arXiv:1406.5145v1 [math.AG] 19 Jun 2014

GEOMETRIC LOWER BOUNDS FOR GENERALIZED RANKS

ZACH TEITLER

ABSTRACT. We revisit a geometric lower bound for Waring rank of polynomials (symmetric rank of symmetric tensors) of [LT10] and generalize it to a lower bound for rank with respect to arbitrary varieties, improving the bound given by the "non-Abelian" catalecticants recently introduced by Landsberg and Ottaviani. This is applied to give lower bounds for ranks of multihomogeneous polynomials (partially symmetric tensors); a special case is the simultaneous Waring decomposition problem for a linear system of polynomials. We generalize the classical Apolarity Lemma to multihomogeneous polynomials and give some more general statements. Finally we revisit the lower bound of [RS11], and again generalize it to multihomogeneous polynomials and some more general settings.

1. INTRODUCTION

Let F be a homogeneous polynomial of degree d in several variables. A **power sum decomposition** of F is an expression $F = c_1\ell_1^d + \cdots + c_r\ell_r^d$ in which the ℓ_i are linear forms and the c_i are scalars. The **length** of a power sum decomposition is the number r of terms. The **Waring rank** of F, denoted $r(F)$, is the least length r of a power sum decomposition of F. A **Waring decomposition** of F is a power sum decomposition of minimal length. For example,

$$xy = \frac{1}{4}\Big((x+y)^2 - (x-y)^2\Big), \tag{1}$$

$$xyz = \frac{1}{24}\Big((x+y+z)^3 - (x+y-z)^3 - (x-y+z)^3 + (x-y-z)^3\Big), \tag{2}$$

so $r(xy) \leq 2$ and $r(xyz) \leq 4$.

In fact, both of these inequalities are actually equalities. Several lower bounds for Waring rank have been developed. The earliest lower bound, and the basis for all the rest, involves *catalecticants*, which were introduced by Sylvester in 1851. We review catalecticants in Section 2. It turns out that the catalecticant lower bound gives $r(xy) \geq 2$ and $r(xyz) \geq 3$. Clearly, an improvement is desirable. One such improvement was given in [LT10], which showed that the catalecticant bound for rank could be improved by adding the dimension of a certain set of singularities. We review this in Section 3; it gives $r(xyz) \geq 4$. Another improvement, originally given in [RS11], is reviewed in Section 5.

Here we are interested in more general notions of rank for which, as we will see, there are well-known (generalized) catalecticant lower bounds. We develop improvements to these lower bounds analogous to the improvements in [LT10]. In some cases we are also able to develop improvements analogous to the one in [RS11].

Date: June 20, 2014.

2010 *Mathematics Subject Classification.* 15A21, 15A69, 14N15.

Key words and phrases. Waring rank, secant varieties.

In the remainder of this introduction we describe, in steps of increasing generality, the notions of rank in which we are interested. In Section 2 we review catalecticants for classical Waring rank (as above) and in the generalized settings. In Section 3 we introduce our improvements to the catalecticant lower bounds generalizing [LT10]. We review Apolarity Lemmas in Section 4. Finally in Section 5 we develop, at least for some cases, improvements to the catalecticant lower bounds generalizing [RS11].

We work over an algebraically closed field $\Bbbk$ of characteristic 0.

1.1. **Classical Waring rank.** Recall that a homogeneous polynomial of degree d in n variables is called an **n-ary d-form** or an **n-ary d-ic**; thus, for example, a homogeneous polynomial of degree 5 in 2 variables is a 2-ary 5-form, or binary quintic.

If F is a quadratic form, the Waring rank of F is equal to its rank as a quadratic form. Also, ranks of binary forms are understood, thanks to 19th century work by Sylvester [Syl51a], [Syl51b], [Syl86], Gundelfinger [Gun86], and others, see [GY10, Ch. 9]; for more recent treatments see for example [Kun86], [CS11], [Rez13b]. Ranks of ternary cubics are well known, probably since the 19th century or early 20th century; see for example [Yer32] and more recently [CM96], [LT10, §8]. Also the theorem of Alexander and Hirschowitz [AH95], [BO08], [Cha01], [Cha02], [Pos12] gives the ranks of *general* forms, meaning those in a dense open subset of the space of forms. Namely, a general d-form in n variables has rank

$$\left\lceil \frac{1}{n}\binom{n+d-1}{n-1} \right\rceil,$$

with a short list of exceptions: when $d=2$, the general rank is n (instead of $(n+1)/2$); when $(n,d)=(3,4),(4,4),(5,4),(5,3)$ the general rank is respectively $6,10,15,8$ (instead of $5,9,14,7$).

However for an arbitrary given form, it is surprisingly nontrivial to determine $r(F)$. There is no known effective way to determine if a given F is general, so that it has the rank given by Alexander–Hirschowitz. Nevertheless, ranks have been worked out in a number of cases. Typically one can express F as a sum of dth powers of linear forms, giving an upper bound on $r(F)$. For example, the power sum decomposition

$$x_1\cdots x_n = \frac{1}{2^{n-1}n!}\sum_{(\epsilon_2,\dots,\epsilon_n)\in\{\pm1\}^{n-1}} \epsilon_2\cdots\epsilon_n(x_1+\epsilon_2x_2+\cdots+\epsilon_nx_n)^n$$

shows $r(x_1\cdots x_n)\le 2^{n-1}$. Computational methods to find power sum decompositions have been developed [BCMT10], [OO13]. Finding good upper bounds is an interesting challenge, see [BBS08], [LT10, §5], [Jel13], [BP13], [BT14].

The problem we are concerned with here, however, is to give a lower bound. We will review some ideas for lower bounds in the following sections; for now, I will simply list all the determinations of Waring ranks of which I am aware.

That $r(x_1\cdots x_n)=2^{n-1}$ was shown for $n=4$ in [LT10] (2010), and for all n in [RS11] (2011), which showed more generally that $r((x_1\cdots x_n)^d)=(d+1)^{n-1}$. The ranks of arbitrary monomials and sums of pairwise coprime monomials were determined in [CCG12] (2012). Plane quartics have been studied [Kle99], [BGI11], [Par13] in great detail. In [TW13], Waring ranks are determined for the defining equations of hyperplane arrangements which are mirror arrangements for complex reflection groups satisfying a certain hypothesis. A few isolated examples have been computed: $r(x(y_1^2+\cdots+y_n^2))=r(x(y_1^2+\cdots+y_n^2+x^2))=2n$ and $r(x_1y_1z_1+\cdots+x_ny_nz_n)=4n$ [LT10, §7], [Ven13]; $r(x_0^2y_0-(x_0+x_1)^2y_1+x_1^2y_2)=9$

[BB13]. See [Rez92] for the forms $(x_1^2 + \cdots + x_n^2)^{d/2}$. Carlini, Catalisano, and Chiantini have shown very recently [CCC14] that $r(F(x_1, \ldots, x_n) + y_1^d + \cdots + y_s^d) = r(F) + s$ and $r(F(x_1, x_2) + G(y_1, y_2)) = r(F) + r(G)$ (a conjecture of Strassen asserts that this should hold for all F and G involving any number of variables). See also [Woo14]. Together with the previously mentioned quadratic and binary forms, ternary cubics, and general forms, this is, as far as I know, a complete list of all forms whose Waring ranks have been determined.

The recency of these results is somewhat surprising given the long history and widespread interest in questions about Waring rank, going back at least to Sylvester and 19th century investigations of apolarity and canonical forms. A wide range of applications has emerged in other areas of mathematics, statistics, engineering, and sciences. See for example [CM96], [DSS09, Chapter 4], [CGLM08], [BCMT10], [BBCM11], [Lan12]. For comprehensive introductions to this subject and its history and applications see [IK99], [Lan12]. We digress to briefly describe some of the applications. The rank of a polynomial may be considered a measure of its complexity, as in the field of *geometric complexity theory* [Lan13]. The linear functional on the space of d-forms corresponding to the inner product with $\ell^d = (a_1 x_1 + \cdots + a_n x_n)^d$ is given by evaluation at the point $(a_1, \ldots, a_n)$ (up to factorial factors), so a power sum decomposition of a polynomial corresponds to a *decomposition of a linear functional into a combination of atomic measures*. See [Rez92] for applications of this idea to number theory, functional analysis, numerical analysis (quadrature problems), and spherical designs. In statistics, a power sum decomposition of a polynomial corresponds to *mixture model of joint distributions of independent identically distributed random variables*. As an extremely simple example of this, a random variable X on a finite set with $P(X = i) = p_i$ for $1 \leq i \leq n$ may be described by the linear form $\ell_X = p_1 x_1 + \cdots + p_n x_n$; then the coefficients of ℓ_X^d give the joint distribution of d independent identically distributed copies of X. For a given random variable Y encoded in the coefficients of a d-form F, a power sum decomposition of F corresponds to an expression of Y as a mixture model of joint distributions of independent identically distributed random variables. (In this context one considers power sum decompositions with the extra constraints that all coefficients must be nonnegative real numbers summing to 1.) This is related to the *PARAFAC/CANDECOMP* decomposition. This basic idea plays a role in applications such as *blind source separation* in signal processing, where an observed signal must be decomposed into simple single sources. Much more discussion and detail may be found in the references above.

Despite this long history and widespread interest we are still left with the simple question: given F, what is $r(F)$?

Example 1.1. Let $\det_n$ be the determinant of an $n \times n$ generic matrix,

$$\det_n = \det \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix},$$

a polynomial of degree n. Since $\det_n$ is a sum of $n!$ monomials each with rank 2^{n-1}, $r(\det_n) \leq 2^{n-1} n!$; in particular $r(\det_3) \leq 24$.

There is a remarkable improvement of this, recently discovered by Derksen [Der13]:

$$
\begin{aligned}
\det_3 = \frac{1}{2}\Big(& (x_{13}+x_{12})(x_{21}-x_{22})(x_{31}+x_{32}) \\
& + (x_{11}+x_{12})(x_{22}-x_{23})(x_{32}+x_{33}) \\
& + 2x_{12}(x_{23}-x_{21})(x_{33}+x_{31}) \\
& + (x_{13}-x_{12})(x_{22}+x_{21})(x_{32}-x_{31}) \\
& + (x_{11}-x_{12})(x_{23}+x_{22})(x_{33}-x_{32})\Big).
\end{aligned} \tag{3}
$$

Each of the 5 terms is a product of 3 linear forms, $\ell_1\ell_2\ell_3$. This has rank 4 by substitution in (2). Thus $r(\det_3) \leq 20$. As noted by Derksen, an improved upper bound for larger determinants follows by Laplace expansion by complementary minors in the first 3 rows. We will see in Example 1.13 that

$$
r(\det_n) \leq \left(\frac{5}{6}\right)^{\lfloor n/3 \rfloor} 2^{n-1} n!. \tag{4}
$$

What about lower bounds for $r(\det_n)$? We will see in Example 2.4 that Sylvester's catalecticant lower bound gives $r(\det_n) \geq \binom{n}{\lfloor n/2 \rfloor}^2$ (so $r(\det_3) \geq 9$). A different lower bound introduced by Ranestad and Schreyer [RS11], together with a result of Masoumeh Sepideh Shafiei [Sha14], gives $r(\det_n) \geq \frac{1}{2}\binom{2n}{n}$ (so $r(\det_3) \geq 10$). The lower bound of [LT10] gives $r(\det_n) \geq \binom{n}{\lfloor n/2 \rfloor}^2 + n^2 - (\lfloor n/2 \rfloor + 1)^2$ (so $r(\det_3) \geq 14$), see Example 3.7. Among these lower bounds for $r(\det_n)$, the Ranestad–Schreyer–Shafiei lower bound grows most quickly, giving the best result for $n \geq 5$. However all three of these lower bounds grow exponentially in n, much more slowly than the factorial upper bound.

It would be quite interesting to determine $r(\det_n)$, or even to give better bounds. See [Sha14], [Sha13] for further discussion of determinants along with permanents (see Example 1.4), Pfaffians, symmetric determinants, etc.

We will revisit the generic determinant in Examples 1.8, 2.4, 2.16, 3.7.

Remark 1.2. The rank of a general form of degree n in n^2 variables is $\lceil \frac{1}{n^2}\binom{n^2+n-1}{n} \rceil$. This is greater than $2^{n-1}n!$ for $n \geq 4$, so $\det_n$ has less than the general rank for $n \geq 4$. However the general rank of a cubic in 9 variables is 19 while $14 \leq r(\det_3) \leq 20$. Even to determine whether $\det_3$ has greater or less than the general rank is an interesting challenge.

Remark 1.3. Note that the coefficients c_i in a power sum decomposition $F = \sum c_i \ell_i^d$ are not necessarily uniquely determined, but they are uniquely determined in the case of a Waring decomposition. Indeed, in a Waring decomposition the ℓ_i^d must be linearly independent, or else the number of terms could be reduced by replacing one of the ℓ_i^d by a linear combination of the others.

1.2. **Simultaneous Waring rank.** As a first step toward full generality, let W be a linear series of homogeneous forms of degree d. A **simultaneous power sum decomposition of** W of length r is a collection of linear forms $\ell_1, \dots, \ell_r$ such that for every $F \in W$ there exist scalars $c_1, \dots, c_r$ yielding a power sum decomposition $F = c_1\ell_1^d + \cdots + c_r\ell_r^d$. That is, $W \subseteq \operatorname{span}\{\ell_1^d, \dots, \ell_r^d\}$. See [Bro33], [Fon02], [CC03], [BL13]. The **simultaneous Waring rank** $r(W)$ is the least length of a simultaneous power sum decomposition of W; a **simultaneous Waring decomposition** is a simultaneous power sum decomposition of minimal length.

Clearly, $r(W) \geq r(F)$ for all $F \in W$, and $r(W) \geq \dim W$. Also clearly, $r(W) \leq \sum r(F_i)$ for a basis $F_1, \ldots, F_n$ of W. Typically it is a difficult problem to determine the maximum and minimum of $r(F)$ for $F \in W$, the set of F on which the maximum is attained, etc.

Example 1.4. Recall that the **permanent** of a $k \times k$ matrix $A = (a_{i,j})$ is

$$\operatorname{per} A = \sum_{\pi \in S_k} \prod_{i=1}^{k} a_{i,\pi(i)},$$

that is, the (un-signed) sum of products with one entry from each row and column of A.

Let $X = (x_{i,j})$, $1 \leq i \leq m$, $1 \leq j \leq n$ be a generic $m \times n$ matrix. Let D_k be the linear series spanned by the k-minors of X, let P_k be spanned by the permanents of $k \times k$ submatrices of X, and let R_k be spanned by the degree k rook-free products in X, that is, products of k distinct entries of X with no two in the same row or column (so that chess rooks placed in those positions would be pairwise non-attacking).

Note that $D_k, P_k \subset R_k$, so $r(D_k), r(P_k) \leq r(R_k)$. Each degree k product of linearly independent factors has rank 2^{k-1}, so

$$r(R_k) \leq 2^{k-1} \dim(R_k) = 2^{k-1} \binom{m}{k} \binom{n}{k} k!. \tag{5}$$

And

$$r(D_k) \leq \left(\frac{5}{6}\right)^{\lfloor k/3 \rfloor} 2^{k-1} k! \dim(D_k) = \left(\frac{5}{6}\right)^{\lfloor k/3 \rfloor} 2^{k-1} \binom{m}{k} \binom{n}{k} k!.$$

As far as I know these are the best known upper bounds for the ranks of R_k and D_k, including in the case $m = n = k$ (where D_k reduces to a single, square determinant). While (5) is also an upper bound for $r(P_k)$, we can do better. First, the Ryser identity [Rys63] (for a square matrix) is the following:

$$\operatorname{per}_k = \operatorname{per}(x_{i,j})_{1 \leq i,j \leq k} = \sum_{S \subseteq \{1,\ldots,k\}} (-1)^{k-|S|} \prod_{i=1}^{k} \sum_{j \in S} x_{i,j}.$$

For example,

$$\begin{aligned}
\operatorname{per}_3 = &(x_{1,1} + x_{1,2} + x_{1,3})(x_{2,1} + x_{2,2} + x_{2,3})(x_{3,1} + x_{3,2} + x_{3,3}) \\
&- (x_{1,1} + x_{1,2})(x_{2,1} + x_{2,2})(x_{3,1} + x_{3,2}) \\
&- (x_{1,2} + x_{1,3})(x_{2,2} + x_{2,3})(x_{3,2} + x_{3,3}) \\
&- (x_{1,1} + x_{1,3})(x_{2,1} + x_{2,3})(x_{3,1} + x_{3,3}) \\
&+ x_{1,1}x_{2,1}x_{3,1} + x_{1,2}x_{2,2}x_{3,2} + x_{1,3}x_{2,3}x_{3,3}.
\end{aligned}$$

This expresses the $k \times k$ permanent as a sum of $2^k - 1$ terms, each of rank 2^{k-1}. Applying this to each basis element gives $r(\operatorname{per}_k) \leq 2^{2k-1} - 2^{k-1}$ and $r(P_k) \leq (2^{2k-1} - 2^{k-1})\binom{m}{k}\binom{n}{k}$. But better, Glynn gives a similar identity [Gly10]:

$$\operatorname{per}(x_{i,j})_{1 \leq i,j \leq k} = \sum_{\substack{\epsilon \in \{\pm 1\}^k \\ \epsilon_1 = 1}} \prod_{i=1}^{k} \sum_{j=1}^{k} \epsilon_i \epsilon_j x_{i,j}. \tag{6}$$

For example,

$$\begin{aligned}\mathrm{per}_3 &= (x_{1,1}+x_{1,2}+x_{1,3})(x_{2,1}+x_{2,2}+x_{2,3})(x_{3,1}+x_{3,2}+x_{3,3})\\ &\quad - (x_{1,1}+x_{1,2}-x_{1,3})(x_{2,1}+x_{2,2}-x_{2,3})(x_{3,1}+x_{3,2}-x_{3,3})\\ &\quad - (x_{1,1}-x_{1,2}+x_{1,3})(x_{2,1}-x_{2,2}+x_{2,3})(x_{3,1}-x_{3,2}+x_{3,3})\\ &\quad + (x_{1,1}-x_{1,2}-x_{1,3})(x_{2,1}-x_{2,2}-x_{2,3})(x_{3,1}-x_{3,2}-x_{3,3}).\end{aligned}$$

This expresses the $k \times k$ permanent as a sum of 2^{k-1} terms, each of rank 2^{k-1}. Therefore $r(\mathrm{per}_k) \le 2^{2k-2}$ and $r(P_k) \le 2^{2k-2}\binom{m}{k}\binom{n}{k}$.

It would be interesting to determine if these natural linear series, especially D_k, admit any simultaneous Waring decomposition shorter than simply decomposing separately each member of a basis for the linear series, or at least if any cleverly chosen basis can do better than the "obvious" defining basis consisting of minors for D_k, permanents for P_k, and products for R_k.

We will revisit these linear series in Examples 2.11, 3.13, 3.14.

Remark 1.5. Note that if $\{\ell_1,\dots,\ell_r\}$ give a simultaneous Waring decomposition of W, then the ℓ_i^d must be linearly independent, or else they would not be a minimal spanning set. Thus for each $F \in W$ the scalar coefficients c_i in the power sum decomposition $F = \sum c_i \ell_i^d$ are uniquely determined, even though it is not necessarily a Waring decomposition.

1.3. **Multihomogeneous polynomials.** More generally, we consider ranks of multihomogeneous polynomials. Fix $s > 0$, positive integers $n_1,\dots,n_s$, and s sets of doubly-indexed variables $x_{i,j}$, $1 \le i \le s$, $1 \le j \le n_i$. A polynomial F in the $x_{i,j}$ is multihomogeneous of multidegree $(d_1,\dots,d_s)$ if for each i, each monomial appearing in F has degree d_i in the ith set of variables, that is, $x_{i,1},\dots,x_{i,n_i}$.

A **multihomogeneous power sum decomposition of** F of length r is an expression $F = \sum_{k=1}^r c_k \ell_{1,k}^{d_1}\cdots\ell_{s,k}^{d_s}$ where each $\ell_{i,k}$ is a linear form in the ith set of variables. As before, the **multihomogeneous Waring rank** of F is the least number of terms in a multihomogeneous power sum decomposition of F and a **multihomogeneous Waring decomposition** of F is a multihomogeneous power sum decomposition of F of minimal length. See [Fon06] (focusing on uniqueness of decompositions).

Ranks of multihomogeneous polynomials generalize several familiar notions. Classical Waring rank is the case $s = 1$. The case when the multidegree is $(1,\dots,1)$ is the usual tensor rank. Tensor rank is very well studied, with applications far too numerous to mention; see [KB09], [Lan12]. If W is a linear series of degree d forms, then the simultaneous Waring rank of W is the rank of a single bihomogeneous polynomial of bidegree $(1,d)$. Namely, let $F_1(x_1,\dots,x_n),\dots,F_s(x_1,\dots,x_n)$ be a basis for W; then $r(W)$ is equal to the rank of the multihomogeneous polynomial $M = \sum t_i F_i$ with multidegree $(1,d)$.

We make this last observation explicit. Let W, F_i, M be as above. For each $j = 1,\dots,s$, let $e_j = (0,\dots,1,\dots,0)$, the s-tuple with 1 in the jth position and all other entries zero. First, if $M = M(t,x) = \sum_{i=1}^r \ell_i(t) m_i(x)^d$ then for each $j = 1,\dots,s$, $F_j = M(e_j,x) = \sum_{i=1}^r \ell_i(e_j) m_i(x)^d$, so the m_i give a simultaneous power sum decomposition of W of length r. Conversely, if $m_1,\dots,m_r$ give a simultaneous power sum decomposition of W, write $F_j = \sum_{i=1}^r c_{i,j} m_i^d$ for each j. The $c_{i,j}$ are uniquely determined by Remark 1.5. For each $1 \le i \le r$, let ℓ_i be the linear form with $\ell_i(e_j) = c_{i,j}$ for $1 \le j \le s$. Then $M = \sum_{i=1}^r \ell_i m_i^d$, giving a multihomogeneous power sum decomposition of M of length r.

In order to distinguish between the classical Waring rank of F as a homogeneous polynomial and the rank of F as a multihomogeneous polynomial, we reserve $r(F)$ for the former and write $r_{MH}(F)$ for the latter, or $r_{MH(d_1,\dots,d_s)}(F)$ if we wish to specify how F is considered to be multihomogeneous.

Example 1.6. $F = x_1 \cdots x_a y_1 \cdots y_b$ is homogeneous of degree $d = a + b$ and has Waring rank 2^{a+b-1}, but it is also bihomogeneous of bidegree (a, b) in the x and y variables, and the bihomogeneous rank of F is at most 2^{a+b-2}. Indeed, a decomposition of this length is given by multiplying decompositions of the separate parts. Let $x_1 \cdots x_a = \sum_{i=1}^{2^{a-1}} \ell_i^a$ and $y_1 \cdots y_b = \sum_{j=1}^{2^{b-1}} m_j^b$ be Waring decompositions. Then $F = \sum_{i,j} \ell_i^a m_j^b$ is a decomposition of F as a bihomogeneous form using 2^{a+b-2} terms.

This is actually the rank when $a = 1$ or $b = 1$—if, say, $a = 1$, then setting $x_1 = 1$ in any decomposition yields a (classical) Waring decomposition of $y_1 \cdots y_b$, which must involve at least 2^{b-1} terms, so as a bihomogeneous form $r_{MH(1,b)}(F) \geq 2^{b-1} = 2^{a+b-2}$.

We will revisit these bihomogeneous products of variables in Example 2.15, in Example 3.21, and in Example 5.11 where we show that in fact $r_{MH}(x_1 \cdots x_a y_1 \cdots y_b) = 2^{a+b-2}$.

Example 1.7. More generally, $r_{MH}(f(X)g(Y)) \leq r(f)r(g)$, with equality if $r(f) = 1$ or $r(g) = 1$.

Example 1.8. The generic determinant $\det_n$ (see Example 1.1) is bihomogeneous of bidegree $(a, n - a)$ in the sets of variables appearing in the first a rows or last $n - a$ rows of the matrix. The rank $r_{MH(a,n-a)}(\det_n)$ is less than or equal to the Waring rank $r(\det_n)$. Also $r_{MH(a,n-a)}(\det_n) \leq 2^{n-2}n!$, since $\det_n$ is a sum of $n!$ monomials each with (multihomogeneous) rank at most 2^{n-2}. Better, Derksen's formula (3) is multihomogeneous in the rows of the matrix, so $r_{MH(a,n-a)}(\det_n) \leq (\frac{5}{6})^{\lfloor n/3 \rfloor} 2^{n-2}n!$, since $\det_n$ is a sum of $(5/6)^{\lfloor n/3 \rfloor} n!$ products of linear forms each with (multihomogeneous) rank at most 2^{n-2}. We can do at least as well by expanding $\det_n$ as an alternating sum of products of maximal minors of the first a rows of the matrix with their complementary minors from the last $n - a$ rows, yielding

$$\begin{aligned} r_{MH(a,n-a)}(\det_n) &\leq \binom{n}{a} r(\det_a) r(\det_{n-a}) \\ &\leq \binom{n}{a} \left(\frac{5}{6}\right)^{\lfloor a/3 \rfloor + \lfloor (n-a)/3 \rfloor} 2^{n-2} a!(n-a)! \\ &= \left(\frac{5}{6}\right)^{\lfloor a/3 \rfloor + \lfloor (n-a)/3 \rfloor} 2^{n-2} n! \\ &\leq \left(\frac{5}{6}\right)^{\lfloor n/3 \rfloor} 2^{n-2} n!. \end{aligned}$$

See also Examples 2.16, 3.22.

1.4. **Generalized rank.** Even more generally, one can define rank with respect to any projective variety. Let the variety $X \subset \mathbb{P}^n$ be nondegenerate, that is, not contained in any hyperplane. For an affine point $q \neq 0$, the **rank of q with respect to X**, denoted $r_X(q)$, is the least r such that there exist some r distinct, reduced affine points $x_1, \dots, x_r$ such that $[x_1], \dots, [x_r] \in X$ and their linear span contains q: that is, $q = c_1 x_1 + \cdots + c_r x_r$ for some scalars c_i. Then the classical Waring rank is rank with respect to a Veronese variety; tensor rank is rank with respect to a Segre variety; the rank of a multihomogeneous polynomial is rank with respect to a Segre-Veronese variety, a product of projective spaces $\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_s-1}$ embedded by the line bundle $\mathcal{O}(d_1,\ldots,d_s)$ where $(d_1,\ldots,d_s)$ is the multidegree of the polynomial.

So this notion includes all the previous ideas, and more. For example, the rank of an alternating tensor — the least length of an expression as a sum of simple wedges — is its rank with respect to a Grassmannian in its Plücker embedding. Ranks with respect to an elliptic normal curve have been studied in [BGI11, Thm. 28].

Example 1.9. Carlini considered "codimension one decompositions" in [Car05]. Given a d-form $F \in S = \Bbbk[x_1,\ldots,x_n]$, such a decomposition is an expression $F = G_1 + \ldots + G_r$, where each G_i is a d-form in a subring generated by $n-1$ linear forms: $G_i \in \Bbbk[\ell_1,\ldots,\ell_{n-1}]$. In [Car05] Carlini determines the number of summands in a codimension one decomposition of a general form. The least number of terms in such a decomposition is given by rank with respect to the variety of forms that depend on (at most) $n-1$ variables, called a **subspace variety**. We will define this more precisely in terms of catalecticants, introduced in the next section. See Definition 2.9.

Example 1.10. Similarly, Carlini considered "binary decompositions" in [Car06a], expressions for F as a sum of binary forms, i.e., forms lying in a subring generated by two linear forms. Again the least number of terms in such an expression is given by rank with respect to a subspace variety, this time parametrizing forms that depend on at most two variables, see Definition 2.9.

Note that classical Waring rank corresponds to decompositions into forms depending on one variable, i.e., homogeneous polynomials in a single linear form. In fact the subspace variety whose points are forms depending on one variable is just the Veronese variety.

Example 1.11. It is interesting to write a form F as a sum of products of linear forms. For example, the determinant and permanent of an $n \times n$ matrix can be written as sums of $n!$ products of linear forms. Derksen's formula improves this, see Example 1.13 below. The least number of terms in such an expression is rank with respect to the variety parametrizing forms which completely factor as products of linear forms, called the **split variety** or the Chow variety of zero-cycles, see for example [AB11].

Let us consider lower bounds for this rank, which we will call **split rank** and denote r_{split}. Since $r(x_1 \cdots x_d) = 2^{d-1}$, we have for any d-form F the relations

$$r_{\mathrm{split}}(F) \le r(F) \le 2^{d-1} r_{\mathrm{split}}(F).$$

Therefore

$$r_{\mathrm{split}}(F) \ge 2^{1-d} r(F),$$

and any lower bound for Waring rank leads to a lower bound for split rank. In Example 1.1 we saw $r(\det_n) \ge \frac{1}{2}\binom{2n}{n}$, so $r_{\mathrm{split}}(\det_n) \ge 2^{-n}\binom{2n}{n}$. From this we get $r_{\mathrm{split}}(\det_3) \ge 3$. The lower bound of [LT10] gives $r(\det_3) \ge 14$, so in fact $r_{\mathrm{split}}(\det_3) \ge 4$.

Similarly, the generic $n \times n$ permanent per_n has Waring rank $\frac{1}{2}\binom{2n}{n} \le r(\mathrm{per}_n) \le 2^{2n-2}$ (see [Sha14]) so $2^{-n}\binom{2n}{n} \le r_{\mathrm{split}}(\mathrm{per}_n)$. And the Glynn identity (6) gives $r_{\mathrm{split}}(\mathrm{per}_n) \le 2^{n-1}$. Thus for instance $3 \le r_{\mathrm{split}}(\mathrm{per}_3) \le 4$. Nathan Ilten has shown that $r_{\mathrm{split}}(\mathrm{per}_3) = 4$ (private communication).

Example 1.12. If F is a d-form and $d = kt$ then F may be written as a sum of kth powers of t-forms. The least number of terms in such an expression is the rank with respect to the variety of d-forms which are kth powers. Classical Waring rank is the case $t = 1$. Sum-of-squares decompositions correspond to $k = 2$ (although in that setting one is usually interested in working over a real field and finding decompositions with nonnegative coefficients, whereas here we work over a closed field). See [CO13], [FOS12], [Rez13a].

Example 1.13. We consider again the remarkable formula in Example 1.1 discovered by Derksen [Der13]. Consider the generic determinant as a multilinear function of the columns, i.e., $\det_n \in (\Bbbk^n)^{\otimes n}$. For this paragraph we denote tensor rank by $r_\otimes$. Naively we have $r_\otimes(\det_n) \leq n!$; for example,

$$\begin{aligned}\det_3 = e_1 \otimes e_2 \otimes e_3 + e_2 \otimes e_3 \otimes e_1 + e_3 \otimes e_1 \otimes e_2 \\ - e_1 \otimes e_3 \otimes e_2 - e_2 \otimes e_1 \otimes e_3 - e_3 \otimes e_2 \otimes e_1,\end{aligned}$$

where $\{e_1, e_2, e_3\}$ is a basis for $\Bbbk^3$. Derksen's formula improves this:

$$\begin{aligned}\det_3 = \frac{1}{2}\Big(&(e_1 + e_2) \otimes (e_2 - e_3) \otimes (e_2 + e_3) \\ &+ (e_1 + e_2) \otimes (e_2 - e_3) \otimes (e_2 + e_3) \\ &+ 2e_2 \otimes (e_3 - e_1) \otimes (e_3 + e_1) \\ &+ (e_3 - e_2) \otimes (e_2 + e_1) \otimes (e_2 - e_1) \\ &+ (e_1 - e_2) \otimes (e_3 + e_2) \otimes (e_3 - e_2)\Big).\end{aligned} \tag{7}$$

This shows $r_\otimes(\det_3) \leq 5$. (One can show $r_\otimes(\det_3) \geq 4$, see below.) As noted by Derksen, Laplace expansion by complementary minors in the first 3 rows gives an improved upper bound for larger determinants.

$$r_\otimes(\det_n) \leq \binom{n}{3} r_\otimes(\det_3) r_\otimes(\det_{n-3}) = \frac{5 \cdot n!}{6 \cdot (n-3)!} r_\otimes(\det_{n-3}),$$

so by induction

$$r_\otimes(\det_n) \leq \left(\frac{5}{6}\right)^{\lfloor n/3 \rfloor} n!.$$

Let us return to considering the determinant as a function of the entries, rather than the columns. We get (3), as in Example 1.11, from (7) by the substitution that replaces $e_i \otimes e_j \otimes e_k$ with $x_{1,i}x_{2,j}x_{3,k}$:

$$\begin{aligned}\det_3 = \frac{1}{2}\Big(&(x_{13} + x_{12})(x_{21} - x_{22})(x_{31} + x_{32}) \\ &+ (x_{11} + x_{12})(x_{22} - x_{23})(x_{32} + x_{33}) \\ &+ 2x_{12}(x_{23} - x_{21})(x_{33} + x_{31}) \\ &+ (x_{13} - x_{12})(x_{22} + x_{21})(x_{32} - x_{31}) \\ &+ (x_{11} - x_{12})(x_{23} + x_{22})(x_{33} - x_{32})\Big)\end{aligned}$$

This shows that $r_{\text{split}}(\det_3) \leq 5$. As in Example 1.11, Laplace expansion shows that

$$r_{\text{split}}(\det_n) \leq \left(\frac{5}{6}\right)^{\lfloor n/3 \rfloor} n!.$$

It follows that Waring rank satisfies

$$r(\det_n) \le 2^{n-1} \left(\frac{5}{6}\right)^{\lfloor n/3 \rfloor} n!,$$

for example $r(\det_3) \le 20$.

Note that these formulas are multihomogeneous: each term involves one factor from each row of the matrix. So Derksen's formula yields similarly improved upper bounds for $r_{MH}(\det_n)$.

Analogously to the substitution that takes (7) to (3), a similar substitution takes any tensor decomposition of $\det_n$ to a similar expression as a sum of products of linear forms. Thus

$$r_\otimes(\det_3) \ge r_{\mathrm{split}}(\det_3) \ge \frac{1}{4} r(\det_3) \ge \frac{14}{4} > 3,$$

which shows $r_\otimes(\det_3) \ge 4$.

The problem of determining rank with respect to an arbitrary variety is at least as hard as the already difficult problems of Waring rank and tensor rank. Nevertheless, a great deal of progress has been made in understanding related questions involving secant varieties, including determining dimensions and equations of secant varieties. Recently, Landsberg and Ottaviani [LO12] introduced a method to generate new equations of secant varieties. They define a notion of catalecticants with respect to a given variety, generalizing Sylvester's catalecticants beyond the case of Veronese varieties as well as the symmetric flattenings or generalized Hankel matrices used for studying secant varieties of Segre varieties (i.e., tensor rank). Their motivation was to find equations for secant varieties, but here we are interested in the lower bounds for rank given by their generalized catalecticants.

2. Catalecticants

For each of the generalizations of Waring rank defined above, there is a reasonably well-known notion of a *catalecticant*. In each case this is a linear map whose rank is a lower bound for (generalized) Waring rank.

For a discussion of the name "catalecticant", see [Rez92, pg. 49–50], [Ger96, Lecture 11], [Mil13].

2.1. **Classical Waring rank.**

If V is a finite dimensional vector space, the dth symmetric power S^dV is the space of symmetric tensors or equivalently homogeneous polynomials of degree d in V. That is, if V has basis $x_1, \dots, x_n$, S^dV is the dth graded piece of the polynomial ring $\Bbbk[V] = \Bbbk[x_1, \dots, x_n]$. We regard S^aV^* as the space of differential operators of order a with constant coefficients, equivalently homogeneous polynomials of degree a in the partial differentiation operators $\partial_1 = \frac{\partial}{\partial x_1}, \dots, \partial_n = \frac{\partial}{\partial x_n}$. That is, S^aV^* is the ath graded piece of the polynomial ring $\Bbbk[V^*] = \Bbbk[\partial_1, \dots, \partial_n]$.

An element of $S^{d-a}V^*$ can be regarded both as a polynomial function on V and as a differential operator on $\Bbbk[V]$. Say $p \in S^{d-a}V^*$ is regarded as a polynomial with corresponding differential operator D_p. For a linear form $\ell = c_1x_1 + \cdots + c_nx_n \in V$ we write $p(\ell) = p(c_1, \dots, c_n)$. We will make frequent use, especially in Section 3, of the following well-known fact: If $p \in S^{d-a}V^*$ and $\ell \in V$ then $D_p(\ell^d) = \frac{d!}{a!}\ell^a p(\ell)$. See for example [IK99, (1.1.10)]. In particular $D_p(\ell^d) = 0$ if and only if $p(\ell) = 0$.

For $0 \leq a \leq d$ the natural map $S^dV \otimes S^{d-a}V^* \to S^aV$ coincides (at least in characteristic zero) with the usual differentiation action, $F \otimes D \mapsto DF$ for a polynomial F of degree d and differential operator D of order $d-a$.

Definition 2.1. Let $F \in S^dV$ and $0 \leq a \leq d$. The a**th catalecticant of** F, denoted C^a_F, is the linear map $C^a_F : S^aV^* \to S^{d-a}V$, $D \mapsto DF$.

This is also called a symmetric flattening or generalized Hankel matrix. (If F is a binary polynomial, i.e. $\dim V = 2$, of degree $d = 2a$, then modulo some binomial factors C^a_F is a Hankel matrix when written in the usual monomial basis for S^aV and S^aV^*.)

The inequality

$$r(F) \geq \operatorname{rank} C^a_F \tag{8}$$

is well known. See [Syl51a], in which Sylvester introduced the catalecticant matrix in 1851. For the reader who is new to this area, we mention a surprisingly quick proof: If $F = c_1\ell_1^d + \cdots + c_r\ell_r^d$, then the image of C^a_F is spanned by the DF for $D \in S^aV^*$, and each DF is contained in the span of $D\ell_1^d, \dots, D\ell_r^d$. Each $D\ell_i^d = c\ell_i^{d-a}$ for some constant c. So the image of C^a_F is contained in the span of $\ell_1^{d-a}, \dots, \ell_r^{d-a}$. Thus the rank of C^a_F is at most $r = r(F)$.

This simple dimension-counting idea carries through in the generalizations as well.

Remark 2.2. Note that C^a_F and C^{d-a}_F are transposes of one another. Certainly they map between the right spaces to be transposes of one another, as $C^a_F : S^aV^* \to S^{d-a}V$ and $C^{d-a}_F : S^{d-a}V^* \to S^aV$. It is easy to see that these maps are actually transposes. Here is a completely elementary argument. Let $A \in S^aV^*$ and $B \in S^{d-a}V^*$. The pairing $\langle C^a_F A, B\rangle$ is given by applying the differentiation operator B to the polynomial $C^a_F A = AF$, yielding BAF. Similarly, the pairing $\langle A, C^{d-a}_F B\rangle$ is given by ABF. These are equal, $BAF = ABF$, and thus C^a_F and C^{d-a}_F are transposes of one another. That is, C^a_F and C^{d-a}_F are transpose because partial differentiation operators commute on polynomials.

Another way to say the same thing is that for any bilinear functional $W_1 \times W_2 \to \Bbbk$, the induced maps $W_1 \to W_2^*$ and $W_2 \to W_1^*$ are transposes. The catalecticant maps C^a_F and C^{d-a}_F arise in this way from the bilinear map $S^aV^* \times S^{d-a}V^* \to S^dV^* \to \Bbbk$ sending $(A, B) \mapsto AB \mapsto ABF$.

Example 2.3. Each derivative of $x_1 \cdots x_n$ is a product of a subset of $\{x_1, \dots, x_n\}$ (or a linear combination of such products). For each a, the products of degree $d-a$ are linearly independent and there are $\binom{n}{a}$ of them. So $r(x_1 \cdots x_n) \geq \operatorname{rank}(C^a_{x_1 \cdots x_n}) = \binom{n}{a}$. This lower bound is maximized when $a = \lfloor n/2 \rfloor$.

For a completely explicit example, $r(xyz) \geq \operatorname{rank}(C^2_{xyz}) = 3$, since the image of C^2_{xyz} is spanned by $\{x = \partial_y\partial_z xyz, y = \partial_x\partial_z xyz, z = \partial_x\partial_y xyz\}$. Compare this with the upper bound $r(xyz) \leq 4$ obtained from the explicit power sum decomposition in the introduction. The non-sharpness of the catalecticant bound in such an undemanding example spurs us to look for improved lower bounds.

See Example 3.6.

Example 2.4. Consider the generic determinant $\det_n$ as in Example 1.1. By the Laplace expansion, $\partial_{i,j}\det_n$ is the complementary $(n-1)$-minor of the matrix. By induction, the image of the ath catalecticant of $\det_n$ is spanned by the $(n-a)$-minors of the matrix. There are $\binom{n}{a}^2$ of these and they are linearly independent (in fact, no two $(n-a)$-minors have any monomials in common, since every monomial appearing in a minor determines the set of rows and columns of the minor). Thus $r(\det_n) \geq \operatorname{rank}(C^a_{\det_n}) = \binom{n}{a}^2$. Again, this lower bound is maximized when $a = \lfloor n/2 \rfloor$.

See Example 3.7.

Remark 2.5. If $n \leq 3$ then the sequence $\operatorname{rank}(C^a_F)$, $a = 0, 1, \dots, d$, is unimodal, so the maximum occurs for $a = \lfloor d/2 \rfloor$ [Sta78]. For $n \geq 5$ the sequence is not necessarily unimodal [BI92], [BL94], [Boi95] and it is not at all clear where the maximum occurs. It is still an open question whether this sequence is unimodal when $n = 4$, see for example [SS12]. It is known to be unimodal when $n = 5$ and $d \leq 15$ [AS13] but surprisingly (to me, at least) this appears to be still open when $n = 4$. (The example of Bernstein–Iarrobino in [BI92] has $n = 5$ and $d = 16$.)

Example 2.6. We briefly present Stanley's nonunimodal example [Sta78, Example 4.3], a form of degree 4 in 13 variables, which we denote $x, y, z, t_1, \dots, t_{10}$. In these variables let $F = x^3 t_1 + x^2 y t_2 + \cdots + xyz t_{10}$, so that the coefficients of the t_i are precisely the monomials of degree 3 in x, y, z. Then $\operatorname{rank} C^0_F = \operatorname{rank} C^4_F = 1$ and $\operatorname{rank} C^1_F = \operatorname{rank} C^3_F = 13$ while $\operatorname{rank} C^2_F = 12$. Thus $r(F) \geq 13$. See Example 3.8.

Example 2.7. We briefly present the nonunimodal example of Bernstein–Iarrobino [BI92], a form of degree 16 in 5 variables, which we denote x, y, z, s, t. Let $F = Gs + Ht$ where G, H are general forms of degree 15 in x, y, z. The ranks of catalecticants of F are as follows:

a	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$\operatorname{rank} C^a_F$	1	5	12	22	35	51	70	91	90	91	70	51	35	22	12	5	1

Thus $r(F) \geq 91$. See Example 3.9.

In fact, the examples of Stanley and Bernstein–Iarrobino are merely the first two members of a much larger family. See [BL94].

Definition 2.8. A polynomial $F \in S^d V$ is called **concise** (with respect to V) if it satisfies the following equivalent conditions:

(1) F cannot be written as a polynomial in a smaller number of variables; that is, if $F \in S^d V'$ for some $V' \subseteq V$, then $V' = V$.
(2) The projective hypersurface $V(F)$ is not a cone.
(3) C^{d-1}_F is surjective, C^1_F is injective.

It would be interesting to have a similar geometric characterization of the condition that C^{d-k}_F is surjective for $k \geq 2$.

Following [Car06b], let the **span** of F, denoted $\langle F \rangle \subseteq V$, be the image of C^{d-1}_F. Then $F \in S^d \langle F \rangle$. Elements of $\langle F \rangle$ are called **essential variables** of F, and F is said to **depend essentially** on k variables if $k = \dim \langle F \rangle$.

Definition 2.9. For each $1 \leq k < n$, the **subspace variety** $\mathrm{Sub}_k \subset \mathbb{P} S^d V$ is the locus of forms depending essentially on k or fewer variables:

$$\mathrm{Sub}_k = \{[f] \mid \dim \langle f \rangle \leq k\} = \{[f] \mid \operatorname{rank} C^{d-1}_F \leq k\}.$$

Note that, upon choosing bases for $S^{d-1} V^*$ and $S^1 V \cong V$, Sub_k is the zero locus of the $(k+1)$-minors of the matrix C^{d-1}_F, whose entries are polynomials in the coefficients of F. So Sub_k is a projective variety.

We have $\nu_d(\mathbb{P}V) = \mathrm{Sub}_1 \subset \mathrm{Sub}_2 \subset \cdots$. Each Sub_k contains, but may be strictly larger than, the secant variety $\sigma_k(\nu_d(\mathbb{P}V))$. Sub_{n-1} is precisely the locus of non-concise forms. See Examples 1.9, 1.10.

2.2. **Simultaneous Waring rank.** Let $W \subseteq S^d V$ be a linear series of degree d forms. Just as the simultaneous Waring rank of W is a special case of multihomogeneous Waring rank, the catalecticants of W are special cases of catalecticants of multihomogeneous forms, which we discuss next. Nevertheless we pause to examine this special case before we go on.

We define two apparently different types of catalecticants (but they will be unified below, see Example 2.14). The $(0,a)$th catalecticant $C_W^{(0,a)}$ is the linear map $S^a V^* \to W^* \otimes S^{d-a} V = \mathrm{Hom}(W, S^{d-a}V)$ sending $D \mapsto (F \mapsto DF)$. The $(1,a)$th catalecticant $C_W^{(1,a)}$ is the linear map $W \otimes S^a V^* \to S^{d-a} V$ sending $F \otimes D \mapsto DF$. Thus the image of $C_W^{(1,a)}$ is the subspace spanned by the images of the C_F^a for $F \in W$.

We have $r(W) \geq \mathrm{rank}(C_W^{(1,a)})$, with essentially the same proof as for classical Waring rank: If W is contained in the span of $\ell_1^d, \dots, \ell_r^d$, then every ath derivative of every element of W is contained in the span of $\ell_1^{d-a}, \dots, \ell_r^{d-a}$.

Similarly $r(W) \geq \mathrm{rank}(C_W^{(0,a)})$: If W is contained in the span of $\ell_1^d, \dots, \ell_r^d$ then $C_W^{(0,a)}(D)$ is determined by $D\ell_1^d, \dots, D\ell_r^d$ which are in turn determined by $D\ell_1^a, \dots, D\ell_r^a \in \Bbbk$, so the image of $C_W^{(0,a)}$ has dimension at most r.

As before, it is easy to see that $C_W^{(1,a)}$ and $C_W^{(0,d-a)}$ are transposes of each other.

Remark 2.10. For a deep investigation of the ranks of catalecticants of linear series, see [GHMS07].

Example 2.11. If D_k is the linear series spanned by k-minors of a generic $m \times n$ matrix, as in Example 1.4, then the image of $C_{D_k}^{(1,a)}$ is spanned by $(k-a)$-minors. That is, the image of $C_{D_k}^{(1,a)}$ is just D_{k-a}. Similarly, when P_k is the linear series spanned by permanents of $k \times k$ submatrices and R_k is the linear series spanned by degree k rook-free products in a generic $m \times n$ matrix, then the image of the $(1,a)$th catalecticant of P_k is P_{k-a} and the image of the $(1,a)$th catalecticant of R_k is R_{k-a}. Thus

$$\mathrm{rank}\, C_{D_k}^{(1,a)} = \mathrm{rank}\, C_{P_k}^{(1,a)} = \binom{m}{k-a}\binom{n}{k-a},$$
$$\mathrm{rank}\, C_{R_k}^{(1,a)} = \binom{m}{k-a}\binom{n}{k-a}(k-a)!.$$

In particular

$$r(D_k) \geq \max_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a}, \qquad r(P_k) \geq \max_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a}, \qquad r(R_k) \geq \max_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a} a!.$$

Note that $\binom{m}{a}\binom{n}{a}$ is maximized when $a = \lceil \frac{mn-1}{m+n+2} \rceil$. When $m = n$ this is $\lceil \frac{n-1}{2} \rceil$. Indeed,

$$\frac{\binom{m}{a+1}\binom{n}{a+1}}{\binom{m}{a}\binom{n}{a}} = \frac{(m-a)(n-a)}{(a+1)^2}$$

and this > 1 if and only if $mn - (m+n)a > 2a + 1$, that is, $(m+n+2)a < mn - 1$. In particular the sequence $\{\binom{m}{a}\binom{n}{a} \mid a = 0, \dots, \min(m,n)\}$ is unimodal.

Similarly, $\binom{m}{a}\binom{n}{a}a!$ is maximized when

$$a = \left\lceil \frac{m+n+1-\sqrt{(m+n+1)^2-4(mn-1)}}{2} \right\rceil .$$

When $m = n$ this is $\left\lceil \frac{2n+1-\sqrt{4n+2}}{2} \right\rceil$.

See Examples 3.13, 3.14.

2.3. **Multihomogeneous polynomials.** At this point it is convenient to introduce multinomial notation.

Notation 2.12. Fix $s \geq 1$ and vector spaces $V_1 \cong \Bbbk^{n_1}, \dots, V_s \cong \Bbbk^{n_s}$. We denote $\mathbf{n} = (n_1, \dots, n_s)$. For $\mathbf{d} = (d_1, \dots, d_s)$, where $0 \leq d_i$ for $1 \leq i \leq s$, we define

$$S^{\mathbf{d}}V = S^{d_1}V_1 \otimes \cdots \otimes S^{d_s}V_s$$

and the dual space $S^{\mathbf{d}}V^*$ similarly. Also, we let $\mathbf{0} = (0, \dots, 0)$, with s entries.

The sum and difference of s-tuples is elementwise. We partially order s-tuples by elementwise comparison; thus $\mathbf{0} \leq \mathbf{a} \leq \mathbf{d}$ means that for each i, $0 \leq a_i \leq d_i$.

Definition 2.13. Fix $\mathbf{n}$ and $\mathbf{d}$ as above. Let $M \in S^{\mathbf{d}}V$ be a multihomogeneous polynomial of multidegree $\mathbf{d}$. For each $\mathbf{0} \leq \mathbf{a} \leq \mathbf{d}$ the $\mathbf{a}$'th catalecticant of M is the linear map

$$C_M^{\mathbf{a}} : S^{\mathbf{a}}V^* \to S^{\mathbf{d}-\mathbf{a}}V,$$

that is

$$C_M^{\mathbf{a}} : \prod_{i=1}^{s} S^{a_i}(V_i^*) \to \prod_{i=1}^{s} S^{d_i - a_i}(V_i)$$

given by the usual contraction or flattening.

In detail, and for the sake of concreteness, suppose for each i we fix a basis $\{x_{i,1}, \dots, x_{i,n_i}\}$ for V_i. We denote the dual basis for V_i^* by $\{\partial_{i,1}, \dots, \partial_{i,n_i}\}$. A polynomial in the $\partial_{i,m}$ acts as a differential operator on $S^{\mathbf{d}}V$ (where each $\partial_{i,m}$ acts as $\partial/\partial x_{i,m}$). The catalecticant $C_M^{\mathbf{a}}$ takes a differential operator D to its evaluation DM on M, regarded as a multihomogeneous polynomial. When D is homogeneous of multidegree $\mathbf{a}$, DM has multidegree $\mathbf{d} - \mathbf{a}$.

The bound

$$r_{MH}(M) \geq \operatorname{rank} C_M^{\mathbf{a}} \tag{9}$$

is well-known. The proof is essentially the same as for classical Waring rank and simultaneous Waring rank. Suppose M is written as a sum of r terms each of the form $\ell_1^{d_1} \cdots \ell_s^{d_s}$, where each $\ell_i \in V_i$. If D is multihomogeneous of multidegree $\mathbf{a}$ then $D(\ell_1^{d_1} \cdots \ell_s^{d_s}) = c\ell_1^{d_1-a_1} \cdots \ell_s^{d_s-a_s}$ for some constant c. So DM is a linear combination of the r terms of the form $\ell_1^{d_1-a_1} \cdots \ell_s^{d_s-a_s}$, which means the image of $C_M^{\mathbf{a}}$ is contained in the linear span of these r terms. Thus $r = r_{MH}(M) \geq \operatorname{rank} C_M^{\mathbf{a}}$.

Example 2.14. Given a linear series $W \subseteq S^dV$, we regard the associated bihomogeneous form M, defined in §1.3, as an element of $S^1W^* \otimes S^dV$. Then the catalacticants of W defined earlier agree with the catalecticants of M defined here: $C_W^{(0,a)} = C_M^{(0,a)}$ and $C_W^{(1,a)} = C_M^{(1,a)}$.

Example 2.15. Let $F = x_1 \cdots x_a y_1 \cdots y_b$, a bihomogeneous form of bidegree $\mathbf{d} = (a, b)$ in the x and y variables, as in Example 1.6. Let $\mathbf{a} = (p, q)$ with $0 \le p \le a$, $0 \le q \le b$. Then the image of $C_F^{\mathbf{a}}$ is spanned by monomials which are products of $a-p$ of the x variables and $b-q$ of the y variables. So $r_{MH(a,b)}(F) \ge \operatorname{rank} C_F^{\mathbf{a}} = \binom{a}{p}\binom{b}{q}$.

In particular, when $a = b = 2$, $r_{MH(2,2)}(x_1x_2y_1y_2) \ge \binom{2}{1}^2 = 4$. We saw before $r_{MH(a,b)}(F) \le 2^{a+b-2}$. Thus $r_{MH(2,2)}(x_1x_2y_1y_2) = 4$.

See Example 3.21.

Example 2.16. The generic determinant $\det_n$ is bihomogeneous of bidegree $\mathbf{d} = (a, n-a)$ in the sets of variables in the first a and last $n-a$ rows, as in Example 1.8. For $\mathbf{a} = (p, q)$ with $0 \le p \le a$, $0 \le q \le n-a$, the image of the catalecticant $C_{\det_n}^{\mathbf{a}}$ is spanned by $(n-p-q)$-minors with $a-p$ rows in the first a rows of the matrix, $n-a-q$ rows in the last $n-a$ rows of the matrix, and any $n-p-q$ columns. Thus $r_{MH(a,n-a)}(\det_n) \ge \operatorname{rank} C_{\det_n}^{\mathbf{a}} = \binom{a}{p}\binom{n-a}{q}\binom{n}{p+q}$.

See Example 3.22.

2.4. **Generalized catalecticants.** Recently Landsberg and Ottaviani [LO12] gave a generalization of Sylvester's catalecticants to the setting of generalized rank. We recall their definition here.

Let X be a projective variety and L a very ample line bundle on X. Let $V = H^0(X, L)^*$, so that L naturally embeds $X \subset \mathbb{P}V$. Fix $v \in V$, $v \ne 0$; we aim to determine, or bound, $r_X(v)$. Let E be a vector bundle on X whose rank (fiber dimension) we denote by $\operatorname{rank} E$. Landsberg and Ottaviani's generalized catalecticant of v with respect to E is a map denoted C_v^E (they use the letter A but we prefer C for *catalecticant*) and defined as follows. From the natural map $E \otimes E^* \to \mathcal{O}_X$ we get $E \otimes (L \otimes E^*) \to L$ and then in turn a multiplication map on global sections, $H^0(E) \otimes H^0(L \otimes E^*) \to H^0(L) = V^*$. We regard $v \in V$ as a map $V^* \to \Bbbk$, so composing gives a bilinear map $H^0(E) \otimes H^0(L \otimes E^*) \to \Bbbk$. The **generalized catalecticant of v with respect to E** is the resulting map $C_v^E : H^0(E) \to H^0(L \otimes E^*)^*$. That is, $C_v^E : s \mapsto (t \mapsto v(st))$.

The catalecticants C_v^E and $C_v^{L \otimes E^*}$ are transposes of one another.

Example 2.17. The classical catalecticant C_F^a for a homogeneous form F corresponds to $X = \mathbb{P}V$, $L = \mathcal{O}_{\mathbb{P}V}(d)$, and $E = \mathcal{O}_{\mathbb{P}V}(a)$. The embedding $X \subset \mathbb{P}H^0(X, L)^* = \mathbb{P}S^dV$ is the Veronese map, taking X to $\nu_d(\mathbb{P}V)$.

The multihomogeneous catalecticant $C_M^{\mathbf{a}}$ for a multihomogeneous form M of multidegree $\mathbf{d} = (d_1, \ldots, d_s)$ corresponds to the following. Let $X = \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_s$ and for each i let $\operatorname{pr}_i : X \to \mathbb{P}V_i$ be the projection onto the ith factor. Then $C_M^{\mathbf{a}}$ corresponds to the line bundles $L = \mathcal{O}_{\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_s}(d_1, \ldots, d_s) = \operatorname{pr}_1^* \mathcal{O}_{\mathbb{P}V_1}(d_1) \otimes \cdots \otimes \operatorname{pr}_s^* \mathcal{O}_{\mathbb{P}V_s}(d_s)$ and $E = \mathcal{O}_{\mathbb{P}V_1 \times \cdots \times \mathbb{P}V_s}(a_1, \ldots, a_s)$. The embedding of X is the Segre–Veronese map, taking X to $\operatorname{Seg}(\nu_{d_1}(\mathbb{P}V_1) \times \cdots \times \nu_{d_s}(\mathbb{P}V_s))$.

The following proposition is essentially just Propositions 5.1.1 and 5.4.1 in [LO12]. (Their Proposition 5.4.1 is more general, working with multipoint restrictions rather than just one point.)

Proposition 2.18. *If $[v] = x \in X$, then $\operatorname{rank} C_v^E \le \operatorname{rank} E$, with equality if and only if both E and $L \otimes E^*$ are globally spanned at x; that is, the maps $H^0(E) \to E_x$ and $H^0(L \otimes E) \to (L \otimes E^*)_x = L_x \otimes E_x^*$ are surjective. For all $0 \ne v \in V$,*

$$r_X(v) \ge \frac{\operatorname{rank} C_v^E}{\operatorname{rank} E}. \tag{10}$$

Proof. If $[v] = x \in X$, C_v^E corresponds to the bilinear functional $H^0(E) \otimes H^0(L \otimes E^*) \to H^0(L) \to \Bbbk$ where the first map is multiplication of global sections and the second map is evaluation at $x \in X$. This commutes with restriction to the fiber at x, namely $E_x \otimes (L \otimes E^*)_x \to L_x \cong \Bbbk$. Therefore the map C_v^E is equal to the composition of the restriction $H^0(E) \to E_x$, followed by the identification $E_x \cong L_x^* \otimes E_x \cong (L \otimes E^*)_x^*$ (up to a choice of scalar, i.e., a basis for $L_x \cong \Bbbk^1$), followed by the transpose of the restriction map, $(L \otimes E^*)_x^* \to H^0(L \otimes E^*)^*$. That is, C_v^E is exactly the natural map

$$H^0(E) \to E_x \cong L_x^* \otimes E_x \cong (L \otimes E^*)_x^* \to H^0(L \otimes E^*)^*.$$

Since $\dim(L \otimes E^*)_x = \dim E_x^* = \operatorname{rank} E$, the composition of these maps has rank at most $\operatorname{rank} E$, with equality if and only if the first map is surjective and the last is injective.

If $v = v_1 + \cdots + v_r$ then $C_v^E = \sum C_{v_i}^E$. This gives (10). □

In particular, when $r_X(v) = r$ then the size $(re+1)$ minors of C_v^E vanish, where $e = \operatorname{rank} E$. Thus the size $(re+1)$ minors of C_v^E vanish on a dense subset of the rth secant variety $\sigma_r(X)$, hence on all of $\sigma_r(X)$. Therefore these minors give equations for $\sigma_r(X)$. One can thus obtain equations for secant varieties by producing suitable bundles E, which is pursued in [LO12] by using representation theory.

More generally, let $\mathcal{E}$ be an $\mathcal{O}_X$-module whose fibers have bounded dimension: that is, there is a positive integer e such that for each $x \in X$, the fiber $\mathcal{E}_x = \mathcal{E} \otimes \Bbbk(x)$ (where $\Bbbk(x) = \mathcal{O}_{X,x}/\mathfrak{m}_x$) has dimension at most e, as a vector space over $\Bbbk(x) \cong \Bbbk$ (since we have assumed $\Bbbk$ is algebraically closed). Let $\mathcal{E}^\vee = \mathcal{H}om_{\mathcal{O}_X}(\mathcal{E}, \mathcal{O}_X)$. Then each fiber of $\mathcal{E}^\vee$ also has vector space dimension at most e. As before, the natural multiplication map $\mathcal{E} \otimes \mathcal{E}^\vee \to \mathcal{O}_X$ yields $\mathcal{E} \otimes (L \otimes \mathcal{E}^\vee) \to L$, which on global sections yields in turn $H^0(X, \mathcal{E}) \otimes H^0(X, L \otimes \mathcal{E}^\vee) \to H^0(L) = V^*$. Once again our fixed $v \in V$ gives a map $V^* \to \Bbbk$, so we have a bilinear functional $H^0(X, \mathcal{E}) \otimes H^0(X, L \otimes \mathcal{E}^\vee) \to \Bbbk$. We define the **generalized catalecticant of v with respect to $\mathcal{E}$** to be $C_v^{\mathcal{E}} : H^0(X, \mathcal{E}) \to H^0(X, L \otimes \mathcal{E}^\vee)^*$.

If $\mathcal{E}$ is reflexive then $C_v^{\mathcal{E}}$ and $C_v^{L \otimes \mathcal{E}^\vee}$ are transposes of one another.

Once again if $[v] = x \in X$ then $C_v^{\mathcal{E}}$ factors through restriction of global sections to the fiber at x:

$$H^0(X, \mathcal{E}) \to \mathcal{E}_x \to ((L \otimes \mathcal{E}^\vee)_x)^* \to H^0(X, L \otimes \mathcal{E}^\vee)^*.$$

Hence if $[v] = x \in X$ then $\operatorname{rank} C_v^{\mathcal{E}} \le \dim \mathcal{E}_x = \dim(L \otimes \mathcal{E}^\vee)_x \le e$, with equality holding in the first inequality if and only if both $\mathcal{E}$ and $L \otimes \mathcal{E}^\vee$ are globally generated at x. Therefore

$$r_X(v) \ge \frac{\operatorname{rank} C_v^{\mathcal{E}}}{e}.$$

Example 2.19. Let $X = \mathrm{Sub}_k \subset \mathbb{P}S^d V$, the subspace variety of forms depending essentially on k or fewer variables. For $[F] \in X$, the catalecticant C_F^{d-1} can be written as a matrix whose entries depend linearly on the coefficients of F. So we have a map of $\mathcal{O}_X$-modules

$$C^{d-1} : S^{d-1}V^* \otimes \mathcal{O}_X(-1) \to V \otimes \mathcal{O}_X$$

given by $C_F^{d-1} : S^{d-1}V^* \to V$ at each $[F] \in X$. The Sub_k are degeneracy loci of the map C^{d-1}. Then the kernel and cokernel of C^{d-1} are naturally associated sheaves, which are vector bundles on each $\mathrm{Sub}_k \setminus \mathrm{Sub}_{k-1}$. It would be interesting to work out the global sections and catalecticants corresponding to these sheaves. Perhaps they might be related to the "Apolarity Lemma" discussed below, see Theorem 4.10.

3. Improved lower bounds

An improvement to the catalecticant lower bound for classical Waring rank was found in [LT10]. It raises the lower bound $\operatorname{rank} C_F^a$ by the dimension of a certain set of singularities of the hypersurface $V(F)$. As our goal is to give a similar improvement more generally we first recall that result.

3.1. **Classical Waring rank.**

Definition 3.1. Let $F \in S^d V$ and $0 \leq a < d$. We define $\Sigma_a(F) \subseteq \mathbb{P}V^*$ to be the common zero locus of all the ath mixed partial derivatives of F; that is, the projective variety defined by the (ideal generated by the) image of the catalecticant C_F^a. As a set, $\Sigma_a(F)$ is the set of points at which F vanishes to order at least $a+1$.

We write $\widehat{\Sigma}_a(F)$ for the affine cone over $\Sigma_a(F)$. Explicitly, $\widehat{\Sigma}_a(F)$ is the common zero locus in V^* of the ath mixed partial derivatives of F; in particular, for $a < d$, the origin $0 \in \widehat{\Sigma}_a(F)$ even if $\Sigma_a(F)$ is empty.

Thus $\Sigma_0(F)$ is the projective hypersurface $V(F)$ and $\Sigma_1(F)$ is the singular locus of $V(F)$, similarly $\widehat{\Sigma}_0(F)$ is the affine hypersurface defined by F and $\widehat{\Sigma}_1(F)$ is its singular locus.

Note that F is concise (Definition 2.8) if and only if $\widehat{\Sigma}_{d-1}(F) = \{0\}$, equivalently $\Sigma_{d-1}(F) = \varnothing$.

Theorem 3.2 (Theorem 1.3 of [LT10])**.** *Let $F \in S^d V$ and let $0 \leq a < d$. If F is concise then*

$$r(F) \geq \operatorname{rank} C_F^{d-a} + \dim \widehat{\Sigma}_a(F). \tag{11}$$

In [LT10] this is stated as

$$r(F) \geq \operatorname{rank} C_F^{d-a} + \dim \Sigma_a(F) + 1,$$

with the understanding that $\dim \varnothing = -1$. Although the affine cone version (11) is simpler (and avoids conventions about negative dimensions), this "projective" one is the version that will generalize.

We review the proof given in [LT10]. First, recall that the differential operator D_p associated to $p \in S^{d-a}V^*$ satisfies $D_p(\ell^d) = \left(\frac{d!}{a!}\right) \ell^a \, p(\ell)$, so $D_p(\ell^d) = 0$ if and only if $p([\ell]) = 0$. We begin with the following well-known statement (see, for example, [ER93, Proposition 4.1]).

Proposition 3.3. *Let $h \in V^*$. Then $h^{d-a} \in \ker C_F^{d-a}$ if and only if $h \in \widehat{\Sigma}_a(F)$.*

Proof. $h^{d-a} \in \ker C_F^{d-a}$ if and only if $h^{d-a}F = 0$, if and only if $\Theta h^{d-a} F = 0$ for all $\Theta \in S^a V^*$, if and only if $h^{d-a}\Theta F = 0$ for all Θ. Since $\deg \Theta F = d-a$, $h^{d-a}\Theta F$ is just the evaluation of ΘF at the point h, $h^{d-a}\Theta F = \Theta F\big|_h$ (up to a factorial factor). So $h^{d-a} \in \ker C_F^{d-a}$ if and only if $\Theta F\big|_h = 0$ for all $\Theta \in S^a V^*$, if and only if F vanishes to order at least $a+1$ at h, if and only if $h \in \widehat{\Sigma}_a(F)$. $\square$

Remark 3.4. Proposition 3.3 verges on tautology: h is in $\widehat{\Sigma}_a(F)$ if and only if the ath derivatives of F vanish at h, if and only if h is a common zero of the forms in the image of C_F^a, if and only if $h^{d-a} \in (\operatorname{img} C_F^a)^\perp = \ker C_F^{d-a}$.

Proof of Theorem 3.2. Suppose $F = \ell_1^d + \cdots + \ell_r^d$. Let $\mathcal{L} = \{p \in S^{d-a}V^* : p([\ell_1]) = \cdots = p([\ell_r]) = 0\}$, the linear series of degree $d-a$ forms vanishing at the points $[\ell_i] \in \mathbb{P}V$. Since $D_p(\ell_i^d) = 0$ if and only if $p([\ell_i]) = 0$, we have $\mathcal{L} \subseteq \ker C_F^{d-a}$. This shows that

$$r \geq \operatorname{codim} \mathcal{L} \geq \operatorname{codim} \ker C_F^{d-a} = \operatorname{rank} C_F^{d-a},$$

the first inequality holding since $\mathcal{L}$ is defined by r linear conditions.

Now we use the hypothesis that F is concise. Because of this, the $[\ell_i]$ can not lie on a hyperplane, or else F could be written in fewer variables. Hence $\mathcal{L}$ can not contain any power $p = h^{d-a}$ of a linear form h; for if $h^{d-a} \in \mathcal{L}$, then $h^{d-a}([\ell_i]) = 0$ for each i, so $h([\ell_i]) = 0$ for each i, and the $[\ell_i]$ would lie on the hyperplane defined by h, a contradiction. Therefore the projectivization $\mathbb{P}\mathcal{L}$ is disjoint from the Veronese $v_{d-a}(\mathbb{P}V^*)$. Since $\mathcal{L} \subseteq \ker C_F^{d-a}$, we disregard everything lying outside this kernel and observe that $\mathbb{P}\mathcal{L}$ is disjoint from $\mathbb{P} \ker C_F^{d-a} \cap v_{d-a}(\mathbb{P}V^*)$. Thus

$$\dim(\mathbb{P}\mathcal{L}) + \dim(\mathbb{P} \ker C_F^{d-a} \cap v_{d-a}(\mathbb{P}V^*)) < \dim \mathbb{P} \ker C_F^{d-a},$$

or else a nonempty intersection would be forced. We subtract each side of this inequality from the ambient dimension $\dim \mathbb{P}S^{d-a}V^*$ and rearrange to get

$$r \geq \operatorname{codim} \mathcal{L} = \operatorname{codim}(\mathbb{P}\mathcal{L}) > \operatorname{rank} C_F^{d-a} + \dim(\mathbb{P} \ker C_F^{d-a} \cap v_{d-a}(\mathbb{P}V^*)).$$

Finally $\mathbb{P} \ker C_F^{d-a} \cap v_{d-a}(\mathbb{P}V^*) \cong \Sigma_a(F)$ by Proposition 3.3. $\square$

Remark 3.5. Since C_F^a and C_F^{d-a} are transposes of one another, we could of course say $r(F) \geq \operatorname{rank} C_F^a + \dim \widehat{\Sigma}_a(F)$. By Remark 2.5, it is impossible (or at least difficult) to say for which a the maximum of $\operatorname{rank} C_F^a$ occurs. On the other hand, since $\Sigma_0(F) \supseteq \Sigma_1(F) \supseteq \cdots$, the dimensions of $\widehat{\Sigma}_a(F)$ are clearly nonincreasing. A priori the maximum value of $\operatorname{rank} C_F^a + \dim \widehat{\Sigma}_a(F)$ could occur at a different a than the maximum value of $\operatorname{rank} C_F^a$, but I do not know any example in which this happens.

Example 3.6. We saw $r(xyz) \leq 4$ and $\operatorname{rank} C_{xyz}^1 = 3$ in Example 2.3. We have $\Sigma_1(xyz) = \operatorname{Sing} V(xyz) = \{[1:0:0],[0:1:0],[0:0:1]\}$ so $\dim \Sigma_1(xyz) = 0$, $\dim \widehat{\Sigma}_1(xyz) = 1$. Therefore $r(xyz) \geq 3 + 1 = 4$.

Example 3.7. We saw $\operatorname{rank} C_{\det_n}^a = \binom{n}{a}^2$ in Example 2.4. And $\det_n$ vanishes to order $a+1$ at a point (matrix) M if and only if every $(n-a)$-minor of M vanishes, that is, M has rank $n-a-1$ or less. The dimension of the locus of matrices of rank $n-a-1$ or less is $n^2 - 1 - (a+1)^2$. So, for each a, $r(\det_n) \geq \binom{n}{a}^2 + n^2 - (a+1)^2$. This is maximized at $a = \lfloor n/2 \rfloor$, so $r(\det_n) \geq \binom{n}{\lfloor n/2 \rfloor}^2 + n^2 - (\lfloor n/2 \rfloor + 1)^2$.

For instance, $r(\det_3) \geq 14$. See [LT10, §9]. For other bounds on $r(\det_n)$ and ranks of permanents, Pfaffians, etc., see [Sha14], [Sha13].

Example 3.8. We consider Stanley's nonunimodal example, see Example 2.6. Since F vanishes to order 3 on the linear subspace $x = y = z = 0$, $\widehat{\Sigma}_2(F) \supseteq V(x,y,z)$, a 10-dimensional subspace. On the other hand, $\widehat{\Sigma}_2(F) \subseteq \widehat{\Sigma}_1(F) \subseteq V(x,y,z)$, since $\partial_{t_1} F = x^3$ so $\partial_{t_1} F = 0$ implies $x = 0$, and so on. So $\widehat{\Sigma}_1(F) = \widehat{\Sigma}_2(F) = V(x,y,z)$. Thus $r(F) \geq \operatorname{rank} C_F^2 + \dim \widehat{\Sigma}_2(F) = 22$, and better, $r(F) \geq \operatorname{rank} C_F^1 + \dim \widehat{\Sigma}_1(F) = 23$.

Example 3.9. We consider the nonunimodal example of Bernstein–Iarrobino, see Example 2.7. Since F vanishes to order 15 on the subspace $x = y = z = 0$, $\dim \widehat{\Sigma}_a(F) \geq 2$ for $0 \leq a \leq 14$.

Now we examine $\widehat{\Sigma}_1(F)$. It contains the 2-plane $x = y = z = 0$. Suppose $(x, y, z) \neq (0, 0, 0)$ and $p = (x, y, z, s, t) \in \widehat{\Sigma}_1(F)$. Since $\partial_s F = \partial_t F = 0$, the projective point $q = [x : y : z]$ must lie in the complete intersection $G = H = 0$, a finite set of 225 points. Furthermore the plane curve $sG + tH$ must be singular at q. We have

$$(\partial_x F, \partial_y F, \partial_z F) = s\nabla G + t\nabla H.$$

Since G, H are general, they intersect transversally at q, with independent gradients. This forces $s = t = 0$. So $\widehat{\Sigma}_1(F)$ is the union of the 2-plane $x = y = z = 0$ and 225 lines with $s = t = 0$, spanned by the points in the complete intersection $G = H = 0$. Therefore $\dim \widehat{\Sigma}_a(F) = 2$ for $0 \leq a \leq 14$.

Hence $r(F) \geq \operatorname{rank} C_F^7 + \dim \widehat{\Sigma}_7(F) = 93$.

3.2. **Simultaneous Waring rank.**

Definition 3.10. Let $W \subseteq S^d V$ be a linear series of degree d forms. For $0 \leq a$, let $\Sigma_{(1,a)}(W) = \bigcap_{F \in W} \Sigma_a(F)$. That is, $\Sigma_{(1,a)}(W)$ is the locus of points in $\mathbb{P}V^*$ at which every member of the linear series W vanishes to order at least $a + 1$. As a scheme, $\Sigma_{(1,a)}(W)$ is defined by the vanishing of all the ath partial derivatives of all the members of W; that is, by the image of the map $C_W^{(1,a)}$.

Definition 3.11. A linear series $W \subset S^d V$ is called **concise** (with respect to V) if it satisfies the following equivalent conditions:

1. (1) W cannot be written as a linear series in a smaller number of variables; that is, if $W \subseteq S^d V'$ for some $V' \subseteq V$, then $V' = V$.
2. (2) If $V' \subset V$ is such that for every $F \in W$, $V(F)$ is a cone over $\mathbb{P}V'$, then $V' = 0$.
3. (3) The set $\widehat{\Sigma}_{(1,d-1)}(W) = \{0\}$.
4. (4) A general $F \in W$ is concise.
5. (5) $C_W^{(1,d-1)}$ is surjective, $C_W^{(0,1)}$ is injective.

Theorem 3.12. *Let $W \subseteq S^d V$ and $0 \leq a < d$. If W is concise then*

$$r(W) \geq \operatorname{rank} C_W^{(0,d-a)} + \dim \widehat{\Sigma}_{(1,a)}(W).$$

Proof. Suppose W is contained in the span of $\ell_1^d, \ldots, \ell_r^d$. Let $\mathcal{L}$ be the linear series of degree $d - a$ forms vanishing at the $[\ell_i]$. Then $\mathcal{L} \subseteq \ker C_W^{(0,d-a)}$. Indeed for $\Theta \in \mathcal{L}$ and $F \in W$, say $F = \sum c_i \ell_i^d$, we have $C_W^{(0,d-a)}(\Theta)(F) = \Theta F = \sum c_i \Theta \ell_i^d = 0$. This already shows $r \geq \operatorname{codim} \mathcal{L} \geq \operatorname{codim} \ker C_W^{(0,d-a)} = \operatorname{rank} C_W^{(0,d-a)}$.

By hypothesis the $[\ell_i]$ do not lie on any hyperplane, so $\mathcal{L}$ does not contain any power of a hyperplane h^{d-a}. Thus $\mathcal{L}$ is disjoint from the Veronese variety $\nu_{d-a}(\mathbb{P}V^*)$. As before, $r \geq \operatorname{codim} \mathcal{L} > \operatorname{codim} \ker C_W^{(0,d-a)} + \dim(\mathbb{P} \ker C_W^{(0,d-a)} \cap \nu_{d-a}(\mathbb{P}V^*))$.

We claim that $\mathbb{P} \ker C_W^{(0,d-a)} \cap \nu_{d-a}(\mathbb{P}V^*) \cong \Sigma_{(1,a)}(W)$. Let $h \in \mathbb{P}V^*$. Then $h^{d-a} \in \ker C_W^{(0,d-a)}$ if and only if $h^{d-a} F = 0$ for all $F \in W$, if and only if for all $F \in W$, F vanishes at $[h]$ to order at least $a + 1$, if and only if $[h] \in \bigcap_{F \in W} \Sigma_a(F) = \Sigma_{(1,a)}(W)$. $\square$

Example 3.13. Fix a generic $m \times n$ matrix $X = (x_{i,j})$, $1 \le i \le m$, $1 \le j \le n$. Let D_k, P_k, R_k be as in Example 1.4. The image of $C^{(1,a)}_{D_k}$ is D_{k-a}, and similarly for P_k and R_k, by Example 2.11, which also contains the ranks of those catalecticants.

Thus $\Sigma_{(1,a)}(D_k)$ is the common zero locus of D_{k-a}, the locus of matrices of rank $\le k-a-1$. This has dimension $mn - (m-(k-a-1))(n-(k-a-1)) = (k-a-1)(m+n-k+a+1)$. Therefore $r(D_k) \ge \binom{m}{k-a}\binom{n}{k-a} + (k-a-1)(m+n-k+a+1)$.

$\Sigma_{(1,a)}(P_k)$ is the common zero locus of P_{k-a}, the locus of matrices all of whose $(k-a)\times(k-a)$ submatrices have permanent zero. This includes at least all matrices with only $k-a-1$ nonzero columns or $k-a-1$ nonzero rows, so $\dim \Sigma_{(1,a)}(P_k) \ge \max(m,n)(k-a-1)$ and $r(P_k) \ge \binom{m}{k-a}\binom{n}{k-a} + \max(m,n)(k-a-1)$. See [Yu99], [LS00], [Kir08], [Stu02, §5.4] for more on the zero locus of permanental ideals.

We treat R_k separately in the next example.

Example 3.14. Fix a generic $m \times n$ matrix $X = (x_{i,j})$, $1 \le i \le m$, $1 \le j \le n$, and let R_k be as in Example 1.4. As in the previous example, $\Sigma_{(1,a)}(R_k)$ is the common zero locus of R_{k-a}. We claim that $\dim \Sigma_{(1,a)}(R_k) = \dim V(R_{k-a}) = \max(m,n)(k-a-1)$.

Note the ideal $\langle R_t \rangle$ is a squarefree monomial ideal, so it can be written as the Stanley-Reisner ideal I_Δ of a simplicial complex Δ on the entries of $X = (x_{i,j})$, $1 \le i \le m$, $1 \le j \le n$. That is, $\langle R_t \rangle$ is generated by the minimal nonfaces of Δ; see for example [MS05, §1.1]. We claim that $\langle R_t \rangle$ is the Stanley-Reisner ideal of the simplicial complex $\Delta(t)$ whose facets are given by unions of $t-1$ rows and columns of X, that is, each facet is the union of a rows and b columns where $a + b = t - 1$.

Granting that $\langle R_t \rangle = I_{\Delta(t)}$, the largest facets of $\Delta(t)$ include $\max(m,n)(t-1)$ entries of X, so their complements define linear subspaces of that dimension; all other facets correspond to smaller components, showing that $\dim V(R_t) = \max(m,n)(t-1)$, as desired.

Clearly $\langle R_1 \rangle = I_{\Delta(1)}$ (trivially). It holds as well for $t = 2$: a square-free product contains two rook-free elements if and only if it does not lie on a “union” of one row or of one column.

A square-free monomial M corresponds to a subset of the entries of the matrix, or to the set of edges in a subgraph $G = G(M)$ of the complete bipartite graph $K_{m,n}$. Recall König's Theorem (see, e.g., [Die10, Thm. 2.1.1]), which asserts that the smallest size of a vertex cover of the bipartite graph G is equal to the largest size of a matching. A matching in G (a set of pairwise non-adjacent edges) corresponds to a rook-free product contained in M. A vertex cover (a set of vertices incident to every edge) corresponds to a set of rows and columns whose union contains M. Thus a monomial M lies in R_t if and only if M contains a rook-free product of degree t, if and only if $G(M)$ has a matching of size at least t, if and only if $G(M)$ has no vertex cover of size less than t, if and only if M is not contained in any union of $t-1$ rows and columns.

In conclusion, $\dim V(R_t) = \max(m,n)(t-1)$,

$$\dim \Sigma_{(1,a)}(R_k) = \dim V(R_{k-a}) = \max(m,n)(k-a-1),$$

and

$$r(R_k) \ge \binom{m}{k-a}\binom{n}{k-a}(k-a)! + \max(m,n)(k-a-1),$$

for $1 \le a \le k$.

3.3. **Simultaneous Waring rank, second version.** Now we describe an improvement to the bound $r(W) \geq \operatorname{rank} C_W^{(0,a)}$. For the first time, we need to consider the projective version of the singular set Σ (rather than the affine version $\widehat{\Sigma}$).

Definition 3.15. Let $W \subseteq S^d V$ be a linear series of degree d forms. For $0 \leq a < d$, let $\Sigma_{(0,a)}(W) = \{([F],[p]) \in \mathbb{P}W \times \mathbb{P}V^* \mid [p] \in \Sigma_a(F)\}$.

Thus $\Sigma_{(0,a)}(W)$ is the scheme defined by the image of the map $C_W^{(0,a)}$. Indeed, the correspondence $\operatorname{Hom}(W, S^{d-a}V) \cong W^* \otimes S^{d-a}V$ takes a map $\Theta : W \to S^{d-a}V$ to $f_\Theta \in W^* \otimes S^{d-a}V$, the function on $W \times S^{d-a}V^*$ defined by $f_\Theta(F,\psi) = \langle \Theta F, \psi \rangle = \psi \Theta F \in \Bbbk$. In particular for $h \in V^*$, $f_\Theta(F, h^{d-a}) = \Theta F\big|_h$, the evaluation of ΘF at the point h (up to a factorial factor). So f_Θ vanishes at a point $(F,p) \in W \times V^* \hookrightarrow W \times S^{d-a}V^*$ if and only if ΘF vanishes at p. In our case, a point $([F],[p])$ is a common zero of every element in the image of $C_W^{(0,a)}$ if and only if $\Theta F\big|_p = 0$ for every $\Theta \in S^a V^*$, equivalently $[p] \in \Sigma_a(F)$.

In order to improve $r(W) \geq \operatorname{rank} C_W^{(1,a)}$, we needed to assume that the map $C_W^{(1,d-1)}$ was surjective, equivalently that W was concise. Now we will have to assume that the map $C_W^{(0,d-1)} : S^{d-1}V^* \to \operatorname{Hom}(W,V) \cong W^* \otimes V$ is surjective. This is considerably stronger. If the map $C_W^{(0,d-1)}$ is surjective, then every nonzero $F \in W$ is concise (for every nonzero linear form $\ell \in V$ there is a linear map $W \to V$ taking $F \mapsto \ell$, and it is realized as $C_W^{(0,d-1)}\Theta$ for some $\Theta \in S^{d-1}V^*$, which means $\Theta F = \ell$; so C_F^{d-1} is surjective). And of course if every nonzero $F \in W$ is concise (and W is nontrivial) then W is concise.

Clearly having W concise does not imply that every $F \in W$ is concise: the linear series in Example 1.4 are concise but are spanned by forms (determinant, permanent, rook-free product) that only depend on the variables in a submatrix. And having every nonzero $F \in W$ be concise does not imply that $C_W^{(0,d-1)}$ is surjective. Let $n = 2$ and let W be the pencil spanned by x^3y^2, x^2y^3. Every member of W is of the form $x^2y^2(ax+by)$ and this is concise because it is not a perfect power. But $C_W^{(0,4)}$ is not surjective. If $\Theta = a_4\partial_x^4 + a_3\partial_x^3\partial_y + \cdots + a_0\partial_y^4$ has $\Theta x^3y^2 = 0$ then $a_3 = a_2 = 0$, and then $\Theta x^2y^3 = 6a_1 x$; so there is no $\Theta \in S^4V^*$ such that $\Theta x^3y^2 = 0$, $\Theta x^2y^3 = y$.

So $C_W^{(0,d-1)}$ being surjective is a strong condition. Nevertheless it can be met. For example, let $n = 2$ and let W be the pencil spanned by x^4y^2, x^2y^4. Then we have

$$\begin{aligned}
&(\partial_x^3\partial_y^2)(x^4y^2) = 48x, && (\partial_x^3\partial_y^2)(x^2y^4) = 0,\\
&(\partial_x^4\partial_y)(x^4y^2) = 48y, && (\partial_x^4\partial_y)(x^2y^4) = 0,\\
&(\partial_x\partial_y^4)(x^4y^2) = 0, && (\partial_x\partial_y^4)(x^2y^4) = 48x,\\
&(\partial_x^2\partial_y^3)(x^4y^2) = 0, && (\partial_x^2\partial_y^3)(x^2y^4) = 48y,
\end{aligned}$$

which shows that $C_W^{(0,5)}$ is surjective onto $\operatorname{Hom}(W,V)$.

Theorem 3.16. *Let $W \subseteq S^dV$ and $0 \leq a < d$. If $C_W^{(0,d-1)}$ is surjective then*

$$r(W) \geq \operatorname{rank} C_W^{(1,d-a)} + \dim \Sigma_{(0,a)}(W) + 1,$$

where $\dim \varnothing = -1$.

Proof. Suppose W is contained in the span of $\ell_1^d, \dots, \ell_r^d$ and these are linearly independent. Every element of $W \otimes S^{d-a}V^*$ can be written in the form

$$\sum_{i=1}^{r} \ell_i^d \otimes \Theta_i,$$

the Θ_i being uniquely determined by the linear independence of the ℓ_i^d. Let

$$\mathcal{L} = \left\{ \sum_{i=1}^{r} \ell_i^d \otimes \Theta_i \in W \otimes S^{d-a}V^* \;\middle|\; \Theta_1([\ell_1]) = \cdots = \Theta_r([\ell_r]) = 0 \right\}.$$

Then $\mathcal{L} \subseteq \ker C_W^{(1,d-a)}$ clearly. Since $\Theta_i([\ell_i]) = 0$ imposes just one condition on the polynomial Θ_i, we have

$$r \geq \operatorname{codim} \mathcal{L} \geq \operatorname{codim} \ker C_W^{(1,d-a)} = \operatorname{rank} C_W^{(1,d-a)}.$$

Next we claim that $\mathcal{L}$ does not contain any nonzero element of the form $F \otimes h^{d-a}$, $F \in W$, $h \in V^*$. For if $F \otimes h^{d-a} \in \mathcal{L}$, say $F = \sum c_i \ell_i^d$, then

$$F \otimes h^{d-a} = \sum \ell_i^d \otimes c_i h^{d-a} \in \mathcal{L},$$

whence $c_1 h([\ell_1]) = \cdots = c_r h([\ell_r]) = 0$. If each $c_i \neq 0$ this implies the $[\ell_i]$ lie on the hyperplane defined by h, and W is not concise. In general (allowing that some c_i may be zero) all we can say is that

$$C_W^{(1,1)}(F \otimes h) = \sum c_i h(\ell_i^d) = 0,$$

which means $C_W^{(1,1)}$ is not injective and $C_W^{(0,d-1)}$ is not surjective, contradicting the hypothesis.

Finally we have $F \otimes h^{d-a} \in \ker C_W^{(1,d-a)}$ if and only if $h^{d-a}(F) = 0$. Proposition 3.3 implies that this happens if and only if $[h] \in \Sigma_a(F)$, equivalently if and only if $([F],[h]) \in \Sigma_{(0,a)}(W)$. The points $F \otimes h^{d-a}$ correspond precisely to points in the Segre embedding of $\mathbb{P}W \times \nu_{d-a}(\mathbb{P}V^*)$ in $\mathbb{P}(W \otimes S^{d-a}V^*)$. We have shown that the intersection of this Segre variety with the kernel of the catalecticant satisfies

$$\mathbb{P} \ker C_W^{(1,d-a)} \cap \operatorname{Seg}(\mathbb{P}W \times \nu_{d-a}(\mathbb{P}V^*)) \cong \Sigma_{(0,a)}(W).$$

Putting this all together we have

$$\begin{aligned} r &\geq \operatorname{codim} \mathcal{L} \\ &> \operatorname{rank} C_W^{(1,d-a)} + \dim \left\{ \mathbb{P} \ker C_W^{(1,d-a)} \cap \operatorname{Seg}(\mathbb{P}W \times \nu_{d-a}(\mathbb{P}V^*)) \right\} \\ &= \operatorname{rank} C_W^{(1,d-a)} + \dim \Sigma_{(0,a)}(W), \end{aligned}$$

as claimed. □

Example 3.17. Here is an example in which Theorem 3.16 gives a better bound than Theorem 3.12.

Let W be the pencil spanned by x^6y^3 and x^4y^5. We have $r(W) \leq r(x^6y^3) + r(x^4y^5) = 7 + 5 = 12$. Better, $r(W) \leq r(x^6y^3 + x^4y^5) + r(x^6y^3 - x^4y^5) = 5 + 5 = 10$.

For lower bounds, we have

a	0	1	2	3	4	5	6	7	8	9
$\operatorname{rank} C_W^{(1,9-a)}$	1	2	3	4	5	6	6	5	4	2

In particular $C_W^{(1,1)}$ is injective. Other than the spanning elements, elements of W are of the form

$$ax^6y^3 + bx^4y^5 = x^4y^3(ax^2 + by^2)$$

which has a quadruple root at $x = 0$, a triple root at $y = 0$, and two simple roots when $ab \neq 0$. Therefore

$$\Sigma_{(1,a)}(W) = \begin{cases} V(x) \cup V(y), & 0 \leq a \leq 2 \\ V(x), & a = 3 \\ \varnothing, & 4 \leq a \end{cases}$$

and

$$\dim \Sigma_{(0,a)}(W) = \begin{cases} 1, & 0 \leq a \leq 3 \\ 0, & 4 \leq a \leq 5 \\ -1, & 6 \leq a \end{cases}$$

We have

a	0	1	2	3	4	5	6	7	8	9
$\operatorname{rank} C_W^{(0,9-a)}$	2	4	5	6	6	5	4	3	2	1
$\dim \widehat{\Sigma}_{(1,a)}(W)$	1	1	1	0	0	0	0	0	0	0

so the lower bound given by Theorem 3.12 is $r(W) \geq 6$. And we have

a	0	1	2	3	4	5	6	7	8	9
$\operatorname{rank} C_W^{(1,9-a)}$	1	2	3	4	5	6	6	5	4	2
$\dim \Sigma_{(0,a)}(W)$	1	1	1	1	0	0	-1	-1	-1	-1

so the lower bound given by Theorem 3.16 is $r(W) \geq 7$.

3.4. **Multihomogeneous polynomials.** $M \in S^{\mathbf{d}}V$ is **concise** if $M \in S^{\mathbf{d}}V' = S^{d_1}V_1' \otimes \cdots \otimes S^{d_s}V_s'$ for $V_1' \subseteq V_1, \ldots, V_s' \subseteq V_s$ implies each $V_i' = V_i$. This is equivalent to $C_M^{\mathbf{d}-\mathbf{e}}$ being surjective, or $C_M^{\mathbf{e}}$ injective, for each $\mathbf{e} = (1, 0, \ldots, 0), \ldots, (0, \ldots, 0, 1)$.

However, we will need a strengthening of conciseness, involving $\mathbf{e}$ with possibly more than one nonzero entry.

Notation 3.18. For a tuple $\mathbf{a} = (a_1, \ldots, a_s)$, we define $\operatorname{supp}(\mathbf{a}) = (e_1, \ldots, e_s)$ where $e_i = 1$ if $a_i > 0$, otherwise $e_i = 0$.

Say that M vanishes to multiorder $\mathbf{b} = (b_1, \ldots, b_s)$ at a point $P \in \prod \mathbb{P}V_i^*$ if, for every differential operator D of multidegree $\mathbf{b}$, $D(M)$ vanishes at P. Equivalently, choosing local coordinates centered at $P = (p_1, \ldots, p_s)$ by choosing local coordinates in each $\mathbb{P}V_i$ centered at p_i and writing M as a (non-homogeneous) polynomial in these coordinates, no monomial appearing in M has multidegree less than or equal to $\mathbf{b}$.

We define $\Sigma_{\mathbf{a}} = \Sigma_{\mathbf{a}}(M) \subset \prod \mathbb{P}V_i^*$ to be the subvariety defined by the image of $C_M^{\mathbf{a}}$, regarded as a set of multihomogeneous polynomials. Equivalently, $\Sigma_{\mathbf{a}}(M)$ is the locus of points at which the multiorder of vanishing of M is not less than or equal to $\mathbf{a}$.

The bound for ranks of multihomogeneous polynomials is the following:

Theorem 3.19. *Let M be a multihomogeneous polynomial as above. If $C_M^{\operatorname{supp}(\mathbf{d}-\mathbf{a})}$ is injective then $r_{MH}(M) \geq \operatorname{rank}(C_M^{\mathbf{d}-\mathbf{a}}) + \dim \Sigma_{\mathbf{a}}(M) + 1$, where $\dim \varnothing = -1$.*

Conciseness would not be enough; it only gives injectivity of $C_M^{\mathrm{supp}(\mathbf{d}-\mathbf{a})}$ when $\mathbf{d}-\mathbf{a}$ has a single nonzero entry.

Proof. Suppose

$$M = \sum_{j=1}^{r} \ell_{j,1}^{d_1} \cdots \ell_{j,s}^{d_s}.$$

For each j, let $P_j = ([\ell_{j,1}], \dots, [\ell_{j,s}]) \in \prod \mathbb{P}V_i$. Let $\mathcal{L} = \{p \in S^{\mathbf{d}-\mathbf{a}}V^* \mid p(P_1) = \cdots = p(P_r) = 0\}$. Denote by D_p the differential operator associated to p. Then

$$D_p(\ell_{j,1}^{d_1} \cdots \ell_{j,s}^{d_s}) = \frac{\mathbf{d}!}{\mathbf{a}!}\, p(\ell_{j,1}, \dots, \ell_{j,s})\, \ell_{j,1}^{a_1} \cdots \ell_{j,s}^{a_s}.$$

Thus for $p \in \mathcal{L}$, $D_p(M) = 0$. This shows $\mathcal{L} \subseteq \ker C_M^{\mathbf{d}-\mathbf{a}}$, hence

$$r \geq \operatorname{codim} \mathcal{L} \geq \operatorname{codim} \ker C_M^{\mathbf{d}-\mathbf{a}} = \operatorname{rank} C_M^{\mathbf{d}-\mathbf{a}}.$$

Now we use the additional hypothesis that $C_M^{\mathrm{supp}(\mathbf{d}-\mathbf{a})}$ is injective. With this assumption, we claim $\mathbb{P}\mathcal{L}$ is disjoint from $\mathbb{P}\ker C_M^{\mathbf{d}-\mathbf{a}} \cap \mathrm{Seg}(\prod_{i=1}^{s} \nu_{d_i-a_i}(\mathbb{P}V_i^*))$, and that this intersection is isomorphic to $\Sigma_{\mathbf{a}}(M)$.

Remark 3.20. The recent paper [AB12] shows the defectivity of certain secant varieties by considering the intersection of a Segre-Veronese variety with the image (rather than kernel) of a catalecticant map.

If the disjointness fails, then $\mathcal{L}$ contains an element $p = h_1^{d_1-a_1} \cdots h_k^{d_k-a_k}$ for some $h_i \in V_i^*$, $h_i \neq 0$, $1 \leq i \leq k$. For each $1 \leq j \leq r$,

$$p(P_j) = p(\ell_{j,1}, \dots, \ell_{j,k}) = h_1(\ell_{j,1})^{d_1-a_1} \cdots h_k(\ell_{j,k})^{d_k-a_k} = 0.$$

We must have $h_i(\ell_{j,i}) = 0$ for some i such that $d_i - a_i > 0$. Let $\mathbf{e} = (e_1, \dots, e_k) = \mathrm{supp}(\mathbf{d}-\mathbf{a})$. Then $h_i(\ell_{j,i})^{e_i} = 0$, so also $h_1(\ell_{j,1})^{e_1} \cdots h_k(\ell_{j,k})^{e_k} = 0$. This holds for all j, so $h_1^{e_1} \cdots h_k^{e_k} \in \ker C_M^{\mathbf{e}}$, contradicting that $C_M^{\mathbf{e}}$ is injective. Hence if $C_M^{\mathrm{supp}(\mathbf{d}-\mathbf{a})}$ is injective then $\mathcal{L}$ is disjoint from the intersection, as claimed.

Next, we claim an element $h_1^{d_1-a_1} \cdots h_k^{d_k-a_k}$ lies in $\ker C_M^{\mathbf{d}-\mathbf{a}}$ if and only if $([h_1], \dots, [h_k]) \in \Sigma_{\mathbf{a}}(M)$. We have $h_1^{d_1-a_1} \cdots h_k^{d_k-a_k} \in \ker C_M^{\mathbf{d}-\mathbf{a}}$ if and only if $(h_1^{d_1-a_1} \cdots h_k^{d_k-a_k})(M) = 0$, if and only if $D(h_1^{d_1-a_1} \cdots h_k^{d_k-a_k})(M) = 0$ for all type $\mathbf{a}$ differential operators $D \in S^{\mathbf{a}}(V^*)$, if and only if $(h_1^{d_1-a_1} \cdots h_k^{d_k-a_k})(D(M)) = 0$ for all such D. Since $D(M)$ is a multihomogeneous polynomial of multidegree $\mathbf{d}-\mathbf{a}$, we have that this latter is equal to the evaluation at the point $([h_1], \dots, [h_k]) \in \prod \mathbb{P}V_i^*$, up to a scalar:

$$(h_1^{d_1-a_1} \cdots h_k^{d_k-a_k})(D(M)) = (\mathbf{d}-\mathbf{a})! \cdot D(M)\big|_{(h_1,\dots,h_k)}.$$

That this vanishes for all D is exactly the condition that $([h_1], \dots, [h_k]) \in \Sigma_{\mathbf{a}}(M)$, as claimed. □

Example 3.21. Let $F = x_1 \cdots x_a y_1 \cdots y_b$, a bihomogeneous form of bidegree (a, b). One may check easily that $C_F^{(1,0)}$, $C_F^{(0,1)}$, and $C_F^{(1,1)}$ are injective. For $\mathbf{a} = (p, q)$ with $0 \leq p \leq a$, $0 \leq q \leq b$, the image of $C_F^{\mathbf{a}}$ is spanned by subproducts of $a-p$ of the x's and $b-q$ of the y's. The common vanishing locus $\Sigma_{\mathbf{a}}(F)$ is points with at least $p+1$ of the x's vanishing or at least $q+1$ of the y's vanishing, equivalently, at most $a-p-1$ of the x's nonvanishing or at most $b-q-1$ of the y's nonvanishing. This is a finite union of products $\mathbb{P}^{a-p-2} \times \mathbb{P}^{b-1}$ and $\mathbb{P}^{a-1} \times \mathbb{P}^{b-q-2}$, with dimension $\max\{a+b-p-3, a+b-q-3\}$. This shows that

$r_{MH}(x_1 \cdots x_a y_1 \cdots y_b) \geq \binom{a}{p}\binom{b}{q} + \max\{a+b-p-3, a+b-q-3\}+1$. For instance, $r_{MH}(x_1x_2x_3y_1y_2) \geq \binom{3}{1}\binom{2}{1} + 3+2-1-3+1 = 8$. Since $r_{MH}(x_1x_2x_3y_1y_2) \leq 2^{3+2-2} = 8$, this determines the multihomogeneous rank.

Example 3.22. The generic determinant $\det_n$ is bihomogeneous of bidegree $(a, n-a)$ in the variables from the first a and last $n-a$ rows. But $C^{(1,1)}_{\det_n}$ is not injective: the kernel is spanned by elements of the following two types.

(1) $\partial_{i_1,j}\partial_{i_2,j}$, a product of two differentials in the same column, with $i_1 \leq a < i_2$.
(2) $\partial_{i,j}\partial_{k,\ell} + \partial_{i,\ell}\partial_{k,j}$, the permanent of a 2×2 submatrix with $i \leq a < k$.

(It is easy to see that these elements are in the kernel, and the row conditions ensure that they have bidegree $(1,1)$. Shafiei's theorem [Sha14] describes $\det_n^\perp$ and implies that the above elements span the kernel of $C^{(1,1)}_{\det_n}$.) So Theorem 3.19 is not applicable to $r_{MH}(\det_n)$.

3.5. **Generalized ranks.** Finally, we give an improved lower bound for generalized ranks.

Theorem 3.23. *Let X be a smooth irreducible variety. Let L be a very ample line bundle on X and let $V = H^0(X,L)^*$ (so $X \hookrightarrow \mathbb{P}V$). Let $v \in V$. Let G be a vector bundle on X, $b > 0$, and $E = G^b$; more generally assume E is a vector bundle on X and there is a bundle map $G^b \to E$ whose kernel has no global sections. Let $e = \operatorname{rank} E$. Assume C^G_v is injective. Then $e\, r_X(v) > \operatorname{rank} C^E_v + \dim \Sigma$, where $\Sigma \subset \mathbb{P}H^0(X,G) \overset{\nu_b}{\hookrightarrow} \mathbb{P}H^0(X,E)$ is the subvariety of $\mathbb{P}H^0(X,G)$ defined by the image of the transpose $(C^E_v)^t = C^{L\otimes E^*}_v : H^0(X, L \otimes E^*) \to H^0(X,E)^*$ (and $\dim \Sigma = -1$ if $\Sigma = \varnothing$).*

(Here linear equations on $\mathbb{P}H^0(X,E)$ induce degree b equations on the Veronese image $\nu_b(\mathbb{P}H^0(X,G))$. Σ is the locus defined by the equations arising from $\operatorname{img}(C^E_v)^t$.)

Proof. Let $v = x_1 + \cdots + x_r$, each $[x_i] \in X$. Let $\mathcal{L} = \{h \in H^0(E) \mid h([x_1]) = \cdots = h([x_r]) = 0\}$. Clearly $C^E_v(h) = 0$ for each $h \in \mathcal{L}$, so $\mathcal{L} \subseteq \ker C^E_v$. Each $h([x_i]) = 0$ imposes e conditions on the global section h, so $\mathcal{L}$ is defined by a system of er equations. This shows $er \geq \operatorname{codim} \mathcal{L} \geq \operatorname{codim} \ker C^E_v = \operatorname{rank} C^E_v$.

If $\mathcal{L}$ contains h^b for any $h \in H^0(X,G)$ then $h \in \ker C^G_v$, contradicting the hypothesis. Thus $\mathbb{P}\mathcal{L}$ is disjoint from $\nu_b(\mathbb{P}H^0(X,G)) \cap \mathbb{P}\ker C^E_v$. Hence

$$er \geq \operatorname{codim} \mathbb{P}\mathcal{L} > \operatorname{codim} \mathbb{P}\ker C^E_v + \dim(\nu_b(\mathbb{P}H^0(X,G)) \cap \mathbb{P}\ker C^E_v).$$

Finally, for $h \in H^0(X,G)$, $h^b \in \mathbb{P}\ker C^E_v$ if and only if h^b is annihilated by each element of the image of the transpose $(C^E_v)^t$. This shows $\nu_b(\mathbb{P}H^0(X,G)) \cap \mathbb{P}\ker C^E_v \cong \Sigma$. □

4. Apolarity Lemmas

In this section we go beyond considering just rank to actually considering the terms that arise in a Waring decomposition. These are related to certain containments of ideals, corresponding to schemes called *apolar schemes*.

4.1. **Classical Waring rank.**

Definition 4.1. Let $S = \Bbbk[x_1, \ldots, x_n]$ and let $T = \Bbbk[\partial_1, \ldots, \partial_n]$ be the dual ring. Let $F \in S_d$. Then $F^\perp = \{\Theta \in T : \Theta F = 0\}$ is a homogeneous ideal, called the **apolar ideal** or **annihilating ideal** of F.

The quotient ring $A^F = T/F^\perp$ is called the **apolar algebra** of F.

Clearly $F^\perp$ contains every form of degree $d+1$ or greater. Correspondingly, A^F is an Artinian algebra. In fact A^F is an Artinian Gorenstein algebra, and every Artinian Gorenstein algebra (finitely generated, standard graded) is isomorphic to an apolar algebra A^F for some F.

Each graded piece $F^\perp_{d-a}$ is the kernel of a catalecticant, $F^\perp_{d-a} = \ker C_F^{d-a}$. The catalecticants are just the graded pieces of the quotient map $T \to A^F$. In particular the rank of C_F^{d-a} is equal to the value of the Hilbert function $h_{A^F}(a) = \dim((A^F)_a)$. The lower bound $r(F) \geq \max_{0\leq a\leq d} \operatorname{rank} C^F_{d-a}$ of Section 2.1 becomes $r(F) \geq \max_{0\leq a\leq d} h_{A^F}(a)$.

Theorem 4.2 (Apolarity Lemma)**.** *Let $F \in S_d$. Let $\ell_1,\dots,\ell_r \in S_1$ and $I = I([\ell_1],\dots,[\ell_r])$. Then there are scalars $c_1,\dots,c_r$ such that $F = c_1\ell_1^d + \cdots + c_r\ell_r^d$ if and only if $I \subset F^\perp$.*

This is the "classical" Apolarity Lemma. See for example [IK99, Theorem 5.3], [RS00, §1.3].

Proof. We have seen for each graded piece $\mathcal{L} = I_k$ and $\Theta \in \mathcal{L}$ that $\Theta\ell_i^d = 0$ for each i, so $\Theta F = 0$ and $\Theta \in \ker C_k^F$. This shows $I \subset F^\perp$.

Conversely if $I \subset F^\perp$, in particular $I_d \subset F_d^\perp$. Note $I_d = \bigcap_{i=1}^r I(\ell_i)_d = \bigcap_{i=1}^r (\ell_i^d)^\perp$. Then $\operatorname{span}\{F\} = (F^\perp)^\perp \subseteq \sum_{i=1}^r \operatorname{span}\{\ell_i^d\} = \operatorname{span}\{\ell_1^d,\dots,\ell_r^d\}$, as desired. □

Here is a version for schemes which is well-known to experts.

Theorem 4.3. *Let $F \in S_d$. Let $Z \subset \mathbb{P}V$ be an arbitrary scheme with saturated homogeneous defining ideal $I = I(Z)$. Let $\nu_d : \mathbb{P}V \to \mathbb{P}S^dV$ be the degree d Veronese map. Then $[F]$ lies in the linear span of the scheme $\nu_d(Z)$ if and only if $I \subset F^\perp$.*

Proof. $[F]$ lies in the linear span of $\nu_d(Z)$ if and only if every linear form vanishing on $\nu_d(Z)$ also vanishes on $[F]$, that is, F is annihilated by the space of linear forms on $\mathbb{P}S^dV$ that vanish on $\nu_d(Z)$; these are precisely the degree d forms on $\mathbb{P}V$ that vanish on Z, i.e., the degree d piece of I. So $[F]$ lies in the linear span of $\nu_d(Z)$ if and only if $F \in (I_d)^\perp$, equivalently $I_d \subset (F^\perp)_d$. For $d = \deg F$, $I_d \subset (F^\perp)_d$ if and only if $I \subset F^\perp$ by for example [BB14, Proposition 3.4(iii)] or [IK99, Lemma 2.15]. □

A scheme $Z \subset \mathbb{P}V$ is called **apolar to** F if its defining ideal is contained in $F^\perp$, equivalently $[F]$ is in the linear span of $\nu_d(Z)$. Thus the Waring rank of F is equal to the least length of a reduced zero-dimensional apolar scheme to F. This leads to obvious generalizations: The **smoothable rank** $sr(F)$ of F is equal to the least length of a smoothable zero-dimensional apolar scheme to F. The **cactus rank** $cr(F)$ of F is equal to the least length of any zero-dimensional apolar scheme to F. Evidently $cr(F) \leq sr(F) \leq r(F)$. For many more notions of rank, see [BBM12].

The colorful name "cactus rank" was coined in [RS11], following [BB14]. However earlier terminology in [IK99, Definition 5.1, Definition 5.66] is as follows: **annihilating scheme** for apolar scheme to F, **scheme length** for cactus rank, and **smoothable scheme length** for smoothable rank.

Remark 4.4. In the next section we will review and generalize a lower bound discovered by Ranestad and Schreyer, which is a lower bound for cactus rank. In contrast, the lower bound in Theorem 3.2 is in fact a lower bound for Waring rank $r(F)$, as opposed to cactus rank or smoothable rank. The proof we have given for Theorem 3.2, in giving a lower bound for the length of a reduced apolar scheme, used the reducedness in the step that concluded, if

h^{d-a} vanished at $[\ell_1], \ldots, [\ell_r]$, then h would also vanish at each $[\ell_i]$. This is precisely the statement that $I = I(\{[\ell_1], \ldots, [\ell_r]\})$ is radical, i.e., the scheme is reduced.

4.2. Simultaneous Waring rank.

Definition 4.5. Let $W \subseteq S_d$. Then

$$W^\perp = \{\Theta \in T : \Theta F = 0 \text{ for all } F \in W\} = \bigcap_{F \in W} F^\perp.$$

The apolar algebra is $A^W = T/W^\perp$.

Note that each graded piece $W_k^\perp$ is the kernel of the catalecticant $C_W^{(0,k)}$. The apolar algebra A^W is a level Artin algebra; these have been studied in for example [GHMS07].

Say a scheme $Z \subset \mathbb{P}V$ is **apolar to** $W \subseteq S^d V$ if $I = I(Z) \subseteq W^\perp$. Then $r(W)$ is the least length of a reduced zero-dimensional apolar scheme. As before we define the smoothable rank $sr(W)$ and cactus rank $cr(W)$: $sr(W)$ is the least length of a smoothable zero-dimensional apolar scheme and $cr(W)$ is the least length of a zero-dimensional apolar scheme.

Here is an apolarity lemma for linear series:

Theorem 4.6. *Let $W \subseteq S_d$. Let $\ell_1, \ldots, \ell_r \in V$ and $I = I([\ell_1], \ldots, [\ell_r])$. Then $W \subseteq \operatorname{span}\{\ell_1^d, \ldots, \ell_r^d\}$ if and only if $I \subset W^\perp$.*

More generally let $Z \subset \mathbb{P}V$ be a scheme with saturated homogeneous defining ideal $I = I(Z)$ and let $\nu_d : \mathbb{P}V \to \mathbb{P}S^d V$ be the degree d Veronese map. Then $\mathbb{P}W \subseteq \operatorname{span}(\nu_d(Z))$ if and only if $I \subset W^\perp$.

Proof. $W \subseteq \operatorname{span}\{\ell_1^d, \ldots, \ell_r^d\}$ if and only if for each $F \in W$, there are constants $c_1, \ldots, c_r$ such that $F = c_1\ell_1^d + \cdots + c_r\ell_r^d$, if and only if for each $F \in W$, $I \subset F^\perp$ (by the usual Apolarity Lemma!), if and only $I \subseteq \bigcap_{F \in W} F^\perp = W^\perp$.

The proof of the scheme version is the same as before. □

4.3. Multihomogeneous polynomial.

Definition 4.7. Let $S = \Bbbk[V_1 \oplus \cdots \oplus V_s]$ and $T = \Bbbk[V_1^* \oplus \cdots \oplus V_s^*]$, multigraded rings. For $M \in S_{\mathbf{d}} = S^{d_1}V_1 \otimes \cdots \otimes S^{d_s}V_s$,

$$M^\perp = \{\Theta \in T : \Theta M = 0\}.$$

The apolar algebra is $A^M = T/M^\perp$.

The apolar algebra A^M is a multigraded Artin algebra. Each multigraded piece $(M^\perp)_{\mathbf{k}}$ is the kernel of the catalecticant $C_M^{\mathbf{k}}$.

A scheme $Z \subset \mathbb{P}V = \prod \mathbb{P}V_i$ is **apolar to** M if $I = I(Z) \subseteq M^\perp$. The smoothable rank and cactus rank are defined as before.

We will state an apolarity lemma for multihomogeneous polynomials using the following notation. For a point $P = (\ell_1, \ldots, \ell_s) \in V_1 \times \cdots \times V_s$, with each $\ell_j \neq 0$, we denote $[P] = ([\ell_1], \ldots, [\ell_s]) \in \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_s$. For each $\mathbf{d} = (d_1, \ldots, d_s)$ let $\nu_{\mathbf{d}} : \prod \mathbb{P}V_i \to \prod \mathbb{P}S^{d_i}V_i$ be the Segre-Veronese map, $\nu_{\mathbf{d}}([\ell_1], \ldots, [\ell_s]) = [\ell_1^{d_1} \cdots \ell_s^{d_s}]$. For convenience let us denote $P^{\mathbf{d}} = \ell_1^{d_1} \cdots \ell_s^{d_s}$. Note that the coordinate ring of $\mathbb{P}V = \mathbb{P}V_1 \times \cdots \times \mathbb{P}V_s$ is $\Bbbk[\mathbb{P}V] = \Bbbk[V^*] = T$.

Theorem 4.8. *Let M be a multihomogeneous polynomial of multidegree $\mathbf{d} = (d_1, \ldots, d_s)$. For $i = 1, \ldots, r$, let $P_i = (\ell_{i,1}, \ldots, \ell_{i,s})$ where each $\ell_{i,j} \in V_j$, $\ell_{i,j} \neq 0$. Let $I = I([P_1], \ldots, [P_r]) \subset T$. Then $M \in \operatorname{span}(P_1^{\mathbf{d}}, \ldots, P_r^{\mathbf{d}})$ if and only if $I \subset M^\perp$.*

More generally let $Z \subset \mathbb{P}V = \prod \mathbb{P}V_i$ be a scheme with defining ideal $I = I(Z)$. Then $[M]$ is in the linear span of $\nu_{\mathbf{d}}(Z)$ if and only if $I \subset M^\perp$.

Proof. If $\Theta \in I_{\mathbf{k}}$ then $\Theta([P_1]) = \cdots = \Theta([P_r]) = 0$ so, as a differential operator, $\Theta(P_1^{\mathbf{d}}) = \cdots = \Theta(P_r^{\mathbf{d}}) = 0$, hence $\Theta(M) = 0$ and $\Theta \in M^\perp$. Conversely, if $I \subset M^\perp$ then in particular $I_{\mathbf{d}} \subseteq M_{\mathbf{d}}^\perp$. For $\Theta \in T_{\mathbf{d}}$, $\Theta \in I_{\mathbf{d}}$ if and only if as a polynomial $\Theta(P_i) = 0$ for each i, if and only if as a differential operator $\Theta(P_i^{\mathbf{d}}) = 0$ for each i, if and only if, in the pairing between the dual spaces $S_{\mathbf{d}}$ and $T_{\mathbf{d}}$, $\Theta \in (P_i^{\mathbf{d}})^\perp$ for each i. So $\bigcap_{i=1}^r (P_i^{\mathbf{d}})^\perp = I_{\mathbf{d}} \subseteq M_{\mathbf{d}}^\perp$. Transposing yields $\operatorname{span}\{M\} \subseteq (I_{\mathbf{d}})^\perp = (\bigcap (P_i^{\mathbf{d}})^\perp)^\perp = \operatorname{span}\{P_1^{\mathbf{d}}, \dots, P_s^{\mathbf{d}}\}$.

The proof of the scheme version is the same as before. $\square$

Remark 4.9. Work in progress by Maciej Gałązka shows that similar statements hold on more general toric varieties.

4.4. **Generalized rank.** I do not know a statement in the full generality of Section 1.4. Here is a statement essentially due to Carlini.

Theorem 4.10 ([Car05], [Car06a]). *Let $F \in \Bbbk[V]$ be a homogeneous form of degree d in $n = \dim V$ variables. Let $W_1, \dots, W_s \subset V$ be linear subspaces and let $I = I(W_1 \cup \cdots \cup W_s)$ be the homogeneous defining ideal of the reduced union of the W_i. Then the following are equivalent:*

1. *There exist $G_i \in \Bbbk[W_i]$, $i = 1, \dots, s$, such that $F \in \operatorname{span}\{G_1, \dots, G_s\}$.*
2. *$I \subset F^\perp$.*

In particular the least number of terms in a "codimension one" decomposition of F, i.e., a decomposition as a sum of forms in $n-1$ variables (see Example 1.9), is equal to the least number of hyperplanes whose union is apolar to F, equivalently, the least degree of a form in $F^\perp$ that factors as a product of distinct linear factors. Similarly, the least number of terms in a decomposition of F as a sum of binary forms (see Example 1.10) is equal to the least number of projective lines whose union is apolar to F. These are the results stated by Carlini, but his proofs actually give the full theorem above. For the convenience of the reader we give here the proof, following Carlini's ideas.

Proof. First suppose $F \in \operatorname{span}\{G_1, \dots, G_s\}$. For each i fix a power sum decomposition of G_i, with terms corresponding to points in W_i. Let J be the homogeneous defining ideal of all the (projective) points, for all the terms that arise for all of the G_i. Since F is a linear combination of the G_i, these terms also give a power sum decomposition of F. By the classical Apolarity Lemma, $J \subset F^\perp$. And since each of the points defined by J lies in one of the W_i, $I \subset J$. So $I \subset F^\perp$.

Conversely suppose $I \subset F^\perp$. For each i choose enough points $\ell_{i,1}, \dots, \ell_{i,k_i} \in W_i$ so that the $\ell_{i,j}^d$ span the space of d-forms on $\mathbb{P}W_i^*$, that is, $\Bbbk[W_i]_d$ (we may take $k_i = \binom{\dim W_i + d - 1}{d}$ and choose the $\ell_{i,j}$ generally in W_i). Let J be the defining ideal of all the points $\ell_{i,j}$. We claim that $J_d \subset I_d$. If $\Theta \in T_d$ is any dual d-form vanishing at each point $\ell_{i,j}$ then, as a differential operator, Θ annihilates each $\ell_{i,j}^d$. Since for each i these span all the d-forms on $\mathbb{P}W_i$, then in fact Θ annihilates all of $\Bbbk[W_i]_d$, for each i. In particular Θ annihilates ℓ^d for every $\ell \in W_i$, for each i. Hence, returning to considering Θ as a polynomial, Θ vanishes at each point $[\ell] \in \mathbb{P}W_i$, for each i. This means $\Theta \in I_d$.

Now from $J_d \subset I_d \subset F^\perp$ we claim $J \subset F^\perp$. For degrees $k > d$, $F_k^\perp = T_k$, so $J_k \subset F_k^\perp$. For degrees $k < d$, $\Theta \in F_k^\perp$ if and only if $\Theta F = 0$, if and only if for every $\Psi \in T_{d-k}$, $\Psi\Theta F = 0$, if and only if for every $\Psi \in T_{d-k}$, $\Psi\Theta \in F^\perp_d$. Meanwhile $\Theta \in J_k$ if and only if $\Theta([\ell_{i,j}]) = 0$ for all i, j, if and only if for every $\Psi \in T_{d-k}$, $\Psi\Theta$ vanishes at each $[\ell_{i,j}]$, if and only if for every $\Psi \in T_{d-k}$, $\Psi\Theta \in J_d$. So if $\Theta \in J_k$ then for every $\Psi \in T_{d-k}$ we have $\Psi\Theta \in J_d \subset F^\perp_d$, which implies that $\Theta \in F^\perp_k$. Therefore $J_k \subset F^\perp_k$, as desired. This proves the claim that $J \subset F^\perp$. Then by the classical Apolarity Lemma, F is in the span of the $\ell^d_{i,j}$, say $F = \sum_i \sum_j c_{i,j}\ell^d_{i,j}$. For each i, set $G_i = \sum_j c_{i,j}\ell^d_{i,j}$. Then $G_i \in \Bbbk[W_i]_d$ and $F \in \operatorname{span}\{G_1, \ldots, G_s\}$. □

The classical Apolarity Lemma is precisely the case $\dim W_1 = \cdots = \dim W_s = 1$.

Note that in contrast to the previous statements, a subideal $I \subset F^\perp$ does not necessarily determine the decomposition; rather it only determines the subspaces $W_1, \ldots, W_s$ over which the decomposition occurs.

It would be interesting to find out if the above theorem, or at least the numerical corollaries discussed before the proof, are related to the catalecticants arising from the sheaves mentioned in Example 2.19.

Remark 4.11. It would be very interesting to have similar statements more generally, or at least for natural situations such as split rank (decomposition as a sum of products of linear forms, see Example 1.11) and, if $d = kt$, then decomposition as a sum of kth powers of t-forms (see Example 1.12).

5. Ranestad–Schreyer bounds

Ranestad and Schreyer gave an extremely elegant lower bound for cactus and Waring rank in [RS11]. We recall their result, then generalize it.

5.1. Classical Waring rank.

Theorem 5.1 ([RS11]). *Let $F \in S^dV$. Let δ be a positive integer such that the homogeneous ideal $F^\perp$ is generated in degrees less than or equal to δ. Then the Waring rank $r(F)$, smoothable rank $sr(F)$, and cactus rank $cr(F)$ satisfy*

$$r(F) \geq sr(F) \geq cr(F) \geq \frac{\ell(A^F)}{\delta},$$

where $\ell(A^F)$ is the length of the apolar algebra A^F.

For the reader's convenience we include the proof given by Ranestad and Schreyer.

Proof. $r(F) \geq sr(F) \geq cr(F)$ is obvious. Let $Z \subset \mathbb{P}V$ be a zero-dimensional apolar scheme of length r. Let $I = I(Z)$ and let $\widehat{Z}$ be the affine variety defined by I. The definition of δ means that the common zero locus in $\mathbb{P}V$ of the linear series $F^\perp_\delta$ is exactly equal to the scheme defined by the ideal $F^\perp$; namely, the empty scheme. So in affine space V, the common zero locus of $F^\perp_\delta$ is (a scheme supported at) the origin. Thus a general δ-form $G \in F^\perp_\delta$ does not vanish at any of the points in the support of Z. Then the affine hypersurface $V(G)$ has proper intersection with $\widehat{Z}$. Since $G \in F^\perp$, $\operatorname{Spec} A^F \subset V(G)$; and since Z is apolar to F, $\operatorname{Spec} A^F \subset \widehat{Z}$. So $\operatorname{Spec} A^F \subseteq V(G) \cap \widehat{Z}$. By Bézout's theorem, then, $\ell(A^F) = \ell(\operatorname{Spec} A^F) \leq \deg(V(G)) \deg(\widehat{Z}) = \delta r$. □

See [RS11] for an application of this to monomials. (And see [CCG12], [BBT13] for more about apolarity of monomials.) See [Sha14], [Sha13] for applications of this to determinants, permanents, Pfaffians, etc. See [TW13] for an application of this to reflection arrangements. See [BBKT13] for a closer examination of this bound, particularly the quantity δ.

In fact this proof shows:

Corollary 5.2. *Let $F \in S^d V$. Let ϵ be a positive integer such that the homogeneous ideal generated by $F^{\perp}_{\leq \epsilon}$ defines a zero-dimensional affine variety, i.e., its common zero locus consists only of the origin in affine space. Then $r(F) \geq \ell(A^F)/\epsilon$.*

Of course if $F^{\perp}$ is generated in degrees less than or equal to δ then $F^{\perp}_{\leq \delta} = F^{\perp}$ whose common zero locus is just the origin because it is a homogeneous Artinian ideal.

Corollary 5.3. *Let $F \in S^d V$. Suppose $F^{\perp}$ is a complete intersection generated in degrees $0 < d_1 \leq \cdots \leq d_n$. Then $sr(F) = cr(F) = d_1 \cdots d_{n-1}$.*

Proof. $\ell(A^F) = d_1 \cdots d_n$ so $sr(F) \geq cr(F) \geq d_1 \cdots d_{n-1}$. Let $F^{\perp} = \langle G_1, \ldots, G_n \rangle$, $\deg G_i = d_i$. Then $\langle G_1, \ldots, G_{n-1} \rangle$ is a one-dimensional complete intersection, so it is smoothable (as every complete intersection is). Hence $sr(F) \leq \deg(T/\langle G_1, \ldots, G_{n-1} \rangle) = d_1 \cdots d_{n-1}$. $\square$

This occurs for monomials and reflection arrangements.

As Ranestad and Schreyer noted, this quantity $\ell(A^F)/\delta$ is in fact a lower bound for smoothable rank and cactus rank, which may be strictly smaller than Waring rank. In this sense, the lower bound Theorem 3.2 has the potential to be better: if, say, for some form F we have

$$\frac{\ell(A^F)}{\delta} \leq cr(F) < \operatorname{rank} C^a_F + \dim \widehat{\Sigma}_a(F) \leq r(F).$$

Example 5.4. Consider for example $F = xy^{d-1}$ with $d \geq 3$. This has cactus rank and smoothable rank equal to $cr(F) = sr(F) = 2$ [RS11]. Note F is concise (since it is not a pure power), $\operatorname{rank} C^{d-2}_F = 2$, and $\Sigma_2(F) = V(y) \neq \varnothing$. Then Theorem 3.2 gives the lower bound

$$r(xy^{d-1}) \geq \operatorname{rank} C^{d-2}_F + \dim \widehat{\Sigma}_2(F) = 2 + 1 = 3.$$

Of course this bound is far from sharp: we know that $r(xy^{d-1}) = d$. But this simple example shows that Theorem 3.2 is not generally a lower bound for cactus rank or smoothable rank.

Nevertheless in practice the Ranestad–Schreyer lower bound seems to be quite a bit better than Theorem 3.2 for many examples of interest.

For $F = \det_3$ and $F = \det_4$ the bound of Theorem 3.2 is better than the bound given by Ranestad–Schreyer–Shafiei (here we credit Ranestad–Schreyer for the bound described above and Shafiei for the determination that $\delta = 2$ when $F = \det_n$), but for $n \geq 5$, the Ranestad–Schreyer–Shafiei bound for $r(\det_n)$ is better than the bound of Theorem 3.2. Incidentally, I do not know the cactus or smoothable ranks of $\det_3$ and $\det_4$.

Remark 5.5. For a general d-form F, the apolar ideal is generated entirely in degree $\delta = (d+2)/2$ if d is even; if d is odd, the apolar ideal is either generated in the two degrees $\lfloor (d+2)/2 \rfloor$, $\delta = \lceil (d+2)/2 \rceil$ or entirely in the degree $\delta = \lfloor (d+2)/2 \rfloor$. (It is easy to see that $F^{\perp}$ must have a generator of degree at most $(d+2)/2$. One can show that $F^{\perp}$ has no generators in degrees strictly less than $\lfloor (d+2)/2 \rfloor$ using results on compressed algebras, see Definition 3.11 and Proposition 3.12 of [IK99]. What takes more work is to show that $F^{\perp}$ is generated in degrees less than or equal to $\lceil (d+2)/2 \rceil$. If d is even this follows from Proposition 4.1B (pg. 362) and Example 4.7 of [Iar84]. For d odd it follows from unpublished notes of Uwe Nagel [personal communication].)

One can then see that the catalecticant lower bound is better than the Ranestad–Schreyer lower bound, and Theorem 3.2 provides no improvement over the catalecticant lower bound.

However the actual rank is known by the Alexander–Hirschowitz theorem and it is strictly greater than the catalecticant lower bound.

5.2. **Simultaneous Waring rank.**

Theorem 5.6. *Let $W \subseteq S^d V$ be a linear series. Suppose $W^\perp$ is generated in degrees less than or equal to δ. Then $r(W) \geq sr(W) \geq cr(W) \geq \ell(A^W)/\delta$.*

The proof is the same.

Example 5.7. Let $X = (x_{i,j})$, $1 \leq i \leq m$, $1 \leq j \leq n$, and consider the linear series D_k, P_k, R_k. First we compute the lengths of the apolar algebras. We have

$$\ell(A^{D_k}) = \ell(A^{P_k}) = \sum_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a}.$$

When $k = m \leq n$ this simplifies to

$$\sum_{0 \leq a \leq m} \binom{m}{m-a}\binom{n}{a} = \binom{m+n}{m}.$$

And

$$\ell(A^{R_k}) = \sum_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a} a!.$$

Shafiei showed that $(\det_k)^\perp$ and $(\mathrm{per}_k)^\perp$ are each generated by quadrics [Sha14]. It is easy to show that $D_k^\perp$ and $P_k^\perp$ are generated by quadrics. Therefore

$$r(D_k), r(P_k) \geq \frac{1}{2} \sum_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a},$$

and in particular when $k = m \leq n$,

$$r(D_m), r(P_m) \geq \frac{1}{2}\binom{m+n}{m}.$$

The same lower bound holds for cactus and smoothable ranks.

It is also easy to see that $R_k^\perp$ is generated by quadrics, namely the products $\partial_{i,j_1}\partial_{i,j_2}$ (two in the same row), $\partial_{i_1,j}\partial_{i_2,j}$ (two in the same column) (including $\partial_{i,j}^2$). So

$$r(R_k) \geq \frac{1}{2} \sum_{0 \leq a \leq k} \binom{m}{a}\binom{n}{a} a!.$$

5.3. **Multihomogeneous polynomials.**

In order to give a generalization of the Ranestad–Schreyer Theorem (5.1) for multihomogeneous polynomials, we will use some elementary facts about finite schemes in multiprojective spaces. While these facts are surely well-known to experts, it was surprisingly difficult to find a comprehensive reference. I am grateful to Adam Van Tuyl for suggesting some (nearly-comprehensive) references, including his dissertation [Tuy01], the dissertation of Lavila-Vidal [LV99], and [Rob98].

Here is a brief list of the particular facts we need.

Lemma 5.8. *If $Z \subset \prod_{i=1}^s \mathbb{P}V_i$ is a zero-dimensional scheme with $I = I(Z)$, then $V(I) \subset \prod_{i=1}^s V_i$ is s-dimensional and $\deg V(I) = \ell(Z)$, where degree is computed with respect to the grading by total degree.*

Proof. The dimension statement is precisely [Tuy01, Prop. 2.2.9] or [LV99, Lemma 1.4.1].

Say $\ell(Z) = r$. This means that for every multidegree $\boldsymbol{\delta}$, $I(Z)_{\boldsymbol{\delta}}$ has codimension at most r in $S_{\boldsymbol{\delta}}$, where $S = \Bbbk[V_1 \oplus \cdots \oplus V_s]$; and there is a multidegree $\boldsymbol{\delta}_0$ such that for every $\boldsymbol{\delta} \geq \boldsymbol{\delta}_0$, $I(Z)_{\boldsymbol{\delta}}$ has codimension equal to r in $S_{\boldsymbol{\delta}}$. Now in the grading by total degree, for $t \gg 0$,

$$\begin{aligned}
\dim(S/I)_t &= \sum_{|\boldsymbol{\delta}|=t} \dim(S_{\boldsymbol{\delta}}/I(Z)_{\boldsymbol{\delta}}) \\
&= \sum_{\substack{|\boldsymbol{\delta}|=t \\ \boldsymbol{\delta}\geq\boldsymbol{\delta}_0}} r + \sum_{\substack{|\boldsymbol{\delta}|=t \\ \boldsymbol{\delta}\not\geq\boldsymbol{\delta}_0}} \dim(S_{\boldsymbol{\delta}}/I(Z)_{\boldsymbol{\delta}}) \\
&= r \cdot \#\{\boldsymbol{\delta}' \geq 0 : |\boldsymbol{\delta}'| = t - |\boldsymbol{\delta}_0|\} + \sum_{\substack{|\boldsymbol{\delta}|=t \\ \boldsymbol{\delta}\not\geq\boldsymbol{\delta}_0}} \dim(S_{\boldsymbol{\delta}}/I(Z)_{\boldsymbol{\delta}}) \\
&= r \cdot \binom{t - |\boldsymbol{\delta}_0| + s - 1}{s-1} + \sum_{\substack{|\boldsymbol{\delta}|=t \\ \boldsymbol{\delta}\not\geq\boldsymbol{\delta}_0}} \dim(S_{\boldsymbol{\delta}}/I(Z)_{\boldsymbol{\delta}}).
\end{aligned}$$

Note

$$\begin{aligned}
\#\{\boldsymbol{\delta} : |\boldsymbol{\delta}| = t, \boldsymbol{\delta} \not\geq \boldsymbol{\delta}_0\} &= \binom{t+s-1}{s-1} - \binom{t - |\boldsymbol{\delta}_0| + s - 1}{s-1} \\
&= \binom{t - |\boldsymbol{\delta}_0| + s - 1}{s-2} + \cdots + \binom{t-1+s-1}{s-2},
\end{aligned}$$

a polynomial $q(t)$ of degree $s-2$. Therefore

$$0 \leq \sum_{\substack{|\boldsymbol{\delta}|=t \\ \boldsymbol{\delta}\not\geq\boldsymbol{\delta}_0}} \dim(S_{\boldsymbol{\delta}}/I(Z)_{\boldsymbol{\delta}}) \leq r \cdot q(t),$$

so this sum is bounded by a polynomial of degree $s-2$; for sufficiently large t it is equal to a polynomial of degree at most $s-2$. Hence for $t \gg 0$, $\dim(S/I)_t = P(t)$ is a polynomial in t of degree $s-1$ with leading coefficient $r/(s-1)!$.

This $P(t)$ is precisely the Hilbert polynomial of $\mathbb{P}V(I) \subset \mathbb{P}(V_1 \oplus \cdots \oplus V_s)$. Hence $\deg V(I) = r = \ell(Z)$ (and we have re-proved the dimension statement). $\square$

Lemma 5.9. *Let $Z \subset \prod_{i=1}^s \mathbb{P}V_i$ be a zero-dimensional scheme with $I = I(Z)$. For each i let $\hat{V}_i = V_1 \times \cdots \times \{0\} \times \cdots \times V_s$, the product of all the V_j with $j \neq i$. Also for each i let $\mathrm{pr}_i : \prod \mathbb{P}V_j \to \prod_{j\neq i} \mathbb{P}V_j$ be the projection onto all the factors except V_i, and let $Z_i = \mathrm{pr}_i(Z)$. Let $I_i = I(Z_i)$.*

Then in $\prod V_i$ we have $V(I) \cap \hat{V}_i = V(I_i)$, an $(s-1)$-dimensional affine variety.

The reader may easily verify this.

Theorem 5.10. *Let $M \in S^{\mathbf{d}}V$. Let $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_s)$ be such that $M^\perp$ is generated in multidegrees less than or equal to $\boldsymbol{\delta}$. Then $r_{MH}(M) \geq sr_{MH}(M) \geq cr_{MH}(M) \geq \ell(A^M)/(\prod \delta_i)$.*

Proof. Let $Z \subset \prod \mathbb{P}V_i$ be a closed zero-dimensional scheme of length r and suppose that $I = I(Z) \subset M^\perp$. The scheme $\operatorname{Spec} A^M$ naturally lies as a closed subscheme in $\operatorname{Spec} T \cong V = \prod V_i$. Let $\widehat{Z} = V(I) \subset V$. Note that $\operatorname{Spec} A^M \subset \widehat{Z}$.

For each $i = 1, \ldots, s$ let $S_i = \Bbbk[\mathbb{P}V_i] \subset \Bbbk[\mathbb{P}V] = S$. The definition of $\boldsymbol{\delta}$ means that $M^\perp \cap S_i$ is an Artinian ideal generated in degrees less than or equal to δ_i. For each i let G_i be a general element of $M^{\perp}_{(0,\dots,\delta_i,\dots,0)}$. Since $M^{\perp}_{(0,\dots,\delta_i,\dots,0)}$ has no basepoints, G_i does not vanish at any of the points of Z. Let $B = V(G_1,\dots,G_s) \subset V$, a complete intersection of codimension s and degree $\prod \delta_i$.

Let $\hat{V}_i = V_1 \times \cdots \times \{0\} \times \cdots \times V_s$, for $1 \le i \le s$. Outside of the union $\bigcup \hat{V}_i$, B is disjoint from $\widehat{Z}$, since none of the G_i vanish at any point of Z. On each $\hat{V}_i$, G_i vanishes identically, but G_j is not identically vanishing for $j \neq i$. So $B \cap \hat{V}_i$ is a complete intersection of codimension $s-1$ in $\hat{V}_i$. By the generality of the G's, $B \cap \hat{V}_i$ does not meet $\widehat{Z} \cap \hat{V}_i$ outside the origin.

Therefore $B \cap \widehat{Z}$ is supported only at the origin. So

$$\ell(A^M) = \ell(\operatorname{Spec} A^M) \le \ell(B \cap \widehat{Z}) = \deg(B)\deg(\widehat{Z}) = \Big(\prod \delta_i\Big) r,$$

which completes the proof. $\square$

Example 5.11. The bihomogeneous form $F = x_1 \cdots x_a y_1 \cdots y_b$ has apolar algebra of length 2^{a+b}. Let $\partial_1, \cdots, \partial_a$ be the dual variables to the x_i and let $\epsilon_1, \dots, \epsilon_b$ be the dual variables to the y_j. Then $F^{\perp} = \langle \partial_1^2, \dots, \partial_a^2, \epsilon_1^2, \dots, \epsilon_b^2 \rangle$. So $F^{\perp}$ is generated in bidegrees $(2,0)$ and $(0,2)$. Set $\boldsymbol{\delta} = (2,2)$. Then

$$r_{MH}(x_1 \cdots x_a y_1 \cdots y_b) \ge \frac{2^{a+b}}{4} = 2^{a+b-2}.$$

Therefore the rank is in fact equal to 2^{a+b-2}. This answers the question raised in Example 1.6.

Example 5.12. More generally let

$$F = (\prod_{j=1}^{n_1} x_{1,j}^{d_1}) \cdots (\prod_{j=1}^{n_s} x_{s,j}^{d_s}),$$

a multihomogeneous form of multidegree $(n_1 d_1, \dots, n_s d_s)$. Then

$$r_{MH}(F) \le \prod_{i=1}^{s} r((x_{i,1} \cdots x_{i,n_i})^{d_i}) = \prod_{i=1}^{s} (d_i + 1)^{n_i - 1}.$$

On the other hand $F^{\perp} = \langle \partial_{1,1}^{d_1+1}, \dots, \partial_{s,n_s}^{d_s+1} \rangle$. Then $\ell(A^F) = \prod_{i=1}^{s} (d_i+1)^{n_i}$ and, with $\boldsymbol{\delta} = (d_1+1, \dots, d_s+1)$, we get $r_{MH}(F) \ge \prod_{i=1}^{s} (d_i+1)^{n_i-1}$, which exactly determines $r_{MH}(F)$.

5.4. **Generalized rank.**

Theorem 5.13. *Let $F \in S^d V$. Fix $k < n = \dim V$. Let $r_k(F)$ be the least number of terms in a decomposition of F as a sum of forms each depending on k or fewer variables. Let $F^{\perp} = (H_1, \dots, H_t)$, $\deg H_i = d_i$, $d_1 \le \cdots \le d_t$, and suppose j is such that $(H_1, \dots, H_j)$ defines a zero-dimensional affine variety. Necessarily $j \ge n$. Then $r_k(F) \ge \ell(A^F)/(d_{j-k+1} \cdots d_j)$.*

Proof. Fix a decomposition $F = F_1 + \cdots + F_r$, each F_i depending on the variables in W_i, a k-dimensional subspace. Let C be the reduced union of the $\mathbb{P}W_i$. Every component of C is $(k-1)$-dimensional.

Let $I = (H_1, \dots, H_j)$. For each $i = 1, \dots, j$, let $G_i \in I_{d_i}$ be general. Then $I = (G_1, \dots, G_j)$. For each $i = 1, \dots, j$, let $C_i = C \cap V(G_i, \dots, G_j)$ and let $I(i) = (G_1, \dots, G_{i-1})$.

Since the linear series I_{d_j} has no basepoints, by generality of G_j and Bertini's theorem, $V(G_j)$ does not contain any component of C. So every component of C_j is $(k-2)$-dimensional. Suppose inductively that every component of C_i is $(k-j+i-2)$-dimensional. Note that

$V(G_1, \dots, G_{i-1})$ is disjoint from $V(G_i, \dots, G_j)$, so the linear series $I(i)_{d_{i-1}}$ (the degree d_{i-1} elements of the ideal $(G_1, \dots, G_{i-1})$) has no basepoints on C_i. By generality of G_{i-1} and Bertini, then, $V(G_{i-1})$ does not include any component of C_i, so every component of C_{i-1} has dimension exactly one less than the dimension of C_i.

Therefore C_{j-k+1} is empty. Let $J = (G_{j-k+1}, \dots, G_j)$ and let $\widehat{Z} = V(J)$ be the affine variety defined by J. Let $\widehat{C} = \bigcup W_i$ be the reduced union of the W_i. Note that every component of $\widehat{C}$ has dimension k, while since J has k generators, every component of $\widehat{Z}$ has codimension at most k. We have just seen that $\widehat{Z}$ intersects $\widehat{C}$ only at the origin; thus $\widehat{Z}$ is a complete intersection, in particular $\widehat{C}$ and $\widehat{Z}$ intersect properly. Since $J \subset F^\perp$, $\operatorname{Spec} A^F \subset \widehat{Z}$. And by the Apolarity Lemma, $I(C) \subset F^\perp$, so $\operatorname{Spec} A^F \subset \widehat{C}$. Therefore

$$\ell(A^F) = \ell(\operatorname{Spec} A^F) \leq \deg(\widehat{C} \cap \widehat{Z}) = \deg(\widehat{C}) \deg(\widehat{Z}) = r(d_{j-k+1} \cdots d_j),$$

as claimed. □

Conjecture 5.14. Let $F = x_1^{d_1} \cdots x_n^{d_n}$ with $0 < d_1 \leq \cdots \leq d_n$. Theorem 5.13 gives

$$r_k(x_1^{d_1} \cdots x_n^{d_n}) \geq \frac{(d_1+1)\cdots(d_n+1)}{(d_{n-k+1}+1)\cdots(d_n+1)} = (d_1+1)\cdots(d_{n-k}+1).$$

Conversely, let

$$x_1^{d_1} \cdots x_{n-k+1}^{d_{n-k+1}} = \sum_{i=1}^m \ell_i^{d_1+\cdots+d_{n-k+1}}$$

be a Waring decomposition, $m = r(x_1^{d_1} \cdots x_{n-k+1}^{d_{n-k+1}}) = (d_2+1)\cdots(d_{n-k+1}+1)$. Then

$$x_1^{d_1} \cdots x_n^{d_n} = \sum_{i=1}^m \ell_i^{d_1+\cdots+d_{n-k+1}} x_{n-k+2}^{d_{n-k+2}} \cdots x_n^{d_n},$$

where each term depends essentially on k variables. Therefore

$$r_k(x_1^{d_1} \cdots x_n^{d_n}) \leq r(x_1^{d_1} \cdots x_{n-k+1}^{d_{n-k+1}}) = (d_2+1)\cdots(d_{n-k+1}+1).$$

We conjecture that this is an equality.

It is trivially true when $k = 1$ (ordinary Waring rank, reducing to the theorem of Carlini–Catalisano–Geramita [CCG12]) or $k = n$. It is true whenever $d_1 = \cdots = d_{n-k+1} \leq d_{n-k+2} \leq \cdots \leq d_n$ because the upper and lower bounds are equal:

$$r_k\big((x_1 \cdots x_{n-k+1})^d x_{n-k+2}^{d_{n-k+2}} \cdots x_n^{d_n}\big) = r\big((x_1 \cdots x_{n-k+1})^d\big) = (d+1)^{n-k}.$$

Acknowledgments

I am grateful to a number of people for helpful conversations. Jarek Buczyński suggested considering the linear series D_k in Example 1.4, alerted me to the work in progress of Maciej Gałązka (Remark 4.9), and found a number of errors. Jeremy Martin and Chris Francisco helped with Example 3.14. Adam Van Tuyl helped tremendously with Section 5.3. Harm Derksen told me about his formula (7) and shared some very helpful comments about where it came from. His suggestions also led me toward Theorem 5.13. Uwe Nagel shared his unpublished results on generation of apolar ideals of general forms, see Remark 5.5.

References

[AB11] Enrique Arrondo and Alessandra Bernardi, *On the variety parameterizing completely decomposable polynomials*, J. Pure Appl. Algebra **215** (2011), no. 3, 201–220. MR 2729216 (2011h:14066)

[AB12] Hirotachi Abo and Maria Chiara Brambilla, *New examples of defective secant varieties of Segre–Veronese varieties*, Collectanea Mathematica **63** (2012), no. 3, 287–297, DOI: 10.1007/s13348-011-0039-7.

[AH95] J. Alexander and A. Hirschowitz, *Polynomial interpolation in several variables*, J. Algebraic Geom. **4** (1995), no. 2, 201–222. MR 1311347 (96f:14065)

[AS13] Jeaman Ahn and Yong Su Shin, *On Gorenstein sequences of socle degrees 4 and 5*, J. Pure Appl. Algebra **217** (2013), no. 5, 854–862. MR 3003310

[BB13] Weronika Buczyńska and Jarosław Buczyński, *On differences between the border rank and the smoothable rank of a polynomial*, arXiv:1305.1726 [math.AG], May 2013.

[BB14] Weronika Buczyńska and Jarosław Buczyński, *Secant varieties to high degree Veronese reembeddings, catalecticant matrices and smoothable Gorenstein schemes*, J. Algebraic Geom. **23** (2014), no. 1, 63–90. MR 3121848

[BBCM11] Alessandra Bernardi, Jérôme Brachat, Pierre Comon, and Bernard Mourrain, *Multihomogeneous polynomial decomposition using moment matrices*, ISSAC 2011—Proceedings of the 36th International Symposium on Symbolic and Algebraic Computation, ACM, New York, 2011, pp. 35–42. MR 2895192

[BBKT13] Weronika Buczyńska, Jarosław Buczyński, Johannes Kleppe, and Zach Teitler, *Apolarity and direct sum decomposability of polynomials*, arXiv:1307.3314 [math.AG], Jul 2013.

[BBM12] Alessandra Bernardi, Jérôme Brachat, and Bernard Mourrain, *A comparison of different notions of ranks of symmetric tensors*, arXiv:1210.8169 [math.AG], Oct 2012.

[BBS08] A. Białynicki-Birula and A. Schinzel, *Representations of multivariate polynomials by sums of univariate polynomials in linear forms*, Colloq. Math. **112** (2008), no. 2, 201–233. MR 2383331 (2009b:12006)

[BBT13] Weronika Buczyńska, Jarosław Buczyński, and Zach Teitler, *Waring decompositions of monomials*, J. Algebra **378** (2013), 45–57. MR 3017012

[BCMT10] Jerome Brachat, Pierre Comon, Bernard Mourrain, and Elias Tsigaridas, *Symmetric tensor decomposition*, Linear Algebra Appl. **433** (2010), no. 11-12, 1851–1872. MR 2736103 (2011k:15047)

[BGI11] Alessandra Bernardi, Alessandro Gimigliano, and Monica Idà, *Computing symmetric rank for symmetric tensors*, Journal of Symbolic Computation **46** (2011), no. 1, 34–53.

[BI92] David Bernstein and Anthony Iarrobino, *A nonunimodal graded Gorenstein Artin algebra in codimension five*, Comm. Algebra **20** (1992), no. 8, 2323–2336. MR 1172667 (93i:13012)

[BL94] Mats Boij and Dan Laksov, *Nonunimodality of graded Gorenstein Artin algebras*, Proc. Amer. Math. Soc. **120** (1994), no. 4, 1083–1092. MR 1227512 (94g:13008)

[BL13] Jarosław Buczyński and J. M. Landsberg, *Ranks of tensors and a generalization of secant varieties*, Linear Algebra Appl. **438** (2013), no. 2, 668–689. MR 2996361

[BO08] Maria Chiara Brambilla and Giorgio Ottaviani, *On the Alexander-Hirschowitz theorem*, J. Pure Appl. Algebra **212** (2008), no. 5, 1229–1251. MR MR2387598 (2008m:14104)

[Boi95] Mats Boij, *Graded Gorenstein Artin algebras whose Hilbert functions have a large number of valleys*, Comm. Algebra **23** (1995), no. 1, 97–103. MR 1311776 (96h:13040)

[BP13] Edoardo Ballico and Alessandro De Paris, *Generic power sum decompositions and bounds for the Waring rank*, arXiv:1312.3494 [math.AG], Dec 2013.

[Bro33] J. Bronowski, *The sums of powers as simultaneous canonical expressions*, Math. Proc. Cambridge Philos. Soc. **29** (1933), no. 2, 245–256.

[BT14] Greg Blekherman and Zach Teitler, *On maximum, typical, and generic ranks*, arXiv:1402.2371 [math.AG], Feb 2014.

[Car05] Enrico Carlini, *Codimension one decompositions and Chow varieties*, Projective varieties with unexpected properties, Walter de Gruyter GmbH & Co. KG, Berlin, 2005, pp. 67–79. MR 2202247 (2007f:14056)

[Car06a] ———, *Binary decompositions and varieties of sums of binaries*, J. Pure Appl. Algebra **204** (2006), no. 2, 380–388. MR 2184818 (2006j:14070)

[Car06b] ———, *Reducing the number of variables of a polynomial*, Algebraic geometry and geometric modeling, Math. Vis., Springer, Berlin, 2006, pp. 237–247. MR 2279854 (2007i:12010)

[CC03] Enrico Carlini and Jaydeep Chipalkatti, *On Waring's problem for several algebraic forms*, Comment. Math. Helv. **78** (2003), no. 3, 494–517. MR 1998391 (2005b:14097)

[CCC14] Enrico Carlini, Maria Virginia Catalisano, and Luca Chiantini, *Progress on the symmetric Strassen conjecture*, arXiv:1405.3721 [math.AG], May 2014.

[CCG12] Enrico Carlini, Maria Virginia Catalisano, and Anthony V. Geramita, *The solution to the Waring problem for monomials and the sum of coprime monomials*, J. Algebra **370** (2012), 5–14.

[CGLM08] Pierre Comon, Gene Golub, Lek-Heng Lim, and Bernard Mourrain, *Symmetric tensors and symmetric tensor rank*, SIAM J. Matrix Anal. Appl. **30** (2008), no. 3, 1254–1279. MR 2447451 (2009i:15039)

[Cha01] Karen A. Chandler, *A brief proof of a maximal rank theorem for generic double points in projective space*, Trans. Amer. Math. Soc. **353** (2001), no. 5, 1907–1920 (electronic). MR 1813598 (2002i:14046)

[Cha02] ———, *Linear systems of cubics singular at general points of projective space*, Compositio Math. **134** (2002), no. 3, 269–282. MR 1943904 (2003i:13023)

[CM96] P. COMON and B. MOURRAIN, *Decomposition of quantics in sums of powers of linear forms*, Signal Processing, Elsevier 53(2), 1996.

[CO13] Enrico Carlini and Alessandro Oneto, *Monomials as sums of k-th powers of forms*, arXiv:1305.4553 [math.AC], May 2013.

[CS11] Gonzalo Comas and Malena Seiguer, *On the rank of a binary form*, Found. Comput. Math. **11** (2011), no. 1, 65–78. MR 2754189

[Der13] Harm Derksen, *On the nuclear norm and the singular value decomposition of tensors*, arXiv:1308.3860 [math.OC], Aug 2013.

[Die10] Reinhard Diestel, *Graph theory*, fourth ed., Graduate Texts in Mathematics, vol. 173, Springer, Heidelberg, 2010. MR 2744811 (2011m:05002)

[DSS09] Mathias Drton, Bernd Sturmfels, and Seth Sullivant, *Lectures on algebraic statistics*, Oberwolfach Seminars, vol. 39, Birkhäuser Verlag, Basel, 2009. MR 2723140 (2012d:62004)

[ER93] Richard Ehrenborg and Gian-Carlo Rota, *Apolarity and canonical forms for homogeneous polynomials*, European J. Combin. **14** (1993), no. 3, 157–181. MR 1215329 (94e:15062)

[Fon02] Claudio Fontanari, *On Waring's problem for many forms and Grassmann defective varieties*, J. Pure Appl. Algebra **174** (2002), no. 3, 243–247. MR 1929406 (2003h:14079)

[Fon06] ———, *On Waring's problem for partially symmetric tensors. Variations on a theme of Mella*, Ann. Univ. Ferrara Sez. VII Sci. Mat. **52** (2006), no. 1, 37–43. MR 2246903 (2007g:14064)

[FOS12] Ralf Fröberg, Giorgio Ottaviani, and Boris Shapiro, *On the Waring problem for polynomial rings*, Proc. Natl. Acad. Sci. USA **109** (2012), no. 15, 5600–5602. MR 2935563

[Ger96] Anthony V. Geramita, *Inverse systems of fat points: Waring's problem, secant varieties of Veronese varieties and parameter spaces for Gorenstein ideals*, The Curves Seminar at Queen's, Vol. X (Kingston, ON, 1995), Queen's Papers in Pure and Appl. Math., vol. 102, Queen's Univ., Kingston, ON, 1996, pp. 2–114.

[GHMS07] Anthony V. Geramita, Tadahito Harima, Juan C. Migliore, and Yong Su Shin, *The Hilbert function of a level algebra*, Mem. Amer. Math. Soc. **186** (2007), no. 872, vi+139. MR 2292384 (2007k:13033)

[Gly10] David G. Glynn, *The permanent of a square matrix*, European J. Combin. **31** (2010), no. 7, 1887–1891. MR 2673027 (2011h:15010)

[Gun86] S. Gundelfinger, *Zur theorie der binären formen*, J. Reine Angew. Math. **100** (1886), 413–424.

[GY10] John Hilton Grace and Alfred Young, *The algebra of invariants*, Cambridge Library Collection, Cambridge University Press, Cambridge, 2010, Reprint of the 1903 original. MR 2850282 (2012g:01022)

[Iar84] Anthony Iarrobino, *Compressed algebras: Artin algebras having given socle degrees and maximal length*, Trans. Amer. Math. Soc. **285** (1984), no. 1, 337–378. MR 748843 (85j:13030)

[IK99] Anthony Iarrobino and Vassil Kanev, *Power sums, Gorenstein algebras, and determinantal loci*, Lecture Notes in Mathematics, vol. 1721, Springer-Verlag, Berlin, 1999, Appendix C by Iarrobino and Steven L. Kleiman. MR 1735271 (2001d:14056)

[Jel13] Joachim Jelisiejew, *An upper bound for the Waring rank of a form*, arXiv:1305.6957 [math.AC], May 2013.

[KB09] Tamara G. Kolda and Brett W. Bader, *Tensor decompositions and applications*, SIAM Rev. **51** (2009), no. 3, 455–500. MR 2535056 (2010j:15027)

[Kir08] George A. Kirkup, *Minimal primes over permanental ideals*, Trans. Amer. Math. Soc. **360** (2008), no. 7, 3751–3770. MR 2386244 (2009a:13017)

[Kle99] Johannes Kleppe, *Representing a homogenous polynomial as a sum of powers of linear forms*, Master's thesis, University of Oslo, 1999, http://folk.uio.no/johannkl/kleppe-master.pdf.

[Kun86] Joseph P. S. Kung, *Gundelfinger's theorem on binary forms*, Stud. Appl. Math. **75** (1986), no. 2, 163–169. MR 859177 (87m:11020)

[Lan12] J. M. Landsberg, *Tensors: geometry and applications*, Graduate Studies in Mathematics, vol. 128, American Mathematical Society, Providence, RI, 2012. MR 2865915

[Lan13] J.M. Landsberg, *Geometric complexity theory: an introduction for geometers*, arXiv:1305.7387 [math.AG], May 2013.

[LO12] J. Landsberg and Giorgio Ottaviani, *Equations for secant varieties of Veronese and other varieties*, Annali di Matematica Pura ed Applicata (2012), 1–38.

[LS00] Reinhard C. Laubenbacher and Irena Swanson, *Permanental ideals*, J. Symbolic Comput. **30** (2000), no. 2, 195–205. MR 1777172 (2001i:13039)

[LT10] J.M. Landsberg and Zach Teitler, *On the ranks and border ranks of symmetric tensors*, Found. Comp. Math. **10** (2010), no. 3, 339–366.

[LV99] Olga Lavila-Vidal, *On the diagonals of a Rees algebra*, Ph.D. thesis, Universitat de Barcelona, 1999.

[Mil13] Jeff Miller, *Earliest known uses of some of the words of mathematics (c)*, Dec 2013, http://jeff560.tripod.com/c.html.

[MS05] Ezra Miller and Bernd Sturmfels, *Combinatorial commutative algebra*, Graduate Texts in Mathematics, vol. 227, Springer-Verlag, New York, 2005. MR 2110098 (2006d:13001)

[OO13] Luke Oeding and Giorgio Ottaviani, *Eigenvectors of tensors and algorithms for Waring decomposition*, Journal of Symbolic Computation **54** (2013), 9–35.

[Par13] Alessandro De Paris, *A proof that the maximal rank for plane quartics is seven*, http://arxiv.org/abs/1309.6475, Sep 2013.

[Pos12] Elisa Postinghel, *A new proof of the Alexander-Hirschowitz interpolation theorem*, Ann. Mat. Pura Appl. (4) **191** (2012), no. 1, 77–94. MR 2886162

[Rez92] Bruce Reznick, *Sums of even powers of real linear forms*, Mem. Amer. Math. Soc. **96** (1992), no. 463, viii+155. MR 1096187 (93h:11043)

[Rez13a] Bruce Reznick, *Forms as sums of powers of lower degree forms*, Slides from a talk at the SIAM Conference on Applied Algebraic Geometry, Fort Collins, Colorado, 2013. http://www.math.uiuc.edu/~reznick/8213f-pm.pdf, Aug 2013.

[Rez13b] ———, *On the length of binary forms*, Quadratic and Higher Degree Forms (New York) (K. Alladi, M. Bhargava, D. Savitt, and P. Tiep, eds.), Developments in Math., vol. 31, Springer, 2013, pp. 207–232.

[Rob98] Paul C. Roberts, *Multiplicities and Chern classes in local algebra*, Cambridge Tracts in Mathematics, vol. 133, Cambridge University Press, Cambridge, 1998. MR 1686450 (2001a:13029)

[RS00] Kristian Ranestad and Frank-Olaf Schreyer, *Varieties of sums of powers*, J. Reine Angew. Math. **525** (2000), 147–181. MR 1780430 (2001m:14009)

[RS11] ———, *On the rank of a symmetric form*, J. Algebra **346** (2011), 340–342. MR 2842085

[Rys63] Herbert John Ryser, *Combinatorial mathematics*, The Carus Mathematical Monographs, No. 14, Published by The Mathematical Association of America, 1963. MR 0150048 (27 #51)

[Sha13] Masoumeh Sepideh Shafiei, *Apolarity for determinants and permanents of generic symmetric matrices*, arXiv:1303.1860 [math.AC], Mar 2013.

[Sha14] ———, *Apolarity for determinants and permanents of generic matrices*, To appear in Journal of Commutative Algebra, 2014.

[SS12] Sumi Seo and Hema Srinivasan, *On unimodality of Hilbert functions of Gorenstein Artin algebras of embedding dimension four*, Comm. Algebra **40** (2012), no. 8, 2893–2905. MR 2968919

[Sta78] Richard P. Stanley, *Hilbert functions of graded algebras*, Advances in Math. **28** (1978), no. 1, 57–83.

[Stu02] Bernd Sturmfels, *Solving systems of polynomial equations*, CBMS Regional Conference Series in Mathematics, vol. 97, Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 2002. MR 1925796 (2003i:13037)

[Syl51a] J.J. Sylvester, *An essay on canonical forms, supplement to a sketch of a memoir on elimination, transformation and canonical forms*, originally published by George Bell, Fleet Street, London, 1851. Paper 34 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973, originally published by Cambridge University Press in 1904., 1851.

[Syl51b] ______, *On a remarkable discovery in the theory of canonical forms and of hyperdeterminants*, originally published in Philosophical Magazine, vol. 2, 1851, pp. 391–410. Paper 41 in *Mathematical Papers*, Vol. 1, Chelsea, New York, 1973, originally published by Cambridge University Press in 1904., 1851.

[Syl86] ______, *Sur une extension d'un théorème de Clebsch relatif aux courbes du quatrième degré*, C.R. Acad. Sci. **102** (1886), 1532–1534.

[Tuy01] Adam Van Tuyl, *Sets of Points in Multi-Projective Spaces and their Hilbert Function*, Ph.D. thesis, Queen's University, 2001.

[TW13] Zach Teitler and Alex Woo, *Apolarity and reflection groups*, arXiv:1304.7202 [math.AG], Apr 2013.

[Ven13] Emanuele Ventura, *A note on the Waring ranks of reducible cubic forms*, arXiv:1305.5394 [math.AG], May 2013.

[Woo14] Youngho Woo, *Some cases on strassen additive conjecture*, arXiv:1406.2213 [math.AG], Jun 2014.

[Yer32] J. Yerushalmy, *On the Configuration of the Nine Base Points of a Pencil of Equianharmonic Cubics*, Amer. J. Math. **54** (1932), no. 2, 279–284. MR MR1506892

[Yu99] Yang Yu, *The permanent rank of a matrix*, J. Combin. Theory Ser. A **85** (1999), no. 2, 237–242. MR 1673948 (99j:15013)

E-mail address: zteitler@boisestate.edu

Department of Mathematics, 1910 University Drive, Boise State University, Boise, ID 83725-1555, USA