Leibniz Universität Hannover

INSTITUTE OF ALGEBRAIC GEOMETRY

The Waring problem for polynomials

Geometry and applications

Master's thesis (online version)

Leo Kayser

Matrikelnr. 10006908

First Examiner: *Prof. Dr. Matthias Schütt*
Second Examiner: *Dr. Emre Sertöz*

January 19, 2022

Contents

Introduction

The classical *Waring problem* in number theory asks the following question:

> *Given $d \in \mathbb{N}$, what is the minimal number $r \in \mathbb{N}$ such that any natural number can be written as a sum of at most r d-th powers of natural numbers?*

For example the question for $d = 2$ is answered by Lagrange's *four-square theorem*. In this thesis we investigate a related question in a more algebro-geometric framework:

> *Given a homogeneous form $F \in \mathbb{C}[x_0, \ldots, x_n]_d$, what is the minimal number $r \in \mathbb{N}$ such that F can be written as a linear combination of r d-th powers of linear forms?*

This number is called the *Waring rank* $\mathrm{WR}(F)$, such a power sum decomposition is called a *Waring decomposition*. Consider for example $F = xy \in \mathbb{C}[x, y]_2$; by expanding $(ax + by)^2$ one easily sees that $\mathrm{WR}(xy) \geq 2$. On the other hand

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

so $\mathrm{WR}(xy) = 2$.

A first concern may be, whether or not such a decomposition always exists. In chapter 1 we will develop the necessary language and machinery to restate this as the geometric fact that the image of the *Veronese map* $\nu_d \colon \mathbb{P}^n \longrightarrow \mathbb{P}^{\binom{n+d}{d}-1}$ is not contained in a hyperplane, and relate the Waring rank to higher secant varieties. But there is also a more straightforward, albeit "appearing-out-of-nowhere" proof:

As any form can be expressed as a linear combination of monomials, it suffices to find decompositions of each monomial $x_0^{d_0} \cdots x_n^{d_n}$. By substitution from a polynomial ring with more variables it suffices to consider the monomial $x_0 \cdots x_n$. For this we can give the formula

$$x_0 \cdots x_n = \sum_{\xi \in \{\pm 1\}^n} \frac{\xi_1 \cdots \xi_n}{2^n n!} \cdot (x_0 + \xi_1 x_1 + \cdots + \xi_n x_n)^n,$$

which is a linear combination of 2^n n-th powers.

Is this expression minimal (i. e. a Waring decomposition)? Can the set of linear forms occurring in this decomposition be anticipated? The answer to these questions is *yes*! In chapter 2 we will meet the *Apolarity Lemma* which relates power sum decomposition of a form F to ideals of schemes of points contained in a certain ideal associated to F. Using this and the additional machinery of *Hilbert functions* we calculate the Waring rank of any monomial.

In the original number-theoretic Waring problem the following phenomenon arises: Every number can be written as a sum of 9 cubes, the bound being sharp (e. g. for $23 = 2 \cdot 2^3 + 7 \cdot 1^3$). But actually, for all natural numbers apart from $\{23, 239\}$, 8 or fewer cubes suffice! So a natural related question is:

> *Given $d \in \mathbb{N}$, what is the minimal number $r \in \mathbb{N}$ such that any* sufficiently large *natural number can be written as a sum of at most r d-th powers of natural numbers?*

In algebraic geometry, an analogous question is to ask for the rank of a form belonging to a *general* form, called the *generic Waring rank*. Formally:

> *Given $n, d \geq 1$, what is the minimal number $r \in \mathbb{N}$ such that all forms in a Zariski-open dense subset $U \subseteq \mathbb{C}[x_0, \ldots, x_n]_d$ have Waring rank $\leq r$?*

While the number-theoretic question is still open[1], the generic Waring rank is completely known! A heuristic argument goes as follows: The set of d-th powers in $\mathbb{C}[x_0, \ldots, x_n]_d$ has dimension $n + 1$. If there were no relations lowering the dimension, one might expect the set of sums of s d-th powers to have dimension $s(n + 1)$. Hence the smallest s such that this set has the dimension of the ambient space $\binom{n+d}{d}$ should be

$$s = \left\lceil \frac{1}{n+1} \binom{n+d}{d} \right\rceil.$$

It turns out that this is indeed the correct answer (apart from a handful of exceptional cases), and a consequence of the celebrated *Alexander-Hirschowitz* Theorem. In chapter 3 we give an overview on how to prove this theorem.

The Waring rank has several interesting connections with other areas of mathematics and science, in chapter 4 we focus on some applications in computer science. We present two algorithms for computing Waring decompositions in the cases $d = 2$ (quadratic forms) and $n = 1$ (binary forms, Sylvester's algorithm).

It is then a natural next question to ask about the inherent complexity of computing the Waring rank. We recall basic notions from complexity theory and discuss a recent result by Shitov showing polynomial time equivalence of the Waring rank problem to that of solving polynomial equations, which is known to be at least NP-hard.

[1] For an overview of this topic I recommend the survey by Vaughan & Wooley [VW02].

We conclude by showcasing a unexpected connection to the theory of parameterized algorithms, recently found by Pratt. He shows how, among other results, short powers sum decompositions of elementary symmetric polynomials yield good algorithms for counting simple closed walks in directed graphs.

The aim of this thesis is to introduce the reader to the interesting and broad topic of Waring rank and related notions such as tensor rank, higher secant varieties, Apolarity and polynomial interpolation. We assume familiarity with projective algebraic varieties and schemes taught in a typical two-semester algebraic geometry course. More specific topics are introduced along the way and references for further reading are given.

Acknowledgment

I would like to thank my supervisor Matthias Schütt for his advice and encouragement. Also thanks to my fellow students Erik, Jan, Johannes & Sophie, whose comments significantly lowered the number of typos in this thesis. Last but not the least I would also like to thank my family and friends for encouraging and supporting me whenever I needed them.

Notation

Throughout this thesis $\Bbbk$ will be an algebraically closed field of characteristic 0 (although many statements and proofs also apply to $\operatorname{char} \Bbbk > d$, the degree of the forms considered).

All schemes considered will be Noetherian and, with the exception of section 1.2, of finite type over $\Bbbk$; mostly embedded in some ambient projective space.

Homogeneous forms whose Waring rank we consider will be typeset as capital letters with lowercase indeterminates such as $F(x_0, \ldots, x_n)$. On the other hand, polynomial functions on some projective space will be denoted as $f(X_0, \ldots, X_n)$. This distinction is useful in chapter 2, where the f are sort of dual to F, on the other hand chapter 4 has to break this convention in some places.

The ring of polynomial functions will be denoted as $S = \Bbbk[X_0, \ldots, X_n]$, or $\Bbbk[\underline{X}]$ if the number of variables is understood. Graded components of graded structures will be denoted as S_d.

The defining homogeneous ideal of a projective scheme $X \subseteq \mathbb{P}^n$ is $I(X)$; the sheaf of ideals is $\mathcal{I}_X$. Conversely, if $I \subseteq S$ is a homogeneous ideal, then the projective scheme it defines will be denoted as $\mathcal{V}(I)$. The defining homogeneous ideal of a closed point $P \in \mathbb{P}^n$ will be denoted by $\mathfrak{m}_P$ (even though this is not a maximal ideal).

The affine cone of a projective variety $X \subseteq \mathbb{P}(\Bbbk^n)$ is $\widehat{X} \subseteq \Bbbk^n$, conversely the set in projective space corresponding to $Y \subseteq \Bbbk^n$ denoted by $\mathbb{P}(Y)$.

Zero-dimensional schemes will be typeset in blackboard bold such as $\mathbb{X}$, their length is $\operatorname{len}(\mathbb{X}) = \dim_{\Bbbk} \mathcal{O}_{\mathbb{X}}(\mathbb{X})$.

We use $\mathbb{N}_0$ and $\mathbb{N}_+$ for the non-negative resp. positive integers. Multi-indices are typeset in bold Greek letters such as $\boldsymbol{\alpha} = (\alpha_0, \ldots, \alpha_n) \in \mathbb{N}_0^n$. We use the following shorthand notation:

$$|\boldsymbol{\alpha}| = \sum_{i=0}^{n} \alpha_i, \qquad \boldsymbol{\alpha}! = \prod_{i=0}^{n} \alpha_i!, \qquad \binom{|\boldsymbol{\alpha}|}{\boldsymbol{\alpha}} = \frac{|\boldsymbol{\alpha}|!}{\alpha_0! \cdots \alpha_n!}, \qquad X^{\boldsymbol{\alpha}} = X_0^{\alpha_0} \cdots X_n^{\alpha_n},$$
$$\boldsymbol{\alpha} \leq \boldsymbol{\beta} \text{ iff } \alpha_i \leq \beta_i \text{ for } i = 0, \ldots, n.$$

We will denote by $\mu_d = \mu_d(\Bbbk) = \{ \zeta \in \Bbbk \mid \zeta^d = 1 \}$ the set of d-th roots of unity in $\Bbbk$.

1

Ranks and secant varieties

In this chapter we formally introduce several notions of ranks such as Waring rank, border rank and tensor rank. These are closely related to higher secant varieties of projective varieties, whose properties are studied. We conclude the chapter by defining the big and little Waring problem, reporting on the progress on these questions. We generally follow the introduction by Carlini, Grieve & Oeding [CGO14] and the excellent expository paper *The Hitchhiker Guide to Secant Varieties and Tensor Decomposition* [Ber+18].

1.1. Waring rank and Border rank

Let's start by introducing the main player of the game: Homogenous polynomials and their *Waring rank*. By a "form" we mean a homogeneous polynomial $F \in \Bbbk[x_0, \dots, x_n] =: \Bbbk[\underline{x}]$ (provided the number of variables $n+1$ is understood). The vector space of homogeneous polynomials of degree d is denoted as $\Bbbk[\underline{x}]_d$.

Definition 1.1. Let $F \in \Bbbk[\underline{x}]_d$ be a form. The *Waring rank* $\mathrm{WR}(F)$ is the least integer $r \geq 0$ such that there exists a decomposition

$$F = \lambda_1 L_1^d + \cdots + \lambda_r L_r^d, \qquad L_1, \dots, L_r \in \Bbbk[\underline{x}]_1 \text{ linear forms, } \lambda_i \in \Bbbk.$$

Any such expression is called a *Waring decomposition* of F.

As mentioned in the introduction the name Waring is borrowed from the motivating classical problem from number theory of expressing natural numbers as sums of d-th powers.

Remark.

- Since $\Bbbk$ is algebraically closed, we can always write $\lambda_i L_i^d = (\sqrt[d]{\lambda_i} L_i)^d$, so we may assume that the coefficients are 1, i. e. $F = L_1^d + \cdots + L_r^d$. But notice that this is not true if we work over non-algebraically closed fields, for example $-x^2$ will never be a sum of squares in $\mathbb{R}[x]$.
- We should make sure that the Waring rank of a form does not depend on the number of variables of the ambient polynomial ring:
  If $F \in \Bbbk[x_0, \dots, x_n]_d \subseteq \Bbbk[x_0, \dots, x_n, \dots, x_{n+m}]_d$, then it should not matter in which ring we seek for a Waring decomposition. Clearly more variables, i. e. more linear forms can only make the decomposition smaller. On the other hand, in any expression

$F = \sum_{i=1}^r L_i(x_0, \ldots, x_{n+m})^d$ we can set $x_{n+1} = \cdots = x_{n+m} = 0$ to obtain a decomposition of the same length in the smaller ring.
- The Waring rank of F and $\lambda \cdot F$, $\lambda \in \Bbbk^\times$ are the same, since
$$F = \sum_{i=1}^r L_i^d \quad \leftrightarrow \quad \lambda \cdot F = \sum_{i=1}^r (\sqrt[d]{\lambda} \cdot L_i)^d.$$
Thus it makes sense to talk about the Waring rank of points of the projectivization $\mathbb{P}(\Bbbk[\underline{x}]_d)$.
- A similar argument shows that a coordinate transformation $\phi \in \mathrm{GL}(n+1, \Bbbk)$ leaves the Waring rank unchanged: $\mathrm{WR}(F \circ \phi) = \mathrm{WR}(F)$.

Example 1.2. Let $F = x_0^d + \cdots + x_k^d$. If there's any fairness in the world, then $\mathrm{WR}(f) = k + 1$.

We may assume $k = n$ (i. e. the ambient polynomial ring has variables $x_0, \ldots, x_n$). Clearly the rank is less than or equal to $n+1$. Assume that $f = \sum_{i=0}^r L_i^d$ with $r < n$. These linear forms span a proper subspace $\langle L_0, \ldots, L_r \rangle_\Bbbk \subsetneq \Bbbk[\underline{x}]_1$; WLOG let $L_0, \ldots, L_l$ be a basis of this subspace. Extend this to a basis
$$y_0 = L_0, \ldots, y_l = L_l, y_{l+1}, \ldots, y_n,$$
then with respect to these coordinates the form F is
$$F(x) = G(y) = y_0^d + \cdots + y_l^d + \left(\sum_{j=0}^l a_{l+1,j} y_j \right)^d + \cdots + \left(\sum_{j=0}^l a_{r,j} y_j \right)^d.$$
From this presentation it is easy to verify that the point $P = [0 : \cdots : 0 : 1]$ is a singularity of the projective hypersurface defined by $\{G(y) = 0\} \subseteq \mathbb{P}_{[y_0:\cdots:y_n]}$, since all partial derivatives with respect to the y_i vanish at P (recall $l < n$). But on the other hand the *Fermat hypersurface* $\mathcal{V}(X_0^d + \cdots + X_n^d) \subseteq \mathbb{P}_{[x_0:\cdots:x_n]}$ is *nonsingular* in characteristic 0 (or $> d$), a contradiction! $\lightning$

Example 1.3 (The degree 2 case). Let $F \in \Bbbk[x_0, \ldots, x_n]_2$ be a quadratic form. It can be identified with the symmetric matrix $A = [a_{ij}] \in \mathrm{Sym}(n+1, \Bbbk)$ such that $F(x) = x^\top A x$.

Claim. $\mathrm{WR}(F) = \operatorname{rank} A$, the usual matrix rank.

Proof of claim. Any symmetric matrix A can be orthogonally diagonalized (over $\Bbbk$ algebraically closed of characteristic 0), i. e. there is an matrix $U \in \mathrm{GL}(n+1, \Bbbk)$ with $UU^\top = \mathbb{1}_{n+1}$ such that $U^{-1}AU = \mathrm{diag}(\lambda_0, \ldots, \lambda_n)$. By the previous remark $\mathrm{WR}(f) = \mathrm{WR}(F \circ U)$, but
$$(F \circ U)(x) = (Ux)^\top A(Ux) = x^\top (U^{-1}AU)x = \sum_{i=0}^n \lambda_i x_i^2.$$
After re-scaling the coordinates/variables we see that the form is projectively equivalent to $x_0^2 + \cdots + x_{k-1}^d$ for $k = \mathrm{rank}(A)$, which has Waring rank k by the previous example. $\square_{\text{claim}}$

The cases $d = 2$ turns out to be quite special in several ways. For example since matrix rank can be checked using the vanishing of certain minors, the set of forms of rank $\leq r$ is a closed subset. This is *not* the general situation; for $d \geq 3$ the opposite is true:

Example 1.4. Consider $F = x_0 x_1^{d-1} \in \mathbb{C}[\underline{x}]_d$, $d \geq 3$. We will see later (example 2.10) that $\mathrm{WR}(F) = d$, but it is easy to see that $\mathrm{WR}(F) \geq 3$: Indeed, since we have the following factorization

$$L_1^d + L_2^d = L_1^d - (\sqrt[d]{-1}L_2)^d = \prod_{\zeta \in \mu_d(\mathbb{C})} (L_1 - \zeta\sqrt[d]{-1}L_2),$$

any form of Waring rank ≤ 2 must either be a single d-th power, or factor into d distinct factors (depending on L_1 and L_2 being linearly dependent or not).

But notice that F has "almost" Waring rank 2 in the sense that

$$F = \frac{1}{d} \cdot \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\big((\varepsilon x_0 + x_1)^d - x_1^d\big).$$

This example shows that the set $\{ F \in \Bbbk[\underline{x}]_d \mid \mathrm{WR}(F) \leq r \}$ may not be Zariski-closed (not even closed in the standard $\mathbb{C}$-topology!). This leads to the following definition

Definition 1.5. Let $F \in \Bbbk[\underline{x}]_d$. The *border rank* $\underline{\mathrm{WR}}(F)$ is the smallest integer $r \geq 0$ such that

$$F \in \overline{\{ G \in \Bbbk[\underline{x}]_d \mid \mathrm{WR}(G) \leq r \}} \subseteq \Bbbk[\underline{x}]_d.$$

Remark. Note that this closure could also be taken in $\mathbb{P}(\Bbbk[\underline{x}]_d)$ by the correspondence of closed subsets of $\mathbb{P}(V)$ and closed cones in V. In particular the border rank is also well-defined in $\mathbb{P}(\Bbbk[\underline{x}]_d)$.

Example 1.6 (Border rank 1). The set of forms of Waring rank 1 is the image of the morphism

$$\mathbb{P}(\Bbbk[\underline{x}]_1) \to \mathbb{P}(\Bbbk[\underline{x}]_d), \qquad [L] \mapsto [L^d].$$

This set is the image of a proper morphism (for details see section 1.3), hence closed, so $\mathrm{WR}(F) = 1$ if and only if $\underline{\mathrm{WR}}(F) = 1$.

Example 1.7. Example 1.4 combined with the previous example shows that $\underline{\mathrm{WR}}(x_0 x_1^d) = 2$, in particular the gap between Waring rank and border rank can get arbitrarily large.

This is bad news! On the one hand we are interested in the Waring rank, on the other hand we have just seen that the set of polynomials of bounded Waring rank W is not closed, so we are not in our beloved land of subvarieties of projective space. Taking the closure $\overline{W}$ gives us subvarieties (or at the very least algebraic sets), but we may fear that this closure is very different from W.

1.2. Constructible sets

To remedy the situation we introduce the notion of a *constructible set*:

Definition 1.8. Let X be a Noetherian scheme. A subset $A \subseteq X$ is *constructible* if it is a finite union of locally closed sets

$$A = \bigcup_{i=1}^{k} \underbrace{(C_i \cap O_i)}_{\text{loc. closed}}, \qquad C_i \text{ closed, } O_i \text{ open.} \tag{$*$}$$

The family of construcible sets is closed under taking intersections, unions and complements, and it is the smallest such family containing the open sets [GW20, Proposition 10.13]. Constructible sets are useful, as the set-theoretic image of a morphism is constructible:

Theorem 1.9 (*Chevalley's theorem*). *Let X, Y be noetherian schemes and $f : X \to Y$ be a morphism of finite type. Let $A \subseteq X$ be constructible, then $f(A) \subseteq Y$ is also constructible.*

The essential part of the proof is to establish the following statement:

Theorem 1.10 ([GW20, Thm. 10.19]). *Let $f : X \to Y$ be a dominant morphism of finite type between Noetherian schemes. Then $f(X)$ contains a dense open subset of Y.*

Proof. 1. Suppose we already showed the theorem for Y irreducible. If $Y = Y_1 \cup \cdots \cup Y_n$ are the irreducible components of Y, then we may apply the theorem to the restrictions $f^{-1}(Y_i) \to Y_i$ and obtain $U_i \subseteq f(X) \cap Y_i$ dense open in Y_i. Then $U_i' = U_i \setminus (\bigcup_{j \neq i} Y_j)$ is open in all Y_i, hence in Y and $U_1' \cup \cdots \cup U_n'$ is dense open in Y.

2. We reduced to the case Y irreducible; as this is a statement about the topology, we may replace X, Y with its reduced structure; hence Y is integral. Since we only need to find *some* dense open subset, we may replace Y and X with affine opens $\operatorname{Spec} A$, $\operatorname{Spec} B$ respectively. Then f is induced by an injective (f dominant) ring homomorphism $\varphi : A \to B = A[b_1, \dots, b_m]$, where A is a domain and B of finite type. To finish the proof, we need to find an $s \in A$ with $D(s) \subseteq f(\operatorname{Spec} B)$, i.e. $\operatorname{Spec} B_s \to \operatorname{Spec} A_s$ surjective.

3. Let $S = A \setminus 0$, $K = S^{-1}A$, then $S^{-1}B$ is of finite type over K, hence we can apply Noether normalization: We can find $\alpha_1, \dots, \alpha_d \in S^{-1}B$ algebraically independent over K such that $S^{-1}B$ is finite over $K[\alpha_1, \dots, \alpha_d]$ and b_i integral over this polynomial ring. Choose a common denominator $s \in S$ of all α_i and the integrality equations of the b_i, then

$$A_s \subseteq A_s[\alpha_1, \dots, \alpha_d] \subseteq B_s$$

Then $\operatorname{Spec} A_s[\underline{\alpha}] \to \operatorname{Spec} A_s$ is surjective ($\mathfrak{p} = \mathfrak{p}^{ec} = \mathfrak{p}[\underline{\alpha}] \cap A$) and $\operatorname{Spec} B_s \to \operatorname{Spec} A_s[\underline{\alpha}]$ is finite, hence also surjective ("lying over"). ∎

Proof of Chevalley's theorem. Write A as in ($*$), then $f(A) = \bigcup_{i=1}^{k} f(C_i \cap O_i)$ and we may assume that A is locally closed. Restricting f to A we may assume $X = A$ altogether.

Let $Z_0 = \overline{f(X)}$, by the previous theorem $f(X)$ contains a dense subset U_0 open in Z_0. Let $Z_1 := \overline{f(X) \cap (Z_0 \setminus U_0)} \subsetneq Z_0$. If $Z_1 = \emptyset$ then we are done, otherwise apply the same reasoning to $f^{-1}(Z_1) \to Z_1$; this yields Z_2 and so on. The process has to stop, as the Z_i are closed and Y is noetherian. $f(X)$ is the constructible set $U_0 \cup \cdots \cup U_l$. ■

Remark. A consequence of the proof is that a constructible set $A \subseteq X$ contains a dense open subset of its closure; although this can be seen more elementary from the definition.

With Chevalley's theorem in our pocket we can return to the Waring problem. We first make the following observation: Let $r, n, d \in \mathbb{N}$, $N = \max\{r, n+1\}$, then

$$\{\, F \in \Bbbk[x_0, \dots, x_n]_d \mid \mathrm{WR}(F) \le r \,\} = \underbrace{(\mathrm{Mat}(N, n+1, \Bbbk) \triangleright x_1^d + \cdots + x_r^d)}_{\subseteq \Bbbk[x_0,\dots,x_N]} \cap \Bbbk[x_0, \dots, x_n]_d,$$

where the action of matrices on forms is given by

$$A \triangleright x_1^d + \cdots + x_r^d = (Ax_1)^d + \cdots + (Ax_r)^d = \sum_{i=1}^{r} (a_{i,1}x_1 + \cdots + a_{i,N}x_N)^d.$$

Any element in the orbit clearly has Waring rank $\le r$, and any Waring decomposition is obtained by choosing a suitable matrix A. With this in mind we can prove the following theorem.

Theorem 1.11. *The following subsets of* $\Bbbk[x_0, \dots, x_n]_d$ *are constructible*

$$W_r = \left\{ F \in \Bbbk[\underline{x}]_d \;\middle|\; \mathrm{WR}(F) = r \right\},$$
$$W_{\le r} = \left\{ F \in \Bbbk[\underline{x}]_d \;\middle|\; \mathrm{WR}(F) \le r \right\}.$$

Proof. By the previous lemma we see that $W_{\le r}$ is the image of the morphism

$$m\colon \mathrm{Mat}(N, n+1, \Bbbk) \longrightarrow \Bbbk[x_1, \dots, x_N]_d, \qquad A \mapsto A \triangleright x_1^d + \cdots + x_r^d$$

intersected with $\Bbbk[x_0, \dots, x_n]_d$. This set is constructible by Chevalley's theorem, and hence $W_r = W_{\le r} \setminus W_{\le (r-1)}$ is also constructible. ■

The fact that $W_{\le r}$ is constructible shows that it is not terribly far away from being a closed subset of $\Bbbk[\underline{x}]_d$. In our most familiar setting with $\Bbbk = \mathbb{C}$ we could also talk about forms in the Euclidean closure, i. e. limits of forms (such as in example 1.4), which clearly is contained in the Zariski-closure. One could worry that the Zariski closure is much bigger than the Euclidean closure, i. e. the intuition of taking limits of forms is misleading. Fortunately, this is not the case!

Theorem 1.12. *Let $A \subseteq \mathbb{C}^n$ be a (nonempty) constructible subset. Then the Zariski closure $\overline{A}$ and Euclidean closure $\overline{A}^{\mathbb{C}}$ coincide.*

Proof. Taken from [Kra84, Satz AI.7.2]. By definition, A is a finite union of locally (Zariski) closed subsets $A = \bigcup_{i=1}^{k} A_i$. In any topological space we have

$$\overline{A_1 \cup A_2} = \overline{A_1} \cup \overline{A_2},$$

so we immediately reduce to the locally closed case, and we may assume the "closed part" to be irreducible. Hence we are in the following situation: $X \subseteq \mathbb{C}^n$ an irreducible affine variety, $\emptyset \neq A \subseteq X$ open, with the aim to show $\overline{A}^{\mathbb{C}} = X$.

Suppose first that X is a curve, let $f : \tilde{X} \to X$ be its normalization; f has finite fibres. Let $x \in M := f^{-1}(X \setminus U)$, as $\tilde{X}$ is a smooth complex curve, there is an Euclidean neighborhood $x \in D \subseteq \tilde{X}$ homeomorphic to a disc in $\mathbb{C}$. Clearly $x \in \overline{D \setminus M}^{\mathbb{C}}$ (M is finite!), and therefore $f(x) \in \overline{f(D \setminus M)}^{\mathbb{C}} = \overline{A}^{\mathbb{C}}$.

Now let X be arbitrary and $x \in X$. Choose an irreducible curve $C \subseteq X$ containing x and not disjoint from A, for example by invoking Bertini's theorem or as in [Kra84, Satz AI.4.5]. Now $A \cap C$ is nonempty open in C; by the first case $x \in \overline{A \cap C}^{\mathbb{C}}$, so $x \in \overline{A}^{\mathbb{C}}$. ■

1.3. The Veronese embedding

As promised in example 1.6 we will now expand on the observation that the (projectivized) set of d-th powers is the image of the morphism of taking d-th powers. These varieties, the *Veronese varieties*, will be of central interest to us, so it's worth studying some of their properties first. We take a *slight* detour in introducing coordinate-free projective space, which yields an elegant description of the Veronese varieties and their tangent space, following the short exposition by Brambilla & Ottavani [BO08].

Let V be $\Bbbk$-vector space of dimension $n+1$, we denote by $S = \mathrm{S}(V) = \bigoplus_{d \geq 0} \mathrm{S}^d V$ the symmetric algebra of V. We get a well-defined perfect pairing

$$\mathrm{S}^d V \times \mathrm{S}^d(V^\vee) \to \Bbbk, \qquad (v_1 \cdots v_d, \ell_1 \cdots \ell_d) \mapsto \sum_{\sigma \in S_d} \prod_{i=1}^{d} \ell_i(v_{\sigma(i)})$$

After choosing a basis $x_0, \ldots, x_n$ of V, let $X_i = x_i^\vee \in V^\vee$ its dual basis, we have

$$S := \mathrm{S}(V^\vee) \cong \Bbbk[X_0, \ldots, X_n], \qquad S_d^\vee := \mathrm{S}^d(V^\vee) \cong \Bbbk[X_0, \ldots, X_n]_d,$$

the homogeneous polynomials of degree d. If $f \in \mathrm{S}^d(V^\vee)$ corresponds to the polynomial

$\sum_{|\boldsymbol{\alpha}|=d} c_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}}$, then for $v = v_0 x_0 + \cdots + v_n x_n \in V$

$$(v^d, f) = \sum_{|\boldsymbol{\alpha}|=d} \sum_{\sigma \in S_d} c_{\boldsymbol{\alpha}} \cdot (X_0(v))^{\alpha_0} \cdots (X_n(v))^{\alpha_n} = d! \sum_{|\boldsymbol{\alpha}|=d} c_{\boldsymbol{\alpha}} \cdot v_0^{\alpha_0} \cdots v_n^{\alpha_n} = d! \cdot f(v_0, \ldots, v_n).$$

Hence $\mathrm{S}(V^\vee)$ is precisely the homogeneous coordinate ring of $\mathbb{P}(V) \cong \operatorname{Proj} \mathrm{S}(V^\vee) \cong \mathbb{P}^n$, and the pairing $(-^d, -)$ corresponds to polynomial evaluation (up to the scalar $d! \in \Bbbk^\times$). For example, the defining ideal of the closed point $[v] \in \mathbb{P}(V)$ is

$$\mathfrak{m}_{[v]} = \left\{ f \in S^\vee \;\middle|\; f(v) := (v^d, f) = 0 \right\} \subseteq \mathrm{S}(V^\vee).$$

Lemma 1.13 (*Veronese embedding*). *Let $d \in \mathbb{N}$. The (well-defined) map*

$$\nu_d \colon \mathbb{P}^n = \mathbb{P}(V) \longrightarrow \mathbb{P}^N = \mathbb{P}(\mathrm{S}^d V), \qquad [v] \mapsto [v^d]$$

defines a closed embedding. Its image is a non-degenerate regular closed (irreducible) subvariety.

By *non-degenerate* we mean that the subvariety is not contained in any hyperplane.

Proof. The map is well-defined, as $[(\lambda v)^d] = [\lambda^d v^d] = [v^d]$. Write $v = \sum_{i=0}^n v_i x_i$, then

$$v^d = (v_0 x_0 + \cdots + v_n x_n)^d = \sum_{|\boldsymbol{\alpha}|=d} \binom{d}{\boldsymbol{\alpha}} \cdot v_0^{\alpha_0} \cdots v_n^{\alpha_n} \cdot x^{\boldsymbol{\alpha}}.$$

Hence up to rescaling with the multinomial coefficients the map is the same as the polynomial map in coordinates with all possible monomials

$$[v_0 : \cdots : v_n] \mapsto [v_0^d : v_0^{d-1} v_1 : \cdots : v_n^d].$$

But this is just the closed embedding given by the *very ample* line bundle $\mathcal{L} = \mathcal{O}_{\mathbb{P}(V)}(d)$. In particular $\nu_d(\mathbb{P}^n) \cong \mathbb{P}^n$ is irreducible, regular and not contained in a hyperplane. ∎

Definition 1.14. The image $V^{d,n} := \nu_d(\mathbb{P}(V))$ is called the *Veronese variety*.

The Veronese variety $V^{d,n}$ is one of our main objects of study. It yields, for example, a simple proof of the finiteness of the Waring rank. Let $V = \Bbbk[\underline{x}]_1$, then $\mathrm{S}^d V = \Bbbk[\underline{x}]_d$.

Lemma 1.15. *The Waring rank of $F \in \Bbbk[\underline{x}]_d$ is the minimal $r \in \mathbb{N}$ such that*

$$\exists P_1, \ldots, P_r \in V^{d,n} : \qquad [F] \in \langle P_1, \ldots, P_r \rangle_{\mathbb{P}}.$$

In particular it is finite and bounded above by $\binom{d+n}{n}$.

Proof. By the previous lemma we know that $\left\langle V^{d,n} \right\rangle_{\mathbb{P}} = \mathbb{P}(\Bbbk[\underline{x}]_d)$, so such an r is guaranteed to exist. This r is the smallest number such that f can be written as a linear combination of r

d-th powers, known to us as the Waring rank. The upper bound is just $\dim \Bbbk[x_0, \ldots, x_n]_d = \binom{d+n}{n}$. ■

We first collect some properties of the Veronese variety. To this end, we transfer the natural group action $G = \mathrm{GL}(V) \circlearrowleft V$ to $\mathrm{S}^d V$ via

$$A \triangleright (v_1 \ldots v_d) := (Av_1) \ldots (Av_d).$$

The corresponding action on the coordinate ring is $\mathrm{S}^d(V^\vee) \ni f \mapsto f \circ A^{-1}$; this is compatible with the natural pairing in the sense

$$(A \triangleright -, A \triangleright -) = (-, -).$$

The map ν_d is $\mathrm{GL}(V)$-equivariant, i.e.

$$A \triangleright \nu_d([v]) = \nu_d(A[v]) \qquad \text{for all } v \in V, A \in \mathrm{GL}(V).$$

Since the action of $\mathrm{GL}(V)$ on $\mathbb{P}(V)$ is transitive, we have

$$V^{d,n} = \nu_d(\mathbb{P}(V)) = \nu_d(\mathrm{GL}(V) \cdot [e_0]) = \mathrm{GL}(V) \triangleright [x_0^d],$$

so $V^{d,n}$ is precisely the $\mathrm{GL}(V)$-orbit of a single d-th power. The following description is attributed to Lasker.

Theorem 1.16 (The tangent space of $V^{d,n}$). (i) *The tangent space $T_{[v^d]}V^{d,n}$ is the subspace*

$$T_{[v^d]}V^{d,n} = \mathbb{P}(v^{d-1}V) = \left\{ [v^{d-1}w] \;\middle|\; w \in V \right\} \subseteq \mathbb{P}(\mathrm{S}^d V).$$

(ii) *Its orthogonal in $\mathrm{S}^d(V^\vee)$ can be described as*

$$(T_{v^d}\widehat{V^{d,n}})^\perp = (\mathfrak{m}_{[v]}^2)_d \subseteq \mathrm{S}^d(V^\vee).$$

Proof. (i) Let $w \in V$, then $t \mapsto [(v+tw)^d]$ parametrizes a curve C in $V^{d,n}$ through $[v^d]$ whose tangent line $T_{[v^d]}C$ is the linearized part $[v^d + td \cdot v^{d-1}w]$. In particular $[v^{d-1}w] \in T_{[v^d]}V^{d,n}$ and hence $\mathbb{P}(v^{d-1}V) \subseteq T_{[v^d]}V^{d,n}$. Both linear spaces have dimension n, hence they are equal

(ii) To calculate the orthogonal, notice that

$$A \triangleright \mathfrak{m}_{[x_0]} = \left\{ f \circ A^{-1} \;\middle|\; (x_0^d, f) = 0 \right\} = \left\{ g \;\middle|\; ((Ax_0)^d, g) = 0 \right\} = \mathfrak{m}_{[Ax_0]}.$$

Hence if we can prove the equality for the point $[x_0^d]$, then for $v = Ax_0 \in V$

$$(T_{v^d}\widehat{V^{d,n}})^\perp = (A \triangleright T_{x_0^d}\widehat{V^{d,n}})^\perp = A \triangleright (T_{x_0^d}\widehat{V^{d,n}})^\perp = A \triangleright (\mathfrak{m}_{[x_0]}^2)_d = (\mathfrak{m}_{[v]}^2)_d.$$

Thus, without loss of generality $v = x_0$. Then $\mathfrak{m}_{[x_0]} = (X_1, \dots, X_n)$, $\mathfrak{m}_{[x_0]}^2 = (X_1^2, X_1X_2 \dots, X_n^2)$ and hence a $\Bbbk$-basis of $(\mathfrak{m}_{[x_0]}^2)_d$ is given by the monomials not divided by X_0^{d-1}. On the other hand by (i) we have $T_{x_0^d}\widehat{V^{d,n}} = \langle X_0^d, X_0^{d-1}X_1, \dots, X_0^{d-1}X_n\rangle_\Bbbk$, so they are clearly orthogonal complements of each other. ■

1.4. Secant varieties

In lemma 1.15 we saw a geometric interpretation of the Waring rank using (higher) secants. This concept can be generalized to any non-degenerate projective variety $X \subseteq \mathbb{P}^N$.

Definition 1.17. Let $X \subseteq \mathbb{P}^N$, then we denote

$$\sigma_s^\circ(X) := \bigcup_{x_1,\dots,x_s \in X} \langle x_1, \dots, x_s\rangle_{\mathbb{P}}, \qquad \sigma_s(X) := \overline{\sigma_s^\circ(X)} \subseteq \mathbb{P}^N.$$

The variety $\sigma_s(X)$ is the s-th *higher secant variety* of X.

Definition 1.18. The X-*rank* of a point $x \in \mathbb{P}^N$ is the minimal number

$$\begin{aligned} \mathrm{R}_X(x) &= \min \{ r \in \mathbb{N} \mid x \in \langle x_1, \dots, x_r\rangle_{\mathbb{P}} \text{ for some } x_1, \dots, x_r \in X \} \\ &= \min \{ r \in \mathbb{N} \mid x \in \sigma_r^\circ(X) \}. \end{aligned}$$

This is well-defined as $\langle X\rangle_{\mathbb{P}} = \mathbb{P}^N$, in particular we have the upper bound $\mathrm{R}_X(x) \le N + 1$. The statement of Lemma 1.15 is that $\mathrm{WR} = \mathrm{R}_{V^{d,n}}$.

We use this opportunity to introduce a very famous relative of the Waring rank.

Definition 1.19. Let $V_1, \dots, V_d$ be $\Bbbk$-vector spaces of dimension $n_1, \dots, n_d \ge 1$. Let $T \in V = V_1 \otimes \dots V_d$ be a tensor, then its *tensor rank* $\mathrm{rank}(T)$ is the minimal $r \ge 0$ such that T is a linear combination of r elementary tensors

$$T = \sum_{i=1}^{r} \lambda_i v_{i,1} \otimes \cdots \otimes v_{i,d}, \qquad \lambda_i \in \Bbbk, v_{i,j} \in V_j.$$

This notion of a rank can also be understood as a rank relative to a projective variety.

Example 1.20. The *Segre embedding* is an embedding of the product of projective spaces into a larger projective space, which can be expressed as

$$\sigma\colon \mathbb{P}(V_1) \times \cdots \times \mathbb{P}(V_d) \to \mathbb{P}(V_1 \otimes \cdots \otimes V_d), \qquad ([v_1], \dots, [v_d]) \mapsto [v_1 \otimes \cdots \otimes v_d].$$

The image X of σ is called a *Segre variety*, and it is easy to see that the tensor rank is precisely the X-rank. For matrices (i. e. tensors of order 2) the tensor rank coincides with the usual matrix rank.

Remark. Let $V_1 = \cdots = V_d = V$. By identifying symmetric tensors with homogeneous forms, it makes sense to talk about the Waring rank of a symmetric tensor and to compare it with its tensor rank. Since powers of linear forms are elementary tensors, we see that $\mathrm{WR}(T) \geq \mathrm{rank}(T)$, and example 1.3 shows that in the case $d = 2$ the two notions coincide. *Comon's conjecture* asks whether this is always the case, and it has been confirmed in many special cases. But there is a counterexample due to Shitov [Shi18], he costructed a symmetric tensor $T \in \mathrm{S}^3\Bbbk^{800}$ with

$$\mathrm{rank}(T) \leq 903 < \mathrm{WR}(T).$$

To study the higher secant varieties, it is useful introduce the *join* of projective varieties, we follow Ådlandsvik [Ådl87] for a particularly elegant proof of Theorem 1.27.
For affine varieties $X_1, \ldots, X_s \subseteq \mathbb{A}^N = \Bbbk^N$ we use the notation

$$X_1 + \cdots + X_s := \left\{ x_1 + \cdots + x_s \in \Bbbk^N \;\middle|\; x_i \in X_i \right\}.$$

The affine cone of a projective variety $X \subseteq \mathbb{P}(\Bbbk^{N+1})$ is denoted by $\widehat{X} \subseteq \Bbbk^{N+1}$, conversely the projective variety corresponding to a cone $Y \subseteq \Bbbk^{N+1}$ is denoted by $\mathbb{P}(Y) \subseteq \mathbb{P}(\Bbbk^{N+1})$.

Definition 1.21. Let $X_1, \ldots, X_s \subseteq \mathbb{P}^N = \mathbb{P}(\Bbbk^{N+1})$ be projective varieties. Their *join*[1] is the variety

$$X_1 \bullet \cdots \bullet X_s := \mathbb{P}\left(\overline{\widehat{X_1} + \cdots + \widehat{X_s}}\right).$$

Observe that for nonempty projective varieties $X_1, \ldots, X_s \subseteq \mathbb{P}^N$

$$\mathbb{P}(\widehat{X_1} + \cdots + \widehat{X_s}) = \bigcup_{x_1 \in X_1, \ldots, x_s \in X_s} \langle x_1, \ldots, x_s \rangle_{\mathbb{P}},$$

in particular $\sigma_s(X) = X \bullet \cdots \bullet X$ (s times).

Remark. This yields another proof of theorem 1.11, in fact for any projective variety the set $\sigma_s^\circ(X) = \mathbb{P}(\widehat{X} + \cdots + \widehat{X})$ is constructible.

1.5. Expected dimensions

The first nontrivial thing to ask about higher secant varieties is its dimension.

Theorem 1.22. *The binary operation $\bullet$ turns the set of irreducible closed subvarieties of $\mathbb{P}^N$ into a commutative semigroup. We have*

$$\dim X_1 \bullet X_2 \leq \dim X_1 + \dim X_2 + 1.$$

Proof. Commutativity is clear, for associativity it suffices to show $X_1 \bullet X_2 \bullet X_3 = (X_1 \bullet X_2) \bullet X_3$.

[1] A common notation is $J(X_1, \ldots, X_s)$, but to emphasize the semigroup structure I chose to use this notation.

Let X, Y, Z be the affine cones, then this amounts to the equality

$$\overline{X+Y+Z} = \overline{\overline{(X+Y)}+Z}.$$

"$\subseteq$" is clear, for the other inclusion consider

$$\begin{array}{ccccc} X \times Y \times Z & \longrightarrow & \overline{(X+Y)} \times Z & \longrightarrow & \overline{\overline{(X+Y)}+Z} \\ (x,y,z) & \longmapsto & (x+y,z) & & \\ & & (w,z) & \longmapsto & (w+z). \end{array}$$

Both maps are dominant, so their composition is dominant, too. Hence the image of $X \times Y \times Z$ is dense in $\overline{\overline{(X+Y)}+Z}$ and we win. The dimension can be calculated as

$$\dim(X_1 \bullet X_2) - 1 = \dim \overline{\widehat{X_1} + \widehat{X_2}} - 2 \leq \dim \widehat{X_1} - 1 + \dim \widehat{X_2} - 1 = \dim X_1 + \dim X_2. \qquad \blacksquare$$

A similar argument shows $X_1 \bullet X_2 \bullet \cdots = X_1 \bullet (X_2 \bullet (\dots))$, in particular $\sigma_s X = X^s$ in this semigroup. We immediately get the *expected dimension* of higher secant varieties:

Definition 1.23. If $X \subseteq \mathbb{P}^N$ is a non-degenerate projective variety of dimension n, then

$$\operatorname{expdim} \sigma_s(X) := \min\{sn + s - 1, N\} \geq \dim \sigma_s X,$$

this number is called the *expected dimension* of $\sigma_s X$. The difference $\delta_s := \operatorname{expdim} \sigma_s X - \dim \sigma_s X$ is the s-*defect*, and X is called s-defective if $\delta_s > 0$.

Corollary 1.24. *If a non-degenerate $X \subseteq \mathbb{P}^N$ is s-defective and $\sigma_{s+1} X \neq \mathbb{P}^N$, then X is also $(s+1)$-defective.*

Proof. From $s \mapsto s+1$ the expected dimension increases by $\Delta = \dim X + 1$, by Theorem 1.22 the actual dimension also incerases at most by Δ, so

$$\dim \sigma_{s+1} X \leq \dim \sigma_s X + \Delta < \operatorname{expdim} \sigma_s X + \Delta = \operatorname{expdim} \sigma_{s+1} X. \qquad \blacksquare$$

The expected dimension increases in steps of Δ (before the ambient space is filled). On might wonder how much smaller the increments of the actual dimension can get. A first answer is given by the following theorem.

Theorem 1.25 ([Ådl87, Corollary 1.4]). *Let $X \subseteq \mathbb{P}^N$ be a non-degenerate subvariety. If $\dim \sigma_{s+1} X \leq \dim \sigma_s X + 1$, then $\sigma_{s+1} X = \mathbb{P}^N$.*

Proof. If $\dim \sigma_{s+1} X = \dim \sigma_s X$, then equality holds for all higher secant varieties and we are done, hence we may assume $\dim \sigma_{s+1} X = \dim \sigma_s X + 1$. We call $y \in \mathbb{P}^N$ a vertex if $J(X, y) = X$, the set of vertices is $\operatorname{Vert}(X)$; one verifies immediately that this is a linear subspace of $\mathbb{P}^n$.

Claim. $X \subseteq \operatorname{Vert}(X^{s+1})$.

Proof of claim. By assumption $X^s \bullet X \not\subseteq X^s$, so X contains a dense open subset of non-vertices $U = X \setminus \operatorname{Vert}(X^s)$. For $x \in U$ we have

$$\begin{aligned} X^s \subsetneq X^s x \subseteq X^{s+1} \quad &\Longrightarrow \quad \dim X^s x = \dim X^{s+1} \\ &\overset{\text{irred.}}{\Longrightarrow} \quad X^s x = X^{s+1} \\ &\Longrightarrow \quad X^{s+1} x = X^s x^2 \overset{x^2=x}{=} X^s x = X^{s+1}. \end{aligned}$$

Hence x is a vertex of X^{s+1}, and thus $X = \overline{U} \subseteq \operatorname{Vert}(X^s)$. $\square_{\text{claim}}$

Now $X \subseteq \operatorname{Vert}(X^{s+1})$ is a linear subspace containing X, so $X, X^2, \ldots, X^{s+1} \subseteq \operatorname{Vert}(X^{s+1})$, but by definition this means $X^{s+1} = X^{s+2} = \cdots = \mathbb{P}^N$. ■

Example 1.26 (Curves are never defective). Let $\mathcal{C} \subseteq \mathbb{P}^N$ be a non-degenerate curve, for example a *rational normal curve* $V^{d,1}$. Then by theorem 1.22 and 1.25

$$\dim \mathcal{C}^{s+1} \leq \dim \mathcal{C}^s + 2, \qquad \dim \mathcal{C}^{s+1} \leq \dim \mathcal{C}^s + 1 \text{ iff } \mathcal{C}^{s+1} = \mathbb{P}^N.$$

Thus, the dimension of the higher secant varieties increase in steps of 2 before it fills $\mathbb{P}^N$, so $\sigma_s \mathcal{C}$ *always* has the expected dimension.

1.6. Terracini's first lemma

If we want to detect s-defectiveness, we need some way to calculate the dimension of the $\sigma_s X$. As $\dim \sigma_s X$ is the dimension of the (Zariski) tangent space of a general point on $\sigma_s X$, it would be sufficient to understand the tangent space of points on secant varieties, or more generally on joins. This will be accomplished by Terracini's lemma below, we give a modern treatment due to Ådlandsvik [Ådl87, Corollary 1.10 & 1.11].

Theorem 1.27 (*Terracini's lemma* - general affine form).
Let $X_1, \ldots, X_s \subseteq \mathbb{A}^N$ be affine varieties, $Y := \overline{X_1 + \cdots + X_s}$. For $x = (x_1, \ldots, x_s) \in \prod_i X_i$, $y := x_1 + \cdots + x_s$ we have

$$T_{x_1} X_1 + \cdots + T_{x_s} X_s \subseteq T_y Y,$$

and there is a dense open subset $U \subseteq Y$ such that equality holds whenever $x_1 + \cdots + x_s \in U$.

Proof. Let $i_j \colon \Bbbk^n \to \Bbbk^n$ be the identity and consider the addition morphism

$$f \colon \prod_{j=1}^{s} \Bbbk^n \to \Bbbk^n, \qquad (a_1, \ldots, a_s) \mapsto i_1(a_1) + \cdots + i_s(a_s) = \sum_{j=1}^{s}.$$

Here $T_{x_i}\mathbb{k}^n = \mathbb{k}^n$ and $\mathrm{d}_x f = \mathrm{d}_x(i_1 + \cdots + i_s) = \mathrm{d}_x i_1 + \cdots + \mathrm{d}_x i_s$ is just addition. By functoriality of d we get $\mathrm{d}_x(f_{|X}) = (\mathrm{d}_x f)_{|Y}$ ($X = \prod_{i=1}^s X_i$), hence we get the first claim by

$$T_{x_1}X_1 + \cdots + T_{x_s}X_s = \mathrm{im}(T_x f_{|X}) \subseteq T_y Y. \tag{$*$}$$

For the second claim recall that by *generic smoothness* in characteristic 0 [Har77, Corollary III.10.7] there is a dense open subset $V \subseteq \mathbb{k}^n$ such that $f\colon f^{-1}(V) \to V$ is a smooth morphism. Smooth morphisms have the property that $T_f\colon T_x \longrightarrow T_{f(x)}$ is surjective [Har77, Proposition III.10.4(iii)], in our case this precisely means equality in ($*$). ∎

By projectivizing everything we immediately obtain the projective version:

Theorem 1.28 (Terracini's lemma - general projective form).
Let $X_1, \ldots, X_s \subseteq \mathbb{P}^N$ be projective varieties. For $(x_1, \ldots, x_s) \in \prod_i X_i$ and $y \in \langle x_1, \ldots, x_s\rangle_{\mathbb{P}}$ we have

$$\langle T_{x_1}X_1, \ldots, T_{x_s}X_s\rangle_{\mathbb{P}} \subseteq T_y(X_1 \bullet \cdots \bullet X_s).$$

There is a dense open subset $U \subseteq Y$ such that equality holds whenever $y \in U$.

Corollary 1.29 (Terracini's lemma for secant varieties).
For a general collection of points $x_1, \ldots, x_r \in X$ and a general point $y \in \langle x_1, \ldots, x_r\rangle_{\mathbb{P}}$ we have

$$T_y\sigma_r(X) = \langle T_{x_1}X, \ldots, T_{x_r}X\rangle_{\mathbb{P}}\,.$$

Remark. Notice that the argument in the proof of theorem 1.27 relies on the crucial fact that $\operatorname{char}\mathbb{k} = 0$. For different, more calculus-style proofs (over $\mathbb{C}$) see [BO08, Lemma 2.2] or [Ber+18, Lemma 1].

To show the usefulness of this theorem we use it to compute the dimension of the secant variety of the *Veronese surface*.

Example 1.30 (The Veronese surface). Consider

$$S = V^{2,2} = \left\{ [L^2] \,\middle|\, L \in \mathbb{k}[x_0, x_1, x_2]_1 \right\} \subseteq \mathbb{P}(\mathbb{k}[x_0, x_1, x_2]_2) \cong \mathbb{P}^5.$$

The expected dimension of $X = \sigma_2 S$ is $\min\{3 \cdot 2 - 1, 5\} = 5$. But the actual dimension turns out to be strictly smaller! The dimension can be calculated as the dimension of the tangent space of a generic point of X, by Terracini's lemma this is given as $\dim\left\langle T_{[v^2]}V^{2,2}, T_{[w^2]}V^{2,2}\right\rangle_{\mathbb{P}}$ for general linear forms v, w. Let $\widehat{S} \subseteq \mathbb{k}[x_0, x_1, x_2]_2$ be the affine cone of S, then

$$\dim\left\langle T_{[v^2]}S, T_{[w^2]}S\right\rangle_{\mathbb{P}} = \dim(T_{v^2}\widehat{S} + T_{w^2}\widehat{S}) - 1 \overset{\text{Lin.Alg.}}{=} 3 + 3 - \dim(T_{v^2}\widehat{S} \cap T_{w^2}\widehat{S}) - 1.$$

By Theorem 1.16(i) we see that

$$T_{v^2}\widehat{S} \cap T_{w^2}\widehat{S} = v\Bbbk[\underline{x}]_1 \cap w\Bbbk[\underline{x}]_1 = \langle vw \rangle_{\Bbbk}$$

which is in fact *not* trivial. Hence X is a four-dimensional irreducible subvariety of $\mathbb{P}^5$, i. e. a hypersurface.

In fact, we knew this already! Recall Example 1.3; there we have seen that the set of quadratic forms of rank $\leq r$ correspond to the symmetric matrices of rank $\leq r$. Here we are considering the projective space $\mathbb{P}(\mathrm{Sym}(3, \Bbbk))$ and the subset of matrices of rank ≤ 2. This is equivalent to not having full rank, i. e. $\det A = 0$. Hence the identification $\mathbb{P}(\Bbbk[\underline{x}]_2) \cong \mathbb{P}(\mathrm{Sym}(3, \Bbbk))$ maps secant variety $\sigma_2 S$ to the hypersurface cut out by $\{\det A = 0\}$.

1.7. The big and little Waring problem

With all the objects and notions defined so far we can finally ask the big (and small) questions.

Problem 1.31 (Little Waring problem). *Given $n, d \geq 1$, what is the maximum Waring rank*

$$g(n, d) := \max \{ \mathrm{WR}(F) \mid F \in \Bbbk[x_0, \ldots, x_n]_d \}?$$

This problem still remains largely unsolved. A typical algebro-geometer[2] might suggest a variation of this problem asking not for the maximal rank, but for the *generic Waring rank*.

Problem 1.32 (Big Waring problem). *Given $n, d \geq 1$, what is the Waring rank $G(n, d)$ of a generic form $F \in \Bbbk[x_0, \ldots, x_n]_d$, i. e. the rank of a dense open set of forms?*

We should ensure first that such a set actually exists. We use the notation W_r and $W_{\leq r}$ from Theorem 1.11.

Lemma 1.33 (Existence of the generic Waring rank). *There is precisely one $G \in \mathbb{N}$ such that W_G contains a dense open subset.*

Proof. We have a chain $\overline{W_{\leq 1}} \subseteq \overline{W_{\leq 2}} \subseteq \overline{W_{\leq 3}} \subseteq \ldots$ (this is the chain of secant varieties), which eventually fills out the ambient space. Let G be the smallest integer with $\overline{W_{\leq G}} = \Bbbk[\underline{x}]_d$. Since $W_{\leq G}$ is constructible, it contains a dense open subset U of its closure $\Bbbk[\underline{x}]_d$. Since $\overline{W_{\leq G-1}}$ is a proper subset of the whole space, we see that W_G contains the dense open subset $U \cap (\overline{W_{\leq G-1}})^{\mathrm{c}}$. Furthermore, all other W_j, $j \neq G$ are contained in the complement of this set, so they do not contain a dense open. ■

It turns out that the big Waring problem has a complete answer: All numbers $G(n, d)$ are known. Even better, we know the dimension of *all* varieties $\sigma_s V^{d,n}$.

[2]Perhaps the generic one.

Theorem 1.34 (Alexander-Hirschowitz, [BO08, Theorem 1.2]). *The higher secant varieties of the Veronese varieties $\sigma_s V^{d,n}$, $n, d, s \geq 1$, have expected dimension*

$$\dim \sigma_s V^{d,n} = \operatorname{expdim} \sigma_s V^{d,n} = \min \left\{ \binom{n+d}{d} - 1, sn + s - 1 \right\}$$

with the following complete list of exceptions

d	n	s	δ_s	$\dim \sigma_s V^{d,n}$
2	≥ 2	$2 \dots n$	$\binom{s}{2}$	$sn + s - 1 - \binom{s}{2}$
3	4	7	1	33
4	2	5	1	14
4	3	9	1	33
4	4	14	1	68

Chapter 3 will be concerned with an outline of a proof of this theorem.

Example 1.35. We already know that the theorem is true in the case $n = 1$, as the higher secant varieties of any non-degenerate curve have expected dimension, see example 1.26.

From this result we obtains the solution to the big Waring problem.

Corollary 1.36 (The solution of the big Waring problem).

$$G(n, d) = \left\lceil \frac{1}{n+1} \binom{n+d}{d} \right\rceil$$

with the following list of exceptions

d	n	$G(n,d)$
2	$\forall$	$n + 1$
3	4	8
4	2	6
4	3	10
4	4	15

Proof. The generic Waring rank is the smallest s such that $\sigma_s V^{d,n}$ fills up its ambient space, i. e. $\dim \sigma_s V^{d,n} = \binom{n+d}{d}$. The Alexander-Hirschowitz theorem gives a complete list of these numbers, so this is simple arithmetic. For example, in the non-exceptional case we want the least s such that

$$sn + s + 1 \geq \binom{n+d}{d} - 1 \qquad \Longleftrightarrow \qquad s \geq \frac{1}{n+1} \binom{n+d}{d}.$$

Thus, rounding up this rational number gives the desired value. For the exceptional cases notice that $G = s+1$ with s the largest number such that $V^{d,n}$ is s-defective. ∎

Example 1.37.

- The case $d = 2$ (which is always exceptional) has been witnessed in example 1.3: The general matrix in $\mathrm{Sym}(n+1, \Bbbk)$ has rank $n+1$.
- For $n = 1$ we have $G(1,d) = \lceil \frac{1}{2}\binom{d+1}{d} \rceil = \lceil \frac{d+1}{2} \rceil$. We know (or learn in example 2.10) that for example $\mathrm{WR}(x_0 x_1^{d-1}) = d$, which is almost twice the generic rank!

Can the ratio $g(n,d)/G(n,d)$ get bigger than in the case $n = 1$? A surprising yet elementary result by Blekherman & Teitler [BT14] tells us that this is not the case.

Theorem 1.38. *For any non-degenerate projective variety $X \subseteq \mathbb{P}^N$ let $g(X)$, $G(X)$ be the maximum and the generic X-rank, defined in the same fashion as in Problem 1.31, 1.32. Then*

$$G(X) \leq g(X) \leq 2 \cdot G(X).$$

Proof. The first inequality is trivial. For the second let $U \subseteq \mathbb{P}^n$ be a dense open subset consisting of elements of generic rank, and let $x \in \mathbb{P}^n$, $u \in U$. Let $\ell = \langle x, u \rangle_{\mathbb{P}}$, then $\ell \cap U$ is dense open in ℓ, hence containing $v \neq u$, so $\ell = \langle u, v \rangle_{\mathbb{P}}$. u and v lie in linear subspaces spanned by $G(X)$ points respectively, so x lies in a subspace generated by $2 \cdot G(X)$ points. ∎

In our case of Waring rank we get (in the non-exceptional cases)

$$g(n,d) \leq 2 \cdot G(n,d) = 2 \cdot \left\lceil \frac{1}{n+1}\binom{n+d}{d} \right\rceil.$$

This is (asymptotically) less than the previously best upper bound

$$g(n,d) \leq \binom{n+d-1}{n} - \binom{n+d-5}{n-2} - \binom{n+d-6}{n-2}.$$

obtained by Ballico & De Paris [BD17].

2

The rank of monomials

In this chapter we will learn about a useful tool to compute the Waring rank of a specific form, the *Apolarity Lemma*. As an application, we compute the Waring rank of all monomials, and even of sums of coprime monomials. Finally, we discuss the Strassen conjecture concerning some additivity property of ranks.

2.1. The apolarity action

In this section we follow [Ber+18, §2.1.4] and [IK99, §1.1].

Let $T = \Bbbk[x_0, \dots, x_n]$, $S = \Bbbk[X_0, \dots, X_n]$ be polynomial rings. We let S act on T by interpreting $X_i \mathrel{\hat{=}} \frac{\partial}{\partial x_i}$, more formally:

Definition 2.1. The *apolarity action* is defined on the monomial basis as

$$\circ\colon S_i \times T_j \longrightarrow T_{j-i}, \qquad X^{\boldsymbol{\alpha}} \circ x^{\boldsymbol{\beta}} := \begin{cases} \frac{\boldsymbol{\beta}!}{(\boldsymbol{\beta}-\boldsymbol{\alpha})!} x^{\boldsymbol{\beta}-\boldsymbol{\alpha}} & \text{if } \boldsymbol{\alpha} \leq \boldsymbol{\beta}, \\ 0 & \text{otherwise.} \end{cases}$$

for multi-indices $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{N}_0^{n+1}$ of length $|\boldsymbol{\alpha}| = i$, $|\boldsymbol{\beta}| = j$.

Remark.

- This action turns T into a S-module, the only thing to verify is

$$(X^{\boldsymbol{\alpha}} \cdot X^{\boldsymbol{\alpha}'}) \circ x^{\boldsymbol{\beta}} = X^{\boldsymbol{\alpha}+\boldsymbol{\alpha}'} \circ x^{\boldsymbol{\beta}} = \begin{cases} \frac{\boldsymbol{\beta}!}{(\boldsymbol{\beta}-\boldsymbol{\alpha}-\boldsymbol{\alpha}')!} x^{\boldsymbol{\beta}-\boldsymbol{\alpha}-\boldsymbol{\alpha}'} & \text{if } \boldsymbol{\alpha}+\boldsymbol{\alpha}' \leq \boldsymbol{\beta}, \\ 0 & \text{otherwise} \end{cases}$$
$$= X^{\boldsymbol{\alpha}} \circ \begin{cases} \frac{\boldsymbol{\beta}!}{(\boldsymbol{\beta}-\boldsymbol{\alpha}')!} x^{\boldsymbol{\beta}-\boldsymbol{\alpha}'} & \text{if } \boldsymbol{\alpha}' \leq \boldsymbol{\beta}, \\ 0 & \text{otherwise} \end{cases} = X^{\boldsymbol{\alpha}} \circ (X^{\boldsymbol{\alpha}'} \circ x^{\boldsymbol{\beta}}).$$

  Thus, we can view T as a set of functions on which the ring of linear differential operators S acts on.
- The fact that the set of differential operators (with composition as multiplication) can be modeled as the commutative (!) polynomial algebra S is motivated by Schwarz' theorem from multivariate calculus.

- The apolarity action induces a *perfect pairing* $S_d \times T_d \to \Bbbk$, in fact this is the pairing from section 1.3 in a different guise!

Definition 2.2. Let $I \subseteq S$ be a homogeneous ideal. The *inverse system* I^{-1} is the set of polynomials annihilated by all elements of I:

$$I^{-1} := \{ F \in T \mid \partial \circ F = 0 \ \forall \partial \in I \}$$

The inverse system I^{-1} is an graded S-submodule of T, but not an ideal (not closed under multiplication). The definition is reminiscent of an orthogonal complement, and indeed we have the following properties

Lemma 2.3. *Let $I, J \subseteq S$ be homogeneous ideals.*

(i) *The graded components of I^{-1} are orthogonal to the graded components of I_d:*

$$(I^{-1})_d = I_d^{\perp} := \{ F \in T_d \mid \partial \circ F = 0 \ \forall \partial \in I_d \}.$$

(ii) *If $I \subseteq J$, then $J^{-1} \subseteq I^{-1}$.*
(iii) $(I+J)^{-1} = I^{-1} \cap J^{-1}$ *and* $(I \cap J)^{-1} = I^{-1} + J^{-1}$.
(iv) $\dim_{\Bbbk} I_d^{-1} = \dim_{\Bbbk} S_d - \dim_{\Bbbk} I_d = \dim_{\Bbbk}(S/I)_d$.

Proof. (i) Clearly $(I^{-1})_d \subseteq I_d^{\perp}$. Conversely let $f \in I_d^{\perp}$, then f is annihilated by all differential operators $\partial \in I_d$, and trivially by operators of higher degree (by definition of $\circ$). Let $\partial \in I_{d-k}$, $k < d$, then $X^{\alpha} \cdot \partial \in I_d$ for all $|\alpha| = k$, since I is an ideal. But

$$0 = (X^{\alpha} \cdot \partial) \circ f = X^{\alpha} \circ (\partial \circ f)$$

implies that no monomial x^{α} can ocur in $\partial \circ f$, hence $\partial \circ f = 0$.

(ii)-(iv) By (i) and the fact that $\circ$ defines a perfect pairing $S_d \times T_d \to \Bbbk$, this is a consequence of the corresponding facts from linear algebra about perfect pairings. ■

Of particular interest is the case of *monomial ideals*. We recall their defining properties:

Lemma 2.4 (Monomial ideals). *For an ideal $I \subseteq \Bbbk[\underline{X}]$ the following are equivalent:*

(i) *I is generated by a (finite) set of monomials;*
(ii) *For any $f = \sum_{\alpha} c_{\alpha} X^{\alpha} \in I$ we have $X^{\alpha} \in I$ whenever $c_{\alpha} \neq 0$;*
(iii) *I is homogeneous and each I_d admits a $\Bbbk$-basis consisting of monomials.*

Example 2.5. If $I \subseteq S$ is a monomial ideal, then Lemma 2.3 and 2.4 show that

$$(I^{-1})_d = I_d^{\perp} = \langle \{ x^{\alpha} \mid X^{\alpha} \notin I_d \} \rangle_{\Bbbk} \subseteq T_d.$$

We can interpret the polynomial ring S as a ring of functions on T_1. Indeed if $L = a_0x_0 + \cdots + a_nx_n$ is a linear form, then similarly to the calculation in section 1.3

$$X^{\boldsymbol{\alpha}} \circ (a_0x_0 + \cdots + a_nx_n)^d = X^{\boldsymbol{\alpha}} \circ \sum_{|\boldsymbol{\beta}|=d} \binom{d}{\boldsymbol{\beta}} a^{\boldsymbol{\beta}} x^{\boldsymbol{\beta}} = a^{\boldsymbol{\alpha}} \binom{d}{\boldsymbol{\alpha}} \cdot \frac{\alpha!}{1} = d! \cdot a_0^{\alpha_0} \cdots a_n^{\alpha_n},$$

so if $\partial = g(X_0, \ldots, X_n)$, then $\partial \circ L^d = d! \cdot g(a_0, \ldots, a_n)$ is function evaluation up to a constant.

With this in mind we can identify S with the homogeneous coordinate ring of $\mathbb{P}(T_1)$, and it makes sense to talk about the vanishing ideal $I(X) \subseteq S$ of a subset $X \subseteq \mathbb{P}(T_1)$.

Example 2.6. If $P = [L] \in \mathbb{P}(T_1)$ is a point, then $(\mathfrak{m}_P^{-1})_d = \Bbbk \cdot L^d$. Indeed, it suffices to show this for a particular form $L = x_0$. Then $\mathfrak{m}_{[x_0]} = (X_1, \ldots, X_n)$ is a monomial ideal with $(\mathfrak{m}[x_0])_d$ consisting of polynomials without X_0^d, by Example 2.5 we conclude $(\mathfrak{m}_{[x_0]}^{-1})_d = \Bbbk x_0^d$.

Definition 2.7. Let $F \in T_d$ be a form. The *apolar ideal* of F is

$$F^\perp := \{ \partial \in S \mid \partial \circ F = 0 \},$$

it is a homogeneous ideal in S.

We are ready to formulate and prove the *Apolarity Lemma*.

Theorem 2.8 (*Apolarity Lemma*). *Let $L_1, \ldots, L_s \in T_1$ be linear forms and $\mathbb{X} = \{[L_1], \ldots, [L_s]\} \subseteq \mathbb{P}(T_1)$. Then for a form $F \in T_d$ the following are equivalent:*

(i) *$F = \lambda_1 L_1^d + \cdots + \lambda_s L_s^d$ for some $\lambda_i \in \Bbbk$;*
(ii) *$I(\mathbb{X}) \subseteq F^\perp$.*

Proof. (i)$\Rightarrow$(ii) Let $\partial \in I(\mathbb{X})$, then $\partial \circ L_i^d = \partial([L_i]) = 0$ for $i = 1, \ldots, s$, and by linearity $\partial \circ F = 0$, so $\partial \in F^\perp$.

(ii)$\Rightarrow$(i) Notice that $I(\mathbb{X}) = \mathfrak{m}_{[L_1]} \cap \cdots \cap \mathfrak{m}_{[L_s]}$. If $I(\mathbb{X}) \subseteq F^\perp$, then

$$F \in (F^\perp)_d^{-1} \subseteq I(\mathbb{X})_d^{-1} \overset{2.3}{=} (\mathfrak{m}_{[L_1]}^{-1})_d + \cdots + (\mathfrak{m}_{[L_1]}^{-1})_d \overset{2.6}{=} \Bbbk L_1^d + \cdots + \Bbbk L_s^d,$$

so F is a linear combination of $L_1^d, \ldots, L_r^d$. ∎

This theorem gives a characterization of the Waring rank.

Corollary 2.9. *Let $0 \neq F \in T$ be a form, then*

$$\mathrm{WR}(F) = \min \left\{ r \in \mathbb{N}_+ \;\middle|\; F^\perp \textit{ contains the ideal of a set of } r \textit{ distinct points} \right\}.$$

Moreover, the linear forms from a Waring decomposition correspond to such points.

Remark. If we know that $F^\perp \supseteq I(\{[L_1], \ldots, [L_r]\})$, then it is not immediately clear how to find coefficients λ_i such that $F = \sum_{i=1}^r \lambda_i L_i^r$. But since we know that there exists *some* solution, we[1] can solve the over-determined system of linear equations in the coefficients given by this equation.

Example 2.10. We can now prove that $F = x_0 x_1^{d-1}$ has Waring rank d. Indeed, we may work in $T = \Bbbk[x_0, x_1]$, $S = \Bbbk[X_0, X_1]$.

$$X_0^i X_1^j \circ x_0 x_1^{d-1} = 0 \text{ iff } i \geq 2 \text{ or } j \geq d \qquad \Longrightarrow \qquad f^\perp = (X_0^2, X_1^d),$$

which contains the ideal $(X_0^d - X_1^d) = \bigcap_{\zeta \in \mu_d}(X_0 - \zeta X_1)$ of d distinct points. Hence $\mathrm{WR}(F) \leq d$.

On the other hand, suppose $F^\perp$ contains the ideal $I(P_1, \dots, P_s)$ of fewer distinct points. Each $\mathfrak{m}_{P_1}$ is a principal prime ideal (∂_i) (height 1 primes in $\Bbbk[X_0, X_1]$), so

$$\partial := \partial_1 \cdots \partial_s \in I(P_1, \dots, P_r) \subseteq (X_0^2, X_1^d).$$

As $s < d$, ∂ must be a multiple of X_0^2, but we assumed the points to be distinct! ↯

A useful relation between the ideal $F^\perp$ and the *colon ideal* $I : J := \{ x \in S \mid xJ \subseteq I \}$ is the following:

Lemma 2.11. *For any form F and any $\partial \in S$ we have $F^\perp : (\partial) = (\partial \circ F)^\perp$.*

Proof. Let $D \in S$, then

$$D \in F^\perp : (\partial) \quad \Leftrightarrow \quad D\partial \in F^\perp \quad \Leftrightarrow \quad D\partial \circ F = 0 \quad \Leftrightarrow \quad D \in (\partial \circ F)^\perp. \qquad \blacksquare$$

2.2. Hilbert functions

Before we can fully utilize the apolarity lemma, we need another tool, the *Hilbert function* of a projective scheme, or more generally, of a graded algebra. As before, let $S = \Bbbk[X_0, \dots, X_n]$.

Definition 2.12. Let $A = \bigoplus_{d \geq 0} A_d$ be a finitely generated graded $\Bbbk$-algebra. The *Hilbert function* is defined as $\mathrm{HF}(A, d) := \dim_\Bbbk A_d$. The *hilbert series*[2] is the formal power series $\mathrm{HS}(A, t) = \sum_{d=0}^\infty (\dim_\Bbbk A_d) t^d \in \mathbb{Z}[[t]]$.

Let $X \subseteq \mathbb{P}^n = \operatorname{Proj} \Bbbk[t_0, \dots, t_n]$ be a projective scheme with homogeneous coordinate ring $A := \Bbbk[\underline{t}]/I(X)$. The *Hilbert function* of X is $\mathrm{HF}(X, d) := \mathrm{HF}(A, d) = \mathrm{HF}(\Bbbk[\underline{t}], d) - \dim_\Bbbk I(X)_d$.

If A is generated by elements $f_1, \dots, f_s$ in degrees $d_1, \dots, d_s$, then $\mathrm{HS}(A, t)$ is a rational function of the form

$$\mathrm{HS}(A, t) = \frac{f(t)}{(1 - t^{d_1}) \cdots (1 - t^{d_s})}$$

[AM69, Theorem 11.1]. In the case of quotients of a polynomial ring $A = S/I$ (so $f_i = X_i$, $d_i = 1$) the denominator simplifies; we will see a particular instance of this in Lemma 2.16.

[1] Or a computer.
[2] Also called *Poincaré series*.

A consequence of this is that the Hilbert function equals a polynomial for $d \gg 0$ [AM69, Corollary 11.2], called the *Hilbert polynomial* $\mathrm{HP}(X,d)$. If $\iota\colon X \hookrightarrow \mathbb{P}^n$ is a projective variety, then by Serre vanishing [Har77, III.5.2(b)] $\mathrm{HF}(X,d)$ becomes the polynomial

$$\mathrm{HF}(X,d) = \chi(\mathbb{P}^n, \iota_*\mathcal{O}_X(d)), \quad d \gg 0, \qquad \chi(\mathbb{P}^n,\mathcal{F}) = \sum_{i\geq 0}(-1)^i h^i(\mathbb{P}^n,\mathcal{F}).$$

One can show that the Hilbert polynomial of X has degree $r = \dim X$. The *degree* of $X \subseteq \mathbb{P}^n$ is defined as $r!$ times the leading coefficient of $\mathrm{HP}(X,d)$ (this is an integer). The degree is the number of intersection points of X with a general linear subspace $L \subseteq \mathbb{P}^n$ of codimension r.

Example 2.13.

- The Hilbert function of the polynomial ring S, and hence that of $X = \mathbb{P}^n$ is

$$\mathrm{HF}(\mathbb{P}^n, d) = \mathrm{HF}(\Bbbk[X_0,\dots,X_n], d) = \binom{n+d}{d},$$

which is a polynomial function for $n \geq 0$. The Hilbert series takes the simple form $\mathrm{HS}(S,t) = (1-t)^{-(n+1)}$.
- If A is Artinian, then $A_d = 0$ for $d \gg 0$, and $\mathrm{HS}(A,1) = \sum_{d=0}^{\infty}\mathrm{HF}(A,d) = \dim_{\Bbbk} A \in \mathbb{N}_0$. In this case the Hilbert polynomial is simply 0.
- If $X, Y \subseteq \mathbb{P}^n$ are projective schemes, then

$$\mathrm{HF}(X\cup Y, d) = \mathrm{HF}(X,d) + \mathrm{HF}(Y,d) - \mathrm{HF}(X\cap Y, d).$$

Indeed, this follows from the short exact sequence of $\Bbbk$-vectorspaces

$$0 \longrightarrow \underbrace{I(X)\cap I(Y)}_{=I(X\cup Y)} \xrightarrow{f\mapsto (f,f)} I(X)\oplus I(Y) \xrightarrow{(f,g)\mapsto f-g} \underbrace{I(X)+I(Y)}_{=I(X\cap Y)} \longrightarrow 0.$$

In particular, the Hilbert *polynomial* is additive for disjoint subschemes and $d \gg 0$, because $S/(I(X)+I(Y))$ is Artinian.
But since $I(X)+I(Y)$ may not span $\mathfrak{m}_+$ entirely, the Hilbert *function* may *not* be additive for each degree! In some sense, this is why chapter 3 on the Alexander-Hirschowitz theorem is non-trivial (and in fact quite involved).

A useful tool for calculating the Hilbert function is the following lemma.

Lemma 2.14. *Let $I \subseteq S$ be a homogeneous ideal and $\partial \in S_1$ not a zero-divisor in S/I. Then*

$$\mathrm{HF}(S/I, d) = \sum_{i=0}^{d} \mathrm{HF}(S/(I+(\partial)), i).$$

Proof. Not being a zero-divisor in S/I means that we have a short exact sequence

$$0 \longrightarrow (S/I)[-1] \xrightarrow{\cdot\overline{\partial}} (S/I) \twoheadrightarrow \underbrace{(S/I)/(\overline{\partial})}_{=S/(I+(\partial))} \longrightarrow 0,$$

where $A[d]$ is the algebra with $A[d]_i = A_{i+d}$. Looking at the dimension of the graded pieces we obtain

$$\mathrm{HF}(S/I, i-1) - \mathrm{HF}(S/I, i) + \mathrm{HF}(S/(I+(\partial)), i) = 0, \qquad i \geq 0.$$

(with $\mathrm{HF}(S/I, -1) = 0$). Summing over $i = 0, \ldots, d$ yields the desired formula. ■

An important special case is the Hilbert function of a *complete intersection*.

Definition 2.15. (i) Let A be a commutative ring (for our purpose $\Bbbk[\underline{t}]$). $f_1, \ldots, f_k \in A$ is an *regular sequence* if f_i is not a zero-divisor in $A/(f_1, \ldots, f_{i-1})$ for $i = 1, \ldots, k$.

(ii) An ideal $I \subseteq A$ is called a *complete intersection* if it is generated by a regular sequence.

Theorem 2.16 (Hilbert series of a complete intersection). *Let $I \subseteq S$ be a complete intersection ideal generated by the regular sequence of homogeneous elements $f_1, \ldots, f_k$, $d_i := \deg f_i \geq 1$. Then*

$$\mathrm{HS}(S/I, t) = \frac{(1-t^{d_1})\cdots(1-t^{d_k})}{(1-t)^{n+1}} = \frac{(1+t+\cdots+t^{d_1-1})\cdots(1+t+\cdots+t^{d_k-1})}{(1-t)^{n+1-k}}.$$

Proof. The proof is similar to the previous lemma. Let $A^j = S/(f_1, \ldots, f_j)$, so $A^0 = S$ and $A^k = S/I$. Since f_j is regular on A^{j-1}, we get short exact sequences

$$0 \longrightarrow A^{j-1}[-d_j] \xrightarrow{\cdot f_j} A^{j-1} \twoheadrightarrow A^j \longrightarrow 0.$$

Taking dimensions of the graded components and forming the corresponding power series we get

$$t^{d_j} \cdot \mathrm{HS}(A^{j-1}, t) - \mathrm{HS}(A^{j-1}, t) + \mathrm{HS}(A^j, t) = 0,$$

or, after rearranging $\mathrm{HS}(A^j, t) = (1-t^{d_j}) \cdot \mathrm{HS}(A^{j-1}, t)$. Thus we can apply induction on j to reduce to the case $A^0 = S$, which has Hilbert series $(1-t)^{-(n+1)}$ by example 2.13:

$$\mathrm{HS}(S/I, t) = \cdots = (1-t^{d_1})\cdots(1-t^{d_k}) \cdot \mathrm{HS}(S, t) = \frac{(1-t^{d_1})\cdots(1-t^{d_k})}{(1-t)^{n+1}}.$$

The second equality from the theorem is just an application of the geometric sum formula $1 - t^d = (1-t)(1+t+\cdots+t^{d-1})$ and canceling the $(1-t)$ terms. ■

Corollary 2.17. *Let I be as in the previous lemma, $A := S/I$.*

(i) *If A is Artinian (this is the case if $k = n+1$, so $\mathcal{V}(I) = \emptyset$), then its $\Bbbk$-dimension is $d_1 \cdots d_k$.*

(ii) *If A is one-dimensional (this is the case if $k = n$, also $\mathbb{X} = \mathcal{V}(I)$ is finite), then* $\deg \mathbb{X} = d_1 \cdots d_k = \operatorname{len}(\mathbb{X})$, *the length of* $\mathbb{X}$.

Recall that the *length of a zero-dimensional* $\Bbbk$*-scheme* $\mathbb{X}$ is defined as

$$\operatorname{len}(\mathbb{X}) = \dim_{\Bbbk} \mathcal{O}_{\mathbb{X}}(\mathbb{X}) = \sum_{P \in |\mathbb{X}|} \dim_{\Bbbk} \mathcal{O}_{\mathbb{X},P}.$$

Proof. (i) As mentioned in example 2.13 we have $\operatorname{HS}(A, 1) = \dim_{\Bbbk} A$. By the previous theorem the Hilbert series of A is actually the polynomial

$$\operatorname{HS}(A, t) = (1 + t + \cdots + t^{d_1 - 1}) \cdots (1 + t + \cdots + t^{d_k - 1}) \in \mathbb{Z}[t],$$

so plugging in $t = 1$ yields $\dim_{\Bbbk} A = d_1 \cdots d_k$.

(ii) Since $\mathbb{X}$ is finite, we can choose a linear form $\partial \in S_1$ not vanishing on any point in the support of $\mathbb{X}$. Then $f_1, \ldots, f_n, \partial$ is a regular sequence and by (i) and Lemma 2.14 for $d \gg 0$ we get

$$\operatorname{HF}(A, d) = \sum_{i=0}^{d} \operatorname{HF}(S/(f_1, \ldots, f_k, \partial), i) = d_1 \cdots d_k.$$

So the Hilbert polynomial of $\mathbb{X}$ is the constant $d_1 \cdots d_k$, which is its degree (by definition). The second equality is a general fact:
Notice that $\mathbb{X}$ lies in the hyperplane $H = \{\partial \neq 0\}$, after a change of coordinates we may assume $\partial = X_0$. Hence $\mathbb{X}$ is (isomorphic to) $\operatorname{Spec} \Bbbk[X_1, \ldots, X_n]/J$ with some dehomogenization of I. Since the Hilbert function becomes constant, for $d \gg 0$ we get that the map

$$(S/I)_d \longrightarrow \Bbbk[X_1, \ldots, X_n]/J, \qquad f \mapsto f_{|H}$$

is an isomorphism, hence $\ell(\mathbb{X}) = \deg \mathbb{X}$. ■

2.3. The Waring rank of monomials

Armed with the Apolarity lemma and some knowledge of Hilbert functions, we are ready to prove

Theorem 2.18. *Let $x_0^{d_0} \cdots x_n^{d_n} \in \Bbbk[\underline{x}]$ be a monomial. After renaming the variables we may assume $1 \leq d_0 \leq \cdots \leq d_n$. Then*

$$\operatorname{WR}(x_0^{d_0} \cdots x_n^{d_n}) = \frac{1}{d_0 + 1} \prod_{i=0}^{n} (d_i + 1).$$

We follow the proof by Carlini, Catalisano & Geramita [CCG12] . We return to the notation from section 1 with $S = \Bbbk[\underline{X}]$ the polynomial ring of differential operators acting on $T = \Bbbk[\underline{x}]$.

Proof of Theorem 2.18. To ease notation, we consider the monomial

$$F = x_0^{d_0-1} \cdots x_n^{d_n-1}, \qquad 2 \le d_0 \le \cdots \le d_n.$$

Just as in example 2.10 we see that

$$X^{\alpha} \circ x_0^{d_0-1} \cdots x_n^{d_n-1} = 0 \text{ iff } \alpha_j \ge d_j \text{ for some } 0 \le j \le n \qquad \Longrightarrow \qquad F^{\perp} = (X_0^{d_0}, \dots, X_n^{d_n}). \tag{$*$}$$

<u>$\mathrm{WR}(F) \le d_1 \cdots d_n$</u>: Consider the points $P_{\xi} = [1 : \xi_1 : \cdots : \xi_n] \in \mathbb{P}^n$, $\xi_i \in \mu_{d_i}$ (d_i-th roots of unity), these are distinct and clearly in the vanishing set of

$$I := (X_1^{d_1} - X_0^{d_1}, \dots, X_n^{d_n} - X_0^{d_n}).$$

I is a complete intersection ideal cutting out a projective subscheme $\mathbb{X} = \operatorname{Proj} S/I$ of dimension 0 and degree $d_1 \cdots d_n$. But $\mathbb{X}$ contains at least the $d_1 \cdots d_n$ distinct points P_{ξ}, so it must be *the* reduced scheme of these points. As $I \overset{(*)}{\subseteq} F^{\perp}$, the Apolarity Lemma yields the number $d_1 \cdots d_n$ as a lower bound for $\mathrm{WR}(F)$.

<u>$\mathrm{WR}(f) \ge d_1 \cdots d_n$</u>: Let $I := I(\mathbb{X}) \subseteq F^{\perp}$ be an ideal of s distinct points $\mathbb{X} = \{P_1, \dots, P_s\}$; our goal is to show that $s \le d_1 \cdots d_n$. Consider the colon ideal

$$I' := I : (X_0) = I(\mathbb{X}) : I(\mathcal{V}(X_0)) = I(\mathbb{X} \setminus \mathcal{V}(X_0)),$$

I' is an ideal of the set of $s' \le s$ points $\mathbb{X}' = \mathbb{X} \setminus \mathcal{V}(X_0)$. We have

$$I' + (X_0) \subseteq \underbrace{(f^{\perp} : (X_0)) + (X_0)}_{=: J} \overset{2.11}{=} (X_0, X_1^{d_1}, \dots, X_n^{d_n}).$$

Notice that by definition of I' the linear form X_0 is not a zero-divisor in S/I'. Therefore we can apply lemma 2.14 to get for $t \gg 0$

$$\begin{aligned} s \ge s' = \mathrm{HF}(S/I', t) &\overset{2.14}{=} \sum_{i=0}^{t} \mathrm{HF}(S/(I' + (X_0)), i) \\ &\ge \sum_{i=0}^{t} \mathrm{HF}(S/J, i) = \sum_{i=0}^{\infty} \mathrm{HF}(S/(X_0, X_1^{d_1}, \dots, X_n^{d_n}), i) \overset{2.17}{=} d_1 \cdots d_n. \end{aligned}$$

■

Example 2.19. Theorem 2.18 reproves example 2.10: $\mathrm{WR}(x_0^1 x_1^{d-1}) = d$.

Also, for $d_0 = \cdots = d_n = 1$ we get $\mathrm{WR}(x_0 \cdots x_n) = 2^n$, the corresponding linear forms being

$x_0 + \xi_1 x_1 + \cdots + \xi_n x_n$, $x_i \in \{\pm 1\}$. This explains the decomposition

$$x_0 \cdots x_n = \frac{1}{2^n n!} \sum_{\xi \in \{\pm 1\}^n} \xi_1 \cdots \xi_n \cdot (x_0 + \xi_1 x_1 + \cdots + \xi_n x_n)^n$$

presented in the introduction.

Remark. We remark that the given result holds true in characteristic $\operatorname{char} \Bbbk > d$. Indeed, the only concern might be that we used that there are $d_i + 1$ distinct $(d_i + 1)$-th roots of unity; this is not true if $\operatorname{char} \Bbbk \mid (d_i + 1)$. But the only possibility for $d_i + 1 > d$ is $d_0 = d$, $F = x_0^d$, and in this case F has trivially Waring rank 1.

2.4. Sums of monomials

We follow Carlini et al [CCG12] again to prove that the Waring rank is additive for sum of coprime monomials, generalizing theorem 2.18.

Theorem 2.20. *Let $F_1, \dots, F_k$ be coprime monomials (i. e. in disjoint sets of variables) of the same degree d. Let $F = F_1 + \cdots + F_k$. If $d = 1$ then $\mathrm{WR}(F) = 1$. If $d \geq 2$, then*

$$\mathrm{WR}(F_1 + \cdots + F_k) = \mathrm{WR}(F_1) + \cdots + \mathrm{WR}(F_k).$$

The strategy resembles the proof of theorem 2.18, also using colon ideals. We need a lemma regarding the Hilbert function first.

Lemma 2.21. *Let $J_1, \dots, J_k \subseteq \mathfrak{m}_+ \subseteq S$ be homogeneous ideals with $(J_1 \cap \cdots \cap J_{j-1}) + J_j = \mathfrak{m}_+$ for $2 \leq j \leq k$, then for $t \geq 0$*

$$\mathrm{HF}(S/(J_1 \cap \cdots \cap J_k, t) = \mathrm{HF}(S/J_1, t) + \cdots + \mathrm{HF}(S/J_k, t) + \delta_{t,0} \cdot (k-1).$$

Proof. We use induction on k. For $k = 1$ there is nothing to prove. In the general case we have a short exact sequence

$$0 \longrightarrow S/(J_1 \cap \cdots \cap J_k) \longrightarrow S/(J_1 \cap \cdots \cap J_{k-1}) \oplus S/J_k \longrightarrow S/(J_1 \cap \cdots \cap J_{k-1} + J_k) \longrightarrow 0$$

The rightmost algebra is just $S/\mathfrak{m}_+ \cong \Bbbk$, so the Hilbert functions satisfy

$$\mathrm{HF}(S/(J_1 \cap \cdots \cap J_k, t) - \mathrm{HF}(S/(J_1 \cap \cdots \cap J_{k-1}), t) - \mathrm{HF}(S/J_k, t) + \underbrace{\mathrm{HF}(\Bbbk, t)}_{= \delta_{t,0}} = 0.$$

Rearranging and using the induction hypothesis yields the result. ∎

Proof of theorem 2.20. We first rule out some low degree cases.

- If $d = 1$, then F is linear and has trivially Waring rank 1.
- If $d = 2$, then we are in the setting of ranks of symmetric matrices, see example 1.3. Then the sum corresponds to symmetric block diagonal matrices, and the rank is additive on block matrices, so the theorem is true in this case.

$\Rightarrow$ We may assume $d > 2$.

The Waring rank of $F = F_1 + \cdots + F_k$ is clearly less than or equal to the sum of the Waring ranks. Thus, by the apolarity lemma we are required to show that if $F^\perp$ contains the ideal $I \subseteq S$ of s distinct points, then $s \geq \sum_{i=1}^k \mathrm{WR}(F_i)$. As in the single monomial case, we will assume

$$F_i = x_{i,0}^{\alpha_{i,0}} \cdots x_{i,n_i}^{\alpha_{i,n_i}}, \qquad 1 \leq \alpha_{i,0} \leq \cdots \leq \alpha_{i,n_i}.$$

Consider the colon ideal $I' = I : (X_{1,0}, \ldots, X_{k,0})$, this is again an ideal of $s' \leq s$ points $P_1, \ldots, P_{s'}$ not lying inside $\mathcal{V}(X_{1,0}, \ldots, X_{r,0})$; we wish to prove that $s' \geq \sum_{i=1}^k \mathrm{WR}(F_i)$.

Claim. There exists a linear polynomial $\partial = a_1 X_{1,0} + \cdots + a_k X_{k,0} \in S_1$, $\alpha_i \in \Bbbk^\times$, which is not a zero-divisor in S/I'.

Proof of claim. ∂ being a zero-divisor in S/I' means that ∂ vanishes on one of the points. Since $P_j \notin \mathcal{V}(X_{1,0}, \ldots, X_{k,0})$, there is some $X_{i_j,0}$ not vanishing on P_j, and hence the following finite union of proper subspaces does not fill the space of linear forms ($\#\Bbbk = \infty$):

$$\left(\bigcup_{j=1}^{s'} \{\, \partial \mid \partial(P_j) = 0 \,\}\right) \cup \left(\bigcup_{i=1}^{k} \langle X_{1,0}, \ldots, \widehat{X_{i,0}}, \ldots, X_{k,0} \rangle_\Bbbk\right) \subsetneq \langle X_{1,0}, \ldots, X_{k,0} \rangle_\Bbbk.$$

Thus there exists a ∂ with the desired property. $\square_{\text{claim}}$

Using $X_{i,0} \circ F = X_{i,0} \circ F_i$ (disjoint sets of variables!), we get

$$\begin{aligned} I' + (\partial) &\subseteq (F^\perp : (X_{1,0}, \ldots, X_{k,0})) + (\partial) \\ &= \big(F^\perp : (X_{1,0})\big) \cap \cdots \cap \big(F^\perp : (X_{k,0})\big) + (\partial) \\ &\overset{2.11}{\subseteq} \underbrace{\big((X_{1,0} \circ F_1)^\perp + (\partial)\big)}_{=:J_1} \cap \cdots \cap \underbrace{\big((X_{k,0} \circ F_k)^\perp + (\partial)\big)}_{=:J_k} \end{aligned} \qquad (*)$$

We can explicitly calculate the J_i. Let w be the number of 1's among the smallest of the exponents of the monomials $\alpha_{1,0}, \ldots, \alpha_{r,0}$; after rearranging the F_i we may assume that $\alpha_{1,0} = \cdots = \alpha_{w,0} = 1$. We make the following observations:

(i) If $1 \leq i \leq w$, then $X_{i,0} \circ F_i = x_{i,1}^{\alpha_{i,1}} \cdots x_{i,n_i}^{\alpha_{i,n_i}}$. In particular, $(X_{i,0} \circ F_i)^\perp$ contains ∂ anyways and we have

$$J_i = (\{\, X_{i',j} \mid i' \neq i, 0 \leq j \leq n_{i'} \,\} \cup \{X_{i,0}, X_{i,1}^{\alpha_{i,1}+1}, \ldots, X_{i,n_i}^{\alpha_{i,n_i}+1}\}).$$

(ii) If $w < i \leq k$, then $X_{i,0} \circ F_i = x_{i,0}^{\alpha_{i,0}-1} x_{i,1}^{\alpha_{i,1}} \cdots x_{i,n_i}^{\alpha_{i,n_i}}$. As $X_{i',0} \in (\partial_{i,0} \circ F_i)^\perp$ for $i' \neq i$, adding ∂ to a set of generators has the same effect as adding $X_{i,0}$. Thus we get *exactly* the same kind of description as in (i).

(iii) Notice that $J_1 \cap \cdots \cap J_{i'-1}$ and $J_{i'}$ together contain *all* $X_{i,j}$'s, so they generate $\mathfrak{m}_+$. This means that we can apply lemma 2.21.

(iv) The intersection $J_1 \cap \cdots \cap J_k$ is generated by the powers of the $X_{i,j}$ in the latter set of generators in (i). The only linear generators among these are $X_{1,0}, \ldots, X_{k,0}$, so

$$\dim_{\Bbbk} (J_1 \cap \cdots \cap J_k)_1 = k.$$

Our estimate of $I' + (\partial)$ as the intersection of the J_i will be *almost* good enough for our purposes, we only need to consider the dimension of the degree 1 component:

Claim. $\dim_{\Bbbk} (I' + (\partial))_1 = 1$, or equivalently, $(I')_1 = 0$.

Proof of claim. Let $\ell = \sum_{i=1}^{k} \sum_{j=0}^{n_i} b_{i,j} X_{i,j} \in (I')_1$. By definition of $I' = I : (X_{1,0}, \ldots, X_{k,0})$ we have

$$\ell \cdot X_{1,0}, \ldots, \ell \cdot X_{k,0} \in I \subseteq F^\perp.$$

This enforces some constraints on the coefficients: Since $d \geq 3$, we can compare coefficients in

$$\begin{aligned} 0 = \ell \circ (X_{i,0} \circ F_i) = \alpha_{i,0} \Big(& b_{i,0} \cdot (\alpha_{i,0} - 1) x_{i,0}^{\alpha_{i,0}-2} x_{i,1}^{\alpha_{i,1}} \cdots x_{i,n_i}^{\alpha_{i,n_i}} \\ & + \sum_{j=1}^{n_i} b_{i,j} \cdot \alpha_{i,j} x_{i,0}^{\alpha_{i,0}-1} x_{i,1}^{\alpha_{i,1}} \cdots x_{i,j}^{\alpha_{i,j}-1} \cdots x_{i,n_i}^{\alpha_{i,n_i}} \Big). \end{aligned}$$

We conclude $b_{i,1} = \cdots = b_{i,n_i} = 0$ for $i = 1, \ldots, k$, and thus $\ell = b_{1,0} X_{1,0} + \cdots + b_{k,0} X_{k,0}$. This shows that ℓ vanishes not only on the s' points in $\mathcal{V}(I')$ as assumed, but also on the remaining points in the subspace $\mathcal{V}(X_{1,0}, \ldots, X_{k,0})$. Thus $\ell \in I \subseteq F^\perp$, but $d = \deg F \geq 2$, so there are no nonzero linear forms in $F^\perp$! $\square_{\text{claim}}$

Now we are ready to complete the proof of the theorem. For $t \gg 0$,

$$
\begin{aligned}
s' = \mathrm{HF}(S/I', T) &= \sum_{i=0}^{t} \mathrm{HF}(S/(I' + (\partial)), i) \\
&\overset{\text{claim}}{=} (\dim_{\Bbbk} S_1 - 1) + \sum_{\substack{i=0 \\ i \neq 1}}^{t} \mathrm{HF}(S/(I' + (\partial)), i) \\
&\overset{(*)}{\leq} (\dim_{\Bbbk} S_1 - 1) + \sum_{\substack{i=0 \\ i \neq 1}}^{t} \mathrm{HF}(S/(J_1 \cap \cdots \cap J_k), i) \\
&\overset{\text{(iv)}}{=} (\dim_{\Bbbk} S_1 - 1) - (\dim_{\Bbbk} S_1 - k) + \sum_{i=0}^{t} \mathrm{HF}(S/(J_1 \cap \cdots \cap J_k), i) \\
&\overset{2.21}{=} (k-1) - (k-1) + \sum_{i=0}^{t} \mathrm{HF}(S/J_1, i) + \cdots + \sum_{i=0}^{t} \mathrm{HF}(S/J_k, i).
\end{aligned}
$$

As the J_i are generated by regular sequences, Corollary 2.17 tells us that

$$
\sum_{i=0}^{\infty} \mathrm{HF}(S/J_1, i) = (\alpha_{i,1} + 1) \cdots (\alpha_{i,n_i} + 1) \overset{2.18}{=} \mathrm{WR}(F_i),
$$

and thus $s \geq s' \geq \mathrm{WR}(F_1) + \cdots + \mathrm{WR}(F_k)$. ■

Example 2.22. If you, like me, are a fan of cracking nuts with sledgehammers, then you will appreciate the fact that Theorem 2.20 proves

$$
\mathrm{WR}(x_1^d + \ldots x_k^d) = k.
$$

The natural next question is to ask about the Waring rank of other families of polynomials related to monomials. An interesting family are the *elementary symmetric polynomials*

$$
e_{n,d} = \sum_{1 \leq i_1 < \cdots < i_d \leq n} x_{i_1} \cdots x_{i_d} \in \Bbbk[x_1, \ldots, x_n]_d, \qquad 1 \leq d \leq n.
$$

If $1 < d < n$, then $e_{n,d}$ is not a sum of coprime monomials, so the previous results can't be applied directly. H. Lee obtained the following results:

Theorem 2.23 ([Lee16]). (i) *For $d = 2k+1$ odd, $n \geq d$, we have a power sum decomposition*

$$
2^{d-1} d! e_{n,d} = \sum_{\substack{I \subseteq \{1,\ldots,n\} \\ |I| \leq k}} (-1)^{|I|} \binom{n-k-|I|-1}{k-|I|} \cdot \left(\delta(I,1)x_1 + \cdots + \delta(I,n)x_n\right)^d,
$$

where $\delta(I,i) = -1$ if $i \in I$ and $+1$ otherwise. This decomposition is optimal, *so*

$$\mathrm{WR}(e_{n,d}) = \sum_{i=0}^{\frac{d-1}{2}} \binom{n}{i}.$$

(ii) *For $d = 2k$ even, $n > d$, then we have a decomposition*

$$2^d(n-d)!e_{n,d} = \sum_{\substack{I \subseteq \{1,\dots,n\} \\ |I| \leq k}} (-1)^{|I|} \binom{n-k-|I|-1}{k-|I|} (n-2|I|) \cdot \left(\delta(I,1)x_1 + \cdots + \delta(I,n)x_n\right)^d,$$

This decomposition is known to be sub-optimal in special cases, but we have the bounds

$$\left(\sum_{i=0}^{d/2} \binom{n}{i}\right) - \binom{n-1}{d/2} \leq \mathrm{WR}(e_{n,d}) \leq \sum_{i=0}^{d/2} \binom{n}{i}.$$

2.5. Strassen's conjecture

Theorem 2.20 shows that when adding monomials in *disjoint* sets of variables, their Waring rank adds up, too. The same is true for all degree 2 forms, since the Waring rank corresponds to matrix rank, which is additive on block diagonal matrices. Thus, one may be tempted to formulate the following

Problem 2.24 (Symmetric direct sum conjecture). *Consider forms $F_i \in \Bbbk[x_{0,i},\dots,x_{n_i,i}]_d$, $d \geq 2$, in disjoints sets of variables. Is it true that*

$$\mathrm{WR}(F_1 + \cdots + F_k) = \mathrm{WR}(F_1) + \cdots + \mathrm{WR}(F_k)?$$

In the tensor rank setting the tensor rank gives a measure of how many multiplications are required to calculate a certain multilinear map. If we have for example maps $F_1\colon V_1 \times W_1 \longrightarrow \Bbbk$ and $F_2\colon V_2 \times W_2 \longrightarrow \Bbbk$, then one may consider the "product map"

$$F_3\colon (V_1 \oplus V_2) \times (W_1 \otimes W_2) \longrightarrow \Bbbk, \qquad F_3((v_1,v_2),(w_1,w_2)) = F_1(v_1,w_1) + F_2(v_2,w_2).$$

In matrix form this corresponds to taking the block diagonal matrix consisting of two blocks for F_1, F_2 respectively. Since the (ordinary) matrix rank is additive on these block diagonals, one may conjecture that any "optimal algorithm" (respectively: minimal tensor decomposition) must respect this direct sum and compute F_1 and F_2 apart from each other. This leads to the original *direct sum conjecture* by Strassen [Str73]

Problem 2.25 (Strassen's direct sum conjecture). *Given tensors $T_i \in \bigotimes_{j=1}^{d} V_{i,j}$, $d \geq 2$ and consider* *the direct sum*

$$T_1 \oplus \cdots \oplus T_k \in \bigotimes_{j=1}^{d} \left(\bigoplus_{i=1}^{k} V_{i,j} \right).$$

Then is it true that $\mathrm{R}(T_1 \oplus \cdots \oplus T_k) = \mathrm{R}(T_1) + \cdots + \mathrm{R}(T_k)$ *?*

This conjecture has been verified in many special cases, but it turns out to be *false* in this generality, as Shitov constructed explicit counterexamples [Shi19]; the symmetric version is still open, though.

The symmetric Strassen conjecture has been proven for other families of forms. Carlini et al. introduced the concept of *e-computability* [Car+15] to bring many of these cases together.

Theorem 2.26. *Let $F_1, \ldots, F_k$ be forms of degree d in disjoint sets of variables. The symmetric Strassen conjecture holds if each F_i is of one of the following forms:*

- *F_i is a monomial;*
- *F_i is a form in ≤ 2 variables;*
- *$F_i = x_0^a(x_1^b + \cdots + x_n^b)$ or $F_i = x_0^a(x_0^b + x_1^b + \cdots + x_n^b)$ with $a + 1 \geq b$;*
- *$F_i = x_0^a(x_1^b + x_2^b)$ or $F_i = x_0^a(x_0^b + x_1^b + x_2^b)$;*
- *$F_i = x_0^a G(x_1, \ldots, x_n)$, where $G^\perp = (g_1, \ldots, g_n)$ is a complete intersection ideal and $\deg(g_j) > a$ for $j = 1, \ldots, n$;*
- *$F_i = \det([x_j^k]_{j,k=0}^n)$ is a Vandermonde determinant.*

3

The Alexander-Hirschowitz Theorem

In this chapter we will sketch the promised proof of the Alexander-Hirschowitz theorem 1.34. The proof in the original paper [AH95] has been substantially simplified; still we will only be able to sketch the main argument due to its length and difficulty. We follow the expositions by Brambilla & Ottavani [BO08] and by Hà & Mantero [HM21], who focus more on the algebro-geometric and the commutative algebra aspect of this theorem respectively.

3.1. Polynomial interpolation

First, a correction: Theorem 1.34 is *not* what is commonly called *the* Alexander-Hirschowitz theorem, but rather an equivalent formulation in the language of secant varieties to Veronese varieties. The *true* Alexander-Hirschowitz theorem is Theorem 3.3, and in this section we will prove the equivalence with theorem 1.34.

Definition 3.1. A *double point* $2P \subseteq \mathbb{P}^n$ supported in $P \in \mathbb{P}^n$ is the non-reduced closed subscheme defined by the homogeneous ideal $\mathfrak{m}_P^2$.

Notice that $2P$ is a scheme of length $\operatorname{len}(2P) = n + 1$. Double points allow us to express when a hypersurface $\mathcal{V}(f) \subseteq \mathbb{P}^n$ is singular at P.

Lemma 3.2. *Let $f \in S_d$ be a polynomial, $2P$ a double point. Then the following are equivalent:*

(i) $f \in I(2P) = \mathfrak{m}_P^2$.
(ii) *$\mathcal{V}(f)$ is singular at $P \in \mathcal{V}(f)$;*
(iii) *All partial derivatives of f vanish at P;*

Proof. (ii)⇔(iii): A projective hypersurface is singular at P if and only if

$$f(P) = \frac{\partial f}{\partial x_0}(P) = \cdots = \frac{\partial f}{\partial x_n}(P) = 0.$$

By the "Euler relation" $d \cdot f = \sum_{i=0}^n X_i \frac{\partial f}{\partial X_0}$ we may omit $f(P) = 0$ in this system of equations.

(i)⇔(ii): Both statements are invariant under a change of coordinates, so we may assume $P = [1 : 0 : \cdots : 0]$. Then $(\mathfrak{m}_P^2)_d$ has a $\Bbbk$-basis of monomials not divided by X_0^{d-1} and it is an explicit calculation to verify that $f \in (\mathfrak{m}_P^2)_d$ is equivalent to the vanishing of f and its first derivatives at P. ∎

So by (iii) f being in $I(2P)_d$ imposes $n+1$ linear relations on the coefficients of f. Similarly, $f \in I(\mathbb{X})_d$, $\mathbb{X}$ consisting of s double points, imposes $s(n+1)$ linear equations on f, so one would expect that $I(\mathbb{X})_d \subseteq S_d$ has codimension $s \cdot (n+1)$ (if it is nonzero). This is the content of the “true” Alexander-Hirschowitz theorem.

Theorem 3.3 (Alexander-Hirschowitz).
Let $\mathbb{X}$ be a general collection of s double points and $I(\mathbb{X})_d = \left\{ f \in S_d \mid f \text{ singular at } \mathbb{X} \right\}$. Then

$$\operatorname{codim} I(\mathbb{X})_d = \min \left\{ s(n+1), \binom{n+d}{d} \right\}$$

except for the following cases

d	n	s	$\operatorname{codim} I(\mathbb{X})_d$
2	≥ 2	$2 \ldots n$	$s(n+1) - \binom{s}{2}$
3	4	7	34
4	2	5	15
4	3	9	34
4	4	14	69

We now show that theorem 3.3 and theorem 1.34 are equivalent to each other. If $\mathbb{X} = \{P_1, \ldots, P_s\}$, then we sometimes write $2\mathbb{X}$ to mean $\{2P_1, \ldots, 2P_s\}$.

Lemma 3.4. *For $d, n, s \geq 1$, and $\mathbb{X} = \{2P_1, \ldots, 2P_s\}$ a scheme of s general double points we have*

$$\sigma_s V^{d,n} + 1 = \operatorname{codim} I(\mathbb{X})_d.$$

In fact, $\mathbb{X} \subseteq \mathbb{P}^n \cong V^{d,n}$ can be any collection of double points $\{P_1, \ldots, P_s\}$ such that the “generality” assumption from Terracini’s lemma 1.29 for $V^{d,n}$ is satisfied and a general point in $\langle P_1, \ldots, P_s \rangle_{\mathbb{P}} \subseteq \sigma_s V^{d,n}$ is regular.

Proof. This is essentially a consequence of Lasker’s theorem 1.16 and Terracini’s lemma 1.27. Let $[v_i] = P_i$ be representatives, then

$$I(\mathbb{X})_d = \bigcap_{i=1}^{s} (\mathfrak{m}_{[v_i]}^2)_d \overset{1.16}{=} \bigcap_{i=1}^{s} (T_{v_i^d} \widehat{V^{d,n}})^{\perp} = \left(\sum_{i=1}^{s} T_{v_i^d} \widehat{V^{d,n}} \right)^{\perp} \overset{1.27}{=} \left(T_{v_1^d + \cdots + v_s^d} \widehat{\sigma_s V^{d,n}} \right)^{\perp}.$$

Thus, taking (co)dimensions and using the generality assumption we get

$$\operatorname{codim}_{\mathbb{k}} I(\mathbb{X})_d = \dim_{\mathbb{k}} I(\mathbb{X})_d^{\perp} = \dim_{\mathbb{k}} T_{v_1^d + \cdots + v_s^d} \widehat{\sigma_s V^{d,n}} = \dim \sigma_s V^{d,n} + 1. \qquad \blacksquare$$

If we impose the condition of passing through a single point on a space of functions, we expect the dimension to decrease by 1. As noticed before, requiring a singularity at a point imposes $n+1$ equations, so we expect the dimension to decrease by this amount. This leads to the following definition:

Definition 3.5. We say that a $\mathbb{X} = \{2P_1, \ldots, 2P_s, Q_1, \ldots, Q_t\}$ collection of ordinary and double points *imposes independent conditions* on $\mathcal{O}_{\mathbb{P}^n}(d)$ if $\operatorname{codim} I(\mathbb{X})_d = \min\left\{(n+1)s + t, \binom{n+d}{d}\right\}$.

More generally, if $\mathbb{X} \subseteq \mathbb{P}^n$ is a zero-dimensional scheme of length $\ell = \operatorname{len}(\mathbb{X})$, then we say that $\mathbb{X}$ is $\mathrm{AH}_n(d)$ if $\operatorname{codim} I(\mathbb{X})_d = \min\left\{\ell, \binom{n+d}{d}\right\}$.

So the statement of the Alexander Hirschowitz theorem is that a general collection of double points in $\mathbb{P}^n$ imposes independent conditions on $\mathcal{O}_{\mathbb{P}^n}(d)$ apart from the listed exceptional cases. This is where the name *polynomial interpolation* comes from, in the one-dimensional case this is very familiar:

Example 3.6. We can reinterpret example 1.35 in this setting. If $\mathbb{X} \subseteq \mathbb{P}^1$ is a scheme of s double points, then we are looking for binary forms $f \in K[X_0, X_1]$ singular at each point $P_1, \ldots, P_s$. We may assume that $[0:1]$ is not among the P_i, then by the Euler relation $\frac{\partial f}{\partial X_0} = X_0 \frac{1}{X_0}(d \cdot f - \frac{\partial f}{\partial X_0})$ we may discard the condition $\frac{\partial f}{\partial X_0} = 0$. After dehomogenization this is equivalent to the following polynomial interpolation problem:

> Given $p_1, \ldots, p_s \in \Bbbk$, find polynomials $f(T) \in \Bbbk[T]_d$ with $f(p_1) = f'(p_1) = 0$ for $i = 1, \ldots, s$.

We can give a closed-form solution to this as

$$f(T) = g(T) \cdot \prod_{i=1}^{s} (T - p_i)^2, \qquad g(T) \in \Bbbk[T]_{d-2s}.$$

This is precisely the expected dimension of $I(\mathbb{X})_d$, so *any* (not just a general) collection of double points imposes independent conditions on $\mathcal{O}_{\mathbb{P}^1}(d)$!

3.2. Specialization of points

In this section we give an overview of the proof strategy. Some obvious difficulties are:

(i) We have 3 degrees of freedom:
   - The dimension n of the ambient space $\mathbb{P}^n$;
   - The number s of general double points;
   - The degree d of hypersurfaces through these points.

(ii) It could be hard to verify if a given configuration of double points is "general enough".

(iii) If one wants to apply induction, it is not obvious how "general position" can be translated to lower-dimensional spaces.

We first take care of the first concern. If $\mathbb{X}, \mathbb{X}' \subseteq \mathbb{P}^n$ are schemes of points with multiplicities, we write $\mathbb{X} \subseteq \mathbb{X}'$ if

$$\mathbb{X} = \{m_1 P_1, \ldots, m_s P_s\}, \quad \mathbb{X}' = \{m_1' P_1, \ldots, m_{s'}' P_{s'}\}, \qquad s' \geq s, m_i' \geq m_i$$

(for us only the case $m_i \in \{1, 2\}$ is relevant). This is equivalent to $I(\mathbb{X}) \supseteq I(\mathbb{X}')$.

Lemma 3.7. *Let $\mathbb{X} \subseteq \mathbb{P}^n$ be a scheme of points with multiplicities that is $\mathrm{AH}_n(d)$.*

(i) *If $\operatorname{codim} I(\mathbb{X})_d = \binom{n+d}{d}$, then any larger scheme $\mathbb{X}' \supseteq \mathbb{X}$ is also $\mathrm{AH}_n(d)$.*
(ii) *If $\operatorname{codim} I(\mathbb{X})_d = \operatorname{len}(\mathbb{X})$, then any smaller scheme $\mathbb{X}' \subseteq \mathbb{X}$ is also $\mathrm{AH}_n(d)$.*
(iii) *In proving the Alexander-Hirschowitz theorem for d, n in a non-exceptional case, it suffices to consider general schemes of double points with the following number of s points:*

$$\underline{\mathfrak{s}} := \left\lfloor \frac{1}{n+1}\binom{n+d}{d} \right\rfloor \leq s \leq \overline{\mathfrak{s}} := \left\lceil \frac{1}{n+1}\binom{n+d}{d} \right\rceil.$$

Proof. (i) Since $\dim_{\Bbbk} I(\mathbb{X})_d = 0$ and $\mathbb{X} \subseteq \mathbb{X}'$ we have $I(\mathbb{X}') \subseteq I(\mathbb{X}) = 0$. Thus $\operatorname{codim}_{\Bbbk} I(\mathbb{X}')_d = \binom{n+d}{d}$ as well, so $\mathbb{X}'$ is $\mathrm{AH}_n(d)$.

(ii) $\mathbb{X}$ imposes a system of $\operatorname{len}(\mathbb{X})$ linear equations on S_d. If $\mathbb{X}$ is $\mathrm{AH}_n(d)$ and has codimension $\ell(X)$, then this system of eqations has maximal rank, so the subsystem of equations coming from $\mathbb{X}$ must also have maximal rank. We conclude that $\operatorname{codim}_{\Bbbk} I(\mathbb{X}') = \operatorname{len}(\mathbb{X}')$, so it also has expected codimension.

(iii) The two (possibly coinciding) numbers $\underline{\mathfrak{s}}, \overline{\mathfrak{s}}$ are the largest resp. smallest possible number of points such that the minimum in $\min\{s(n+1), \binom{n+d}{d}\}$ is attained in the first resp. second argument. If general subsets of s double points are $\mathrm{AH}_n(d)$, then by (i) and (ii) any other number of general double points is $\mathrm{AH}_n(d)$. ■

Remark. While the precise statement is slightly different/stronger, morally speaking this is a translation of corollary 1.24 into this setting, so solving the big Waring problem is *essentially the same* as proving the Alexander-Hirschowitz theorem.

This basically removes the parameter s from the regular cases of the theorem. Notice how we can still apply this method to the exceptional cases, for example while $d = 3, n = 4$ $s = 7$ is exceptional, to prove $\mathrm{AH}_4(3)$ for all other choices of s it suffices to consider $s \in \{6, 8\}$.

The next simplification is a very important one, and it relies on the following *semi-continuity* property of the Hilbert function. Let $\Bbbk(\underline{z})$ be the purely transcendental field extension of $\Bbbk$ obtained by adjoining the $s(n+1)$ indeterminates $\underline{z} = \{z_{ij}, i = 1, \ldots, s,\ j = 0, \ldots, n\}$. The set of *generic double points* is the set

$$\mathbb{X} = \{2\boldsymbol{P}_1, \ldots, 2\boldsymbol{P}_s\} \subseteq \mathbb{P}^n_{\Bbbk(\underline{z})}, \qquad \boldsymbol{P}_i = [z_{i0} : \cdots : z_{in}].$$

For any $\lambda \in \mathbb{A}_{\Bbbk}^{s(n+1)}$ with $\lambda_{i0}, \dots, \lambda_{in}$ not all zero for each i we get the *specialization*

$$\mathbb{X}(\lambda) = \{2\boldsymbol{P}_1(\lambda), \dots, 2\boldsymbol{P}_s(\lambda)\}, \qquad \boldsymbol{P}_i(\lambda) = [\lambda_{i0} : \dots : \lambda_{in}].$$

Theorem 3.8 (Semi-continuity of the Hilbert function). *Let $\mathbb{X}$ be the scheme of s generic double points and $\mathbb{Y}$ any scheme of s double points. Then for any $d \geq 0$,*

$$\dim_{\Bbbk} I(\mathbb{Y})_d \geq \dim_{\Bbbk} I(\mathbb{X})_d.$$

Furthermore, for general $\mathbb{Y}$ equality holds.

The proof of this statement (even slightly more general) can be found in the paper by Ha & Mantero [HM21, Theorem D.3], we sketch the important ideas.

Idea of proof. Consider the following set:

$$G_t = \left\{ \lambda \in \mathbb{A}_{\Bbbk}^{s(n+1)} \,\middle|\, \dim_{\Bbbk} I(\mathbb{X}(\lambda))_d \geq t \right\}.$$

Claim 1. The sets G_t are closet subsets of $\mathbb{A}_{\Bbbk}^{s(n+1)}$.

Proof of claim. One may proceed as follows: Let $\underline{c} = \{ c_{\boldsymbol{\alpha}} \mid |\boldsymbol{\alpha}| = d \}$ be a set of indeterminates and let

$$f = \sum_{|\boldsymbol{\alpha}|=d} c_{\boldsymbol{\alpha}} X^{\boldsymbol{\alpha}} \in S[\underline{c}]$$

be the generic polynomial with coefficients $\underline{c}$. Consider the matrices

$$D = \left[\frac{\partial X^{\boldsymbol{\alpha}}}{\partial X_j}\right]_{j=0,\dots,n,\,|\boldsymbol{\alpha}|=d}, \qquad B(\lambda) = \begin{pmatrix} D(\boldsymbol{P}_1(\lambda)) \\ \vdots \\ D(\boldsymbol{P}_s(\lambda)) \end{pmatrix}$$

where $D(\boldsymbol{P}_i)$ is obtained from D by plugging in the values of the points into the variables. Then the form f is in $I(\mathbb{X}(\lambda))_d$ if and only its coefficients $c_{\boldsymbol{\alpha}}$ satisfy

$$B(\lambda) \cdot (c_{(d,\dots,0)}, \dots, c_{\boldsymbol{\alpha}}, \dots, c_{(0,\dots,d)})^{\mathsf{T}} = 0$$

This is a linear system of equations, so $\lambda \in G_t$ if and only the kernel of $B(\lambda)$ has dimension at least t. This is equivalent to some condition on the rank of $B(\lambda)$, which is a closed condition. $\square_{\text{claim}}$

Claim 2. Let $t_0 := \dim_{\Bbbk} I(\mathbb{X})_d$, then B_{t_0} contains a dense open subset.

Proof of claim. Let $f_1, \dots, f_{t_0} \in I(\mathbb{X})_d \subseteq \Bbbk(\underline{z})[X_0, \dots, X_n]$ be a $\Bbbk(\underline{z})$-basis, WLOG $f_i \in \Bbbk[\underline{z}][\underline{X}]$. Let M be the matrix indexed by $i = 1, \dots, t_0$ and $|\boldsymbol{\alpha}| = d$ whose $(i, \boldsymbol{\alpha})$-th entry is the coefficient of X^α in f_i. Since they form a basis, M has maximal rank t_0, so at least one minor does not vanish identically. So there exists a dense open set $U \subseteq \mathbb{A}^{s(n+1)}$ of points $\lambda \in U$ such that this minor doesn't vanish in the specialization $M(\lambda)$. In particular the specializations of the f_i are linearly independent forms in $I(\mathbb{X}(\lambda))_d$. $\square_{\text{claim}}$

These two claims give the proposition, because B_{t_0} is a closed set containing a dense open, hence all schemes of double points satisfy

$$\dim_{\Bbbk} I(\mathbb{Y})_d \geq t_0 = \dim_{\Bbbk} I(\mathbb{X})_d.$$

Also, the dense open set U proves that we have equality for general $\mathbb{Y}$. ■

This has the following surprising consequence:

Corollary 3.9. *Fix n, d, s. The following assertions are equivalent:*

(i) *Some specific collection of s double points is* $\mathrm{AH}_n(d)$.
(ii) All *general collections of s double points are* $\mathrm{AH}_n(d)$.

Proof. Indeed, if a specific collection of double points $\mathbb{Y}$ is $\mathrm{AH}_n(d)$, then

$$\max\{\tbinom{n+d}{d} - s(n+1), 0\} = \dim_{\Bbbk} I(\mathbb{Y})_d \leq \dim_{\Bbbk} I(\mathbb{X})_d \leq \max\{\tbinom{n+d}{d} - s(n+1), 0\},$$

so we have equality. But the theorem shows that a general collection of double points has the same Hilbert function value at d, so it must be also $\mathrm{AH}_n(d)$! ■

This leads to the following algorithmic approach to the theorem.

Algorithm 1 Randomized algorithm for proving cases of the Alexander-Hirschowitz theorem

Require: Integers $n, d, s \geq 1$, Bound on coordinates of points $range$.
Ensure: $\mathbb{X}$ = set of points such that $2\mathbb{X}$ is $\mathrm{AH}_n(d)$.
for $i = 1, \ldots, s$ **do**
  Sample random integers $a_0, \ldots, a_n \in \{-range, \ldots, range\}$.
  $P_i \leftarrow [a_0 : \cdots : a_n]$ //One should ensure that not all $a_i = 0$.
  $I_i \leftarrow \mathfrak{m}_{P_i}$
end for
$I \leftarrow I_1^2 \cap \cdots \cap I_s^2$
$expdim \leftarrow \max\{\tbinom{n+d}{d} - s(n+1), 0\}$
if $\dim_{\Bbbk} I_d = expdim$ **then**
  $\mathbb{X} \leftarrow \{P_1, \ldots, P_s\}$
else
  return "Could not verify $\mathrm{AH}_n(d)$ on this $\mathbb{X}$."
end if

By Corollary 3.9 if this algorithm returns successfully for any random collection of points, then we know that any general such collection of points is $\mathrm{AH}_n(d)$! A sample implementation in Sage [The21] can be found in appendix A.

Example 3.10. While this algorithmic approach cannot resolve the exceptional cases, it can give upper bounds on the defect. For example, the following randomly sampled set of 7 points show that the defect is at most 1:

$$[1:0:1:0:0],\quad [1:1:0:1:1],\quad [1:1:1:1:1],\quad [1:0:1:-1:1],$$
$$[-1:-1:0:1:-1],\quad [0:1:1:0:-1],\quad [1:0:0:1:1].$$

3.3. Terracini's second lemma and the case $n = 2$

Terracini's first lemma 1.29 gives us a method to determine the dimension of the secant variety. Terracini's second lemma extends on this to give a *necessary* criterion for a collection of points to *not* impose independent conditions.

Theorem 3.11. *Assume that a general collection $\mathbb{X} = \{2P_1, \dots, 2P_s\}$ of double points does* not *imposing independent conditions on $\mathcal{O}_{\mathbb{P}^n}(d)$. Let $x_i = \nu_d(P_i) \in V^{d,n}$ be their images.*

(i) *There exists a positive-dimensional closed subvariety $Y \subseteq V^{d,n}$ through $x_1, \dots, x_s$ such that for all $y \in Y$ we have*
$$T_y V^{d,n} \subseteq \left\langle T_{x_1} V^{d,n}, \dots, T_{x_s} V^{d,n} \right\rangle_{\mathbb{P}}.$$

(ii) *Any hypersurface $\mathcal{V}(f) \subseteq \mathbb{P}^n$ of degree d singular at $P_1, \dots, P_n$ is singular at each point of $\nu_d^{-1}(Y)$.*

In the proof we use the *abstract secant variety* of a projective non-degenerate variety $X \subseteq \mathbb{P}^N$

$$\sigma^s X = \overline{\{\, (z, x_1, \dots, x_s) \mid z \in \langle x_1, \dots, x_s \rangle_{\mathbb{P}} \,\}} \subseteq \mathbb{P}^N \times X \times \cdots \times X.$$

This variety has dimension $s \cdot \dim X + \min\{s-1, N\}$: The last s factors contribute $s \cdot \dim X$ dimensions, and the general collection of s points on X will span an $(s-1)$-dimensional projective subspace or the whole space [1]. Notice that if $\pi_0 \colon \mathbb{P}^N \times \prod_i X \to \mathbb{P}^n$ is the projection onto the first factor, then $\pi_0(\sigma^s X) = \sigma_s X$.

Proof. Let $z \in \langle x_1, \dots, x_s \rangle_{\mathbb{P}}$ be such that $T_z \sigma_s V^{d,n} = \left\langle T_{x_1} V^{d,n}, \dots, T_{x_s} V^{d,n} \right\rangle_{\mathbb{P}}$ (this is the only "generality" assumption we need in this proof). Let $\Sigma_z := \pi_0^{-1}(z) \subseteq \sigma^s V^{d,n}$, we claim that $Y := \pi_1(\Sigma_z) \subseteq V^{d,n}$ is a subvariety with the desired properties.

- Notice first that Σ_z is invariant under permutations by $\mathfrak{S}_s$ of component $1, \dots, s$.

[1] This is because "being linearly dependent" is a closed property and by non-degeneracy there has to be *some* set of points spanning a $\min\{s-1, N\}$-dimensional subspace.

- By assumption $\dim \sigma^s V^{d,n} > \dim \sigma_s V^{d,n}$, so $\dim \Sigma_z \geq 1$. Clearly the projection of Σ_z onto $V^{d,n} \times \cdots \times V^{d,n}$ is also positive-dimensional. As this set is invariant under permutations, its projection onto the first (or any) factor cannot be finite, so $\dim Y > 0$.
- By the choice of z we have $(z, x_1, \ldots, x_s) \in \Sigma_z = \pi_0^{-1}(z)$, by permutation-invariance we have $x_1, \ldots, x_s \in Y$.
- Let $y_1 \in Y$, be the image of $(y, y_2, \ldots, y_s, z) \in \Sigma_z$ under π_1, then

$$T_{y_1} V^{d,n} \subseteq \left\langle T_{y_1} V^{d,n}, \ldots, T_{y_s} V^{d,n} \right\rangle_{\mathbb{P}} \subseteq T_z \sigma_s V^{d,n} = \left\langle T_{x_1} V^{d,n}, \ldots, T_{x_s} V^{d,n} \right\rangle_{\mathbb{P}}.$$

This establishes (i). For (ii) let $\mathcal{V}(f) \subseteq \mathbb{P}^n$ be a hypersurface singular at $P_1, \ldots, P_s$, then

$$f \in (\mathfrak{m}_{P_1}^2 \cap \cdots \cap \mathfrak{m}_{P_s}^2)_d \overset{1.16}{=} \left\langle T_{P_1} V^{d,n}, \ldots, T_{P_s} V^{d,n} \right\rangle_{\mathbb{P}}^{\perp} \subseteq (T_y V^{d,n})^{\perp} \overset{1.16}{=} (\mathfrak{m}_{\nu_d^{-1}(y)}^2)_d.$$

for any $y \in Y$, i. e. $\mathcal{V}(f)$ is singular along $\nu_d^{-1}(Y)$. ■

Before applying Terracini's second lemma we recall a basic version of *Bézout's theorem*.

Theorem 3.12. *Let $C, C' \subseteq \mathbb{P}^2$ be plane curves of degree d, d', then exactly one of the following is true:*

(i) *C and C' share an irreducible component;*
(ii) *$C \cap C'$ is a 0-dimensional scheme of length $C.C' = d \cdot d'$, in particular $\#(C \cap C') \leq dd'$.*

Theorem 3.13 (The case $n = 2$). *A general collection of s double points $\mathbf{X}_s \subseteq \mathbb{P}^2$ imposes independent conditions on degree d plane curves except in the following two cases:*

- *$d = 2$, $s = 2$;*
- *$d = 4$, $s = 5$.*

Proof. We may always assume $s \geq 2$. We check $d = 1, \ldots, 4$ first.

(1) Lines are never singular, and the expected dimension of $I_{\mathbf{X}}(1)$ is always 0.
(2) Any singular quadratic curve C is necessarily a union of two lines, and if $\#\operatorname{Sing}(C) \geq 2$, then C must be the double-line through these two points. So $\operatorname{codim} I_{\mathbf{X}_2}(2) = 6 - 1$, one less than expected. For $s \geq 3$ the expected dimension (0) is correct.
(3) For $s = 3$, let ℓ_1, ℓ_2, ℓ_3 be the lines through the three points. Any cubic singular at these points must contain each of the ℓ_i (as $3 = C.\ell_i < 2 \cdot 2$, using Bézout), so it must be *the* cubic $\ell_1 \cup \ell_2 \cup \ell_3$. So the expected dimension of 1 is correct, and hence this covers the cases $s = 1, 2, 3$; for 4 the expectation that there are no cubics is also verified.
(4) For $s = 4$, $\operatorname{expdim} I_{\mathbf{X}_4}(4) = \binom{2+4}{4} - 3 \cdot 3 = 3$. We first show that any quartic C with 4 singularities contains a (smooth) quadric Q through these points. Indeed, take a fifth point on C, let Q be a quadric trough these points, then $8 = C.Q < 2 \cdot 4 + 1$ shows that by Bézout C must contain Q. Thus, requiring the quartic to contain two additional general (single) points forces $C = Q \cup Q'$, each unique (!) quadric containing the four singular

points and one of the additional points. This shows $\dim I_{\mathbb{X}_3}(4) - 2 = 1$ as desired, this implies the remaining cases $s = 1, 2, 3, 4$ (Lemma 3.7).

For $s = 5$ we expect $\dim I_{\mathbb{X}_5}(4) = 5$, but as the previous case there actually *does* exists a unique quartic singular in 5 general points, namely the double quadric through these points.

Now let $d \geq 5$ and assume that a general collection of s double points does *not* impose independent conditions on plane curves of degree d. Let C be a plane curve through $\mathbb{X}_s$, by Terracini's second lemma C must contain an irreducible plane curve Y of degree l in its singular locus.

- This means $2Y \subseteq C$, so $2l \leq d$.
- The family of curves through s general points is nonempty only if $\dim \Bbbk[X_0, X_1, X_2]_l > s$.
- The largest s' such that $s'(n+1) = \min\{s'(2+1), \binom{2+d}{2}\}$ is $s' = \lfloor \frac{1}{3}\binom{d+2}{2} \rfloor$. Suppose $s < s'$, then $s+1$ would also be an exceptional case, so we may assume $s \geq s'$.

Summarizing, we have the following three inequalities

$$\underbrace{2l \leq d}_{2Y \subseteq C}, \qquad \underbrace{s \leq \binom{l+2}{2} - 1 = \frac{l(l+3)}{2}}_{C \text{ contains } s \text{ general points}}, \qquad \left\lfloor \frac{1}{3}\binom{d+2}{2} \right\rfloor \leq s.$$

Combining these yields

$$\left\lfloor \frac{(d+2)(d+1)}{6} \right\rfloor \leq s \leq \frac{d}{4}\left(\frac{d}{2} + 3\right)$$

which we can evaluate to admit integer solutions $d = 0, 2, 3, 4, 6$. So the only interesting case is $d = 6$, which gives $\lfloor \frac{56}{6} \rfloor \leq s \leq 9$, so $s = 9$ and $l = 3$. So C is necessarily a double cubic, more specifically *the unique* cubic through 9 general points. This incidentally matches the expected dimension $\binom{2+6}{2} - 9 \cdot 3 = 1$, so this case is also not exceptional. ■

3.4. The exceptional cases

Theorem 3.14. *For $s \leq n+1$ we have*

$$\dim \sigma_s V^{d,n} = \dim \{\, S \in \mathrm{Sym}(n+1, \Bbbk) \mid \mathrm{rank}(S) \leq s \,\} - 1 = sn + s - 1 - \binom{s}{2}.$$

In particular $V^{2,n}$ is s-defective for $s = 2, \ldots, n$ with $\delta_s = \binom{s}{2}$.

Proof. The first equality is immediate from the discussion in example 1.3. There we also discussed the group action

$$G = \mathrm{GL}(n, \Bbbk) \circlearrowleft \mathrm{Sym}(n, \Bbbk), \qquad M \triangleright S := MSM^{\mathsf{T}}.$$

Under this action the symmetric matrices of given rank form a single orbit

$$\{ S \in \mathrm{Sym}(n, \Bbbk) \mid \mathrm{rank}(S) = s \} = G \triangleright \underbrace{\left[\begin{array}{c|c} \mathbb{1}_s & 0 \\ \hline 0 & 0 \end{array}\right]}_{=: D_s} = f(G),$$

where $f\colon G \longrightarrow \mathrm{Sym}(n, \Bbbk)$, $f(M) = M \triangleright D_s$. The fibre of $S = f(N)$ is

$$\begin{aligned} f^{-1}(S) &= \left\{ M \in G \;\middle|\; M D_s M^\top = S = N D_s N^\top \right\} \\ &= \left\{ M \in G \;\middle|\; (N^{-1}M) D_s (N^{-1}M)^\top = D_s \right\} = N \cdot f^{-1}(D_s) \cong \mathrm{Stab}_G(D_s). \end{aligned}$$

Since f is dominant, the dimension of a general fibre is $\dim G - \dim f(G)$, but all fibres are isomorphic, so it suffices to calculate the dimension of $\mathrm{Stab}_G(D_s)$. We use block matrix notation to calculate

$$\begin{aligned} M D_s M^\perp &= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \left[\begin{array}{c|c} \mathbb{1}_s & 0 \\ \hline 0 & 0 \end{array}\right] \left[\begin{array}{c|c} A^\top & C^\top \\ \hline B^\top & D^\top \end{array}\right] \\ &= \left[\begin{array}{c|c} AA^\top & AC^\top \\ \hline CA^\top & CC^\top \end{array}\right] \overset{!}{=} \left[\begin{array}{c|c} \mathbb{1}_s & 0 \\ \hline 0 & 0 \end{array}\right] = D_s \end{aligned}$$

This implies $A \in \mathrm{O}(s, \Bbbk)$ and $C = 0$ (as $A^\top$ is invertible). As $M \in \mathrm{GL}(n, \Bbbk)$, this in turn implies $D \in \mathrm{GL}(n-s, \Bbbk)$ (and $B \in \mathrm{Mat}(s, n-s, \Bbbk)$, but we knew this already). Conversely, any matrix of this form fixes D_s, so

$$\mathrm{Stab}_G(D_s) = \left[\begin{array}{c|c} \mathrm{O}(s, \Bbbk) & \mathrm{Mat}(s, n-s, \Bbbk) \\ \hline 0 & \mathrm{GL}(n-s, \Bbbk) \end{array}\right], \quad \dim \mathrm{Stab}_G(D_s) = \underbrace{\binom{s}{2}}_{\dim \mathrm{O}(s,\Bbbk)} + s(n-s) + (n-s)^2.$$

With this knowledge we can compute

$$\dim f(G) = \dim G - \dim \mathrm{Stab}_G(D_s) = n^2 - \binom{s}{2} - n(n-s) = sn - \binom{s}{2}.$$

Translating this back to secant varieties gives

$$\dim \sigma_s V^{2,n} = \dim \{ S \in \mathrm{Sym}(n+1, \Bbbk) \mid \mathrm{rank}(S) = s \} - 1 = s(n+1) - 1 - \binom{s}{2}. \qquad \blacksquare$$

Lemma 3.15 (The exceptional quartic cases). *Let $d = 4$,*

$$(n, s) \in \{(2,5), (3,9), (4,14)\},$$

then $\mathbb{X}_s$ does not impose independent conditions of on quartic hypersurfaces in $\mathbb{P}^n$. More specifically there exists a quartic hypersurface in $\mathbb{P}^n$ through s general double points, even though the expected dimension of $I_{\mathbb{X}_s}(4)$ is zero.

Proof. The values in question are

n	s	$\binom{n+4}{4}$	$\binom{n+4}{4} - s \cdot (n+1)$
2	5	15	0
3	9	35	−1
4	14	70	0

so we expect there to be no quartics through $\mathbb{X}_s$ at all.

But since $s = \binom{n+2}{2} - 1 < \dim_{\Bbbk} \Bbbk[X_0, \dots, X_n]_2$, there exists a unique quadratic hypersurface Q through s general points. So this double quadric $2Q$ is a singular quartic through $\mathbb{X}_s$ double points. One may use the randomized algorithm to prove that the defect is 1. ∎

Lemma 3.16 (The exceptional cubic case). *$\sigma_7 V^{3,4}$ is defective, i. e. there exists a cubic hypersurface in $\mathbb{P}^4$ through 7 general double points even though the expected dimension of $I_{\mathbb{X}_7}(3)$ is zero.*

Recall that a *rational normal curve* $C_d \subseteq \mathbb{P}^d$ is any curve projectively equivalent to $V^{d,1} \subseteq \mathbb{P}^d$.[2] We use the following basic facts, which can be found for example in the book by Harris [Har92]:

- For any $n+3$ points in $\mathbb{P}^n$ in general linear position (i. e. no $n+1$ lie in a hyperplane) there exists a rational normal curve through these points [Har92, Theorem 1.18].
- The rational normal curve C_d corresponding to the image of $[t_0 : t_1] \mapsto [t_0^d : t_0^{d-1}t_1 : \dots : t_1^d]$ can be written as the set of points $[x_0 : \dots : x_d] \in \mathbb{P}^n$ such that [Har92, Example 1.16]

$$\operatorname{rank} \begin{bmatrix} x_0 & x_1 & \dots & x_k \\ x_1 & x_2 & \dots & x_{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d-k} & x_{d-k+1} & \dots & x_d \end{bmatrix} = 1$$

Proof of lemma 3.16. By the remark about rational normal curves there exists a rational normal curve C through $4+3$ general points. After a change of coordinates we may assume that this is given as

$$C = \left\{ [x_0 : \dots : x_4] \in \mathbb{P}^4 \;\middle|\; \operatorname{rank} \begin{bmatrix} x_0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \end{bmatrix} \leq 1 \right\}.$$

[2] More intrinsically, after some identification $\mathbb{P}^d \cong \mathbb{P}(\Bbbk[x_0, x_1]_d)$, C is the image of $\mathbb{P}^1$ under $x \mapsto [f_0(d), \dots, f_d(x)]$ for some basis of $\Bbbk[x_0, x_1]_d$.

Points on secants to $\mathcal{C}$ hence correspond to linear combinations of rank 1 matrices of the previous shape, which have certainly rank ≤ 2. Hence we get the inclusion (which is actually an equality, but we dont't need this)

$$\sigma_2 C \subseteq X := \left\{ [x_0 : \cdots : x_4] \in \mathbb{P}^4 \;\middle|\; \operatorname{rank} \begin{bmatrix} x_0 & x_1 & x_2 \\ x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \end{bmatrix} \leq 2 \right\} = \mathcal{V}\left(\det \begin{bmatrix} X_0 & X_1 & X_2 \\ X_1 & X_2 & X_3 \\ X_2 & X_3 & X_4 \end{bmatrix} \right).$$

So $X \subseteq \mathbb{P}^4$ is a degree 3 hypersurface containing the 7 points, and it is an elementary (but tedious) calculation that $\mathcal{C} \subseteq \operatorname{Sing}(X)$. This shows that $I(\mathbb{X})_3 \neq 0$, despite the expected dimension being 0. ■

This concludes all exceptional cases from the list.

3.5. Traces, residues and Terracini's inductive argument

In this section we provide a way to reduce the property $\mathrm{AH}_n(d)$ of a scheme of double points to that of a scheme in a smaller projective space or with smaller degree. This will enable us to use induction in the proof of the Alexander-Hirschowitz theorem.

Definition 3.17. Let $X \subseteq \mathbb{P}^n$ be a scheme, $\mathcal{I}_X$ its corresponding ideal sheaf and let $H \subseteq \mathbb{P}^n$ be a hyperplane.

(i) The scheme $\mathrm{Tr}_H(X) := X \cap H$ (formally: the fibre of X under the inclusion $H \hookrightarrow \mathbb{P}^n$) is called the *trace* of X with respect to H.
(ii) The scheme $\mathrm{Res}_H(X)$ defined by the ideal sheaf $\mathcal{I}_X : \mathcal{I}_H$ is called the *residue* of X with respect to H.

This has the following interpretation:

Lemma 3.18 (Trace and residue). *Let $\mathbb{X} = \{2P_1, \ldots, 2P_s\}$ be a collection of double points, of which the first $u \leq s$ are supported on a hyperplane $\iota \colon H \hookrightarrow \mathbb{P}^n$.*

(i) $\mathrm{Tr}_H(\mathbb{X})$ *is the scheme of u double points $\{2P_1, \ldots, 2P_u\} \subseteq H \cong \mathbb{P}^{n-1}$.*
(ii) $\mathrm{Res}_H(\mathbb{X})$ *is the scheme $\{P_1, \ldots, P_u, 2P_{u+1}, \ldots, 2P_s\} \subseteq \mathbb{P}^n$.*
(iii) *We have the short exact sequence of ideal sheaves*

$$0 \longrightarrow \mathcal{I}_{\mathrm{Res}_H(\mathbb{X})}(-1) \longrightarrow \mathcal{I}_{\mathbb{X}} \longrightarrow \iota_* \mathcal{I}_{\mathrm{Tr}_H(\mathbb{X})} \longrightarrow 0$$

Proof. After a change of coordinates we may assume that $H = \{X_0 = 0\}$.

(i) We distinguish two cases.

- If $P \notin H$, then $\mathcal{V}(\mathfrak{m}_P^2 + (X_0)) = \emptyset$.

- If $P \in H$ then without loss of generality let $P = [0 : \cdots : 0 : 1]$, then

$$((X_0, \ldots, X_{n-1})^2 + (X_0))/(X_0) = \underbrace{(\overline{X_1}, \ldots, \overline{X_{n-1}})^2}_{= \mathfrak{m}_P^2 \text{ (in } \mathbb{P}^{n-1})} \subseteq \Bbbk[\overline{X_1}, \ldots, \overline{X_n}].$$

Hence

$$I(\mathrm{Tr}_H(\mathbb{X})) = I(\{2P_1, \ldots, 2P_u\}) \subseteq \Bbbk[\overline{X_1}, \ldots, \overline{X_n}]$$

which establishes the statement.

(ii) • Let $P \notin H$, WLOG $P = [1 : 0 : \cdots : 0]$, then $f \in \mathfrak{m}_P^2 : (X_0)$ means $Xf = 0$ and all its derivatives vanishes at P. This implies the same for f, so $f \in \mathfrak{m}_P^2$, and in particular $\mathfrak{m}_P^2 : (X_0) = \mathfrak{m}_P^2$.

- Let $P \in H$, i. e. $X_0 \in \mathfrak{m}_P$, then $\mathfrak{m}_P \subseteq \mathfrak{m}_P^2 : (X_0)$. If $f \in \mathfrak{m}_P^2 : (X_0)$, then $\frac{\partial f}{\partial X_0}(P) = f(P) = 0$, so $f \in \mathfrak{m}_P$ and we have equality.

This establishes the desired equality

$$I(\mathbb{X}) : (X_0) = (\mathfrak{m}_{P_1}^2 : (X_0)) \cap \cdots \cap (\mathfrak{m}_{P_s}^2 : (X_0)) = \mathfrak{m}_{P_1} \cap \cdots \cap \mathfrak{m}_{P_u} \cap \mathfrak{m}_{P_{u+1}}^2 \cap \cdots \cap \mathfrak{m}_{P_s}^2.$$

(iii) It suffices to check exactness of the following sequence of (homogeneous) ideals

$$0 \longrightarrow I(\mathbb{X}) : (X_0) \xrightarrow{\cdot X_0} I(\mathbb{X}) \longrightarrow (I(\mathbb{X}) + (X_0))/(X_0) \longrightarrow 0,$$

but this is straightforward from the definitions. ■

Twisting with $\mathcal{O}(d)$ and taking global sections of the exact sequence from the previous lemma yields

Corollary 3.19 (Castelnuovo exact sequence). *For $\mathbb{X}$ and H as before we have the left-exact sequence*

$$0 \longrightarrow I(\mathrm{Res}_H(\mathbb{X}))_{d-1} \longrightarrow I(\mathbb{X})_d \longrightarrow I(\mathrm{Tr}_H(\mathbb{X}))_d.$$

In particular we have the Castelnuovo inequality

$$\dim_{\Bbbk} I(\mathbb{X})_d \leq \dim_{\Bbbk} I(\mathrm{Res}_H(\mathbb{X}))_{d-1} + \dim_{\Bbbk} I(\mathrm{Tr}_H(\mathbb{X}))_d.$$

This estimate allows for a crucial reduction step in the proof of the Alexander-Hirschowitz theorem.

Theorem 3.20 (Terracini's inductive argument). *Let $\mathbb{X} \subseteq \mathbb{P}^n$ be a union of s double points, u of them contained in a hyperplane H. Assume*

- $\mathrm{Tr}_H(\mathbb{X})$ *imposes independent condition on $\mathcal{O}_{\mathbb{P}^{n-1}}(d)$;*
- $\mathrm{Res}_H(\mathbb{X})$ *imposes independent conditions on $\mathcal{O}_{\mathbb{P}^n}(d-1)$;*

- *One of the following two pairs of inequalities hold true*
  (a) $un \geq \binom{n-1+d}{n-1}$ *and* $s(n+1) - un \geq \binom{d+n-1}{n}$,
  (b) $un \leq \binom{n-1+d}{n-1}$ *and* $s(n+1) - un \leq \binom{d+n-1}{n}$.

Then $\mathbf{X}$ *does impose independent conditions on* $\mathcal{O}_{\mathbb{P}^n}(d)$.

Remark. The two sets of inequalities can be expressed in the single statement

"un lies between the numbers $\binom{n-1+d}{n-1}$ and $s(n+1) - \binom{n-1+d}{n}$."

Proof. We know that $\dim_{\Bbbk} I(\mathbf{X})_d \geq \max\{\binom{n+d}{n} - s(n+1), 0\}$, our goal is to show equality. The first two assumptions ensure (notice that $\operatorname{Res}_H(\mathbf{X})$ is a mixture of u single and $s-u$ double points!)

$$\begin{aligned}\dim_{\Bbbk} I(\operatorname{Tr}_H \mathbf{X})_d &= \max\left\{\binom{n-1+d}{n-1} - un, 0\right\},\\ \dim_{\Bbbk} I(\operatorname{Res}_H \mathbf{X})_{d-1} &= \max\left\{\binom{n+d-1}{n} - (s-u)(n+1) - u, 0\right\}.\end{aligned}$$

We can apply these identities to the Castelnuovo inequality in the two cases.

(a) Here the two maxima are 0, so Castelnuovo yields $I(\mathbf{X})_d = 0$.
(b) Here the maxima are attained at first term, so using the addition theorem for the binomial coefficients we get

$$\dim_{\Bbbk} I(\mathbf{X})_d \leq \binom{n-1+d}{n-1} - un + \binom{n-1+d}{n} - (s-u)(n+1) - u = \binom{n+d}{n} - s(n+1).$$

Thus, in both cases $\dim_{\Bbbk} I(\mathbf{X})_d \leq \max\{\binom{n+d}{n} - s(n+1), 0\}$ as desired. ■

Example 3.21. We consider the case $n = 3$ and try to use Theorem 3.20 for induction on $d \geq 3$. By Lemma 3.7 we need to consider the values

$$\frac{1}{3+1}\binom{3+d}{d} = \frac{(d+3)(d+2)(d+1)}{4!} \quad \rightsquigarrow$$

d	3	4	5	6
$\overline{s}$	5	9	14	21
$\underline{s}$	5	8	14	21

- The base case $d = 3$ can be explicitly evaluated on the "star configuration"

$$P_0 = [1:0:0:0], \quad P_1 = [0:1:0:0], \quad P_2 = [0:0:1:0], \quad P_3 = [0:0:0:1],$$
$$P_4 = [1:1:1:1].$$

  One immediately verifies that $I(\{2P_0, \ldots, 2P_4\})_3 = 0$, which is expected.

- Next is $d = 4$, $\underline{s} = 8$. Consider a union $\mathbb{X}$ of 8 double points, in order to satisfy the hypotheses of Theorem 3.20 we need to choose $0 \le u \le 8$ such that

$$12 = 8(3+1) - \binom{4+3-1}{3} \le 3u \le \binom{3-1+4}{3-1} = 15 \quad \Longrightarrow \quad u \in \{4, 5\}.$$

  We take[3] $u = 4$ and specialize 4 of the 8 double points on a hyperplane $H \subseteq \mathbb{P}^3$ such that they are general in $H \cong \mathbb{P}^2$, then this scheme of double points is $\mathrm{AH}_2(4)$ by theorem 3.13.
  Arrange the remaining 4 points $\mathbb{Y}$ in $\mathbb{P}^3$ in general position, then $\mathbb{Y}$ is $\mathrm{AH}_3(3)$ by the induction hypothesis, i.e. $\dim_{\mathbb{k}} I(\mathbb{Y})_3 = 4$. If the single points on H are in general position, then forcing the cubic hypersurfaces to pass through these points reduces $I(\mathrm{Res}_H \mathbb{X})_3$ to 0. Indeed, otherwise there would be a cubic which is a union of a quartic and H, singular along four general points, which is impossible (because the singular locus of this quartic would be a linear subspace containing 4 points in $\mathbb{P}^3$ in general position).
  Thus we can apply Theorem 3.20 and see that a general scheme of 8 (and hence also $s \le 8$) double points is $\mathrm{AH}_3(4)$.
- For $d = 5$, $s = 14$, the inequalities force $u = 7$. A general scheme of 7 double points in the plane is $\mathrm{AH}_2(5)$, and the scheme $\mathbb{Y}$ of the remaining 7 double points in $\mathbb{P}^3$ is also $\mathrm{AH}_3(4)$ by induction. Since there is no cubic through 7 general double points, the dimension $\dim_{\mathbb{k}} I(\mathbb{Y})_4 = 7$ can be reduced to 0 by imposing the additional condition of passing through the 7 single points on H.
  Again we can apply the Theorem and see that a general scheme of 14 (and hence any number!) double points is $\mathrm{AH}_3(5)$.
- The case $d = 6$, $s = 21$ is where we first run into trouble. The inequalities give $28 \le 3u \le 28$, so Theorem 3.20 can *not* be applied. The case $s = 20$ admits the choices $u \in \{8, 9\}$, and it is possible to carry out the argument in this case.

We see that it is easier to consider $s < \underline{s}$, since then the interval between $\binom{n-1+d}{n-1}$ and $s(n+1) - \binom{n-1+d}{n}$ will always contain some un, $u \in \mathbb{N}_0$. To adress the "top" case, i.e. $\mathrm{AH}_3(6)$ for 21 general double points, is *much* more difficult, and requires a different, more refined strategy.

3.6. The cubic case $d = 3$

In the case of cubics we cannot use Terracini's inductive argument, since virtually all cases for $d = 2$ are defective.

[3] Notice that $u = 5$ would *not* work here, because no general scheme of 5 double points in the plane is $\mathrm{AH}_2(4)$ (Theorem 3.13).

We use the following notation:

$$s_n := \underline{s} = \begin{cases} \frac{(n+3)(n+2)}{6} & \text{if } n \equiv 0, 1 \bmod 3, \\ \frac{(n+4)(n+1)}{6} & \text{if } n \equiv 2 \bmod 3, \end{cases}$$

$$\ell_n := \binom{n+3}{3} - s_n(n+1)$$

The cases $n = 2$ (Theorem 3.13) and $n = 3$ (previous example) have been discussed. The proof for the general case is split into two cases depending on $n \bmod 3$. We need to show that a set of s_n general double points in $\mathbb{P}^n$ is $\mathrm{AH}_n(3)$. In order to apply induction *and* to stay in this particular case, it is necessary to do steps of size 3, i. e. reduce the statement from $\mathbb{P}^n$ to a linear subspace of codimension 3. The key tool to this is

Theorem 3.22 ([BO08, Proposition 5.4]). *Let $n \geq 3$, $n \neq 4$, and let $L \subseteq \mathbb{P}^n$ be a subspace of codimension 3.*

(i) *If $n \equiv 0, 1 \bmod 3$, then there are no cubic hypersurfaces in $\mathbb{P}^3$ which*
  - *contain L,*
  - *contain $s_{n-3} = \frac{n(n-1)}{6}$ general double points supported in L and*
  - *contain $s_n - s_{n-3} = n + 1$ general double points in $\mathbb{P}^n$.*

(ii) *If $n \equiv 2 \bmod 3$, then there are no cubic hypersurfaces in $\mathbb{P}^3$ which*
  - *contain L,*
  - *contain $s_{n-3} = \frac{(n+1)(n-2)}{6}$ general double points supported in L,*
  - *contain $s_n - s_{n-3} = n + 1$ general double points in $\mathbb{P}^n$ and*
  - *contain a general scheme η of length $\mathrm{len}(\eta) = \ell_n$ supported at a point in L such that and $\mathrm{len}(\eta \cap L) = \ell_n - 1$.*

Idea of proof. The main three steps are the following:

- Consider first three general codimension 3 hypersurfaces $L, M, N \subseteq \mathbb{P}^n$ ($n \geq 5$) with 3 general double points on each of them. The first claim is that no cubic hypersurface contains this constellation, which is proven by induction on n using the isomorphism

$$\underbrace{I(L \cup M \cup N)_3}_{\text{over } \mathbb{P}^n} \cong \underbrace{I((L \cup M \cup N) \cap H)_3}_{\text{over } H \cong \mathbb{P}^{n-1}}$$

  for a general hyperplane $H \subseteq \mathbb{P}^n$.[4]
- This fact is used to prove a similar statement about two linear subspaces: Let $L, M \subseteq \mathbb{P}^n$ be general codimension 3 hypersurfaces and consider $n - 2$ general double points

[4] The argument by Brambill & Ottavani contains an incorrect statement, fixed by Ha & Mantero [HM21, Proof of Claim 5.1.3].

on each of L, M, and additionally three general single points in $\mathbb{P}^n$, then no cubic hypersurface contains this scheme. This statement is proven using inducion on $n-3 \mapsto n$ by introducing a third codimension 3 space N and specializing the points in a clever way to the intersections.

- Finally, the statement in the theorem is proven by introducing a second codimension 3 subspace M and using induction on $n-3 \mapsto n$. In this (and also in the previous steps) the base cases of the induction have to be checked manually (using a computer). ■

Corollary 3.23 (The case $d = 3$). *A general collection of s double points in $\mathbb{P}^n$ is $\mathrm{AH}_n(3)$ except in the case $n = 4$ (Lemma 3.16). More specifically:*

(i) *If $n \equiv 0, 1 \bmod 3$, $n \neq 4$, then s_n general double points in $\mathbb{P}^n$ are $\mathrm{AH}_n(3)$,*
(ii) *If $n \equiv 2 \bmod 3$, then s_n general double points and a zero-dimensional scheme of length ℓ_n impose independent conditions on cubics.*

Proof. Notice that the second part implies the first part. Indeed, for $n \equiv 0, 1 \bmod 3$ $s_n = \underline{s} = \overline{s}$ is the only case which one needs to check. In the second case, the argument in Lemma 3.7(ii) shows that a general scheme of s_n double points is $\mathrm{AH}_n(3)$. Moreover, a general scheme of $\overline{s} = s_n + 1$ double points contains a (general) scheme of s_n double points and a scheme of the prescribed length (inside the $(n+1)$-th double point), so the argument from Lemma 3.7(i) applies.

We now turn to the proof of the two statements. Consider the exact sequence

$$0 \longrightarrow \mathcal{I}_L(3) \longrightarrow \mathcal{O}_{\mathbb{P}^n}(3) \longrightarrow \iota_* \mathcal{O}_L(3) \longrightarrow 0$$

Taking global sections and looking at the subspace of functions passing through the scheme at hand, we can apply induction on n (in steps of 3).

(i) The induction start is $n = 3$ (previous section) and $n = 7$ (has to be done manually). Assume the statement is true for $\mathbb{P}^{n-3}$.

Specialize $s_{n-3} = \frac{n(n-1)}{6}$ of the s_n double points on L in general position, and the remaining $n+1$ points in general position in $\mathbb{P}^n$. Then Theorem 3.22 shows that no cubics contain L and are singular at the given points, while the induction hypothesis shows that no cubics in L are singular on the $\frac{n(n-1)}{6}$ points in L. This yields the theorem for the whole configuration.

(ii) In this case the starting point is $n = 2$ (Theorem 3.13).

Specialize $s_{n-3} = \frac{(n+1)(n-2)}{6}$ double points on L and arrange the remaining $s_n - s_{n-3} = n+1$ double points in $\mathbb{P}^n$. Then specialize the scheme η of length ℓ_n on L such that $\eta \cap L$ has length $\ell_n - 1$. Then again Theorem 3.22 and the induction hypothesis can be applied to the exact sequence, concluding the proof. ■

3.7. La methode d'Horace differentielle

Now that all exceptional cases (from the table of Theorem 3.3) have been explained, and all cases for $n \le 2$ and $d \le 3$ have been discussed, we are ready to apply the crucial induction step. This clever refinement of Terracini's induction step 3.20 (also known as the "methode d'Horace") is called the "methode d'Horace differentielle", an explaination of the name is given in the paper by Bernardi et al. [Ber+18, Section 2.2.1].

Theorem 3.24 ([HM21, Theorem 2.9]). *Fix $n \ge 2$, $d \ge 4$ and $\underline{s} \le s \le \overline{s}$. Let $q \in \mathbb{Z}$, $\varepsilon \in \{0, \dots, n-1\}$ be defined via the following expression:*

$$nq + \varepsilon = s(n+1) - \binom{n+d-1}{n}.$$

Assume that

(i) *q general double points in $\mathbb{P}^{n-1}$ are $\mathrm{AH}_{n-1}(d)$,*
(ii) *$s - q$ general double points in $\mathbb{P}^n$ are $\mathrm{AH}_n(d-1)$ and*
(iii) *$s - q - \varepsilon$ general double points in $\mathbb{P}^n$ are $\mathrm{AH}_n(d-2)$,*

then s general double points in $\mathbb{P}^n$ are $\mathrm{AH}_n(d)$.

Using the cases $d \le 3$ and $n \le 2$, one can see that Theorem 3.24 can be used to prove all but finitely may cases for (n, d) in the Alexander-Hirschowitz theorem; the remaining cases can be checked manually [HM21, Proof of Theorem 2.10].

We give a rough outline of the proof. Recall that the value of the Hilbert function of a zero-dimensional scheme $\mathbb{X}$ is the length of $\mathbb{X}$ for $d \gg 0$. This motivates the following definition:

Definition 3.25. Let $\mathbb{X} \subseteq \mathbb{P}^n$ be a zero-dimensional scheme.

(i) $\mathbb{X}$ is called *multiplicity d-independent* if $\mathrm{HF}(\mathbb{X}, d) = \mathrm{len}(\mathbb{X})$.
(ii) If $V \subseteq \Bbbk[X_0, \dots, X_n]_d$ is a vector space, then the *Hilbert function with respect to V* is

$$\mathrm{HF}(\mathbb{X}, V) := \dim_{\Bbbk} V - \dim_{\Bbbk}(I(\mathbb{X}) \cap V).$$

(iii) $\mathbb{X}$ is called *multiplicity V-independent* if $\mathrm{HF}(\mathbb{X}, V) = \mathrm{len}(\mathbb{X})$.

Definition 3.26 ([HM21, Appendix E]). A zero-dimensional scheme $\mathbb{X}$ is said to be *curvilinear* if one of the following equivalent conditions is satisfied:

(i) Locally, $\mathbb{X}$ can be embedded into a smooth curve
(ii) For every $P \in \mathbb{X}$, $\dim T_P\mathbb{X} \le 1$
(iii) $\mathbb{X}$ is the (disjoint) union of schemes of the form $\operatorname{Spec} \Bbbk[t]/(t^l)$, $l \in \mathbb{N}_+$.

Curvilinear schemes are useful to us for the following two reasons: Firstly, they form a dense open subset of the Hilbert scheme $\mathrm{Hilb}_s(\mathbb{P}^n)$ of zero-dimensional projective schemes of length s [HM21, Proposition E.7]. The second reason is the following theorem [HM21, Lemma 2.7], allowing to reduce computations to curvilinear schemes.

Theorem 3.27 (Curvilinear Lemma). *Let $\mathbb{X} \subseteq \mathbb{P}^n$ be a zero-dimensional scheme contained in a finite union of double points. Let $V \subseteq \Bbbk[\underline{X}]_d$ be a vector space. Then the following are equivalent:*

(i) *$\mathbb{X}$ is multiplicity V-independent;*
(ii) *every curvilinear subscheme of $\mathbb{X}$ is multiplicity V-independent.*

Now we can outline the proof of Theorem 3.24.

Proof of Theorem 3.24 (outline). One proceeds in 4 steps.

Step 1. Fix a hyperplane $H \subseteq \mathbb{P}^n$. Choose

- a general collection 2Ψ of $s - q - \varepsilon$ double points in $\mathbb{P}^n \setminus H$,
- general collections 2Λ, 2Γ of q and ε double points in H.

Step 2. By hypothesis (ii), $2\Lambda \cup 2\Gamma$ is $\mathrm{AH}_n(d)$. Taking the trace of 2Γ with respect to H, one obtains that $2\Lambda \cup \mathrm{Tr}_H(2\Gamma)$ and $\Psi \cup 2\Lambda \cup \mathrm{Tr}_H(2\Gamma)$ have maximal Hilbert function in degree d, (that is: $\min\{\dim_{\Bbbk} S_d, \mathrm{len}(\dots)\}$; this is always an upper bound). From this point on the proof is split into two cases depending on whether $s = \overline{s}$ or $s = \underline{s} < \overline{s}$. Then one verifies that it suffices to prove that 2Γ is multiplicity $V := I(2\Lambda \cup 2\Psi)$-independent. For example, in the case $s = \underline{s}$ the goal of the theorem is to show that

$$\mathrm{HF}(2\Gamma \cup 2\Lambda \cup 2\Psi, d) = s(n+1).$$

One can show that $\mathrm{HF}(2\Lambda \cup 2\Psi, d) = (n+1)(s-\varepsilon)$, then if 2Γ were multiplicity V-independent, we get the desired result

$$\begin{aligned}\mathrm{HF}(2\Gamma \cup 2\Lambda \cup 2\Psi, d) &= \dim_{\Bbbk} S_d - \dim_{\Bbbk}\left(I(2\Gamma)_d \cap I(2\Lambda \cup 2\Psi)_d\right)\\ &= \mathrm{HF}(2\Lambda \cup 2\Psi, d) + \mathrm{len}(2\Gamma) = (n+1)(s-\varepsilon) + \varepsilon(n+1) = s(n+1).\end{aligned}$$

Step 3. In order to show that 2Γ is multiplicity $I(2\Lambda \cup 2\Psi)$-independent, a degeneration argument is used. For $t = (t_1, \dots, t_\varepsilon) \in \Bbbk^\varepsilon$ we take a flat family of general points $\Gamma_t = \{\gamma_{1,t_1}, \dots, \gamma_{\varepsilon,t_\varepsilon}\} \subseteq \mathbb{P}^n$ and a family of hyperplanes $\{H_{t_1}, \dots, H_{t_\varepsilon}\}$ such that

(1) $\gamma_{i,t_i} \in H_{t_i}$ for $i = 1, \dots, \varepsilon$
(2) $\gamma_{i,t_i} \notin H$ for any $t_i \neq 0$ and $i = 1, \dots, \varepsilon$
(3) For $t = 0$, $H_0 = H$ and $\gamma_{i,0} = \gamma_i$ (the points in Γ) for $i = 1, \dots, \varepsilon$.

So we have a parametrized family of (single) points on hyperplanes, which "converges" to $\Gamma \subseteq H$ for $t \to 0$. A stronger version of the semi-continuity of the Hilbert function [HM21, Theorem D.9] shows that in order to prove that 2Γ is multiplicity V-independent, it suffices to prove that $2\Gamma_t$ is multiplicity V-independent for *some* (!) t.

Step 4. One proceeds to argue by contradiction that no such t exists. By the Curvilinear Lemma 3.27, for each t there exists a curvilinear scheme $\Theta_t = \theta_1 \cup \cdots \cup \theta_\varepsilon \subseteq 2\Gamma_t$ supported in Γ_t which is also not multiplicity V-independent, i. e.

$$\mathrm{HF}(2\Lambda \cup 2\Psi \cup \Theta_t, d) < \mathrm{HF}(2\Lambda \cup 2\Psi, d) + \mathrm{len}(\Theta_t).$$

As the curvilinear schemes are dense in the Hilbert scheme, there exists a limit Θ_0 where $t \to 0$. Finally, one can use Θ_0, the semicontinuity of the Hilbert function and Castelnuovos inequality 3.19 to arrive at a contradiction. ■

Example 3.28. We return to the case $n = 3$, $d = 6$, $s = 21$ from example 3.21, where the easy induction step could not be applied. Here $q = 9$, $\varepsilon = 1$, and (according to the proof) we choose a scheme $\mathbb{X}$ consisting of

- a general collection 2Ψ of 11 double points outside H,
- a general collections 2Λ of 9 double points in H and a single double point 2γ in H.

The problem was that the point 2γ of length 4 could neither be placed in the trace nor in the residue of $\mathbb{X}$ with respect to H. The degeneration argument enables us to substitute 2γ by a curvilinear scheme θ, having length $\mathrm{len}(\mathrm{Tr}_H(\theta)) = 1$, so it fits in the trace.

Remark. Postinghel published a completely different proof of the Alexander-Hirschowitz theorem using degenerations [Pos12].

4

Applications

In this chapter we are going to look into some computational aspects of the Waring problem and related notions. First we will discuss some algorithms to calculate a Waring decomposition of a given form. Then we will study the complexity of the Waring rank in general and sketch a proof by Shitov that the Waring problem is NP-hard. Finally, we give some application of the Waring Rank problem to the problem of

4.1. Algorithms for the Waring rank

First we start with a not so serious example.

Example 4.1 (An extremely fast, but useless algorithm). Consider the following algorithm.

Algorithm 2 The Alexander-Hirschowitz algorithm

Require: A form $F \in \Bbbk[x_0, \ldots, x_n]_d$.
Ensure: $r = \mathrm{WR}(F)$.
1: $r \leftarrow G(n, d)$, obtained from Corollary 1.36.

While this algorithm doesn't even look at the form, it will be correct on virtually all inputs from $\Bbbk[x_0, \ldots, x_n]_d$, because a dense open set of forms has generic rank. But it will be certainly wrong for any form of non-generic rank, so it's hard to argue that this algorithm even solves the problem.

A more useful example can be considered in the case $d = 2$. In this case example 1.3 tells us that the Waring rank of $F = xAx^\mathsf{T}$ equals the rank of the symmetric matrix A. More specifically:

If $U = [u_{ij}] \in \mathrm{O}(n+1, \Bbbk)$ with $U^{-1}AU = \mathrm{diag}(\lambda_0, \ldots, \lambda_n)$, then let $I = \{\, i \in \{0, \ldots, n\} \mid \lambda_i \neq 0 \,\}$

$$F(x) = \sum_{i \in I} \lambda_i L_i^2, \qquad L_i = x_i \circ U^{-1} = \sum_{j=0}^{n} U_{ji} x_i.$$

This leads to the following algorithm.

Algorithm 3 Waring decomposition using orthogonal diagonalization

Require: A form $F = \sum_{1\leq i,j\leq n} b_{ij} \in \Bbbk[x_1,\dots,x_n]_2$.
Ensure: $r = \mathrm{WR}(F)$, $\lambda_1,\dots,\lambda_r \in \Bbbk$, $L_1,\dots,L_r \in \Bbbk[x_1,\dots,x_n]_1$ with $F = \sum_{i=1}^r \lambda_i L_i^2$

```
A ← [(b_ij + b_ji)/2]_{i,j=1}^n.
U ← orthogonal_eigenbasis(A).
D ← U^T A U        //This is a diagonal matrix of eigenvalues of A.
r ← rank(D)
{i_1, ..., i_r} ← { i ∈ {1, ..., n} | D_ii ≠ 0 }
for j = 1, ..., r do
    λ_j ← D_{i_j,i_j}
    L_i ← U_{i_j,1} x_1 + ··· + U_{i_j,n} x_n
end for
```

We also consider the case of *binary forms*, i.e. forms in two variables $F \in \Bbbk[x_0, x_1]$, which was already considered by Sylvester in the 1800's. We give a modern treatment, following [BGI11, Section 3].

We recall the basics about Grassmannians [Har92, Lecture 6].

Definition 4.2. Let V be a $\Bbbk$-vector space of dimension n.

(i) The *Grassmannian of subspaces* of dimension $0 \leq k \leq n$ is the set

$$G(k, V) := \{ U \subseteq V \mid U \text{ is a vector subspace of dimension } k \}.$$

(ii) The *Grassmannian of projective subspaces* of $\mathbb{P}(V)$ of dimension $0 \leq k \leq n-1$ is the set

$$\mathbb{G}(k, \mathbb{P}(V)) := \{ U \subseteq \mathbb{P}(V) \mid U \text{ is a linear subspace of dimension } k \}.$$

The Grassmannians are projective varieties in a natural way: We have the *Plücker embedding*

$$\iota \colon G(k, V) \to \mathbb{P}(\bigwedge^k V), \qquad U = \langle v_1, \dots, v_k \rangle_{\Bbbk} \mapsto [v_1 \wedge \cdots \wedge v_k]$$

This is well-defined, as a different choice of a basis of U changes the wedge product only by the determinant of some base change matrix, so this does not change the representative in $\mathbb{P}(\bigwedge^k V)$. The map ι is injective and its image is a projective variety, thus we made $G(k, V) \cong G(k-1, \mathbb{P}(V))$ into a projective variety. Since subspaces $U \subseteq V$ of dimension k are in natural correspondence with subspaces of dimension $n-d$ in $V^\vee$, we have a natural isomorphism

$$\alpha_{k,n-k} \colon G(k, V) \cong G(n-k, V^\vee).$$

We cite the following Lemma. Let $V = \Bbbk[x_0, x_1]_1$ be the two-dimensional vector space of linear forms.

Lemma 4.3 ([BGI11, Lemma 19]). *For $d \geq r \geq 1$ consider the map*

$$\phi_{r,d-r+1} \colon \mathbb{P}(\mathrm{S}^r V) \to G(d-r+1, \mathrm{S}^d V), \qquad [F] \mapsto F \cdot \mathrm{S}^{d-r} V.$$

(i) *The image of $\iota \circ \phi_{r,d-r+1}$ in $\mathbb{P}(\bigwedge^{d-r+1} \mathrm{S}^d V) \cong \mathbb{P}(\mathrm{S}^{d-1+r}(\mathrm{S}^r V))$ is the $(d-r+1)$-th Veronese embedding of $\mathbb{P}(\mathrm{S}^r V)$.*

(ii) *Identifying $G(d-r+1, \mathrm{S}^d V) \cong \mathbb{G}(r-1, \mathbb{P}(\mathrm{S}^d V^\vee))$, this Veronese variety is the set of linear spaces r-secant to the rational normal curve $\mathcal{C}_d = \nu_d(\mathrm{S}^1 V^\vee) \subseteq \mathbb{P}(\mathrm{S}^d V^\vee)$. More precisely, the image of $[F]$ is the r-secant spanned by the linear factors of F.*

After identifying $\mathbb{P}(\mathrm{S}^d V^\vee)$ with $\mathbb{P}(\mathrm{S}^d V)$, we get the following maps:

$$\begin{tikzcd} \mathbb{P}(\mathrm{S}^r V) \arrow[r, "\phi_{r,d-r+1}"] \arrow[rrr, bend right, "\gamma"'] & \mathbb{G}(d-r, \mathbb{P}(\mathrm{S}^d V)) \arrow[r, "\alpha_{k,n-k}", "\cong"'] & \mathbb{G}(r-1, \mathbb{P}(\mathrm{S}^d V^\vee)) \arrow[r, "\cong"] & \mathbb{G}(r-1, \mathbb{P}(\mathrm{S}^d V)) \end{tikzcd}$$

The content of lemma 4.3 can be rephrased as follows: For a projective subspace $\mathbb{P}(W) \subseteq \mathbb{P}(\mathrm{S}^d V)$ the following are equivalent:

- $\mathbb{P}(W)$ is a r-secant to $\mathcal{C}_d$ in r distinct points;
- $\mathbb{P}(W) = \gamma([F])$ for some form F with d distinct roots.

Therefore we get the following description of the Waring rank of a binary form.

Corollary 4.4. *The Waring rank of a form $0 \neq F \in \mathrm{S}^d V$ is the smallest $r \in \mathbb{N}_+$ such that*

- *$F \in \mathbb{P}(W) \subseteq \mathbb{P}(\mathrm{S}^d V)$ for some $\mathbb{P}(W) \in \mathrm{im}(\gamma) \subseteq \mathbb{G}(r-1, \mathbb{P}(\mathrm{S}^d V))$ and*
- *there exists a form $F_0 \in \mathrm{S}^r V$ with r distinct roots and $\gamma([F_0]) = \mathbb{P}(W)$.*

To describe the image of $\phi_{r,d-r+1}([F_0])$ notice that if $F_0 = \sum_{i=0}^r u_i x_0^i x_1^{r-i}$, then a basis $F_0 \cdot \Bbbk[x_0, x_1]_{d-r}$ is given by

$$\left\{ \sum_{i=0}^r u_i x_0^{i+j} x_1^{d-i-j} \;\middle|\; j = 0, \dots, d-r \right\}.$$

With respect to the monomial basis, this is the subspace spanned by the rows of the following $(d-r+1 \times d+1)$-matrix:

$$\begin{bmatrix} u_0 & u_1 & \dots & u_r & 0 & \dots & 0 & 0 \\ 0 & u_0 & u_1 & \dots & u_r & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \ddots & \vdots & \\ 0 & \dots & 0 & u_0 & u_1 & \dots & u_r & 0 \\ 0 & 0 & \dots & 0 & u_0 & u_1 & \dots & u_r \end{bmatrix}$$

Let $Z_0, \dots, Z_d$ be the dual basis of the monomial basis (which is the monomial basis with an extra factor of $\binom{d}{i}$, see the apolarity section), then $\alpha_{k,n-k}$ maps such a span to the intersection of the hyperplanes

$$\begin{cases} H_0 & : u_0 Z_0 + \cdots + u_r Z_r = 0 \\ H_1 & : u_0 Z_1 + \cdots + u_r Z_{r+1} = 0 \\ \vdots & \\ H_{d-r} & : u_0 Z_{d-r} + \cdots + u_r Z_d = 0 \end{cases}$$

In the view of Corollary 4.4 we see that $F = \sum_{i=0}^{d} a_i \binom{d}{i} x_0^i x_1^{d-i} \in S^d V$ belongs to the subspace $\mathbb{P}(W) = \gamma([F_0])$ if and only if $F \in H_0 \cap \cdots \cap H_{d-r}$ defined using the coefficients of F_0 as before. This can be described as the linear system of equations in $u = (u_0, \ldots, u_r)$

$$\begin{bmatrix} a_0 & a_1 & \ldots & a_r \\ a_1 & a_2 & \ldots & a_{r+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{d-r} & a_{d-r} & \ldots & a_d \end{bmatrix} \cdot \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{$*$}$$

Definition 4.5. The previous matrix corresponding to the binary form $F = \sum_{i=0}^{d} a_i \binom{d}{i} x_0^i x_1^{d-i}$ is called the *catalecticant matrix* $\mathrm{Cat}_{r,d-r}(F)$.

There are two cases for the equation ($*$).

- If $r + 1 > d - r + 1$ then we have more variables than equations, so there is always a nontrivial solution.
- If $r \leq 2d$ then there is a nontrivial solution if and only if $\mathrm{Cat}_{r,d-r}(F)$ has rank $\leq r$

We can finally state Sylvester's algorithm.

Algorithm 4 Sylvester's algorithm

Require: A binary form $0 \neq F = \sum_{i=0}^{d} a_i \binom{d}{i} x_0^i x_1^{d-i} \in \mathbb{C}[x_0, x_1]_d$.
Ensure: $r = \mathrm{WR}(F)$, $F = \sum_{j=1}^{r} \lambda_i L_i^d$ a Waring decomposition.

1: $r \leftarrow 1$.
2: **while** $\mathrm{rank}\, \mathrm{Cat}_{r,d-r}(F)$ is maximal **do**
3: $\quad r \leftarrow r + 1$
4: **end while**
5: Take any nontrivial element $0 \neq F_0 \in \ker \mathrm{Cat}_{r,d-r}(F)$.
6: Compute the roots $(\alpha_i, \beta_i) \in \mathbb{C}^2$ of F_0, $i = 1, \ldots, r$.
7: **if** the roots are not distinct in $\mathbb{P}(\mathbb{C}^2)$ **then**
8: $\quad$ go to step 2 $\quad$ //i. e. increase r further
9: **else**
10: $\quad$ Construct the set of linear forms $\{L_i = \alpha_i x_0 + \beta_i x_1\}$.
11: $\quad$ Solve the linear system of equations $F = \sum_{i=1}^{r} \lambda_i L_i^d$.
12: $\quad$ **return** the Waring decomposition $F = \sum_{i=1}^{r} \lambda_i L_i^d$.
13: **end if**

Example 4.6 (Adapted from [BGI11, Example 11]). We compute a Waring recomposition of

$$F = 2x_0^4 - 4x_0^3x_1 + 30x_0^2x_1^2 - 28x_0x_1^3 + 17x_1^4.$$

Factoring out the binomial coefficients, we get $(a_0, \dots, a_4) = (2, -1, 5, -7, 17)$. The first two Catalecticant matrices are

$$\mathrm{Cat}_{1,3}(F) = \begin{bmatrix} 2 & -1 \\ -1 & 5 \\ 5 & -7 \\ -7 & 17 \end{bmatrix}, \qquad \mathrm{Cat}_{2,2}(F) = \begin{bmatrix} 2 & -1 & 5 \\ -1 & 5 & -7 \\ 5 & -7 & 17 \end{bmatrix}$$

The first Catalecticant matrix has rank 2, while the second matrix has rank $2 < 3$. A generator of the kernel is $(2, -1, -1)^{\mathsf{T}}$ corresponding to the form

$$F_0 = 2X_0^2 - X_0X_1 - X_1^2 = (X_0 - X_1)(2X_0 + X_1).$$

Thus the roots $(\alpha_1, \beta_1) = (1, 1)$ and $(\alpha_2, \beta_2) = (1, -2)$ are distinct. Thus the linear forms of a Waring decomposition are given as $L_1 = x_0 + x_1$, $L_2 = x_0 - 2x_1$, and the equation

$$2x_0^4 - 4x_0^3x_1 + 30x_0^2x_1^2 - 28x_0x_1^3 + 17x_1^4 = \lambda_1(x_0 + x_1)^4 + \lambda_2(x_0 - 2x_1)^4$$

gives $\lambda_1 = \lambda_2 = 1$.

Remark. This algorithm iterates over all $r = 1, \dots, \mathrm{WR}(F)$. It is possible to skip this iterative process by first calculating the Border rank

$$r = \underline{\mathrm{WR}}(F) = \mathrm{rank}\, \mathrm{Cat}_{\lfloor \frac{d}{2} \rfloor, \lceil \frac{d}{2} \rceil}(F),$$

and this constrains $\mathrm{WR}(F) \in \{r, d-r+2\}$. This result was obtained by studying the structure of the *stratification* of $\Bbbk[x_0, x_1]_d$ by Waring rank, for details see the paper by Bernardi, Gimigliano and Idà [BGI11, Section 3].

4.2. Some background on complexity theory

Now that we looked at some algorithms for the Waring rank, it is natural to ask about general complexity theoretic results about the Waring rank and related notions. In this section we recall some notions from complexity theory to make this precise, a good reference is the modern treatment by Arora & Barak [AB09]. We assume familiarity with the notion of a Turing machine.

So far our (pseudocode) algorithms have not been concerned with the question whether a computer (with finite memory) can actually execute such an algorithm. For example if $\Bbbk = \mathbb{C}$ then we cannot even store most numbers in memory, let alone do arithmetic with them.

Definition 4.7. Let $\mathcal{R}$ be a ring. We say that one can *compute effectively* in $\mathcal{R}$ if the following are satisfied:

- Elements $a \in \mathcal{R}$ can be encoded as finite strings $\langle a \rangle$ over some finite alphabet Σ, formally
$$\mathrm{enc}\colon \mathcal{R} \hookrightarrow \Sigma^* = \{\,\text{strings } x_1 \dots x_n \mid n \geq 0, x_i \in \Sigma\,\}, \qquad a \mapsto \langle a \rangle.$$
- There exist deterministic Turing machines M_{add}, M_{mult}, which on input $\langle a \rangle \# \langle b \rangle$ produce $\langle a + b \rangle$ and $\langle a \cdot b \rangle$ respectively.
- The running time of M_{add}, M_{mult} is polynomial in its input bit length, i. e. the number of computation steps on input x is $O(|x|^c)$ for some constant c.

These notions ensure that a computer can do basic arithmetic in $\mathcal{R}$ efficiently. For example, $\mathcal{R}$ could be $\mathbb{Z}, \mathbb{Q}, \mathbb{F}_q$ (any finite field), but not $\mathbb{R}$ or $\mathbb{C}$ (because they are uncountable). Also, a polynomial ring $\mathcal{R}[T_1, \dots, T_n]$ can be represented as linear combinations of monomials, this again yields an efficiently computable ring.

Definition 4.8. We fix some alphabets Σ, Δ.

- A *formal language* (or *problem*) is a subset $A \subseteq \Sigma^*$.
- The complexity class P consists of formal languages A such that there exists a deterministic Turing machine deciding A in polynomial time.
- The complexity class NP consists of formal languages A such that there exists a nondeterministic Turing machine deciding A in polynomial time.
- A language $A \subseteq \Sigma^*$ can be polynomial time many-one reduced to $B \subseteq \Delta^*$, in symbols $A \leq_{\mathrm{m}}^{\mathsf{P}} B$, if there exists a function $f \colon \Sigma^* \to \Delta^*$ computable by a polynomial time Turing machine, such that
$$w \in A \quad \Longleftrightarrow \quad f(w) \in B \qquad \forall w \in \Sigma^*.$$
If $A \leq_{\mathrm{m}}^{\mathsf{P}} B$ and $B \leq_{\mathrm{m}}^{\mathsf{P}} A$, then A and B are said to be *polynomial time equivalent*, in symbols $A \equiv_{\mathrm{m}}^{\mathsf{P}} B$.
- A language A is NP-*hard* if every NP problem can be polynomial time reduced to A, i. e. $B \leq_{\mathrm{m}}^{\mathsf{P}} A$ for all $B \in \mathsf{NP}$. A is NP-*complete* if $A \in \mathsf{NP}$ and it is NP-hard.

Example 4.9. Probably the most well-known NP-complete problem is
$$\mathtt{SAT} = \left\{\, \langle \varphi \rangle \;\middle|\; \varphi \text{ is a Boolean formula and } \varphi(a_1, \dots, a_n) = 1 \text{ for some } a \in \{0,1\}^n \,\right\}.$$
this is a consequence of the famous Cook-Levin theorem. The $\mathtt{SAT}$ problem is NP-complete even when restricting to Boolean formulae in 3-conjunctive normal form (this variant is called

3SAT), i. e. those of the following form:

$$\varphi = \bigwedge_{i=1}^{k} (l_{i,1} \vee l_{i,2} \vee l_{i,3}), \qquad l_{i,j} \in \{x_1, \dots, x_n, \neg x_1, \dots, \neg x_n\}. \tag{**}$$

For us the following problem is of interest. Fix integral domains $\mathcal{R} \subseteq \mathcal{S}$, $\mathcal{R}$ being efficiently computable. Then we can define the problem of deciding solvability in $\mathcal{S}$ of a system of polynomial equations defined over $\mathcal{R}$

$$\texttt{NULLSTELLENSATZ}_{\mathcal{S}/\mathcal{R}} = \left\{ \langle f_1, \dots, f_m \rangle \;\middle|\; \begin{array}{c} f_1, \dots, f_m \in \mathcal{R}[T_1, \dots, T_n] \\ \text{have a common root in } \mathcal{S}^n. \end{array} \right\}$$

A polynomial time reduction from 3SAT can be given as follows:

For a formula φ as in (**), consider the following set of polynomials in $\mathcal{R}[y_1, \dots, y_n]$:

$$b_j = y_j(y_j - 1), \qquad j = 1, \dots, n;$$

$$c_i = v(l_{i,1}) \cdot v(l_{i,2}) \cdot v(l_{i,3}), \qquad i = 1, \dots, k, \qquad \text{where } v(l) = \begin{cases} y_j - 1 & \text{if } l = x_j \\ y_j & \text{if } l = \neg x_j \end{cases}$$

The expanded (!) version of these polynomials can be computed in polynomial time.

Claim. The binary vector $a = (a_1, \dots, a_n) \in \{0, 1\}^n$ represents a satisfying assignment of φ if and only if it is a solution to the polynomial system of equations $b_1 = \dots = b_n = c_1 = \dots = c_k = 0$.

Proof of claim. Indeed, the equations $b_1 = \dots = b_n = 0$ force any solution to consist of 0's and 1's (here integrality is used). We have $c_i(a_1, \dots, a_n) = 0$ if and only of one of the terms $v(l_{i,1})$ vanishes, which is equivalent to $l_{i,1}$ being satisfied. This proves the claim. $\square_{\text{claim}}$

Since 3SAT is NP-hard (and $\leq_{\mathrm{m}}^{\mathrm{P}}$ is transitive), we see that $\texttt{NULLSTELLENSATZ}_{\mathcal{S}/\mathcal{R}}$ is NP-hard. This problem is in fact *much* harder, at least for $\mathcal{R} = \mathcal{S} = \mathbb{Z}$, since the (negative) solution to Hilbert's tenth problem shows that the problem of solvability of Diophantine equations is undecidable.

4.3. The NP-hardness of the Waring rank

We are ready talk about the complexity of Tensor rank and Waring rank. Since we are interested in algebraically closed fields such as $\mathbb{C}$, it makes sense to distinguish between the ring $\mathcal{R}$ over which a given tensor/form is defined and the ring $\mathcal{S}$ over which the rank has to be calculated. For example one could ask about the complex Waring rank of forms with rational coefficients ($\mathcal{R} = \mathbb{Q} \subsetneq \mathcal{S} = \mathbb{C}$).

Notice that the definition of Waring rank (1.1) and tensor rank (1.19) make sense over arbitrary integral domains (replacing "vector space" by "free module"). Hence we consider the following decision problems:

$$\begin{aligned}\texttt{TENSOR\_RANK}_{\mathcal{S}/\mathcal{R}} &= \{\, \langle T, r\rangle \mid r \in \mathbb{N}_0, \text{T is a tensor defined over } \mathcal{R},\ \operatorname{rank}_{\mathcal{S}}(T) \le r \,\}\\ \texttt{WARING\_RANK}_{\mathcal{S}/\mathcal{R}} &= \left\{ \langle F, r\rangle \;\middle|\; r \in \mathbb{N}_0,\ F \in \mathcal{R}[\underline{x}]_d,\ \mathrm{WR}_{\mathcal{S}}(F) \le r \right\}.\end{aligned}$$

In a paper from 2016 Shitov proved the following result characterizing the complexity of $\texttt{TENSOR\_RANK}_{\mathcal{S}/\mathcal{R}}$:

Theorem 4.10 ([Shi16, Theorem 3]). *The problem of deciding if a tensor T with entries in $\mathcal{R}$ has rank at most r is polynomial time equivalent to the problem of deciding solvability in $\mathcal{S}$ of systems of polynomials defined over $\mathcal{R}$. In short:*

$$\texttt{TENSOR\_RANK}_{\mathcal{S}/\mathcal{R}} \equiv_m^{\mathsf{P}} \texttt{NULLSTELLENSATZ}_{\mathcal{S}/\mathcal{R}}.$$

In fact, it suffices to consider only tensors of degree 3, i. e. in $\mathcal{R}^I \otimes \mathcal{R}^J \otimes \mathcal{R}^K$.

The proof can be summarized as follows: Reducing tensor rank to a system of polynomial equations is fairly easy: If $T \in \mathcal{R}^I \otimes \mathcal{R}^J \otimes \mathcal{R}^K$, then the statement $\operatorname{rank}_{\mathcal{S}}(T) \le r$ can be rephrased as

$$\exists \left\{ \begin{array}{c} u_1, \ldots, u_r \in \mathcal{S}^I; \\ v_1, \ldots, v_r \in \mathcal{S}^J; \\ w_1, \ldots, w_r \in \mathcal{S}^K; \\ \lambda_1, \ldots, \lambda_r \in \mathcal{S} \end{array} \right\} \quad : \quad T = \sum_{i=1}^{r} \lambda_i \cdot (u_i \otimes v_i \otimes w_i).$$

This is a system of $I \cdot J \cdot K$ polynomial equations in $r \cdot (I + J + K + 1)$ variables. A similar strategy allows for reducing the Waring rank problem to a system of polynomial equations.

In order to give a reduction of polynomial equations to a rank problem, Shitov first considers the problem of *(minimal) rank matrix completion*:

$$\begin{aligned}&\texttt{MATRIX\_COMPLETION\_RANK}_{r,\mathcal{S}/\mathcal{R}}\\ &\qquad = \left\{ \langle A\rangle \;\middle|\; \begin{array}{c} A \in \mathrm{Mat}(m, n, \mathcal{R} \cup \{*\}), \text{ one can assign the } * \text{ with elements} \\ \text{from } \mathcal{S} \text{ such that the completed matrix has } \mathcal{S}\text{-rank}^1 \le r \end{array} \right\}\end{aligned}$$

He constructs a polynomial time reduction

$$\texttt{NULLSTELLENSATZ}_{\mathcal{S}/\mathcal{R}} \le_{\mathrm{m}}^{\mathsf{P}} \texttt{MATRIX\_COMPLETION\_RANK}_{3,\mathcal{S}/\mathcal{R}}, \qquad \langle f_1, \ldots, f_m\rangle \mapsto \langle \mathcal{B}\rangle$$

which implies that $\texttt{MATRIX\_COMPLETION\_RANK}_{r,\mathcal{S}/\mathcal{R}}$, for $r \ge 3$, is polynomial time equivalent

[1] Here, matrix rank over an integral domain is understood to be tensor rank, coinciding with the usual rank in the case of fields.

to $\texttt{NULLSTELLENSATZ}_{\mathcal{S}/\mathcal{R}}$ for any $r \geq 3$.

He then continues to construct from the incomplete matrix $\mathcal{B}$ with τ many $*$'s a tensor T which has tensor rank $\operatorname{rank}_{\mathcal{S}} T \leq \tau + 3$ if and only if $\mathcal{B}$ admits a completion of rank ≤ 3.

Shitov then proceeds to relate the tensor rank problem (of degree 3 tensors) to the Waring rank problem. Assume $\mathcal{S} = \mathcal{F}$ is a field, let I, J, K be disjoint index sets of size n, and consider a tensor $T \in \mathcal{F}^{I \times J \times K}$. A tensor can be viewed as a 3-dimensional array of elements of $\mathcal{F}$ indexed by $I \times J \times K$, and we adopt the notation $T(i|j|k)$ to denote the entries of the tensor.

We can consider a symmetrization $S(T) \in \mathcal{F}^{H \times H \times H}$, $H := I \cup J \cup K$, given as

$$S(T)(\alpha|\beta|\gamma) = \begin{cases} T(i|j|k) & \text{if } (\alpha, \beta, \gamma) \text{ is a permutation of } (i, j, k) \in I \times J \times K, \\ 0 & \text{otherwise.} \end{cases}$$

Locally, we use the notation $I^2_{\leq} = \{ (i_1, i_2) \mid i_1, i_2 \in I,\ i_1 \leq i_2 \}$. Let $S \in \mathcal{F}^{H \times H \times H}$ be a tensor, $\mathcal{H} := H \cup I^2_{\leq} \cup J^2_{\leq} \cup K^2_{\leq}$, then we can enlarge S to a tensor $\mathcal{T}(S) \in \mathcal{F}^{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}$ by the following procedure:

- Let $\pi = (i_1, i_2) \in I^2_{\leq}$, then the "$\pi$-th unit" is the tensor $M \in \mathcal{F}^{H \times H}$ with

$$M(\alpha|\beta) = \begin{cases} 1 & \text{if } \alpha, \beta \in \{i_1, i_2\} \\ 0 & \text{otherwise.} \end{cases}$$

- $S \rightsquigarrow S' \in \mathcal{F}^{\mathcal{H} \times H \times H}$ is obtained by adjoining π-th unit tensors to S for each $\pi \in I^2_{\leq} \cup J^2_{\leq} \cup K^2_{\leq}$.
- Similarly, $S' \rightsquigarrow S'' \in \mathcal{F}^{\mathcal{H} \times \mathcal{H} \times H}$ is obtained by adjoining π-th unit tensors (with the appropriate index set) to S' in the second dimension, for each $\pi \in I^2_{\leq} \cup J^2_{\leq} \cup K^2_{\leq}$.
- Finally, $S'' \rightsquigarrow \mathcal{T}(S) \in \mathcal{F}^{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}$ is constructed in a similar fashion by adjoining π-th unit tensors in the third dimension.

Notice that $\mathcal{T}(S)$ is symmetric if S is symmetric, and both constructions $T \mapsto S(T)$ and $S \mapsto \mathcal{T}(S)$ can be carried out over the base ring $\mathcal{R}$ in polynomial time. The crucial point of this construction is that we can relate the tensor rank of a tensor $T \in \mathcal{F}^{I \times J \times K}$ to the symmetric tensor rank of $\mathcal{T}(S(T)) \in \mathcal{F}^{\mathcal{H} \times \mathcal{H} \times \mathcal{H}}$ (i. e. the Waring rank of the corresponding form) by the following formula (under the mild assumption $|\mathcal{F}| \geq 9$):

$$\mathrm{WR}_{\mathcal{F}}\, \mathcal{T}(S(T)) = \operatorname{rank}_{\mathcal{F}} T + 9 \cdot \binom{n+1}{2}.$$

This shows that the tensor rank problem in $\mathcal{F}^{I \times J \times K}$ can be polynomial-time reduced to the Waring problem for degree 3 forms in $|\mathcal{H}| = 3n + 3\binom{n+1}{2}$ variables. As a consequence, for any field extension $K/\mathbb{Q}$ we have

$$\texttt{WARING\_RANK}_{K/\mathbb{Q}} \equiv_m^{\mathsf{P}} \texttt{NULLSTELLENSATZ}_{K/\mathbb{Q}},$$

even when restricting to degree 3 forms, and in particular, the Waring rank problem is NP-hard.

Remark. This has the (unfortunate) consequence that, if $\texttt{NULLSTELLENSATZ}_{\mathbb{C}/\mathbb{Q}}$ turns out to be undecidable, then there is no general algorithm calculating the Waring rank *at all*!

4.4. Parametrized algorithms

The previous section suggests that it is infeasible to find an efficient general purpose algorithm for the Waring problem! In order to end this thesis on a positive note, we turn our attention to the recent work of Kevin Pratt [Pra18], in which he relates Waring decompositions of certain polynomials to improved parameterized and exact algorithms for various interesting problems.

In this section we work over the field of complex numbers, in particular all Waring ranks are understood to be over $\mathbb{C}$. Let $G = (V, E)$ be a directed graph, $V = \{v_1, \dots, v_n\}$.

Definition 4.11. (i) A *walk* of length d in G is a sequence $w = (v_{i_0}, \dots, v_{i_d})$ with $(v_{i_{j-1}}, v_{i_j}) \in E$ for $j = 1, \dots, d$.
(ii) If $i_d = i_0$, then w is a *closed walk*. If, additionally, all nodes in w are pairwise distinct (apart from $v_{i_0} = v_{i_d}$), then w is called a *simple cycle*.

Problem 4.12. *Describe an algorithm which on input $\langle G, d\rangle$ calculates the number of simple cycles in G of length d.*

Consider the following *symbolic adjacency matrix* and *graph walk polynomial*[2]

$$A_G := [a_{ij}] \in \operatorname{Mat}(n, \mathbb{C}[\underline{x}]_1), \qquad a_{ij} = \begin{cases} x_i & \text{if } (v_i, v_j) \in E; \\ 0 & \text{otherwise.} \end{cases}$$

$$F_G := \operatorname{tr}(A_G^d) \in \mathbb{C}[\underline{x}]_d.$$

We first prove a combinatorial lemma relating F_G to closed cycles and walks in G.

Lemma 4.13. (i) *The terms of F_G represent closed walks of length d in G:*

$$F_G = \sum_{\substack{\text{closed walks} \\ (v_{i_0}, \dots, v_{i_d})}} x_{i_0} \cdots x_{i_{d-1}}.$$

[2] These are not established names.

(ii) *The number of simple cycles of length d in G is given by the constant value*

$$e_{n,d}\left(\tfrac{\partial}{\partial x_1},\dots,\tfrac{\partial}{\partial x_n}\right)F_G.$$

Here, we meet our action of the polynomial ring of differential operators in $X_i = \frac{\partial}{\partial x_i}$ again, with the notation from chapter 2 we write this expression as $e_{n,d} \circ F_G$.

Proof. (i) By the definition of matrix multiplication, the (i,k)-th entry of A_G^2 consists of the sum $\sum_j x_i x_j$, where j runs over all paths (v_i, v_j, v_k). Inductively, the (i,j)-th entry of A_G^d is

$$\sum_{\substack{\text{paths from } i_0 = i \text{ to } i_d = j \\ (v_{i_0},\dots,v_{i_d})}} x_{i_0}\cdots x_{i_{d-1}}.$$

Hence the polynomials on the diagonal of A_G^d correspond to closed walks, and taking the trace yields the desired identity.

(ii) Let $g = X^{\alpha}$ be a monomial, then for any form F of degree d, $g \circ F = \alpha! \cdot c_{\alpha}$, or just c_{α} if g has no repeated factors. As $e_{n,d}$ consists of all degree d monomials with no repeated factors, by linearity and (i) we get

$$e_{n,d} \circ F_G = \#\{\text{simple closed walks in } G\}. \qquad \blacksquare$$

So finding an algorithm counting the number of closed cycles in G is synonymous with an algorithm evaluating $e_{n,d}\left(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\right)F_G$! Here our knowledge from section 2.1 is useful:

Lemma 4.14. *Let $F \in \mathbb{C}[x_1,\dots,x_n]_d$, $g \in \mathbb{C}[X_1,\dots,X_n]_d$*

(i) *We can "switch" the roles of F and g in the apolarity action, i.e. we have the identity*

$$g(\underline{X}) \circ F(\underline{x}) = F(\underline{X}) \circ g(\underline{x}).$$

(ii) *If $F = \lambda_1 L_1^d + \dots + \lambda_s L_s^d$, where $L_i = c_{i,1}x_1 + \dots + c_{i,n}x_n \in \mathbb{C}[\underline{x}]_1$, then*

$$g \circ F = d! \cdot \sum_{i=1}^{r} \lambda_i g(c_{i,1},\dots,c_{i,n}).$$

Proof. (i) By bilinearity of $\circ$ this breaks down to the fact that for $\deg F = \deg g$,

$$X^{\alpha} \circ x^{\beta} = \begin{cases} \alpha! & \text{if } \alpha = \beta \\ 0 & \text{otherwise.} \end{cases}$$

(ii) This is an immediate consequence of the discussion after Example 2.5. $\blacksquare$

Continuing our discussion, this means that evaluation of $e_{n,d}\left(\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\right)F_G$ can be realized as evaluation of $e_{n,d}$ on the coefficients of a power sum decomposition of F. This observation itself is not so helpful, since the algorithm would have to find a (sufficiently short) power sum decomposition for F_G, which is (probably) a hard problem. But the Lemma also allows us to switch the roles of F_G and $e_{n,d}$, so if

$$e_{n,d} = \sum_{i=1}^{s} \lambda_i L_i^d, \qquad L_i = \sum_{i=1}^{s} c_{i,1}X_1 + \cdots + c_{i,n}X_n$$

is a power sum decomposition, then for *any* G we get

$$\#\{\text{simple cycles in } G\} = d! \sum_{i=1}^{s} \lambda_i F_G(c_{i,1}, \ldots, c_{i,n}).$$

We obtain an algorithm which counts the number of simple cycles in G by using $\mathrm{WR}(e_{n,d})$ many *black box* evaluations of F_G! Theorem 2.23 describes such decompositions, so for d odd we get the number of simple cycles in G as (using the notation from 2.23)

$$\sum_{\substack{I\subseteq\{1,\ldots,n\}\\ |I|\leq\lfloor d/2\rfloor}} \frac{(-1)^{|I|}}{2^{d-1}} \binom{n-\lfloor d/2\rfloor-|I|-1}{\lfloor d/2\rfloor-|I|} \cdot F_G\big(\delta(I,1),\ldots,\delta(I,n)\big).$$

Corollary 4.15. *This Formula yields a* $\binom{n}{\lfloor d/2\rfloor}\,\mathrm{poly}(n)$ *time and* $\mathrm{poly}(n)$ *space algorithm for counting simply cycles.*

This discussion has the following interesting consequence: Given polynomials F, g with only black-box access to F. Then one can evaluate $g \circ F$ using $\mathrm{WR}(g)$ many queries to F. Pratt proves that this bound is actually optimal:

Theorem 4.16 ([Pra18, Theorem 6]). *Fix $g \in \mathbb{C}[\underline{x}]$ and let $F \in \mathbb{C}[\underline{x}]$ be given as a black-box. The minimum number of queries to F needed to compute $g(\frac{\partial}{\partial x})F$ is $\mathrm{WR}(g)$, assuming unit-cost arithmetic operations*[3]*.*

Sketch of proof. The upper bound has already been established.

Claim. Let $m < \mathrm{WR}(g)$. For $\mathbb{X} = \{P_1, \ldots, P_m\} \subseteq \mathbb{P}(\mathbb{C}^n)$ there exists $f \in \mathbb{C}[\underline{X}]_d$ such that $f \in I(\mathbb{X})$ but $g(\frac{\partial}{\partial X})f \neq 0$

Proof of claim. Indeed, assume the contrary, then $I(\mathbb{X}) \subseteq g^{\perp}$, but the Apolarity Lemma 2.8 then implies that $\mathrm{WR}(g) \leq m$. $\square_{\text{claim}}$

So assume an algorithm for $g(\frac{\partial}{\partial x})F$ makes m queries to F, let $v_1, \ldots, v_m \in \mathbb{C}^n$ be these points. Let f be as in the claim for $\mathbb{X} = \{[v_1], \ldots, [v_m]\}$ (possibly discarding identical projective

[3]This means that we only count the number of queries, discarding the cost of actually evaluating F and summing up the results.

points), then

$$(F + f)(v_i) = F(v_i), \qquad i = 1, \ldots, m.$$

Since the algorithm only has black-box access, this shows that it cannot distinguish F from $F + f$. But by construction

$$g(\tfrac{\partial}{\partial X})(F + f) \neq g(\tfrac{\partial}{\partial X})F,$$

a contradiction! ↯ ■

Conclusion

In this thesis we introduced the Waring problem for homogeneous forms and related problems. We discussed both special cases, such as quadratic forms, binary forms and sums of monomials, and general statements, i. e. the Alexander-Hirschowitz theorem.

There are many other interesting related topics, open problems and applications of the Waring problem which we could not discuss here. We give some hints to the literature.

- One possible generalization of the discussion in this thesis is to allow for a base field of positive characteristic. In order to get an interesting theory for cases such as $0 < \operatorname{char}(\Bbbk) \leq d$, one usually replaces the ring of polynomials by the *ring of divided powers* (which coincides with the polynomial ring in the characteristic 0 case) and asks for power sum decompositions in this ring instead. This is the point of view taken in the book [IK99], where the generic Waring problem is related to the Alexander-Hirschowitz theorem in the case $\operatorname{char}(\Bbbk) \nmid d$.
- Another generalization is to ask for the Waring rank over non-algebraically closed fields such as number fields or the real numbers. In this case some arithmetic subtleties enter, for example Reznick [Rez13] shows that the form $F = 3x^5 - 20x^3y^2 + 10xy^4$ has the following ranks over some fields:

$$\mathrm{WR}_{\mathbb{Q}(\sqrt{-1})}(F) = 3, \qquad \mathrm{WR}_{\mathbb{Q}(\sqrt{-2})}(F) = 4, \qquad \mathrm{WR}_{\mathbb{R}}(F) = 5.$$

  The real Waring rank is particularly of practical interest, in this case (taking the euclidean topology) there might be several generic rank (i. e. r such that the set of forms of rank r has nonempty interior) [Ber+18, Section 5.7].
- Instead of asking for *some* Waring decomposition of a given form, one can try to study the set of *all* Waring decompositions at once. If $F = \lambda_1 L_1^d + \cdots + \lambda_s L_s^d$ is a power sum decomposition of $F \in \Bbbk[x_0, \ldots, x_n]_d$, then the linear forms $[L_1], \ldots, [L_s]$ are points in $(\mathbb{P}^n)^\vee$ (as in chapter 2), so they define a point in the *Hilbert scheme* $\mathrm{Hilb}_s((\mathbb{P}^n)^\vee)$. The *variety of power sums* is the scheme

$$\mathrm{VSP}(F, s) = \overline{\left\{ \{[L_1], \ldots, [L_s]\} \in \mathrm{Hilb}_s((\mathbb{P}^n)^\vee) \;\middle|\; \exists \lambda \in \Bbbk^s : F = \lambda_1 L_1^d + \cdots + \lambda_s L_s^d \right\}}$$

  For example, Silvester already noticed that for binary forms $F \in \Bbbk[x_0, x_1]_d$, d odd with and $\mathrm{WR}(F) = (d+1)/2$, the Waring decomposition is *unique* up to scaling and reordering the forms. In this case the variety of power sums is a single point. Questions about dimension, degree, normality/smoothmess are discussed in the book by Iarrobino & Kanev [IK99] and in the paper by Ranestad & Schreyer [RS98].

- The topic of tensor rank and its wide range of connections into different fields of science is a whole topic on its own. An introduction to these applications is given in the expository paper by Kolda & Bader [KB09]. Both tensor rank and symmetric rank have imtimate connections to complexity theory, for example the exponent of matrix multiplication, i. e. the lowest ω such that $N \times N$ matrix multiplication can be computed with $O(N^{\omega})$ scalar operations (currently the best known bound is $\omega \leq 2.3728596$).

A

Sage Code

Here is the Sage code I used to test some cases of the Alexander-Hirschowitz theorem. This code is by no means optimized, for example instead of taking the intersection one should solve the linear system of equations imposed by the set of double points, otherwise the algorithm becomes very slow for larger values of s.

```
from sage.misc.prandom import randint

def random_point_ideal(S, max_denominator=10):
    X = S.gens()
    num_coordinates = len(X)
    i_nonzero = randint(0, num_coordinates-1)

    # Sample random point
    p = [randint(-max_denominator, max_denominator) for i in range(num_coordinates)]
    p[i_nonzero] = randint(1, max_denominator)
    linear_generators = []

    # Calculate defining ideal
    for j in range(0,num_coordinates):
        if j == i_nonzero:
            continue
        linear_generators.append(p[j]*X[i_nonzero] - p[i_nonzero]*X[j])

    m = ideal(linear_generators)
    return m, p

def random_double_point_ideal(S, s, max_denominator=10):
    if s == 0:
        return S.unit_ideal()
    m, p = random_point_ideal(S, max_denominator)
    I = m^2
    XX = [p]
    if s == 1:
        return I, XX
    ideals_to_intersect = []
    for i in range(s-1):
        m, p = random_point_ideal(S, max_denominator)
```

```
            ideals_to_intersect.append(m^2)
            XX.append(p)
    I = I.intersection(*tuple(ideals_to_intersect))
    return I,XX

def expected_codimension(n,d,s):
    N = binomial(n+d,d)
    return min(s*(n+1), N)

# Calculate the Hilbert function by using the Hilbert series
def hf(ideal, d):
    if d < 0:
        return 0
    t = PowerSeriesRing(QQ, 't', default_prec=(d+1)).gen()
    coefficients = ideal.hilbert_series()(t).coefficients()
    return coefficients[d]

# Try to prove the Alexander-Hirschowitz theorem by randomly picking points
def check_AH(n,d,s, max_trials=10, max_denominator=10):
    S = PolynomialRing(QQ, n+1, "X")
    expdim = expected_codimension(n,d,s)
    for i in range(max_trials):
        I,XX = random_double_point_ideal(S, s, max_denominator)
        if hf(I,d) == expdim:
            print(f"Found␣scheme␣of␣{s}␣double␣points␣satisfying␣AH_{n}({d}).")
            return XX
        else:
            print(f"Attempt␣{i}␣failed.␣Defect␣is␣<={expdim␣-␣hf(I,d)}")

    print(f"Could␣not␣verify␣AH_{n}({d})␣on␣{s}␣double␣points.")
    return None

def check_AH_on_relevant_s(n,d, max_trials=10, max_denominator=10):
    l = []
    for s in relevant_cases(n,d):
        X = check_AH(n,d,s, max_trials, max_denominator)
        if not X:
            return
        l.append(X)
    return l

def relevant_cases(n,d):
    q = binomial(n+d,d)/(n+1)
    return set([floor(q), ceil(q)])
```

Index

Bibliography

[AB09] Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge: Cambridge University Press, 2009. ISBN: 9780521424264. DOI: 10.1017/CBO9780511804090 (cit. on p. 61).

[Ådl87] Bjørn Ådlandsvik. "Joins and higher secant varieties". In: *Mathematica Scandinavica* 61 (1987), pp. 213–222. ISSN: 0025-5521. URL: https://www.jstor.org/stable/24492046 (visited on 07/19/2021) (cit. on pp. 16–18).

[AH95] J. Alexander and A. Hirschowitz. "Polynomial interpolation in several variables". In: *Journal of Algebraic Geometry* 4 (Jan. 1995) (cit. on p. 37).

[AM69] Michael Francis Atiyah and I. G. MacDonald. *Introduction to commutative algebra.* Addison-Wesley-Longman, 1969, pp. I–IX, 1–128. ISBN: 978-0-201-40751-8 (cit. on pp. 26, 27).

[BD17] Edoardo Ballico and Alessandro De Paris. "Generic Power Sum Decompositions and Bounds for the Waring Rank". In: *Discrete & Computational Geometry* 57.4 (Mar. 2017), pp. 896–914. ISSN: 1432-0444. DOI: 10.1007/s00454-017-9886-7 (cit. on p. 22).

[Ber+18] Alessandra Bernardi et al. "The Hitchhiker Guide to: Secant Varieties and Tensor Decomposition". In: *Mathematics* 6.12 (Dec. 2018), p. 314. DOI: 10.3390/math6120314 (cit. on pp. 7, 19, 23, 54, 71).

[BGI11] Alessandra Bernardi, Alessandro Gimigliano, and Monica Idà. "Computing symmetric rank for symmetric tensors". In: *Journal of Symbolic Computation* 46.1 (2011), pp. 34–53. ISSN: 0747-7171. DOI: https://doi.org/10.1016/j.jsc.2010.08.001 (cit. on pp. 58, 59, 61).

[BO08] Maria Chiara Brambilla and Giorgio Ottaviani. "On the Alexander–Hirschowitz theorem". In: *Journal of Pure and Applied Algebra* 212.5 (2008), pp. 1229–1251. ISSN: 0022-4049. DOI: https://doi.org/10.1016/j.jpaa.2007.09.014 (cit. on pp. 12, 19, 21, 37, 52).

[BT14] Greg Blekherman and Zach Teitler. "On maximum, typical and generic ranks". In: *Mathematische Annalen* 362 (2014), pp. 1021–1031 (cit. on p. 22).

[Car+15] E. Carlini et al. *Symmetric tensors: rank, Strassen's conjecture and e-computability*. 2015. arXiv: 1506.03176 [math.AC] (cit. on p. 36).

[CCG12] Enrico Carlini, Maria Virginia Catalisano, and Anthony V. Geramita. “The solution to the Waring problem for monomials and the sum of coprime monomials”. In: *Journal of Algebra* 370 (2012), pp. 5–14. ISSN: 0021-8693. DOI: https://doi.org/10.1016/j.jalgebra.2012.07.028 (cit. on pp. 29, 31).

[CGO14] Enrico Carlini, Nathan Grieve, and Luke Oeding. “Four Lectures on Secant Varieties”. In: *Connections Between Algebra, Combinatorics, and Geometry*. Ed. by Susan M. Cooper and Sean Sather-Wagstaff. Vol. 76. Springer, New York, NY, 2014. DOI: 10.1007/978-1-4939-0626-0_2 (cit. on p. 7).

[GW20] Ulrich Görtz and Torsten Wedhorn. *Algebraic Geometry I: Schemes: With Examples and Exercises*. en. 2nd ed. Springer Studium Mathematik - Master. Springer Spektrum, 2020. ISBN: 9783658307325. DOI: 10.1007/978-3-658-30733-2 (cit. on p. 10).

[Har77] Robin Hartshorne. *Algebraic Geometry. Corr. 8th printing*. English. Vol. 52. Springer, New York, NY, 1977. ISBN: 3-540-90244-9 (cit. on pp. 19, 27).

[Har92] Joe Harris. *Algebraic Geometry*. Springer New York, 1992. 352 pp. ISBN: 144193099X (cit. on pp. 47, 58).

[HM21] Huy Tai Ha and Paolo Mantero. *The Alexander-Hirschowitz theorem and related problems*. 2021. arXiv: 2101.09762 [math.AC] (cit. on pp. 37, 41, 52, 54, 55).

[IK99] Anthony Iarrobino and Vassil Kanev. *Power Sums, Gorenstein Algebras, and Determinantal Loci*. Springer Berlin Heidelberg, 1999. 388 pp. ISBN: 3540667660 (cit. on pp. 23, 71, 72).

[KB09] Tamara G. Kolda and Brett W. Bader. “Tensor Decompositions and Applications”. In: *SIAM REVIEW* 51.3 (2009), pp. 455–500 (cit. on p. 72).

[Kra84] Hanspeter Kraft. *Geometrische Methoden in der Invariantentheorie*. 1st ed. Aspects of Mathematics. Vieweg+Teubner Verlag, 1984. ISBN: 978-3-528-08525-4 (cit. on p. 12).

[Lee16] Hwangrae Lee. “Power sum decompositions of elementary symmetric polynomials”. In: *Linear Algebra and its Applications* 492 (2016), pp. 89–97. ISSN: 0024-3795. DOI: https://doi.org/10.1016/j.laa.2015.11.018 (cit. on p. 34).

[Pos12] Elisa Postinghel. “A new proof of the Alexander-Hirschowitz interpolation theorem”. In: *Annali di Matematica* 191 (2012). DOI: 10.1007/s10231-010-0175-9 (cit. on p. 56).

[Pra18] Kevin Pratt. “Waring Rank, Parameterized and Exact Algorithms”. In: *IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)* (July 17, 2018). arXiv: 1807.06194v4 [cs.DS] (cit. on pp. 66, 68).

[Rez13] Bruce Reznick. "On the Length of Binary Forms". In: *Quadratic and Higher Degree Forms*. Ed. by Krishnaswami Alladi et al. New York, NY: Springer New York, 2013, pp. 207–232. ISBN: 978-1-4614-7488-3. DOI: 10.1007/978-1-4614-7488-3_8 (cit. on p. 71).

[RS98] Kristian Ranestad and Frank-Olaf Schreyer. "Varieties of sums of powers". In: *Journal fur die Reine und Angewandte Mathematik* 525 (Jan. 1998). DOI: 10.1515/crll.2000.064 (cit. on p. 72).

[Shi16] Yaroslav Shitov. *How hard is the tensor rank?* 2016. arXiv: 1611.01559 [math.CO] (cit. on p. 64).

[Shi18] Yaroslav Shitov. "A Counterexample to Comon's Conjecture". In: *SIAM J. Appl. Algebra Geom.* 2 (2018), pp. 428–443 (cit. on p. 16).

[Shi19] Yaroslav Shitov. "Counterexamples to Strassen's direct sum conjecture". In: *Acta Mathematica* 222.2 (2019), pp. 363–379. DOI: 10.4310/ACTA.2019.v222.n2.a3 (cit. on p. 36).

[Str73] Volker Strassen. "Vermeidung von Divisionen." In: *Journal für die reine und angewandte Mathematik* 264 (1973), pp. 184–202. URL: http://eudml.org/doc/151394 (cit. on p. 35).

[The21] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 9.3)*. 2021. URL: https://www.sagemath.org (cit. on p. 43).

[VW02] Robert C. Vaughanv and Trevor Wooley. *Waring's Problem: A Survey*. May 2002 (cit. on p. 2).