Kunz had given examples in which there wasn't an asymptotic formula. So I looked into Kunz's papers, discovering that he had considered such questions before me. So it was only proper to call the function the Hilbert–Kunz function. And call the associated limiting value the Hilbert–Kunz multiplicity, even though Kunz had thought that it needn't exist!"

To prove the existence of the Hilbert–Kunz multiplicity, we will consider modules as well as rings. We use a somewhat different treatment than the paper of Monsky [47], organizing our approach through a lemma proved by Dutta [20], which is not only interesting in its own right, but has the additional benefit that we can directly apply it to show the existence of the F-signature as well. However, in the end, all the approaches use that the map from R to $R^{1/p}$ is essentially the same as $R^{1/q}$ to $R^{1/qp}$, and that the sum of the reciprocals of the powers of p converges.

Lemma 4 (Dutta [20, see proof of Proposition, p. 428]). *Let $(R, \mathfrak{m}, k)$ be a local Noetherian domain of dimension d and prime characteristic p. Assume that R is F-finite. Then there exists a constant C and a fixed finite set of nonzero primes, $\{Q_1, \ldots, Q_n\}$ such that for every $q = p^e$, the R-module $R^{1/q}$ has a prime filtration having at most $Cq^{d+\alpha(R)}$ copies of R/Q_i for $i \geq 1$ and $q^{d+\alpha(R)}$ copies of R.*

Proof. The proof we give, similar to Dutta's proof, was shown to me by Karen Smith, and is essentially found in Appendix 2 of [35], proof of Exercise 10.4.

Use induction on d; the $d = 0$ case is trivial.

Fix a maximal rank free submodule G of $R^{1/p}$. We know that the rank of G is $p^{d+\alpha(R)}$. Let T be the cokernel of the inclusion $G \subset R^{1/p}$. Fix a prime cyclic filtration of T, and extend it by G to a filtration of $R^{1/p}$:

$$0 \subset G = M_0 \subset M_1 \subset M_2 \subset \cdots \subset M_t = R^{1/p}.$$

Because G is maximal rank, the prime cyclic factors $M_{i+1}/M_i = R/\mathfrak{A}_i$ all have dimension strictly less than the dimension of R. Let C_i be the constant which (by induction) works for $R/\mathfrak{A}_i$, let C be twice the sum of all the C_i, and let Ω be the collection of the (finite) sets of primes appearing in the filtrations of all the $(R/\mathfrak{A}_i)^{1/q}$, as well as the prime (0). We claim that Ω and C satisfy the conclusion of the problem.

By induction on q, we prove that $R^{1/q}$ has a prime filtration using primes from Ω, with at most $\frac{C}{2}(1 + 1/p + \ldots + 1/q)q^{d+\alpha(R)}$ copies of each one. Assume this is true for q. Take $p^e = q$ roots of all the modules above. We have a prime cyclic filtration (except at zeroth spot, where it is obvious how to extend to one) of $R^{1/q}$ modules

$$0 \subset G^{1/q} = M_0^{1/q} \subset M_1^{1/q} \subset M_2^{1/q} \subset \cdots \subset M_t^{1/q} = R^{1/qp},$$

where each factor has the form $(R/\mathfrak{A}_i)^{1/q} = R^{1/q}/\mathfrak{A}_i^{1/q}$.

To make this into a prime cyclic filtration of R modules, we simply refine each inclusion $M_i^{1/q} \subset M_{i+1}^{1/q}$ of R modules by a prime cyclic filtration. This amounts to filtering $M_{i+1}^{1/q}/M_i^{1/q} = (R/\mathfrak{A}_i)^{1/q}$ by $R/\mathfrak{A}_i$ prime cyclic modules. By induction on d, this can be done with only primes from Ω, and appearing with multiplicities at most $\leq C_i q^{d-1+\alpha(R/\mathfrak{A}_i)} = C_i q^{d-1+\alpha(R)}$. Thus the primes appearing in this prime cycle filtration of $R^{1/qp}/G^{1/q}$ all come from Ω, and each one appears at most $(\sum_i C_i)q^{d-1+\alpha(R)}$ times.

To refine the R submodule $G^{1/q}$ into a prime filtration we deal with each of the free summands $R^{1/q}$ separately. By induction there are only primes from Ω appearing, and the multiplicity of R/Q_i in $G^{1/q}$ is no more than $(\text{rank} G)(\frac{C}{2})(1 + 1/p + \ldots + 1/q)(q^{d+\alpha(R)})$. The total number is then at most $(\frac{C}{2})(1 + 1/p + \ldots + 1/q)((qp)^{d+\alpha(R)}) + \frac{C}{2}q^{d-1+\alpha(R)} \leq \frac{C}{2}(1 + \ldots + 1/(qp))(qp)^{d+\alpha(R)} \leq C(qp)^{d+\alpha(R)}$. □

Lemma 5. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic $p > 0$. Let M be a finitely generated R-module. There exists a constant $C > 0$ such that for all $e \geq 0$ and any $\mathfrak{m}$-primary ideal I of R with $\mathfrak{m}^{[q]} \subseteq I$, where $q = p^e$, we have that*

$$\lambda(R/I \otimes_R M) \leq C q^{\dim M}.$$

Proof. Set $t = \mu(\mathfrak{m})$. Since $\mathfrak{m}^{tq} \subseteq \mathfrak{m}^{[q]}$, we see that $R/\mathfrak{m}^{tq} \otimes_R M$ surjects onto $R/I \otimes_R M$. Therefore $\lambda(R/I \otimes_R M) \leq \lambda(R/(\mathfrak{m}^{tq}) \otimes_R M)$. The Hilbert polynomial of M with respect to $\mathfrak{m}^t$ has degree $\dim(M)$. If the leading coefficient of this polynomial is c, it is clear that any $C >> c$ satisfies the desired bound. □

Lemma 6. *Let $(R, \mathfrak{m}, k)$ be a local ring of dimension d and prime characteristic p. If T is a finitely generated torsion R-module then there exists a constant D such that for all $q = p^e$, and for all I containing $\mathfrak{m}^{[q]}$, $\lambda(\text{Tor}_1^R(R/I, T)) \leq Dq^{d-1}$.*

Proof. Choose a nonzero divisor $c \in R$ which annihilates T, and consider an $R/(c) = A$ presentation of T:

$$\dots \longrightarrow A^s \longrightarrow A^r \longrightarrow T \longrightarrow 0.$$

Let N be the kernel of the surjection of A^r onto T. Tensoring with R/I, we obtain an exact sequence,

$$\mathrm{Tor}_1^R(A^r, R/I) \longrightarrow \mathrm{Tor}_1^R(T, R/I) \longrightarrow N/IN \longrightarrow (A/I)^r \longrightarrow T/IT \longrightarrow 0.$$

Since N is torsion, Lemma 5 implies that the length of N/IN is bounded above by Eq^{d-1}, for some fixed constant E depending only on N. Thus it suffices to bound the length of $\mathrm{Tor}_1^R(A^r, R/I)$. Notice that r does not depend upon q or I. Hence it suffices to bound the length of $\mathrm{Tor}_1^R(A, R/I)$. From the exact sequence $0 \longrightarrow R \xrightarrow{c} R \longrightarrow A \longrightarrow 0$, we obtain after tensoring with R/I that $\mathrm{Tor}_1^R(A, R/I) \cong (I : c)/I$. However, the length of $(I : c)/I$ is the same as the length of $R/(I, c)$, and by Lemma 5, this length is bounded by Gq^{d-1} for some constant G depending only on A. □

Exercise 7. Prove Lemma 6 with the modification that $\lambda(\mathrm{Tor}_1^R(R/I, T)) \leq Dq^{\dim(T)}$ (this is not so easy).

These lemmas have the following crucial consequence, which is a key point in the paper of Tucker [68, Corollary 3.5]:

Corollary 8. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local domain of dimension d and prime characteristic p. Assume that R is F-finite. There exists a constant C such that for all $q = p^e$ and all $q' = p^{e'}$ and for all ideals I containing $\mathfrak{m}^{[q]}$,*

$$|\lambda(R/I^{[q']}) - (q')^{d+\alpha(R)}\lambda(R/I)| \leq C(q')^{d+\alpha(R)}q^{d-1}.$$

Proof. Fix the constant C and the primes $\{Q_1, \dots, Q_n\}$ as in the statement of Lemma 4. Then for all q' there is an exact sequence,

$$0 \longrightarrow R^{(q')^{d+\alpha(R)}} \longrightarrow R^{1/q'} \longrightarrow T \longrightarrow 0,$$

where T has a prime filtration by at most $C(q')^{d+\alpha(R)}$ copies of each R/Q_i. Tensoring with R/I, we see that the difference of lengths, $|\lambda(R/I^{[q']}) - (q')^{d+\alpha(R)}\lambda(R/I)|$, is bounded by the sum of $\lambda(T/IT) + \lambda(\mathrm{Tor}_1^R(T, R/I))$. This sum in turn is bounded by

$$\sum_{i=1}^{n} C(q')^{d+\alpha(R)}(\lambda(R/(Q_i, I)) + \lambda(\mathrm{Tor}_1^R(R/Q_i, R/I)).$$

To prove the corollary it suffices to prove that there is a constant D, not depending on q, q', or I such that $\lambda(R/(Q_i, I)) \leq Dq^{d-1}$ for each i, and $\lambda(\mathrm{Tor}_1^R(R/Q_i, R/I)) \leq Dq^{d-1}$. The existence of such a constant D follows from Lemmas 5 and 6, respectively. □

Remark 9. We can now easily prove that the Hilbert–Kunz multiplicity exists for the ring itself and arbitrary $\mathfrak{m}$-primary ideals I in the case R is an F-finite domain. To do the general case, however, requires a little more work which one needs in any case to deal with additivity properties of the Hilbert–Kunz multiplicity. However, it is worth seeing this easy case deduced from the corollary. We may assume that k is algebraically closed. Set $c_q = \lambda(R/I^{[q]})/q^d$. Apply Corollary 8 with I replaced by $I^{[q]}$. Divide by $(q'q)^d$. We obtain that for all q, q',

$$|c_{qq'} - c_q| \leq \frac{C}{q}.$$

This inequality forces the set of c_q to be a Cauchy sequence, and hence they converge.

Lemma 10. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local reduced ring of dimension d and prime characteristic $p > 0$. Let $P_1, \ldots, P_m$ be those minimal primes of R with $\dim(R/P_i) = d$. If M and N are finitely generated R-modules such that $M_{P_i} \cong N_{P_i}$ for each i, then there exists a positive constant C such that for all $e \geq 0$ and for every ideal I of R with $\mathfrak{m}^{[q]} \subseteq I$, where $q = p^e$, we have $|\lambda(R/I \otimes_R M) - \lambda(R/I \otimes_R N)| \leq Cq^{d-1}$.*

Proof. Let $W = R \setminus (\cup_i P_i)$, so that $R_W \cong R_{P_1} \times \cdots \times R_{P_m}$, and we have that $M_W \cong N_W$. Since $(\mathrm{Hom}_R(M, N))_W \cong \mathrm{Hom}_{R_W}(M_W, N_W)$, there is some $\phi \in \mathrm{Hom}_R(M, N)$ such that ϕ_W is an isomorphism. Since $\mathrm{coker}(\phi)$ satisfies $\mathrm{coker}(\phi)_W = 0$ and thus has dimension strictly smaller than d, we can find a positive constant C such that for all $e \geq 0$ and for any ideal I of R which contains $\mathfrak{m}^{[q]}$, we have that $|\lambda(R/I \otimes_R R/\mathrm{coker}(\phi))| \leq Cq^{d-1}$. □

We use some well-known notation in the next few results. Let $f, g : \mathbb{N} \to \mathbb{R}$ be functions from the nonnegative integers to the real numbers. Recall that $f(n) = O(g(n))$ if there exists a positive constant C such that $|f(n)| \leq Cg(n)$ for all $n \gg 0$, and we write $f(n) = o(g(n))$ if $\lim_{n\to\infty} f(n)/g(n) = 0$.

Proposition 11. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic $p > 0$. Let $0 \to N \to M \to K \to 0$ be a short exact sequence of finitely generated R-modules. Then,*

$$\lambda(M/I^{[q]}M) = \lambda(N/I^{[q]}N) + \lambda(K/I^{[q]}K) + O(q^{d-1}).$$

Proof. First suppose that R is reduced. Then M and $N \oplus K$ have isomorphic localizations at each minimal prime of R, and the claim follows from Lemma 10.

If R is not reduced, choose q' such that $(\mathrm{nilrad}(R))^{[q']} = 0$, and consider the same exact sequence as a sequence of $R^{q'}$-modules. This ring is reduced and applying the reduced case with the ideal $I^{[q']} \cap R^{q'}$ yields that

$$\lambda(M/I^{[qq']}M) = \lambda(N/I^{[qq']}N) + \lambda(K/I^{[qq']}K) + O(q^{d-1}).$$

Since $O(q^{d-1}) = O((qq')^{d-1})$, the proposition is proved. □

We are now able to prove the existence of the Hilbert–Kunz multiplicity:

Theorem 12. *Let* $(R, \mathfrak{m}, k)$ *be a Noetherian local ring of dimension* d *and prime characteristic* $p > 0$. *Let* M *be a finitely generated* R*-module, and let* I *be an* $\mathfrak{m}$*-primary ideal. There is a real constant* $\alpha = e_{HK}(I, M) \geq 1$ *such that* $\lambda(M/I^{[q]}M) = \alpha q^d + O(q^{d-1})$. *If*

$$0 \to N \to M \to K \to 0$$

is a short exact sequence of finitely generated R*-modules, then*

$$e_{HK}(I, M) = e_{HK}(I, K) + e_{HK}(I, N).$$

Proof. By making a faithfully flat extension there is no loss of generality in assuming that R is a complete local ring with algebraically closed residue field. By taking a prime filtration of M and using Proposition 11 it suffices to do the case in which $M = R/P$ for some prime P of R. Thus there is no loss of generality in assuming that R is an F-finite domain and $M = R$ in proving the first assertion. The second assertion follows immediately from the first assertion and Proposition 11.

To prove the existence, we are now in the case of Remark 9, which finishes the proof. □

We often suppress the R in $e_{HK}(I, R)$ and just write $e_{HK}(I)$. When $I = \mathfrak{m}$, we set $e_{HK}(M) = e_{HK}(\mathfrak{m}, M)$, and refer to this value as the Hilbert–Kunz multiplicity of M.

Example 13. Unlike the usual multiplicity, the Hilbert–Kunz multiplicity is typically not an integer. The Hilbert–Kunz function can appear quite bizarre, at least to begin with. For example, let $R = \mathbb{Z}/5\mathbb{Z}[x_1, x_2, x_3, x_4]/(x_1^4 + \cdots + x_4^4)$, then with $I = (x_1, \ldots, x_4)$, $\lambda(R/I^{[5^e]}) = \frac{168}{61}(5^{3e}) - \frac{107}{61}(3^e)$ by [28]. Note that R is a 3-dimensional Gorenstein ring with isolated singularity.

Just as in the theory of usual multiplicity, it is now easy to prove some basic remarks on the behavior of the Hilbert–Kunz multiplicity. In particular, the following additivity theorem is highly useful.

Theorem 14. *Let* $(R, \mathfrak{m}, k)$ *be a local Noetherian ring of dimension* d *and prime characteristic* p. *let* I *be an* $\mathfrak{m}$*-primary ideal, and let* M *be a finitely generated* R*-module. Let* Λ *be the set of minimal prime ideals* P *of* R *such that* $\dim(R/P) = \dim(R)$. *Then*

$$e_{HK}(I, M) = \sum_{P \in \Lambda} e_{HK}(I, R/P)\lambda(M_P).$$

Proof. By Theorem 11, Hilbert–Kunz multiplicity is additive on short exact sequences. Fix a prime filtration of M, say

$$0 = M_0 \subseteq M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n = M$$

where $M_{i+1}/M_i \cong R/P_i$ (P_i a prime) for all $0 \le i \le n-1$. As $e_{HK}(I, R/Q) = 0$ if $\dim(R/Q) < \dim(R)$, the additivity of multiplicity applied to this filtration shows that $e_{HK}(I, M)$ is a sum of the $e_{HK}(I, R/P)$ for $P \in \Lambda$, counted as many times as R/P appears as some M_{i+1}/M_i. We can count this by localizing at P. In this case, we have a filtration of M_P, where all terms collapse except for those in which $(M_{i+1}/M_i)_P \cong (R/P)_P$, and the number of such copies is exactly the length of M_P. □

Corollary 15. *Let $(R, \mathfrak{m}, k)$ be a local Noetherian domain of dimension d and prime characteristic p. Let I be an $\mathfrak{m}$-primary ideal of R and M a finitely generated R-module. Then $e_{HK}(I, M) = e_{HK}(I, R) \operatorname{rank}_R M$.*

Proof. Recall that the rank of M is by definition the dimension of $M \otimes_R K$ over K, where K is the field of fractions of R. We apply Lemma 10 with $W = R \setminus 0$: if we set $r = \operatorname{rank}_R M$, then $W^{-1}M \cong K^r \cong W^{-1}R^r$, and the corollary follows. □

Theorem 16. *Let $(R, \mathfrak{m}, k)$ be a d-dimensional local Noetherian domain of prime characteristic p, with field of fractions K, and let I be an $\mathfrak{m}$-primary ideal. Let S be a module-finite extension domain of R with field of fractions L. Then*

$$e_{HK}(I, R) = \sum_{Q \in \operatorname{Max}(S), \dim S_Q = d} \frac{e_{HK}(IS_Q, S_Q)[S/Q : k]}{[L : K]}.$$

Proof. Since $W^{-1}S \cong W^{-1}R^{[L:K]}$, we can apply Lemma 10 to conclude that $e_{HK}(I, S) = e_{HK}(I, R)[L : K]$. On the other hand,

$$e_{HK}(I, S) = \lim_{q \to \infty} \lambda_R(S/I^{[q]}S)/q^d.$$

As every maximal ideal Q of S contains $\mathfrak{m}S$, the Chinese Remainder Theorem implies that $S/I^{[q]}S \cong \prod_{Q \in \operatorname{Max}(S)} S_Q/I^{[q]}S_Q$. In particular, $\lambda_R(S/I^{[q]}S) = \sum_{Q \in \operatorname{Max}(S)} \lambda_R(S_Q/I^{[q]}S_Q) = \sum_{Q \in \operatorname{Max}(S)} \lambda_{S_Q}(S_Q/I^{[q]}S_Q)[S/Q : k]$. Therefore, $e_{HK}(I, S)$ equals

$$\lim_{q \to \infty} \sum_{Q \in \operatorname{Max}(S)} \lambda_{S_Q}(S_Q/I^{[q]}S_Q)[S/Q : k]/q^d$$
$$= \lim_{q \to \infty} \sum_{\dim S_Q = d} \lambda_{S_Q}(S_Q/I^{[q]}S_Q)[S/Q : k]/q^d.$$

Hence

$$e_{HK}(I, R) = \sum_{Q \in \operatorname{Max}(S), \dim S_Q = d} \frac{e_{HK}(IS_Q, S_Q)[S/Q : k]}{[L : K]}.$$ □

Example 17. Consider the Veronese subring R defined by

$$R = k[[X_1^{i_1} \cdots X_d^{i_d} \mid i_1, \ldots, i_d \ge 0, \sum i_j = r]].$$

Applying Theorem 16 to $R \hookrightarrow S = k[[X_1, \dots, X_d]]$, we get

$$e_{HK}(R) = \frac{1}{r}\binom{d+r-1}{r}. \tag{1}$$

In particular, if $d = 2$, $r = e(A)$, then $e_{HK}(R) = \frac{e(R)+1}{2}$.

For other examples, consider the quotient singularities.

Example 18. See [73, Theorem 5.4]. Let S be a regular local ring and suppose that G is a finite group of automorphisms of S with invariant ring R with maximal ideal $\mathfrak{m}$. By Theorem 16 and Exercise 3, one sees that $e_{HK}(R) = \frac{1}{|G|}\lambda(S/\mathfrak{m}S)$.

This formula is used, together with a lot more work, by Watanabe and Yoshida to give the following formulas for the Hilbert–Kunz multiplicities of the famous double points below: Let $(R, \mathfrak{m}) = k[[x, y, z]]/(f)$ where f is one of the following:

type	equation	char R	$e_{HK}(R)$	
(A_n)	$f = xy + z^{n+1}$	$p \geq 2$	$2 - 1/(n+1)$	$(n \geq 1)$
(D_n)	$f = x^2 + yz^2 + y^{n-1}$	$p \geq 3$	$2 - 1/4(n-2)$	$(n \geq 4)$
(E_6)	$f = x^2 + y^3 + z^4$	$p \geq 5$	$2 - 1/24$	
(E_7)	$f = x^2 + y^3 + yz^3$	$p \geq 5$	$2 - 1/48$	
(E_8)	$f = x^2 + y^3 + z^5$	$p \geq 7$	$2 - 1/120$	

Each of these hypersurfaces is the invariant subring by a finite subgroup $G \subseteq SL(2, k)$ which acts on the polynomial ring $k[x, y]$. We have that $e_{HK}(R) = 2 - 1/|G|$; see [73, Theorem 5.1].

Example 19. Let $S = k[x, y, z]$ where k is a field of characteristic at least five. Let $h \in S$ be homogeneous of degree 3. Set $R = S/(h)$, and let $\mathfrak{m} = (x, y, z)R$. If h is smooth, then $e_{HK}(\mathfrak{m}) = \frac{9}{4}$, while if h is a nodal or cuspidal cubic, $e_{HK}(\mathfrak{m}) = \frac{7}{3}$. This has been done in various ways. Pardue in his thesis did the nodal cubic; see also Buchweitz and Chen [17], Brenner [13], Monsky [51], and Trivedi [65] and in characteristic 2, [48].

Here are a few more examples, which we leave as an exercise:

Exercise 20. We consider quadric hypersurfaces in $\mathbb{P}^3$. Let k be a field of characteristic $p > 2$, and let Let R be one of the following rings:

$$\begin{cases} k[[X, Y, Z, W]]/(X^2), & \text{if } \operatorname{rank}(q) = 1, \\ k[[X, Y, Z, W]]/(X^2 - YZ), & \text{if } \operatorname{rank}(q) = 2, \\ k[[X, Y, Z, W]]/(XY - ZW), & \text{if } \operatorname{rank}(q) = 3. \end{cases} \tag{2}$$

Prove that $e_{HK}(R) = 2$, $\frac{3}{2}$, or $\frac{4}{3}$, respectively.

For a long time it was thought that the Hilbert–Kunz multiplicity would always be a rational number. All the known examples were rational, e.g., for rings of finite Cohen–Macaulay type (see [59]) or more generally F-finite type [62, 79], for many computed hypersurfaces, for binomial hypersurfaces [19], and for graded normal rings of dimension two [12, 65]. However, in recent years, Monsky has given convincing evidence that this will not be true, though as of the writing of this chapter, there is only overwhelming evidence, but not a proof. One example given by Monsky is the following:

Example 21. Let k be a finite field of characteristic 2 and $h = x^3 + y^3 + xyz \in k[[x, y, z]]$. Then Monsky conjectures, with a huge amount of evidence, that the Hilbert–Kunz multiplicity of the hypersurface $uv + h = 0$ is $\frac{4}{3} + \frac{5}{14\sqrt{7}}$. Even more, it appears that transcendental Hilbert–Kunz multiplicities exist. We refer to [53, 54] for details.

4 Hilbert–Kunz Multiplicity Equal to One

We begin this section with an easy, but crucial, estimate on the size of Hilbert–Kunz functions which was observed independently in [73, Lemma 4.2] and [29].

Lemma 1. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic $p0$. Let $I \subseteq J$ be two ideals with I $\mathfrak{m}$-primary (we allow $J = R$). Then $\lambda(R/I^{[q]}) \leq \lambda(J/I) \cdot \lambda(R/\mathfrak{m}^{[q]}) + \lambda(R/J^{[q]})$.*

Proof. Set $s = \lambda(J/I)$. Take a filtration of $I \subseteq J \subseteq R$:

$$I = J_0 \subsetneq J_1 \subsetneq J_2 \subsetneq \cdots \subsetneq J_s = J \subseteq R$$

so that $\lambda(J_i/J_{i-1}) = 1$, i.e., $J_i/J_{i-1} \cong R/\mathfrak{m}$, $\forall i = 1, 2, \ldots, s$. That is to say $J_i = (J_{i-1}, x_i)$ for some $x_i \in J_i$ such that $J_{i-1} : x_i = \mathfrak{m}$.

For every $q = p^e$, there is a corresponding filtration of $I^{[q]} \subseteq J^{[q]} \subseteq R$:

$$I^{[q]} = J_0^{[q]} \subseteq J_1^{[q]} \subseteq J_2^{[q]} \subseteq \cdots \subseteq J_s^{[q]} = J^{[q]} \subseteq R,$$

where $J_i^{[q]}/J_{i-1}^{[q]} \cong R/(J_{i-1}^{[q]} : x_i^q)$, which is a homomorphic image of $R/m^{[q]}$, for every $i = 1, 2, \ldots, s$. So $\lambda(J_i^{[q]}/J_{i-1}^{[q]}) \leq \lambda(R/\mathfrak{m}^{[q]})$. Therefore $\lambda(R/I^{[q]}) \leq \lambda(J/I) \cdot \lambda(R/\mathfrak{m}^{[q]}) + \lambda(R/J^{[q]})$. □

Corollary 2. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic p. Let I be a $\mathfrak{m}$-primary ideal of R. Then $\lambda(R/I^{[q]}) \leq \lambda(R/I) \cdot \lambda(R/\mathfrak{m}^{[q]})$. If $I \subseteq J$ then $e_{HK}(I, R) \leq \lambda(J/I) e_{HK}(R) + e_{HK}(J, R)$.*

Proof. To prove the first statement, we take $J = R$ and apply Lemma 1. For the second statement, the corollary follows from Lemma 1 by dividing by q^d and then taking the limits. □

Our goal is to prove that regularity is characterized by the Hilbert–Kunz multiplicity being one, if the ring is formally unmixed . This condition is necessary by the easy exercise below. Our treatment is taken directly from [37].

Exercise 3. Let $R = k[[x, y, z]]/(xz, xy)$, where k is a field of characteristic p. Prove that $e_{HK}(R) = 1$.

Theorem 4. *Let $(R, \mathfrak{m})$ be a Noetherian local ring of dimension d and prime characteristic p. Let J be an ideal such that* $\dim R/J = 1$ *and* height $J = d - 1$. *Assume that $x \in R$ is a nonzero divisor in R/J and set $I = (J, x)$. Assume that R_P is regular for every minimal prime P containing J. Then $e_{HK}(I, R) \geq \lambda(R/I)$.*

Proof. Use the properties of the usual multiplicity of parameter ideals and the associativity formula for the usual multiplicity, and we have

$$\begin{aligned}
e_{HK}(I, R) &= \lim_{q\to\infty} \frac{1}{q^d} \cdot \lambda(R/I^{[q]}) = \lim_{q\to\infty} \frac{1}{q^d} \cdot \lambda(R/(J^{[q]}, x^q)) \\
&\geq \lim_{q\to\infty} \frac{1}{q^d} \cdot e(x^q; R/J^{[q]}) = \lim_{q\to\infty} \frac{q}{q^d} \cdot e(x; R/J^{[q]}) \\
&= \lim_{q\to\infty} \frac{1}{q^{d-1}} \cdot e(x; R/J^{[q]}) \\
&= \lim_{q\to\infty} \frac{1}{q^{d-1}} \cdot \sum_{P\in\min(R/J)} e(x; R/P) \cdot \lambda_{R_P}(R_P/J_P^{[q]}) \\
&= \lim_{q\to\infty} \frac{1}{q^{d-1}} \cdot \sum_{P\in\min(R/J)} e(x; R/P) \cdot q^{d-1} \cdot \lambda_{R_P}(R_P/J_P) \\
&= \lim_{q\to\infty} \sum_{P\in\min(R/J)} e(x; R/P) \cdot \lambda_{R_P}(R_P/J_P) \\
&= \sum_{P\in\min(R/J)} e(x; R/P) \cdot \lambda_{R_P}(R_P/J_P) \\
&= e(x; R/J) = \lambda(R/(J, x)) = \lambda(R/I).
\end{aligned}$$

□

Observe that after we prove that $e_{HK}(R) = 1$ implies the regularity of R, then regularity forces $e_{HK}(I) = \lambda(R/I)$ for all $\mathfrak{m}$-primary ideals I, by using the work above.

A critical step in proving the main result of this section is in constructing an $\mathfrak{m}$-primary ideal $I \subseteq \mathfrak{m}^{[p]}$ such that $e_{HK}(I) \geq \lambda(R/I)$. This was proved by Watanabe and Yoshida [73, Theorem 1.5] but in a different way than is done here.

Theorem 5. *Let $(R, \mathfrak{m}, k)$ be a formally unmixed Noetherian local ring of dimension d and prime characteristic p. Then $e_{HK}(R) = 1$ if and only if R is regular.*

Proof. We have already observed that if R is regular, then the Hilbert–Kunz multiplicity is one. We prove the converse. Since the Hilbert–Kunz multiplicity of R is the same as that of its completion, we may assume R is complete. The additivity formula for Hilbert–Kunz multiplicity Theorem 14 shows that $e_{HK}(R) = \sum_P e_{HK}(R/P) \cdot \lambda(R_P)$ where the sum is over all minimal primes of maximal dimension. Since $e_{HK}(R) = 1$, we deduce that R can have only one minimal prime P and R_P has to be field, i.e., $P_P = 0$. Hence $P = 0$ since $R \setminus P$ consists of non-zero divisors. Thus R is a domain.

It suffices to prove that $\lambda(R/\mathfrak{m}^{[p]}) \leq p^d$ (where $d = \dim(R)$) as then Theorem 1 gives that R must be regular.

The singular locus of R is closed and not equal to $\operatorname{Spec}(R)$. It follows that we can choose a prime P such that $\dim(R/P) = 1$ and R_P is regular, which we leave as an exercise for the reader. Since the intersection of the symbolic powers of P is zero and R is complete, Chevalley's lemma gives that some sufficiently large symbolic power of P lies inside $\mathfrak{m}^{[p]}$. Call this symbolic power J. Choose $x \in \mathfrak{m}^{[p]}$ such that $x \notin P$. The ideal $I = (J, x)$ lies in $\mathfrak{m}^{[p]}$ and satisfies the hypothesis of Theorem 4. Hence

$$e_{HK}(I) \geq \lambda(R/I).$$

On the other hand we have $e_{HK}(I, R) \leq \lambda(\mathfrak{m}^{[p]}/I) \cdot e_{HK}(R) + e_{HK}(\mathfrak{m}^{[p]}, R) = \lambda(\mathfrak{m}^{[p]}/I) + e_{HK}(\mathfrak{m}^{[p]}, R) \leq \lambda(\mathfrak{m}^{[p]}/I) + \lambda(R/\mathfrak{m}^{[p]})$, by Lemma 1 and Corollary 2.

That is to say

$$\begin{aligned} \lambda(\mathfrak{m}^{[p]}/I) + \lambda(R/\mathfrak{m}^{[p]}) =& \lambda(R/I) \leq e_{HK}(I, R) & (3)\\ \leq& \lambda(\mathfrak{m}^{[p]}/I) + e_{HK}(\mathfrak{m}^{[p]}, R) & (4)\\ \leq& \lambda(\mathfrak{m}^{[p]}/I) + \lambda(R/\mathfrak{m}^{[p]}), & (5) \end{aligned}$$

which forces $\lambda(R/\mathfrak{m}^{[p]}) = e_{HK}(\mathfrak{m}^{[p]}, R)$. However,

$$e_{HK}(\mathfrak{m}^{[p]}, R) = \varinjlim \frac{\lambda(R/\mathfrak{m}^{[pq]})}{q^d} = \varinjlim \frac{p^d \cdot \lambda(R/\mathfrak{m}^{[pq]})}{(pq)^d} = p^d \cdot e_{HK}(R) = p^d.$$

Together the equalities imply that $\lambda(R/\mathfrak{m}^{[p]}) = p^d$, which implies that R is regular by Theorem 1. □

The basic filtration lemmas, together with Kunz's theorem, already give a better result, *provided* the ring is Cohen–Macaulay. In fact, this is one of the more subtle and difficult points, to prove that Hilbert–Kunz multiplicity near one should imply that the ring is Cohen–Macaulay. A crucial step is provided by results of Goto and Nakamura; see [27]. They prove the following beautiful generalization of the result of Serre which proves that the multiplicity of a parameter ideal is its colength if and only if the ring is Cohen–Macaulay.

Theorem 6 (Goto–Nakamura [27]). *Let $(R, \mathfrak{m}, k)$ be an unmixed Noetherian local ring of prime characteristic p which is the homomorphic image of a Cohen–Macaulay local ring. Let J be an ideal generated by a system of parameters. Then $e(J) \geq \lambda(R/J^*)$ with equality if and only if R is F-rational (and therefore is Cohen–Macaulay).*

The general philosophy is that the closer the Hilbert–Kunz multiplicity is to one, the better the singularities of the ring. The following proposition was proved by Blickle and Enescu, using results of Goto and Nakamura and Watanabe and Yoshida to first obtain that the ring is Cohen–Macaulay. We state the full result here, but only give the proof assuming Cohen–Macaulay.

Proposition 7 (Blickle–Enescu [8]). *Let $(R, \mathfrak{m})$ be a Noetherian local ring of dimension d and prime characteristic p. If R is not regular, then $e_{HK}(R) > 1 + \frac{1}{p^d d!}$.*

Proof. We give the proof assuming that R is Cohen–Macaulay. Let $e(R)$ be the multiplicity of R. We may assume the residue field is infinite. Fix a minimal reduction K of the maximal ideal. We apply Corollary 2 with $I = K^{[p]}$ and $J = \mathfrak{m}^{[p]}$. This gives that $e(R)p^d = e_{HK}(K^{[p]}) \leq \lambda(\mathfrak{m}^{[p]}/K^{[p]})e_{HK}(R) + e_{HK}(\mathfrak{m}^{[p]}) = \lambda(\mathfrak{m}^{[p]}/K^{[p]})e_{HK}(R) + p^d e_{HK}(R)$. Note that the first equality follows from the formula of Lech [63, Theorem 11.2.10]. By Theorem 1, $\lambda(\mathfrak{m}^{[p]}/K^{[p]}) = e(R)p^d - \lambda(R/\mathfrak{m}^{[p]}) \leq e(R)p^d - (p^d + 1)$ because R is not regular. Putting these inequalities together and cancelling terms yields that $e(R)p^d \leq (e(R)p^d - 1)e_{HK}(R)$ or $1 + \frac{1}{e(R)p^d - 1} \leq e_{HK}(R)$. Since $e(R)/d! \leq e_{HK}(R)$, if $e(R) > d!$, then $1 + \frac{1}{d!} < e_{HK}(R)$, a stronger statement than what we claim. Otherwise, $e(R)p^d - 1 < p^d d!$, and the proposition follows. □

The reader should ask themselves where the assumption that R is Cohen–Macaulay is used in the above proof.

The methods in this section also give a proof of a result of Kunz concerning the behavior of Hilbert–Kunz multiplicity under specialization. It is still an open problem whether or not Hilbert–Kunz multiplicity is upper semicontinuous. See, however, the interesting papers of Shepherd-Barron [60] (but be careful—Corollary 2 is not quite correct) and Enescu and Shimomoto [21].

Proposition 8 (Kunz [41, Corollary 3.8]). *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic p, and let P be a prime ideal of R such that* $\text{height}(P) + \dim(R/P) = \dim(R)$. *Then $e_{HK}(R_P) \leq e_{HK}(R)$. In fact, if $t = \dim(R/P)$, then $q^t \cdot \lambda_{R_P}((R/P^{[q]})_P) \leq \lambda(R/\mathfrak{m}^{[q]})$ for every $q = p^e$.*

Proof. By induction, it is enough to prove the case where $\text{height}(P) = \dim(R) - 1$. Notice it suffices to prove the second inequality.

Choose $f \in \mathfrak{m} - P$. Then, using the properties of the usual multiplicity of parameter ideals, the associativity formula for the usual multiplicity, we have, for all $q = p^e$,

$$\lambda(R/(P,f)^{[q]}) = \lambda(R/(P^{[q]}, f^q)) \tag{6}$$
$$\geq e(f^q; R/P^{[q]}) \tag{7}$$
$$= \lambda_{R_P}((R/P^{[q]})_P) \cdot e(f^q; R/P) \tag{8}$$
$$= \lambda_{R_P}((R/P^{[q]})_P) \cdot q \cdot \lambda(R/(f,P)). \tag{9}$$

By Corollary 2, we know that $\lambda(R/(f,P)) \cdot \lambda(R/\mathfrak{m}^{[q]}) \geq \lambda(R/(P,f)^{[q]})$. Hence $\lambda(R/\mathfrak{m}^{[q]}) \geq q \cdot \lambda_{R_P}((R/P^{[q]})_P)$ for every $q = p^e$. □

5 Hilbert–Kunz Multiplicity and Tight Closure

There is almost an exact parallel between the relationship of integral closure to the usual Hilbert–Samuel multiplicity and the relationship between tight closure to Hilbert–Kunz multiplicity. Just as in the case of the Hilbert–Samuel multiplicity, this relationship is important both theoretically and necessary to fully understand multiplicity. We use a key result of Aberbach [1] to make the proofs easier than the original proof in [31].

Let R^o denote the complement of the union of all minimal primes of a ring R. The definition of tight closure for ideals is:

Definition 1. Let R be a Noetherian ring of prime characteristic p. Let I be an ideal of R. An element $x \in R$ is said to be in the tight closure of I if there exists an element $c \in R^o$ such that for all large $q = p^e$, $cx^q \in I^{[q]}$.

There is also a definition of the tight closure of submodules of finitely generated R-modules, which we do not use in these notes. Of particular interest are rings in which every ideal is tightly closed.

Definition 2. A Noetherian ring in which every ideal is tightly closed is called *weakly F-regular*. A Noetherian ring R such that R_W is weakly F-regular for every multiplicative system W is called *F-regular*.

We list a few of the main properties satisfied by tight closure.

Proposition 3. *Let R be a Noetherian ring of prime characteristic p, and let I be an ideal:*

(1) $(I^*)^* = I^*$. If $I_1 \subseteq I_2 \subseteq R$, then $I_1^* \subseteq I_2^*$.
(2) If R is reduced or if I has positive height, then $x \in R$ is in I^* if and only if there exists $c \in R^o$ such that $cx^q \in I^{[q]}$ for all $q = p^e$.
(3) An element $x \in R$ is in I^* iff the image of x in R/P is in the tight closure of $(I+P)/P$ for every minimal prime P of R.

Proof. Part (1) and (2) follow immediately from the definition.

We prove (3). One direction is clear: if $x \in I^*$, then this remains true modulo every minimal prime of R since $c \in R^o$. Let $P_1, \dots, P_n$ be the minimal primes of R. If $c_i' \in R/P_i$ is nonzero, we can always lift c_i' to an element $c_i \in R^o$ by using the prime avoidance theorem. Suppose that $c_i' \in R/P_i$ is nonzero and such that $c_i' x_i^q \in I_i^{[q]}$ for all large q, where x_i (respectively I_i) represent the images of x (respectively I) in R/P_i. Choose a lifting $c_i \in R^o$ of c_i'. Then $c_i x^q \in I^{[q]} + P_i$ for every i. Choose elements t_i in all the minimal primes except P_i. Set $c = \sum_i c_i t_i$. It is easy to check that $c \in R^o$. Choose $q' \gg 0$ so that $N^{[q']} = 0$, where N is the nilradical of R. Then $cx^q \in I^{[q]} + N$, and so $c^{q'} x^{qq'} \in I^{[qq']}$, which proves that $x \in I^*$. □

One direction of our main result of this section is quite easy from the definition:

Proposition 4. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic p. Let I be an $\mathfrak{m}$-primary ideal, and suppose that $I \subseteq J \subseteq I^*$. Then $e_{HK}(I) = e_{HK}(J)$.*

Proof. By assumption there is an element $c \in R^o$ such that c annihilates the modules $J^{[q]}/I^{[q]}$ for all large $q = p^e$. These modules have a bounded number of generators, say t, given by the number of generators of J. In particular, $(R/(c, I^{[q]}))^t$ maps onto $J^{[q]}/I^{[q]}$, so that the length is at most $t \cdot \lambda(R/(c, I^{[q]}))$. However, the length of $R/(c, I^{[q]})$ is at most $O(q^{d-1})$ since the dimension of $R/(c)$ is $d - 1$. It follows that $|\lambda(R/J^{[q]}) - \lambda(R/I^{[q]})| = O(q^{d-1})$, and so $e_{HK}(I) = e_{HK}(J)$. □

The main result of this section is the following:

Theorem 5. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d and prime characteristic p which is formally unmixed. Let $I \subseteq J$ be $\mathfrak{m}$-primary ideals. Then $e_{HK}(I) = e_{HK}(J)$ if and only if $J \subseteq I^*$.*

Proof. One direction has already been done. To prove the other, we first observe that for $\mathfrak{m}$-primary ideals K, $e_{HK}(K) = e_{HK}(\widehat{K})$ and $(\widehat{K})^* = \widehat{K^*}$. We leave this latter equality as an exercise (see also [31, Proposition 4.14]). Hence we may assume that R is complete. Suppose that $e_{HK}(I) = e_{HK}(J)$. We need to prove that $J \subseteq I^*$. If not, there exists a minimal prime P of R such that the image of J in R/P is not in the tight closure of the image of I in R/P, by Proposition 3. By the additivity formula for Hilbert–Kunz multiplicity, Proposition 14, as well as our assumption that R is formally unmixed, we must have that $e_{HK}((I + P)/P) = e_{HK}((J + P)/P)$. Hence we may assume that R is a complete local domain.

Suppose by way of contradiction that J is not in I^*. We may assume that $J = (x, I)$ for some $x \notin I^*$. We now use a result of Aberbach [1]: since $x \notin I^*$, there exists a fixed integer k such that for all $q = p^e$, $I^{[q]} : x^q \subseteq \mathfrak{m}^{\lfloor q/k \rfloor}$. But now for all large enough q, $\lambda(R/I^{[q]}) - \lambda(R/(I^{[q]}, x^q)) = \lambda(R/(I^{[q]} : x^q)) \geq \lambda(R/\mathfrak{m}^{\rfloor q/k \lfloor}) \geq \delta q^d$, where δ is any positive real strictly less than $\frac{e(R)}{d!k}$, where $e(R)$ is the multiplicity of R. This proves that $e_{HK}(I) \neq e_{HK}(J)$, a contradiction. □

With this tight closure characterization of the Hilbert–Kunz multiplicity, we can give an important estimate on it in the case the ring is not F-rational, meaning that systems of parameters are not tightly closed. The is due to Blickle and Enescu [8], and later strengthened in [6, Corollary 3.5].

Proposition 6. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local unmixed ring of dimension d and prime characteristic p which is not F-rational. Assume that R is the homomorphic image of a Cohen–Macaulay ring. Set $e(R)$ equal to the multiplicity of R. Then $e_{HK}(R) \geq 1 + \frac{1}{e(R)-1}$.*

Proof. We may assume that the residue field is infinite. Choose a minimal reduction of the maximal ideal and let J be the ideal generated by that reduction. Since R is not F-rational, $J^* \neq J$. We use Lemma 1 to see that

$$e(R) = e_{HK}(J) = e_{HK}(J^*) \leq \lambda(R/J^*)e_{HK}(R) \leq (e(R) - 1)e_{HK}(R)$$

giving the result. Here the first equality is from the formula of Lech, [63, Theorem 11.2.10], the second is from the tight closure characterization of the Hilbert–Kunz multiplicity, and the third inequality is from Theorem 6. □

If $e(R) > d!$, then since $e_{HK}(R) \geq e(R)/d!$, we see that $e_{HK}(R) \geq 1 + \frac{1}{d!}$. On the other hand, if $e(R) \leq d!$, then $e(R) - 1 < d!$, and Proposition 6 shows that in the case R is not F-rational, we have the same estimate that $e_{HK}(R) \geq 1 + \frac{1}{d!}$.

Remark 7. It is worth noting that the relationship between the Hilbert–Kunz multiplicity of ideals and the tight closure was an important idea in the construction by Brenner and Monsky [15] of a counterexample to the localization problem in tight closure theory.

6 F-Signature

The work of Hochster and Roberts on the Cohen–Macaulayness of rings of invariants [30] focused attention on the splitting properties of the map from R to $R^{1/p}$. If R is F-finite, then this map splits as a homomorphism of R-modules if and only if R is F-pure, i.e., the Frobenius homomorphism is a pure map. Thus the idea of splitting copies of R out of $R^{1/p}$ clearly had something to say about the singularities of R. This idea was further explored during the development of tight closure, with the concept of strong F-regularity. In [62], Smith and Van den Bergh studied the asymptotic behavior of summands of $R^{1/q}$ for rings of finite F-representation type which are strongly F-regular. Yao [78] later removed the assumption of strong F-regularity from their work. For free summands, in [36], the idea of the F-signature was introduced as a way to asymptotically key track of the number such summands of $R^{1/q}$ as q varies. As it turns out, almost the exact same ideas were introduced at the same time by Watanabe and Yoshida [76] in their study of minimal relative Hilbert–Kunz multiplicity. The F-signature provides delicate information about the singularities of R, as we shall see. One immediate problem was to show that a limit exists in this asymptotic construction. When R is Gorenstein, this was done in [36], and we reproduce that argument here since it is not difficult and has the additional benefit of expressing the F-signature as a difference of the Hilbert–Kunz multiplicities of two ideals. The case when R is not Gorenstein proved to be considerably harder. After many partial results (see, e.g., [2, 79]) Kevin Tucker recently proved the limit always exists. We give a modified version of his proof here.

We first set up the basic ideas. Let $(R, \mathfrak{m}, k)$ be a d-dimensional reduced Noetherian local ring with prime characteristic p and residue field k. We assume that R is F-finite. By a_q we denote the largest rank of a free R-module appearing in a direct sum decomposition of $R^{1/q}$, where as usual $q = p^e$. We write $R^{1/q} \cong R^{a_q} \oplus M_q$ as an R-module, where M_q has no free direct summands. The number a_q is called the *eth Frobenius splitting number* of R.

Definition 1. The F-signature of R, denoted $s(R)$, is $s(R) = \varinjlim \frac{a_q}{q^{d+\alpha(R)}}$, the limit taken as q goes to infinity, provided the limit exists.

We first prove that the limit exists in the Gorenstein case, partly due to the ease of the proof, and partly due to the fact that it gives a precise value for the F-signature in terms of Hilbert–Kunz multiplicities. This theorem is found in [36].

Theorem 2. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local reduced Gorenstein ring of dimension d and prime characteristic p. Then $\varinjlim \frac{a_q}{q^{d+\alpha(R)}}$ exists and is equal to the difference between the Hilbert–Kunz multiplicity of the ideal I generated by a system of parameters, and the Hilbert–Kunz multiplicity of the ideal $I : \mathfrak{m}$.*

Proof. Let $I = (x_1, \ldots, x_d)$ be generated by a system of parameters. We claim that the difference $\lambda(M/IM) - \lambda(M/(I : \mathfrak{m})M)$ is zero for all maximal Cohen–Macaulay modules M without a free summand. We state this as a separate lemma.

Lemma 3. *Let $(R, \mathfrak{m})$ be a Gorenstein local ring and let M be a maximal Cohen–Macaulay R-module without a free summand. Let I be an ideal generated by a system of parameters for R, and let $\Delta \in R$ be a representative for the socle of R/I. Then $\Delta M \subseteq IM$.*

Proof. Choose generators $\{m_1, \ldots, m_n\}$ for M and define a homomorphism $R \to M^n$ by $1 \mapsto (m_1, \ldots, m_n)$. Let N be the cokernel, so that we have an exact sequence

$$0 \to R \to M^n \to N \to 0.$$

Since M has no free summands, this exact sequence is nonsplit. This implies, since R is Gorenstein, that N is not Cohen–Macaulay. When we kill I, therefore, there is a nonzero Tor:

$$0 \to \mathrm{Tor}_1^R(N, R/I) \to \overline{R} \to \overline{M}^n \to \overline{N} \to 0.$$

Since the map $\overline{R} \to \overline{M}^n$ has a nonzero kernel, we must have $\overline{\Delta} \mapsto 0$. Since the elements $m_1, \ldots, m_n$ generate M, this says precisely that $\Delta M \subseteq IM$. □

Returning to the proof of Theorem 2, we write $R^{1/q} = R^{a_q} \oplus M_q$, where M_q is a maximal Cohen–Macaulay module without free summands. Applying Lemma 3, we then see that $q^{\alpha(R)}(\lambda(R/I^{[q]}) - \lambda(R/(I, \Delta)^{[q]})) = a_q$ and therefore

$$e_{HK}(I, R) - e_{HK}((I, \Delta), R) = s(R). \qquad \square$$

Remark 4. The proof above shows that the F-signature of a Gorenstein local ring is 0 if and only if for some (or equivalently for all) ideals I generated by a system of parameters, $e_{HK}(I) = e_{HK}(I : \mathfrak{m})$. As we have seen, this equality holds if and only if I and $I : \mathfrak{m}$ have the same tight closure, which is true if and only if I is not tightly closed, since every ideal properly containing I must contain $I : \mathfrak{m}$. Thus the F-signature is positive in this case if and only if R is F-rational (and then is strongly F-regular, as R is Gorenstein.) Aberbach and Leuschke [3] proved in general that the F-signature is positive if and only if R is strongly F-regular. In fact the ideas of the proof above extend to prove something a little less than strong F-regularity, namely, that [36, Theorem 11] if the lim sup of a_q/q^d is positive, then R must be weakly F-regular, and in particular is Cohen–Macaulay and integrally closed. Thus, if R is not weakly F-regular, $s(R)$ exists and is 0. We prove this important fact next. For graded rings, it is known that strong and weak F-regularity are equivalent [44].

Remark 5. Watanabe and Yoshida [76] systematically studied minimal possible difference between the Hilbert–Kunz multiplicity of two $\mathfrak{m}$-primary ideals. They go further, and introduced the notion of minimal relative Hilbert–Kunz multiplicity $mHK(R)$. By their definition, $mHK(R) = \liminf \lambda_R(R/ann_R Rz^q)$, where z is a generator of the socle of the injective hull $E_R(k)$. They prove that $mHK(R) \leq e_{HK}(I) - e_{HK}(I')$ for $\mathfrak{m}$-primary ideals $I \subset I'$ with $\lambda_R(I'/I) = 1$. If R is Gorenstein, they prove the minimal relative Hilbert–Kunz multiplicity is in fact $e_{HK}(J) - e_{HK}(J : \mathfrak{m})$ for any parameter ideal J of R. As an example, we quote one of their theorems: Let k be a field of characteristic $p > 0$, and let $R = k[x_1, \ldots, x_d]^G$ be the invariant subring by a finite subgroup G of $GL(d, k)$ with $(p, |G|) = 1$. Also, assume that G contains no pseudo-reflections. Then the minimal relative Hilbert–Kunz multiplicity is $1/|G|$.

Lemma 6. *Assume that $(R, \mathfrak{m})$ is a reduced F-finite local ring containing a field of prime characteristic p, and let $d = \dim R$. We adopt the notation from the beginning of this section. If $s(R) > 0$, then R is weakly F-regular.*

Proof. Assume that $s(R) > 0$, but R is not weakly F-regular, that is, not all ideals of R are tightly closed. By [32, Theorem 6.1] R has a test element, and then [31, Proposition 6.1] shows that the tight closure of an arbitrary ideal in R is the intersection of $\mathfrak{m}$-primary tightly closed ideals. Since R is not weakly F-regular, there exists an $\mathfrak{m}$-primary ideal I with $I \neq I^*$. Choose an element Δ of $I : \mathfrak{m}$ which is not in I^*:

$$q^{\alpha(R)}(\lambda(R/I^{[q]}) - \lambda(R/(I,\Delta)^{[q]})) = \lambda(R^{1/q}/IR^{1/q}) - \lambda(R^{1/q}/(I,\Delta)R^{1/q}) \geq a_q.$$

Dividing by $q^{d+\alpha(R)}$ and taking the limit gives on the left-hand side a difference of Hilbert–Kunz multiplicities:

$$e_{HK}(I) - e_{HK}((I,\Delta)) \geq s(R).$$

But by Theorem 5, this difference is zero, showing that $s(R) = 0$. □

The beautiful idea of Tucker's proof that the F-signature exists in general is to represent it as a limit of certain normalized Hilbert–Kunz multiplicities, which are decreasing. To capture this, we first discuss some general facts about free summands of modules.

Discussion 6.7. Let $(R, \mathfrak{m})$ be a Noetherian local reduced ring, and let M be a torsion-free R-module. We can always write $M = N \oplus F$, where F is free and N has no free summands. We define a submodule M_{nf} of M to be $N + \mathfrak{m}F$. On the face of it, this submodule depends on the choice of N. However, we can also describe this submodule by the following:

$$\{x \in M \mid \phi(x) \in \mathfrak{m} \,\forall\, \phi \in \operatorname{Hom}_R(M, R)\}.$$

To see that these are the same, simply note that clearly M_{nf} is inside the above submodule (note it is a submodule!), and conversely, if x is in the submodule, then $x \in M_{nf}$; otherwise we can write $x = n + y$, where y is a minimal generator of F and where $n \in N$. The submodule Ry of M clearly splits off as a free summand, so there is a $\phi : M \longrightarrow R$ such that $\phi(y) = 1$. Then $\phi(x) = 1 + \phi(n) \notin \mathfrak{m}$, a contradiction. Note that M/M_{nf} is a vector space of dimension equal to the rank of F.

Definition 8. Let $(R, \mathfrak{m}, k)$ be a reduced local Noetherian ring of prime characteristic p. For $q = p^e$, we let $I_q := (R^{1/q})^{[q]}_{nf}$, an ideal in R.

This ideal was considered in work of Yongwei Yao [78] as well as Florian Enescu and Ian Aberbach [5]. Observe that Tucker defines it as follows, which from the discussion above is equivalent to our definition:

$$I_q = \{r \in R \mid \phi(r^{1/q}) \in \mathfrak{m} \,\forall\, \phi \in \operatorname{Hom}_R(R^{1/q}, R)\}.$$

We group some basic remarks about these ideals in the following proposition:

Proposition 9. *Let $(R, \mathfrak{m}, k)$ be a reduced local Noetherian ring of prime characteristic p. Then $\mathfrak{m}^{[q]} \subseteq I_q$ for all $q = p^e$. Furthermore, $I_q^{[q']} \subseteq I_{qq'}$ for all $q = p^e$ and $q' = p^{e'}$. If the residue field is perfect, $\lambda(R/I_q) = a_q$.*

Proof. Since $\mathfrak{m}R^{1/q} \subseteq (R^{1/q})_{nf}$, it is immediate from the definition that $\mathfrak{m}^{[q]} \subseteq I_q$. To prove the second statement, let $r \in I_q$, so that $r^{1/q} \in (R^{1/q})_{nf}$. Then $(r^{q'})^{1/qq'} =$

$r^{1/q} \in R^{1/qq'}$ is clearly $I_{qq'}$ by the second description of these ideals, since if $\phi : R^{1/qq'} \longrightarrow R$ was such that $\phi(r^{1/q}) \notin \mathfrak{m}$, restricting ϕ to $R^{1/q}$ would give the contradiction that $r \notin I_q$. The last statement of the proposition follows since $\lambda(R/I_q) = \lambda(R^{1/q}/I_q^{1/q}R^{1/q}) = \lambda(R^{1/q}/(R^{1/q})_{nf}) = a_q$. □

We are ready to prove Tucker's theorem:

Theorem 10 (Tucker [68, Theorem 4.9]). *Let* $(R, \mathfrak{m}, k)$ *be a Noetherian local ring of dimension* d *and prime characteristic* p. *Assume that* R *is F-finite. Then* $s(R) = \varinjlim \frac{a_q}{q^{d+\alpha(R)}}$ *exists.*

Proof. We can complete R and extend the residue field to assume that R is complete with perfect residue field. By Lemma 6 if R is not weakly F-regular, then $s(R) = 0$. Hence we may assume that R is weakly F-regular, and is in particular a Cohen–Macaulay domain. We use Corollary 8. We have that there is a constant C such that for all q, q',

$$|\lambda(R/I_q^{[q']}) - (q')^d \lambda(R/I_q)| \leq C(q')^d q^{d-1}.$$

Dividing by $(q')^d$ we obtain that

$$|\lambda(R/I_q^{[q']})/(q')^d - \lambda(R/I_q)| \leq Cq^{d-1}.$$

Taking the limit as q' goes to infinity, we see that

$$|e_{HK}(I_q) - a_q| \leq Cq^{d-1}.$$

Dividing by q^d shows that the F-signature exists if and only if the limit of $e_{HK}(I_q)/q^d$ exists. This follows by noting that $I_q^{[p]} \subseteq I_{qp}$ for all q, so that $e_{HK}(I_{qp}) \leq e_{HK}(I_q^{[p]}) = p^d e_{HK}(I_q)$, so that dividing through by qp shows that the sequence $\{e_{HK}(I_q)/q^d\}$ is decreasing, and thus has a limit, necessarily equal to $s(R)$. □

Example 11. We return to Example 18, where the Hilbert–Kunz multiplicity of simple quotient singularities were given. Let $(R, \mathfrak{m})$ be a two-dimensional complete Cohen–Macaulay ring. Assume that R is F-finite and is Gorenstein and F-rational. Then R is a double point and is isomorphic to $k[[x, y, z]]/(f)$, where f is one of the following:

type	equation	char R	$s(R)$	
(A_n)	$f = xy + z^{n+1}$	$p \geq 2$	$1/(n+1)$	$(n \geq 1)$
(D_n)	$f = x^2 + yz^2 + y^{n-1}$	$p \geq 3$	$1/4(n-2)$	$(n \geq 4)$
(E_6)	$f = x^2 + y^3 + z^4$	$p \geq 5$	$1/24$	
(E_7)	$f = x^2 + y^3 + yz^3$	$p \geq 5$	$1/48$	
(E_8)	$f = x^2 + y^3 + z^5$	$p \geq 7$	$1/120$	

As in Example 18, in each of these examples a minimal reduction J of the maximal ideal $\mathfrak{m}$ has the property that $\mathfrak{m}/J$ is a vector space of dimension 1. Hence $e_{HK}(J) - e_{HK}(R) = s(R)$ by Theorem 2. Since J is generated by a regular sequence and is a reduction of $\mathfrak{m}$, $e_{HK}(J) = e(J) = e(\mathfrak{m}) = 2$. On the other hand, Example 18 gives the Hilbert–Kunz multiplicity for each of these examples, and in each case it is $2 - 1/|G|$, where each ring is the invariant ring of a finite group G acting on a power series ring, giving our statement. Notice that the F-signature is exactly $1/|G|$. The same reasoning applies to Example 21 to show that if the Hilbert–Kunz multiplicity is irrational in this example, as expected, then so is the F-signature in the same example.

7 A Second Coefficient

In this section we take up a more careful study of the Hilbert–Kunz function, showing that a second coefficient exists in great generality. This was proved in [38], and further improved in [34]. The approach we give in this chapter is a bit different than those appearing elsewhere, following an alternate proof developed by Moira McDermott and myself, but not previously published. The proof in [38] relies on the theory of divisors associated to modules. The approach here rests on growth of Tor modules. In some ways this method is less transparent than that in [38], but this author believes it has considerable value nonetheless. We are aiming to prove:

Theorem 1. *Let $(R, \mathfrak{m}, k)$ be an excellent, local, normal ring of characteristic p with a perfect residue field and* $\dim R = d$. *Let I be an $\mathfrak{m}$-primary ideal. Then* $\lambda(M/I^{[q]}M) = \alpha q^d + \beta q^{d-1} + O(q^{d-2})$ *for some α and β in $\mathbb{R}$.*

In [34] the condition that R be normal is weakened to just assuming that R satisfies Serre's condition R_1.

One could hope that this theorem could be generalized to prove that there exists a constant γ such that $\lambda(M/I^{[q]}M) = \alpha q^d + \beta q^{d-1} + \gamma q^{d-2} + O(q^{d-3})$ whenever R is non-singular in codimension two. However, this will not be true. For instance, see Example 13.

We first discuss the growth of Tor modules, expanding on what we did in earlier sections.

Lemma 2. *Let $(R, \mathfrak{m}, k)$ be a local ring of characteristic p. If T is a finitely generated torsion R-module with* $\dim T = \ell$, *then* $\lambda(\mathrm{Tor}_1(T, R/I^{[q]})) \le O(q^{\ell})$.

Proof. Set $d = \dim R$. Choose a system of parameters $(x_1, \ldots, x_d) \subseteq I$. We induct on $\lambda(I/(x_1, \ldots, x_d))$. If $\lambda(I/(x_1, \ldots, x_d)) > 0$, then there exists $J \subset I$ with $\lambda(I/J) = 1$ so that we may write $I = (J, u)$ with $J : u = \mathfrak{m}$. For every $q = p^n$ there is an exact sequence:

$$0 \to R/J^{[q]}{:}u^q \to R/J^{[q]} \to R/I^{[q]} \to 0.$$

Tensor with T and look at the following portion of the long exact sequence:

$$\cdots \to \mathrm{Tor}_1(R/J^{[q]}, T) \to \mathrm{Tor}_1(R/I^{[q]}, T) \to \mathrm{Tor}_0(R/J^{[q]}{:}u^q, T) \to \cdots.$$

We have $\lambda(\mathrm{Tor}_1(R/J^{[q]}, T)) \leq O(q^\ell)$ by induction. Also, since $J{:}u = \mathfrak{m}$, we have $\mathfrak{m}^{[q]} \subseteq J^{[q]}{:}u^q$ and $\lambda(\mathrm{Tor}_0(R/J^{[q]}{:}u^q, T)) \leq \lambda(\mathrm{Tor}_0(R/\mathfrak{m}^{[q]}, T))$. But $\lambda(\mathrm{Tor}_0(R/\mathfrak{m}^{[q]}, T))$ is the Hilbert–Kunz function for T, so $\lambda(\mathrm{Tor}_0(R/\mathfrak{m}^{[q]}, T)) \leq O(q^\ell)$.

We have reduced to the case where $\lambda(I/(x_1, \ldots, x_d)) = 0$. We need a theorem which is implicitly in Roberts [57] and explicitly given as Theorem 6.2 in [32]:

Theorem 3. *Let (R, m) be a local ring of characteristic p and let $G_\bullet$ be a finite complex*

$$0 \to G_n \to \ldots \to G_0 \to 0$$

of length n such that each G_i is a finitely generated free module, and suppose that each $H_i(G_\bullet)$ has finite length. Suppose that M is a finitely generated R-module. Let $d = \dim M$. Then there is a constant $C > 0$ such that $\ell(H_{n-t}(M \otimes_R F^e(G_\bullet)) \leq C q^{\min(d,t)}$ for all $t \geq 0$ and all $e \geq 0$, where $q = p^e$.

Consider $K_\bullet((\underline{x}); R)$, the Koszul complex on $(x_1, \ldots, x_d)$. Let $H_\bullet((\underline{x}); R)$ denote the homology of the Koszul complex. We apply the above theorem to conclude that there exists a constant $C > 0$ such that $\lambda(H_{d-t}(T \otimes F^e(K_\bullet))) \leq C q^{\min\{\ell,t\}}$ for all t and for all e. Hence $\lambda(H_i(T \otimes F^e(K_\bullet))) \leq O(q^\ell)$ for all i. In general, $H_1(T \otimes F^e(K_\bullet)))$ maps onto $\mathrm{Tor}_1(T, R/I^{[q]}))$, which gives the stated result. ◻

Next we study the growth of Tor_2.

Lemma 4. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of dimension d satisfying Serre's condition S_2 and having prime characteristic p. Let T be an R-module with $\dim T \leq d-2$. Then $\lambda(\mathrm{Tor}_2(T, R/I^{[q]})) = O(q^{d-2})$.*

Proof. Pick a regular sequence x, y contained in the annihilator of T. There is an exact sequence:

$$0 \to T' \to (R/(x, y))^n \to T \to 0$$

Note $\dim T' = d-2$. Next tensor with $R/I^{[q]}$ and consider the following portion of the long exact sequence:

$$\cdots \to \mathrm{Tor}_2(R/(x, y), R/I^{[q]})^{\oplus n} \to \mathrm{Tor}_2(T, R/I^{[q]}) \to \mathrm{Tor}_1(T', R/I^{[q]}) \to \cdots.$$

Since x, y is regular sequence, we know $\sum_{i=0}^{2} \lambda(\mathrm{Tor}_i(R/(x, y), R/I^{[q]})) = 0$. Also, $\lambda(\mathrm{Tor}_1(R/(x, y), R/I^{[q]})) = O(q^{d-2})$ by Lemma 2. Then $\lambda(\mathrm{Tor}_2(R/(x, y), R/I^{[q]})) = O(q^{d-2})$ as well. We also know that $\lambda(\mathrm{Tor}_1(T', R/I^{[q]})) = O(q^{d-2})$ by Lemma 2. From the long exact sequence above, we conclude that $\lambda(\mathrm{Tor}_2(T, R/I^{[q]})) = O(q^{d-2})$. ◻

The main surprise is the next lemma, which shows that for the first Tor, modules which are torsion-free have slower growth than those which are torsion!

Lemma 5. *Let* $(R, \mathfrak{m}, k)$ *be a normal local ring of dimension* d *and prime characteristic* p*. Let* M *be a torsion-free* R*-module. Then* $\lambda(\operatorname{Tor}_1(M, R/I^{[q]})) = O(q^{d-2})$.

Proof. Consider the following exact sequence where $M^* = \operatorname{Hom}_R(M, R)$:

$$0 \to M \xrightarrow{\theta} M^{**} \to T \to 0.$$

Note that θ is an isomorphism in codimension one and consequently T is a torsion module with $\dim T \leq d-2$. We obtain the following long exact sequence:

$$\cdots \to \operatorname{Tor}_2(T, R/I^{[q]}) \to \operatorname{Tor}_1(M, R/I^{[q]}) \to \operatorname{Tor}_1(M^{**}, R/I^{[q]})$$
$$\to \operatorname{Tor}_1(T, R/I^{[q]}) \to \cdots .$$

From this we conclude that

$$\begin{aligned} &|\lambda(\operatorname{Tor}_1(M, R/I^{[q]})) - \lambda(\operatorname{Tor}_1(M^{**}, R/I^{[q]}))| \\ &\quad \leq \lambda(\operatorname{Tor}_2(T, R/I^{[q]})) + \lambda(\operatorname{Tor}_1(T, R/I^{[q]})) \\ &\quad = O(q^{d-2}). \end{aligned}$$

The last inequality follows from Lemmas 2 and 4. So we may replace M by M^{**} and assume that M has depth 2. Therefore, M is S_2 and M_P is free for all height one primes P.

We can choose a regular sequence x, y such that they kill all $\operatorname{Tor}_i^R(M,)$ for $i \geq 1$. This can be done in many ways. For example, we leave as an exercise that there exists a sequence, x, y, which is a regular sequence on R and on M such that multiplication by x on M factors through a free module $F = R^r$ and multiplication by y on M also factors through F. These multiplications then induce homotopies which can be used to prove our claim.

We let $\ldots \longrightarrow F_2 \longrightarrow F_1 \longrightarrow F_0 \longrightarrow M \longrightarrow 0$ be the start of a minimal free resolution of M. We tensor with $R/I^{[q]}$, and write $'$ for images after tensoring. Let Z_q be the kernel of the induced map from F_1' to F_0', and B_q be the image of the induced map from F_2' to F_1'. Thus, $\operatorname{Tor}_1(M, R/I^{[q]}) = Z_q/B_q$. Consider the short exact sequence:

$$0 \longrightarrow \operatorname{Tor}_1(M, R/I^{[q]}) \longrightarrow F_1'/B_q \longrightarrow N/I^{[q]}N \longrightarrow 0,$$

where N is the kernel of the map from F_0 onto M. We tensor with $R/(x, y)$ and use that both x and y annihilate $\operatorname{Tor}_1(M, R/I^{[q]})$ to see that the length of this Tor is at most $\lambda(\operatorname{Tor}_1(R/(x, y), N/I^{[q]}N)) + \lambda(F_1'/(B_q + (x, y)F_1')) \leq$

$\lambda(\mathrm{Tor}_1(R/(x,y), N/I^{[q]}N)) + \lambda(R/((x,y)+I^{[q]}))\cdot \mathrm{rank}(F_1)$. If we had the term $R/I^{[q]}$ in the first Tor module instead of $N/I^{[q]}N$, we could apply Lemma 2 directly to see the sum is $O(q^{d-2})$. We leave it to the reader to show that this change does not affect the order of growth. □

We record the following two corollaries to Lemma 5.

Corollary 6. *Let* $(R,\mathfrak{m},k)$ *be a local, normal ring of characteristic* p *with* $\dim R = d$. *Let* M *be a finitely generated* R*-module. Then for all* $i \geq 2$, $\lambda(\mathrm{Tor}_i(M, R/I^{[q]})) = O(q^{d-2})$.

Proof. Consider the exact sequence $0 \to \Omega^1(M) \to F \to M \to 0$ where F is free. Hence $\lambda(\mathrm{Tor}_i(M, R/I^{[q]})) \cong \lambda(\mathrm{Tor}_{i-1}(\Omega^1(M), R/I^{[q]}))$. It follows that to prove the lemma, we need only consider the case $i = 2$, and in this case since $\Omega^1(M)$ is torsion-free, the lemma above implies that $\lambda(\mathrm{Tor}_1(\Omega^1(M), R/I^{[q]})) = O(q^{d-2})$, giving that $\lambda(\mathrm{Tor}_2(M, R/I^{[q]})) = O(q^{d-2})$. □

The next corollary shows that $\lambda(\mathrm{Tor}_1(-, R/I^{[q]}))$ is additive on short exact sequences of torsion modules, up to $O(q^{d-2})$.

Corollary 7. *If* T_1, T_2, *and* T_3 *are torsion* R*-modules and* $0 \to T_1 \to T_2 \to T_3 \to 0$ *is exact, then* $|\sum_{i=1}^{3}(-1)^{i+1}\lambda(\mathrm{Tor}_1(T_i, R/I^{[q]}))| = O(q^{d-2})$.

Proof. After tensoring the exact sequence with $R/I^{[q]}$ we obtain the following long exact sequence:

$$\cdots \to \mathrm{Tor}_2(T_3, R/I^{[q]}) \to \mathrm{Tor}_1(T_1, R/I^{[q]}) \to \mathrm{Tor}_1(T_2, R/I^{[q]}) \to \mathrm{Tor}_1(T_3, R/I^{[q]})$$
$$\to \mathrm{Tor}_0(T_1, R/I^{[q]}) \to \mathrm{Tor}_0(T_2, R/I^{[q]}) \to \mathrm{Tor}_0(T_3, R/I^{[q]}) \to 0.$$

We examine the cokernel at one spot in the previous sequence. Consider

$$\to \mathrm{Tor}_2(T_3, R/I^{[q]}) \to \mathrm{Tor}_1(T_1, R/I^{[q]}) \to \mathrm{Tor}_1(T_2, R/I^{[q]})$$
$$\to \mathrm{Tor}_1(T_3, R/I^{[q]}) \to C \to 0.$$

We know that $\lambda(\mathrm{Tor}_2(T_3, R/I^{[q]})) = O(q^{d-2})$ by Corollary 6. It is therefore enough to show that $\lambda(C) = O(q^{d-2})$. We also have the exact sequence

$$0 \to C \to \mathrm{Tor}_0(T_1, R/I^{[q]}) \to \mathrm{Tor}_0(T_2, R/I^{[q]}) \to \mathrm{Tor}_0(T_3, R/I^{[q]}) \to 0.$$

Since the T_i are torsion modules, $\dim T_i \leq d-1$, and there are constants $c_i \geq 0$ such that $\lambda(\mathrm{Tor}_0(T_i, R/I^{[q]})) = c_i q^{d-1} + O(q^{d-2})$ so that

$$\lambda(C) = c_1 q^{d-1} - c_2 q^{d-1} + c_3 q^{d-1} + O(q^{d-2}).$$

But since the Hilbert–Kunz multiplicity is additive on short exact sequences, $c_2 = c_1 + c_3$, and hence $\lambda(C) = O(q^{d-2})$. □

The next result refines Lemma 2 by proving the existence of a coefficient giving the growth pattern.

Theorem 8. *Let* $(R, \mathfrak{m}, k)$ *be an excellent, local, normal ring of characteristic* p *with perfect residue field and with* $\dim R = d$. *Let* N *be a torsion* R*-module. Then there exists* $\gamma(N) \in \mathbb{R}$ *such that* $\lambda(\operatorname{Tor}_1(N, R/I^{[q]})) = \gamma(N)q^{d-1} + O(q^{d-2})$.

Proof. We may complete R and henceforth assume R is complete. Hence R is F-finite.

By Corollary 7, it is enough to prove the result for $N = R/Q$ where Q is a prime of R. If $\dim N \le d-2$, we know that $\lambda(\operatorname{Tor}_1(N, R/I^{[q]})) = O(q^{d-2})$ by Lemma 2 and $\lambda(\operatorname{Tor}_2(N, R/I^{[q]})) \le O(q^{d-2})$ by Lemma 4. Hence, it suffices to prove the case in which $N = R/Q$ where Q is a height one prime of R. Consider the following exact sequence:

$$0 \to (R/Q)^{p^{d-1}} \to (R/Q)^{1/p} \to T \to 0.$$

Tensor with $R/I^{[q]}$ and look at the following portion of the corresponding long exact sequence:

$$\begin{aligned}&\to \operatorname{Tor}_2(T, R/I^{[q]}) \to \operatorname{Tor}_1(R/Q, R/I^{[q]})^{p^{d-1}} \to \operatorname{Tor}_1((R/Q)^{1/p}, R/I^{[q]})\\ &\quad\to \operatorname{Tor}_1(T, R/I^{[q]}) \to .\end{aligned}$$

From this we see that

$$|p^{d-1}\lambda(\operatorname{Tor}_1(R/Q, R/I^{[q]})) - \lambda(\operatorname{Tor}_1((R/Q)^{1/p}, R/I^{[q]}))| = O(q^{d-2}). \quad (10)$$

Next consider the exact sequence $0 \to Q^{1/p} \to R^{1/p} \to (R/Q)^{1/p} \to 0$. First note that $\lambda(\operatorname{Tor}_1(R^{1/p}, R/I^{[q]})) = O(q^{d-2})$ by Lemma 5. From the usual long exact sequence on Tor we observe that

$$\begin{aligned}\lambda(\operatorname{Tor}_1((R/Q)^{1/p}, R/I^{[q]})) &\le \lambda(\operatorname{Tor}_0(Q^{1/p}, R/I^{[q]})) - \lambda(\operatorname{Tor}_0(R^{1/p}, R/I^{[q]}))\\ &\quad + \lambda(\operatorname{Tor}_0((R/Q)^{1/p}, R/I^{[q]})) + O(q^{d-2})\\ &\le \lambda(\operatorname{Tor}_0(Q, R/I^{[pq]})) - \lambda(\operatorname{Tor}_0(R, R/I^{[pq]}))\\ &\quad + \lambda(\operatorname{Tor}_0(R/Q, R/I^{[pq]})) + O(q^{d-2}).\end{aligned}$$

Now consider the sequence $0 \to Q \to R \to R/Q \to 0$. After tensoring with $R/I^{[pq]}$, it is clear from the usual long exact sequence that

$$\begin{aligned}\lambda(\operatorname{Tor}_1(R/Q, R/I^{[pq]})) &= \lambda(\operatorname{Tor}_0(Q, R/I^{[pq]})) - \lambda(\operatorname{Tor}_0(R, R/I^{[pq]}))\\ &\quad + \lambda(\operatorname{Tor}_0(R/Q, R/I^{[pq]})).\end{aligned}$$

Combining this with the previous inequality shows that

$$\lambda(\mathrm{Tor}_1((R/Q)^{1/p}, R/I^{[q]})) \leq \lambda(\mathrm{Tor}_1(R/Q, R/I^{[pq]})) + O(q^{d-2}).$$

Combining (10) and the previous inequality yields

$$p^{d-1}\lambda(\mathrm{Tor}_1(R/Q, R/I^{[q]})) - \lambda(\mathrm{Tor}_1(R/Q, R/I^{[pq]})) \leq O(q^{d-2}).$$

Recall that $q = p^e$. Define $\delta_q = \lambda(\mathrm{Tor}_1(R/Q, R/I^{[q]}))/q^{d-1}$. We claim that $\{\delta_q\}$ is a Cauchy sequence. We use the previous inequality to observe that

$$\begin{aligned}\delta_{pq} - \delta_q &= \lambda(\mathrm{Tor}_1(R/Q, R/I^{[pq]}))/(pq)^{d-1} - p^{d-1}\lambda(\mathrm{Tor}_1(R/Q, R/I^{[q]}))/p^{d-1}q^{d-1}\\ &= O(1/q).\end{aligned}$$

The sequence $\{\delta_q\}$ converges to some $\gamma(R/Q) \in \mathbb{R}$. A simple argument shows further that $|\delta_q - \gamma(R/Q)| = O(q^{-1})$. Hence $\lambda(\mathrm{Tor}_1(R/Q, R/I^{[q]})) = \gamma(R/Q)q^{d-1} + O(q^{d-2})$. □

Proposition 9. *Let $(R, \mathfrak{m}, k)$ be an excellent, local, normal ring of characteristic p with $\dim R = d$. Let M be a torsion-free R-module of rank r. Then there exists $\gamma(M) \in \mathbb{R}$ such that $\lambda(\mathrm{Tor}_0(M, R/I^{[q]})) - r\lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma(M)q^{d-1} + O(q^{d-2})$.*

Proof. We may complete R and henceforth assume R is complete. Since M is torsion-free of rank r as an R-module, we can choose an embedding $R^r \to M$ such that the cokernel T is a torsion module over R, and so $\dim T \leq d-1$. We have the following exact sequence: $0 \to R^r \to M \to T \to 0$. Tensor with $R/I^{[q]}$ and consider the usual long exact sequence:

$$\begin{aligned}0 &\to \mathrm{Tor}_1(M, R/I^{[q]}) \to \mathrm{Tor}_1(T, R/I^{[q]}) \to \mathrm{Tor}_0(R, R/I^{[q]})^{\oplus r}\\ &\to \mathrm{Tor}_0(M, R/I^{[q]}) \to \mathrm{Tor}_0(T, R/I^{[q]}) \to 0.\end{aligned}$$

We know that $\lambda(\mathrm{Tor}_1(M, R/I^{[q]})) = O(q^{d-2})$ by Lemma 5 and

$$\lambda(\mathrm{Tor}_1(T, R/I^{[q]})) = \gamma(T)q^{d-1} + O(q^{d-2})$$

by Theorem 8. Also, $\lambda(\mathrm{Tor}_0(T, R/I^{[q]}))$ is the Hilbert–Kunz function for T, and therefore there is a constant $C \geq 0$ such that $\lambda(\mathrm{Tor}_0(T, R/I^{[q]})) = Cq^{d-1} + O(q^{d-2})$. Thus,

$$\lambda(\mathrm{Tor}_0(M, R/I^{[q]})) - r\lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma(M)q^{d-1} + O(q^{d-2})$$

for some $\gamma(M) \in \mathbb{R}$. □

Corollary 10. *Let R be an excellent, local, normal ring of characteristic p with perfect residue field and* $\dim R = d$. *Then there exists $\gamma = \gamma(R^{1/p}) \in \mathbb{R}$ such that*

$$\lambda(\mathrm{Tor}_0(R, R/I^{[pq]})) - p^d \lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma q^{d-1} + O(q^{d-2}).$$

Proof. We complete R and assume it is complete. Then $R^{1/p}$ is a finitely generated R-module of rank p^d. Thus,

$$\lambda(\mathrm{Tor}_0(R^{1/p}, R/I^{[q]})) - p^d \lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma q^{d-1} + O(q^{d-2})$$

for some $\gamma \in \mathbb{R}$ by Proposition 9. As $\lambda(\mathrm{Tor}_0(R^{1/p}, R/I^{[q]})) = \lambda(\mathrm{Tor}_0(R, R/I^{[pq]}))$, we have

$$\lambda(\mathrm{Tor}_0(R, R/I^{[pq]})) - p^d \lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma(R^{1/p}) q^{d-1} + O(q^{d-2}). \quad \square$$

The next two theorems are the main content in [38]. As mentioned earlier, the approach in this chapter is through divisors attached to modules, rather than the growth of the length of Tor modules. See [43] for further analysis of the second coefficient.

Theorem 11. *Let $(R, \mathfrak{m}, k)$ be an excellent, local, normal ring of dimension d and prime characteristic p with a perfect residue field. Then there exists $\beta(R) \in \mathbb{R}$ such that $\lambda(R/I^{[q]}) = e_{HK}(R)q^d + \beta(R)q^{d-1} + O(q^{d-2})$.*

Proof. We may complete R and henceforth assume R is complete. Define $\epsilon_q := \lambda(R/I^{[q]}) - (\gamma(R^{1/p})/(p^{d-1} - p^d))q^{d-1}$. Recall that $q = p^e$. We claim that $\{\epsilon_q/q^d\}$ is a Cauchy sequence. Corollary 10 shows that $\epsilon_{pq} - p^d\epsilon_q = O(q^{d-2})$. Hence $|\epsilon_{pq}/(pq)^d - \epsilon_q/q^d| = O(q^{-2})$. The sequence $\{\epsilon_q/q^d\}$ converges to some $\alpha(R) \in \mathbb{R}$. Another simple geometric series argument shows that $|\epsilon_q/q^d - \alpha(R)| = O(q^{-2})$ and so $\epsilon_q = \alpha(R)q^d + O(q^{d-2})$. In other words, $\lambda(R/I^{[q]}) = \alpha(R)q^d + \beta(R)q^{d-1} + O(q^{d-2})$ where $\beta(R) = \gamma(R^{1/p})/(p^{d-1} - p^d)$. Clearly $\alpha(R) = e_{HK}(R)$ is forced. $\square$

Theorem 12. *Let $(R, \mathfrak{m}, k)$ be an excellent, local, normal ring of dimension d and prime characteristic p with a perfect residue field. Let M be finitely generated R-module. Then there exists $\beta(M) \in \mathbb{R}$ such that $\lambda(M/I^{[q]}M) = e_{HK}(M)q^d + \beta(M)q^{d-1} + O(q^{d-2})$.*

Proof. We may complete R and henceforth assume R is complete. Suppose M is a torsion-free R-module of rank r. We know that $\lambda(\mathrm{Tor}_0(M, R/I^{[q]})) - r\lambda(\mathrm{Tor}_0(R, R/I^{[q]})) = \gamma(M)q^{d-1} + O(q^{d-2})$ for some $\gamma(M) \in \mathbb{R}$ by Proposition 9. By Theorem 11 we know that $\lambda(R/I^{[q]}) = \alpha(R)q^d + \beta(R)q^{d-1} + O(q^{d-2})$. Combining these two results gives:

$$\lambda(\mathrm{Tor}_0(M, R/I^{[q]})) - r(\alpha(R)q^d + \beta(R)q^{d-1} + O(q^{d-2})) = \gamma(M)q^{d-1} + O(q^{d-2})$$

$$\lambda(\mathrm{Tor}_0(M, R/I^{[q]})) = r\alpha(R)q^d + (r\beta(R) + \gamma(M))q^{d-1} + O(q^{d-2}).$$

If M is not torsion-free, then we have the following exact sequence where $\overline{M}$ is torsion-free:

$$0 \to T \to M \to \overline{M} \to 0.$$

Tensor with $R/I^{[q]}$ and consider the usual long exact sequence

$$\cdots \to \mathrm{Tor}_1(\overline{M}, R/I^{[q]}) \to T/I^{[q]}T \to M/I^{[q]}M \to \overline{M}/I^{[q]}\overline{M} \to 0.$$

We know $\lambda(\mathrm{Tor}_1(\overline{M}, R/I^{[q]})) = O(q^{d-2})$ by Lemma 5. Also, $\lambda(T/I^{[q]}T) = e_{HK}(T)q^{\dim T} + O(q^{\dim T-1})$ and $\dim T \leq d-1$. Hence the result for M follows from the result for $\overline{M}$. □

8 Estimates on Hilbert–Kunz Multiplicity

In this section we discuss estimates of the Hilbert–Kunz multiplicity. A key motivating idea in this process was introduced in the paper of Blickle and Enescu [8] which proved that for rings which are not regular, the Hilbert–Kunz multiplicity is bounded away from 1 uniformly. This is the content of Proposition 7, which gives the lower bound of $1 + \frac{1}{p^d d!}$ for formally unmixed non-regular rings. However, it was felt that the presence of the characteristic p in the formula bounding the Hilbert–Kunz multiplicity away from 1 should not be necessary. Watanabe and Yoshida [77] made this explicit with the following conjecture:

Conjecture 1. Let $d \geq 1$ be an integer and $p > 2$ a prime number. Put $R_{p,d} := \overline{F_p}[[x_0, x_1, \ldots, x_d]]/(x_0^2 + \cdots + x_d^2)$. Let $(R, \mathfrak{m}, k)$ be a d-dimensional unmixed local ring with $k = \overline{F_p}$, an algebraic closure of the field with p-elements. Then the following statements hold:

(1) If R is not regular, then $e_{HK}(R) \geq e_{HK}(R_{p,d}) \geq 1 + a_d$, where a_d is the dth coefficient of the power series expansion of $\sec(x) + \tan(x)$ around 0.
(2) If $e_{HK}(R) = e_{HK}(R_{p,d})$, then the $\mathfrak{m}$-adic completion $\widehat{R}$ of R is isomorphic to $R_{p,d}$ as local rings.

There are several methods which have been used to estimate the Hilbert–Kunz multiplicity. Perhaps the most effective method is due to Watanabe and Yoshida, the method of estimation by computing volumes. Closely related ideas were also introduced by Hanes [29]. We illustrate this method in the simplest case where R is a Cohen–Macaulay local ring of dimension 2. Higher dimensional cases are of course more difficult, but the basic volume estimates are similar. The point is to estimate $l_A(\mathfrak{m}^{[q]}/J^{[q]})$ (where J is a minimal reduction of $\mathfrak{m}$) using volumes in $\mathbb{R}^d$. In a later paper, Watanabe and Yoshida use the methods, somewhat refined, to study higher dimension. In [77], they prove their conjecture up to dimension four. Aberbach and Enescu [7] have extended this by verifying the first part of the conjecture up to dimension six. Dimension seven is open as of the time this chapter was written.

We need the following lemma to prove Theorem 3. Just as in [73], it is convenient to adopt the following notation: if t is a real number, then $I^t := I^{\lfloor t \rfloor}$.

Lemma 2. *Let $(R, \mathfrak{m}, k)$ be an unmixed local ring of $\dim R = 2$, of prime characteristic p, with infinite residue field. Let J be a parameter ideal of R. Let $1 \leq s < 2$. Then we have the following limits:*

$$\lim_{q\to\infty} \frac{\lambda(R/J^{sq})}{q^2} = \frac{e(J)s^2}{2}, \qquad \lim_{q\to\infty} \lambda\left(\frac{J^{sq} + (J^*)^{[q]}}{J^{[q]}}\right) = e(J) \cdot \frac{(2-s)^2}{2}$$

Proof. We leave these for the reader as an exercise. The first follows from the usual Hilbert–Samuel multiplicity, while the second can be immediately reduced to the case in which R is a power series ring and the parameters are regular parameters. In this case the second limit can be thought of as computing a certain volume. We will describe the d-dimensional case after proving the theorem. □

Theorem 3 (Watanabe–Yoshida [73, Corollary]). *Let $(R, \mathfrak{m}, k)$ be a two-dimensional Cohen–Macaulay local ring of prime characteristic p. Put $e = e(R)$, the multiplicity of R. Then the following statements hold*:

(1) $e_{HK}(R) \geq \frac{e+1}{2}$.

(2) Suppose that $k = \overline{k}$. Then $e_{HK}(R) = \frac{e+1}{2}$ holds if and only if the associated graded ring $gr_{\mathfrak{m}}(R)$ is isomorphic to the Veronese subring $k[X, Y]^{(e)}$.

Proof. We will only prove the first statement. We claim that

$$e_{HK}(R) \geq \frac{r+2}{2r+2} e,$$

where e is the multiplicity of R, and r is the minimal number of generators of $\mathfrak{m}/J^*$. The theorem follows easily from this inequality, since the fact that $e \geq r - 1$ implies that $\frac{e+1}{2} \leq \frac{r+2}{2r+2} e$.

To prove the above claim, we let s be a real number, $1 \leq s < 2$. We may assume that the residue field is infinite, and we then choose a minimal reduction J of the maximal ideal. Note that $\lambda(\mathfrak{m}^{[q]}/(J^*)^{[q]}) = eq^2 - e_{HK}(R)q^2 + O(q)$, by the tight closure characterization of the Hilbert–Kunz multiplicity, Theorems 5 and 12.

We have the following:

$$\begin{aligned}
&\lambda(\mathfrak{m}^{[q]}/(J^*)^{[q]}) \\
&\quad \leq \lambda((\mathfrak{m}^{[q]} + \mathfrak{m}^{sq})/((J^*)^{[q]} + \mathfrak{m}^{sq})) + \lambda(((J^*)^{[q]} + \mathfrak{m}^{sq})/((J^*)^{[q]} + J^{sq}) \\
&\qquad + \lambda(((J^*)^{[q]} + J^{sq})/J^{[q]}).
\end{aligned}$$

The middle term in this sum is negligible, since J is a reduction of $\mathfrak{m}$, so that there is a fixed power of $\mathfrak{m}$ annihilating these modules, and the number of generators of a power of $\mathfrak{m}$ grows as $O(q)$. Hence the entire term is $O(q)$.

We prove that

$$\lambda((\mathfrak{m}^{[q]} + \mathfrak{m}^{sq})/((J^*)^{[q]} + \mathfrak{m}^{sq})) \le r \cdot \lambda(R/J^{(s-1)q}) + O(q).$$

By our assumption, we can write as $\mathfrak{m} = J^* + Ru_1 + \cdots + Ru_r$. Since $J^{(s-1)q}u_i^q \subseteq \mathfrak{m}^{sq} \subseteq \mathfrak{m}^{sq} + (J^*)^{[q]}$, we have

$$\lambda\left(\frac{\mathfrak{m}^{[q]} + \mathfrak{m}^{sq}}{(J^*)^{[q]} + \mathfrak{m}^{sq}}\right) \le \sum_{i=1}^{r} \lambda\left(R/((J^*)^{[q]} + \mathfrak{m}^{sq}) : u_i^q\right) \le r \cdot \lambda(R/J^{(s-1)q}).$$

Also, we have $\lambda((J^*)^{[q]}/J^{[q]}) = O(q^{d-1})$ by Theorem 5. Hence,

$$\lambda(\mathfrak{m}^{[q]}/(J^*)^{[q]}) \le r \cdot \lambda(R/J^{(s-1)q}) + \lambda\left(\frac{(J^*)^{[q]} + J^{sq}}{J^{[q]}}\right) + O(q).$$

Dividing by q^2 and letting q go to infinity, it follows from Lemma 2 that

$$e_{HK}(J) - e_{HK}(\mathfrak{m}) \le r \cdot e \cdot \frac{(s-1)^2}{2} + e \cdot \frac{(2-s)^2}{2}.$$

Setting $s = \frac{r+2}{r+1}$ proves the claim and finishes the proof of the theorem. □

The more general situation is as follows. We take the next discussion directly from [77]. For any positive real number s, we put

$$v_s := \mathrm{Vol}\left\{(x_1, \ldots, x_d) \in [0,1]^d \;\middle|\; \sum_{i=1}^{d} x_i \le s\right\}, \quad v_s' := 1 - v_s,$$

where $\mathrm{Vol}(W)$ denotes the volume of $W \subseteq \mathbb{R}^d$. With this notation, a key theorem in the work of Watanabe and Yoshida is the following:

Theorem 4. *Let $(R, \mathfrak{m}, k)$ be an unmixed local ring of characteristic $p > 0$. Put $d = \dim R \ge 1$. Let J be a minimal reduction of $\mathfrak{m}$, and let r be an integer with $r \ge \mu_R(\mathfrak{m}/J^*)$, where J^* denotes the tight closure of J. Also, let $s \ge 1$ be a rational number. Then we have*

$$e_{HK}(R) \ge e(R)\left\{v_s - r \cdot \frac{(s-1)^d}{d!}\right\}. \qquad (11)$$

This has been extended in [7].

Example 5 (cf. [17,73]). Let $(R, \mathfrak{m}, k)$ be a hypersurface local ring of characteristic $p > 0$ with $d = \dim R \ge 1$. Then

$$e_{HK}(R) \ge \beta_{d+1} \cdot e(R),$$

where β_{d+1} is given by the formula:

$$\mathrm{Vol}\left\{\underline{x} \in [0,1]^d \;\middle|\; \frac{d-1}{2} \le \sum x_i \le \frac{d+1}{2}\right\} = 1 - \nu_{\frac{d-1}{2}} - \nu'_{\frac{d+1}{2}}.$$

The first few values of β_{d+1}, beginning at $d = 0$, are the following: $1, 1, \frac{3}{4}, \frac{2}{3}, \frac{115}{192}$, and for $d = 5$, $\frac{11}{20}$.

Exercise 6 (Watanabe–Yoshida [73, Theorem (2.15)]). Let $(R, \mathfrak{m}, k)$ be a local ring of characteristic $p > 0$. Let $G = gr_{\mathfrak{m}}(R)$ the associated graded ring of R with respect $\mathfrak{m}$ as above. Then $e_{HK}(R) \le e_{HK}(G_{\mathfrak{M}}) \le e(R)$. Give an example to show that equality does not necessarily hold. (In fact, it seldom holds.)

Our final bounds rest on another technique, due to Aberbach and Enescu, as refined by Celikbas, Dao, Huneke, and Zhang, which allows one to give a uniform lower bound on the Hilbert–Kunz functions of non-regular rings. The basic idea of Aberbach and Enescu is to adjoin roots of elements in some fixed minimal reduction of the maximal ideal. In a bounded number of steps of such adjunctions, one reaches a ring which is not F-rational. In this case as we have seen, there are good lower bounds for the Hilbert–Kunz multiplicity. This reduces the problem to understanding the relationship between Hilbert–Kunz multiplicity of a ring and the ring adjoined some root. At this point the estimates in [18] are helpful. The first uniform bound was given in [6]:

Theorem 7 (Aberbach–Enescu). *Let $(R, \mathfrak{m}, k)$ be an unmixed ring of dimension $d \ge 2$ and prime characteristic p. If R is not regular, then*

$$e_{HK}(R) \ge 1 + \frac{1}{d(d!(d-1)+1)^d}.$$

This bound was improved in the paper [18] as we describe below. The essential new idea is in the following proposition:

Proposition 8. *Let R be a local Noetherian domain, and let $I = (J, u)$ where J is an integrally closed $\mathfrak{m}$-primary ideal of R and $u \in Soc(J)$. If M is a finitely generated torsion-free R-module, then*

$$\ell(IM/JM) \ge \operatorname{rank} M.$$

Proof. Set $N = (JM :_M u)$. Since $\dfrac{M}{N} \cong \dfrac{(J,u)M}{JM}$ and $\mathfrak{m}M \subseteq N$, we can write $M = N + N'$ with $\mu(N') = \ell\left(\dfrac{IM}{JM}\right)$. Thus it suffices to prove $\mu(N') \ge \operatorname{rank}(M)$. Since $u(M/N') \subseteq J(M/N')$, it follows from the determinantal trick [63, 2.1.8] that there is an element $r = u^n + j_1 \cdot u^{n-1} + \cdots + j_n$ with $j_i \in J^i$ for all i such that $rM \subseteq N'$. Observe that $r \ne 0$ since J is integrally closed and $u \notin J$. Since $M_r = N'_r$, this implies that $\mu(N') \ge \operatorname{rank}(N') = \operatorname{rank}(M)$. □

Given two ideals I and J with $J \subseteq I$, $\overline{\ell}(I/J)$ will denote the longest chain of integrally closed ideals between J and I.

Corollary 9. *Let R be a Noetherian local domain. Let J be an integrally closed $\mathfrak{m}$-primary ideal of R and let I be an ideal containing J. If M is a finitely generated torsion-free R-module, then*

$$\ell(IM/JM) \geq \overline{\ell}(I/J) \cdot \mathrm{rank}(M).$$

Proof. Set $n = \overline{\ell}(I/J)$. Then there is a chain of ideals

$$J = K_0 \subset K_1 \subset \ldots \subset K_{n-1} \subset K_n = I$$

with $\overline{K_i} = K_i$ for all i. Then

$$\ell\,(IM/JM) \geq \sum_{j=0}^{n} \ell(K_{j+1}M/K_jM) \geq \sum_{j=0}^{n} \ell((K_j, u_j)M/K_jM)$$

for some $u_j \in K_{j+1} \cap Soc(K_j)$. Thus the result follows from Proposition 8. ☐

One of the important ideas in proving that Hilbert–Kunz multiplicity equal to one implies regularity was showing an inequality $e_{HK}(I) \geq \lambda(R/I)$ for a suitable $\mathfrak{m}$-primary ideal I. Recall that we have equality if R is regular. This idea was developed in [73, 2.17], where the following questions were raised:

Let R be a Cohen–Macaulay local ring of characteristic $p > 0$. Then for any $\mathfrak{m}$-primary ideal I, do we have (1) $e_{HK}(I) \geq \ell(R/I)$? (2) If $pd_R(R/I) < \infty$, is $e_{HK}(I) = \ell(R/I)$?

The answer to both questions turns out to be negative; for example, see the paper of Kurano [42]. The next exercise shows that (1) is true for many $\mathfrak{m}$-primary ideals [18]:

Exercise 10. Assume R is an excellent normal ring with an algebraically closed residue field. If I is an integrally closed $\mathfrak{m}$-primary ideal of R, then

$$e_{HK}(I) \geq \ell(R/I) + e_{HK}(R) - 1.$$

If I is an $\mathfrak{m}$-primary ideal such that there is an integrally closed ideal $K \subset I$ with $\ell(I/K) = 1$, then

$$e_{HK}(I) \geq \ell(R/I).$$

(Hint: Use [72, 2.1] and Corollary 9.)

We turn to better uniform lower bounds for the Hilbert–Kunz multiplicity. An important point is the following, which we leave as an exercise (see [18]):

Exercise 11. Assume R is Cohen–Macaulay and normal, and let $x \in \mathfrak{m} - \mathfrak{m}^2$ be part of a minimal reduction of $\mathfrak{m}$. Let $S = R[y]$ with $y^n = x$. Then $\mathfrak{m}S + (y^i)$ is integrally closed for any nonnegative integer i.

Corollary 12. *Assume that* $(R, \mathfrak{m}, k)$ *is a Cohen–Macaulay normal local ring of prime characteristic* p *with infinite residue field. Let* $x \in \mathfrak{m} - \mathfrak{m}^2$ *be part of a minimal reduction of* $\mathfrak{m}$ *and let* $S = R[y]$ *with* $y^n = x$*. Then*

$$e_{HK}(R) - 1 \geq \frac{e_{HK}(S) - 1}{n}.$$

Proof. It follows from Proposition 11 and Corollary 9 that

$$e_{HK}(\mathfrak{m}S) \geq \ell(S/\mathfrak{m}S) + e_{HK}(S) - 1$$

Note that $S/\mathfrak{m}S \cong k[y]/(y^n)$. So $\ell(S/\mathfrak{m}S) = n$. Moreover, $e_{HK}(\mathfrak{m}S) = n \cdot e_{HK}(R)$ by Theorem 16. Therefore,

$$n \cdot e_{HK}(R) \geq n + e_{HK}(S) - 1$$

and hence the result follows. □

We can now give a rough lower bound on the Hilbert–Kunz multiplicity of a non-regular local ring, which depends only upon the dimension of the ring. This is an improvement of the bound of Aberbach and Enescu, Theorem 7.

Theorem 13. *Let* $(R, \mathfrak{m}, k)$ *be a formally unmixed Noetherian local ring of prime characteristic* p*, multiplicity* $e > 1$*, and dimension* d*. Then* $e_{HK}(R) \geq 1 + \frac{1}{d!d^d}$*.*

Proof. If $e_{HK}(R) \geq 1 + 1/d!$, there is nothing to prove. Hence we may assume that $e_{HK}(R) < 1 + 1/d!$, and then R is F-regular and Gorenstein by [6, 3.6] (see Proposition 6 as well). Thus we may assume that R is F-rational and Gorenstein.

Let $(\underline{x}) = (x_1, \cdots, x_d)$ be a minimal reduction of $\mathfrak{m}$. Consider the set of overrings $S = R[x_1^{1/n}, \ldots, x_i^{1/n}] = R_{i,n}$ which are not F-rational. Choose n and i such that we attain $\min\{n^i : R_{i,n} \text{ is not F-rational}\}$. Set $S = R_{i,n}$. Then by Proposition 6 applied to $x_1^{1/n}, \ldots, x_i^{1/n}, x_{i+1}, \ldots, x_d$,

$$e_{HK}(S) \geq \frac{e(S)}{e(S) - 1}.$$

However, since $S/(x_1^{1/n}, \ldots, x_i^{1/n}, x_{i+1}, .., x_d) \cong R/(\underline{x})$, we have $e(S) = e$. Therefore, $e_{HK}(S) \geq 1 + \frac{1}{e-1}$.

Let $R_0 = R$, and for each $i \geq j \geq 1$, let $R_j = R_{j-1}[x_j^{1/n}]$, then by Corollary 12,

$$e_{HK}(R_j) - 1 \geq \frac{e_{HK}(R_{j-1}) - 1}{n}.$$

Since $e - 1 < d!$, it remains to prove that

$$\min\{n^i : R_{i,n} \text{ is not F-regular}\} \leq d^d.$$

To do this we note that it suffices to prove that $R[x_1^{1/d}, \ldots, x_d^{1/d}]$ is not F-regular. Set $y_i = x_i^{1/d}$. Then a socle representative of $S/(\underline{x})$ is $u \cdot y_1^{d-1} \ldots y_d^{d-1}$, where u generates the socle of $(\underline{x}R)$. Let v be any discrete valuation centered on the maximal ideal of S. Then we claim that

$$v(u \cdot y_1^{d-1} \ldots y_d^{d-1}) \geq d v(\mathfrak{m}).$$

Since $v(u) \geq v(\mathfrak{m})$, this is clear.

It follows that $u \cdot y_1^{d-1} \ldots y_d^{d-1} \in \overline{(\mathfrak{m}S)^d}$. By the tight closure Briançon–Skoda theorem [31, Section 5] this implies that $(x_1, \ldots, x_d)S$ is not tightly closed, which gives the desired conclusion. □

Another approach, closely related to the volume methods of Watanabe and Yoshida, was given by Douglas Hanes in [29]. We close this survey with some of his results. See in particular [29, Theorem 2.4] and [29, Corollary 2.8].

Theorem 14. *Let $(R, \mathfrak{m}, k)$ be a Noetherian local ring of prime characteristic p, and dimension $d \geq 2$. Let I be an $\mathfrak{m}$-primary ideal, and set $t = \mu(I)$. Then,*

$$e_{HK}(I) \geq \frac{e(I)}{d!} \cdot \frac{t}{(t^{1/(d-1)} - 1)^{d-1}}.$$

Proof. We note that $I^{[q]} \subseteq I^q$ for all $q = p^e$ and $\mu(I^{[q]}) \leq t$ for all q. Hence, for all $q = p^e$ and any $s \in \mathbb{N}$, $\lambda((I^{[q]} + I^{q+s})/I^{q+s}) \leq t \cdot \lambda(R/I^s)$. Therefore, for all $q = p^e$ and any $s \in \mathbb{N}$, we see that

$$\lambda(R/I^{[q]}) \geq \lambda(R/(I^{[q]} + I^{q+s})) \geq \lambda(R/I^{q+s}) - t \cdot \lambda(R/I^s).$$

Just as in the work of Watanabe and Yoshida, the key point is to choose s carefully. Set $s = q\alpha$. We obtain that

$$\left(\frac{e(I)}{d!}\right)[(q + q\alpha)^d - t(q\alpha)^d] \leq \lambda(R/I^{[q]}) + O(q^{d-1}).$$

Ignoring the $O(q^{d-1})$ term and computing the maximal value of the function on the left-hand side of this equation, we obtain that a maximum is achieved when $\alpha = \frac{1}{(t^{1/(d-1)}-1)^{d-1}}$. The best lower bound for $e_{HK}(I)$ is obtained by setting $s = \lfloor \frac{q}{(t^{1/(d-1)}-1)^{d-1}} \rfloor$. Note that $s > 0$, since $t \geq d \geq 2$. We may write $s = q(\alpha - \epsilon)$ where $\epsilon < 1/q$. Applying the equations above with this value of s gives us that

$$\lambda(R/I^{[q]}) \geq \left(\frac{e(I)}{d!}\right) q^d [(1 + \alpha - \epsilon)^d - t(\alpha - \epsilon)^d] + O(q^{d-1}).$$

Dividing through by q^d, and letting q go to infinity (and ϵ toward 0), we obtain the estimate

$$e_{HK}(I) \geq \left(\frac{e(I)}{d!}\right)[(1+\alpha)^d - t(\alpha)^d],$$

from which the theorem follows. □

Corollary 15. *Let $(R, \mathfrak{m}, k)$ be a d-dimensional hypersurface ring of prime characteristic p, where $d \geq 3$. Then $e_{HK}(R) \geq e(R)2^{d-1}/d!$.*

Proof. Apply the previous theorem. Notice that the function $F(t) = \frac{t}{(t^{1/(d-1)}-1)^{d-1}}$ is decreasing, and $F(2^{d-1}) = 2^{d-1}$. As long as $\mu(\mathfrak{m}) \leq 2^{d-1}$ we can then apply the theorem. Since $\mu(\mathfrak{m}) \leq d+1$ and $d \geq 3$, the inequality holds. □

Acknowledgements The author was partially supported by NSF grant DMS-1063538. I thank them for their support.

References

1. Aberbach, I.M.: Extensions of weakly and strongly F-rational rings by flat maps. J. Algebra **241**, 799–807 (2001)
2. Aberbach, I.M.: The existence of the F-signature for rings with large Q-Gorenstein locus. J. Algebra **319**, 2994–3005 (2008)
3. Aberbach, I.M., Leuschke, G.: The F-signature and strong F-regularity. Math. Res. Lett. **10**, 51–56 (2003)
4. Aberbach, I.M., Enescu, F.: The structure of F-pure rings. Math. Z. **250**, 791–806 (2005)
5. Aberbach, I.M., Enescu, F.: When does the F-signature exist? Ann. Fac. Sci. Toulouse Math. **15**, 195–201 (2006)
6. Aberbach, I.M., Enescu, F.: Lower bounds for the Hilbert–Kunz multiplicities in local rings of fixed dimension. Michigan Math. J. **57**, 1–16 (2008)
7. Aberbach, I.M., Enescu, F.: New estimates of Hilbert–Kunz multiplicities for local rings of fixed dimension.
8. Blickle, M., Enescu, F.: On rings with small Hilbert–Kunz multiplicity. Proc. Amer. Math. Soc. **132**, 2505–2509 (2004)
9. Blickle, M., Schwede, K., Tucker, K.: F-signature of pairs and the asymptotic behavior of Frobenius splittings. Adv. Math. **231**, 3232–3258 (2012)
10. Blickle, M., Schwede, K., Tucker, K.: F-signature of pairs: Continuity, p-fractals and minimal log discrepancies. arXiv:1111.2762
11. Brenner, H.: A linear bound for Frobenius powers and an inclusion bound for tight closure. Michigan Math. J. **53**, 585–596 (2005)
12. Brenner, H.: The rationality of the Hilbert–Kunz multiplicity in graded dimension two. Math. Ann. **334**, 91–110 (2006)
13. Brenner, H.: The Hilbert–Kunz function in graded dimension two. Comm. Algebra **35**, 3199–3213 (2007)
14. Brenner, H., Li, J., Miller, C.: A direct limit for limit Hilbert–Kunz multiplicity for smooth projective curves. arXiv:1104.2662
15. Brenner, H., Monsky, P.: Tight closure does not commute with localization. Ann. Math. **171**, 571–588 (2010)

16. Bruns, W., Herzog, J.: Cohen–Macaulay rings. Cambridge Studies in Advanced Mathematics, vol. 39. Cambridge University Press, Cambridge (1993)
17. Buchweitz, R.-O., Chen, Q.: Hilbert–Kunz functions of cubic curves and surfaces. J. Algebra **197**, 246–267 (1997)
18. Celikbas, O., Dao, H., Huneke, C., Zhang, Y.: Bounds on the Hilbert–Kunz multiplicity. Nagoya Math. J. **205**, 149–165 (2012)
19. Conca, A.: Hilbert–Kunz function of monomial ideals and binomial hypersurfaces. Manuscripta Math. **90**, 287–300 (1996)
20. Dutta, S.: Frobenius and multiplicities. J. Algebra **85**, 424–448 (1983)
21. Enescu, F., Shimomoto, K.: On the upper semi-continuity of the Hilbert–Kunz multiplicity. J. Algebra **285**, 222–237 (2005)
22. Epstein, N., Yao, Y.: Some extensions of Hilbert–Kunz multiplicity. arXiv:1103.4730
23. Eto, K.: Multiplicity and Hilbert–Kunz multiplicity of monoid rings. Tokyo J. Math. **25**, 241–245 (2002)
24. Eto, K., Yoshida, K.: Notes on Hilbert–Kunz multiplicity of Rees algebras. Comm. Algebra **31**, 5943–5976 (2003)
25. Fakhruddin, N., Trivedi, V.: Hilbert–Kunz functions and multiplicities for full flag varieties and elliptic curves. J. Pure Appl. Algebra **181**, 23–52 (2003)
26. Gessel, I., Monsky, P.: The limit as $p \longrightarrow$ infinity of the Hilbert–Kunz multiplicity of sum$(x_i^{d_i})$. arXiv:1007.2004
27. Goto, S., Nakamura, Y.: Multiplicity and tight closures of parameters. J. Algebra **244**, 302–311 (2001)
28. Han, C., Monsky, P.: Some surprising Hilbert–Kunz functions. Math. Z. **214**, 119–135 (1993)
29. Hanes, D.: Notes on the Hilbert–Kunz function. J. Algebra **265**, 619–630 (2003)
30. Hochster, M., Roberts, J.: Rings of invariants of reductive groups acting on regular rings are Cohen–Macaulay. Adv. Math. **13**, 115–175 (1974)
31. Hochster, M., Huneke, C.: Tight closure, invariant theory and the Briançon–Skoda theorem. J. Amer. Math. Soc. **3**, 31–116 (1990)
32. Hochster, M., Huneke, C.: Phantom homology. Mem. Amer. Math. Soc. **103**(490) (1993)
33. Hochster, M., Huneke, C.: F-regularity, test elements, and smooth base change. Trans. Amer. Math. Soc. **346**, 1–62 (1994)
34. Hochster, M., Yao, Y.: Second coefficients of Hilbert–Kunz functions for domains. Preprint
35. Huneke, C.: Tight closure and its applications. With an appendix by Melvin Hochster. CBMS Regional Conference Series in Mathematics, vol. 88 (1996)
36. Huneke, C., Leuschke, G.: Two theorems about maximal Cohen–Macaulay modules. Math. Ann. **324**, 391–404 (2002)
37. Huneke, C., Yao, Y.: Unmixed local rings with minimal Hilbert–Kunz multiplicity are regular. Proc. Amer. Math. Soc. **130**, 661–665 (2002)
38. Huneke, C., McDermott, M., Monsky, P.: Hilbert–Kunz functions for normal rings. Math. Res. Lett. **11**, 539–546 (2004)
39. Kreuzer, M.: Computing Hilbert–Kunz functions of 1-dimensional graded rings. Univ. Iagel. Acta Math. **45**, 81–95 (2007)
40. Kunz, E.: Characterizations of regular local rings for characteristic p. Amer. J. Math. **91**, 772–784 (1969)
41. Kunz, E.: On Noetherian rings of characteristic p. Amer. J. Math. **98**, 999–1013 (1976)
42. Kurano, K.: On Roberts rings. J. Math. Soc. Japan **53**, 333–355 (2001)
43. Kurano, K.: The singular Riemann–Roch theorem and Hilbert–Kunz functions. J. Algebra **304**, 487–499 (2006)
44. Lyubeznik, G., Smith, K.E.: Strong and weak F-regularity are equivalent for graded rings. Amer. J. Math. **121**, 1279–1290 (1999)
45. McDonnell, L.: Hilbert–Samuel and Hilbert–Kunz functions of zero-dimensional ideals. Ph.D. Thesis, The University of Nebraska, Lincoln (2011)
46. Miller, L., Swanson, I.: Hilbert–Kunz functions of 2×2 determinantal rings. arXiv:1206.1015

47. Monsky, P.: The Hilbert–Kunz function. Math. Ann. **263**, 43–49 (1983)
48. Monsky, P.: The Hilbert–Kunz function of a characteristic 2 cubic. J. Algebra **197**, 268–277 (1997)
49. Monsky, P.: Hilbert–Kunz functions in a family: point-S_4 quartics. J. Algebra **208**, 343–358 (1998)
50. Monsky, P.: Hilbert–Kunz functions in a family: line-S_4 quartics. J. Algebra **208**, 359–371 (1998)
51. Monsky, P.: Hilbert–Kunz functions for irreducible plane curves. J. Algebra **316**, 326–345 (2007)
52. Monsky, P.: Rationality of Hilbert–Kunz multiplicities: a likely counterexample. Special volume in honor of Melvin Hochster. Michigan Math. J. **57**, 605–613 (2008)
53. Monsky, P.: Algebraicity of some Hilbert–Kunz multiplicities (modulo a conjecture). arXiv:0907.2470
54. Monsky, P.: Transcendence of some Hilbert–Kunz multiplicities (modulo a conjecture). arXiv:0908.0971
55. Monksy, P., Teixeira, P.: p-fractals and power series. I. Some 2 variable results. J. Algebra **280**, 505–536 (2004)
56. Monsky, P., Teixeira, P.: p-fractals and power series. II. Some applications to Hilbert–Kunz theory. J. Algebra **304**, 237–255 (2006)
57. Roberts, P.: Le théorème d'intersection. C. R. Acad. Sci. Paris Sér. I Math. **304**, 177–180 (1987)
58. Schwede, K., Tucker, K.: A survey of test ideals. Progress in Commutative Algebra 2, Closures, Finiteness and Factorization, pp. 39–99. Walter de Gruyter GmbH & Co. KG, Berlin (2012)
59. Seibert, G.: The Hilbert–Kunz function of rings of finite Cohen–Macaulay type. Arch. Math. **69**, 286–296 (1997)
60. Shepherd-Barron, N.I.: On a problem of Ernst Kunz concerning certain characteristic functions of local rings. Arch. Math. (Basel) **31**, 562–564 (1978/79)
61. Singh, A.K.: The F-signature of an affine semigroup ring. J. Pure Appl. Algebra **196**, 313–321 (2005)
62. Smith, K.E., Van den Bergh, M.: Simplicity of rings of differential operators in prime characteristic. Proc. London Math. Soc. **75**, 32–62 (1997)
63. Swanson, I., Huneke, C.: Integral Closure of Ideals, Rings, and Modules. Cambridge University Press, Cambridge (2006)
64. Teixeira, P.: Syzygy gap fractals I. Some structural results and an upper bound. J. Algebra **350**, 132–162 (2012)
65. Trivedi, V.: Semistability and Hilbert–Kunz multiplicities for curves. J. Algebra **284**, 627–644 (2005)
66. Trivedi, V.: Strong semistability and Hilbert–Kunz multiplicity for singular plane curves. Commutative algebra and algebraic geometry, pp. 165–173. Contemporary Mathematics, vol. 390. American Mathematical Society, Providence (2005)
67. Trivedi, V.: Hilbert–Kunz multiplicity and reduction mod p. Nagoya Math. J. **185**, 123–141 (2007)
68. Tucker, K.: F-signature exists. Inventiones
69. Upadhyay, S.: The Hilbert–Kunz function for binomial hypersurfaces. arXiv:1101.5936
70. Von Korff, M.: F-Signature of affine toric varieties. arXiv:1110.0552
71. Vraciu, A.: Drops in joint Hilbert–Kunz multiplicities and projective equivalence of ideals. Preprint (2012)
72. Watanabe, K.-i.: Chains of integrally closed ideals. In: Commutative algebra (Grenoble/Lyon, 2001). Contemporary Mathematics, vol. 331, pp. 353–358. American Mathematical Society, Providence (2003)
73. Watanabe, K.-i., Yoshida, K.-i.: Hilbert–Kunz multiplicity and an inequality between multiplicity and colength. J. Algebra **230**, 295–317 (2000)
74. Watanabe, K.-i., Yoshida, K.: Hilbert–Kunz multiplicity of two-dimensional local rings. Nagoya Math. J. **162**, 87–110 (2001)

75. Watanabe, K.-i., Yoshida, K.: Hilbert–Kunz multiplicity, McKay correspondence and good ideals in two-dimensional rational singularities. Manuscripta Math. **104**, 275–294 (2001)
76. Watanabe, K.-i., Yoshida, K.-i.: Minimal relative Hilbert–Kunz multiplicity. Illinois J. Math. **48**, 273–294 (2004)
77. Watanabe, K.-i., Yoshida, K.-i.: Hilbert–Kunz multiplicity of three-dimensional local rings. Nagoya Math. J. **177**, 47–75 (2005)
78. Yao, Y.: Modules with finite F-representation type. J. London Math. Soc. **72**, 53–72 (2005)
79. Yao, Y.: Observations on the F-signature of local rings of characteristic p. J. Algebra **299**, 198–218 (2006)

Pure O-Sequences: Known Results, Applications, and Open Problems

Juan Migliore, Uwe Nagel, and Fabrizio Zanello

1 Introduction

Pure O-sequences are fascinating objects that arise in several mathematical areas. They have been the subject of extensive research, yet our knowledge about pure O-sequences is limited. The goal of this note is to survey some of the known results and to motivate further investigations by pointing out connections to various interesting problems. Much of the material for this chapter has been drawn from the recent monograph [3] by the authors with Mats Boij and Rosa Miró–Roig, cited in this article as BMMNZ. We often quote results giving the name of the author(s) and the year the result was published so as to also hint at the history of ideas in the development of the theory of pure O-sequences.

The multifaceted interest in pure O-sequences is already indicated in their definition. On the one hand, a pure O-sequence can be defined as the vector whose entries record the number of monomials of a fixed degree in an order ideal generated by monomials of the same degree. On the other, a pure O-sequence is the Hilbert function of a finite-dimensional graded algebra that is level, i.e., its socle is concentrated in one degree, and has monomial relations. There is an

J. Migliore (✉)
Department of Mathematics, University of Notre Dame,
Notre Dame, IN 46556, USA
e-mail: migliore.1@nd.edu

U. Nagel
Department of Mathematics, University of Kentucky,
715 Patterson Office Tower, Lexington, KY 40506-0027, USA
e-mail: uwe.nagel@uky.edu

F. Zanello
Department of Mathematical Sciences, Michigan Technological University,
Houghton, MI 49931, USA
e-mail: zanello@math.mit.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_16,

extensive literature on monomial ideals and an extensive literature on level algebras. Pure O-sequences form a bridge between the two theories, and we will outline the work on pure O-sequences from this point of view. But more than that, these sequences have a broad array of applications and occur in many settings, largely combinatorial ones.

In Sect. 2, we review some of the basic results in the theory of pure O-sequences and focus on qualitative aspects of their shape. For instance, in some cases, pure O-sequences are known to be unimodal; that is, they are first weakly increasing, and once the peak is reached, they are weakly decreasing. However, pure O-sequences may fail to be unimodal, even with arbitrarily many "valleys". We include a discussion of recent results on conditions that force unimodality.

Connections to various combinatorial problems are the subject of Sect. 3. Face vectors of pure simplicial complexes are examples of pure O-sequences. In particular, the existence of certain block designs, such as Steiner systems, is related to that of some pure O-sequences. As a special case, the existence of finite projective planes is equivalent to the existence of particular pure O-sequences.

Another challenging problem is Stanley's conjecture that the h-vector of any matroid complex is a pure O-sequence. We discuss some recent progress. However, the conjecture remains open in general.

In Sect. 4, we describe results on the enumeration of pure O-sequences, with a focus on asymptotic properties. In particular, it follows that, when the number of variables is large, "almost all" pure O-sequences are unimodal.

We conclude this chapter with a collection of open problems, most of which are mentioned in the earlier sections.

2 Monomial Level Algebras

A finite, nonempty set X of (monic) monomials in the indeterminates $y_1, \ldots, y_r$ is called a *monomial order ideal* if whenever $M \in X$ and N is a monomial dividing M, then $N \in X$. The *h-vector* of X is defined to be the vector $\underline{h} = (h_0 = 1, h_1, \ldots, h_e)$ counting the number of monomials of X in each degree. A monomial order ideal, X, is called *pure* if all maximal monomials of X (in the partial ordering given by divisibility) have the same degree. A *pure O-sequence* is the h-vector of a pure monomial order ideal. For reasons that will be clear shortly, we call e the *socle degree* of $\underline{h}$. The *type* of a pure O-sequence is the number of maximal monomials.

Notice that if we think of $y_1, \ldots, y_r$ as the indeterminates of a polynomial ring $\mathcal{R} = K[y_1, \ldots, y_r]$ over a field, the question of whether a given sequence is or is not a pure O-sequence does not depend on the choice of K. Thus, for many of our results, it does not matter what we choose for K. However, in some situations, choosing a "nice" field K allows us to use special algebraic tools to say something about pure O-sequences, so in this case we make whatever additional assumptions we need for K.

We now let $R = K[x_1, \ldots, x_r]$, where K is an infinite field. We will consider standard graded artinian K-algebras $A = R/I$, where I will usually be a monomial ideal. Without loss of generality we will assume that I does not contain nonzero linear forms, so we will define r to be the *codimension* of A.

Let $\mathcal{R} = K[y_1, \ldots, y_r]$ and consider the action of R on monomials of $\mathcal{R}$ by contraction. By this we mean the action generated by

$$x_i \circ y_1^{a_1} y_2^{a_2} \cdots y_r^{a_r} = \begin{cases} y_1^{a_1} y_2^{a_2} \cdots y_i^{a_1-1} \cdots y_r^{a_r}, & \text{if } a_i > 0, \\ 0, & \text{if } a_i = 0. \end{cases}$$

For a monomial ideal $I \subset R$, we define the *inverse system* to be the R-module $I^{\perp} = \mathrm{ann}_{\mathcal{R}}(I) \subset \mathcal{R}$. One can check that $I^{\perp}$ consists of the monomials not in I (identifying x_i with y_i), and as such it can be viewed as a monomial order ideal. Recalling that for a standard graded algebra R/I, the *Hilbert function* is defined to be $h_{R/I}(t) = \dim[R/I]_t$, we observe that the h-vector (as defined above) of the order ideal $I^{\perp}$ coincides with the Hilbert function of R/I.

Furthermore, $I^{\perp}$ is a pure monomial order ideal if and only if R/I is a *level* algebra; that is, the socle of R/I (i.e., the annihilator of the homogeneous maximal ideal of R/I) is concentrated in one degree, called the *socle degree* of R/I; it is necessarily the degree of the maximal monomials of $I^{\perp}$. The dimension of the socle as a K-vector space is equal to the type of the pure O-sequence. See [25, 39] for more details on inverse systems.

Thus the study of pure O-sequences boils down to a study of the possible Hilbert functions of artinian monomial level algebras. In some cases we can rule out candidates for pure O-sequences by showing that there is not even a level algebra with that Hilbert function, but more often we need to use the structure of monomial algebras themselves.

The most basic tool is Macaulay's theorem to determine if the sequence is even an O-sequence, that is, to determine if it is the Hilbert function of *some* artinian algebra. We refer to [8, 47] for details of Macaulay's theorem, but we recall the statement. Let n and d be positive integers. There exist uniquely determined integers $k_d > k_{d-1} > \cdots > k_\delta \geq \delta \geq 1$ such that

$$n = n_{(d)} = \binom{k_d}{d} + \binom{k_{d-1}}{d-1} + \cdots + \binom{k_\delta}{\delta}.$$

This is called the *d-binomial expansion of n*. We set

$$(n_{(d)})_1^1 = \binom{k_d + 1}{d+1} + \binom{k_{d-1} + 1}{d} + \cdots + \binom{k_\delta + 1}{\delta + 1}$$

and $(0_{(d)})_1^1 = 0$, for each d. Then Macaulay's theorem is the following.

Theorem 1 (Macaulay [47]). *Let A be a standard graded algebra with Hilbert function $h_A(t) := h_t$. Then for all $t \geq 1$, $h_{t+1} \leq ((h_t)_{(t)})^1_1$.*

An *O-sequence* is a (possibly infinite) sequence of integers $(1, h_1, h_2, \dots)$ that satisfies the growth condition of Theorem 1 for every value of t. Thus the O-sequences are the sequences that occur as the Hilbert function of some standard graded algebra.

Example 2. The sequence $(1, 3, 6, 8, 8, 10)$ is not a pure O-sequence because it is not even an O-sequence (the growth from degree 4 to degree 5 is too big). Similarly, $(1, 3, 5, 5, 4, 4)$ is an O-sequence but it is not a pure O-sequence because it is not the Hilbert function of a *level* algebra (see [26]). Finally, $h = (1, 3, 6, 10, 15, 21, 28, 27, 27, 28)$ is the Hilbert function of a level algebra, but it is not a pure O-sequence [4] because there is no *monomial* level algebra with this Hilbert function. In fact, h has been the first nonunimodal level Hilbert function discovered in codimension 3 (see the third author [83]), and Boyle [4] has shown that this is in fact the smallest possible such Hilbert function.

So the challenge is to determine what additional conditions on an O-sequence are imposed by requiring that it be the Hilbert function of an *artinian level monomial* algebra. The first result, due originally to Stanley [67] with subsequent proofs given by Watanabe [78], Ikeda [65], Reid, Roberts, and Roitman [63], Herzog and Popescu [33], and Lindsey [45], concerns monomial complete intersections. (Note though that an equivalent property was proven earlier by de Bruijn, van Ebbenhorst Tengbergen, and Kruyswijk [21].) This result requires us to introduce here the notion of the Weak and Strong Lefschetz properties. The consequence of this result for pure O-sequences of type 1 is perhaps not so critical, as alternative proofs could be given. But its influence in the study of the Strong Lefschetz property (SLP) and the Weak Lefschetz property (WLP) in general and on related topics in commutative algebra, can hardly be overstated. In the result below and throughout the chapter, we will call a sequence *unimodal* if it is nondecreasing up to some degree and then nonincreasing past that degree. We will call it *strictly unimodal* if it is strictly increasing up to some degree, then possibly constant for some range, then strictly decreasing until it reaches zero, and then zero past that point. Unimodality is a central concept not only in combinatorics and combinatorial commutative algebra but also in other branches of mathematics. See for instance the classical surveys of Stanley ([68], 1989) and Brenti ([6], 1994).

Theorem 3 (see above for sources). *Let $R = K[x_1, \dots, x_r]$, where K has characteristic zero, and let I be an artinian monomial complete intersection, i.e.,*

$$I = \langle x_1^{a_1}, \dots, x_r^{a_r} \rangle.$$

Let L be a general linear form. Then for any positive integers d and i, the homomorphism induced by multiplication by L^d,

$$\times L^d : [R/I]_i \to [R/I]_{i+d},$$

has maximal rank. (In particular, this is true when $d = 1$.) As a consequence, a pure O-sequence of type 1 is strictly unimodal.

When I is an arbitrary artinian homogeneous ideal, the above maximal rank property for all d and i is called the *SLP*, and the case $d = 1$ is called the *WLP*. A consequence of the WLP is that in the range where $(\times L)$ is injective, we have a short exact sequence:

$$0 \to [R/I]_i \to [R/I]_{i+1} \to [R/(I, L)]_{i+1} \to 0.$$

Thus, the *first difference* $\Delta h_{R/I}(t) = h_{R/I}(t) - h_{R/I}(t-1)$ is the Hilbert function of a standard graded algebra in this range; that is, it is again an O-sequence. We then say that $h_{R/I}$ is a *differentiable O-sequence* in this range. It also follows from this sequence that if the WLP holds, then once the peak of the Hilbert function is reached, the Hilbert function will be nonincreasing, and therefore the whole Hilbert function is unimodal.

Remark 4. Simple examples show that Theorem 3 may fail in positive characteristic. It turns out that the question in which positive characteristics a monomial complete intersection has the WLP or SLP leads to unexpected connections to the problem of determining the number of certain plane partitions, lozenge tilings, or families of lattice paths (see [11, 16, 18, 19, 43]).

One of the early important results on pure O-sequences is due to Hibi [34].

Theorem 5 (Hibi [34]). *Let $\underline{h}$ be a pure O-sequence of socle degree e. Then*

$$h_i \le h_j$$

whenever $0 \le i \le j \le e-i$. This has the following two important consequences:

(a) $\underline{h}$ is flawless, i.e., $h_i \le h_{e-i}$ for all $0 \le i \le \lfloor \frac{e}{2} \rfloor$.
(b) The "first half" of $\underline{h}$ is nondecreasing:

$$1 = h_0 \le h_1 \le h_2 \le \cdots \le h_{\lfloor \frac{e}{2} \rfloor}.$$

This latter result was later improved by the following algebraic *g-theorem* of Hausel [32]:

Theorem 6 (Hausel [32]). *Let A be a monomial Artinian level algebra of socle degree e. If the field K has characteristic zero, then for a general linear form L, the induced multiplication*

$$\times L : A_j \to A_{j+1}$$

is an injection for all $j = 0, 1, \ldots, \lfloor \frac{e-1}{2} \rfloor$. In particular, over any field, the sequence

$$1, h_1 - 1, h_2 - h_1, \ldots, h_{\lfloor \frac{e-1}{2} \rfloor + 1} - h_{\lfloor \frac{e-1}{2} \rfloor}$$

is an O-sequence, i.e., the "first half" of $\underline{h}$ is a differentiable O-sequence.

We have the following additional results on differentiability from [3]:

Theorem 7 (Boij et al. [3]).

(a) Every finite differentiable O-sequence $\underline{h}$ is the "first half" of some pure O-sequence. (This is the converse of Hausel's theorem.)
(b) In particular, any finite differentiable O-sequence is pure (by truncation).
(c) Any nondecreasing pure O-sequence of socle degree ≤ 3 is differentiable.

It turns out that (c) is the best possible result in this direction:

Proposition 8 (Boij et al. [3]). *There exist nondecreasing pure O-sequences of any socle degree $e \geq 4$ that are not differentiable.*

Example 9. We illustrate the preceding result with an example from [3]. Observe first that the h-vector $\underline{h}' = (1, 4, 10, 20, 35)$ is a pure O-sequence since it is the h-vector of the truncation of a polynomial ring in four variables, w, x, y, z; the pure order ideal arises using all 35 monomials of degree 4 in w, x, y, z. The h-vector $\underline{h}'' = (1, 4, 6, 4, 1)$ is also a pure O-sequence, since it is the order ideal generated by a monomial $abcd$ in four new variables. Now we work in a polynomial ring in the eight variables w, x, y, z, a, b, c, d, and we consider the pure order ideal generated by the above 36 monomials of degree 4. The resulting h-vector h is

$$\begin{array}{rrrrrr} & 1 & 4 & 10 & 20 & 35 \\ + & & 4 & 6 & 4 & 1 \\ \hline & 1 & 8 & 16 & 24 & 36 \end{array}.$$

Since the first difference of $h = (1, 8, 16, 24, 36)$ is $(1, 7, 8, 8, 12)$, which is not an O-sequence (because $12 > (8_{(3)})^1_1 = 10$), we have constructed the desired example.

Putting aside the class of nondecreasing pure O-sequences, we now turn to the question of unimodality. There are three factors that go into whether a nonunimodal example will exist: the codimension (i.e., the number of variables), the socle degree, and the type. An easy application of Macaulay's theorem gives that any standard graded algebra of codimension two has unimodal Hilbert function, and in fact it is not hard to show that if $K[x, y]/I$ is level (monomial or not) then the Hilbert function is strictly unimodal (see, e.g., Iarrobino [38], 1984). Hence the interesting questions arise for codimension $r \geq 3$.

We will begin with some results involving the socle degree (some of which will also bring in the codimension). It follows from Hibi's result on flawlessness that any pure O-sequence of socle degree ≤ 3 is unimodal. The next case, socle degree 4, already is not necessarily unimodal, again thanks to an example from [3]:

Example 10. Observe that $\underline{h}' = (1, 5, 15, 35, 70)$ is a pure O-sequence, since it is the h-vector of the truncation of a polynomial ring in five variables, and as before $\underline{h}'' = (1, 4, 6, 4, 1)$ is also pure, since it corresponds to the maximal monomial

$abcd \in K[a,b,c,d]$. Hence, reasoning as above, we now consider one copy of $\underline{h}'$ and eleven copies of $\underline{h}''$ as h-vectors of pure O-sequences in twelve different rings, and we work in the tensor product of those rings. It follows that

$$\underline{h} = (1, 5, 15, 35, 70) + 11 \cdot (0, 4, 6, 4, 1) = (1, 49, 81, 79, 81)$$

is a nonunimodal pure O-sequence of socle degree 4.

A natural question, then, is what is the smallest codimension for which nonunimodal pure O-sequences exist with socle degree 4. This remains open. We do have the following results from [4], however:

Theorem 11 (Boyle [4]).

(a) All pure O-sequences of socle degree ≤ 9 in three variables are unimodal.
(b) All pure O-sequences of socle degree ≤ 4 in four variables are unimodal.
(c) In four or more variables, there exist nonunimodal pure O-sequences in all socle degrees ≥ 7.

In [2], Boij and the third author gave a nonunimodal pure O-sequence of codimension 3 and socle degree 12. This is the smallest known example in codimension 3. It follows from this and Boyle's result that in codimension 3, the only open cases are socle degrees 10 and 11. Notice that, in codimension 4, the previous theorem leaves only open the socle degrees 5 and 6.

Now we turn to questions involving the type. Of course it follows immediately from Theorem 3 that pure O-sequences of type 1 are unimodal. What else can be deduced about pure O-sequences using the WLP? A collection of results was obtained in [3] which showed, in some sense, the limits of the WLP in the study of pure O-sequences.

Theorem 12 (Boij et al. [3]). *Over a field of characteristic zero the following hold:*

(a) Any monomial artinian level algebra of type 2 in three variables has the WLP. Thus a pure O-sequence of type 2 and codimension 3 is differentiable until it reaches its peak, is possibly constant, and then is nonincreasing until it reaches zero.
(b) Fix two positive integers r and d. Then all monomial artinian level algebras of codimension r and type d possess the WLP if and only if at least one the following is true:
   (i) $r = 1$ or 2
   (ii) $d = 1$
   (iii) $r = 3$ and $d = 2$

The proof of (a) was surprisingly long and intricate. The main point of (b) is that in all other cases, we were able to show that artinian monomial level algebras exist that do *not* have the WLP.

Notice that Theorems 3, 6, and 12(a) require that K have characteristic zero. The statements about injectivity and surjectivity require this property of the characteristic, and indeed a great deal of research has been carried out to see what happens when K has positive characteristic; we refer to [52] for an overview of these results. The consequences on the shape of the pure O-sequences are indeed characteristic free, as has been noted above, since the Hilbert function of a monomial ideal does not depend on the characteristic.

If one is studying all artinian level monomial algebras of fixed type, Theorem 12 is a serious limitation on the usefulness of the WLP. However, in the study of pure O-sequences (e.g., to determine combinations of codimension, type, and socle degree that force unimodality) it is still conceivable that the WLP will play a useful role. As a trivial example, we know that monomial complete intersections in any codimension possess the WLP, thanks to Theorem 3. It is not known (except in codimension $\leq$ 3, as noted below) whether *all* complete intersections have this property, but the knowledge in the special case of monomial ideals is enough to say that all complete intersection Hilbert functions are unimodal. Perhaps a similar phenomenon will allow the WLP to continue to play a role in the study of pure O-sequences. A first approach using this philosophy was obtained by Cook and the second author [17], where they *lifted* a monomial ideal to an ideal of a reduced set of points in one more variable, showed that the general artinian reduction has the WLP, and concluded that the Hilbert function of the original monomial algebra is unimodal, regardless of whether it has the WLP or not.

For this reason we mention a useful tool in studying the WLP that was introduced by the first and second authors with Harima and Watanabe in [30] and whose study was continued by Brenner and Kaid in [5]. This is the study of the *syzygy bundle* and the use of the Grauert–Mülich theorem (see [58]). It was used in [30] to show that *any* complete intersection I in $K[x,y,z]$ over a field of characteristic zero has the WLP. The idea is to restrict to a general line, say one defined by a general linear form L. The key is that the restricted ideal $(I,L)/(L)$ (which now has codimension 2, hence has a Hilbert–Burch matrix) should have minimal syzygies in two consecutive degrees, at most. The idea of [30] was that for height 3 complete intersections, this information can be obtained by considering the module of syzygies of I, sheafifying it to obtain the syzygy bundle, and applying the Grauert–Mülich theorem to the general line defined by L. Of course this introduces questions about the semistability of the syzygy bundle, which we mostly omit here. In a more general setting, the following result from [5] summarizes the idea nicely, at least for codimension three.

Theorem 13 (Brenner–Kaid [5]). *Let $I = \langle f_1, \ldots, f_k\rangle \subset K[x,y,z] = R$ be an artinian homogeneous ideal whose syzygy bundle $\mathcal{S}$ is semistable on $\mathbb{P}^2$. Then*

(a) If the restriction of $\mathcal{S}$ splits on a general line L as

$$\mathcal{S}_L \cong \bigoplus_{i=1}^{s} \mathcal{O}_L(a+1) \oplus \bigoplus_{i=s+1}^{k-1} \mathcal{O}_L(a),$$

then $A = R/I$ has the WLP.

(b) If the restriction of $\mathcal{S}$ splits on a general line L as

$$\mathcal{S}_L \cong \mathcal{O}_L(a_1) \oplus \cdots \oplus \mathcal{O}_L(a_{n-1})$$

with $a_1 \geq a_2 \geq \cdots \geq a_{n-1}$ and $a_1 - a_{n-1} \geq 2$, then $A = R/I$ does not have the WLP.

Other applications of this approach, for higher codimension, can be found for instance in [55].

Returning to unimodality questions, tools other than the WLP will also be needed. We have the following theorems of Boyle [4], which relied on decomposition results and an analysis of complete intersection Hilbert functions but did not use the WLP.

Theorem 14 (Boyle [4]).

(a) In codimension 3, all pure O-sequences of type 3 are strictly unimodal.
(b) In codimension 4, all pure O-sequences of type 2 are strictly unimodal.

A natural question is whether all pure O-sequences of type 2 and arbitrary codimension are unimodal. Also, for any fixed codimension, it is an interesting problem to determine which types force unimodality. The "record" for the smallest known nonunimodal example in codimension 3 is type 14, given in [3].

Finally, one can ask "how nonunimodal" a pure O-sequence can be. The answer is "as nonunimodal as you want." We have

Theorem 15 (Boij et al. [3]). *For any integers $M \geq 2$ and $r \geq 3$, there exists a pure O-sequence in r variables which is nonunimodal and has exactly M maxima.*

Of course the "price" in Theorem 15 is paid in having a large socle degree and type.

In fact, even *Cohen–Macaulay f-vectors* (i.e., the face vectors of Cohen–Macaulay simplicial complexes, which are a much smaller subset of pure O-sequences) can be nonunimodal with arbitrarily many peaks (see [61]). This result considerably extends Theorem 15, even though, unlike for arbitrary pure O-sequences, here the number of variables becomes necessarily very large as the number of peaks increases.

A good topic to build a bridge between the algebraic and the combinatorial sides of the theory of pure O-sequences is the *Interval Conjecture* (ICP).

The *Interval Property* (IP) was introduced in 2009 by the third author [84], where he conjectured its existence for the set of Hilbert functions of level—and, in a suitably symmetric way, Gorenstein—algebras. Namely, the IP says that if two (not necessarily finite) sequences, h and h', of a class S of integer sequences coincide in all entries but one, say

$$h = (h_0, \ldots, h_{i-1}, h_i, h_{i+1}, \ldots) \text{ and } h' = (h_0, \ldots, h_{i-1}, h_i + \alpha, h_{i+1}, \ldots),$$

for some index i and some positive integer α, then the sequences

$$(h_0, \dots, h_{i-1}, h_i + \beta, h_{i+1}, \dots)$$

are also in S, for all $\beta = 1, 2, \dots, \alpha - 1$.

Given that level and Gorenstein Hilbert functions are nearly impossible to characterize, the IP appears to be both a very natural property and one of the strongest structural results that we might hope to achieve for the set of such sequences.

For example, it is proved in [84] that the IP holds for all Gorenstein Hilbert functions of socle degree 4. Since Gorenstein Hilbert functions are symmetric, this means that, for any fixed r, $(1, r, a, r, 1)$ is Gorenstein if and only if a ranges between some minimum possible value, say $f(r)$, and $\binom{r+1}{2}$. This latter is the maximum allowed by a polynomial ring in r variables and is achieved by the so-called *compressed* Gorenstein algebras (see, e.g., the 1984 papers of Fröberg–Laksov and Iarrobino [24, 38] or the third author's works [81, 82]).

Notice that the existence of the IP for these Hilbert functions is especially helpful in view of the fact that, for most codimensions r, the value of $f(r)$ is not known. In fact, such Gorenstein Hilbert functions are "highly" nonunimodal. Asymptotically, we have

$$\lim_{r \to \infty} \frac{f(r)}{r^{2/3}} = 6^{2/3},$$

as proved by these three authors in [53] (2008), solving a long-standing conjecture of Stanley [69] (see also [54], 2009, for some broad generalizations).

The Interval Property is still wide open today for both level and Gorenstein algebras.

In a more combinatorial direction, in BMMNZ 2012 the IP has then also been conjectured for 1) pure O-sequences (under the name "ICP") and 2) the f-vectors of pure simplicial complexes, a topic of discussion of the next section.

As for pure O-sequences, the ICP has been proved in a number of special cases. Most importantly, it is known when the socle degree is at most 3 in any number of variables.

Theorem 16 (Boij et al. [3]). *The ICP holds for the set of all pure O-sequences of socle degree $e \leq 3$.*

Thanks to this result, Hà, Stokes, and the third author [27] have recently developed a new approach leading to a proof of Stanley's matroid h-vector conjecture in Krull-dimension 3, as we will see in the next section.

While the ICP remains open in most instances—e.g., in three variables—it must be pointed out that, just recently, it has been *disproved* in the four-variable case by Constantinescu and Varbaro (see [15, Remark 1.10]), who found the following counterexample.

Example 17 (Constantinescu–Varbaro [15]). Consider the pure order ideals generated by $\{x^3y^2z, x^3yt^2, x^3z^2t\}$ and $\{x^4y^2, x^3yzt, x^2z^2t^2\}$. Their h-vectors are the pure O-sequences $(1, 4, 10, 13, 12, 9, 3)$ and $(1, 4, 10, 13, 14, 9, 3)$. However, an exhaustive computer search over all sets of three monomials of degree 6 in four variables reveals that the sequence $(1, 4, 10, 13, 13, 9, 3)$ is not pure, contrary to the ICP.

It is worth remarking that $h = (1, 4, 10, 13, 13, 9, 3)$ is, however, a level h-vector, and so this does not provide a counterexample to the IP for arbitrary level algebras. Indeed, h is the h-vector of a level algebra in four variables whose inverse system is generated by two sums of sixth powers of six general linear forms each and the sixth power of one general linear form.

As for pure f-vectors, the IP is still wide open, and little progress has been made so far.

In general, at this time, it is still unclear what the exact scope of the Interval Property is and if it can also be of use in other areas of combinatorial algebra or even enumerative combinatorics. It is well known to hold, e.g., for the set of Hilbert functions of graded algebras of any Krull-dimension (see Macaulay's theorem), the f-vectors of arbitrary simplicial complexes (the Kruskal–Katona theorem), and the f-vectors of Cohen–Macaulay complexes (BMMNZ 2012). Instead, the IP fails quite dramatically, for example, for matroid h-vectors, which are conjecturally another subset of pure O-sequences, as we will see in the next section (we refer to BMMNZ 2012 and [69] for details). Stanley and the third author [72] recently looked at the IP in the context of *r-differential posets*, a class of ranked posets generalizing the Young lattice of integer partitions and the Young–Fibonacci lattice. Here even though the IP fails in general, it might be a reasonable property to conjecture, for instance, for $r = 1$, which is the most natural class of differential posets.

3 Pure O-Sequences and Combinatorics

Much of the motivation for the study of pure O-sequences comes from combinatorics, and in this section we give an overview of this side of the theory. In order to put in context the definition of a pure O-sequence given in the introduction, we quickly recall the notion of posets and order ideals. For an introduction to this theory, we refer to Chapter 3 of Stanley's new edition of "EC1" ([71], 2012). A *poset* (short for *partially ordered set*) is a set S equipped with a binary relation, "$\leq$," that is (1) reflexive (i.e., $a \leq a$ for all $a \in S$), (2) antisymmetric ($a \leq b \leq a$ implies $a = b$), and(3) transitive ($a \leq b \leq c$ implies $a \leq c$).

An *order ideal* in a poset S is a subset I of S that is closed with respect to "$\leq$." That is, if $t \in I$ and $s \leq t$, then $s \in I$. Thus, our monomial order ideals are the (finite) order ideals of the poset P of all monomials in the polynomial ring $R = K[y_1, \ldots, y_r]$, where the binary relation of P is divisibility. Notice that P

is a *ranked* poset, where the rank of a monomial is its degree in R. Therefore, Macaulay's O-sequences are exactly the possible rank functions of the order ideals of P, since every Hilbert function satisfying Macaulay's theorem can be achieved by a monomial algebra. Similarly, as we have seen, a pure O-sequence is the rank function of some monomial order ideal whose *generators* (i.e., the antichain of maximal monomials) are all of the same degree.

Another fundamental class of order ideals are those contained in the *Boolean algebra* B_r, the poset of all subsets of $\{1, 2, \ldots, r\}$, ordered by inclusion. By identifying the integer i with a vertex v_i, the order ideals of B_r are usually called *simplicial complexes* (on r vertices).

The elements of a simplicial complex Δ are dubbed *faces*, and the maximal faces are the *facets* of Δ. The *dimension of a face* is its cardinality minus 1, and the *dimension of* Δ is the largest of the dimensions of its faces.

Notice that if we identify v_i with a variable $y_i \in R = K[y_1, \ldots, y_r]$, then simplicial complexes also coincide with the order ideals of P generated by *squarefree* monomials. In particular, if we define as *pure* those simplicial complexes whose facets have all the same dimension, then clearly their rank vectors, called *pure f-vectors*, are the special subset of pure O-sequences that can be generated by squarefree monomials.

Example 1. The simplicial complex

$$\Delta = \{\{1,2,3\}, \{2,3,4\}, \{1,2\}, \{1,3\}, \{2,3\}, \{2,4\}, \{3,4\}, \{1\}, \{2\}, \{3\}, \{4\}, \emptyset\}$$

is the order ideal of B_4 generated by $\{1, 2, 3\}$ and $\{2, 3, 4\}$.

Thus, Δ is a pure complex of dimension 2, whose pure f-vector is $f_\Delta = (1, 4, 5, 2)$.

Equivalently, f_Δ is the pure O-sequence generated by the two squarefree monomials $y_1 y_2 y_3$ and $y_2 y_3 y_4$.

Similarly to Macaulay's theorem for arbitrary O-sequences, we know a characterization of the class of pure f-vectors thanks to the classical *Kruskal–Katona theorem* (see, e.g., [69]). However, analogously, things become dramatically more complicated (hopeless, we should say) when it comes to attempting a characterization of *pure* f-vectors.

In the last section of BMMNZ 2012, we have begun a study of pure f-vectors, but still very little is known today beyond what is known for arbitrary pure O-sequences.

Besides their obvious intrinsic importance—simplicial complexes are a central object in algebraic combinatorics, combinatorial algebra, and topology, just to name a few subjects—pure f-vectors also carry fascinating applications. It is on their connections to finite geometries and design theory that we want to focus in the next portion of this section.

It will follow from our discussion, as probably first observed by Björner ([1], 1994), that a characterization of pure simplicial complexes and their f-vectors would imply, for instance, a characterization of all Steiner systems and, as a further special case, a classification of all finite projective planes, which is one of the major open problems in geometry.

A *Steiner system* $S(l, m, r)$ is an r-element set V, together with a collection of m-subsets of V, called *blocks*, such that every l-subset of V is contained in exactly one block. Steiner systems are a special family of the so-called *block designs*. We refer our reader to the two texts [14, 44], where she can find a truly vast amount of information on combinatorial designs. For instance, a Steiner *triple* system (STS) is a Steiner system $S(2, 3, r)$, while $S(3, 4, r)$ is dubbed a Steiner *quadruple* system, where r is the *order* of the system.

Since we are dealing with maximal sets of the same cardinality (m, in this case), it is clear that if we identify each element of V with a variable y_i, then the existence of Steiner systems (and similarly for other block designs) will be equivalent to the existence of certain pure f-vectors.

Example 2. Let us consider STS's of order 7, i.e., $S(2, 3, 7)$. Constructing such a design is tantamount to determining a family of squarefree degree 3 monomials of $R = K[y_1, y_2, \dots, y_7]$, say $M_1, M_2, \dots, M_t$, such that each squarefree degree 2 monomial of R divides exactly one of the M_i.

Clearly, since there are $\binom{7}{2} = 21$ squarefree degree 2 monomials in R, if the M_i exist, then $t = 21/\binom{3}{2} = 7$. In other words, $S(2, 3, 7)$ exists if and only if

$$f = (1, 7, 21, 7)$$

is a pure f-vector.

Notice also that f exists as a pure f-vector if and only if it exists as a pure O-sequence, since for seven degree 3 monomials to have a total of 21 degree 2 divisors, each needs to have exactly three linear divisors, i.e., it must be squarefree.

It is easy to see that an STS of order 7 and so the pure f-vector $f = (1, 7, 21, 7)$ do indeed exist using the monomial order ideal generated by

$$y_1y_2y_3,\ y_3y_4y_5,\ y_3y_6y_7,\ y_1y_4y_7,\ y_2y_4y_6,\ y_2y_5y_7,\ y_1y_5y_6.$$

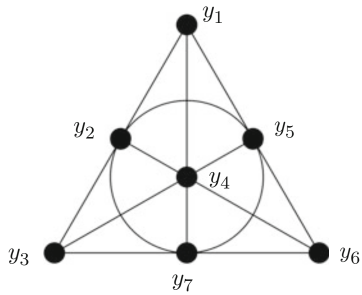
Some simple numerical observations show that a necessary condition for an STS of order r to exist is that r be congruent to 1 or 3 modulo 6, and it is a classical result of Kirkman ([40], 1847) that this is also sufficient. In other words,

$$f = \left(1, r, \binom{r}{2}, \binom{r}{2}/3\right)$$

is a pure f-vector, if and only if it is a pure O-sequence, if and only if r is congruent to 1 or 3 modulo 6.

A different and more challenging problem is the classification of all Steiner systems, up to isomorphism. Even the existence of particular systems sometimes brings into the story a nontrivial amount of interesting algebra. As an illustration, we mention here the case of the Steiner systems $S(4,5,11)$, $S(5,6,12)$, $S(3,6,22)$, $S(4,7,23)$, and $S(5,8,24)$, which are intimately connected to the first sporadic finite simple groups ever discovered, called the *Mathieu groups* (see Mathieu [48, 49], 1861 and 1873). These five groups—denoted respectively by M_{11}, M_{12}, M_{22}, M_{23}, and M_{24}—in fact arise as the *automorphism groups* of the above Steiner systems (i.e., the transformations of the systems that preserve the blocks).

There exists only one STS of order 7, which is called the *Fano plane* (see the figure below) for reasons that will be clear in a minute. In other words, the seven monomials of the previous example are, up to isomorphism, the only possible set of generators for a pure order ideal in $K[y_1, y_2, \ldots, y_7]$ with $(1, 7, 21, 7)$ as its pure O-sequence.



Also for $r = 9$, there exists a unique STS. However, it is reasonable to believe that the number of nonisomorphic STS increases extremely quickly for r large. For instance, there are 80 nonisomorphic STS of order 15, and there are 11,084,874,829 of order 19.

Similarly, the possible orders of Steiner quadruple systems are known and nicely characterized, since the obvious necessary conditions again turn out to be also sufficient: $S(3,4,r)$ exists if and only if its order r is congruent to 2 or 4 modulo 6, as proved by Hanani ([28], 1960). In other words, reasoning as above, since there are four possible 3-subsets of any given 4-set, we have that

$$f = \left(1, r, \binom{r}{2}, \binom{r}{3}, \binom{r}{3}/4\right)$$

is a pure f-vector if and only if it is a pure O-sequence, if and only if r is congruent to 2 or 4 modulo 6.

However, as we increase the cardinality m of the blocks, things become more and more obscure. This is due to the high complexity of computing combinatorial designs over a large vertex set, as well as to the lack of a general theory.

Already for Steiner quintuple systems, the trivial necessary conditions (r congruent to 3 or 5, but not to 4, modulo 6) are no longer sufficient. For example, no Steiner quintuple system $S(4,5,17)$ exists (see [59], 2008). The smallest value of

r for which the existence of $S(4,5,r)$ is currently open is 21. In other words, it is unknown whether

$$\left(1, 21, \binom{21}{2}, \binom{21}{3}, \binom{21}{4}, \binom{21}{4}/5\right) = (1, 21, 210, 1330, 5985, 1197)$$

is a pure O-sequence.

Perhaps the best-known family of examples of Steiner systems is that of finite projective planes, so they deserve a special mention here. Recall that a *projective plane* is a collection of points and lines such that any two lines "intersect at" exactly one point and any two points "lie on" exactly one line (one also assumes that there exist four points no three of which are collinear, in order to avoid uninteresting pathological situations).

If the projective plane is finite, it can easily be seen that the number of points is equal to the number of lines and that this number must be of the form $q^2 + q + 1$, where the integer q is the *order* of the plane. Further, in a projective plane of order q, any line contains exactly $q + 1$ points, and by duality, any point is at the intersection of exactly $q + 1$ lines. In other words, finite projective planes are the Steiner systems $S(2, q+1, q^2+q+1)$. The reader may want to consult, e.g., [23] for an introduction to this area.

Thus, the above example of a Steiner system $S(2,3,7)$, the Fano plane, is the unique smallest possible projective plane.

Similarly to how we argued earlier in terms of pure f-vectors, one can show that a projective plane of order q exists if and only if

$$h = \left(1, q^2+q+1, (q^2+q+1)\binom{q+1}{2}, (q^2+q+1)\binom{q+1}{3}, \dots, (q^2+q+1)\binom{q+1}{q+1}\right)$$

is a pure f-vector, if and only if it is a pure O-sequence.

A major open problem in geometry asks for a classification of all finite projective planes, or even just of the possible values that q may assume. Conjecturally, q is always the power of a prime, and it is a standard algebraic exercise, using finite field theory, to construct a projective plane of any order $q = p^n$.

The best general necessary condition known today on q is still the following theorem from [7, 13]:

Theorem 3 (Bruck–Ryser–Chowla, 1949 and 1950). *If q is the order of a projective plane and q is congruent to 1 or 2 modulo 4, then q is the sum of two squares.*

Thus, for instance, as a consequence of the Bruck–Ryser–Chowla theorem, no projective plane of order 6 exists. In other words,

$$(1, 43, 903, 1505, 1505, 903, 301, 43)$$

is not a pure O-sequence. However, already ruling out the existence of projective planes of order 10 has required a major computational effort (see Lam [41], 1991). The case $q = 12$ is still open.

Notice that, at least for certain values of q, the number of nonisomorphic projective planes of order q can be very large, and a general classification seems entirely out of reach. The smallest q for which there exists more than one nonisomorphic projective plane is 9, where the four possible cases were already known to Veblen ([77], 1907).

The second important application of pure O-sequences that we want to discuss brings our attention to a very special class of simplicial complexes called *matroid* complexes. Matroids are ubiquitous in mathematics, where they often show up in surprising ways (see [56, 60, 79, 80]).

The algebraic theory of matroids began in the same 1977 seminal paper of Stanley [66] that introduced pure O-sequences. A finite matroid can be naturally identified with a pure simplicial complex over $V = \{1, 2, \ldots, r\}$, such that its restriction to any subset of V is also a pure complex.

One associates, to any given simplicial complex Δ over $V = \{1, 2, \ldots, r\}$, the following squarefree monomial ideal in $S = K[x_1, \ldots, x_r]$, where K is a field:

$$I_\Delta = \left\langle x_F = \prod_{i \in F} x_i \mid F \notin \Delta \right\rangle.$$

I_Δ is called the *Stanley–Reisner ideal* of the complex Δ, and the quotient algebra S/I_Δ is its *Stanley–Reisner ring*.

It is a standard fact of combinatorial commutative algebra (see, e.g., [69]) that the Stanley–Reisner ring S/I_Δ of a matroid complex is Cohen–Macaulay and level, although of course of positive Krull-dimension (except in degenerate cases). Thus, the h-vector of S/I_Δ is level, since it is the h-vector of an artinian reduction of S/I_Δ. However, even though S/I_Δ is presented by monomials, notice that its artinian reductions will in general be far from monomial for they require taking quotients by "general enough" linear forms.

The following spectacularly simple conjecture of Stanley ([66], 1977) predicts that, for any matroid complex Δ, we can nonetheless find *some* artinian monomial level algebra having the h-vector of S/I_Δ as its h-vector:

Conjecture 3.4. *Any matroid h-vector is a pure O-sequence.*

The problem of characterizing matroid h-vectors appears to be once again hopeless, and Conjecture 3.4 has motivated much of the algebraic work done on matroids over the past 35 years (see, as a highly nonexhaustive list, [9, 10, 15, 22, 31, 50, 57, 62, 64, 73–75]).

The main approach to Conjecture 3.4 has been, given the h-vectors of a certain class of matroids, to explicitly produce some pure monomial order ideals having those matroid h-vectors as their pure O-sequences. Recently, Hà, Stokes, and the third author [27] introduced a "more abstract" approach to Conjecture 3.4. Their main idea, inspired by the latest progress on pure O-sequences made in BMMNZ 2012 and in particular the proof of the ICP in socle degree 3, has been to try to reduce Stanley's conjecture, as much as possible, to one on the properties of pure O-sequences, thus avoiding explicit construction of a monomial ideal for each matroid h-vector.

The approach of [27] has already led to a proof of Conjecture 3.4 for all matroid complexes of Krull-dimension at most 2 (the dimension 1 case that had been the focus of a large portion of the thesis [73] simply followed in a few lines).

Theorem 5 (Hà–Stokes–Zanello, 2010). *All matroid h-vectors $(1, h_1, h_2, h_3)$ are pure O-sequences.*

More generally, the following is a first concrete, if still tentative, general approach to Conjecture 3.4 (see [27]).

Assuming Conjecture 3.4 holds for all matroid complexes whose deletions with respect to any vertex are cones (which may not be too difficult to show with the techniques of paper [27]), Conjecture 3.4 is true in general under the following two natural (but still too bold?) assumptions:

(a) *Any matroid h-vector is differentiable for as long as it is nondecreasing.* (In fact, incidentally, would a *g-element* that Hausel [32] and Swartz [74] proved to exist in the "first half" of a matroid carry on all the way?)
(b) *Suppose that the* shifted sum*, $h'' = (1, h_1 + 1, h_2 + h'_1, \ldots, h_e + h'_{e-1})$, of two pure O-sequences h and h' is differentiable for as long as it is nondecreasing. Then h'' is also a pure O-sequence.*

4 Enumerations of Pure O-Sequences

As we have seen above, pure O-sequences arise in several areas, yet their properties are not well understood, and there are other important questions that should be addressed even if a classification is not available. For example, one would like to estimate the number of pure O-sequences of given codimension and socle degree. What happens asymptotically? Moreover, we have seen that pure O-sequences can be as far from being unimodal as we want. Nevertheless, one may ask: What are the odds for a pure O-sequence to be unimodal?

In order to discuss such questions, let us denote by $O(r, e)$, $P(r, e)$, and $D(r, e)$ the sets of O-sequences, pure O-sequences, and differentiable O-sequences, respectively, that have codimension r and socle degree e. Recall that given two functions $f, g : \mathbb{R} \to \mathbb{R}$, one says that f is *asymptotic* to g and writes $f(r) \sim_r g(r)$ if $\lim_r f(r)/g(r) = 1$. All limits are taken for r approaching infinity.

Consider now an O-sequence $(1, r-1, h_2, \ldots, h_e)$ in $O(r-1, e)$. Integrating it, that is, passing to $(1, r, r-1+h_2, \ldots, h_{e-1}+h_e)$, provides a differentiable

O-sequence in $D(r,e)$. Thus, since finite differentiable O-sequences are pure by Theorem 7(b), we have the following inclusions:

$$O(r-1,e) \hookrightarrow D(r,e) \subset P(r,e) \subset O(r,e).$$

Results by Linusson (see [46]) imply that, for r large, the cardinalities of $O(r-1,e)$ and $O(r,e)$ are asymptotically equal. It follows that in large codimensions almost all O-sequences are pure. More precisely, one has

Theorem 1 (Boij et al. [3]). *Fix a positive integer e. Then, for r large, almost all O-sequences of socle degree e are differentiable. Namely,*

$$\#O(r,e) \sim_r \#P(r,e) \sim_r \#D(r,e) \sim_r c_e \cdot r^{\binom{e+1}{2}-1},$$

where

$$c_e = \frac{\prod_{i=0}^{e-2} \binom{\binom{e+1}{2} - \binom{i+1}{2} - 1}{i}}{\left(\binom{e+1}{2} - 1\right)!}.$$

Since pure O-sequences are Hilbert functions of level algebras, we immediately get the following consequence.

Corollary 2 (Boij et al. [3]). *Fix a positive integer e. Let $L(r,e)$ be the set of level Hilbert functions of codimension r and socle degree e. Then, for r large, almost all level sequences are pure and unimodal, and*

$$\#L(r,e) \sim_r c_e r^{\binom{e+1}{2}-1}.$$

The outcome changes drastically if we fix as additional parameter the socle type t. Since each monomial of degree t is divisible by at most t distinct variables, we observe

Proposition 3 (Boij et al. [3]). *Let $P(r,e,t)$ be the set of pure O-sequences of codimension r, socle degree e, and type t. Then $\#P(r,e,t) = 0$ for $r > te$, this bound being sharp.*

Note that, in contrast to this result, there is no analogous restriction on level Hilbert functions. For example, Gorenstein algebras have type 1 and admit any positive socle degree and codimension. In fact, asymptotically, the number of their Hilbert functions is known.

Recall that an *SI-sequence* of socle degree e is an O-sequence that is symmetric about $\frac{e}{2}$ and differentiable up to degree $\lfloor \frac{e}{2} \rfloor$. Initially, Stanley and Iarrobino (see [69]) had hoped that all Hilbert functions of Gorenstein algebras were SI-sequences. Although this is not true (see [67] for the first counterexample), it is *almost true*! In fact, any differentiable O-sequence in $D(r, \lfloor \frac{e}{2} \rfloor)$ can be extended to a symmetric sequence, so that the result is an SI-sequence. Moreover, every SI-sequence is the Hilbert function of some Gorenstein algebra (see [12, 29, 51]). Obviously, the first half of the Hilbert function of a Gorenstein algebra is an O-sequence. Taken together, it follows that the number of Gorenstein Hilbert functions that are not SI-sequences is negligible:

Theorem 4 (Boij et al. [3]). *Fix a positive integer e. Let $G(r, e)$ be the set of Gorenstein Hilbert functions of codimension r and socle degree e, and let $SI(r, e)$ be the set of SI-sequences of codimension r and socle degree e. Then, for r large, almost all Gorenstein Hilbert functions are SI-sequences. More precisely,*

$$\#G(r, e) \sim_r \#SI(r, e) \sim_r c_{\lfloor e/2 \rfloor} r^{\binom{\lfloor e/2 \rfloor + 1}{2} - 1}.$$

Returning to pure O-sequences, it would be very interesting to determine, or at least to find a good estimate of the number of pure O-sequences of codimension r, socle degree e, and type t. This seems a difficult problem. However, in the simplest case, where $t = 1$, there is an easy combinatorial answer. In fact, any pure O-sequence of type 1 is the Hilbert function of a (complete intersection) algebra whose inverse system is a monomial of the form $y_1^{a_1} \cdots y_r^{a_r}$, where $a_1 + \cdots + a_r = e$ and $a_i \geq 1$ for all i. We may assume that $a_1 \geq a_2 \geq \cdots \geq a_r$ so that $(a_1, \ldots, a_r)$ is a *partition* of e. Since it is easy to see that distinct partitions lead to different Hilbert functions, we arrive at the following result.

Proposition 5 (Boij et al. [3]). *$\#P(r, e, 1) = p_r(e)$, the number of partitions of the integer e having exactly r parts.*

We now consider a slightly different asymptotic enumeration question. Fix positive integers e and t. Since the number of monomials dividing t monomials of degree e is finite, there are only finitely many pure O-sequences of socle degree e and type t. Denote their set by $P(*, e, t)$. Determining $\#P(*, e, t)$ exactly seems out of reach. However, one may hope to at least find its order for t large. The first interesting case, namely, $e = 3$, has been settled.

Theorem 6 (Boij et al. [3]). *Let $\#P(*, 3, t)$ denote the number of pure O-sequences of socle degree 3 and type t. Then*

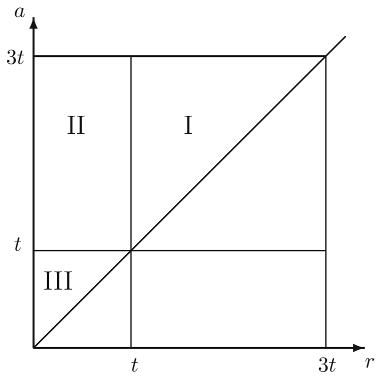
$$\lim_{t \to \infty} \frac{\#P(*, 3, t)}{t^2} = \frac{9}{2}.$$

Its proof gives some further information. Consider a pure O-sequence $(1, r, a, t)$. Then

$$r \leq a \leq 3t$$

by Hibi's theorem and the fact that t monomials of degree 3 are divisible by at most $3t$ distinct quadratic monomials. It follows that $\#P(*, 3, t) \leq \frac{9}{2}t^2$. To see that the two functions are in fact asymptotically equal, notice that, for fixed t, the possible values of r and a fall into one of the following three regions, illustrated in the figure below:

Region I: $t \le r \le a \le 3t$
Region II: $0 < r < t \le a \le 3t$
Region III: $0 < r \le a < t$



Using superscripts to denote the sets of pure O-sequences in each region, it is shown in BMMNZ that

$$\lim_{t\to\infty} \frac{\#P^{(I)}(*,3,t)}{t^2} = 2, \quad \lim_{t\to\infty} \frac{\#P^{(II)}(*,3,t)}{t^2} = 2, \quad \text{and}$$
$$\lim_{t\to\infty} \frac{\#P^{(III)}(*,3,t)}{t^2} = \frac{1}{2}.$$

The arguments use in a crucial way the Interval Property for pure O-sequences of socle degree 3. Unfortunately, this property fails in general. Nevertheless it would be very interesting to extend the above results to socle degree $e \ge 4$.

5 Open Problems

In this section we collect a few interesting problems that remain open in the area of pure O-sequences. Most of them have been discussed in the previous sections, but some are related problems that were not addressed above.

1. What is the largest type t for which all pure O-sequences are unimodal (independently of the codimension or socle degree)? Even proving that $t \ge 2$, i.e., that pure O-sequences of type 2 are unimodal in any codimension, would be very interesting.
2. For a fixed codimension r, what is the largest type t for which all pure O-sequences are unimodal?

3. For a fixed codimension r, what is the largest socle degree e for which all pure O-sequences are unimodal?
4. What is the smallest codimension r for which there exists a nonunimodal pure O-sequence of socle degree 4?
5. Determine asymptotically the number $\#P(*, e, t)$ of pure O-sequences of socle degree e and type t for t large. What is the order of magnitude of $\#P(*, e, t)$?
6. The first example of a nonunimodal pure O-sequence (due to Stanley [66], 1977) was $(1, n = 505, 2065, 3395, 3325, 3493)$, which is in fact the f-vector of a Cohen–Macaulay simplicial complex, hence in particular a pure f-vector. What is the smallest number of variables n (i.e., the number of vertices of the complex) allowing the existence of a nonunimodal pure f-vector?

   Tahat [76] has recently determined the sharp lower bound $n = 328$ for nonunimodal *Cohen–Macaulay* f-vectors of socle degree 5 (the socle degree of Stanley's original example) and has produced examples in larger socle degree with n as low as 39.
7. Stanley's Twenty-Fifth Problem for the year 2000 IMU Volume "Mathematics: Frontiers and Perspectives" [70] asks, among a few other things: are all matroid f-vectors unimodal or even log-concave? What about matroid h-vectors?

   Notice that matroid f-vectors are a much smaller subset of Cohen–Macaulay f-vectors. For a major recent breakthrough on this problem, see Huh [35] and Huh and Katz [37]; as a consequence of their work, log-concavity (hence unimodality) is now known for all f-vectors (see Lenz [42], 2012) and h-vectors (Huh [36]) of *representable* matroids. Notice also that, in general, the log-concavity of matroid h-vectors would imply the log-concavity of matroid f-vectors, as proved by Dawson [20] (1984) and Lenz [42].

Acknowledgments We wish to thank David Cook II and Richard Stanley for their helpful comments. The third author also thanks Jürgen Bierbrauer for an interesting discussion on the connections between group theory and Steiner systems. We thank the referee for a careful reading of the chapter.

The work for this chapter was done while the first author was sponsored by the National Security Agency under Grant Number H98230-12-1-0204 and by the Simons Foundation under grant #208579. The work for this chapter was done while the second author was sponsored by the National Security Agency under Grant Number H98230-12-1-0247 and by the Simons Foundation under grant #208869.

References

1. Björner, A.: Nonpure shellability, F-vectors, subspace arrangements and complexity. In: Formal Power Series and Algebraic Combinatorics. Series Formelles et Combinatoire Algébrique (1994). DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 24, pp. 25–54 (1994)
2. Boij, M., Zanello, F.: Level algebras with bad properties. Proc. Amer. Math. Soc. **135**(9), 2713–2722 (2007)
3. Boij, M., Migliore, J., Miró-Roig, R., Nagel, U., Zanello, F.: On the shape of a pure O-sequence. Mem. Amer. Math. Soc. **218**(2024), vii + 78 pp (2012)

4. Boyle, B.: On the unimodality of pure O-sequences. Ph.D. Thesis, University of Notre Dame, Indiana (2012)
5. Brenner, H., Kaid, A.: Syzygy bundles on $\mathbb{P}^2$ and the weak Lefschetz property. Illinois J. Math. **51**(4), 1299–1308 (2007)
6. Brenti, F.: Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update. In: Jerusalem Combinatorics '93. Contemporary Mathematics, vol. 178, pp. 71–89 (1994)
7. Bruck, R.H., Ryser, H.J.: The nonexistence of certain finite projective planes. Canadian J. Math. **1**, 88–93 (1949)
8. Bruns, W., Herzog, J.: Cohen–Macaulay rings. Revised edition, Cambridge Studies in Advanced Mathematics, vol. 39. Cambridge University Press, Cambridge (1998)
9. Chari, M.K.: Matroid inequalities. Discrete Math. **147**(1–3), 283–286 (1995)
10. Chari, M.K.: Two decompositions in topological combinatorics with applications to matroid complexes. Trans. Amer. Math. Soc. **349**(10), 3925–3943 (1997)
11. Chen, C., Guo, A., Jin, X., Liu, G.: Trivariate monomial complete intersections and plane partitions. J. Commut. Algebra **3**(4), 459–490 (2011)
12. Cho, Y.H., Iarrobino, A.: Hilbert Functions and Level Algebras. J. Algebra **241**(2), 745–758
13. Chowla, S., Ryser, H.J.: Combinatorial problems. Canadian J. Math. **2**, 93–99 (1950)
14. Colbourn, C.J., Dinitz, J.H. (eds.): Handbook of Combinatorial Designs. CRC Press, Boca Raton (1996)
15. Constantinescu, A., Varbaro, M.: h-vectors of matroid complexes. Preprint (2012)
16. Cook, D. II: The Lefschetz properties of monomial complete intersections in positive characteristic. J. Algebra **369**, 42–58 (2012)
17. Cook, D. II, Nagel, U.: Hyperplane sections and the subtlety of the Lefschetz properties. J. Pure Appl. Algebra **216**(1), 108–114 (2012)
18. Cook, D. II, Nagel, U.: The weak Lefschetz property, monomial ideals, and lozenges. Illinois J. Math. (to appear). (arXiv:0909.3509)
19. Cook, D. II, Nagel, U.: Enumerations deciding the weak Lefschetz property. Preprint. (arXiv:1105.6062)
20. Dawson, J.E.: A collection of sets related to the Tutte polynomial of a matroid. In: Graph theory, Singapore (1983). Lecture Notes in Mathematics, vol. 1073, pp. 193–204. Springer, Berlin (1984)
21. de Bruijn, N.G., van Ebbenhorst Tengbergen, Ca., Kruyswijk, D.: On the set of divisors of a number. Nieuw Arch. Wiskunde (2) **23**, 191–193 (1951)
22. De Loera, J.A., Kemper, Y., Klee, S.: h-vectors of small matroid complexes. Electron. J. Combin. **19**, P14 (2012)
23. Dembowski, P.: Finite geometries. Reprint of the 1968 original, Classics in Mathematics, Springer, Berlin, xii + 375 pp (1997)
24. Fröberg, R., Laksov, D.: Compressed Algebras. Conference on Complete Intersections in Acireale. In: Lecture Notes in Mathematics, vol. 1092, pp. 121–151. Springer, Berlin (1984)
25. Geramita, A.V.: Inverse systems of fat points: Waring's problem, secant varieties and Veronese varieties and parametric spaces of Gorenstein ideals. Queen's Papers in Pure and Applied Mathematics, vol. 102. The Curves Seminar at Queen's, vol. X, pp. 3–114 (1996)
26. Geramita, A.V., Harima, T., Migliore, J., Shin, Y.: The Hilbert function of a level algebra. Mem. Amer. Math. Soc. **186**(872), vi + 139 pp (2007)
27. Hà, H.T., Stokes, E., Zanello, F.: Pure O-sequences and matroid h-vectors. Ann. Comb. (to appear) (arXiv:1006.0325)
28. Hanani, M.: On Quadruple Systems. Canad. J. Math. **12**, 145–157 (1960)
29. Harima, T.: Characterization of Hilbert functions of Gorenstein Artin algebras with the weak Stanley property. Proc. Amer. Math. Soc. **123**(29), 3631–3638 (1995)
30. Harima, T., Migliore, J., Nagel, U., Watanabe, J.: The weak and strong Lefschetz properties for Artinian K-algebras. J. Algebra **262**(1), 99–126 (2003)
31. Hausel, T., Sturmfels, B.: Toric hyperkähler varieties. Doc. Math. **7**, 495–534 (2002)
32. Hausel, T.: Quaternionic geometry of matroids. Cent. Eur. J. Math. **3**(1), 26–38 (2005)

33. Herzog, J., Popescu, D.: The strong Lefschetz property and simple extensions. Preprint (arXiv:math/05065537)
34. Hibi, T.: What can be said about pure O-sequences? J. Combin. Theory Ser. A **50**(2), 319–322 (1989)
35. Huh, J.: Milnor numbers of projective hypersurfaces and the chromatic polynomial of graphs. J. Amer. Math. Soc. **25**, 907–927 (2012)
36. Huh, J.: h-vectors of matroids and logarithmic concavity. Preprint (arXiv:1201.2915)
37. Huh, J., Katz, E.: Log-concavity of characteristic polynomials and the Bergman fan of matroids. Math. Ann. **354**, 1103–1116 (2012)
38. Iarrobino, A.: Compressed Algebras: Artin algebras having given socle degrees and maximal length. Trans. Amer. Math. Soc. **285**(1), 337–378 (1984)
39. Iarrobino, A., Kanev, V.: Power sums, Gorenstein algebras, and determinantal loci. Springer Lecture Notes in Mathematics, vol. 1721. Springer, Heidelberg (1999)
40. Kirkman, T.P.: On a Problem in Combinatorics, Cambridge Dublin Math. J. **2**, 191–204 (1847)
41. Lam, C.W.H.: The search for a finite projective plane of order 10. Amer. Math. Monthly **98**(4), 305–318 (1991)
42. Lenz, M.: The f-vector of a realizable matroid complex is strictly log-concave. Combinatorics Probability, and Computing (to appear). (arXiv:1106.2944)
43. Li, J., Zanello, F.: Monomial complete intersections, the Weak Lefschetz Property and plane partitions. Discrete Math. **310**(24), 3558–3570 (2010)
44. Lindner, C.C., Rodger, C.A.: Design Theory. CRC Press, Boca Raton (1997)
45. Lindsey, M.: A class of Hilbert series and the strong Lefschetz property. Proc. Amer. Math. Soc. **139**(1), 79–92 (2011)
46. Linusson, S.: The number of M-sequences and f-vectors. Combinatorica **19**(2), 255–266 (1999)
47. Macaulay, F.H.S.: Some properties of enumeration in the theory of modular systems. Proc. London Math. Soc. **26**(1), 531–555 (1927)
48. Mathieu, E.: Mémoire sur l'étude des fonctions de plusieurs quantités, sur la manière de les former et sur les substitutions qui les laissent invariables. J. Math. Pures Appl. (Liouville) (2) **VI**, 241–323 (1861)
49. Mathieu, E.: Sur la fonction cinq fois transitive de 24 quantités. Liouville J. (2) **XVIII**, 25–47 (1873)
50. Merino, C.: The chip firing game and matroid complexes. Discrete models: combinatorics, computation, and geometry (2001). Discrete Mathematics and Theoretical Computer Science Proceedings, AA, Maison de l'informatique et des mathématiques discrétes (MIMD), Paris. pp. 245–255 (2001)
51. Migliore, J., Nagel, U.: Reduced arithmetically Gorenstein schemes and simplicial polytopes with maximal Betti numbers. Adv. Math. **180**(1), 1–63 (2003)
52. Migliore, J., Nagel, U.: A tour of the weak and strong Lefschetz properties. Preprint (arXiv:1109.5718)
53. Migliore, J., Nagel, U., Zanello, F.: A lower bound on the second entry of a Gorenstein h-vector and a conjecture of Stanley. Proc. Amer. Math. Soc. **136**(8), 2755–2762 (2008)
54. Migliore, J., Nagel, U., Zanello, F.: Bounds and asymptotic minimal growth for Gorenstein Hilbert functions. J. Algebra **210**(5), 1510–1521 (2009)
55. Migliore, J., Miró-Roig, R., Nagel, U.: Monomial ideals, almost complete intersections, and the weak Lefschetz property. Trans. Amer. Math. Soc. **363**(1), 229–257 (2011)
56. Neel, D.L., Neudauer, N.A.: Matroids you have known. Math. Mag. **82**(1), 26–41 (2009)
57. Oh, S.: Generalized permutohedra, h-vectors of cotransversal matroids and pure O-sequences. Preprint (arXiv:1005.5586)
58. Okonek, C., Schneider, M., Spindler, H.: Vector bundles on complex projective spaces. Corrected reprint of the 1988 edition. With an appendix by S.I. Gelfand. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel (2011)
59. Östergard, P.R.J., Pottonen, O.: There exists no Steiner system S(4,5,17). J. Combin. Theory Ser. A **115**(8), 1570–1573 (2008)
60. Oxley, J.G.: Matroid Theory. Oxford University Press, Oxford (2006)

61. Pastine, A.. Zanello, F.: Two unfortunate properties of pure f-vectors, in preparation
62. Proudfoot, N.: On the h-vector of a matroid complex. unpublished note.
63. Reid, L., Roberts, L., Roitman, M.: On complete intersections and their Hilbert functions. Canad. Math. Bull. **34**(4), 525–535 (1991)
64. Schweig, J.: On the h-Vector of a lattice path matroid. Electron. J. Combin. **17**(1), N3 (2010)
65. Sekiguchi, H.: The upper bound of the Dilworth number and the Rees number of Noetherian local rings with a Hilbert function. Adv. Math. **124**(2), 197–206 (1996)
66. Stanley, R.: Cohen–Macaulay complexes. In: Aigner, M. (ed.) Higher Combinatorics, pp. 51–62. Reidel, Dordrecht (1977)
67. Stanley, R.: Hilbert functions of graded algebras. Adv. Math. **28**(1), 57–83 (1978)
68. Stanley, R.: Log-concave and unimodal sequences in algebra, combinatorics, and geometry. Ann. New York Acad. Sci. **576**, 500–535 (1989)
69. Stanley, R.: Combinatorics and commutative algebra. Progress in Mathematics, vol. 41, 2nd edn. Birkhäuser, Boston (1996)
70. Stanley, R.: Positivity problems and conjectures in algebraic combinatorics. In Mathematics: Frontiers and Perspectives, pp. 295–319. American Mathematical Society, Providence (2000)
71. Stanley, R.: Enumerative Combinatorics, vol. I, 2nd edn. Cambridge University Press, Cambridge (2012)
72. Stanley, R., Zanello, F.: On the rank function of a differential poset. Electron. J. Combin. **19**(2), p. 13, 17 (2012)
73. Stokes, E.: The h-vectors of matroids and the arithmetic degree of squarefree strongly stable ideals. Ph.D. Thesis, University of Kentucky, UK (2007)
74. Swartz, E.: g-elements of matroid complexes. J. Combin. Theory Ser. B **88**(2), 369–375 (2003)
75. Swartz, E.: g-elements, finite buildings, and higher Cohen–Macaulay connectivity. J. Combin. Theory Ser. A **113**(7), 1305–1320 (2006)
76. Tahat, A.: On the nonunimodality of Cohen-Macaulay f-vectors. M.S. Thesis, Michigan Technological University, in preparation
77. Veblen, O., Wedderburn, J.H.M.: Non-Desarguesian and non-Pascalian geometries. Trans. Amer. Math. Soc. **8**(3), 379–388 (1907)
78. Watanabe, J.: The Dilworth number of Artinian rings and finite posets with rank function. In: Commutative Algebra and Combinatorics, Advanced Studies in Pure Mathematics, vol. 11, pp. 303–312. Kinokuniya Co. North Holland, Amsterdam (1987)
79. White, N. (ed.): Theory of matroids. Encyclopedia of Mathematics and Its Applications, vol. 26. Cambridge University Press, Cambridge (1986)
80. White, N. (ed.): Matroids Applications. In: Encyclopedia of Mathematics and Its Applications, vol. 40. Cambridge University Press, Cambridge (1992)
81. Zanello, F.: Extending the idea of compressed algebra to arbitrary socle-vectors. J. Algebra **270**(1), 181–198 (2003)
82. Zanello, F.: Extending the idea of compressed algebra to arbitrary socle-vectors, II: cases of non-existence. J. Algebra **275**(2), 730–748 (2004)
83. Zanello, F.: A non-unimodal codimension 3 level h-vector. J. Algebra **305**(2), 949–956 (2006)
84. Zanello, F.: Interval Conjectures for level Hilbert functions. J. Algebra **321**(10), 2705–2715 (2009)

Bounding Projective Dimension

Jason McCullough and Alexandra Seceleanu

1 Introduction

The use of algorithms in algebra as well as the study of their complexity was initiated before the advent of modern computers. Hermann [25] studied the ideal membership problem, i.e determining whether a given polynomial is in a fixed homogeneous ideal, and found a doubly exponential bound on its computational complexity. Later Mayr and Meyer [31] found examples which show that her bound was nearly optimal. Their examples were further studied by Bayer and Stillman [2] and Koh [28] who showed that these ideals also had syzygies whose degrees are doubly exponential in the number of variables of the ambient ring.

This survey addresses a different measure of the complexity of an ideal, approaching the problem from the perspective of computing the minimal free resolution of the ideal. Among invariants of free resolutions, we focus on the projective dimension, which counts the number of steps one needs to undertake in finding a minimal resolution; the precise definition of projective dimension is given in Sect. 2. In this chapter we discuss estimates on the projective dimension of cyclic graded modules over a polynomial ring in terms of the degrees of the minimal generators of the defining ideal. We also establish connections to another well-known invariant, namely, regularity.

J. McCullough (✉)
Department of Mathematics, The Mathematical Sciences Research Institute,
17 Gauss Way, Berkeley, CA 94720, USA
e-mail: jmccullough@msri.org

A. Seceleanu
Department of Mathematics, University of Nebraska-Lincoln, 203 Avery Hall,
Lincoln, NE 68588, USA
e-mail: aseceleanu2@math.unl.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_17,

The investigation of this problem was initiated by Stillman who posed the following question:

Stillman's Question 1.1 (Peeva–Stillman [36, Problem 3.14]). *Let R be any standard graded polynomial ring, suppose $I = (f_1, \dots, f_n) \subset R$ is a homogeneous ideal, and $f_1, \dots, f_n$ is a minimal set of generators of I. Is there a bound on the projective dimension of R/I depending only on $d_1, \dots, d_n$, where $d_i = \deg(f_i)$ for $i = 1, \dots, n$?*

Note that the degrees d_i of the generators as well as the number n of generators of I are part of the data with which we may bound the projective dimension $\mathrm{pd}(R/I)$; however Stillman's question asks for a bound independent of the number of variables.

To completely answer Stillman's question one would ideally like to describe:

1. A bound for $\mathrm{pd}(R/I)$ in terms of $d_1, \dots, d_n$ which is always valid
2. Examples of ideals I where the bound in (1) is the best possible
3. Much better bounds for $\mathrm{pd}(R/I)$ valid if I satisfies special conditions

In this survey we gather recent results which partially answer (2) and (3). We remark that question (1) is still wide open. We hope this chapter serves as a convenient survey of these results and spurs future work in this area.

In the next section, we fix notation for the remainder of this chapter and explain the equivalence of Stillman's problem to the analogous problem on bounding Castelnuovo–Mumford regularity. In Sect. 3, we summarize the main results that explore special cases of Stillman's question, including a sketch of the bound for ideals generated by three quadratics, three cubics, and arbitrarily many quadratics. In Sect. 4, we present several examples of ideals with large projective dimension giving large lower bounds on possible answers to Stillman's question. In Sect. 5, we summarize some related bounds on projective dimension that are distinct from Stillman's question. We close in Sect. 6 with some questions and possible approaches to Stillman's question.

2 Background and an Equivalent Problem

For the rest of this chapter, we stick to the following conventions: We use $R = K[x_1, \dots, x_N]$ to denote a polynomial ring over an arbitrary field K in N variables and we let $\mathfrak{m} = (x_1, x_2, \dots, x_N)$ denote the graded maximal ideal. We consider R as a standard graded ring with $\deg(x_i) = 1$ for all $i = 1, \dots, N$. We call a homogeneous polynomial a *form*. We denote by R_i the K-vector space of degree-i forms in R. Hence $R = \oplus_{i \geq 0} R_i$ as a K-vector space. We also denote by $R(-d)$ the rank-one free module with generator in degree d so that $R(-d)_i = R_{i-d}$. Given any finitely generated R-module M, a *free resolution* $F_\bullet$ of M is an exact sequence of the form

$$\mathbf{F}_\bullet : F_0 \xleftarrow{\partial_1} F_1 \xleftarrow{\partial_2} \cdots \xleftarrow{\partial_{s-1}} F_{s-1} \xleftarrow{\partial_s} F_s \xleftarrow{\partial_{s+1}} \cdots$$

where F_i is a free module and $M = F_0/\operatorname{Im}(\partial_1)$. The length of a resolution $F_\bullet$ is the greatest integer n such that $F_n \neq 0$, if such an integer exists; otherwise the length is infinite. We then define the *projective dimension* of M, denoted $\operatorname{pd}(M)$, to be the minimum of the lengths of all free resolutions of M. When M is graded, we require that the free resolution of M be graded, ∂_i is a graded map for all i. Moreover, $\mathbf{F}_\bullet$ is called *minimal* if $\operatorname{Im}(\partial_i) \subset \mathfrak{m}F_{i-1}$ for $i \geq 1$. The minimal graded free resolution of M is unique up to isomorphism, and it follows that $\operatorname{pd}(M)$ is the length of any minimal graded free resolution of M.

The projective dimension can be thought of as a measure of how far M is from being a free module, since finitely generated modules with projective dimension 0 are free. We note that over R every finitely generated graded projective module is free. This explains why the length of a free resolution is called the projective dimension.

It was Hilbert [26] who first studied free resolutions associated to graded modules over a polynomial ring. His syzygy theorem shows that every graded module over a polynomial ring has a finite, graded free resolution. (See [14] for a proof.)

Theorem 1 (Hilbert [26]). *Every finitely generated graded module M over the ring $K[x_1, \ldots, x_N]$ has a graded free resolution of length $\leq N$. Hence* $\operatorname{pd}(M) \leq N$.

In this survey, we shall consider the projective dimension of homogeneous ideals $I \subset R$, with the exception of Sect. 3.2, where the homogeneity assumption is not required. By convention, we study the projective dimension of cyclic modules $\operatorname{pd}(R/I)$ rather than that of ideals, noting that $\operatorname{pd}(R/I) = \operatorname{pd}(I) + 1$ for all ideals I. Hilbert's syzygy theorem shows that $\operatorname{pd}(R/I) \leq N$ for all ideals I. Even for ideals, this bound is tight. The graded maximal ideal $\mathfrak{m} = (x_1, \ldots, x_N)$ defines a cyclic module $K \cong R/\mathfrak{m}$ with $\operatorname{pd}(R/\mathfrak{m}) = N$. In fact, the Koszul complex on the variables $x_1, \ldots, x_N$ gives a minimal free resolution of $R/\mathfrak{m}$ of length N.

For a graded free resolution $\mathbf{F}_\bullet$ of M, we write $F_i = \oplus_{j \in \mathbb{Z}} R(-j)^{\beta_{i,j}(M)}$. The integers $\beta_{i,j}(M)$ are invariants of M and are called the *Betti numbers* of M. We often record these in a matrix, called the *Betti table* of M. By convention, we write $\beta_{i,j}$ in column i and row $j - i$.

	0	1	2	$\cdots$	i	$\cdots$
0:	$\beta_{0,0}(M)$	$\beta_{1,1}(M)$	$\beta_{2,2}(M)$	$\cdots$	$\beta_{i,i}(M)$	$\cdots$
1:	$\beta_{0,1}(M)$	$\beta_{1,2}(M)$	$\beta_{2,3}(M)$	$\cdots$	$\beta_{i,i+1}(M)$	$\cdots$
2:	$\beta_{0,2}(M)$	$\beta_{1,3}(M)$	$\beta_{2,4}(M)$	$\cdots$	$\beta_{i,i+2}(M)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	
j:	$\beta_{0,j}(M)$	$\beta_{1,j+1}(M)$	$\beta_{2,j+2}(M)$	$\cdots$	$\beta_{i,i+j}(M)$	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	

Another way to measure the complexity of M is to look at the degrees of the generators of the free modules F_i. We define the *Castelnuovo–Mumford regularity* of M (or just the regularity of M) to be

$$\operatorname{reg}(M) = \max\{j \mid \beta_{i,i+j}(M) \neq 0 \text{ for some } i\}.$$

Hence $\operatorname{pd}(M)$ is the index of the last nonzero column and $\operatorname{reg}(M)$ is the index of the last nonzero row in the Betti table of M. Regularity has many connections with algebraic geometry and Hilbert functions, for which we refer the interested reader to [15] and [34]. In general, both $\operatorname{reg}(R/I)$ and $\operatorname{pd}(R/I)$ may depend on the characteristic of the base field K, even in the case where I is a monomial ideal. (See, e.g., [34, Example 12.4].) A complete answer to Stillman's question should be independent of the base field, although an answer in any characteristic would be very interesting.

Example 2. Let $R = K[w,x,y,z]$ and let $I = (w^2, x^2, wy + xz)$. Then R/I has minimal graded free resolution:

$$R \xleftarrow{\begin{pmatrix} w^2 & x^2 & wy+xz \end{pmatrix}} R(-2)^3 \xleftarrow{\begin{pmatrix} -x^2 & 0 & -wy-xz & -xy & -y^2 \\ w^2 & -wy-xz & 0 & -wz & z^2 \\ 0 & x^2 & w^2 & wx & wy-xz \end{pmatrix}} R(-4)^5 \longleftarrow$$

$$\xleftarrow{\begin{pmatrix} y & z & 0 & 0 \\ w & 0 & 0 & z \\ 0 & -x & -y & 0 \\ -x & w & z & -y \\ 0 & 0 & w & x \end{pmatrix}} R(-5)^4 \xleftarrow{\begin{pmatrix} -z \\ y \\ -x \\ w \end{pmatrix}} R(-6) \longleftarrow 0$$

The Betti table of M is then

	0	1	2	3	4
0:	1	-	-	-	-
1:	-	3	-	-	-
2:	-	-	5	4	1

Therefore, we have $\operatorname{pd}(R/I) = 4$ and $\operatorname{reg}(R/I) = 2$.

One could pose a similar question to Question 1.1 by replacing the words "projective dimension" with "regularity" and asking for a bound on $\operatorname{reg}(R/I)$ purely in terms of the degrees of the generators.

Question 2.3 (Peeva–Stillman [36, Problem 3.15]). Let R be any polynomial ring and suppose $I = (f_1, \ldots, f_n) \subset R$ is an ideal. Is there a bound on $\operatorname{reg}(R/I)$ depending only on $d_1, \ldots, d_n$, where $d_i = \deg(f_i)$ for $i = 1, \ldots, n$?

One of the primary motivations for studying Question 1.1 is that it is equivalent to Question 2.3.

Theorem 4 (Caviglia [34, Theorem 29.5], [17, Section 1.3]). *Let K be a field. Fix a sequence of natural numbers $d_1 \leq \ldots \leq d_n$. The following are equivalent:*

(1) There exists a function $B(n, d_1, \dots, d_n)$ *such that* $\mathrm{pd}_R(R/I) \le B(n, d_1, \dots, d_n)$ *if* R *is any polynomial ring over* K *and* $I \subset R$ *is any graded ideal with a minimal system of homogeneous generators of degrees* $d_1 \le \dots \le d_n$.

(2) There exists a function $C(n, d_1 \dots d_n)$ *such that* $\mathrm{reg}_R(R/I) \le C(n, d_1, \dots, d_n)$ *if* R *is any polynomial ring over* K *and* $I \subset R$ *is any graded ideal with a minimal system of homogeneous generators of degrees* $d_1 \le \dots \le d_n$.

We outline a proof of this result below. First we recall a related bound on regularity. Similar to the existence of a bound on projective dimension given by the Hilbert basis theorem, there is a doubly exponential bound for the regularity of an ideal I expressed in terms of the number of variables of the ambient ring and the maximal degree of a minimal generator of I. This bound can be deduced from work of Galligo [22] and Giusti [23] in characteristic zero, as was observed by Bayer and Mumford [3, Theorem 3.7]. It was later proved in all characteristics by Caviglia and Sbarra [9].

Theorem 5. *Let* $R = K[x_1, \dots, x_N]$ *be a polynomial ring over a field* K. *Let* I *be a graded ideal in* R *and let* r *be the maximal degree of an element in a minimal system of homogeneous generators of* I. *Then* $\mathrm{reg}(I) \le (2r)^{2^{N-2}}$.

We use this bound to prove Theorem 4.

Proof of Theorem 4. Let $R = K[x_1, \dots, x_N]$ throughout. Assume (1) holds and fix an ideal I with minimal generators of degrees $d_1 \le \dots \le d_n$. Let $p = \mathrm{pd}(R/I) \le B(n, d_1, \dots, d_n)$. By the Auslander–Buchsbaum theorem, $\mathrm{depth}(R/I) = N - p$, which means that one can find a regular sequence of linear forms $l_1, \dots, l_{N-p}$ on R/I. If ℓ is a linear nonzero divisor on R/I, one obtains a short exact sequence of the form

$$0 \longrightarrow \frac{R}{I}(-1) \xrightarrow{\cdot \ell} \frac{R}{I} \longrightarrow \frac{R}{I + (\ell)} \longrightarrow 0.$$

The mapping cone construction now yields $\mathrm{reg}\,\frac{R}{I} = \mathrm{reg}\,\frac{R}{I+(\ell)}$ and by induction

$$\mathrm{reg}\,\frac{R}{I} = \mathrm{reg}\,\frac{R}{I + (l_1, \dots, l_{N-p})}.$$

Set $\bar{R} = R/(l_1, \dots, l_{N-p})$. Then $R/(I + (l_1, \dots, l_{N-p})) = \bar{R}/I\bar{R}$ can be viewed as a quotient algebra of the polynomial ring $\bar{R}$. The ring $\bar{R}$ is isomorphic to a polynomial ring in p variables; hence by applying Theorem 5 an upper bound on $\mathrm{reg}\,R/I$ is $(2d_n)^{2^{p-2}}$. One may now set

$$C(n, d_1, \dots, d_n) = (2d_n)^{2^{B(n,d_1,\dots,d_n)-2}}.$$

Conversely assume (2) holds and fix an ideal $I \subset R$. Denote by $\mathrm{gin}(I)$ the generic initial ideal of I with respect to the degree reverse lexicographic ordering on the monomials of R. This term order has good properties with respect to both projective dimension and regularity. In particular, Bayer and Stillman [2, Corollaries 19.11 and 20.21] proved that $\operatorname{pd}(R/\operatorname{gin}(I)) = \operatorname{pd}(R/I)$ and $\operatorname{reg}(R/\operatorname{gin}(I)) = \operatorname{reg}(R/I)$. Moreover, the projective dimension of $R/\operatorname{gin}(I)$ can be read off directly from a minimal set of generators as the largest index among the indices of variables appearing in the minimal generators. (Equivalently, the projective dimension of $R/\operatorname{gin}(I)$ can be interpreted as the number of distinct variables appearing in the unique minimal generating set of the ideal.)

Assume that I has minimal generators of degrees $d_1, \ldots, d_n$. The relation between the projective dimensions of R/I and $R/\operatorname{gin}(I)$ allows us to bound $\operatorname{pd}(R/I)$ in terms of $C(n, d_1 \ldots d_n)$ and the number of generators of $\operatorname{gin}(I)$ as follows:

$$\begin{aligned}
\operatorname{pd}(R/I) &= \operatorname{pd}(R/\operatorname{gin}(I)) \\
&= \text{number of variables appearing in generators of } \operatorname{gin}(I) \\
&\leq \text{sum of the degrees of generators of } \operatorname{gin}(I) \\
&\leq \text{number of generators of } \operatorname{gin}(I) \cdot \text{max generator degree of} \operatorname{gin}(I) \\
&\leq \text{number of generators of } \operatorname{gin}(I) \cdot \operatorname{reg}(R/\operatorname{gin}(I)) \\
&= \text{number of generators of } \operatorname{gin}(I) \cdot \operatorname{reg}(R/I) \\
&\leq \text{number of generators of } \operatorname{gin}(I) \cdot C(n, d_1, \ldots, d_n)
\end{aligned}$$

Hence it is sufficient to bound the number of generators of $\operatorname{gin}(I)$ in terms of $d_1, \ldots, d_n$. Since we may assume a bound on $\operatorname{reg}(I) = \operatorname{reg}(\operatorname{gin}(I))$ is given by $C(n, d_1, \ldots, d_n)$, this means that the degrees of minimal generators of $\operatorname{gin}(I)$ are at most $C(n, d_1, \ldots, d_n)$. Note that we may assume I is already written in generic coordinates since a linear change of coordinates does not change the values of the input parameters $d_1, \ldots, d_n$.

To estimate the number of generators of the initial ideal of I, we need to understand the algorithmic procedure that produces a Gröbner basis of I, i.e., a set of elements of I whose leading terms generate the initial ideal $in(I)$. This algorithm was given by Buchberger and it involves enlarging a generating set of I by adding to the set at each step reductions of S-polynomials obtained using pairs f, g of polynomials from the preceding step's output set. The S-polynomial of a pair f, g is defined as

$$S(f, g) = \frac{lcm(LT(f), LT(g))}{LT(f)} f - \frac{lcm(LT(f), LT(g))}{LT(g)} g.$$

Here $LT(f), LT(g)$ are the leading terms of the polynomials f, g, respectively, and $lcm(m_1, m_2)$ denotes the least common multiple of the monomials m_1 and m_2. Note that the degree of $S(f, g)$ is always greater or equal to the maximum of the degrees of the polynomials f, g and that equality is attained if and only if $in(f) \mid in(g)$ or $in(g) \mid in(f)$, in which case this S-polynomial need not be included in a reduced Gröbner basis. Recall that a Gröbner basis is called reduced if no monomial in any element of the basis is in the ideal generated by the leading terms of the other elements of the basis. Hence at each stage in Buchberger's algorithm the maximum degree of the polynomials obtained strictly increases. Thus the number of steps in Buchberger's algorithm is bounded by the regularity of $in(I)$, hence also by $C(n, d_1, \dots, d_n)$.

Now starting with n minimal generators of I, in the first step in Buchberger's algorithm, one computes at most $\binom{n}{2}$ S-polynomials. Similarly if one denotes the number of S-polynomials computed at the i^{th} step of Buchberger's algorithm by n_i, then $n_{i+1} \le \binom{n_i}{2}$ and $n_1 \le \binom{n}{2}$. Hence there is a polynomial $P(d_1, \dots, d_N)$ such that $\sum_{i=1}^{C(d_1,\dots,d_n)} n_i \le P(d_1, \dots, d_n)$. Finally one may set in this case

$$B(n, d_1, \dots, d_n) = P(d_1, \dots, d_n) \cdot C(n, d_1, \dots, d_n).$$

□

It is worth noting that the bounds achieved for Questions 1.1 and 2.3 are likely quite different.

3 Upper Bounds and Special Cases

In this section we summarize the cases where the answer to Stillman's question is known to be affirmative. In some simple cases one easily sees that a bound on projective dimension is possible. However, even with three-generated ideals in degree two, producing such a bound is nontrivial.

In the simplest case, that of $I = (f)$ being a principal ideal, $\text{pd}(R/I) = 1$. If $I = (f, g)$ is minimally generated by two forms, either $\text{ht}(I) = 1$ or 2. If $\text{ht}(I) = 2$, then f, g is a regular sequence and $R/(f, g)$ is resolved by the Koszul complex on f and g. So $\text{pd}(R/I) = 2$. If $\text{ht}(I) = 1$, then there exist c, f', and g' with $f = cf'$, $g = cg'$, and $\text{ht}(f', g') = 2$, so again $\text{pd}(R/I) = 2$. We consider the complications for the three-generated case in the following section.

We also note here that when $I = (m_1, m_2, \dots, m_n)$ is generated by n monomials, $\text{pd}(R/I) \le n$. This is clear when $n = 1$ and follows by induction by considering the short exact sequence

$$0 \longrightarrow R/((m_1, m_2, \dots, m_{n-1}) : m_n) \xrightarrow{m_n} R/(m_1, m_2, \dots, m_{n-1}) \longrightarrow R/I \longrightarrow 0.$$

Since $((m_1, m_2, \dots, m_{n-1}) : m_n)$ is a monomial ideal generated by the $n - 1$ monomials $\frac{\text{lcm}(m_i, m_n)}{m_n}$ for $i = 1$ to $n - 1$, the projective dimension of the first two terms is at most $n - 1$ by induction. Hence we have $\text{pd}(R/I) \le n$, say by [18, Lemma 1]. Alternatively, one could use the fact that the Taylor resolution (see, e.g., [34, Construction 26.5]) of I is a possibly non-minimal free resolution of R/I of length n. Hence the projective dimension of R/I is no larger than n.

In general, if I is generated by n polynomials each with at most m terms of degree d, then it takes at most mnd linear forms to express those n generators. Thus the projective dimension of such an ideal is at most mnd, independent of the number of variables in the ring. So all interesting cases for Stillman's question occur when we do not assume a bound on the number of terms in each minimal generator of I.

For the rest of the section we consider the next simplest cases: three-generated ideals in low degree and ideals generated by quadratic polynomials.

3.1 Three-Generated Ideals

In this section we consider the projective dimension of R/I where I is minimally generated by three quadratic or three cubic forms in $K[x_1, x_2, \dots, x_N]$. In the case of three quadratic forms, Eisenbud and Huneke proved the following in unpublished work:

Theorem 1 (Eisenbud–Huneke). *Let $I = (f, g, h)$ where $f, g,$ and h are homogeneous minimal generators of degree 2 in a polynomial ring R. Then* $\operatorname{pd}(R/I) \leq 4$.

We will need several results to prove this theorem. Since $\operatorname{pd}(R/I)$ does not change after tensoring with an extension of the field of coefficients, we may assume that K is infinite. First we show that the multiplicity of I is at most 3. Then we handle the multiplicity 1, 2, and 3 cases separately. We begin by defining the multiplicity of an ideal and recalling related results.

For a graded module M, the *Hilbert series* $H_M(t) = \sum_{i \geq 0} dim_K M_i t^i$ can be written as a rational function of the form $H_M(t) = \frac{h(t)}{(1-t)^d}$, where d is the dimension of M and h is a polynomial of degree at most N. We define the *multiplicity* of a graded R-module M to be the value $e(M) = h(1)$. For an artinian module M, the multiplicity is equal to the *length* of the module defined as $\lambda(M) = \sum_{i \geq 0} dim_K M_i$. By convention, for a homogeneous ideal I, we refer to $e(R/I)$ as the multiplicity of I.

Next, we recall the associativity formula for multiplicity. (See, e.g., [30, Theorem 14.7].) For an ideal J of R,

$$e(R/J) = \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R) \\ \dim(R/\mathfrak{p}) = \dim(R/J)}} e(R/\mathfrak{p}) \lambda(R_\mathfrak{p}/J_\mathfrak{p}).$$

Let I^{un} denote the unmixed part of I, defined as the intersection of minimal primary components of I with height equal to $\operatorname{ht}(I)$. For every $\mathfrak{p} \in \operatorname{Spec}(R)$ with $\dim(R/\mathfrak{p}) = \dim(R/I)$, we have that $I^{un}_\mathfrak{p} = I_\mathfrak{p}$. Hence

$$e(R/I^{un}) = \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R) \\ \dim(R/\mathfrak{p}) = \dim(R/I^{un})}} e(R/\mathfrak{p}) \lambda(R_\mathfrak{p}/I^{un}_\mathfrak{p}) = \sum_{\substack{\mathfrak{p} \in \operatorname{Spec}(R) \\ \dim(R/\mathfrak{p}) = \dim(R/I)}} e(R/\mathfrak{p}) \lambda(R_\mathfrak{p}/I_\mathfrak{p}) = e(R/I).$$

We will often pass to the unmixed part of I and use the fact that the multiplicity does not change, as in the following result.

Proposition 2. *Using the notation in Theorem 1, if* $\operatorname{ht}(I) = 2$*, then* $e(R/I) \leq 3$.

Proof. By passing to the unmixed part of I and using structure theorems on ideals with small multiplicity, we can finish off the proof. We may assume that f, g form a regular sequence of quadratic forms. Thus $e(R/(f, g)) = 4$. We have the series of containments $(f, g) \subset I \subset I^{un}$. Note that (f, g) and I^{un} are unmixed ideals of height two. If $e(R/(f, g)) = e(R/I^{un})$, then $(f, g) = I^{un}$ by [18, Lemma 8]. But this would force $(f, g) = (f, g, h)$, contradicting that h is a minimal generator of I. Thus $4 = e(R/(f, g)) > e(R/I^{un}) = e(R/I)$. □

Following Engheta [17], we introduce the following notation to keep track of the possibilities for the associated primes of minimal height of an ideal J.

Definition 3. We say J is of type $\langle e = e_1, e_2, \ldots, e_m | \lambda = \lambda_1, \lambda_2, \ldots, \lambda_m \rangle$ if J has m associated primes $\mathfrak{p}_1, \ldots, \mathfrak{p}_m$ of minimal height with $e(R/\mathfrak{p}_i) = e_i$ and with $\lambda(R_{\mathfrak{p}_i}/J_{\mathfrak{p}_i}) = \lambda_i$, for $i = 1, \ldots, m$. (In which case, we have $e(R/J) = \sum_{i=1}^{m} e_i \lambda_i$ by the associativity formula.)

We also need the following structure theorem for ideals of height two and multiplicity two.

Proposition 4 (Engheta [18, Proposition 11]). *Let J be a height two unmixed ideal of multiplicity two. Then* $\operatorname{pd}(R/J) \leq 3$ *and J is one of the following:*

(1) $(x, y) \cap (w, z) = (xw, xz, yw, yz)$ *with independent linear forms* w, x, y, z.
(2) (x, yz) *with independent linear forms* x, y, z.
(3) A prime ideal generated by a linear form and an irreducible quadratic.
(4) (x, y^2) *with independent linear forms* x, y.
(5) $(x, y)^2 + (ax + by)$ *with independent linear forms* x, y *and* $a, b \in \mathfrak{m}$ *such that* x, y, a, b *form a regular sequence.*

The proof of this proposition uses the associativity formula to divide the possibilities into one of three cases, namely, ideals of type $\langle e = 2 | \lambda = 1 \rangle$ (Case (3)), type $\langle e = 1 | \lambda = 2 \rangle$ (Cases (4) and (5)), and type $\langle e = 1, 1 | \lambda = 1, 1 \rangle$ (Cases (1) and (2)). Finally, one checks by hand that $\operatorname{pd}(R/J) \leq 3$ in each of the resulting cases.

We also need the following result, obtained originally using residual intersection techniques by Huneke and Ulrich [27] and later using more elementary homological algebra techniques by Fan [21].

Theorem 5 (Huneke–Ulrich [27, p. 20], Fan [21, Corollary 1.2]). *Let R be a regular local ring and let I be a three-generated ideal of height 2. If R/I^{un} is Cohen–Macaulay (i.e.,* $\operatorname{pd}(R/I^{un}) = 2$*), then* $\operatorname{pd}(R/I) \leq 3$.

For a proof, we refer the reader to [21]. This result allows us to focus only on unmixed ideals with fixed multiplicity. Using the results above, we are now ready to prove Theorem 1.

Proof of Theorem 1. By Krull's (generalized) principal ideal theorem [30, Theorem 13.5], $\mathrm{ht}(I) \leq 3$. If $\mathrm{ht}(I) = 1$, then there are linear forms c, f', g', h' with $f = cf'$, $g = cg'$, and $h = ch'$. Hence $I \cong (f', g', h')$, and so $\mathrm{pd}(R/I) = 3$.

If $\mathrm{ht}(I) = 3$, then f, g, h form a regular sequence and the Koszul complex on f, g, h forms a minimal free resolution of R/I. Again $\mathrm{pd}(R/I) = 3$. Hence we may assume that $\mathrm{ht}(I) = 2$. Moreover, we may assume that f, g form a regular sequence.

Now by Proposition 2, $e(R/I) = 1$, 2, or 3. If $e(R/I) = 1$, then by the associativity formula, I^{un} is primary to a height two prime ideal $\mathfrak{p}$ of multiplicity one. Such a prime ideal is generated by two linear forms, say $\mathfrak{p} = (x, y)$. Since $\lambda(R_{\mathfrak{p}}/I_{\mathfrak{p}}) = 1$ and I is $\mathfrak{p}$-primary, $I^{un} = \mathfrak{p}$. Clearly $\mathrm{pd}(R/\mathfrak{p}) = 2$. It then follows from Theorem 5 that $\mathrm{pd}(R/I) \leq 3$.

If $e(R/I) = 3$, consider the short exact sequence

$$0 \to R/((f, g) : I) \xrightarrow{h} R/(f, g) \to R/I \to 0.$$

Since f, g form a regular sequence of quadratic forms, we have $e(R/(f, g)) = 4$. Since multiplicity is additive in short exact sequences, $e(R/((f, g) : I)) = 1$. As $(f, g) : I$ is unmixed, we have $(f, g) : I = (x, y)$ for independent linear forms x and y. Therefore, $\mathrm{pd}(R/((f, g) : I)) = 2$. Since $\mathrm{pd}(R/(f, g)) = 2$, it follows that $\mathrm{pd}(R/I) \leq 3$.

Finally, in the case where $e(R/I) = 2$, we use the same exact sequence above. In this case $(f, g) : I$ is an unmixed, height two ideal of multiplicity two. By Proposition 4, $\mathrm{pd}(R/((f, g) : I))) \leq 3$. It follows that $\mathrm{pd}(R/I) \leq 4$. This completes the proof. □

We see from Example 2 that this bound is indeed tight. The next case to consider, that of an ideal minimally generated by three cubics, is significantly more complicated. In his thesis [17], and subsequently in a sequence of papers [18, 20], Engheta proved the following:

Theorem 6 (Engheta [20, Theorem 5]). *If f, g, h are three cubic forms in a polynomial ring R over a field, then* $\mathrm{pd}(R/(f, g, h)) \leq 36$.

The outline of the proof is similar to that given above for the case of three quadratic forms. Engheta first shows that the multiplicity of such an ideal is at most 8. (In characteristic zero, Engheta shows that the multiplicity can be at most 7. See [19].) He then analyzes each case separately, often using techniques from linkage theory and the structure theorems for unmixed ideals of small multiplicity. Unfortunately there is currently no complete structure theorem of unmixed ideals of multiplicity three. In those remaining cases, he shows that such ideals can be expressed in terms of a fixed number of linear or quadratic forms. A similar technique was later used by Ananyan and Hochster to study all ideals generated by linear and quadratic polynomials. For more details see Sect. 3.2.

We also note that the bound of 36 is likely not tight. The largest known projective dimension for an ideal minimally generated by three cubic forms is just 5. The first such example was given by Engheta. (See [20, Section 3].) The following simple example was discovered by the first author in [32].

Example 7. Let $R = K[a, b, c, x, y]$, where K is any field. Let $\mathfrak{m}$ denote the graded maximal ideal. Consider the ideal $I = (x^3, y^3, x^2a + xyb + y^2c)$. Then $x^2y^2 \in (I : \mathfrak{m}) - I$. It follows that $\operatorname{depth}(R/I) = 0$ and, by the Auslander–Buchsbaum theorem, that $\operatorname{pd}(R/I) = 5$.

3.2 Ideals Generated by Quadratic Polynomials

In a certain sense, ideals generated by quadratic polynomials are ubiquitous. In [33, Thorem 1], Mumford shows that any projective variety of degree d can be re-embedded (via the d-uple embedding) as a variety cut out by an ideal generated by quadratic forms.

In [1], Ananyan and Hochster propose a method of analyzing the projective dimension of ideals generated by polynomials of degree at most two, which need not be homogeneous. We review their idea of using a specific standard form as well as the derived recursive bound on projective dimension. Since the techniques of Ananyan and Hochster can be applied when the minimal generators are non-homogeneous, we reserve the use of the term quadratic form for a homogeneous polynomial of degree two, and we call a possibly not homogeneous polynomial of degree two a quadratic polynomial. We then illustrate the techniques of [1] for the case of ideals generated by three homogeneous quadratics.

We begin with describing the focus of interest of this section: the standard form associated to an ideal generated by linear and quadratic polynomials. We note in Remark 10 that standard forms are by no means unique; however we shall often pick a convenient standard form and refer to it by abuse of terminology as *the* standard form associated to a certain ideal.

Definition 8. Let I be an ideal generated by m linear polynomials and n quadratic polynomials in a standard graded polynomial ring $R = K[x_1, \dots, x_N]$. A standard form associated to the ideal I is given by a partition of a K-basis $\{x_1, \dots, x_N\}$ of R_1 into subsets which satisfy the properties listed below together with a set $\{F_1, \dots, F_{m+n}\}$ of generators of I written in a manner compatible with this partition. In the following we shall refer to the elements $\{x_1, \dots, x_N\}$ as variables. We describe the properties required by the standard form first for the variables:

(1) The first m variables $x_1, \dots x_m$, called *leading variables*, are the linear minimal generators of I.
(2) The next $h = ht(I/(x_1, \dots x_m))$ variables $x_{m+1}, \dots x_{m+h}$, called *front variables*, are chosen such that the images $f_1, \dots, f_h$ of a maximal regular sequence $F_1, \dots, F_h$ of quadratic forms in $I/(x_1, \dots, x_m)$ under the projection $\pi : R \longrightarrow K[x_{m+1}, \dots, x_{m+h}]$ continue to form a regular sequence.

(3) The next r variables $x_{m+h+1}, \dots x_{m+h+r}$, called *primary coefficient variables*, are the coefficients of the leading and front variables when $F_1, \dots, F_n$ are viewed as polynomials in $K[x_{m+h+1}, \dots, x_N][x_1, \dots, x_{m+h}]$.
(4) The next s variables $x_{m+h+r+1}, \dots x_{m+h+r+s}$, called *secondary coefficient variables*, are the coefficients of the primary coefficient variables in the images of $F_1, \dots, F_n$ under the projection $\pi' : R \longrightarrow K[x_1, \dots, x_{m+h}]$, viewed as polynomials in $K[x_{m+h+r+1}, \dots, x_N][x_{m+h+1}, \dots, x_{m+h+r}]$.
(5) The variables $x_{m+h+r+1}, \dots, x_N$ are called the *tail variables*.

In practice, a maximal regular sequence $x_1, \dots x_m, F_1, \dots F_h$ of elements of I is chosen first, and the variables $x_{m+h+1}, \dots x_N$ are obtained by extending this sequence to a system of parameters on R. From this point on we view the generators F_i as being written in terms of the variables described above. The term monomial henceforth will be used for monomials in the variables $x_1, \dots, x_N$. Next we list the properties required by the standard form for the polynomials F_i:

1. $F_{n+i} = x_i$ for $1 \leq i \leq m$ are the linear generators of I.
2. $F_1, \dots, F_h$ form a maximal regular sequence in I.
3. $F_1, \dots, F_n$ contain no terms written using leading variables only.
4. Some of the F_i may be 0 and we require that these appear last.

Partitioning the set of variables of the ring R into the various categories appearing above yields natural partitions of the monomials appearing in the generators F_i. Recall that the projection map onto the smaller polynomial ring generated by the front variables takes F_i to f_i. We call the f_i *front polynomials*. Similarly, define g_i to be the image of F_i via projecting onto the polynomial ring generated by the tail variables. We call g_i *tail polynomials*. Therefore $F_i = f_i + e_i + g_i$, where e_i is the sum of mixed terms in the leading and front coefficient variables, front and primary coefficient variables, leading and primary coefficient variables, or primary and secondary coefficient variables and quadratic terms in the primary coefficients.

The following estimates are a clear consequence of Definition 8 and will prove crucial toward establishing the bound in Theorem 16.

Proposition 9 (Size estimates for the types of variables). *Given I an ideal generated by m linear polynomials and n quadratic polynomials, the number of variables needed to write I in a standard form is bounded by the sum of the following estimates:*

(1) Exactly m leading variables, $x_1, \dots x_m$
(2) Exactly $h = ht(I/(x_1, \dots x_m))$ front variables
(3) At most $n(m + h)$ primary coefficients
(4) The total number of variables needed to write the ideal of tail polynomials $g_1, \dots, g_n$ in standard form

To understand some of the subtleties of the standard form algorithm, we exhibit two examples which fit in the framework of three-generated ideals. In particular, we wish to illustrate the following:

Remark 10. The invariants h, m, n in Definition 8 are uniquely determined by I (for n to be uniquely determined one needs to assume that the set of generators of I was minimal to begin with). However, the parameter $h' = \mathrm{ht}(g_1, \ldots, g_n)$ may vary among different standard forms associated to the same ideal I. We note that since $(g_1, \ldots, g_n) = I/(x_1, \ldots, x_{m+h+r})$, one always has $h' \le h$.

Example 11 (The twisted cubic). Let $I_C \subset K[x_1, x_2, x_3, x_4]$ be the ideal of maximal minors of the matrix $\begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \end{bmatrix}$. A computation shows $I_C = (x_2^2 - x_1x_3, x_3^2 - x_2x_4, x_2x_3 - x_1x_4)$ is a prime ideal of height 2 and multiplicity 3. When thought of as an ideal in $K[x_1, x_2, x_3, x_4]$, I cuts out a curve $C \subset \mathbb{P}^3$ known as the twisted cubic.

To find a standard form one may pick x_2, x_3 as front variables. We underline the front variables in all examples for ease of parsing the respective standard forms. An inspection of the defining equations of I reveals that with respect to this choice of front variables, x_1, x_4 become primary coefficients and there are no tail variables. Following the notation introduced in 8, we write

$$\begin{aligned} F_1 &= \underbrace{\underline{x_2^2}}_{f_1} - \underbrace{\underline{x_3}x_1}_{e_1} \\ F_2 &= \underbrace{\underline{x_3^2}}_{f_2} - \underbrace{\underline{x_2}x_4}_{e_2} \\ F_3 &= \underbrace{\underline{x_2x_3}}_{f_3} - \underbrace{x_1x_4}_{e_3}. \end{aligned}$$

Note that $(g_1, g_2, g_3) = (0)$ and consequently $h' = 0$.

In the following we show that, regardless of the choice of the different types of variables, $h' = 0$ for any standard form of the ideal of the twisted cubic. Assume F_1, F_2 is a maximal regular sequence inside I_C. (Necessarily, F_1, F_2 will be quadratic polynomials.) Since (F_1, F_2) generate a complete intersection of multiplicity 4, $(F_1, F_1) \subset I_C$ and I_C is a prime ideal of multiplicity 3, the primary decomposition of (F_1, F_2) must be $(F_1, F_2) = I_C \cap I_L$, where I_L is a prime ideal of multiplicity 1 and height 2, i.e., the defining ideal of a line in $\mathbb{P}^3$. Let $I_L = (\ell_1, \ell_2)$ (for the choice of F_1, F_2 listed in the example above, $I_L = (x_2, x_3)$).

To extend F_1, F_2 to a maximal regular sequence on R one must pick variables $y_3, y_4 \notin I_L$. The set l_1, l_2, y_3, y_4 is a basis for R_1 and we shall for the moment think of the equations of I_C written in terms of this basis. Since $(F_1, F_2) \subset I_L$, F_1, F_2 are linear combinations of terms divisible either by l_1 or by l_2 and since (F_1, F_2) is not contained in I_L^2, some of these terms must be of the form $l_i y_j$ $(1 \le i \le 2, 3 \le j \le 4)$. In fact, since (F_1, F_2) is not a cone (it is in fact the union of the twisted cubic curve C and the line L), it cannot be written in terms of 3 variables only; hence both x_3 and x_4 must appear in the cross terms. We now consider any choice of front variables which will be of the form

$$y_1 = a_1 l_1 + a_2 l_2 + a_3 y_3 + a_4 y_4,\ \ y_2 = b_1 l_1 + b_2 l_2 + b_3 y_3 + b_4 y_4,$$

with $a_i, b_i \in K$ and $\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}$ of rank 2. The cross terms described above yield terms of the form $y_i y_j$ ($1 \leq i \leq 2, 3 \leq j \leq 4$) which witness the fact that y_3, y_4 are primary coefficients. Therefore there are no tail variables and no tail polynomials in the standard form.

The reader may wonder what ideals exhibit standard forms with other values of h'. We prove a statement regarding the case $h' = 2$.

Lemma 12. *The only homogeneous height 2 ideal of* $R = K[x_1, x_2, x_3, x_4]$ *which admits a standard form with* $h' = 2$ *is* $I = (x_1^2 + x_3^2, x_2^2 + x_4^2, x_1x_2 + x_3x_4)$.

Proof. If I has linear generators, then the height of I modulo the leading variables is $h < \mathrm{ht}(I) = 2$. Since $h' \leq h < 2$, this contradicts the assumption $h' = 2$. Therefore there are no leading variables. Assume, up to relabeling the variables, that x_1, x_2 are the front variables. Note that the front polynomials are contained in the K-span of the monomials $\{x_1^2, x_1x_2, x_2^2\}$. Let Q be the kernel of the ring homomorphism

$$K[T_1, T_2, T_3] \stackrel{(f_1, f_2, f_3)}{\longrightarrow} K[x_1, x_2].$$

If the vector space dimension of the K-span of f_1, f_2, f_3 is 3, then by taking suitable linear combinations of the f_i (and corresponding linear combinations of the original generators F_i) we may assume $f_1 = x_1^2, f_2 = x_2^2$, and $f_3 = x_1x_2$. In this case $Q = (T_1T_2 - T_3^2)$. By [1, Key Lemma (c)] the tail polynomials satisfy the front relations, i.e., $g_1g_2 - g_3^2 = 0$. Since $h' = \mathrm{ht}(g_1, g_2, g_3) = 2$, the tail polynomials cannot satisfy additional relations. Thus the tail polynomials must be of the form $g_1 = l_1, g_2 = l_2, g_3 = l_1l_2$ with $l_1, l_2 \in K[x_3, x_4]_1$. But now one makes l_1, l_2 the tail variables and recovers the desired form of the ideal I.

If the vector space dimension of the K-span of f_1, f_2, f_3 is 2, then we may assume that f_1, f_2 form a regular sequence and $f_3 = 0$; hence $T_3 \in Q$. Another application of [1, Key Lemma (c)] guarantees that $g_3 = 0$. It is important to note that $h' = 2$ means at least 2 tail variables are present. The assumption that the ambient ring has exactly four variables ensures there are no primary coefficients, so that $e_1 = e_2 = e_3 = 0$. But now $F_3 = f_3 + e_3 + g_3 = 0$, a contradiction. □

We are thus led to a closer examination of the ideal in the previous lemma. This example exhibits a contrasting behavior to the ideal of the twisted cubic: it allows standard forms with $h' = 0, 1$, and 2, respectively.

Example 13. Let $I = (x_1^2 + x_3^2, x_2^2 + x_4^2, x_1x_2 + x_3x_4) \subset \mathbb{C}[x_1, x_2, x_3, x_4]$. For this example we shall study what possible values can occur for h':

(1) Our first choice will be to take x_1, x_2 as front variables. This makes x_3, x_4 tail variables and there are no primary (or secondary) coefficients. With the notation of Definition 8, we have

$$\begin{aligned} F_1 &= \underbrace{\underline{x_1^2}}_{f_1} - \underbrace{x_3^2}_{g_1} \\ F_2 &= \underbrace{\underline{x_2^2}}_{f_2} - \underbrace{x_4^2}_{g_2} \\ F_3 &= \underbrace{\underline{x_1 x_2}}_{f_3} - \underbrace{x_3 x_4}_{g_3}. \end{aligned}$$

In this case $(g_1, g_2, g_3) = (x_3^2, x_4^2, x_3x_4)$ is an ideal of height $h' = 2$.

(2) Our second choice will be to take $x_1' = x_1 + ix_3$ and x_2 as front variables. Rewriting the ideal I by substituting $x_1 = x_1' - ix_3$ yields

$$\begin{aligned} I &= ((\underline{x_1'} - ix_3)^2 + x_3^2, \underline{x_2}^2 + x_4^2, (\underline{x_1'} - ix_3)\underline{x_2} + x_3x_4) \\ &= (\underline{x_1'}^2 - 2i \cdot \underline{x_1}x_3,\ \underline{x_2}^2 + x_4^2,\ \underline{x_1'x_2} - i \cdot \underline{x_2}x_3 + x_3x_4). \end{aligned}$$

With the notation of Definition 8 we have

$$\begin{aligned} F_1 &= \underbrace{\underline{x_1'}^2}_{f_1} \underbrace{-2i \cdot \underline{x_1'}x_3}_{e_1} \\ F_2 &= \underbrace{\underline{x_2}^2}_{f_2} + \underbrace{x_4^2}_{g_2} \\ F_3 &= \underbrace{\underline{x_1'x_2}}_{f_3} \underbrace{-i \cdot \underline{x_2}x_3 + x_3x_4}_{e_3}. \end{aligned}$$

In this case $(g_1, g_2, g_3) = (x_4^2)$ which is an ideal of height $h' = 1$.

(3) Our last choice will be to take $y_1 = x_1 + ix_3$, $y_2 = x_2 + ix_4$ as front variables, and we shall rename $y_3 = x_3$, $y_4 = x_4$. Rewriting the ideal I with respect to the linear change of coordinates from the x variables to the y variables yields

$$\begin{aligned} I &= ((\underline{y_1} - iy_3)^2 + y_3^2, (\underline{y_2} - y_4)^2 + y_4^2, (\underline{y_1} - iy_3)(\underline{y_2} - iy_4) + y_3y_4) \\ &= (\underline{y_1}^2 - 2i \cdot \underline{y_1}y_3,\ \underline{y_2}^2 - 2i \cdot \underline{y_2}y_4,\ \underline{y_1y_2} - i \cdot \underline{y_1}y_4 - i \cdot \underline{y_2}y_3). \end{aligned}$$

With the notation of Definition 8 we have

$$\begin{aligned} F_1 &= \underbrace{\underline{y_1}^2}_{f_1} \underbrace{-2i \cdot \underline{y_1}y_3}_{e_1} \\ F_2 &= \underbrace{\underline{y_2}^2}_{f_2} \underbrace{-2i \cdot \underline{y_2}y_4}_{e_2} \\ F_3 &= \underbrace{\underline{y_1y_2}}_{f_3} \underbrace{-i \cdot \underline{y_1}y_4 + i \cdot \underline{y_2}y_3}_{e_3}. \end{aligned}$$

In this case $(g_1, g_2, g_3) = 0$ and consequently $h' = 0$. This behavior of the standard form can be expected since the primary decomposition

$$I = (x_1 + ix_3, x_2 - ix_4) \cap (x_1 - ix_3, x_2 + ix_4) \cap (x_1, x_2, x_3, x_4)$$

shows $I \subset (y_1, y_2)$, which means there cannot be any tail polynomials with respect to choosing y_1, y_2 as front variables.

We now sketch how the construction of standard form associated to an ideal I is applied to finding bounds for the projective dimension of R/I in [1]. The main idea is that the use of standard forms allows one to find a suitable polynomial algebra A generated by linear and quadratic forms that contains the ideal I while having a number of generators that can be bounded in terms of m, n, h. The leading, front, and primary coefficient variables are all included as generators of A. Proposition 9 provides estimates for the respective sizes of these sets of variables. Note that it remains only to find a suitable ambient algebra for the ideal $(g_1, \dots g_n)$. The rest of the generators of A are iteratively determined, reducing the height of the ideal being analyzed.

The case when $h' < h$ allows one to use induction to complete the process. The case when $h' = h$ requires another application of the standard form for the ideal $(x_1, \dots, x_{m+h+r+s}, g_1, \dots g_n)$, which yields new front polynomials $\alpha_1, \dots, \alpha_n$ and new tail polynomials $\beta_1, \dots, \beta_n$. Let h'' be the height of the ideal $(\beta_1, \dots, \beta_n)$. If $h'' < h$ then A will be generated by the leading, front, and primary coefficient variables of I together with the generators of the algebra containing $(\beta_1, \dots, \beta_n)$. If $h'' = h$ then the ideals $(f_1, \dots, f_n)$, $(g_1, \dots, g_n)$, $(\alpha_1, \dots, \alpha_n)$, and $(\beta_1, \dots, \beta_n)$ are proven to be linearly presented. Toward this end one uses Lemma 14, the proof of which can be found in [1]. From here one deduces that there were exactly h nonzero g_i which, together with the leading, front, primary, and secondary coefficient variables in the standard form of I, are the generators of A.

The following table summarizes the notations introduced. The last column refers to estimates discussed in Proposition 9.

Ideal	Height	Ambient ring	Number of variables
$(F_1, \dots, F_n)$	h	$S = K[x_{m+1}, \dots, x_N]$	$N - m$
$(f_1, \dots, f_n)$	h	$K[x_{m+1}, \dots, x_{m+h}]$	h
$(g_1, \dots, g_n)$	h'	$K[x_{m+h+r+s+1}, \dots, x_N]$	$N - m - h - r - s$
$(\alpha_1, \dots, \alpha_n)$	h'	$K[x_{m+h+r+s+1}, \dots, x_{m+h+r+s+h'}]$	h'
$(\beta_1, \dots, \beta_n)$	h''	$K[x_{m+h+r+s+h'+1}, \dots, x_N]$	$\leq N - m - h - r - s - h'$

Lemma 14 (Ananyan–Hochster [1, Lemma 4]). *Let $\{\alpha_1, \dots, \alpha_n\}$ and $\{\beta_1, \dots, \beta_n\}$ be homogeneous polynomials in two disjoint sets of indeterminates. Assume that the two sets of polynomials satisfy the same relations and denote the ideal of relations by P. Furthermore assume that the ideal of relations on the polynomials $\{\alpha_1 + \beta_1, \dots, \alpha_N + \beta_n\}$ contains P. Then P is generated by linear forms.*

We wish to give a flavor of the recursive argument by Ananyan–Hochster [1] that allows one to estimate the number of generators of A by applying their arguments to the case of ideals generated by three quadratic forms.

Proposition 15. *Let I be a height 2 ideal minimally generated by three quadratic forms:*

(1) If $h' = 0$, then I is an ideal in a polynomial ring generated by at most 26 linear forms.
(2) If $h' = 1$, then I is an ideal in a polynomial ring generated by at most 30 linear forms or 26 linear forms and a quadratic form.
(3) If $h' = 2$, then I is an ideal in a polynomial ring generated by at most 296 forms.

Proof. (1) If $h' = 0$, then the polynomials e_i are expressible as linear combinations of x_1, x_2 with indeterminate (primary) coefficients, quadratic terms in the primary coefficients, and mixed terms in the primary and secondary coefficients. The K-vector space V of primary coefficients has dimension at most 6 (at most 2 primary coefficients appear in each of the 3 defining equations), and consequently the vector space W spanned by secondary coefficients has dimension at most $3 \cdot 6 = 18$. Since all the tail polynomials vanish, I is an ideal of the polynomial ring on variables x_1, x_2 and the union of bases of V and W.

(2) If $h' = 1$, then the previous considerations on the polynomials e_i hold and furthermore (g_1, g_2, g_3) is an ideal of height one. Therefore $(g_1, g_2, g_3) = (yy_1, yy_2, yy_3)$ for some linear forms y, y_1, y_2, y_3 (some of the y_i could be 0) or $(g_1, g_2, g_3) = (q)$ where q is an irreducible quadratic form. In the first situation I can be written in terms of x_1, x_2, at most 6 primary coefficients, at most 18 secondary coefficients, and at most 4 linear forms y, y_1, y_2, y_3; in the second case, I can be written in terms of x_1, x_2, at most 6 primary coefficients, at most 18 secondary coefficients, and one quadratic form q.

(3) If $h' = 2$, then one proceeds by putting g_1, g_2, g_3 in standard form with respect to a set of at most 18 leading variables consisting of the secondary coefficients in the standard form of the ideal I. This produces two new front variables, at most $3 \cdot 20 = 60$ new primary coefficients, and $60 \cdot 3 = 180$ secondary coefficients, a new set of front polynomials $\alpha_1, \alpha_2, \alpha_3$ and a new set of tail polynomials $\beta_1, \beta_2, \beta_3$. Let $h'' = \mathrm{ht}(\beta_1, \beta_2, \beta_3)$.

(3a) In case $h'' \leq 1$, by cases (1) and (2), the polynomials $\beta_1, \beta_2, \beta_3$ can be written in terms of at most 30 algebraically independent forms. Together with x_1, x_2, the first 6 primary coefficients, the 18 new leading variables, the 60 new primary coefficients, and the 180 new secondary coefficients, one counts 296 algebraically independent forms.

(3b) In case $h'' = 2$, Lemma 14 yields that there are exactly 2 nonzero g_i. We count the quantities needed to write the generators of I as follows: x_1, x_2, the first 6 primary coefficients, the 18 secondary coefficients, and the two nonzero g_i. That amounts to at most 28 algebraically independent forms. □

Applying induction on the height of the ideal in a similar manner to the proof of the proposition above, Ananyan and Hochster obtain:

Theorem 16. *Let I be an ideal generated by m linear and n quadratic polynomials with* $\operatorname{ht} I = h$. *Then there exists a function $B(m,n,h)$ recursively defined by*

$$B(m,n,h) = (m+h)(n^3+n^2+n+1) + h(n+1) + B((m+h)n^2, n, h-1)$$

such that I can be viewed as an ideal in a polynomial ring of at most $B(m,n,h)$ variables.

Based on this theorem and an asymptotic analysis carried out in [1] on the growth of the function $B(m,n,h)$, we conclude:

Corollary 17. *Stillman's question 1.1 has a positive answer in the case of ideals generated by linear and quadratic polynomials. In this case, there exists a bound on projective dimension with asymptotic order of magnitude $2(m+n)^{2(m+n)}$, where m and n are the number of linear and quadratic generators of the ideal, respectively.*

We review the specific values of $B(0,3,2)$ found in Proposition 15.

Case	h	h'	h''	Linear forms	Quadratic forms	Total forms
(1)	2	0	0	≤26	0	≤26
(2)	2	1	0	≤30	≤1	≤30
(3a)	2	2	1	≤296	≤1	≤296
(3b)	2	2	2	≤26	≤2	≤28

The bounds that are found using Theorem 16 are not tight. For example, compare the estimates in the previous table with the exact bound of 4 for the projective dimension of ideals generated by three quadratic forms found in Proposition 2.

To illustrate how the idea of counting algebraically independent variables can be improved by deeper knowledge of certain parameters associated to the ideal I, we give better bounds on the projective dimension of ideals generated by three quadratic forms using knowledge of the structure of associated primes of ideals of low multiplicity. The reader is encouraged to contrast the previous bounds to the following table in which columns 2 to 4 refer to the number of algebraically independent parameters needed to write I. The first row of the table comes from the easy observation that if $I \subset (x,y)$ with x,y linear forms, then $I = (xl_{1,1} + yl_{1,2}, xl_{2,1} + yl_{2,2}, xl_{3,1} + yl_{3,2})$ with $l_{i,j}$ linear forms (possibly 0). The second row stems from the observation that if $I \subset (x,q)$ with x a linear form and q a quadratic form, then $I = (xl_1 + aq, xl_2 + bq, xl_3 + cq)$ with l_i linear forms (possibly 0) and $a,b,c \in K$. Finally if $I \subset I_C$, the defining ideal of the twisted cubic, and I is minimally generated by three quadric forms, then $I = I_C$. The three rows exhaust the possible types of minimal associated primes of height two ideals of multiplicity at most 3 (see Proposition 4 for multiplicity 2 and [16] for multiplicity 3).

$Ass(I)$	Linear forms	Quadratic forms	Total forms	$\mathrm{pd}(R/I)$
$(x, y) \in Ass(I)$	≤ 8	0	≤ 8	≤ 8
$(x, q) \in Ass(I)$	≤ 4	1	≤ 5	≤ 5
$I = I_C$	4	0	4	2

This provides heuristic evidence that the bound in Theorem 16 is far from being tight. We ask in Sect. 6 if there is a polynomial bound on the projective dimension of ideals generated by quadratic polynomials.

4 Lower Bounds and Examples

In most cases, excepting the special cases from the previous section, there is little indication of whether the answer to Stillman's question is affirmative or what the resulting bound would look like. One way to gain intuition into the question is to look for families of ideals with large projective dimension relative to the degrees of the generators. We present several such families in this section. Note that they neither prove nor disprove Stillman's question, but they do provide large lower bounds on any possible answer.

An early motivation for studying Question 1.1 comes from the study of three-generated ideals. Burch [7] proved the following theorem in the local case, which was extended by Kohn [29] to the global case. We state the polynomial ring case here.

Theorem 1 (Burch [7], Kohn [29, Theorem A]). *Let $N \in \mathbb{N}$. There exists a polynomial ring $R = K[x_1, \ldots, x_{2N}]$ and an ideal $I = (f, g, h)$ with three generators with* $\mathrm{pd}(R/I) = N + 2$.

Hence we cannot hope to find a bound on $\mathrm{pd}(R/I)$ purely in terms of the number of generators. However, if one applies this construction to a polynomial ring, the degrees of the generators grow linearly with respect to the projective dimension. Engheta computed the degrees of the three generators in the Burch–Kohn construction must be at least N, N, and $2N-2$, respectively. (See [17, Section 1.2.2].) Here is one such choice of generators.

Example 2. Let $R = K[x_1, \ldots, x_N, y_1, \ldots, y_N]$ and let

$$f = \prod_{i=1}^{N} x_i, \quad g = \prod_{i=1}^{N} y_i, \quad h = \sum_{i=1}^{N} \prod_{\substack{j=1 \\ j \neq i}}^{N} x_j y_j.$$

Then $I = (f, g, h)$ satisfies $\mathrm{pd}(R/I) = N + 2$. For example, when $N = 3$, we get the ideal

$$I = (x_1x_2x_3, y_1y_2y_3, x_2y_2x_3y_3 + x_1y_1x_3y_3 + x_1y_1x_2y_2),$$

in which case $\text{pd}(R/I) = 5$ and R/I has resolution

$$R \leftarrow R^3 \leftarrow R^9 \leftarrow R^{12} \leftarrow R^6 \leftarrow R \leftarrow 0.$$

A stronger result was later proved by Bruns. He shows [6, Satz 3] that any projective resolution is the projective resolution of some three-generated ideal after modifying the first three modules in the resolution. In practice, if one constructs a three-generated ideal with the same projective dimension as one with more generators, the degrees of the generators grow. If one could bound the growth of the generators when finding the "Brunsification" of an ideal, one could reduce the study of Stillman's question to that of three-generated ideals.

4.1 Ideals with Large Projective Dimension

The following example was given by the first author in [32]. A similar construction was given by Whieldon in [37].

Fix integers $m \geq 1, n \geq 0, d \geq 2$. Let $p = \frac{(m+d-2)!}{(m-1)!(d-1)!}$. Let $Z_1, \ldots, Z_p$ denote all the degree $d-1$ monomials in the variables $x_1, \ldots, x_m$, ordered arbitrarily. Set $R = K[x_1, \ldots, x_m, y_{1,1}, \ldots, y_{p,n}]$, a polynomial ring with $m + pn$ variables over any field K. Finally, define the ideal $I_{m,n,d}$ as

$$I_{m,n,d} = \left(x_1^d, x_2^d, \ldots, x_m^d, \sum_{i=1}^{p} Z_i y_{i,1}, \sum_{i=1}^{p} Z_i y_{i,2}, \ldots, \sum_{i=1}^{p} Z_i y_{i,n} \right).$$

Note that $I_{m,n,d}$ has $m+n$ homogeneous generators all of degree d. The following result gives a formula for the projective dimension in terms of m, n, and d.

Theorem 3 (McCullough [32, Theorem 3.3]). *With the notation above,*

$$\text{pd}(R/I_{m,n,d}) = m + np = m + n\frac{(m+d-2)!}{(m-1)!(d-1)!}.$$

The proof uses the graded Auslander–Buchsbaum theorem. (See, e.g., [14, Theorem 19.9].) One shows that $\text{depth}(R/I_{m,n,d}) = 0$, and hence $\text{pd}(R/I_{m,n,d})$ is as large as possible.

For certain choices of m, n, d, this construction yields ideals with very large projective dimension. However, the three-generated case, where $m = 2$ and $n = 1$, yields only linear growth of projective dimension.

Example 4. Let $d \in \mathbb{N}$, let $R = K[x_1, x_2, y_1, y_2, \ldots, y_d]$, and consider the ideals

$$I_{2,1,d} = \left(x_1^d, x_2^d, x_1^{d-1}y_1 + x_1^{d-2}x_2y_2 + \cdots + x_2^{d-1}y_d\right).$$

By the above theorem, $\text{pd}(R/I_{2,1,d}) = d+2$. Note that the cases $d = 2$ and $d = 3$ are given in Examples 2 and 7, respectively.

Example 5. Fix $d = 2, m = n \geq 2$ and let $R = K[x_1, x_2, \ldots, x_n, y_{1,1}, \ldots, y_{n,n}]$. Now consider the ideals

$$I_{n,n,2} = \left(x_1^2, x_2^2, \ldots, x_n^2, \sum_{i=1}^{n} x_i y_{i,1}, \sum_{i=1}^{n} x_i y_{i,2}, \ldots, \sum_{i=1}^{n} x_i y_{i,n} \right).$$

Then $I_{n,n,2}$ is generated by $2n$ quadratic polynomials and satisfies $\operatorname{pd}(R/I_{n,n,2}) = n^2 + n$. To the best of our knowledge, these are the largest projective dimension examples known for ideals generated by quadratics. So we get a lower bound of $\frac{N^2+2N}{4}$ on an answer to Stillman's question for ideals generated by N quadratic forms—much smaller than the exponential bound achieved by Ananyan and Hochster. It would be interesting to know how close either of these bounds are to being tight.

4.2 *Ideals with Larger Projective Dimension*

In this section, we construct a family of ideals with exponentially growing projective dimension relative to the degrees of the generators, even in the three-generated case. This construction can be considered as an inductive version of the family in the previous section. The family was constructed in joint work by the two authors along with Beder, Núñez–Betancourt, Snapp, and Stone in [4].

Fix integers $g \geq 2$ and a tuple of integers $m_1, \ldots, m_n$ with $m_n \geq 0$, $m_{n-1} \geq 1$, and $m_i \geq 2$ for $i = 1, \ldots, n-2$. We set $d = 1 + \sum_{i=1}^{n} m_i$. Now we define a family of sets of matrices as follows. For each $k = 0, \ldots, n$, define $\mathcal{A}_k$ to be the set of $g \times n$ matrices satisfying the following properties:

1. All entries are nonnegative integers.
2. For $i \leq k$, column i sums to m_i.
3. For $i > k$, column i contains all zeros.
4. For $i \leq \min\{k, n-1\}$, column i contains at least two nonzero entries.

These matrices are used in the definition of an ideal in the standard graded ring

$$R = K[x_{i,j}, y_{\mathsf{A}} \mid 1 \leq i \leq g, 1 \leq j \leq n, \mathsf{A} \in \mathcal{A}_n].$$

Let $\mathsf{X} = (x_{i,j})$ denote a $g \times n$ matrix of variables, and for every matrix $\mathsf{A} \in \mathcal{A}_n$, set $\mathsf{X}^{\mathsf{A}} = \prod_{i=1}^{g} \prod_{j=1}^{n} x_{i,j}^{a_{i,j}}$. We define the ideal $I_{g,(m_1,\ldots,m_n)}$ to be

$$I_{g,(m_1,\ldots,m_n)} = \left(x_{1,1}^d, \ldots, x_{m,1}^d, f \right),$$

$$\text{where } f = \sum_{k=1}^{n-1} \sum_{\mathsf{A} \in \mathcal{A}_{k-1}} \sum_{j=1}^{g} \mathsf{X}^{\mathsf{A}} x_{j,k}^{m_k} x_{j,k+1}^{d_{k+1}} + \sum_{\mathsf{B} \in \mathcal{A}_n} \mathsf{X}^{\mathsf{B}} y_{\mathsf{B}}.$$

With this notation, we have the following formula for the projective dimension.

Theorem 6 (Beder et al. [4, Corollary 3.3]). *Using the notation above, we have*

$$\operatorname{pd}(R/I_{g,(m_1,\dots,m_n)}) = \prod_{i=1}^{n-1}\left(\frac{(m_i+g-1)!}{(g-1)!(m_i)!} - g\right)\left(\frac{(m_n+g-1)!}{(g-1)!(m_n)!}\right) + gn.$$

As a result, one can define ideals with three generators in degree d and with projective dimension larger than $\sqrt{d}^{\sqrt{d}-1}$.

Corollary 7 (Beder et al. [4, Corollary 3.5]). *Over any field K and for any positive integer p, there exists an ideal I in a polynomial ring R over K with three homogeneous generators in degree p^2 such that* $\operatorname{pd}(R/I) \geq p^{p-1}$.

Proof. This follows directly from Theorem 6 by taking the ideal

$$I = I_{2,(\underbrace{p+1,\dots,p+1}_{p-1 \text{ times}},0)}.$$

□

Here we give an example of a three-generated ideal I with $d = 5$ and with $\operatorname{pd}(R/I) = 8$.

Example 8. Consider the ideal $I_{2,(3,1)}$. Since $g = 2$ and $(m_1, m_2) = (3, 1)$, this is an ideal with three degree 5 generators. We compute the sets $\mathcal{A}_k$ first:

$$\mathcal{A}_0 = \left\{\begin{pmatrix}0&0\\0&0\end{pmatrix}\right\},$$

$$\mathcal{A}_1 = \left\{\begin{pmatrix}2&0\\1&0\end{pmatrix}, \begin{pmatrix}1&0\\2&0\end{pmatrix}\right\},$$

$$\mathcal{A}_2 = \left\{\begin{pmatrix}2&1\\1&0\end{pmatrix}, \begin{pmatrix}1&1\\2&0\end{pmatrix}, \begin{pmatrix}2&0\\1&1\end{pmatrix}, \begin{pmatrix}1&0\\2&1\end{pmatrix}\right\}.$$

Then our ring is

$$R = K\left[x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2}, y_{\left(\begin{smallmatrix}2&1\\1&0\end{smallmatrix}\right)}, y_{\left(\begin{smallmatrix}1&1\\2&0\end{smallmatrix}\right)}, y_{\left(\begin{smallmatrix}2&0\\1&1\end{smallmatrix}\right)}, y_{\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)}\right],$$

and the ideal is

$$I_{2,(3,1)} = (x_{1,1}^5, x_{2,1}^5, f),$$

where

$$\begin{aligned} f &= x_{1,1}^3x_{1,2}^2 + x_{1,1}^3x_{1,2}^2 + \mathsf{X}^{\left(\begin{smallmatrix}2&1\\1&0\end{smallmatrix}\right)} y_{\left(\begin{smallmatrix}2&1\\1&0\end{smallmatrix}\right)} + \mathsf{X}^{\left(\begin{smallmatrix}1&1\\2&0\end{smallmatrix}\right)} y_{\left(\begin{smallmatrix}1&1\\2&0\end{smallmatrix}\right)} + \mathsf{X}^{\left(\begin{smallmatrix}2&0\\1&1\end{smallmatrix}\right)} y_{\left(\begin{smallmatrix}2&0\\1&1\end{smallmatrix}\right)} + \mathsf{X}^{\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)} y_{\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)} \\ &= x_{1,1}^3x_{1,2}^2 + x_{1,1}^3x_{1,2}^2 + x_{1,1}^2x_{1,2}x_{2,1}y_{\left(\begin{smallmatrix}2&1\\1&0\end{smallmatrix}\right)} + x_{1,1}x_{1,2}x_{1,2}^2y_{\left(\begin{smallmatrix}1&1\\2&0\end{smallmatrix}\right)} \\ &\quad + x_{1,1}^2x_{2,1}x_{2,2}y_{\left(\begin{smallmatrix}2&0\\1&1\end{smallmatrix}\right)} + x_{1,1}x_{2,1}^2x_{2,2}y_{\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)}. \end{aligned}$$

Then $\mathrm{pd}(R/I_{2,(3,1)}) = 8$ and the Betti table for $R/I_{2,(3,1)}$ is shown below.

	0	1	2	3	4	5	6	7	8
Total:	1	3	53	184	287	248	124	34	4
0:	1	-	-	-	-	-	-	-	-
1:	-	-	-	-	-	-	-	-	-
2:	-	-	-	-	-	-	-	-	-
3:	-	-	-	-	-	-	-	-	-
4:	-	3	-	-	-	-	-	-	-
5:	-	-	-	-	-	-	-	-	-
6:	-	-	-	-	-	-	-	-	-
7:	-	-	-	-	-	-	-	-	-
8:	-	-	3	-	-	-	-	-	-
9:	-	-	3	4	-	-	-	-	-
10:	-	-	13	46	68	56	28	8	1
11:	-	-	33	132	218	192	96	26	3
12:	-	-	1	2	1	-	-	-	-

Another interesting characteristic of this family of ideals is that it subsumes two other constructions. The construction in the previous section of ideals of the form $I_{m,1,d}$ for positive integers m, d corresponds to the ideal $I_{m,(d-1)}$ from this section, up to a relabeling of the variables. In fact, the ideals $I_{g,(m_1,\ldots,m_n)}$ may be thought of as an inductive version of the ideals in the previous section. This new family of ideals also subsumes a family of ideals studied by Caviglia [8]. Let $R = K[w, x, y, z]$ and let $C_d = \left(w^d, x^d, wy^{d-1} + xz^{d-1}\right)$. Caviglia showed that $\mathrm{reg}(R/C_d) = d^2-2$. We first note that the ideals C_d correspond to the ideals $I_{2,(1,d-2)}$, again with a relabeling of the variables. It is also noted in [4] that some of the ideals $I_{g,(m_1,m_2,\ldots,m_n)}$ have regularity larger than $d^2 - 2$. It would be interesting to compute the regularity of this new family of ideals as this would give insight into the regularity version of Stillman's question.

5 Related Bounds

While this survey is primarily concerned with Stillman's question, we want to mention some similar results that bound projective dimension in terms of data other than the degrees of the generators. This section is independent of the preceding sections.

Let I be an ideal of $R = K[x_1, \ldots, x_N]$. A *monomial support* of I is the collection of monomials that appear as terms in a set of minimal generators of I. Note that a monomial support of an ideal is not unique. Related to Stillman's question, Huneke asked if $\mathrm{pd}(R/I)$ was bounded by the number of monomials in a monomial support of I. If I is a monomial ideal generated by m monomials, then the monomial support of I has size m and the Taylor resolution of R/I has length m. Hence $\mathrm{pd}(R/I) \le m$. So Huneke's question has a positive answer for monomial ideals.

In [10], Caviglia and Kummini answer Huneke's question in the negative by constructing a family of binomial ideals whose projective dimension grows exponentially relative to the size of a monomial support. In particular, for each pair of integers $n \geq 2$ and $d \geq 2$, they construct an ideal supported by $2(n-1)(d-1)+n$ monomials with projective dimension n^d. Hence they show that any upper bound for the projective dimension of an ideal supported on m monomials counted with multiplicity is at least $2^{m/2}$. These ideals also provide lower bounds on possible answers to Stillman's question, but we present stronger examples in Sect. 4.

Several bounds on projective dimension for edge ideals are proven by Dao et al. [13]. Most notably, they prove that the projective dimension of the edge ideal of a graph with n vertices and maximal vertex degree d is bounded above by $n\left(1 - \frac{1}{2d}\right)$ [13, Corollary 5.6]. They also prove a logarithmic bound on the projective dimension of squarefree monomial ideals of height 2 satisfying Serre's condition S_k for some $k \geq 2$. (See [13, Corollary 4.10].) Several other bounds in terms of other graph parameters are given in [12].

Finally, we mention the following result of Peeva and Sturmfels. Below $R = K[x_1, \ldots, x_N]$, $\mathcal{L}$ is a sublattice of $\mathbb{Z}^n$, and $I_{\mathcal{L}}$ is the associated lattice ideal in R, that is,

$$I_{\mathcal{L}} = \left\langle \mathbf{x}^{\mathfrak{a}} - \mathbf{x}^{\mathfrak{b}} \mid \mathfrak{a}, \mathfrak{b} \in \mathbb{N}^n \text{ and } \mathfrak{a} - \mathfrak{b} \in \mathcal{L} \right\rangle.$$

In this setting, the projective dimension of $R/I_{\mathcal{L}}$ is bounded by an expression depending only the height of $I_{\mathcal{L}}$.

Theorem 1 (Peeva–Sturmfels [35, Theorem 2.3]). *The projective dimension of $R/I_{\mathcal{L}}$ as an R-module is at most* $2^{\mathrm{ht}(I_{\mathcal{L}})} - 1$.

Note that this instantly gives an answer to Stillman's question for lattice ideals since $\mathrm{ht}(I)$ is always at most the number of minimal generators of I by Krull's generalized principal ideal theorem [30, Theorem 13.5]. However, we cannot expect such a bound in terms of $\mathrm{ht}(I)$ in general. The construction by Burch–Kohn or any of the examples in Sect. 4 provide examples of ideals with fixed height and unbounded projective dimension.

6 Questions

We close by posing some specific open problems related to Stillman's question.

Question 6.1. We note that the case of an ideal I generated by three quadratics has a tight upper bound of 4 on the projective dimension of R/I. Engheta's upper bound of 36 in the case of an ideal generated by three cubics is likely far from tight. In fact, one expects that 5 is the upper bound. Can one prove this? Such a reduction will likely involve strong structure theorems on unmixed ideals of height two and low multiplicity.

Question 6.2. Similarly, Ananyan's and Hochster's exponential bound on $\mathrm{pd}(R/I)$ for ideals I generated by quadratic polynomials is likely not tight. Can one find a smaller, perhaps even polynomial bound on $\mathrm{pd}(R/I)$ where I is generated by n quadratics?

Question 6.3. There are several reductions that might make Stillman's question more tractable. Given an ideal, can one bound the degrees of the generators of the corresponding three-generated ideal produced by Bruns' theorem? If so, one could focus exclusively on three-generated ideals.

Question 6.4. Can one bound the projective dimension of all unmixed ideals of a given height and multiplicity? The structure theorems for ideals of height two and small multiplicity indicate that this might be possible and would provide information about Stillman's question.

Question 6.5. Finally, we note that there are several results showing that under certain hypotheses on an ideal, one can achieve very good bounds on the regularity of the ideal in terms of the degrees of the generators. (See, e.g., [3, 5, 11].) Are any corresponding bounds possible for projective dimension?

Acknowledgements Thanks are also due to David Eisenbud, Craig Huneke, Mircea Mustata, and Claudia Polini for organizing the Mathematical Research Community in 2010, where we were first introduced to Stillman's question. We also thank the AMS for their generous support of the MRC program. We wish to thank Jesse Beder, Bahman Engheta, Luis Núñez–Betancourt, Mike Stillman, and Gwyneth Whieldon for several useful conversations. All computations needed for the development of this chapter have been performed using *Macaulay2* [24]. Finally, we thank the anonymous referee for many useful suggestions. The second author acknowledges partial support through NSF grant DMS-0803082.

References

1. Ananyan, T., Hochster, M.: Ideals generated by quadratic polynomials. Math. Res. Lett. **19**(1), 233–244 (2012)
2. Bayer, D., Stillman, M.: On the complexity of computing syzygies. J. Symbolic Comput. **6**(2–3), 135–147 (1988). Computational aspects of commutative algebra.
3. Bayer, D., Mumford, D.: What can be computed in algebraic geometry? Computational algebraic geometry and commutative algebra (Cortona, 1991). Symposium on Mathematics, vol. XXXIV. Cambridge University Press, Cambridge, pp. 1–48 (1993)
4. Beder, J., McCullough, J., Betancourt L.N., Seceleanu, A., Snapp, B., Stone, B.: Ideals with larger projective dimension and regularity. J. Sym. Comp. **46**, 1105–1113 (2011)
5. Bertram, A., Ein, L., Lazarsfeld, R.: Vanishing theorems, a theorem of Severi, and the equations defining projective varieties. J. Amer. Math. Soc. **4**(3), 587–602 (1991)
6. Bruns, W.: "Jede" endliche freie Auflösung ist freie Auflösung eines von drei Elementen erzeugten Ideals. J. Algebra **39**(2), 429–439 (1976)
7. Burch, L.: A note on the homology of ideals generated by three elements in local rings. Proc. Cambridge Philos. Soc. **64**, 949–952 (1968)
8. Caviglia, G.: Kozul algebras, Castelnuovo-Mumford regularity and generic initial ideals. Ph.D. thesis, University of Kansas, Lawrence (2004)

9. Caviglia, G., Sbarra, E.: Characteristic-free bounds for the Castelnuovo-Mumford regularity. Compos. Math. **141**(6), 1365–1373 (2005)
10. Caviglia, G., Kummini, M.: Some ideals with large projective dimension. Proc. Amer. Math. Soc. **136**(2), 505–509 (2008)
11. Chardin, M., Ulrich, B.: Liaison and Castelnuovo-Mumford regularity. Amer. J. Math. **124**(6), 1103–1124 (2002)
12. Dao, H., Schweig, J.: Projective dimension, graph domination parameters, and independence complex homology. (preprint) arXiv: 1110.2841v2
13. Dao, H., Huneke, C., Schweig, J.: Bounds on the regularity and projective dimension of ideals associated to graphs. J. Algebraic Combin. 1–19 (2012)
14. Eisenbud, D.: Commutative algebra. In: Graduate Texts in Mathematics, vol. 150, Springer, New York (1995). With a view toward algebraic geometry
15. Eisenbud, D.: The geometry of syzygies. In: Graduate Texts in Mathematics, vol. 229, Springer, New York (2005). A second course in commutative algebra and algebraic geometry
16. Eisenbud, D., Harris, J.: On varieties of minimal degree. Proc. Sympos. Pure Math. **46**, 3–13 (1987)
17. Engheta, B.: Bounds on projective dimension. Ph.D. thesis, University of Kansas (2005)
18. Engheta, B.: On the projective dimension and the unmixed part of three cubics. J. Algebra **316**(2), 715–734 (2007)
19. Engheta, B.: Bound on the multiplicity of almost complete intersections. Comm. Algebra **37**(3), 948–953 (2009)
20. Engheta, B.: A bound on the projective dimension of three cubics. J. Symbolic Comput. **45**(1), 60–73 (2010)
21. Fan, C.-T.: Primary decompositions of three-generated ideals. J. Pure Appl. Algebra **105**(2), 167–182 (1995)
22. Galligo, A.: Théorème de division et stabilité en géométrie analytique locale. Ann. Inst. Fourier (Grenoble) **29**(2), vii, 107–184 (1979)
23. Giusti, M.: Some effectivity problems in polynomial ideal theory. EUROSAM 84 (Cambridge, 1984). In: Lecture Notes in Computer Science, vol. 174, pp. 159–171. Springer, Berlin (1984)
24. Grayson, D.R., Stillman, M.E.: Macaulay2, a software system for research in algebraic geometry. Available at http://www.math.uiuc.edu/Macaulay2/
25. Hermann, G.: Die Frage der endlich vielen Schritte in der Theorie der Polynomideale. Math. Ann. **95**(1), 736–788 (1926)
26. Hilbert, D.: Ueber die Theorie der algebraischen Formen. Math. Ann. **36**(4), 473–534 (1890)
27. Huneke, C., Ulrich, B.: Residual intersections. J. Reine Angew. Math. **390**, 1–20 (1988)
28. Koh, J.: Ideals generated by quadrics exhibiting double exponential degrees. J. Algebra **200**(1), 225–245 (1998)
29. Kohn, P.: Ideals generated by three elements. Proc. Amer. Math. Soc. **35**, 55–58 (1972)
30. Matsumura, H.: Commutative ring theory. In: Cambridge Studies in Advanced Mathematics, vol. 8, 2nd edn. Cambridge University Press, Cambridge (1989). Translated from the Japanese by M. Reid
31. Mayr, E.W., Meyer, A.R.: The complexity of the word problems for commutative semigroups and polynomial ideals. Adv. Math. **46**(3), 305–329 (1982)
32. McCullough, J.: A family of ideals with few generators in low degree and large projective dimension. Proc. Amer. Math. Soc. **139**(6), 2017–2023 (2011)
33. Mumford, D.: Varieties defined by quadratic equations. Questions on Algebraic Varieties (C.I.M.E. III Ciclo, Varenna 1969), Edizioni Cremonese, Rome, pp. 29–100 (1970)
34. Peeva, I.: Graded syzygies. In: Algebra and Applications, vol. 14. Springer, London (2011). MR 2560561 (2011j:13015)
35. Peeva, I., Sturmfels, B.: Syzygies of codimension 2 lattice ideals. Math. Z. **229**(1), 163–194 (1998)
36. Peeva, I., Stillman, M.: Open problems on syzygies and Hilbert functions. J. Commut. Algebra **1**(1), 159–195 (2009)
37. Whieldon, G.: Bounding invariants of Stanley–Reisner ideals. Ph.D. thesis, Cornell University, New York (2011)

Brauer–Thrall Theory for Maximal Cohen–Macaulay Modules

Graham J. Leuschke and Roger Wiegand

1 Introduction

The Brauer–Thrall Conjectures first appeared in a 1957 paper by Thrall's student Jans [25]. They say, roughly speaking, that if a finite-dimensional algebra A over a field k has infinite representation type, then A has lots of big indecomposable finitely generated modules. Recall that A has *finite representation type* provided there are only finitely many indecomposable finitely generated A-modules up to isomorphism, *bounded representation type* provided there is a bound on the k-dimensions of the indecomposable finitely generated A-modules, and *strongly unbounded representation type* provided there is an infinite strictly increasing sequence (n_i) of positive integers such that A has, for each i, infinitely many non-isomorphic indecomposable modules of k-dimension n_i. Here are the conjectures:

Conjecture 1 (First Brauer–Thrall Conjecture (BT1)). If A has bounded representation type, then A has finite representation type.

Conjecture 2 (Second Brauer–Thrall Conjecture (BT2)). If A has unbounded representation type and k is infinite, then A has strongly unbounded representation type.

Under mild hypotheses, both conjectures are now theorems. Roĭter [39] verified (BT1), and Nazarova and Roĭter [35] proved (BT2) for perfect fields k. See [38] or [20] for some history on these results.

G.J. Leuschke
Department of Mathematics, Syracuse University, Syracuse, NY 13244, USA
e-mail: gjleusch@math.syr.edu

R. Wiegand (✉)
Department of Mathematics, University of Nebraska–Lincoln, Lincoln, NE 68588, USA
e-mail: rwiegand@math.unl.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_18,

When we move from Artinian rings to local rings $(R, \mathfrak{m}, \mathsf{k})$ of positive dimension, the first thing we need to do is to decide on the right class of modules. If R is not a principal ideal ring, constructions going back to Kronecker [29] and Weierstraß [41] show that R has indecomposable modules requiring arbitrarily many generators. Moreover, if k is infinite, then for every n there are $|\mathsf{k}|$ non-isomorphic indecomposable modules each of which requires exactly n generators. (See [33, Theorem 3.3 and Exercise 3.25].) Thus imposing finiteness or boundedness conditions on the class of *all* modules does not lead to anything interesting.

Restricting to torsion-free modules yields a more robust theory, at least in dimension one. In the 1960s Jacobinski [24] and, independently, Drozd and Roĭter [16] studied orders in algebraic number fields and, more generally, rings essentially module-finite over the ring of integers, and classified the rings having only finitely many indecomposable finitely generated torsion-free modules up to isomorphism.

In dimensions greater than one, there are just too many torsion-free modules. Indeed, Bass [5] proved in 1962 that every local domain of dimension two or more has indecomposable finitely generated torsion-free modules of arbitrarily large rank.

The maximal Cohen–Macaulay (MCM) property, a higher-dimensional form of torsion-freeness, turns out to give a fruitful class of modules to study. The equality of a geometric invariant (dimension) with an arithmetic one (depth) makes MCM modules easy to work with, simultaneously ensuring that in some sense they faithfully reflect the structure of the ring. For example, a Cohen–Macaulay local ring has no non-free MCM modules if and only if it is a regular local ring, so the rings that are the simplest homologically are also simple in this sense. Imposing finiteness or boundedness conditions on MCM modules over a Cohen–Macaulay local ring leads to classes of rings that are large enough to include interesting examples, but small enough to study effectively. The seminal work of Herzog [22], Artin and Verdier [1], Auslander [3], and Buchweitz, Greuel, Knörrer, and Schreyer [8, 28] supports this assertion. For example, the main result of [8, 28] is that a complete equicharacteristic hypersurface singularity over an algebraically closed field of characteristics zero has only finitely many indecomposable MCM modules up to isomorphism if and only if it is a simple singularity in the sense of V. I. Arnol′d, that is, one of the (A_n), (D_n), (E_6), (E_7), or (E_8) hypersurface singularities.

Next, we have to decide what invariant should be used to measure the size of a finitely generated module M. Two obvious choices are $\mu_R(M)$, the minimal number of generators required for M, and $\mathrm{e}_R(M)$, the multiplicity of M. We choose multiplicity.

Definition 3. Let $(R, \mathfrak{m}, \mathsf{k})$ be a local ring:

(i) R has *finite* CM type provided R has, up to isomorphism, only finitely many indecomposable MCM modules.
(ii) R has *bounded* CM type provided there is a bound on the multiplicities of the indecomposable MCM R-modules.
(iii) R has *strongly unbounded* CM type provided there is an increasing sequence $n_1 < n_2 < \cdots$ of positive integers such that, for every i, there are infinitely many indecomposable MCM modules of multiplicity n_i.

Here, then, are the Brauer–Thrall Conjectures for MCM modules:

Conjecture 4 (First Brauer–Thrall Conjecture for MCM modules (BTM1)). If a local ring $(R, \mathfrak{m}, \mathsf{k})$ has bounded CM type, then R has finite CM type.

Conjecture 5 (Second Brauer–Thrall Conjecture for MCM modules (BTM2)). If a local ring $(R, \mathfrak{m}, \mathsf{k})$ has unbounded CM type and k is infinite, then R has strongly unbounded CM type.

For MCM modules, multiplicity and number of generators enjoy a linear relationship:

$$\mu_R(M) \leq \mathrm{e}_R(M) \leq \mathrm{e}(R) \cdot \mu_R(M)\,, \tag{1}$$

for every MCM R-module. (See [33, Corollary A.24] for a proof of the first inequality.) It follows that we could replace multiplicity by number of generators in Definition 3 without changing the class of rings satisfying bounded (respectively, strongly unbounded) CM type.

In fact, Conjecture 4 is false, the designation "conjecture" being merely a convenient nod to history. The first counterexample was given by Dieterich in 1980 [14]. Let k be a field of characteristic two, let $A = \mathsf{k}[\![x]\!]$, and let G be the two-element group. Then AG has bounded but infinite CM type. Of course AG is isomorphic to $\mathsf{k}[\![x, y]\!]/(y^2)$, which, as we will see in the next section, has bounded but infinite CM type for *any* field k.

Conventions and Notation 1.6. Throughout, R will be a local ring (always assumed to be commutative and Noetherian). The notation $(R, \mathfrak{m}, \mathsf{k})$ indicates that $\mathfrak{m}$ is the maximal ideal of R and that k is the residue field $R/\mathfrak{m}$. All modules are assumed to be finitely generated. The $\mathfrak{m}$-adic completion of R is $\widehat{R}$, and the integral closure of R in its total quotient ring $K := \{\text{non-zerodivisors}\}^{-1}R$ is $\overline{R}$. A module M is *maximal Cohen–Macaulay* (abbreviated "MCM") provided $\mathrm{depth}(M) = \dim(R)$. We will denote the multiplicity $\mathrm{e}(\mathfrak{m}, M)$ of the maximal ideal on M simply by $\mathrm{e}_R(M)$, and we write $\mathrm{e}(R)$ instead of $\mathrm{e}_R(R)$. (See [34, Chap. 14].) The modifier "Cohen–Macaulay," when applied to the ring R, will often be abbreviated "CM." Our standard reference for matters commutative algebraic will be [34], and for representation theory we refer to [33] or [47].

2 Dimension One

Before getting started, let us observe that both conjectures are true for local Artinian rings. In this case all finitely generated modules are MCM modules. If $(R, \mathfrak{m})$ is an Artinian principal ideal ring with $\mathfrak{m}^t = 0$, the indecomposable modules are $R/\mathfrak{m}^i$, $1 \leq i \leq t$. We have already observed that if R is not a principal ideal ring, then there exist, for each $n \geq 1$, indecomposable modules requiring exactly n generators and, if k is infinite, $|\mathsf{k}|$ of them.

Now, on to dimension one! We recall the characterization of one-dimensional rings of finite CM type:

Theorem 1. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a Cohen–Macaulay local ring of dimension one. Then R has finite CM type if and only if:*

(i) R is reduced.
(ii) $\mu_R(\overline{R}) \le 3$.
(iii) $\frac{\mathfrak{m}\overline{R}+R}{R}$ is cyclic as an R-module.

Items (i) and (ii) are equivalent to the condition that $\widehat{R}$ is reduced and $\mathrm{e}(R) \le 3$. Conditions (ii) and (iii) are often called the "Drozd–Roĭter conditions" [11] to recognize the 1966 paper [16] where they first appeared and were shown to characterize the rings of finite CM type among local rings essentially module-finite over $\mathbb{Z}$. The work of Drozd and Roĭter was clarified considerably in 1978 by Green and Reiner [19], who used explicit matrix reductions to verify finiteness of CM type in the presence of the Drozd–Roĭter conditions. In 1989 Wiegand [42] adapted constructions in [16] to prove the "only if" direction in general. A separable base-change argument in [42] and the matrix decompositions of Green and Reiner verified the "if" direction, except in the case of an imperfect residue field of characteristic two or three. In [44] Wiegand took care of the case of characteristic three. Finally, Çimen, in his Ph.D. dissertation [9], completed the proof of Theorem 1 via difficult matrix reductions. (Cf. [10].)

Although we will not say much about non-CM rings, we record the following result from [44], which, together with Theorem 1, characterizes the one-dimensional local rings with finite CM type:

Theorem 2. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a one-dimensional local ring, not necessarily Cohen–Macaulay, and let N be the nilradical of R. Then R has finite CM type if and only if (i) R/N (which* is *CM) has finite CM type, and (ii) $N \cap \mathfrak{m}^n = 0$ for some positive integer n.*

The proof of the "only if" direction in Theorem 1 (necessity of the Drozd–Roĭter conditions) in [42] actually shows more and confirms BTM1 in the analytically unramified case. We will say that a finitely generated module M over a CM local ring R has *constant rank n* provided $K \otimes_R M \cong K^{(n)}$, where K is the total quotient ring. Equivalently, $M_\mathfrak{p}$ is a free $R_\mathfrak{p}$-module of rank n for every minimal prime ideal $\mathfrak{p}$ of R. In this case $\mathrm{e}(M) = n\mathrm{e}(R)$.

Theorem 3 (BTM1 when $\widehat{R}$ is reduced, [42]). *Let $(R, \mathfrak{m}, \mathsf{k})$ be a one-dimensional local ring with reduced completion. If R has infinite CM type then for every n there is an indecomposable MCM R-module of constant rank n. In particular, R has unbounded CM type.*

We have already seen that BTM1 can fail if there are nilpotents. We showed in 2005 [32, Theorem 2.4] that there are essentially only three counterexamples to BTM1 in dimension one. Recall that the complete (A_∞) and (D_∞) curve singularities are, respectively, the rings $\mathsf{k}[\![x,y]\!]/(y^2)$ and $\mathsf{k}[\![x,y]\!]/(xy^2)$. They arise as the respective limits of the (A_n) singularities $\mathsf{k}[\![x,y]\!]/(y^2+x^{n+1})$ and the (D_n) singularities $\mathsf{k}[\![x,y]\!]/(xy^2+x^{n-1})$ as $n \longrightarrow \infty$.

Theorem 4 (Failure of BTM1, [32]). *Let $(R,\mathfrak{m},\mathsf{k})$ be an equicharacteristic, one-dimensional, Cohen–Macaulay local ring, with k infinite. Then R has bounded but infinite CM type if and only if the completion $\widehat{R}$ is isomorphic to one of the following:*

(i) $\mathsf{k}[\![x,y]\!]/(y^2)$*, the ($A_\infty$) singularity*
(ii) $T := \mathsf{k}[\![x,y]\!]/(xy^2)$*, the ($D_\infty$) singularity*
(iii) $E := \mathrm{End}_T(\mathfrak{m}_T)$*, the endomorphism ring of the maximal ideal of T*

The ring E has a presentation $E \cong \mathsf{k}[\![X,Y,Z]\!]/(XY,YZ,Z^2)$.

The assumption that k be infinite is annoying. It is tempting to try to eliminate this assumption via the flat local homomorphism $R \longrightarrow S := R[z]_{\mathfrak{m}R[z]}$, where z is an indeterminate. The problem would be to show that if R has unbounded CM type then so has S. While it is rather easy to show that finite CM type descends along flat local homomorphisms (as long as the closed fiber is CM) [45, Theorem 1.6], it is not known (at least to us) whether an analogous result holds for descent of *bounded* CM type. In fact, it is not even known, in higher dimensions, whether bounded CM type descends from the completion. Such descent was a crucial part of the proof of Theorem 4, but the proof of descent was based not on abstract considerations but on the precise presentations, in [8], of the indecomposable $\widehat{R}$-modules in each of the three cases. Using these presentations, we were able to say exactly which MCM $\widehat{R}$-modules are extended from R-modules, and thereby deduce that R itself has bounded CM type. Part of the difficulty in proving a general statement of this form is that there may be no uniform bound on the number of indecomposable MCM $\widehat{R}$-modules required to decompose the completion of an indecomposable MCM R-module (see [33, Example 17.11]).

At this point we have shown, for CM local rings of dimension one, that BTM1 holds in the analytically unramified case but fails (just a little bit) in general. We turn now to BTM2 for CM local rings of dimension one.

Theorem 5. *Let $(R,\mathfrak{m},\mathsf{k})$ be a one-dimensional local Cohen–Macaulay ring with unbounded CM type and with k infinite. Assume either (i) $\widehat{R}$ is reduced Or (ii) R contains a field.*
Then, for each positive integer n, R has $|\mathsf{k}|$ pairwise non-isomorphic indecomposable MCM modules of constant rank n. In particular, BTM2 holds for one-dimensional CM local rings that satisfy either (i) or (ii).

Karr and Wiegand [26, Theorem 1.4] proved this in the analytically unramified case (i). Later Leuschke and Wiegand modified that proof, using ideas from [31,32], to prove the result in the equicharacteristic case (ii). See [33, Theorem 17.10]. The rest of this section is devoted to a sketch of the main ideas of the proof of Theorem 5.

Assume, for the rest of this section, that $(R, \mathfrak{m}, \mathsf{k})$ is a one-dimensional CM local ring satisfying the hypotheses of Theorem 5. In particular, k is infinite and R has unbounded CM type. The first step, proved by Bass [6] in the analytically unramified case, appears as Theorem 2.1 of [31]:

Lemma 6. *Suppose* $\mathrm{e}(R) \leq 2$. *Then every indecomposable MCM R-module is isomorphic to an ideal of R and hence has multiplicity at most two.*

Thus we may assume that $\mathrm{e}(R) \geq 3$. Then R has a finite birational extension S (an intermediate ring between R and its total quotient ring K such that S is finitely generated as an R-module) with $\mu_R(S) = \mathrm{e}(R)$. Although we will need to choose S with some care, we note here that $S := \bigcup_{n \geq 1} \mathrm{End}_R(\mathfrak{m}^n)$ has the right number of generators. (See [31, Lemma 2.6].) In the analytically unramified case, one typically takes $S = \overline{R}$. (Notice that none of this works if R is not CM, since $R = K$ in that case!) Let $\mathfrak{f}$ be the conductor, that is, the largest ideal of S that is contained in R. Putting $A = R/\mathfrak{f}$, $B = S/\mathfrak{f}$, and $D = B/\mathfrak{m}B$, we obtain a commutative diagram

$$\begin{array}{ccc} R & \longrightarrow & S \\ \downarrow & & \downarrow \\ A & \longrightarrow & B \\ \downarrow & & \downarrow \\ \mathsf{k} & \longrightarrow & D \end{array} \tag{2}$$

in which the top square is a pullback and D is a k-algebra of dimension $\mathrm{e}(R)$.

Now let n be a fixed positive integer, and let $t \in \mathsf{k}$. We wish to build a family, parametrized by t, of indecomposable MCM R-modules of constant rank n. The following construction [42, Construction 2.5], [33, Construction 3.13] is based on work of Drozd and Roĭter [16]. Let I be the $n \times n$ identity matrix and H be the nilpotent $n \times n$ Jordan block with 1's on the superdiagonal and 0's elsewhere. Let α and β be elements of D such that $\{1, \alpha, \beta\}$ is linearly independent over k. (Eventually we will have to impose additional restrictions on α and β.) Let V_t be the k-subspace of $D^{(n)}$ spanned by the columns of the $n \times 2n$ matrix:

$$\Psi_t := [I \quad \alpha I + \beta(tI + H)] \ . \tag{3}$$

Let $\pi \colon S^{(n)} \twoheadrightarrow D^{(n)}$ be the canonical surjection, and define M_t by the following pullback diagram:

$$\begin{array}{ccc} M_t & \longrightarrow & S^{(n)} \\ \downarrow & & \downarrow \pi \\ V_t & \hookrightarrow & D^{(n)} \end{array} \tag{4}$$

Then M_t is an MCM R-module of constant rank n, and it is indecomposable if the *pair* $V_t \subseteq D^{(n)}$ is indecomposable in the following sense: There is no idempotent endomorphism ε of $D^{(n)}$, other than 0 and the identity, such that $\varepsilon(V_t) \subseteq V_t$. Moreover if $tu \in \mathsf{k}$, and $M_t \cong M_u$, then the pairs $(V_t \subseteq D^{(n)})$ and $(V_u \subseteq D^{(n)})$ are isomorphic, in the sense that there is an automorphism φ of $D^{(n)}$ such that $\varphi(V_t) \subseteq V_u$. Our goal, then, is to choose α and β so that we get $|\mathsf{k}|$ non-isomorphic indecomposable pairs $(V_t \subseteq D^{(n)})$.

Suppose first that $\mathrm{e}(R) = 3$. We need to choose a finite birational extension $R \subset S$ such that

$$\mu_R(S) = 3 \quad \text{and} \quad \mu_R\left(\frac{\mathfrak{m}S + R}{R}\right) \geq 2\,. \tag{5}$$

If R is analytically unramified, the assumption that R has unbounded (hence infinite) CM type implies failure of the second Drozd–Roĭter condition (iii) in Theorem 1, and we can take $S = \overline{R}$. If R is analytically ramified but contains a field, the fact that $\widehat{R}$ is *not* one of the three exceptional rings of Theorem 4 leads, after substantial computation, to the right choice for S. (See the proof of [32, Theorem 1.5] or the proofs of Theorems 17.6 and 17.9 in [33].)

Now, with our carefully chosen birational extension $R \longrightarrow S$, we have

$$\dim_{\mathsf{k}}(B/\mathfrak{m}B) = 3 \quad \text{and} \quad \dim_{\mathsf{k}}\left(\frac{\mathfrak{m}B + A}{\mathfrak{m}^2B + A}\right) \geq 2\,, \tag{6}$$

for the Artinian rings A and B in the diagram (2). Put $C = \mathfrak{m}B + A$, and choose elements $x, y \in \mathfrak{m}B$ so that their images in $\frac{\mathfrak{m}B+A}{\mathfrak{m}^2B+A}$ are linearly independent. Since $C/\mathfrak{m}C$ maps onto $\frac{\mathfrak{m}B+A}{\mathfrak{m}^2B+A}$, the images α and β of x and y in $C/\mathfrak{m}C$ are linearly independent. By [33, Lemmas 3.10 and 3.11] it suffices to build the requisite pairs $(V_t \subseteq (C/\mathfrak{m}C)^{(n)})$, since these will yield, via extension, non-isomorphic indecomposable pairs $(V_t \subseteq D^{(n)})$. Moreover, with this choice of α and β, we have the relations

$$\alpha^2 = \alpha\beta = \beta^2 = 0\,. \tag{7}$$

Returning to the general case $\mathrm{e}(R) \geq 3$, we may assume that either $\dim_{\mathsf{k}}(D) \geq 4$ or else D contains elements α and β satisfying (7). In order to show that there are enough values of t that produce non-isomorphic indecomposable pairs $(V_t \subseteq D^{(n)})$, we let t and u be elements of k, not necessarily distinct, and suppose that φ is a k-endomorphism of $D^{(n)}$ that carries V_t into V_u. We regard φ as an $n \times n$ matrix with entries in D. Recalling that V_t is the column space of the matrix Ψ_t in (3), we see that the condition $\varphi V_t \subseteq V_u$ yields a $2n \times 2n$ matrix θ over k satisfying the equation

$$\varphi\Psi_t = \Psi_u\theta\,. \tag{8}$$

Write $\theta = \left[\begin{smallmatrix} E & F \\ P & Q \end{smallmatrix}\right]$, where E, F, P, and Q are $n \times n$ blocks. Then (8) gives the following two equations:

$$\begin{aligned} \varphi &= E + \alpha P + \beta(uI + H)P \\ \alpha\varphi + \beta\varphi(tI + H) &= F + \alpha Q + \beta(uI + H)Q\,. \end{aligned} \tag{9}$$

Substituting the first equation into the second and combining terms, we get the following equation:

$$\begin{aligned} -F + \alpha(E - Q) + \beta(tE - uQ + EH - HQ) + (\alpha + t\beta)(\alpha + u\beta)P \\ + \alpha\beta(HP + PH) + \beta^2(HPH + tHP + uPH) = 0\,. \end{aligned} \tag{10}$$

Suppose there exist elements α and β satisfying Equation (7). With this choice of α and β, (10) collapses:

$$-F + \alpha(E - Q) + \beta(tE - uQ + EH - HQ) = 0\,. \tag{11}$$

Since all capital letters in (11) represent matrices over k and since $\{1, \alpha, \beta\}$ is linearly independent over k, we get the equations

$$F = 0\,, \qquad E = Q\,, \qquad \text{and} \qquad (t - u)E + EH - HE = 0\,.$$

After a bit of fiddling (see [33, Case 3.14] for the details) we reach two conclusions:

(i) If φ is invertible, then $t = u$. Thus the modules are pairwise non-isomorphic.
(ii) If $t = u$ and φ is idempotent, then φ is either 0 or I. Thus all of the modules are indecomposable.

The key issue in these computations is that the matrix H is non-derogatory, so that its commutator in the full matrix ring is just the local ring $\mathsf{k}[H] \cong \mathsf{k}[X]/(X^n)$.

We may therefore assume that $\dim_{\mathsf{k}}(D) \geq 4$. With a little luck, the algebra D might contain an element α that does *not* satisfy a nontrivial quadratic relation over k. In this case, we choose any element $\beta \in D$ so that $\{1, \alpha, \alpha^2, \beta\}$ is linearly independent, and we set

$$G = \{t \in \mathsf{k} \mid \{1, \alpha, \beta, (\alpha + t\beta)^2\} \text{ is linearly independent}\}\,.$$

This set is nonempty and Zariski-open and hence cofinite in k. For $t \in G$, put

$$G_t = \{u \in G \mid \{1, \alpha, \beta, (\alpha + t\beta)(\alpha + u\beta)\} \text{ is linearly independent}\}\,.$$

Then G_t is cofinite in G for each $t \in G$. Moreover, one can check the following, using the mess (10):

(i) If t and u are distinct elements of G with $u \in G_t$, then φ is not an isomorphism.
(ii) If $t = u \in G$ and φ is idempotent, then φ is either 0 or I.

The desired conclusions follow easily. (See [33, Case 3.16] for the details.)

The remainder of the proof [33, (3.17)–(3.21)] is a careful analysis of the k-algebras D in which every element is quadratic over k. (The fact that k is infinite obviates consideration of the last case [33, Case 3.22], where our construction does not work and Dade's construction [13] is used to produce *one* indecomposable of rank n.)

In studying direct-sum decompositions over one-dimensional local rings, it is important to know about indecomposable MCM modules of nonconstant rank. (See [43], where the first author determined exactly how badly Krull–Remak–Schmidt uniqueness can fail.) If $(R, \mathfrak{m}, \mathsf{k})$ is a one-dimensional, analytically unramified local ring with minimal prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_s$, we define the *rank* of a module to be the s-tuple $(r_1, \dots, r_s)$, where r_i is the dimension of $(M_{\mathfrak{p}_i})$ as a vector space over the field $R_{\mathfrak{p}_i}$. Crabbe and Saccon [12] have recently proved the following:

Theorem 7. *Let $(R, \mathfrak{m}, \mathsf{k})$ be an analytically unramified local ring of dimension one, with minimal prime ideals $\mathfrak{p}_1, \dots, \mathfrak{p}_s$. Assume that $R/\mathfrak{p}_1$ has infinite CM type. Let $\underline{r} := (r_1, \dots, r_s)$ be an arbitrary s-tuple of nonnegative integers with $r_1 \geq r_i$ for each i and with $r_1 > 0$. Then there is an indecomposable MCM R-module with* $\operatorname{rank}(M) = \underline{r}$*, and $|\mathsf{k}|$ non-isomorphic ones if k is infinite.*

3 Brauer–Thrall I for Hypersurfaces

In Theorem 4 we saw that there are just two plane curve singularities that contradict BTM1. Here we promote this result to higher-dimensional hypersurfaces, with the following theorem from [31] (cf. [33, Theorem 17.5]):

Theorem 1. *Let k be an algebraically closed field of characteristic different from two, and let $R = \mathsf{k}[\![x_0, \dots, x_d]\!]/(f)$, where f is a nonzero element of $(x_0, \dots, x_d)$ and $d \geq 2$. Then R has bounded but infinite CM type if and only if $R \cong \mathsf{k}[\![x_0, \dots, x_d]\!]/(g + x_2^2 + \cdots + x_d^2)$, where g is a polynomial in $\mathsf{k}[x_0, x_1]$ defining either an (A_∞) or (D_∞) curve singularity.*

This theorem and its proof are modeled on the beautiful result of Buchweitz, Greuel, Knörrer, and Schreyer, where "bounded but infinite" is replaced by "finite," and the singularities in the conclusion are the *simple* or *ADE* singularities, [33, §4.3].

The "if" direction of Theorem 1 hinges on the following result (see [33, Theorem 17.2]):

Lemma 2 (Knörrer [28]). *Let k be a field, and put $S = \mathsf{k}[\![x_0, \dots, x_d]\!]$. Let f be a non zero non-unit of S, $R = S/(f)$ and $R^\# = S[\![z]\!]/(f + z^2)$.*

(i) If $R^\#$ has finite (respectively, bounded) CM type, so has R.
(ii) Assume R has finite (respectively, bounded) CM type and $\operatorname{char}(\mathsf{k}) \neq 2$. *Then $R^\#$ has finite (respectively, bounded) CM type. More precisely, if $\mu_R(M) \leq B$ for every indecomposable MCM R-module M, then $\mu_{R^\#}(N) \leq 2B$ for every indecomposable MCM $R^\#$-module N.*

For the "only if" direction, we need Lemma 2 and the following result due to Kawasaki [27, Theorem 4.1]:

Lemma 3. *Let $(R, \mathfrak{m})$ be a d-dimensional abstract hypersurface (a local ring whose completion $\widehat{R}$ has the form $S/(f)$, where $(S, \mathfrak{n})$ is a regular local ring and $f \in \mathfrak{n}$). Let n be any positive integer, and let M be the $(d+1)^{st}$ syzygy of $R/\mathfrak{m}^n$. If* $\mathrm{e}(R) > 2$, *then M is an indecomposable MCM R-module, and $\mu_R(M) \geq \binom{d+n-1}{d-1}$. In particular, if $d \geq 2$ then R has unbounded CM type.*

If, now, $d \geq 2$ and R (as in Theorem 1) has bounded but infinite CM type, then $\mathrm{e}(R) \leq 2$. Using the Weierstraß Preparation Theorem and a change of variables, we can put f into the form $g + x_d^2$, with $g \in \mathsf{k}[\![x_0, \ldots, x_{d-1}]\!]$. Then $\mathsf{k}[\![x_0, \ldots, x_{d-1}]\!]/(g)$ has bounded but infinite CM type, by Lemma 2. We repeat this process till we get down to dimension one, and then we invoke Theorem 4.

4 Brauer–Thrall I for Excellent Isolated Singularities

The starting point here is the Harada–Sai Lemma [21, Lemmas 11 and 12], sharpened by Eisenbud and de la Peña in 1998 [17]. By a *Harada–Sai sequence*, we mean a sequence

$$M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{s-1}} M_s$$

of R-homomorphisms satisfying:

(i) Each M_i is indecomposable of finite length.
(ii) No f_i is an isomorphism.
(iii) The composition $f_{s-1} f_{s-2} \cdots f_1$ is nonzero.

Lemma 1 (Harada–Sai Lemma). *With the notation above, suppose $\ell_R(M_i) \leq b$ for each i. Then $s \leq 2^b - 1$.*

In fact, Eisenbud and de la Peña [17] characterized the integer sequences that can occur in the form $(\ell_R(M_1), \ldots, \ell_R(M_s))$ for some Harada–Sai sequence over some R. In order to apply Harada–Sai to MCM modules, we need to reduce modulo a suitable system of parameters to get down to the Artinian case. Of course, an arbitrary system of parameters will not work, since indecomposability and non-isomorphism will not be preserved. What we need is a *faithful system of parameters*, that is, a system of parameters $\underline{x} = x_1, \ldots, x_d$ such that $\underline{x}\operatorname{Ext}^1_R(M, N) = 0$ for every MCM R-module M and every finitely generated R-module N. Here are some useful properties of faithful systems of parameters (where we write $\underline{x}^2$ for the system of parameters $(x_1^2, \ldots, x_d^2)$):

Lemma 2. *Let $\underline{x}$ be a faithful system of parameters for a CM local ring R.*

(i) Let M and N be MCM R-modules, and suppose $\varphi\colon M/\underline{x}^2M \longrightarrow N/\underline{x}^2N$ is an isomorphism. There is an isomorphism $\tilde{\varphi}\colon M \longrightarrow N$ such that $\tilde{\varphi}\otimes_R (R/(\underline{x})) = \varphi \otimes_R (R/(\underline{x}))$.
(ii) Let s : $0 \longrightarrow N \longrightarrow E \longrightarrow M \longrightarrow 0$ be a short exact of MCM R-modules. Then s splits if and only if $s \otimes_R (R/(\underline{x}^2))$ splits.
(iii) Assume R is Henselian, and let M be an indecomposable MCM R-module. Then $M/\underline{x}^2M$ is indecomposable.

Using these properties, one obtains the Harada–Sai Lemma for MCM modules [33, Theorem 15.19].

Lemma 3. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a CM, Henselian local ring and $\underline{x}$ a faithful system of parameters. Let*

$$M_1 \xrightarrow{f_1} M_2 \xrightarrow{f_2} \cdots \xrightarrow{f_{s-1}} M_s$$

be a sequence of R-homomorphisms, with each M_i indecomposable and MCM. Assume that

$$(f_{s-1} f_{s-2} \cdots f_1) \otimes_R (R/(\underline{x}^2)) \neq 0\,.$$

If $\ell_R(M_i/\underline{x}M_i) \le b$ for all i, then $s \le 2^b - 1$.

Suppose, instead, that we have a bound, say, B, on the multiplicities $\mathrm{e}(M_i)$. Choosing t such that $\mathfrak{m}^t \subseteq (\underline{x}^2)$, we get a bound $b := Bt^d$ on the lengths of the modules $M_i/\underline{x}^2M_i$. A walk around the AR quiver of R then proves BTM1. (See [47, Chap. 6] or [33, §15.3].) Of course, none of this does any good unless the ring R *has* a faithful system of parameters. The big theorem here is due to Yoshino [46] (cf. [33, Theorem 15.8]):

Theorem 4. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a complete CM local ring containing a field. Assume k is perfect and that R has an isolated singularity. Then R has a faithful system of parameters.*

Putting all of this stuff together, we obtain the following theorem, proved independently by Dieterich [15] and Yoshino [46]:

Theorem 5. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a complete, equicharacteristic local ring with perfect residue field k. Then R has finite CM type if and only if:*

(i) R has bounded CM type.
(ii) R has an isolated singularity.

The main thrust is the "if" direction, the converse being a consequence of Auslander's famous theorem [2] that complete CM rings with finite CM type must be isolated singularities.

In 2005, Leuschke and Wiegand used ascent and descent techniques to prove the following generalization [32, Theorem 3.4]:

Theorem 6. *Let* $(R, \mathfrak{m}, \mathsf{k})$ *be an excellent, equicharacteristic local ring with perfect residue field* k*. Then* R *has finite CM type if and only if:*

(i) R *has bounded CM type.*
(ii) R *has an isolated singularity.*

This time, for the "only if" direction, one needs the Huneke–Leuschke version [23] of Auslander's theorem, stating that *every* CM ring of finite CM type has an isolated singularity.

Without the word "excellent," Theorem 6 would be false. For example, the ring $\mathbb{C}[\![x, y]\!]/(y^2)$ is the completion of an integral domain $(R, \mathfrak{m})$, by Lech's Theorem [30]. Theorem 4 implies that R has bounded but infinite CM type, and of course R has an isolated singularity.

5 Brauer–Thrall II

In Sect. 2 we proved a strong form of BTM2 for one-dimensional CM local rings, assuming only that the ring is either analytically unramified or equicharacteristic. In higher dimensions, no such general results are known. One problem, already mentioned, is that there is no general result showing descent of bounded CM type along flat local homomorphisms. Typically, one restricts to complete (or at least excellent Henselian) isolated singularities with algebraically closed residue field, in order to make use of the Auslander–Reiten quiver.

The following result was proved by Dieterich [15, Theorem 20] in 1987, for characteristics different from two. The case $\mathrm{char}(\mathsf{k}) = 2$ was proved by Popescu and Roczen [37] in 1991.

Theorem 1. *Let* $R = \mathsf{k}[\![x_0, \ldots, x_d]\!]/(f)$ *be a hypersurface isolated singularity, with* k *algebraically closed. If* R *has infinite CM type, then* R *has strictly unbounded CM type.*

Using Elkik's theorem [18] on modules extended from the Henselization, one can generalize this result to excellent Henselian rings (cf. [36]):

Corollary 2. *Let* $(R, \mathfrak{m}, \mathsf{k})$ *be an excellent, equicharacteristic, Henselian local ring whose completion is a hypersurface. Assume that* R *has an isolated singularity and that* k *is algebraically closed. If* R *has infinite CM type, then* R *has strictly unbounded CM type. (In particular, both BTM1 and BTM2 hold for these rings.)*

Excellence guarantees that the completion $\widehat{R}$ is an isolated singularity too. (In fact, all one needs is that the inclusion $R \longrightarrow \widehat{R}$ be a regular homomorphism (see [33, Proposition 10.9]).) If N is an MCM $\widehat{R}$-module, then N is free on the punctured spectrum of $\widehat{R}$ and hence, by [18], is extended from an R-module.

This means that the map $M \mapsto \widehat{M}$, from MCM R-modules to MCM $\widehat{R}$-modules, is bijective on isomorphism classes. Since $\mathrm{e}_R(M) = \mathrm{e}_{\widehat{R}}(\widehat{M})$, the corollary follows from the theorem.

The main thing we want to talk about in this section is Smalø's remarkable result [40] that produces, from an infinite family of indecomposable MCM modules of *one fixed* multiplicity n, an integer $n' > n$ and an infinite family of indecomposable MCM modules of multiplicity n'. In principle, this ought to make proofs of BTM2 lots easier. We will give two such applications and also point out some limitations to this approach. Here is Smalø's theorem, proved in 1980 for Artin algebras:

Theorem 3. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a complete CM-isolated singularity with k algebraically closed. Suppose $\{M_i\}_{i\in I}$ is an infinite family of pairwise non-isomorphic indecomposable MCM R-modules, all of the same multiplicity n. There exist an integer $n' > n$, a subset J of I with $|J| = |I|$, and a family $\{N_j\}_{j\in J}$ of pairwise non-isomorphic indecomposable MCM R-modules, each of multiplicity n'.*

The basic ideas of Smalø's proof survive transplantation to the MCM context remarkably well. The proof uses the Harada–Sai Lemma 3 as well as a couple of lemmas that control multiplicity as one wanders around the AR quiver. One of these [4, Lemma 4.2.7] bounds the growth of the Betti numbers $\beta_i(M)$ of a MCM module M over a CM local ring of multiplicity e: $\beta_{i+1} \le (e-1)\beta_i$ for all i. Another gives a linear bound between the multiplicities of the source and target of an irreducible homomorphism: With R as in the theorem, there is a positive constant c such that $\mathrm{e}_R(M) \le c\,\mathrm{e}_R(N) \le c^2\mathrm{e}_R(M)$ whenever $M \longrightarrow N$ is an irreducible homomorphism of indecomposable MCM R-modules. We refer the reader to [33, Sect. 15.4] for the details.

Here is an obvious corollary of Smalø's theorem:

Corollary 4. *Let $(R, \mathfrak{m}, \mathsf{k})$ be a complete CM-isolated singularity, with k algebraically closed. If R has uncountable CM type, then there is an sequence $n_1 < n_2 < n_3 < \dots$ of positive integers such that R has, for each i, uncountably many non-isomorphic indecomposable MCM modules of multiplicity n_i.*

As another application, one can give a proof of BTM2 in dimension one that is much less computational than the one given in Sect. 2, at least in an important special case. Suppose that $(R, \mathfrak{m}, \mathsf{k})$ is a complete, reduced local ring of dimension one, and assume R has infinite CM type. Then the Drozd–Roĭter conditions ((ii) and (iii) of Theorem 1) fail. It is now a comparatively simple matter (see [42, §4]) to show that R has an infinite family of pairwise non-isomorphic ideals. We decompose each of these ideals into indecomposable summands, noting that $\mathrm{e}(R)$ bounds the number of summands of each ideal. This yields infinitely many pairwise non-isomorphic indecomposable MCM modules, each with multiplicity bounded by $\mathrm{e}(R)$, and hence an infinite subfamily consisting of modules of fixed multiplicity. Now Smalø's theorem shows that BTM2 holds for these rings.

Do not be misled by this example. In higher dimensions there is no hope of starting the inductive hypothesis with modules of rank one, in view of the following theorem due to Bruns [7, Corollary 2]:

Theorem 5. *Let A be any commutative Noetherian ring and M a finitely generated R-module of constant rank r. Let N be a second syzygy of M, and let s be the (constant) rank of N. If M is not free, then the codimension of its non-free locus is at most $r + s + 1$.*

Corollary 6. *Let $(R, \mathfrak{m})$ be a d-dimensional isolated singularity whose completion is a hypersurface. Let M be a non-free MCM R-module of constant rank r. Then $r \geq \frac{1}{2}(d-1)$.*

This bound is probably much too low. In fact, Buchweitz, Greuel, and Schreyer [8] conjecture that $r \geq 2^{d-1}$. Nonetheless, the bound given in the corollary rules out MCM ideals once the dimension exceeds three.

6 Open Questions

Here we list a few open questions, some of which have already been mentioned at least implicitly.

Question 6.1. Are there *any* counterexamples to BTM2? Of course this is the same as asking whether BTM2 is true, but let us not even assume that $(R, \mathfrak{m}, \mathsf{k})$ is CM. What if $\dim(R) = 1$? What if $\dim(R) = 1$ and R is not CM? The list goes on....

Question 6.2. Can one delete the assumption, in Theorem 4, that k be infinite?

Question 6.3. If $(R, \mathfrak{m})$ is a local CM ring whose completion $\widehat{R}$ has bounded CM type, must R have bounded CM type? More generally, let $R \longrightarrow S$ be a flat local homomorphism with CM closed fiber. If S has bounded CM type, must R have bounded CM type?

Question 6.4. Can one delete the assumption, in Theorem 6, that k be perfect?

Question 6.5. Can we improve Corollary 6, getting better lower bounds for the rank, or multiplicity, of a non-free MCM module?

Acknowledgments Research for this work was partially supported by NSF grant DMS-0902119 (GJL) and by a Simons Foundation Collaboration Grant (RW).

References

1. Artin, M., Verdier, J.-L. Reflexive modules over rational double points. Math. Ann. **270**(1), 79–82 (1985). MR 769609
2. Auslander, M.: Isolated singularities and existence of almost split sequences. Representation Theory, II (Ottawa, Ont., 1984). Lecture Notes in Mathematics, vol. 1178, pp. 194–242. Springer, Berlin (1986). MR 842486

3. Auslander, M.: Rational singularities and almost split sequences. Trans. Amer. Math. Soc. **293**(2), 511–531 (1986). MR 816307
4. Avramov, L.L.: Infinite free resolutions. Six Lectures on Commutative Algebra (Bellaterra, 1996). Progress in Mathematics, vol. 166, pp. 1–118. Birkhäuser, Basel (1998). MR 1648664
5. Bass, H.: Torsion free and projective modules. Trans. Amer. Math. Soc. **102**, 319–327 (1962). MR 0140542
6. Bass, H.: On the ubiquity of Gorenstein rings. Math. Z. **82**, 8–28 (1963). MR 0153708
7. Bruns, W.: The Eisenbud-Evans generalized principal ideal theorem and determinantal ideals. Proc. Amer. Math. Soc. **83**(1), 19–24 (1981). MR 0619972
8. Buchweitz, R.-O., Greuel, G.-M., Schreyer, F.-O.: Cohen–Macaulay modules on hypersurface singularities. II. Invent. Math. **88**(1), 165–182 (1987). MR 877011
9. Çimen, N.: One-dimensional rings of finite Cohen–Macaulay type. Ph.D. thesis, University of Nebraska–Lincoln, Lincoln, NE (1994). Ph.D. Thesis, University of Nebraska, Lincoln, p. 105. MR 2691778
10. Çimen, N.: One-dimensional rings of finite Cohen-Macaulay type. J. Pure Appl. Algebra **132**(3), 275–308 (1998). MR 1642094
11. Çimen, N., Wiegand, R., Wiegand, S.: One-dimensional rings of finite representation type. Abelian groups and modules (Padova, 1994). Mathematical Applications, vol. 343, pp. 95–121. Kluwer Academic Publishers, Dordrecht (1995) MR 1378192
12. Crabbe, A., Saccon, S.: Ranks of indecomposable modules over rings of infinite Cohen–Macaulay type. Comm. Algebra, to appear, Available at arXiv:1201.314v1.
13. Dade, E.C.: Some indecomposable group representations. Ann. Math. (2) **77**, 406–412 (1963). MR 0144981
14. Dieterich, E.: Representation types of group rings over complete discrete valuation rings. Integral representations and applications (Oberwolfach, 1980). Lecture Notes in Mathematics, vol. 882, pp. 369–389. Springer, Berlin (1981). MR 646112
15. Dieterich, E.: Reduction of isolated singularities. Comment. Math. Helv. **62**(4), 654–676 (1987). MR 920064
16. Drozd, Y.A., Roĭter, A.V.: Commutative rings with a finite number of indecomposable integral representations. Izv. Akad. Nauk SSSR Ser. Mat. **31**, 783–798 (1967). MR 0220716
17. Eisenbud, D., Peña, J.A.: Chains of maps between indecomposable modules. J. Reine Angew. Math. **504**, 29–35 (1998). MR 1656826
18. Elkik, R.: Solutions d'équations à coefficients dans un anneau hensélien. Ann. Sci. École Norm. Sup. (4) **6**, 553–603 (1974). MR 0345966
19. Green, E.L., Reiner, I.: Integral representations and diagrams. Michigan Math. J. **25**(1), 53–84 (1978). MR 497882
20. Gustafson, W.H.: The history of algebras and their representations. Representations of algebras (Puebla, 1980). Lecture Notes in Mathematics, vol. 944, pp. 1–28. Springer, Berlin (1982). MR 672114
21. Harada, M., Sai, Y.: On categories of indecomposable modules I. Osaka J. Math. **8**, 309–321 (1971)
22. Herzog, J.: Ringe mit nur endlich vielen Isomorphieklassen von maximalen, unzerlegbaren Cohen–Macaulay–Moduln. Math. Ann. **233**(1), 21–34 (1978). MR 0463155
23. Huneke, C., Leuschke, G.J.: Two theorems about maximal Cohen–Macaulay modules. Math. Ann. **324**(2), 391–404 (2002). MR 1933863
24. Jacobinski, H.: Sur les ordres commutatifs avec un nombre fini de réseaux indécomposables. Acta Math. **118**, 1–31 (1967). MR 0212001
25. Jans, J.P.: On the indecomposable representations of algebras. Ann. Math. (2) **66**, 418–429 (1957). MR 0088485
26. Karr, R., Wiegand, R.: Direct-sum behavior of modules over one-dimensional rings. Commutative algebra—Noetherian and non-Noetherian perspectives, 251–275, Springer, New York, 2011. MR 2762514
27. Kawasaki, T.: Local cohomology modules of indecomposable surjective-Buchsbaum modules over Gorenstein local rings. J. Math. Soc. Japan **48**(3), 551–566 (1996). MR 1389995

28. Knörrer, H.: Cohen–Macaulay modules on hypersurface singularities. I. Invent. Math. **88**(1), 153–164 (1987). MR 877010
29. Kronecker, L.: Über die congruenten Transformationen der bilinearen Formen. In: Hensel, K. (ed.) Leopold Kroneckers Werke, vol. I. Monatsberichte Königl. Preuß. Akad. Wiss. Berlin, pp. 423–483 (1874). (German)
30. Lech, C.: A method for constructing bad Noetherian local rings. Algebra, algebraic topology and their interactions (Stockholm, 1983). Lecture Notes in Mathematics, vol. 1183, pp. 241–247. Springer, Berlin (1986). MR 846452
31. Leuschke, G.J., Wiegand, R.: Hypersurfaces of bounded Cohen-Macaulay type. J. Pure Appl. Algebra **201**(1–3), 204–217 (2005). MR 2158755
32. Leuschke, G.J., Wiegand, R.: Local rings of bounded Cohen-Macaulay type. Algebr. Represent. Theory **8**(2), 225–238 (2005). MR 2162283
33. Leuschke, G.J., Wiegand, R.: Cohen–Macaulay representations. Mathematical Surveys and Monographs, American Mathematical Society, Providence (2012). MR 2292367
34. Matsumura, H.: Commutative ring theory. In: Cambridge Studies in Advanced Mathematics, vol. 8, 2nd edn. Cambridge University Press, Cambridge (1989). Translated from the Japanese by M. Reid. MR 1011461
35. Nazarova, L.A., Roĭter, A.V.: Kategornye matrichnye zadachi i problema Brauera-Trella. Izdat. “Naukova Dumka”, Kiev (1973). MR 0412233
36. Popescu, D., Roczen, M.: Indecomposable Cohen-Macaulay modules and irreducible maps. Compositio Math. **76**(1–2), 277–294 (1990). Algebraic geometry (Berlin, 1988). MR 1078867
37. Popescu, D., Roczen, M.: The second Brauer-Thrall conjecture for isolated singularities of excellent hypersurfaces. Manuscripta Math. **71**(4), 375–383 (1991). MR 1104991
38. Ringel, C.M.: On algorithms for solving vector space problems. I. Report on the Brauer-Thrall conjectures: Rojter’s theorem and the theorem of Nazarova and Rojter. Representation theory, I (Proceedings of the Workshop, Carleton University, Ottawa, Ontoria (1979). Lecture Notes in Mathematics, vol. 831, pp. 104–136. Springer, Berlin (1980). MR 607142
39. Roĭter, A.V.: Unboundedness of the dimensions of the indecomposable representations of an algebra which has infinitely many indecomposable representations. Izv. Akad. Nauk SSSR Ser. Mat. **32**, 1275–1282 (1968). MR 0238893
40. Smalø, S.O.: The inductive step of the second Brauer–Thrall conjecture. Canad. J. Math. **32**(2), 342–349 (1980). MR 571928
41. Weierstrass, K.: On the theory of bilinear and quadratic forms. (Zur Theorie der bilinearen und quadratischen Formen.) Monatsberichte Königl. Preuß. Akad. Wiss. Berlin, 1868. (German)
42. Wiegand, R.: Noetherian rings of bounded representation type. Commutative algebra (Berkeley, CA, 1987). Mathematical Science Research Institute Publications, vol. 15, pp. 497–516. Springer, New York (1989). MR 1015536
43. Wiegand, R.: Curve singularities of finite Cohen–Macaulay type. Ark. Mat. **29**(2), 339–357 (1991). MR 1150382
44. Wiegand, R.: One-dimensional local rings with finite Cohen–Macaulay type. Algebraic geometry and its applications (West Lafayette, IN, 1990), pp. 381–389. Springer, New York (1994). MR 1272043
45. Wiegand, R.: Local rings of finite Cohen-Macaulay type. J. Algebra **203**(1), 156–168 (1998). MR 1620725
46. Yoshino, Y.: Brauer–Thrall type theorem for maximal Cohen–Macaulay modules. J. Math. Soc. Japan **39**(4), 719–739 (1987). MR 905636
47. Yoshino, Y.: Cohen–Macaulay modules over Cohen–Macaulay rings. In: London Mathematical Society Lecture Note Series, vol. 146. Cambridge University Press, Cambridge (1990). MR 1079937

Tight Closure's Failure to Localize - a Self-Contained Exposition

Paul Monsky

1 Introduction: Brenner's Insight

Holger Brenner and I have given a negative solution to the localization problem for tight closure [1]. The argument involves the Hilbert-Kunz theory of plane curves (and in particular [2]) together with results of Brenner, Hochster, and Huneke on test elements and local cohomology.

But most of this machinery, useful as it is for understanding our counterexample, may be dispensed with; in this chapter I give a treatment of the example, using only linear algebra, material from an introductory abstract algebra course, and a little local cohomology developed ab initio. The reader doesn't need to know anything about Hilbert-Kunz theory, homological algebra, vector bundles, or tight closure. Though the arguments are largely drawn from [1,2], everything is proved here from scratch.

Definition 1. If A is a Noetherian domain of characteristic $p > 0$, q is a power of p, and I is an ideal of A, $I^{[q]}$ is the ideal generated by all v^q, v in I.

Definition 2. u is in the tight closure, I^*, of I if for some $d \neq 0$, $du^q \in I^{[q]}$ for all q.

Suppose now that $S \subset A$ is multiplicatively closed, $0 \notin S$. Then $S^{-1}I$ is an ideal of $S^{-1}A$, and we can form the ideal $(S^{-1}I)^*$. The localization problem asks whether $(S^{-1}I)^*$ is always equal to $S^{-1}(I)^*$. In other words, suppose that $f \in (S^{-1}I)^*$. Must there exist an s in S such that $sf \in I^*$? After giving positive solutions to the localization problem in some special cases, Brenner realized that the study of a 1-parameter family would give a negative answer provided the family satisfied a certain counter-intuitive condition. I'll explain this insight of Brenner's in the

P. Monsky (✉)
Brandeis University, Waltham, MA 02454-9110, USA
e-mail: monsky@brandeis.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_19,

context of a 1-parameter family of projective plane curves. Let L be algebraically closed of characteristic p, and let P and P_1 in $L[x, y, z]$ be homogeneous of the same degree. For α in L set $g_\alpha = P + \alpha P_1$ and $R_\alpha = L[x, y, z]/g_\alpha$. Let R_{gen} be the ring $L(t)[x, y, z]/(P + tP_1)$. Fix f in $L[x, y, z]$ and an ideal I in $L[x, y, z]$. Now I generates ideals in each R_α and in R_{gen}. Abusing language we call all these ideals I.

Theorem 3 (Brenner). *Suppose that:*

(a) $f \in I^*$ *in* R_{gen}.
(b) There exist infinitely many α *in* L *for which* R_α *is a domain and* $f \notin I^*$ *in* R_α.

Then the localization problem has a negative answer.

Proof. Take $A = L[x, y, z, t]/(P + tP_1)$, and let $S \subset A$ be $L[t] - \{0\}$. Note that $S^{-1}A$ and $A/(t - \alpha)$ identify with R_{gen} and R_α, respectively. $I \subset L[x, y, z]$ generates an ideal in A that we again call I. Since R_{gen} identifies with $S^{-1}A$, we see from (a) that $f \in (S^{-1}I)^*$ in $S^{-1}A$. Suppose however that $sf \in I^*$ for some $s = s(t)$. Then for some $d \neq 0$ in A, we have $ds^q f^q \in I^{[q]}$ for all q. Now by (b) there are infinitely many α in L with $A/(t-\alpha)$ a domain and $f \notin I^*$ in $A/(t-\alpha)$. The corresponding ideals, $(t - \alpha)$, are distinct height 1 primes in A, and so cannot all contain ds. Fix one such $t - \alpha$ with $ds \notin (t - \alpha)$. If $\bar{d}$ is the image of d in $A/(t - \alpha) = R_\alpha$, then $\bar{d}s(\alpha)^q f^q \in I^{[q]}$ in R_α for all q. But $\bar{d} \neq 0$ and $s(\alpha)$ is a nonzero element of L. We conclude that $\bar{d} f^q \in I^{[q]}$ in R_α for all q, contradicting the choice of α. □

How is a 1-parameter family satisfying (a) and (b) to be found? In [2], I had studied a 1-parameter linear family of plane quartics in characteristic 2, obtaining counter-intuitive results suggestive of (a) and (b). (This was done in ignorance of tight closure; my goal was to calculate the "Hilbert-Kunz multiplicities" of the curves in this family.) It turned out that the matrix calculations in [2], slightly extended and combined with a suitable "test element theorem," were exactly what was needed to produce the example. In the following three sections I describe these calculations. The final two sections use some simple algebra to complete the proof.

Throughout, L will be a field of characteristic 2 and P the element $z^4 + xyz^2 + (x^3 + y^3)z$ of $L[x, y, z]$. If $\alpha \in L, \alpha \neq 0, g_\alpha = P + \alpha x^2 y^2$. It's easy to see that g_α is irreducible, so that $R_\alpha = L[x, y, z]/g_\alpha$ is a domain. Fix a power, Q, of 2. Let O be the graded L-algebra $L[x, y, z]/(x^{4Q}, y^{4Q}, z^{4Q})$. Multiplication by g_α gives a map $O_j \to O_{j+4}$ for each j. The key to establishing (a) and (b) of Theorem 3 is the close study of the kernel N_{6Q-5} of $g_\alpha : O_{6Q-5} \to O_{6Q-1}$, both when α is transcendental over $Z/2$ and $Q \geq 2$ and when α is algebraic over $Z/2$ and Q is a certain power of 2 attached to α. When $Q \geq 2$, O_{6Q-5} and O_{6Q-1} have dimensions $12Q^2 - 12$ and $12Q^2$, and one might expect $N_{6Q-5} = (0)$ for every choice of Q. This is true for transcendental α (see Theorem 13) but false for algebraic α (Corollary 6, with Q as in Definition 4).

2 Some Identities Involving P

We begin by defining some elements of $Z/2[x, y]$.

Definition 1. If r is a power of 2, then:

(1) A_r (resp. B_r) is $\sum x^i y^j$, the sum extending over all pairs (i, j) with $i \equiv j \quad (3)$ and $i + j = 4r - 2$ (resp. $4r - 1$).
(2) $C_1 = 1$ and $C_{2r} = A_r^2$.

Each monomial $x^i y^j$ appearing in A_{2r} has $i \equiv j \quad (2)$. Those monomials with i (and j) even sum to B_r^2, while those with i (and j) odd sum to xyA_r^2. So $A_{2r} = B_r^2 + xyA_r^2$. A similar argument shows that $B_{2r} = (x^3 + y^3)A_r^2$.

Lemma 2. *The following identities hold in $Z/2[x, y, z]$: When Q is a power of 2, then $z^{4Q} = A_Q z^2 + B_Q z + \sum_{rs=Q}(C_r P)^s$.*

Proof. Since $A_1 = xy$, $B_1 = x^3 + y^3$, and $C_1 = 1$, the case $Q = 1$ follows from the definition of P. In general we argue by induction, squaring the identity for Q, replacing z^4 by $xyz^2 + (x^3 + y^3)z + P$, and using the identities following Definition 1. □

Now let L, Q, and $O = L[x, y, z]/(x^{4Q}, y^{4Q}, z^{4Q})$ be as in the last section.

Definition 3. R_Q is the element $A_Q z^2 + B_Q z$ of O_{4Q}, while Δ in O_{6Q-1} is $\sum x^i y^j z^k$, the sum extending over all triples (i, j, k) with $i + j + k = 6Q - 1$, $i \not\equiv j \quad (3)$, and $k = 1$ or 2.

Lemma 4. *Suppose $i + j = 2Q - 1$. Then, in O, $(x^i y^j + x^j y^i)R_Q$ is 0 if $i \equiv j \quad (3)$ and is Δ otherwise.*

Proof. Definition 1 shows that $(x^3+y^3)A_Q$ and $(x^3+y^3)B_Q$ both lie in (x^{4Q}, y^{4Q}). So $x^3 R_Q = y^3 R_Q$ in O. It follows immediately that when $i + j = 2Q - 1$ then $x^i y^j R_Q$ only depends on i mod 3. This gives the first part of Lemma 4 and shows that when $i \not\equiv j \quad (3)$, $(x^i y^j + x^j y^i)R_Q = (x^{Q-1}y^Q + x^Q y^{Q-1})R_Q$. But this last element is easily seen to be Δ. □

Theorem 5.

(1) In O, $R_Q = \sum_{rs=Q}(C_r P)^s$.
(2) Suppose that $i + j = 2Q - 1$. Then in O, $(x^i y^j + x^j y^i)P^Q = \varepsilon\Delta + (x^i y^j + x^j y^i)\sum_{\substack{rs=Q\\ s\neq Q}}(C_r P)^s$, where ε is 0 if $i \equiv j \quad (3)$ and is 1 otherwise.

Proof. Combining Lemma 2 with the definition of R_Q, noting that $z^{4Q} = 0$ in O, we get (1). Since $C_1 = 1$, $P^Q = R_Q + \sum_{\substack{rs=Q\\ s\neq Q}}(C_r P)^s$. Multiplying by $x^i y^j + x^j y^i$ and applying Lemma 4 gives (2). □

Lemma 6. *Suppose* $i + j = 2Q - 1$. *The coefficient of* $x^{4Q-2}y^{Q-2}$ *in* $(x^i y^j + x^j y^i)(x^3 + y^3)^{Q-1}$ *is 0 if* $i \not\equiv j \quad (3)$ *and is 1 otherwise.*

Proof. The first assertion is clear. For the second note that the coefficient in question is the sum of the coefficients of $x^{4Q-2-i}y^{Q-2-j}$ and $x^{4Q-2-j}y^{Q-2-i}$ in $(x^3 + y^3)^{Q-1} = x^{3Q-3} + x^{3Q-6}y^3 + \cdots + y^{3Q-3}$. The first of these coefficients is 1 when i is both $\geq Q + 1$ and $\equiv j \quad (3)$, while the second is 1 when j is both $\geq Q + 1$ and $\equiv i \quad (3)$. Since precisely one of i and j is $\geq Q + 1$ (they cannot be Q and $Q - 1$) we get the lemma. □

3 The Spaces X and Y: The Case of Transcendental α

L, P, and g_α are as in the final paragraph of the introduction. $Q \geq 2$ is a power of 2, while O is the graded L-algebra $L[x, y, z]/(x^{4Q}, y^{4Q}, z^{4Q})$, and N_{6Q-5} is the kernel of $g_\alpha : O_{6Q-5} \to O_{6Q-1}$.

Definition 1.

(1) $[i, j] = x^i y^j + x^j y^i$.
(2) $X \subset O_{6Q-5}$ is spanned by the $[i, j]P^k$ with $i + j + 4k = 6Q - 5$ and $k = 0, 1, \ldots, Q - 1$.
(3) $Y \subset O_{6Q-1}$ is spanned by the $[i, j]P^k$ with $i + j + 4k = 6Q - 1$ and $k = 0, 1, \ldots, Q - 1$.

Theorem 2. *Let* (Y, Δ) *be the subspace of* O_{6Q-1} *spanned by* Y *and the element* Δ *of Definition 3. Then* $g_\alpha \cdot X \subset (Y, \Delta)$.

Proof. Evidently $(x^2 y^2) \cdot X \subset Y$. It remains to show that $P \cdot X \subset (Y, \Delta)$. This will follow if we can prove that $P \cdot [i, j] \cdot P^{Q-1} \in (Y, \Delta)$ whenever $i + j = 2Q - 1$. By Theorem 5 it suffices to show that each $[i, j]C_r^s P^s$ is in Y when $rs = Q$ and $s < Q$. This is easy: $[i, j] \cdot C_r^s$ is a symmetric form in x and y of (odd) degree $(2Q - 1) + s(4r - 4) = 6Q - 1 - 4s$. □

The $[i, j]P^k$ with $i + j + 4k = 6Q - 5$, $i, j < 4Q$, $k < Q$, and i odd evidently span X. Noting that each such element has the form $(x^i y^j + x^j y^i)z^{4k} +$ terms of lower degree in z, with i odd and j even, we see that these elements are a basis of X. One constructs a basis of Y similarly and finds that $\dim X = \dim Y$; both dimensions are in fact $\frac{3Q^2}{2}$. A basis of (Y, Δ) is given by the $[i, j]P^k$ with $i + j + 4k = 6Q - 1$, $i, j < 4Q$, $k < Q$, and i odd, together with Δ.

Note that the kernel of the map $g_\alpha : X \to (Y, \Delta)$ of Theorem 2 is just $N_{6Q-5} \cap X$. We'll get a better understanding of this space by replacing X and (Y, Δ) by certain quotients.

Definition 3. D is the graded L-algebra $L[x, y]/(x^{4Q}, y^{4Q})$. For $1 \leq i \leq Q$, let E_i be the element $[2i - 1, 2Q - 2i]$ of D_{2Q-1}, and F_i be the element $x^{2Q}y^{2Q}E_i$ of D_{6Q-1}. Let D_{2Q-1}^{sym} and D_{6Q-1}^{sym} be the Q-dimensional subspaces of D_{2Q-1} and D_{6Q-1} spanned by the E_i and F_i, respectively.

Definition 4. $X \to D^{sym}_{2Q-1}$ is the map taking $[i, j]P^k$ to 0 when $k < Q-1$ and to $[i, j]$ when $k = Q-1$ (in which case $i + j = 2Q-1$).

Definition 5. $Y \to D^{sym}_{6Q-1} \oplus L$ takes

$$[i, j]P^k \ \text{ to } \ \big([i, j](\alpha x^2 y^2)^k, 0\big)$$
$$\Delta \ \text{ to } \ (0, 1).$$

Using the bases of X and (Y, Δ) we've constructed, we see that these L-linear maps are well defined. They are evidently onto.

Lemma 6. *Let X_0 and Y_0 be the kernels of the maps of Definitions 4 and 5. Then g_α maps X_0 bijectively to Y_0.*

Proof. Our description of a basis of X shows that X_0 is spanned by the $[i, j]P^k$ with $i + j + 4k = 6Q-5$ and $k = 0, 1, \ldots, Q-2$. So a nonzero element, u, of X_0 has the form $A(x, y)z^k+$ terms of lower degree in z, where $A(x, y) \neq 0$ in D and $k < 4Q-4$. Then $g_\alpha u = A(x, y)z^{k+4} + \cdots \neq 0$; we conclude that g_α maps X_0 injectively. If $k \le Q-2$, then $g_\alpha[i, j] \cdot P^k = [i, j]P^{k+1} + \alpha[i+2, j+2]P^k$. Both terms on the right map to $\big([i, j](\alpha x^2 y^2)^{k+1}, 0\big)$ under the map of Definition 5, and we conclude that $g_\alpha(X_0) \subset Y_0$. Note also that the maps of Definitions 4 and 5 are onto, that $\dim X = \dim Y$, and that $\dim D^{sym}_{2Q-1} = \dim D^{sym}_{6Q-1} = Q$. This tells us that $\dim X_0 = \dim Y_0$, so that $g_\alpha \cdot X_0 = Y_0$. □

In view of Lemma 6, $(N_{6Q-5}) \cap X$ identifies with the kernel of the map $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ induced by $g_\alpha : X \to (Y, \Delta)$. With respect to the bases $E_1, \ldots, E_Q$ of D^{sym}_{2Q-1} and $F_1, \ldots, F_Q, 1$ of $D^{sym}_{6Q-1} \oplus L$, the matrix of this induced map has the form $\left(\frac{M}{b}\right)$ where M is a Q by Q matrix and $b = (b_1, \ldots, b_Q)$ is a row vector. We shall use Theorem 5 to write down M and b.

Lemma 7. *The map $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ induced by $g_\alpha : X \to (Y, \Delta)$ takes E_j to $\left(E_j \cdot \left(\sum_{rs=Q} \alpha^s C_r^s x^{2s} y^{2s}\right), b_j\right)$, where $b_j = 0$ if $2j-1 \equiv 2Q-2j \quad (3)$, and is 1 otherwise.*

Proof. E_j pulls back to $E_j \cdot P^{Q-1}$ in X. Multiplication by g_α takes this to $E_j \cdot (\alpha x^2 y^2 P^{Q-1} + P^Q)$. By Theorem 5 this is

$$b_j \Delta + E_j \left(\alpha x^2 y^2 P^{Q-1} + \sum_{\substack{rs=Q \\ s \neq Q}} (C_r P)^s\right).$$

Under the map of Definition 5, the first term above goes to $(0, b_j)$, while the second goes to $E_j \cdot \sum_{rs=Q} (C_r \alpha x^2 y^2)^s$, giving the lemma. □

Theorem 8. *Situation as in Lemma 7. The image of E_j is $\left(\sum \alpha^s F_i, b_j\right)$ where the sum extends over all pairs (s, i) with s/Q and $i \equiv j \quad (3s)$.*

Proof. Using the definitions of A_r and C_r we find that $C_r x^2 y^2 = \sum x^{2r+2k} y^{2r-2k}$, the sum extending over all k in $(-r, r)$ with $k \equiv 0 \quad (3)$. So $C_r^s x^{2s} y^{2s} = \sum x^{2Q+2l} y^{2Q-2l}$, the sum extending over all l in $(-Q, Q)$ with $l \equiv 0 \quad (3s)$. Then $E_j(C_r^s x^{2s} y^{2s})$ is $\sum F_i$, the sum extending over all $i \equiv j \quad (3s)$, and Lemma 7 gives the result. □

Corollary 9. *Let $b_i = 0$ if $2i - 1 \equiv 2Q - 2i \quad (3)$ and $b_i = 1$ otherwise. Then the matrix of the induced map $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ with respect to the bases introduced earlier is $\left(\frac{M}{b(Q)}\right)$ where $m_{i,j} = \sum \alpha^s$, the sum extending over all s/Q with $i \equiv j \quad (3s)$, and $b(Q) = (b_1, \ldots, b_Q)$.*

Corollary 10. *Suppose that $\alpha \in L$ is transcendental over $Z/2$. Then $(N_{6Q-5}) \cap X = (0)$.*

Proof. The matrix M of Corollary 9 has entries in $Z/2[\alpha]$. Each $m_{i,i}$ is a degree Q polynomial in α while the other entries have degree $< Q$. Since α is transcendental over $Z/2$, $\det M \neq 0$, $\left(\frac{M}{b(Q)}\right)$ has rank Q and $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ is 1–1. But the kernel of this map identifies with $(N_{6Q-5}) \cap X$. □

For the rest of this section we assume α transcendental over $Z/2$. We'll use Corollary 10 to show that N_{6Q-5} is (0). Any $u \neq 0$ in O may be written as $A(x, y) \cdot z^r +$ lower degree terms in z, where $A(x, y) \neq 0$ in D, and $r < 4Q$. We say that u has z-degree r.

Lemma 11. *Suppose that $u \in O_{6Q-5}$ is fixed by $(x, y) \to (y, x)$ and has z-degree $\leq 4Q - 4$. Then if $g_\alpha u \in g_\alpha X$, $u \in X$.*

Proof. We argue by induction on the z-degree of u. If the z-degree is 0, then u, being fixed by $(x, y) \to (y, x)$, is a linear combination of $x^i y^j + x^j y^i$ with $i + j = 6Q - 5$ and so is in X. If $u = A \cdot z^{4k} + \cdots$, with $k > 0$, let $v = u + AP^k$. Then the z-degree of v is $< 4k$, and $g_\alpha v \in g_\alpha X$. By induction, $v \in X$, and so $u \in X$. Suppose finally that $u = A \cdot z^r + \cdots$ with $r \not\equiv 0 \quad (4)$ and $r < 4Q - 4$. Then $g_\alpha u = A \cdot z^{r+4} + \cdots$ has z-degree that is neither divisible by 4 nor equal to 2. As $g_\alpha u \in g_\alpha X \subset (Y, \Delta)$, our description given earlier of a basis of (Y, Δ) shows this to be impossible. □

Lemma 12. *If $u \in N_{6Q-5}$ has z-degree $\leq 4Q - 4$ then $u = 0$.*

Proof. The linear automorphism $(x, y, z) \to (y, x, z)$ of $L[x, y, z]$ fixes g_α. So the automorphism of O that it induces stabilizes N_{6Q-5}. Let $\bar{u}$ be the image of u under this automorphism. Lemma 11 applied to $u + \bar{u}$ shows that $u + \bar{u} \in X$. Since $u + \bar{u} \in N_{6Q-5}$, Corollary 10 shows that $u = \bar{u}$. Applying Lemma 11 to u we find that $u \in X$. Another application of Corollary 10 completes the proof. □

Theorem 13. $N_{6Q-5} = (0)$.

Proof. Replacing L by a larger field, if necessary, we may assume that L contains some ω with $\omega^2 + \omega + 1 = 0$. We make use of 3 linear automorphisms of $L[x, y, z]$:

$$\sigma : (x, y, z) \to (x, y, z + x + y).$$
$$\tau : (x, y, z) \to (x, y, z + \omega x + \omega^2 y).$$
$$\rho : (x, y, z) \to (x, y, z + \omega^2 x + \omega y).$$

Since $P = z(z + x + y)(z + \omega x + \omega^2 y)(z + \omega^2 x + \omega y)$, these automorphisms fix P as well as x and y. So they fix g_α, and the automorphisms of O that they induce stabilize N_{6Q-5}.

Suppose now that $u = Az^r + \cdots$ is an element of N_{6Q-5} of z-degree r. By Lemma 12, $r = 4Q - 3$, $4Q - 2$, or $4Q - 1$. Suppose first that $r = 4Q - 3$. Then $u^\sigma + u = A(x + y) \cdot z^{4Q-4} + \cdots$. Since A is a nonzero element of D_{2Q-2}, $A \cdot (x + y) \neq 0$ in D. This contradicts Lemma 12 applied to the element $u^\sigma + u$ of N_{6Q-5}. Suppose next that $u = Az^{4Q-2} + Bz^{4Q-3} + \cdots$ has z-degree $4Q - 2$. Then,

$$u^\tau + u = \big(A(\omega x + \omega^2 y)^2 + B(\omega x + \omega^2 y)\big) \cdot z^{4Q-4} + \cdots$$
$$u^\rho + u = \big(A(\omega^2 x + \omega y)^2 + B(\omega^2 x + \omega y)\big) \cdot z^{4Q-4} + \cdots$$

Lemma 12 applied to $u^\tau + u$ and $u^\rho + u$ shows that both are 0. This immediately tells us that $(x^3 + y^3) \cdot A$ is 0 in D. Since A is a nonzero element of D_{2Q-3} this is impossible. Finally if $u = Az^{4Q-1} + \cdots$ has z-degree $4Q - 1$, then $u^\sigma + u = A(x + y)z^{4Q-2} + \cdots$, and we get an element of N_{6Q-5} of z-degree $4Q - 2$; we've shown this can't happen. □

Corollary 14. *Let $R_\alpha = L[x, y, z]/g_\alpha$ where $\alpha \in L$ is transcendental over $Z/2$. Let f be any degree 6 element of R_α and I be the ideal (x^4, y^4, z^4) of R_α. Then $xyf^Q \in I^{[Q]}$ for all Q. Consequently, $f \in I^*$ in R_α.*

Proof. We may assume $Q > 1$. O_{12Q-3} is 1-dimensional, spanned by $(xyz)^{4Q-1}$. If $i + j = 12Q - 3$, multiplication gives a bilinear pairing $O_i \times O_j \to L$, and one sees immediately that the pairing is nondegenerate. Multiplication by g_α gives maps $O_{6Q-5} \to O_{6Q-1}$ and $O_{6Q-2} \to O_{6Q+2}$ that are dual under the above pairings. By Theorem 13 the first of these maps is 1–1. So the second is onto, and in particular xyf^Q lies in its image. In other words, $xyf^Q \in (x^{4Q}, y^{4Q}, z^{4Q}, g_\alpha)$ in $L[x, y, z]$. Passing to R_α we get the result. □

Theorem 15. *Let L be an algebraically closed field of characteristic 2, while $P = z^4 + xyz^2 + (x^3 + y^3)z$ and $P_1 = x^2y^2$. Let R_{gen} be as in Sect. 1. Let f be any degree 6 element of $L[x, y, z]$ and I be the ideal (x^4, y^4, z^4) of $L[x, y, z]$. Then, in the language of Theorem 3, $f \in I^*$ in R_{gen}.*

Proof. $R_{gen} = L(t)[x, y, z]/g_t$, and we use Corollary 14 with L replaced by $L(t)$. □

4 Matrix Calculations: The Case of Algebraic α

Definition 1. Suppose $Q \geq 2$ is a power of 2. Let $M = |m_{i,j}|$ be a matrix with entries in L, where $1 \leq i, j \leq Q$. We say that M is a "special Q-matrix" if the following hold:

(1) $m_{i,j} = 0$ if $i \not\equiv j \quad (3)$ or $\quad i = j$.
(2) If $i \equiv j \quad (3)$ and $i \neq j$, then $m_{i,j} \neq 0$ and depends only on $\mathrm{ord}_2(i - j)$.

Theorem 2. *A special Q-matrix has rank $Q - 2$.*

Proof. We argue by induction on Q. When $Q = 2$, $M = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. When $Q \geq 4$, write M as

$$\begin{pmatrix} M_1 & M_2 & M_3 \\ M_4 & N & M_5 \\ M_6 & M_7 & M_8 \end{pmatrix},$$

where M_1 and M_8 are $\frac{Q}{4}$ by $\frac{Q}{4}$ matrices. Using the fact that M is a special Q-matrix we find that $M_1 = M_8$, $M_2 = M_7$, $M_3 = M_6$, and $M_4 = M_5$, that $M_1 + M_3$ is a nonzero scalar matrix, and that N is a special $\frac{Q}{2}$-matrix. So we may write M as

$$\begin{pmatrix} M_1 & D & M_3 \\ E & N & E \\ M_3 & D & M_1 \end{pmatrix}.$$

Making elementary row and column operations we get

$$\begin{pmatrix} M_1 & D & M_1 + M_3 \\ E & N & 0 \\ M_1 + M_3 & 0 & 0 \end{pmatrix}.$$

Since $M_1 + M_3$ is a nonzero scalar, further elementary operations yield

$$\begin{pmatrix} 0 & 0 & M_1 + M_3 \\ 0 & N & 0 \\ M_1 + M_3 & 0 & 0 \end{pmatrix}.$$

Then $\mathrm{rank}\, M = \mathrm{rank}\, N + 2\left(\frac{Q}{4}\right)$ which is $Q - 2$ by the induction assumption. □

Now let $b(Q) = (b_1, \ldots, b_Q)$ be the row vector of Corollary 9; $b_i = 0$ if $2i - 1 \equiv 2Q - 2i \quad (3)$ and is 1 otherwise. Let $b^*(Q) = (b_1^*, \ldots, b_Q^*)$ be defined as follows: $b_i^* = 1$ if $2i - 1 \equiv 2Q - 2i \quad (3)$ and is 0 otherwise. In other words, $b_i^* = 1 + b_i$. We need a modification of Theorem 2.

Theorem 3. *Let M be a special Q-matrix. Then the $Q + 1$ by Q and $Q + 2$ by Q matrices*

$$\left(\frac{M}{b(Q)}\right) \qquad and \qquad \left(\begin{array}{c} M \\ \hline b(Q) \\ \hline b^*(Q) \end{array}\right)$$

have rank $Q-1$ and Q, respectively.

Proof. Again we argue by induction on Q. When $Q = 2$, $b(Q) = (1,0)$ and $b^*(Q) = (0,1)$. Suppose $Q \geq 4$. Write $b(Q)$ as a concatenation $(F_0|F_1|F_2)$ where F_0 and F_2 have length $Q/4$. Since $b_{i+3} = b_i$, $F_0 = F_2$, and one verifies that $F_1 = b\left(\frac{Q}{2}\right)$. As in the proof of Theorem 2 we may write

$$\left(\frac{M}{b(Q)}\right) \qquad as \qquad \begin{pmatrix} M_1 & D & M_3 \\ E & N & E \\ M_3 & D & M_1 \\ F & b\left(\frac{Q}{2}\right) & F \end{pmatrix}.$$

The same elementary row and column operations that were performed in the proof of Theorem 2 take this matrix to

$$\begin{pmatrix} 0 & 0 & M_1 + M_3 \\ 0 & N & 0 \\ M_1 + M_3 & 0 & 0 \\ 0 & b\left(\frac{Q}{2}\right) & 0 \end{pmatrix}.$$

The rank of this matrix is

$$\operatorname{rank}\left(\frac{N}{b\left(\frac{Q}{2}\right)}\right) + 2\,(Q/4)\,,$$

which is $Q-1$ by the induction assumption. The calculation of the rank of

$$\left(\begin{array}{c} M \\ \hline b(Q) \\ \hline b^*(Q) \end{array}\right)$$

is entirely similar. □

Suppose now that $\alpha \in L^*$ is algebraic over $Z/2$. We attach to α a Q as follows:

Definition 4. Write $\alpha = \lambda^2 + \lambda$ and let $m = m(\alpha)$ be the degree of λ over $Z/2$. Then $Q = 2^{m-1}$. (Since $\alpha \neq 0$, $\lambda \notin Z/2$, and consequently $Q \geq 2$.)

Theorem 5. *Let Q be as in Definition 4. Then the matrix M of Corollary 9 is a special Q-matrix.*

Proof. $m_{i,i} = \sum_{s/Q} \alpha^s = \sum_{s/Q}(\lambda^s + \lambda^{2s}) = \lambda + \lambda^{2^m}$. As the degree of λ over $Z/2$ is m, each $m_{i,i}$ is 0. When $i \not\equiv j \quad (3)$ there are no s such that $i \equiv j \quad (3s)$, and so $m_{i,j} = 0$. When $i \equiv j \quad (3)$, $i \neq j$, let $l = 1 + \mathrm{ord}_2(i - j)$. Then $m_{i,j} = \sum_{s/2^l-1} \alpha^s = \lambda + \lambda^{2^l}$. Now $\mathrm{ord}_2(i - j) < \mathrm{ord}_2(Q)$, and so $l < m$. Thus $m_{i,j} \neq 0$ and only depends on l. □

Corollary 6. *In the situation of Theorem 5, $(N_{6Q-5}) \cap X$ is a 1-dimensional space.*

Proof. Theorems 3 and 5 show that the matrix

$$\left(\frac{M}{b(Q)}\right)$$

of Corollary 9 has rank $Q - 1$. So the induced map $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ of the last section has 1-dimensional kernel. As we've seen, this kernel identifies with $(N_{6Q-5}) \cap X$. □

Now let u be a generator of $(N_{6Q-5}) \cap X$. Our next goal is to show that the coefficient of $x^{4Q-2}y^{Q-2}z^{Q-1}$ in u is nonzero.

Lemma 7. *If u is in X_0, no monomial appearing in u can have the exponent of z equal to $Q - 1$.*

Proof. It's enough to show that no monomial appearing in P^k when $0 \leq k \leq Q - 2$ can have the exponent of z equal to $Q - 1$. Write k as $\sum_1^l b_i$ where the b_i are distinct powers of 2. Since $k < Q - 1$, l is at most $m - 2$. Now P^k is the product of $\left(z^4 + xyz^2 + (x^3 + y^3)z\right)^{b_i}$. This is a sum of terms, each of the form (an element of $Z/2[x, y]) \cdot z^{\sum a_i b_i}$ with each $a_i = 1, 2$, or 4. So if the result fails, $Q - 1$ is a sum of $m - 2$ or fewer powers of 2. Then $Q - 1$ is a sum of $m - 2$ or fewer distinct powers of 2, which is impossible. □

Definition 8. If $v \in D^{sym}_{2Q-1}$, then $\rho(v)$ is the coefficient of $x^{4Q-2}y^{Q-2}z^{Q-1}$ in a pullback of v to X under the map of Definition 4; by Lemma 7 this is independent of the choice of the pullback.

Now a pullback of E_i to X is $E_i P^{Q-1} = E_i \left((x^3 + y^3)z + xyz^2 + z^4\right)^{Q-1}$. So $\rho(E_i)$ is the coefficient of $x^{4Q-2}y^{Q-2}$ in $E_i(x^3 + y^3)^{Q-1}$. By Lemma 6 this is just the b_i^* defined after Theorem 2.

Theorem 9. *If u is a generator of $(N_{6Q-5}) \cap X$, the coefficient of $x^{4Q-2}y^{Q-2}z^{Q-1}$ in u is $\neq 0$.*

Proof. Combining the map $D^{sym}_{2Q-1} \to D^{sym}_{6Q-1} \oplus L$ induced by $X \to (Y, \Delta)$ with ρ, we get a map $D^{sym}_{2Q-1} \to (D^{sym}_{6Q-1} \oplus L) \oplus L$. The discussion above, combined with Theorem 5, shows that with respect to the obvious bases the matrix of this map is

$$\left(\begin{array}{c} M \\ \hline b(Q) \\ \hline b^*(Q) \end{array}\right),$$

where M is a special Q-matrix. By Theorem 3 this matrix has rank Q; consequently $D_{2Q-1}^{sym} \to (D_{6Q-1}^{sym} \oplus L) \oplus L$ is 1–1. The image, $\bar{u}$, of u in D_{2Q-1}^{sym} is $\neq 0$. Since the image of $\bar{u}$ in $D_{6Q-1}^{sym} \oplus L$ is 0, $\rho(\bar{u}) \neq 0$, giving the theorem. □

Theorem 10. *Suppose $\alpha \neq 0$ is algebraic over $Z/2$, that $R_\alpha = L[x, y, z]/g_\alpha$, that $f = y^3z^3$, and that $I = (x^4, y^4, z^4)$. Then there is a Q such that $xyf^Q \notin I^{[Q]}$ in R_α.*

Proof. Take Q as in Definition 4. Let u be as in Theorem 9. Then the coefficient of $(xyz)^{4Q-1}$ in $uxyf^Q$ is the coefficient of $x^{4Q-2}y^{Q-2}z^{Q-1}$ in u, which is $\neq 0$ by Theorem 9. So $uxyf^Q \neq 0$ in O. Since $g_\alpha u = 0$, $xyf^Q \notin g_\alpha O$. In other words, $xyf^Q \notin (x^{4Q}, y^{4Q}, z^{4Q}, g_\alpha)$ in $L[x, y, z]$. Now pass to R_α. □

5 Test Elements

Definition 1. $c \neq 0$ in R_α is a "test element" if whenever J is an ideal of R_α and $h \in J^*$, then $ch^q \in J^{[q]}$ for all q.

Remark 1. Suppose that for each $\alpha \neq 0$, xy is a test element in R_α. Then the localization problem has a negative solution. To see this, take L algebraically closed of characteristic 2. Let $P = z^4 + xyz^2 + (x^3 + y^3)z$, $P_1 = x^2y^2$, $I = (x^4, y^4, z^4)$, and $f = y^3z^3$. We saw in Sect. 3 that $f \in I^*$ in R_{gen}. If $\alpha \neq 0$ in L is algebraic over $Z/2$ then R_α is a domain, and Theorem 10 shows that $xyf^Q \notin I^{[Q]}$ for some Q. Since xy is a test element in R_α, $f \notin I^*$ in R_α. As there are infinitely many such α, Brenner's Theorem 3 gives the result.

Remark 2. In fact, xy is a test element in each R_α. Since each g_α, $\alpha \neq 0$, defines a smooth plane quartic, this is a special case of the following deep result of Brenner. Let $A = L[x, y, z]/g$ where $\operatorname{char} L = p$ and g is a form of degree r defining a smooth projective plane curve. Let J be an ideal of A and $h \in J^*$. Then if $c \in A$ is homogeneous of degree $> r - 3 + \frac{r-3}{p}$, $ch^q \in J^{[q]}$ for all q.

But the proof of this result is deep, using homological algebra, vector bundle theory, and an ampleness criterion of Hartshorne and Mumford. It can't be a part of any short self-contained treatment of our counterexample, and in this exposition, I'll take another route. For clarity write θ for the image of z in $A = R_\alpha = L[x, y, z]/g_\alpha$, so that $A = L[x, y, \theta]$.

Lemma 2. *For each power q of 2 and each j, $(x^3+y^3)^{q-1}\theta^j \in L[x,y,\theta^q]$.*

Proof. $q=1$ is clear. If $q=2$, we may assume $j=1$. But $(x^3+y^3)\theta = \alpha x^2y^2 + xy\theta^2 + \theta^4$, giving the result. Taking q^{th} powers we find that $(x^3+y^3)^q \cdot (\theta^q)^j \in L[x,y,\theta^{2q}]$. We can now prove the lemma by induction on q. Evidently $(x^3+y^3)^{2q-1}\cdot\theta^j = (x^3+y^3)^q \cdot \big((x^3+y^3)^{q-1}\theta^j\big) \in (x^3+y^3)^q \cdot L[x,y,\theta^q]$. But each $(x^3+y^3)^q\cdot(\theta^q)^j$ is in $L[x,y,\theta^{2q}]$. □

Now assume that L is algebraically closed. The arguments that follow are made working in an algebraic closure of the field $L(x,y,\theta)$.

Lemma 3. *Suppose $d \neq 0$ is in $L[x,y]$. Then for each large power, r, of 2 there is an A-linear map $\gamma : L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta^{\frac{1}{r}}] \to A$ taking $d^{\frac{1}{r}}$ to x^3+y^3.*

Proof. $(x^3+y^3)^r\cdot\theta^j \in L[x,y,\theta^r]$. So $(x^3+y^3)\theta^{\frac{j}{r}} \in L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta]$, and $(x^3+y^3)L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta^{\frac{1}{r}}] \subset L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta]$. Now the $x^{\frac{i}{r}}y^{\frac{j}{r}}$, i and $j<r$, form a basis of $L[x^{\frac{1}{r}}, y^{\frac{1}{r}}]$ over $L[x,y]$. Since θ is separable over $L[x,y]$ they are also a basis of $L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta]$ over $L[x,y,\theta]$ (and of $L(x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta^{\frac{1}{r}}) = L(x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta)$ over $L(x,y,\theta)$). We may assume that some monomial appearing in d has coefficient 1. Since r is large, $d^{\frac{1}{r}}$ is an L-linear combination of our basis elements $x^{\frac{i}{r}}y^{\frac{j}{r}}$; also one of the projection maps $p : L(x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta) \to L(x,y,\theta)$ takes $d^{\frac{1}{r}}$ to 1. Let γ be the map $u \to p\big((x^3+y^3)u\big)$. Then $\gamma(d^{\frac{1}{r}}) = x^3+y^3$. Since $(x^3+y^3)\cdot L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta^{\frac{1}{r}}] \subset L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta]$, and each of the projection maps takes this last ring into $L[x,y,\theta] = A$, we're done. □

Lemma 4. *If L is algebraically closed, x^3+y^3 is a test element in $A = R_\alpha$.*

Proof. Suppose J is an ideal of R_α and $h \in J^*$. Then $dh^q \in J^{[q]}$ for some $d \neq 0$ and all q. We may replace d by any A-multiple and may assume $d \neq 0$ is in $L[x,y]$. Choose r and γ as in Lemma 3. Then $dh^{qr} \in J^{[qr]}$, and so $d^{\frac{1}{r}}h^q \in J^{[q]}\cdot L[x^{\frac{1}{r}}, y^{\frac{1}{r}}, \theta^{\frac{1}{r}}]$. Applying γ we find that $(x^3+y^3)h^q \in J^{[q]}$ for all q. □

In the next section we'll use the elementary Lemma 4 in place of Brenner's test element theorem to complete the exposition of the counterexample.

6 The Module H^2: Completion of the Proof

Our goal is as follows:

Lemma 1. *Suppose L is algebraically closed. Let I be the ideal (x^4, y^4, z^4) of $A = R_\alpha = L[x,y,z]/g_\alpha$, where $\alpha \neq 0$. Suppose that c and f are homogeneous elements of R_α of degrees 2 and 6. Then if $cf^Q \notin I^{[Q]}$ for some Q, $f \notin I^*$.*

Note that Lemma 1 and our earlier results provide the negative solution to the localization problem. For we may argue as in Remark 1 following Definition 1, using Theorem 10 and Lemma 1 to see that $(y^3) * (z^3) \notin I^*$ in R_α when α is algebraic over $Z/2$.

Let T be the graded L-algebra $A/(x^{4Q}, y^{4Q}) = L[x, y, z]/(x^{4Q}, y^{4Q}, g_\alpha)$. We develop a few properties of T. Evidently 1, z, z^2, and z^3 form a basis of T over $L[x, y]/(x^{4Q}, y^{4Q})$. So an L-basis of T consists of the $x^i y^j z^k$ with $i, j < 4Q$ and $k < 4$. In particular, T_{8Q+1} is 1-dimensional, spanned by $(xy)^{4Q-1}z^3$. Also, the subspace of T annihilated by x and y is 4-dimensional, spanned by the $(xy)^{4Q-1}z^k$, with $k = 0, 1, 2, 3$. It follows that an element of T is annihilated by x, y, and z if and only if it lies in T_{8Q+1}.

Lemma 2. *If $i + j = 8Q + 1$ the pairing $T_i \times T_j \to L$ induced by multiplication is nondegenerate.*

Proof. We show the left kernel is (0), arguing by induction on j. The case $i = 8Q + 1$, $j = 0$ is trivial. Suppose $i < 8Q + 1$ and $u \in T_i$ annihilates T_j. Then xu, yu, and zu annihilate T_{j-1}. By induction xu, yu, and zu are 0, and since $i < 8Q+1$, $u = 0$. □

For the rest of the section we fix Q with $cf^Q \notin I^{[Q]}$. We shall assume that $f \in I^*$ and get a contradiction.

Lemma 3. *There exists a w in A_{2Q-1} with:*

(1) $z^{4Q}w \in (x^{4Q}, y^{4Q})$.
(2) $f^Q w \notin (x^{4Q}, y^{4Q})$.

Proof. Multiplication by z^{4Q} induces maps $T_{2Q+2} \to T_{6Q+2}$ and $T_{2Q-1} \to T_{6Q-1}$. These maps are dual under the pairings of Lemma 2. Now $cf^Q \notin I^{[Q]}$ in A; consequently cf^Q is not in the image of the first map. So there is a w in the kernel of the second map with $wcf^Q \neq 0$ in T. Thinking of w as an element of A_{2Q-1} we find that $f^Q w \notin (x^{4Q}, y^{4Q})$. Now $w \to 0$ in T_{6Q-1}, and so $z^{4Q}w \in (x^{4Q}, y^{4Q})$. □

Now let K be the field of fractions of A. Then $A\left[\frac{1}{xy}\right]$, $A\left[\frac{1}{x}\right]$, and $A\left[\frac{1}{y}\right]$ are A-submodules of K. Let H^2 be the quotient module $A\left[\frac{1}{xy}\right] / \left(A\left[\frac{1}{x}\right] + A\left[\frac{1}{y}\right]\right)$. ($H^2$ is a local cohomology module but we won't use any machinery from that theory.) Note that H^2 is Z-graded; when u is in A_l, $\frac{u}{x^i y^j}$ has degree $l - i - j$. Using the fact that 1, z, z^2, and z^3 are a basis of A over $L[x, y]$ we find

1. $\frac{1}{x^i y^j}, \frac{z}{x^i y^j}, \frac{z^2}{x^i y^j}$, and $\frac{z^3}{x^i y^j}$, $i, j > 0$ are an L-basis of H^2.
2. $\frac{u}{x^i y^j}$ is 0 in H^2 if and only if $u \in (x^i, y^j)$ in A.

The map $u \to u^2$, $K \to K$ stabilizes $A\left[\frac{1}{xy}\right]$, $A\left[\frac{1}{x}\right]$, and $A\left[\frac{1}{y}\right]$ and so induces an additive function $\Phi : H^2 \to H^2$. Evidently $\Phi(H^2_l) \subset H^2_{2l}$; furthermore $\Phi(aU) = a^2\Phi(U)$. If $q = 2^n$ we abbreviate $\Phi^n(U)$ to $U^{[q]}$.

Lemma 4. *There is a U in H^2_{-1}, $U \neq 0$, such that $(x^3+y^3)\cdot U^{[q]} = 0$ for all q.*

Proof. Take w as in Lemma 3 and let W be the element $\frac{w}{x^{4Q}y^{4Q}}$ of H^2; set $U = f^Q W$. The degree of U is $(2Q-1)-8Q+6Q = -1$. Now $f^Q w \notin (x^{4Q}, y^{4Q})$, and so $U \neq 0$. Also $x^{4Q}w$, $y^{4Q}w$, and $z^{4Q}w$ are all in (x^{4Q}, y^{4Q}), and so $I^{[Q]}\cdot W = (0)$. Applying Φ repeatedly we find that $I^{[qQ]}W^{[q]} = (0)$.

We are assuming that $f \in I^*$. Since x^3+y^3 is a test element in A, $(x^3+y^3)f^{qQ} \in I^{[qQ]}$. So $(x^3+y^3)f^{qQ}W^{[q]} = 0$. But $f^{qQ}W^{[q]} = U^{[q]}$. □

Lemma 5. *Suppose $\alpha \neq 1$ and U is a nonzero element of H^2_{-1}. Then $(x^3+y^3)U^{[8]} \neq 0$.*

Proof (sketch). Since U has degree -1 it is an L-linear combination of $\frac{z}{xy}$, $\frac{z^2}{x^2y}$, $\frac{z^2}{xy^2}$, $\frac{z^3}{x^3y}$, $\frac{z^3}{x^2y^2}$, and $\frac{z^3}{xy^3}$. I'll assume first that U is an L-linear combination of $\frac{z}{xy}$, $\frac{z^2}{x^2y}$, and $\frac{z^2}{xy^2}$. We know from Lemma 2 that $(x^3+y^3)z^{4Q} \equiv (x^3+y^3)\sum_{rs=Q}(C_rP)^s \bmod (x^{4Q}, y^{4Q})$ in the polynomial ring $L[x,y,z]$. So, mod $(x^{4Q}, y^{4Q}, g_\alpha)$, we have $(x^3+y^3)z^{4Q} \equiv (x^3+y^3)\sum_{rs=Q}(C_r\alpha x^2y^2)^s$. Taking $Q=2$ we get

$$(x^3+y^3)\left(\frac{z}{xy}\right)^{[8]} = \frac{x^3+y^3}{x^8y^8}(\alpha^2x^4y^4+\alpha x^4y^4) = (\alpha^2+\alpha)\left(\frac{1}{x^4y}+\frac{1}{xy^4}\right).$$

Taking $Q=4$ we get

$$(x^3+y^3)\left(\frac{z^2}{x^2y}\right)^{[8]} = \frac{x^3+y^3}{x^{16}y^8}(\alpha^4x^8y^8+\alpha^2x^8y^8+\alpha(x^{14}y^2+x^8y^8+x^2y^{14})).$$

So

$$(x^3+y^3)\left(\frac{z^2}{x^2y}\right)^{[8]} = \frac{x^3+y^3}{x^{16}y^8}\cdot\alpha x^{14}y^2 = \frac{\alpha}{x^2y^3}.$$

Similarly,

$$(x^3+y^3)\left(\frac{z^2}{xy^2}\right)^{[8]} = \frac{\alpha}{x^3y^2}.$$

Since $\alpha^2+\alpha \neq 0$, and no L-linear combination of $\frac{1}{x^4y}+\frac{1}{xy^4}$, $\frac{1}{x^2y^3}$, and $\frac{1}{x^3y^2}$ can be 0, we're done. When U is an L-linear combination of all 6 basis elements of H^2_{-1}, one may proceed by making a similar but more elaborate calculation. Alternatively one may use the automorphisms σ, τ, and ρ of Theorem 13, which act on H^2, to construct a nonzero V in H^2_{-1}, with $V^{[8]} = 0$, which is an L-linear combination of $\frac{z}{xy}$, $\frac{z^2}{x^2y}$, and $\frac{z^2}{xy^2}$. We leave details to the reader. □

We can now complete the proof of Lemma 1; we cannot simultaneously have $f \in I^*$ and $cf^Q \notin I^{[Q]}$ for some Q. If $\alpha \neq 1$ this follows from Lemmas 4 and 5. If $\alpha = 1$ we modify the proof of Lemma 5 to show that $(x^3+y^3)U^{[16]} \neq 0$, once again contradicting Lemma 4.

References

1. Brenner, H., Monsky, P.: Tight closure does not commute with localization. Ann. Math. **171**(1), 571–588 (2010)
2. Monsky, P.: Hilbert-Kunz functions in a family: Point S_4-quartics. J. Algebra **208**, 343–358 (1998)

Introduction to the Hyperdeterminant and to the Rank of Multidimensional Matrices

Giorgio Ottaviani

1 Introduction

The classical theory of determinants was placed on a solid basis by Cayley in 1843. A few years later, Cayley himself elaborated a generalization to the multidimensional setting [6] in two different ways. There are indeed several ways to generalize the notion of determinant to multidimensional matrices. Cayley's second attempt has a geometric flavour and was very fruitful. This invariant constructed by Cayley is named today hyperdeterminant (after [15]) and reduces to the determinant in the case of square matrices, which will be referred to as the classical case. The explicit computation of the hyperdeterminant presented from the very beginning exceptional difficulties. Even today explicit formulas are known only in some cases, like the so-called boundary format case and in a few others. In general one has to invoke elimination theory. Maybe for this reason the theory was forgotten for almost 150 years. Only in 1992, thanks to a fundamental paper by Gelfand, Kapranov and Zelevinsky, the theory was placed in the modern language, and many new results have been found. The book [15], of the same three authors, is the basic source on the topic. Also Chap. 9 of [33] is a recommended reading, a bit more advanced, see also [3]. Two sources about classical determinants are [22, 27] (the second one has also a German translation). The extension of the determinant to the multidimensional setting contained in these two sources is based on the formal extension of the formula computing the classical determinant summing over all the permutations (like in Cayley first attempt), and they have different properties from the hyperdeterminant studied in [15] (the paper [16] glimpses a link between the two approaches in the $2 \times 2 \times 2$ case).

G. Ottaviani (✉)
Dipartimento di Matematica, Università di Firenze, viale Morgagni 67/A, 50134 Firenze, Italy
e-mail: ottavian@math.unifi.it

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_20,

In this survey we introduce the hyperdeterminants and some of its properties from scratch. Our aim is to provide elementary arguments when they are available. The main tools we use are the biduality theorem and the language of vector bundles. We will use Geometric Invariant Theory only in Sect. 7. Essentially no results are original, but the presentation is more geometric than the standard one. In particular the basic computation of the dimension of the dual to the Segre variety is performed by describing the contact locus in the Segre varieties.

I wish to thank an anonymous referee for careful reading and several useful suggestions.

2 Multidimensional Matrices and the Local Geometry of Segre Varieties

Let V_i be complex vector spaces of dimension $k_i + 1$ for $i = 0, \dots, p$.

We are interested in the tensor product $V_0 \otimes \dots \otimes V_p$, where the group $GL(V_0) \times \dots \times GL(V_p)$ acts in a natural way.

Once a basis is fixed in each V_i, the tensors can be represented as multidimensional matrices of format $(k_0 + 1) \times \dots \times (k_p + 1)$.

There are $p + 1$ ways to cut a matrix of format $(k_0 + 1) \times \dots \times (k_p + 1)$ into parallel slices (Fig. 1), generalizing the classical description of rows and columns for $p = 1$.

The classical case $p = 1$ is much easier than the case $p \geq 2$ mainly because there are only finitely many orbits for the action of $GL(V_0) \times GL(V_1)$.

Let

$$D_r = \{f \in V_0 \otimes V_1 | \mathrm{rk}\, f \leq r\}. \tag{1}$$

We have that $D_r \setminus D_{r-1}$ are exactly the orbits of this action, and in particular the maximal rank matrices form the dense orbit.

Note that D_1 is isomorphic to the Segre variety $\mathbb{P}(V_0) \times \mathbb{P}(V_1)$ (they were introduced in [29]) and that it coincides with the set of decomposable tensors, which have the form $v_0 \otimes v_1$ for $v_i \in V_i$.

The first remark is

Lemma 1. *The rank of f coincides with the minimum number r of summands in a decomposition $f = \sum_{i=1}^{r} t_i$ with $t_i \in D_1$.*

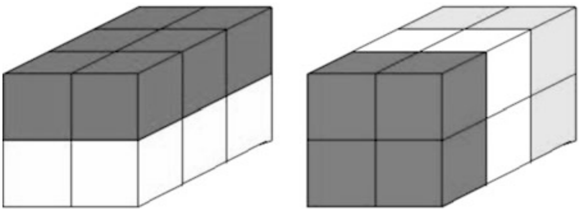


Fig. 1 Two ways to cut a $3 \times 2 \times 2$ matrix into parallel slices

Proof. Acting with the group $GL(V_0) \times GL(V_1)$ f takes the form $f = \sum_{i=1}^{r} v_0^i \otimes v_1^i$, where $\{v_0^i\}$ is a basis of V_0 and $\{v_1^i\}$ is a basis of V_1, corresponding to the matrix

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

In this form the statement is obvious. □

The kth secant variety $\sigma_k(X)$ of a projective irreducible variety X is the Zariski closure of the union of the projective span $< x_1, \ldots x_k >$ where $x_i \in X$. We have a chain of inclusions

$$X = \sigma_1(X) \subset \sigma_2(X) \subset \ldots$$

With this definition, Lemma 1 reads

Corollary 2.

$$\sigma_k(D_1) = D_k.$$

Let us state also, for future reference, the celebrated "Terracini lemma" (see e.g. [35]), whose proof is straightforward by a local computation.

Theorem 3 (Terracini Lemma). *Let X be a projective irreducible variety and let $z \in < x_1, \ldots, x_k >$ be a general point in $\sigma_k(X)$. Then*

$$T_z\sigma_k(X) = < T_{x_1}X, \ldots, T_{x_k}X >.$$

The tangent spaces $T_{x_i}X$ appearing in the Terracini lemma are the projective tangent spaces. Sometimes, we will denote by the same symbol the affine tangent spaces, this abuse of notation should not create any serious confusion.

We illustrate a few properties of the Segre variety $\mathbb{P}(V_0) \times \ldots \times \mathbb{P}(V_p)$.
It is, in a natural way, a projective variety according to the Segre embedding

$$\begin{array}{ccc} \mathbb{P}(V_0) \times \ldots \times \mathbb{P}(V_p) & \longrightarrow & \mathbb{P}(V_0 \otimes \ldots \otimes V_p) \\ (v_0, \ldots v_p) & \mapsto & v_0 \otimes \ldots \otimes v_p. \end{array}$$

In this embedding, the Segre variety coincides with the projectivization of the set of *decomposable* tensors. The proof of the following proposition is straightforward (by induction on p) and we omit it.

Proposition 4. *Every $\phi \in V_0 \otimes \ldots \otimes V_p$ induces for any $i = 0, \ldots, p$ the contraction map*

$$C_i(\phi) \colon V_1^\vee \otimes \ldots \widehat{V_i^\vee} \ldots \otimes V_p^\vee \longrightarrow V_i,$$

where the i th factor is dropped from the source space. The tensor ϕ is decomposable if and only if $rk(C_i(\phi)) \leq 1$ for every $i = 0, \ldots, p$.

Fig. 2 The three flattenings of the matrix in Fig. 1. If the 2-minors of two of them vanish, then the matrix corresponds to a decomposable tensor (a point in the Segre variety)

The previous proposition gives equations of the Segre variety as 2×2 minors of the contraction maps $C_i(\phi)$. These maps are called flattenings, because they are represented by bidimensional matrices obtained like in Fig. 2.

Remark 5. In Proposition 4 it is enough that the rank conditions are satisfied for all $i = 0, \dots, p$ except one.

A feature of the Segre variety is that it contains a lot of linear subspaces.

For any point $x = v_0 \otimes \dots \otimes v_p$, the linear space $v_0 \otimes \dots V_i \dots \otimes v_p$ passes through x for $i = 0, \dots, p$; it can be identified with the fiber of the projection

$$\pi_i : \mathbb{P}^{k_0} \times \dots \times \mathbb{P}^{k_p} \longrightarrow \mathbb{P}^{k_0} \times \dots \widehat{\mathbb{P}^{k_i}} \dots \times \mathbb{P}^{k_p}.$$

We will denote the projectivization of the linear subspace $v_0 \otimes \dots V_i \dots \otimes v_p$. as $\mathbb{P}^{k_i}_x$.

These linear spaces have important properties described by the following proposition.

Proposition 6. *Let $x \in X = \mathbb{P}^{k_0} \times \dots \times \mathbb{P}^{k_p}$.*

(i) The tangent space at x is the span of the $p+1$ linear spaces $\mathbb{P}^{k_i}_x$ that is, T_pX is the projectivization of $\oplus_i v_0 \otimes \dots V_i \dots \otimes v_p$.
(ii) The tangent space at x meets X in the union of the $p+1$ linear spaces $\mathbb{P}^{k_i}_x$.
(iii) Any linear space in X passing through x is contained in one of the $p+1$ linear spaces $\mathbb{P}^{k_i}_x$.

Proof. The tangent vector to a path $v_0(t) \otimes \dots \otimes v_p(t)$ for $t = 0$ is $\sum_{i=0}^{p} v_0(0) \otimes \dots v_i'(0) \dots \otimes v_p(0)$. Since $v_i'(0)$ may be chosen as an arbitrary vector, the statement (i) is clear.

(ii) Fix a basis $\{e_j^0, \dots, e_j^{k_j}\}$ of V_j for $j = 0, \dots, p$ and let $\{e_{j,0}, \dots, e_{j,k_j}\}$ be the dual basis. We may assume that x corresponds to $e_0^0 \otimes \dots \otimes e_p^0$. Consider a decomposable tensor ϕ in the tangent space at x, so $\phi = v_0 \otimes \dots \otimes e_p^0 + \dots + e_0^0 \otimes \dots \otimes v_p$ for some v_i. We want to prove that v_i and e_i^0 are linearly independent for at most one index i. Otherwise we may assume $\dim(v_0, e_0^0) = 2$, $\dim(v_1, e_1^0) = 2$. Consider the contraction

$$C_0(\phi)(e_{1,0} \otimes e_{2,0} \otimes \dots \otimes e_{p,0}) = v_0 + (\dots)e_0^0$$

$$C_0(\phi)(e_{1,1} \otimes e_{2,0} \otimes \dots \otimes e_{p,0}) = \left(e_{1,1}(v_1) + \sum_{j=2}^{p} e_{j,0}(v_j)\right) e_0^0.$$

Since we may assume also $e_{1,1}(v_1) \neq 0$, by replacing $e_{1,1}$ with a scalar multiple we have also $\left(e_{1,1}(v_1) + \sum_{j=2}^{p} e_{j,0}(v_j)\right) \neq 0$. This implies that rank $C_0(\phi) \geq 2$ which is a contradiction. For an alternative approach generalizable to any homogeneous space see [19].

(iii) A linear space in X passing through x is contained in the tangent space at x, hence the statement follows from (ii). □

3 The Biduality Theorem and the Contact Loci in the Segre Varieties

The projective space $\mathbb{P}(V)$ consists of linear subspaces of dimension one of V. The dual space $\mathbb{P}(V^\vee)$ consists of linear subspaces of codimension one (hyperplanes) of V. Hence the points in $\mathbb{P}(V^\vee)$ are exactly the hyperplanes of $\mathbb{P}(V)$.

Let us recall the definition of dual variety. Let $X \subset \mathbb{P}(V)$ be a projective irreducible variety. A hyperplane H is called *tangent* to X if H contains the tangent space to X at some nonsingular point $x \in X$.

The *dual variety* $X^\vee \subset \mathbb{P}(V^\vee)$ is defined as the Zariski closure of the set of all the tangent hyperplanes. Part of the biduality theorem below says that $X^{\vee\vee} = X$, but more is true. Consider the incidence variety $\mathcal{V}$ given by the closure of the set

$$\{(x, H) \in X \times \mathbb{P}(V^\vee) | x \text{ is a smooth point and } T_x X \subset H\}.$$

$\mathcal{V}$ is identified in a natural way with the projective bundle $\mathbb{P}(N(-1)^\vee)$, where N is the normal bundle to X (see Remark 5).

Theorem 1 (Biduality Theorem). *Let $X \subset \mathbb{P}(V)$ be an irreducible projective variety. We have*

$$X^{\vee\vee} = X. \tag{2}$$

Moreover if x is a smooth point of X and H is a smooth point of $X^\vee$, then H is tangent to X at x if and only if x, regarded as a hyperplane in $\mathbb{P}(V^\vee)$, is tangent to $X^\vee$ at H. In other words the diagram

$$\begin{array}{ccc} & \mathcal{V} & \\ \swarrow^{p_1} & & \searrow^{p_2} \\ X & & X^\vee \end{array} \tag{3}$$

is symmetric.

For a proof, in the setting of symplectic geometry, we refer to [15], Theorem 1.1.

Note, as a consequence of the biduality theorem, that the fibers of both the projections of $\mathcal{V}$ over smooth points are linear spaces. This is trivial for the left projection, but it is not trivial for the right one. Let us record this fact.

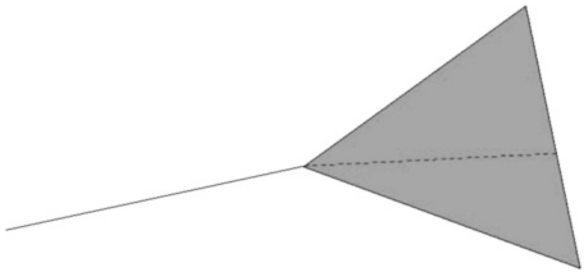


Fig. 3 The tangent space at a point $x \in X = \mathbb{P}^1 \times \mathbb{P}^2$ cuts X into two linear spaces meeting at x; the general hyperplane tangent at x is tangent along a line (*dotted* in the figure)

Corollary 2. *Let X be smooth and let H be a general tangent hyperplane (corresponding to a smooth point of $X^\vee$). Then $\{x \in X | T_x X \subseteq H\}$ is a linear subspace (this is called the contact locus of H in X).*

As a first application we compute the dimension of the dual to a Segre variety.

Theorem 3 (Contact loci in Segre varieties). *Let $X = \mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p}$.*

(i) If $k_0 \geq \sum_{i=1}^{p} k_i$ then a general hyperplane tangent at x is tangent along a linear space of dimension $k_0 - \sum_{i=1}^{p} k_i$ contained in the fiber $\mathbb{P}^{k_0}_x$. In this case the codimension of $X^\vee$ is $1 + k_0 - \sum_{i=1}^{p} k_i$.
(ii) If $k_0 \leq \sum_{i=1}^{p} k_i$ then a general hyperplane tangent at x is tangent only at x. In this case $X^\vee$ is a hypersurface.
(iii) The dual variety $X^\vee$ is a hypersurface if and only if the following holds:

$$\max k_i = k_0 \leq \sum_{i=1}^{p} k_i.$$

Proof. We remind that, by Proposition 6 (i), a hyperplane H is tangent at x if and only if it contains the $p+1$ fibers through x. By Corollary 2 a general hyperplane is tangent along a linear variety. By Proposition 6 (iii) a linear variety in X is contained in one of the fibers. Let H be a general hyperplane tangent at x (Fig. 3). We inspect the fibers through y when $y \in \mathbb{P}^{k_0}_x$. The locus where H contains the fiber $\mathbb{P}^{k_i}_y$ is a linear space in $\mathbb{P}^{k_0}_x$ of codimension k_i indeed the fibers can be globally parametrized by y plus other k_i independent points. This description proves (i), because the variety V in (3) has the same dimension of a hypersurface in $\mathbb{P}(V^\vee)$, and we just computed the general fibers of p_2. Also (ii) follows by the same argument because the conditions are more than the dimension of the space. (iii) is a consequence of (i) and (ii). □

Definition 4. A format $(k_0+1) \times \ldots \times (k_p+1)$ with $k_0 = \max_j k_j$ is called a boundary format if $k_0 = \sum_{i=1}^{p} k_i$. In other words, the boundary format corresponds to the equality in (iii) of Theorem 3.

Remark 5. According to [21], Theorem 3 says that for a Segre variety with normal bundle N, the twist $N(-1)$ is ample if and only if the inequality $\max k_i = k_0 \leq \sum_{i=1}^{p} k_i$ holds.

Note that for $p = 1$ the dual variety to $D_1 = \mathbb{P}^{k_0} \times \mathbb{P}^{k_1}$ is a hypersurface if and only if $k_0 = k_1$ (square case). This is better understood by the following result.

Theorem 6. *Let $k_0 \geq k_1$. In the projective spaces of $(k_0 + 1) \times (k_1 + 1)$ matrices the dual variety to the variety D_r (defined in formula (1)) is D_{k_1+1-r}.*

When $k_0 = k_1$ (square case) the determinant hypersurface is the dual of D_1.

In order to prove the Theorem 6 we need the following proposition.

Proposition 7. *Let X be a irreducible projective variety. For any k*

$$(\sigma_{k+1}(X))^{\vee} \subset (\sigma_k(X))^{\vee} .$$

Proof. The proposition is almost a tautology after the Terracini lemma. The dual to the $(k + 1)$th secant variety $(\sigma_{k+1}(X))^{\vee}$ is defined as the closure of the set of hyperplanes H containing $T_z\sigma_{k+1}(X)$ for z being a smooth point in $\sigma_{k+1}(X)$, so $z \in < x_1, \ldots, x_{k+1} >$ for general $x_i \in X$. By the Terracini lemma (Proposition 3) H contains $T_{x_1}, \ldots T_{x_{k+1}}$ hence H contains $T_{z'}\sigma_k(X)$ for the general $z' \in < x_1, \ldots, x_k >$ (removing the last point). □

Proof of Theorem 6. Due to Proposition 7 and Corollary 2 we have the chain of inclusions

$$D_1^{\vee} \supset D_2^{\vee} \supset \ldots \supset D_{k_1}^{\vee}.$$

By the biduality theorem any inclusion must be strict. Since the $D_i^{\vee}$ are $GL(V_0) \times GL(V_1)$-invariant and the finitely many orbit closures are given by D_i, the only possible solution is that the above chain coincides with

$$D_{k_1} \supset \ldots \supset D_1. \qquad \square$$

Example 8. When $X \subset \mathbb{P}^n$ is the rational normal curve, $\sigma_k(X)$ consists of polynomials which are sums of k powers, while $\sigma_k(X)^{\vee}$ consists of polynomials having k double roots. We get that $\sigma_k(X)^{\vee} = Chow_{2^k,1^{n-2k}}(\mathbb{P}^1)$ according to the notations of Sect. 8.

Remark 9. A common misunderstanding after Theorem 6 is that $X \subset Y$ implies the converse inclusion $X^{\vee} \supset Y^{\vee}$. This is in general false. The simplest counterexample is to take X to be a point of a smooth (plane) conic Y. Here $X^{\vee}$ is a line and $Y^{\vee}$ is again a smooth conic.

Remark 10. The proof of Theorem 6 is short, avoiding local computations, but rather indirect.

We point out the elegant proof of Theorem 6 given by Eisenbud in Prop. 1.7 of [12], which gives more information. Eisenbud considers $V_0 \otimes V_1$ as the space of linear maps $Hom(V_0^\vee, V_1)$ and its dual $Hom(V_1, V_0^\vee)$. These spaces are dual under the pairing $< f, g >:= tr(fg)$ for $f \in Hom(V_0^\vee, V_1)$ and $g \in Hom(V_1, V_0^\vee)$. Eisenbud proves that if $f \in D_r \setminus D_{r-1}$ then the tangent hyperplanes at f to D_r are exactly the g such that $fg = 0$, $gf = 0$. These conditions force the rank of g to be $\leq k_1 + 1 - r$. Conversely any g of rank $\leq k_1 + r - 1$ satisfies these two conditions for some f of rank r, proving Theorem 6.

The above proposition is important because it gives a geometric interpretation of the determinant, as the dual of the Segre variety. This is the notion that better generalizes to multidimensional matrices.

Definition 11. Let

$$\max k_i = k_0 \leq \sum_{i=1}^{p} k_i .$$

The equation of the dual variety to $\mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p}$ is the hyperdeterminant.

A point deserves a clarification. Since the dual variety lives in the dual space, we have defined the hyperdeterminant in the dual space to the space of matrices, and not in the original space of matrices. Although there is no canonical isomorphism between the space of matrices and its dual space this apparent ambiguity can be solved by the invariance.

Indeed we may construct infinitely many isomorphisms between the space $V_0 \otimes \ldots \otimes V_p$ and its dual $V_0^\vee \otimes \ldots \otimes V_p^\vee$ by fixing a basis constructed from the bases of the single spaces V_i. Any function on the space of matrices which is invariant with respect to the action of $SL(V_0) \times \ldots \times SL(V_p)$ induces, by using any isomorphisms $V_i \simeq V_i^\vee$, a function on the dual space, and, due to the invariance, it does not depend on the chosen isomorphism.

4 Degenerate Matrices and the Hyperdeterminant

From now on we will refer to a multidimensional matrix A as simply a matrix, and write $Det(A)$ for its hyperdeterminant (when it exists).

Definition 1. A matrix A is called degenerate if there exists a nonzero $(x^0 \otimes x^1 \otimes \ldots \otimes x^p) \in V_0 \otimes \ldots \otimes V_p$ such that

$$A(x^0, x^1, \ldots, V_i, \ldots, x^p) = 0 \qquad \forall i = 0, \ldots, p. \tag{4}$$

The "kernel" $K(A)$ is by definition the variety of nonzero $(x^0 \otimes x^1 \otimes \ldots \otimes x^p) \in V_0 \otimes \ldots \otimes V_p$ such that (4) is satisfied.

So, by Proposition 6 (i), a matrix A is degenerate if and only if A corresponds to a hyperplane tangent in $K(A)$. This gives an algebraic reformulation of the definition of hyperdeterminant.

We get

Proposition 2. *(i) The (projectivization of the) variety of degenerate matrices of format* $(k_0+1)\times\ldots\times(k_p+1)$ *is the dual variety of the Segre variety* $\mathbb{P}^{k_0}\times\ldots\times\mathbb{P}^{k_p}$.

(ii) Let $\max k_i = k_0 \le \sum_{i=1}^{p} k_i$. *A matrix A is degenerate if and only if* $Det(A)=0$.

Proof. Part (i) is a reformulation of Proposition 6 (i). Part (ii) is a reformulation of Theorem 3 (iii) and Definition 11. □

Degenerate matrices A such that $K(A)$ consists of a single point are exactly the smooth points of the hypersurface.

The degree of the hyperdeterminant can be computed by means of a generating function. Let $N(k_0,\ldots,k_p)$ be the degree of the hyperdeterminant of format $(k_0+1)\times\ldots\times(k_p+1)$, and let $k_0=\max_j k_j$. Set $N(k_0,\ldots,k_p)=0$ if $k_0>\sum_{i=1}^{p} k_i$ (in this case we may set $Det=1$).

Theorem 3 (Gelfand et al. [15] Theorem XIV 2.4).

$$\sum_{k_0,\ldots,k_p\ge 0} N(k_0,\ldots,k_p) z_0^{k_0}\cdots z_p^{k_p} = \frac{1}{\left(1-\sum_{i=2}^{p+1}(i-1)x_i(z_0,\ldots,z_p)\right)^2},$$

where x_i *is the ith elementary symmetric function.*

For example, for $p=2$,

$$\sum_{k_0,k_1,k_2\ge 0} N(k_0,k_1,k_2) z_0^{k_0} z_1^{k_1} z_2^{k_2} = \frac{1}{(1-(z_0z_1+z_0z_2+z_1z_2)-2z_0z_1z_2)^2}.$$

We list a few useful degree of hyperdeterminants, corresponding to formats (a,b,c) with $a+b+c\le 12$.

Note that the degree of format $(2,b,b+1)$ is smaller than the degree of its subformat $(2,b,b)$ (for $b\ge 4$). Therefore a Laplace expansion cannot exist, at least not in a naive way.

In the boundary format case the degree simplifies to $\frac{(k_0+1)!}{k_1!\ldots k_p|}$, as we will see in Sect. 6.

Let us see what happens to the hyperdeterminant after swapping two parallel slices (Fig. 4).

Theorem 4. *Let N be the degree of hyperdeterminant of format* $(k_0+1)\times\ldots\times(k_p+1)$.

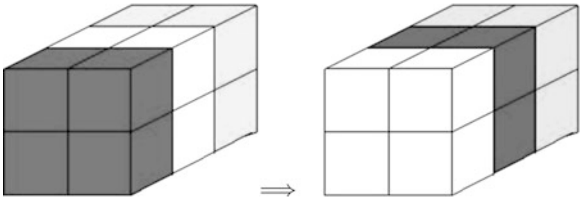


Fig. 4 Swapping two *vertical* slices in the format $3 \times 2 \times 2$ leaves the hyperdeterminant invariant

Format	Degree	Boundary format
(2,2,2)	4	
(2,2,3)	6	*
(2,3,3)	12	
(2,3,4)	12	*
(2,4,4)	24	
(2,4,5)	20	
(3,3,3)	36	
(3,3,4)	48	
(3,3,5)	30	*
(3,4,4)	108	
(3,4,5)	120	
(4,4,4)	272	
(2,b,b)	2b(b-1)	
(2,b,b+1)	b(b+1)	*
(a,b,a+b-1)	$\frac{(a+b-1)!}{(a-1)!(b-1)!}$	*

(i) $\frac{N}{k_i+1}$ *is an integer.*

(ii) After swapping two parallel slices of format $(k_0+1)\times\ldots\widehat{(k_i+1)}\ldots\times(k_p+1)$ *the hyperdeterminant changes its sign if* $\frac{N}{k_i+1}$ *is odd and remains invariant if* $\frac{N}{k_i+1}$ *is even.*

(iii) A matrix with two proportional parallel slices has hyperdeterminant equal to zero.

Proof. It is clear from its definition that the hyperdeterminant is a relative invariant for the group $G = GL(k_0+1)\times\ldots\times GL(k_p+1)$. Moreover the hyperdeterminant is homogeneous on each slice, and by the action, the degree has to be the same on parallel slices. Since there are $(k_i + 1)$ parallel slices, the hyperdeterminant is homogeneous of degree $\frac{N}{(k_i+1)}$ with respect to each slice, which proves (i). Hence for $g \in GL(k_i+1)$ we get $Det(A*g) = Det(A)\cdot(\det(g))^{N/(k_i+1)}$. If g is a permutation matrix, it acts on the parallel slices by permuting them (Fig. 5). In particular if g swaps two slices it satisfies $\det g = -1$; hence (ii) follows. (iii) follows because a convenient g acting on slices makes a whole slice equal to zero. □

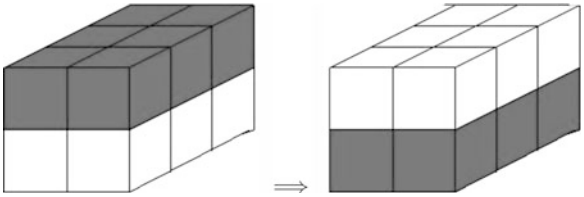


Fig. 5 Swapping two *horizontal* slices in the format $3 \times 2 \times 2$, the hyperdeterminant changes sign

5 Schläfli Technique of Computation

For simplicity we develop Schläfli technique in this section only for three-dimensional matrices, although a similar argument works in any dimension (see [15]).

Let A be a matrix of format $a \times b \times b$. It can be seen as a $b \times b$ matrix with entries which are linear forms over V_0 and we denote it by $\tilde{A}(x)$ with $x \in V_0 \simeq \mathbb{C}^a$ (Fig. 6).

Then we can compute $\det \tilde{A}(x)$ and we get a homogeneous form over V_0 that we can identify with a hypersurface in $\mathbb{P}V_0$.

Schläfli main remark is the following.

Theorem 1 (Schläfli). *Let A be degenerate and let $v_0 \otimes v_1 \otimes v_2 \in K(A)$. Then the hypersurface* $\det \tilde{A}(x)$ *is singular at v_0.*

Proof. After a linear change of coordinates we may assume that $v_0 \otimes v_1 \otimes v_2$ corresponds to the corner of the matrix A, meaning that we choose a basis of V_i where v_i is the first element. Let $\tilde{A} = \sum_{i=0}^{k_1} x_i A_i$ where A_i are the $b \times b$ slices. After a look at Fig. 7, we get that the assumption that $v_0 \otimes v_1 \otimes v_2 \in K(A)$ means that the first slice has the form

$$A_0 = \begin{bmatrix} 0 & 0 & \ldots, & 0 \\ 0 & * & \ldots & * \\ 0 & \vdots & & \vdots \\ 0 & * & \ldots & * \end{bmatrix},$$

and the successive slices have the form

$$A_i = \begin{bmatrix} 0 & * & \ldots, & * \\ * & * & \ldots & * \\ * & \vdots & & \vdots \\ * & * & \ldots & * \end{bmatrix}.$$

Note that $\det A_0 = 0$ and v_0 has coordinates $(1, 0, \ldots 0)$. In an affine coordinate system centred at this point we may assume $x_0 = 1$.

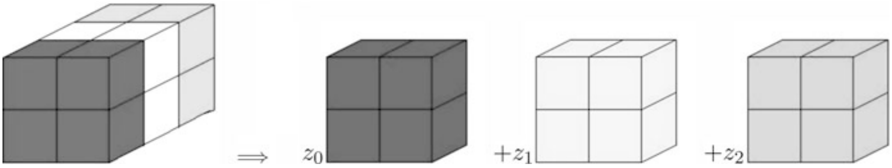
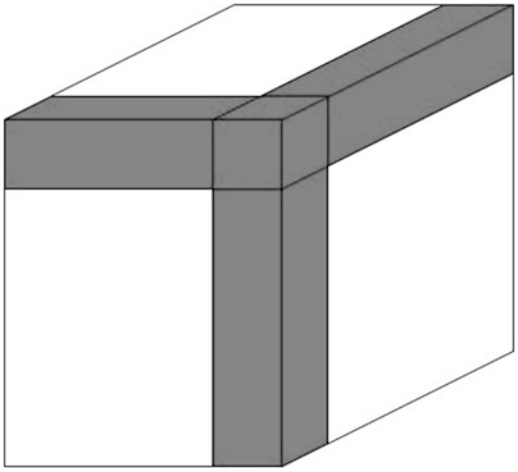


Fig. 6 A tridimensional matrix gives a bidimensional matrix with linear forms as coefficients

Fig. 7 A matrix A is degenerate if and only if it becomes zero on the coloured part after a linear change of coordinates. The kernel $K(A)$ corresponds to the coloured vertex



In conclusion we get

$$\tilde{A} = \begin{bmatrix} 0 & \sum_i l_{1,i}x_i & \dots, & \sum_i l_{k_1,i}x_i \\ \sum_i m_{1,i}x_i & * & \dots & * \\ \vdots & \vdots & & \vdots \\ \sum_i m_{k_1,i}x_i & * & \dots & * \end{bmatrix}.$$

where $l_{j,i}$ and $m_{j,i}$ are scalars.

Now, by expanding the determinant on the first row and then on the first column, we get that $\det \tilde{A}$ has no linear terms in x_i, and then the origin is a singular point, as we wanted. □

The conclusion of Theorem 1 is that for a degenerate matrix A, the discriminant of the polynomial $\det \tilde{A}(x)$ has to vanish. In other words the hyperdeterminant of A divides the discriminant of $\det \tilde{A}(x)$. The following proposition characterizes exactly the cases where the converse holds, and it was used by Schläfli to compute the hyperdeterminant in these cases.

Proposition 2. *(i) For matrices A of format* $2\times b\times b$ *and* $3\times b\times b$ *the discriminant of* $\det \tilde{A}(x)$ *coincides with the hyperdeterminant of A.*

(ii) For matrices A of format $4 \times b \times b$ *the discriminant of* $\det \tilde{A}(x)$ *is equal to the product of the hyperdeterminant and an extra factor which is a square.*
(iii) For matrices A of format $a \times b \times b$ *with* $a \geq 5$*, the discriminant of* $\det \tilde{A}(x)$ *vanishes identically.*

Partial proof (see [15] for a complete proof) (iii) follows because the locus of matrices of rank $\leq b-2$ has codimension 4 in the space of $b \times b$ matrices, then every determinant hypersurface in a space of projective dimension at least four contains points where the rank drops at least by 2, and these are singular points of the hypersurface. In particular for $a \geq 5$, $\det \tilde{A}(x)$ is singular.

Let us prove (i) in the format case $2 \times b \times b$. From Theorem 1 we get that the hyperdeterminant of A divides the discriminant of $\det \tilde{A}(x)$. The degree of the discriminant of a polynomial of degree b is $2(b-1)$; hence the degree of the discriminant of $\det \tilde{A}(x)$ is $2b(b-1)$ with respect to the coefficients of A. Once we know that the degree of the hyperdeterminant of format $2 \times b \times b$ is $2b(b-1)$ the proof is completed. This was stated in the table in the previous section but it was not proved. We provide an alternative argument for $\det \tilde{A}(x)$ singular implies A is degenerate. As a consequence this provides also a proof of the degree formula. After a linear change of coordinates we may assume as in the proof of Theorem 1 that

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & B_0 \end{bmatrix}$$

and

$$A_1 = \begin{bmatrix} c_1 & * \\ * & * \end{bmatrix}.$$

Note that $\det A_0 = 0$ and v_0 has coordinates $(1, 0, \dots 0)$. In an affine coordinate system centred at this point we may assume $x_0 = 1$.

Then $\det \tilde{A} = \det \begin{bmatrix} c_1 x_1 & x_1 \cdot (*) \\ x_1 \cdot (*) & B_0 + x_1 \cdot (*) \end{bmatrix} = c_1 x_1 \det(B_0) + (\text{higher terms in } x_i)$.
If $\det(B_0) \neq 0$ we get $c_1 = 0$, hence A is degenerate. If $\det(B_0) = 0$ we can rearrange A_0 in such a way that the first two rows and two columns vanish. In this case it is easy to arrange a zero in the upper left 2×2 block of A_1 so that again A is degenerate. This pattern can be extended to the $3 \times b \times b$ case. □

Example 3. Regarding (ii) of Proposition 2, we describe the format $4 \times 3 \times 3$. From the table in the previous section, the hyperdeterminant of format $4 \times 3 \times 3$ has degree 48. The discriminant of cubic surfaces has degree 32, and $32 \cdot 3 = 96$. So there is an extra factor of degree 48. It is the square of the invariant coming from the invariant I_8, which is the invariant of minimal degree for cubic surfaces [10].

This gives an interpretation of the invariant I_8. Namely, a smooth cubic surface has I_8 vanishing if and only if it has a determinantal representation as a $4 \times 3 \times 3$ matrix which is degenerate.

Example 4. Let us see the $3 \times 2 \times 2$ example, which can be computed explicitly in at least three different ways. The first way is an application of Schläfli technique and it is described here. The second way is by looking at multilinear systems and it is described in Theorem 3. The third way is described in Example 7.

A matrix A of format $3 \times 2 \times 2$ defines the following 2×2 matrix:

$$\begin{bmatrix} a_{000}x_0 + a_{100}x_1 + a_{200}x_2 & a_{001}x_0 + a_{101}x_1 + a_{201}x_2 \\ a_{010}x_0 + a_{110}x_1 + a_{210}x_2 & a_{011}x_0 + a_{111}x_1 + a_{211}x_2 \end{bmatrix}.$$

The determinant of this matrix defines the following projective conic in the variables x_0, x_1, x_2,

$$x_0^2 \det \begin{bmatrix} a_{000} & a_{001} \\ a_{010} & a_{011} \end{bmatrix} + x_0 x_1 \left(\det \begin{bmatrix} a_{000} & a_{101} \\ a_{010} & a_{111} \end{bmatrix} + \det \begin{bmatrix} a_{100} & a_{001} \\ a_{110} & a_{011} \end{bmatrix} \right) + \ldots = \tag{5}$$

$$= (x_0, x_1, x_2) \cdot C \cdot (x_0, x_1, x_2)^t, \tag{6}$$

where C is a 3×3 symmetric matrix.

By Proposition 2 the hyperdeterminant of A is equal to the determinant of C hence the hyperdeterminant vanishes if and only if the previous conic is singular.

Example 5. An interesting example is given by matrices of format $3 \times 3 \times 3$. The hyperdeterminant of format $3 \times 3 \times 3$ has degree 36. The discriminant of cubic curves has degree 12, and $12 \cdot 3 = 36$.

The three determinants obtained by the three possible directions give three elliptic cubic curves. When one of the three is smooth, then all three are smooth (this happens if and only if $Det \neq 0$), and all three are isomorphic (see [32] and also [23] Proposition 1). So we get three different determinantal representations of the same cubic curve. The Theorem 1 of [23] says that the moduli space of $3 \times 3 \times 3$ matrices under the action of $GL(3) \times GL(3) \times GL(3)$ is isomorphic to the moduli space of triples (C, L_1, L_2) where C is an elliptic curve and L_1, L_2 are two (non isomorphic) line bundles of degree three induced by the pullback of $\mathcal{O}(1)$ under the other two determinantal representations.

Example 6. The $2 \times 2 \times 2$ case gives with an analogous computation the celebrated Cayley formula for A of format $2 \times 2 \times 2$ from the discriminant of the polynomial

$$\det \begin{bmatrix} a_{000}x_0 + a_{010}x_1 & a_{001}x_0 + a_{011}x_1 \\ a_{100}x_0 + a_{110}x_1 & a_{101}x_0 + a_{111}x_1 \end{bmatrix} = 0$$

which is

$$Det(A) = \left(\begin{vmatrix} a_{000} & a_{011} \\ a_{100} & a_{111} \end{vmatrix} + \begin{vmatrix} a_{010} & a_{001} \\ a_{110} & a_{101} \end{vmatrix} \right)^2 - 4 \begin{vmatrix} a_{000} & a_{001} \\ a_{100} & a_{101} \end{vmatrix} \cdot \begin{vmatrix} a_{010} & a_{011} \\ a_{110} & a_{111} \end{vmatrix}.$$

The previous formula expands with exactly 12 summands which have a nice symmetry and are the following: [6]

$$
\begin{gathered}
Det(A) = \left(a_{000}^2 a_{111}^2 + a_{001}^2 a_{110}^2 + a_{010}^2 a_{101}^2 + a_{011}^2 a_{100}^2\right) + \\
-2\left(a_{000}a_{001}a_{110}a_{111} + a_{000}a_{010}a_{101}a_{111} + a_{000}a_{011}a_{100}a_{111} + \right. \\
\left. + a_{001}a_{010}a_{101}a_{110} + a_{001}a_{011}a_{110}a_{100} + a_{010}a_{011}a_{101}a_{100}\right) + \\
+4\left(a_{000}a_{011}a_{101}a_{110} + a_{001}a_{010}a_{100}a_{111}\right).
\end{gathered}
$$

The first four summands correspond to the four line diagonals of the cube the following six summands correspond to the six plane diagonals while the last two summands correspond to the vertices of the even and odd tetrahedra inscribed in the cube.

The $2 \times 2 \times 2$ hyperdeterminant is homogeneous of degree two in each slice and remains invariant under swapping two slices. The $2 \times 2 \times 2$ hyperdeterminant appeared several times recently in the physics literature see for example [11].

Remark 7. According to [17], the hyperdeterminant of format $2 \times 2 \times 2 \times 2$ is a polynomial of degree 24 containing $2,894,276$ terms. A more concise expression has been found in [31], in the setting of algebraic statistics, by expressing the hyperderminant in terms of cumulants. The resulting expression has $13,819$ terms.

6 Multilinear Systems, Matrices of Boundary Format

A square matrix A is degenerate if and only if the linear system $A \cdot x = 0$ has a nonzero solution. This section explores how this notion generalizes to the multidimensional setting by replacing the linear system with a multilinear system (we borrowed some extracts from [26], and I thank Vallès for his permission). The answer is that the hyperdeterminant captures the condition of existence of nontrivial solutions only in the boundary format case.

Let $k_0 = \max_j\{k_j\}$. A matrix A of format $(k_0 + 1) \times \ldots \times (k_p + 1)$ defines the linear map $C_0(A) \in Hom(V_1^\vee \otimes \ldots V_p^\vee, V_0)$ (see Proposition 4) which in turn defines a multilinear system. A nontrivial solution of this system is given by nonzero $x^i \in V_i^\vee$ such that

$$C_0(A)(x^1 \otimes \ldots \otimes x^p) = 0.$$

This is equivalent to the case $i = 0$ of the definition of degenerate (see (4) in Definition 1), namely, to

$$\exists \text{ nonzero } x^1 \otimes \ldots \otimes x^p \in V_1 \otimes \ldots \otimes V_p \text{ such that } A(V_0, x^1, \ldots, x^p) = 0. \quad (7)$$

We say that A satisfying (7) is 0-degenerate (Fig. 8).

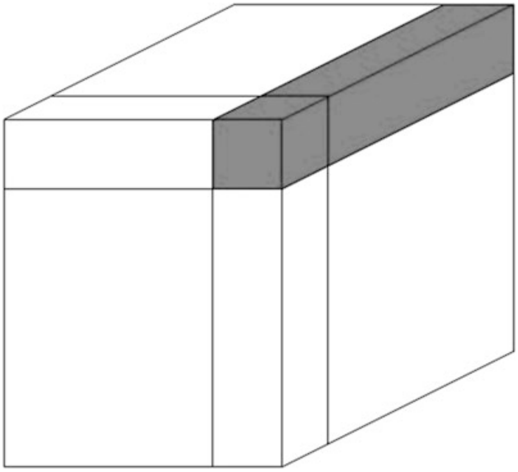


Fig. 8 A is 0-degenerate if it becomes zero on the coloured part after a linear change of coordinates

In the language of [12], the 0-degenerate matrices correspond exactly to the matrices which are not 1-generic. The condition to be not 0-degenerate can be expressed indeed as a Chow condition imposing that the linear subspace $\ker A$ meets the Segre variety. For computations of Chow conditions, see [13].

Theorem 1 (Gelfand et al. [15], Theorem XIV 3.1). *Let $k_0 \geq \sum_{i=1}^{p} k_i$ (in particular in the boundary format). A is 0-degenerate if and only if it is degenerate.*

Proof. If A is degenerate it is obviously 0-degenerate. Let us assume conversely that A is 0-degenerate. By assumption there is a nonzero $(x^1 \otimes \ldots \otimes x^p) \in V_1 \otimes \ldots \otimes V_p$ such that (7) holds.

Consider in the unknown x^0 the linear system (which is (4) for $i = 1$)

$$A(x^0, V_1, x^2, \ldots, x^p) = 0.$$

It consists of k_1 independent equations with respect to (7), because one of the equations is already contained in (7). For the same reason the linear system given by (4) for general i consists of k_i independent equations with respect to (7). Altogether we have $\sum_{i=1}^{p} k_i$ linear equations in the $k_0 + 1$ unknowns which are the coordinates of x^0, since the unknowns are more than the number of equations, we get a nonzero solution as we wanted. □

Theorem 2 (Gelfand et al. [15] Theorem XIV 1.3). *(triangular inequality).*

(i) *If $k_0 \geq \sum_{i=1}^{p} k_i$ the variety of 0-degenerate matrices has codimension $1 + k_0 - \sum_{i=1}^{p} k_i$ in $M(k_0 + 1, \ldots, k_p + 1)$.*
(ii) *If $k_0 < \sum_{i=1}^{p} k_i$ all matrices are 0-degenerate.*
(iii) *The variety of 0-degenerate matrices has codimension 1 in $M(k_0+1, \ldots, k_p+1)$ exactly when the equality $k_0 = \sum_{i=1}^{p} k_i$ holds, that is, in the boundary format case.*

Proof. If $k_0 \geq \sum_{i=1}^{p} k_i$ the codimension is the same as that of the variety of degenerate matrices by Theorem 1. This codimension has been computed in Theorem 3 and Proposition 2, and it is 1 only when the equality holds.

If $k_0 < \sum_{i=1}^{p} k_i$, all matrices are 0-degenerate. Indeed the kernel of $A \in Hom(V_1^\vee \otimes \ldots V_p^\vee, V_0)$ meets the Segre variety by dimensional reasons. □

We get the second promised expression for the $3 \times 2 \times 2$ case.

Theorem 3 (Cayley). *Let A be a matrix of format $3 \times 2 \times 2$ and let A_{00}, A_{01}, A_{10}, A_{11} be the 3×3 submatrices obtained from*

$$\begin{bmatrix} a_{000} & a_{001} & a_{010} & a_{011} \\ a_{100} & a_{101} & a_{110} & a_{111} \\ a_{200} & a_{201} & a_{210} & a_{211} \end{bmatrix}$$

by removing respectively the first column (00), the second (01), the third (10) and the fourth (11). The multilinear system $A(x \otimes y) = 0$ given by $\sum a_{ikj} x_j y_k = 0$ has nontrivial solutions if and only if $\det A_{01} \det A_{10} - \det A_{00} \det A_{11} = 0$ which coincides with $Det(A)$.

Proof. We may assume that the 3×4 matrix has rank 3; otherwise we get a system with only two independent equations which has always a nontrivial solution. A solution $(x_0y_0, x_0y_1, x_1y_0, x_1y_1)$ has to be proportional to $(\det A_{00}, -\det A_{01}, \det A_{10}, -\det A_{11})$ by the Cramer rule. Now the condition is just the equation of the Segre quadric $\mathbb{P}^1 \times \mathbb{P}^1 \subset \mathbb{P}^3$. □

Let $A \in V_0 \otimes \ldots \otimes V_p$ be of boundary format and let $m_j = \sum_{i=1}^{j-1} k_i$ with the convention $m_1 = 0$.

We remark that the definition of m_i depends on the order we have chosen among the k_j's (see Remark 6).

With the above notations the vector spaces $V_0^\vee \otimes S^{m_1}V_1 \otimes \ldots \otimes S^{m_p}V_p$ and $S^{m_1+1}V_1 \otimes \ldots \otimes S^{m_p+1}V_p$ have the same dimension $N = \frac{(k_0+1)!}{k_1!\ldots k_r!}$.

Theorem 4 (and definition of ∂_A). *Let $k_0 = \sum_{i=1}^{p} k_i$. Then the hypersurface of 0-degenerate matrices has degree $N = \frac{(k_0+1)!}{k_1!\ldots k_r!}$, and its equation is given by the determinant of the natural morphism*

$$\partial_A : V_0^\vee \otimes S^{m_1}V_1 \otimes \ldots \otimes S^{m_p}V_p \longrightarrow S^{m_1+1}V_1 \otimes \ldots \otimes S^{m_p+1}V_p.$$

Proof. If A is 0-degenerate then we get $A(v_1 \otimes \ldots \otimes v_p) = 0$ for some $v_i \in V_i^\vee$, $v_i \neq 0$ for $i = 1, \ldots, p$. Then $(\partial_A)^\vee \left(v_1^{\otimes m_1+1} \otimes \ldots \otimes v_p^{\otimes m_p+1}\right) = 0$.

Conversely if A is non-0-degenerate we get a surjective natural map of vector bundles over $X = \mathbb{P}(V_2) \times \ldots \times \mathbb{P}(V_p)$

$$V_0^\vee \otimes \mathcal{O}_X \xrightarrow{\phi_A} V_1 \otimes \mathcal{O}_X(1, \ldots, 1).$$

Indeed, by our definition, ϕ_A is surjective if and only if A is non-0-degenerate.

We construct a vector bundle S over $\mathbb{P}(V_2) \times \ldots \times \mathbb{P}(V_p)$ whose dual $S^\vee$ is the kernel of ϕ_A so that we have the exact sequence

$$0 \longrightarrow S^\vee \longrightarrow V_0^\vee \otimes \mathcal{O} \longrightarrow V_1 \otimes \mathcal{O}(1, \ldots, 1) \longrightarrow 0. \quad (8)$$

After tensoring by $\mathcal{O}(m_2, \ldots, m_p)$ and taking cohomology we get

$$H^0(S^\vee(m_2, m_3, \ldots, m_p)) \longrightarrow V_0^\vee \otimes S^{m_1} V_1 \otimes \ldots \otimes S^{m_p} V_p \xrightarrow{\partial_A} S^{m_1+1} V_1 \otimes \ldots \otimes S^{m_p+1} V_p,$$

and we need to prove

$$H^0(S^\vee(m_2, m_3, \ldots, m_p)) = 0. \quad (9)$$

Let $d = \dim\left(\mathbb{P}(V_2) \times \ldots \times \mathbb{P}(V_p)\right) = \sum_{i=2}^p k_i = m_{p+1} - k_1$.

Since $\det(S^\vee) = \mathcal{O}(-k_1 - 1, \ldots, -k_1 - 1)$ and $rank\ S^\vee = d$, it follows, by using the natural identification $S^\vee \simeq \wedge^{d-1} S \otimes \det(S^\vee)$, that

$$S^\vee(m_2, m_3, \ldots, m_p) \simeq \wedge^{d-1} S \otimes \det(S^\vee)(m_2, m_3, \ldots, m_p),$$

hence

$$S^\vee(m_2, m_3, \ldots, m_p) \simeq \wedge^{d-1} S(-1, -k_1 - 1 + m_3, \ldots, -k_1 - 1 + m_p). \quad (10)$$

Hence, by taking the $(d-1)$st wedge power of the dual of the sequence (8), and using Künneth's formula to calculate the cohomology as in [14], the result follows. □

Corollary 5. *Let $k_0 = \sum_{i=1}^p k_i$. The hyperdeterminant of $A \in V_0 \otimes \ldots \otimes V_p$ is the usual determinant of ∂_A, that is,*

$$Det(A) := det \partial_A, \quad (11)$$

where $\partial_A = H^0(\phi_A)$ and $\phi_A : V_0^\vee \otimes \mathcal{O}_X \xrightarrow{\phi_A} V_1 \otimes \mathcal{O}_X(1, \ldots, 1)$ is the sheaf morphism associated to A. In particular

$$degDet = \frac{(k_0 + 1)!}{k_1! \ldots k_r!}.$$

This is theorem 3.3 of chapter 14 of [15].

Remark 6. Any permutation of the p numbers $k_1, \ldots, k_p$ gives different m_i's and hence a different map ∂_A. As noticed by Gelfand, Kapranov and Zelevinsky, in all cases the determinant of ∂_A is the same by Theorem 4.

Example 7. **The** $3 \times 2 \times 2$ **case (third computation).** In this case the morphism $V_0^\vee \otimes V_1 \to S^2V_1 \otimes V_2$ is represented by a 6×6 matrix, which, with the obvious notations, is the following:

$$\begin{bmatrix} a_{000} & a_{100} & a_{200} & 0 & 0 & 0 \\ a_{001} & a_{101} & a_{201} & 0 & 0 & 0 \\ a_{010} & a_{110} & a_{210} & a_{000} & a_{100} & a_{200} \\ a_{011} & a_{111} & a_{211} & a_{001} & a_{101} & a_{201} \\ 0 & 0 & 0 & a_{010} & a_{110} & a_{210} \\ 0 & 0 & 0 & a_{011} & a_{111} & a_{211} \end{bmatrix}. \tag{12}$$

The hyperdeterminant is the determinant of this matrix. Note that this determinant is symmetric with respect to the second and the third index, but this is not apparent from the above matrix.

Example 8. In the case $4 \times 3 \times 2$ the hyperdeterminant can be obtained as the usual determinant of one of the following two maps:

$$V_0^\vee \otimes V_1 \to S^2V_1 \otimes V_2,$$
$$V_0^\vee \otimes S^2V_2 \to V_1 \otimes S^3V_2.$$

Remark 9. The fact that the degree of the hypersurface of 0-degenerate matrices is $N = \frac{(k_0+1)!}{k_1!\ldots k_p!}$ could be obtained in an alternative way. We know that A is 0-degenerate iff the corresponding $\ker A$ meets the Segre variety.

Hence the condition is given by a polynomial $P(x_1, \ldots, x_m)$ in the variables $x_i \in \mathbb{P}\left(\wedge^{k_0+1}(V_1 \otimes \ldots \otimes V_p)\right)$ of degree equal to the degree of the Segre variety which is $\frac{k_0!}{k_1!\ldots k_p!}$. Since x_i have degree $k_0 + 1$ in terms of the coefficients of A, the result follows.

Let $A = (a_{i_0,\ldots,i_p})$ a matrix of boundary format $(k_0 + 1) \times \ldots \times (k_p + 1)$ and $B = (b_{j_0,\ldots,j_q})$ of boundary format $(l_0+1)\times\ldots\times(l_q+1)$; if $k_p = l_0$ the convolution (or product) $A*B$ (see [15]) of A and B is defined as the $(p+q)$-dimensional matrix C of format $(k_0 + 1) \times \ldots \times (k_{p-1} + 1) \times (l_1 + 1) \times \ldots \times (l_q + 1)$ with entries

$$c_{i_0,\ldots,i_{p-1},j_1,\ldots,j_q} = \sum_{h=0}^{k_p} a_{i_0,\ldots,i_{p-1},h} b_{h,j_1,\ldots,j_q}.$$

Note that C has again boundary format. The following analogue of the *Cauchy–Binet formula* holds.

Theorem 10. *If $A \in V_0 \otimes \ldots \otimes V_p$ and $B \in W_0 \otimes \ldots \otimes W_q$ are boundary format matrices with $dimV_i = k_i + 1$, $dimW_j = l_j + 1$ and $W_0^\vee \simeq V_p$ then $A * B$ satisfies*

$$Det(A * B) = (DetA)^{\binom{l_0}{l_1,\ldots,l_q}} (DetB)^{\binom{k_0+1}{k_1,\ldots,k_{p-1},k_p+1}}. \tag{13}$$

Proof. [9]. □

Corollary 11. *A and B are nondegenerate if and only if $A * B$ is nondegenerate.*

Example 12. From Corolllary 5 the degree of the hyperdeterminant of a boundary format $(k_0 + 1) \times \ldots \times (k_p + 1)$ matrix A is given by

$$N_A = \frac{(k_0 + 1)!}{k_1! \ldots k_p!}.$$

Thus, (13) can be rewritten as

$$Det(A * B) = [(DetA)^{N_B} (DetB)^{N_A}]^{\frac{1}{l_0+1}}.$$

7 Link with Geometric Invariant Theory in the boundary Format Case

In the boundary format case it is well defined as a unique "diagonal" given by elements $a_{i_0 \ldots i_p}$ satisfying $i_0 = \sum_{j=1}^{p} i_j$ (see Fig. 9). We will see in this section how these matrices behave under the action of $SL(V_0) \times \ldots \times SL(V_p)$ in the setting of the geometric invariant theory.

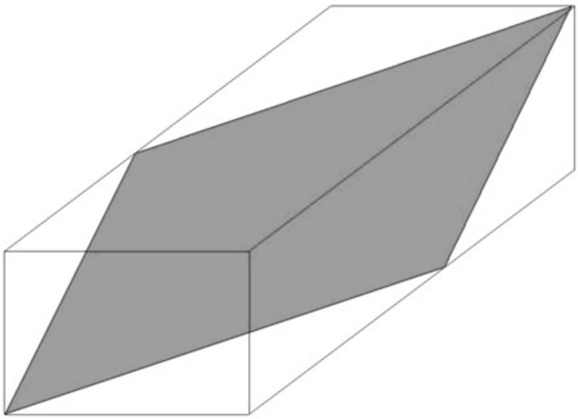


Fig. 9 The diagonal of a boundary format matrix. A triangular matrix fits one of the two half-spaces cut by the diagonal

Definition 1. A $p+1$-dimensional matrix of boundary format $A \in V_0 \otimes \ldots \otimes V_p$ is called triangulable if there exist bases in V_j such that $a_{i_0,\ldots,i_p} = 0$ for $i_0 > \sum_{t=1}^{p} i_t$.

Definition 2. A $p+1$-dimensional matrix of boundary format $A \in V_0 \otimes \ldots \otimes V_p$ is called diagonalizable if there exist bases in V_j such that $a_{i_0,\ldots,i_p} = 0$ for $i_0 \neq \sum_{t=1}^{p} i_t$.

Definition 3. A $p+1$-dimensional matrix of boundary format $A \in V_0 \otimes \ldots \otimes V_p$ is an identity if one of the following equivalent conditions holds:

(i) There exist bases in V_j such that

$$a_{i_0,\ldots,i_p} = \begin{cases} 0 & \text{for} \quad i_0 \neq \sum_{t=1}^{p} i_t, \\ 1 & \text{for} \quad i_0 = \sum_{t=1}^{p} i_t. \end{cases}$$

(ii) There exist a vector space U of dimension 2 and isomorphisms $V_j \simeq S^{k_j} U$ such that A belongs to the unique one-dimensional $SL(U)$-invariant subspace of $S^{k_0}U \otimes S^{k_1}U \otimes \ldots \otimes S^{k_p}U$.

The equivalence between (i) and (ii) follows easily from the following remark: the matrix A satisfies the condition ii) if and only if it corresponds to the natural multiplication map $S^{k_1}U \otimes \ldots \otimes S^{k_p}U \to S^{k_0}U$ (after a suitable isomorphism $U \simeq U^\vee$ has been fixed).

The definitions of triangulable, diagonalizable and identity apply to elements of $\mathbb{P}(V_0 \otimes \ldots \otimes V_p)$ as well. In particular all identity matrices fill a distinguished orbit in $\mathbb{P}(V_0 \otimes \ldots \otimes V_p)$.

The function Det is $SL(V_0) \times \ldots \times SL(V_p)$-invariant in particular if $Det\ A \neq 0$ then A is semistable for the action of $SL(V_0) \times \ldots \times SL(V_p)$. We denote by $Stab\ (A) \subset SL(V_0) \times \ldots \times SL(V_p)$ the stabilizer subgroup of A and by $Stab\ (A)^0$ its connected component containing the identity. The main results are the following.

Theorem 4 (Ancona–Ottaviani [2]). *Let $A \in \mathbb{P}(V_0 \otimes \ldots \otimes V_p)$ of boundary format such that $Det\ A \neq 0$. Then*

$$A \text{ is triangulable} \iff A \text{ is not stable for the action of } SL(V_0) \times \ldots \times SL(V_p)$$

Theorem 5 (Ancona–Ottaviani [2]). *Let $A \in \mathbb{P}(V_0 \otimes \ldots \otimes V_p)$, be of boundary format such that $Det\ A \neq 0$. Then*

$$A \text{ is diagonalizable} \iff Stab(A) \text{ contains a subgroup isomorphic to } \mathbb{C}^*.$$

The proof of the above two theorems relies on the Hilbert–Mumford criterion. The proof of the following theorem needs more geometry.

Theorem 6 (Ancona–Ottaviani [2] for $p = 2$, Dionisi [8] for $p \geq 3$). *Let $A \in \mathbb{P}(V_0 \otimes V_1 \otimes \ldots \otimes V_p)$ of boundary format such that $Det\ A \neq 0$. Then there exists a 2-dimensional vector space U such that $SL(U)$ acts over $V_i \simeq S^{k_i}U$, and* *according to this action on* $V_0 \otimes \ldots \otimes V_p$ *we have* $Stab\ (A)^0 \subset SL(U)$*. Moreover the following cases are possible:*

$$Stab\ (A)^0 \simeq \begin{cases} 0 & \textit{(trivial subgroup)}, \\ \mathbb{C}, & \\ \mathbb{C}^*, & \\ SL(2) & \textit{(this case occurs if and only if } A \textit{ is an identity).} \end{cases}$$

Remark 7. When A is an identity then $Stab\ (A) \simeq SL(2)$.

Example 8. From the expression we have seen in the $3\times2\times2$ case one can compute that the hyperderminant of a diagonal matrix is

$$Det(A) = a_{000}^2 a_{110} a_{101} a_{211}^2.$$

In general the hyperdeterminant of a diagonal matrix is given by a monomial involving all the coefficients on the diagonal with certain exponents; see [34].

8 The Symmetric Case

We analyse the classical *pencil of quadrics* from the point of view of hyperdeterminants.

Let A be a $2 \times n \times n$ matrix with two symmetric $n \times n$ slices A_0, A_1.

Proposition 1. *If* A *is degenerate, then its kernel contains an element of the form* $z \otimes x \otimes x \in \mathbb{C}^2 \otimes \mathbb{C}^n \otimes \mathbb{C}^n$.

Proof. By assumption there are nonzero $z = (z_0, z_1)^t \in \mathbb{C}^2$ and $x, y \in \mathbb{C}^n$ such that $x^t A_i y = 0$ for $i = 0, 1$, $x^t(z_0 A_0 + z_1 A_1) = 0$, $(z_0 A_0 + z_1 A_1)y = 0$. We may assume $z_1 \neq 0$ and $y^t A_0 y = 0$. Pick λ such that $x^t A_0 x + \lambda^2 y^t A_0 y = 0$. By the assumptions we get $z_0(x^t A_0 x + \lambda^2 y^t A_0 y) + z_1(x^t A_1 x + \lambda^2 y^t A_1 y) = 0$ which implies $x^t A_1 x + \lambda^2 y^t A_1 y = 0$. We get $(x + \lambda y)^t A_i (x + \lambda y) = 0$ for $i = 0, 1$, as we wanted. □

Theorem 2. *Let* A *be a* $2 \times n \times n$ *matrix with two symmetric* $n \times n$ *slices* A_0, A_1. *The following are equivalent:*

(i) $Det(A) \neq 0$.
(ii) The characteristic polynomial $\det(A_0 + tA_1)$ *has n distinct roots.*
(iii) The codimension two subvariety intersection of the quadrics A_0, A_1 *is smooth.*

Proof. (i) $\Longrightarrow$ (ii) Assume that $t = 0$ is a double root of $\det(A_0 + tA_1) = 0$. We may assume that A_0, A_1 have the same shape than in the proof of Proposition 2, and the same argument shows that A is degenerate.

(ii) $\Longrightarrow$ (iii) Assume that the point $(1, 0, \ldots, 0)$ belongs to both quadrics, and their tangent spaces in this point do not meet transversely, so we may assume they are both equal to $x_1 = 0$. Hence both matrices have the form

$$A_i = \begin{bmatrix} 0 & a_i & 0 \\ a_i & * & * \\ 0 & * & 0 \end{bmatrix}$$

and $\det(A_0 + tA_1)$ contains the factor $(a_0 + ta_1)^2$, against the assumption.

(iii) $\Longrightarrow$ (i) If A is degenerate, by the Proposition 1, we may assume that the two slices have the same shape as in the proof of Theorem 1. Hence A_0 is singular at the point $(1, 0, \ldots, 0)$ which is common to A_1, which gives a contradiction. □

Proposition 3. *Let A be a $2 \times n \times n$ matrix with two symmetric $n \times n$ slices A_0, A_1. If $Det(A) \neq 0$ then the two quadrics A_i are simultaneously diagonalizable (as quadratic forms), that is, there is an invertible matrix C such that $C^t A_i C = D_i$ with D_i diagonal. The columns of C correspond to the n distinct singular points found for each root of* $\det(A_0 + tA_1)$.

Proof. We may assume that A_0, A_1 are both nonsingular and from Theorem 2 we get distinct λ_i for $i = 1, \ldots, n$ such that $\det(A_0 + \lambda_i A_1) = 0$. For any $i = 1, \ldots n$ we obtain nonzero $v_i \in \mathbb{C}^n$ such that $(A_0 + \lambda_i A_1)v_i = 0$. From these equations we get $\lambda_i(v_j^t A_1 v_i) = v_j^t A_0 v_i = \lambda_j(v_j^t A_1 v_i)$. Hence for $i \neq j$ we get $v_j^t A_1 v_i = 0$ and also $v_j^t A_0 v_i = 0$. Let C be the matrix having v_i as columns. The identities found are equivalent to $C^t A_i C = D_i$ where D_i has $v_1^t A_i v_1, \ldots, v_n^t A_i v_n$ on the diagonal. It remains to show that v_i are independent. This follows because $v_i^t A_0 v_i \neq 0$ for any i; otherwise, $A_0 v_i = 0$ and A_0 should be singular. □

Remark 4. One may assume that $D_0 = diag(\lambda_1, \ldots, \lambda_n)$ and $D_1 = diag(\mu_1, \ldots, \mu_n)$, and in this case $Det(A)$ is proportional to

$$\prod_{i<j} (\lambda_i \mu_j - \lambda_j \mu_i)^2.$$

So there are simultaneously diagonalizable pairs of quadrics with vanishing hyperdeterminant in other words, the converse to Proposition 3 does not hold. The condition for a pair of quadrics to be simultaneously diagonalizable is more subtle. For two smooth conics in the plane ($n = 3$) one has to avoid just the case that the two conics touch in a single point (they can touch in two distinct points and still being simultaneously diagonalizable).

Oeding considers in [24] the case of homogeneous polynomials of degree d in $n + 1$ variables; they give a symmetric tensor of format $(n + 1) \times \ldots \times (n + 1)$ (d times), corresponding to the embedding $S^d\mathbb{C}^{n+1} \subset \otimes^d \mathbb{C}^{n+1}$. The coefficients $a_{i_1,\ldots,i_d}$ of the multidimensional matrix satisfy

$$a_{i_1,\ldots,i_d} = a_{\sigma(i_1),\ldots,\sigma(i_d)}$$

for every permutation σ. Do not confuse this notion with the determinantal representation like in Example 5. For example, the Fermat cubic $x_0^3+x_1^3+x_2^3$ defines a $3 \times 3 \times 3$ tensors with only three entries equal to 1, and its three determinantal representations are all equal to $x_0x_1x_2$, corresponding to three lines.

The first easy result is the following.

Theorem 5. *The discriminant of $f \in S^d\mathbb{C}^{n+1}$ divides the hyperdeterminant of the multidimensional matrix in $\otimes^d\mathbb{C}^{n+1}$ corresponding to f.*

Proof. Let f be singular at v_0; we get that f correspond to a multilinear map $A_f : \mathbb{C}^{n+1} \times \ldots \times \mathbb{C}^{n+1} \to \mathbb{C}$ such that $A_f(v_0, \ldots, v_0) = 0$ and $A_f(\mathbb{C}^{n+1}, v_0, \ldots, v_0) = A_f(v_0, \mathbb{C}^{n+1}, v_0, \ldots, v_0) = \ldots = 0$. Hence the kernel of A_f contains $v_0 \otimes \ldots \otimes v_0$, A_f is degenerate and it has zero hyperdeterminant. □

Oeding proves that the converse is true only in two cases: the square case $n \times n$ and the $2 \times 2 \times 2$ case.

In all the other cases the hyperdeterminant of a symmetric tensor has the discriminant as a factor but contains interesting extra terms.

For example, in the $3 \times 3 \times 3$ case, the hyperdeterminant has degree 36, and it is the product $D \cdot S^6$, where D is the discriminant of degree 12 and S is the Aronhold invariant, which vanishes on plane cubics which are sum of three cubes of linear forms. In other words S is the equation of (the closure of) the $SL(3)$-orbit of the Fermat cubic $x_0^3 + x_1^3 + x_2^3$.

In order to describe the extra terms, let us consider any partition $\lambda = (\lambda_1, \ldots, \lambda_s)$ of d, that is, $d = \lambda_1 + \ldots + \lambda_s$; we may assume $\lambda_1 \geq \ldots \geq \lambda_s$. For any partition λ of d define $Chow_\lambda(\mathbb{P}^n)$ as the closure in $\mathbb{P}(S^d\mathbb{C}^{n+1})$ of the set of polynomials of degree d which are expressible as $l_1^{\lambda_1} \cdots l_s^{\lambda_s}$ where l_i are linear forms.

Theorem 6 (Oeding [24]). *The dual variety $Chow_\lambda(\mathbb{P}^n)^\vee$ is a hypersurface except for the two cases:*

(i) $n = 1$ *and* $\lambda_s = 1$.
(ii) $n \geq 2$ *and* $\lambda = (d-1, 1)$

Let $\Theta_{\lambda,n}$ be the equation of $Chow_\lambda(\mathbb{P}^n)^\vee$ when it is a hypersurface (see Theorem 6).

Theorem 7 (Oeding [24]). *The hyperdeterminant of a symmetric matrix of format $n \times \ldots \times n$ (d times) splits as the product*

$$\prod_\lambda \Theta_{\lambda,n}^{m_\lambda},$$

where m_λ is the multinomial coefficient $\binom{d}{\lambda_1,\ldots\lambda_s}$ and the product is extended over all the partitions such that $Chow_\lambda(\mathbb{P}^n)^\vee$ is a hypersurface, classified by Theorem 6.

There is always the factor corresponding to the trivial partition d. The Chow variety $Chow_d(\mathbb{P}^n)$ is the Veronese variety, and its dual variety is the discriminant (according to Theorem 5), appearing in the product with exponent one. Indeed one sees immediately from Theorem 6 that this is the only factor just in the square case $n \times n$ ($d = 2$) and in the $2 \times 2 \times 2$ case.

9 Weierstrass Canonical Form and Kac's Theorem

Note that the only format $2 \times b \times c$ where the hyperdeterminant exists (so that the triangular inequality is satisfied) are $2 \times k \times k$ and $2 \times k \times (k + 1)$.

The $2 \times k \times k$ case has the same behaviour as the symmetric case considered in Sect. 8. We record the main classification result in the nondegenerate case.

Theorem 1 (Weierstrass). *Let A be a $2 \times k \times k$ matrix and let A_0, A_1 be the two slices. Assume that $Det(A) \neq 0$. Under the action of $GL(K) \times GL(K)$, A is equivalent to a matrix where A_0 is the identity and $A_1 = diag(\lambda_1, \ldots, \lambda_k)$. In this form the hyperdeterminant of A is equal to $\prod_{i<j}(\lambda_i - \lambda_j)^2$.*

The other case $2 \times k \times (k + 1)$ has boundary format, and it was also solved by Weierstrass.

Theorem 2 (Weierstrass). *All nondegenerate matrices of type $2 \times k \times (k + 1)$ are $GL(K) \times GL(K + 1)$ equivalent to the identity matrix having the two slices*

$$\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}.$$

Proof. Let A, A' be two such matrices. Since they are nondegenerate they define two exact sequences on $\mathbb{P}^1$:

$$0 \to \mathcal{O}(-k) \longrightarrow \mathcal{O}^{k+1} \xrightarrow{A} \mathcal{O}(1)^k \to 0.$$

$$0 \to \mathcal{O}(-k) \longrightarrow \mathcal{O}^{k+1} \xrightarrow{A'} \mathcal{O}(1)^k \to 0.$$

We want to show that there is a commutative diagram

$$\begin{array}{ccccccc} 0 \to & \mathcal{O}(-k) & \longrightarrow & \mathcal{O}^{k+1} & \xrightarrow{A} & \mathcal{O}(1)^k & \to 0 \\ & \downarrow{\scriptstyle 1} & \searrow & \downarrow{\scriptstyle f} & & & \\ 0 \to & \mathcal{O}(-k) & \longrightarrow & \mathcal{O}^{k+1} & \xrightarrow{A'} & \mathcal{O}(1)^k & \to 0. \end{array}$$

In order to show the existence of f we apply the functor $Hom(-, \mathcal{O}^{k+1})$ to the first row. We get

$$Hom(\mathcal{O}^{k+1}, \mathcal{O}^{k+1}) \xrightarrow{g} Hom(\mathcal{O}(-k), \mathcal{O}^{k+1}) \to Ext^1(\mathcal{O}(1)^k, \mathcal{O}^{k+1}) \simeq H^1(\mathcal{O}(-1)^{k(k+1)}) = 0.$$

Hence, g is surjective and f exists. Now it is straightforward to complete the diagram with a morphism $\phi\colon \mathcal{O}(1)^k \to \mathcal{O}(1)^k$, which is a isomorphism by the snake lemma. □

Let (x_0, x_1) be homogeneous coordinates on $\mathbb{P}^1$. The identity matrix appearing in Theorem 2 corresponds to the morphism of vector bundles given by

$$I_k(x_0, x_1) := \begin{pmatrix} x_0 & x_1 & & \\ & \ddots & \ddots & \\ & & x_0 & x_1 \end{pmatrix}.$$

It is interesting, and quite unexpected, that the format $2 \times k \times (k+1)$ is a building block for all the other formats $2 \times b \times c$. The canonical form illustrated by the following theorem is called the Weierstrass canonical form (there is an extension in the degenerate case that we do not pursue here).

Theorem 3 (Kronecker). *Let $2 \le b < c$. There exist unique $n, m, q \in \mathbb{N}$ satisfying*

$$\begin{cases} b = nq & + m(q+1) \\ c = n(q+1) + m(q+2) \end{cases}$$

such that the general tensor $t \in \mathbb{C}^2 \otimes \mathbb{C}^b \otimes \mathbb{C}^c$ decomposes under the action of $GL(b) \times GL(c)$ as n blocks $2 \times q \times (q+1)$ and m blocks $2 \times (q+1) \times (q+2)$ in Weierstrass form.

Kac has generalized this statement to the format $2 \le w \le s \le t$ satisfying the inequality $t^2 - wst + s^2 \ge 1$. Note that in these cases the hyperdeterminant does not exist (for $w \ge 3$). The result is interesting because it gives again a canonical form.

Given w, define by the recurrence relation $a_0 = 0, a_1 = 1, a_j = wa_{j-1} - a_{j-2}$. For $w = 2$ get $0, 1, 2, \ldots$ and Kronecker's result. For $w = 3$ get $0, 1, 3, 8, 21, 55, \ldots$ (odd Fibonacci numbers).

Theorem 4 (Kac [18]). *Let $2 \le w \le s \le t$ satisfying the inequality $t^2 - wst + s^2 \ge 1$. Then there exist unique $n, m, j \in \mathbb{N}$ satisfying*

$$\begin{cases} s = na_j & + ma_{j+1} \\ t = na_{j+1} + ma_{j+2} \end{cases}$$

such that the general tensor $t \in \mathbb{C}^w \otimes \mathbb{C}^s \otimes \mathbb{C}^t$ decomposes under the action of $GL(s) \times GL(t)$ as n blocks $w \times a_j \times a_{j+1}$ and m blocks $w \times a_{j+1} \times a_{j+2}$ which are denoted "Fibonacci blocks" (Fig. 10). They can be described by representation theory (see [4]).

The original proof of Kac [18] uses representations of quivers. In [4] there is an independent proof in the language of vector bundles.

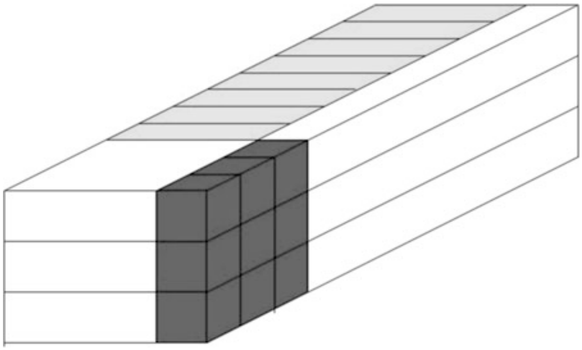


Fig. 10 A decomposition in two Fibonacci blocks

10 The Rank and the k-Secant Varieties

The classical determinant is the equation of the dual variety of the variety of decomposable tensors. This property of the determinant has been chosen as definition of hyperdeterminant in the multidimensional setting.

The determinant gives the condition for a homogeneous linear system to have nontrivial solutions. We have discussed this second property in the multidimensional setting in Sect. 6.

A third property of the determinant is that it vanishes precisely on matrices not of maximal rank.

This property generalizes in the multidimensional setting in a completely different manner, and it is no more governed by the hyperdeterminant.

The rank r of a multidimensional matrix A of format $(k_0 + 1) \times \ldots \times (k_p + 1)$ is the minimum number of decomposable summands $t_i = x_0^i \otimes \ldots \otimes x_n^i$ needed to express it, that is, $A = \sum_{i=1}^{r} t_i$ is minimal.

So the rth secant variety $\sigma_r(\mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p})$ parametrizes multidimensional matrices of rank r and their limits.

The problem of describing these secant varieties is widely open (for the first properties see [5]). Even their dimension is not known in general, although it is conjectured that it coincides with the expected value $r(1 + \sum_{i=1}^{r} k_i) - 1$ (when this number is smaller than the dimension of the ambient space) unless a list of well-understood cases [1]. The analogous result for symmetric multidimensional matrices has been proved by Alexander and Hirschowitz.

The first attempt to find equations of $\sigma_r(\mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p})$ is through the minors of the flattening maps defined in Proposition 4.

Indeed Raicu proved in [28] that the ideal of $\sigma_2(\mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p})$ is generated by the 3×3-minors of $C_i(\phi)$, so proving a conjecture by Garcia, Stillmann and Sturmfels.

These varieties σ_2 are never hypersurfaces, unless the trivial case of $\sigma_2(\mathbb{P}^2 \times \mathbb{P}^2)$, which is given by the classical 3×3 determinant.

The first nontrivial case is given by $\sigma_i(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2)$, which can be described in a uniform way for $i = 1, \dots 4$ by the $2i + 1$-minors of the more general flattening (Young flattening)

$$C_0(\phi)\colon V_0^\vee \otimes V_1 \to \wedge^2 V_1 \otimes V_2$$

for $\phi \in V_0 \otimes V_1 \otimes V_2$ and by the $2i + 1$-minors of the analogous flattening $C_1(\phi)$ and $C_2(\phi)$ obtained by permutations.

For $i = 4$ the secant variety $\sigma_4(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2)$ is a hypersurface in this case $\det C_i(\phi)$ are independent from i and they all give the same hypersurface of degree 9 (see [25] Sect. 3). This hypersurface was found first by Strassen in [30], with a slightly different construction. We emphasize that this hypersurface is different from the hyperdeterminant, which has degree 36, constructed in Example 5.

If A_i are the three slices of A, the matrix of $C_0(\phi)$ can be depicted as

$$\begin{bmatrix} 0 & A_2 & -A_1 \\ -A_2 & 0 & A_0 \\ A_1 & -A_0 & 0 \end{bmatrix}.$$

When A is symmetric, $C_i(\phi)$ are independent from i and appear to be skew-symmetric. In this case the pfaffians of order $2i + 2$ of $C_0(\phi)$ define the ith secant variety of the Veronese variety given by the 3-embedding of $\mathbb{P}^2$ in $\mathbb{P}(S^3\mathbb{C}^3)$. For $i = 3$ we get the Aronhold invariant of degree 4 which is the equation of the orbit of the Fermat cubic (see [20]) that we encountered in Sect. 8.

11 Open Problems

There are a lot of interesting (and difficult) open problems on the subject, starting from looking for equations of secant varieties to the Segre varieties.

Here we propose three problems on the hyperdeterminant that seem tractable (at least the first two) and interesting to me.

Problem 1. Find the equations for the dual varieties to Segre varieties when they are not hypersurfaces, so when $k_0 = \max_j k_j > \sum k_i$ (when they are hypersurfaces the single equation is the hyperdeterminant).

Let us see the example of format $4\times2\times2$. In this case the dual variety to $\mathbb{P}^3\times\mathbb{P}^1\times\mathbb{P}^1$ has codimension 2, by Theorem 3 (i). By Theorem 1 the dual variety consists of matrices which are not 0-degenerate. One sees that if A has format $4 \times 2 \times 2$, the multilinear system $\tilde{A}\colon \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^4$ has a nontrivial solution if and only if the following two conditions hold:

(i) The hyperdeterminant of every $3 \times 2 \times 2$ submatrix of A vanishes.
(ii) $\det(\tilde{A}) = 0$ where $\tilde{A}$ is seen as a 4×4 matrix.

So the equations in (i) and (ii) give the answer to this problem for the format $4\times2\times2$.

Note that the equations of the dual of $\mathbb{P}^{k_0} \times \ldots \times \mathbb{P}^{k_p}$ when $k_0 > \sum k_i$ must contain all the hyperdeterminants of submatrices of boundary format (with $k_0' = \sum k_i$).

Problem 2. Compute the irreducible factors of the hyperdeterminant of a skew-symmetric tensor in $\wedge^d \mathbb{C}^n \subset \otimes^d \mathbb{C}^n$.

This means to extend Oeding Theorem 7 to the skew-symmetric case. In this case even in the square case the classical determinant is not irreducible; indeed it is the square of the pfaffian. The dual varieties to Grassmann varieties will play into the game.

Problem 3. This question is a bit more vague. The definition of hyperdeterminant of boundary format with the linear map ∂_A (compare with Theorem 4 and Corollary 5) can be generalized to other cases where the codimension of the degenerate matrices is bigger than one. Specifically, if $k_0, \ldots, k_p$ are nonnegative integers satisfying $k_0 = \sum_{i=1}^{p} k_i$ then we denote again $m_j = \sum_{i=1}^{j-1} k_i$ with the convention $m_1 = 0$, like in Sect. 6.

Assume we have vector spaces $V_0, \ldots, V_p$ and a positive integer q such that $\dim V_0 = q(k_0 + 1), \dim V_1 = q(k_1 + 1)$ and $\dim V_i = (k_i + 1)$ for $i = 2, \ldots, p$. Then the vector spaces $V_0 \otimes S^{m_1} V_1 \otimes \ldots \otimes S^{m_p} V_p$ and $S^{m_1+1} V_1 \otimes \ldots \otimes S^{m_p+1} V_p$ still have the same dimension, and there is an analogous invariant given by $\det(\partial_A)$. The question is to study the properties of this invariant.

Although this construction can seem artificious, it found an application in the first case $q = p = 2$, leading to the proof [7] that the moduli space of instanton bundles on $\mathbb{P}^3$ is affine.

References

1. Abo, H., Ottaviani, G., Peterson, C.: Induction for secant varieties of Segre varieties. Trans. Amer. Math. Soc. **361**(2), 767–792 (2009)
2. Ancona, V., Ottaviani, G.: Unstable hyperplanes for Steiner bundles and multidimensional matrices. Adv. Geom. **1**(2), 165–192 (2001)
3. Boffi, G., Weyman, J.: Koszul complexes and hyperdeterminants. J. Alg. **230**, 68–88 (2000)
4. Brambilla, C.: Cokernel bundles and Fibonacci bundles. Mathematische Nachrichten **281**(4), 499–516 (2008)
5. Catalisano, M.V., Geramita, A.V., Gimigliano, A.: Ranks of tensors, secant varieties of Segre varieties and fat points. Linear Algebra and Appl. **355**, 263–285 (2002); Linear Algebra Appl. **367**, 347–348 (2003)
6. Cayley, A.: On the theory of linear transformations. Cambridge Math. J. **4**, 1–16 (1845)
7. Costa, L., Ottaviani, G.,Nondegenerate multidimensional matrices and instanton bundles. Trans. Amer. Math. Soc. **355**(1), 49–55 (2003)
8. Dionisi, C.: Stabilizers for nondegenerate matrices of boundary format and Steiner bundles. Revista Matematica Complutense **17**(2), 459–469 (2004)
9. Dionisi, C., Ottaviani, G.: The Cauchy-Binet formula for hyperdeterminants of boundary format. J. Algebra **259**(1), 87–94 (2003)
10. Dolgachev, I.: Lectures on invariant theory. In: London Mathematical Society Lecture Note Series, vol. 296. Cambridge University Press, Cambridge (2003)

11. Duff, M.J.: String triality, black hole entropy, and Cayley's hyperdeterminant. Phys. Rev. D **76**(2), 025017, 4 pp (2007)
12. Eisenbud, D.: Linear sections of determinantal varieties. Amer. J. Math. **110**(3), 541–575 (1988)
13. Eisenbud, D., Schreyer, F.O., Weyman, J.: Resultants and Chow forms via exterior syzygies. J. AMS **16**(3), 537–579 (2003)
14. Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V.: Hyperdeterminants. Adv. Math. **96**, 226–263 (1992)
15. Gelfand, I.M., Kapranov, M.M., Zelevinsky, A.V.: Discriminants, Resultants and Multidimensional Determinants. Birkhäuser, Boston (1994)
16. Gherardelli, F.: Osservazioni sugli iperdeterminanti. Istituto Lombardo (Rend. Sc.) A **127**, 107–113 (1993)
17. Huggins, P., Sturmfels, B., Yu, J., Yuster, D.: The hyperdeterminant and triangulations of the 4-cube. Math. Comput. **77**, 1653–1680 (2008)
18. Kac, V.: Infinite root systems, representations of graphs and invariant theory. Invent. Math. **56**(1), 57–92 (1980)
19. Landsberg, J.M., Manivel, L.: On the projective geometry of rational homogeneous varieties. Comment. Math. Helv. **78**(1), 65–100 (2003)
20. Landsberg, J.M., Ottaviani, G.: Equations for secant varieties of Veronese and other varieties. Annali di Matematica Pura e Applicata. To appear arXiv:1111.4567. DOI 10.1007/s10231-011-0238-6
21. Lazarsfeld, R.: Positivity in algebraic geometry. I. Classical setting: line bundles and linear series. Ergebnisse der Mathematik und ihrer Grenzgebiete, vol. 48. Springer, Berlin (2004)
22. Muir, T.: The Theory of the Determinant in the Historical Order of Development, Vol. 1–4. Dover, New York (1960)
23. Ng, K.O.: The classification of $(3,3,3)$ trilinear forms. J. Reine Angew. Math. **468**, 49–75 (1995)
24. Oeding, L.: The hyperdeterminant of a symmetric tensor. Adv. Math. **231**(3-4), 1308–1326 (2012) arXiv:1107.4659
25. Ottaviani, G.: Symplectic bundles on the plane, secant varieties and Lüroth quartics revisited. In: Casnati, G., Catanese, F., Notari, R. (eds.) Quaderni di Matematica, vol. 21. Vector bundles and Low Codimensional Subvarieties: State of the Art and Recent Developments. Aracne (2008)
26. Ottaviani, G., Vallès, J.: Moduli of vector bundles and group action. Notes for Autumn school in Algebraic Geometry. Wykno (Poland) (2001)
27. Pascal, E.: I determinanti. Hoepli, Milano (1897)
28. Raicu, C.: Secant varieties of Segre–Veronese varieties. arXiv:1011.5867
29. Segre, C.: Sulle varietà che rappresentano le coppie di punti di due piani o spazi. Rend. Circ. Mat. palermo, **V**, 192–204 (1891). Opere, **I**, 173–184, ed. Cremonese, Roma (1957)
30. Strassen, V.: Rank and optimal computation of generic tensors. Linear Algebra Appl. **52/53**, 645–685 (1983)
31. Sturmfels, B., Zwiernik, P.: Binary cumulant varieties. Ann. Combin., To appear. arXiv:1103.0153
32. Thrall, R.M., Chanler, H.: Ternary trilinear forms in the field of complex numbers. Duke Math. J. **4**(4), 678–690 (1938)
33. Weyman, J.: Cohomology of Vector Bundles and Syzygies. Cambridge Tracts in Mathematics, vol. 149. Cambridge (2003)
34. Weyman, J., Zelevinsky, A.V. Singularities of hyperdeterminants of boundary format. Ann. Inst. Fourier, Grenoble **46**(3), 591–644 (1996)
35. Zak, F.: Tangents and secants of algebraic varieties. Translations of Mathematical Monographs, vol. 127. American Mathematical Society, Providence (1993)

Commutative Algebra of Subspace and Hyperplane Arrangements

Hal Schenck and Jessica Sidman

1 Introduction

Let V be a vector space of dimension n over a field $\mathbb{K}$ and let $X_1, \ldots, X_d$ be linear subspaces of V defined by ideals $I_1, \ldots, I_d$. Then $X = X_1 \cup \cdots \cup X_d$ is a (reducible) variety with ideal $I = I_1 \cap \cdots \cap I_d \subset S = \mathbb{K}[x_1, \ldots, x_n]$ in the affine space V or its projectivization $\mathbb{P} = \mathbb{P}(V)$. We call the variety X an *arrangement of linear subspaces*, or *subspace arrangement*.

Arrangements of linear subspaces have connections with a wealth of mathematical objects in areas as diverse as topology, invariant theory, combinatorics, algebraic geometry, and statistics. Arrangements have also recently played a prominent role in applied mathematics, appearing as key players in data mining and generalized principal component analysis [59], in the study of the topological complexity of robot motion planning [21, 44], and in the study of configuration spaces and the Gaudin model of mathematical physics [98]. We give an overview of a number of problems having close connections to commutative algebra and algebraic geometry; the field is very broad, so this survey is selective.

This chapter has two main sections. The first focuses on questions about the ideal of an arrangement of linear subspaces. Although it is just the intersection of a finite collection of ideals generated by linear forms, little was known about even the most basic numerical invariants of a general such ideal until about 10 years ago. Indeed, as recently as 1999, not even a bound on the degrees of the minimal generators of the ideal of an arrangement of d linear spaces was known [29].

H. Schenck (✉)
Mathematics Department, University of Illinois, Urbana, IL 61801, USA
e-mail: schenck@math.uiuc.edu

J. Sidman
Department of Mathematics and Statistics, Mount Holyoke College,
South Hadley, MA 01075, USA
e-mail: jsidman@mtholyoke.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_21,

The second section focuses on hyperplane arrangements. The ideal of a hyperplane arrangement is principal, so interesting results relating commutative algebra and hyperplane arrangements concern ancillary algebraic structures associated to them.

We give an overview of what can be said about the ideals of arbitrary subspace arrangements in Sect. 2.1. We take the point of view that interesting open problems about subspace arrangements are in one of two directions. In the first, the numerical invariants of sufficiently generic subspace arrangements are studied as in Sect. 2.2. By contrast, Sect. 2.3 concerns special subspace arrangements whose ideals are amenable to combinatorial constructions. We set the stage for this discussion in Sect. 2.3.1 by briefly highlighting some methods for studying coordinate subspace arrangements which are defined by square-free monomial ideals. We then describe how subspace arrangements whose ideals are generated by products of binomial linear forms have come up in many interesting contexts and pose several questions regarding them in Sect. 2.3.2. Finally, we discuss arrangements of lines called graph curves (Sect. 2.3.3) and their higher-dimensional analogs (Sect. 2.3.4).

The study of hyperplane arrangements is more well developed than that of arbitrary subspace arrangements, and there are several open questions with long histories. We begin with a quick overview of the area and then focus on three major open questions. In Sect. 3.1 we study the module of derivations tangent to the arrangement and Terao's famous conjecture that freeness of this module depends only on combinatorics. In Sect. 3.2, we introduce the Orlik–Solomon algebra, which is the cohomology ring of the arrangement complement, and we close in Sect. 3.3 with the Orlik–Terao algebra and connections to classical problems on fatpoints and blowups of $\mathbb{P}^2$.

2 Subspace Arrangements

2.1 The Ideal of a Subspace Arrangement

If we wish to understand the ideals of subspace arrangements it is natural to begin by discussing numerical invariants such as the degrees of minimal generators, (Castelnuovo–Mumford) regularity, and the Hilbert function. In this section we will survey what is known generally and then turn to the consideration of subspace arrangements satisfying genericity conditions that simplify the analysis of the ideal in Sect. 2.2.

For a survey of results on free resolutions of the ideal of a subspace arrangement see [89]. There is another interesting line of inquiry relating the Cohen–Macaulay property and subspace arrangements. See, for example, [46, 78, 103].

A natural first approximation to understanding the complexity of an ideal is to bound the degrees of its minimal generators.

Conjecture 1 (Derksen [29]). If I is the ideal of an arrangement of d linear subspaces, then I is generated by homogeneous forms of degree less than or equal to d.

The motivation for Conjecture 1 came from invariant theory. In [29], Derksen showed that if G is a reductive linear group acting on a finite-dimensional vector space V then the generators of the ideal of a certain subspace arrangement give rise to generators of the invariant ring $\mathbb{K}[V]^G$. This construction would close the so-called "Noether gap," providing a bound on the degrees of the generators of $\mathbb{K}[V]^G$ regardless of the characteristic of $\mathbb{K}$ if the degrees of the generators of the ideal of the subspace arrangement were all less than or equal to d.

Sturmfels extended the conjecture in a natural way, asking if the regularity of I is bounded by d. Derksen and the second author proved the following theorem using an idea of Conca and Herzog who were studying the regularity of products of ideals generated by linear forms in [22].

Theorem 2 (Derksen–Sidman [31, 87]). *The ideal of an arrangement of d linear spaces has regularity bounded by d.*

The Conjecture 1 follows as a corollary. Moreover, the regularity bound is sharp.

Example 3. Suppose that I is the ideal of d points that all lie on a line L in $\mathbb{P}^2$. There must be a minimal generator f of I that does not contain L, as the line is not in the arrangement. But any such polynomial f vanishes at d points on the line L so must have at least degree d.

However, it is easy to find examples of arrangements with regularity much lower than the bound in Theorem 2. In fact, arrangements whose ideals have regularity 2 are characterized combinatorially in [38] by associating to each subspace arrangement X a weighted graph $G_X = (V, E)$. There is a vertex v_i in V for each subspace X_i with weight $w(v_i) = 1 + \dim X_i$, and an edge $e \in E$ with weight $w(e) = 1 + \dim(X_i \cap X_j)$ between v_i and v_j if $X_i \cap X_j$ is nonempty. If $\chi_w(G) = \sum_{v \in V} w(v) - \sum_{e \in E} w(e)$, then we have:

Theorem 4 (Theorem 5.1 in [38]). *Let $\mathbb{K}$ be an algebraically closed field. A subspace arrangement X is 2-regular if and only if G_X has a spanning forest F with $\chi_w(F) = 1 + \dim(\operatorname{span} X)$.*

2.2 Numerical Invariants of Generic Subspace Arrangements

To go beyond Theorem 2 we need to put additional restrictions on the subspace arrangements that we consider. In this section we discuss what can be said for generic arrangements. Definition 5 describes one way of characterizing genericity.

Definition 5 (Derksen [30]). If $A \subset [d]$, let $c_A := \operatorname{codim} \cap_{i \in A} X_i$ (where $c_i := c_{\{i\}} = \operatorname{codim} X_i$). We say that the subspaces in X have *transverse* intersections if $c_A = \min\{n, |\sum_{i \in A} c_i|\}$ for all $A \subset [d]$. In [22] arrangements satisfying this condition are said to be *linearly general*.

If Definition 5 is satisfied, then the ideal I and the products of ideal $I_1 \cdots I_d$ define the same projective variety scheme theoretically.

Theorem 6 (Proposition 3.4 in [22]). *Define $J = I_1 \cdots I_d$. If the subspaces in the arrangement defined by $I = I_1 \cap \ldots \cap I_d$ intersect transversally then the ideals I and J agree in degrees $\geq d$.*

Conca and Herzog prove Theorem 6 by first describing the primary decomposition of an arbitrary product of ideals generated by linear forms, and then specializing to the case in which the ideals define subspaces that intersect transversally. The point is that the only embedded prime associated to J is the irrelevant maximal ideal. Thus, the saturation of J is equal to I. The result also follows from the proof of Corollary 4.7 in [30], which shows the equality of the Hilbert functions of I and J in degrees d and higher.

This simplifies the analysis of I in high degrees because the numerical invariants of J are completely combinatorial and may be derived from the codimensions of the nontrivial intersections of subsets of the subspaces.

Theorem 7 (Corollary 4.8 in [30]). *The Hilbert function of $J = I_1 \cdots I_d$ is given by*

$$h_J(m) = \sum_{A \subset [d], c_A < n} (-1)^{|A|} \binom{m + n - 1 - c_A}{n - 1 - c_A}.$$

Of course, if the subspaces in I meet transversally, $h_I(m) = h_J(m)$ for all $m \geq d$, and this gives a formula for the Hilbert function of I in degree at least d. Derksen also gives a formula for the Hilbert series of J and notes that as J has regularity d and is generated by forms of degree d, it must have a linear minimal free resolution. Therefore, the graded Betti numbers of J are always combinatorial invariants of any subspace arrangement.

Arrangements whose subspaces meet transversally are the natural generalization of sets of generic points or lines. For both generic sets of points and generic sets of lines in $\mathbb{P}^n$, it is known that the Hilbert function and Hilbert polynomial agree in all degrees m for which the value of the Hilbert polynomial is less than $\binom{n+m}{m}$. (For lines this is due to [51].) Work in [17] is motivated by the following question:

Question 8. Let X be a subspace arrangement with ideal I and let H_I denote the Hilbert polynomial of I. If the subspaces in X meet transversally, is $h_I(m) = \min\{\binom{n+m}{m}, H_I(m)\}$ for all m?

First steps in understanding the Hilbert function of these generic subspace arrangements in low degrees are given in [17]. The Hilbert function and polynomial of configurations of lines in "grids" is given in [49].

2.3 Arrangements Defined by Products of Linear Forms

If I is the ideal of a subspace arrangement it may be difficult to tell from looking at generators alone that $V(I)$ is an arrangement of linear subspaces. However, if the generators of I factor completely into products of linear forms, then it is easy to see that all of the irreducible components of $V(I)$ must be linear. When this phenomenon occurs it often indicates that the combinatorics of the arrangement is particularly rich, and in some cases it may be used to understand the ideal and the invariants discussed in Sect. 2.1.

In Sect. 2.3.1, we begin by recalling some basic facts about Stanley–Reisner theory which gives a dictionary between ideals of coordinate subspace arrangements, which are generated by square-free monomials, and simplicial complexes. We discuss some alternatives to Stanley–Reisner theory in the study of square-free monomial ideals and discuss ideals generated by products of linear forms that are binomials in Sect. 2.3.2.

2.3.1 Coordinate Subspace Arrangements

Let $S = \mathbb{K}[x_1, \ldots, x_n]$. If the ideals $I_1, \ldots, I_d$ are generated by subsets of the variables in the ring S, we say that the corresponding arrangement X is a *coordinate* or *Boolean* subspace arrangement. The intersection $I = I_1 \cap \cdots \cap I_d$ is generated by square-free monomials. Thus, we see that both the ideals of the individual subspaces and the generators of the ideal of the arrangement may be specified by giving subsets of variables. Hence, there are many possible combinatorial constructions that we might use to describe the generators and minimal primes of I. We will discuss Stanley–Reisner rings, edge ideals, and facet ideals.

Note that the matrices in a minimal free graded resolution of I have only square-free monomials as nonzero entries, so the nonzero Betti numbers live in square-free multidegrees. If we let $\deg x_i = e_i$, the ith standard basis element in $\mathbb{R}^n$, then the minimal free graded resolution of I is $\mathbb{N}^n$ graded, with *multigraded* Betti numbers $\beta_{i,\mathbf{a}}$ for $\mathbf{a} \in \mathbb{N}^n$ satisfying $\sum_{|\mathbf{a}|=j} \beta_{i,\mathbf{a}} = \beta_{i,j}$. Thus, we might suspect that the graded Betti numbers are also purely combinatorial objects. However, they may depend on the characteristic of the ground field. As we shall see in Hochster's theorem below, the graded Betti numbers may be computed using combinatorics once the ground field is fixed.

Stanley–Reisner Rings The most general, well-developed, and widely known program for studying square-free monomial ideals is the theory of Stanley–Reisner ideals. From this point of view, monomial generators correspond to the nonfaces of a simplicial complex Δ, and the topology of Δ can be used to describe numerical invariants of I.

To describe the correspondence more precisely, we set some notation. Let $[n] = \{1, \ldots, n\}$. If $\sigma = \{i_1, \ldots, i_k\} \subset [n]$, then $\overline{\sigma} = [n] \backslash \sigma$. We use the shorthand $\mathbf{x}^\sigma := x_{i_1} \cdots x_{i_k}$ for monomials and for ideals we let $I_\sigma := \langle x_{i_1}, \ldots, x_{i_k} \rangle$.

Given a simplicial complex Δ with vertex set $[n]$, the corresponding Stanley–Reisner ideal $I_\Delta = \langle \mathbf{x}^{\sigma_1}, \ldots, \mathbf{x}^{\sigma_j} \rangle$, where $\sigma_1, \ldots, \sigma_j$ are the minimal nonfaces of Δ. Alternatively, if we are given an ideal $I = \langle \mathbf{x}^{\sigma_1}, \ldots, \mathbf{x}^{\sigma_j} \rangle$, and $I = I_{\tau_1} \cap \cdots \cap I_{\tau_k}$, where $I_{\tau_i} \not\subseteq I_{\tau_j}$ for all i and j we may define Δ to be the simplicial complex with vertex set $[n]$ whose maximal faces are $\overline{\tau_1}, \ldots, \overline{\tau_k}$. The subspaces in the arrangement defined by I_Δ are given by the ideals I_{τ_i}.

Hochster [53] showed that the multigraded Betti numbers may be computed combinatorially, once the characteristic of the field is fixed. We give a reinterpretation of the theorem due to Bayer and Charalambous [9]. To state the theorem we introduce two additional concepts. The *Alexander dual* of a simplicial complex Δ is a simplicial complex $\Delta^* = \{\overline{\sigma} \mid \sigma \notin \Delta\}$. If σ is a face in Δ we define the *link* of σ to be

$$\mathrm{link}_\Delta(\sigma) = \{\tau \in \Delta \mid \tau \cup \sigma \in \Delta, \tau \cap \sigma = \emptyset\}.$$

Theorem 9 (Hochster [53], Theorem 2.4 [9]). *If $I = I_\Delta$ then*

$$\beta_{i,\sigma}(S/I) = \dim_{\mathbb{K}} \tilde{H}_{i-1}(\mathrm{link}_{\Delta^*}(\overline{\sigma}); \mathbb{K}).$$

We can also use the graded Betti numbers of a Stanley–Reisner ideal to compute topological invariants of the complement of a real or complex subspace arrangement, which is a classical problem. See [11] for an introduction. Let $V_\Delta \subset \mathbb{R}^n$ denote the real coordinate subspace arrangement associated to the Stanley–Reisner ideal I_Δ. Peeva et al. [72] use the *lcm-lattice* of least common multiples of monomials ordered by divisibility to show that the cohomology of the complement M_Δ, of V_Δ, is given by the graded Betti numbers of S/I_{Δ^*}.

Theorem 10 (Theorem 3.1 in [72]). *Let Δ be a simplicial complex on $[n]$. If $S = \mathbb{K}[x_1, \ldots, x_n]$ then*

$$\dim \tilde{H}^i(M_\Delta; \mathbb{K}) = \sum_{j \geq 0} \beta_{j,i+j}(S/I_{\Delta^*}).$$

Stanley–Reisner rings are also closely connected to toric varieties—in [10], Bifet, De Concini, and Procesi prove the equivariant cohomology ring of a smooth, complete toric variety is isomorphic (as an algebra) to a Stanley–Reisner ring, and there are analogous results [15, 24, 54] in the non-equivariant setting. See Sect. 12.4 of [23] for an expository account.

Remark 11. Under suitable hypotheses, Goresky and Macpherson [47] have shown that the equivariant cohomology ring of a complex projective algebraic variety with a torus action is isomorphic to the coordinate ring of a subspace arrangement. In [60] examples of the subspace arrangements that arise in the study of Springer fibers are computed. It would be interesting to try to understand additional examples explicitly and to study the combinatorics and commutative algebra of the arrangements that arise from this circle of ideas.

Edge Ideals and Facet Ideals If the generators of I are all square-free *quadratic* monomial ideals, then I may be interpreted as the *edge ideal* $I(G)$ of a graph G with vertex set $[n]$. There is an edge between vertices i and j in G if and only if the monomial $x_i x_j$ is in I. See [100] by Villareal for foundational material. Thinking of a graph as a 1-dimensional simplicial complex, Faridi [41] generalized this notion by defining the *facet ideal* $F(\Delta)$ associated to an arbitrary simplicial complex Δ to be the ideal generated by $\mathbf{x}^\sigma$ for all facets σ in Δ.

To see the subspaces in $V(F(\Delta))$, we need the primary decomposition of $F(\Delta)$, which is given in terms of *minimal vertex covers* in Δ. We say that a set $\sigma \subset V$ is a *vertex cover* in Δ if every facet in Δ is incident to a vertex in σ. A vertex cover is minimal if no proper subset is a vertex cover.

Theorem 12 (Proposition 1 in [41] and Proposition 6.1.16 in [100] for edge ideals). *If Δ is a simplicial complex and $C(\Delta)$ is the set of all minimal vertex covers in Δ, then*

$$I(\Delta) = \cap_{\sigma \in C(\Delta)} I_\sigma.$$

As graded Betti numbers of monomial ideals may depend on the ground field, in general they cannot be computed solely in terms of combinatorics. However, there are situations in which the minimal resolution has a nice description. If G is a graph, define the *complement* G^c to be the graph on the same vertex set, with an edge between vertices i and j if and only if there is no edge between them in G.

Theorem 13 (Fröberg [45]). *The edge ideal of a graph G has a linear resolution if and only if every minimal cycle in G^c has length 3.*

There is a large dictionary of results that translates results about edge ideals into results about graphs and vice versa. Thus the stage is set for proving (or reproving) big results from one category in terms of the other.

2.3.2 Arrangements Defined by Binomial Linear Forms

We saw that there are various ways of associating graphs and simplicial complexes to the ideals of coordinate subspace arrangements and that these associations allow communication between the combinatorial and algebraic perspectives. In this section we see that the combinatorics of subspace arrangements defined by binomial linear forms is also very rich and that there are many open questions regarding these arrangements.

We state our results in terms of partitions of the set $[n] = \{1, \ldots, n\}$ following [13]. We say that a partition π of $[n]$ has shape $\lambda = (\lambda_1, \ldots, \lambda_k)$ where the λ_i are weakly decreasing if its blocks are arranged in weakly decreasing order by size and the ith block has size λ_i. If i and j are in $[n]$, then $i \equiv_\pi j$ if they are in the same block of π. To each partition π we associate the ideal of a linear subspace $I_\pi = \langle x_i - x_j \mid i \equiv_\pi j \rangle$ and a product of linear forms $f_\pi = \Pi_{i \equiv_\pi j, i<j}(x_i - x_j)$. If we fix a shape λ then X_λ is the variety defined by $I_\lambda = \cap_\pi I_\pi$, where the intersection is over all π with shape λ. The varieties X_λ also appear in the survey on subspace arrangements by Björner [11].

A priori, the ideal I_λ need not be generated by products of binomials, but this is indeed the case in Theorems 14 and 15. These results inspired the work in [13] in which the duality between them was formalized and extended.

Theorem 14 (Li–Li [56]). *Let $\lambda = (p, 1, \ldots, 1)$ so that $X_\lambda \subset \mathbb{K}^n$ is the set of all subspaces of $\mathbb{K}^n$ in which at most p-coordinates are equal. Then I_λ is generated by polynomials of the form f_π where π has $p - 1$ blocks.*

Theorem 15 (Lovász [58]). *Let $X = \cup X_\lambda$ where λ has $p - 1$ blocks. Then $I(X)$ is generated by polynomials of the form f_π and π has shape $(p, 1, \ldots, 1)$.*

The original motivation for both of these results came from graph theory. Given a graph G on vertex set $[n]$ define $f_G = \Pi_{i<j,\{i,j\}\in G}(x_i - x_j)$. A subset of the vertices of a graph G is an independence set if no two are joined by an edge. The *independence number* of a graph G is the maximum size of an independence set. Then a subset U of $[n]$ is an independence set if and only if the polynomial f_G is nonzero when all of the variables x_i with $i \in U$ are set equal to each other.

Corollary 16 (Li–Li [56]). *G has independence number at most $p + 1$ if and only if f_G is in the ideal in Theorem 14.*

We say that a graph on $[n]$ is p-colorable if its vertices can be partitioned into p blocks having the property that no block contains a pair of vertices joined by an edge.

Corollary 17 (Lovász [58]). *A graph G fails to be $(p - 1)$ colorable if and only if f_G is in the ideal given in Theorem 15.*

The varieties in Theorems 14 and 15 have connections to representation theory. The symmetric group S_n acts on the intersection lattice L of the braid arrangement $V(\Pi_{i<j}(x_i - x_j))$ by permuting the coordinates of $\mathbb{K}^n$, and the orbits of S_n on L correspond to X_λ where λ is a partition of $[n]$. The arrangement in Theorem 14 is an orbit and the arrangement in Theorem 15 is a union of orbits. Up to sign the polynomials f_π are the Garnir polynomials of [70] which generate Specht modules that are indecomposable representations of the symmetric group if $\operatorname{char} \mathbb{K} \neq 2$. Arrangements X_λ where $\lambda = (m, \ldots, m, 1, \ldots, 1)$ appear in Theorem 5.10 [40] connected with a filtration of a certain module over the rational Cherednik algebra.

Theorems 14 and 15 were generalized to arrangements in the intersection lattices of other reflection groups.

Theorem 18 (Sidman [88]). *Let $X \subset \mathbb{C}^n$ be the set of all points for which the mth power of at most p-coordinates are equal. Then $I(X)$ is generated by $\Pi_{i\equiv_\lambda j, i<j}(x_i^m - x_j^m)$ and λ has $p - 1$ blocks.*

Theorem 19 (Sidman [88]). *Let $X \subset \mathbb{C}^n$ be the set of all subspaces of $\mathbb{C}^n$ in which the mth powers of at least p-coordinates are equal. Then $I(X)$ is generated by polynomials of the form $\Pi_{i\equiv_\lambda j, i<j}(x_i^m - x_j^m)$ and λ has shape $(p + 1, 1, \ldots, 1)$.*

There are many interesting questions in commutative algebra related to arrangements such as these. As the ideals are not monomial ideals, determining a Gröbner basis is an interesting problem. In [25] De Loera showed that the generators given in Theorem 15 form a Gröbner basis with respect to *any* term ordering.

Problem 20. Let X be an arrangement that is a union of orbits in the intersection lattice of the braid arrangement:

(1) Find generators for $I(X)$.
(2) Determine a universal Gröbner basis for $I(X)$.
(3) Compute the Gröbner fan of $I(X)$.
(4) Study the (nonreduced) coordinate subspace arrangements associated to Gröbner degenerations of $I(X)$. Are there Gröbner degenerations that are square-free monomial ideals?
(5) Compute the Hilbert function and Hilbert polynomial of $S/I(X)$.
(6) Compute the graded Betti numbers of $S/I(X)$.
(7) Construct a (minimal) free resolution of the $I(X)$. Will an analog of the lcm-lattice construction of [72] be possible?

Work of Haiman and Woo [50] contains partial progress towards answering questions (1), (2), and (5).

2.3.3 Graph Curves

In this section we discuss subspace arrangements that are unions of lines whose pairwise intersections are determined by the data of a graph. We begin by introducing *abstract graph curves* which are unions of projective lines that give rise to subspace arrangements when they are embedded in $\mathbb{P}^n$ so that each projective line is a linear subspace.

Definition 21. Let $G = (V, E)$ be a graph on $[d]$ in which each vertex has degree less than or equal to three. The *abstract graph curve* associated to G, $C(G)$, is defined to be a union of projective lines L_v, for $v \in V$ in which all singularities are nodes. The line L_v intersects L_u in a node if and only if there is an edge between u and v. We may assume that the lines intersect at $0, 1$, and ∞, as the automorphism group of $\mathbb{P}^1$ can take any three points to these three.

The foundational work on graph curves appeared in [8, 69]. We highlight their results which connect to the theme of our paper—combinatorial descriptions and numerical invariants of ideals.

A given abstract curve may be embedded into projective space in many different ways and thus may give rise to many different subspace arrangements. Since the degree and genus of the image of the embedding are constant after we fix G, the Hilbert polynomial is also constant. However, the more refined numerical invariants, the graded Betti numbers may vary with the embedding.

It is possible to impose a numerical condition on the graph G that gives us a "canonical" ring to study. If each vertex of G has degree exactly three, then the genus of $C(G)$ is given by $g = \frac{d}{2}+1$ and $C(G)$ is the union $2g-2$ lines. Theorem 22 describes when the canonical divisor class gives an embedding of trivalent graph curves in terms of its topology. We say that G is k-connected if the removal of any set of edges of size less than k does not increase the number of its connected components.

Theorem 22 (Proposition 2.5 [8]). *The canonical series on a trivalent graph $C(G)$ is very ample if and only if G is 3-connected.*

Bayer and Eisenbud were interested in Green's conjecture for canonical curves which relates the shape of the graded Betti diagram of the canonical ring of a curve to properties of line bundles on C. Since each component has three nodes, the automorphism group of $C(G)$ is finite, and $C(G) \subset \mathbb{P}^{g-1}$ is a stable curve. The idea was that if Green's conjecture could be proven for the "special" class of graph curves, then this would shed light on the result for general curves. Green's conjecture is stated in terms of the *Clifford index* of a smooth curve C.

Definition 23. If C is a smooth curve of genus g and L is a line bundle on C with $h^0(L), h^1(L) \geq 2$, then $c(L) = \deg L - 2(h^0(L)-1)$ and the *Clifford index*, $\operatorname{cliff} C$, is the minimum value of $c(L)$ over all such L.

Conjecture 24 ([34, 48]). Assume that $\mathbb{K}$ is algebraically closed and $\operatorname{char} \mathbb{K} \neq 2$. Let C be a smooth nonhyperelliptic curve of genus g canonically embedded in $\mathbb{P}^{g-1}$. Then the length of the 2-linear strand in the minimal free resolution of $S/I(C) = g - 2 - \operatorname{cliff} C$.

For trivalent graph curves Bayer and Eisenbud were able to give explicit monomial ideal generators in Proposition 3.1 and showed that if G is also 3-connected then $C(G)$ is arithmetically Cohen–Macaulay. To approach Green's conjecture they needed to extend the definition of the Clifford index to graph curves.

Definition 25. Let G be a connected graph in which each vertex has degree ≤ 3. Let f_G be the minimum of all $|\omega|$ such that $\omega \subset E$ and $G \backslash \omega$ has two connected components that each have nontrivial homology. The *combinatorial Clifford index* of G is $\operatorname{cliff} G = f_G - 2$.

Bayer and Eisenbud proved Green's conjecture for planar graphs.

Theorem 26 (Proposition 7.3 in [8]). *If G is a trivalent planar graph then the length of the 2-linear strand in the minimal free resolution of the homogeneous coordinate ring of $C(G)$ in its canonical embedding is $g - 2 - \operatorname{cliff} G$.*

Bayer and Eisenbud stated many problems about graph curves at the end of their paper, most of which are still open. In particular, they asked what a higher-dimensional analog of a graph curve would look like. In Sect. 2.3.4 we suggest one possibility (which likely does not agree with what they had in mind).

Graph curves have appeared in connection to other ideas in algebraic geometry. Some questions focus more on describing the line bundles on $C(G)$ and Brill–Noether theory than on explicitly algebraic matters [6, 7]. The surjectivity of the Gauss map for general curves of genus 10 or ≥ 12 was proved via degenerations to graph curves in [19]. (See also [62, 63].) Ciliberto and Miranda related planar map colorings to the existence of certain 1-forms on $C(G)$ in [18].

2.3.4 Secant Varieties of Graph Curves

If X is a variety in $\mathbb{P}^n$, its kth secant variety Σ_k is defined to be the closure of the set of all k-planes that meet X in at least $k+1$ points. Thus, the secant variety of a graph curve is a higher-dimensional subspace arrangement in which the subspaces in the arrangement are obtained by taking the linear span of subsets of lines in the graph curve. In this section we discuss results about high-degree graph curves and pose several open questions about graph curves and their secant varieties.

Work of Vermeire [99] and Sidman–Vermeire [90] suggests that the graded Betti diagrams of the (higher) secant varieties of smooth high-degree curves follow a regular pattern and that the theorems that we know for curves are special cases of this general picture. Noting that the secant varieties of graph curves are equidimensional subspace arrangements, Eisenbud and Stillman asked if a similar picture would hold for "high-degree" graph curves.

With this motivation, high-degree graph curves were investigated by Burnham et al. [16]. If $C(G)$ is a graph curve and G has $\geq 2g+2$ vertices, then there is no canonical embedding of $C(G)$ as an arrangement of lines. However, in [16] the authors construct an embedding of $C(G)$ into $\mathbb{P}^{d-g}$ whose ideal has an explicit description in terms of products of monomials and binomial linear forms (Theorem 1.7 in [16]).

Cycles in G impose constraints on the embeddings of $C(G)$ that prevent the graded Betti diagrams of high-degree graph curves from behaving like the graded Betti diagrams of their smooth counterparts. For example, if C is a smooth curve of degree $d \geq 2g+1+p$, then property N_p is satisfied. However, N_p fails for $C(G)$ if G has girth $\leq p+2$ (Corollary 3.3 in [16]), and it is not always possible to construct a graph on d vertices with genus g so that $d \geq 2g+1+p$ and girth at least $p+2$. The girth of G also constrains syzygies in the 3-linear strand of the resolution of the first secant variety of $C(G)$.

Here are several open problems about graph curves and their secant varieties.

Question 27. Is there a nice combinatorial description of the secant varieties of a graph curve? Is there a description that not only encodes the pairwise intersections of spaces but also the entire intersection lattice?

One possibility for encoding the combinatorics of equidimensional arrangements was given in [3] in which Abo, Kley, and Peterson use liaison to relate 2-plane arrangements to certain smooth surfaces of general type in $\mathbb{P}^4$. Two m-dimensional subspaces of $\mathbb{P}^n$ generically have intersection of dimension $n-2m$, so one might propose to record only the intersections whose dimension exceeds the expected dimension.

Definition 28 (Definition 3.8 in [3]). If X is an arrangement of d subspaces of dimension m in $\mathbb{P}^r$, the *incidence graph* of X has vertex set $[d]$ and vertices i and j are joined by an edge if $\dim X_i \cap X_j > n-2m$.

As noted in Remark 3.9 of [3], the incidence graph only captures pairwise intersections, so the triangle is the incidence graph of three 2-planes in $\mathbb{P}^4$ meeting in a line as well as three 2-planes that meet pairwise in three distinct lines. Moreover, the incidence graph does not capture any information about the dimension of the intersections of subspaces. The incidence graph associated to two 2-planes in $\mathbb{P}^5$ that meet in a line is the same as the incidence graph associated to two 2-planes in $\mathbb{P}^5$ meeting in a point.

Question 29 (Question 4.2 in [16]). Give conditions ensuring that the secant varieties of a graph curve have ideals generated by products of linear forms.

Problem 30 (Compare to Conjecture 4.3 in [16]). Determine the conditions under which the kth secant variety of C_G is Cohen–Macaulay and has regularity equal to $2k+1$.

Conjecture 31 (Conjecture 4.4 in [16]). Let G be a graph with d vertices embedded in $\mathbb{P}^n$ via Theorem 1.3 in [16]. If n is the girth of G, $d=2g+p+1$, and $n-2 \leq p$ then property N_p fails and $\beta_{n-2,n}$ counts the number of cycles of length n in G.

3 Hyperplane Arrangements

The simplest class of subspace arrangements are hyperplane arrangements: every subspace $X_i = \mathbf{V}(l_i)$ is a hyperplane. Since the ideal $I_X = \prod_{i=1}^{d} l_i \subseteq S = Sym(V^*) = \mathbb{K}[x_1, \ldots x_n]$ is principal, rather than investigating I, the study of hyperplane arrangements focuses on a host of subtle and intricate questions on the geometry and topology of the complement $M = V \setminus X$, where V is a $\mathbb{K}$-vector space, with $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. The original motivating question was to determine the fundamental group and cohomology ring of M. Since projectivizing and adding the hyperplane at infinity leaves M is unchanged, henceforth we assume the arrangement is central and essential. Central simply means that $X_i = \mathbf{V}(l_i)$, with l_i homogeneous, while essential means that after a change of coordinates, X contains the coordinate hyperplanes $\mathbf{V}(x_1), \ldots, \mathbf{V}(x_n)$; hence $M \not\simeq M' \times \mathbb{K}^m$. The relationship between the projective and affine complements is easy to see: $M \simeq k^* \times \mathbb{P}(M)$, with the $\mathbb{K}^*$ factor corresponding to lines through the origin. In keeping with the literature, in this section we write

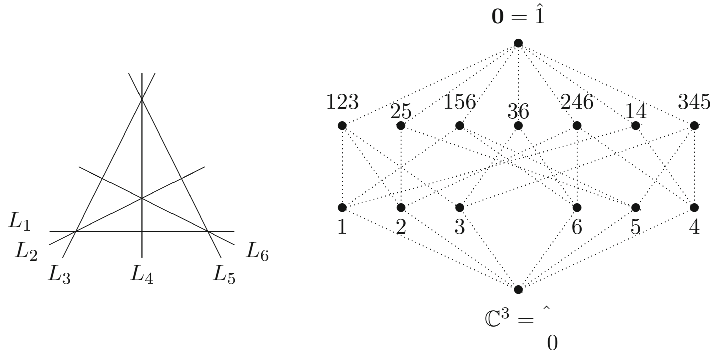


Fig. 1 The braid arrangement $A_3 \subseteq \mathbb{P}^2$ and $L(A_3) \subseteq \mathbb{C}^3$

$$X = \mathcal{A} = \bigcup_{i=1}^{d} H_i \subseteq V$$

and assume unless noted otherwise that $\mathbb{K} = \mathbb{C}$. Orlik and Solomon [65] discovered a beautiful combinatorial description of the cohomology ring $H^*(M, \mathbb{Z})$, which is determined entirely in terms of the intersection lattice $L(\mathcal{A})$. This lattice (in the graded poset sense) consists of the intersections of elements of $\mathcal{A}$, ordered by reverse inclusion. The ambient vector space $V = \hat{0}$, rank-one elements of $L_{\mathcal{A}}$ are hyperplanes, and the origin is $\hat{1}$.

Example 1. The six equations $x_i - x_j = 0$, where $1 \leq i < j \leq 4$, define a hyperplane arrangement in $\mathbb{K}^4$. This is difficult to visualize, but note that the arrangement is not essential: the common intersection is $W = \text{Span}(1, 1, 1, 1)$. Projecting to $W^{\perp}$ results in an essential, central arrangement in $\mathbb{K}^3$. A real version is depicted on the left below, as a collection of lines $\mathcal{A}$ in $\mathbb{P}^2$. The right-hand figure depicts the intersection lattice $L(\mathcal{A}) \subseteq \mathbb{K}^3$ (Fig. 1).

This arrangement is known as the braid arrangement; it is relevant to mathematical physics as a configuration space for non-colliding points.

The last two ingredients we need from combinatorics are the Möbius function and Poincaré polynomial.

Definition 2. The Möbius function $\mu : L(\mathcal{A}) \longrightarrow \mathbb{Z}$ is given by

$$\begin{aligned} \mu(\hat{0}) &= 1 \\ \mu(t) &= -\sum_{s<t} \mu(s), \text{ if } \hat{0} < t. \end{aligned}$$

The Poincaré polynomial $\pi(\mathcal{A}, t)$ is defined as

$$\pi(\mathcal{A}, t) = \sum_{x \in L(\mathcal{A})} \mu(x) \cdot (-t)^{\mathrm{rank}(x)}.$$

The triple points of A_3 have $\mu = 2$, and the double points have $\mu = 1$, so $\pi(A_3, t) = 1 + 6t + 11t^2 + 6t^3$. With the combinatorial preliminaries in hand, we now describe three central problems on the commutative algebra of hyperplane arrangements.

3.1 Freeness of the Module of Derivations $D(\mathcal{A})$

The most famous open conjecture in the theory of arrangements is Terao's conjecture that the freeness of the module of derivations depends only on $L(\mathcal{A})$.

Definition 3. The module $D(\mathcal{A})$ is defined via

$$D(\mathcal{A}) = \{\theta \in Der_{\mathbb{K}}(S) | \theta(l_i) \in \langle l_i \rangle \ \forall \ \mathbf{V}(l_i) \in \mathcal{A}\}.$$

$\mathcal{A}$ is said to be *free* if $D(\mathcal{A})$ is a free module.

Let $F = \prod_{i=1}^{d} l_i$ be the defining equation for $\mathcal{A}$. Since F is homogeneous, the Euler derivation $E = \sum_{i=1}^{n} x_i \partial/\partial x_i$ is in $D(\mathcal{A})$. Our assumption that $\mathcal{A}$ is central means that $D(\mathcal{A})$ is graded, with grading given by the coefficients of θ, so the Euler derivation is of degree one. In fact, as long as $char(\mathbb{K}) = 0$, the map $\theta \mapsto \theta(F)$ gives a surjection from $D(\mathcal{A}) \to \langle F \rangle$, with kernel consisting of $\sum_{i=1}^{n} f_i \partial F/\partial x_i = 0$. Hence, we have a short exact sequence

$$0 \longrightarrow D_0(\mathcal{A}) \longrightarrow D(\mathcal{A}) \longrightarrow \langle F \rangle \longrightarrow 0,$$

where $D_0(\mathcal{A})$ consists of syzygies on the Jacobian ideal J_F of F. In particular, $D(\mathcal{A})$ splits as $D_0(\mathcal{A}) \oplus E$ and is free iff $D_0(\mathcal{A})$ is free. By the Hilbert–Burch theorem, this follows iff J_F is Cohen–Macaulay. In the study of singularities, Saito [79] first formulated a local version of this: $D(\mathcal{A})$ is free if there exist derivations $\theta_1, \ldots, \theta_n$ such that the determinant of the matrix of coefficients is a nonzero multiple of F.

Example 4. In Example 1, if instead of projecting to $\mathbb{C}^3$ we keep the original defining equations, then for $k \in \{1, 2, 3, 4\}$,

$$\theta_k = \sum_{i=1}^{4} x_i^{k-1} \partial/\partial x_i$$

satisfies

$$\theta_k(x_i - x_j) = x_i^{k-1} - x_j^{k-1} \in \langle x_i - x_j \rangle,$$

and by Saito's criterion the arrangement is free. Notice that there is a derivation of degree zero, reflecting the fact that the arrangement is not essential. Taking the three-dimensional projection in $W^{\perp}$ yields an essential arrangement which is still free, with generators in degrees $1, 2$, and 3. These degrees also appear in the Poincaré polynomial

$$\pi(A_3, t) = 1 + 6t + 11t^2 + 6t^3 = (1 + t)(1 + 2t)(1 + 3t),$$

which is explained by Theorem 5.

Theorem 5 (Terao [95]). *If* $D(\mathcal{A}) \simeq \bigoplus_{i=1}^{n} S(-a_i)$, *then*

$$\pi(\mathcal{A}, t) = \prod (1 + a_i t).$$

Factorization of $\pi(\mathcal{A}, t)$ does not imply freeness: if $\mathcal{A}$ consists of five lines in $\mathbb{P}^2$ through a point p and two lines through a point $q \neq p$, such that all intersections save at p are normal crossing, then $\pi(\mathcal{A}, t) = (1 + t)(1 + 3t)^2$, but $\mathcal{A}$ is not free. In fact, most arrangements are not free; for example, adding any line to Example 1 results in an arrangement which is not free, *unless* the line connects two points with $\mu = 1$. For another example, a generic arrangement with more than n hyperplanes will have a Poincaré polynomial which does not factor, so cannot be free. On the other hand, the arrangement consisting of the $d = n$ coordinate hyperplanes has $L(\mathcal{A})$ isomorphic to the Boolean lattice and is easily checked to be free, with all generators of $D(\mathcal{A})$ of degree one. The two most interesting classes which are known to be free are the class of reflection arrangements (the hyperplanes are fixed by a finite reflection group) [94] and arrangements with supersolvable intersection lattice ($L(\mathcal{A})$ has a maximal chain of modular elements) [96]. The following question of Terao has been open for over twenty years:

Conjecture 6 (Orlik–Terao [66]). Freeness of $D(\mathcal{A})$ depends only on $L(\mathcal{A})$.

As shown by Ziegler in [111], the hypothesis that $char(\mathbb{K}) = 0$ is necessary for this: Ziegler gives a pair of examples $\mathcal{A}_1$, $\mathcal{A}_2$ with $L(\mathcal{A}_1) \simeq L(\mathcal{A}_2)$, where $\mathcal{A}_1$ is free and $\mathcal{A}_2$ is not. However, the arrangements are defined over fields of different positive characteristics. We close by discussing what is known about the conjecture as well as giving some indications of related results.

1. $L(\mathcal{A})$ *does not determine the graded Betti numbers of* $D(\mathcal{A})$. An example of Ziegler [109] shows that the graded Betti numbers of $D(\mathcal{A})$ depend on nonlinear geometry. Consider two arrangements, each consisting of nine lines in $\mathbb{P}^2$, with six lines bounding an symmetric hexagon and three lines connecting opposite vertices. This means there are six triple points and eighteen double points. The behavior of $D(\mathcal{A})$ depends on if the triple points lie on a smooth conic or not.

2. *Freeness is an open condition.* A result of Yuzvinsky [104] shows that freeness is an open condition in an appropriate parameter space. For a fixed Poincaré polynomial $\pi(\mathcal{A}, t) = (1 + t)(1 + at)(1 + (d - 1 - a)t)$, only a finite number of intersection lattices can arise, and the largest number of combinatorial types for $L(\mathcal{A})$ occurs when $\pi(\mathcal{A}, t) = (1 + t)(1 + \lfloor \frac{d-1}{2} \rfloor t)(1 + \lceil \frac{d-1}{2} \rceil t))$. For d odd, $D_0(\mathcal{A})$ is semistable, and [80] shows that if $\mathcal{A}$ is not free, then $D_0(\mathcal{A})$ has generators in degree $> \frac{d-1}{2}$ and degree $< \frac{d-1}{2}$.
3. $\pi(A, t)$ *can be expressed in terms of Hilbert series.* In [91], Solomon–Terao show that there is a version of Theorem 5 which applies to all arrangements, in the sense that $\pi(\mathcal{A}, t)$ may be written in terms of the Hilbert series of modules $D^p(\mathcal{A})$, which are higher-order analogs of $D(\mathcal{A})$. When $\mathcal{A}$ is free, these modules are simply $\Lambda^p(D(\mathcal{A}))$, but this is not the case in general. In [64], Mustaţă–Schenck give a generalization of Theorem 5 when $\mathcal{A}$ is locally free; further generalizations appear in Denham–Schulze [27]. Very little is known about the modules $D^p(\mathcal{A})$.
4. *An inductive method for freeness.* A key method of proving that an arrangement is free involves an inductive operation known as deletion–restriction: fix $H \in \mathcal{A}$, and define $A' = \mathcal{A} \setminus H,\ A'' = \mathcal{A}|_H$. Then there is a left exact sequence

$$0 \longrightarrow D(\mathcal{A}')(-1) \xrightarrow{\cdot H} D(\mathcal{A}) \longrightarrow D(\mathcal{A}|_H)$$

which is typically not right exact. However, Terao [93] shows that any two of the following imply the third:

(1) $D(\mathcal{A}) \simeq \oplus_{i=1}^{n} S(-b_i)$.
(2) $D(\mathcal{A}') \simeq S(-b_n + 1) \oplus_{i=1}^{n-1} S(-b_i)$.
(3) $D(\mathcal{A}'') \simeq \oplus_{i=1}^{n-1} S/H(-b_i)$.

5. *The generalization to arrangements of curves is false.* The deletion–restriction method can be generalized to smooth plane curve arrangements [84], if the singularities are quasihomogeneous. Terao's conjecture turns out to be special to the case of hyperplane arrangements: [83] exhibits a pair of arrangements of lines and conics, with isomorphic intersection posets and only ordinary singularities, but where one arrangement is free and the other is not. Can [84] be generalized to hypersurface arrangements?
6. *Multiarrangements*. One promising new technique for attacking Conjecture 6 involves multiarrangements:

Definition 7. A multiarrangement $(\mathcal{A}, \mathbf{m})$ consists of an arrangement $\mathcal{A}$, along with a multiplicity $m_i \in \mathbb{N}$ for each $H \in \mathcal{A}$:

$$D(\mathcal{A}, \mathbf{m}) = \{\theta \mid \theta(l_i) \in \langle l_i^{m_i} \rangle\ \forall\ \mathbf{V}(l_i) \in \mathcal{A}\}.$$

In [102], Yoshinaga shows that a line arrangement $\mathcal{A} \subseteq \mathbb{P}^2$ is free if $\pi(\mathcal{A}, t) = (1+t)(1+at)(1+bt)$ and $D(\mathcal{A}|_H, \mathbf{m}) \simeq S/H(-a) \oplus S/H(-b)$. In [110] Ziegler showed these conditions are also necessary. Generalizations to higher dimension appear in [85,101], and multiarrangement versions of Theorem 5 appear in [1,2]. In general, even less is known about $D(\mathcal{A}, \mathbf{m})$ than about $D(\mathcal{A})$.

3.2 Resonance Varieties and Orlik–Solomon Algebra

In [65], Orlik–Solomon show that the cohomology ring of the complement $M = \mathbb{C}^n \setminus \bigcup_{i=1}^d H_i$ has presentation $H^*(M, \mathbb{Z}) = \bigwedge(\mathbb{Z}^d)/I$, with generators $e_1, \ldots, e_d$ in degree 1 and

$$I = \langle \sum_q (-1)^{q-1} e_{i_1} \cdots \widehat{e_{i_q}} \cdots e_{i_r} \mid \operatorname{codim} H_{i_1} \cap \cdots \cap H_{i_r} < r \rangle.$$

Let A denote the Orlik–Solomon algebra $H^*(M, \mathbb{Z})$. While $L(\mathcal{A})$ determines A, the fundamental group $\pi_1(M)$ is not determined by $L(\mathcal{A})$. A pair of conjectures of Suciu [92] connect the LCS and Chen ranks of $\pi_1(M)$ to commutative algebra and subtle combinatorics of $L(\mathcal{A})$. First, we need some definitions. Since A is a quotient of an exterior algebra, multiplication by an element $a \in A^1$ gives a degree one differential on A, yielding a cochain complex (A, a):

$$(A,a): \quad 0 \longrightarrow A^0 \xrightarrow{\wedge a} A^1 \xrightarrow{\wedge a} A^2 \xrightarrow{\wedge a} \cdots \xrightarrow{\wedge a} A^\ell \longrightarrow 0 \,. \tag{1}$$

The complex (A, a) was first studied by Aomoto in [4] and appears subsequently in work of Esnault–Schechtman–Viehweg [39]. In [105], Yuzvinsky shows that (A, a) is exact as long as $\sum_{i=1}^d a_i \neq 0$.

Definition 8. The resonance variety $R^k(\mathcal{A})$ consists of points $a = \sum_{i=1}^n a_i e_i \leftrightarrow (a_1 : \cdots : a_n)$ in $\mathbb{P}(A^1) \cong \mathbb{P}^{d-1}$ for which $H^k(A, a) \neq 0$.

In [42], Falk introduced the concept of a *neighborly partition*: a partition Π of $\mathcal{A}$ is neighborly if, for any rank-two flat $Y \in L_2(\mathcal{A})$ and any block π of Π, $\mu(Y) \leq |Y \cap \pi| \implies Y \subseteq \pi$. Falk proved that all components of $R^1(A)$ arise from such partitions of subarrangements of $\mathcal{A}$ and conjectured that $R^1(\mathcal{A})$ was a subspace arrangement. This was proved by Cohen–Suciu [20] and Libgober–Yuzvinsky [57].

Example 9. For Example 1, since A_3 is an arrangement in $\mathbb{C}^3$, all four tuples are dependent and $A = \Lambda(\mathbb{Z}^6)/I$, where

$$\begin{aligned} I &= \langle \partial(e_1e_2e_3), \partial(e_1e_5e_6), \partial(e_2e_4e_6), \partial(e_3e_4e_5), \partial(e_ie_je_ke_l) \rangle \\ &= \langle e_1 \wedge e_2 - e_1 \wedge e_3 + e_2 \wedge e_3, \cdots \rangle. \end{aligned}$$

In this example, I is actually generated by quadrics. Each triple point determines a subarrangement and neighborly partition consisting of a single block. The partition $|14|25|36|$ is also neighborly, and a computation [92] shows that $R^1(A_3)$ is the disjoint union of five lines, corresponding to the five neighborly partitions.

Recall that the LCS ranks $\phi_i(G)$ of a group G are given by

$$\phi_i(G) = \dim G_i/G_{i+1} \otimes \mathbb{Q}, \text{ where } G = G_1 \text{ and } G_{i+1} = [G, G_i],$$

and the Chen ranks $\theta_k(G)$ are the LCS ranks of $G/[[G,G],[G,G]]$. For the fundamental group $\pi_1(M)$, using results of Brieskorn [14], Kohno [55], Shelton–Yuzvinsky [86], and Priddy [75], Peeva shows in [71] that

$$\prod_{k=1}^{\infty} \frac{1}{(1-t^k)^{\phi_k(\pi_1(M))}} = \sum_{i=0}^{\infty} \dim_{\mathbb{Q}} Tor_i^A(\mathbb{Q},\mathbb{Q})_i t^i.$$

In [82], Schenck–Suciu use the Bernstein–Gelfand–Gelfand correspondence to show the $\theta_k(\pi_1(M))$ are also expressed in terms of a Tor:

$$\sum_{i \geq 2} \theta_i(\pi_1(M)) t^i = \sum_{i \geq 2} \dim_{\mathbb{Q}} Tor_i^E(A,\mathbb{Q})_{i+1} t^i. \tag{2}$$

The second major commutative algebra conjecture in arrangements is

Conjecture 10 (Suciu [92]). Let h_r be the number of components of $R^1(\mathcal{A})$ of dimension r. Then for $k \gg 0$,

$$\theta_k(\pi_1(M)) = (k-1) \sum_{r \geq 1} h_r \binom{r+k-1}{k}.$$

Furthermore, if $\phi_4(\pi_1(M)) = \theta_4(\pi_1(M))$, then

$$\prod_{k \geq 1} (1-t^k)^{\phi_k(\pi_1(M))} = \prod_{r \geq 1} (1-(r+1)t)^{h_r}.$$

The connection to commutative algebra comes by considering the a_i in (1) as variables in a symmetric algebra R and (A,a) as a complex of R-modules. Using Fox calculus and a linearization technique, Cohen–Suciu [20] show that if B is the R-module corresponding to the cokernel of the transpose of the map

$$A^1 \xrightarrow{\wedge a} A^2,$$

then $R^1(A) = \mathbf{V}(\mathrm{ann}(B))$, and in [68], Papadima–Suciu prove that the Hilbert series of B is $\sum_{i \geq 2} \theta_i t^i$.

The next step comes from a result of Eisenbud–Popescu–Yuzvinsky [37], showing that the resolution of $\mathrm{Hom}_E(A, E)$ is linear, and as a consequence the complex (A, a) of R-modules is exact, save at the last position. Let $F(\mathcal{A})$ denote the cokernel of the last map in (A, a). Since (A, a) gives a free resolution of $F(\mathcal{A})$, there is an exact sequence

$$0 \longrightarrow Ext_R^{n-1}(F(A), R) \longrightarrow B \longrightarrow \mathfrak{m} \longrightarrow 0,$$

and the Bernstein–Gelfand–Gelfand correspondence connects B to the $\mathrm{Tor}_i^A(A, \mathbb{Q})_{i+1}$ in (2); [82] proves that for $k \gg 0$,

$$\theta_k \geq (k-1) \sum_{r \geq 1} h_r \binom{r+k-1}{k}.$$

Proving the remaining direction of Conjecture 10 consists of determining the Hilbert polynomial of $Ext_R^{n-1}(F(A), R)$. Perhaps the most striking of the results above are those of [37], which are proved using upper semicontinuity of Betti numbers and Alexander duality. These methods have broad applicability, and we close with a sketch of the proof.

Theorem 11 (Eisenbud et al. [37]). *The free resolution of A^* over E is linear:*

$$\cdots \longrightarrow E^{r_2}(n-2-d) \longrightarrow E^{r_1}(n-1-d) \longrightarrow E^{r_0}(n-d) \longrightarrow A^* \to 0,$$

where $(1-t)^n \sum_{i=0}^{\infty} r_i t^i = (-t)^d \pi(\mathcal{A}, \frac{-1}{t})$.

First, since the $E-$ module structure on A^* comes from the identification $A^* = \mathrm{Hom}_{\mathbb{K}}(A, \mathbb{K})$ and by [35], Theorem 21.1,

$$\mathrm{Hom}_{\mathbb{K}}(A, \mathbb{K}) \cong \mathrm{Hom}_E(A, E)$$

$\mathrm{Hom}_E(E/\operatorname{in}(I), E)$ is a flat deformation of A^*. Since graded Betti numbers are upper semicontinuous under flat deformation (Theorem 3.6 of [52]), it suffices to prove the result for $\mathrm{Hom}_E(E/\operatorname{in}(I), E)$. Giving a homomorphism of E-modules $E/\operatorname{in}(I) \to E$ is equivalent to giving an element $a \in E$ such that $a \cdot \operatorname{in}(I) = 0$; hence

$$\mathrm{Hom}_E(E/\operatorname{in}(I), E) \simeq \operatorname{ann}_E(\operatorname{in}(I)).$$

Let $J = \operatorname{ann}_E(\operatorname{in}(I))$. Since we are working in E, $\operatorname{in}(I)$ will be square-free and hence define an exterior Stanley–Reisner ideal.

In [32], Eagon–Reiner show that a square-free monomial ideal I_Δ has a linear resolution iff the Stanley–Reisner ring of the Alexander dual Δ^* is Cohen–Macaulay, and in [5], Aramova–Avramov–Herzog prove an exterior version of this result. Recall that the Alexander dual Δ^* of a simplicial complex Δ consists of the complements of the nonfaces of Δ. To prove Theorem 11, [37] shows that $\operatorname{in}(I)$ and

J are Alexander dual and applies [5]. Fix an ordering $\prec$ on the hyperplanes of $\mathcal{A}$. Björner–Ziegler [12] show that in lexicographic order with $e_i < e_j$ if $H_i \prec H_j$, $\mathrm{in}(I)$ consists of broken circuits, which are dependent sets with smallest element removed.

Example 12. In Example 1, the circuits are

$$\{\{1,2,3\},\{1,5,6\},\{2,4,6\},\{3,4,5\},\{i,j,k,l\}\}.$$

Hence,

$$\mathrm{in}(I) = \langle e_2 \wedge e_3, e_5 \wedge e_6, e_4 \wedge e_6, e_4 \wedge e_5 \rangle.$$

Considering $\mathrm{in}(I)$ as the Stanley–Reisner ideal of Δ, the maximal faces of Δ^* are $\{1,4,5,6\},\{1,2,3,4\},\{1,2,3,5\}$, and $\{1,2,3,6\}\}$; hence

$$I_{\Delta^*} = \langle e_2 \wedge e_5 \wedge e_6, e_2 \wedge e_4 \wedge e_6, e_2 \wedge e_4 \wedge e_5, e_3 \wedge e_5 \wedge e_6, e_3 \wedge e_4 \wedge e_6, e_3 \wedge e_4 \wedge e_5 \rangle$$

Lemma 13. *For a simplicial complex Δ, $I_\Delta = \mathrm{ann}(I_{\Delta^*}) \subseteq E$.*

Proof. It is easy to see [61] that the Stanley–Reisner ideal of the Alexander dual I_{Δ^*} is obtained by monomializing the primary decomposition of I_Δ. Since each monomial $x_{i_1} \cdots x_{i_k} \in I_\Delta$ corresponds to a prime component $\langle x_{i_1}, \ldots, x_{i_k} \rangle$ in the primary decomposition of I_{Δ^*}, every monomial in I_Δ has at least one variable from every monomial in I_{Δ^*}. In the exterior algebra, this implies that I_Δ annihilates I_{Δ^*}, and a check shows that any monomial in I_Δ has this form. □

Proof. (Theorem 11) By the lemma,

$$\mathrm{in}(I) = I_\Delta = \mathrm{ann}_E\, I_{\Delta^*}.$$

In [77], Provan and Billera prove that the simplicial complex associated to the broken circuit complex of a matroid is shellable, hence Cohen–Macaulay. The result of [5] now implies that $I_{\Delta^*} = J$ has linear free resolution. □

While Theorem 11 provides the ranks in the free resolution, it would be interesting to have a complete description of the differentials. The first two are determined in [28].

Corollary 14 (Eisenbud et al. [37]). *(A, a) is exact except at the last step.*

Proof. Apply the functor $\mathrm{Hom}_E(\bullet, E)$ to the free resolution of A^*. Since E is Gorenstein, injective resolutions over E are duals of free resolutions, so Theorem 11 yields an injective resolution of A. Let $F(\mathcal{A})$ be the S-module which maps to this injective resolution. Then the Bernstein–Gelfand–Gelfand correspondence and Theorem 3.7 of [36] imply that (A, a) is a free resolution of $F(\mathcal{A})$. □

In [26], Denham–Schenck show that the higher resonance varieties have a similar description

$$R^k(A) = \bigcup_{k' \leq k} \mathbf{V}(\operatorname{ann} \operatorname{Ext}^{n-k'}(F(A), S)).$$

Does there exist a combinatorial description of components of $R^k(A)$, similar to the description of $R^1(A)$ via neighborly partitions?

3.3 The Orlik–Terao Algebra, Nets and Multinets

The final problem on arrangements comes from a symmetric analog of the Orlik–Solomon algebra and ties to the classical theory of nets [74].

Definition 15. Let $\mathcal{A} = \cup_{i=1}^d \mathbf{V}(l_i) \subseteq \mathbb{P}^n$ and $R = \mathbb{K}[y_1, \ldots, y_d]$. For each linear dependency $\Lambda = \sum_{j=1}^k c_{i_j} l_{i_j} = 0$, define

$$f_\Lambda = \sum_{j=1}^k c_{i_j} (y_{i_1} \cdots \hat{y}_{i_j} \cdots y_{i_k}),$$

and let I be the ideal generated by the f_Λ. The Orlik–Terao algebra $C(\mathcal{A}) = R/I$, and the Artinian Orlik–Terao algebra $AOT(\mathcal{A}) = C(\mathcal{A})/\langle y_1^2, \ldots, y_d^2 \rangle$.

Orlik and Terao introduced the Artinian version in [67] to answer a question posed by Aomoto and showed the Hilbert series satisfies

$$\operatorname{HS}(AOT(\mathcal{A}), t) = \pi(\mathcal{A}, t),$$

while in [97], Terao showed that the non-Artinian version satisfies

$$\operatorname{HS}(C(\mathcal{A}), t) = \pi\Big(\mathcal{A}, \frac{t}{1-t}\Big).$$

Let $F = l_1 \cdot l_2 \cdots l_d$, and consider the rational map

$$\mathbb{P}^{n-1} \xrightarrow{\left[\frac{F}{l_1} : \frac{F}{l_2} : \cdots : \frac{F}{l_d}\right]} \mathbb{P}^{d-1} \tag{3}$$

Then [83] shows that $C(\mathcal{A})$ is the coordinate ring of the closure of the image of this map. For $n = 3$, let Y be the blowup of $\mathbb{P}^2$ at $\operatorname{sing}(\mathcal{A})$, E_i the exceptional curves, and consider the divisor

$$D = (d-1)E_0 - \sum_{p_i \in \operatorname{sing}(\mathcal{A})} \mu(p_i) E_i$$

on Y. In this case, the map (3) is an isomorphism on $\mathbb{P}^2 \setminus \mathcal{A}$, blows up the intersection points, and blows down the lines of $\mathcal{A}$ [81].

Example 16. For Example 1, the resulting surface in $\mathbb{P}^5$ is the intersection of the Segre threefold $\Sigma_{2,1}$ with a quadric hypersurface, and the graded Betti numbers are

Total	1	4	5	2
0	1	–	–	–
1	–	4	2	–
2	–	–	3	2

The surface has degree

$$D^2 = (d-1)^2 - \sum_{p_i \in \operatorname{sing}(\mathcal{A})} \mu(p_i)^2 = 6$$

and has six singular points, corresponding to the contracted lines of $\mathcal{A}$.

Little is known about $C(\mathcal{A})$: Proudfoot–Speyer [76] show that it is Cohen–Macaulay and it follows that $C(\mathcal{A})$ is at most n-regular. The algebra $C(\mathcal{A})$ detects subtle geometry invisible to the Orlik–Solomon algebra; for example, [83] shows that $C(\mathcal{A})$ can distinguish between the two Ziegler arrangements [109] discussed earlier.

A first question is to determine the minimal free resolution of $C(\mathcal{A})$, and [81] makes some progress by connecting the quadratic strand to $R^1(A)$ via factorizations $D = F + G$ with $h^0(F) = 2$, so F defines a pencil. This ties in to the work of Libgober–Yuzvinsky [57], Pereira–Yuzvinsky [73], and Yuzvinsky [106–108] relating the classical geometry of pencils and nets to $R^1(\mathcal{A})$.

Definition 17. Let $3 \leq k \in \mathbb{Z}$. A k-*net* in $\mathbb{P}^2$ is a pair $(\mathcal{A}, Z)$ where $\mathcal{A}$ is a finite set of distinct lines partitioned into k subsets $\mathcal{A} = \bigcup_{i=1}^{k} \mathcal{A}_i$ and Z is a finite set of points, such that:

(1) For every $i \neq j$ and every $L \in \mathcal{A}_i,\ L' \in \mathcal{A}_j, L \cap L' \in Z$
(2) For every $p \in Z$ and every $i \in \{1, \ldots, k\}$, $\exists$ a unique $L \in A_i$ containing Z

For a k-net, $|A_i| = |L \cap Z|$ for any block A_i and line $L \in \mathcal{A}$; $m = |A_i|$ is the order of the net. A k-net of order m is called a (k, m)-net.

Example 18. In Example 1, the partition $|14|25|36|$ defines a $(3, 2)$ net. The divisor D decomposes as

$$F = 2E_0 - \sum_{\mu(p)=2} E_p, \text{ and } G = 3E_0 - \sum_{p \in \operatorname{sing}(\mathcal{A})} E_p.$$

Since the four triple points are not collinear, the dimension of the space of conics through the four reduced points is two. It is easy to check that $h^0(G) = 3$, and it follows from results of Eisenbud [33] that the image of Y lies on the scroll $\Sigma_{2,1}$.

In [106] Yuzvinsky shows that a net must have $k \in \{3, 4, 5\}$ and that any finite subgroup of a two-dimensional torus can be realized (the blocks can be identified as the multiplication table of the group) as a 3-net. In [107], Yuzvinsky shows there are no 5-nets. The only known example of a 4-net is the Hessian arrangement (Example 6.30 of [66]) whose twelve lines are the four degenerate cubics

$$x^3 + y^3 + x^3 - 3axyz = 0, \text{ with } a^3 = 1 \text{ or } a = \infty.$$

Conjecture 19 (Yuzvinsky [108]). The only 4-net is the Hessian.

It is known that there are no $(4, m)$-nets with $m \in \{4, 5, 6\}$. In [43], Falk and Yuzvinsky introduce the concept of a multinet, where lines may occur with multiplicity. Multinets are much more complicated than nets, and we refer to [43] or [108] for many open questions.

Acknowledgements The authors would like to thank Tom Braden, Stephen Griffeth, Steven Sam, Julianna Tymoczko, Alexander Woo, and Adam Van Tuyl for helpful conversations. Schenck supported by NSF 1068754, NSA H98230-11-1-0170.

References

1. Abe, T., Terao, H., Wakefield, M.: The characteristic polynomial of a multiarrangement. Adv. Math. **215**, 825–838 (2007)
2. Abe, T., Terao, H., Wakefield, M.: The Euler multiplicity and addition-deletion theorems for multiarrangements. J. Lond. Math. Soc. **77**, 335–348 (2008)
3. Abo, H., Kley, H., Peterson, C.: Petersen plane arrangements and a surface with multiple 7-secants. Adv. Geom. **9**, 349–369 (2009)
4. Aomoto, K.: Un théorème du type de Matsushima–Murakami concernant l'intégrale des fonctions multiformes. J. Math. Pures Appl. **52**, 1–11 (1973)
5. Aramova, L., Avramov, L., Herzog, J.: Resolutions of monomial ideals and cohomology over exterior algebras. Trans. Amer. Math. Soc. **352**, 579–594 (2000)
6. Ballico, E.: On the gonality of graph curves. Manuscripta Math. **129**, 169–180 (2009)
7. Ballico, E.: Graph curves with gonality one. Rend. Semin. Mat. Univ. Politec. Torino **68**, 17–28 (2010)
8. Bayer, D., Eisenbud, D.: Graph curves. Adv. Math. **86**, 1–40 (1991)
9. Bayer, D., Charalambous, H., Popescu, S.: Extremal Betti numbers and applications to monomial ideals. J. Algebra **221**, 497–512 (1999)
10. Bifet, E., De Concini, C., Procesi, C.: Cohomology of regular embeddings. Adv. Math. **82**, 1–34 (1990)
11. Björner, A.: Subspace arrangements. In: Progress in Mathematics, vol. 119, pp. 321–370. Birkhäuser, Basel (1994)
12. Björner, A., Ziegler, G.: Broken circuit complexes: factorizations and generalizations. J. Combin. Theory Ser. B **51**, 96–126 (1991)
13. Björner, A., Peeva, I., Sidman, J.: Subspace arrangements defined by products of linear forms. J. London Math. Soc. **71**, 273–288 (2005)
14. Brieskorn, E.: Sur les groupes de tresses [d'après V. I. Arnol' d]. Séminaire Bourbaki, 24ème année, pp. 21–44. Springer, Berlin (1973)
15. Buchstaber, V., Panov, T.: Torus actions and their applications in topology and combinatorics. In: University Lecture Series, vol. 24. American Mathematical Society, Providence (2002)
16. Burnham, G., Rosen, Z., Sidman, J., Vermeire, P.: Line arrangements modeling curves of high degree: equations, syzygies and secants. Preprint (2012)

17. Carlini, E., Catalisano, M., Geramita, A.: Subspace arrangements, configurations of linear spaces and the quadrics containing them. J. Algebra **362**, 70–83 (2012)
18. Ciliberto, C., Miranda, R., Graph curves, colorings, and matroids. Zero-Dimensional Schemes, Ravello, 1992. de Gruyter, Berlin, pp. 89–111 (1994)
19. Ciliberto, C., Harris, J., Miranda, R.: On the surjectivity of the Wahl map. Duke Math. J. **57**, 829–858 (1988)
20. Cohen, D., Suciu, A.: Characteristic varieties of arrangements. Math. Proc. Cambridge Philos. Soc. **127**, 33–53 (1999)
21. Cohen, D., Farber, M.: Topological complexity of collision-free motion planning on surfaces. Compos. Math. **147**, 649–660 (2011)
22. Conca, A., Herzog, J.: Castelnuovo-Mumford regularity of products of ideals. Collect. Math. **54**, 137–152 (2003)
23. Cox, D.A., Little, J.B., Schenck, H.K.: Toric varieties. In: Graduate Studies in Mathematics. American Mathematical Society, Providince (2011)
24. Danilov, V.: The geometry of toric varieties. Uspekhi Mat. Nauk **33**, 85–134 (1978)
25. de Loera, J.: Gröbner bases and graph colorings. Beiträge Algebra Geom. **36**, 89–96 (1995)
26. Denham, G., Schenck, H.: The double Ext spectral sequence and rank varieties. Preprint (2012)
27. Denham, G., Schulze, M.: Complexes, duality and Chern classes of logarithmic forms along hyperplane arrangements. Adv. Stud. Pure Math. **62**, 27–58 (2012)
28. Denham, G., Yuzvinsky, S.: Annihilators of Orlik-Solomon relations. Adv. Appl. Math. **28**, 231–249 (2002)
29. Derksen, H.: Computation of invariants for reductive groups. Adv. Math. **141**, 366–384 (1999)
30. Derksen, H.: Hilbert series of subspace arrangements. J. Pure Appl. Algebra **209**, 91–98 (2007)
31. Derksen, H., Sidman, J.: A sharp bound for the Castelnuovo–Mumford regularity of subspace arrangements. Adv. Math. **172**, 151–157 (2002)
32. Eagon, J., Reiner, V., Resolutions of Stanley–Reisner rings and Alexander duality. J. Pure Appl. Algebra **130**, 265–275 (1998)
33. Eisenbud, D.: Linear sections of determinantal varieties. Amer. J. Math. **110**, 541–575 (1988)
34. Eisenbud, D.: Green's conjecture: an orientation for algebraists. In: Research Notes in Mathematics, vol. 2, pp. 51–78. Jones and Bartlett, Boston (1992)
35. Eisenbud, D.: Commutative algebra. In: Graduate Texts in Mathematics, vol. 150. Springer, NY (1995)
36. Eisenbud, D., Fløystad, G., Schreyer, F.-O.: Sheaf cohomology and free resolutions over exterior algebras. Trans. Amer. Math. Soc. **355**, 4397–4426 (2003)
37. Eisenbud, D., Popescu, S., Yuzvinsky, S.: Hyperplane arrangement cohomology and monomials in the exterior algebra. Trans. Amer. Math. Soc. **355**, 4365–4383 (2003)
38. Eisenbud, D., Green, M., Hulek, K., Popescu, S.: Small schemes and varieties of minimal degree. Amer. J. Math. **128**, 1363–1389 (2006)
39. Esnault, H., Schechtman, V., Viehweg, E.: Cohomology of local systems on the complement of hyperplanes. Invent. Math. **109**, 557–561 (1992)
40. Etingof, P., Stoica, E. (appendix S. Griffeth): Unitary representations of rational Cherednik algebras. Represent. Theory **13**, 349–370 (2009)
41. Faridi, S.: The facet ideal of a simplicial complex. Manuscripta Math. **109**, 159–174 (2002)
42. Falk, M.: Arrangements and cohomology. Ann. Comb. **1**, 135–157 (1997)
43. Falk, M., Yuzvinsky, S.: Multinets, resonance varieties, and pencils of plane curves. Compos. Math. **143**, 1069–1088 (2007)
44. Farber, M., Tabachnikov, S., Yuzvinsky, S.: Topological robotics: motion planning in projective spaces. Int. Math. Res. Not., **34**, 1853–1870 (2003)
45. Fröberg, R. : On Stanley–Reisner rings. Banach Center Publ. **26**, 57–70 (1988)
46. Geramita, A., Weibel, C.: On the Cohen–Macaulay and Buchsbaum property for unions of planes in affine space. J. Algebra **92**, 413–445 (1985)

47. Goresky, M., MacPherson, R.: On the spectrum of the equivariant cohomology ring. Canad. J. Math. **62**, 262–283 (2010)
48. Green, M.: Koszul cohomology and the geometry of projective varieties. J. Differential Geom. **19**, 125–171 (1984)
49. Guida, M., Orecchia, F.: Algebraic properties of grids of projective lines. J. Pure Appl. Algebra **208**, 603–615 (2007)
50. Haiman, M., Woo, A.: Garnir ideals, Springer fibers, and Ellingsrud-Stromme cells on the Hilbert scheme of points. Preprint (2012)
51. Hartshorne, R., Hirschowitz, A.: Droites en position générale dans l'espace projectif. In: Lecture Notes in Mathematics, vol. 961, pp. 169–188. Springer, Berlin (1982)
52. Herzog, J.: Finite free resolutions. In: NATO Science Series III Computer and Systems Sciences, vol. 196, pp. 118–144. IOS, Amsterdam (2005)
53. Hochster, M.: Cohen–Macaulay rings, combinatorics, and simplicial complexes. In: Ring Theory, II. pp. 171–223. Dekker, New York (1977)
54. Jurkiewicz, J.: Chow ring of projective nonsingular torus embedding. Colloq. Math. **43**, 261–270 (1980)
55. Kohno, T.: On the holonomy Lie algebra and the nilpotent completion of the fundamental group of the complement of hypersurfaces. Nagoya Math. J. **92**, 21–37 (1983)
56. Li, S.-Y.R., Li, W.C.W.: Independence numbers of graphs and generators of ideals. Combinatorica **1**, 55–61 (1981)
57. Libgober, A., Yuzvinsky, S.: Cohomology of the Orlik–Solomon algebras and local systems. Compositio Math. **121**, 337–361 (2000)
58. Lovász, L.: Stable sets and polynomials. Discrete Math. **124**, 137–153 (1994)
59. Ma, Y., Yang, A., Derksen, H., Fossum, R.: Estimation of subspace arrangements with applications in modeling and segmenting mixed data. SIAM Rev. **50**, 413–458 (2008)
60. MacPherson, R.: Equivariant invariants and linear geometry. In: IAS/Park City Mathematical Series, vol. 13, pp. 317–388. American Mathematical Society, Providence (2007)
61. Miller, E., Sturmfels, B.: Combinatorial commutative algebra. Graduate Texts in Mathematics, vol. 227. Springer, New York (2005)
62. Miranda, R.: Graph Curves and Curves on $K3$ Surfaces. World Scientific Publishing, Teaneck (1989)
63. Miranda, R.: The Gaussian map for certain planar graph curves. In: Contemporary Mathematics, vol. 116, pp. 115–124. American Mathematical Society, Providence (1991)
64. Mustaţă, M., Schenck, H.: The module of logarithmic p-forms of a locally free arrangement. J. Algebra **241**, 699–719 (2001)
65. Orlik, P., Solomon, L.: Combinatorics and topology of complements of hyperplanes. Invent. Math. **56**, 167–189 (1980)
66. Orlik, P., Terao, H.: Arrangements of Hyperplanes. Springer, Berlin (1992)
67. Orlik, P., Terao, H.: Commutative algebras for arrangements. Nagoya Math. J. **134**, 65–73 (1994)
68. Papadima, S., Suciu, A.: Chen Lie algebras. Int. Math. Res. Not. **21**, 1057–1086 (2004)
69. Park, S.W.: A combinatorial study of graph curves. Ph.D. Thesis, Brandeis University, Waltham, MA (1990)
70. Peel, M.: Specht modules and symmetric groups. J. Algebra **36**, 88–97 (1975)
71. Peeva, I.: Hyperplane arrangements and linear strands in resolutions. Trans. Amer. Math. Soc. **355**, 609–618 (2003)
72. Peeva, I., Reiner, V., Welker, V.: Cohomology of real diagonal subspace arrangements via resolutions. Compositio Math. **117**, 99–115 (1999)
73. Pereira, J., Yuzvinsky, S.: Completely reducible hypersurfaces in a pencil. Adv. Math. **219**, 672–688 (2008)
74. Pereira, J., Pirio, L.: An Invitation to Web Geometry. IMPA Mathematical Publications, Rio de Janeiro (2009)
75. Priddy, S.: Koszul resolutions. Trans. Amer. Math. Soc. **152**, 39–60 (1970)
76. Proudfoot, N., Speyer, D.: A broken circuit ring. Beiträge Algebra Geom. **47**, 161–166 (2006)

77. Provan, J.S., Billera, L.: Decompositions of simplicial complexes related to diameters of convex polyhedra. Math. Oper. Res. **5**, 576–594 (1980)
78. Reisner, G.: Cohen-Macaulay quotients of polynomial rings. Adv. Math. **21**, 30–49 (1976)
79. Saito, K.: Theory of logarithmic differential forms and logarithmic vector fields. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **27**, 265–291 (1980)
80. Schenck, H.: Elementary modifications and line configurations in $\mathbb{P}^2$. Comment. Math. Helv. **78**, 447–462 (2003)
81. Schenck, H.: Resonance varieties via blowups of $\mathbb{P}^2$ and scrolls. Int. Math. Res. Not. **20**, 4756–4778 (2011)
82. Schenck, H., Suciu, A.: Resonance, linear syzygies, Chen groups, and the Bernstein–Gelfand–Gelfand correspondence. Trans. Amer. Math. Soc. **358**, 2269–2289 (2006)
83. Schenck, H., Tohǎneanu, S.: Freeness of conic-line arrangements in $\mathbb{P}^2$. Comment. Math. Helv. **84**, 235–258 (2009)
84. Schenck, H., Terao, H., Yoshinaga, M.: Logarithmic vector fields for quasihomogeneous curve configurations in $\mathbb{P}^2$. Preprint (2012)
85. Schulze, M.: Freeness and multirestriction of hyperplane arrangements. Compos. Math. **148**, 799–806 (2012)
86. Shelton, B., Yuzvinsky, S.: Koszul algebras from graphs and hyperplane arrangements. J. London Math. Soc. **56**, 477–490 (1997)
87. Sidman, J.: On the Castelnuovo-Mumford regularity of subspace arrangements. Ph.D. Thesis, University of Michigan, Michigan (2002)
88. Sidman, J.: Defining equations of subspace arrangements embedded in reflection arrangements. Int. Math. Res. Not., 713–727 (2004)
89. Sidman, J.: Resolutions and subspace arrangements. In: Lecture Notes in Pure Applied Mathematics, vol. 254, pp. 249–265. Chapman & Hall/CRC, Boca Raton (2007)
90. Sidman, J., Vermeire, P.: Syzygies of the secant variety of a curve. Algebra Number Theory **3**, 445–465 (2009)
91. Solomon, L., Terao, H.: A formula for the characteristic polynomial of an arrangement. Adv. Math. **64**, 305–325 (1987)
92. Suciu, A.: Fundamental groups of line arrangements: enumerative aspects. In: Contemporary Mathematics, vol. 276, pp. 43–79. American Mathematical Society, Providence (2001)
93. Terao, H.: Arrangements of hyperplanes and their freeness. I. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **27**, 293–312 (1980)
94. Terao, H.: Free arrangements of hyperplanes and unitary reflection groups. Proc. Japan Acad. Ser. A Math. Sci. **56**, 389–392 (1980)
95. Terao, H.: Generalized exponents of a free arrangement of hyperplanes and Shepherd-Todd-Brieskorn formula. Invent. Math. **63**, 159–179 (1981)
96. Terao, H.: Modular elements of lattices and topological fibration. Adv. Math. **62**, 135–154 (1986)
97. Terao, H.: Algebras generated by reciprocals of linear forms. J. Algebra **250**, 549–558 (2002)
98. Varchenko, A.: Special functions, KZ type equations, and representation theory. CBMS Regional Conference Series in Mathematics, vol. 98 (2003)
99. Vermeire, P.: Regularity and normality of the secant variety to a projective curve. J. Algebra **319**, 1264–1270 (2008)
100. Villarreal, R.H.: Monomial algebras. In: Monographs and Textbooks in Pure and Applied Mathematics, vol. 238. Marcel Dekker, New York (2001)
101. Yoshinaga, M.: Characterization of a free arrangement and conjecture of Edelman and Reiner. Invent. Math. **157**, 449–454 (2004)
102. Yoshinaga, M.: On the freeness of 3-arrangements. Bull. London Math. Soc. **37**, 126–134 (2005)
103. Yuzvinsky, S.: Cohen-Macaulay seminormalizations of unions of linear subspaces. J. Algebra **132**, 431–445 (1990)
104. Yuzvinsky, S.: The first two obstructions to the freeness of arrangements. Trans. Amer. Math. Soc. **335**, 231–244 (1993)

105. Yuzvinsky, S.: Cohomology of the Brieskorn-Orlik-Solomon algebras. Comm. Algebra **23**, 5339–5354 (1995)
106. Yuzvinsky, S.: Realization of finite abelian groups by nets in $\mathbb{P}^2$. Compos. Math. **140**, 1614–1624 (2004)
107. Yuzvinsky, S.: A new bound on the number of special fibers in a pencil of curves. Proc. Amer. Math. Soc. **137**, 1641–1648 (2009)
108. Yuzvinsky, S.: Resonance varieties of arrangement complements. Adv. Stud. Pure Math. **62**, 553–570 (2012)
109. Ziegler, G.: Combinatorial construction of logarithmic differential forms. Adv. Math. **76**, 116–154 (1989)
110. Ziegler, G.: Multiarrangements of hyperplanes and their freeness. In: Contemporary Mathematics, vol. 90, pp. 345–359. American Mathematical Society, Providence (1989)
111. Ziegler, G.: Matroid representations and free arrangements. Trans. Amer. Math. Soc. **320**, 525–541 (1990)

Cohomological Degrees and Applications

Wolmer V. Vasconcelos

MSC 2010: Primary: 13H10, 13H15. Secondary: 13D02, 13D07.

Introduction

It helps to keep in mind that when a ring $\mathbf{R}$ is viewed as some set of generic functions on a space X, its $\mathbf{R}$-modules should be looked at as a kind of representations of $\mathbf{R}$. To add a measure of control in the wildness the full gadgetry of homological algebra has to be brought in to generate various schemes of classification, say by attaching numerical invariants to a ring or module: Krull dimension, projective dimension, Castelnuovo–Mumford regularity, and Hilbert functions [in graded structures]. Several of these invariants arise from cohomological calculations providing metrics that capture the complexity—sometimes the deviation from smoothness of the ring/space. The class of objects called *Cohen–Macaulay* forms a paradigm for what might be called the "good guys." The corresponding geometric objects are not always smooth, but their cohomology is slimmed-down when compared to the wilder singular spaces. It is often the case that computation on them run faster. Interestingly enough major examples of rings of invariants, long after the solution by Nagata of Hilbert's 14th problem, were shown by Hochster to be Cohen–Macaulay. They are more efficiently packaged, and many geometric constructions that hold in smooth spaces can be carried out also on them. Of obvious interest is how such objects come about—that is, identifying processes, e.g., extensions or modifications—leading to them. An associated problem is that of devising criteria to

W.V. Vasconcelos (✉)
Department of Mathematics, Rutgers University, 110 Frelinghuysen Rd, Piscataway, NJ 08854-8019, USA
e-mail: vasconce@math.rutgers.edu

I. Peeva (ed.), *Commutative Algebra: Expository Papers Dedicated to David Eisenbud on the Occasion of His 65th Birthday*, DOI 10.1007/978-1-4614-5292-8_22,

detect the presence of the property and more economically determine their invariants without having to go through cumbersome cohomological calculations.

Probably the single major metric for these objects is their *multiplicity*. There are various methods to study it, some specific to special classes of objects. One can argue for the need for an extension of the multiplicity applied in broader contexts. In general, the multiplicity function of algebraic geometry and commutative algebra has wide usage in sizing up a module or an algebra. Its drawbacks are that it ignores the lower dimensional components of the structure and does not behave too well with respect to hyperplane sections. To account for the components, Bayer and Mumford [1] introduced the *arithmetic degree*. To account for a fuller behavior of hyperplane sections, *extended degrees* were introduced. They all coincide on Cohen–Macaulay rings and algebras. Unlike multiplicities, that have a combinatorial foundation, extended degrees are both combinatorial and very homological. The need to express local behavior in global terms presents many technical challenges for their computation.

Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring (or a Noetherian graded algebra) and let $\mathcal{M}(\mathbf{R})$ be the category of finitely generated $\mathbf{R}$-modules (or the appropriate category of graded modules). A *degree function* is simply a numerical function

$$\mathbf{d} : \mathcal{M}(\mathbf{R}) \mapsto \mathbb{N}.$$

The more interesting of them initialize on modules of finite length, are additive on *certain* short exact sequences, and have mechanisms that control how the functions behave under generic hyperplane sections. Some of these functions are classical degree (multiplicity) and Castelnuovo–Mumford regularity. Most degree functions $\mathbf{d}$ are derived from the leading coefficients of Hilbert polynomials of certain filtrations. Refinements involve assembling $\mathbf{d}$ by adding several of these coefficients, so that a value such as $\mathbf{d}(A)$ may capture several elements of the structure of A.

The most demanding requirement on these functions are those regarding their behavior with respect to generic hyperplane sections. When A has positive depth and $h \in \mathbf{R}$ is a generic hyperplane section, the comparison

$$\mathbf{d}(A) \leftrightarrow \mathbf{d}(A/hA)$$

is the principal divider among the degrees. For those directly derived from the classical multiplicity, one always has

$$\mathbf{d}(A) \leq \mathbf{d}(A/hA).$$

For the other family of degrees, the cohomological multiplicities, the relationship is reversed,

$$\mathbf{d}(A) \geq \mathbf{d}(A/hA).$$

It is this feature that makes them appealing as complexity benchmarks, motivating our interest on cohomological (or extended) degrees [6], particularly on the homological degree and bdeg (the so-called big Degs).

1 Cohomological Degrees

Introduction

Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring (or a standard graded algebra over an Artinian local ring) of infinite residue field. We denote by $\mathcal{M}(\mathbf{R})$ the category of finitely generated $\mathbf{R}$-module (or the corresponding category of graded $\mathbf{R}$-modules).

Cohomological degrees are extensions of the multiplicity that agree with it for Cohen–Macaulay modules. On Artinian modules they are just the ordinary length function. In dimension one there is just one such function, but in higher dimension there is an infinite set of them. A general class of these functions was introduced in [6], while a prototype was defined earlier in [18]. In his thesis [9], Gunston carried out a more formal examination of such functions in order to introduce his own construction of a new cohomological degree.

A convenient approach to build and study the properties of these degrees is via recursion. Therefore one of the points that must be taken care of is that of an appropriate *generic hyperplane* section.

1.1 The Ordinary Multiplicity

The premier example of a degree (vector space dimension excluded) is that of the *multiplicity* of a module. Let $(\mathbf{S}, \mathfrak{m})$ be a local ring and let M be a finitely generated $\mathbf{S}$-module. The Hilbert–Samuel function of M is

$$H_M : n \mapsto \lambda(M/\mathfrak{m}^{n+1}M), \tag{1}$$

which for $n \gg 0$ is given by a polynomial:

$$H_M(n) = \frac{\deg(M)}{d!} n^d + \text{lower order terms.}$$

The integer $d = \dim M$ is the *dimension* of M, and the integer $\deg(M)$ is its *multiplicity*. Under certain conditions, $\deg(M)$ can be interpreted as the *volume* of a manifold, or of a polytope.

This *degree* arises in the setting of finitely generated graded modules over graded $\mathbf{R}$-algebra $\mathbf{A}$, for an Artinian ring $\mathbf{R}$: From $\mathbf{A} = \mathbf{R}[z_1, \ldots, z_d]$ and $M = \bigoplus_{n\in\mathbb{Z}} M_n$ (more general grading will be considered later), under fairly broad conditions the function, referred as its *Hilbert function*,

$$H_M : n \mapsto \lambda(M_n)$$

provides a wealth of information, with a *degree* coding a great deal of it. In reality the Hilbert functions mentioned above are distinct, but are closely related, that is, we will call either $H_M(n)$ or

$$\sum_{k\leq n} H_k(M)$$

as the Hilbert function of M, or refer to the second kind as the Hilbert–Samuel function.

One of the paths we employ to obtain new degrees from old ones is the following. Let **R** be a Noetherian ring and **A** a finitely generated graded **R**-algebra. Usually we will require that **A** be a standard graded algebra, but later we shall discuss more general gradings. Suppose we are equipped with a numerical function λ on the category of finitely generated **R**-modules. Let M be a finitely generated graded **A**-modules, to which we associate the formal power series

$$\mathbb{P}_M(\mathbf{t}) = \sum_{n\in\mathbb{Z}} \lambda(M_n)\mathbf{t}^n.$$

Under some conditions, this function is a rational function,

$$\mathbb{P}_M(\mathbf{t}) = \frac{h_M(\mathbf{t})}{(1-\mathbf{t})^d},$$

$h_M(\mathbf{t}) \in \mathbb{Z}[\mathbf{t}, \mathbf{t}^{-1}]$. From this representation is extracted the Hilbert polynomial

$$P_M(n) = \sum_{i=0}^{d-1}(-1)^i \mathrm{e}_i(M)\binom{n+d-i-1}{d-i-1},$$

with the property that $\lambda(M_n) = P_M(n)$ for $n \gg 0$.

There is an obvious degradation in the information about M when we move from $\mathbb{P}_M(\mathbf{t})$ to one of the individual coefficients of $P_M(\mathbf{t})$:

$$\mathbb{P}_M(\mathbf{t}) \to P_M(\mathbf{t}) \to \{\mathrm{e}_0(M), \mathrm{e}_1(M), \ldots, \mathrm{e}_{d-1}(M)\} \to \mathrm{e}_i(M).$$

Nevertheless it will be the $\mathrm{e}_i(M)$, particularly on $\mathrm{e}_0(M)$ and $\mathrm{e}_1(M)$, that we look for carriers of information on M. The reasons are both for the practicality it involves but also justified by the surprising amount of information these coefficients may pack.

In the study of Hilbert functions associated to a filtration $\mathcal{M} = \{M_n, n \in \mathbb{Z}\}$ of a module M, it is common to consider the function

$$H_{\mathcal{M}}(n) = \lambda(M/M_{n+1}).$$

There will be (under appropriate finiteness conditions) a rational function

$$\mathbb{P}_{\mathcal{M}}(\mathbf{t}) = \frac{h_{\mathcal{M}}(\mathbf{t})}{(1-\mathbf{t})^{d+1}},$$

$h_{\mathcal{M}}(\mathbf{t}) \in \mathbb{Z}[\mathbf{t}, \mathbf{t}^{-1}]$. From this representation is extracted the Hilbert polynomial

$$P_{\mathcal{M}}(n) = \sum_{i=0}^{d} (-1)^i \mathrm{e}_i(M) \binom{n+d-i}{d-i},$$

with the property that $\lambda(M/M_{n+1}) = P_{\mathcal{M}}(n)$ for $n \gg 0$.

Hyperplane Sections and Hilbert Polynomials

We need rules to compute these coefficients. Typically they involve the so-called *superficial elements* or *filter regular elements*. We keep the terminology of *generic hyperplane section*, even when dealing with Samuel's multiplicity with respect to an $\mathfrak{m}$-primary ideal I and its Hilbert coefficients $\mathrm{e}_i(M) = \mathrm{e}_i(I, M)$. Hopefully this usage will not lead to undue confusion. We say that $h \in I$ is a *parameter* for M, if $\dim_{\mathbf{R}} M/hM < \dim_{\mathbf{R}} M$.

Let us begin with the following.

Lemma 1. *Let* $(\mathbf{R}, \mathfrak{m})$ *be a Noetherian local ring, I an $\mathfrak{m}$-primary ideal of* $\mathbf{R}$*, and M a finitely generated* $\mathbf{R}$*-module. Let $h \in I$ and suppose that $\lambda(0 :_M h) < \infty$. Then we have the following:*

(a) $\lambda(0 :_M h) \leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM))$.
(b) h is a parameter for M, if $\dim_{\mathbf{R}} M > 0$.
(c) If $\dim_{\mathbf{R}} M > 1$ *and M/hM is Cohen–Macaulay, then M is Cohen–Macaulay.*

Proof. Suppose that $\dim_{\mathbf{R}} M > 0$ and let $\mathfrak{p} \in \mathrm{Supp}_{\mathbf{R}} M$ with $\dim \mathbf{R}/\mathfrak{p} = \dim_{\mathbf{R}} M$. Then

$$(0) :_{M_{\mathfrak{p}}} h = (0),$$

since $\mathfrak{p} \neq \mathfrak{m}$. As $\dim_{\mathbf{R}_{\mathfrak{p}}} M_{\mathfrak{p}} = 0$, we get $h \notin \mathfrak{p}$. Hence h is a parameter for M, if $\dim_{\mathbf{R}} M > 0$.

We look at the exact sequence

$$0 \to (0) :_M h \to \mathrm{H}^0_{\mathfrak{m}}(M) \xrightarrow{h} \mathrm{H}^0_{\mathfrak{m}}(M) \xrightarrow{\varphi} \mathrm{H}^0_{\mathfrak{m}}(M/hM) \to \mathrm{H}^1_{\mathfrak{m}}(M) \xrightarrow{h} \mathrm{H}^1_{\mathfrak{m}}(M) \to \mathrm{H}^1_{\mathfrak{m}}(M/hM) \to \cdots$$

of local cohomology modules derived from the exact sequence

$$0 \to (0) :_M h \to M \xrightarrow{h} M \to M/hM \to 0$$

of $\mathbf{R}$-modules. We then have

$$\lambda((0) :_M h) = \lambda(\mathrm{Im}\,\varphi) = \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM)) - \lambda((0) :_{\mathrm{H}^1_{\mathfrak{m}}(M)} h) \leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM)).$$

Therefore, if $\dim_{\mathbf{R}} M > 1$ and M/hM is Cohen–Macaulay, then h is M-regular, and hence M is Cohen–Macaulay as well. □

We will make repeated use of [7, (12.1)], [14, (22.6)], and [13, Sect. 3].

Proposition 2. *Let* $(\mathbf{R}, \mathfrak{m})$ *be a Noetherian local ring,* I *an* $\mathfrak{m}$*-primary ideal of* $\mathbf{R}$*, and* M *a finitely generated* $\mathbf{R}$*-module with* $r = \dim_{\mathbf{R}} M > 0$.

(a) Let $h \in I$ *and assume that* h *is superficial for* M *with respect to* I *(in particular* $h \in I \setminus \mathfrak{m}I$*). Then the Hilbert coefficients of* M *and* M/hM *satisfy*

$$\mathrm{e}_i(M) = \mathrm{e}_i(M/hM) \ \textit{for} \ 0 \le i < r-1 \ \textit{and}$$
$$\mathrm{e}_{r-1}(M) = \mathrm{e}_{r-1}(M/hM) + (-1)^r \lambda(0 :_M h).$$

(b1) Let $0 \to A \to B \to C \to 0$ *be an exact sequence of finitely generated* $\mathbf{R}$*-modules. If* $t = \dim A < s = \dim B$*, then* $\mathrm{e}_i(B) = \mathrm{e}_i(C)$ *for* $0 \le i < s-t$*. In particular, if* $t = 0$ *and* $s \ge 2$*, then* $\mathrm{e}_1(B) = \mathrm{e}_1(C)$.

(b2) If $0 \to A \to B \to C \to 0$ *is an exact sequence of modules of the same dimension, then*

$$\mathrm{e}_0(B) = \mathrm{e}_0(A) + \mathrm{e}_0(C) \ \textit{and}$$
$$\mathrm{e}_1(B) \ge \mathrm{e}_1(A) + \mathrm{e}_1(C).$$

(c) If M *is a module of dimension* 1 *and* $I = (h)$ *is a parameter ideal for* M*, then*

$$\mathrm{e}_1(M) = -\lambda(\mathrm{H}^0_{\mathfrak{m}}(M)).$$

(d) If M *is a module of dimension* 2 *and* $I = (h, b)$ *is a parameter ideal for* M*, then*

$$\begin{aligned}\mathrm{e}_1(M) &= \mathrm{e}_1(M/hM) + \lambda(0 :_M h) = -\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM)) + \lambda(0 :_M h)\\ &= -\lambda((0) :_{\mathrm{H}^1_{\mathfrak{m}}(M)} h).\end{aligned}$$

Proof. See Proof of Lemma 1 for assertion (d). □

1.2 Extended Degree Functions

Throughout we suppose that the residue field k of $\mathbf{R}$ is infinite. Moreover, in some of the discussions, we will assume that $\mathbf{R}$ is a quotient of a Gorenstein ring. Both conditions are realized by considering two changes of rings: $\mathbf{R} \to \mathbf{S} = \mathbf{R}[X]_{\mathfrak{m}[X]} \to \widehat{\mathbf{S}}$, the latter being the completion relative to the maximal ideal.

Definition 3. If $(\mathbf{R}, \mathfrak{m})$ is a local ring, a *notion of genericity* on $\mathcal{M}(\mathbf{R})$ is a function

$$U : \{\text{isomorphism classes of } \mathcal{M}(\mathbf{R})\} \longrightarrow \{\text{nonempty subsets of } \mathfrak{m}/\mathfrak{m}^2\}$$

subject to the following conditions for each $A \in \mathcal{M}(\mathbf{R})$:

(i) If $f - g \in \mathfrak{m}^2$ then $f \in U(A)$ if and only if $g \in U(A)$.
(ii) The set $\overline{U(A)} \subset \mathfrak{m}/\mathfrak{m}^2$ contains a nonempty Zariski-open subset.
(iii) If depth $A > 0$ and $f \in U(A)$, then f is regular on A.

There is a similar definition for graded modules. We shall usually switch notation, denoting the algebra by $\mathbf{S}$.

Another extension is that associated to an $\mathfrak{m}$-primary ideal I [12]: a notion of genericity on $\mathcal{R}$ with respect to I is a function

$$U : \{\text{isomorphism classes of } \mathcal{M}(\mathbf{R})\} \longrightarrow \{\text{nonempty subsets of } I/\mathfrak{m}I\}$$

subject to the following conditions for each $A \in \mathcal{M}(\mathbf{R})$:

(i) If $f - g \in \mathfrak{m}I$ then $f \in U(A)$ if and only if $g \in U(A)$.
(ii) The set $\overline{U(A)} \subset I/\mathfrak{m}I$ contains a nonempty Zariski-open subset.
(iii) If depth $A > 0$ and $f \in U(A)$, then f is regular on A.

Fixing a notion of genericity $U(\cdot)$ one has the following extension of the classical multiplicity.

Definition 4. A *cohomological degree*, or *extended multiplicity function*, is a function

$$\operatorname{Deg}(\cdot) : \mathcal{M}(\mathbf{R}) \mapsto \mathbb{N}$$

that satisfies the following conditions.

(i) If $L = \mathrm{H}^0_{\mathfrak{m}}(M)$ is the submodule of elements of M that are annihilated by a power of the maximal ideal and $\overline{M} = M/L$, then

$$\operatorname{Deg}(M) = \operatorname{Deg}(\overline{M}) + \lambda(L), \tag{2}$$

where $\lambda(\cdot)$ is the ordinary length function.

(ii) (Bertini's rule) If M has positive depth, there is $h \in \mathfrak{m} \setminus \mathfrak{m}^2$, such that

$$\operatorname{Deg}(M) \geq \operatorname{Deg}(M/hM). \tag{3}$$

(iii) (The calibration rule) If M is a Cohen–Macaulay module, then

$$\operatorname{Deg}(M) = \deg(M), \tag{4}$$

where $\deg(M)$ is the ordinary multiplicity of M.

These functions will be referred to as **big Degs**. If $\dim \mathbf{R} = 0$, $\lambda(\cdot)$ is the unique Deg function. For $\dim \mathbf{R} = 1$, the function $\mathrm{Deg}(M) = \lambda(L) + \deg(M/L)$ is the unique extended degree. When $d \geq 2$, there are several big Degs. An explicit Deg, for all dimensions, was introduced in [18]. We refer to h as a *superficial* element relative to M and Deg. We will abuse the terminology when h is chosen to be superficial for M/L.

To define such functions on general local rings one makes use of standard changes of rings. For example, if $(\mathbf{R}, \mathfrak{m})$ is a Noetherian local ring and $\mathbf{X}$ is a set of indeterminates, $\mathbf{S} = \mathbf{R}[\mathbf{X}]_{\mathfrak{m}[\mathbf{X}]}$ is a Noetherian local ring of dimension d, and the flat change of rings $\mathbf{R} \to \mathbf{S}$ preserves multiplicity. Its large residue field permits the development of notions of genericity and the definition of extended degrees.

In one special case this notion becomes familiar.

Proposition 5. *Let* $\mathbf{R}$ *be a local Noetherian ring and* M *a finitely generated* $\mathbf{R}$-*module of Krull dimension one. For any extended multiplicity function* $\mathrm{Deg}(\cdot)$*,*

$$\mathrm{Deg}(M) = \deg(M) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)) = \mathrm{adeg}(M).$$

If I is an $\mathfrak{m}$-primary ideal, in the proposition above, replacing $\deg(M)$ by Samuel's multiplicity $\mathrm{e}(I; M)$ would result in the extended multiplicity $\mathrm{Deg}_I(M)$. This usually means that M is a finitely generated graded module over a ring such as $\mathrm{gr}_I(\mathbf{R})$.

Functions with the Bertini Rule

Both the notion of genericity and the Bertini property above occur in limited ways. We will encounter functions

$$\mathbf{T} : \mathcal{M}(\mathbf{R}) \to \mathbb{Z},$$

such that for certain classes of modules M satisfies

$$\mathbf{T}(M/hM) \leq \mathbf{T}(M),$$

for generic elements h. For modules of positive Krull dimension if $\mathbf{T}$ is a multiplicity function the inequality goes the other way.

1.3 General Properties of Extended Degrees

Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and let $\mathbf{A}$ be a standard graded algebra over $\mathbf{R}$. We are going to derive basic properties of extended degree functions defined over $\mathbf{R}$ or $\mathbf{A}$. Throughout Deg is an unspecified extended degree function. We examine some relationships between the $\mathrm{Deg}(M)$ of a module and some data expressed in its projective resolution, typically the degrees and the ranks of the higher modules of syzygies.

Remark 6. The axiom (ii) above admits variations. Let M be a finitely generated $\mathbf{R}$-module and set $H = \mathrm{H}^0_{\mathfrak{m}}(M)$ and $N = M/H$. Suppose h is superficial with regard N and Deg. Then we have a short exact sequence

$$0 \to H/hH \longrightarrow M/hM \longrightarrow N/hN \to 0,$$

and therefore

$$\mathrm{Deg}(M/hM) = \mathrm{Deg}(H/hH) + \mathrm{Deg}(N/hN) \leq \mathrm{Deg}(H) + \mathrm{Deg}(N) = \mathrm{Deg}(M).$$

Cohen–Macaulay Deficiency

The following assertion is a justification for the terminology.

Proposition 7. *For any extended degree* Deg*, the Cohen–Macaulay deficiency*

$$I(M) = \mathrm{Deg}(M) - \deg(M)$$

vanishes if and only if M is Cohen–Macaulay.

Proof. We begin by observing that the axioms guarantee that $\mathrm{Deg}(M) \geq \deg(M)$. Let us use induction on the dimension of M. If $\dim M \leq 1$, the assertion holds since $\mathrm{Deg}(M) = \deg(M) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(M))$.

If $\dim M \geq 2$, $L = \mathrm{H}_{\mathfrak{m}}(M)$ and h is a superficial element for $\mathrm{Deg}(M')$, $M' = M/L$, then

$$\mathrm{Deg}(M) = \lambda(L) + \mathrm{Deg}(M') \geq \lambda(L) + \mathrm{Deg}(M'/hM').$$

The assumption implies that $L = 0$ and $\mathrm{Deg}(M'/hM') = \deg(M') = \deg(M)$. By induction M' is Cohen–Macaulay and therefore M is Cohen–Macaulay as well. □

Betti Numbers

We first want to emphasize the control that extended degrees have over the Betti numbers of the modules.

Theorem 8. *Let M be a module of dimension d and let $\mathbf{x} = \{x_1, \ldots, x_d\}$ be a superficial sequence relative to M and the extended degree* Deg*. Then*

$$\lambda(M/(\mathbf{x})M) \leq \mathrm{Deg}(M).$$

Proof. This is a straightforward calculation. If $M_0 = \mathrm{H}^0_{\mathfrak{m}}(M)$ and $M' = M/H$, the exact sequence

$$0 \to H/x_1H \longrightarrow M/x_1M \longrightarrow M'/x_1M' \to 0$$

gives

$$\mathrm{Deg}(M') = \mathrm{Deg}(M) - \lambda(H) \leq \mathrm{Deg}(M) - \lambda(H/x_1H),$$

and allows an easy induction on the dimension of the module. □

We assume that $(\mathbf{R}, \mathfrak{m}, k = \mathbf{R}/\mathfrak{m})$ is a Cohen–Macaulay local ring. For any finitely generated $\mathbf{R}$-module A, we denote by $\beta_i^{\mathbf{R}}(A)$ its ith Betti number and by $\mu_i(A)$ its ith *Bass number*:

$$\beta_i^{\mathbf{R}}(A) = \dim_k \mathrm{Tor}_i^{\mathbf{R}}(k, A)$$
$$\mu_i(A) = \dim_k \mathrm{Ext}^i_{\mathbf{R}}(k, A).$$

We recall a classical result on the Betti numbers of the residue field of $\mathbf{R}$ under the change of rings $\mathbf{R} \to \mathbf{R}/(\mathbf{x})$, where $\mathbf{x}$ is a regular element of $\mathbf{R}$. This is best expressed in term of the *Poincaré series* of $\mathbf{R}$:

$$\mathbf{P}(\mathbf{R}) = \sum_{i \geq 0} \beta_i^{\mathbf{R}}(k)\mathbf{t}^i.$$

Theorem 9 (Gulliksen–Levin [8, Corollary 3.4.2]). *Let* $\mathbf{x}$ *be a regular element of* $\mathbf{R}$. *Put* $\mathbf{R}' = \mathbf{R}/(\mathbf{x})$. *Then*

(1) ([Tate]) *If* $\mathbf{x} \in \mathfrak{m}^2$ *then*
$$\mathbf{P}(\mathbf{R}') = \frac{\mathbf{P}(\mathbf{R})}{1 - \mathbf{t}^2}.$$
(2) If $\mathbf{x} \in \mathfrak{m} \setminus \mathfrak{m}^2$ *then*
$$\mathbf{P}(\mathbf{R}') = \frac{\mathbf{P}(\mathbf{R})}{1 + \mathbf{t}}.$$

Theorem 10. *Let A be a finitely generated* $\mathbf{R}$*-module. For any* $\mathrm{Deg}_{\mathfrak{m}}(\cdot)$ *function and any integer* $i \geq 0$*,*

$$\beta_i^{\mathbf{R}}(A) \leq \beta_i^{\mathbf{R}}(k) \cdot \mathrm{Deg}_{\mathfrak{m}}(A),$$
$$\mu_i(A) \leq \mu_i(k) \cdot \mathrm{Deg}_{\mathfrak{m}}(A),$$

in particular, $\nu(A) \leq \mathrm{Deg}_{\mathfrak{m}}(A)$.

Proof. If L is the submodule of A of finite support, the exact sequence

$$0 \to L \longrightarrow A \longrightarrow A' \to 0$$

gives $\beta_i(A) \leq \beta_i(L) + \beta_i(A')$ (for simplicity we set $\beta_i(A) = \beta_i^{\mathbf{R}}(A)$). For the summand $\beta_i(L)$, by induction on the length of L, one has that $\beta_i(L) \leq \lambda(L)\beta_i(k)$. For the other summand, one chooses a sufficiently generic hyperplane section (good for both A' and $\mathbf{R}$ if need be, in particular an element $\mathbf{x} \in \mathfrak{m} \setminus \mathfrak{m}^2$), and setting $\mathbf{R}' = \mathbf{R}/(\mathbf{x})$ gives

$$\beta_i(A') = \beta_i^{\mathbf{R}'}(A'/\mathbf{x}A')).$$

Now by an induction on the dimension of A', one has

$$\beta_i^{\mathbf{R}'}(A'/\mathbf{x}A') \leq \beta_i^{\mathbf{R}'}(k)\mathrm{Deg}(A'/\mathbf{x}A').$$

Finally, since $\mathrm{Deg}(A'/\mathbf{x}A') \leq \mathrm{Deg}_{\mathfrak{m}}(A')$, we appeal to Theorem 9(2) (saying that $\beta_i^{\mathbf{R}'}(k) \leq \beta_i^{\mathbf{R}}(k)$) to prove the assertion.

One can use a similar argument for $\mu_i(A)$. □

Remark 11. The bound is different when we use general Samuel multiplicities. Consider a case of a system of parameters $\mathbf{z} = \{\mathbf{x}_1, \ldots, \mathbf{x}_r, \mathbf{y}_1, \ldots, \mathbf{y}_s\}$, $r + s = \dim A$, where $\mathbf{x}_i \in \mathfrak{m}^2$ and the $\mathbf{y}_j$ form a subset of a minimal set of generators of $\mathfrak{m}$. We leave as an exercise the proof that

$$P(A) = \sum_{i\geq 0} \beta_i^{\mathbf{R}}(A)\mathbf{t}^i \leq \mathrm{Deg}_{\mathbf{z}}(A)\frac{\mathbf{P(R)}}{(1-\mathbf{t}^2)^r}.$$

2 Degs and Castelnuovo–Mumford Regularity

Introduction

This is one of most useful of the degree functions and has excellent treatments in [7]. It has several equivalent formulations, one of which is the following: Let $\mathbf{R} = k[x_1, \ldots, x_d]$ be a ring of polynomials over the field k with the standard grading, and let $\mathbf{A}$ be a finitely generated graded $\mathbf{R}$-module with a minimal graded resolution

$$0 \to F_n \longrightarrow F_{n-1} \longrightarrow \cdots \longrightarrow F_1 \longrightarrow F_0 \to 0,$$

$$F_j = \bigoplus_j \mathbf{R}[-a_{ij}].$$

Then $\mathrm{reg}(\cdot)$ is defined by (see below for a more precise formulation)

$$\mathrm{reg}(\mathbf{A}) = \sup\{a_{i,j} - i\}.$$

Before we point out a more abstract formulation of the Castelnuovo–Mumford regularity, or simply *regularity*, one of its important properties shows up: the Hilbert–Poincaré series of $\mathbf{A}$,

$$\mathbb{P}_{\mathbf{A}}(t) = \frac{\sum_{ij}(-1)^j t^{a_{ij}}}{(1-t)^d},$$

encodes the Hilbert function $H_{\mathbf{A}}(t)$ and Hilbert polynomial $\mathcal{P}_{\mathbf{A}}(t)$, and from the expression for $\mathbb{P}_{\mathbf{A}}(t)$, one has

$$H_{\mathbf{A}}(n) = \mathcal{P}_{\mathbf{A}}(n), \quad n \geq \operatorname{reg}(\mathbf{A}).$$

Another use of *regularity* is for the estimation of various indices attached to algebraic structures. Consider for instance a standard graded algebra $\mathbf{A}$ over the infinite field k. If $\dim \mathbf{A} = d$, by Noether Normalization, there are subalgebras, $\mathbf{B} = k[z_1, \ldots, z_d]$, $z_i \in A_1$, $d = \dim \mathbf{A}$, such that $\mathbf{A}$ is finite over $\mathbf{B}$. This implies that

$$A_{n+1} = (z_1, \ldots, z_d)A_n.$$

The smallest such degree n is called the *reduction number of* $\mathbf{A}$ *relative to* $\mathbf{B}$, $\operatorname{red}_{\mathbf{B}}(\mathbf{A})$. Its value may vary with the choice of $\mathbf{B}$, but the inequality

$$\operatorname{red}_{\mathbf{B}}(\mathbf{A}) \leq \operatorname{reg}(\mathbf{A})$$

will always hold.

Cohomological Formulation

Let $\mathbf{R} = \bigoplus_{n\geq 0} R_n = R_0[R_1]$ be a finitely generated graded algebra over the Noetherian ring $\mathbf{R}_0$. For any graded $\mathbf{R}$-module F, define

$$\alpha(F) = \begin{cases} \sup\{n \mid F_n \neq 0\} & \text{if } F \neq 0, \\ -\infty & \text{if } F = 0. \end{cases}$$

Let $\mathbf{R} = \bigoplus_{n\geq 0} R_n$ be a standard graded ring of Krull dimension d with irrelevant maximal ideal $M = (\mathfrak{m}, \mathbf{R}_+)$ (note that $(R_0, \mathfrak{m})$ is a local ring), and let E be a finitely generated graded $\mathbf{R}$-module. Some local cohomology modules $\mathrm{H}^i_J(E)$ of E often give rise to graded modules with $\alpha(\mathrm{H}^i_J(E)) < \infty$. This occurs, for instance, when $J = \mathbf{R}_+$ or $J = M$. This fact gives rise to several numerical measures of the cohomology of E.

Definition 1. For any finitely generated graded $\mathbf{R}$-module F, and for each integer $i \geq 0$, the integer

$$a_i(F) = \alpha(\mathrm{H}^i_M(F))$$

is the ith *a-invariant* of F.

We shall also make use of the following complementary notion.

Definition 2. For any finitely generated graded **R**-module F, and for each integer $i \geq 0$, the integer

$$\underline{a}_i(F) = \alpha(\mathrm{H}^i_{\mathbf{R}_+}(F))$$

is the ith $\underline{a}$*-invariant* of F.

By abuse of terminology, if F has Krull dimension d, we shall refer to $a_d(F)$ as simply the a-invariant of F. In the case $F = \mathbf{R}$, if $\omega_{\mathbf{R}}$ is the canonical module of **R**, by local duality it follows that

$$a(\mathbf{R}) = -\inf\{\, i \mid (\omega_{\mathbf{R}})_i \neq 0 \,\}. \tag{5}$$

The $\underline{a}_i$-invariants are usually assembled into the *Castelnuovo–Mumford regularity* of F

$$\mathrm{reg}(F) = \sup\{\underline{a}_i(F) + i \mid i \geq 0\}.$$

2.1 Basic Comparisons

The role of the regularity in the connection between the Hilbert series and the Hilbert polynomial of modules over polynomial rings was already observed. The following is more precise and general [2, Theorem 4.4.3].

Theorem 3. *Let* **R** *be a local Artinian ring,* **A** *a standard graded algebra over* **R**, *and M a finitely generated graded* **A**-*module. Then*

$$H_M(n) - \mathcal{P}_M(n) = \sum_{j \geq 0} (-1)^j \lambda(\mathrm{H}^j_{\mathbf{A}_+}(M)_n),$$

in particular $H_M(n) = \mathcal{P}_M(n)$ for $n >$ $\mathrm{reg}(M)$.

2.2 Hyperplane Sections

We briefly describe the behavior of $\mathrm{reg}(\cdot)$ with regard to some exact sequences.

Proposition 4. *Let* **R** *be a standard graded algebra, and let*

$$0 \to A \longrightarrow B \longrightarrow C \to 0$$

be an exact sequence of finitely generated **R**-*modules* (*and homogeneous homomorphisms*), *then*

$$\mathrm{reg}(B) \leq \mathrm{reg}(A) + \mathrm{reg}(C).$$

Similarly,

$$\operatorname{reg}(A) \leq \operatorname{reg}(B) + \operatorname{reg}(C), \quad \text{and} \quad \operatorname{reg}(C) \leq \operatorname{reg}(A) + \operatorname{reg}(B).$$

Proof. The functions $\underline{a}_i(\cdot) = \alpha(\mathrm{H}^i_{\mathbf{R}_+}(\cdot))$ clearly satisfy $\underline{a}_i(B) \leq \underline{a}_i(A) + \underline{a}_i(C)$. Thus $a_i(B) + i \leq \operatorname{reg}(A) + \operatorname{reg}(C)$ for each i.

The other assertions have similar proofs. □

The function $\operatorname{reg}(\cdot)$ can also be characterized by its initialization on modules of finite length (see [7, Proposition 20.20]):

Proposition 5. *If h is a linear form of $\mathbf{R}$ whose annihilator $0 :_A h$ has finite length, then*

$$\operatorname{reg}(A) = \max\{\operatorname{reg}(0 :_A h), \operatorname{reg}(A/hA)\}.$$

The proof goes through an analysis of the exact sequence

$$0 \to (0 :_A h)[-1] \longrightarrow A[-1] \longrightarrow A \longrightarrow A/hA \to 0.$$

As a long exercise, the reader can verify that $\operatorname{reg}(\cdot)$ is the unique degree function satisfying these rules.

Let $\mathbf{R}$ be a standard graded algebra and M a finitely generated graded $\mathbf{R}$-module. Suppose $\alpha(M)$ is the maximum of the degrees of a minimal generating set of M. If M is a module of finite length, it is clear that

$$\operatorname{reg}(M) \leq \alpha(M) + \lambda(M) - 1.$$

We will prove an extension of this inequality where $\lambda(M)$ is replaced by $\operatorname{Deg}(M)$ [15, Theorem 2.5].

Theorem 6. *Let $\mathbf{R} = k[x_1, \ldots, x_n]$ be a ring of polynomials over the infinite field k, and let M be a finitely generated graded $\mathbf{R}$-module of positive dimension. Then we have for every sufficiently general linear form h and for any extended degree* Deg,

$$\operatorname{Deg}(M/hM) - \operatorname{reg}(M/hM) \leq \operatorname{Deg}(M) - \operatorname{reg}(M).$$

Proof. By Proposition 5, we have

$$\operatorname{reg}(M) = \max\{\operatorname{reg}(0 :_M h), \operatorname{reg}(M/hM)\},$$

so that if $\operatorname{reg}(M) = \operatorname{reg}(M/hM)$ we can make use of Remark 6 asserting that $\operatorname{Deg}(M/hM) \leq \operatorname{Deg}(M)$ for all sufficiently generic $h's$.

We may thus assume $\operatorname{reg}(M/hM) < \operatorname{reg}(M)$. For any module A, denote $A^* = A/\mathrm{H}^0_{\mathfrak{m}}(A)$. Consider the exact sequence

$$0 \to H \longrightarrow M \longrightarrow N \to 0,$$

where $H = \mathrm{H}^0_{\mathfrak{m}}(M)$. Note that $\dim N > 0$. Since h is regular on N, we have the exact sequence

$$0 \to H/hH \longrightarrow M/hM \longrightarrow N/hN \to 0.$$

Taking the local cohomology of this sequence we obtain the exact sequences

$$0 \to H/hH \longrightarrow \mathrm{H}^0_{\mathfrak{m}}(M/hM) \longrightarrow \mathrm{H}^0_{\mathfrak{m}}(N/hN) \to 0$$

and the isomorphism

$$\mathrm{H}^i_{\mathfrak{m}}(M/hM) \simeq \mathrm{H}^i_{\mathfrak{m}}(N/hN), \quad i \geq 1,$$

$(M/hM)^* \simeq (N/hN)^*$. In particular we have that $\mathrm{reg}(N/hN) \leq \mathrm{reg}(M/hM)$.

We now process $\mathrm{Deg}(M/hM)$:

$$\begin{aligned} \mathrm{Deg}(M/hM) &= \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM) + \mathrm{Deg}((M/hM)^*) \\ &= \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM) + \mathrm{Deg}((N/hN)^*) \\ &\leq \sum_{j \leq \mathrm{reg}(M/hM)} \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)_j) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(N/hN)_j) + \mathrm{Deg}((N/hN)^*) \\ &= \sum_{j \leq \mathrm{reg}(M/hM)} \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)_j) + \mathrm{Deg}(N/hN) \\ &\leq \sum_{j \leq \mathrm{reg}(M/hM)} \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)_j) + \mathrm{Deg}(N). \end{aligned}$$

To complete the proof it suffices to add to this expression the observation

$$\mathrm{reg}(M) - \mathrm{reg}(M/hM) \leq \sum_{j=\mathrm{reg}(M/hM)+1}^{\mathrm{reg}(M)} \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)_j).$$

Summing up we obtain the desired inequality. □

Theorem 7. *Let* $\mathbf{A}$ *be a finitely generated graded* $\mathbf{R}$*-module, generated by elements of degree at most* $\alpha(\mathbf{A})$*. Then for any extended degree function* Deg*,*

$$\mathrm{reg}(\mathbf{A}) \leq \mathrm{Deg}(\mathbf{A}) + \alpha(\mathbf{A}).$$

2.3 Regularity of Some Derived Functors

Let $\mathbf{R} = k[x_1, \ldots, x_n]$ be a polynomial ring in n indeterminates over the field k. We briefly describe some results of Chardin, Ha, and Hoa [3] on regularity bounds for derived functors of $\mathrm{Hom}(\mathbf{X}, \mathbf{Y})$ and $\mathbf{X} \otimes \mathbf{Y}$. The first of their results [3, Lemma 2.1] is a beautiful cohomological calculation:

Theorem 8. *Let $\mathcal{F}_\bullet$ be a graded complex of free $\mathbf{R}$-modules with*

$$\mathcal{F}_i = \bigoplus_{f_i \leq j \leq b_i} \mathbf{R}[-j]^{\beta_{ij}}.$$

Set $T_i = \sum_j \beta_{ij}$. Then

$$\mathrm{reg}(\mathrm{H}_i(\mathcal{F}_\bullet)) \leq \max\{b_i, b_{i+1}, [T_{i+1}(b_i - f_{i+1})]^{2^{n-2}} + f_{i+1} + 2, [T_i(b_{i-1} - f_i)]^{2^{n-2}} + f_i\}.$$

For two finitely generated graded M and N, it leads directly to estimates for $\mathrm{reg}(\mathrm{Tor}_i^{\mathbf{R}}(M, N))$. Let us state their result on $\mathrm{reg}(\mathrm{Ext}_{\mathbf{R}}^i(M, N))$ [3, Theorem 2.3(2)]. For a finitely generated graded $\mathbf{R}$-module P, set the following notation for the Betti number, initial degree (indeg), and regularity of $\mathrm{Tor}_i^{\mathbf{R}}(P, k)$:

$$\begin{aligned} \beta_i(P) &= \dim_k(\mathrm{Tor}_i^{\mathbf{R}}(P, k)) \\ f_i(P) &= \mathrm{indeg}(\mathrm{Tor}_i^{\mathbf{R}}(P, k)) \\ \mathrm{reg}_i(P) &= \mathrm{reg}(\mathrm{Tor}_i^{\mathbf{R}}(P, k)). \end{aligned}$$

Theorem 9. *Let M and N be finitely generated graded modules over the polynomial ring $\mathbf{R}$. With the notation above, set $T_i = \sum_{p-q=i} \beta_p(M)\beta_q(N)$, $r_M = \mathrm{reg}(M) - \mathrm{indeg}(M)$, $r_N = \mathrm{reg}(N) - \mathrm{indeg}(N)$, and $\delta = \mathrm{indeg}(N) - \mathrm{indeg}(M)$. Then*

$$\mathrm{reg}(\mathrm{Ext}_{\mathbf{R}}^i(M, N)) + i \leq (r_M + r_N + 1)^{2^{n-2}} \max\{T_i, T_{i+1}\}^{2^{n-2}} + 1 - \delta.$$

This establishes the existence of polynomials of $\mathrm{reg}(M)$ and $\mathrm{reg}(N)$ and of the various Betti numbers that bound the regularity of the Exts and Tors.

3 Homological Degree

Introduction

To establish the existence of cohomological degrees in arbitrary dimensions, we describe in some detail one such function introduced in [18].

3.1 Construction

If a and b are integers, we set $\binom{a}{b} = 0$ if $a < b$, and $\binom{a}{b} = 1$ if $a = b$ and possibly negative.

Definition 1. Let M be a finitely generated graded module over the graded algebra $\mathbf{A}$ and $\mathbf{S}$ a Gorenstein graded algebra mapping onto $\mathbf{A}$, with maximal graded ideal $\mathfrak{m}$. Set $\dim \mathbf{S} = r$, $\dim M = d$. The *homological degree* of M is the integer

$$\operatorname{hdeg}(M) = \deg(M) + \sum_{i=r-d+1}^{r} \binom{d-1}{i-r+d-1} \cdot \operatorname{hdeg}(\operatorname{Ext}^i_{\mathbf{S}}(M, \mathbf{S})). \tag{6}$$

This expression becomes more condensed when $\dim M = \dim \mathbf{S} = d > 0$:

$$\operatorname{hdeg}(M) = \deg(M) + \sum_{i=1}^{d} \binom{d-1}{i-1} \cdot \operatorname{hdeg}(\operatorname{Ext}^i_{\mathbf{S}}(M, \mathbf{S})). \tag{7}$$

The definition of hdeg can be extended to any Noetherian local ring $\mathbf{S}$ by setting $\operatorname{hdeg}(M) = \operatorname{hdeg}(\widehat{\mathbf{S}} \otimes_{\mathbf{S}} M)$. On other occasions, we may also assume that the residue field of $\mathbf{S}$ is infinite, an assumption that can be realized by replacing $(\mathbf{S}, \mathfrak{m})$ by the local ring $\mathbf{S}[X]_{\mathfrak{m}\mathbf{S}[X]}$. In fact, if X is any set of indeterminates, the localization is still a Noetherian ring, so the residue field can be assumed to have any needed cardinality, as we shall assume in the proof.

Remark 2. Consider the case when $\dim M = 2$; we assume that $\dim \mathbf{S} = 2$ also. The expression for $\operatorname{hdeg}(M)$ is now

$$\operatorname{hdeg}(M) = \deg(M) + \operatorname{hdeg}(\operatorname{Ext}^1_{\mathbf{S}}(M, \mathbf{S})) + \operatorname{hdeg}(\operatorname{Ext}^2_{\mathbf{S}}(M, \mathbf{S})).$$

The last summand, by duality, is the length of the submodule $\mathrm{H}^0_{\mathfrak{m}}(M)$. The middle term is a module of dimension at most one, so can be described according to Proposition 5 by the equality

$$\operatorname{hdeg}(\operatorname{Ext}^1_{\mathbf{S}}(M, \mathbf{S})) = \deg(\operatorname{Ext}^1_{\mathbf{S}}(\operatorname{Ext}^1_{\mathbf{S}}(M, \mathbf{S}), \mathbf{S})) + \deg(\operatorname{Ext}^2_{\mathbf{S}}(\operatorname{Ext}^1_{\mathbf{S}}(M, \mathbf{S}), \mathbf{S})).$$

There are alternative ways to define the hdeg of an $\mathbf{R}$-module M in a manner that does not require the direct presence of the Gorenstein ring $\mathbf{S}$. For simplicity we assume $\dim \mathbf{S} = \dim M$. If we take for $\mathbf{S}$ the completion of $\mathbf{R}$ and denote by $E = E_{\mathbf{S}}(\mathbf{S}/\mathfrak{m})$ the injective envelope of its residue field, for each integer i define

$$M_i = \operatorname{Hom}_{\mathbf{S}}(\mathrm{H}^i_{\mathfrak{m}}(M), E).$$

Note that M_i is a finitely generated $\mathbf{S}$-module.

We leave to the reader to show that:

Proposition 3. *For* $\mathbf{R}$ *and* M *as above,*

$$\mathrm{hdeg}(M) = \begin{cases} \lambda(M), & \text{if } \dim M = 0 \\ \deg(M) + \sum_{i=0}^{d-1} \binom{d-1}{i} \cdot \mathrm{hdeg}(M_i), & \text{if } \dim M > 0. \end{cases} \tag{8}$$

Theorem 4. *The function* $\mathrm{hdeg}(\cdot)$ *is a cohomological degree.*

The proof requires a special notion of generic hyperplane sections that fits the concept of *genericity* defined earlier. Let $\mathbf{S}$ be a Gorenstein standard graded ring with infinite residue field and M a finitely generated graded module over $\mathbf{S}$. We recall that a *superficial* element of order r for M is an element $z \in S_r$ such that $0:_M z$ is a submodule of M of finite length.

Definition 5. A *special hyperplane section* of M is an element $h \in \mathbf{S}_1$ that is superficial for all the iterated Exts

$$M_{i_1,i_2,\ldots,i_p} = \mathrm{Ext}^{i_1}_{\mathbf{S}}(\mathrm{Ext}^{i_2}_{\mathbf{S}}(\cdots(\mathrm{Ext}^{i_{p-1}}_{\mathbf{S}}(\mathrm{Ext}^{i_p}_{\mathbf{S}}(M,\mathbf{S}),\mathbf{S}),\cdots,\mathbf{S}))),$$

and all sequences of integers $i_1 \geq i_2 \geq \cdots \geq i_p \geq 0$.

By local duality it follows that, up to shifts in grading, there are only finitely many such modules. Actually, it is enough to consider those sequences in which $i_1 \leq \dim S$ and $p \leq 2 \cdot \dim S$, which ensures the existence of such 1-forms as h. It is clear that this property holds for generic hyperplane sections.

The following result establishes $\mathrm{hdeg}(\cdot)$ as a *bona fide* cohomological degree.

Theorem 6. *Let* $\mathbf{S}$ *be a standard Gorenstein graded algebra and* M *a finitely generated graded module of depth at least* 1. *If* $h \in \mathbf{S}_1$ *is a generic hyperplane section on* M, *then*

$$\mathrm{hdeg}(M) \geq \mathrm{hdeg}(M/hM).$$

Proof. This will require several technical reductions. We assume that h is a regular, generic hyperplane section for M that is regular on $\mathbf{S}$. We also assume that $\dim M = \dim \mathbf{S} = d$, and derive several exact sequences from

$$0 \to M \xrightarrow{h} M \longrightarrow N \to 0. \tag{9}$$

For simplicity, we write $M_i = \mathrm{Ext}^i_{\mathbf{S}}(M,\mathbf{S})$, and $N_i = \mathrm{Ext}^{i+1}_{\mathbf{S}}(N,\mathbf{S})$ in the case of N. (The latter because N is a module of dimension $\dim \mathbf{S} - 1$ and $N_i = \mathrm{Ext}^i_{\mathbf{S}/(h)}(N, \mathbf{S}/(h))$.) □

Using this notation, in view of the binomial coefficients in the definition of $\mathrm{hdeg}(\cdot)$, it will be enough to prove:

Lemma 7. *For M and h as above,*

$$\text{hdeg}(N_i) \leq \text{hdeg}(M_i) + \text{hdeg}(M_{i+1}), \text{ for } i \geq 1.$$

Proof. The sequence (9) gives rise to the cohomology long exact sequence

$$0 \to M_0 \longrightarrow M_0 \longrightarrow N_0 \longrightarrow M_1 \longrightarrow M_1 \longrightarrow N_1 \longrightarrow M_2 \longrightarrow \cdots$$
$$\cdots \longrightarrow M_{d-2} \longrightarrow M_{d-2} \longrightarrow N_{d-2} \longrightarrow M_{d-1} \longrightarrow M_{d-1} \longrightarrow N_{d-1} \to 0,$$

which are broken up into shorter exact sequences as follows:

$$0 \to L_i \longrightarrow M_i \longrightarrow \widetilde{M}_i \to 0 \tag{10}$$
$$0 \to \widetilde{M}_i \longrightarrow M_i \longrightarrow G_i \to 0 \tag{11}$$
$$0 \to G_i \longrightarrow N_i \longrightarrow L_{i+1} \to 0. \tag{12}$$

□

We note that all L_i have finite length, because of the condition on h. For $i = 0$, we have the usual relation $\deg(M) = \deg(N)$. When $\widetilde{M}_i$ has finite length, then M_i, G_i, and N_i have finite length, and

$$\begin{aligned}\text{hdeg}(N_i) = \lambda(N_i) &= \lambda(G_i) + \lambda(L_{i+1})\\ &\leq \text{hdeg}(M_i) + \text{hdeg}(M_{i+1}).\end{aligned}$$

It is a similar relation that we want to establish for all other cases.

Proposition 8. *Let* **S** *be a Gorenstein graded algebra and let*

$$0 \to A \longrightarrow B \longrightarrow C \to 0$$

be an exact sequence of graded modules. Then

(a) If A is a module of finite length, then

$$\text{hdeg}(B) = \text{hdeg}(A) + \text{hdeg}(C).$$

(b) If C is a module of finite length, then

$$\text{hdeg}(B) \leq \text{hdeg}(A) + \text{hdeg}(C).$$

(c) Moreover, in the previous case, if $\dim B = d$, *then*

$$\text{hdeg}(A) \leq \text{hdeg}(B) + (d-1)\text{hdeg}(C).$$

(d) If C is a module of finite length and depth $B \geq 2$*, then*

$$\text{hdeg}(A) = \text{hdeg}(B) + \text{hdeg}(C).$$

Proof. They are all clear if B is a module of finite length so we assume that $\dim B = d \geq 1$.

(a) This is immediate since $\deg(B) = \deg(C)$ and the cohomology sequence gives

$$\begin{aligned}\text{Ext}^i_{\mathbf{S}}(B,\mathbf{S}) &= \text{Ext}^i_{\mathbf{S}}(C,\mathbf{S}),\ 1 \leq i \leq d-1,\ \text{and}\\ \lambda(\text{Ext}^d_{\mathbf{S}}(B,\mathbf{S})) &= \lambda(\text{Ext}^d_{\mathbf{S}}(A,\mathbf{S})) + \lambda(\text{Ext}^d_{\mathbf{S}}(C,\mathbf{S})).\end{aligned}$$

(b) Similarly we have

$$\text{Ext}^i_{\mathbf{S}}(B,\mathbf{S}) = \text{Ext}^i_{\mathbf{S}}(A,\mathbf{S}),\ 1 \leq i < d-1,$$

and the exact sequence

$$0 \to \text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S}) \to \text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S}) \to \text{Ext}^d_{\mathbf{S}}(C,\mathbf{S}) \to \text{Ext}^d_{\mathbf{S}}(B,\mathbf{S}) \to \text{Ext}^d_{\mathbf{S}}(A,\mathbf{S}) \to 0. \tag{13}$$

If $\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S})$ has finite length, then

$$\begin{aligned}\text{hdeg}(\text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S})) &\leq \text{hdeg}(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S}))\\ \text{hdeg}(\text{Ext}^{d}_{\mathbf{S}}(B,\mathbf{S})) &\leq \text{hdeg}(\text{Ext}^{d}_{\mathbf{S}}(A,\mathbf{S})) + \text{hdeg}(\text{Ext}^{d}_{\mathbf{S}}(C,\mathbf{S})).\end{aligned}$$

Otherwise, $\dim \text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S}) = 1$, and

$$\text{hdeg}(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S})) = \deg(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S})) + \lambda(\text{H}^0_{\mathfrak{m}}(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S}))).$$

Since we also have

$$\begin{aligned}\deg(\text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S})) &= \deg(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S})),\\ \lambda(\text{H}^0_{\mathfrak{m}}(\text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S}))) &\leq \lambda(\text{H}^0_{\mathfrak{m}}(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S}))),\end{aligned}$$

we again obtain the stated bound.

(c) In the sequence (13), if $\dim \text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S}) = 0$, then

$$\lambda(\text{Ext}^{d-1}_{\mathbf{S}}(A,\mathbf{S})) \leq \lambda(\text{Ext}^{d-1}_{\mathbf{S}}(B,\mathbf{S})) + \lambda(C), \tag{14}$$

and also

$$\lambda(\text{Ext}^{d}_{\mathbf{S}}(A,\mathbf{S})) \leq \lambda(\text{Ext}^{d}_{\mathbf{S}}(B,\mathbf{S})).$$

When taken into the formula for $\text{hdeg}(A)$, the binomial coefficient $\binom{d-1}{d-2}$ gives the desired factor for $\lambda(C)$.

On the other hand, if $\dim \mathrm{Ext}_{\mathbf{S}}^{d-1}(B,\mathbf{S}) = 1$, we also have

$$\mathrm{hdeg}(\mathrm{Ext}_{\mathbf{S}}^{d-1}(A,\mathbf{S})) \leq \mathrm{hdeg}(\mathrm{Ext}_{\mathbf{S}}^{d-1}(B,\mathbf{S})) + \lambda(C),$$

the dimension one case of (14).

(d) This follows by applying the definition of hdeg to the exact sequence.

□

Suppose that $\dim \widetilde{M_i} \geq 1$. From Proposition 8(b) we have

$$\mathrm{hdeg}(N_i) \leq \mathrm{hdeg}(G_i) + \lambda(L_{i+1}). \tag{15}$$

We must now relate $\mathrm{hdeg}(G_i)$ to $\deg(M_i)$. Apply the functor $\mathrm{H}^0_{\mathfrak{m}}(\cdot)$ to the sequence (11) and consider the commutative diagram

$$\begin{array}{ccccccc} 0 \to & \widetilde{M}_i & \longrightarrow & M_i & \longrightarrow & G_i & \to 0 \\ & \uparrow & & \uparrow & & \uparrow & \\ 0 \to & \mathrm{H}^0_{\mathfrak{m}}(\widetilde{M}_i) & \longrightarrow & \mathrm{H}^0_{\mathfrak{m}}(M_i) & \longrightarrow & \mathrm{H}^0_{\mathfrak{m}}(G_i) & \end{array},$$

in which we denote by H_i the image of the natural map

$$\mathrm{H}^0_{\mathfrak{m}}(M_i) \longrightarrow \mathrm{H}^0_{\mathfrak{m}}(G_i).$$

Through the snake lemma, we obtain the exact sequence

$$0 \to \widetilde{M_i}/\mathrm{H}^0_{\mathfrak{m}}(\widetilde{M}_i) \xrightarrow{\alpha} M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i) \longrightarrow G_i/H_i \to 0. \tag{16}$$

Furthermore, from (10) there is a natural isomorphism

$$\beta : M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i) \cong \widetilde{M_i}/\mathrm{H}^0_{\mathfrak{m}}(\widetilde{M}_i),$$

while from (11) there is a natural injection

$$\widetilde{M_i}/\mathrm{H}^0_{\mathfrak{m}}(\widetilde{M}_i) \hookrightarrow M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i),$$

whose composite with β is induced by multiplication by h on $M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i)$. We may thus replace $\widetilde{M_i}/\mathrm{H}^0_{\mathfrak{m}}(\widetilde{M}_i)$ by $M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i)$ in (16) and take α as multiplication by h:

$$0 \to M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i) \xrightarrow{h} M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i) \longrightarrow G_i/H_i \to 0.$$

Observe that since

$$\mathrm{Ext}_{\mathbf{S}}^{j}(M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i),\mathbf{S}) = \mathrm{Ext}_{\mathbf{S}}^{j}(M_i,\mathbf{S}),\ \ j < \dim S,$$

h is still a regular, generic hyperplane section for $M_i/\mathrm{H}^0_{\mathfrak{m}}(M_i)$. By induction on the dimension of the module, we have

$$\text{hdeg}(M_i/\text{H}^0_{\mathfrak{m}}(M_i)) \geq \text{hdeg}(G_i/H_i).$$

Now from Proposition 8(a), we have

$$\text{hdeg}(G_i) = \text{hdeg}(G_i/H_i) + \lambda(H_i).$$

Since these summands are bounded, by $\text{hdeg}(M_i/\text{H}^0_{\mathfrak{m}}(M_i))$ and $\lambda(\text{H}^0_{\mathfrak{m}}(M_i))$, respectively (in fact, $\lambda(H_i) = \lambda(L_i)$) we have

$$\text{hdeg}(G_i) \leq \text{hdeg}(M_i/\text{H}^0_{\mathfrak{m}}(M_i)) + \lambda(\text{H}^0_{\mathfrak{m}}(M_i)) = \text{hdeg}(M_i),$$

the last equality by Proposition 8(a) again. Finally, taking this estimate into (15), we get

$$\begin{aligned} \text{hdeg}(N_i) &\leq \text{hdeg}(G_i) + \lambda(L_{i+1}) \\ &\leq \text{hdeg}(M_i) + \text{hdeg}(M_{i+1}) \end{aligned} \tag{17}$$

to establish the claim. □

Remark 9. That equality does not always hold is shown by the following example. Suppose that $\mathbf{R} = k[x, y]$ and $M = (x, y)^2$. Then $\text{hdeg}(M) = 4$, but $\text{hdeg}(M/hM) = 3$ for any hyperplane section h. To get an example of a ring one takes the idealization of M.

Remark 10. It should be emphasized that the *weight* binomial coefficients in the definition of hdeg were chosen to enable the Bertini property at the expense of its behavior on other short exact sequences. Suppose $\mathbf{R} = k[x, y, z]$, and $M = R \oplus R/(x, y)$. Then

$$\text{hdeg}(M) = \deg(\mathbf{R}) + \binom{3-1}{2-1}\deg(\mathbf{R}/(x, y)) = 1 + 2 = 3.$$

Example 11. This example shows how the function hdeg captures important aspects of the structure of the module. We recall the notion of a *sequentially* Cohen–Macaulay module. This is a module M having a filtration

$$0 = M_0 \subset M_1 \subset \cdots \subset M_r = M,$$

with the property that each factor M_i/M_{i-1} is Cohen–Macaulay and

$$\dim M_i/M_{i-1} < \dim M_{i+1}/M_i, \ i = 1, \ldots, r-1.$$

If L is the leftmost nonzero submodule in such a chain, it follows easily that $N = M/L$ is also sequentially Cohen–Macaulay (L is Cohen–Macaulay by hypothesis) and we have

$$\mathrm{Ext}^p_{\mathbf{S}}(M, \mathbf{S}) = \mathrm{Ext}^p_{\mathbf{S}}(L, \mathbf{S}),\ p = \dim L$$
$$\mathrm{Ext}^i_{\mathbf{S}}(M, \mathbf{S}) = \mathrm{Ext}^i_{\mathbf{S}}(N, \mathbf{S}),\ i < p$$

with the other Ext's vanishing. These properties and induction lead to a formula for $\mathrm{hdeg}(M)$ in terms of the $\deg(M_{i+1}/M_i)$:

- If $d = \dim M$ note that $r \leq d$. Padding the filtration with repeated terms (giving rise to trivial factors) we may assume $r = d$.
- For each i we have a short exact sequence

$$0 \to M_i/M_{i-1} \longrightarrow M/M_{i-1} \longrightarrow M/M_i \to 0,$$

  where M_i/M_{i-1} is Cohen–Macaulay of dimension i, or it is (0), while M/M_i, by induction, is a module of depth at least $i+1$. In particular $\mathrm{Ext}^{d-i}_{\mathbf{S}}(M/M_{i-1}, \mathbf{S}) = \mathrm{Ext}^{d-i}_{\mathbf{S}}(M_i/M_{i-1}, \mathbf{S})$ and $\mathrm{Ext}^j_{\mathbf{S}}(M/M_{i-1}, \mathbf{S}) = 0$ for $j \geq d - i + 1$.
- Descending induction will lead to $\mathrm{Ext}^{d-i}_{\mathbf{S}}(M, \mathbf{S}) = \mathrm{Ext}^{d-i}_{\mathbf{S}}(M_i/M_{i-1}, \mathbf{S})$.

$$\begin{aligned}
\mathrm{hdeg}(M) &= \deg(M) + \sum_{i=1}^{d} \binom{d-1}{i-1} \cdot \mathrm{hdeg}(\mathrm{Ext}^i_{\mathbf{S}}(M, \mathbf{S})) \\
&= \deg(M) + \sum_{i=1}^{d} \binom{d-1}{i-1} \cdot \mathrm{hdeg}(\mathrm{Ext}^i_{\mathbf{S}}(M_{d-i}/M_{d-i-1}, \mathbf{S})) \\
&= \deg(M) + \sum_{i=1}^{d} \binom{d-1}{i-1} \cdot \deg(M_{d-i}/M_{d-i-1}).
\end{aligned}$$

3.2 Buchsbaum Modules

The expression for the function $\mathrm{hdeg}(\cdot)$ arrives in known territory if M is a generalized Cohen–Macaulay module.

Definition 12. Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and M a finitely generated $\mathbf{R}$-module. M is a *generalized Cohen–Macaulay module* if $M_{\mathfrak{p}}$ is Cohen–Macaulay for all prime ideals $\mathfrak{p} \neq \mathfrak{m}$.

A typical characterization is:

Theorem 13. *Let* $(\mathbf{R}, \mathfrak{m})$ *be a Noetherian local ring and* M *a finitely generated* $\mathbf{R}$*-module.* M *is a generalized Cohen–Macaulay module if and only if the modules* $\mathrm{H}^i_{\mathfrak{m}}(M)$ *are finitely generated for all* $i < \dim M$.

For these modules the expression hdeg(M), by local duality, converts into:

Proposition 14. *If M is a generalized Cohen–Macaulay module of dimension d, then*

$$\operatorname{hdeg}(M) = \deg(M) + \sum_{i=0}^{d-1} \binom{d-1}{i} \cdot \lambda(\mathrm{H}^i_{\mathfrak{m}}(M)).$$

Buchsbaum rings and modules were introduced by Stückrad and Vogel. They are studied in great detail in [17].

Definition 15. Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and M a finitely generated $\mathbf{R}$-module. M is a *Buchsbaum module* if the partial Euler characteristic, $\chi_1(\mathbf{x}; M)$, is independent of the system of parameters $\mathbf{x}$ for M. In other words, $\lambda(M/(\mathbf{x})M) - \deg_{\mathbf{x}}(M)$ is independent of $\mathbf{x}$. This integer is called the *Buchsbaum invariant* of M.

Two of its properties are:

Theorem 16. *Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and M a finitely generated $\mathbf{R}$-module. Then*

1. *M is a Buchsbaum module if and only if every system of parameters for M is a d-sequence relative to M.*
2. *If M is a Buchsbaum module then $\mathrm{H}^i_{\mathfrak{m}}(M)$ are $\mathbf{R}/\mathfrak{m}$-vector spaces for $i < \dim M$. (In particular they are generalized Cohen–Macaulay modules.) The converse does not hold true.*

Corollary 17. *If M is a Buchsbaum module, then*

$$\operatorname{hdeg}(M) = \deg(M) + I(M),$$

where $I(M)$ is the Buchsbaum invariant of M.

An effective test for the property was found by Yamagishi [20].

Remark 18. A related theme is how to decide whether an $\mathbf{R}$-module M is Cohen–Macaulay on the punctured spectrum. If $\mathbf{R}$ is a regular local ring and the annihilator of M is equidimensional a criterion can be cast as follows. Suppose $\operatorname{codim} M = r$ and

$$\cdots \longrightarrow F_{r+1} \stackrel{\varphi}{\longrightarrow} F_r \longrightarrow \cdots$$

is a projective resolution of M. For $\mathfrak{p} \neq \mathfrak{m}$, $M_{\mathfrak{p}}$ is Cohen–Macaulay if and only if $\operatorname{proj\,dim}_{\mathbf{R}_{\mathfrak{p}}} M_{\mathfrak{p}} = r$, a condition equivalent to saying that the image of $\varphi_{\mathfrak{p}}$ splits off $(F_r)_{\mathfrak{p}}$. In other words the ideal $I(\varphi)$ of maximal minors of φ is $\mathfrak{m}$-primary.

It would be less cumbersome to find a more amenable module and make use of the rigidity of Tor since $\mathbf{R}$ is a regular local ring. Let us sketch this with a non-amenable module! The following is an equivalence:

$$M \text{ is free on the punctured spectrum} \Leftrightarrow \mathrm{Tor}^{\mathbf{R}}_{r+1}(M, M) \text{ has finite support.}$$

3.3 Cohomological Degrees and Samuel Multiplicities

There are variations of cohomological degree functions obtained by using Samuel's notion of multiplicity.

Definition 19. Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and I an $\mathfrak{m}$-primary ideal. We denote by $\mathrm{hdeg}_I(\cdot)$ the function obtained by replacing, in the definition of $\mathrm{hdeg}(\cdot)$, $\deg(M)$ by Samuel's $\mathrm{e}(I; M)$.

Let us make a rough comparison between $\mathrm{hdeg}(M)$ and $\mathrm{hdeg}_I(M)$.

Proposition 20. *Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and let I be an $\mathfrak{m}$-primary ideal. Suppose $\mathfrak{m}^r \subset I$. If M is an $\mathbf{R}$-module of dimension d, then*

$$\mathrm{hdeg}_I(M) \leq r^d \cdot \deg(M) + r^{d-1} \cdot (\mathrm{hdeg}(M) - \deg(M)).$$

Proof. If r is the index of nilpotency of $\mathbf{R}/I$, for any $\mathbf{R}$-module L of dimension s,

$$\lambda(L/(\mathfrak{m}^r)^n L) \geq \lambda(L/I^n L).$$

The Hilbert polynomial of L gives

$$\lambda(L/(\mathfrak{m}^r)^n L) = \deg(M)\frac{r^s}{s!}n^s + \text{lower terms}.$$

We now apply this estimate to the definition of $\mathrm{hdeg}(M)$, taking into account that its terms are evaluated at modules of decreasing dimension. □

For later reference we rephrase Proposition 32 for use with Samuel's multiplicities.

Theorem 21. *Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring, let I be an $\mathfrak{m}$-primary ideal, and let M be a finitely generated $\mathbf{R}$-module of dimension $d \geq 1$. Let $\mathbf{x} = \{x_1, \ldots, x_r\}$ be a superficial sequence in I relative to M and* hdeg_I. *Then*

$$\mathrm{hdeg}_I(M/(\mathbf{x})M) \leq \mathrm{hdeg}_I(M).$$

Moreover, if $r < d$ then

$$\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/(\mathbf{x})M)) \leq \mathrm{hdeg}_I(M) - \mathrm{e}(I; M).$$

Corollary 22. *Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring of dimension $d > 0$ and infinite residue field. For an $\mathfrak{m}$-primary I there is a minimal reduction J such that*

$$\lambda(\mathbf{R}/J) \leq \mathrm{hdeg}_I(\mathbf{R}).$$

In particular, the set of all $\lambda(\mathbf{R}/J)$ for all parameter ideals with the same integral closure is bounded.

3.4 Homological Torsion

There are other combinatorial expressions of the terms defining $\text{hdeg}(M)$, $\text{hdeg}(\text{Ext}^i_{\mathbf{S}}(M,\mathbf{S}))$, that behave well under hyperplane section. It will turn out to be useful in certain characterizations of generalized Cohen–Macaulay modules and in the study of Hilbert coefficients.

Definition 23. Let $\mathbf{S}$ be a ring of dimension r and M an $\mathbf{S}$-module of dimension $d \geq 2$. Its *homological torsion* is the integer

$$\mathbf{T}(M) = \sum_{i=r-d+1}^{r-1} \binom{r-2}{i-d+r-1} \cdot \text{hdeg}(\text{Ext}^i_{\mathbf{S}}(M,\mathbf{S})).$$

If I is an $\mathfrak{m}$-primary, the integer obtained using hdeg_I will be denoted $\mathbf{T}_I(M)$.

If $r = d$, this formula becomes

$$\mathbf{T}(M) = \sum_{i=1}^{d-1} \binom{d-2}{i-1} \cdot \text{hdeg}(\text{Ext}^i_{\mathbf{S}}(M,\mathbf{S})).$$

If $d = 2$, $\mathbf{T}(M) = \text{hdeg}(\text{Ext}^1_{\mathbf{S}}(M,\mathbf{S})) = \deg(\text{Ext}^1_{\mathbf{S}}(M,\mathbf{S})) + \deg(\text{Ext}^2_{\mathbf{S}}(\text{Ext}^1_{\mathbf{S}}(M,\mathbf{S}),\mathbf{S}))$.

The restriction to $d \geq 2$ explanation: $\lambda(M)$ if $d = 0$, or $\lambda(\text{H}^0_{\mathfrak{m}}(M))$ if $d = 1$ are quantities we liken to the ordinary torsion.

Example 24. If M is a generalized Cohen–Macaulay module of dimension $d \geq 2$,

$$\mathbf{T}_I(M) \leq \sum_{i=1}^{d-1} \binom{d-2}{i-1} \lambda(\text{H}^i_{\mathfrak{m}}(M)),$$

with equality if M is Buchsbaum.

Theorem 25. *Let M be a module of dimension $d \geq 3$ and let h be a generic hyperplane section. Then $\mathbf{T}(M/hM) \leq \mathbf{T}(M)$.*

Proof. We use induction on d. If $d \geq 3$, using the notation of Lemma 7, in particular setting $N = M/hM$, we have

$$\mathbf{T}(N) = \sum_{i=1}^{d-1} \binom{d-2}{i-1} \cdot \text{hdeg}(\text{Ext}^i_{\mathbf{S}}(N,\mathbf{S})).$$

By Lemma 7,

$$\text{hdeg}(\text{Ext}^i_{\mathbf{S}}(N,\mathbf{S})) \leq \text{hdeg}(\text{Ext}^i_{\mathbf{S}}(M,\mathbf{S})) + \text{hdeg}(\text{Ext}^{i+1}_{\mathbf{S}}(N,\mathbf{S})),$$

which by rearranging gives $\mathbf{T}(N) \leq \mathbf{T}(M)$, as desired. $\square$

Higher order *homological torsion* functions can be similarly defined. For this reason it is appropriate to write $\mathbf{T}^{(1)}(M) := \mathbf{T}(M)$.

Definition 26. Let M be an $\mathbf{S}$-module of dimension $d \geq 3$. Its *second-order homological torsion* is the integer

$$\mathbf{T}^{(2)}(M) = \sum_{i=1}^{d-2} \binom{d-3}{i-1} \cdot \operatorname{hdeg}(\operatorname{Ext}^i_{\mathbf{S}}(M, \mathbf{S})).$$

If I is an $\mathfrak{m}$-primary, the integer obtained using hdeg_I will be denoted $\mathbf{T}^{(2)}_I(M)$.

More generally, let M be a module of dimension d. For each $j \leq d$, set

$$\mathbf{T}^{(j)}(M) = \sum_{i=0}^{d-j} \binom{d-j-1}{i-1} \cdot \operatorname{hdeg}(\operatorname{Ext}^i_{\mathbf{S}}(M, \mathbf{S})).$$

Theorem 27. *Let M be a module of dimension $d \geq 4$ and let h be a generic hyperplane section. Then $\mathbf{T}^{(2)}(M/hM) \leq \mathbf{T}^{(2)}(M)$. A similar assertion holds for all $\mathbf{T}^{(j)}$.*

Corollary 28. *Let M be a module of dimension $r = d + 1 \geq 3$. Then*

$$\operatorname{hdeg}(M) > \mathbf{T}^{(1)}(M) \geq \mathbf{T}^{(2)}(M) \geq \cdots \geq \mathbf{T}^{(d)}(M).$$

Remark 29. There are a number of rigidity questions about the values of the Hilbert coefficients $e_i(I)$. Typically they have the form

$$|e_1(I)| \geq |e_2(I)| \geq \cdots \geq |e_d(I)|.$$

Two of these cases are (i) parameter ideals and (ii) normal ideals, or more generally the case of the normalized filtration. Because values of $\mathbf{T}^{(i)}_I(\mathbf{R})$ have been used, in a few cases, to bound the $e_i(I)$, the descending chain in Corollary 28 argues for an underlying rigidity. This runs counter to the known formulas for the values of the $e_i(I)$ for general ideals (e.g., [16, Theorem 4.1]).

Generalized Cohen–Macaulay Modules

These modules can be characterized in terms of their homological torsions.

Proposition 30. *Let $(\mathbf{R}, \mathfrak{m})$ be a Gorenstein local ring of dimension $d \geq 2$ and M a finitely generated $\mathbf{R}$-module. Then M is a generalized Cohen–Macaulay module if and only if $T_I(M)$ is bounded for every $\mathfrak{m}$-primary ideal I. More precisely, if $I = (x_1, \ldots, x_d)$ is a parameter ideal, and $I_n = (x_1^n, \ldots, x_d^n)$, it suffices that all $T_{I_n}(M)$ be bounded.*

Proof. If $d = 2$, $T(M) = \text{hdeg}(\text{Ext}^1_{\mathbf{R}}(M, \mathbf{R}))$. This is a module of dimension at most one. If the dimension is 1, $\deg_{I_n}(\text{Ext}^1_{\mathbf{R}}(M, \mathbf{R})$ is a polynomial of degree 1 for large n.

If $d \geq 3$, let h be a generic hyperplane section. It will suffice to show that M/hM is a generalized Cohen–Macaulay module. For that we can use Theorem 25 and induction.

□

3.5 Castelnuovo–Mumford Regularity and Homological Degree

The following result of Chardin, Ha, and Hoa [3] provides a polynomial bound for the homological degree in terms of the Castelnuovo–Mumford regularity. (For $d = 2$, a similar bound had been found by Gunston in his thesis [9].)

Theorem 31. *Let M be a nonzero finitely generated graded $\mathbf{R}$-module of dimension $d > 0$. Denote by n the minimal number of generators of M and by $\alpha(M)$ the maximal degree of the a minimal set of homogeneous generators. Then*

$$\text{hdeg}(M) \leq \left[n \binom{\text{reg}(M) - \alpha(M) + n}{n} \right]^{2^{(d-1)^2}}. \tag{18}$$

3.6 Specialization and Torsion

One of the uses of extended degrees is the following. Let M be a module and $\mathbf{x} = \{x_1, \ldots, x_r\}$ be a superficial sequence for the module M relative to an extended degree Deg. How to estimate the length of $\text{H}^0_{\mathfrak{m}}(M)$ in terms of the initial data of M?

Let us consider the case of $r = 1$. Let $H = \text{H}^0_{\mathfrak{m}}(M)$ and write

$$0 \to H \longrightarrow M \longrightarrow M' \to 0. \tag{19}$$

Reduction modulo x_1 gives the exact sequence

$$0 \to H/x_1H \longrightarrow M/x_1M \longrightarrow M'/x_1M' \to 0. \tag{20}$$

From the first sequence we have $\text{Deg}(M) = \text{Deg}(H) + \text{Deg}(M')$, and from the second

$$\text{Deg}(M/x_1M) - \text{Deg}(H/x_1H) = \text{Deg}(M'/x_1M') \leq \text{Deg}(M').$$

Taking local cohomology of the second exact sequence yields the short exact sequence

$$0 \to H/x_1H \longrightarrow \mathrm{H}^0_{\mathfrak{m}}(M/x_1M) \longrightarrow \mathrm{H}^0_{\mathfrak{m}}(M'/x_1M') \to 0,$$

from which we have the estimation

$$\begin{aligned}\mathrm{Deg}(\mathrm{H}^0_{\mathfrak{m}}(M/x_1M)) &= \mathrm{Deg}(H/x_1H) + \mathrm{Deg}(\mathrm{H}^0_{\mathfrak{m}}(M'/x_1M'))\\ &\leq \mathrm{Deg}(H/x_1H) + \mathrm{Deg}(M'/x_1M')\\ &\leq \mathrm{Deg}(H) + \mathrm{Deg}(M') = \mathrm{Deg}(M).\end{aligned}$$

We resume these observations as:

Proposition 32. *Let M be a module and let $\{x_1, \ldots, x_r\}$ be a superficial sequence relative to M and* Deg. *Then*

$$\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/(x_1, \ldots, x_r)M)) \leq \mathrm{Deg}(M).$$

Now we derive a more precise formula using hdeg. It will be of use later.

Theorem 33. *Let M be a module of dimension $d \geq 2$ and let $\mathbf{x} = \{x_1, \ldots, x_{d-1}\}$ be a superficial sequence for M and* hdeg. *Then*

$$\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/(\mathbf{x})M)) \leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)) + \mathbf{T}(M).$$

Proof. Consider the exact sequence

$$0 \to H = \mathrm{H}^0_{\mathfrak{m}}(M) \longrightarrow M \longrightarrow M' \to 0.$$

We have $\mathrm{Ext}^i_{\mathbf{S}}(M, \mathbf{S}) = \mathrm{Ext}^i_{\mathbf{S}}(M', \mathbf{S})$ for $d > i \geq 0$, and therefore $T(M) = T(M')$. On the other hand, reduction mod $\mathbf{x}$ gives

$$\begin{aligned}\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/(\mathbf{x})M)) &\leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M'/(\mathbf{x})M')) + \lambda(H/(\mathbf{x})H)\\ &\leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M'/(\mathbf{x})M')) + \lambda(H),\end{aligned}$$

which shows that it is enough to prove the assertion for M'.

If $d > 2$, we apply Theorem 25, to pass to M'/x_1M'. This reduces all the way to the case $d = 2$. Let M be a module of positive depth. Write $h = \mathbf{x}$. The assertion requires that $\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM)) \leq \mathrm{hdeg}(\mathrm{Ext}^1_{\mathbf{S}}(M, \mathbf{S}))$. We have the cohomology exact sequence

$$\mathrm{Ext}^1_{\mathbf{S}}(M, \mathbf{S}) \xrightarrow{h} \mathrm{Ext}^1_{\mathbf{S}}(M, \mathbf{S}) \longrightarrow \mathrm{Ext}^2_{\mathbf{S}}(M/hM, \mathbf{S}) \longrightarrow \mathrm{Ext}^2_{\mathbf{S}}(M, \mathbf{S}) = 0,$$

where
$$\lambda(\mathrm{H}^0_{\mathfrak{m}}(M/hM)) = \mathrm{hdeg}(\mathrm{Ext}^2_{\mathbf{S}}(M/hM, \mathbf{S})).$$

If $\mathrm{Ext}^1_{\mathbf{S}}(M, \mathbf{S})$ has finite length the assertion is clear. Otherwise $L = \mathrm{Ext}^1_{\mathbf{S}}(M, \mathbf{S})$ is a module of dimension 1 over a discrete valuation domain V with h for its parameter. By the fundamental theorem for such modules,

$$V = V^r \oplus (\bigoplus_{j=1}^{s} V/h^{e_j}V),$$

so that multiplication by h gives

$$\lambda(L/hL) = r + s \leq r + \sum_{j=1}^{s} e_j = \mathrm{hdeg}(L).$$

An alternative argument at this point is to consider the exact sequence (we may assume $\dim \mathbf{S} = 1$)

$$0 \to L_0 \longrightarrow L \stackrel{h}{\longrightarrow} L \longrightarrow L/hL \to 0,$$

where both L_0 and L/hL have finite length. If F denotes the image of the multiplication by h on L, we have the exact sequences $0 \to L_0 \longrightarrow L \longrightarrow F \to 0$ and $0 \to F \longrightarrow L \longrightarrow L/hL \to 0$. Dualizing we have $\mathrm{Hom}_{\mathbf{S}}(L, \mathbf{S}) = \mathrm{Hom}_{\mathbf{S}}(F, \mathbf{S})$ and the exact sequence

$$0 \to \mathrm{Hom}_{\mathbf{S}}(L, \mathbf{S}) \stackrel{h}{\longrightarrow} \mathrm{Hom}_{\mathbf{S}}(L, \mathbf{S}) \longrightarrow \mathrm{Ext}^1_{\mathbf{S}}(L/hL, \mathbf{S}) \longrightarrow \mathrm{Ext}^1_{\mathbf{S}}(F, \mathbf{S}),$$

which shows that

$$\lambda(L/hL) \leq \deg(L) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(F)) \leq \deg(L) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(L)) = \mathrm{hdeg}(L),$$

as desired. □

4 Bdeg: The Extreme Cohomological Degree

Introduction

The content of this section is mostly the work of Tor Gunston and Kia Dalili. In his thesis [9], Gunston introduced the following cohomological degree:

Definition 1. Let **R** be a Noetherian local ring with infinite residue field or a standard graded algebra over an infinite field. For a finitely generated **R**-module (graded if required) M, $\mathrm{bdeg}(M)$ is the integer

$$\mathrm{bdeg}(M) = \min\{\mathrm{Deg}(M) \mid \mathrm{Deg} \text{ is a cohomological degree function}\}.$$

This function is well-defined since cohomological degrees exist (e.g., hdeg). It is obviously a cohomological degree itself. Gunston proved that [9, Theorem 3.1.3]):

Theorem 2 (Gunston). *Suppose that M is a finitely generated* **R***-module of positive depth. Then for a generic hyperplane section h,*

$$\mathrm{bdeg}(M) = \mathrm{bdeg}(M/hM).$$

Corollary 3. *If M is a finitely generated* **R***-module of dimension d there is a generic superficial sequence $\mathbf{x} = \{x_1, \ldots, x_d\}$ for M such that*

$$\mathrm{bdeg}(M) \leq \lambda(\mathrm{H}^0_{\mathfrak{m}}(M)) + \sum_{i=1}^{d} \lambda(\mathrm{H}^0_{\mathfrak{m}}(M/(x_1, \ldots, x_i)M)) \leq (d+1)\mathrm{hdeg}(M) - d\,\deg(M).$$

Proof. The first inequality is a direct consequence of Theorem 2, the second from Theorem 21. □

4.1 *Rules of Computation of* bdeg

One of the difficulties of computing cohomological degrees lies on their behavior on short exact sequences. The degree bdeg has more amenable properties, according to the following rules.

Proposition 4. *Let* **R** *be a standard graded algebra and A a finitely generated graded* **R***-module with Hilbert function $H_A(\mathbf{t})$. Then*

$$\mathrm{bdeg}(A) \leq \sum_{j=\alpha(A)}^{\mathrm{reg}(A)} H_A(j).$$

Proof. Set $L = \mathrm{H}^0_{\mathfrak{m}}(A)$ and consider the exact sequence

$$0 \to L \longrightarrow A \longrightarrow A' = A/L \to 0.$$

We may assume $\dim A \geq 1$. Let h be a generic hyperplane section. The proof follows by induction from the following inequalities:

$$\begin{aligned}
\operatorname{reg}(A') &= \operatorname{reg}(A'/hA'),\\
\alpha(A) &\geq \alpha(A'/hA'),\\
H_A(j) &= H_L(j) + H_{A'}(j) \geq H_L(j) + H_{A'/hA'}(j),\\
\operatorname{bdeg}(A) &= \lambda(L) + \operatorname{bdeg}(A') = \lambda(L) + \operatorname{bdeg}(A'/hA'),\\
\operatorname{reg}(A) &= \max\{\operatorname{reg}(L), \operatorname{reg}(A')\} = \max\{\operatorname{reg}(L), \operatorname{reg}(A'/hA')\}.
\end{aligned}$$

□

Proposition 5 (Gunston [9, Proposition 3.2.2]). *Suppose A, B, and C are finitely generated* $\mathbf{R}$*-modules and*

$$0 \to A \longrightarrow B \longrightarrow C \to 0$$

is an exact sequence. If $\lambda(C) < \infty$, then

$$\operatorname{bdeg}(B) \leq \operatorname{bdeg}(A) + \lambda(\mathrm{H}^0_{\mathfrak{m}}(B)) - \lambda(\mathrm{H}^0_{\mathfrak{m}}(A)).$$

In particular, we have

$$\operatorname{bdeg}(B) \leq \operatorname{bdeg}(A) + \lambda(C), \quad \text{and if depth} B > 0, \text{ then } \operatorname{bdeg}(B) \leq \operatorname{bdeg}(A).$$

A similar calculation in [4, Proposition 3.2] then leads to:

Proposition 6. *Suppose A, B, and C are finitely generated* $\mathbf{R}$*-modules and*

$$0 \to A \longrightarrow B \longrightarrow C \to 0$$

is an exact sequence. Then

1. $\operatorname{bdeg}(B) \leq \operatorname{bdeg}(A) + \operatorname{bdeg}(C)$;
2. $\operatorname{bdeg}(A) \leq \operatorname{bdeg}(B) + (\dim A - 1)\operatorname{bdeg}(C)$.

Proof. We prove only (1). The case where A has finite length is a consequence of bdeg being an extended degree. For the other cases, one makes use of indiction on the dimension, repeated use of snake lemma, and the following special case.

Lemma 7. *Suppose A, B, and C are finitely generated* $\mathbf{R}$*-modules and*

$$0 \to A \longrightarrow B \longrightarrow C \to 0$$

is an exact sequence. If C has finite length then

$$\operatorname{bdeg}(B) \leq \operatorname{bdeg}(A) + \operatorname{bdeg}(C).$$

Proof. For a finitely generated **R**-module, denote by A_0 its submodule of finite support; set $A' = A/A_0$. If $A = A_0$, the assertion follows from the additivity of the length function. Clearly we may replace A by A/A_0 and B by B/A_0; thus changing notation we may assume that A has positive depth. It follows that B_0 embeds in C. We may now replace B by B/B_0 and C by C/B_0.

Changing notation again, we may assume that both A and B have positive depth. Let h be an appropriate hyperplane section for A and B, that is, $\operatorname{bdeg}(A) = \operatorname{bdeg}(A/hA)$ and $\operatorname{bdeg}(B) = \operatorname{bdeg}(B/hB)$. Reduction mod h gives the exact sequence

$$0 \to E \longrightarrow A/hA \longrightarrow B/hB \longrightarrow C/hC \to 0,$$

and thus by induction (E has finite length and $\operatorname{bdeg}(E) = \operatorname{bdeg}(C/hC)$),

$$\begin{aligned} \operatorname{bdeg}(B) = \operatorname{bdeg}(B/hB) &\leq \operatorname{bdeg}((A/hA)/E) + \operatorname{bdeg}(C/hC) \\ &\leq \operatorname{bdeg}(A) + \operatorname{bdeg}(C), \end{aligned}$$

as desired. □

We continue with the proof of Proposition 6. If C has positive depth, $A_0 = B_0$ and as in the Lemma we replace A and B by A/A_0 and B/B_0, respectively. Picking an appropriate hyperplane section h, reduction gives an exact sequence with the same bdegs but in lower dimension. Consider the diagram

$$\begin{array}{ccccccccc}
 & & & & & & 0 & & \\
 & & & & & & \downarrow & & \\
 & & 0 & & & & C_0 & & \\
 & & \downarrow & & & & \downarrow & & \\
0 & \longrightarrow & A & \longrightarrow & B & \longrightarrow & C & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & A_1 & \longrightarrow & B & \longrightarrow & C/C_0 & \longrightarrow & 0 \\
 & & \downarrow & & & & \downarrow & & \\
 & & C_0 & & & & 0 & & \\
 & & \downarrow & & & & & & \\
 & & 0 & & & & & &
\end{array} .$$

Since C/C_0 has positive depth (or vanishes),

$$\mathrm{bdeg}(B) \leq \mathrm{hdeg}(A_1) + \mathrm{bdeg}(C/C_0) = \mathrm{hdeg}(A_1) + \mathrm{bdeg}(C) - \mathrm{bdeg}(C_0),$$

while the Lemma gives

$$\mathrm{bdeg}(A_1) \leq \mathrm{bdeg}(A) + \mathrm{bdeg}(C_0),$$

which proves the assertion. □

Remark 8. Unlike syzygies, the delicate task of bounding $\mathrm{bdeg}(C)$ in terms of $\mathrm{bdeg}(A)$ and $\mathrm{bdeg}(B)$ is not possible: just consider a module with a free resolution

$$0 \to F \xrightarrow{\varphi} F \longrightarrow C \to 0.$$

Then $\mathrm{hdeg}(C) = \deg(\mathbf{R}/(\det(\varphi)))$, which is independent of $\mathrm{bdeg}(F)$.

4.2 Monomial Ideals

Let $\mathbf{R} = k[x_1, \dots, x_n]$ and I a monomial ideal. The rules in Proposition 6 can be used for estimating $\mathrm{bdeg}(\mathbf{R}/I)$ and $\mathrm{bdeg}(I)$. Let us consider some cases.

Example 9. Let G be a graph on a vertex V set indexed by the set of variables and edge set E and set $I = I(G)$ its edge ideal.

- Let I be monomial ideal defined by the complete graph K_n: $I = (x_i x_j, i < j)$. Consider the exact sequence (a so-called *deconstruction sequence* in the terminology of R. Villarreal)

$$0 \to \mathbf{R}/I : x_n \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}/(I, x_n) \to 0.$$

  $I : x_n = (x_1, \dots, x_{n-1})$ and $\mathbf{R}/(x_n, I) = \mathbf{R}'/I'$, where $\mathbf{R}' = k[x_1, \dots, x_{n-1}]$ and I' is the ideal corresponding to the graph K_{n-1}. Using (6) and induction we get

$$\mathrm{bdeg}(\mathbf{R}/I) = \deg(\mathbf{R}/I) = n, \quad \mathrm{bdeg}(I) \leq 1 + n(n-1).$$

- For a graph with a minimal vertex cover $\{x_1, \dots, x_c\}$—in other words $(x_1, \dots, x_c)$ is a minimal prime of I of maximal dimension—one has to repeat the deconstruction step above $c - 1$ times: from

$$I = x_1 I_1 + x_2 I_2 + \cdots + x_c I_c,$$

$$0 \to \mathbf{R}/J_1 \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}/(I, x_1) \to 0$$

  that is

$$0 \to \mathbf{R}/J_1 \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}/(I, x_1) \to 0$$

$J_1 = (I_1 + x_2 I_2 + \cdots + x_c I_c)$.

$$\mathrm{bdeg}(\mathbf{R}/I) \leq c+1, \quad \mathrm{bdeg}(I) \leq 1 + (n-1)(c+1).$$

Proposition 10. *Let I be a monomial ideal of $\mathbf{R}$ generated by the set of monomials $\{m_1, \ldots, m_r\}$ of degree $\leq d$. Then*

$$\mathrm{bdeg}(\mathbf{R}/I) \leq \binom{n+d-1}{d}.$$

Proof. Denote $f(p,q)$ a bound for $\mathrm{bdeg}(\mathbf{R}/I)$ valid for ideals generated by monomials of degree $p < d$ or in $p < n$ variables. The reconstruction sequence and Proposition 6 assert that we can take

$$f(d,n) = \binom{n+d-1}{d}.$$

As the case of graph ideals indicates this often overstates the value of $\mathrm{bdeg}(\mathbf{R}/I)$. We use induction on the number r of monomials. If x_n is a variable present in one of the monomials m_i, say $m_i = x_n m_i'$, $i \leq s$ and $m_i \notin (x_n)$ for $i > s$, $I : x_n = (m_1', \ldots, m_s')$, we have the exact sequence

$$\begin{aligned} 0 \to \mathbf{R}/I : x_n = \mathbf{R}/(m_1', \ldots, m_s') \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}/(I, x_n) \\ = \mathbf{R}'/(m_{s+1}, \ldots, m_r) \to 0. \end{aligned}$$ □

This gives, using Theorem 10, the following uniform bound for the Betti numbers of monomial ideals.

Corollary 11. *Let $I = (m_1, \ldots, m_r)$ be a monomial ideal of $\mathbf{R} = k[x_1, \ldots, x_n]$,* deg $m_{\leq} d$. *Then the Betti numbers of $\mathbf{R}/I$ are bounded by*

$$\beta_i(\mathbf{R}/I) \leq \mathrm{bdeg}(\mathbf{R}/I) \cdot \binom{n}{i} \leq \binom{n+d-1}{d} \cdot \binom{n}{i}.$$

A more interesting calculation is:

Theorem 12. *Let $\mathbf{R}$ be a Buchsbaum local ring with infinite residue field and $\mathbf{x} = \{x_1, \ldots, x_d\} \subset \mathfrak{m} \setminus \mathfrak{m}^2$ be a system of parameters. Then*

$$\mathrm{bdeg}_{(\mathbf{x})}(\mathbf{R}) = \lambda(\mathbf{R}/(\mathbf{x})).$$

Proposition 13. *Suppose that $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ is a d-sequence in $\mathbf{R}$. Then we have the following:*

(a) The images of $x_2, x_3, \ldots, x_n$ in $\mathbf{R}/(x_1)$ form a d-sequence.

(b) $[(0){:}\,x_1] \cap (x_1, x_2, \ldots, x_n) = (0)$.
(c) *The images of* $x_1, x_2, \ldots, x_n$ *in* $\mathbf{R}/[(0){:}\,x_1]$ *form a d-sequence.*
(d) $\mathbf{x}$ *is a regular sequence if and only if* $(x_1, \ldots, x_{n-1}) : x_n = (x_1, \ldots, x_{n-1})$.
(e) *If* $\mathbf{x}$ *forms a d-sequence then the sequence* $\mathbf{x}' = \{x_1^*, \ldots, x_n^*\}$ *of* $\mathrm{gr}_{\mathbf{x}}(\mathbf{R})$ *is a d-sequence.*

Proof. In (d), the argument is a straightforward calculation. We may assume $n > 1$. If $(x_1, \ldots, x_{n-1}) : x_n = (x_1, \ldots, x_{n-1})$, we claim that $(x_1, \ldots, x_{n-2}) : x_{n-1} = (x_1, \ldots, x_{n-2})$. From $ax_{n-1} \in (a_1, \ldots, x_{n-2})$, we have $x_n x_{n-1} a \in (x_1, \ldots, x_{n-2})$, so

$$a \in (x_1, \ldots, x_{n-2}) : x_{n-1}x_n = (x_1, \ldots, x_{n-2}) : x_n \subset (x_1, \ldots, x_{n-1})$$
$$: x_n = (x_1, \ldots, x_{n-1}).$$

Thus $a = a_1x_1 + \cdots + a_{n-1}x_{n-1}$ and from $x_{n-1}a \in (x_1, \ldots, x_{n-2})$ we get $a_{n-1}x_{n-1}^2 \in (x_1, \ldots, x_{n-2})$ and therefore $a_{n-1}x_{n-1} \in (x_1, \ldots, x_{n-2})$, as desired.

(e) is proved in [10, Theorem 1.2], converse in [11]. □

An interesting consequence is:

Corollary 14. *Let* $\mathbf{R}$ *be a Noetherian local ring and* M *a finitely generated* $\mathbf{R}$*-module. If* $\mathbf{x}$ *is a system of parameters that is a d-sequence on the module* M *then* $\mathbf{x}$ *is a superficial sequence with respect to* M*.*

Definition 15. Let $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ be a sequence of elements in $\mathbf{R}$. $\mathbf{x}$ is a proper sequence if

$$x_{i+1}\mathrm{H}_j(x_1, x_2, \ldots, x_i) = 0, \quad \text{for } i = 0, 1, \ldots, n-1,\ j > 0,$$

where $\mathrm{H}_j(x_1, x_2, \ldots, x_i)$ is the Koszul homology associated to the subsequence $\{x_1, x_2, \ldots, x_i\}$.

Proposition 16. *Suppose that* $\mathbf{R}$ *is a Noetherian ring. Let* $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ *be a sequence in* $\mathbf{R}$*.*

(a) *If* $\mathbf{x}$ *is a d-sequence, then* $\mathbf{x}$ *is a proper sequence.*
(b) *If* $\mathbf{x}$ *is a proper sequence, then*

$$x_k\mathrm{H}_j(x_1, x_2, \ldots, x_i) = 0, \quad \textit{for } i = 0, 1, \ldots, n-1,\ j > 0,\ k > i.$$

(c) *Suppose that* $\mathbf{R}$ *is a local ring of dimension* $n > 0$*. If* $\mathbf{x}$ *is a proper sequence that is also a system of parameters in* $\mathbf{R}$*, then* $\mathrm{H}_j(x_1, x_2, \ldots, x_i)$ *is a module of finite length for* $i = 0, 1, \ldots, n$ *and* $j > 0$*.*

The following provides for a ready source of these sequences.

Proposition 17. *Let* $\mathbf{R}$ *be Noetherian local and* $\mathbf{x} = \{x_1, \ldots, x_d\}$ *a parameter ideal. If* $(x_1, \ldots, x_{d-1})$ *has grade* $d - 1$*, then*

1. $\mathbf{x}$ *is a proper sequence.*
2. $\mathbf{x}$ *is a d-sequence if and only if* $(x_1, \ldots, x_{d-1}) : x_d^2 = (x_1, \ldots, x_{d-1}) : x_d$.

Proof. We use induction with basic properties of d-sequences (Proposition 13):

- Set $I = (\mathbf{x})$. Since $\mathrm{H}^0_{\mathfrak{m}}(\mathbf{R}) = 0 : x_1$, consider the exact sequence

$$0 \to (0 : x_1) \longrightarrow \mathbf{R} \longrightarrow \mathbf{R}' \to 0.$$

  From (13) and basic properties of Buchsbaum rings (see [17]), we have that $(0 : x_1) \cap I = 0$ and $\mathbf{R}'$ is a Buchsbaum ring of positive depth.
- In the equality $\mathrm{bdeg}(\mathbf{R}) = \lambda(\mathrm{H}^0_{\mathfrak{m}}(\mathbf{R})) + \mathrm{bdeg}(\mathbf{R}')$, we make use of another property $\mathrm{bdeg}(\cdot)$: there exists a generic element $x \in I' = I\mathbf{R}'$ such that

$$\mathrm{bdeg}(\mathbf{R}') = \mathrm{bdeg}(\mathbf{R}'/(x)).$$

- If $d > 1$, $\mathbf{R}'$ is a Buchsbaum ring of dimension $d - 1$. Thus by induction $\mathrm{bdeg}(\mathbf{R}'/(x)) = \lambda(\mathbf{R}'/I')$.
- Let us return to the exact sequence above (and the case $d = 1$). Tensoring it modulo I, we get the complex

$$(0 : x_1)/I(0 : x_1) \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}'/I' \to 0.$$

  Since $(0 : x_1) \cap I$, we obtain the exact sequence

$$0 \to (0 : x_1) \longrightarrow \mathbf{R}/I \longrightarrow \mathbf{R}'/I' \to 0,$$

which gives the formula $\mathrm{bdeg}_I(\mathbf{R}) = \lambda(\mathbf{R}/I)$. □

Corollary 18. *Let* $\mathbf{R}$ *be a Buchsbaum local ring. For any system of parameters* $\mathbf{x} = \{x_1, \ldots, x_d\} \subset \mathfrak{m} \setminus \mathfrak{m}^2$, $\lambda(\mathbf{R}/(\mathbf{x}))$ *depends only on the integral closure of* $(\mathbf{x})$.

The following comes from a general property of Deg:

Corollary 19. *Let* $\mathbf{S}$ *be a regular local ring of dimension* n *and* $\mathbf{R} = \mathbf{S}/L$ *a Buchsbaum ring of dimension* d. *If* $\mathbf{x} = \{x_1, \ldots, x_d\} \subset \mathfrak{m} \setminus \mathfrak{m}^2$ *is a system of parameters for* $\mathbf{R}$, *then the Betti numbers of* $\mathbf{R}$ *are bounded by*

$$\beta_i(\mathbf{R}) \leq \lambda(\mathbf{S}/(L, \mathbf{x})) \cdot \binom{n}{i}.$$

4.3 Multiplicity-Based Complexity of Derived Functors

Let $(\mathbf{R}, \mathfrak{m})$ be a Noetherian local ring and let A and B be finitely generated $\mathbf{R}$-modules. Motivated by the occurrence of the derived functors of $\mathrm{Hom}_{\mathbf{R}}(A, \cdot)$ and

$A \otimes_{\mathbf{R}} \cdot$ in several constructions based on A we seek to develop gauges for the sizes for these modules. In the case of graded modules, a rich **degree**-based theory has been developed centered on the notion of Castelnuovo regularity. It is particularly well suited to handle complexity properties of tensor products and modules of homomorphisms.

Let us recall two general questions regarding the modules $\mathbf{C} = \mathrm{Hom}_{\mathbf{R}}(A, B)$ and $\mathbf{D} = A \otimes_{\mathbf{R}} B$.

- Can the minimal number of generators $\nu(\mathbf{C})$ be estimated in terms of $\nu(A)$ and $\nu(B)$ and other properties of A and B? If $\mathbf{R} = \mathbb{Z}$, or one of its localizations, the answer requires information derived from the structure theory for those modules. Since this is not available for higher dimensional rings, one can argue that an answer requires knowledge of the cohomology of the modules.
- In contrast the [minimal] number of generators of $\mathbf{D}$ is simply $\nu(\mathbf{D}) = \nu(A) \cdot \nu(B)$. What is hard about $\mathbf{D}$ is to gather information about its torsion, more precisely about its associated primes. Consider its submodule of finite support

$$\mathrm{H}^0_{\mathfrak{m}}(\mathbf{D}) = \mathrm{H}^0_{\mathfrak{m}}(A \otimes_{\mathbf{R}} B),$$

  and denote its length by $\mathrm{h}^0(\mathbf{D})$. Can one estimate $\mathrm{h}^0(\mathbf{D})$ in terms of A and B?
- (The HomAB problem) One formulation of these questions is the following. Let $\mathbf{R}$ be a Noetherian local ring. A question is whether there is a polynomial $\mathbf{f}(\mathbf{x}, \mathbf{y}) \in \mathbb{Q}[\mathbf{x}, \mathbf{y}]$ such that for any finitely generated $\mathbf{R}$-modules A and B,

$$\nu(\mathrm{Hom}(A, B)) \leq \mathbf{f}(\mathrm{hdeg}(A), \mathrm{hdeg}(B)).$$

  A test case asks for uniform bounds for $\nu(\mathrm{Hom}(A, A))$, where A is a Cohen–Macaulay module.
- Muddling the issues is how to account for the interaction between A and B. One attempt, that of replacing A and B by their direct sum $A \oplus B$, is only a temporary fix as it poses the question of what are the "self-interactions" of a module?
- This brings us full circle: which properties of A, B and of their interaction can we bring to the table? We shall refer to these questions as the *HomAB* and *TorsionAB* conjectures. They make sense even as *pure* questions of homological algebra, but we have in view *applied* versions.

Before we discuss specific motivations linking these questions to the other topics of these notes, we want to highlight the importance of deriving ring-theoretic properties of the ring $\mathbf{C} = \mathrm{Hom}_{\mathbf{R}}(A, A)$. Instances of $\mathbf{C}$ as a non-commutative desingularization of $\mathrm{Spec}(\mathbf{R})$ are found in the recent literature, and $\nu(\mathbf{C})$ may play a role as an embedding dimension. In these cases, A is a Cohen–Macaulay module, but we still lack estimates. This is obviously a stimulating question.

The sought-after estimations are polynomial functions on $\mathrm{hdeg}(A)$ and $\mathrm{hdeg}(B)$, whose coefficients are given in terms of invariants of $\mathbf{R}$. The first of these questions was treated in [4, 5], who refer to it as the HomAB question. It asks for uniform estimates for the number of generators of $\mathrm{Hom}_{\mathbf{R}}(A, B)$ in terms of invariants of $\mathbf{R}$,

A, and B. The extended question asks for the estimates of the number of generators of $\mathrm{Ext}^i_{\mathbf{R}}(A, B)$ (or of other functors).

In addition to the appeal of the question in basic homological algebra, such modules of endomorphisms appear frequently in several constructions, particularly in the algorithms that seek the integral closure of algebras (see [19, Chap. 6]). The algorithms employed tend to use rounds of operations of the form:

- $\mathrm{Hom}_{\mathbf{R}}(E, E)$: ring extension
- $A \to \tilde{A}$: S_2-ification of A
- $I : J$.

Remark 20. Before we outline a solution of the HomAB problem for graded modules, let us recall some explicit calculations from [4, 5].

1. ([4, Theorem 5.3]) If $\mathbf{R}$ is a Gorenstein local ring of dimension d then

$$\nu(\mathrm{Hom}(A, \mathbf{R})) \leq (\deg(\mathbf{R}) + d(d-1)/2)\mathrm{hdeg}(A).$$

2. ([4, Theorem 6.9]) If $\mathbf{R}$ is a Gorenstein local ring of dimension d and A is a vector bundle of finite projective dimension then for any $\mathbf{R}$-module B

$$\nu(\mathrm{Hom}(A, B)) \leq (\deg(\mathbf{R}) + d(d-1)/2 + \sum_{i=1}^{d} \beta_i^{\mathbf{R}}(k))\mathrm{hdeg}(A)\mathrm{hdeg}(B).$$

Kia Dalili has made use of the work of Chardin, Ha, and Hoa [3] to give an affirmative answer to the HomAB question and the related question on tensor products.

Theorem 21 (Dalili). *Let* $\mathbf{R} = k[x_1, \ldots, x_n]$ *be a standard graded algebra over the field* k. *There exist polynomials* $\mathbf{f}_i$ *and* $\mathbf{g}_i$ *such that for any two finitely generated graded modules* A *and* B*:*

- $\mathrm{bdeg}(\mathrm{Ext}^i_{\mathbf{R}}(A, B)) \leq \mathbf{f}_i(\mathrm{bdeg}(A), \mathrm{bdeg}(B), \alpha(A), \alpha(B))$
- $\mathrm{bdeg}(\mathrm{Tor}_i^{\mathbf{R}}(A, B)) \leq \mathbf{g}_i(\mathrm{bdeg}(A), \mathrm{bdeg}(B), \alpha(A), \alpha(B))$

Proof. We will just sketch his argument for $\mathrm{bdeg}(\mathrm{Hom}_{\mathbf{R}}(A, B))$.

- According to Theorem 9, there is a polynomial $\mathbf{f}_0$ so that

$$\mathrm{reg}(\mathrm{Hom}_{\mathbf{R}}(A, B)) \leq \mathbf{f}_0(\mathrm{reg}(A), \mathrm{reg}(B), \text{ Betti numbers of } A \text{ and } B).$$

On the other hand, by Theorem 6,

$$\mathrm{reg}(P) \leq \mathrm{Deg}(P) + \alpha(P)$$

for any extended degree Deg. Since each Betti number of a graded module P satisfies

$$\beta_i(P) \leq \mathrm{Deg}(P)\binom{n}{i},$$

the inequality leads to another polynomial inequality

$$\mathrm{reg}(\mathrm{Hom}_{\mathbf{R}}(A, B) \leq \mathbf{f}(\mathrm{Deg}(A), \mathrm{Deg}(B), \alpha(A), \alpha(B)).$$

- We now apply to properties of bdeg. If $F_0 \to A$ is a minimal free presentation of A, $\mathrm{Hom}_{\mathbf{R}}(A, B) \subset \mathrm{Hom}_{\mathbf{R}}(F_0, B)$, by Proposition 4,

$$\mathrm{bdeg}(\mathrm{Hom}_{\mathbf{R}}(A, B)) \leq \sum_{j=0}^{\mathrm{reg}(\mathrm{Hom}_{\mathbf{R}(A,B))}} H_{\mathrm{Hom}_{\mathbf{R}(F_0,B)}}(j).$$

□

5 Some Open Questions

If $\mathbf{R}$ is a Noetherian local ring a *semiadditive degree function* is a mapping

$$\mathbf{d} : \mathcal{M}(\mathbf{R}) \to \mathbb{Q}$$

such that if $0 \to A \to B \to C \to 0$ is a short exact sequence of modules in $\mathcal{M}(\mathbf{R})$ then

$$\mathbf{d}(B) \leq \mathbf{d}(A) + \mathbf{d}(C).$$

Such functions are amenable, as the example above indicates, to the derivation of binomial bounds. We have seen several degrees with this property, starting with $\nu(\cdot)$ (and more generally Bass and Betti numbers) and including $\mathrm{e}_1(\mathbf{x}; \cdot)$ (the first Hilbert coefficient), $\chi_1(\mathbf{x}; \cdot)$ (the first partial Euler characteristic), $\mathrm{reg}(\cdot)$, and $\mathrm{bdeg}(\cdot)$.

Question 1. Is $\mathrm{hdeg}(\cdot)$ semiadditive?

Question 2. Let $\mathbf{R}$ be a Cohen–Macaulay local ring of dimension d. Consider the set of rational numbers

$$\frac{\mathrm{hdeg}(M) - \mathrm{hdeg}(M/hM)}{\mathrm{deg}(M)}$$

over all $M \in \mathcal{M}(\mathbf{R})$ and all generic hyperplane sections. Is this set finite, or, more generally, bounded? Can it be expressed as an invariant of $\mathbf{R}$?

Question 3. Let $\mathbf{x}$ be a system of parameters of the Noetherian local ring $\mathbf{R}$. If $\mathbf{x}$ is a d-sequence, find an estimation for $\mathrm{hdeg}_{(\mathbf{x})}(\mathbf{R})$.

Question 4. Let $\mathbf{R} = k[\Delta]$ be the Stanley–Reisner ring of the simplicial complex Δ. Find estimations for $\text{hdeg}(\mathbf{R})$ and $\text{bdeg}(\mathbf{R})$.

Acknowledgements The author was partially supported by the NSF.

References

1. Bayer, D., Mumford, D.: What can be computed in Algebraic Geometry? In: Eisenbud, D., Robbiano, L. (eds.) Computational Algebraic Geometry and Commutative Algebra. Proceedings, Cortona 1991, pp. 1–48. Cambridge University Press, Cambridge (1993)
2. Bruns, W., Herzog, J.: Cohen–Macaulay Rings. Cambridge University Press, Cambridge (1993)
3. Chardin, M., Ha, D.T., Hoa, L.T.: Castelnuovo–Mumford regularity of Ext modules and homological degree. Trans. Amer. Math. Soc. **363**, 3439–3456 (2011). arXiv: math.AC/0903.4535
4. Dalili, K.: On the number of generators of modules of homomorphisms. J. Algebra **311**, 463–491 (2007)
5. Dalili, K., Vasconcelos, W.V.: Cohomological degrees and the HomAB conjecture. In: Corso, A., Migliore, J., Polini, C. (eds.) Algebra, Geometry and Their Interactions. Contemporary Mathematics, vol. 448, pp. 43–61 (2007)
6. Doering, L.R., Gunston, T., Vasconcelos, W.V.: Cohomological degrees and Hilbert functions of graded modules. American J. Math. **120**, 493–504 (1998)
7. Eisenbud, D.: Commutative Algebra with a View Toward Algebraic Geometry. Springer, Berlin (1995)
8. Gulliksen, T., Levin, G.: Homology of local rings. Queen's Paper in Pure and Applied Mathematics, vol. 20. Queen's University, Kingston, Ontario (1969)
9. Gunston, T.: Cohomological degrees, Dilworth numbers and linear resolution. Ph.D. Thesis, Rutgers University (1998). arxiv:1008.3711[math.AC]
10. Huneke, C.: Symbolic powers of prime ideals and special graded algebras. Comm. Algebra **9**, 339–366 (1981)
11. Kühl, M.: Über die symmetrische Algebra eines Ideals. Dissertation, Universität Essen, Germany (1981)
12. Linh, C.H.: Upper bound for the Castelnuovo–Mumford regularity of associated graded modules. Comm. Algebra **33**, 1817–1831 (2005)
13. Mandal, M., Singh, B., Verma, J.K.: On some conjectures about the Chern numbers of filtrations. J. Algebra **325**, 147–162 (2011)
14. Nagata, M.: Local Rings. Interscience, New York (1962)
15. Nagel, U.: Comparing Castelnuovo–Mumford regularity and extended degree: the borderline cases. Trans. Amer. Math. Soc. **357**, 3585–3603 (2005)
16. Rossi, M.E., Trung, N.V., Valla, G.: Castelnuovo–Mumford regularity and extended degree. Trans. Amer. Math. Soc. **355**, 1773–1786 (2003)
17. Stückrad, J., Vogel, W.: Buchsbaum Rings and Applications. Springer, New York (1986)
18. Vasconcelos, W.V.: The homological degree of a module. Trans. Amer. Math. Soc. **350**, 1167–1179 (1998)
19. Vasconcelos, W.V.: Integral closure. In: Springer Monographs in Mathematics. Springer, New York (2005)
20. Yamagishi, K.: Bass number characterization of surjective Buchsbaum modules. Math. Proc. Cambridge Phil. Soc. **110**, 261–279 (1991)