Computation and Complexity in Algebraic Geometry

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

D I S S E R T A T I O N

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr.rer.nat.)

im Fachgebiet

Mathematik

vorgelegt

von M.Sc. Elisabeth Leonie Kayser
geboren am 20. Februar 1998 in Hannover

Die Annahme der Dissertation wurde empfohlen von:

1. Prof. Dr. Jan Draisma (Universität Bern)
2. Dr. Simon Telen (Max-Planck-Institut für Mathematik in den Naturwissenschaften)

Die Verleihung des akademischen Grades erfolgt mit Bestehen der Verteidigung am 9. Juni 2026 mit dem Gesamtprädikat magna cum laude.

Acknowledgements

> Hofstadter's law: *It always takes longer than you expect, even when you take into account Hofstadter's law.*
>
> Douglas Hofstadter [Hof99]

I would like to start by saying a heartfelt thank you to my advisor Simon. Your guidance and encouragement have been so important to throughout my journey as a a Ph.D. student. You showed me how to become a research mathematician, introduced me to the world of numerical algebraic geometry, and believed in me through all the highs and lows. When it comes to mentorship, I also want to thank Fulvio for his support as my TAC mentor. I have only seen you a few times in person, but our hours of discussions on Zoom have been so enjoyable and have greatly influenced my mathematical interest over the years. Thank you Bernd for introducing me to nonlinear algebra, offering me many opportunities and connecting me with many of the people I now call friends. A heartfelt thank you to Mirke, Saskia, Katharina, Jörg, Thomas, Jana and to the entire administration team. Your help with so many things has been invaluable, and you lifted a lot of weight off my shoulders.

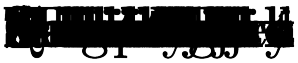
I'd like to thank Florian, Fulvio, Jan, Julian, Sibren, and Simon for spotting typos and providing helpful comments on this thesis. Your feedback greatly improved the exposition and has hopefully reduced its complexity. May I never make the mistake of looking at this document again after submission.

I am grateful for the many scientific collaborations that I had the privilege of being a part of. Thank you Alexander, Andreas, Barbara, Bernhard, Celia, Fulvio, Javier, Mireille, Sibren, and Simon for the productive, but most importantly, enjoyable time we spent together thinking, drafting, mathing and laughing. May our conjectures prove to be correct!

As the last of the Youngling to become a Post-Padawan: Thank you Barbara, Dmitrii, Laura and Max for being with me on this journey. We had joyful evenings together enjoying Italian dinner, playing Mario Kart, eating pancakes and just genuinely being there for each other. May the Force be with us! I had the honor of sharing the position of Ph.D. representative with Anton and Barbara. We were the best representatives the MPI ever had in the year 2024! Thank you, Celia, Mireille, Neza, Pierre, Laura, and Skye for co-organizing the Queer in Math Day 2024 and 2025. These events warmed my heart like no others. Thank you for all the game nights, retreats, IMPaRtieS, COMBOs, conferences, and all the other wonderful distractions that the MPI crowd has to offer.

Contrary to popular belief, there is life outside math, and Leipzig has offered me many reasons to leave the Inselstraße and explore the city. I am grateful to have been a member of the Chorlektiv Leipzicals for the last three years. Our rehearsals, concerts, retreats and karaoke nights have brought a joy to my life that only music can. The RosaLinde e.V. often provided a great reason to bike to Lindenau and participate in their support groups. The Sunday brunches and evening discussions helped me grow as a person — not only wider, but hopefully also wiser. I would also like to thank the Leipziger Stadtwerke for keeping the swimming pools open until 10 p.m. in the evenings, to stop me from drowning in office work.

The number of wonderful people I met at the institute is nearly uncountable. Thank you all for making the MPI MiS more than just a research institute, but a place I called home. To avoid embarrassing myself by forgetting to mention one of you, here are all of you beautiful people:



I would like to thank my parents for their never-ending love and support. And last but not least, thank you Nadja for sticking with me, with all the long-distance traveling and DB adventures. My life is better with your company.

Summary

In this thesis, we examine systems of polynomial equations and their solutions through the lens of computation and complexity. We explore problems from mathematics, science and engineering, driven by the desire to design efficient algorithms and to uncover structural complexity.

After a preliminary Chapter 0 on projective geometry, we investigate the problem of numerical symmetric tensor decomposition in Chapter 1. The usage of eigenvalue methods for polynomial system solving leads us to the study of Hilbert functions of non-saturated ideals of points. Specifically, we predict the exact values of the Hilbert function of the ideal generated by the degree d component of the ideal of a general set of points in projective space. We prove this conjecture in several infinite families and obtain unconditional bounds on the regularity of the ideal.

In Chapter 2 we solve a "formidable challenge" posed by Geiger et al. related to Mukai's construction of canonical curves of genus 8. Given a generic self-dual point configuration of 14 points in $\mathbb{P}^6$, we construct a 6-dimensional projective subspace in $\mathbb{P}^{14}$ whose intersection with the Grassmannian $X_8 = \mathrm{Gr}(2, \mathbb{C}^6) \hookrightarrow \mathbb{P}^{14}$ in its Plücker embedding realizes that point configuration as a linear slice. Our approach relies on a novel parametrization of the moduli space of self-dual points using skew-symmetric matrices.

In Chapter 3, we study the computational complexity of the subalgebra membership problem. We prove that the general problem is complete for the class EXPSPACE, while the homogeneous variant is complete for PSPACE. For the first time, we also provide upper bounds on the degree of a subalgebra membership certificate. Furthermore, we investigate the semilinear structure of initial algebras of certain invariant rings, revealing a new finite description for infinitely generated algebras.

Motivated by maximum likelihood estimation and particle physics, we consider the complex critical points of a product of affine-linear forms $\ell_0^{u_0} \cdots \ell_n^{u_n}$ in Chapter 4. The logarithmic discriminant characterizes exponents $u \in \mathbb{P}^n$ for which the function has a degenerate critical point in the hyperplane arrangement complement $\mathbb{C}^d \setminus \{ \ell_0 \cdots \ell_n = 0 \}$. We prove that for arrangements in general position the logarithmic discriminant is an irreducible hypersurface and study its degree and positivity.

In Chapter 5 we address the Euclidean distance minimization problem to secant varieties of the rational normal curve $\mathbb{P}\mathcal{X}_{d,r} = \sigma_r \nu_d \mathbb{P}^1$. Using a novel approach, we calculate the algebraic degree of this optimization problem for generic and specific inner products. Our research unifies and generalizes previous results and leads to new conjectures about Euclidean distance degrees for the Bombieri norm. We also consider the case of $d \to \infty$, leading to new insights into the model order reduction problem in dynamical systems.

Zusammenfassung

In dieser Dissertation studieren wir polynomielle Gleichungssysteme unter den Gesichtspunkten der Computerberechnung und Komplexität. Wir erkunden Probleme aus Mathematik, Naturwissenschaften und Ingenieurswissenschaften, stets angetrieben durch das Bedürfnis, effiziente Algorithmen zu gestalten und strukturelle Komplexität aufzudecken.

Nach einem einführenden Kapitel 0 über projektive Geometrie untersuchen wir in Kapitel 1 das Problem der numerischen symmetrischen Tensorzerlegung. Ein Ansatz via Eigenwertmethoden zur Lösung polynomieller Gleichungen führt uns zum Studium der Hilbertfunktionen von nicht-saturierten Polynomidealen von Punkten. Genauer geben wir eine Vermutung über die exakten Werte der Hilbertfunktion des Ideals $\mathrm{I}(\mathbb{X})_{\langle d \rangle}$, welches erzeugt wird von der d-ten Komponente des Ideals einer allgemeinen Menge von Punkten im projektiven Raum. Wir beweisen diese Vermutung in mehreren unendlichen Familien und erhalten obere Schranken bezüglich der Regularität des Ideals.

In Kapitel 2 bezwingen wir eine von Geiger et al. gestellte "formidable challenge" im Zusammenhang mit Mukais Konstruktion kanonischer Kurven von Genus 8. Gegeben eine allgemeine selbstduale Punktkonfiguration von 14 Punkten im $\mathbb{P}^6$, konstruieren wir einen 6-dimensionalen projektiven Unterraum im $\mathbb{P}^{14}$, dessen Schnittmenge mit der Grassmannschen $X_8 = \mathrm{Gr}(2, \mathbb{C}^6)$ in ihrer Plückereinbettung diese Punktkonfiguration als linearen Schnitt realisiert. Unser Ansatz beruht auf einer neuen Parametrisierung des Modulraums der selbstdualen Punktkonfigurationen mittels schiefsymmetrischer Matrizen.

In Kapitel 3 studieren wir die Berechnungskomplexität des Zugehörigkeitsproblems für polynomielle Unteralgebren. Wir beweisen, dass das Problem in voller Allgemeinheit vollständig für die Klasse EXPSPACE ist, während die homogene Variante vollständig für PSPACE ist. Erstmals geben wir auch obere Schranken für den Grad eines Unteralgebrazugehörigkeitszertifikats. Weiterhin untersuchen wir die semilineare Struktur von Initialalgebren bestimmter Invariantenringe, welche eine neuartige endliche Beschreibung unendlich erzeugter Unteralgebren ermöglicht.

Motiviert durch Maximum-Likelihood-Schätzung und Teilchenphysik studieren wir in Kapitel 4 die komplexen kritischen Punkte eines Produkts von affin-linearen Funktionen $\ell_0^{u_0} \cdots \ell_n^{u_n}$. Die logarithmische Diskriminante charakterisiert Exponenten $u \in \mathbb{P}^n$ für welche die Funktion einen ausgearteten kritischen Punkt im Komplement des Hyperebenenarrangements $\mathbb{C}^d \setminus \{ \ell_0 \cdots \ell_n = 0 \}$ hat. Wir beweisen für Arrangements in allgemeiner Lage, dass die logarithmische Diskriminante eine irreduzible Hyperfläche ist, und studieren ihren Grad und Positivität.

In Kapitel 5 widmen wir uns dem Minimierungsproblem des euklidischen Abstands zu Sekantenvarietäten der rationalen Normalenkurve $\mathbb{P}\mathcal{X}_{d,r} = \sigma_r \nu_d \mathbb{P}^1$. Mittels eines neuartigen Ansatzes berechnen wir den algebraischen Grad dieses Optimierungsproblems

sowohl für generische als auch für spezifische Skalarprodukte. Unsere Forschung vereint und erweitert vorherige Resultate und führt zu neuen Vermutungen zu Euclidean distance degrees bezüglich des Bombieri-Skalarprodukts. Wir betrachten ebenfalls die Situation für $d \to \infty$, welche zu neuen Erkenntnissen zum Model Order Reduction Problem in der Theorie der dynamischen Systeme führt.

Authorship

The results presented in this thesis are based on four published articles of myself [GKT25; BK25; Kay25; KKT25] as well as currently unpublished research. Many of the results were obtained in collaboration with coauthors. All five chapters 1 to 5 start with one or two **Background** sections of expository nature; these sections are written by myself and contain no original research results. In the remaining sections, authorship is attributed as follows:

Chapter 0 is an introductory chapter containing background material on commutative algebra and algebraic geometry. The exposition is written by myself and contains no original research with the exception of Propositions 0.3.9 and 0.4.15, the former appearing in the article [GKT25, Proposition A.3].

Chapter 1 is based on the article *Hilbert Functions of Chopped Ideals* [GKT25] coauthored with Fulvio Gesmundo and Simon Telen. All authors contributed equally to all aspects of the publication and mathematical results presented in the article. I am responsible for the `Macaulay2` code, Simon Telen has written the `Julia` code.

Chapter 2 is based on the article *Mukai Lifting of Self-Dual Points in* $\mathbb{P}^6$ [BK25] coauthored with Barbara Betti. Both authors contributed equally to all aspects of the publication, implementation and mathematical results presented in the article. The content of the article also appears in Barbara Betti's Ph.D. thesis [Bet25, Chapter 4].

Chapter 3 is based on the article *Computational Complexity of Polynomial Subalgebras* [Kay25], except for Section 3.5.3, which is based on currently unpublished joint work with Bernhard Reinke. Both authors contributed equally to the research presented in that section.

Chapter 4 is based on the article *Logarithmic Discriminants of Hyperplane Arrangements* [KKT25] coauthored with Andreas Kretschmer and Simon Telen. All authors contributed equally to all aspects of the publication, implementation and mathematical results presented in the article.

Chapter 5 is based on currently unpublished joint work with Sibren Lagauw. Both authors contributed equally to the research. Parts of the work in this chapter have also appeared in Sibren Lagauw's Ph.D. thesis [Lag26, Chapter 5 & 6].

All figures in this thesis have been created by the author unless specified otherwise. The only exception is Figure 2.1, which is used with permission from silviana amethyst. The images have been created using Asymptote, GeoGebra, `mathcha.io`, Matlab, Mathematica, and TikZ/pgfplots.

Contents

List of Symbols

Numbers and sets

$\mathbb{N}, \mathbb{N}_+$	Natural numbers $\{0, 1, 2, \dots\}$, positive integers $\{1, 2, 3, \dots\}$
$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	Integers, rational, real, and complex numbers
$\mathcal{P}(X)$	Powerset of a set X
$[n]$	n-element set $\{0, \dots, n-1\}$
$\binom{t}{k}$	Binomial coefficient $\frac{t(t-1)\cdots(t-(k-1))}{k!}$
$\binom{X}{k}$	Set of k-element subsets of the set X
i	Imaginary unit, $\sqrt{-1}$

Linear algebra

V	A $\mathbb{K}$-vector space
e_i	Standard basis vectors of $\mathbb{K}^n$ or $\mathbb{K}^{n+1}$
$V^\vee$	Dual vector space
$\dim_{\mathbb{K}} V$	Dimension of V
$\operatorname{Sym}^d V$	Space of symmetric tensor of degree d over V
$\mathrm{S}(V)$	Symmetric algebra generated by V
$\mathbb{P}V$	Projectivization of V
$[v]$	Class of $0 \neq v \in V$ in $\mathbb{P}(V)$

Rings, ideals, and modules

$\mathbb{K}$	A field, algebraically closed unless specified otherwise
R	Finite type $\mathbb{K}$-algebra, often a polynomial ring in n variables
S	Standard graded polynomial ring, usually in $n+1$ variables
$\langle f_1, \dots, f_r \rangle$	Ideal generated by $f_1, \dots, f_s$
$\mathrm{hf}_M, \mathrm{HS}_M, \mathrm{HP}_M$	Hilbert function, series, and polynomial of M
$\operatorname{reg}_{\mathrm{H}}(M)$	Hilbert regularity of M
J^{sat}	Saturation of the ideal J
$\operatorname{Ann}_S M$	Annihilator ideal of the S-module M
$\operatorname{length} M$	Length of module of finite length
$\gcd(f, g)$	Greatest common divisor of the polynomials f, g
$I_k(\varphi)$	Ideal generated by the k-minors of the matrix φ

Term orders and initial forms

$\mathsf{Mon}(\underline{x})$	Monoid of monomials in the variables $\underline{x} = \{x_1, \dots\}$
$\prec, \prec_{\text{lex}}$	Monomial order, lexicographic monomial order
$<_w$	Partial order defined by w
$\operatorname{in}_\prec f$, $\operatorname{in}_\prec I$, $\operatorname{in}_\prec A$	$\prec$-initial term/ideal/algebra of $f/I/A$
$\operatorname{in}_w f$, $\operatorname{in}_w I$, $\operatorname{in}_w A$	w-initial term/ideal/algebra of $f/I/A$
$\operatorname{nf}_\prec(f, J)$	$\prec$-normal form of f against the ideal I

Homological algebra

$\operatorname{reg}_{\text{CM}}(M)$	Castelnuovo–Mumford regularity of M
$\operatorname{Syz}(m_1, \dots, m_b)$	Module of syzygies of the generators $m_1, \dots, m_b \in M$
$\beta_{ij}(M)$	Graded Betti number $\dim_{\mathbb{K}} \operatorname{Tor}_i^S(M, \mathbb{K})_j$
$\operatorname{p.dim} M$	Projective dimension of M
$\operatorname{depth}_I M$	I-depth of the module M
$\operatorname{H}^i_{\mathfrak{m}}(M)$	i-th local cohomology module of M

Algebraic geometry

$\mathbb{A}^n, \mathbb{P}^n$	Affine and projective space of dimension n over $\mathbb{K}$
$\mathbb{V}(\mathcal{F})$	Vanishing scheme of the polynomial equations $\mathcal{F}$
$\operatorname{I}(X)$	Ideal of the subset/subscheme
X^{red}	Underlying variety of a scheme X
$\dim X$, $\dim M$	Krull dimension of the variety/scheme X or $\mathbb{V}(\operatorname{Ann}_S(M))$
$\deg X$, $\deg S/I$	Degree of the scheme X/graded ring S/I
$\widehat{X}$	Affine cone over a projective variety
$\operatorname{Spec} R$	Affine variety with coordinate ring R
$\operatorname{Proj} R$	Projective variety with homogeneous coordinate ring R
$\Omega_{X/Y}$	Sheaf of relative differentials of $X \to Y$
J_f	Jacobi matrix of the morphism $f\colon X \to Y$
$\operatorname{Gr}(k, V)$	Grassmannian of k-dim. subspaces of the vector space V
$\mathbb{G}(k, \mathbb{P}^n)$	Grassmannian of k-dimensional projective subspaces of $\mathbb{P}^n$
$\sigma_r X$	r-th secant variety of the projective variety X
ν_d	Veronese embedding of degree d

Intersection theory

$\operatorname{A}(X)$, $\operatorname{A}^d(X)$	Chow ring of X and its d-th graded component
$[Y]$	Class of the subscheme $Y \subseteq X$ in the Chow ring $\operatorname{A}(X)$
$c(\mathcal{F})$	Chern class of the bundle $\mathcal{F}$
$s(\mathcal{F})$	Segre class of the bundle $\mathcal{F}$
$D_r(\varphi)$	Degeneracy locus where $\operatorname{rank} \varphi_x \leq r$

Chapter 1

Z	A general set of r points in $\mathbb{P}^n$
$\operatorname{Sat}(J)$	The saturation index set of J
$d = d_{n,r}$	$\min\{\, j \in \mathbb{N} \mid \binom{n+j}{n} \,\}$
$I_{\langle d \rangle} = \langle I_d \rangle$	The chopped ideal in degree d
$\gamma_n(d, r)$	The saturation gap of r points in $\mathbb{P}^n$
μ_j	The multiplication map $\mu_j \colon I(Z)_d \otimes_{\mathbb{K}} S_j \to I(Z)_{d+j}$
B	Hilbert–Burch matrix
$\mathrm{frö}_{d,s}(t)$	The expected Hilbert function from Fröberg's conjecture
U_{genHF}	The open set in $(\mathbb{P}^n)^r$ of points with generic Hilbert function
$\mathfrak{c}$	Chopping map
$\Delta \operatorname{hf}_Z$	First difference of the function hf_Z

Chapter 2

$H(x; u)$	Parametric polynomial system
$\operatorname{mvol}(P_1, \ldots, P_n)$	Mixed volume of n convex polyhedra in $\mathbb{R}^n$
Γ	Set of $2n$ self-dual points in $\mathbb{P}^{n-1}$
$\mathcal{C}$	Cyley transform
$\mathcal{P}_n^m$	Moduli space of ordered m-tuples of points in $\mathbb{P}^n$ modulo $\mathrm{SL}(n)$
$\mathcal{A}_n$	Moduli space of self-dual points in $\mathbb{P}^n$
X_8	Mukai Grassmannian of genus 8 $X_8 = \mathrm{Gr}(2, \mathbb{C}^6)$
$L, \mathbb{L}$	Linear embedding $\mathbb{P}^6 \to \mathbb{P}^{14}$ and its image
p_i, q_j	The Plücker coordinates and relations for $\mathrm{Gr}(2, \mathbb{C}^6)$

Chapter 3

$\mathrm{A} \leq_{\mathrm{m}}^{\mathsf{L}} \mathrm{B}$	Decision problem A is log-space many-one reducible to B
$h = (h_i)_i$	Ideal membership certificate
p	Subalgebra membership certificate
$\deg GB(J)$	Largest degree of an element in the min. Gröbner basis of J
$\mathbb{K}[M]$	Semigroup ring of an affine semigroup $M \subseteq \mathbb{Z}^n$
$\mathbb{K}[M]^G$	Invariant subring of G
ϕ, $L(\phi)$	A Presburger formula and the corresponding semilinear set

Chapter 4

$\mathrm{Fit}_k(M)$	k-th Fitting ideal
$\mathrm{Ram}(f), \mathrm{Branch}(f)$	Ramification and branch locus of $f\colon X \to Y$
$\mathcal{Z}_i(X)$	i-th associated hypersurface of X
Ch_X, Hu_X	Chow form and Hurwitz form of the projective variety X
$\mathcal{L}_u(x)$	Log-likelihood function
$\mathcal{L}_X^\circ, \mathcal{L}_X$	Likelihood correspondence in $X_{\mathrm{reg}} \times \mathbb{P}^n$ and its closure in $X \times \mathbb{P}^n$
$f\colon \mathcal{L}_X \to \mathbb{P}^n$	The likelihood projection
$\mathrm{MLdeg}(X)$	Maximum likelihood degree of X
$\mathcal{A}, \ell$	Hyperplane arrangement $\mathcal{A} \subseteq \mathbb{C}^d$ defined by $\ell_0, \dots, \ell_n$
X	Very-affine variety, usually $\mathbb{C}^d \setminus \mathcal{A}$
L, A, b	The matrix $L^\mathsf{T} = [b \mid A]$ such that $\ell = Ax + b$
$\chi_{\mathcal{A}}$	Characteristic polynomial of $\mathcal{A}$
$\mathbb{T}^{n+1}, \mathbb{T} = \mathbb{T}(\mathbb{P}^n)$	$n+1$ dimensional algebraic torus and the torus in $\mathbb{P}^n$
$\nabla_{\log}(X), \Delta_{\log}(X)$	Logarithmic discriminant and its defining equation
$\nabla_{\mathrm{Hu}}(X)$	Hurwitz discriminant of X
$\mathcal{R}_L$	Reciprocal linear space associated to $L \in \mathbb{C}^{(d+1)\times(n+1)}$
$\mathcal{M}_{0,m}$	The moduli space of m-marked curves of genus 0
$\Delta_{\log}^w(X)$	w-initial form of $\Delta_{\log}(X)$

Chapter 5

Λ, Q	Symmetric non-degenerate weight matrix, $Q(x) = x^\mathsf{T}\Lambda x$
$\mathrm{EDD}_\Lambda(X), \mathrm{EDD}_{\mathrm{gen}}(X)$	Euclidean distance degree of X with resp. to Λ, generic EDD
$\mathrm{EDdef}_\Lambda(X)$	ED defect
$y, \hat{y}$	Data point in $\mathbb{C}^n$, approximant $\hat{y}$
ℓ	Vector of Lagrange multipliers
$\mathcal{E}_X$	ED correspondence
$\mathcal{X}_{d,r}$	Affine variety of Hankel matrices of rank $\leq r$
μ_x	Milnor number of isolated singularity
$\mathcal{Q}_\Lambda$	Isotropic quadric
a, $a(z)$	Linear recurrence and its characteristic polynomial
$\mathrm{vand}(\omega), \mathrm{vand}^{(j)}(\omega)$	Vandermonde vector and confluent Vandermonde vector
$T_d(a)$	Toeplitz matrix
$H_r(y)$	Hankel matrix
$\mathcal{Z}\{y\}$	$\mathcal{Z}$-transform of sequence y
$\mathcal{L}_y(\hat{y}, a, \ell)$	Lagrange function
$\mathcal{B}_\Lambda$, $\mathcal{G}$	Bad and good locus
$\mathcal{E}_{d,r}$	Substituted ED correspondence
$\mathbf{1}, F, \Theta$	Unit, Frobenius and Bombieri inner product
$\mathcal{W}$	Walsh variety

List of Tables

List of Figures

Introduction

> For he who seeks for methods without having a definite problem in mind seeks for the most part in vain.
>
> DAVID HILBERT [Hil02]

Algebraic geometry is a centuries-old branch of mathematics that studies polynomial equations and their solutions using algebraic techniques. Nevertheless, an increasing number of students learning modern algebraic geometry who have never tried to solve a polynomial system of equations. In a sense, the relatively young field of computational algebraic geometry goes back to the roots of the discipline, answering concrete questions that the grand-grandparents of algebraic geometry would appreciate just as well. How many solutions are there? Can we parametrize them? How do they vary as coefficients change?

What is computation in algebraic geometry?

Let $f_1, \dots, f_s$ be a given set of polynomials in $R = \mathbb{C}[x_1, \dots, x_n]$. Assume that the ideal I generated by the polynomials is zero-dimensional, that is, the set

$$Z = \mathbb{V}(I) = \{\, z \in \mathbb{C}^n \mid f_1(z) = \cdots = f_s(z) = 0 \,\}$$

is finite. What does it mean to *solve* this polynomial system?

Using a computer algebra system such as `Macaulay2` [GS] or `OSCAR` [OSC24], and an exact representation of the f_i (say, with rational coefficients), one can manipulate the ideal I symbolically. Operations on I such as computing the radical, primary decomposition and variable elimination reveal information about the solutions. For example, one can determine number of solutions, whether the points are defined over the coefficient field, or the minimal algebraic equations that the coordinates satisfy.

Another interpretation of "solving" is to compute numerical approximations (floating-point numbers) of the roots. This can be accomplished for example using the homotopy continuation approach: Put the given equations f in a suitable family of equations $F(x; u)$ where $F(x; 0) = f$ and $F(x; 1)$ is a simpler system with known solutions. Then one can try to track the solutions of the simple system to our target system as u approaches 0. To satisfy the desire for theoretical guarantees, one can also certify the existence of a solution z in a small neighborhood using interval arithmetic.

One can also transform the task of solving a polynomial system into the problem of computing eigenvalues/eigenvectors of certain matrices. The key idea is that R/I is a finite-dimensional vector space on which the coordinates x_i act as linear endomorphisms; their eigendecompositions reveals the coordinates of the solutions. This approach benefits from fast and robust numerical eigenvalue solvers and shifts the difficulty to the problem of constructing the matrices.

All of these three approaches have their uses, but none are "free": Symbolic computation is often slow in practice, even for moderately-sized systems. Homotopy continuation may have to track millions of paths even if only a few of them converge. Constructing the multiplication matrices for the eigenvalue solver is a science of its own. Why is that?

What is complexity in algebraic geometry?

Whenever we encounter a projective variety, we ask ourselves:

"What is your dimension? What is your degree? What are your defining equations?"

These properties of a variety are a measure of the structural complexity, and they help us understand the object at hand. The degree and size of a Gröbner basis provide additional insight into the complexity of the object at hand, and limit what can be computed symbolically in algebraic geometry [BM93]. Similarly, the degree of an optimization problem — that is, the number of complex critical points of the objective function — gives insight into the optimization landscape and can help find global optima when all complex critical points can be computed in a certified way.

Gröbner bases and other symbolic methods for the study of polynomial ideals suffer from the curse of a worst-case (doubly-)exponential growth. Computational complexity theory can help explain this worst-case behavior and presents a general framework of what kind of problems are solvable in theory and practice. The analysis of algorithms in computer algebra, such as Buchberger's algorithm [Buc06], is not an easy task; it is closely related to the algebraic and combinatorial complexity of polynomial ideals.

Moving from symbolic methods in computer algebra to numerical algebraic geometry reveals different aspects of complexity. Aside from the algebraic degree (the number of solutions to check or track), another important aspect are *discriminant loci* in parameter spaces. These loci, which are often hypersurfaces, constitute of parameters where the qualitative nature of the solution set changes: Solutions merge to form singular solutions, or positive-dimensional components appear. Path-tracking algorithms work best when the solution path stays far away from the discriminant, which allows for larger iterative step sizes and improved runtime. The degree of the discriminant locus is a measure of how difficult it is to avoid. It plays a fundamental role in the theoretical and practical efficiency of numerical polynomial system solving [BC13].

What is this thesis about?

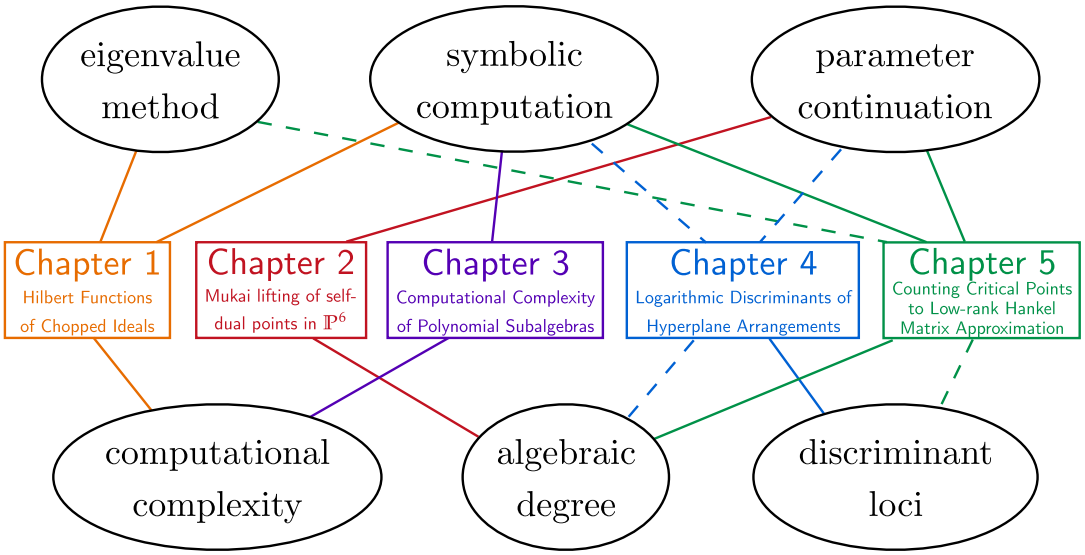
Whenever there is the desire to compute and understand, there is a measure of complexity. Although complexity may eem like the antagonist of computation, both are necessary to reveal the structure of a mathematical problem. Even more importantly, both can lead to the discovery of beautiful mathematical results! The goal of this thesis is to uncover some of this beauty hidden in problems from applied algebraic geometry:

(1) Solve challenging polynomial systems from mathematics and the sciences.

(2) Study the complexity surrounding the computational problem.

(3) Uncover beautiful geometric and combinatorial structures.

The following diagram highlights connections of the chapters of this thesis with the themes of computation and complexity.



We now discuss the main research highlights presented in this thesis.

Chapter 0 is of introductory nature, presenting a useful toolkit from commutative algebra and algebraic geometry to be leveraged in the following chapters. A careful review of topics such as Hilbert functions and free resolutions (used in Chapter 1), Gröbner and SAGBI bases (Chapters 2 and 3), differentials (Chapter 4) and basic intersection theory (Chapters 4 and 5) form a solid foundation for the main body of the thesis.

In Chapter 1, our point of departure is the problem of symmetric tensor decomposition: Given $F \in \mathbb{C}[X_0, \ldots, X_n]_D$, find linear forms $L_1, \ldots, L_r$ such that $F = L_1^D + \cdots + L_r^D$. Under genericity assumptions and in a range of not-too large r, the linear forms $[L_1], \ldots, [L_r] \in \mathbb{P}^n$ are uniquely determined by F, and equations of certain degree vanishing on these points can be obtained via apolarity theory [IK99]. These equations constitute a *heavily* over-determined polynomial system in $\mathbb{P}^n$. We propose to use the

eigenvalue method to solve these systems, motivated by previous success for other tensor decomposition problems [TV22]. This approach requires knowledge about the Hilbert function of the ideal spanned by the equations, which happen to be the subideal spanned by all equations vanishing on the points in a low degree.

In this way, we are naturally lead to the following, abstract algebraic problem: Given the ideal of a general set of points $Z \subseteq \mathbb{P}^n$, what is the nature of the Hilbert function of the *chopped ideal* generated by the degree d component of $\mathrm{I}(Z)$? We present a precise conjecture about the values of the Hilbert function, backed by substantial evidence both of computational and theoretical nature. We prove the conjecture in several infinite families, and provide an implementation of the symmetric tensor decomposition procedure using this method.

In Chapter 2, we turn our focus to the homotopy continuation approach to solve a challenging problem from algebraic geometry. In 1988, Mukai proved that the intersection of the Grassmannian $\mathrm{Gr}(2,6) \subseteq \mathbb{P}^{14}$ in its Plücker embedding with a general linear subspace of dimension 7 is a general curve of genus 8 in its canonical embedding [Muk88; Muk92]. Slicing this Grassmannian with linear spaces of other dimensions yields other interesting varieties: K3 surfaces, Fano 3-folds and, when slicing with general $\mathbb{P}^6$, self-dual point configurations. The Mukai lifting problem, posed by Geiger et al. [Gei+23] as a "formidable challenge", concerns the inverse task: Given such a variety, realize it as a linear section of the Grassmannian. In 2009 Petrakiev showed that general self-dual point configurations of 14 points in $\mathbb{P}^6$ indeed arise as complementary linear sections of $\mathrm{Gr}(2,6)$ [Pet09].

We turn Petrakiev's existence result into a practical algorithm: Using a homotopy continuation approach, we compute, for a given self-dual point configuration, a linear section of the Grassmannian that realizes this configuration. Our method relies on a novel parametrization of the space of self-dual points using skew-symmetric matrices. Along the way, we also experimentally study the nature of real solutions to the slicing problem.

In Chapter 3, we take a closer look at polynomial subalgebras and their computational complexity. The complexity of polynomial ideals and Gröbner bases has been studied extensively at least since the 1980's, with doubly-exponential worst-case complexity results, but also many more fine-grained results depending on the geometry of the ideal [MT17]. Computation with polynomial subalgebras have become a prominent subject in pure and applied algebra in recent years [Rei03; Bur+24]. In contrast, the computational complexity of polynomial subalgebras has not been systematically studied before in the literature.

We fill this gap by proving complexity-theoretic completeness of the subalgebra membership problem for various classes of polynomials. We prove EXPSPACE-completeness for the general problem, as well as PSPACE-completeness for the class of homogeneous polynomials. Furthermore, worst-case degree bounds are presented for subalgebra membership certificates. We also study the structure of infinitely generated subalgebras, revealing a semilinear structure in the initial algebras of certain invariant rings.

In Chapter 4, we turn our attention to an optimization problem from algebraic statistics: The maximum likelihood estimation problem (MLE) for linear models concerns minimizing a product of linear forms $\ell_0(x)^{u_0} \cdots \ell_n^{u_n}(x)$ over the complement of their vanishing locus. In the context of positive geometry, the critical equations of this problem appear as the scattering equations for bi-adjoint scalar ϕ^3-theories in particle physics. The likelihood geometry [HS14] is closely linked to the combinatorial structure of the associated hyperplane arrangement given by the linear forms.

We study the discriminant locus of the MLE problem, consisting of complex parameters for which the critical points become degenerate. We show that for general linear forms, this *logarithmic discriminant* is an irreducible hypersurface in the parameter space. We study its degree and positivity, using the projective geometry of reciprocal linear spaces. We also study the logarithmic discriminant of specific hyperplane arrangements complement representing the moduli space of marked genus 0 curves $\mathcal{M}_{0,m}$, drawing connections to soft limits in particle physics.

In the final Chapter 5, we join forces with our engineering friends and study approximation problems from dynamical systems through the lens of algebraic geometry. In mathematical language, this corresponds to a Euclidean distance minimization problem to the variety of low-rank Hankel matrices $\mathcal{X}_{d,r}$, also known as the r-secant variety of the rational normal curve in $\mathbb{P}^d$. The search for globally optimal solutions leads us to consider the algebraic degree of this problem, also known as the Euclidean distance degree (EDD).

Using tools from intersection theory, we study the structure of the critical equations, recovering and extending previous results on the EDD of $\mathcal{X}_{d,r}$. Extensive computer experiments lead us to conjectures about specific EDDs with respect to the Bombieri inner product, which were previously out of reach for computation. We also consider an "infinite-dimensional" analogue of the problem as $d \to \infty$, which corresponds to a rational function approximation problem. We give a new proof of finiteness for the critical locus and present a formula for the algebraic degree of this problem.

Despite all the complexities involved, I hope to also achieve the following fourth goal:

(4) Make the theory of computational algebra and geometry accessible and enjoyable through ready-made implementations and examples.

Chapter 0

A projective geometry toolbox

Contents

> "Should you just be an algebraist or a geometer?" is like saying "Would you rather be deaf or blind?".
>
> MICHAEL ATIYAH [Ati02]

In this preliminary chapter we collect various results from commutative algebra and projective geometry that will be used in the later chapters. None of the statements or proofs in this chapter are claimed to be original unless indicated otherwise (specifically, Propositions 0.3.9 and 0.4.15). We provide proof ideas when deemed illustrative, and otherwise defer to standard references such as [Eis95; Mat87; Eis05] for commutative algebra and [Har92; Rus16; Ful98; EH16] for projective geometry and intersection theory.

The chapter is structured as follows: We start Section 0.1 by introducing basic results about Hilbert functions and free resolutions of graded modules. We then take a quick homological tour through Cohen–Macaulay rings and Castelnuovo–Mumford regularity. Next, in Section 0.2 we consider Gröbner and SAGBI bases, hinting at their use in computational algebraic geometry. We provide several results on linear slices of projective varieties in Section 0.3, and finish with a brief introduction to intersection theory and degeneracy loci in Section 0.4.

Throughout this chapter let $\mathbb{K}$ be an algebraically closed field, except for Section 0.2, where it is useful to allow "small fields" such as $\mathbb{Q}$ and $\mathbb{F}_q$ for computational purposes. All varieties and schemes will be quasi-projective over the algebraically closed field $\mathbb{K}$. A variety is a reduced (quasi-projective) scheme, potentially reducible. Irreducible varieties will be referred to as such. $S = \mathbb{K}[x_0, \dots, x_n]$ will always denote a standard graded polynomial ring in $n+1$ variables over $\mathbb{K}$. All rings considered will be Noetherian and can be assumed to be finitely generated $\mathbb{K}$-algebras or localization thereof.

0.1 Commutative algebra

In this section we introduce basic notions from the theory of Hilbert functions of graded modules, as well as their graded free resolutions. We also introduce the class of Cohen–Macaulay rings that enjoy very pleasant features connecting homological and geometric properties. We also briefly introduce the module of relative differentials.

0.1.1 Hilbert functions of graded modules

We begin with collecting the basic facts about Hilbert functions of graded modules.

Definition 0.1.1. The *Hilbert function* of a graded $\mathbb{K}$-vector space $V = \bigoplus_{d\in\mathbb{Z}} V_d$ with finite-dimensional components V_d is

$$\mathrm{hf}_V \colon \mathbb{Z} \to \mathbb{N}, \qquad \mathrm{hf}_V(d) = \dim_{\mathbb{K}} V_d.$$

If $V_d = 0$ for $d < -N$, then the *Hilbert series* HS_V is the formal power series

$$\mathrm{HS}_V(T) = \sum_{d=-N}^{\infty} \mathrm{hf}_V(d) T^d \in T^{-N}\mathbb{Z}[\![T]\!] \subseteq \mathbb{Q}(\!(T)\!).$$

If there is a d_0 and a polynomial $p(t) \in \mathbb{Q}[t]$ such that $\mathrm{hf}_V(d) = p(d)$ for $d \geq d_0$, then this polynomial is uniquely determined by V and is called the *Hilbert polynomial* HP_V.

Example 0.1.2. The polynomial ring $S = \mathbb{K}[x_0, \ldots, x_n]$ has the Hilbert function

$$\mathrm{hf}_S(d) = \dim_{\mathbb{K}} S_d = \#\{\text{monomials of degree } d\} = \begin{cases} \binom{d+n}{n} & d \geq 0 \\ 0 & d < 0. \end{cases}$$

This function agrees with the polynomial $p(t) = \binom{n+t}{n}$ for $t \geq -n$. The Hilbert series has the simple form

$$\mathrm{HS}_S(T) = \sum_{d=0}^{\infty} \binom{n+d}{n} T^d = \frac{1}{(1-T)^{n+1}}. \qquad \triangleleft$$

A finite graded module $M = \bigoplus_{d \in \mathbb{Z}} M_d$ over the polynomial ring S has finite-dimensional components and indeed has a Hilbert polynomial. More precisely, its Hilbert series is a rational function.

Lemma 0.1.3. *For a finite graded module M the Hilbert series is rational of the form*

$$\mathrm{HS}_M(T) = \frac{c_l T^l + \cdots + c_r T^r}{(1-T)^{n+1}}, \qquad l \leq 0 \leq r, \quad c_l, \ldots, c_r \in \mathbb{Z}.$$

The Hilbert function and polynomial have the form

$$\mathrm{hf}_M(d) = \sum_{i=l}^{r} c_i \cdot \mathrm{hf}_S(d-i), \qquad \mathrm{HP}_M(t) = \sum_{i=l}^{r} c_i \cdot \binom{t-i+n}{n}.$$

With this in mind, the following definition makes sense:

Definition 0.1.4 (Hilbert regularity)**.** The *Hilbert regularity* of a finite graded S-module is

$$\mathrm{reg}_{\mathrm{H}} M = \min\{\, d_0 \in \mathbb{Z} \mid \mathrm{hf}_M(d) = \mathrm{HP}_M(d) \text{ for all } d \geq d_0 \,\}.$$

We will mainly be interested in Hilbert functions and Hilbert polynomials of graded ideals $I \subseteq S$ and their quotients. We briefly recall the ideal-variety (or rather, ideal-subscheme) correspondence in projective space.

Definition 0.1.5 (Saturation)**.** The *saturation* of an ideal $I \subseteq S$ is

$$I^{\mathrm{sat}} = \left\{ f \in S \;\middle|\; \mathfrak{m}^k f \subseteq I \text{ for some } k \geq 1 \right\} = \bigcup_{k \geq 1} (I : \mathfrak{m}^k).$$

I is saturated if $I = I^{\mathrm{sat}}$.

The ideals $I \subseteq I^{\mathrm{sat}}$ differ only in finitely many degrees. Ideals with the same saturation define the same projective scheme. Denote by X^{red} the underlying reduced subvariety of X.

Proposition 0.1.6 (Projective Nullstellensatz)**.** *The operations* $\mathbb{V}(-)$ *and* $\mathrm{I}(-)$ *induce a inclusion-reversing bijection between projective schemes* $X \subseteq \mathbb{P}^n$ *and saturated ideals* $I \subseteq S$*. Here, reduced subsets correspond to radical ideals and irreducible subvarieties correspond to prime ideals. To be precise: For a graded ideal* $I \subseteq S$ *we have*

$$\mathrm{I}(\mathbb{V}(I)) = I^{\mathrm{sat}}, \qquad \mathrm{I}(\mathbb{V}(I)^{\mathrm{red}}) = \begin{cases} \sqrt{I} & \text{if } \mathbb{V}(I) \neq \emptyset \text{ or } I = S, \\ \mathfrak{m} & \text{otherwise.} \end{cases}$$

Recall the degree of a projective scheme $X \subseteq \mathbb{P}^n$ of dimension k is the length of the zero-dimensional scheme $X \cap H_1 \cap \cdots \cap H_k$ for general hyperplanes H_i. When X is (generically) reduced, this intersection is a finite set of reduced points.

Lemma 0.1.7 (Dimension and degree via Hilbert polynomial)**.** *Let* $I \subseteq S$ *be an ideal with* $\emptyset \neq X = \mathbb{V}(I) \subseteq \mathbb{P}^n$*. Write* $\mathrm{HS}_{S/I}(T) = \frac{q(T)}{(1-T)^{d+1}}$ *with* $(1-T) \nmid q(T)$ *(i. e.* $q(1) \neq 0$*). Then*

$$d = \deg \mathrm{HP}_{S/I} = \dim X.$$

Moreover, the leading term of $\mathrm{HP}_{S/I}(t)$ *is* $\frac{1}{d!}(\deg X)t^d$*.*

0.1.2 Syzygies and free resolution

Let M be a graded S-module. We make the following definitions only in the graded case, but most of the theory also makes sense in the local or, to some extent, ungraded case. Denote by $S[-d]$ the S-module with shifted grading $S[-d]_j = S_{j-d}$.

Definition 0.1.8. (i) A minimal generating set of M is a set of homogeneous $m_1, \ldots, m_b$ generating M as a module, such that no proper subset is a generating set.

(ii) Let $m_1, \ldots, m_b \in M$ be homogeneous of degrees $\deg m_i = d_i$. The module of syzygies of $(m_1, \ldots, m_b)$ is the set of its annihilating coefficient vectors

$$\mathrm{Syz}(m_1, \ldots, m_b) \coloneqq \left\{ (f_1, \ldots, f_b) \in \bigoplus_{i=1}^{b} S[-d_i] \;\middle|\; f_1 m_1 + \cdots + f_b m_b = 0 \right\}.$$

(iii) If $m_1, \ldots, m_b$ is a minimal generating set of M, then $\mathrm{Syz}(M) \coloneqq \mathrm{Syz}(m_1, \ldots, m_b)$ is the first syzygy module of M.

As the kernel of a graded map of S-modules $\bigoplus_{i=1}^{b} S[-d_i] \to M$, the syzygy module is graded. It is also well-defined, as the following lemma shows:

Lemma 0.1.9. (i) *The number of minimal generators and their degrees are uniquely determined. More specifically for any minimal homogeneous generating set* $G \subseteq M$*,*

$$\beta_{0,j}(M) \coloneqq \#\{\, m \in G \mid \deg m = j \,\} = \dim_{\mathbb{K}} M_j / S_1 M_{j-1}.$$

(ii) *If $n_1, \dots, n_b \in M$ is a second set of minimal generators of degrees e_i, then there is a graded isomorphism identifying the syzygy submodules*

$$\psi\colon \bigoplus_{i=1}^{b} S[-d_i] \xrightarrow{\sim} \bigoplus_{i=1}^{b} S[-e_i], \quad \psi(\mathrm{Syz}(m_1, \dots, m_b)) = \mathrm{Syz}(n_1, \dots, n_b).$$

Proof. (i) The number of minimal generators of lowest degree $d \in \mathbb{Z}$ is necessarily $\dim_{\mathbb{K}} M_d$. These elements generate a $\mathbb{K}$-subspace $S_1 M_d \subseteq M_{d+1}$. The minimal generators of degree $d+1$ form a basis of a complement of this subspace, hence their number is $\dim_{\mathbb{K}} M_{d+1}/S_1 M_d$. The general statement follows by induction.

(ii) After reordering, we may assume $d_i = e_i$ for each generator and they are sorted by degree in ascending order. The main difficulty is of notational nature.

Write n_i as a linear combination of the m_j with homogeneous coefficients $a_{ij} \in S_{d_i - d_j}$, in particular $a_{ij} = 0$ if $d_i < d_j$. Then define

$$\psi\colon F = \bigoplus_{i=1}^{b} S[-d_i]\mathrm{e}_i \to F, \qquad \psi(\mathrm{e}_i) = \sum_{j=1}^{b} a_{ij}\mathrm{e}_j,$$

this is a homomorphism of free graded S-modules. By construction, ψ makes the following diagrams commutative:

$$\begin{array}{ccccc} \mathrm{Syz}(m) & \hookrightarrow & F & \xrightarrow{m} & M \\ \downarrow{\scriptstyle \psi} & & \downarrow{\scriptstyle \psi} & & \| \\ \mathrm{Syz}(n) & \hookrightarrow & F & \xrightarrow{n} & M \end{array} \qquad \begin{array}{ccc} \mathbb{K}^{\beta_{0,j}} = F_j/S_1F_{j-1} & \xrightarrow{m} & M_j/S_1M_{j-1} \\ \downarrow{\scriptstyle \overline{\psi_j}} & & \| \\ \mathbb{K}^{\beta_{0,j}} = F_j/S_1F_{j-1} & \xrightarrow{n} & M_j/S_1M_{j-1} \end{array} \qquad j \in \mathbb{Z}.$$

The matrix $M(\psi) \in S^{b \times b}$ with respect to the given basis has upper triangular block form

$$M(\psi) = \begin{bmatrix} A_{d,d} & A_{d,d+1} & \cdots & A_{d,d_{\max}} \\ 0 & A_{d+1,d+1} & \ddots & \vdots \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & A_{d_{\max},d_{\max}} \end{bmatrix}, \qquad A_{\gamma,\delta} \in S_{\delta-\gamma}^{\beta_{0,\delta} \times \beta_{0,\gamma}}.$$

The diagonal blocks $A_{\gamma,\gamma} \in \mathbb{K}^{\beta_{0,\gamma} \times \beta_{0,\gamma}}$ are invertible, since modulo the lower order generators, they map a basis $\subseteq m$ of $M_\gamma/S_1M_{\gamma-1}$ to a basis $\subseteq n$. Thus ψ is an isomorphism, and by definition of Syz it maps the syzygy submodules into each other. □

Definition 0.1.10. A *graded free resolution* is a (potentially infinite) chain complex of finite free S-modules

$$\cdots \xrightarrow{\varphi_2} F_2 \xrightarrow{\varphi_1} F_1 \xrightarrow{\varphi_0} F_0 \longrightarrow 0 \longrightarrow \cdots, \qquad F_i \cong \bigoplus_{j \in \mathbb{Z}} S[-j]^{\beta_{ij}}$$

together with a map $\varepsilon\colon F_0 \to M$, such that

$$\cdots \xrightarrow{\varphi_2} F_2 \xrightarrow{\varphi_1} F_1 \xrightarrow{\varphi_0} F_0 \xrightarrow{\varepsilon} M \longrightarrow 0$$

is an exact sequence of graded S-modules.

A free resolution is *minimal* if every non-zero entry of the matrices φ_i are non-constant (so of positive degree). If $(F_i)_i \to M$ is a minimal free resolution of M, then the $\beta_{ij} =: \beta_{ij}(M)$ are the *graded Betti numbers* of M.

Some basic properties of (minimal) free resolutions are summarized in the following lemma.

Lemma 0.1.11. (i) *Every finite module M has a minimal free resolution.*

(ii) *The following conditions are equivalent for a free resolution $(F_\bullet)$ of M.*

(a) *$(F_\bullet)$ is minimal in the sense of the previous definition.*

(b) *$(F_\bullet)$ is minimal among all free resolutions partially ordered by $(F_\bullet) \prec_{\text{lex}} (G_\bullet)$. Here $(F_\bullet) \prec_{\text{lex}} (G_\bullet)$ if there is a $i_0 \geq 0$ with $\operatorname{rank}_S F_i = \operatorname{rank}_S G_i$ for $i < i_0$ and $\operatorname{rank}_S F_{i_0} < \operatorname{rank}_S G_{i_0}$.*

(c) *$(F_\bullet)$ is* not *of the form $(F_\bullet) \cong (G_\bullet) \oplus (E_\bullet)$, where $(G_\bullet)$ is a free resolution of M and $(E_\bullet) = (\cdots \to 0 \to S[-d] \xrightarrow{\cdot\lambda} S[-d] \to 0 \to \cdots)$, $\lambda \in \mathbb{K}^\times$.*

(iii) *The graded Betti numbers are independent of the minimal free resolution, they satisfy*

$$\beta_{ij}(M) = \dim_{\mathbb{K}} \operatorname{Tor}_i^S(M, \mathbb{K})_j.$$

The projective dimension $\operatorname{p.dim} M$, the minimal $\delta \in \mathbb{N} \cup \{\infty\}$ such that M has a resolution by (non-necessarily graded) projective S-modules, equals the largest i such that some $\beta_{ij}(M) \neq 0$.

(iv) *In fact, any two minimal free resolutions $(F_i)_i, (F'_i)_i$ are isomorphic, i. e. there exist graded isomorphisms $\psi_i \colon F_i \to F'_i$ making the composite diagram commutative.*

Proof. (i) Let $m_1, \ldots, m_{\beta_0}$ be a homogeneous minimal generating set of M and $\varepsilon := m \colon F_0 := \bigoplus_i S[-\deg(m_i)] \twoheadrightarrow M$. Set $K := \operatorname{Ker} \varepsilon = \operatorname{Syz}(m)$ and pick minimal generators k_i of K. Set $F_1 := \bigoplus_i S[-\deg(k_i)]$ and $\varphi_0 := k \colon F_1 \to K \subseteq F_0$. Repeating this yields a free resolution, and it is minimal: If $\varphi_i \not\equiv 0$ modulo $\mathfrak{m}$, then the generators of the i-th syzygy module would be $\mathbb{K}$-linearly dependent.

(ii) See [Eis05, Section 1B] for a discussion.

(iii) The graded Tor modules can be calculated using any graded free resolution $(F_\bullet)$, by $\operatorname{Tor}_i^S(M, \mathbb{K}) = \operatorname{H}_i(F_\bullet \otimes_S S/\mathfrak{m}) = \operatorname{H}_i(F_\bullet/\mathfrak{m}F_\bullet)$. Note that the differentials of $F_\bullet/\mathfrak{m}F_\bullet$ are identically zero by minimality, so

$$\operatorname{Tor}_i^S(M, \mathbb{K}) = \operatorname{H}_i(F_\bullet/\mathfrak{m}F_\bullet) = F_i/\mathfrak{m}F_i = \bigoplus_{j \in \mathbb{Z}} \mathbb{K}[-j]^{\beta_{ij}}.$$

A minimal free resolution is a projective resolution, showing that its length bounds $\operatorname{p.dim} M$ from above. On the other hand, since $\operatorname{Tor}_i^S(M, \mathbb{K})$ can be computed by any projective resolution, the other inequality follows as well.

(iv) We can construct isomorphisms $\psi_0, \psi_1, \dots$ iteratively via Lemma 0.1.9(ii). □

Corollary 0.1.12. *If $\beta_{i,j}(M) = 0$ for some i and all $j \le j_0$, then $\beta_{i+1,j+1}(M) = 0$ for all $j \le j_0$.*

Proof. By assumption, we have $F_i = \bigoplus_{j>j_0} S[-j]^{\beta_{ij}}$. Since the entries in the map $\varphi_i \colon F_{i+1} \to F_i$ are of positive degree (by minimality), the smallest degree of a generator in F_{i+1} must be at least $j_0 + 1$. □

Example 0.1.13 (The graded Koszul complex)**.** Let $f_1, \dots, f_s \in S = \mathbb{K}[x_0, \dots, x_n]$ be homogeneous elements of positive degree $d_i = \deg f_i$. Consider the complex of graded free modules

$$F_1 = \bigoplus_{i=1}^{s} S[-d_i]\mathrm{e}_i, \qquad F_k := \bigwedge^k F_1 \cong \bigoplus_{1 \le i_1 < \dots < i_k \le s} S[-d_{i_1} - \dots - d_{i_k}]\mathrm{e}_{i_1} \wedge \dots \wedge \mathrm{e}_{i_k}.$$

with differentials

$$\varphi_{k-1} \colon \bigwedge^k F_1 \to \bigwedge^{k-1} F_1, \qquad \mathrm{e}_{i_1} \wedge \dots \wedge \mathrm{e}_{i_k} \mapsto \sum_{j=1}^{k} (-1)^{j-1} f_{i_j} \cdot \mathrm{e}_{i_1} \wedge \dots \widehat{\mathrm{e}_{i_j}} \dots \wedge \mathrm{e}_{i_k}.$$

Then the *Koszul complex* $(F_\bullet)$ is a minimal complex (in the sense that the φ_i have entries of positive degrees), and it is exact if and only if $f_1, \dots, f_s$ is a regular sequence (see Definition 0.1.19 below).

A simple special case is $n+1 = s = 2$, the syzygies of two binary forms $f, g \in \mathbb{K}[x_0, x_1]$. Assume first that f, g are coprime, then the minimal syzygy of f and g is given by $g \cdot f - f \cdot g = 0$. Let $d := \deg f$, $e := \deg g$, then the minimal free resolution of $\langle f, g \rangle_S$ is given by

$$0 \longleftarrow \langle f, g \rangle_S \xleftarrow{[f\ g]} S[-d] \oplus S[-e] \xleftarrow{\begin{bmatrix} g \\ -f \end{bmatrix}} S[-d-e] \longleftarrow 0.$$

In particular, $\operatorname{Syz}(f, g) = S \cdot (g, -f) \subseteq S[-d] \oplus S[-e]$. This leads to a homogeneous version of *Bézout's lemma.* ◁

Lemma 0.1.14 (Bézout)**.** *Let $f, g \in \mathbb{K}[x_0, x_1]$ be of degree d, e and $c := \deg \gcd(f, g)$. Then $\operatorname{Syz}(f, g)$ is spanned by the element $(\frac{g}{\gcd(f,g)}, \frac{-f}{\gcd(f,g)})$ and*

$$\dim_{\mathbb{K}} \operatorname{Syz}(f, g)_k = \begin{cases} 0 & k < d + e - c \\ k - d - e + c + 1 & k \ge d + e - c. \end{cases}$$

Over non-polynomial rings, minimal free resolutions may be infinitely long:

Example 0.1.15. Let $R = \mathbb{K}[\varepsilon]/\langle \varepsilon^2 \rangle = \mathbb{K} \oplus \mathbb{K}\varepsilon$ and $M = R/\langle \varepsilon \rangle$. Then the minimal free resolution is given by

$$\dots \xrightarrow{\cdot\varepsilon} R[-2] \xrightarrow{\cdot\varepsilon} R[-1] \xrightarrow{\cdot\varepsilon} R \twoheadrightarrow M \to 0.$$

In particular, there is no finite free resolution and the non-zero graded Betti numbers are $\beta_{i,i} = 1$ for $i \ge 0$. ◁

We are only interested in modules over polynomial rings, which have a *finite* free resolution:

Theorem 0.1.16 (Hilbert syzygy theorem)**.** *Let $S = \mathbb{K}[x_0, \dots, x_n]$ and M a finite graded S-module. Then M has a graded free resolution of length $\leq n+1$.*

Proof idea. The statement is true for complete intersections by Example 0.1.13. In particular, the maximal ideal $\mathfrak{m} = \langle x_0, \dots, x_n \rangle$ has a graded free resolution of length $n+1$. The graded Betti numbers $\beta_{i,j}(M) = \dim_{\mathbb{K}} \operatorname{Tor}_i^S(M, \mathbb{K})_j$ detect the length of the free resolution by Lemma 0.1.11(iii). Since $\operatorname{Tor}_i^S(-,-)$ is symmetric, the Betti numbers can be computed by tensoring the minimal free resolution of $\mathbb{K} = S/\mathfrak{m}$ by M. Hence $\beta_{i,j}(M) = 0$ for $i > n+1$. □

Thus the Betti numbers of a finite graded S-module can be arranged in a finite two-way list

Definition 0.1.17 (Betti table)**.** The Betti table of a graded module M is the integer table indexed by rows $d = 0, 1, \dots$ and columns $i = 0, 1, 2, \dots$, whose (d, i)-th entry is $\beta_{i,d+i}$.

This convention partially for the unavoidable linear degree growth with increasing homological degree (Corollary 0.1.12). For example, an ideal has a linear free resolution if the Betti table of I is concentrated in a single row.

Definition 0.1.18. The *Castelnuovo–Mumford regularity* of a graded module M is

$$\operatorname{reg}_{\mathrm{CM}}(M) \coloneqq \sup_i b_i(M) - i, \qquad b_i(M) \coloneqq \max\{\, j \mid \beta_{ij}(M) \neq 0 \,\}.$$

In the next section we provide characterizations of $\operatorname{reg}_{\mathrm{CM}}(M)$ and s connection to the Hilbert regularity.

0.1.3 Cohen–Macaulay Rings and local cohomology

Definition 0.1.19 (Depth, Cohen–Macaulay ring)**.** Let R be a ring and M a R-module.

(i) A sequence $f_1, \dots, f_d \in I$ is a M-regular sequence if f_i is a non-zerodivisor on $M/\langle f_1, \dots, f_{i-1}\rangle M$ for $i = 1, \dots, d$ and $M/\langle f_1, \dots, f_{i-1}\rangle M \neq 0$.

(ii) Let $I \subseteq R$ be an ideal. If $IM \neq M$, then the I-depth of M is

$$\operatorname{depth}_I(M) = \sup\{\, d \mid \exists M\text{-regular sequence } f_1, \dots, f_d \in I \,\} \in \mathbb{N} \cup \{\infty\}$$

By convention, if $IM = M$ then $\operatorname{depth}_I(M) \coloneqq \infty$.

(iii) If R is a Noetherian local ring and M is finite over R, then M is Cohen-Macaulay if $\operatorname{depth}_{\mathfrak{m}}(M) = \dim M \coloneqq \dim R/\operatorname{Ann}_R(M)$.

If R is not (necessarily) local, then M is *Cohen–Macaulay* if $M_{\mathfrak{p}}$ is Cohen–Macaulay over $R_{\mathfrak{p}}$ for all primes $\mathfrak{p} \in \operatorname{Spec} R$. A ring is Cohen–Macaulay if it is Cohen–Macaulay as a module over itself.

If R is local or a standard graded algebra with maximal ideal $\mathfrak{m}$, then we write $\operatorname{depth}(M)$ instead of $\operatorname{depth}_{\mathfrak{m}}(M)$.

Example 0.1.20. Regular local rings are Cohen–Macaulay, as their maximal ideals are generated by regular sequences. The polynomial ring over a field $S = \mathbb{K}[x_0, \dots, x_n]$ is Cohen–Macaulay. ◁

The invariants $\operatorname{depth} M$ and $\operatorname{p.dim} M$ are related by the following formula, giving an alternative characterization of the Cohen–Macaulay property.

Theorem 0.1.21 (Auslander–Buchsbaum formula)**.** *Let R be a Noetherian local ring and M a finitely generated R-module with $\operatorname{p.dim}_R M < \infty$, then*

$$\operatorname{depth} R = \operatorname{depth} M + \operatorname{p.dim}_R M.$$

Corollary 0.1.22. *If R is a regular local ring, then M is Cohen–Macaulay if and only if $\operatorname{p.dim}_R M = \dim R - \dim M$.*

We are mainly interested in modules over $S = \mathbb{K}[x_0, \dots, x_n]$, in which case one can test the Cohen–Macaulay property "globally" using homogeneous regular elements. For more characterizations of standard graded Cohen–Macaulay rings, see [Cho92].

Proposition 0.1.23. *If M is a finite graded S-module, then M is Cohen–Macaulay if and only if the graded depth, the longest homogeneous M-regular sequence in $\mathfrak{m}$, equals $\dim M$. Moreover, the Auslander–Buchsbaum formula holds in the graded case as well.*

Regular sequences behave particularly well in Cohen–Macaulay rings.

Theorem 0.1.24 (Regular sequences in CM rings)**.** *Let R be a local or standard graded CM ring and $f_1, \dots, f_k \in \mathfrak{m}$ (homogeneous in the graded case). Then the following are equivalent:*

(a) $\dim R/\langle f_1, \dots, f_i\rangle = \dim R - i$ *for* $i = 1, \dots, k$;

(b) $f_1, \dots, f_k$ *is a R-regular sequence.*

If this is the case, then $R/\langle f_1, \dots, f_k\rangle$ is also a CM ring.

By the previous theorem, complete intersections in Cohen-Macaulay rings are again Cohen–Macaulay. A deep generalization is the following [BV88, Theorem 2.1 & 2.8]

Theorem 0.1.25 (Determinantal rings are Cohen–Macaulay)**.** *Let R be an equidimensional CM ring and $\varphi \in R^{e\times f}$ a matrix. Let $0 \le r \le \min\{e, f\}$ and consider the* determinantal ideal

$$I = I_{r+1}(\varphi) \coloneqq \langle (r+1)\text{-minors of } \varphi\rangle \subseteq R.$$

Then $\dim R/I \ge \dim R - (e-r)(f-r)$, and if equality holds then R/I is Cohen-Macaulay.

One pleasant feature of Cohen–Macaulay rings is that they have no embedded components:

Theorem 0.1.26 (Unmixedness theorem)**.** *Let $I \subseteq R$ be an ideal such that R/I is Cohen–Macaulay, then I has no embedded components. In particular, I is reduced if and only if this generically holds on every irreducible component of $\mathbb{V}(I) \subseteq \operatorname{Spec} R$.*

Proof. The first part is [Eis95, Corollary 18.14]. For the second part, note that if I were generically reduced but not reduced, then the locus $\mathfrak{p} \in \operatorname{Spec} R$ where I is not reduced would be an embedded component, contradicting the first part. □

To present the relation to saturation and regularity, we briefly introduce the notion of *local cohomology*. Our presentation loosely follows [Eis05, Appendix 1].

Definition 0.1.27. The local cohomology functors $\mathrm{H}^i_{\mathfrak{m}}(-)\colon S\mathrm{Mod} \to S\mathrm{Mod}$ are the right-derived functors of the local sections functor

$$\mathrm{H}^0_{\mathfrak{m}}(M) = \left\{ m \in M \;\middle|\; \exists k\colon \mathfrak{m}^k m = 0 \right\}.$$

If M is graded, then so are its local cohomology modules.

Example 0.1.28. If $I \subseteq S$ is a homogeneous ideal, then $\mathrm{H}^0_{\mathfrak{m}}(S/I) = I^{\mathrm{sat}}/I$. In particular, $\mathrm{H}^0_{\mathfrak{m}}(M)$ is a finite length S-module. This is true for any finitely generated S-module: $\mathrm{H}^0_{\mathfrak{m}}(M) \subseteq M$ is finitely generated, hence annihilated by a power $\mathfrak{m}^k$. ◁

The higher local cohomology modules are usually not finitely generated, but still Artinian. Hence these modules have a finite *end* $\operatorname{end}(N) \coloneqq \sup\{\, d \in \mathbb{Z} \mid N_d \neq 0 \,\} < \infty$.

Theorem 0.1.29 (Characterizations of regularity)**.** *Let M be a finite graded S-module. The following three numbers coincide:*

(i) $\operatorname{reg}_{\mathrm{CM}}(M) = \max_i(\max\{\, j \mid \beta_{ij}(M) \neq 0 \,\} - i)$;

(ii) $\max_i(\operatorname{end}(\mathrm{H}^i_{\mathfrak{m}}(M)) + i)$;

(iii) $\min\{\, d \in \mathbb{Z} \mid M_{\geq d} \text{ has a linear resolution} \,\}$.

Moreover, these numbers are also an upper bound *on*

(iv) $\operatorname{reg}_{\mathrm{H}}(M) + \operatorname{depth}(M) - 1$.

If M is Cohen–Macaulay, then equality holds.

For a discussion of this result and questions about $\operatorname{reg}_{\mathrm{CM}}$, see [Cha07].

Proposition 0.1.30. *Let M be a finite S-module, then*

$$\operatorname{depth} M = \inf\left\{ i \in \mathbb{Z} \;\middle|\; \mathrm{H}^i_{\mathfrak{m}}(M) \neq 0 \right\} \qquad \dim M = \max\left\{ i \in \mathbb{Z} \;\middle|\; \mathrm{H}^i_{\mathfrak{m}}(M) \neq 0 \right\}.$$

Corollary 0.1.31. *A finitely graded S-module is Cohen–Macaulay if and only if all but a single $\mathrm{H}^d_{\mathfrak{m}}(M)$ vanish (in which case $d = \dim M = \operatorname{depth} M$).*

Lemma 0.1.32. *Let $I \subseteq S$ be a homogeneous ideal.*

(i) *I is saturated if and only if* $\operatorname{depth}_S S/I > 0$.

(ii) *The Hilbert function of a saturated ideal* $\operatorname{hf}_{S/I}$ *is (weakly) increasing.*

Proof. The first statement is just a combination of Example 0.1.28 and Proposition 0.1.30. For the second statement note since a saturated ideal does not have $\mathfrak{m}$ as an associated component, no primary component contains all of S_1. Since $\mathbb{K}$ is infinite, there is hence a non-zerodivisor in S_1, and so $\operatorname{hf}_{S/I}(t) - \operatorname{hf}_{S/I}(t-1) = \operatorname{hf}_{S/(I+\langle \ell \rangle)}(t) \geq 0$. □

0.1.4 Modules of differentials

We briefly discuss the module of relative differentials, for reference see [Sta25, Section 00RM].

Definition 0.1.33. Let B be an A-algebra and M a B-module. An *A-derivation* is an A-linear map $\mathrm{d}\colon B \to M$ that satisfies the Leibniz rule

$$\mathrm{d}(b \cdot b') = b\,\mathrm{d}b' + b'\,\mathrm{d}b \qquad b, b' \in B.$$

The set of A-derivations with values in M is a B-module denoted by $\operatorname{Der}_A(B, M)$.

The *module of relative differentials* is a B-module $\Omega^1_{B/A}$ with a universal A-derivation $\mathrm{d}\colon B \to \Omega^1_{B/A}$ characterized by the universal property $\operatorname{Hom}_B(\Omega^1_{B/A}, M) \cong_B \operatorname{Der}_A(B, M)$ functorially in M.

The module of relative differentials $\Omega^1_{B/A}$ exists and is unique up to unique isomorphism by standard arguments. It enjoys the following basic properties:

Lemma 0.1.34. (i) *If $S \subseteq B$ is multiplicatively closed, then $\Omega^1_{S^{-1}B/A} = S^{-1}\Omega^1_{B/A}$.*

(ii) *For rings $A \to B \to C$ there is a canonical right-exact sequence of C-modules*

$$C \otimes_B \Omega^1_{B/A} \longrightarrow \Omega^1_{C/A} \longrightarrow \Omega^1_{C/B} \longrightarrow 0.$$

(iii) *If $I \subseteq B$ is an ideal, then we have a right-exact sequence of $C := B/I$-modules*

$$I/I^2 \longrightarrow \Omega^1_{B/A} \otimes_B C \longrightarrow \Omega^1_{C/A} \longrightarrow 0,$$

where the first map is given by $b + I^2 \mapsto \mathrm{d}b \otimes 1$.

(iv) *If $A \to A'$ is another algebra and $B' := A' \otimes_A B$, then we have a canonical isomorphism $B' \otimes_B \Omega^1_{B/A} = \Omega_{B'/A'}$ of B'-modules.*

Example 0.1.35. If B is generated by $b_1, \ldots, b_n$ as an A-algebra, then it is easy to see from the Leibniz rule that $\Omega^1_{B/A}$ is generated as a B-module by $\mathrm{d}b_1, \ldots, \mathrm{d}b_n$.

If $B = A[x_1, \ldots, x_n]$, then this surjective map $B^n \to \Omega^1_{B/A}, (f_i)_i \mapsto \sum_i f_i\,\mathrm{d}x_i$ is in fact an isomorphism: It is easy to check that $B^n = \bigoplus_{i=1}^n B\,\mathrm{d}x_i$ satisfies the universal property,

as A-derivations $d\colon B \to M$ are uniquely determined by an (arbitrary!) choice of images $dx_1, \dots, dx_n$.

In the general case of an A-algebra $B = A[x_1, \dots, x_n]/I$, $I = \langle f_1, \dots, f_s \rangle$, one can apply the previous calculation and Lemma 0.1.34(iii) to obtain

$$I/I^2 \longrightarrow B^n \longrightarrow \Omega^1_{B/A} \longrightarrow 0.$$

The image of the first arrow is generated by $\mathrm{d}f_j = \sum_i \frac{\partial f_j}{\partial x_i}\,\mathrm{d}x_i$. In particular, we have a presentation of $\Omega^1_{B/A}$ as the cokernel of the transpose of the Jacobi matrix J_f (as a matrix over B)

$$B^s \xrightarrow{J_f^\mathsf{T}} B^n \longrightarrow \Omega^1_{B/A} \longrightarrow 0. \qquad \triangleleft$$

Lemma 0.1.34(i) shows that the construction of $\Omega^1_{B/A}$ globalizes to the sheaf of relative differentials $\Omega^1_{X/Y}$ on X of a map of varieties (or schemes) $X \to Y$. We restate part of the previous lemma in the language of sheaves for future reference.

Lemma 0.1.36. (i) *If $f\colon X \to Y$, $g\colon Y \to Z$ are morphisms, then we have an exact sequence of $\mathcal{O}_X$-modules*

$$g^*\Omega^1_{Y/Z} \longrightarrow \Omega^1_{X/Z} \longrightarrow \Omega^1_{X/Y} \longrightarrow 0.$$

(ii) *If $f\colon X \to Y$ and $t\colon Y' \to Y$ is another morphism, let $X' \coloneqq X \times_Y Y'$ and let $t'\colon X' \to X$ be the base change. Then we have an identification $(t')^*\Omega^1_{X/Y} = \Omega^1_{X'/Y'}$.*

Example 0.1.35 shows that in the case $X = \operatorname{Spec}\mathbb{K}$, $\Omega^1_{X/\mathbb{K}}$ is the dual of the kernel of the Jacobi matrix, that is, it is the *cotangent sheaf*. In particular, $\Omega^1_{X/k}$ is locally free if and only if X is a smooth $\mathbb{K}$-variety. Moreover, if $f\colon X \to Y$ is a map of smooth varieties, then

$$f^*\Omega^1_{Y/\mathbb{K}} \xrightarrow{J_f^\mathsf{T}} \Omega^1_{X/\mathbb{K}} \longrightarrow \Omega_{X/Y} \longrightarrow 0,$$

so the sheaf of relative differentials enjoys a finite presentation via the cotangent bundles of X and Y, connected by the Jacobi matrix.

Remark 0.1.37. In this sense, the failure of $\Omega^1_{X/Y}$ to be locally free around a point x measures that the morphism $X \to Y$ is singular at $x \in X$. This point of view will be elaborated in Section 4.1.1, where we discuss discriminants. $\triangleleft$

0.2 Gröbner and SAGBI bases

In this section we introduce the fundamental notions of Gröbner bases for polynomial ideals $I \subseteq \mathbb{K}[\underline{x}] \coloneqq \mathbb{K}[x_1, \dots, x_n]$ and their relatives, SAGBI bases for polynomial subalgebras. Standard references for computational algebra are the books [CLO16; Joa17]; for subalgebra bases see [RS90].

0.2.1 Orders, bases, and normal forms

For a set of variables $\underline{x} = \{x_1, \dots, x_n\}$ we denote by $\mathsf{Mon}(\underline{x})$ the set of monomials in $\underline{x}$.

Definition 0.2.1. A (global) *monomial order* $\prec$ is a total order on the set of monomials $\mathsf{Mon}(\underline{x})$ stable under multiplication ($x^\alpha \prec x^\beta$ implies $x^\alpha x^\gamma \prec x^\beta x^\gamma$) and such that $1 = x^0$ is the minimal element.

The *initial term* $\mathrm{in}_\prec(f)$ of a non-zero polynomial $f = \sum_\alpha f_\alpha x^\alpha$ is the non-zero term $f_\alpha x^\alpha$ with the $\prec$-largest monomial. For convenience we define $\mathrm{in}_\prec(0) = 0$.

Given an ideal I or a subalgebra A, its initial ideal (or initial algebra, respectively) is

$$\mathrm{in}_\prec(I) \coloneqq \langle \{ \mathrm{in}_\prec(f) \mid f \in I \} \rangle_{\mathbb{K}[\underline{x}]}, \qquad \mathrm{in}_\prec(A) \coloneqq \mathbb{K}[\{ \mathrm{in}_\prec(f) \mid f \in A \}].$$

Monomial orders are well-orders, hence any nonempty set of monomials has a minimal element. In the following, fix such a monomial order, for example, the *lexicographic order* $\prec_{\mathrm{lex}}$:

$$\begin{aligned} x^\alpha \prec_{\mathrm{lex}} x^\beta \quad \iff \quad & \alpha_1 < \beta_1 \text{ or } (\alpha_1 = \beta_1 \text{ and } \alpha_2 < \beta_2) \\ & \text{or } (\alpha_1 = \beta_1 \text{ and } \alpha_2 = \beta_2 \text{ and } \alpha_3 < \beta_3) \\ & \text{or } \dots \end{aligned}$$

We will always order the variables such that $x_1 > \cdots > x_n$.

Definition 0.2.2. A *Gröbner basis* of an ideal I is a *finite* subset $G \subseteq I$ such that $\mathrm{in}_\prec(I) = \langle \{ \mathrm{in}_\prec(g) \mid g \in G \} \rangle$. A Gröbner basis is *reduced* if $0 \notin G$, the leading coefficients are 1 and no term of $g \in G$ is in $\langle \mathrm{in}_\prec(G \setminus \{g\}) \rangle$.

On the other hand, a *SAGBI basis* of a subalgebra A is *any* subset $S \subseteq A$ such that $\mathrm{in}_\prec(A) = \mathbb{K}[\{ \mathrm{in}_\prec(f) \mid f \in S \}]$.

In both cases, such bases generate I and A respectively, hence the (historical) name "basis". The reduced Gröbner basis of an ideal I is unique and has the smallest number of elements among all Gröbner bases of I (with respect to the fixed monomial order). For more background on Gröbner bases see for example [CLO16, Chapter 2], [Joa17, Chapter 21] or [KR00, Chapter 2].

Lemma 0.2.3. (i) *Every ideal has a Gröbner basis, every subalgebra has a (countable) SAGBI basis.*

(ii) *Gröbner/SAGBI bases generate their respective ideal/subalgebra.*

(iii) *Reduced Gröbner bases of an ideal exist and are unique.*

On the other hand, finite SAGBI bases may not exist; we will come back to this topic in Section 3.5.

Proof. (i) The initial ideal of a polynomial ideal $\mathrm{in}_\prec I$ is finitely generated since $\mathbb{K}[x_1, \dots, x_n]$ is Noetherian. It is generated by initial monomials of polynomials in I, hence there is a finite set $\{g_1, \dots, g_s\} \subseteq I$ with $\langle \mathrm{in}_\prec g_1, \dots, \mathrm{in}_\prec g_s \rangle = \mathrm{in}_\prec I$.

A (rather trivial) SAGBI basis of A is given by A itself (there is no requirement of finiteness). To construct a countable SAGBI basis, pick a polynomial $f_\alpha \in A$ with $\mathrm{in}_\prec f = x^\alpha$ for every $x^\alpha \in \mathrm{in}_\prec A$, this is a countable set.

(ii) Let G be a Gröbner basis and assume $\langle G \rangle \subsetneq I$. Pick $f \in I \setminus \langle G \rangle$ such that $\mathrm{in}_\prec f$ is minimal among $\{ \mathrm{in}_\prec f \mid f \in I \setminus \langle G \rangle \}$. Pick $g \in G$ such that $\mathrm{in}_\prec f = x^\beta \cdot \mathrm{in}_\prec g$, then $f - x^\beta g \in I$ has a smaller leading term, hence is in $\langle G \rangle$. This is a contradiction to $f \notin \langle G \rangle$. The argument for subalgebras is similar.

(iii) Existence: First construct a minimal Gröbner basis G by picking for each minimal generator $x^\alpha \in \mathrm{in}_\prec I$ a generator $g \in I$ with $\mathrm{in}_\prec g = x^\alpha$ with leading coefficient 1. Then iterate through $g \in G$ and successively cancel (non-leading) terms of g that are in $\mathrm{in}_\prec I = \langle \mathrm{in}_\prec G \rangle$.

Uniqueness: It is easy to see from the reducedness property that the leading terms of a reduced Gröbner basis necessarily form a minimal generating set of the initial ideal. Let G, H be reduced Gröbner bases of I, by the previous remark $\#G = \#H =: s$ and we can pair up $g_i \in G, h_i \in H$ such that $\mathrm{in}_\prec g_i = \mathrm{in}_\prec h_i$, $i = 1, \dots, s$. Assume $g_i - h_i \neq 0$, then $\mathrm{in}_\prec(g_i - h_i)$ is a term from either g_i or h_i (or both), hence indivisible by $\mathrm{in}_\prec g_j, \mathrm{in}_\prec h_j$ for all j by reducedness. This implies $g_i - h_i = 0$, and hence $G = H$. □

In order to construct a Gröbner basis for a given ideal, Buchberger's algorithm and its refinements can be used [CLO16, Section 2.7]. In this thesis, we will only use theoretical properties of Gröbner bases, though many of the computer algebra examples we discuss are internally based on Gröbner basis computations.

Definition 0.2.4. Given an ideal $I \subseteq \mathbb{K}[\underline{x}]$, the *normal form* $\mathrm{nf}_\prec(f, I)$ of $f \in \mathbb{K}[\underline{x}]$ against I is the unique polynomial $r \in f + I$ such that no monomial of r is in $\mathrm{in}_\prec(I)$.

Lemma 0.2.5. *The normal form exists and is unique as claimed. It is the unique polynomial in $f + I$ with minimal support.*

Here finite sets of monomials $S, T \subseteq \mathsf{Mon}(\underline{x})$ are compared by checking them "lexicographically" in the total order given by $\prec$:

$$S \prec_{\mathfrak{P}} T \quad \text{iff} \quad \max_\prec(S \setminus T) \prec \max_\prec(T \setminus S) \qquad (\text{Convention: } \max_\prec \emptyset := -\infty).$$

Proof. For existence, modify f by subtracting elements from I whose leading term equals a term in f; this must terminate since it decreases the support in $\prec_{\mathfrak{P}}$ and this is a well-order as well.

If there are two normal forms r, r', then $r - r'$ is a normal form for 0. But clearly 0 is the only normal form for 0. This also shows that $\mathrm{nf}_\prec(-, I)$ is a $\mathbb{K}$-linear map.

If $r \in f + I$ has minimal support, then none of its terms can be canceled by an element in $\mathrm{in}_\prec(I)$. Hence it is the unique normal form, in particular the only polynomial in $f + I$ with minimal support. □

In particular, one can test ideal membership with the normal form: $\mathrm{nf}_\prec(f, I) = 0$ if and only if $f \in I$. A practical way to compute normal forms is the division algorithm, displayed in Figure 1. This iterative process terminates, as $\mathrm{in}_\prec f_{i+1} \prec \mathrm{in}_\prec f_i$ and $\prec$ is a well-order.

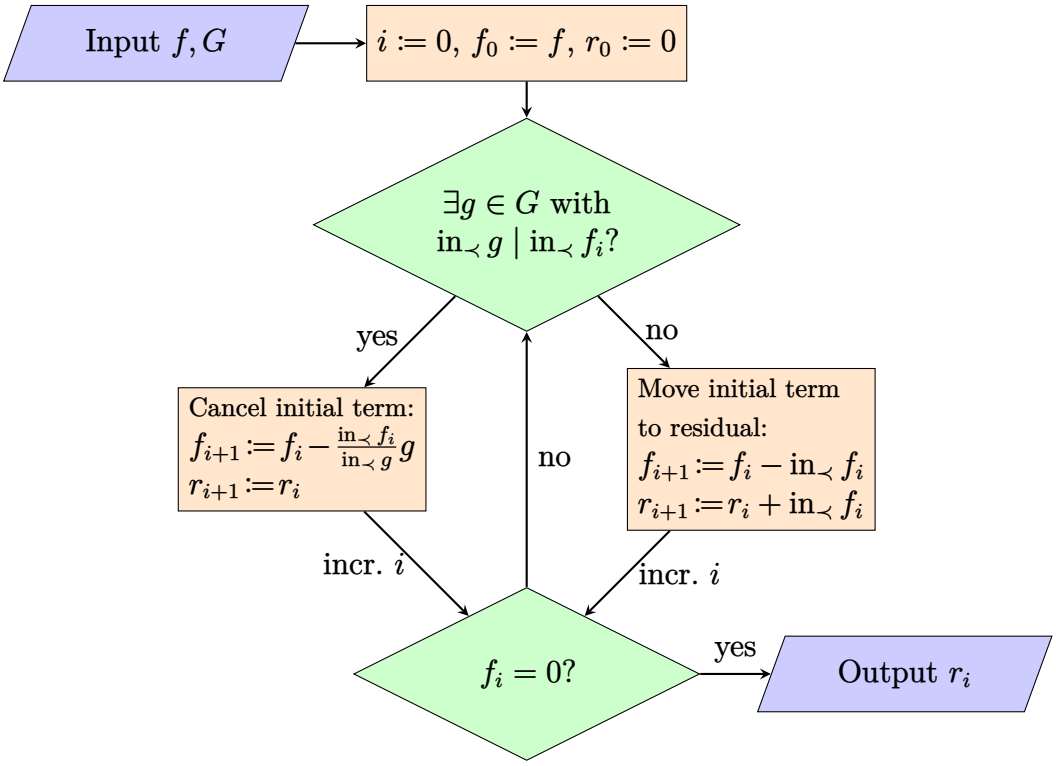


Figure 1: Multivariate division with remainder.

Theorem 0.2.6 ([KR00, Theorem 2.4.1]). *The set G is a Gröbner basis of I if and only if the result of the division algorithm is independent of the choices $g \in G$ for any f. Moreover, in that case the end result is $r = \mathrm{nf}_\prec(f, I)$.*

0.2.2 Initial degenerations

The exposition here follows [ES94, Section 15.8]. Given an ideal $I \subseteq R = \mathbb{K}[\underline{x}]$, the goal is to describe a flat family over $\mathbb{K}[t]$ whose generic fibers $t = \alpha \in \mathbb{K} \neq 0$ are isomorphic to R/I, while the special fiber $t = 0$ is isomorphic to $R/\mathrm{in}_\prec I$. An analogous construction for initial algebras is considered — assuming a finite SAGBI basis exists.

Definition 0.2.7 (Weight vector, initial form). A *weight vector* $0 \neq w \in (\mathbb{R}^n)^\vee$ defines a partial order on $\mathsf{Mon}(\underline{x})$ via $x^\alpha >_w x^\beta$ if $w \cdot \alpha > w \cdot \beta$.

The *initial form* $\mathrm{in}_w f$ of a polynomial $f = \sum_\alpha f_\alpha x^\alpha$ is

$$\sum_{w\cdot\alpha = b} f_\alpha x^\alpha, \qquad b = b_w(f) = \max\{\, w \cdot \alpha \mid f_\alpha \neq 0 \,\} \in \mathbb{R} \cup \{-\infty\}.$$

If I is an ideal, then its w-initial ideal is $\operatorname{in}_w I = \langle \operatorname{in}_w f \mid f \in I \rangle$. A finite subset $G \subseteq I$ is a w-Gröbner basis for I if $\langle G \rangle = I$ and $\operatorname{in}_w I = \langle \operatorname{in}_w G \rangle$.

Example 0.2.8. If $n = 1$ and $w > 0$, then $\operatorname{in}_w f = x^{\deg f}$, while for $w < 0$, $\operatorname{in}_w f$ is the lowest-degree term. Similarly for general n, if $w = (1, \dots, 1)$, then the initial form $\operatorname{in}_w f$ is the sum of all terms whose total degree coincides with the total degree of f. ◁

The initial form can be characterized (over $\mathbb{C}$) via

$$\operatorname{in}_w f = \lim_{t \to 0} \underbrace{t^b f(t^{-w_1} x_1, \dots, t^{-w_n} x_n)}_{=: f_t^w}. \tag{0.1}$$

Over arbitrary fields and $w \in \mathbb{Z}^n$, consider this expression in $\mathbb{K}[\underline{x}][t]$ and then set $t \mapsto 0$.

As the example shows, initial forms are not necessarily single terms/monomials. However, a given monomial order can be approximated in the following sense:

We say that w is *compatible* with a given monomial order $\prec$ on a set $\mathcal{S} \subseteq \mathbb{K}[\underline{x}]$ if $\operatorname{in}_\prec f = \operatorname{in}_w f$ for all $f \in \mathcal{S}$.

Proposition 0.2.9. (i) *[ES94, Exercise 15.12] For every finite set of polynomials $\mathcal{S}$ there exists a weight vector $w \in \mathbb{N}^n$ such that w is compatible with $\prec$ on $\mathcal{S}$.*

(ii) *[ES94, Proposition 15.16] Let G be a Gröbner basis of an ideal I with respect to $\prec$. There exists a finite set $\mathcal{S}$ of binomials such that if $w \in \mathbb{Z}^n$ is compatible with $\prec$ on $\mathcal{S}$, then G is a w-Gröbner basis and $\operatorname{in}_w I = \operatorname{in}_\prec I$.*

This proposition allows us to describe the desired flat family as follows: Let w be a weight vector and consider $\tilde{I} := \langle f_t^w \mid f \in I \rangle \subseteq R[t]$, where f_t^w is constructed as in (0.1). A generating set of $\tilde{I}$ is given by $\{ g_t^w \mid g \in G \}$ where G is any w-Gröbner basis for I [ES94, Exercise 15.25].

Theorem 0.2.10 ([ES94, Theorem 15.17])**.** *For every ideal I and every weight vector w, the $\mathbb{K}[t]$-algebra $R[t]/\tilde{I}$ is free (and hence flat). We have*

$$R[t]/\tilde{I} \otimes_{\mathbb{K}[t]} \mathbb{K}[t]/\langle t - \kappa \rangle = R[t]/\langle \tilde{I}, t - \kappa \rangle \cong \begin{cases} R/I & \kappa \in \mathbb{K}^\times \\ R/\operatorname{in}_w I & \kappa = 0. \end{cases}$$

We now briefly sketch an analogous construction in the context of subalgebras. Let $A \subseteq R$ be a subalgebra and for a weight vector w, let $\tilde{A} := \mathbb{K}[f_t^w \mid f \in A] \subseteq R[t]$. If $K \subseteq A$ is a SAGBI basis for $\prec$ and if $<_w$ represents $\prec$ on K, then $\tilde{A}$ is generated by g_t^w, $g \in K$. In this case, $\tilde{A} \leftarrow \mathbb{K}[t]$ is a flat family whose general fibers $t = \kappa \in \mathbb{K}^\times$ are isomorphic to A and whose special fiber $t = 0$ is the initial algebra $\operatorname{in}_\prec A$. We will see this in action for the homogeneous coordinate ring of the Grassmannian in Section 2.3.2. For more details and a vast generalization to the context of Khovanskii bases, see [And13].

Remark 0.2.11. In the homogeneous case, the flatness can also be deduced from the fact that a family of projective schemes is flat if and only if the Hilbert polynomial is constant, together with the observation that an algebra/ideal and its initial ideal/algebra (with respect to a monomial order) have the same Hilbert function. ◁

Geometrically, initial degenerations of ideals degenerate a variety into a union of coordinate subspaces (potentially with multiplicities and embedded components). Analogously, initial degenerations of subalgebras degenerate a variety into a toric variety $\operatorname{Spec}(\operatorname{in}_{\prec} A)$ (if a finite SAGBI basis exists).

0.3 Linear sections of varieties

In this section we collect some useful results about projective subspaces of $\mathbb{P}^n$ and their intersection with projective varieties.

0.3.1 A variant of Bézout's theorem

Bézout's theorem, in its basic form, states that if $X_1, \dots, X_n \subseteq \mathbb{P}^n$ are hypersurfaces and if $X_1 \cap \cdots \cap X_n$ is a finite set (or rather, scheme), then

$$\deg X_1 \cap \cdots \cap X_n = \deg X_1 \cdots \deg X_n.$$

We will formulate a vast generalization in the language of intersection theory in Theorem 0.4.7 below. However, many "Bézout-like" results require at least dimensional transversality of the intersection, that is, all intersections have expected dimensions (for a more precise definition see below).

To remedy this, we also formulate a version using the following variant of degree, which is useful in algebraic complexity theory [Bür98, Section 8.2].

Definition 0.3.1 (Component-wise degree)**.** Let X be a locally closed subset of $\mathbb{P}^n$ and $X_1, \dots, X_k$ its irreducible components. The *component-wise degree* of X is

$$\operatorname{c.deg} X \coloneqq \sum_{i=1}^{k} \deg \overline{X_i}.$$

For closed subsets $X \subseteq \mathbb{P}^n$, $\deg X \leq \operatorname{c.deg} X$ with equality iff X is equidimensional. With this notation we can state a variant of Bézout's theorem allowing for improper intersections [Ful98, Example 8.4.6] or [Bür98, Exercise 8.18].

Theorem 0.3.2 (Component-wise Bézout Inequality)**.** *Let $X, Y \subseteq \mathbb{P}^n$ be locally closed subsets, then*

$$\operatorname{c.deg}(X \cap Y) \leq \operatorname{c.deg} X \cdot \operatorname{c.deg} Y.$$

Applying this to the intersection of a variety and a complementary linear space yields the following useful degree bound.

Corollary 0.3.3. *If $X \subseteq \mathbb{P}^n$ is a variety of codimension c and $\Lambda \subseteq \mathbb{P}^n$ a linear space of dimension c, then the number of irreducible components of $X \cap \Lambda$ is* at most *the degree of X.*

0.3.2 Secant varieties

Let $\mathbb{K}$ be algebraically closed. For a set of points $p_1, \dots, p_r \in \mathbb{P}^n$, denote by $\langle p_1, \dots, p_r \rangle_{\mathbb{P}}$ the projective subspace spanned by the points.

Definition 0.3.4 (Secant variety)**.** Let $X \subseteq \mathbb{P}^n$ be a projective variety. The r-th secant variety $\sigma_r X \subseteq \mathbb{P}^n$ is

$$\sigma_r X \coloneqq \overline{\bigcup_{p_1,\dots,p_r \in X} \langle p_1, \dots, p_r \rangle_{\mathbb{P}}} \subseteq \mathbb{P}^n.$$

Example 0.3.5. Let $V^{n,d} = \nu_d(\mathbb{P}^n) \subseteq \mathbb{P}^{\binom{n+d}{n}-1} = \mathbb{P}(\operatorname{Sym}^d \mathbb{K}^{n+1})$ be the n-dimensional Veronese variety. Points in $V^{n,d}$ correspond to simple symmetric tensors $v^{\otimes d}$, or, if $\operatorname{char} \mathbb{K} > d$, powers of linear forms $\frac{1}{d!} L^d$. Here $\sigma_r(V^{n,d})$ corresponds to symmetric tensors of border rank at most r, that is, limits of sums of r d-th powers [Ber+18, Section 2.1.3]. We will study the problem of finding such decompositions in more detail in Section 1.2. ◁

In the definition of the secant variety, we take the union of all secant linear spaces spanned by (at most) r points. There is no harm in parametrizing the same linear space multiple times, but it turns out that for most multisecant linear spaces, there is a *unique* set of points on X spanning them.

Theorem 0.3.6 (Multisecant lemma [Rus16, Proposition 1.4.3])**.** *Let $X \subseteq \mathbb{P}^n$ be an irreducible projective variety and assume $\operatorname{char} \mathbb{K} = 0$. Let $1 \leq r \leq \operatorname{codim} X$, then there is a dense open set $U \subseteq X^{\times r}$ such that for all $(p_1, \dots, p_r) \in U$ one has*

$$\langle p_1, \dots, p_r \rangle_{\mathbb{P}} \cap X = \{p_1, \dots, p_r\}^{\mathrm{red}}.$$

Note that for $r = \operatorname{codim} X + 1$, most r-secants will intersect X in $\deg(X)$ many points.

0.3.3 Linear non-degeneracy

Definition 0.3.7. A projective subscheme $X \subseteq \mathbb{P}^n$ is *non-degenerate* if it is not contained in a hyperplane, equivalently if $\mathrm{I}(X)_1 = 0$.

The following lemma states that a hyperplane section of a connected non-degenerate variety is again non-degenerate.

Lemma 0.3.8 ([BL13, Lemma 8.1])**.** *Let $X \subseteq \mathbb{P}^n$ be a connected variety and $H \subseteq \mathbb{P}^n$ a hyperplane not containing any irreducible component of X. If X is non-degenerate in $\mathbb{P}^n$, then the scheme $X \cap H$ is non-degenerate in $H \cong \mathbb{P}^{n-1}$.*

The conclusion can fail if the variety is not connected (pick X as two skew lines in $\mathbb{P}^3$ and $H \subseteq \mathbb{P}^2$ a general plane) or if the intersection is considered set-theoretically (pick a tangent line H of a plane conic $X \subseteq \mathbb{P}^2$).

The following proposition is a higher codimension analogue to Lemma 0.3.8; it appeared as [GKT25, Proposition A.3] with the additional assumption $\operatorname{char} \mathbb{K} = 0$.

Proposition 0.3.9. *Let $X \subseteq \mathbb{P}^N$ be a non-degenerate irreducible variety of dimension c. Let Λ be a linear space with $\operatorname{codim}\Lambda = c$. Suppose $X \cap \Lambda$ is a set of reduced points. Then $X \cap \Lambda$ is non-degenerate in Λ.*

We present a proof different from [GKT25, Proposition A.3], independent of the characteristic of the base field.

Proof. Let $V_1, V_2 \subseteq \mathbb{P}^n$ be irreducible varieties with $V_1 \cup V_2$ non-degenerate. By a result of Lazarsfeld [Ful98, Example 12.3.5], in the notation of Definition 0.3.1,

$$\operatorname{c.deg}(V_1 \cap V_2) \leq \deg(V_1)\deg(V_2) - e, \qquad e \coloneqq \dim(V_1 \cap V_2) - \dim V_1 - \dim V_2 + n. \tag{0.2}$$

Assume that $X \cap \Lambda$ is degenerate in Λ, then there is a hyperplane $L \subseteq \Lambda$ with $X \cap \Lambda = X \cap L$. We apply (0.2) to $V_1 = X$, $V_2 = L$ to arrive at a contradiction:

On the one hand, by reducedness of $X \cap \Lambda$ we know that the intersection $X \cap \Lambda$ is transversal, so it consists of $\deg X$ many points. So the left hand side of (0.2) is $\deg X$. On the other hand, the excess dimension is $e = 1$ here, so the right hand side reads $\deg X - 1$, which violates the inequality. □

0.4 Intersection theory

In this section we introduce basic definitions and results from intersection theory. The standard reference is [Ful98], for an introduction see [EH16].

0.4.1 Chow rings

Let $X \supset Y$ be irreducible varieties, Y of codimension 1 in X. In this setting, the ring $R \coloneqq \mathcal{O}_{X,Y}$ is a Noetherian local domain of dimension 1. Thus, for any non-zero $a \in R$, the quotient $R/\langle a\rangle$ is zero-dimensional (or the zero ring), and hence has finite length as a R-module. If $\mathfrak{m}_R = \langle u\rangle$ is principal (so R is a discrete valuation ring) and $a = u^k$, then $\operatorname{length}(R/\langle a\rangle) = k$. In the non-regular case, the length is still a useful measure of multiplicity:

Definition 0.4.1. Let $f \in K(X)^\times$ be a rational function on X, and write $f = a/b$ with $a, b \in R \coloneqq \mathcal{O}_{X,Y}$. The *order of vanishing* of f along Y is

$$\operatorname{ord}_Y(f) = \operatorname{length}_R(R/\langle a\rangle) - \operatorname{length}_R(R/\langle b\rangle).$$

This is well-defined and provides a surjective group homomorphism $K(X)^\times \to \mathbb{Z}$. Moreover, for a fixed $f \in K(X)^\times$, there are only finitely many codimension 1 subvarieties Y with $\operatorname{ord}_Y(f) \neq 0$ [Ful98, Appendix A.3 & B.4.3]. Thus, the following definition makes sense:

Definition 0.4.2 (Cycles, principal divisor, Chow group)**.** Let X be an irreducible variety. The group of *k-cycles* in X is the free abelian group $Z_k(X)$ generated by irreducible subvarieties of dimension k

$$Z_k(X) \coloneqq \bigoplus_{Y \subsetneq X \text{ dim } k} \mathbb{Z} \cdot Y.$$

The *principal divisor* associated to $f \in K(X)^\times$ is

$$\operatorname{div}(f) \coloneqq \sum_{Y \subsetneq X \text{ codim } 1} \operatorname{ord}_Y(f) \cdot Y \in Z_{\dim X - 1}(X)$$

A cycle $A \in Z_k(X)$ is rationally equivalent to 0 if there exist irreducible subvarieties $X_1, \dots, X_m \subseteq X$ of dimension $k+1$ such that

$$A = \operatorname{div}(f_1) + \cdots + \operatorname{div}(f_m), \qquad f_i \in K(X_i)^\times.$$

The k-th *Chow group* $\mathrm{A}_k(X)$ is the quotient of $Z_k(X)$ modulo the subgroup of cycles rationally equivalent to 0. Write $\mathrm{A}(X) \coloneqq \bigoplus_{k=0}^{\dim X} \mathrm{A}_k(X)$.

If $Y \subseteq X$ is a closed subscheme with irreducible components $Y_1, \dots, Y_m$, then we associate to Y the Chow class

$$[Y] \coloneqq \sum_{i=1}^{m} \operatorname{length}_{\mathcal{O}_{Y,Y_i}}(\mathcal{O}_{Y,Y_i}) \cdot Y_i \in \mathrm{A}(X).$$

Example 0.4.3. The Chow group $\mathrm{A}_{\dim X - 1}(X)$ is also known as the *(Weil) divisor class group*. If X is smooth (or locally factorial), then it is naturally isomorphic to the Picard group of line bundles on X up to isomorphism.

For example, $\mathrm{A}_{n-1}(\mathbb{P}^n) \cong \mathbb{Z}$, associating to a hypersurface its degree, and all hypersurfaces of the same degree are rationally equivalent.

More generally, $\mathrm{A}_k(\mathbb{P}^n) = \mathbb{Z} \cdot [L] \cong \mathbb{Z}$, where L is the class of any linear subspace of dimension k. The integer d for $[A] = d \cdot [L]$ is the degree of the k-cycle A, which recovers our previous notion of degree: If $Y \subseteq \mathbb{P}^n$ is a subscheme of equidimension k, then $[Y] = (\deg Y) \cdot [L]$ ◁

A main goal of intersection theory is to provide a useful intersection product on cycle classes. If two subvarieties $Y_1, Y_2 \subseteq X$ intersect transversally, then one may try to define $[Y_1] \cdot [Y_2] = [Y_1 \cap Y_2]$. This is made formal by the following:

Definition 0.4.4. Two irreducible subvarieties $Y_1, Y_2 \subseteq X$ *intersect properly* if $Y_1 \cap Y_2$ has dimension $\dim Y_1 + \dim Y_2 - \dim X$ (this number is always a lower bound). They *intersect (generically) transversally* if they intersect properly and each irreducible component of $Y_1 \cap Y_2$ is (generically) smooth.

Two cycles $A = \sum_i a_i A_i$, $B = \sum_j b_j B_j$ intersect properly resp. (generically) transversally if all pairs A_i, B_j intersect properly resp. (generically) transversally.

Theorem 0.4.5 (Moving lemma)**.** *Let A, B be cycles in a smooth variety X.*

(i) *There exists a rationally equivalent cycle* $[A'] = [A]$ *such that* A' *and* B *intersect properly. In* $\operatorname{char} \mathbb{K} = 0$ *they can be chosen to intersect generically transversally.*

(ii) *If* A', B *intersect generically transversally, then the class* $[A' \cap B] \coloneqq \sum_{ij} a'_i b_j [A'_i \cap B_j]$ *is independent of* A'.

The moving lemma can be used to construct a ring structure on the Chow group $\mathrm{A}(X)$. For a smooth irreducible variety X we use the notation $\mathrm{A}^d(X) \coloneqq \mathrm{A}_{\dim X - d}(X)$.

Theorem 0.4.6. *Let* X *be a smooth projective variety. There exists a graded ring structure on* $\mathrm{A}(X) = \bigoplus_{d=0}^{\dim X} \mathrm{A}^d(X)$ *given by* $[A] \cdot [B] = [A \cap B]$ *for generically transversally intersecting cycles* A, B.

Given a proper intersection of subvarieties $A \cap B$, one might hope that the class of $[A] \cdot [B]$ is always given by $[A \cap B]$ (here $A \cap B$ is understood to be scheme-theoretic intersection, and $[-]$ as in Definition 0.4.2). This is not true in general (see for example [EH16, Example 2.6]), and a corrected version is given by *Serre's intersection formula* [EH16, Theorem 2.7]. However, it is true under mild hypotheses [Ful98, Proposition 8.2], and can be seen as a refinement of Theorem 0.1.24.

Theorem 0.4.7 (Generalized Bézout's theorem)**.** *Let* X *be a smooth variety and consider* $Y_1, \dots, Y_m \subseteq X$ *irreducible subschemes. Assume that the* Y_i *intersect properly, that is,*

$$\dim(Y_1 \cap \cdots \cap Y_m) = \dim X - \sum_i \operatorname{codim}_X Y_i.$$

If for each irreducible component $Z \subseteq Y_1 \cap \cdots \cap Y_m$*, all* Y_i *are generically Cohen–Macaulay along* Z *(that is,* $\mathcal{O}_{Y_i, Z}$ *is Cohen–Macaulay), then*

$$[Y_1 \cap \cdots \cap Y_m] = [Y_1] \cdots [Y_m].$$

In particular, this is always satisfied when the Y_i are Cohen–Macaulay, for example hypersurfaces in smooth varieties.

We are mainly working in the Chow ring of projective and bi-projective space, their ring structure is closely related to Bézout's theorem:

Theorem 0.4.8. (i) $\mathrm{A}(\mathbb{P}^n) \cong \mathbb{Z}[\alpha] / \langle \alpha^{n+1} \rangle$*, generated by a hyperplane class* $\alpha = [H] \in \mathrm{A}^1(\mathbb{P}^n)$.

(ii) $\mathrm{A}(\mathbb{P}^m \times \mathbb{P}^n) = \mathbb{Z}[\alpha, \beta] / \langle \alpha^{m+1}, \beta^{n+1} \rangle$*, generated by* $\alpha = [H_1 \times \mathbb{P}^n]$ *and* $\beta = [\mathbb{P}^m \times H_2]$*,* $H_1 \subseteq \mathbb{P}^m$*,* $H_2 \subseteq \mathbb{P}^n$ *hyperplanes.*

For a proof see [EH16, Theorem 2.10] or [Ful98, Example 8.4.2].

0.4.2 Thom–Porteous formula

Theorem 0.4.7 allows for easy computation of the class of the intersection of several hypersurfaces in a smooth variety X — provided it has the expected dimension. A similar statement can be made for determinantal schemes which have the expected dimension.

Let $\mathcal{E}, \mathcal{F}$ be vector bundles of rank e, f on the smooth variety X, and $\varphi\colon \mathcal{E} \to \mathcal{F}$ a homomorphism of vector bundles. Equivalently, $\mathcal{E}, \mathcal{F}$ are locally free sheaves, and on an open cover we can locally identify φ with a matrix $\mathcal{O}_X(U)^e \to \mathcal{O}_X(U)^f$. The locus $D_r(\varphi) \subseteq X$ where φ has rank $\leq r$ is naturally a closed subscheme locally defined by the $(r+1)$-minors of φ. By Theorem 0.1.25, every irreducible component of $D_r(\varphi)$ has codimension at least $(e-r)(f-r)$, and if equality holds, then $D_r(\varphi)$ is Cohen–Macaulay. The content of Thom–Porteous formula is an expression of the Chow class $[D_r(\varphi)] \in \mathrm{A}^{(e-r)(f-r)}(X)$ in terms of invariants of $\mathcal{E}$ and $\mathcal{F}$.

Definition 0.4.9 (Chern class)**.** There exist a unique way to assign to each vector bundle $\mathcal{F}$ on a smooth variety X a class $c(\mathcal{F}) \coloneqq c_0(\mathcal{F}) + \cdots + c_{\dim X}(\mathcal{F})$, $c_i(\mathcal{F}) \in \mathrm{A}^i(X)$, such that:

(i) If D is a divisor, then $c(\mathcal{O}_X(D)) = 1 + [D]$.

(ii) *Whitney's formula*: For every short exact sequence

$$0 \longrightarrow \mathcal{E} \longrightarrow \mathcal{F} \longrightarrow \mathcal{G} \longrightarrow 0$$

one has $c(\mathcal{F}) = c(\mathcal{E})c(\mathcal{G}) \in \mathrm{A}(X)$.

(iii) If $f\colon X \to Y$ is a morphism of smooth varieties, then $f^*(c(\mathcal{F})) = c(f^*(\mathcal{F}))$.

The sum $c(\mathcal{F})$ is the *total Chern class* of $\mathcal{F}$.

Here $f^*\colon \mathrm{A}(Y) \to \mathrm{A}(X)$ is the pullback of Chow groups, X, Y smooth. It satisfies $f^*([A]) = [f^{-1}(A)]$ if $A \subseteq Y$ is a subvariety such that $f^{-1}(A)$ is generically reduced and $\operatorname{codim}_X f^{-1}(A) = \operatorname{codim}_Y A$.

Example 0.4.10. Let $X = \mathbb{P}^n$ and $\mathcal{F} = \bigoplus_{i=1}^f \mathcal{O}_{\mathbb{P}^n}(d_i)$, then by the Whitney sum formula

$$c(\mathcal{F}) = (1 + d_1[H]) \cdots (1 + d_f[H]), \qquad c_i = e_i(d_1, \ldots, d_f)[H]^i.$$

Here e_i is the i-th elementary symmetric polynomial. ◁

Since $c_0(\mathcal{F}) = 1$, any total Chern class is an invertible element in the ring $\mathrm{A}(X)$, so quotients of the form $c(\mathcal{F})/c(\mathcal{E})$ make sense in $\mathrm{A}(X)$.

Definition 0.4.11 (Segre class)**.** The *Segre class* of a bundle $\mathcal{F}$ is $s(\mathcal{F}) = 1/c(\mathcal{F})$.

To state the formula for the class of the degeneracy locus $[\Sigma_r(\varphi)]$, we need the following notation. For an element a in the graded ring $\mathrm{A}(X)$ with homogeneous components a_i, define the matrix

$$\mathbb{D}^e_f(a) = \begin{bmatrix} a_f & a_{f+1} & \cdots & a_{f+e-1} \\ a_{f-1} & a_f & \ddots & a_{f+e-2} \\ \vdots & \ddots & \ddots & \vdots \\ a_{f-e+1} & a_{f-e+2} & \cdots & a_f \end{bmatrix} \in \mathrm{A}(X)^{e \times e}.$$

Observe that the determinant of such a matrix is homogeneous of degree ef.

Theorem 0.4.12 (Thom–Porteous)**.** *Let $\varphi\colon \mathcal{E} \to \mathcal{F}$ be a morphism of vector bundles on the smooth variety X and assume that $\dim D_r = \dim X - (e-r)(f-r)$, then its Chow class is given as*

$$[D_r(\varphi)] = \det \mathbb{D}^{e-r}_{f-r}\left(\frac{c(\mathcal{F})}{c(\mathcal{E})}\right).$$

For a proof see [Ful98, Chapter 14] or [EH16, Chapter 12]. We are mainly interested in the case where $r = \min\{e, f\} - 1$. In many cases this gives an alternative geometric description of Chern classes.

Corollary 0.4.13. *Under the same hypotheses, if $r = e - 1 < f$, then $D_r(\varphi)$ describes the locus of non-injectivity of φ. Its class is given by*

$$[D_r(\varphi)] = \left\{\frac{c(\mathcal{F})}{c(\mathcal{E})}\right\}^{f-e+1},$$

where $\{-\}^k$ is the degree k part. If $\mathcal{E} = \mathcal{O}_X^e$ is trivial, then $[D_r(\varphi)] = c_{f-e+1}(\mathcal{F})$.

Example 0.4.14. If $X = \mathbb{P}^n$, $\mathcal{E} = \mathcal{O}_X$, $\mathcal{F} = \bigoplus_{i=1}^f \mathcal{O}_{\mathbb{P}^n}(d_i)$ and $\varphi = (\varphi_1, \dots, \varphi_f)^\mathsf{T}$, $\varphi_i \in \mathrm{H}^0(\mathcal{O}_{\mathbb{P}^n}(d_i)) = S_{d_i}$, then the Thom–Porteous formula recovers the proper intersection of f hypersurfaces in projective space, hence Bézout's theorem. ◁

0.4.3 Reducedness of determinantal schemes

Our main application of the Thom–Porteous formula is to count points in projective space. For this purpose it is useful to have a criterion for the scheme structure of the degeneracy locus $D_r(\varphi) \subseteq X$ to be reduced. As this is a local question, we only state it in the affine case.

Let X be an affine scheme with coordinate ring R and

$$\varphi = \begin{bmatrix} \varphi_{10} & \cdots & \varphi_{1n} \\ \vdots & \ddots & \vdots \\ \varphi_{m0} & \cdots & \varphi_{mn} \end{bmatrix} \in R^{m\times(n+1)}, \quad m \geq n+1$$

a tall matrix over R. The *kernel vector incidence* is the scheme defined by

$$\mathcal{K} \coloneqq \left\{ (x, [v]) \;\middle|\; \varphi(x)\cdot v = \left[\textstyle\sum_{j=0}^n \varphi_{ij}(x)\cdot v_j\right]_{i=1}^m = 0 \right\} \subseteq X \times \mathbb{P}^n.$$

The scheme-theoretic image $Z \subseteq X$ of $\mathcal{K}$ under the projection to X has the same underlying set as $D_n(\varphi)$. In particular, $\sqrt{I_{n+1}(\varphi)} = \sqrt{\mathrm{I}(Z)}$. The following lemma gives a sufficient criterion for this to be already an equality without radicals.

Proposition 0.4.15. *If for all $x \in X$ we have $\operatorname{rank}\varphi(x) \in \{n, n+1\}$, then $I_{n+1}(\varphi) = \mathrm{I}(Z)$. In particular, if $\mathcal{K}$ is reduced, then so is $D_n(\varphi)$.*

Proof. By the assumption, for every $x \in X$ there is a n-minor of φ *not* vanishing. Covering X by the complements of the vanishing sets of these minors and arguing locally, we may assume without loss of generality that the submatrix $N = [\varphi_{ij}]_{i,j=1}^{n}$ is invertible. By multiplying from the left by $N^{-1} \oplus \mathrm{Id}_{m-n}$ we can assume $N = \mathrm{Id}_n$; this doesn't change the structure of the incidence nor the ideal of minors (for example by the Cauchy–Binet formula). So we are in the situation

$$\varphi \cdot v = \left[\begin{array}{ccc} u_1 & 1 & \\ \vdots & & \ddots \\ u_n & & 1 \\ \hline w_1 & & \\ \vdots & & C \\ w_{m-n} & & \end{array}\right] \cdot \begin{pmatrix} v_0 \\ \vdots \\ v_n \end{pmatrix} \stackrel{!}{=} 0$$

Let $\tilde{v} = (v_1, \dots, v_n)^\mathsf{T}$, then the ideal of $\mathcal{K}$ is given by

$$\langle v_0 u + \tilde{v}, C\tilde{v} + v_0 w \rangle \subseteq R[v_0, \dots, v_n],$$

potentially after saturation with $\langle v_0, \dots, v_n \rangle$. This can be rearranged to $\langle v_0 u + \tilde{v}, v_0(Cu - w) \rangle$. On the chart $v_0 = 1$ this is $\langle u + \tilde{v}, Cu - w \rangle$, on the hyperplane $v_0 = 0$ this is $\langle v_1, \dots, v_n \rangle$ having no solution. So $\mathcal{K}$ is supported in $X \times \mathbb{A}^n$ and the ideal of Z is

$$\mathrm{I}(Z) = \langle \tilde{v} + u, Cu - w \rangle_{R[\tilde{v}]} \cap R = \langle Cu - w \rangle_R.$$

The first $(n+1)$-minor of φ is

$$\det \begin{bmatrix} u_1 & 1 & & \\ \vdots & & \ddots & \\ u_n & & & 1 \\ w_1 & c_{11} & \cdots & c_{1n} \end{bmatrix} = (-1)^{n+1}(c_{11}u_1 + \cdots + c_{1n}u_n - w_1).$$

Analogous expansions for other rows show the inclusion $\mathrm{I}(Z) \subseteq I_{n+1}(\varphi)$.

The other inclusion holds (without conditions on the rank) because it holds in the "universal" case

$$R = \mathbb{C}[\{u_i\}^n, \{w_j\}^{m-n}, \{c_{ij}\}^{(m-n)\times n}], \qquad \varphi = \left[\begin{array}{ccc} u_1 & 1 & \\ \vdots & & \ddots \\ u_n & & 1 \\ \hline w_1 & & \\ \vdots & & C \\ w_{m-n} & & \end{array}\right].$$

There $\mathrm{I}(Z) = \langle Cu - w \rangle_R$ is a prime ideal ($R/\mathrm{I}(Z) \cong \mathbb{C}[\{u_i\}^n, \{c_{ij}\}^{(m-n)\times n}]$) and hence

$$I_{n+1}(\varphi) \subseteq \sqrt{I_{n+1}(\varphi)} = \sqrt{\mathrm{I}(Z)} = \mathrm{I}(Z). \qquad \square$$

Chapter 1

Hilbert Functions of Chopped Ideals

Contents

> Aim for the chopping block. If you aim for the wood, you will have nothing. Aim past the wood, aim through the wood; aim for the chopping block.
>
> ANNIE DILLARD [Dil90]

In this first chapter, we study the problem of solving a homogeneous zero-dimensional polynomial system $Z = \{\mathcal{F} = 0\} \subseteq \mathbb{P}^n$ using the eigenvalue method. This method requires knowledge about the Hilbert function of the ideal $J \subseteq S = \mathbb{C}[x_0, \dots, x_n]$ generated by the given equations, in particular when $\mathrm{hf}_{S/J}(j) = \deg Z$. Motivated by an application to symmetric tensor decomposition, we are interested in the case where J is the *chopped ideal* $J = \langle \mathrm{I}(Z)_d \rangle_S$ of a general set of r points Z. We present a precise conjecture for the Hilbert function of S/J below in Expected Syzygy Conjecture 1.3.10, and prove it in many cases:

Main Results 1 (1.2.3, 1.3.13, 1.5.4, 1.5.6, 1.6.1, 1.6.5)**.** *Conjecture 1.3.10 is true in the following cases:*

(i) $r_{\max} = \binom{n+d}{n} - (n+1)$ *for all* d *in all dimensions* n*.*

(ii) *In the plane for* $r_{\min} = \frac{1}{2}(d+1)^2$ *when* d *is odd.*

(iii) *In a large number of individual cases in low dimension (Table 1.1).*

The Hilbert regularity of J *is bounded above by* $\mathrm{reg}_{\mathrm{H}}(S/J) \leq nd - (n+1)$*. Our method is relevant for symmetric tensor decomposition using apolarity theory.*

This chapter is based on the paper *Hilbert Functions of Chopped Ideals* [GKT25] coauthored with Fulvio Gesmundo and Simon Telen. It is structured as follows:

In Section 1.1 we start by introducing the eigenvalue method for polynomial system solving. In Section 1.2, we study the computational problem of symmetric tensor decomposition via the apolarity method. This naturally leads us to consider Hilbert functions of chopped ideals of general sets of points in $\mathbb{P}^n$. With this in mind, in Section 1.3 we look at conjectures about homological properties of general sets of points. We derive the Expected Syzygy Conjecture 1.3.10 (ESC) about the Hilbert function of the chopped ideal, and verify it using computer algebra in a large number of cases. In Section 1.4, we take a closed look at the saturation gap of chopped ideals, study when it is finite, and relate it to the regularity of the ideal. In Section 1.5, we prove ESC for points in $\mathbb{P}^2$ for two infinite families of points, illustrating these methods in the example of 18 points in the plane. Finally, in Section 1.6 we prove ESC for $r = \binom{n+d}{n} - (n+1)$ points in $\mathbb{P}^n$ using a liaison argument, and provide a general upper bound on the length of the saturation gap.

Throughout the chapter $\mathbb{K}$ is an algebraically closed field, sometimes specified to be of characteristic 0, and $S = \mathbb{K}[x_0, \dots, x_n]$.

1.1 Background: The eigenvalue method for polynomial systems

In the first section, we introduce the eigenvalue method for polynomial system solving. This method transforms the nonlinear system of polynomial equation into a large eigenvalue problem, which can then be solved using numerical linear algebra techniques. We present the affine and the projective version, introducing the saturation set as a measure of complexity for this approach. A standard reference for the eigenvalue method is [DE05, Chapter 2 & 3], for a more recent treatment see also [Tel20].

1.1.1 Affine version

Let $R = \mathbb{K}[x_1, \dots, x_n]$ and let $I \subseteq R$ be zero-dimensional ideal, that is, $\mathbb{V}(I) \subseteq \mathbb{K}^n$ is a finite set, equivalently $\dim_{\mathbb{K}} R/I =: \ell < \infty$. Pick $g \in R$ and consider the $\mathbb{K}$-linear multiplication map

$$M_g \colon R/I \to R/I \qquad f \mapsto g \cdot f.$$

Let $\mathbb{V}(\sqrt{I}) = \{z_1, \dots, z_r\}$ be the solution set of I, neglecting multiplicities. Let

$$\mathrm{ev}_{z_i} \colon R/I \to \mathbb{K}, \qquad f \mapsto f(z_i)$$

be the evaluation maps, well-defined as $z_i \in \mathbb{V}(I)$.

Theorem 1.1.1 (Affine eigenvalue-eigenvector theorem)**.** *Let $g \in R$.*

(i) *The eigenvalues of the endomorphism M_g are $g(z_1), \dots, g(z_r)$ (possibly with repetition or multiplicites).*

(ii) *More precisely, the dual map $M_g^\vee \colon (R/I)^\vee \to (R/I)^\vee$ has eigenvalue-eigenvector pairs*

$$(g(z_i), \mathrm{ev}_{z_i} \colon f \mapsto f(z_i)), \qquad i = 1, \dots, r.$$

(iii) *The characteristic polynomial of M_g factors as*

$$\mathrm{charpol}_{M_g}(t) = \prod_{i=1}^{r} (t - g(z_i))^{\ell_i},$$

where $\ell_i = \dim_{\mathbb{K}} (R/I)_{z_i}$ is the multiplicity of I at the points z_i.

Proof. To prove (ii) and one inclusion in (i), we verify that ev_{z_i} is an eigenvector of $M_g^\vee$ with eigenvalue $g(z_i)$. Indeed, for any $f \in R/I$,

$$(M_g^\vee(\mathrm{ev}_{z_i}))(f) = (\mathrm{ev}_{z_i} \circ M_g)(f) = \mathrm{ev}_{z_i}(g \cdot f) = g(z_i) f(z_i) = (g(z_i) \cdot \mathrm{ev}_{z_i})(f).$$

Next, decompose R/I as a product of local Artin algebras, that is, if $\mathfrak{m}_i = \mathrm{I}(z_i) \subseteq R$, then

$$R/I \cong R/J_1 \times \cdots \times R/J_r, \qquad J_i \ \mathfrak{m}_i\text{-primary}.$$

In this way we identify the factors R/J_i as direct summands of R/I, and M_g restricts to each R/J_i: If $e_i \in R/I$ is the idempotent with $e_i(R/I) = R/J_i$, then

$$M_g(R/J_i) = M_g(e_i \cdot R/I) = g \cdot e_i \cdot R/I = e_i \cdot g \cdot R/I \subseteq e_i \cdot R/I = R/J_i.$$

Since the characteristic polynomial of M_g is the product of those of $M_g|_{R/J_i}$, this reduces (iii) and the remainder of (i) to the case where $\mathbb{V}(I) = \{z\}$ is a single point of multiplicity $\ell = \dim_{\mathbb{K}} R/I$.

By the Nullstellensatz $g - g(z) \in \mathrm{I}(\{z\}) = \sqrt{I}$, so $(g-g(z))^m \in I$ for some $m > 0$. Thus $0 = M_{(g-g(z))^m} = (M_g - g(z)\,\mathrm{Id}_{R/I})^m$, so $M_g - g(z)\,\mathrm{Id}_{R/I}$ is nilpotent. This establishes that the characteristic polynomial is $\mathrm{charpol}_{M_g}(t) = (t - g(z))^\ell$. □

This lemma yields the desired linear algebra approach to solve the polynomial system.

Corollary 1.1.2. *The coordinates of the solutions $z_j = (x_1(z_j), \ldots, x_n(z_j)) \in \mathbb{K}^n$ can be recovered as the eigenvalues of $M_{x_1}, \ldots, M_{x_n}$ with simultaneous left eigenvectors ev_{z_j}.*

Example 1.1.3. Consider the univariate case, where

$$I = \langle f \rangle \subseteq \mathbb{K}[x], \qquad f = x^\ell + f_{\ell-1}x^{\ell-1} + \cdots + f_1 x + f_0.$$

We can pick the basis $\overline{1}, \overline{x}, \ldots, \overline{x^{\ell-1}}$ of R/I. For $g \coloneqq x$, the multiplication map is represented by the companion matrix

$$M_x \triangleq \begin{bmatrix} 0 & 0 & \cdots & 0 & -f_0 \\ 1 & 0 & \cdots & 0 & -f_1 \\ 0 & 1 & \cdots & 0 & -f_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -f_{\ell-1} \end{bmatrix} \in \mathbb{K}^{\ell \times \ell}.$$

The eigenvalue-eigenvector theorem states the well-known fact that f is the characteristic polynomial of its own companion matrix. ◁

Pick a term order $\prec$ on R. The *standard monomials* of I, which consist of the finite set of monomials *not* in the initial ideal $\mathrm{in}_\prec I$ (Section 0.2.1), form a basis $\overline{m_1}, \ldots, \overline{m_\ell}$ of R/I. Constructing the matrix M_g in this basis is essentially the same as computing the normal forms of $\mathrm{nf}(m_i g, I)$ for $i = 1, \ldots, \ell$. For more information on the role of normal forms in the context of eigenvalue methods, see [TMB18].

1.1.2 Projective version

Let $J \subseteq S = \mathbb{K}[x_0, \ldots, x_n]$ be a graded ideal with $Z \coloneqq \mathbb{V}(J)^{\mathrm{red}} = \{z_1, \ldots, z_r\} \subseteq \mathbb{P}^n$ a finite set. The Hilbert function $\mathrm{hf}_{S/J}(t)$ is eventually constant with value

$$\ell = \mathrm{HP}_{S/J}(t) = \deg \mathbb{V}(J).$$

As the coordinates of the points are only defined up to rescaling, we can only hope to evaluate expressions $g(z)/h(z)$ with $g, h \in S_e$ homogeneous of the same degree. The

multiplication map M_g here becomes a graded map $(M_g)_d \colon (S/J)_d \to (S/J)_{d+e}$. To state the projective version of the eigenvalue-eigenvector theorem, we need the concept of the saturation index set.

Definition 1.1.4. The *saturation index set* $\mathrm{Sat}(J)$ is the set of indices

$$\mathrm{Sat}(J) = \left\{ d \in \mathbb{N} \;\middle|\; J_d = (J^{\mathrm{sat}})_d \text{ and } \mathrm{hf}_{S/J}(d) = \ell \right\}.$$

Example 1.1.5. If $I = I^{\mathrm{sat}}$ is saturated, then the Hilbert function is increasing, so $\mathrm{Sat}(I) = \mathbb{Z}_{\geq \mathrm{reg_H}(I)}$, see Lemma 1.3.1 below. In practice, often only a non-saturated subideal $J \subsetneq \mathrm{I}(Z)$ is accessible. Such is the case in Section 1.2.3 below. ◁

Since saturation only alters finitely many components of J, there is a d_0 such that $\mathbb{Z}_{\geq d_0} \subseteq \mathrm{Sat}(J)$. Degrees in the saturation index set are useful since they allow us to invert the multiplication map.

Lemma 1.1.6. *If $h \in S_e$ with $h(z) \neq 0$ for $z \in Z$ and $d, d+e \in \mathrm{Sat}(J)$, then $(M_h)_d \colon (S/J)_d \to (S/J)_{d+e}$ is an isomorphism.*

Proof. Since both domain and target degrees are saturated by assumption, we may assume without loss of generality that $J = J^{\mathrm{sat}}$. S/J is a Cohen–Macaulay ring of equidimension 1, hence a homogeneous element $h \in S_e$ is regular on S/J if and only if $\mathbb{V}(J + \langle h \rangle) = \emptyset$ (Theorem 0.1.24). This is the case by the assumption $\mathbb{V}(h) \cap Z = \emptyset$. Regularity implies that M_f and hence $(M_f)_d$ is injective. By the assumption $d, d+e \in \mathrm{Sat}(J)$, source and target have dimension ℓ, hence $(M_h)_d$ is an isomorphism. □

Theorem 1.1.7 (Projective eigenvalue-eigenvector theorem)**.** *Let $d, d+e \in \mathrm{Sat}(J)$ and $g, h \in S_e$ with $h(z) \neq 0$ for $z \in Z$. Let*

$$M \coloneqq (M_h)_d^{-1} \circ (M_g)_d \colon (S/J)_d \xrightarrow{M_g} (S/J)_{d+e} \xrightarrow{M_h^{-1}} (S/J)_d,$$

which is well-defind by the previous lemma.

(i) *The eigenvalues of the endomorphism M are $\frac{g}{h}(z_1), \ldots, \frac{g}{h}(z_r)$ (possibly with repetition or multiplicites).*

(ii) *The characteristic polynomial of M factors as*

$$\mathrm{charpol}_{M_g}(t) = \prod_{i=1}^{r} \left(t - \frac{g}{h}(z_i) \right)^{\ell_i},$$

where ℓ_i is the multiplicity of J at the points z_i.

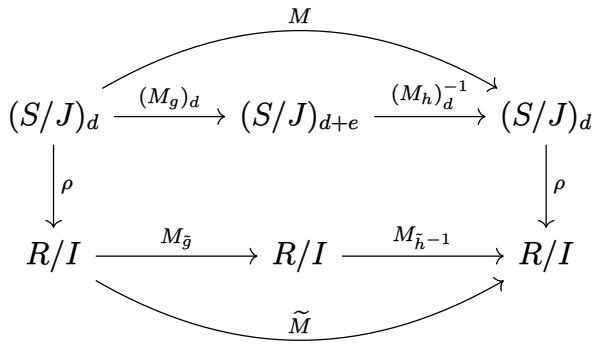
Remark 1.1.8. The left- and right eigenvectors of M can also be described, though slightly less canonically compared to the affine case [Tel20, Definition 3.2.5]. ◁

Proof. Again, we may assume that J is saturated. Without loss of generality, assume $\mathbb{V}(x_0) \cap Z = \emptyset$, and consider the affine chart

$$Z \subseteq \{\, (1 : x_1 : \cdots : x_n) \mid x \in \mathbb{A}^n \,\} \subseteq \mathbb{P}^n.$$

Our goal is to reduce the projective theorem to the affine version on this chart.

Let $R = \mathbb{K}[x_1, \ldots, x_n]$ and $I = \mathrm{I}(Z|_{\mathbb{A}^n}) \subseteq R$ the vanishing ideal in the coordinate ring of this chart. Let $\tilde{g} = g(1, x_1, \ldots, x_n) \in R/I$ be the dehomogenization of g and similarly for $\tilde{h}$. The element $\tilde{h} \in R/I$ is invertible since it vanishes nowhere, therefore it makes sense to define $\widetilde{M} \coloneqq M_{\tilde{g}\cdot\tilde{h}^{-1}} \in \mathrm{End}_{\mathbb{K}}(R/I)$. Then the following diagram is commutative



where the vertical arrows are dehomogenizations $\rho\colon (S/J)_d \to R/I$. ρ is an isomorphism since it is injective (this uses that J is the homogenization of I) and both have the same dimension ℓ. This shows that $M = \rho^{-1}\widetilde{M}\rho$, so M and $\widetilde{M}$ have the same eigenvalues and characteristic polynomial. Now apply the affine eigenvalue-eigenvector theorem. □

The main computational effort in using these eigenvalue methods lies in the task to construct the multiplication matrices M_f themselves. This can be accomplished using the help of *Macaulay matrices* [Tel20, Section 3.4.2], which are indexed by monomials of increasing degree. The computational cost grows with the size of the chosen integers $d, d+e \in \mathrm{Sat}(J)$, it grows like the Hilbert function of the ambient polynomial ring $\mathrm{hf}_S(d+e) = \binom{n+d+e}{n}$. Thus, it is advantageous to chose $d, d+e \in \mathrm{Sat}(J)$ *as small as possible.*

Remark 1.1.9. For practical implementations of eigenvalue solvers, see for example [BT25] or [Ver25]. ◁

1.2 Symmetric Tensors Decomposition

In this section we consider symmetric tensors and the computational problem of symmetric tensor decomposition. After introducing the theory of symmetric tensors, the symmetric algebra and the apolar action, we show how symmetric tensor decomposition can be rephrased as a non-saturated polynomial system of equations. This leads us to the definition of chopped ideal, and we formulate a conjecture about the saturation index set of chopped ideals of general sets of points.

1.2.1 Symmetric tensors

Recall that for vector spaces $V_1, \dots, V_d$ and any permutation $\sigma \in \mathfrak{S}_d$, there is an induced isomorphism

$$V_1 \otimes \cdots \otimes V_d \xrightarrow{\sim} V_{\sigma(1)} \otimes \cdots \otimes V_{\sigma(n)}, \qquad v_1 \otimes \cdots \otimes v_d \mapsto v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}.$$

In particular, if all $V_i = V$ for a finite $\mathbb{K}$-vector space, then the symmetric group acts on the tensor power $V^{\otimes d}$.

Definition 1.2.1 (Symmetric tensors, divided power monomials)**.** A *symmetric tensor* $F \in V^{\otimes d}$ is a tensor invariant under the action of the symmetric group $\mathfrak{S}_d$ on the d tensor factors. The vector space of all degree d symmetric tensors with its induced $\mathrm{GL}(V)$-action is denoted by $\mathrm{Sym}^d V$.

If $X_0, \dots, X_n$ is a basis of V, the *divided power monomials* are the symmetric tensors

$$X^{[\beta]} \coloneqq \sum_{\substack{I \in [n+1]^d \\ \#\{j \mid I_j = i\} = \beta_i}} X_{I_1} \otimes \cdots \otimes X_{I_d} \in \mathrm{Sym}^d V \qquad \beta \in \mathbb{N}^{n+1},\ |\beta| = d.$$

Choosing a basis of V turns a tensor into a hypercube of numbers $F = (F_{i_1,\dots,i_d}) \in \mathbb{K}^{(n+1)^d}$, then symmetric tensors are those unchanged under shuffling the indices:

$$(F_{i_{\sigma(1)},\dots,i_{\sigma(d)}}) = (F_{i_1,\dots,i_d}) \qquad \text{for all } \sigma \in \mathfrak{S}_d.$$

The divided power monomial $X^{[\beta]}$ corresponds to the tensor which is the sum over the $\mathfrak{S}_d$-orbit of $X_1^{\otimes \beta_1} \otimes \cdots \otimes X_n^{\beta_n}$ *without repetition*:

$$(X^{[\beta]})_{\boldsymbol{i}} = \begin{cases} 1 & \text{if in } \boldsymbol{i} = (i_1, \dots, i_d) \in [n+1]^d,\ \beta_1 \text{ entries are } 1,\ \beta_2 \text{ entries are } 2, \dots; \\ 0 & \text{otherwise.} \end{cases}$$

Example 1.2.2 (Symmetric matrices are symmetric tensors)**.** A tensor $T \in V^{\otimes d}$ correspond to a multilinear form (on the dual space)

$$V^\vee \times \cdots \times V^\vee \to \mathbb{K}, \qquad (f_1, \dots, f_d) \mapsto (f_1 \otimes \cdots \otimes f_d)(F).$$

In this perspective, a tensor is symmetric if the value of this multilinear map is independent of reordering the inputs.

Choosing a basis and specializing to $d = 2$, a tensor thus can be identified with a bilinear map on V, hence represented by a matrix $A \in \mathbb{K}^{(n+1)\times(n+1)}$. The permutation action corresponds to taking the transpose, so A is symmetric (as a tensor) if it is a symmetric matrix, $A^\mathsf{T} = A$. For example, in the case of dimension 2, we have

$$X^{[2,0]} = X_0^{\otimes 2} \mathrel{\hat{=}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad X^{[1,1]} = X_0 \otimes X_1 + X_1 \otimes X_0 \mathrel{\hat{=}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad X^{[0,2]} = X_1^{\otimes 2} \mathrel{\hat{=}} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad \triangleleft$$

A related concept to symmetric tensors is the *symmetric algebra*, the free polynomial algebra generated by V

$$\mathrm{S}(V) = \bigoplus_{d\geq 0} \mathrm{S}^d V = \frac{\bigoplus_{d\geq 0} V^{\otimes d}}{\langle v\otimes w - w\otimes v \mid v,w\in V\rangle} \cong \mathbb{K}[X_0,\dots,X_n].$$

These two concepts are related in that they are precisely dual to each other: For $e \geq d \geq 0$ consider the bilinear *contraction map*

$$(V^\vee)^{\otimes d} \times V^{\otimes e} \to V^{\otimes e-d}, \quad (\ell_1\otimes\cdots\otimes\ell_d, v_1\otimes\cdots\otimes v_e) \mapsto \left(\prod_{i=1}^{d} \ell_i(v_i)\right) v_{d+1}\otimes\cdots\otimes v_e.$$

This map restricts to an action of the symmetric algebra on symmetric tensors, and identifies them as dual to each other. The main properties are summarized in the following lemma:

Lemma 1.2.3 (The apolar pairing)**.** (i) *Tensor contraction restricts to a well-defined pairing*

$$\mathrm{S}^d(V^\vee) \times \mathrm{Sym}^e V \to \mathrm{Sym}^{e-d} V, \qquad (f,G) \mapsto f\bullet G.$$

(ii) *Under this action, if $x_1,\dots,x_n \in V^\vee$ is the basis dual to $X_1,\dots,X_n \in V$, then*

$$x^\alpha \bullet X^{[\beta]} = \begin{cases} X^{[\beta-\alpha]} & \text{if } \alpha_i \leq \beta_i \text{ for all } i, \\ 0 & \text{otherwise.} \end{cases}$$

(iii) *The divided power monomials $\{X^{[\beta]}\}$ form a basis of $\mathrm{Sym}^d V$, and they are the basis dual to the monomial basis $\{x^\alpha\}$ of $\mathrm{S}^d V^\vee$. In particular, the pairing is perfect in degree $d = e$.*

(iv) *The action from* (i) *turns $\mathrm{Sym}\, V$ into a graded $\mathrm{S}(V^\vee)$ module.*

(v) *If $v = (v_0,\dots,v_n) \in V$ and $f \in \mathrm{S}^d(V^\vee) = \mathbb{K}[x_0,\dots,x_n]_d$, then*

$$f\bullet v^{\otimes d} = f(v_1,\dots,v_n) \in \mathbb{K}.$$

In this way, $\mathrm{S}(V^\vee)$ is the homogeneous coordinate ring of the projective space $\mathbb{P}(V)$.

Proof. (i) For well-definedness notice that if one picks two representatives of $f \in \mathrm{S}^d(V^\vee)$, they differ by a tensor that is a linear combination of vectors of the form $\ell_1\otimes\cdots(\ell_i\otimes\ell_{i+1} - \ell_{i+1}\otimes\ell_i)\cdots\otimes\ell_d$. Such elements act trivially on $\mathrm{Sym}^d V$, as witnessed by the transposition swapping i and $i+1$. It is easy to see that the contraction of a symmetric tensor is again symmetric.

(ii) Immediate from the definition of $X^{[\beta]}$ and the relation $x_i(X_j) = \delta_{ij}$.

(iii) It is easy to see that every symmetric tensor has a unique representation as a linear combination of DP monomials, for example by successively subtracting orbits of simple tensors. This shows that the DP monomials form a basis of $\operatorname{Sym}^d V$.

If $d = e$, then the pairing gives a bilinear form $\mathrm{S}^d(V^\vee) \times \operatorname{Sym}^d V \to \mathbb{K}$, and by (ii) we see that x^β is precisely dual to $X^{[\beta]}$. This also establishes that the pairing is perfect.

(iv) To check that this is a well-defined graded module structure, we only need to check associativity: $f \bullet (g \bullet H) = (f \cdot g) \bullet H$. This can be read off from (ii).

(v) The identity is immediate from the definition of the pairing. In this way we can evaluate polynomial functions in vectors $v \in V$, which precisely gives the action of $\mathbb{K}[x_0, \dots, x_n]$ on $\mathbb{K}^{n+1}$ (and hence $\mathbb{P}^n(\mathbb{K})$). □

For complex-minded people, the space of symmetric tensors is often identified with the symmetric algebra. This is justified by the following lemma, relating the apolar action to partial differentiation. Recall the multi-index notation $\alpha! \coloneqq \alpha_0! \cdots \alpha_n!$ and $\binom{d}{\alpha} \coloneqq d!/\alpha!$ where $d = |\alpha| = \sum_i \alpha_i$.

Lemma 1.2.4. *The* $\mathrm{GL}(V)$*-equivariant linear map*

$$V^{\otimes d} \to \operatorname{Sym}^d V, \qquad T \mapsto \sum_{\sigma \in \mathfrak{S}_d} \sigma(T)$$

is well-defined when restricted to $\psi_d \colon \mathrm{S}^d V \to \operatorname{Sym}^d V$*. It is an isomorphism if (and only if)* $\operatorname{char} \mathbb{K} = 0$ *or* $> d$*. Under this identification we have*

$$\psi_d(X^\alpha) = \alpha! \cdot X^{[\alpha]}, \qquad x^\alpha \bullet \psi_e(f(X)) = \psi_{e-d}\left(\frac{\partial^d}{\partial X_1^{\alpha_0} \cdots \partial X_n^{\alpha_n}} f(X)\right).$$

with $|\alpha| = d$*,* $f(X) \in \mathrm{S}^e V$*. In particular, in characteristic* 0 *the induced apolar action of* $\mathrm{S}(V^\vee)$ *on* $\mathrm{S}(V)$ *is the action by partial differentiation,* $x_i \mathrel{\hat{=}} \frac{\partial}{\partial X_i}$.

The isomorphism $\psi_\bullet$ identifies $X^{[\alpha]}$ with $\frac{1}{\alpha!} X^\alpha$; this explains the name "divided power monomial". See also [Eis95, Chapter A2.4] or [IK99, Appendix A] for more information about the divided power structure.

Proof. The map is well-defined on the quotient since it maps elements in the ideal $\langle v \otimes w - w \otimes v \rangle_{T(V)}$ to zero. By the orbit-stabilizer formula, the map

$$\mathfrak{S}_d \to V^{\otimes d}, \qquad \sigma \mapsto \sigma(X_1^{\otimes \beta_1} \otimes \cdots \otimes X_n^{\beta_n})$$

meets every permutation of this elementary tensor exactly $d!/\binom{d}{\alpha} = \alpha_1! \cdots \alpha_n!$ times, hence $\psi_d(X^\alpha) = \alpha! X^{[\alpha]}$. If $d! \in \mathbb{K}^\times$, then the formula shows that ψ_d is an isomorphism. On the other hand, if $d! = 0$, then $\psi(X_1^{\otimes d}) = 0$.

To see that $\psi_\bullet$ turns the apolar action into partial differentiation, we only need to consider $f = X^\beta$ by bilinearity. In that case

$$x^\alpha \bullet \psi_e(X^\beta) = \beta! \cdot x^\alpha \bullet X^{[\beta]} = \frac{\beta!}{(\beta-\alpha)!}(\beta-\alpha)! \cdot X^{[\beta-\alpha]}$$
$$= \psi_{e-d}\left(\frac{\beta!}{(\beta-\alpha)!}X^{\beta-\alpha}\right) = \psi_{e-d}\left(\frac{\partial^d}{\partial X^\alpha}X^\beta\right). \qquad \square$$

In conclusion, for the most part there is no harm in identifying symmetric tensors $\operatorname{Sym}^d V$ with polynomials $\mathrm{S}(V) = \mathbb{K}[X_0, \dots, X_n]$, unless one is interested in fields of small characteristic. The apolar action by the dual polynomial ring $\mathrm{S}(V^\vee) = \mathbb{K}[x_0, \dots, x_n]$ is then identified with the partial differentiation action.

1.2.2 Tensor decomposition and the apolarity lemma

We now turn to the problem of symmetric tensor decomposition and its relation to the apolar action. For more information about (symmetric) tensor decomposition, see [Lan12; Ber+18; IK99].

Definition 1.2.5 (Symmetric tensor rank)**.** A *simple symmetric tensor* is of the form $F = v^{\otimes d}$. The *symmetric tensor rank* of $F \in \operatorname{Sym}^d V$ is the least $r \in \mathbb{N}$ such that F is a linear combination of simple symmetric tensors:

$$R(F) = \min\left\{ r \in \mathbb{N} \;\middle|\; F = \lambda_1 v_1^{\otimes d} + \cdots + \lambda_r v_r^{\otimes d}, \lambda_i \in \mathbb{K}, v_i \in V \right\}.$$

Note that since $\mathbb{K}$ is algebraically closed, we may always pick $\lambda_i = 1$ for all i. We will see in a moment in Corollary 1.2.10 that $R(F) < \infty$. In characteristic 0, simple symmetric tensors correspond to powers of linear forms under the identification $\operatorname{Sym}^d V \cong \mathbb{K}[\underline{X}]$; in that context symmetric tensor decompositions are also called *Waring decompositions*.

Example 1.2.6 (Symmetric tensor rank generalizes matrix rank)**.** Let $\operatorname{char}\mathbb{K} \neq 2$, then every symmetric matrix M is congruent to a diagonal matrix $A = P^\mathsf{T} M P$, $P \in \mathrm{GL}(\mathbb{K}^n)$. For such a matrix we exhibit the symmetric tensor decomposition

$$A = \begin{bmatrix} \lambda_1 & & & & \\ & \ddots & & & \\ & & \lambda_r & & \\ & & & 0 & \\ & & & & \ddots \end{bmatrix} = \lambda_1 e_1 e_1^\mathsf{T} + \cdots + \lambda_r e_r e_r^\mathsf{T} = \sum_{i=1}^r \lambda_i e_i^{\otimes 2},$$

so $R(A) \leq \operatorname{rank} A$ (the latter being ordinary matrix rank). On the other hand, every matrix of the form vv^T has rank 1, so if A is a linear combination of r simple symmetric tensors, then $\operatorname{rank} A \leq r$.

Thus, in this special case symmetric tensor rank coincides with the familiar matrix rank. $R(-)$ can be seen as a generalization to higher order symmetric tensors. Note however that here $R(A) = \operatorname{rank} A \leq \dim V$, which is very false for higher order tensors as we will see. $\triangleleft$

The apolar action introduced in the previous section provides a powerful tool to study symmetric tensor decompositions using commutative algebra. The connection is given by the *Apolarity Lemma*. For this we need to introduce the apolar ideal of a symmetric tensor.

Definition 1.2.7. The *apolar ideal* of $0 \neq F \in \operatorname{Sym}^d V$ is the graded ideal

$$\operatorname{Ann}_S F = \{\, f \mid f \bullet F = 0 \,\} \subseteq S = \mathbb{K}[x_0, \dots, x_n].$$

Each graded component is given by the kernel $(\operatorname{Ann}_S F)_j = \operatorname{Ker} F_{j,d-j}$ of the *apolar map*

$$F_{j,d-j} \colon S_j \to \operatorname{Sym}^{d-j} V, \qquad f \mapsto f \bullet F.$$

For a set of points $z_1, \dots, z_m \in \mathbb{P}(V)$ denote by $\mathrm{I}(\{z_1, \dots, z_m\}) \subseteq S$ their homogeneous vanishing ideal.

Theorem 1.2.8 (Apolarity Lemma)**.** *Let $0 \neq F \in \operatorname{Sym}^d V$, $v_1, \dots, v_r \in V \setminus 0$ and set $Z = \{[v_1], \dots, [v_r]\} \subseteq \mathbb{P}V$. The following are equivalent:*

(a) $F = \lambda_1 v_1^{\otimes d} + \cdots + \lambda_r v_r^{\otimes d}$ *for some* $\lambda_1, \dots, \lambda_r \in \mathbb{K}$*;*

(b) $\mathrm{I}(Z) \subseteq \operatorname{Ann}_S F$*;*

(c) $\mathrm{I}(Z)_d \subseteq (\operatorname{Ann}_S F)_d$.

We will use for a subspace $U \subseteq W$ the notation $U^\perp \coloneqq \{\, \ell \in W^\vee \mid \ell(U) = 0 \,\}$.

Proof. For (a)⇒(b), let $f \in \mathrm{I}(\mathbb{X})$, then by Lemma 1.2.3(v) $0 = f(v_i) = f \bullet v_i^{\otimes d}$ for each i and hence

$$f \bullet F = \sum_{i=1}^{r} \lambda_i f \bullet v_i^{\otimes d} = 0.$$

This establishes $\mathrm{I}(Z) \subseteq \operatorname{Ann}_S F$. Trivially, (b) implies (c).

For the remaining implication, assume (c). Using the duality of $\operatorname{Sym}^d V$ and S_d, by Lemma 1.2.3 we have

$$F \in (\operatorname{Ann}_S F)_d^\perp \overset{(c)}{\subseteq} (\mathrm{I}(Z)_d)^\perp = \left(\bigcap_{z \in \mathbb{Z}} \mathrm{I}(z)_d \right)^\perp = \sum_{z \in Z} (\mathrm{I}(z)_d)^\perp.$$

For each point $z = [v]$ we have

$$\mathrm{I}(z)_d = \left\{ f \in S_d \;\middle|\; 0 = f(v) = f \bullet v^{\otimes d} \right\} = \operatorname{Ann}_S(v^{\otimes d})_d$$

and hence $(\mathrm{I}(z)_d)^\perp = \mathbb{K} \cdot v^{\otimes d}$. Plugging this into the previous equation yields

$$F \in \mathbb{K} v_1^{\otimes d} + \cdots + \mathbb{K} v_r^{\otimes d}. \qquad \square$$

Thus, the task of finding a symmetric tensor decomposition of length r is equivalent to finding an ideal of r points inside $\operatorname{Ann}_S F$. We illustrate this in the case of binary forms:

Example 1.2.9 (Binary forms)**.** If $V = \mathbb{C}^2$ with basis X, Y, then symmetric tensor decomposition corresponds to the task of writing $F \in \operatorname{Sym}^d V \mathrel{\hat{=}} \mathbb{C}[X,Y]_d$ as a sum of d-th powers of linear forms $L_i \in V \mathrel{\hat{=}} \mathbb{C}[X,Y]_1$.

The apolar ideal $\operatorname{Ann}_S(F) = \langle f, g\rangle_S$ is generated by two forms $f, g \in \mathbb{C}[x,y]$ of degree $\deg f = r$, $\deg g = r' := d+2-r$ [IK99, Theorem 1.44iv]. Here, r is the smallest index such that $(\operatorname{Ann}_S F)_r \neq 0$.

By the apolarity lemma, we are looking for ideals of points inside $\operatorname{Ann}_S(F)$. In other words, as $\dim \mathbb{P}(V) = 1$, we are looking for polynomials without repeated roots. There are two cases:

(i) If f has simple roots $[v_1], \dots, [v_r] \in \mathbb{P}^1$, then F has symmetric tensor rank $R(F) = r$ and a minimal decomposition is given by the v_i.

(ii) If f has repeated roots, then there is no reduced homogeneous polynomial of degree $< r'$ in $\operatorname{Ann}_S(F)$. On the other hand, $(\operatorname{Ann}_S F)_{r'} \subseteq S_{r'}$ is a base-point free linear system on $\mathbb{P}^1$ as $\mathbb{V}(f,g) = \emptyset$. By Bertini's theorem, a general polynomial in $(\operatorname{Ann}_S F)_{r'}$ has reduced zero-set. So in this case $R(F) = r'$.

This indicates a strategy to determine $R(F)$: First, compute f as a non-zero element of $\operatorname{Ker} F_{r,d-r}$ for the smallest r where this kernel is non-trivial. If f has simple roots ($\operatorname{disc}(f) \neq 0$), then $R(F) = r$, otherwise $R(F) = d+2-r$. To actually compute a decomposition, one needs to split f (or a random element of $(\operatorname{Ann}_S F)_{r'}$ in the second case) into linear factors. ◁

We can deduce from the apolarity lemma that every symmetric tensor has a symmetric tensor decomposition:

Corollary 1.2.10. *Every $F \in \operatorname{Sym}^d V$ has a symmetric tensor decomposition using at most $\binom{n+d}{n}$ summands.*

Proof. Pick a general set of $\binom{n+d}{n} = \dim_{\mathbb{K}} S_d$ points $Z \subseteq \mathbb{P}(V)$, then $\mathrm{I}(Z)_d = 0 \subseteq (\operatorname{Ann}_S F)_d$. By the apolarity lemma, F is a linear combination of $v^{\otimes d}$, $[v] \in Z$. □

Remark 1.2.11. Much better bounds on the symmetric rank are known. In fact, by the celebrated Alexander–Hirschowitz theorem, a general symmetric tensor has rank $\lceil \frac{1}{n+1}\binom{n+d}{n} \rceil$ (at least if $\operatorname{char} \mathbb{K} \nmid d$) up to a few exceptional cases [IK99, Section 2.1]. ◁

1.2.3 Tensor decomposition using eigenvalue methods

In general, tensor decompositions are not unique, and often difficult to compute. However, for general symmetric tensors of rank r not too big, apolarity *can* be used to find the decomposition. A precise statement is the following, see [IK99, Lemma 1.19 and Theorem 2.6ii].

Proposition 1.2.12. *Assume $F = \sum_{i=1}^r \lambda_i v_i^{\otimes d}$, and let $Z = \{[v_1], \dots, [v_r]\} \subseteq \mathbb{P}^{n-1}$.*

(i) *Let $j \geq 0$ be such that $\dim S_{d-j}/\mathrm{I}(Z)_{d-j} = r$ (this is satisfied for general Z and $j \leq d/2$, see Lemma 1.3.2), then $\mathrm{I}(Z)_j = (\operatorname{Ann}_S F)_j$.*

(ii) *If* $r \leq \binom{n+\lceil d/2 \rceil}{n} - (n+1)$ *and* Z *is general, then the decomposition of* F *is unique up to rescaling and permuting the* v_i*.*

This gives the following pseudo-algorithm for computing the symmetric tensor decomposition of a general symmetric tensor F with $R(F) \leq \binom{n+\lceil d/2 \rceil}{n} - (n+1)$.

(1) Compute the apolar ideal in degree(s) $j \leq \lfloor d/2 \rfloor$

$$\mathcal{F} \coloneqq (\operatorname{Ann}_S F)_j = \operatorname{Ker} F_{j,d-j} \subseteq S = \mathbb{K}[x_0, \dots, x_n]$$

(2) Find the solutions to the polynomial systems $Z = \mathbb{V}(\mathcal{F}) \subseteq \mathbb{P}^{n-1}$

(3) Find the scalars $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ such that $F = \sum_{i=1}^r \lambda_i v_i^{\otimes d}$, $[v_i] \in Z$.

The first and third step amount to solving a linear system over $\mathbb{K}$, while the crucial second step is a nonlinear system of equations. A caveat is that we have yet to verify that the equations $\mathcal{F} = \mathrm{I}(Z)_j$ *actually* determine the set of points, that is,

$$\mathbb{V}(\mathrm{I}(Z)_j) \stackrel{?}{=} Z \qquad Z \subseteq \mathbb{P}^{n-1} \text{ } r \text{ general points.}$$

For numerically solving the polynomial system $\mathcal{F}$ we propose using the eigenvalue method as opposed to, say, homotopy continuation methods. The reason is that $\mathcal{F}$ is almost always overdetermined, and so a homotopy continuation approach would (usually) resort to solving a square sub system which has $d^n \gg r$ solutions (and then filtering using the remaining equations). More refined homotopy continuation approaches have been proposed in [Ber+17], but they rely on the knowledge of certain information about secant varieties of the Veronese variety, which is currently out of reach. In fact, previous work on (non-symmetric) tensor decomposition has already successfully applied eigenvalue methods, outperforming homotopy continuation methods [TV22]. An implementation of this method can be found at https://mathrepo.mis.mpg.de/ChoppedIdeals/.

In order to successfully apply the eigenvalue method, we need information about the saturation index set (Section 1.1.2) and, hence, the Hilbert function of the apolar ideal. Under suitable genericity assumptions from Proposition 1.2.12, this is a question purely about ideals of general sets of points:

Problem 1.2.13. Let $Z = \{z_1, \dots, z_r\} \subseteq \mathbb{P}^n$ be a general set of points and $I = \mathrm{I}(Z) \subseteq S = \mathbb{K}[x_0, \dots, x_n]$.

(i) What is the nature of $\mathbb{V}(\mathrm{I}(Z)_d) \subseteq \mathbb{P}^n$ depending on n, d, r? Does it equal Z?

(ii) What is the saturation index set of the ideal $J = \langle I_d \rangle_S$?

(iii) More generally, what is the Hilbert function of J? ⍰

Note that in this formulation, d is the degree of the polynomials in the (apolar/point) ideal, not the degree of the symmetric tensor (!).

Definition 1.2.14 (Chopped ideal)**.** Let $I \subseteq S = \mathbb{K}[x_0, \dots, x_n]$ be a homogeneous ideal and $d \geq 0$. The *chopped ideal* in degree d associated to I is $I_{\langle d \rangle} \coloneqq \langle I_d \rangle_S$.

The notation $I_{\langle d \rangle}$ goes back to Herzog & Hibi [HH99], where the authors study ideals whose chopped ideals have linear resolutions. Phrased differently, Problem 1.2.13 is about chopped ideals of the ideal of a general set of points. Before we propose an answer to these questions, we give a detailed case study of the smallest interesting case.

Example 1.2.15 (18 points in the plane)**.** Consider the planar case $n = 2$. The smallest case in which Proposition 1.2.12 applies, but the apolar ideal generators in the algorithm do not generate the entire ideal of points, is a general symmetric tensor of degree 10 and rank 18. The set of such tensors in $\mathrm{Sym}^{10}\,\mathbb{K}^3$ has dimension 54, for comparison, a general symmetric tensor has rank 22. In this case, our polynomial system can access all equations of degree ≤ 5 in the ideal of the 18 points in the decomposition.

Let Z be a set of 18 general points in $\mathbb{P}^2$. The lowest degree elements of $\mathrm{I}(Z)$ are in degree 5 and, a priori, $\mathrm{I}(Z)$ can have minimal generators in degree 5 and 6 (Lemma 1.3.2). It turns out that $\mathrm{I}(Z)$ is generated by three quintics and one sextic. In particular, the chopped ideal $J = I_{\langle 5 \rangle} = \langle \mathrm{I}(Z)_5 \rangle_S$ does *not* coincides with $\mathrm{I}(Z)$. We provide a simple snippet of code in `Macaulay2` [GS] using the convenient package `Points.m2` [Sti+] to compute $\mathrm{I}(Z), J$ and the corresponding Hilbert functions:

```
loadPackage "Points"
I = randomPoints(2,18);
J = ideal super basis(5,I);
for t to 10 list {t, hilbertFunction(t,I), hilbertFunction(t,Ichop)}
```

The two Hilbert functions are recorded below:

t	0	1	2	3	4	5	6	7	8	9	...
$\mathrm{hf}_{S/\mathrm{I}(Z)}(t)$	1	3	6	10	15	18	18	18	18	18	...
$\mathrm{hf}_{S/\mathrm{I}(Z)_{\langle 5 \rangle}}(t)$	1	3	6	10	15	18	**19**	18	18	18	...

We observe that the Hilbert polynomials are the same: they are equal to the constant 18. However, the Hilbert function of S/J overshoots the Hilbert polynomial in degree 6, and then falls back to 18 in degree 7. Here, the saturation index set is $\{5\} \cup \mathbb{N}_{\geq 7}$. This specific example is explained in detail in Section 1.5.1. ◁

We now propose an answer to Problem 1.2.13(ii). For $Z \subseteq \mathbb{P}^n$, let $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle_S$ and assume $d \in \mathrm{Sat}(\mathrm{I}(Z))$ (hence also $d \in \mathrm{Sat}(I_{\langle d \rangle})$). Define the *saturation gap* as

$$\gamma(d, Z) \coloneqq \min \left\{ e > 0 \;\middle|\; \mathrm{hf}_{S/I_{\langle d \rangle}}(d+e) = r \right\}.$$

For general Z, the value $\gamma(d, Z) = \gamma_n(d, r)$ only depends on n, d, r (Proposition 1.3.14).

Conjecture 1.2.16 (Saturation Gap Conjecture). Let $n, d, r \in \mathbb{N}_{>0}$ be integers with $r < \mathrm{hf}_S(d) - n$, then

$$\gamma_n(d, r) = \min \left\{ e > 0 \;\middle|\; \sum_{k=0}^{n-3} (-1)^k \,\mathrm{hf}_S(d + e - kd) \binom{\mathrm{hf}_S(d) - r}{k} \leq r \right\}. \qquad ⍰$$

Remark 1.2.17. We note here the subtle distinction between the functions $\binom{n+t}{n} = \mathrm{HP}_S(t)$ and $\mathrm{hf}_S(t)$. The former is a polynomial on $\mathbb{Z}$, non-zero for $t < -n$. The latter agrees with the former for $t \geq 0$, but is identically 0 for $t < 0$. ◁

The implementation in [KKT] uses this bound in the construction of the multiplication map M from the (projective) eigenvalue theorem. We will give a heuristic derivation of this formula in Section 1.3.2. In Corollary 1.6.5, we will prove the upper bound

$$\gamma_n(d, r) \leq (n-1)d - (n+1),$$

at least when $\operatorname{char} \mathbb{K} = 0$. This bound is sharp when $r = \mathrm{hf}_S(d) - n - 1$ (Theorem 1.6.1), in which case it indeed coincides with the expression from Conjecture 1.2.16.

1.3 The expected syzygy conjecture

In this section we will derive the Expected Syzygy Conjecture 1.3.10 (ESC) as the conjectural answer to Problem 1.2.13. We start by reviewing properties of ideals of general points, namely the Ideal Generation Conjecture 1.3.3 and the Minimal Resolution Conjecture 1.3.4. We then connect ESC to Fröberg's Conjecture 1.3.8, and prove ESC in many small cases using computer algebra.

1.3.1 General points in projective space

Let $Z \subseteq \mathbb{P}^n$ be a finite subscheme of length r, such as a set of r points. We start by examining basic homological properties of the ideal $I = \mathrm{I}(Z) \subseteq S$ [IK99, Theorem 1.69].

Lemma 1.3.1. *The Hilbert function* $\mathrm{hf}_{S/I}$ *is weakly increasing towards the constant Hilbert polynomial* $\mathrm{HP}_Z(t) = r$*. We have*

$$\begin{aligned} \operatorname{depth} S/I = \dim S/I &= 1 \\ \operatorname{p.dim} S/I &= n \\ \operatorname{reg}_{\mathrm{CM}} S/I = \operatorname{reg}_{\mathrm{H}} S/I &= \min\{\, d \mid \dim_{\mathbb{K}}(S/I)_d = r \,\}. \end{aligned}$$

In particular, Z is arithmetically Cohen–Macaulay.

Proof. Since $\dim Z = 0$ and $I = I_Z$ is saturated, $1 \leq \operatorname{depth} S/I \leq \dim S/I = \dim Z + 1 = 1$, so S/I arithmetically Cohen-Macaulay. The statement about the Hilbert function follows from Lemma 0.1.32. The projective dimension is readily computed using the Auslander–Buchsbaum formula (Theorem 0.1.21), and the two regularities coincide by Theorem 0.1.29. □

We now focus on the case that $Z = \{z_1, \ldots, z_r\} \subseteq \mathbb{P}^n$ is a collection of general points.

Lemma 1.3.2. *Let* $d = d_{r,n} := \min\{\, j \in \mathbb{N} \mid \binom{n+j}{n} \geq r \,\}$.

(i) *We have* $\operatorname{reg}_{\mathrm{CM}} S/I = d$ *and*

$$\operatorname{hf}_Z(t) = \min\{\operatorname{hf}_S(t), r\} = (1, n+1, \tbinom{n+2}{n}, \dots, \tbinom{n+d-1}{n}, r, r, \dots). \tag{1.1}$$

(ii) *The minimal generators of* $\mathrm{I}(Z)$ *are concentrated in degree* $d, d+1$ *and*

$$\beta_{1,d} = \binom{n+d}{n} - r, \qquad \beta_{1,d+1} = \binom{n+d+1}{n} - r - \operatorname{rank} \mu_1,$$

where $\mu_1 \colon S_1 \otimes_{\mathbb{K}} I_d \to I_{d+1}$ *is the multiplication map.*

(iii) *For all* $i \geq 1$*, at most two Betti numbers are nonzero:* $\beta_{ij} = 0$*,* $j \notin \{d+i-1, d+i\}$.

Proof. A simple induction on r shows that r general points impose independent conditions on degree t forms: $\dim_{\mathbb{K}} \mathrm{I}(Z)_t = \max\{\operatorname{hf}_S(t) - r, 0\}$. Thus, the claimed values for the Hilbert function are correct for $t \leq d$. By monotonicity (Lemma 1.3.1) they are thus correct for all t. The aforementioned lemma then also shows $\operatorname{reg}_{\mathrm{CM}} S/I = \operatorname{reg}_{\mathrm{H}} S/I = d$.

To see (ii), by Definition 0.1.18 of the Castelnuovo–Mumford regularity we have

$$\begin{aligned} d = \operatorname{reg}_{\mathrm{CM}} S/I &\geq \max\{\, j \mid \beta_{1j}(S/I) \neq 0 \,\} - 1 \\ &= \max\{\, j \mid I \text{ has min. generator in deg. } j \,\} - 1. \end{aligned}$$

This together with the observation $I_t = 0$ for $t < d$ shows that all minimal generators are in degree $d, d+1$. The minimal generators in degree d form a basis of I_d, while the minimal generators in degree $d+1$ span a complement of $\operatorname{Im}(\mu_1)$ in I_{d+1}.

Finally, we can deduce (iii), using that $\beta_{ij} = 0$ implies $\beta_{i+k,j+k} = 0$ for all $k \geq 0$:

$$\begin{aligned} d + i - 1 \leq \min\{\, j \mid \beta_{1,j-(i-1)} \neq 0 \,\} &\leq \min\{\, j \mid \beta_{ij} \neq 0 \,\} \\ &\leq \max\{\, j \mid \beta_{ij} \neq 0 \,\} \leq \operatorname{reg}_{\mathrm{CM}} S/I + i = d + i. \end{aligned}$$

□

If the multiplication map $\mu_1 \colon S_1 \otimes_{\mathbb{K}} I_d \to I_{d+1}$ is surjective, then all minimal generators are in degree d. On the other hand, if the multiplication map is injective, then

$$\beta_{1,d+1} = \binom{n+d+1}{n} - (n+1)\left(\binom{n+d}{n} - r\right).$$

The long-standing Ideal Generation Conjecture [GO82] claims that these two are the only possibilities:

Conjecture 1.3.3 (Ideal Generation Conjecture (IGC)). For general Z and $d = \operatorname{reg}_{\mathrm{H}} Z$, the multiplication map $\mu_1 \colon S_1 \otimes_{\mathbb{K}} \mathrm{I}(Z)_d \to \mathrm{I}(Z)_{d+1}$ has maximal rank. ⁇

In other words, the conjecture predicts that the minimal generators in degree d have no linear syzygies unless they already generate the full ideal, that is, either $\beta_{1,d+1} = 0$ or $\beta_{2,d+1} = 0$ (or both).

The Minimal Resolution Conjecture [Lor93] is an extension of the IGC to higher syzygies:

Conjecture 1.3.4 (Minimal Resolution Conjecture (MRC)). For general Z and all $i \geq 1$, $\beta_{i,d+i} \cdot \beta_{i+1,d+i} = 0$. ⍰

Note that this vanishing behavior together with the knowledge of hf_Z determines the entire Betti table. The MRC is known to be true for $n = 2$ [GM84; GGR86], for $n = 3$ [Bal87] and for $n = 4$ [Wal95; OM17]. Moreover, it has been proved for all n when $r \gg n$ is sufficiently large [HS96] and in many other specific cases, for which we refer to [EP96]. It is however false in general, the smallest known counterexample being 11 general points in $\mathbb{P}^6$ [Eis+02]. There are no known counterexamples to the IGC.

We record here the statement in the case of $\mathbb{P}^2$, where the Hilbert–Burch theorem [Eis95, Theorem 3.2] dictates the structure of the minimal free resolution.

Proposition 1.3.5 (MRC in $\mathbb{P}^2$). *For a general collection of r points $Z \subseteq \mathbb{P}^2$ with $\mathrm{reg}_{\mathrm{H}}(Z) = d$, the minimal free resolution of $S/\mathrm{I}(Z)$ has the form*

$$0 \longleftarrow S/\mathrm{I}(Z) \longleftarrow S \longleftarrow \begin{matrix} S[-d]^{\beta_{1,d}} \\ \bigoplus \\ S[-(d+1)]^{\beta_{1,d+1}} \end{matrix} \xleftarrow{B} \begin{matrix} S[-(d+1)]^{\beta_{2,d+1}} \\ \bigoplus \\ S[-(d+2)]^{\beta_{2,d+2}} \end{matrix} \longleftarrow 0.$$

Here $\beta_{1,d} = \mathrm{hf}_S(d) - r$, $\beta_{1,d+1} = \max\{0, \mathrm{hf}_S(d+1) - (n+1)\beta_{1,d} - r\}$ and either

(i) *$\beta_{1,d+1} > 0$, in which case $\beta_{2,d+1} = 0$ and $\beta_{2,d+2} = \beta_{1,d+1} + \beta_{1,d+2} - 1$, or*

(ii) *$\beta_{1,d+1} = 0$, in which case $\beta_{2,d+1} = d(d+2) - 2r$ and $\beta_{2,d+2} = \beta_{1,d} - \beta_{2,d+1} - 1$.*

For a proof of this particular case, see for example [Sau85, Prop. 1.7]. In the paper the proof goes via polarization of monomial ideals.

Example 1.3.6. A general set Z of 18 points in $\mathbb{P}^2$ has regularity $d = 5$. In accordance with the aforementioned conjectures, it has 3 minimal generators in degree 5 and one minimal generator in degree 6. The minimal free resolution has the shape

$$0 \longleftarrow S/\mathrm{I}(Z) \longleftarrow S \longleftarrow S[-5]^3 \oplus S[-6] \xleftarrow{B} S[-7]^3 \longleftarrow 0. \qquad \triangleleft$$

1.3.2 The Fröberg conjecture and expected syzygies

Recall that we are interested in the chopped ideal $J = \mathrm{I}(Z)_{\langle t \rangle}$, where $Z = \{z_1, \ldots, z_r\} \subseteq \mathbb{P}^n$ is a general set of points. Using $d = d_{n,r} = \mathrm{reg}_{\mathrm{H}} Z$, Lemma 1.3.1 shows that $I_{\langle t \rangle} = 0$ for $t < d$ while $(I_{\langle t \rangle})_j = I_j$ for all $j \geq t$ if $t > d$. Hence considering the chopped ideal is only interesting in degree d.

The Ideal Generation Conjecture 1.3.3 describes $(I_{\langle d \rangle})_{d+1}$ by claiming that the multiplication map $\mu_1 \colon S_1 \otimes_{\mathbb{K}} I_d \to I_{d+1}$ has maximal rank. One might ask whether this continues to hold for multiplication with higher degree forms, that is, whether

$$\mu_e \colon S_e \otimes_{\mathbb{K}} I_d \to I_{d+e}, \qquad g \otimes f \mapsto g \cdot f$$

is either injective or surjective. This turns out to be false in general.

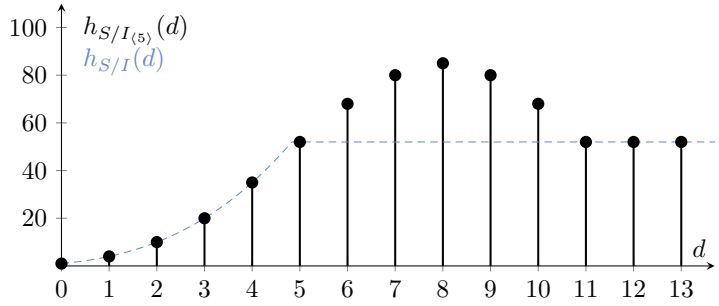


Figure 1.1: The Hilbert function of $\mathrm{I}(Z)_{\langle 5\rangle}$ where Z are 52 general points in $\mathbb{P}^3(\mathbb{C})$.

Example 1.3.7. A general set of 52 points in $\mathbb{P}^3(\mathbb{C})$ has 4 minimal generators $f_1, \dots, f_4$ in degree 5 and 16 further minimal generators in degree 6. The chopped ideal has the following Hilbert function (compare Figure 1.1)

j	1	2	3	4	5	6	7	8	9	10	11	…
$\mathrm{hf}_{S/I}(j)$	4	10	20	35	52	52	52	52	52	52	52	…
$\mathrm{hf}_{S/I_{\langle 5\rangle}}(j)$	4	10	20	35	52	68	80	85	80	**68**	52	…
$\mathrm{hf}_S(j) - 4\,\mathrm{hf}_S(j-5)$	4	10	20	35	52	68	80	85	80	**62**	(28)	…

The multiplication map $\mu_5 \colon S_5 \otimes I_5 \cong \mathbb{C}^{224} \to I_{10} \cong \mathbb{C}^{234}$ is *neither* injective nor surjective. A basis of the 6-dim'l kernel is given by the Koszul syzygies $f_j \otimes f_j - f_j \otimes f_i$, $1 \le i < j \le 4$. ◁

This example illustrates that for the higher degree multiplication maps μ_e, syzygies have to be considered to correct the Hilbert function. Let $s = \mathrm{hf}_I(d) = \binom{n+d}{n} - r$ be the number of minimal generators in degree d. The kernel of μ_d always contains at least the $\binom{s}{2}$ Koszul syzygies from degree $2d$ onwards. These syzygies themsemselves will produce $\binom{s}{3}$ syzygies in degree $3d$ and so on (see Example 0.1.13 for more about the Koszul complex). A reasonable second guess is therefore

$$\mathrm{hf}_{S/I_{\langle d\rangle}}(t) \overset{?}{=} \mathrm{hf}_S(t) - \underbrace{s \cdot \mathrm{hf}_S(t-d)}_{\text{gen's of } I_d} + \underbrace{\binom{s}{2} \cdot \mathrm{hf}_S(t-2d)}_{\text{Koszul syzygies}} - \underbrace{\binom{s}{3} \cdot \mathrm{hf}_S(t-3d)}_{\text{Koszul syzygy syzygies}} \pm \dots$$

This is indeed (essentially) the Hilbert function of an ideal generated by s general forms, at least conjecturally:

Conjecture 1.3.8 (Fröberg [Frö85]). Let $J \subseteq \mathbb{K}[x_0, \dots, x_n]$ be an ideal generated by s general forms of degree d, then

$$\mathrm{hf}_{S/J}(t) = \mathrm{frö}_{d,s}(t) \coloneqq \begin{cases} \sum_{k=0}^{s} (-1)^k \binom{s}{k} \cdot \mathrm{hf}_S(t-kd) & t < t_0 \\ 0 & t \ge t_0, \end{cases}$$

where $t_0 \in \mathbb{N} \cup \{\infty\}$ is minimal such that the summation is less than or equal to 0. ⍰

In the same paper, Fröberg proved the following lower bound. Consider the *lexicographic ordering* on functions $h, h' \colon \mathbb{Z}_{\geq 0} \to \mathbb{Z}$:

$$h \leq_{\text{lex}} h' \quad \text{if and only if} \quad \inf\{\, t \mid h(t) < h'(t) \,\} \leq \inf\{\, t \mid h(t) > h'(t) \,\}.$$

In other words, $h <_{\text{lex}} h'$ if $h(t) = h'(t)$ for $t < t_0$ and $h(t_0) < h'(t_0)$. This is a total order on functions h, and a pointwise inequality $h(t) \leq h'(t)$ for all $t \geq 0$ implies $h \leq_{\text{lex}} h'$.

Theorem 1.3.9 ([Frö85])**.** *For any ideal $J \subseteq S$ of depth 0 generated by $s \geq n+1$ elements in degree d one has $\operatorname{hf}_{S/J}(t) \geq_{\text{lex}} \operatorname{frö}_{d,s}(t)$. If Fröberg's Conjecture 1.3.8 holds (for particular n, d, s), the lex-inequality is upgraded to a pointwise inequality:*

$$\operatorname{hf}_{S/J}(t) \geq \operatorname{frö}_{d,s}(t) \qquad \text{for all } t \geq 0.$$

We now return to our ideal $I = \mathrm{I}(Z)$ of r general points. The chopped ideal $I_{\langle d \rangle}$ can not have the same Hilbert function as an ideal generated by generic forms, since the latter has Hilbert polynomial 0. On the other hand, our chopped ideal $I_{\langle d \rangle}$ is a subideal of I, and every component I_t, $t > d$, generates the entire ideal I in degrees $\geq t$. In particular, as soon as $\operatorname{hf}_{S/I_{\langle t \rangle}}(t) \leq r$, the Hilbert function of $S/I_{\langle t \rangle}$ must stabilize at r as well. Combining these two observations leads us to conjecture the following:

Conjecture 1.3.10 (Expected Syzygy Conjecture). Let Z be a general set of r points in $\mathbb{P}^n$, let $d \in \mathbb{N}$ be the smallest integer for which $s := \binom{n+d}{n} - r \geq 0$. Then for any $t \geq 0$

$$\operatorname{hf}_{S/I_{\langle d \rangle}}(t) = \begin{cases} \displaystyle\sum_{k=0}^{s} (-1)^k \binom{s}{k} \cdot \operatorname{hf}_S(t - kd) & t < t_0 \\ r & t \geq t_0, \end{cases} \tag{1.2}$$

where $t_0 > d$ is the smallest integer such that the sum is less than or equal to r. ⍰

We show that this function is a *lexicographic lower bound* on the actual Hilbert function in the following sense:

Proposition 1.3.11. *Let $I \subseteq S$ be of dimension 1, degree $r < \operatorname{hf}_S(d) - n$ and Hilbert function (1.1). If $\mathrm{ESC}_{n,r}(t)$ is the Hilbert function predicted by Conjecture 1.3.10, then*

$$\operatorname{hf}_{S/I_{\langle d \rangle}}(t) \geq_{\text{lex}} \mathrm{ESC}_{n,r}(t)$$

If Fröberg's Conjecture 1.3.8 is true for $\operatorname{hf}_S(d) - r$ general forms in S_d, then pointwise inequality holds.

Proof. By [Frö85, Lem. 1], $\operatorname{hf}_{S/I_{\langle d \rangle}}(t) \geq \operatorname{hf}_{S/J}(t)$ where J is generated by $\operatorname{hf}_I(d)$ general forms of degree d. Applying Fröberg's Theorem 1.3.9 from above to J, which has dimension and depth 0, we obtain $\operatorname{hf}_{S/I_{\langle d \rangle}} \geq_{\text{lex}} \operatorname{frö}_{d,\operatorname{hf}_I(d)}$. If Fröberg's conjecture is true, then pointwise inequality holds.

Furthermore, if $\operatorname{hf}_{S/I_{\langle d \rangle}}(t') \leq r$ for some minimal $t_0 > d$, then $(I_{\langle d \rangle})_{t_0} = I_{t_0}$. Since the minimal generators of $\mathrm{I}(Z)$ are in degree at most $d+1 \leq t_0$, we have $(I_{\langle d \rangle})_t = I_t$ for $t \geq t_0$ and hence $\operatorname{hf}_{S/I_{\langle d \rangle}}$ "sticks" to r from that point onwards. □

Remark 1.3.12. The Expected Syzygy Conjecture 1.3.10 predicts that for Z general, the Hilbert function $\mathrm{hf}_{S/\mathrm{I}(Z)_{\langle d\rangle}}$ satisfies Proposition 1.3.11 with equality. In particular, the multiplication map $\mu_e : I_d \otimes_{\mathbb{K}} S_e \to I_{d+e}$ is either surjective onto I_{d+e}, or it achieves the largest conceivable rank from Proposition 1.3.11:

$$\mathrm{hf}_{I_{\langle d\rangle}}(d+e) = \sum_{k=1}^{s}(-1)^{k+1} \cdot \mathrm{hf}_S(d+e-kd) \cdot \binom{\mathrm{hf}_S(d)-r}{k}$$

until this sum rises above $\mathrm{hf}_I(d+e)-r$, from which point on $\mathrm{hf}_{I_{\langle d\rangle}}(d+e) = \mathrm{hf}_I(d+e)$. In this sense, (ESC) can be considered as a generalization of the Ideal Generation Conjecture 1.3.3, which only concerns the first multiplication map μ_1. ◁

1.3.3 Verifying (ESC) using computer algebra

Before investigating the Expected Syzygy Conjecture 1.3.10 in more detail, we first conduct a computational study to verify the conjecture in many cases.

Theorem 1.3.13. *Conjecture 1.3.10 holds for a set of r general points in $\mathbb{P}^n(\mathbb{C})$, for all values of n, r displayed in Table 1.1.*

n	2	3	4	5	6	7	8	9	10
r	≤ 2343	≤ 2296	≤ 1815	≤ 1272	≤ 908	≤ 767	≤ 479	≤ 207	≤ 267

Table 1.1: Values for which Conjecture 1.3.10 is proven to be correct using computer algebra.

Proof. For every (n, r) of interest, we exhibit an instance $Z \in (\mathbb{P}^n(\mathbb{Q}))^r$ with generic Hilbert function (1.1) and for which the Hilbert function of $\mathrm{I}(Z)_{\langle d\rangle}$ satisfies (1.2). These instances can be found online at

https://mathrepo.mis.mpg.de/ChoppedIdeals/.

This guarantees that the corresponding open set U of the following Proposition 1.3.14 is non-empty, and therefore it is Zariski-dense. This shows that for a general instance $Z \in (\mathbb{P}^n(\mathbb{C}))^r$, the Hilbert function of Z satisfies Conjecture 1.3.10 holds. □

Proposition 1.3.14 (Semi-continuity of chopped ideals)**.** *Let $r < \mathrm{hf}_S(d) - n$ and let $U_{\mathrm{genHF}} \subseteq (\mathbb{P}^n)^r$ be the collections of points with generic Hilbert function (1.1).*

The set $U_{k,e} = \{ Z \in U_{\mathrm{genHF}} \mid \mathrm{hf}_{\mathrm{I}(Z)_{\langle d\rangle}}(d+e) \geq k \}$ is Zariski open in $(\mathbb{P}^n)^r$. In particular, the set $U = \{ Z \in (\mathbb{P}^n)^r \mid \mathrm{hf}_{\mathrm{I}(Z)_{\langle d\rangle}}$ satisfies (1.2)$\}$ is Zariski open.

We postpone the proof to the next section 1.4.1. We now describe in more detail how to verify that a (randomly chosen) collection of points $(\mathbb{P}^n(\mathbb{Q}))^r$ satisfies (1.2).

Notice that it suffices to check that $\mathrm{hf}_{\mathrm{I}(Z)_{\langle d\rangle}}(d+e)$ agrees with the expected Hilbert function (1.2) for e up to the expected saturation gap of Conjecture 1.2.16 (see the proof of Proposition 1.3.11). To speed up the computations, we make the following

observation. Let $I \subseteq S = \mathbb{C}[x_0, \ldots, x_n]$ be an ideal generated by polynomials $f_1, \ldots, f_s$ with coefficients in $\mathbb{Z} \subseteq \mathbb{C}$. Then

$$\dim_{\mathbb{C}} I_t = \dim_{\mathbb{Q}}(I \cap \mathbb{Q}[x_0, \ldots, x_n])_t \geq \dim_{\mathbb{F}_p}(I_{\mathbb{F}_p})_t.$$

Here $I_{\mathbb{F}_p} \subseteq \mathbb{F}_p[x_0, \ldots, x_n]$ is the reduction modulo p of the ideal I. Checking that the upper bound (1.2) is attained is much easier over a finite field, and it guarantees the bound is attained over $\mathbb{Q}$, hence over $\mathbb{C}$. This leads to the following strategy for proving Theorem 1.3.13.

First implement the expected Hilbert function and the expected gap size according to Conjecture 1.3.10 and Conjecture 1.2.16, here called `expectedHF`(n, r, t) and `expectedGapSize`(n, r). Next, execute the following steps for given (n, r):

(1) Determine $d := \min\{\, t \mid \mathrm{hf}_S(t) \geq r \,\}$.

(2) Sample r points $Z \subseteq \mathbb{P}^n(\mathbb{F}_p)$ (represented by homogeneous integer coordinates).

(3) Calculate the ideal $I := \mathrm{I}(Z)$ and set $J := \langle I_d \rangle_{\mathbb{F}_p[x_0,\ldots,x_n]}$.

(4) Calculate the Hilbert function of S/J up to $d +$ `expectedGapSize`(n, r).

(5) Check if the values match with `expectedHF`(n, r, t).

Example 1.3.15. The following code in `Macaulay2`, assuming an implementation of `expectedGapSize` and `expectedHF`, demonstrates the procedure. In this example 41 random points in $\mathbb{P}^2(\mathbb{F}_{1009})$ are chosen and the Hilbert series of their chopped ideal is computed.

```
loadPackage "Points"
n = 2; r = 41;
d = 8; -- determined by (n,r)
I = points randomPointsMat(ZZ/1009[x_0..x_n], r);
e = expectedGapSize(n,r) -- 3
J = ideal select(first entries gens I, f -> degree f == {d});
hs = hilbertSeries(J, Order=>d+e+1) --
↪  1 + 3T + 6T^2 + 10T^3 + 15T^4 + 21T^5 + 28T^6 + 36T^7 + 41T^8 + 43T^9 + 42T^10 + 41T^11
T = (ring hs)_0;
for t to d+e do assert (coefficient(T^t, hs) == expectedHF(n,r,t))
```

This computation terminates within a fraction of a second. Using the same code to test 198 points in $\mathbb{P}^6(\mathbb{F}_{1009})$ ($d = 4$, gap size $e_0 = 5$), it takes about 3 seconds to construct I and 141 seconds to expand $\mathrm{HS}_{S/I_{\langle 4 \rangle}}(T)$ to degree 9. Larger instances from Theorem 1.3.13 can take much longer to compute and the verification was performed on a compute server. ◁

We briefly discuss some additional speed-ups. In the cases where all minimal generators are expected to be in degree d, Conjecture 1.3.10 is equivalent to the IGC. Here it is *much* faster to calculate `mingens`$(\mathrm{I}(Z))$ and compare with the expected first graded Betti numbers $\beta_{1,d}$, $\beta_{1,d+1}$. Also, much computation time is spent constructing the ideal of points. For large r, a significant speedup is obtained when using the methods implemented in `Points.m2` [Sti+].

We conclude with a variation on the computer experiment. Instead of computing ideals of (random) points one can also try to find monomial ideals certifying Conjecture 1.3.10. This method, for instance, can be used to prove the MRC in $\mathbb{P}^2$ [Sau85; GGR86]. An exhaustive search is possible in $\mathbb{P}^2$ for small values of r and leads to the following result:

Proposition 1.3.16. *Let $n = 2$, $S = \mathbb{K}[x, y, z]$.*

(i) *For $r = 18$ there is a unique, up to permutation of the variables, monomial ideal $I = \langle x^3y^2, y^3z^2, z^3x^2, x^2y^2z^2\rangle_S$ with Hilbert function (1.1), which satisfies Conjecture 1.3.10.*

(ii) *For $r \in \{25, 32, 33\}$ there are* no *monomial ideals satisfying Conjecture 1.3.10.*

1.4 The saturation gap

In this section we take a closed look at the saturation gap e_0, the smallest positive integer such that $(I_{\langle d\rangle})_{d+e_0} = I_{d+e_0}$ (where I is the ideal of a general set of points Z as before). We answer the question about when this gap is finite using the projective geometry of the Veronese variety. We then study the regularity of the chopped ideal.

In the previous section, we determined the integer $d = \mathrm{reg}_{\mathrm{H}}(Z) = \min\{\, j \mid \mathrm{hf}_S(j) \geq j \,\}$ from n and r. In this section, we switch the roles: We fix n, d and study the chopped ideal $I(Z)_{\langle d\rangle}$ depending on r.

1.4.1 Geometry of the chopping map

Denote by $U_{\mathrm{genHF}} \subseteq (\mathbb{P}^n)^r$ the dense Zariski open subset of $(\mathbb{P}^n)^r$ consisting of r-tuples satisfying (1.1). We focus on the chopped ideals $I_{\langle d\rangle}$ where $I = \mathrm{I}(Z)$ for some $Z \in U_{\mathrm{genHF}}$. Moreover, we are interested in the case where Z can be computed from its chopped ideal. To this end, we determine the values of r for which $\mathrm{I}(Z)$ and $I_{\langle d\rangle}$ define the same subscheme of $\mathbb{P}^n$.

Theorem 1.4.1. *Let $d \geq 1$, let Z be a general set of $r \leq \mathrm{hf}_S(d)$ points in $\mathbb{P}^n$ and $I_{\langle d\rangle} \coloneqq \mathrm{I}(Z)_{\langle d\rangle}$.*

(i) *If $r = \mathrm{hf}_S(d) - n$, then $\mathbb{V}(I_{\langle d\rangle})$ is a set of d^n reduced points.*

(ii) *If $r \geq \mathrm{hf}_S(d) - n$, then $\mathbb{V}(I_{\langle d\rangle})$ is a complete intersection of codimension $\mathrm{hf}_S(d) - r$.*

(iii) *Assume $\mathrm{char}\,\mathbb{K} = 0$. If $r < \mathrm{hf}_S(d) - n$, then $\mathbb{V}(I_{\langle d\rangle})$ is the reduced scheme Z.*

For the proof, we consider a geometric version of the operation of *chopping* an ideal.

Definition 1.4.2 (Chopping map $\mathfrak{c}$)**.** For given d with $\operatorname{hf}_S(d) \geq r$, the *chopping map* is

$$\mathfrak{c} \colon U_{\text{genHF}} \to \operatorname{Gr}(\operatorname{hf}_S(d) - r, S_d), \qquad Z \mapsto [\mathrm{I}(Z)_d].$$

The chopping map is a morphism of varieties. In fact, there is a commutative diagram involving the Veronese embedding $\nu_d \colon \mathbb{P}^n \hookrightarrow \mathbb{P}(S_d^\vee)$

$$\begin{array}{ccc} \mathbb{P}(S_d^\vee)^r \supseteq (\nu_d(\mathbb{P}^n))^r & \xrightarrow{\text{span}} & \operatorname{Gr}(r, S_d^\vee) \\ {\scriptstyle (\nu_d)^{\times r}}\uparrow & & \downarrow{\scriptstyle \cong} \\ (\mathbb{P}^n)^r \supseteq U_{\text{genHF}} & \xrightarrow{\mathfrak{c}} & \operatorname{Gr}(\operatorname{hf}_S(d) - r, S_d). \end{array}$$

For a linear space $T \subseteq S_d$, the scheme $\mathbb{V}(T) \subseteq \mathbb{P}^n$ is the intersection $\nu_d(\mathbb{P}^n) \cap \mathbb{P}(T^\perp) \subseteq \mathbb{P}(S_d^\vee)$ under the identification induced by the Veronese embedding, see, e.g. [Lan12, Prop. 4.4.1.1]. Notice that $\mathfrak{c}$ is invariant under permutation of the factors of $(\mathbb{P}^n)^r$, therefore it induces a map on the quotient $\tilde{\mathfrak{c}} \colon U_{\text{genHF}}/\mathfrak{S}_r \to \operatorname{Gr}(\operatorname{hf}_S(d) - r, S_d)$.

Proposition 1.4.3 (Fibers of $\mathfrak{c}$)**.** *Let $r, n, d > 0$ be integers with $r \leq \operatorname{hf}_S(d)$.*

(i) *If $r \geq \operatorname{hf}_S(d) - n$, then $\mathfrak{c}$ is dominant. The general fibers of $\mathfrak{c}$ have dimension $nr - r(\operatorname{hf}_S(d) - r)$.*

(ii) *Assume $\operatorname{char} \mathbb{K} = 0$. If $r \leq \operatorname{hf}_S(d) - n$, then $\mathfrak{c}$ is generically finite. The induced map $\tilde{\mathfrak{c}}$ has degree $\binom{d^n}{r}$ if $r = \operatorname{hf}_S(d) - n$, it is generically injective otherwise.*

Proof of Theorem 1.4.1 and Proposition 1.4.3. First consider the case $r = \operatorname{hf}_S(d) - n$, where $\Lambda = \mathfrak{c}(Z) \subseteq \mathbb{P}(S_d^\vee)$ is a complementary-dimensional subspace to $\nu_d(\mathbb{P}^n)$. A general linear space $\Lambda \in \operatorname{Gr}(r, S_d^\vee)$ intersects $\nu_d(\mathbb{P}^n)$ in a non-degenerate set of reduced points. Picking r points on $\nu_d(\mathbb{P}^n)$ spanning Λ, we see that the map $\mathfrak{c}$ is dominant. Furthermore, $\Lambda \cap \nu_d(\mathbb{P}^n)$ consists of $\deg \nu_d(\mathbb{P}^n) = d^n$ reduced points. By genericity, any subset of r points span Λ, hence $\tilde{\mathfrak{c}}^{-1}(\Lambda^\perp)$ consists of $\binom{d^n}{r}$ points in $U_{\text{genHF}}/\mathfrak{S}_r$.

Next, let $r > r'$ where $r' \coloneqq \operatorname{hf}_S(d) - n$, and let $U' \subseteq (\mathbb{P}^n)^{r'}$ be the open set from the previous case. For any set of r points $Z \in U_{\text{genHF}}$ containing a subset Z' belonging to U', we must have $\dim \mathbb{V}(\mathrm{I}(Z)_d) = \#(Z \setminus Z')$. Indeed, modulo $\mathrm{I}(Z)_d$, there are $\#(Z \setminus Z')$ linearly independent elements in $\mathrm{I}(Z')_d$: if $\dim \mathbb{V}(\mathrm{I}(Z)_d) > \#(Z \setminus Z')$ then these additional equations could not cut out the 0-dimensional set of points Z', in contradiction with the previous part of the proof. Since S is graded Cohen-Macaulay, this implies that a basis of $\mathrm{I}(Z)_d$ is a regular sequence (Theorem 0.1.24). This shows that for any general enough Z, the variety $\mathbb{V}(\mathrm{I}(Z)_d)$ is a complete intersection of dimension $r - (\operatorname{hf}_S(d) - n)$. Proving that the chopping map is dominant is done exactly as in the previous case, the claim about the fiber dimension is a dimension count.

Finally, consider $r < \operatorname{hf}_S(d) - n$. We give a proof valid in characteristic 0. The Multisecant Lemma 0.3.6 states that for a projective variety $X \subseteq \mathbb{P}^N$ and for almost

all $(p_0, \dots, p_k) \in X^{k+1}$, $k < \operatorname{codim} X$, the $(k+1)$-secant spanned by the points scheme-theoretically intersects X only in the p_i:

$$\langle p_0, \dots, p_k \rangle_{\mathbb{P}} \cap X = \{p_0, \dots, p_k\}.$$

Applying this to the Veronese variety, for a general set of $r < \operatorname{hf}_S(d) - n$ points Z, we have

$$\nu_d(\mathbb{V}(\mathrm{I}(Z)_d)) = \nu_d(\mathbb{P}^n) \cap \mathbb{P}(\mathrm{I}(Z)_d^{\perp}) = \nu_d(\mathbb{P}^n) \cap \langle \nu_d(Z) \rangle_{\mathbb{P}} = \nu_d(Z).$$

This shows that $\mathbb{V}(\mathrm{I}(Z)_d) = Z$ for general Z. In particular, $\tilde{\mathfrak{c}}$ is generically injective. □

This answers our question on when Z can be recovered from the chopped ideal.

Corollary 1.4.4. *With notation as before, for general Z the following are equivalent:*

(i) $\mathbb{V}(\mathrm{I}(Z)_d) = Z$, *as reduced schemes;*

(ii) $(I_{\langle d \rangle})^{\mathrm{sat}} = \mathrm{I}(Z)$, *so* $(I_{\langle d \rangle})_t = I_t$ *for* $t \gg 0$;

(iii) $r < \operatorname{hf}_S(d) - n$ *or* $r = 1$ *or* $(n, r) = (2, 4)$.

Proof. By the projective Nullstellensatz, we have $J^{\mathrm{sat}} = \mathrm{I}(\mathbb{V}(J))$, this shows the equivalence of (i) and (ii). If $r < \operatorname{hf}_S(d) - n$, then by Theorem 1.4.1 $\mathbb{V}(\mathrm{I}(Z)_{\langle d \rangle}) = Z$. If $r \geq \operatorname{hf}_S(d) - n$, then $\mathbb{V}(\mathrm{I}(Z)_{\langle d \rangle}) \supseteq Z$ is a complete intersection, which is the case for general Z only if Z is a single point or four points in $\mathbb{P}^2$. □

Inspecting the proof of Theorem 1.4.1 and 1.4.3, we make a useful observation.

Corollary 1.4.5. *For n, d, r such that $\operatorname{hf}_S(d) - n \leq r < \operatorname{hf}_S(d)$ and Z general, a general collection of polynomials $f_1, \dots, f_s \in \mathrm{I}(Z)_d$ is a regular sequence when $s \leq \operatorname{hf}_S(d) - r$. If $s = n = \operatorname{hf}_S(d) - r$, then this complete intersection is reduced according to Theorem 1.4.1*(i).

Remark 1.4.6. The assumption $\operatorname{char} \mathbb{K} = 0$ is only needed to appeal to the Multisecant Lemma. If the Multisecant Lemma holds for $\operatorname{hf}_S(d) - n - 1$ points on $\nu_d(\mathbb{P}^n) \subseteq \mathbb{P}(S_d^\vee)$, then the results holds in arbitrary characteristic. In fact, this will be the only time in this chapter where we actively use a characteristic 0 assumption. However, some results below rely on Theorem 1.4.1, which forces us to add the characteristic 0 assumption. ◁

We use chopping map $\mathfrak{c}$ to complete the proof of Theorem 1.3.13 by proving the semi-continuity of the chopped ideal Hilbert function. The key idea is to realize the condition $\operatorname{hf}_{I_{\langle d \rangle}}(d+e) \leq k$ as a degeneracy condition of a vector bundle map.

Proof of Proposition 1.3.14. Let $\mathrm{Gr}(\operatorname{hf}_S(d) - r, S_d)$ be the Grassmannian of planes of dimension $\operatorname{hf}_S(d) - r$ in S_d. Consider the vector bundle $\mathcal{E} := \operatorname{Hom}(\mathcal{S} \otimes S_e, S_{d+e})$, where $\mathcal{S}$ denotes the subspace bundle over $\mathrm{Gr}(\operatorname{hf}_S(d) - r, S_d)$: the fiber of $\mathcal{E}$ at a plane $[L]$ is $\mathcal{E}_L = \operatorname{Hom}(L \otimes_{\mathbb{K}} S_e, S_{d+e})$. The multiplication map $\mu_e \colon S_d \otimes_{\mathbb{K}} S_e \to S_{d+e}$ defines a global section of $\mathcal{E}$ via restriction. The pull-back $\mathfrak{c}^*\mathcal{E}$ of $\mathcal{E}$ via the chopping

map defines a bundle over U_{genHF} and $\mathfrak{c}^*(\mu_e)$ defines a global section of $\mathfrak{c}^*\mathcal{E}$. The set $U_{k,e} = \{ Z \in U_{\text{genHF}} \mid \operatorname{hf}_{\mathrm{I}(Z)_{\langle d \rangle}}(d+e) \geq k \}$ is the complement of the degeneracy locus

$$V_{k,e} \coloneqq \left\{ Z \in U_{\text{genHF}} \;\middle|\; \operatorname{rank}\left(\mu_e \colon \mathrm{I}(Z)_d \otimes_{\mathbb{K}} S_e \to S_{d+e}\right) < k \right\} = D_{k-1}(\mathfrak{c}^*\mu_e).$$

This shows that $V_{k,e}$ is Zariski closed, hence $U_{k,e}$ is Zariski open.

The set $U = \{ Z \in (\mathbb{P}^n)^r \mid (1.2) \text{ holds} \}$ is the intersection of U_{genHF} with finitely many open sets $U_{k_e,e}$ for $e = 1, \dots, m_r$. Here m_r is any upper bound on the saturation gap, for instance $m_r = (n-1)d - (n+1)$ from Corollary 1.6.5; k_e is the expected value of $\operatorname{hf}_{I_{\langle d \rangle}}(d+e)$ in (1.2). This concludes the proof. □

1.4.2 Regularity of the chopped ideal

We now introduce what we call the *interesting range* for given n, d. Corollary 1.4.4 tells us that for given n, d, a set of r general points can be recovered from the degree d component of its ideal if and only if $r < \operatorname{hf}_S(d) - n$ (assuming $\operatorname{char} \mathbb{K} = 0$). On the other hand, by Lemma 1.3.2, $\mathrm{I}(Z)$ has generators in degree $d+1$ if $\operatorname{hf}_S(d+1) - r - (\operatorname{hf}_S(d) - r)(n+1) > 0$, or equivalently

$$r > \frac{(n+1)\operatorname{hf}_S(d) - \operatorname{hf}_S(d+1)}{n}.$$

Lemma 1.4.7 (The interesting range)**.** *Assume* $\operatorname{char} \mathbb{K} = 0$ *and let* $n, d > 0$*. If*

$$\frac{(n+1)\operatorname{hf}_S(d) - \operatorname{hf}_S(d+1)}{n} < r < \operatorname{hf}_S(d) - n, \tag{1.3}$$

then a general set of r *points in* $\mathbb{P}^n$ *has Hilbert regularity* d*,* $\mathbb{V}(\mathrm{I}(Z)_d) = Z$ *and* $I_{\langle d \rangle} \subsetneq I$*. If the Ideal Generation Conjecture 1.3.3 holds (for* n, r*), then the lower bound is tight.*

From now on we will only consider r in the interesting range.

Remark 1.4.8. Note that $\frac{(n+1)\operatorname{hf}_S(d) - \operatorname{hf}_S(d+1)}{n} \geq \operatorname{hf}_S(d-1)$. In particular, in the interesting range, equations of degree d are equations of minimal degree, as in Section 1.3. ◁

The Expected Syzygy Conjecture 1.3.10 predicts the precise value of the saturation gap e_0, or equivalently, of the saturation degree

$$\operatorname{sat.deg} I_{\langle d \rangle} = \min \left\{ t_0 \;\middle|\; (I_{\langle d \rangle})_t = I_t \text{ for all } t \geq t_0 \right\} = d + e_0$$

It is worthwhile to compare this bound to general saturation bounds, in this direction a useful result is the following result by Ein, Hà and Lazarsfeld:

Theorem 1.4.9 ([EHL22, Theorem A])**.** *Assume* $J = \langle f_0, \dots, f_p \rangle \subseteq \mathbb{C}[x_0, \dots, x_n]$ *is generated by forms of degree* $d_0 \geq \cdots \geq d_p$ *and assume that the scheme* $\mathbb{V}(J) \subseteq \mathbb{P}^n$ *is smooth. Then the saturation degree of* J *is at most* $d_0 + \cdots + d_n - n$*.*

In our situation, J is generated by $h_S(d) - r$ forms in degree d, and by Corollary 1.4.4 the chopped ideal defines the reduced points (at least in characteristic 0). Here the theorem yields the saturation degree bound $d(n+1) - n$, so $e_0 \leq (d-1)n$. Below in

Corollary 1.6.5, we will prove the sharper bound $e_0 \leq (d-1)n - (n+1)$ for our chopped ideals, which is optimal for $r = \mathrm{hf}_S(d) - (n+1)$.

We close this section by connecting the Hilbert regularity of $S/I_{\langle d \rangle}$ to its Castelnuovo–Mumford regularity.

Theorem 1.4.10. *Let $J \subseteq S$ be a one-dimensional graded ideal, then*

$$\operatorname{reg}_{\mathrm{CM}} S/J \;=\; \max\{\operatorname{reg}_{\mathrm{H}} S/J - 1, \operatorname{reg}_{\mathrm{H}} S/J^{\mathrm{sat}}\}. \tag{1.4}$$

Applying this theorem to a chopped ideal of a general set of points, we obtain:

Corollary 1.4.11. *Assume* $\operatorname{char} \mathbb{K} = 0$ *and let n, r, d satisfy (1.3). Then for a general set of r points*

$$\operatorname{reg}_{\mathrm{CM}} S/I_{\langle d \rangle} \;=\; \operatorname{reg}_{\mathrm{H}} S/I_{\langle d \rangle} - 1.$$

The Conjecture 1.2.16 predicts the Hilbert regularity of $I_{\langle d \rangle}$, hence this conjecture is directly related to Castelnuovo-Mumford regularity.

Proof of Theorem 1.4.10. The proof relies on local cohomology. Let $\mathfrak{m} = \langle x_0, \ldots, x_n \rangle_S$. The 0-th local cohomology group measures non-saturatedness as

$$\mathrm{H}^0_{\mathfrak{m}}(S/J) = \left\{ x + J \in S/J \;\middle|\; \mathfrak{m}^k x \subseteq J,\ k \gg 0 \right\} = J^{\mathrm{sat}}/J.$$

The Krull dimension of S/I can be characterized as the largest $i \geq 0$ with $\mathrm{H}^i_{\mathfrak{m}}(S/I) \neq 0$ (Proposition 0.1.30), so all cohomology groups $\mathrm{H}^i_{\mathfrak{m}}(S/J)$ vanish for $i \geq 2$.

We next provide a description of the remaining first local cohomology group. Let $I = J^{\mathrm{sat}}$ and let $Z = \operatorname{Proj} S/I \subseteq \mathbb{P}^n$ be the scheme defined by I. The quotients S/I and S/J have the same higher local cohomology. The comparison sequence for local and sheaf cohomology

$$0 \longrightarrow S/I \longrightarrow \bigoplus_{d \in \mathbb{Z}} \underbrace{\mathrm{H}^0(\mathbb{P}^n, \mathcal{O}_Z(d))}_{\cong \mathbb{K}^{\deg Z}} \longrightarrow \mathrm{H}^1_{\mathfrak{m}}(S/I) \longrightarrow 0.$$

shows that $\mathrm{H}^1_{\mathfrak{m}}(S/I)_d$ is the cokernel of $(S/I)_d \hookrightarrow \mathrm{H}^0(\mathbb{P}^n, \mathcal{O}_Z(d))$. Introducing the notation $\operatorname{end}(N) := \sup\{\, t \in \mathbb{Z} \mid N_t \neq 0 \,\}$, this shows that $\operatorname{end}(\mathrm{H}^1_{\mathfrak{m}}(S/J)) + 1 = \operatorname{reg}_{\mathrm{H}} S/I =: d$.

Now the Castelnuovo-Mumford regularity $\operatorname{reg}_{\mathrm{CM}} M$ can be expressed in terms of local cohomology as $\max\{\operatorname{end}(\mathrm{H}^i_{\mathfrak{m}}(M)) + i \mid i \in \mathbb{Z}\}$ (Theorem 0.1.29). For $M = S/J$ using vanishing in degree $i \geq \dim S/J = 1$, this gives

$$\operatorname{reg}_{\mathrm{CM}} S/J \;=\; \max\{\operatorname{end}(I/J), d\}. \tag{1.5}$$

To relate this to the maximum in equation (1.4), we distinguish two cases:

- If $\operatorname{end}(I/J) \geq d$, then $\mathrm{hf}_{S/J}(t) > \mathrm{hf}_{S/I}(t)$ for some $t \geq d$. In this case $\operatorname{end}(I/J) = \operatorname{reg}_{\mathrm{H}} S/J - 1$.
- Otherwise $\operatorname{end}(I/J) + 1 \leq d$, then $\operatorname{reg}_{\mathrm{H}} S/J \leq \operatorname{reg}_{\mathrm{H}} S/I$ and the maximum in (1.5) is attained at d. □

1.5 Points in the plane

When $n = 1$, Z is a set of points on the projective line. In this case $\mathrm{I}(Z)$ is a principal ideal and it always coincides with its chopped ideal. In particular, Conjecture 1.3.10 and Conjecture 1.2.16 hold trivially, as well as the IGC and the MRC.

This section studies the case $n = 2$, that is when $Z \in (\mathbb{P}^2)^r$ is a collection of r general points in the plane. Figure 1.2 shows the saturation gaps for some values of r. In each case we use the chopped ideal $I_{\langle d \rangle}$ in degree $d = \min\{\, t \mid \mathrm{hf}_S(t) \geq r \,\}$. The gap is only plotted in cases where $I_{\langle d \rangle}$ defines Z scheme-theoretically, following Corollary 1.4.4.

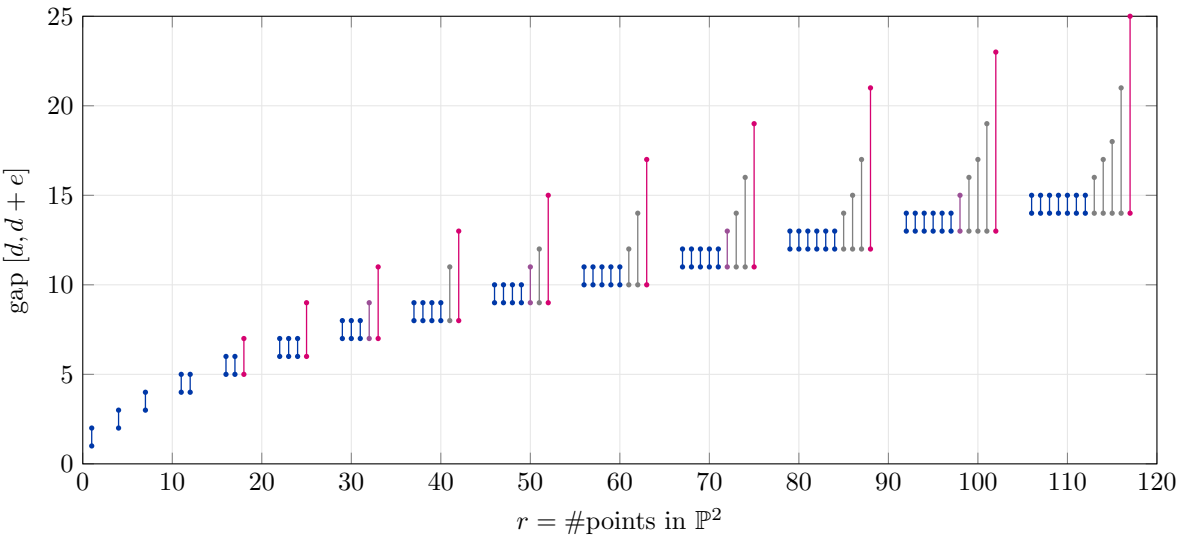


Figure 1.2: Saturation gaps for chopped ideals of points in the plane.

Since the IGC is known to be true in $\mathbb{P}^2$, the interesting range from Lemma 1.4.7 provides exactly the range where $I_{\langle d \rangle} \neq \mathrm{I}(Z)$, and they both define the set Z as a scheme:

$$\frac{d(d+2)}{2} < r < \frac{(d+2)(d+1)}{2} - 2. \tag{1.6}$$

If $d < 5$, this range is empty and the corresponding gaps in Figure 1.2 have length 1. For $d = 5$, the only integer solution to (1.6) is $r = 18$. This is the leftmost length-two gap in Figure 1.2. Hence, the simplest interesting chopped ideal is that of *three quintics passing through 18 general points in the plane*, first encountered in Example 1.2.15. In Section 1.5.1, we thoroughly work out this instructive example. For $d \geq 5$, write

$$r_{d,\min} := \left\lfloor \frac{d(d+2)}{2} \right\rfloor + 1 \qquad \text{and} \qquad r_{d,\max} = \frac{(d+2)(d+1)}{2} - 3$$

for the extremal values in the range (1.6). Section 1.5.2 proves Conjecture 1.3.10 for $r = r_{d,\max}$, and Section 1.5.3 proves it for $r = r_{d,\min}$, when d is odd. In Figure 1.2, $r_{d,\max}$ corresponds to the pink rightmost gaps of each group, while $r_{d,\min}$ corresponds to the purple leftmost gap of length > 1. Throughout this section, let $S = \mathbb{K}[x_0, x_1, x_2]$.

1.5.1 A tale of 18 points

> The art of doing mathematics consists in finding that special case which contains all the germs of generality.
>
> DAVID HILBERT [DR88]

Let $Z = (z_1, \dots, z_{18}) \in (\mathbb{P}^2)^{18}$ be a configuration of 18 general points in $\mathbb{P}^2$. Equation (1.1) guarantees that Z has no equations of degree 4 and exactly $3 = 21 - 18$ equations of degree 5. Hence $\mathrm{I}(Z)_5 = \langle f_0, f_1, f_2 \rangle_{\mathbb{K}}$ for three linearly independent quintics $f_i \in S_5$, and the chopped ideal is $I_{\langle 5 \rangle} = \langle f_0, f_1, f_2 \rangle_S$.

Notice that $\mathrm{hf}_{S/I_{\langle 5 \rangle}}(6) \geq 28 - 3 \cdot 3 = 19$ and equality holds if and only if the three quintics f_0, f_1, f_2 do not have linear syzygies. Since the IGC is true for $n = 2$, this is indeed the case. Moreover, the computation in Example 1.2.15 proves that $\mathrm{hf}_{S/I_{\langle 5 \rangle}}(7) = 18$, so the saturation gap is 2. The minimal resolution of $\mathrm{I}(Z)$ according to Proposition 1.3.5 is

$$0 \longleftarrow \mathrm{I}(Z) \longleftarrow \begin{matrix} S[-5]^3 \\ \bigoplus \\ S[-6] \end{matrix} \xleftarrow{B} S[-7]^3 \longleftarrow 0.$$

The minimal generators of $\mathrm{I}(Z)$ are the maximal minors of the Hilbert-Burch matrix B, which is a 4×3-matrix with three rows of quadrics and one row of linear forms. As a result, the maximal minors are three quintics, spanning the linear space $\mathrm{I}(Z)_5$, and one sextic, which is an element of $\mathrm{I}(Z)_6 \setminus (I_{\langle 5 \rangle})_6$. The existence of the sextic is predicted by Proposition 1.3.11 and the gap $\gamma_n(d, r) = \gamma_2(5, 18) = 2$ agrees with Conjecture 1.2.16.

It was observed in Example 1.2.15 that $(I_{\langle 5 \rangle})_7 = \mathrm{I}(Z)_7$. A dimension count shows that $\mathrm{hf}_{S/I_{\langle 5 \rangle}}(7) = 18$ if and only if the forms f_0, f_1, f_2 do not have quadratic syzygies.

Proposition 1.5.1. *Let $Z \subseteq \mathbb{P}^2$ be a set of 18 general points and let $f_0, f_1, f_2 \in \mathrm{I}(Z)_5 \subseteq S_5$ be linearly independent. Then f_0, f_1, f_2 have no quadratic syzygies. That is, $\mathrm{hf}_{S/I_{\langle 5 \rangle}}(7) = 18$.*

We present two different proofs of Proposition 1.5.1, to illustrate the idea of the more general proofs of Theorems 1.5.4, 1.5.6 and 1.6.1.

First proof. The Hilbert-Burch matrix B has the form

$$B = \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \\ \ell_0 & \ell_1 & \ell_2 \end{pmatrix} \in S^{4 \times 3},$$

where $q_{ij} \in S_2$ are quadrics, and $\ell_i \in S_1$ are linear forms. The quadratic syzygies of f_0, f_1, f_2 are the $\mathbb{K}$-linear combinations of the columns of B whose last entry is zero. If such a non-trivial $\mathbb{K}$-linear combination exists, the linear forms ℓ_0, ℓ_1, ℓ_2 are linearly dependent.

Hence, $\mathbb{V}(\ell_0, \ell_1, \ell_2) \subseteq \mathbb{P}^2$ is non-empty. The quintics f_0, f_1, f_2 are the 3×3 minors of B involving the last row, so that $\mathbb{V}(\ell_0, \ell_1, \ell_2) \subseteq \mathbb{V}(f_0, f_1, f_2)$, showing $\mathbb{V}(\ell_0, \ell_1, \ell_2)$ is one of the points in Z. The genericity of Z, together with the fact that the construction is invariant under permutation of the 18 points, leads to a contradiction; we refer to the proof of Theorem 1.5.6 for details on this construction. Hence ℓ_0, ℓ_1, ℓ_2 are linearly independent, and one concludes that f_0, f_1, f_2 do not have quadratic syzygies. □

Second proof. Alternatively, one can prove Proposition 1.5.1 via a classical liaison argument. Suppose $s_0 f_0 + s_1 f_1 + s_2 f_2 = 0$ is a quadratic syzygy of f_0, f_1, f_2, for some $s_j \in S_2$. Assume f_0, f_1, f_2 are chosen generically in $\mathrm{I}(Z)_5$. Let $K \subseteq \mathbb{P}^2$ be the set of points defined by the ideal $\langle f_1, f_2 \rangle_S$. By Bézout's theorem, and Corollary 1.4.5, K is a complete intersection of 25 reduced points, and Z is a subset of K. In other words, $K = Z \dot\cup Z'$ where Z' is a set of 7 points, called the *liaison of Z in K*. For every $z \in Z'$, we have $s_0(z) f_0(z) = 0$. We have $f_0(z) \neq 0$, otherwise $z \in Z$, and we conclude that $s_0 \in \mathrm{I}(Z')$. On the other hand, the theory of liaison guarantees that Z' has no nonzero quadratic equations; the necessary results to prove this statement are presented in Section 1.6. We conclude that $s_0 = 0$, and analogously for $s_1 = s_2 = 0$. □

We now discuss how to find the *missing sextic*, which is uniquely determined modulo the 9-dimensional linear space $(I_{\langle 5 \rangle})_6 \subseteq \mathrm{I}(Z)_6$. We provide a way to compute an element of $\mathrm{I}(Z)_6 \setminus (I_{\langle 5 \rangle})_6$ from Z.

Proposition 1.5.2. *Let $Z \subseteq \mathbb{P}^2$ be a general set of 18 points. For every partition $Z = Z_1 \dot\cup Z_2$ into two sets of 9 points, the form $g = g_1 g_2$ is in $\mathrm{I}(Z)_6 \setminus \mathrm{I}(Z)_{\langle 5 \rangle}$, where $\mathrm{I}(Z_i)_3 = \mathbb{K} g_i$.*

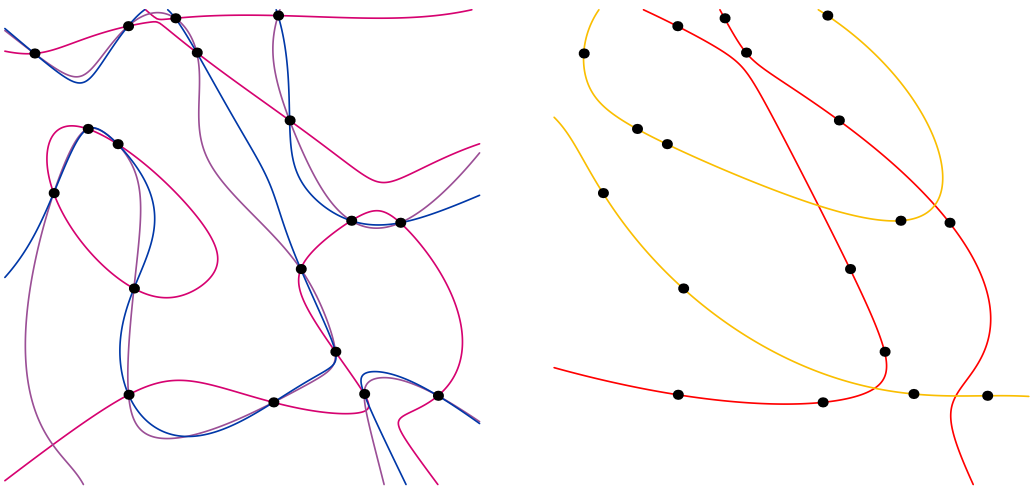


Figure 1.3: Three quintics passing through 18 general points in the plane (left) and a "missing" split sextic with two cubic components passing through 9 of the points each (right).

The construction is illustrated in Figure 1.3. To prove the result, we will use Proposition 0.3.9 about non-degeneracy of transversal sections, and the following elementary lemma, which is at the foundation of monodromy techniques:

Lemma 1.5.3. *Let $f\colon X \to Y$ be a dominant generically finite map of irreducible varieties. Let $Z \subseteq X$ be a closed subset which intersects general fibers of f. Then $Z = X$.*

Proof. Since Z intersects general fibers of f, we have that $f|_Z$ is dominant too. Since $f|_Z$ is dominant and Z is a subvariety of X, we have $\dim X \geq \dim Z \geq \dim Y$. On the other hand, since f is dominant and finite, we have $\dim X = \dim Y$. Therefore $\dim Z = \dim X$ and since X is irreducible, we conclude $Z = X$. □

We will also use that the degree of the variety

$$\mathcal{P}_{3,3} \coloneqq \{\, g \in \mathbb{P}S_6 \mid g = g_1 g_2 \text{ for some } g_1, g_2 \in S_3 \,\} \subseteq \mathbb{P}S_6$$

is $\frac{1}{2}\binom{18}{9} = 24310$. This can be computed using elementary intersection theory; see, e.g., [EH16, Sec. 2.2.2] for a similar calculation.

Proof of Proposition 1.5.2. We proceed in two steps: We first show that there is *some* partition for which $g = g_1 g_2 \in \mathrm{I}(Z)_6 \setminus \mathrm{I}(Z)_{\langle 5\rangle}$. Then, we use Lemma 1.5.3 to show that the same must hold for all partitions.

Since Z is general, we have $\dim_{\mathbb{K}} \mathrm{I}(Z)_6 = 10$, $\dim_{\mathbb{K}}(\mathrm{I}(Z)_{\langle 5\rangle})_6 = 9$. Notice $\dim \mathcal{P}_{3,3} = 9 + 9 = 18 = \operatorname{codim}_{\mathbb{P}S_6} \mathbb{P}\mathrm{I}(Z)_6$. Let $W \coloneqq \mathcal{P}_{3,3} \cap \mathbb{P}\mathrm{I}(Z)_6 \subseteq \mathbb{P}S_6$.

For every $g = g_1 g_2 \in W$, observe that $Z_i = Z \cap \{g_i = 0\}$ defines a bipartition of Z into two subsets of 9 points. Indeed, $Z = Z_1 \cup Z_2$, and no subset of 10 points in Z has a cubic equation because of the genericity assumption. On the other hand, any bipartition of $Z = Z_1 \dot\cup Z_2$ into two subsets of 9 points gives rise to an element $g = g_1 g_2 \in W$; by genericity, all these elements are distinct. This shows that $W = \mathcal{P}_{3,3} \cap \mathbb{P}\mathrm{I}(Z)_6$ is a set of $\frac{1}{2}\binom{18}{2} = \deg \mathcal{P}_{3,3}$ points. In particular, W is reduced.

By Proposition 0.3.9, W is linearly non-degenerate in $\mathbb{P}\mathrm{I}(Z)_6$, so it is not contained in the hyperplane $\mathbb{P}(\mathrm{I}(Z)_{\langle 5\rangle})_6$. This shows that at least one element of W is not contained in the chopped ideal $\mathrm{I}(Z)_{\langle 5\rangle}$.

We now show that $W \subseteq \mathbb{P}\mathrm{I}(Z)_6 \setminus \mathbb{P}(\mathrm{I}(Z)_{\langle 5\rangle})_6$. Consider the varieties

$$\mathcal{Y} \coloneqq \overline{\left\{ (\mathrm{I}(Z)_{\langle 5\rangle})_6 \in \mathrm{Gr}(9, S_6) \;\middle|\; Z \subseteq \mathbb{P}^2 \text{ is a set of 18 points in general position} \right\}},$$

$$\mathcal{X} \coloneqq \overline{\left\{ (Z_1, Z_2, g_1, g_2) \in (\mathbb{P}^2)^9 \times (\mathbb{P}^2)^9 \times \mathbb{P}S_3 \times \mathbb{P}S_3 \;\middle|\; \begin{array}{c} Z_1 \cap Z_2 = \emptyset \\ Z_1, Z_2, Z_1 \cup Z_2 \text{ in gen'l position} \\ \mathrm{I}(Z_i)_3 = \langle g_i \rangle \end{array} \right\}}.$$

These are irreducible subvarieties of the Grassmannian $\mathrm{Gr}(9, S_6)$ and of the product $(\mathbb{P}^2)^9 \times (\mathbb{P}^2)^9 \times \mathbb{P}S_3 \times \mathbb{P}S_3$ respectively. Define the dominant generically finite rational map

$$\phi\colon \mathcal{X} \dashrightarrow \mathcal{Y}, \qquad (Z_1, Z_2, g_1, g_2) \mapsto (\mathrm{I}(Z_1 \cup Z_2)_{\langle 5\rangle})_6.$$

Let $\mathcal{Z}$ be the subvariety of $\mathcal{X}$ defined by

$$\mathcal{Z} \coloneqq \overline{\left\{ (Z_1, Z_2, g_1, g_2) \;\middle|\; g_1 g_2 \in (\mathrm{I}(Z_1 \cup Z_2)_{\langle 5 \rangle})_6 \right\}} \subseteq \mathcal{X}.$$

If $W \cap \mathbb{P}((\mathrm{I}(Z)_{\langle 5 \rangle})_6) \neq \emptyset$, then $\mathcal{Z}$ would intersect general fibers of ϕ. In this case, Lemma 1.5.3 guarantees $\mathcal{X} = \mathcal{Z}$. This implies that for every $(Z_1, Z_2, g_1, g_2) \in \mathcal{Z}$, we have $g_1 g_2 \in (\mathrm{I}(Z_1 \cup Z_2)_{\langle 5 \rangle})_6$. In other words $W \subseteq \mathbb{P}(\mathrm{I}(Z)_{\langle 5 \rangle})_6$, which contradicts what we saw above. We conclude $W \cap \mathbb{P}(\mathrm{I}(Z)_{\langle 5 \rangle})_6 = \emptyset$. In other words, every $g = g_1 g_2$ satisfies $g \notin \mathrm{I}(Z)_{\langle 5 \rangle}$ as desired. □

1.5.2 The case $r = r_{d,\max}$

In this section, we prove the Expected Syzygy Conjecture 1.3.10 for the maximal number of points $r = r_{d,\max}$ in the plane. Fix $d \geq 5$ and let $Z \in (\mathbb{P}^2)^{r_{d,\max}}$ be a general collection of $r_{d,\max}$ points. By construction, $\mathrm{I}(Z)$ is zero in degree smaller than d and the chopped ideal $\mathrm{I}(Z)_{\langle d \rangle}$ has 3 generators of degree d. By Corollary 1.4.4, $\mathrm{I}(Z)_{\langle d \rangle}$ defines Z scheme-theoretically. Since the IGC holds in $\mathbb{P}^2$, the three generators of degree d have no linear syzygies and $\mathrm{I}(Z)$ has $d-4$ minimal generators of degree $d+1$. By Proposition 1.3.5, the minimal free resolution of $\mathrm{I}(Z)$ is

$$0 \longleftarrow \mathrm{I}(Z) \longleftarrow \begin{matrix} S[-d]^3 \\ \oplus \\ S[-(d+1)]^{d-4} \end{matrix} \xleftarrow{B} S[-(d+2)]^{d-2} \longleftarrow 0. \tag{1.7}$$

Conjecture 1.3.10 predicts the value for $\mathrm{hf}_{I_{\langle d \rangle}}$ in degree $d+e$:

$$\mathrm{hf}_{\mathrm{I}(Z)_{\langle d \rangle}}(d+e) = \min\left\{ 3 \cdot \binom{e+2}{2}, \binom{d+e+2}{2} - \binom{d+2}{2} + 3 \right\}.$$

For $e = d-3$, both arguments give the minimum:

$$3 \cdot \binom{d-1}{2} = \frac{3d^2 - 9d + 6}{2} = \binom{2d-1}{2} - \binom{d+2}{2} + 3.$$

Hence, we expect the map $\mu_{d-3} \colon S_{d-3} \otimes \mathrm{I}(Z)_d \to \mathrm{I}(Z)_{2d-3}$ to be an isomorphism and all intermediate multiplication maps to be injectire.

Theorem 1.5.4. *Assume* $\operatorname{char} \mathbb{K} = 0$. *Let* $Z \in (\mathbb{P}^2)^{r_{d,\max}}$ *be a collection of* $r_{d,\max}$ *general points, and let* $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle$ *be its chopped ideal. The Hilbert function of* $I_{\langle d \rangle}$ *satisfies*

$$\begin{aligned} \mathrm{hf}_{I_{\langle d \rangle}}(t) &= 0 && \text{if } t < d, \\ \mathrm{hf}_{I_{\langle d \rangle}}(t) &= 3 \cdot \mathrm{hf}_S(t-d) && \text{if } d \leq t \leq 2d-3, \\ \mathrm{hf}_{I_{\langle d \rangle}}(t) &= \mathrm{hf}_S(t) - r_{d,\max} && \text{if } t \geq 2d-3. \end{aligned}$$

In this case, Conjecture 1.3.10 and Conjecture 1.2.16 hold and $\gamma_2(d, r_{d,\max}) = d-3$.

Proof. It suffices to show that the three generators f_0, f_1, f_2 of the chopped ideal do not have any syzygies in degree $d-3$. This guarantees that the inequality in Proposition 1.3.11 is a pointwise upper bound, and in turn that equality holds. Suppose (s_0, s_1, s_2) is a syzygy of degree $d-3$:

$$s_0 f_0 + s_1 f_1 + s_2 f_2 = 0 \qquad \text{for some } s_0, s_1, s_2 \in S_{d-3}.$$

We are going to prove that $s_0 = s_1 = s_2 = 0$. By Corollary 1.4.5 we may assume that f_1, f_2 generate a complete intersection ideal defining a set K of d^2 distinct points in $\mathbb{P}^2$. The set K contains Z, and we write $Z' = K \setminus Z$ for the complement of Z in K.

It suffices to show that $\mathrm{I}(Z')_{d-3} = 0$. Indeed, for every $z \in Z'$ we have $f_1(z) = f_2(z) = 0$, which implies $f_0(z)s_0(z) = 0$. But $f_0(z) \neq 0$ because by Corollary 1.4.4 the chopped ideal defines Z scheme-theoretically and $z \notin Z$. Hence $s_0 \in \mathrm{I}(Z')_{d-3}$. If $\mathrm{I}(Z')_{d-3} = 0$, we obtain $s_0 = 0$. This implies $s_1 = s_2 = 0$ as well, because s_1, s_2 defines a syzygy of the complete intersection $\langle f_1, f_2 \rangle_S$, which does not have non-trivial syzygies in degree smaller than d.

We are left with showing that $\mathrm{I}(Z')_{d-3} = 0$. The Hilbert-Burch matrix B' of Z' can be obtained from the Hilbert-Burch matrix B of Z in (1.7) as follows; see, e.g., [Sau85, Prop. 1.3]. The entries of B in its first three rows are quadrics, and the ones in the remaining $d-4$ rows are linear forms. The second and third row of B correspond to f_1, f_2. The Hilbert-Burch matrix B' is the transpose of the submatrix obtained from B by removing the two rows corresponding to f_1, f_2. Therefore B' is a $(d-2) \times (d-3)$ matrix whose first column consists of quadratic forms, and the remaining $d-4$ columns consist of linear forms. The maximal minors of B' have degree $d-2$, and they are minimal generators of $\mathrm{I}(Z')$ by the Hilbert-Burch Theorem. In particular, $\mathrm{I}(Z')_{d-3} = 0$, as desired. □

Remark 1.5.5. Theorem 1.5.4 is a special version of Theorem 1.6.1, whose proof resorts to liaison theory in higher dimension and is less explicit. Therefore, we chose to include both proofs. ◁

1.5.3 The case $r = r_{d,\min}$ for odd d

Let $d = 2\delta + 1$ be odd, and let Z be a set of $r = r_{d,\min} = 2(\delta+1)^2$ general points in $\mathbb{P}^2$. By Proposition 1.3.5, $\mathrm{I}(Z)$ is generated by $\delta + 1$ forms of degree d and 1 form of degree $d+1$:

$$\mathrm{I}(Z) \;=\; \langle f_0, \ldots, f_\delta, g \rangle_S,$$

with $f_0, \ldots, f_\delta \in S_d$ and $g \in S_{d+1}$. In this section we prove the following result.

Theorem 1.5.6. *Assume* $\operatorname{char} \mathbb{K} = 0$. *Let* $d = 2\delta+1$ *and let* $Z \in (\mathbb{P}^2)^{r_{d,\min}}$ *be a collection of* $r_{d,\min} = 2(\delta+1)^2$ *general points. The Hilbert Function of* $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle_S$ *satisfies*

$$\mathrm{hf}_{S/I_{\langle d \rangle}}(d) = \begin{cases} r_{d,\min} & t = d \\ r_{d,\min} + 1 & t = d+1 \\ r_{d,\min} & t \geq d+2 \end{cases}$$

In this case, Conjecture 1.3.10 and Conjecture 1.2.16 hold and $\gamma_2(d, r_{d,\min}) = 2$.

The proof uses the minimal free resolution of $\mathrm{I}(Z)$, obtained from Proposition 1.3.5

$$0 \longleftarrow \mathrm{I}(Z) \longleftarrow \begin{matrix} S[-d]^{\delta+1} \\ \bigoplus \\ S[-(d+1)] \end{matrix} \xleftarrow{B} S[-(d+2)]^{\delta+1} \longleftarrow 0.$$

The Hilbert-Burch matrix B has the following form:

$$B = \begin{pmatrix} q_{00} & \cdots & q_{0\delta} \\ \vdots & & \vdots \\ q_{\delta 0} & \cdots & q_{\delta\delta} \\ \ell_0 & \cdots & \ell_\delta \end{pmatrix},$$

for some quadratic forms $q_{ij} \in S_2$ and linear forms $\ell_j \in S_1$. The degree e syzygies of $f_0, \ldots, f_\delta$ are the elements of the $\mathbb{K}$-vector space

$$\mathrm{Syz}(f_0, \ldots, f_\delta)_e = \left\{ (s_0, \ldots, s_\delta) \in (S_e)^{\delta+1} \;\middle|\; s_0 f_0 + \cdots + s_\delta f_\delta = 0 \right\}.$$

Proof of Theorem 1.5.6. The fact that $\mathrm{hf}_{S/I_{\langle d \rangle}}(d) = r_{d,\min}$ and $\mathrm{hf}_{S/I_{\langle d \rangle}}(d+1) = r_{d,\min} + 1$ follow since the IGC holds for $\mathbb{P}^2$.

For the statement on $\mathrm{hf}_{S/I\langle d \rangle}(t)$ for $t \geq d+2$, observe that $(\delta+1) \cdot \mathrm{hf}_S(2) = \mathrm{hf}_S(d+2) - r_{d,\min} + (\delta - 2)$. Hence, it suffices to show that the forms $f_0, \ldots, f_\delta$ generating the chopped ideal $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle_S$ have exactly $\delta - 2$ quadratic syzygies, i.e.

$$\dim_{\mathbb{K}} \mathrm{Syz}(f_0, \ldots, f_\delta)_2 = \delta - 2.$$

It is clear that $\dim_{\mathbb{K}} \mathrm{Syz}(f_0, \ldots, f_\delta)_2 \geq \mathrm{hf}_I(d) \, \mathrm{hf}_S(2) - \mathrm{hf}_I(d+2) = \delta - 2$; this also follows from Proposition 1.3.11. We show that $f_0, \ldots, f_\delta$ cannot have $\delta - 1$ syzygies.

The linear span $L_Z := \langle \ell_0, \ldots, \ell_\delta \rangle_{\mathbb{K}}$ of the linear forms in the last row of B does not depend on the choice of the minimal free resolution by Lemma 0.1.11(iv).

If $f_0, \ldots, f_\delta$ have $\delta - 1$ quadratic syzygies for a general choice of Z, then $\dim_{\mathbb{K}} L_Z$ is at most 2, which implies that the variety $\mathbb{V}(L_Z)$ is non-empty. Moreover, by the Hilbert-Burch Theorem, the ideal generated by L_Z contains $I_{\langle d \rangle} = \langle f_0, \ldots, f_\delta \rangle_S$, which cuts out Z scheme-theoretically by Proposition 1.4.3. Hence $\dim_{\mathbb{K}} L_Z = 2$ and $\mathbb{V}(L_Z)$ is one of the points in Z. Define

$$\psi \colon (\mathbb{P}^2)^r \dashrightarrow \mathbb{P}^2, \qquad \psi(Z) := \mathbb{V}(L_Z).$$

This is a rational map: the coordinates of L_Z, and hence of the point it defines, can be expressed as polynomial functions of the coordinates of the configuration $Z \in (\mathbb{P}^2)^r$, for instance via successive applications of Cramer's rule. Moreover, we showed that $\psi(Z) \in Z$. By definition, ψ is invariant under the action of the symmetric group $\mathfrak{S}_r$ permuting the factors of $(\mathbb{P}^2)^r$. Consider the subvarieties of $(\mathbb{P}^2)^r$ defined by

$$Y_j := \overline{\{ Z = (z_1, \ldots, z_r) \in \mathrm{Dom}(\psi) \mid \psi(Z) = z_j \}}.$$

We have $(\mathbb{P}^2)^r = \bigcup_{j=1}^r Y_j$. Since $(\mathbb{P}^2)^r$ is irreducible, we have $Y_j = (\mathbb{P}^2)^r$ for at least one j. On the other hand, $\mathfrak{S}_r$ invariance implies that if $Y_j = (\mathbb{P}^2)^r$ for one j, then this must be true for all j's. But any two Y_j are distinct because generically Z consists of distinct elements. This gives a contradiction showing that the map ψ cannot exist. We obtain that $\dim_{\mathbb{K}} L_Z = 3$ for general Z and this concludes the proof. □

1.6 The largest possible saturation gap

In this section, we consider a set Z of $r = \binom{d+n}{n} - (n+1)$ general points in $\mathbb{P}^n$. Corollary 1.4.4 guarantees this is the largest possible number of points such that the chopped ideal $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle_S$ in the ring $S = \mathbb{K}[x_0, \ldots, x_n]$ defines Z scheme-theoretically. We will show that the Expected Syzygy Conjecture 1.3.10 is true in this case:

Theorem 1.6.1. *Assume* $\operatorname{char} \mathbb{K} = 0$. *Let* n, d *be positive integers and let* $Z \subseteq \mathbb{P}^n$ *be a set of* $r = \mathrm{hf}_S(d) - (n+1)$ *general points. The Hilbert function of the chopped ideal* $I_{\langle d \rangle} = \langle \mathrm{I}(Z)_d \rangle_S$ *satisfies*

$$\mathrm{hf}_{S/I_{\langle d \rangle}}(d+e) = \sum_{k \geq 0} (-1)^k \cdot \mathrm{hf}_S(d+e-kd) \cdot \binom{n+1}{k},$$

for $e \leq (n-1)d - (n+1)$. *Conjecture 1.3.10 and Conjecture 1.2.16 hold with* $\gamma_n(d,r) = (n-1)d - (n+1)$.

As a consequence, we will obtain an upper bound on the saturation gap $\gamma_n(d,r)$ for all $r < \mathrm{hf}_S(d) - n$ in Corollary 1.6.5. The proof relies on higher-codimension liaison theory. In the first section we review the necessary results, and then proceed with the proof of Theorem 1.6.1 in the final section of this chapter.

1.6.1 Liaison

The proof of Theorem 1.6.1 relies on a fundamental fact in the theory of liaison. Given the Hilbert function of a set of points Z, one can compute the Hilbert function of the complementary set of points $Z' = K \setminus Z$ in a complete intersection $K \supseteq Z$. We record this fact in Proposition 1.6.4 below. In order to state it precisely, we introduce the following notation. The function

$$\Delta \mathrm{hf}_Z(t) \coloneqq \mathrm{hf}_Z(t) - \mathrm{hf}_Z(t-1)$$

is the *first difference* of the Hilbert function of Z. Often, $\Delta \mathrm{hf}_Z(t)$ is called the h-vector of Z and it is recorded as the sequence of its non-zero values. We record some properties immediate from Lemmas 0.1.32 and 1.3.1.

Lemma 1.6.2. *For a finite set of points* $Z \subseteq \mathbb{P}^n$, *we have*

(i) $\mathrm{hf}_Z(t) = \sum_{t'=0}^{t} \Delta \mathrm{hf}_Z(t')$;

(ii) $\operatorname{reg}_{\mathrm{H}}(Z) = \max\{\, t \mid \Delta\operatorname{hf}_Z(t) > 0 \,\}$;

(iii) *if* $\mathrm{I}(Z)_t = 0$ *then* $\Delta\operatorname{hf}_Z(t') = \binom{t'+n-1}{n-1}$ *for* $t' \leq t$.

For complete intersections K, the function $\Delta\operatorname{hf}_K$ has a symmetry property [Eis95, Ch. 17]. This can be explained using the fact that an Artininan reduction of $S/I(K)$ is a standard graded Gorenstein algebra.

Lemma 1.6.3. *Let $K \subseteq \mathbb{P}^n$ be a set of $d_1 \cdots d_n$ points whose ideal is generated by a regular sequence $f_1, \ldots, f_n$, with $\deg(f_i) = d_i$. Then $\operatorname{reg}_{\mathrm{H}}(K) = d_1 + \cdots + d_n - n$. In particular, if $d_1 = \cdots = d_n = d$, $\operatorname{reg}_{\mathrm{H}}(K) = n(d-1)$. Moreover, $\Delta\operatorname{hf}_K$ is symmetric, that is, setting $\rho = \operatorname{reg}_{\mathrm{H}}(K)$, we have $\Delta\operatorname{hf}_K(t) = \Delta\operatorname{hf}_K(\rho - t)$.*

The theory of liaison studies the relation between the distinct irreducible components (or union of such) of a complete intersection; we only illustrate one result in the context of ideals of points. We refer to [MN02; Mig98] for an extensive exposition of the subject. The following is a consequence of [Mig98, Prop. 5.2.1].

Proposition 1.6.4. *Let $Z \subseteq \mathbb{P}^n$ be a set of points and let $f_1, \ldots, f_n \in \mathrm{I}(Z)$ be homogeneous polynomials of degree $d_1, \ldots, d_n$ defining a smooth complete intersection K of degree $d_1 \cdots d_n$. Let $Z' = K \setminus Z$ be the complement of Z in K. Let $\rho = \operatorname{reg}_{\mathrm{H}}(K) = d_1 + \cdots + d_n - n$. Then $\Delta\operatorname{hf}_Z(t) + \Delta\operatorname{hf}_{Z'}(\rho - t) = \Delta\operatorname{hf}_K(t)$.*

In words, Proposition 1.6.4 says that $\Delta\operatorname{hf}_{Z'}$ equals the sequence $\Delta\operatorname{hf}_K - \Delta\operatorname{hf}_Z$, in the reversed order. The proof of this result uses a construction known as the *mapping cone* in homological algebra. The key fact is that the resolution of $\mathrm{I}(Z')$ can be obtained from that of $\mathrm{I}(Z)$ using the fact the resolution of $\mathrm{I}(K)$ is the Koszul complex of $f_1, \ldots, f_n$.

1.6.2 Proof of Theorem 1.6.1

We now have all the ingredients to give a proof of the main theorem of this section.

Proof of Theorem 1.6.1. By construction, we have $\dim_{\mathbb{K}}(I_{\langle d \rangle})_d = \dim_{\mathbb{K}} \mathrm{I}(Z)_d = n + 1$. Removing $z_0 \in Z$, a general set $f_1, \ldots, f_n \in \mathrm{I}(Z \setminus \{z_0\})_d$ defines a reduced complete intersection $K \subseteq \mathbb{P}^n$ of d^n points by Corollary 1.4.5. We add a generator f_0 such that $I_{\langle d \rangle} = \langle f_0, \ldots, f_n \rangle_S$.

The Hilbert function $\operatorname{hf}_K(t)$ has the following compact form, which can be computed directly from the dimension of the syzygy modules in the Koszul complex (Example 0.1.13):

$$\operatorname{hf}_K(t) \;=\; \sum_{k=0}^{n} (-1)^k \operatorname{hf}_S(t - kd) \binom{n}{k}.$$

In particular $\operatorname{hf}_K(t) = \operatorname{hf}_Z(t)$ for $t \leq d-1$, and $\operatorname{hf}_K(d) = \operatorname{hf}_Z(d) + 1$. Recall from Lemma 1.6.3 that $\operatorname{reg}_{\mathrm{H}}(K) = n(d-1)$. Set $\rho \coloneqq n(d-1)$ and let $Z' \coloneqq K \setminus Z$ be the complement of Z in K. By Lemma 1.6.3 and Proposition 1.6.4, we have

$$\Delta\operatorname{hf}_Z(t) + \Delta\operatorname{hf}_{Z'}(\rho - t) \;=\; \Delta\operatorname{hf}_K(t) \;=\; \Delta\operatorname{hf}_K(\rho - t).$$

Since $\Delta \operatorname{hf}_Z(t) = 0$ for $t \geq d+1$, we deduce $\Delta \operatorname{hf}_{Z'}(t) = \Delta \operatorname{hf}_K(t)$ for $t \leq \rho - (d+1)$. Since $\mathrm{I}(Z') \subseteq \mathrm{I}(K)$, this implies $\mathrm{I}(Z')_t = \mathrm{I}(K)_t$ for $t \leq \rho - (d+1)$. In particular, we obtain that

$$\dim_{\mathbb{K}} \mathrm{I}(Z')_t \;=\; \operatorname{hf}_S(t) - \operatorname{hf}_K(t) \;=\; \sum_{k\geq 1} (-1)^{k+1} \operatorname{hf}_S(t-kd) \binom{n}{k},$$

for $t \leq (n-1)d - (n+1)$. See Figure 1.4 for a schematic representation.

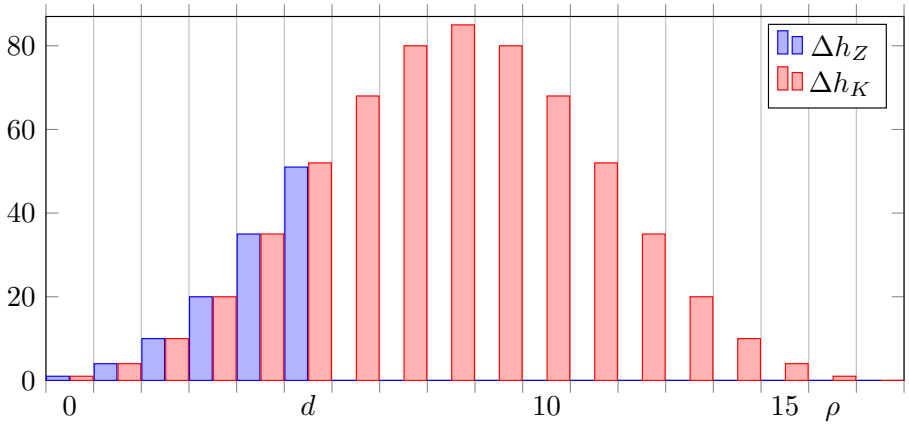


Figure 1.4: The sequences $\Delta \operatorname{hf}_Z$ and $\Delta \operatorname{hf}_K$ for $r = 121$ points in $\mathbb{P}^4$. Their difference, in reversed order, equals $\Delta \operatorname{hf}_{Z'}$. Note $\Delta \operatorname{hf}_K(5) = \Delta \operatorname{hf}_Z(5) + 1$.

Now, fix $e \leq nd - n - (d+1)$ and let $\operatorname{Syz}_e \subseteq S_e^{n+1}$ be the subspace of syzygies of degree e; i.e. the tuples $(s_0, \ldots, s_n)$ with $s_0 f_0 + \cdots + s_n f_n = 0$.

Consider the projection onto the 0-th component S_e, that is

$$\pi_0 \colon \operatorname{Syz}_e \to S_e, \qquad \pi_0(s_0, \ldots, s_n) = s_0$$

We observe that the image of this map is contained in $\mathrm{I}(Z')$. Indeed, since $Z' \subseteq K$, for every $p \in Z'$ we have $f_j(p) = 0$ for $j = 1, \ldots, n$. This implies $s_0(p) f_0(p) = 0$ for $p \in Z'$. Since $I_{\langle d \rangle}$ defines Z scheme-theoretically, we deduce that $f_0(p) \neq 0$ for every $p \in Z'$; hence $s_0(p) = 0$ for $p \in Z'$, as desired. Therefore, the image of the map π_0 is contained in $\mathrm{I}(Z')_e$.

The kernel of π_0 consists of tuples $(0, s_1, \ldots, s_n) \in S_e^{n+1}$ such that $s_1 f_1 + \cdots + s_n f_n = 0$; hence $(s_1, \ldots, s_n)$ defines a syzygy of $f_1, \ldots, f_n$. Since the ideal $\langle f_1, \ldots, f_n \rangle_S$ is a complete intersection, its only syzygies in degree e are generated by the Koszul syzygies, and we deduce

$$\dim_{\mathbb{K}}(\operatorname{Ker} \pi_0) \;=\; \sum_{k\geq 2} (-1)^k \operatorname{hf}_S(d+e-kd) \binom{n}{k}.$$

Since $\dim_{\mathbb{K}} \operatorname{Syz}_e = \dim_{\mathbb{K}}(\operatorname{Ker}\pi_0) + \dim_{\mathbb{K}}(\operatorname{Im}\pi_0) \le \dim_{\mathbb{K}}(\operatorname{Ker}\pi_0) + \dim_{\mathbb{K}} \mathrm{I}(Z')_e$, the standard identity $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ yields

$$\begin{aligned}\dim_{\mathbb{K}} \operatorname{Syz}_e &\le \sum_{k\ge 1}(-1)^{k+1} \operatorname{hf}_S(e-kd)\binom{n}{k+1} + \sum_{k\ge 1}(-1)^{k+1}\operatorname{hf}_S(e-kd)\binom{n}{k} \\ &= \sum_{k\ge 1}(-1)^{k+1}\operatorname{hf}_S(e-kd)\binom{n+1}{k+1}.\end{aligned}$$

We conclude that for $e \le (n-1)d-(n+1)$,

$$\operatorname{hf}_{I_{\langle d\rangle}}(d+e) \;=\; (n+1)\cdot \operatorname{hf}_S(e) - \dim_{\mathbb{K}} \operatorname{Syz}_e \;\ge\; \sum_{k\ge 1}(-1)^{k+1}\operatorname{hf}_S(d+e-kd)\binom{n+1}{k}.$$

The reverse inequality for $\operatorname{hf}_{I_{\langle d\rangle}}(d+e)$ follows from Fröberg's Theorem 1.3.9 which is a pointwise lower bound since Fröberg's conjecture is true in this case [Frö85, Section 3.1].

It remains to show that the inequality $\operatorname{hf}_{S/I_{\langle d\rangle}}(d+e) \ge \operatorname{hf}_S(d+e)-(n+1)$ is an equality for $e_0 \coloneqq (n-1)d-(n+1)$ and strict for e_0-1 (Proposition 1.3.11), thus establishing the length of the saturation gap e_0. We have shown that $\operatorname{hf}_{S/I_{\langle d\rangle}}(t) = \operatorname{frö}_{d,n+1}(t)$ for $t \le nd-(n+1)$, where the latter is the Hilbert function of any ideal generated by a regular sequence of $n+1$ elements in degree d. Hence it suffices to check the (in)equalities for any such complete intersection, for example the monomial ideal $M \coloneqq \langle x_0^d, \dots, x_n^d\rangle_S$. The Hilbert function of S/M has the combinatorial description

$$\operatorname{hf}_{S/M}(t) \;=\; \#\left\{ \boldsymbol{x}^\alpha \in S_t \;\middle|\; \boldsymbol{x}^\alpha \text{ divides } x_0^{d-1}\cdots x_n^{d-1} \right\}.$$

The involution $\boldsymbol{x}^\alpha \mapsto x_0^{d-1}\cdots x_n^{d-1}/\boldsymbol{x}^\alpha$ shows the symmetry $\operatorname{hf}_{S/M}(t) = \operatorname{hf}_{S/M}((n+1)(d-1)-t)$. Thus

$$\begin{aligned}\operatorname{hf}_{S/M}(d+e_0) \;&=\; \operatorname{hf}_{S/M}(n(d-1)-1) \overset{\text{sym}}{=} \operatorname{hf}_{S/M}(d) \;=\; \operatorname{hf}_S(d)-(n+1) \\ \operatorname{hf}_{S/M}(d+e_0-1) &\overset{\text{sym}}{=} \operatorname{hf}_{S/M}(d+1) \;=\; \operatorname{hf}_S(d+1)-(n+1)n\end{aligned}$$

using $d \ge 2$ in the last equality. To see that the last term is strictly larger than $\operatorname{hf}_S(d)-(n+1)$, we calculate

$$\begin{aligned}\big(\operatorname{hf}_S(d+1)-(n+1)n\big) - \big(\operatorname{hf}_S(d)-(n+1)\big) \;&=\; \binom{n+d}{n-1} - (n+1)(n-1) \\ \overset{d\ge 2}{\ge}\; \frac{(n+2)(n+1)n}{6} - (n+1)(n-1) \;&=\; \frac{1}{6}(n+1)(n^2-4n+6) \;>\; 0. \qquad \square\end{aligned}$$

This leads to the following bound on the saturation gap.

Corollary 1.6.5. *For $r \le \operatorname{hf}_S(d)-(n+1)$ general points in the plane, the saturation gap $\gamma_n(d,r)$ is at most $(n-1)d-(n+1)$. The sum in Conjecture 1.3.10 reduces to $n-1$ terms.*

Proof. The proof is by reverse induction on r. The base case is $r = \mathrm{hf}_S(d) - (n+1)$, for which the statement follows from Theorem 1.6.1. We are going to show that $\gamma_n(d, r-1) \leq \gamma_n(d,r)$. Let $Z_{r-1} = (z_1, \ldots, z_{r-1}) \in (\mathbb{P}^n)^{r-1}$ be general, let z_r be one additional general point and set $Z_r = (z_1, \ldots, z_r)$. Let $e_0 \coloneqq \gamma_n(d,r)$, we have $\mathrm{hf}_{S/\mathrm{I}(Z_r)_{\langle d \rangle}}(d+e_0) = r$ and need to show $\mathrm{hf}_{S/\mathrm{I}(Z_{r-1})_{\langle d \rangle}}(d+e_0) = r-1$. For this is suffices to show

$$(\mathrm{I}(Z_{r-1})_{\langle d \rangle})_{d+e_0} \;=\; S_{e_0} \cdot \mathrm{I}(Z_{r-1})_d \;\supsetneq\; S_{e_0} \cdot \mathrm{I}(Z_r)_d \;=\; (\mathrm{I}(Z_r)_{\langle d \rangle})_{d+e_0}.$$

By genericity of Z_r, we can pick $f \in \mathrm{I}(Z_{r-1})_d \setminus \mathrm{I}(Z_r)_d$ and $h \in S_{e_0}$ not vanishing on z_r, then $fh \in S_{e_0} \cdot \mathrm{I}(Z_{r-1})_d$, but $fh \notin \mathrm{I}(Z_r) \supseteq \mathrm{I}(Z_r)_{\langle d \rangle}$. □

Chapter 2

Mukai lifting of self-dual points in $\mathbb{P}^6$

Contents

> What we face may look insurmountable. But I learned something from all those years of training and competing. I learned something from all those sets and reps when I didn't think I could lift another ounce of weight. What I learned is that we are always stronger than we know.
>
> ARNOLD SCHWARZENEGGER [Sch14]

A set of $2n$ points in $\mathbb{P}^{n-1}$ is self-dual if it is invariant under the Gale transform. Motivated by Mukai's work on canonical curves [Muk88; Muk92], Petrakiev showed that a general self-dual set of 14 points in $\mathbb{P}^6$ arises as the intersection of the Grassmannian in its Plücker embedding $\mathrm{Gr}(2,6) \subseteq \mathbb{P}^{14}$ with a complementary linear space [Pet09]. The *Mukai lifting Problem 2.2.14*, introduced in [Gei+23, Section 4] as a "formidable challenge", concerns the inverse task:

Main Results 2 (2.2.11, 2.3.2, 2.3.3, 2.4.1)**.** *We use numerical parameter continuation to tackle the Mukai lifting problem and implement an algorithm in* `Julia` *to solve it. We use a novel parametrization of the moduli space of self-dual point configurations using skew-symmetric matrices. Along the way, we also implement the forward problem of slicing Grassmannians and use it to experimentally study the real solutions to this problem.*

This chapter is based on the paper *Mukai Lifting of Self-Dual Points in* $\mathbb{P}^6$ [BK25] coauthored with Barbara Betti. The full algorithm is implemented in `Julia` [Bez+12] and is available at

`https://mathrepo.mis.mpg.de/MukaiLiftP6`

This chapter is structured as follows: In Section 2.1 we introduce the parameter continuation method for polynomial system solving, and highlight some useful start systems. In Section 2.2 we give a primer on self-dual point configurations in projective space, present characterizations and develop the skew normal form. We look at the the Mukai construction on canonical curves and its relation to moduli of self-dual points and we formulate the Mukai lifting and slicing problems. In Section 2.3 we solve the slicing problem using a toric degeneration of the Grassmannian and we study the distribution of real solutions. In Section 2.4 we implement the lifting problem using the skew-normal form, and we indicate how our machinery can be used to test Petrakiev's conjecture. Lastly, Section 2.5 discusses an extended example showing how to use the implementation.

Throughout the chapter $\mathbb{K}$ is an algebraically closed field, later specialized $\mathbb{K} = \mathbb{C}$, and $S = \mathbb{K}[x_1, \ldots, x_n]$.

2.1 Background: Solving polynomial systems using parameter continuation

In this section we discuss the fundamentals of using path tracking algorithms to solve polynomial systems F. The main idea is to realize the system as a member of a parametric family $H(x;u)$ ($H(x;0) = F$) and track the (hopefully easier to obtain) solutions of the start system $H(x;1)$ along $u \to 0$. We also look at several start systems used in practice. A standard reference for homotopy-based polynomial system solving is the book by Sommese & Wampler [SW05], see also Sascha Timme's Ph.D. thesis [Tim21] for details specifically about the software `HomotopyContinuation.jl` that we use.

2.1.1 Path tracking

Homotopy continuation is a method used to numerically solve a multivariate polynomial system $F(x) = 0$ with a zero-dimensional set of solutions over the complex numbers. It relies on embedding $F(x)$ into a family of polynomial systems that is easier to solve.

Example 2.1.1. Consider a sufficiently general inhomogeneous system of n polynomial equations $f_1, \dots, f_n \in \mathbb{C}[x_1, \dots, x_n]$ of degrees $\deg f_i = d_i$. By Bézout's theorem, the system $F = (f_1, \dots, f_n) = 0$ has $D \coloneqq d_1 \cdots d_n$ solutions in $\mathbb{C}^n$. To make our life easier, we first consider a very special polynomial system whose solutions are easy to find:

$$G = (g_1, \dots, g_n), \qquad g_i = x_i^{d_i} - 1.$$

It is immediate to describe all D distinct solutions of $\{G = 0\}$ as tuples of roots of unity. The trick is now to interpolate between these two systems by defining

$$H(x;t) = tG(x) + (1-t)F(x) \in \mathbb{C}[x_1, \dots, x_n; t], \qquad t \in [0,1].$$

For $t = 1$ we know all solutions x_{start} to the system $H(x;1) = 0$; moreover $t = 1$ is a regular value of the incidence

$$\{\, (x;u) \mid H(x;u) = 0 \,\} \longrightarrow \mathbb{A}^1_u.$$

For all but finitely many points in $\mathbb{A}^1_t$, the fiber in $\mathbb{C}^n$ consists of a set of D reduced points. By traveling in the parameter space from $1 \rightsquigarrow 0$ avoiding the discriminant locus (Section 4.1.1), we can track the start solutions x_{start} from the *start system* G to (hopefully) obtain the solutions for the *target system* F. ◁

The key theoretical insight is the Parameter Continuation theorem [SW05, Theorem 7.1.4] which guarantees the existence of continuous non-singular solution paths on a Zariski-open subset of the parameter space. The paths $\gamma(t)$ are locally defined by the implicit function theorem or, in other words, by the differential equation

$$\gamma(t)\colon [0,1] \to \mathbb{C}^n, \qquad H_x(\gamma(t);t) \cdot \gamma'(t) + H_t(\gamma(t),t) = 0.$$

The path tracking can be performed using standard numerical continuation methods such as Newton's method [Tim21, Chapter 4]. If the parametrized polynomial system

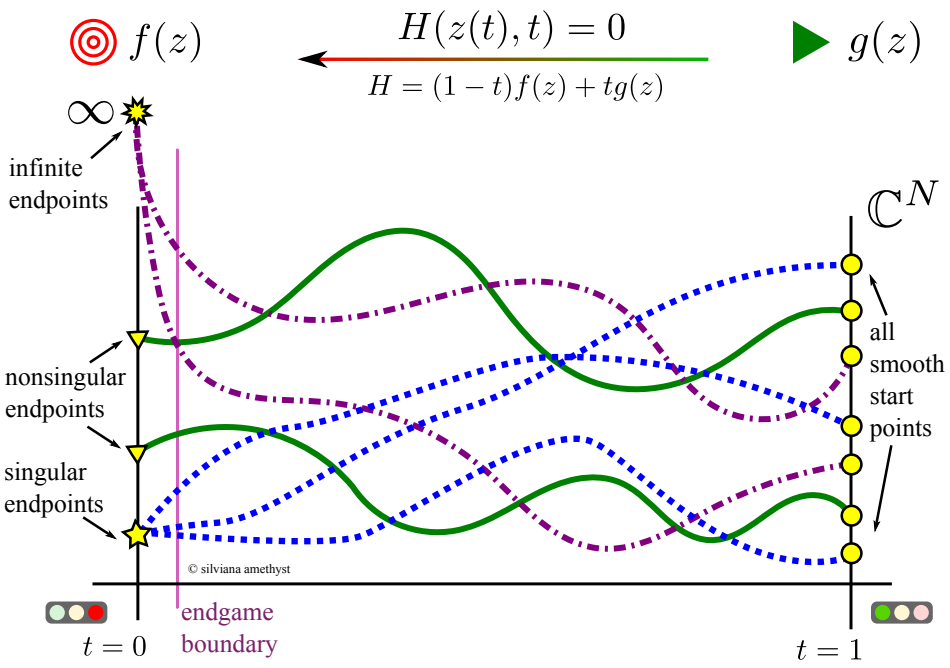


Figure 2.1: Path tracking in homotopy continuation. Used with permission from silviana amethyst [ame25].

is singular at the target parameter $t = 0$ (that is, if 0 is a branch point), then more refined *endgame strategies* are reqiured [SW05, Chapter 10]. The general situation is illustrated in Figure 2.1: Some paths may converge to non-singular endpoints, some paths come together at a singular endpoint, and some paths may diverge (in the given affine coordinate system).

We will use the implementation by Breiding and Timme [BT18] in the `Julia` package `HomotopyContinuation.jl`. The following snippet shows how to compute the target solution at the parameter `t_target` given start solution(s) `start_sol` at the paramter `t_start`.

```
using HomotopyContinuation as HC
@var x[1:n] t[1:m]
f = [f_1,...,f_l] # Define equations f_i(x,t) appropriately
F = System(f, variables = x[1:n], parameters = t[1:m])
res = HC.solve(F, start_sol; start_parameters = t_start, target_parameters = t_target)
```

For a particular polynomial system there are many possible start systems and homotopies. Different choices may have a different impact on the computational time of the `solve()` function.

Definition 2.1.2. If the start system has the same number of solutions as $F(x)$ and all tracked paths arrive at distinct endpoints, we say that we have found an *optimal parameter homotopy*.

Optimal start systems are preferred because they require to track the *minimal* number of paths necessary to still get all solutions of $F(x) = 0$.

2.1.2 Start systems

In this section we describe different parameter homotopies useful in practice. By the discussion of the previous section, the parametrized system $H(x;t)$ should have a constant number of complex solution for $t \in (0,1]$ that are regular. In other words, that $t \mapsto \mathbb{V}(H(x;t))$ defines a flat, or rather, smooth family outside $t = 0$.

One source of flat families comes from *initial degenerations*: By the discussion in Section 0.2.2, a variety $X \subseteq \mathbb{C}^N$ whose coordinate ring $R = \mathcal{O}(X)$ admits a finite SAGBI basis (with respect to a monomial order) admits a flat degeneration to the toric variety $X_0 = \operatorname{Spec}(\operatorname{in}_\prec R)$ (Theorem 0.2.10 and the discussion afterwards). This approach has been used in [BSW23] to construct *Khovanskii homotopies*; we will see this homotopy in action in the later sections of this chapter.

Another commonly used family of polynomials takes advantage of sparsity of the target system. Let $A \subseteq \mathbb{Z}^n$ be a finite set of exponent vectors and denote the space of Laurent polynommials with support in A as

$$\mathbb{K}^A = \left\{ f = \sum_{\alpha \in A} f_\alpha x^\alpha \;\middle|\; f_\alpha \in \mathbb{K} \right\} \subseteq \mathbb{K}[x_1^{\pm 1}, \dots, x_n^{\pm 1}].$$

The Bernstein–Khovanskii–Kushnirenko Theorem 2.1.5 [Ber79] gives a precise bound on the number of zeros $z \in (\mathbb{K}^\times)^n$ of a square polynomial system for general $(f_1, \dots, f_n) \in \mathbb{K}^{A_1} \times \cdots \times \mathbb{K}^{A_n}$. For the precise statement, we need the notion of *mixed volume*. Let $P_1, \dots, P_r \subseteq \mathbb{R}^n$ be convex polytopes and consider the function

$$f \colon \mathbb{R}^r_{\geq 0} \to \mathbb{R}, \qquad f(t_1, \dots, t_r) = \operatorname{vol}(t_1 \cdot P_1 + \cdots + t_r \cdot P_r)$$

where $P + P' = \{ v + v' \mid v \in P, v' \in P' \}$ denotes the Minkowski sum. The function f turns out to be a homogeneous polynomial of degree n [Mon21, Theorem V.39].

Definition 2.1.3 (Mixed volume)**.** The *mixed volume* $\operatorname{mvol}(K_1, \dots, K_n)$ of $r = n$ convex polytopes in $\mathbb{R}^n$ is the coefficient of $t_1 \cdots t_n$ in f.

Example 2.1.4. If $P_1 = \cdots = P_n = P$, then $t_1 P + \cdots + t_n K = (t_1 + \cdots + t_n)P$ has (ordinary) volume $(t_1 + \cdots + t_n)^n \operatorname{vol}(P)$. Consequently, the coefficient of $t_1 \cdots t_n$ is $\operatorname{mvol}(P, \dots, P) = n! \operatorname{vol}(P)$. In this sense, mixed volume generalizes ordinary volume. It is in fact the unique such function that is symmetric and multilinear in its arguments. ◁

Theorem 2.1.5 (Bernstein–Khovanskii–Kushnirenko (BKK))**.** *Let $A_1, \dots, A_n \subseteq \mathbb{Z}^n$ be finite sets and $P_i := \operatorname{Conv}(A_i) \subseteq \mathbb{R}^n$ their Newton polytopes. If $f_i \in \mathbb{K}^{A_i}$ are such that $\mathbb{V}_{(\mathbb{K}^\times)^n}(f_1, \dots, f_n)$ is finite, then*

$$\operatorname{length} \mathbb{V}_{(\mathbb{K}^\times)^n}(f_1, \dots, f_n) \leq n! \operatorname{mvol}(P_1, \dots, P_n).$$

For general such $f_1, \dots, f_n$, the intersection is reduced and equality holds.

The condition that $f_1, \dots, f_n$ are *BKK-general* in the sense of the theorem is an open condition and can be expressed in the combinatorics of the coefficients [Mon21, Section VII.3.2]. For an introduction to the BKK theorem, extensions and the relation to toric and convex geometry, consider the book [Mon21].

To solve a sparse polynomial systems $(f_1, \dots, f_n) \in \prod_{i=1}^n \mathbb{C}^{A_i}$, one can therefore use a straight-line Homotopy $H(x,u) = tg(x) + (1-t)f(x)$, where $(g_1, \dots, g_n)$ are BKK-general. The BKK theorem then guarantees that for almost all $u \in \mathbb{C}$, the system $H(x,u)$ has the same number of non-degenerate solutions and that they vary smoothly.

Sturmfels and Huber [HS95] constructed a BKK-general polyhedral start system using mixed subdivisions of the Newton polytopes. This start system implemented in `HomotopyContinuation.jl` as `polyhedral` and is the default start system used when using the `solve` command.

A third useful source of start systems is to take advantage of an incidence structure: Informally, consider an incidence correspondence of the form

$$\begin{array}{ccccc} & & \mathcal{I} = \{\, f(x,y) = 0 \,\} & & \\ & \overset{\mathrm{pr}_X}{\swarrow} & & \overset{\mathrm{pr}_Y}{\searrow} & \\ X & & & & Y \end{array}$$

Usually, one is interested in the fibers of one side, say $\mathrm{pr}_X^{-1}(x_0)$, but this may be a priori difficult problem. On the other hand, it might be much easier to find for given $y \in Y$ a point x_1 in the fiber $\mathrm{pr}_Y^{-1}(y)$. If this is the case, then one can first compute such a pair $(x_1, y) \in \mathcal{I}$ and then use this as a start system to move $x_1 \rightsquigarrow x_0$.

We will utilize this strategy in this chapter to solve the Mukai lifting Problem 2.2.14 below; in fact all three methods mentioned in this section are used: To solve the "forward problem" (the slicing problem) we will perform an initial degeneration of the Grassmannian $\mathrm{Gr}(2,6)$ to a toric variety. A general system of linear equations can easily be solved on the toric variety using the polyhedral start system. Once the slicing problem is solved, we use it as a start system for the lifting problem.

2.2 Self-dual point configurations

In this section we give an introduction to the theory of self-dual points. A set of $2n$ points in $\mathbb{P}^{n-1}$ is self-dual if it is invariant under Gale duality, we present a series of equivalent characterizations. We define the moduli space $\mathcal{A}_n$ of self-dual points in $\mathbb{P}^n$ and consider rational parametrizations of $\mathcal{A}_n$. We also discuss the connection of self-dual points with Mukai's construction of canonical curves of low genus [Muk88] and state the *Mukai Lifting Problem 2.2.14*.

2.2.1 Characterizations of self-dual points

Let $S = \mathbb{K}[X_1, \dots, X_n]$ be the homogeneous coordinate ring of $\mathbb{P}^{n-1} = \mathbb{P}(\mathbb{K}^n)$ and let $\mathrm{I}(X)$ be the homogeneous ideal of a given set $X \subseteq \mathbb{P}^{n-1}$. Recall that a set of points

$\Gamma \subseteq \mathbb{P}^{n-1}$ is non-degenerate if they are not contained in a hyperplane, namely, $\mathrm{I}(\Gamma)_1 = 0$. A set of points is *linearly general* if any subset of at most n points is linearly independent. Let $\mathrm{Diag}(n) \subseteq \mathrm{GL}(n)$ be the set of invertible diagonal matrices and let $\Gamma \subseteq \mathbb{P}(\mathbb{K}^n)$ be a non-degenerate ordered set of $2n$ distinct points, represented by the columns of a $n \times 2n$ matrix.

Lemma 2.2.1. *The following are equivalent for* Γ*:*

(i) *All subsets of* $2n-1$ *points impose the same number of conditions on quadrics as* Γ*, in symbols* $I(\Gamma)_2 = I(\Gamma \setminus \{\gamma\})_2$ *for all* $\gamma \in \Gamma$*;*

(ii) *There exists an invertible diagonal matrix* $\Lambda \in \mathrm{Diag}(2n)$ *such that* $\Gamma \cdot \Lambda \cdot \Gamma^\mathsf{T} = \mathbf{0}$*;*

(iii) *There exists a partition* $\Gamma = \Gamma_1 \dot\cup \Gamma_2$ *and a non-degenerate symmetric bilinear form* $Q \in \mathrm{Sym}(n)$ *such that both* Γ_i *are orthogonal bases with respect to* Q*, i. e.* $\Gamma_i^\mathsf{T} \cdot Q \cdot \Gamma_i \in \mathrm{Diag}(n)$*.*

If Γ *is linearly general, these are equivalent to*

(iii') *For any partition* $\Gamma = \Gamma_1 \dot\cup \Gamma_2$ *into* $n+n$ *points there is a non-degenerate bilinear form* $Q \in \mathrm{Sym}(n)$ *such that both* Γ_i *are orthogonal bases with respect to* Q*.*

Proof. (ii)⇒(i) Let $A = \{\, \alpha \in \mathbb{N}^n \mid |\alpha| = 2 \,\}$, for a point $x \in \mathbb{P}(\mathbb{K}^n)$ define $x^A \coloneqq (x^\alpha)_{\alpha \in A}$ as the row vector of monomials X^α evaluated in z. Then the linear map $E\colon \mathbb{K}[X_1, \dots, X_n]_2 \to \mathbb{K}^{2n}$ given by evaluation of quadrics in Γ is represented by

$$E = \left[(\gamma_i)^\alpha\right]_{1 \le i \le 2n, |\alpha|=2} = \begin{bmatrix} -(\gamma_1)^A- \\ \vdots \\ -(\gamma_{2n})^A- \end{bmatrix}, \qquad \Gamma = [\gamma_1 | \dots | \gamma_{2n}].$$

Its kernel $\operatorname{Ker} E \subseteq \mathbb{K}[X_1, \dots, X_n]_2$ consists of quadrics passing through Γ. We need to show that the rank of E does not change when deleting any of the $2n$ rows from E. For a point γ, the vector γ^A consists of the upper half of the matrix $\gamma \cdot \gamma^\mathsf{T}$, and by (ii) we have $\sum_i \Lambda_{ii} \gamma_i \gamma_i^\mathsf{T} = \mathbf{0}$. By assumption all Λ_{ii} are nonzero, so any row of E is a linear combination of the other rows, proving the rank condition.

(i)⇒(ii) Reversing the argument, by (i) every $\gamma_j \gamma_j^\mathsf{T}$ is a linear combination

$$\gamma_j \gamma_j^\mathsf{T} = \sum_{i \neq j} \beta_i^{(j)} \gamma_i \gamma_i^\mathsf{T}.$$

Thus the system of equations $\Gamma \Lambda \Gamma^\mathsf{T} = \mathbf{0}$, linear in diagonal matrices Λ, has solutions $\mathrm{diag}(\beta^{(j)})$ with $\beta_j^{(j)} = -1 \neq 0$. Since $\mathbb{K}$ is infinite, there exists a solution with all entries nonzero.

(ii)⇒(iii): After rescaling the columns by $\gamma_i \mapsto 1/\sqrt{\Lambda_{ii}}\gamma_i$ we may assume that $\Lambda = \mathrm{Id}_{2n}$. After reordering the columns we may assume that $\Gamma = [\Gamma_1|\Gamma_2]$ where Γ_1 is invertible. Consider the non-degenerate symmetric matrix $Q \coloneqq \Gamma_1^{-\mathsf{T}}\Gamma_1^{-1}$, then using

$$\mathbf{0} = \Gamma\Gamma^{\mathsf{T}} = \Gamma_1\Gamma_1^{\mathsf{T}} + \Gamma_2\Gamma_2^{\mathsf{T}}$$

we see that Γ_2 is also invertible and $Q = -\Gamma_2^{-\mathsf{T}}\Gamma_2^{-1}$, and hence $\Gamma_i^{\mathsf{T}}Q\Gamma_i = \pm\,\mathrm{Id}_n$.

(iii)⇒(ii): This argument is reversible after rescaling Γ to match $\Gamma_i^{\mathsf{T}}Q\Gamma_i = \pm\,\mathrm{Id}_n$.

(ii)⇔(iii'): The same argument clearly applies to any partition if Γ is linearly general. □

Definition 2.2.2. A set of $2n$ points in $\mathbb{P}^{n-1}$ is *self-dual* or self-associated if it satisfies any of the equivalent conditions (i),(ii),(iii) from the previous lemma.

The name "self-dual" is related to the *Gale transform* on the space of point-configurations, which we will introduce in Section 2.2.3. The fixed points of the Gale transform are exactly self-dual point configurations. A deeper characterization of self-dual points related to the Gorenstein property of $S/I(\Gamma)$ is given by Eisenbud and Popescu [EP00]:

A finite-dimensional $\mathbb{K}$-algebra A is Gorenstein if it is injective as a module over itself. An equivalent characterization is the existence of a $\mathbb{K}$-linear map $e\colon A \to \mathbb{K}$ such that $(f,g) \mapsto e(fg)$ is a perfect pairing. A standard graded $\mathbb{K}$-algebra S_X of Krull dimension n is Gorenstein if there is a regular sequence $\ell_1,\dots,\ell_n \in (S_X)_1$ such that $S_X/\langle \ell_1,\dots,\ell_n\rangle$ is Gorenstein.

Example 2.2.3 (Apolar algebras)**.** If $F \in \mathrm{Sym}^d \mathbb{K}^n$ is a symmetric tensor, then the *apolar algebra* $A_F = S/\operatorname{Ann}_S(F)$ is graded Gorenstein (notation as in Section 1.2). The perfect pairing is given by multiplication of forms in degree j with those of degree $d-j$, which land in $(A_F)_d \cong \mathbb{K}F \cong \mathbb{K}$. ◁

Theorem 2.2.4 ([EP00, Theorem 7.1 & 7.3])**.** *For a non-degenerate set of* $2n$ *points* $\Gamma \subseteq \mathbb{P}^{n-1}$ *the following are equivalent:*

(i) Γ *is self-dual and fails by* 1 *to impose independent conditions on quadrics;*

(ii) *every subset of* $2n-1$ *points imposes independent conditions on quadrics, but* Γ *does not impose independent conditions;*

(iii) Γ *is arithmetically Gorenstein, i. e. the graded ring* $S_\Gamma = S/I(\Gamma)$ *is Gorenstein.*

2.2.2 Normal forms for self-dual points

We consider the problem of finding normal forms for general self-dual point configurations modulo the action of $\mathrm{GL}(n)$. Lemma 2.2.1(iii) suggests a normal form using orthogonal matrices over $\mathbb{K}$

$$\mathrm{SO}(n) \coloneqq \{\, P \in \mathrm{SL}(n) \mid P \cdot P^{\mathsf{T}} = \mathrm{Id}_n \,\}.$$

Theorem 2.2.5 (Orthogonal normal form). *Let $\Gamma = [\gamma_1, \dots, \gamma_{2n}] \subseteq \mathbb{P}(\mathbb{K}^n)$ be an ordered set of $2n$ points with $\gamma_1, \dots, \gamma_n$ linearly independent. Then the following hold:*

(i) *Γ is self-dual if and only if there exists a matrix $A \in \mathrm{PGL}(n)$ such that, after rescaling the columns of Γ, we have $A \cdot \Gamma = [\mathrm{Id}_n \,|P]$ for some $P \in \mathrm{SO}(n)$.*

(ii) *Moreover, the orthogonal matrix P is unique up to multiplying the columns with $\varepsilon \in \{\pm 1\}^n$ such that $\prod_i \varepsilon_i = 1$.*

Proof. (i) If $A \cdot \Gamma = [\mathrm{Id}_n \,|P]$, then Γ is self-dual by Lemma 2.2.1(iii), for the converse we provide a construction of A and P. By Lemma 2.2.1(ii) there exists an invertible diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_{2n})$ such that $\Gamma \cdot \Lambda \cdot \Gamma^\mathsf{T} = 0$. Rescaling the columns, i. e. multiplying Γ on the right with $\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}, \sqrt{-\lambda_{n+1}}, \cdots, \sqrt{-\lambda_{2n}})$, we can assume that $\Lambda = \mathrm{Id}_n \oplus -\mathrm{Id}_n$. Since the first n points of $\Gamma = [\Gamma_1 | \Gamma_2]$ are linearly independent, the matrix $A := \Gamma_1$ is invertible and $A^{-1} \cdot \Gamma = [\mathrm{Id}_n \,|P]$ is self-dual with respect the same diagonal matrix as Γ. Then

$$\begin{bmatrix} \mathrm{Id}_n & P \end{bmatrix} \cdot \begin{bmatrix} \mathrm{Id}_n & \\ & -\mathrm{Id}_n \end{bmatrix} \cdot \begin{bmatrix} \mathrm{Id}_n \\ P^\mathsf{T} \end{bmatrix} = \mathrm{Id}_n - P \cdot P^\mathsf{T} = 0,$$

which shows that P is an orthogonal matrix. If P has determinant -1, we can rescale the last column of Γ and P to obtain a special orthogonal matrix.

(ii) The orthogonal matrix P only depends on the choice of the square roots in the rescaling with $\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{-\lambda_{2n}})$, and hence is unique up to signs. □

From now on assume that $\mathrm{char}(\mathbb{K}) \neq 2$, there is no harm in restricting to $\mathbb{K} = \mathbb{C}$. A normal form more suitable for computation relies on the parametrization of the orthogonal group via the Cayley transform [Cay46]. Let $U \subseteq \mathbb{K}^{n \times n}$ be the set of matrices A with $I_n + A$ invertible, then the *Cayley transform* is the self-inverse isomorphism of varieties

$$\mathcal{C} \colon U \to U, \qquad \mathcal{C}(A) = (I_n - A)(I_n + A)^{-1}.$$

Noting that the product $(I_n - A)(I_n + A)^{-1}$ commutes, a straightforward computation shows that $\mathcal{C}$ induces a birational map between skew-symmetric and special orthogonal matrices:

$$\mathcal{C} \colon \mathrm{Skew}(n) \cap U \leftrightarrow \mathrm{SO}(n) \cap U.$$

Hence the orthogonal normal form $[I_n \mid P]$ of a general set of points (in particular, for general P) can be rewritten as $[I_n \mid \mathcal{C}(S)]$ for $S \in \mathrm{Skew}(n)$.

Remark 2.2.6. Over any field, an orthogonal matrix of odd size *always* has an eigenvalue equal to 1 or -1 (if its determinant is, respectively, 1 or -1). If the matrix is generic in $\mathrm{SO}(n)$, then we have that -1 is *not* an eigenvalue, thus the Cayley transform is well-defined for generic $P \in \mathrm{SO}(n)$. ◁

Since we consider point configurations up to projective linear transformations, we can apply the transformation $I_n + S$ to obtain

$$[I_n \mid \mathcal{C}(S)] = [I_n \mid (I_n - S)(I_n + S)^{-1}] \quad \rightsquigarrow \quad [I_n + S \mid I_n - S]. \tag{2.1}$$

This step was suggested to us by Simon Telen and we refer to the representation in (2.1) as a *skew normal form* (SNF). Informally, skew normal forms provide an efficient way to represent "general" self-dual point configurations. In the next section we introduce the moduli space $\mathcal{A}_{n-1}$ of self-dual point configurations, allowing us to make this a precise statement.

2.2.3 The space of point configurations

Following [MFK94, Chapter 3] or [DO88, Chapter II], the moduli space $\mathcal{P}_n^m$ of ordered sets of m points in $\mathbb{P}^n$ can be constructed as a GIT quotient as follows:

- The product $(\mathbb{P}^n)^m$ admits group actions of $G = \mathrm{SL}(n+1)$ and $\mathfrak{S}_m$ via

  $$g \cdot (x_1, \ldots, x_m) = (gx_1, \ldots, gx_m), \qquad \sigma \cdot (x_1, \ldots, x_m) = (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(m)}).$$

- Consider the very ample line bundle on $(\mathbb{P}^n)^m$

  $$\mathcal{L} \coloneqq \mathcal{O}_{(\mathbb{P}^n)^m}(1, \ldots, 1) = \bigotimes_{i=1}^{m} \pi_i^* \mathcal{O}_{\mathbb{P}^n}(1).$$

  This enables us to embed $(\mathbb{P}^n)^m$ as a projective variety in $\mathbb{P}^{(n+1)^m - 1}$ with homogeneous coordinate ring $\bigoplus_{k \geq 0} H^0((\mathbb{P}^n)^m, \mathcal{L}^{\otimes k})$. The bundle $\mathcal{L}$ linearizes with respect to the action of $\mathrm{SL}(n+1)$ in the sense of [MFK94].

- With respect to this linearized action, a tuple $Z \in (\mathbb{P}^n)^m$ is semi-stable if and only if

  $$\dim \mathrm{Span}_{\mathbb{P}^n}(Y) + 1 \geq \frac{(n+1) \cdot \#Y}{m} \qquad \text{for all } Y \subseteq Z.$$

  The tuple Z is stable if and only if this inequality is always strict [DO88, Chapter II Thm. 1].

- Consider the graded ring of invariants $R_n^m \coloneqq \bigoplus_{k \geq 0} H^0((\mathbb{P}^n)^m, \mathcal{L}^{\otimes k})^G$, then

  $$\mathcal{P}_n^m \coloneqq \mathrm{Proj}\, R_n^m = ((\mathbb{P}^n)^m)^{\mathrm{ss}} /\!/_{\mathcal{L}} G.$$

The moduli space of ordered point configurations $\mathcal{P}_n^m$ is a natural domain for the Gale transformation, which we now introduce, a reference is [DO88, Chapter III].

Definition 2.2.7. The *Gale transform* of $\Gamma \in \mathcal{P}_n^m$ is given by any ordered set $G(\Gamma) \in \mathcal{P}_{m-n-2}^m$ of m points in $\mathbb{P}^{m-n-2}$, represented by a matrix of size $(m-n-1) \times m$ such that there exists an invertible diagonal matrix $\Lambda \in \mathrm{Diag}(m)$ verifying

$$\Gamma \cdot \Lambda \cdot G(\Gamma)^{\mathsf{T}} = 0. \tag{2.2}$$

In other words, the *Gale transform* is the map between moduli spaces of ordered set of points

$$G\colon \mathcal{P}_n^m \longrightarrow \mathcal{P}_{m-n-2}^m, \qquad \Gamma \longmapsto G(\Gamma).$$

The Gale transform of Γ is given by the transpose of a kernel matrix of Γ. Defining the moduli space $\mathcal{P}_n^m$ as above makes this construction well-defined for any semi-stable configuration. Assuming that $\Gamma = [\mathrm{Id}_{n+1} | A]$ and taking the diagonal matrix Λ in the definition as $\mathrm{Id}_m \oplus -\mathrm{Id}_m$, it follows that the Gale transform of Γ is given by the configuration $G(\Gamma) \in \mathcal{P}_{m-n-2}^m$ represented by the columns of the matrix $[A^\mathsf{T} | \mathrm{Id}_{m-(n+1)}]$.

Proposition 2.2.8. *$\Gamma \in \mathcal{P}_{n-1}^{2n}$ is self-dual if and only if it is fixed under the Gale transform.*

Proof. Inspecting Equation (2.2), we see that $\Gamma = G(\Gamma)$ in $\mathcal{P}_{n-1}^{2n}$ is a reformulation to condition (ii) of Lemma 2.2.1. □

The characterization using the Gale transform finally leads to the definition of the moduli space of self-dual points.

Definition 2.2.9. The *moduli space of un-ordered self-dual point configurations* in $\mathbb{P}^n$ is the fixed locus of the Gale transform G on $\mathcal{P}_n^{2n+2}/\mathfrak{S}_{2n+2}$. It will be denoted by $\mathcal{A}_n$.

The variety $\mathcal{A}_{n-1}$ is an irreducible rational variety of dimension $\binom{n}{2}$ [DO88, Theorem III.4]. With this definition, we can formally state that the orthogonal normal form and the skew normal form yield parametrizations of the general self-dual points.

Proposition 2.2.10. *Assume that* $\mathrm{char}(\mathbb{K}) \neq 2$. *The orthogonal normal form yields a surjective morphism* $\mathrm{SO}(n) \twoheadrightarrow \mathcal{A}_{n-1}$ *generically finite of degree* $2^{2n-2}(2n)!$.

A previously published version of this proposition contains a mistake in the degree computation, we thank Alessio Caminata for pointing this out to us.

Proof. The orthogonal normal form $P \mapsto \Gamma = [I_n | P]$ is surjective by the previous theorem and the characterization of semi-stable points. This normal form of an ordered set of self-dual points is unique up to the action of 2^{n-1} ways of multiplying columns of $P \in \mathrm{SO}(n)$ by ± 1 (action of $(\mathbb{K}^\times)^{2n}$ on columns of Γ) and the 2^{n-1} ways of multiplying the rows of P by ± 1 (action of $\mathrm{SL}(n)$ on rows); here we use $-1 \neq 1$. Moreover, for a general Γ, all its permutations by $\sigma \in \mathfrak{S}_{2n}$ have distinct normal forms, hence the rational map $\mathrm{SO}(n) \to \mathcal{A}_{n-1}$ is generically finite of degree $2^{2n-2} \cdot (2n)!$. □

As noticed in Equation (2.1), we can also use skew-symmetric matrices in the parametrization. This has the advantage that it avoids the delicate step of inverting a matrix in an implementation of this parametrization, reducing computation time and numerical error.

Theorem 2.2.11 (Skew normal form)**.** *A general configuration of self-dual points $\Gamma \in \mathcal{A}_{n-1}$ admits a skew normal form. Moreover, the dominant rational map*

$$\mathrm{Skew}(n) \dashrightarrow \mathcal{A}_{n-1}, \qquad S \mapsto [I_n + S \mid I_n - S]$$

is generically finite of degree $(2n)!2^{2n-2}$.

Since a skew-symmetric matrix $S \in \mathrm{Skew}(n)$ is uniquely determined by the $\binom{n}{2}$ entries in the upper triangular part, a suitable parameter space for $\mathcal{A}_{n-1}$ is $\mathbb{C}^{\binom{n}{2}}$:

$$(s_1, \ldots, s_{\binom{n}{2}}) \mapsto \left[\begin{array}{cccc|cccc} 1 & s_1 & \cdots & s_{n-1} & 1 & -s_1 & \cdots & -s_{n-1} \\ -s_1 & 1 & \ddots & \vdots & s_1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & s_{\binom{n}{2}} & \vdots & \ddots & \ddots & -s_{\binom{n}{2}} \\ -s_{n-1} & \cdots & -s_{\binom{n}{2}} & 1 & s_{n-1} & \cdots & s_{\binom{n}{2}} & 1 \end{array}\right].$$

Example 2.2.12. In our situation of interest, 14 general self-dual points in $\mathbb{P}^6$ have exactly

$$2^{12} \cdot 14! = 357.082.280.755.200$$

distinct orthogonal or skew normal forms. ◁

2.2.4 Mukai-type results for $\mathcal{A}_n$

For small values of n, self-dual configurations in $\mathbb{P}^n$ are well-studied. A related landmark result is Mukai's description of polarized K3 surfaces and (canonical) curves of genus g as complete intersections on homogeneous varieties X_g [Muk88; Muk92]. He proved that a general polarized K3 surface $(S, \mathcal{L})$ of genus g, meaning $\mathcal{L}$ is ample and $\mathcal{L}^2 = 2g - 2$, arises as the intersection of homogeneous varieties $X_g \subseteq \mathbb{P}^N$ with general hypersurfaces (in most cases hyperplanes) if $g \leq 10$. Similarly, general canonical curves $C \subseteq \mathbb{P}^{g-1}$, embedded by their very ample canonical bundle ω_C, arise as sections of X_g if $g \leq 9$.

In dimension zero, general self-dual point configurations in $\mathbb{P}^{g-2}$ arise as sections of X_g if $g \leq 8$. This is classical for $g \leq 6$ (see [EP00, Part II] for a modern account) and was proven by Petrakiev for $g = 7, 8$ [Pet09]. In particular, all self-dual point configurations in $\mathbb{P}^n$ for $n + 2 \leq 8$ arise as linear sections of canonical curves by Mukai's result. For $n + 2 = g \geq 9$ this is no longer true by a simple moduli count. A summary of these results is given in the following theorem.

Theorem 2.2.13. (i) *All sets of four points in $\mathbb{P}^1$ are self-dual.*

(ii) *Six points in $\mathbb{P}^2$ are self-dual if and only if they are the intersection of a quadric and cubic.*

(iii) *A general set in $\mathcal{A}_3$ is a complete intersection of three quadric surfaces in $\mathbb{P}^3$.*

(iv) *A general set in $\mathcal{A}_4$ is a quadric section of an elliptic normal curve in $\mathbb{P}^5$; equivalently it is the intersection of $\mathrm{Gr}(2,5) \subseteq \mathbb{P}^9$ with a quadric and a linear space.*

(v) *A general set in $\mathcal{A}_5$ is a linear section of the Lagrangean Grassmannian $X_7 = \mathrm{LG}_+(5,10) \subseteq \mathbb{P}^{15}$.*

(vi) *A general set in $\mathcal{A}_6$ is a linear section of the Grassmannian $X_8 = \mathrm{Gr}(2,6) \subseteq \mathbb{P}^{14}$.*

In cases (ii)–(vi)*, an equivalent statement is that the set $\Gamma \subseteq \mathbb{P}^{n-1}$ is a hyperplane section of a canonical curve of genus $n+1$ in $\mathbb{P}^n$.*

This leads us to consider the *Mukai lifting problem* mentioned in the beginning of the chapter. For convenience, we only state in the case of $\mathbb{P}^6$.

Problem 2.2.14. (i) **Slicing:** Given a complementary linear subspace $\mathbb{P}^6 \cong \mathbb{L} \subseteq \mathbb{P}^{14}$, find the self-dual point configuration $\mathbb{L} \cap \mathrm{Gr}(2,6) \hookrightarrow \mathbb{P}^6$.

(ii) **Lifting:** Given a general self-dual set $\Gamma \subseteq \mathbb{P}^6$, find a linear embedding $L\colon \mathbb{P}^6 \xrightarrow{\sim} \mathbb{L} \subseteq \mathbb{P}^{14}$ such that

$$L(\Gamma) = \mathbb{L} \cap \mathrm{Gr}(2,6). \tag{2.3}$$

(?)

Part (iv)–(vi) in Theorem 2.2.13 all correspond to a non-trivial lifting problem. For reference, we record a precise statement of Petrakiev's result on $\mathcal{A}_6$. The group $\mathrm{SL}(6)$ acts on $\mathbb{C}^{15} = \bigwedge^2 \mathbb{C}^6$ and hence induces an action on $\mathrm{Gr}(7, \mathbb{C}^{15})$, leaving $X_8 \subseteq \mathbb{P}(\mathbb{C}^{15})$ invariant.

Theorem 2.2.15 ([Pet09, Thm. 3.3])**.** *The slicing map*

$$f\colon \mathrm{Gr}(7, \mathbb{C}^{15}) / \mathrm{SL}(6) \dashrightarrow \mathcal{A}_6, \qquad \mathbb{L} \mapsto \mathbb{L} \cap X_8$$

defines a generically finite dominant rational map to the space of self-dual points $\mathcal{A}_6$.

In particular, the slicing Problem 2.2.14 has a solution for general configurations.

2.3 Slicing using homotopy continuation

In this section we solve the Slicing Problem 2.2.14 using a toric degeneration of $\mathrm{Gr}(2,6)$. As an application, we study the distribution of real points among uniformly chosen real linear spaces from $\mathrm{Gr}(7, \mathbb{R}^{15})$.

Remark 2.3.1. In this section we compute a set of self-dual points Γ as a given 6-dimensional linear subspace $\mathbb{L}$ intersecting the Grassmannian $X_8 := \mathrm{Gr}(2,6) \subseteq \mathbb{P}^{14}$. By construction, the linear space $\mathbb{L} \subseteq \mathbb{P}^{14}$ is a *solution* to the Mukai lifting problem for Γ. Later in Section 2.4, we will set up a homotopy in which Γ serves to get the start parameter $\underline{a}_{\mathrm{start}}$ and $\mathbb{L}$ serves to get the start solution. ◁

2.3.1 Computing a linear section of the Grassmannian

By definition, the variety $X_8 = \mathrm{Gr}(2,6)$ parameterizes 2-dimensional subspaces in $\mathbb{C}^6$. We represent an element in an affine chart of X_8 as the row span of the 2×6 matrix:

$$H = \begin{bmatrix} 1 & 0 & t_1 & t_2 & t_3 & t_4 \\ 0 & 1 & t_5 & t_6 & t_7 & t_8 \end{bmatrix}. \tag{2.4}$$

The Plücker embedding $p\colon \mathrm{Gr}(2,6) \hookrightarrow \mathbb{P}^{14}$, $p(H) = (p_0 : \cdots : p_{14}) \in \mathbb{P}^{14}$ is given by taking the 2×2 minors of H, these are the 15 Plücker coordinates in $\mathbb{C}[t_1, \ldots, t_8]$. In our affine chart we have $p_0 = 1$.

Lemma 2.3.2. *Let $A \in \mathbb{C}^{8\times 15}$ be a general matrix and $\mathbb{L} = \mathbb{P}(\operatorname{Ker} A)$. The polynomial system in $\underline{t}$*

$$A \cdot \big(p_0(\underline{t}), \ldots, p_{14}(\underline{t})\big)^{\mathsf{T}} = 0 \tag{2.5}$$

has exactly 14 solutions $\underline{t}_1, \ldots, \underline{t}_{14}$ in $\mathbb{C}^8$. The set $Z \coloneqq \{\, p(\underline{t}_k) \mid k = 1, \ldots, 14 \,\}$ is the image of a configuration of self-dual points $\Gamma \subseteq \mathbb{P}^6$ under a linear map $L\colon \mathbb{P}^6 \xrightarrow{\sim} \mathbb{L} \subseteq \mathbb{P}^{14}$.

Equation (2.5) is a *linear* equation in the Plücker coordinates of $\underline{t}$.

Proof. The Grassmannian X_8 is a 8-dimensional variety of degree 14 [MS21, Thm 5.13], so its intersection with 8 linear hyperplanes in general position contains exactly 14 points. Let L be a kernel matrix for A so $\mathrm{Im}(L) = \mathbb{L}$. If $\underline{t}_k$ is a solution to (2.5), then $p(\underline{t}_k)$ is in the kernel of A and hence it lies in the image $\mathbb{L}$ of the linear map $L\colon \mathbb{P}^6 \to \mathbb{P}^{14}$. Since $X_8 \subseteq \mathbb{P}^{14}$ is arithmetically Gorenstein, so is $Z \subseteq \mathbb{L}$ and by Theorem 2.2.4(iii) it corresponds to a self dual set of points $\Gamma \subseteq \mathbb{P}^6$ such that $L(\Gamma) = P$. To compute such a configuration it is enough to compute a solution Γ to:

$$L \cdot \Gamma = \begin{bmatrix} p_0(\underline{t}_1) & & p_0(\underline{t}_{14}) \\ \vdots & \cdots & \vdots \\ p_{14}(\underline{t}_1) & & p_{14}(\underline{t}_{14}) \end{bmatrix}. \qquad \square$$

We point out that the previous lemma gives a polynomial system to construct the self-dual configuration Γ in a linear section $\mathbb{L}$ defined by 8 hyperplanes $\sum_{j=1}^{15} a_{ij} x_j = 0$, $i = 1, \ldots, 8$. After fixing a random matrix A, we use a toric degeneration of the Grassmannian to compute solutions $\underline{t}_k$ for (2.5). We dedicate the following section to describe the construction of the degeneration.

2.3.2 Toric degeneration of the Grassmannian

The homogeneous coordinate ring $R \coloneqq S_X \subseteq \mathbb{K}[x_1, \ldots, x_n]$ of a projective variety $X \subseteq \mathbb{P}^{n-1}$ can be studied by looking at the *initial algebra* of R with respect to a term order $\prec$, that is, the algebra generated by the initial terms with respect to $\prec$ of every element in R

$$\mathrm{in}_\prec(R) \coloneqq \mathbb{K}[\mathrm{in}_\prec(f) \mid f \in R].$$

Although the coordinate ring of a projective variety is a finitely generated algebra, the corresponding initial algebra may not be finitely generated for a given term order. The initial algebra is finitely generated if R admit a finite SAGBI basis for that term order, that is, a finite set $\mathcal{F} \subseteq R$ whose initial terms generate the initial algebra. We refer to Section 0.2 for more background on initial terms, ideals, and algebras.

Remark 2.3.3. SAGBI bases have been introduced by Robbiano and Sweedler under this name in [RS06]. They are called *Canonical Subalgebra bases* by Sturmfels in [Stu96]. Later, Kaveh and Manon in [KM19] gave a more general definition introducing the terminology *Khovanskii basis*. SAGBI/Khovanskii bases are the analog of Gröbner bases for ideals to subalgebras and they have been effectively used in computer algebra in [BPT25]. ◁

When the coordinate ring R admits a finite SAGBI basis, the initial algebra is a finitely generated monomial algebra, hence it can be viewed as the coordinate ring of a toric variety. In this case we say that X admits a *toric degeneration* to the projective toric variety $\operatorname{Proj}(\operatorname{in}_\prec(R))$.

Definition 2.3.4. A *toric degeneration* of $X \subseteq \mathbb{P}^n$ is a variety $Y \subseteq \mathbb{P}^n \times \mathbb{A}^1$ such that the projection $\pi\colon Y \to \mathbb{A}^1$ is flat, $\pi^{-1}(0) \hookrightarrow \mathbb{P}^n$ is a toric variety and $\pi^{-1}(\lambda) \cong X$ for every $\lambda \neq 0$.

Degenerations are used to formalize the concept of limit of a family of varieties as the *special fiber* $\pi^{-1}(0)$ can be thought of as the limit of $\pi^{-1}(t)$ for $t \to 0$.

Now we demonstrate this degeneration for the Grassmannian X_8. Recall that a weight vector w represents a term order $\prec$ for a finite set of polynomials $\mathcal{F}$ if $\operatorname{in}_w(f) = \operatorname{in}_\prec(f)$ for every $f \in \mathcal{F}$. A *diagonal term order* is a term order for which the initial terms of the Plücker coordinates p_i are the corresponding monomial in a main diagonal of (2.4).

Proposition 2.3.5. *[Stu96, Proposition 11.8] The Plücker coordinates are a finite SAGBI basis for the coordinate ring R of the Grassmannian X_8 with respect to any diagonal term order $\prec_d$, which is represented by the weight vector $w = (3,2,1,0,0,1,2,3)$ on the Plücker coordinates:*

$$\operatorname{in}_w(R) = \mathbb{K}[\operatorname{in}_w(p_0), \dots, \operatorname{in}_w(p_{14})].$$

We introduce a new variable u, that will be the parameter of the degeneration. For every polynomial $f = \sum_\alpha c_\alpha t^\alpha \in R$, we define $b_w(f) \coloneqq \max\{ w \cdot \alpha \mid c_\alpha \neq 0 \}$ and

$$\begin{aligned} f_u &\coloneqq \sum_\alpha c_\alpha t^\alpha \cdot u^{b_w(f) - \omega\cdot\alpha} \in \mathbb{K}[t_1, \dots, t_8, u], \\ R_u &\coloneqq \mathbb{K}[f_u \mid f \in R] \subseteq \mathbb{K}[t_1, \dots, t_8, u]. \end{aligned} \tag{2.6}$$

We recover the coordinate ring and its initial algebra as $R_0 = \operatorname{in}_w(R)$ and $R_1 = R$. By moving the parameter u from 1 to 0 we degenerate the coordinate ring R into a monomial algebra. The latter is the coordinate ring of a toric variety and this corresponds to degenerating X_8 into the toric variety $\operatorname{Spec}(R_0)$.

Remark 2.3.6. The previous construction is an application of the more general result by Anderson [And13] regarding toric degenerations of projective varieties whose coordinate ring has a full-rank valuation with finitely generated value semigroup. Taking w as before, the full-rank valuation

$$\nu\colon R \setminus 0 \to \mathbb{Z}^8, \qquad \nu\Big(\sum_\alpha c_\alpha t^\alpha\Big) = \operatorname{argmax}\{\, \omega \cdot \alpha \mid c_\alpha \neq 0 \,\}.$$

induces a term order represented by the vector w. The existence of a finite SAGBI basis for this term order is equivalent to having a finitely generated value semigroup for ν by [KM19]. ◁

Going back to Equation (2.5), we introduce the parameter u to degenerate it to a toric ideal as follows. We denote by $\underline{a}$ the entries of the matrix A from Lemma 2.3.2 and define the system $F(\underline{t}, u, \underline{a})$ with 8 equations by:

$$F(\underline{t}, u, \underline{a}) = \{A \cdot (p_{0,u}(\underline{t}), \ldots, p_{14,u}(\underline{t}))^\intercal = 0\}. \tag{2.7}$$

The variables $u, \underline{a}$ are the parameters in the homotopy, the equations in (2.5) correspond to the target system $F(\underline{t}, 1, \underline{a})$. Fixing random values $\underline{a}_\mathrm{s}$ we obtain the start system $F(\underline{t}, 0, \underline{a}_\mathrm{s})$ that involves only linear relations among monomials, i.e. the leading terms of the Plücker coordinates. `HomotopyContinuation.jl` can easily solve this by automatically choosing a polyhedral homotopy. Then we track the solutions to obtain the solutions of $F(\underline{t}, 1, \underline{a}_\mathrm{s})$. Given a matrix A with entries $\underline{a}_\mathrm{t}$, we solve $F(\underline{t}, 1, \underline{a}_\mathrm{t})$ as the following snippet of `Julia` code shows.

The equations in (2.7) are encoded in the vector `F`. The first homotopy computes the solution on the toric degeneration by moving the hyperplanes that intersect it and the parameter $u = 0$ is fixed. The second homotopy transports the solution from the toric degeneration to the Grassmannian X_8 by increasing the parameter u from 0 to 1 and fixing the parameters $\underline{a}$. The third homotopy finally solves $F(\underline{t}, 1, \underline{a}_\mathrm{t})$ by tracking the solution of $F(\underline{t}, 1, \underline{a}_\mathrm{s})$.

```
C          = System(F, variables = x, parameters = [u;a[:]])

A_rand     = randn(ComplexF64, length(a));
a0_start   = [0;A_rand]
toric_sol = solutions(HC.solve(C,              # solutions of F(t,0,a_s)
                target_parameters = a0_start))

a1_start   = [1;A_rand]
start_sol = solutions(HC.solve(C, toric_sol; # solutions of F(t,1,a_s)
                start_parameters = a0_start, target_parameters = a1_start))

a1_target = [1;A] # given A
X8_cap_A   = solutions(HC.solve(C, start_sol; # solutions of F(t,1,a_t)
                start_parameters = a1_start, target_parameters = a1_target))
```

Once we have the solutions P to $F(\underline{t}, 1, \underline{a}_t)$, we obtain a configuration of self-dual points Γ that is embedded in $\mathbb{P}^{14}$ by solving the linear system $L(\Gamma) = P$, where $L \coloneqq \mathrm{Ker}([\underline{a}_t])$ is a kernel basis. This is illustrated in the following snippet of code, the function `numerical_plücker` (numerically) evaluates the Plücker embedding on points in the affine chart $H \cong \mathbb{C}^8$.

```
Z = reshape(vcat(numerical_plücker.(X8_cap_A)...),15,14) # Embed points in P^14
L = LinearAlgebra.nullspace(A) # Obtain matrix L whose image is Ker(A)
Gamma = L\Z # Take the inverse image of Z under L
```

Remark 2.3.7. This illustrates the general principle in homotopy continuation methods that once you have a full solution set for one general parameter in your family (here $\underline{a}_s$), you can then use this to quickly solve other system in this family (here $\underline{a}_t$) without the initial *overhead* of the two-step degeneration (polyhedral+degeneration). ◁

2.3.3 Real solutions

For a transversal slicing of the Grassmannian X_8 with 8 complex hyperplanes, as described in Equation (2.5), we get 14 complex solutions. We study how many of these are real for random choices of real hyperplanes. The regions in $\mathrm{Gr}(7, \mathbb{R}^{15})$ corresponding to transversal linear spaces are partitioned into open semi-algebraic cells by the number of real intersection points.

We compute the solutions with `HomotopyContinuation.jl` using $F(\underline{t}, 1, \underline{a}_t)$ as start system, for which we computed a start solution `X8_cap_A` previously. The real hyperplanes are chosen as follows: we draw the coefficients of the defining equations from a standard normal distribution with mean 0 and standard deviation 1. This gives the uniform distribution on the Grassmannian $\mathrm{Gr}_{\mathbb{R}}(7, 15)$ with respect to the Haar measure from the compact real Lie group $\mathrm{O}(15)$ [Chi03, Thm. 2.2.1]. In Table 2.1 we report the result of the sampling after ten million iterations.

Real points	0	2	4	6	8	10	12	14	succ	fail
Count	178244	2277545	4079864	2547002	790369	116267	10508	192	9999991	9
	1.78%	22.78%	40.80%	25.47%	7.90%	1.16%	0.11%	0.002%	100%	-

Table 2.1: Slicing X_8 with uniformly sampled real $\mathbb{L} \subseteq \mathbb{P}^{14}(\mathbb{R})$ 10^7 times.

We observe that the case in which all 14 solutions are real is very rare but it does happen. The columns *succ* and *fail* report the number of times the computations succeeded in computing all the 14 distinct solutions and the number of times it failed.

2.4 Lifting using homotopy continuation

In the previous section we computed pairs $(\mathbb{L}, Z)$ with $X_8 \cap \mathbb{L} = Z \subseteq \mathbb{P}^{14}$. In this section these will be turned into the start solution of a parameterized polynomial system solving the lifting problem Problem 2.2.14.

2.4.1 The skew normal form homotopy

As mentioned, an important step for using Homotopy Continuation is parameterizing the space of self-dual points $\mathcal{A}_6$ in $\mathbb{P}^6$. The main problem is that this space is not convex (in an obvious way) and, more generally, it is crucial to find a path between two self-dual configurations that stays *inside* the space of self-dual points. Using the parametrization given by $\mathbb{C}^{21} \dashrightarrow \mathcal{A}_6$ that exploits the skew normal form in Theorem 2.2.11, we solve this problem efficiently: after computing normal forms for start and target configuration, we can choose a straight line through the parameter space $\mathbb{C}^{21}$. If s_{start} is chosen sufficiently randomly, this path will avoid the discriminant (branch locus) as it has real codimension 2.

With this structure, the Mukai lifting problem becomes the following: For $S \in \text{Skew}(7) \cong \mathbb{C}^{21}$ we want to recover the embedding $L\colon \mathbb{P}^6 \xrightarrow{\sim} \mathbb{L} \subseteq \mathbb{P}^{14}$ such that

$$L \cdot \{\gamma \in [I_7 + S \mid I_7 - S]\} = \mathbb{L} \cap X_8.$$

The following lemma describes a polynomial system of equations using the Plücker relations which accomplishes this.

Lemma 2.4.1. *Let γ_i be the i-th column of a set Γ in skew normal form and let $q_1, \dots q_{15}$ be the Plücker relations of X_8, i. e. the generators of the defining ideal of $X_8 \subseteq \mathbb{P}^{14}$. The embedding $L = [\ell_{ij}]\colon \mathbb{P}^6 \xrightarrow{\sim} \mathbb{L} \subseteq \mathbb{P}^{14}$ is a solution for the Mukai lifting problem for Γ if and only if*

$$q_k(L\gamma_i) = 0 \qquad \text{for all } i = 1, \dots, 14,\ k = 1, \dots, 15. \tag{2.8}$$

Proof. The only condition that L has to satisfy is that the image of Γ lies inside the Grassmannian X_8. This is ensured by the vanishing of the Plücker relations on $L \cdot \gamma_i$ for every $i = 1, \dots, 14$. Indeed if Equation (2.8) holds, then $L(\Gamma) \subseteq \mathbb{L} \cap X_8$ and we have equality as the latter is a set of 14 points by Lemma 2.3.2 and L is injective. □

We observe that the system in (2.8) has degree 2 in the 105 variables ℓ_{ij} and has $14 \cdot 15 = 210$ equations, so it is heavily over-determined. Additionally, this system is not zero-dimensional, which is a necessary condition for applying homotopy continuation methods. Indeed, Petrakiev's Theorem 2.2.15 implies that the rational map

$$f\colon \text{Gr}(7, \mathbb{C}^{15}) \dashrightarrow \mathcal{A}_6, \qquad \mathbb{L} \mapsto \mathbb{L} \cap X_8$$

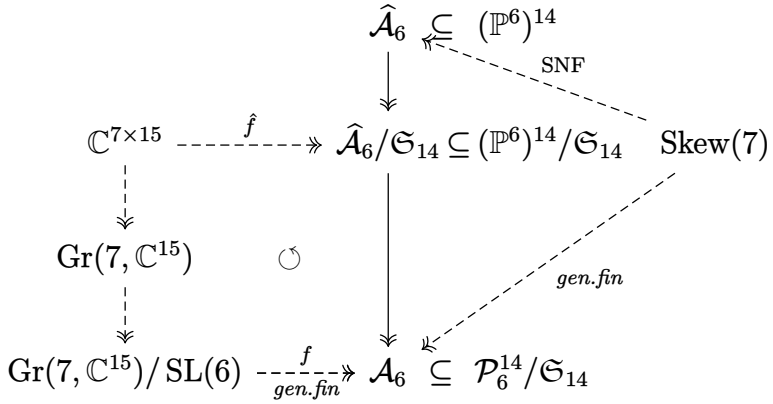
generically has 35-dim'l fibers (unions of orbits of the SL(6) action on $\mathbb{C}^{15} = \bigwedge^2 \mathbb{C}^6$).

Let $\widehat{\mathcal{A}}_6 \subseteq (\mathbb{P}^6)^{14}$ be the locally closed set of semistable self-dual point configuration. In Section 2.3 we considered the related rational map (note that we considered kernel matrices instead):

$$\hat{f}\colon \mathbb{C}^{7\times 15} \dashrightarrow \widehat{\mathcal{A}}_6/\mathfrak{S}_{14}, \qquad L \mapsto L^{-1}(\text{Im}(L) \cap X_8) \subseteq \mathbb{P}^6.$$

The relationship between f, $\hat{f}$ and the skew normal forms can be summarized as follows.

Theorem 2.4.2. *We have a commutative diagram*



In particular, $\hat{f}$ is dominant with general fiber of dimension $\dim \hat{f}^{-1}(\Gamma) = 36$.

Proof. Commutativity of the diagram follows from the previous discussion. Since f is dominant (Theorem 2.2.15) and the image of $\hat{f}$ is invariant under SL(7), $\hat{f}$ is dominant too. The dimension of a general fiber is

$$\dim \hat{f}^{-1}(\Gamma) = \dim \mathbb{C}^{7\times 15} - \dim \widehat{\mathcal{A}}_6 = 105 - 69 = 36. \qquad \square$$

To get a system with a finite set of solutions we need to impose a number of conditions on the entries of L equal to the dimension of a general fiber of $\hat{f}$. Hence we can construct L as follows using 69 variables ℓ_i, $i = 1, \dots, 69$:

As described in Section 2.3 we consider a self-dual start configuration Γ_{start} embedded in the linear space $\mathbb{L}_{\text{start}} \coloneqq L_{\text{start}}(\mathbb{P}^6) \subseteq \mathbb{P}^{14}$. By adding random linear combinations of the 69 variables to each entry of L_{start}, given by random matrices $A_k \in \mathbb{C}^{7\times 15}$ for $k = 1, \dots, 69$, we construct L as:

$$L = L_{\text{start}} + \sum_{k=1}^{69} \ell_k \cdot A_k$$

With this interpretation, the start solution for `HomotopyContinuation.jl` is a vector of zeros $\ell_{\text{start}} = (0, \dots, 0)$ and the start parameters are the entries of the skew-symmetric matrix S_{start} in a skew normal form of Γ_{start}.

The number of solutions to this system could be larger than the degree of the slicing map f (which is conjecturally birational), because we are intersecting SL(6)-orbits with linear spaces. Since we only want to construct *some* Mukai lift, we do not investigate this subtlety further, though it would be interesting to study this degree in future work.

We finally describe part of the construction of Equation (2.8) in `Julia`. We concatenate three different systems as follows:

- First we define `L_system` to be the system containing the equations of the image of a linear map $L\colon \mathbb{P}^6 \to \mathbb{P}^{14}$, as in Lemma 2.4.1 and L as described above in the variables $\mathtt{l} \mathrel{\hat{=}} (\ell_1, \dots, \ell_{69})$. The variables `x` are the coordinates of a point in $\mathbb{P}^6$.

- Then we evaluate the Plücker relations in these expressions, obtaining the system `Q_system` of 15 equations in `l` and `x`.
- Finally, we evaluate the equations of `Q_system` in each of the points in $\Gamma_{\mathtt{s}} = [I_7 + S | I_7 - S]$, keeping the variables $\mathtt{s} \mathrel{\hat{=}} S \in \mathrm{Skew}(7)$ as parameters.

A relevant part of the construction is summarized in the following snippet:

```
L_system = System(L*x, variables=[x;l])
plücker_system = System([oscar_to_HC_Q(plück_oscar[i], q) for i=1:15], variables=q)
Q = plücker_system(expressions(L_system));
Q_system = System(Q, variables=[x;l]);
equations = vcat([Q_system([Gamma[:,i];l]) for i in 1:14]...);
parameterized_system = System(equations, variables=l, parameters=s);
```

The full implementation together with the example notebook presented in the next section can be found at

https://mathrepo.mis.mpg.de/MukaiLiftP6

2.4.2 Towards testing Petrakiev's Conjecture

Petrakiev's dominant rational map $f\colon \mathrm{Gr}(7, \mathbb{C}^{15})/\mathrm{SL}(6) \dashrightarrow \mathcal{A}_6$ from Theorem 2.2.15 is generically finite. In the same paper the author conjectures that f is actually birational [Pet09, Remark 2.11]. In this section we indicate how one can use our implementation to test this conjecture.

Given a general point configuration $\Gamma \in \mathcal{A}_6$, the question is whether for any two $\mathbb{L}_1, \mathbb{L}_2 \in \mathrm{Gr}(\mathbb{P}^6, \mathbb{P}^{14})$ with $\mathbb{L}_i \cap \mathrm{Gr}(2,6) = \Gamma$ (in the sense that there exist linear isomorphisms $\mathbb{P}^6 \xrightarrow{\sim} \mathbb{L}_i$ sending Γ to $\mathbb{L}_i \cap \mathrm{Gr}(2,6)$), there is a matrix $G \in \mathrm{SL}(6)$ with $G \cdot \mathbb{L}_1 = \mathbb{L}_2$. Here, the action of $\mathrm{SL}(6)$ is the one induced by $\mathbb{P}^{14} = \mathbb{P}(\bigwedge^2 \mathbb{C}^6)$ described in Section 2.2.4. We can separate this test into the following steps:

(0) Pick a general $\Gamma \in \mathcal{A}_6$.

(1) Compute two preimages $\mathbb{L}_1, \mathbb{L}_2$ of Γ under $\mathrm{Gr}(\mathbb{P}^6, \mathbb{P}^{14}) \dashrightarrow \mathcal{A}_6$.

(2) Try to compute a matrix $G \in \mathrm{SL}(6)$ with $G \cdot \mathbb{L}_1 = \mathbb{L}_2$.

Step (0) can be achieved using any of the normal forms discussed in Section 2.2.2, most conveniently a skew normal form obtained from a random complex skew matrix.

To get one preimage one can run the lifting algorithm described in the previous section. To obtain a second preimage $\mathbb{L}_2$, one can run a monodromy-based approach on this parametrized polynomial system [MR17]. Since the graph of $\mathrm{Gr}(7, \mathbb{C}^{15})/\mathrm{SL}(6) \dashrightarrow \mathcal{A}_6$ is irreducible, one can obtain any other element in the fiber by tracking the first element along a homotopy with start and end point Γ. To obtain such paths, one can take any path connecting two skew normal forms of Γ, for example by piecing together line

segments in Skew(7). Figure 2.2 describes one such path: From an initial starting parameter S_0, move to S_1 and then in a triangle $S_1 \to S_2 \to S_3 \to S_1$.

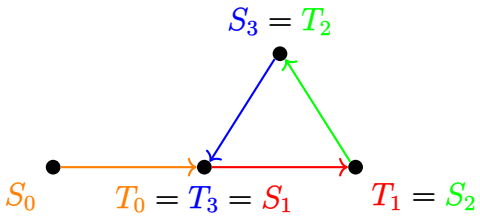


Figure 2.2: A schematic monodromy loop.

Recall that our algorithm doesn't directly work on $\mathrm{Gr}(7, \mathbb{C}^{15})/\mathrm{SL}(5)$, but rather in an affine subspace of $\mathbb{C}^{7\times 15}$ which maps generically finitely to $\widehat{\mathcal{A}}_6/\mathfrak{S}_{14}$ under $\hat{f}$ (Theorem 2.4.2). This enables us to compute suitable $\mathbb{L}_2$, therefore making step (1) possible. An example of two such matrices are also included in the aforementioned MathRepo page.

Step (2) on the other hand, is an instance of an *orbit membership problem*: Representing $\mathbb{L}_i$ as matrices $L_i \in \mathbb{C}^{15\times 7} \cong (\bigwedge^2 \mathbb{C}^6) \otimes \mathbb{C}^7$, one needs to decide the existence

$$\exists G \in \mathrm{SL}(6), B \in \mathrm{GL}(7) \qquad \text{such that} \qquad B \cdot L_1 \cdot \wedge^2 G = L_2 \quad ?$$

This is a difficult algorithmic problem, and has been studied intensively in the context of orbit closure problems in computational invariant theory [DK02]. We leave the computational solution of this numerical algebraic geometry problem for future work.

2.5 Implementation and extended example

In this final section we explain the implementation through an instructive example. We consider the self-dual configuration $\Gamma \subseteq \mathbb{P}^6$ given by the columns of the following matrix:

$$\Gamma = \begin{bmatrix} 7 & -2 & 6 & -1 & -6 & 1 & -9 & 7 & 0 & 6 & 1 & 8 & -3 & 7 \\ -1 & 2 & -5 & -2 & 0 & -4 & 3 & -3 & -4 & -3 & 4 & -2 & 4 & -1 \\ 1 & 4 & -1 & -5 & -3 & 6 & 8 & -1 & -8 & -3 & 5 & 1 & -6 & -8 \\ 3 & -6 & 4 & -3 & -4 & 6 & 0 & 5 & 8 & 2 & 3 & 2 & -8 & 0 \\ 1 & -2 & 1 & 0 & -4 & 2 & 2 & 3 & 4 & -1 & 2 & 2 & -2 & -2 \\ 0 & -6 & -5 & 6 & 3 & 7 & -3 & 2 & 8 & -7 & -6 & -3 & -5 & 5 \\ -3 & 3 & -4 & 1 & 4 & 3 & 2 & -3 & -6 & -4 & -3 & -4 & -1 & -2 \end{bmatrix}$$

To verify that Γ is indeed self-dual, we use condition (ii) of Lemma 2.2.1 and compute an invertible diagonal matrix Λ verifying $\Gamma \cdot \Lambda \cdot \Gamma^\intercal = 0$. This is a linear system in 14 unknowns (the entries of Λ) and $\binom{7+1}{2} = 28$ equations, namely $\sum_{i=1}^{14} \lambda_i \gamma_i \gamma_i^\intercal = 0$. The function `certify_selfdual` returns a floating point approximation of the (exact) normed solution

$$\Lambda = \frac{1}{\sqrt{13.0625}} \operatorname{diag}(-1,-1,-1,-1,-1,-1,-1,1,\tfrac{1}{4},1,1,1,1,1).$$

```
Lambda = certify_selfdual(Gamma)
norm( Gamma * Lambda * transpose(Gamma), Inf ) #2.6867397195928788e-14
```

We bring Γ into an orthogonal normal form Γ_{ONF} as in Theorem 2.2.5, essentially using linear algebra. Applying an appropriate scaling Λ_{scale} and linear map A we have $A \cdot \Gamma_{\mathrm{ONF}} = \Gamma \cdot \Lambda_{\mathrm{scale}}$.

$$\Gamma_{\mathrm{ONF}} = [I_7 \mid P] = \begin{bmatrix} 1&0&0&0&0&0&0&-0.3399&0.06&-0.0924&0.235&0.7564&-0.3033&0.3911 \\ 0&1&0&0&0&0&0&-0.2052&-0.8506&0.0197&-0.1977&0.268&0.1942&-0.2922 \\ 0&0&1&0&0&0&0&0.1008&0.2302&0.8118&-0.2422&0.3148&-0.1311&-0.3208 \\ 0&0&0&1&0&0&0&0.0948&0.319&-0.4084&-0.722&0.3291&0.3032&-0.0306 \\ 0&0&0&0&1&0&0&-0.7734&0.0615&0.1552&-0.4091&-0.3747&-0.1609&0.2009 \\ 0&0&0&0&0&1&0&0.2627&-0.1669&-0.2427&-0.2875&-0.072&-0.8558&-0.1552 \\ 0&0&0&0&0&0&1&-0.3951&0.2942&-0.2868&0.2752&0.055&-0.0546&-0.7703 \end{bmatrix}$$

$$A = i \begin{bmatrix} 3.6821 & -1.052 & 3.1561 & -0.526 & -3.1561 & 0.526 & -4.7341 \\ -0.526 & 1.052 & -2.63 & -1.052 & 0.0 & -2.104 & 1.578 \\ 0.526 & 2.104 & -0.526 & -2.63 & -1.578 & 3.1561 & 4.2081 \\ 1.578 & -3.1561 & 2.104 & -1.578 & -2.104 & 3.1561 & 0.0 \\ 0.526 & -1.052 & 0.526 & 0.0 & -2.104 & 1.052 & 1.052 \\ 0.0 & -3.1561 & -2.63 & 3.1561 & 1.578 & 3.6821 & -1.578 \\ -1.578 & 1.578 & -2.104 & 0.526 & 2.104 & 1.578 & 1.052 \end{bmatrix} \qquad \Lambda_{\mathrm{scale}} = \sqrt{\Lambda[1{:}7]} \oplus -\sqrt{\Lambda[8{:}14]}$$

```
Gamma_ONF, A, Lambda_scale = orthogonal_normal_form(Gamma)
norm( A * Gamma_ONF - Gamma * Lambda_scale, Inf ) # 2.6645352591003757e-15
```

We compute a skew normal form Γ_{SNF} by applying the Cayley transform to P to obtain the skew-symmetric matrix S and we compute $\Gamma_{\mathrm{SNF}} = [I_7 + S | I_7 - S]$. The implementation ensures that $P \in \mathrm{SO}(7)$ so that the Cayley transform is defined, as pointed out in Remark 2.2.6. The output is a floating point approximation of the following exact matrix:

$$S = \begin{bmatrix} 0&-1&1&0&-1&4&2 \\ 1&0&3&-2&-2&10&12 \\ -1&-3&0&2&1&-1&-2 \\ 0&2&-2&0&-1&-10&-6 \\ 1&2&-1&1&0&-4&-4 \\ -4&-10&1&10&4&0&-6 \\ -2&-12&2&6&4&6&0 \end{bmatrix} \qquad \Gamma_{\mathrm{SNF}} = \begin{bmatrix} 1&-1&1&0&-1&4&2&1&1&-1&0&1&-4&-2 \\ 1&1&3&-2&-2&10&12&-1&1&-3&2&2&-10&-12 \\ -1&-3&1&2&1&-1&-2&1&3&1&-2&-1&1&2 \\ 0&2&-2&1&-1&-10&-6&0&-2&2&1&1&10&6 \\ 1&2&-1&1&1&-4&-4&-1&-2&1&-1&1&4&4 \\ -4&-10&1&10&4&1&-6&4&10&-1&-10&-4&1&6 \\ -2&-12&2&6&4&6&1&2&12&-2&-6&-4&-6&1 \end{bmatrix}$$

```
P = Matrix{Float64}(Gamma_ONF[:,8:14])
det(P) #1.0000000000000009
S = cayley(P)
Gamma_SNF = [I+S I-S]
```

We compute the linear embedding $\widehat{L}\colon \mathbb{P}^6 \dashrightarrow \mathbb{L} \subseteq \mathbb{P}^{14}$ such that $\widehat{L}(\Gamma_{\mathrm{SNF}}) = \mathbb{L} \cap X_8$. The parametrized polynomial system in Equation (2.8) has been precomputed as `F` as described in Section 2.4 together with start solution `l_start` $= 0$ and start parameters `S_start`. The target parameter of the homotopy are the entries of the matrix S. We executed this computation on a single thread of a 512 GB RAM machine using 2x 12-Core Intel Xeon E5-2680 v3 processor working at 2.50 GHz in 3172 seconds.

```
S_target = skew_to_vector(S) # unrolls the upper right half of S into its 21 entries
@time result = HomotopyContinuation.solve(F, l_start;
               start_parameters=S_start, target_parameters=S_target)
sol = solutions(result)[1]
L_hat = L_start + sum(sol[i]*A_rand[i] for i in eachindex(l))
```

$$\widehat{L} = \begin{bmatrix}
0.11+1.48i & -0.31-0.31i & -2.09-0.9i & 2.23-1.62i & -0.66-4.3i & 1.75-0.77i & 0.33-0.75i \\
2.37-1.0i & -0.66+0.3i & -0.77+1.73i & -0.59-1.0i & 0.5-0.2i & 0.52-1.16i & -0.02+0.15i \\
-0.37-1.9i & -0.34+0.46i & 1.07+0.61i & -0.27-0.56i & -0.91-1.21i & -0.37-0.29i & -0.17-0.63i \\
-1.74-1.36i & 0.57-0.67i & 0.71+0.34i & 0.77+1.58i & 3.09+0.01i & -0.6+1.76i & 1.07-0.91i \\
-0.86-0.15i & 0.42+0.27i & -0.09-0.18i & -0.19-0.36i & -2.18+0.17i & -0.65-0.3i & -0.09-0.15i \\
-0.26+2.25i & 1.12-0.89i & 0.92-1.52i & 4.07+1.13i & 0.87-0.98i & -1.68+1.66i & 1.67-0.93i \\
1.34+0.68i & -0.12-0.38i & -0.91-0.23i & -0.12-1.23i & 0.95-1.19i & 1.04+1.04i & -0.42-0.64i \\
0.19-1.7i & -1.69+0.8i & -0.39-0.27i & -2.99-0.63i & 0.67+2.89i & 0.45+0.98i & -1.3+1.11i \\
-0.43-0.07i & -0.93-0.57i & -1.16-1.07i & -0.12-1.08i & 1.06-2.44i & 2.42+0.22i & -0.73-0.21i \\
-0.34-0.7i & -0.09+0.37i & 0.49+0.02i & -1.14+0.16i & -0.27-0.16i & 0.47-0.24i & -0.65+0.31i \\
0.79+1.62i & -0.44-1.02i & 0.04-0.1i & 2.02-2.51i & -0.53-0.65i & 0.58+2.37i & -0.52-1.89i \\
0.54+0.89i & -0.2+0.19i & -1.1+0.96i & -0.45+0.8i & 0.53+0.7i & 0.35-1.09i & 0.18+0.8i \\
1.57-0.58i & -0.55+0.14i & -0.14+0.07i & -1.58-0.16i & 0.19-0.88i & 1.54-0.12i & -0.93+0.57i \\
0.58-0.81i & -0.46+0.06i & 0.6+0.77i & 0.04-0.56i & 0.59-0.73i & -0.25+0.0i & 0.14-0.51i \\
0.72-0.79i & 0.92-1.27i & 0.33+0.38i & 1.78-1.0i & 0.64+1.39i & -1.98+1.87i & 1.05-1.57i
\end{bmatrix}.$$

To obtain the map $L\colon \mathbb{P}^6 \hookrightarrow \mathbb{P}^{14}$ with $L(\Gamma) \subseteq X_8$, one only needs to compose $\widehat{L}$ with the previous coordinate transformations $\Gamma \rightsquigarrow \Gamma_{\text{SNF}}$:

$$\mathbb{L} \cap X_8 = \widehat{L}(\Gamma_{\text{SNF}}) = \widehat{L}\left((I_7 + S) \cdot \Gamma_{\text{ONF}}\right) = \widehat{L}\left((I_7 + S) \cdot A^{-1} \cdot \Gamma \cdot \Lambda_{\text{scale}}\right).$$

Since the configuration Γ does not change when multiplying on the right by a diagonal matrix, we obtain the solution for the Mukai lifting problem for Γ as:

$$L = \widehat{L} \cdot (I_7 + S) \cdot A^{-1}.$$

We can verify $L(\Gamma) = \mathbb{L} \cap X_8$ by evaluating $L(\gamma_i)$ in the Plücker relations

```
L = L_hat*(I+S)*inv(A)
maximum([ norm(q(L*Gamma[:,i]), Inf) for i = 1:14]) #7.907992711624597e-12
```

Chapter 3

Computational Complexity of Polynomial Subalgebras

Contents

> The core challenge for computing science is hence a conceptual one, viz. what (abstract) mechanisms we can conceive without getting lost in the complexities of our own making.
>
> EDSGER W. DIJKSTRA [Dij86]

In the previous two chapters, we have studied two computationally challenging problems from algebraic geometry. The slicing problem lead us to consider calculations in the Plücker ring, the projective coordinate ring of the Grassmannian $\mathrm{Gr}(2, \mathbb{C}^6)$, generated by the Plücker coordinates. Computation in this polynomial subalgebra significantly simplified the problem at hand, and subalgebras themselves have become a more prominent subject in pure and applied algebra in recent years [Rei03; Bur+24].

In this chapter, we take a step back and consider the inherent computational complexity of a related problem: The *subalgebra membership problem*, and its complexity-theoretic relation to the *ideal membership problem*. The complexity of polynomial ideals and Gröbner bases has been studied extensively at least since the 1980's [MM82; May89; MT17]. Their complexity has been placed in the class EXPSPACE, with more fine-grained results available. The main goal of this chapter is to parallel and contrast this with subalgebra membership:

Main Results 3 (3.3.4, 3.3.6, 3.3.8, 3.4.3, 3.5.12, 3.5.13)**.** *Let $\mathbb{K}$ be a well-endowed field.* $\textsc{AlgMem}_{\mathbb{K}}$ *is* EXPSPACE*-complete. The degree of a membership certificate p for $g \in \mathbb{K}[f_1, \dots, f_s]$, that is, $g = p(f_1, \dots, f_s)$ can be bounded by*

$$\deg p \leq \deg(g)^{n+1} \cdot d^{O(s^2(n+s)^2 2^n)}.$$

The problem is PSPACE*-complete when restricted to homogeneous polynomials, and* NP*-complete when restricted to monomial algebras.*

We also propose the class of semilinear monomial algebras, and show that infinitely generated initial algebras of certain invariant rings have semilinear initial algebras.

This chapter is based on the paper *Computational Complexity of Polynomial Subalgebras* [Kay25], as well as currently unpublished joint work with Bernhard Reinke in section Section 3.5.3. It is structured as follows:

Section 3.1 features a crash course in computational complexity theory, introducing relevant notions such as complexity classes and reductions. In Section 3.2 we recall the key results regarding the complexity of $\textsc{IdealMem}_{\mathbb{K}}$, focussing on the results that will be useful for our analysis. Our contributions start in the heart of the chapter in Section 3.3, where we prove degree bounds and complexity upper bounds on subalgebra membership using normal form calculations. In Section 3.4 we provide matching complexity lower bounds, closing the circle of reductions. We finish in Section 3.5 by considering the special case of monomial subalgebras and begin to investigate some (conjectural) structure of infinitely generated initial algebras.

Throughout the chapter, $\mathbb{K}$ denotes an arbitrary field; often specified to allow for efficient computation of arithmetic operations (Definition 3.2.1).

3.1 Background: Computational complexity theory

Computational complexity studies, among other things, the amount of computational resources required to solve a given mathematical problem. The precise answer depends on the nature of the problem, the model of computation and the kind of resource we are measuring. We will be (mostly) using the Turing model of computation, and consider the memory required to compute the answer to the membership question (a *decision problem*).

In this section we briefly introduce the necessary concepts from complexity theory to state and understand the known results on the complexity of polynomial ideals and to state and prove our results on subalgebra membership. A classical reference for computational complexity theory is the book by Hopcroft and Ullman [HU79], a more modern treatment is the book by Arora and Barak [AB09], for a brief introduction with a view towards computer algebra, see also [Joa17, Section 25.8].

3.1.1 Problems and algorithms

Denote by $\mathcal{P}(X)$ the powerset of a set X, consisting of all subsets of X.

Definition 3.1.1 (Decision/function problem)**.** An *alphabet* is a finite set Σ. The set of all finite character strings is $\Sigma^* = \bigsqcup_{n\geq 0} \Sigma^n$. The length of a string $w \in \Sigma^n$ is $|w| = n$.

Let Σ, Δ be alphabets. A *decision problem* is a subset $A \subseteq \Sigma^*$. A *function problem* is a function $F\colon \Sigma^* \to \mathcal{P}(\Delta^*)$ which assigns to every input $w \in \Sigma^*$ a set of valid outputs $F(w) \subseteq \Delta^*$.

One should think of a decision problem as a question which has for every $x \in \Sigma^*$ a yes/no answer; the decision problem being the subset of "yes"-instances. Similarly, a function problem asks for every input $x \in \Sigma^*$ to compute a valid output $y \in \Delta^*$, $y \in f(x)$.

Example 3.1.2 (Boolean satisfiability)**.** A common type of decision problem is checking whether a given formula has a valid solution. The most important one in theoretical computer science is satisfiability of Boolean formulas: A boolean formula φ is a well-formed expression over the alphabet $\Sigma = \{x, \mathtt{0}, \mathtt{1}, (,), \wedge, \vee, \neg\}$, such as

$$\varphi = x\mathtt{0} \wedge (\neg x\mathtt{1} \vee x\mathtt{10}) \mathrel{\hat{=}} x_0 \text{ and } ((\text{not } x_1) \text{ or } x_{10}).$$

Here, well-formed means that it is a syntactically correct formula (parentheses match, no $\wedge$ after $\vee$, indices only on x, ...). When assigning truth values ($\mathtt{0} \mathrel{\hat{=}}$ true, $\mathtt{1} \mathrel{\hat{=}}$ false) to the variables x_i, a formula φ is satisfied if it evaluates to $\mathtt{1}$ under the usual semantics. For example, the previous formula evaluates to $\mathtt{1}$ under the assignment $x_0 \mapsto \mathtt{1}$, $x_1 \mapsto \mathtt{0}$, $x_{10} \mapsto \mathtt{0}$. φ is in 3-conjunctive normal form (3CNF) if it is of the form

$$\varphi = (\ell_{11} \vee \ell_{12} \vee \ell_{13}) \wedge \cdots \wedge (\ell_{k1} \vee \ell_{k2} \vee \ell_{k3}),$$

where the ℓ_{ij} are *literals*

$$\ell_{ij} = x\mathrm{bin}(a_{ij}) \text{ or } \neg x\mathrm{bin}(a_{ij}), \quad a_{ij} \in \mathbb{N}.$$

The computational problem of satisfiability can be declared in the following way:

3SAT	
Input:	$\varphi(x_1, \ldots, x_n)$ a Boolean formula in 3CNF
Output:	Is there a satisfying assignment $\{x_1, \ldots, x_n\} \to \{0, 1\}$?

1IN3SAT	
Input:	$S_1, \ldots, S_k \subseteq \mathbb{N}$ sets, $\lvert S_i \rvert \leq 3$
Output:	Is there a set $T \subseteq \mathbb{N}$ such that $\lvert T \cap S_i \rvert = 1$ for $i = 1, \ldots, k$?

Here 1IN3SAT is a variant where all literals are without negation, but *exactly* one of the variables in a clause $\ell_{i1} \vee \ell_{i2} \vee \ell_{i3}$ evaluates to 1 (compare [Sch78], also known as *positive/monotone* 1IN3SAT). ◁

Note that we have not completely specified the encoding of the mathematical objects involved (formulas, integers, sets, ...), nor what a "well-formed" formula is. While technically required, the actual details are not important for the complexity of the problems. More specifically, complexity classes are intentionally defined in a way that is insensitive to the details of the encoding, or even to smaller changes to the computational model used [AB09, Chapter 1]. For this reason we will only comment on the actual representation of the mathematical objects when the encoding has an effect on the computational complexity.

Example 3.1.3 (Diophantine equations)**.** A classical question in mathematics is the solvability of polynomial equations in the integers. Here the input is a polynomial (of a particular form), and the problem is to decide whether it has a root in the integers.

HILBERT10	
Input:	$f(x_1, \ldots, x_n) \in \mathbb{Z}[x_1, \ldots, x_n]$
Output:	Is $f(c_1, \ldots, c_n) = 0$ for some $c_1, \ldots, c_n \in \mathbb{Z}$?

UNBOUNDEDSUBSETSUM	
Input:	$a_1, \ldots, a_n, b \in \mathbb{N}$
Output:	Is $\sum_{i=1}^{n} c_i a_i = b$ for some $c_1, \ldots, c_n \in \mathbb{N}$?

Famously, Hilbert's tenth problem is undecidable by the Matiyasevich–Robinson–Davis–Putnam theorem [Mat70], while linear diophantine equations are comparatively easy to solve. ◁

Next, we briefly introduce the Turing model of computation and some variants, to formally define what "computation" means to us. The formal definition of a Turing machine and its transition behavior are quite technical, so we invite the reader to gain intuition from Figure 3.1 and the following Example 3.1.6.

Definition 3.1.4 (Turing machine)**.** A k-tape Turing machine is a tuple $M = (\Gamma, Q, \delta)$, where

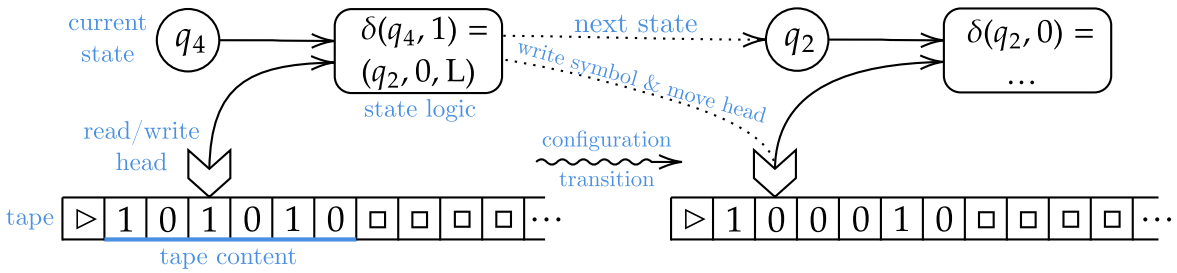


Figure 3.1: Visualization of a 1-tape Turing machine doing one configuration transition.

(i) Γ is a finite set of *tape symbols*, containing a (distinguished) *blank symbol* $\square \in \Gamma$ and a (left) boundary symbol $\triangleright \in \Gamma$;

(ii) Q is a finite set of *states*, containing a starting state q_0 and a halting state q_{halt}.

(iii) δ is the (non-deterministic) transition function which, given the current state and tape symbols, returns possible successive states and behavior of the read/write head

$$\delta \colon Q \times \Gamma^k \to \mathcal{P}(Q \times \Gamma^k \times \{\mathrm{L}, \mathrm{R}\}).$$

The machine is *deterministic* if δ actually maps to $Q \times \Gamma^k \times \{\mathrm{L}, \mathrm{R}\}$ (instead of a set of such tuples).

Definition 3.1.5 (Configuration transition, accepting)**.** A *configuration* is a tuple

$$(q, (i^{(j)}, \triangleright b_1^{(j)} b_2^{(j)} \dots)_{j=1,\dots,k}) \in Q \times (\mathbb{Z} \times \Gamma^*)^k.$$

Here, $q \in Q$ is the current state, $i^{(1)}, \dots, i^{(k)} \in \mathbb{Z}$ are the positions of the read/write heads, and $\triangleright b_1^{(j)} b_2^{(j)} \dots$ is the content of tape j.

The transition relation $\vdash_M^*$ is the transitive hull of the "successive configuration relation" (for notational convenience only spelled out in the single-tape case)

$$(q, i, \triangleright b_1 \dots b_{i-1} b_i b_{i+1} \dots) \vdash_M \begin{cases} (q', i-1, \dots b_{i-1} c b_{i+1} \dots) & \text{if } \delta(q, b_i) \ni (q', c, \mathrm{L}), \\ (q', i+1, \dots b_{i-1} c b_{i+1} \dots) & \text{if } \delta(q, b_i) \ni (q', c, \mathrm{R}). \end{cases}$$

The tape content is understood to be infinitely padded to the right with the blank symbol ($\triangleright b_1 \cdots b_l \square\square \dots$), so that the case $i > |b|$ never occurs. The left boundary markers $\triangleright$ may never be overwritten and the head may not pass over it (e.g. $\delta(q, \triangleright) \ni (q', \triangleright, \mathrm{L})$ is not allowed). The machine M *halts* on the input $w = w_1 \cdots w_n$ if

$$\big(q_0, ((1, \triangleright w_1 \cdots w_n), (1, \triangleright), \dots, (1, \triangleright))\big) \vdash_M^* (q_{\text{halt}}, (i^{(j)}, b^{(j)})_j)$$

for some *final configuration* $(i^{(j)}, b^{(j)})_j$. In this context, the first tape is the *input tape* and the last tape is the *output tape*. A decision problem $\mathrm{A} \subseteq \Sigma^*$ is *decided* by M if M halts on every input from Σ^* and

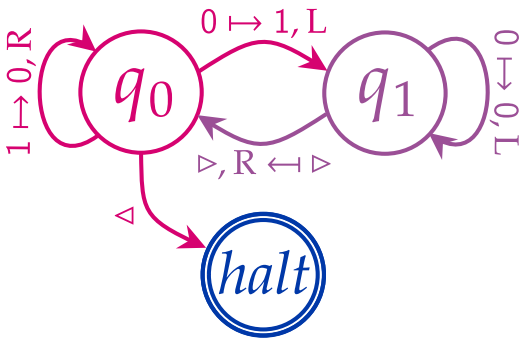


Figure 3.2: The linearly bounded automaton from Example 3.1.6.

(0) if $w \notin A$, write 0 on the output tape in *every* final configuration;

(1) if $w \in A$, writes 1 on the output tape in *some* final configuration.

Example 3.1.6. The following table shows the transition function δ of a deterministic single-tape Turing machine ($\Gamma = \{0, 1, \triangleright, \triangleleft\}, Q = \{q_0, q_1, q_{\text{halt}}\}, \delta$) ($\triangleleft$ is the blank symbol):

b	$\triangleright$	0	1	$\triangleleft$
q_0		$(q_1, 1, L)$	$(q_0, 0, R)$	halt
q_1	$(q_0, \triangleright, R)$	$(q_1, 0, L)$		

Transitions that do not occur (on input $w \in \{0,1\}^*$) have been left unspecified for simplicity. The transition function can be visualized as a decorated graph as in Figure 3.2. On input $w \coloneqq 0 \ldots 0$ (n zeros), this machine will count in binary to $1 \ldots 1$ and then halt. An example of the configuration sequence on input 000 can be found at mathrepo.mis.mpg.de/ComplexityOfSubalgebras/configurations.html. ◁

Remark 3.1.7. The difference between deterministic and non-deterministic Turing machines is that in the deterministic model, at any point in the computation there is at most one next transition step possible, while in the non-deterministic model, a finite set of transitions is possible. Acceptance in the non-deterministic setting means that there *exists* a sequence of possible transitions that reach a final configuration which accepts. ◁

The machine in the previous example does not leave the input area $\triangleright w \triangleleft$. Such a machine is called a *linearly bounded automaton* (LBA), and its computational capabilities are closely tied to *context-sensitive languages*.

Example 3.1.8 (Halting problems)**.** A source of interesting computational problems are halting problems for various models of computation. Turing proved that the problem whether a Turing machine halts (on an empty input tape) can *not* be decided by a Turing machine. We state the halting problem for linearly bounded automata for later reference.

LBAHALT	
Input:	A deterministic LBA $M = (Q, \Gamma = \{0, 1, \triangleright, \triangleleft\}, \delta)$ and an input $w \in \{0,1\}^*$
Output:	Does M halt on input w?

Note that this problem is indeed decidable by a Turing machine: One can simulate the transition sequence of M, noting which configurations have been encountered, and reject if a loop is detected. ◁

3.1.2 Complexity classes

Now that we have a computational model, we can mathematically reason about the difficulty of solving (deciding) a problem with a Turing machine. Recall the Landau notation for functions $f, g\colon \mathbb{N} \to \mathbb{N}$:

$$f \in O(g) \iff \limsup_{n\to\infty} f(n)/g(n) < \infty \iff \exists c, n_0 \text{ s.t. } f(n) \leq c \cdot g(n) \text{ for all } n \geq n_0$$
$$f \in o(g) \iff \lim_{n\to\infty} f(n)/g(n) = 0.$$

Definition 3.1.9 (Time and space)**.** Let $f\colon \mathbb{N} \to \mathbb{N}$ be a function and $\mathrm{A} \subseteq \Sigma^*$ a decision problem.

(i) $\mathrm{A} \in \mathsf{TIME}(f)$ if there exists a deterministic Turing machine M which on input w decides $w \in \mathrm{A}$ using $O(f(|w|))$ steps.

(ii) $\mathrm{A} \in \mathsf{SPACE}(f)$ if there exists a deterministic Turing machine M which on input w decides $w \in \mathrm{A}$ and visiting $O(f(|w|))$ cells[1].

$\mathsf{NTIME}(f)$ and $\mathsf{NSPACE}(f)$ are defined analogously, dropping the requirement of determinacy. We define polynomial and exponential time and space complexity classes as

$$\mathsf{P} \coloneqq \mathsf{TIME}(n^{O(1)}) = \bigcup_{c\geq 1} \mathsf{TIME}(n^c), \qquad \mathsf{NP} \coloneqq \mathsf{NTIME}(n^{O(1)}), \qquad \mathsf{EXP} \coloneqq \mathsf{TIME}(2^{n^{O(1)}})$$
$$\mathsf{L} \coloneqq \mathsf{SPACE}(\log(n)), \qquad \mathsf{PSPACE} \coloneqq \mathsf{SPACE}(n^{O(1)}), \qquad \mathsf{EXPSPACE} \coloneqq \mathsf{SPACE}(2^{n^{O(1)}}).$$

One has the following chain of inclusions

$$\mathsf{L} \subseteq \mathsf{P} \subseteq \mathsf{NP} \subseteq \mathsf{PSPACE} \subseteq \mathsf{EXP} \subseteq \mathsf{EXPSPACE}.$$

The space hierarchy theorem says that if $f \in o(g)$, then $\mathsf{SPACE}(f) \subsetneq \mathsf{SPACE}(g)$ (under mild technical conditions on g). Thus the inclusions of the space classes above are strict; similarly there is a time hierarchy theorem. We note an alternative characterization of NP, which is useful to show that a problem is in NP [AB09, Thm. 2.6].

Lemma 3.1.10 (NP = P-verifiable)**.** *A problem $\mathrm{A} \subseteq \Sigma^*$ is in NP if and only if there is a deterministic polynomial time Turing machine M such that a word $w \in \Sigma^*$ is in A if and only if there exists a* certificate *$y \in \Sigma^*$ of size $|y| = |w|^{O(1)}$ such that M accepts (w, y).*

[1] Technically, if $f(n) < n$, one should understand the input tape as read-only and only count the cells visited on the other tapes (otherwise it would be impossible to completely read the input).

Later, we will deal with function problems whose outputs can be extremely long compared to the input, such as Gröbner basis computation. To fairly judge the space complexity of such a problem, we only count the "internally used" working memory:

Definition 3.1.11 (Function complexity)**.** A function problem $F\colon \Sigma^* \to \Delta^*$ is computable in space g if there is a deterministic k-tape Turing machine ($k \geq 2$) M with a designated input and output tape, which, given $x \in \Sigma^*$, writes a $y \in F(x)$ on the output tape, and visits at most $g(|x|)$ cells on the tapes excluding input and output tape. Here, the input tape is read-only and the output tape is write-only.

Example 3.1.12 (Not all long problems are difficult)**.** Consider the function problem

$$F\colon \{0,1\}^* \to \{0,1\}^*, \qquad F(\mathrm{bin}(n)) = \mathrm{bin}(2^n).$$

The length of $F(x)$ is exponential in $|x|$, hence F can not be computed in polynomial space in the naive sense. However, since we don't count the space used to write the result to the output tape, one can simply count in binary from 0 to x and write as many 0's to the output, all in space $O(|x|)$. ◁

We return to the task of classifying the complexity of computational problems.

- A *complexity upper bound* for a decision problem A and a complexity class C is an algorithm solving A and satisfying the resource constraints of C.
- A *lower bound* or C-*hardness* result for A is a proof that the computation of any problem A' $\in$ C can be reduced to the problem A.

We use the notion of log-space many-one reductions to formalize this:

Definition 3.1.13 (Reduction, completeness)**.** Let A $\subseteq \Sigma^*$, B $\subseteq \Delta^*$, be decision problems. A is *log-space many-one reducible* to B, in symbols A $\leq_{\mathrm{m}}^{\mathsf{L}}$ B, if there is a function $f\colon \Sigma^* \to \Delta^*$, computable in logarithmic working space, such that $x \in$ A if and only if $f(x) \in$ B.

A problem A is C-*hard*, if A$' \leq_{\mathrm{m}}^{\mathsf{L}}$ A for all A$'$ in C. It is C-*complete* if it is also in C.

Example 3.1.14. All decision problems introduced so far are complete for some class:

- The famous *Cook–Levin theorem* states that 3SAT is NP-complete, and so is 1IN3SAT [Sch78]. The proof is a *generic reduction*, meaning one fixes a $A \in$ NP decided by a non-deterministic polynomial time machine M, and gives a reduction $A \leq_{\mathrm{m}}^{\mathsf{L}}$ 3SAT which expresses the computation of M as a large Boolean formula.
- UNBOUNDEDSUBSETSUM is also NP-complete, but only if the integers are encoded in binary — the unary version is significantly easier (see Example 3.1.21). For NP-membership one can use Lemma 3.1.10; the certificate y are the coefficients.
- LBAHALT is complete for PSPACE. The apparent limitation of the "linear boundedness" of LBAs still allows us to reduce any PSPACE-computation to LBAHALT by (artificially) increasing the length of the LBA in the reduction.

These examples also illustrate that the ability of log-space computable functions to increase length polynomially is a feature and not a bug. If one restricts to linear log-space reductions (which require $|f(x)| = O(|x|)$), then the above completeness results have to be modified. For example, LBAHALT is then "only" $\mathsf{NSPACE}(n)$-complete. ◁

3.1.3 Circuit complexity

We end this crash course in computational complexity theory by briefly recalling some notions from circuit complexity required to define well-endowed fields. A standard reference for circuit complexity is [Vol99].

For a vertex v of a directed graph, we denote by $\deg_{\text{in}}(v)$ and $\deg_{\text{out}}(v)$ the number of ingoing and outgoing edges, respectively.

Definition 3.1.15 (Boolean circuit)**.** Let F be a set of Boolean functions $f\colon \{0,1\}^{\text{ar}(f)} \to \{0,1\}$, where $\text{ar}(f)$ is the arity of f.

A *circuit* is a directed and acyclic graph C decorated with node labels ℓ such that:

(i) Input: There are n source nodes ($\deg_{\text{in}} = 0$), labeled with $x_1, \dots, x_n$.

(ii) Internal: Each non-source node v is labeled with a function f such that $\text{ar}(f) = \deg_{\text{in}}(v)$.

(iii) Output: There are m output nodes ($\deg_{\text{out}} = 0$), labeled (additionally) with $1, \dots, m$.

A circuit computes a Boolean function $f_C\colon \{0,1\}^n \to \{0,1\}^m$ in a natural way. The number of nodes is denoted by $\text{size}(C)$ and $\text{depth}(C)$ is the maximal length of a path from an input note to an output node.

Example 3.1.16. A common set of Boolean functions to use in a circuit are the bounded fan-in gates $\mathcal{B} = \{\texttt{AND}_2, \texttt{OR}_2, \texttt{NOT}_1\}$ (here the indices indicate the arity), or the unbounded fan-in gates $\mathcal{UB} = \{\texttt{NOT}_1\} \cup \{\, \texttt{AND}_n, \texttt{OR}_n \mid n \geq 2 \,\}$.

While both sets of boolean functions are complete (in the sense that every function can be composed from these building blocks), circuits using $\mathcal{B}$ are (at least heuristically) bigger in size and depth than those using $\mathcal{UB}$. For example, building $\texttt{AND}_n$ with $\mathcal{B}$ requires a circuit with $n-1$ internal nodes and which has depth $O(\log(n))$ using a balanced tree.

We also note the additional set $\mathcal{UBM} = \mathcal{UB} \cup \{\, \texttt{maj}_n \mid n \geq 2 \,\}$, where

$$\texttt{maj}_n(x_1, \dots, x_n) = \begin{cases} 1 & \text{if } \#\{\, i \mid x_i = 1 \,\} > \frac{n}{2} \\ 0 & \text{otherwise.} \end{cases}$$ ◁

Definition 3.1.17 (Uniform circuit family)**.** A family of boolean circuits $\mathcal{C} = \{C_n\}_{n \in \mathbb{N}}$ over F is a sequence of circuits each over F such that $\text{ar}(f_{C_n}) = n$. The family computes a function $f_{\mathcal{C}}\colon \{0,1\}^* \to \{0,1\}^*$ via $f_{\mathcal{C}}(x) = f_{C_{|x|}}(x)$.

If each circuit has exactly one output node, then f_C is a map to $\{0,1\}$ and $L(\mathcal{C}) \coloneqq f_{\mathcal{C}}^{-1}(1) \subseteq \{0,1\}^*$ is the set decided by C.

The family is *log-uniform* if on input n a representation of (C_n, ℓ_n) can be computed in working space $\log(n)$.

Definition 3.1.18 (Circuit complexity classes)**.** A function or decision problem A is in NC^k if it can be decided with a log-uniform family of Boolean circuits over $\mathcal{B}$, such that $\operatorname{depth}(C_n) = O(\log^k(n))$ and $\operatorname{size}(C_n) \in n^{O(1)}$.

Similarly, A is in AC^k if the same condition holds with circuits over $\mathcal{UB}$, and in TC^k with circuits over $\mathcal{UBM}$.

Remark 3.1.19. Technically, the uniform classes with $k \leq 1$ should be defined with a stricter resource bound ($\mathsf{DLOGTIME}$) on the Turing machine, since $\mathsf{NC}^1 \subseteq \mathsf{L}$. ◁

Example 3.1.20. Addition of two n-bit integers is in AC^0 using a carry lookahead adder. It is not in NC^0, since the leading digit depends on all input bits, while functions in NC^0 depend only on a bounded number of input bits.

Addition of n n-bit integers is in NC^1 (here, there are circuits for each bit in the output, and the circuits only needs to be specified for square integers). Using the elementary school multiplication method, multiplication of two n-bit numbers reduces to the addition of $n \leq 2n$-bit numbers, hence multiplication is also in NC^1. ◁

Example 3.1.21 (Unary subset sum is in TC^0)**.** We mentioned in Example 3.1.14 that UnboundedSubsetSum is NP-complete, even when restricting to the bounded version. However, this only holds true if the numbers are encoded in binary. Kane [Kan17] showed that (bounded) unary subset sum is in L, and even in TC^0. ◁

To relate the circuit complexity classes (for decision problems) to our previously defined ones, we have the following inclusions:

$$\cdots \subseteq \mathsf{NC}^i \subseteq \mathsf{AC}^i \subseteq \mathsf{TC}^i \subseteq \mathsf{NC}^{i+1} \subseteq \cdots$$

$$\mathsf{NC}^0 \subsetneq \mathsf{AC}^0 \subsetneq \mathsf{TC}^0 \subseteq \mathsf{NC}^1 \subseteq \mathsf{L} \subseteq \mathsf{NL} \subseteq \mathsf{AC}^1 \subseteq \cdots \subseteq \mathsf{NC} = \mathsf{NC}^{O(1)} \subseteq \mathsf{P} \subseteq \cdots$$

3.2 Background: The ideal world

In this section, we give an overview of the complexity results regarding ideal membership, representation degree, and Gröbner basis degrees. Let C be a class of polynomials in several variables, such as Homog, the class of all homogeneous polynomials, or Mon, the class of monomials. Here, the number of variables is not fixed and may depend on the input (unless, for example, $\mathsf{C} = \mathbb{K}[x_1, \ldots, x_n]$ for a particular n). The two central decision problems are the following:

IdealMem$_{\mathbb{K}}(\mathsf{C})$, IdealMem$_{\mathbb{K}}$	
Input:	$f_1, \ldots, f_s \in \mathsf{C}$ (or $\mathbb{K}[\underline{x}]$); $g \in \mathbb{K}[\underline{x}]$
Output:	Is $g \in \langle f_1, \ldots, f_s \rangle$?

AlgMem$_{\mathbb{K}}(\mathsf{C})$, AlgMem$_{\mathbb{K}}$	
Input:	$f_1, \ldots, f_s \in \mathsf{C}$ (or $\mathbb{K}[\underline{x}]$); $g \in \mathbb{K}[\underline{x}]$
Output:	Is $g \in \mathbb{K}[f_1, \ldots, f_s]$?

Our results on the complexity of AlgMem$_{\mathbb{K}}(\mathsf{C})$ are based on refined complexity bounds of IdealMem$_{\mathbb{K}}(\mathsf{C})$ and related problems, which is why we give a brief overview in this section. A more comprehensive discussion of the complexity of ideals can be found in [MT17].

3.2.1 Representing rings and polynomials

In the problems $\textsc{IdealMem}_{\mathbb{K}}$ and $\textsc{AlgMem}_{\mathbb{K}}$, the input consists of encoded polynomials over $\mathbb{K}$, for example as a string of characters. In order to allow for efficient calculations with polynomials, we require that the coefficient field itself supports efficient arithmetic as well. This has been formalized as the concept of a *well-endowed ring* [BCP83].

Informally, elements x of a well-endowed ring A of "length" $\alpha(x) \leq n$ are represented, possibly redundantly, by strings $\{0,1\}^{l(n)}$ of polynomial size. It is required that representations of elements of any length, addition, subtraction and multiplication can be efficiently computed.

Definition 3.2.1 (Well-endowed ring)**.** Let (A, α) be a ring with a *length function* $\alpha\colon A \to \mathbb{N}$, satisfying for $x, y \in A$

$$\begin{aligned} \alpha(x+y) &\leq \max\{\alpha(x), \alpha(y)\} + O(1) \\ \alpha(xy) &\leq \quad \alpha(x) + \alpha(y) \quad + O(\log \max\{\alpha(x), \alpha(y)\}). \end{aligned}$$

A *succinct representation* of (A, α) is a pair $(l\colon \mathbb{N} \to \mathbb{N}, r = (r_n\colon \{0,1\}^{l(n)} \to A)_n)$, $l(n) \in n^{O(1)}$, such that $\{\, x \in A \mid \alpha(x) \leq n \,\} \subseteq r_n(\{0,1\}^{l(n)})$.

The ring (A, α) is *well-endowed* if it has a uniform (details omitted) succinct representation such that a valid representation of $x + y$ in $\{0,1\}^*$ can be computed from a valid representation of x, y in NC^0, and similarly for $-x$ (in NC^0) and xy (in NC^1).

By slight abuse of notation, a field is well-endowed if it is the fraction field of a well-endowed integral domain.

For more details consider also [Rit12, Section I.3.3].

Example 3.2.2. Finite fields are trivially well-endowed (for any length function), as the underlying set is finite. Despite addition of binary numbers not being in NC^0 (Example 3.1.20), there is a different, more redundant representation of the ring $\mathbb{Z}$ which turns it into a well-endowed ring:

Let $p \geq 3$ be an integer, let $\alpha(x) = \lceil 1 + \log_p(|x| + 1) \rceil$; this is a length function for $\mathbb{Z}$. We represent an integer as a tuple $(a_0, \ldots, a_n) \in \{-p+1, \ldots, p-1\}^{n+1}$ via

$$r_n(a_0, \ldots, a_n) = a_0 + a_1 p + \cdots + a_n p^n.$$

Every integer $x \in \mathbb{Z}$ with $\alpha(x) \leq n$ can be represented in this way by $n+1$ digits, and each digit is represented using $\lceil \log_2(2p-1) \rceil$ bits. In this way, $(l(n) = \lceil \log_2(2p-1) \rceil n, (r_n)_n)$ is a succinct uniform representation. One can show that in this representation, addition can be performed in NC^0 (this uses $p \geq 3$) [Rit12, Lemma 3.17]. Since multiplication is repeated addition (Example 3.1.20), multiplication is in NC^1 and $\mathbb{Z}$ is a well-endowed ring. With our convention, this makes $\mathbb{Q}$ into a well-endowed field (but not necessarily into a well-endowed ring). ◁

Many rings used in computer algebra are well-endowed:

Theorem 3.2.3 ([BCP83])**.** *If A is well-endowed, then so are $A[x]$, A^k, $\mathrm{Mat}_n(A)$, and all finite-dimensional A-algebras. In particular, finite fields, algebraic integers in number fields, and polynomial and matrix rings over them are well-endowed.*

The main reason why we care about such a strict efficiency requirement on the coefficient ring is that well-endowed rings allow for very efficient linear algebra calculations over their fraction field:

Theorem 3.2.4 ([BCP83]). *Let $\mathbb{K}$ be a fraction field of a well-endowed domain. Let $M \in \mathbb{K}^{m\times n}$, $m \leq n$, be a matrix whose entries have q-bit numerators and denominators. Then the computation of its rank, adjoint, and (if $m = n$) determinant and characteristic polynomial is in* $\mathsf{SPACE}(\log^2(nq))$.

Now that we have a notion of efficiently computable coefficient ring $\mathbb{K}$, we return to the question on how to encode polynomials. Generally, one has to list its coefficients and exponent vectors, and hence the input length is bounded below by the number of terms, the number of variables occurring, and, depending on the encoding, the coefficients and the (logarithm of the) degree.

Definition 3.2.5. The encoding of a polynomial $f = \sum_{|\alpha|\leq d} c_\alpha x^\alpha$ is *dense* if it lists all monomials with their coefficients (α, c_α) up to its degree, and *sparse* if it only lists those terms with non-zero coeffient. The representation of a monomial $x^\alpha \mathrel{\hat{=}} (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ is in *binary* if each α_i is encoded in binary, and *unary* if it is encoded in unary.

Example 3.2.6. The polynomial $f = 42x_1^3x_3^6 - 3x_2 \in \mathbb{Q}[x_1, x_2, x_3]$ could be encoded with binary exponents as $((\mathtt{11}, \mathtt{0}, \mathtt{110}), \mathrm{enc}(42)), ((\mathtt{0}, \mathtt{1}, \mathtt{0}), \mathrm{enc}(-3))$, and with unary exponents as $((\mathtt{111}, \epsilon, \mathtt{111111}), \mathrm{enc}(42)), ((\epsilon, \mathtt{1}, \epsilon), \mathrm{enc}(-3))$ (ϵ representing the empty string). In dense encoding, a representation of f would have to list all $\binom{9+3}{3} = 220$ monomials up to degree 9 (almost all having coefficient 0), while a sparse encoding only has to list these two terms. ◁

All complexity results in this chapter will hold with respect to both dense and sparse encodings and binary or unary monomial representation (with the exception of Theorem 3.5.2), as long as we make the reasonable assumption that the field $\mathbb{K}$ is well-endowed. For this reason, we will not elaborate further on the encoding.

3.2.2 Complexity of Polynomial Ideals

The major result about the complexity of polynomial ideals is the EXPSPACE worst-case complexity of $\textsc{IdealMem}_{\mathbb{K}}$.

Definition 3.2.7 (Representation degree)**.** Let $f_1, \ldots, f_s, g \in \mathbb{K}[\underline{x}]$ with $g \in \langle f_1, \ldots, f_s\rangle_{\mathbb{K}[\underline{x}]}$. A *representation* of this ideal membership is a vector $h_1, \ldots, h_s \in R$ with $g = h_1f_1 + \cdots + h_sg_s$. The *representation degree* of $(f_1, \ldots, f_s; g)$ is the smallest possible $D = \max_i \deg h_i$ over all representations (h_i).

Theorem 3.2.8. *Let $\mathbb{K}$ be a well-endowed field.*

(i) (Mayr & Mayer [MM82], Mayr [May89]) *The problem $\textsc{IdealMem}_{\mathbb{K}}$ is complete for* EXPSPACE. *A representation of ideal membership can be computed in exponential working space.*

(ii) (Mayr [May97]) *The problem* $\textsc{IdealMem}_{\mathbb{K}}(\mathsf{Homog})$ *is complete for* PSPACE. *The general problem is also in* PSPACE *when bounding the number of variables.*

All hardness results already hold for ideals generated by binomials $x^\alpha - x^\beta$.

This is an instance where the phrase "compute in exponential working space" has to be taken carefully, as in Example 3.1.12. Here, the length of a representation (bit-length, number of terms, degree) may be doubly exponential in the input length, thus does not fit into the working memory.

Next, we note bounds on the representation degree, also in the special case of complete intersections.

Theorem 3.2.9. *Let* $g \in \langle f_1, \dots, f_s \rangle \subseteq \mathbb{K}[x_1, \dots, x_n]$, $d \coloneqq \max_i \deg f_i$, *and write* $g = \sum_{i=1}^s h_i f_i$ *with* $h_i \in K[\underline{x}]$ *of minimal representation degree* $D = \max_i \deg h_i$.

(i) (Hermann [Her26], Mayr & Meyer [MM82]) $D \leq \deg g + (ds)^{2^n}$.

(ii) (Mayr & Meyer [MM82]) *There exists a sequence of ideals* $I_n \subseteq \mathbb{K}[x_1, \dots, x_{O(n)}]$ *generated by* $O(n)$ *polynomials of degree* $d + O(1)$ *and* g *of degree* 1 *such that* $D \geq d^{2^n}$.

(iii) (Dickenstein, Fitchas, Giusti & Sessa [Dic+91]) *If* $f_1, \dots, f_s$ *define a complete intersection ideal, then* $D \leq \deg g + d^s$.

Both from a theoretical and computational point of view, effective degree bounds on *Gröbner bases* are also important. We recall the classical *Dubé bound* as well as a dimension-refined version

Theorem 3.2.10. *Let* $I \subseteq \mathbb{K}[x_1, \dots, x_n]$ *be an ideal generated by* $f_1, \dots, f_s$ *of degree* $d_i \coloneqq \deg f_i$, $d_1 \geq \cdots \geq d_s$, *and consider any monomial order.*

(i) (Dubé [Dub90]) *The degree of the reduced Gröbner basis* GB *of* I, *i. e. the largest degree of an element of* GB, *is bounded by*

$$\deg GB \leq 2\left(\frac{d_1^2}{2} + d_1\right)^{2^{n-1}}.$$

(ii) (Mayr & Ritscher [MR13]) *If* I *has dimension* r, *then the degree is bounded by*

$$\deg GB \leq 2\left(\frac{1}{2}(d_1 \cdots d_{n-r})^{2(n-r)} + d_1\right)^{2^r} \leq 2\left(\frac{1}{2}d_1^{2(n-r)^2} + d_1\right)^{2^r}.$$

These upper bounds are asymptotically sharp.

We also note a complexity result on polynomial normal form computations, whose proof idea will be used in the complexity analysis of subalgebra membership.

Theorem 3.2.11 ([KM96]). *The normal form* $\operatorname{nf}_\prec(g, \langle f_1, \dots, f_s\rangle)$ *can be computed in exponential space over a well-endowed field.*

Proof idea. Let $I \coloneqq \langle f_1, \dots, f_s\rangle_{\mathbb{K}[\underline{x}]}$ and $\prec \coloneqq \prec_{\text{lex}}$ the lexicographic order (for simplicity). Let $GB = GB(I)$ be the reduced Gröbner basis of I, then Kühnle & Mayr show

$$\deg \operatorname{nf}_\prec(g, I) \leq \deg(g) + (\deg(GB) + 1)^{n^2+1} \deg(g)^n.$$

Let D be an upper bound on the degree of the coefficients $h_i \in K[\underline{x}]$ in a representation $g - \operatorname{nf}_\prec(g, I) = \sum_{i=1}^s h_i f_i$, for example the Hermann bound from Theorem 3.2.9, then Kühnle & Mayr reduce the normal form calculation into a linear algebra problem of size $\operatorname{poly}(D) = D^{O(1)}$. Using efficient parallel algorithms for matrix rank and the parallel computation hypothesis (Theorem 3.2.4), this yields an algorithm in space $\operatorname{polylog}(D) = (\log D)^{O(1)}$. Using the Dubé bound (Theorem 3.2.10) for $\deg(GB)$ and the Hermann bound (Theorem 3.2.9) for D yields the mentioned exponential space algorithm for normal form calculation. For details see [KM96, Section 2] or [Rit12, Lemma III.6.3]. □

3.3 Upper bounds on subalgebra membership

In this section we provide various complexity upper bounds, that is, algorithms for variants of AlgMem$_\mathbb{K}$, which place these problems in PSPACE and EXPSPACE. Our complexity analysis of the membership problem for a subalgebra relies on a classical elimination approach using tag variables, attributed to Spear [SS88; Spe77]. We then study special classes of polynomials for which better complexity bounds hold.

3.3.1 Subalgebra membership using normal forms

Definition 3.3.1 (Subalgebra membership certification degree). Let $f_1, \dots, f_s, g \in \mathbb{K}[\underline{x}]$ with $g \in \mathbb{K}[f_1, \dots, f_s]$. A *certificate* of this subalgebra membership is a polynomial $p \in \mathbb{K}[t_1, \dots, t_s]$ with $g = p(f_1, \dots, f_s)$. The *certification degree* of $(f_1, \dots, f_s; g)$ is the smallest possible $\deg p$ over all certificates p.

Let $\underline{t} = \{t_1, \dots, t_s\}$ be additional variables, and consider an *elimination order* on $\mathbb{K}[\underline{x}, \underline{t}]$, i. e. a monomial order such that $x_i \succ t^\alpha$ for all x_i and $t^\alpha \in \mathbb{K}[\underline{t}]$. For example, use the lexicographical order $\prec_{\text{lex}}$ if the monomial involves x_i, and a fixed order on $\underline{t}$ otherwise.

Lemma 3.3.2 (Shannon & Sweedler [SS88]). *Let* $f_1, \dots, f_s, g \in \mathbb{K}[\underline{x}]$ *and define the ideal* $J \coloneqq \langle f_1 - t_1, \dots, f_s - t_s\rangle \subseteq \mathbb{K}[\underline{x}, \underline{t}]$, *then*

$$g \in \mathbb{K}[f_1, \dots, f_s] \qquad \textit{if and only if} \qquad p \coloneqq \operatorname{nf}_\prec(g, J) \in \mathbb{K}[\underline{t}].$$

If this is the case, then $g = p(f_1, \dots, f_s)$, *interpreting* p *as a polynomial in* $t_1, \dots, t_s$.

Proof. If $p = \mathrm{nf}_\prec(g, J) \in \mathbb{K}[\underline{t}]$, then write $g = p + \sum_{i=1}^{s} h_i(f_i - t_i)$. Since g does not depend on $\underline{t}$, replacing $t_i \mapsto f_i$ in this equation yields $g = p(f_1, \ldots, f_s)$. This shows the backwards direction and the last part.

Conversely, assume $g = P(f_1, \ldots, f_s)$ for $P \in \mathbb{K}[\underline{t}]$, then in the same way $P \in g + J$. By Lemma 0.2.5, p is the element with minimal support in $g + J$. Since $\prec$ eliminates $\underline{x}$ and P involves no $\underline{x}$, this means that $p \in \mathbb{K}[\underline{t}]$. □

Note that there are infinitely many distinct certificates P unless $f_1, \ldots, f_s$ are algebraically independent. However, the normal form yields a minimal certificate in the following sense:

Corollary 3.3.3. *If $g \in A$, then $\mathrm{nf}_\prec(g, J)$ is the unique certificate with minimal support (with respect to $\prec_\mathrm{p}$, in the sense of Lemma 0.2.5). In particular, if $\prec$ is degree-dominated when restricted to $\mathbb{K}[\underline{t}]$ (so $t^\alpha \prec t^\beta$ if $\deg t^\alpha < \deg t^\beta$), then $\mathrm{nf}_\prec(g, J)$ is a subalgebra membership certificate of lowest degree.*

Subalgebra membership can thus be reduced to the task of calculating the normal form of a polynomial against an ideal in a larger polynomial ring. This approach is well-known, for example as a fall-back method in the `Macaulay2` package `SubalgebraBases` [Bur+24]. Applying Kühnle & Mayr's exponential space algorithm for computing the normal form (Theorem 3.2.11) thus yields our first complexity upper bound.

Theorem 3.3.4. *For any well-endowed field $\mathbb{K}$, $\textsc{AlgMem}_\mathbb{K}$ is in* EXPSPACE. *Moreover, for any monomial order on $\underline{t}$, the minimal certificate $p \in \mathbb{K}[t_1, \ldots, t_s]$ such that $g = p(f_1, \ldots, f_s)$ can be output in exponential working space.*

Proof. Given $f_1, \ldots, f_s, g$, the algorithm computes the normal form $p := \mathrm{nf}_\prec(g, J) \in \mathbb{K}[\underline{x}, \underline{t}]$ using Kühnle & Mayr's algorithm [KM96]. If there is a nonzero term in p involving some x_i, then $g \notin \mathbb{K}[f_1, \ldots, f_s]$. Otherwise, p is a certificate of subalgebra membership by Lemma 3.3.2, which can be enumerated to the output tape using single exponential working space.

Note that it might not be possible to fit the normal form in exponential working space as mentioned in Example 3.1.12. Instead one has to enumerate the terms of p, which fit in exponential working space using and check for occurrences of x_i individually. □

3.3.2 Special classes of polynomials and a degree bound

Our proof of Theorem 3.3.4 did not use anything particular about the structure of the ideal J. However, since $J = \langle t_1 - f_1, \ldots, t_s - f_s \rangle$ is a compelete intersection ideal, we have more refined upper bounds available regarding representation and Gröbner basis degrees. We apply this to the case of a fixed number of variables n and to the case of a fixed subalgebra A.

Theorem 3.3.5. *For a fixed subalgebra $A = \mathbb{K}[f_1, \ldots, f_s] \subseteq \mathbb{K}[\underline{x}]$ over a well-endowed field, the membership problem is in* PSPACE *(with respect to the input length of g). In fact, for a* fixed *number of variables n, we have $\textsc{AlgMem}_\mathbb{K}(\mathbb{K}[x_1, \ldots, x_n]) \in$* PSPACE.

Proof. The algorithm is the same as in Theorem 3.3.4, only the complexity analysis is slightly more elaborate. The fixed algebra case is implied by the fixed variable case. Alternatively, normal form calculation with respect to a fixed ideal is possible in polynomial space, as remarked by Kühnle & Mayr [KM96, Section 6].

If only the number of variables $x_1, \dots, x_n$ is fixed, then more care has to be taken, as the computation takes place in the ring $\mathbb{K}[x_1, \dots, x_n, t_1, \dots, t_s]$, whose number of variables $n + s$ still depends on the input length. However, the ideal $J = \langle t_1 - f_1, \dots, t_s - f_s \rangle$ is a complete intersection of dimension n, hence using Theorem 3.2.10(ii) its Gröbner basis degree is bounded above by

$$\deg GB(J) \leq G := 2\left(\frac{1}{2}d^{2s^2} + d\right)^{2^n}.$$

Furthermore, the representation degree of $g - \operatorname{nf}_\prec(g, J)$ is bounded above by $D := G + d^s$ by Theorem 3.2.9(iii). Since n is fixed, we see that D is single exponential in the input length, hence polylog(D) is polynomial. This shows that the Meyer & Kühnle algorithm works in polynomial space for a bounded number of variables. □

Corollary 3.3.6. *If $g \in \mathbb{K}[f_1, \dots, f_s] \subseteq \mathbb{K}[\underline{x}]$, then there exists a certificate $p \in \mathbb{K}[t_1, \dots, t_s]$ with $g = p(f_1, \dots, f_s)$ of degree*

$$\deg p \leq \deg(g) + \left(\left(\tfrac{1}{2}d^{2s^2} + d\right)^{2^n} + 1\right)^{(n+s)^2+1} \deg(g)^{n+s}.$$

Proof. This is the degree bound on the normal form by Kühnle & Mayr evaluated for the ideal J (note that the ambient polynomial ring has $n + s$ variables). □

Remark 3.3.7. This degree can probably be improved by studying the derivation of the normal form degree bounds in this particular situation. We suspect that a "nicer" upper bound akin to the Hermann bound (Theorem 3.2.9) should hold. ◁

Of great interest is the case of subalgebras generated by homogeneous polynomials. In this case, A is graded compatibly with the standard grading of the ambient ring.

Theorem 3.3.8. *For any well-endowed field $\mathbb{K}$,* $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Homog})$ *is in* PSPACE.

Proof. As a consequence of the grading, we see that a certificate p of lowest degree (if it exists) has degree at most $\deg g$. Thus either $\deg \operatorname{nf}_\prec(g, J) > \deg g$, in which case $g \notin K[f_1, \dots, f_s]$, or $\deg \operatorname{nf}_\prec(g, J) \leq \deg g$ and $g \in \mathbb{K}[f_1, \dots, f_s]$ if and only if $\operatorname{nf}_\prec(g, J) \in K[\underline{t}]$. This leads to a variation of the algorithm of Theorem 3.3.4 except we only attempt to compute the normal form up to degree $\deg g$. The complexity analysis is analogous to Theorem 3.3.5, noting that here a representation degree bound is $D := \deg g + d^s$. □

Remark 3.3.9. Degree bounds on certificates for elimination problems have been studied before, for example by Galligo & Jelonek [GJ20]. Their results are not necessarily applicable here, as our ideal J has dimension n in a ring of $n + s$ variables, while the results of Galligo & Jelonek only apply to intersections of an ideal with $K[x_1, \dots, x_n]$ if the ideal has dimension at least $n + 1$. ◁

3.4 Lower bounds on subalgebra membership

In this section, we prove a matching lower bound for the subalgebra membership problem and its homogeneous variant. The hardness results for $\textsc{IdealMem}_{\mathbb{K}}$ were constructed by embedding the word problem for commutative semigroups and its homogeneous variant into binomial ideals [MM82; May97]. We continue this chain of reductions, and give a concrete example of subalgebra membership instances with complicated certificates.

3.4.1 The reduction $\textsc{IdealMem}_{\mathbb{K}} \leq_{\mathrm{m}}^{\mathsf{L}} \textsc{AlgMem}_{\mathbb{K}}$

We start by describing how to embed ideal membership into subalgebra membership.

Theorem 3.4.1. *Let $f_1, \dots, f_s, g \in \mathbb{K}[x_1, \dots, x_n]$ and let t be a new variable. The following are equivalent:*

(a) $g \in \langle f_1, \dots, f_s \rangle \subseteq \mathbb{K}[\underline{x}]$;

(b) $tg \in \mathbb{K}[tf_1, \dots, tf_s, x_1, \dots, x_n] \subseteq \mathbb{K}[t, x_1, \dots, x_n]$.

Proof. If $g = h_1 f_1 + \cdots + h_s f_s$ with $h_i \in \mathbb{K}[\underline{x}]$, then

$$tg = h_1(tf_1) + \cdots + h_s(tf_s),$$

which certifies that tg is in $\mathbb{K}[tf_1, \dots, tf_s, \underline{x}]$. Conversely, assume

$$tg = p(tf_1, \dots, tf_s, x_1, \dots, x_n), \qquad p \in R \coloneqq \mathbb{K}[u_1, \dots, u_s, x_1, \dots, x_n].$$

Declare a (non-standard) grading on R and $S \coloneqq \mathbb{K}[t, x_1, \dots, x_n]$ by assigning $\deg t = \deg u_j = 1$ and $\deg x_i = 0$, so

$$R_d \coloneqq \bigoplus_{\substack{\alpha \in \mathbb{N}^s \\ |\alpha| = d}} u_1^{\alpha_1} \cdots u_s^{\alpha_s} \mathbb{K}[x_1, \dots, x_n], \qquad S_d \coloneqq t^d \mathbb{K}[x_1, \dots, x_n].$$

The evaluation homomorphism $R \to S$ substituting $x_i \mapsto x_i$, $u_j \mapsto t_j f_j$ respects the grading. Since $tg \in S_1$ is homogeneous of degree 1, decomposing p into its graded components $p = \sum_d p_d$ ($p_d \in R_d$) reveals that

$$p_d(tf_1, \dots, tf_s, x_1, \dots, x_n) = \begin{cases} p(tf_1, \dots, x_n) = tg & \text{if } d = 1 \\ 0 & \text{if } d \neq 1. \end{cases}$$

Thus by replacing p by p_1, we may assume that every term in p involves exactly one u_j. Setting $t = 1$, we read

$$p(f_1, \dots, f_s, x_1, \dots, x_n) = g,$$

which, by the structure of $p \in R_1$, is a representation for $g \in \langle f_1, \dots, f_s \rangle$. □

Corollary 3.4.2. *Let $A \coloneqq \mathbb{K}[tf_1, \dots, tf_s, x_1, \dots, x_n]$.*

(i) *The representation degree of* $g \in \langle f_1, \dots, f_s \rangle$ *is one less than the certification degree of* $tg \in A$.

(ii) *The smallest possible number of non-zero terms of all* h_i *in a representation* $g = h_1 f_1 + \cdots + h_s f_s$ *equals the smallest possible number of terms of a certificate for* $tg \in A$.

Proof. The forward direction in the previous proof gives that the total degree and number of terms for representation of ideal membership are a lower bound for the respective values for the subalgebra (the total degree increases by 1 by attaching the u_j to the h_j). Conversely, the grading argument shows that if p is a certificate, then so is its graded component p_1, which has a lower or equal degree and number of terms. Finally, setting $t \mapsto 1$ in p_1 gives a representation for $g \in \langle f_1, \dots, f_s \rangle$, showing the desired equalities. □

With this result we can deduce matching complexity lower bounds for subalgebra membership and its homogeneous variant. We can rephrase Theorem 3.4.1 as a complexity-theoretic reduction.

Theorem 3.4.3. *The map*

$$(f_1, \dots, f_s; g \in \mathbb{K}[\underline{x}]) \;\mapsto\; (tf_1, \dots, tf_s, x_1, \dots, x_n; tg \in \mathbb{K}[t, \underline{x}])$$

provides a reduction $\textsc{IdealMem}_{\mathbb{K}} \leq_{\mathrm{m}}^{\mathsf{L}} \textsc{AlgMem}_{\mathbb{K}}$. *More generally, it provides a reduction* $\textsc{IdealMem}_{\mathbb{K}}(\mathsf{C}) \leq_{\mathrm{m}}^{\mathsf{L}} \textsc{AlgMem}_{\mathbb{K}}(\mathsf{C})$ *for any class of polynomials* C *containing single variables and closed under multiplication with variables.*

3.4.2 Hard instances for subalgebra membership

With this we can complete the proof of the main results of this chapter regarding completeness. Let Binom be the class of polynomials which are monomials or binomials. The circle of complexity-theoretic reductions used is displayed in Figure 3.3.

Theorem 3.4.4. *Let* $\mathbb{K}$ *be a well-endowed field.* $\textsc{AlgMem}_{\mathbb{K}}$ *is complete for* $\mathsf{EXPSPACE}$ *and* $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Homog})$ *is complete for* PSPACE. *The same results hold for* Binom *and* $\mathsf{Binom} \cap \mathsf{Homog}$.

Proof. The complexity-theoretic upper bounds have been established in Theorem 3.3.4 and Theorem 3.3.8. The lower bounds now follow from the hardness of $\textsc{IdealMem}_{\mathbb{K}}$ and its homogeneous variant (Theorem 3.2.8) in conjunction with the reduction in Theorem 3.4.3. The hardness result only uses binomial ideals, hence holds for Binom already. □

As mentioned in the beginning of this section, the complexity lower bounds for ideal membership were achieved by studying the word problem for commutative semigroups. We recall part of this construction to indicate what our hard $\textsc{AlgMem}_{\mathbb{K}}$ instances may look like.

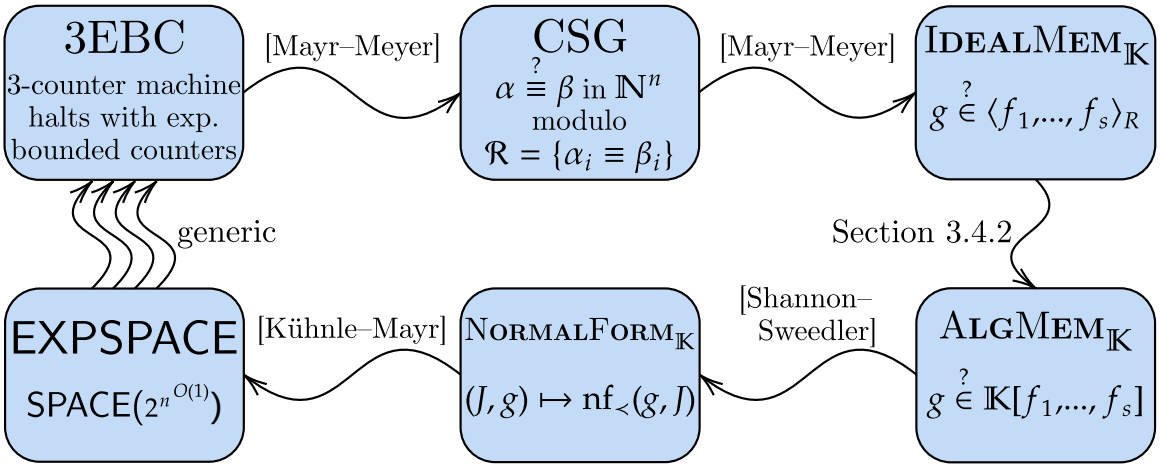


Figure 3.3: A circle of reductions connecting AlgMem$_{\mathbb{K}}$ with the class EXPSPACE.

Let $\mathcal{R} = \{(x^{\alpha_i}, x^{\beta_i})\}_{i=1}^{s}$ be a set of pairs of monomials, called *replacement rules*. Let $\equiv_{\mathcal{R}}$ be the congruence relation on $\mathsf{Mon}(\underline{x})$ generated by

$$x^{\gamma}x^{\alpha} \equiv_{\mathcal{R}} x^{\gamma}x^{\beta} \qquad (x^{\alpha}, x^{\beta}) \in \mathcal{R},\ x^{\gamma} \in \mathsf{Mon}(\underline{x}).$$

Concretely, two monomials m, m' are equivalent if one can be turned into the other by repeatedly replacing a factor $x^{\alpha} \mid m$ by the corresponding x^{β} or by replacing $x^{\beta} \mid m$ by x^{α}.

Word problem for Commutative Semigroups, CSG	
Input:	Replacement rules $\mathcal{R} = \{(x^{\alpha_i}, x^{\beta_i})\}_{i=1}^{s}$, Monomials $m, m' \in \mathsf{Mon}(\underline{x})$
Output:	Is $m \equiv_{\mathcal{R}} m'$?

Mayr & Meyer [MM82] prove EXPSPACE-hardness of CSG and provide a reduction CSG $\leq_{\mathrm{m}}^{\mathsf{L}}$ IdealMem$_{\mathbb{K}}$(Binom) via the map

$$(\mathcal{R}, m, m') \ \mapsto\ (x^{\alpha_1} - x^{\beta_1}, \dots, x^{\alpha_s} - x^{\beta_s}; m - m').$$

More precisely, the minimal number of terms in a representation for this ideal membership problem equals the minimal number of replacement steps necessary in showing the congruence $m \equiv_{\mathcal{R}} m'$. Using the known double-exponential examples from [MM82], we can prove worst-case certification degree lower bounds for subalgebra membership. For an accessible exposition of the double-exponential lower bounds compare [BS88] or the author's Master's thesis [Kay22].

Theorem 3.4.5. *There exist polynomials $f_1, \dots, f_{s(k)}, g \in \mathbb{K}[x_1, \dots, x_{n(k)}]$ of degree ≤ 6, $n = O(k)$, $s = O(k)$, such that $g \in \mathbb{K}[f_1, \dots, f_s]$, but every certificate $p \in \mathbb{K}[t_1, \dots, t_s]$ has degree at least 2^{2^k} and at least 2^{2^k} terms. These f_i and g can be chosen as binomials and monomials.*

Proof. The CSG instances from [MM82] lead to ideal membership of $O(k)$ binomials in $n = O(k)$ variables of degree ≤ 5 (to obtain this exact degree bound, use the slightly improved degree bound from [BS88]). In these such that every representation has degree and number of terms bounded below by 2^{2^k}. Applying the reduction from Theorem 3.4.3, we obtain an instance of $\textsc{AlgMem}_{\mathbb{K}}$, whose smallest certificate p has the same number of terms and greater total degree by Corollary 3.4.2. □

Remark 3.4.6. Following [BS88], this construction can be easily adapted to prove lower bounds of the form d^{2^k} at the expense of increasing the algebra generator degree to $d + O(1)$. This asymptotically matches the double-exponential upper bounds from Corollary 3.3.6. It would be interesting to get more refined lower bounds in terms of other properties of the algebra. For example, what could analogous statements of Theorem 3.2.9 and Theorem 3.2.10 be? ◁

Next, we look at homogeneous algebras, where the certification degree can not blow up. However, we can still produce $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Homog})$ instances that require exponentially many terms in their certificates.

Theorem 3.4.7. *There exist a subalgebra $A = \mathbb{K}[f_1, \dots, f_{5n}] \subseteq \mathbb{K}[x_1, \dots, x_{3n+3}]$ generated by single variables and homogeneous binomials of degree ≤ 3, and a binomial g of degree $n+2$ with the following property: $g \in A$, but every polynomial $p \in \mathbb{K}[\underline{t}]$ with $g = p(f_1, \dots, f_{5n})$ has at least 2^{n+1} terms.*

Proof. The construction is based on the binary-counting LBA from Example 3.1.6. In [May97, Theorem 17] Mayr describes how to reduce LBAHALT to commutative semigroups/ideals, proving PSPACE-hardness of $\textsc{IdealMem}_{\mathbb{K}}(\mathsf{Homog})$. The configuration $(q, i, b_0, \dots, b_{n+1})$ of the LBA is represented by the monomial $qh_i x_{0,\triangleright} x_{1,b_1} \cdots x_{n,b_n} x_{n+1,\triangleleft}$ and the replacement rules/binomial generators mimic the transition function δ. Applying our reduction (Theorem 3.4.3), we obtain a binary-counting subalgebra.

We can simplify the resulting algebra significantly: We can remove redundant transitions (for example, we don't actually need $\triangleright, \triangleleft$ since the head position variable knows the current position), and the new variable t introduced in Theorem 3.4.1 can also be omitted, as the combined variables qh_i have a similar effect. This leads to the following subalgebra:

$$\begin{aligned}
\mathcal{T} = \{q_0, q_1\} \dot\cup \{h_0, \dots, h_n\}, \quad \underline{x} &= \{x_{1,0}, x_{1,1}, \dots, x_{n,0}, x_{n,1}\}, \\
\mathcal{R} &\coloneqq \{\, q_0 h_i x_{i,0} - q_1 h_{i-1} x_{i,1} \mid 1 \leq i \leq n \,\} \\
&\quad \cup \{\, q_0 h_i x_{i,1} - q_0 h_{i+1} x_{i,0} \mid 1 \leq i \leq n-1 \,\} \\
&\quad \cup \{\, q_1 h_i x_{i,0} - q_1 h_{i-1} x_{i,0} \mid 1 \leq i \leq n \,\} \\
&\quad \cup \{\, q_1 h_0 - q_0 h_1 \,\}, \\
A &\coloneqq \mathbb{K}[f_1, \dots, f_{5n}] \coloneqq \mathbb{K}[\mathcal{R} \cup \underline{x}] \subseteq \mathbb{K}[\mathcal{T} \cup \underline{x}], \\
g &\coloneqq q_0 h_1 x_{1,0} \cdots x_{n,0} - q_0 h_n x_{1,0} \cdots x_{n-1,0} x_{n,1}.
\end{aligned}$$

Then $g \in A$ by construction, since the binary counter will go from $(q_0, 1, \triangleright 0 \dots 0 \triangleleft)$ to $(q_0, 1, \triangleright 1 \dots 1 \triangleleft)$ and then walk to the right, erasing the 1's until it reaches the configuration $(q_0, n, \triangleright 0 \dots 01 \triangleleft)$. On the way it writes every number $0, \dots, 2^{n-1}$ on the tape, taking at

least 2 steps each time, for a total of $\geq 2^{n+1}$ steps. By Corollary 3.4.2 any certificate for $g \in A$ must essentially contain this derivation, hence has at least 2^{n+1} terms. □

An implementation of the homogeneous binary-counting subalgebra in `Macaulay2` can be found at `mathrepo.mis.mpg.de/ComplexityOfSubalgebras`.

3.5 Complexity of initial algebras

In this final section, we consider the complexity of monomial algebra membership and pose related questions on the complexity of SAGBI bases. We outline an interesting connection between infinitely generated initial algebras and the theory of semilinear sets.

3.5.1 Monomial subalgebras

Monomial subalgebras A are $\mathbb{N}^n$-graded in the sense that

$$A = \bigoplus_{\substack{\alpha \in \mathbb{N}^n \\ x^\alpha \in A}} \mathbb{K} x^\alpha \subseteq \mathbb{K}[\underline{x}].$$

This has the useful consequence that a polynomial $\sum_\alpha c_\alpha x^\alpha$ is in A if and only if every monomial x^α, $c_\alpha \neq 0$, is in A. Furthermore, monomial algebras are related to linear programming and linear Diophantine equations [RS90, Remark 1.9], as

$$x^\beta \in K[x^{\alpha_1}, \ldots, x^{\alpha_s}] \quad \iff \quad \exists c \in \mathbb{N}^s \text{ s.t. } \beta = \sum_{i=1}^{s} c_i \alpha_s. \tag{3.1}$$

Theorem 3.5.1. *For any $\mathbb{K}$ the problem $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Mon})$ is NP-complete. This is true even when restricting to square-free monomials.*

Proof. For NP membership one immediately reduces to the case where the polynomial to test f is a monomial due to the $\mathbb{N}^n$-grading. In eq. (3.1) we have $c_j \leq \max_i \beta_i$, so the bit length of c_j is bounded by the bit length of β. Hence, non-deterministically guessing c yields a NP-algorithm.

For NP-hardness, one can reduce from the NP-complete problem $\textsc{1in3Sat}$ (Example 3.1.2). Indeed, given sets $S_1, \ldots, S_n \subseteq \{1, \ldots, s\}$, then let $\alpha_1, \ldots, \alpha_s \in \{0,1\}^n$ be the integer vectors with $(\alpha_i)_j = 1$ when $i \in S_j$. Set $\beta = (1, \ldots, 1) \in \mathbb{N}^n$. Then eq. (3.1) encodes exactly the question for $\textsc{1in3Sat}$, a solution corresponding to $T = \{\, i \mid c_i = 1 \,\}$. In this construction, all monomials x^{α_i}, x^β are square-free. □

We see that $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Mon})$ is NP-complete even for polynomials of bounded degree. The same is true if we instead bound the number of variables $\mathsf{Mon}_n := \mathsf{Mon}(x_1, \ldots, x_n)$, as long as the exponents are encoded in binary.

Theorem 3.5.2. *The problem $\textsc{AlgMem}_{\mathbb{K}}(\mathsf{Mon}_n)$ for fixed $n \geq 1$ is NP-complete for binary exponent encoding and in TC^0 with unary encoding.*

Proof. Encoding the exponents as binary, the unary case $n = 1$ is a direct translation of the UnboundedSubsetSum problem, which is NP-hard.

On the other hand, if the monomials are encoded in unary, then a generating-function approach as in [Kan17] provides a family of circuits in TC^0 deciding $\textsc{AlgMem}_K(\mathsf{Mon}_n)$. In more detail, Kane describes TC^0 circuits deciding the unary vector-valued subset sum problem, which corresponds to solutions of Equation (3.1) with $c_i \in \{0, 1\}$. We can reduce the case of arbitrary $c_i \in \mathbb{N}$ to the $\{0, 1\}$ case by replacing every generator x^{α_i} by $x^{\alpha_i}, x^{2\alpha_i}, x^{4\alpha_i}, \ldots, x^{2^\kappa \alpha_i}$, $\kappa = \lfloor \log_2 |\beta| \rfloor$. Since the numbers are encoded in unary, the input size is proportional to $C = |\beta| + \sum_{i=1}^{s} |\alpha_i|$, and this larger generating set has size $O(C^2)$.

Alternatively, one can apply a TC^0-variant of Courcelle's theorem to obtain TC^0-membership, compare [EJT12, Theorem 13]. □

Remark 3.5.3. The NP-hardness results from Theorems 3.5.1 and 3.5.2 are in stark contrast to the analogous results for monomial *ideals*: Monomial ideal membership is computationally trivial, as $x^\beta \in \langle x^{\alpha_1}, \ldots, x^{\alpha_s} \rangle$ if and only if component-wise $\beta \geq \alpha_i$ for some i. ◁

3.5.2 Finiteness of SAGBI bases

We return to a general subalgebra $A = \mathbb{K}[f_1, \ldots, f_s] \subseteq \mathbb{K}[\underline{x}]$ equipped with a monomial order $\prec$. In Section 0.2 we defined the concept of SAGBI bases $S \subseteq A$ as subsets with $\mathbb{K}[\{ \operatorname{in}_\prec f \mid f \in S \}] = \operatorname{in}_\prec A$.

The previous results in this section provide a modest complexity lower bound of deciding subalgebra membership in the presence of SAGBI bases. The author is not aware of an analogous result for ideal membership given a Gröbner basis.

Corollary 3.5.4. *The problem* $\textsc{AlgMem}_K$ *is* NP*-hard, even if the input polynomials* $f_1, \ldots, f_s$ *form a finite SAGBI basis of* $\mathbb{K}[f_1, \ldots, f_s]$.

Proof. Subalgebras generated by a finite set S of monomials have S as a SAGBI basis. Hence Theorem 3.5.1 provides the NP lower bound. □

A major difference to the theory of Gröbner bases is that not every finitely generated subalgebra has a *finite* SAGBI basis.

Example 3.5.5. The subalgebra $\mathbb{K}[x, xy - y^2, xy^2] \subseteq \mathbb{K}[x_1, x_2]$ has the infinitely generated initial algebra $\mathbb{K}[\{ x_1 x_2^k \mid k \geq 0 \}]$ for any monomial order with $x_1 \succ x_2$ [RS90, Example 4.11]. ◁

We therefore propose the study of the following two decision problems related to the initial algebra.

SAGBIFINITE$_{\mathbb{K}}$(C), SAGBIFINITE$_{\mathbb{K}}$	
Input:	$f_1, \dots, f_s \in \mathsf{C}$ (or $\mathbb{K}[\underline{x}]$)
Output:	Does $\mathbb{K}[f_1, \dots, f_s]$ have a finite SAGBI basis?

INALGMEM$_{\mathbb{K}}$(C), INALGMEM$_{\mathbb{K}}$	
Input:	$f_1, \dots, f_s \in \mathsf{C}$ (or $\mathbb{K}[\underline{x}]$), $x^\alpha \in \mathsf{Mon}(\underline{x})$
Output:	Is $x^\alpha \in \mathrm{in}_\prec \mathbb{K}[f_1, \dots, f_s]$?

Robbiano & Sweedler showed that an algorithm somewhat analogous to Buchberger's algorithm for Gröbner bases can be used to enumerate a SAGBI basis, which will terminate if (and only if) A has a *finite* SAGBI basis. A procedure that returns `yes` if the output is yes, but never terminates if the output is no, is called a *semi-algorithm*, and a problem is *semi-decidable* or *recursively enumerable* if there is a semi-algorithm "solving" it.

Proposition 3.5.6 (Robbiano & Sweedler [RS90])**.** INALGMEM$_{\mathbb{K}}$ *and* SAGBIFINITE$_{\mathbb{K}}$ *are semi-decidable over a computable field* $\mathbb{K}$*.*

We are not aware of any general better complexity bounds, or even just if these problems are computable at all (though a negative answer would be quite surprising). A first, modest complexity result in this direction is:

Theorem 3.5.7. INALGMEM$_{\mathbb{K}}$(Homog) *is* PSPACE*-complete.*

Proof. For membership, use the homogeneity to enumerate the initial algebra in the required degree. For hardness, use the same reduction as before. □

3.5.3 The semilinearity conjecture

We have seen in the previous section that the initial algebra of a finitely generated algebra can be infinitely generated. Despite this, the subalgebra from Example 3.5.5 still has a relatively simple structure, as displayed in Figure 3.4. In this final section, we aim to provide a finite combinatorial description for some initial algebras which are themselves infinitely generated.

The underlying initial monoid $M \subseteq \mathbb{N}^2$ is the union of the sets $e_1 + \mathbb{N}^2$ and $\{0\}$. This is an example of a semilinear set [GS66].

Definition 3.5.8. Let $X \subseteq \mathbb{Z}^n$ be a subset.

(i) X is *affine-linear* if there exist $b_1, \dots, b_m, c \in \mathbb{Z}^n$ such that $X = c + \langle b_1, \dots, b_m \rangle_{\mathbb{N}}$.

(ii) X is *semilinear* if $X = X_1 \cup \cdots \cup X_k$ with X_i affine-linear.

We say that a monomial algebra is semilinear if its monoid of exponents $\subseteq \mathbb{Z}^n$ is semilinear.

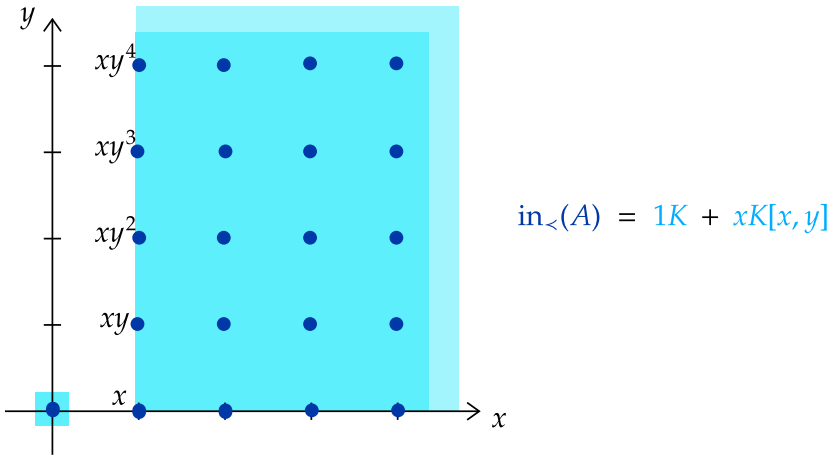


Figure 3.4: $\mathrm{in}_\prec \mathbb{K}[x, xy - y^2, xy^2]$ is the union of two affine-linear subsets in $\mathbb{N}^2$.

Example 3.5.9. A submonoid $M \subseteq \mathbb{Z}^n$ is finitely generated (an *affine semigroup*) if and only if it is affine-linear as a set. Figure 3.4 shows that the class of semilinear monoids is strictly larger than that of affine semigroups. ◁

Looking at basic examples known in the literature, one may suspect that all initial algebras are semilinear. This turns out to be false.

Theorem 3.5.10. *There exist finitely generated subalgebras whose initial algebra is not semilinear. In fact, the invariant ring $\mathbb{K}[x, y, z]^{\mathfrak{A}_3}$ has this property for "most" term orders.*

Here, $\mathfrak{S}_n \supset \mathfrak{A}_n$ are the symmetric and alternating group on n letters, respectively.

Proof. It is known that for any term order $\prec$ and any subgroup $G < \mathfrak{S}_n$ that is not a product $\mathfrak{S}(N_1) \times \cdots \times \mathfrak{S}(N_k)$ $(N_1 \dot\cup \cdots \dot\cup N_k = \{1, \ldots, n\})$, the initial algebra of $\mathbb{K}[x_1, \ldots, x_n]^G$ is not finitely generated. Moreover, varying $\prec$ over all real weight orders (even restricting to monomial orders), there are uncountably many distinct initial subalgebras [Kur02]. The smallest example in this family is $\mathbb{K}[x, y, z]^{\mathfrak{A}_3}$.

On the other hand, the set of semilinear sets is countable, hence for uncountably many orders $\prec$, the initial algebra cannot be semilinear. □

This proof is highly non-constructive, and we will see in a moment that for reasonable term orders the initial monoid of $\mathbb{K}[x, y, z]^{\mathfrak{A}_3}$ is indeed semilinear. Here is one way of defining a "resonable term order".

Definition 3.5.11. A toral order $\prec$ on $\mathsf{Mon}(x_1^{\pm 1}, \ldots, x_n^{\pm 1})$ is a *rational weight order* if there exist $A \in \mathbb{Q}^{m \times n}$ such that $\alpha \prec \beta$ if and only if $A\alpha \leq_{\mathrm{lex}} A\beta$, $\alpha, \beta \in \mathbb{N}^n$.

Recall that every monomial order is represented by a real weight matrix, but this includes very wild term orders (most of them being not computable!). All orders commonly encountered (lexicographic, degree (reverse) lexicographic, block orders, ...) are rational weight orders.

Conjecture 3.5.12 (Semilinearity conjecture). Let $\prec$ be a rational weight order and $A \subseteq \mathbb{K}[x_1, \dots, x_n]$ a finitely generated subalgebra. Is $\mathrm{in}_\prec(A)$ semilinear? ⍰

An interesting class of infinitely generated initial algebras that adhere to this conjecture is given by the following theorem.

Theorem 3.5.13. *Let $G \subset \mathrm{GL}(\mathbb{Z}^n)$ be a finite subgroup. Let $M \subset \mathbb{Z}^n$ be an affine semigroup with $gM = M$ for all $g \in G$. Let $\mathbb{K}[M]^G$ be the invariant subalgebra and let $\prec$ be a rational monomial order. Then $\mathrm{in}_\prec \mathbb{K}[M]^G$ is semilinear.*

Before proving this theorem, we highlight the connection of semilinear sets with Presburger arithmetic.

Definition 3.5.14 (Presburger formula)**.** A Presburger formula in the variables X is an expression built from the following rules:

- Atoms: $a_1 + \cdots + a_m < b_1 + \cdots + b_n$ is a formula, where each a_i, b_j is either in X or a constant in $\mathbb{N}$.
- Connectives: Logical connections of formulas via $\vee, \wedge, \neg$ are formulas.
- Quantifiers: If ϕ is a formula and $x \in X$, then $\exists x\, \phi$ and $\forall x\, \phi$ is a formula.

If ϕ is a formula and $x_1, \dots, x_n$ are the variables in ϕ *not* bound to a $\exists$- or $\forall$-quantifier, then $x_1, \dots, x_n$ are the *free variables*; we write $\phi(x_1, \dots, x_n)$ to emphasize this.

We can evaluate Presburger formulas with n free variables in $\mathbb{Z}^n$. It turns out that solution sets $L(\phi) \coloneqq \{\, v \in \mathbb{Z}^n \mid \phi(v_1, \dots, v_n) \,\}$ of such formulas describe semilinear sets [GS66], for a different proof see [Kra03, Section 2.5].

Proposition 3.5.15 (Ginsburg–Spanier)**.** *For a set $S \subseteq \mathbb{Z}^n$ the following are equivalent:*

(a) *S is semilinear;*

(b) *There is a Presburger formula $\phi(x_1, \dots, x_n)$ with $S = L(\phi)$;*

(c) *There is a quantifier-free Presburger formula $\phi'(x_1, \dots, x_n)$ with modulo constraints (also allow $\sum_i a_i \equiv_m \sum_j b_j$ for $m \in \mathbb{N}$ as atoms) with $S = L(\phi')$.*

This proposition provides some motivation for the semilinearity conjecture: Presburger arithmetic is a weak model of computation/logic which is still strong enough to express (some aspects) of rational weight orders, for example:

Lemma 3.5.16. *With notation as in Theorem 3.5.13, for $x^u \in \mathsf{Mon}(x_1, \dots, x_n)$ the following are equivalent:*

(a) *$x^u \in \mathrm{in}_\prec \mathbb{K}[M]^G$;*

(b) *$u \in M$ and x^u is the largest monomial in its G-orbit $\mathcal{O}_u \coloneqq \{\, x^{g \cdot u} \mid g \in G \,\}$.*

Proof. If $x^u \in M$ is the largest monomial in its orbit $\mathcal{O}_u$, then x^u is the leading monomial of $f = \sum_{x^v \in \mathcal{O}_u} x^v \in \mathbb{K}[M]$. By construction, f is G-invariant, hence $x^u = \mathrm{in}_\prec f$ is in $\mathrm{in}_\prec \mathbb{K}[M]^G$.

Conversely, if $x^u \in \mathrm{in}_\prec \mathbb{K}[M]^G$, then $x^u = \mathrm{in}_\prec f$ for some $f \in \mathbb{K}[M]^G$. Since f is G-invariant, all monomials from $\mathcal{O}_u$ must be terms in f, and hence smaller than x^u itself. □

Proof of Theorem 3.5.13. By Proposition 3.5.15 and Lemma 3.5.16 it suffices to show that the condition (b) of the previous lemma can be expressed in Presburger arithmetic.

If $M = \langle m_1, \dots, m_s \rangle_\mathbb{N}$, $m_i = (v_i^j)_{j=1}^n - (w_i^j)_{j=1}^n$, $v_i, w_i \in \mathbb{N}^n$, then the condition $u \in M$ can be expressed as

$$\phi_1(u_1, \dots, u_n) = \exists y_1 \dots \exists y_m \bigwedge_{j=1}^{n} (v_1^j \cdot y_1 + \dots + v_s^j \cdot y_s = w_1^j \cdot y_1 + \dots + w_s^j \cdot y_s + u_j).$$

Here $a \cdot y$ for $a \in \mathbb{N}$ is understood as $\underbrace{y + \dots + y}_{a \text{ times}}$.

The condition of being maximal in its orbit can be expressed as follows: Let A be a weight matrix without loss of generality with integer entries. Then the following Presburger formula describes $\{ (u, v) \in \mathbb{Z}^n \times \mathbb{Z}^n \mid u \prec_A v \}$:

$$\phi_\prec(u_1, \dots, u_n, v_1, \dots, v_n) = \bigvee_{i=1}^{m} \left(\left(\bigwedge_{j=1}^{i-1} (a_j \cdot u = a_j \cdot v) \right) \wedge (a_i \cdot u < a_i \cdot v) \right).$$

Next, note that if $\phi(x_1, \dots, x_n)$ is a Presburger formula and $\ell_i(y) = \sum_{j=1}^m a_{ij} y_j$ are linear forms with $a_{ij} \in \mathbb{Z}$, then $\phi(\ell_1(y), \dots, \ell_n(y))$ can be transformed into a valid Presburger formula by moving any negative coefficients to the other side of the respective (in)equality.

In this way we can describe a formula that describes $\{ u \mid u = \max_\prec \mathcal{O}_u \}$:

$$\phi_2(u_1, \dots, u_n) = \bigwedge_{g \in G} ((gu = u) \vee \phi_\prec(gu, u)).$$

Then the initial monoid is described by the Presburger formula $(\phi_1 \wedge \phi_2)(u_1, \dots, u_n)$. □

The proof of Theorem 3.5.13 is constructive in the sense that it gives an explicit Presburger formula that defines the initial algebra. One could feed this into a computer program that computes a semilinear description of the monoid.

Remark 3.5.17. Conjecture 3.5.12 asks whether the initial monoid of every finitely generated algebra is semilinear (for rational weight orders). It would also be interesting to study which semilinear monoids appear as initial monoids of finitely generated algebras, even if the conjecture is false in general. ◁

Chapter 4

Logarithmic Discriminants of Hyperplane Arrangements

Contents

> Discrimination is not done by villains.
> It's done by us.
>
> VIVIENNE MING [And16]

A recurring task in particle physics and algebraic statistics is to compute the complex critical points of a product of powers of affine-linear functions $\mathcal{L}_u(x) = \ell_0^{u_0} \cdots \ell_n^{u_n}$. The logarithmic discriminant characterizes exponents u for which such a function has a degenerate critical point in the corresponding hyperplane arrangement complement $X = \mathbb{C}^d \setminus \mathbb{V}(\ell_0 \cdots \ell_n)$. We study properties of this discriminant, exploiting its connection with the Hurwitz form of a reciprocal linear space. Our main result can be summarized as follows:

Main Results 4 (4.3.9, 4.4.12, 4.5.6, 4.5.10, 4.5.11)**.** *Let $\ell_0, \dots, \ell_n \in \mathbb{C}[x_1, \dots, x_d]_{\leq 1}$ be general linear forms. If $n = d$, then the logarithmic discriminant variety is empty. If $n \geq d+1$, then $\nabla_{\log}(X) \subseteq \mathbb{P}^n$ is an irreducible hypersurface, cut out by a homogeneous polynomial of degree $2d\binom{n-1}{d}$. If the ℓ_i are real, then this polynomial can be chosen to take positive values on $\mathbb{R}^{n+1}_+$.*

This chapter is based on joint work with Andreas Kretschmer and Simon Telen [KKT25]. It is structured as follows: In Section 4.1 we introduce the theoretical framework of branch loci and look at various classical discriminants, including the Hurwitz form. In Section 4.2 we formally define the logarithmic discriminant of a very affine variety in the context of likelihood geometry. In Section 4.3, we introduce the main objects of study in this chapter, hyperplane arrangement complements and their logarithmic discriminants. In Section 4.4, we draw a connection of the logarithmic discriminant to the Hurwitz form of the reciprocal linear space associated with the arrangement. As a first application, we show that in many cases the logarithmic discriminant polynomial of a real arrangement is positive on the positive orthant. In Section 4.5, we consider generic arrangements and finally prove Main Results 4. More specifically, we first give a full picture for arrangements of points on the line ($d = 1$) in Section 4.5.1, and then consider irreducibility and degree of the logarithmic discriminant in Sections 4.5.2 and 4.5.3. In the final Section 4.6, we present a case study where $X = \mathcal{M}_{0,m}$, which is related to CHY theory for particle scattering.

4.1 Background: Discriminants

Informally, discriminants measure where a genericity statement in algebraic geometry breaks down. For example, among univariate polynomials $f \in \mathbb{P}(\mathbb{C}[z]_{\leq n}) = \mathbb{P}^n$, almost all f have n distinct roots in $\mathbb{C}$ (or $\mathbb{P}^1$). The (classical) discriminant is the closed subset $\nabla \subseteq \mathbb{P}^n$ of polynomials with a multiple root. In this section we introduce discriminants in the context of branch loci of morphisms and highlight several important examples. We also introduce the Hurwitz form, the discriminant of the genericity statement "a

projective variety $X \subseteq \mathbb{P}^n$ intersects a general complementary linear subspace in $\deg X$ reduced points". The Hurwitz form will be of particular importance in our study of the logarithmic discriminant.

4.1.1 Ramification and branch loci

We begin with a general definition of ramification and branch locus, which we then unravel in the special situation that we are interested in. A reference is [EH06, Chapter V].

Intuitively, the ramification locus of a morphism $f\colon X \to Y$ consists of the points in X that are singular in their (scheme-theoretic) fiber. This can be detected using the Jacobi criterion of smoothness, requiring that the Jacobi matrix at that point drops rank. The sheaf of relative differentials is locally the cokernel of the transpose of the Jacobi matrix (Section 0.1.4), so its support $\operatorname{supp}\Omega^1_{X/Y} = \{\, x \in X \mid (\Omega^1_{X/Y})_x \neq 0 \,\}$ coincides with the ramification locus. To give this closed subset of X a scheme structure in a global and functorial way, we use the Fitting ideal of $\Omega^1_{X/Y}$ to define the scheme structure.

If M is a finitely presented module over a ring R, then the k-th Fitting ideal $\operatorname{Fit}_k(M) \subseteq R$ is the ideal of $(n-k)$-minors of any presentation matrix

$$R^m \xrightarrow{A} R^n \longrightarrow M \longrightarrow 0.$$

This is independent of the presentation of M, and globalizes to the Fitting ideal $\operatorname{Fit}_k(\mathcal{F}) \subset \mathcal{O}_X$ of a coherent sheaf $\mathcal{F}$ on a variety X.

Definition 4.1.1 (Ramification locus, branch locus)**.** Let $f\colon X \to Y$ be a morphism of irreducible varieties and $d := \dim X - \dim Y$. The *ramification locus* $\operatorname{Ram}(f) \subseteq X$ is the scheme locally defined by the d-th fitting ideal $\operatorname{Fit}_d(\Omega_{X/Y}) \subseteq \mathcal{O}_X$.

The *Branch locus* or *discriminant* $\operatorname{Branch}(f) \subseteq Y$ is the image $\overline{f(\operatorname{Ram}(f))}$.

One usually applies this definition only when f is flat of constant relative dimension d, or when f is finite (or both).

Remark 4.1.2. There are different conventions about the scheme structure of $\operatorname{Ram}(f)$, whose underlying set is the *support* of $\Omega_{X/Y}$. For example, one could take the annihilator ideal of $\Omega_{X/Y}$ or the reduced scheme structure. We follow the convention from [EH06, Definition V-21], where the ramification locus is called *singular scheme of* f. We are slightly inconsistent in that we should then define $\operatorname{Branch}(f)$ as the *Fitting image* of f, defined by $\operatorname{Fit}_0(f_*\mathcal{O}_X) \subseteq \mathcal{O}_Y$, but this will not matter for our purposes. ◁

This general definition also applies to morphisms $f\colon X \to Y$. However, we will only apply it in the special case where X, Y are both irreducible smooth varieties with $\dim X \leq \dim Y$ (most often equality, so f is generically finite). In this case the cotangent sheaves $\Omega^1_{X/\mathbb{C}}$ and $\Omega^1_{Y/\mathbb{C}}$ are locally free sheaves, respectively, and Lemma 0.1.36(i) gives the presentation

$$f^*\Omega^1_{Y/\mathbb{C}} \xrightarrow{\delta_f} \Omega^1_{X/\mathbb{C}} \longrightarrow \Omega^1_{X/Y} \longrightarrow 0.$$

Here δ_f is the dual of the differential, locally represented by the transpose of the Jacobi map. From this presentation we see that $\operatorname{Fit}_0(\Omega^1_{X/Y}) = I_{\dim Y}(J_f) \subseteq \mathcal{O}_X$ is ideal of minors

of the Jacobi matrix locally. We obtain a simple special case of the *Zariski–Nagata purity theorem* [Sta25, Lemma 0BMB].

Theorem 4.1.3 (Purity theorem)**.** *Let $f\colon X \to Y$ be a morphism of smooth irreducible varieties of the same dimension. The subscheme $\operatorname{Ram}(f) \subseteq X$ is a locally principal divisor in X, that is, it is locally defined by a single equation. As a set, we have*

$$\operatorname{Ram}(f) = \left\{ x \in X \;\middle|\; f^{-1}(f(x)) \text{ is non-reduced or positive-dimensional in } x \right\}.$$

Additionally, if f is finite of degree d, then the Branch locus is also of codimension 1 and

$$\operatorname{Branch}(f) = \left\{ y \in Y \;\middle|\; f^{-1}(y) \text{ not reduced} \right\} = \left\{ y \in Y \;\middle|\; \#f^{-1}(y) < d \right\}.$$

Proof. The first part follows from the previous discussion. For the set-theoretic description, note that $x \in \operatorname{Ram}(f)$ if and only if $\operatorname{rank} J_f(x) < n$, which is equivalent to $\dim_{\mathbb{C}} T_x f^{-1}(f(x)) > 0$, which happens if and only if $\dim_x f^{-1}(f(x)) > 0$ or $\dim_x f^{-1}(f(x)) = 0$ and $f^{-1}(f(x))$ is not reduced at x.

If f is finite, then its fibers have the same dimension 0 on all of Y, hence f is flat. Thus, $\operatorname{length}(f^{-1}(y)) = d$ for all $y \in Y$, and so $f^{-1}(y)$ is non-reduced if and only if it consists of fewer than d distinct points (some of which are then necessarily non-reduced). □

Under the assumption of finiteness (as in the second part of the purity theorem), the discriminant $\operatorname{Branch}(f)$ gives exactly the locus where f is *not* a locally trivial fibration:

Theorem 4.1.4 (Ehresmann [Ehr51])**.** *Let $f\colon M \to N$ be a surjective smooth map of smooth manifolds and $q \in N$. Assume*

(i) *q is a regular value of f: For all $p \in f^{-1}(q)$, $\mathrm{d}f_p\colon T_pM \to T_qN$ is surjective;*

(ii) *f is locally* (topologically) proper *around q: There is a neighborhood $q \in V \subseteq N$ such $f^{-1}(K)$ is compact for $K \subseteq V$ compact.*

Then f is a a fiber bundle in a neighborhood of q, i. e. of the form $M \cong N \times F \xrightarrow{\mathrm{pr}_F} N$, where $F = f^{-1}(q)$.

This leads to the following interpretation of the discriminant of a finite (in particular proper) morphism of smooth irreducible varieties of the same dimension:

- Outside $\operatorname{Branch}(f) \subseteq Y$, the map $X \to Y$ is an (analytic-)locally trivial fibration with fibers $f^{-1}(y)$ consisting of $\deg f$ many distinct reduced points.
- When the parameter $y \in Y$ approaches the discriminant, then two points come together to form a non-reduced point ($y \in \operatorname{Branch}(f)$ is not a regular value).
- If X, Y and f are defined over $\mathbb{R}$, then applying Ehresmann's theorem to the real points $X(\mathbb{R}) \to Y(\mathbb{R})$, we also see that in each connected component of $Y(\mathbb{R}) \setminus \operatorname{Branch}(f)$ the number of real solutions is constant.

We also note the following useful base-change property:

Lemma 4.1.5. *Given a cartesian diagram of varieties (that is, $X' = X \times_Y Y'$)*

$$\begin{array}{ccccc} \mathrm{Ram}(f') & \subseteq & X' & \xrightarrow{\varphi'} & X \supseteq \mathrm{Ram}(f) \\ & & \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ \mathrm{Branch}(f') & \subseteq & Y' & \xrightarrow{\varphi} & Y \supseteq \mathrm{Branch}(f) \end{array}$$

we have that $\mathrm{Ram}(f') = \varphi'^{-1}(\mathrm{Ram}(f))$ *and* $\mathrm{Branch}(f') \subseteq \varphi^{-1}(\mathrm{Branch}(f))$ *as subschemes of X' and Y' respectively.*

Proof. By Lemma 0.1.36 we have $\varphi'^*\Omega^1_{X/Y} = \Omega^1_{X'/Y'}$. The formation of fitting ideals is compatible with base change as well [Sta25, Lemma 0C3D], that is,

$$\langle \varphi'^{-1} \operatorname{Fit}_0(\Omega^1_{X/Y}) \rangle_{\mathcal{O}_{Y'}} = \operatorname{Fit}_0(\varphi'^*\Omega^1_{X/Y}) \subseteq \mathcal{O}_{Y'}.$$

This shows the statement about the ramification loci. The statement about branch loci follows as we have scheme-theoretic inclusions

$$f'(\mathrm{Ram}(f')) = f'(\varphi'^{-1}(\mathrm{Ram}(f))) \subseteq \varphi^{-1}(f(\mathrm{Ram}(f))). \qquad \square$$

Remark 4.1.6. A cartesian diagram such as in Lemma 4.1.5 is also cartesian in the category of sets. If $R \subseteq X$, then it *is true* as sets that $f'(\varphi'^{-1}(R)) = \varphi^{-1}(f(R))$. The major difference in the category of varieties or schemes (apart from multiplicities) is that the scheme-theoretic image is the *closure* of the set-theoretic image, hence the preimage φ^{-1} may introduce additional components. We will see examples of this below when comparing the logarithmic and Hurwitz discriminant in Section 4.4. This problem disappears at least set-theoretically if one assumes f to be proper. $\triangleleft$

4.1.2 Examples of discriminants and the Hurwitz form

A good reference for discriminants and related topics is the book [GKZ08]. All discriminants for us will be of the following form: There is a smooth irreducible parameter variety Y and an incidence variety $X \subseteq Z \times Y$ such that X is irreducible and smooth as well (often a fiber bundle over Z). In this setup, the discriminant ∇ is the branch locus of the projection $\mathrm{pr}_Y \colon X \to Y$.

Example 4.1.7 (Univariate discriminant)**.** Let $Z = \mathbb{P}^1$ and $Y = \mathbb{P}(\mathbb{C}[z_0, z_1]_n) = \mathbb{P}^n$. Consider the incidence

$$X = \{\, (f, p) \mid f(p) = 0 \,\} \subseteq \mathbb{P}^n \times \mathbb{P}^1.$$

The fiber over every $p \in \mathbb{P}^1$ is the linear space of codimension 1 of polynomials vanishing in p. Hence X is a projective bundle over $\mathbb{P}^1$, and thus irreducible and smooth of dimension n. The ramification locus consists of the the subset

$$R \coloneqq \{\, (f, p) \mid p \text{ is a multiple root of } f \,\} \subseteq \mathbb{P}^n \times \mathbb{P}^1$$

and hence the branch locus is the locus of polynomials having a multiple root. This locus is a hypersurface in $\mathbb{P}^n_a$ defined by a single polynomial $\operatorname{disc}_n = \operatorname{disc}_n(f)(a_0, \dots, a_n)$ in the coefficients of the polynomial $f = \sum_{i=0}^n a_i z_0^{n-i} z_1^i$, the *discriminant* of f. For example,

$$\operatorname{disc}_2 = a_1^2 - 4\,a_0 a_2, \qquad \operatorname{disc}_3 = a_1^2 a_2^2 - 4\,a_0 a_2^3 - 4\,a_1^3 a_3 + 18\,a_0 a_1 a_2 a_3 - 27\,a_0^2 a_3^2$$

and in general $\deg \operatorname{disc}_n = 2n - 2$. When f depends on several variables, we write $\operatorname{disc}_x(f) \coloneqq \operatorname{disc}_{\deg_x f}(f(x))$ to emphasize the variable dependence. ◁

When the incidence X is of smaller dimension than the parameter space Y, the Fitting ideal is trivial (by convention). In this case the discriminant is just the image of the incidence onto the parameter variety.

Example 4.1.8 (Resultant)**.** The resultant is the discriminant of the statement that two general polynomials f, g have no common root. The subset of $\mathbb{P}(\mathbb{K}[z_0, z_1]_d) \times \mathbb{P}(\mathbb{K}[z_0, z_1]_e)$ of polynomials with a common root is a hypersurface of bidegree (d, e). It is defined by the polynomial $\operatorname{res}(f, g)$ which can be constructed as the determinant of the linear map

$$S_{e-1} \oplus S_{d-1} \to S_{d+e-1}, \qquad (h_1, h_2) \mapsto h_1 f + h_2 g,$$

compare Bézout's Lemma 0.1.14. Here the incidence

$$X = \{\, (f, g, p) \mid f(p) = g(p) = 0 \,\} \subseteq \mathbb{P}^d \times \mathbb{P}^e \times \mathbb{P}^1$$

has dimension $d + e - 1$, hence the Branch locus is the image of the incidence. Comparing with the previous example, $\operatorname{res}(\frac{\partial}{\partial z_0} f, \frac{\partial}{\partial z_1} f)$ recovers the discriminant $\operatorname{disc}(f)$. ◁

Of particular importance for us are discriminants measuring contact of a projective variety with linear spaces of "unexpected" tangency.

Example 4.1.9 (Chow forms)**.** Let $X \subseteq \mathbb{P}^n$ be an irreducible variety of dimension k and $\mathbb{G} \coloneqq \mathbb{G}(n - k - 1, \mathbb{P}^n)$. Most projective subspaces $\Lambda \in \mathbb{G}$ of dimension $n - k - 1$ do *not* intersect X, consider the incidence of linear subspaces that do:

$$W \coloneqq \{\, (x, \Lambda) \mid x \in \Lambda \,\} \subseteq X \times \mathbb{G}\,.$$

The fiber over $x \in X$ is isomorphic to a Grassmannian $\mathbb{G}(n - k - 2, \mathbb{P}(\mathbb{K}^{n+1}/\mathbb{K}x))$, hence W is an irreducible variety of dimension

$$\dim X + \dim \mathbb{G}(n - k - 2, \mathbb{P}^{n-1}) = k + (n - k - 1)(k + 1) = \dim \mathbb{G} - 1.$$

Consequently, the image $\mathcal{Z}_0(X)$ of the projection to $\mathbb{G}$ is an irreducible subvariety. The projection is generically finite, for example by the Multisecant Lemma 0.3.6 (actually birational), so $\mathcal{Z}_0(X)$ is of codimension 1. The Plücker ring $R = S(\mathbb{G})$, the homogeneous coordinate ring of the Grassmannian in the Plücker embedding, is a unique factorization domain [GKZ08, Proposition 3.2.1]. It follows that $\mathcal{Z}_0(X)$ is defined by a *single* irreducible element in R_d for some d. This is the *Chow form* $\operatorname{Ch}_X \in R_d$, unique up to scaling. This definition can be extended to equidimensional varieties, and, by linearity, to all effective k-cycles in $\mathbb{P}^n$, where $\operatorname{Ch}_{\sum_i m_i X_i} = \prod_i (\operatorname{Ch}_{X_i})^{m_i}$.

It turns out that the degree of $\mathrm{Ch}_X \in R_d$ coincides with the degree of X as a projective variety (or, in general, as a k-cycle). Furthermore, one can recover the variety (or cycle) from its chow form, so the Chow form provides a parameter space for effective k-cycles of degree d. For an introduction to Chow forms see [GKZ08, Chapter 3 & 4] or the introductory article [DS95]. ◁

The Chow form is the first in a sequence of associated varieties

$$\mathcal{Z}_i(X) = \overline{\{\, \Lambda \mid \exists x \in X_{\mathrm{reg}} \cap \Lambda \text{ s.t. } \dim \mathbb{T}_x X \cap \Lambda \geq i \,\}} \subseteq \mathbb{G}(n - \dim X - 1 + i, \mathbb{P}^n)$$

interpolating between the Chow form ($i = 0$) and the dual variety ($i = \dim X$). Of particular interest to us is the first associated hypersurface, which we consider next.

4.1.3 Hurwitz forms

Let $X \subseteq \mathbb{P}^n$ be an irreducible variety of dimension k and set $\mathbb{G} \coloneqq \mathbb{G}(n-k, \mathbb{P}^n)$. Most projective subspaces $\Lambda \in \mathbb{G}$ of dimension $n-k$ intersect X transversally in $\deg X$ points. By Theorem 4.1.3, the locus of pairs (x, Λ) such that $x \in X_{\mathrm{reg}} \cap \Lambda$ but Λ does *not* intersect X transversally at x is thus the *ramification locus* W of the projection

$$\{\, (x, \Lambda) \in X_{\mathrm{reg}} \times \mathbb{G} \mid x \in X_{\mathrm{reg}} \cap \Lambda \,\} \to \mathbb{G}$$

of smooth irreducible varieties of the same dimension $k(n+1-k)$. To see that the ramification locus W is an irreducible subset, consider the incidence

$$W' \coloneqq \{\, (x, L, \Lambda) \mid x \in X_{\mathrm{reg}},\ L \in \mathbb{G}(1, \mathbb{T}_x X),\ x \in L \subseteq \Lambda \,\}.$$

The fiber over $x \in X_{\mathrm{reg}}$ consists of flags $x \in L \subseteq \Lambda$, where L is restricted to come from the tangent space $\mathbb{T}_x X$. Thus, W' is smooth and irreducible, and so is its projection onto $X_{\mathrm{reg}} \times \mathbb{G}$, the set W. We obtain the first associated hypersurface $\mathcal{Z}_1(X) \subseteq \mathbb{G}$ as the image closure of W under the projection $W \to \mathbb{G}$.

Theorem 4.1.10 ([Stu17, Theorem 1.1])**.** *Let $X \subseteq \mathbb{P}^n$ be an irreducible variety that is not a linear space. Then $\mathcal{Z}_1(X) \subseteq \mathbb{G}(n-k, \mathbb{P}^n)$ is an irreducible hypersurface, defined by a homogeneous element $\mathrm{Hu}_X \in S(\mathbb{G}(n-k, \mathbb{P}^n))$. If the singular locus of X has codimension ≥ 2, then*

$$\deg \mathrm{Hu}_X = 2 \deg X + 2 g_{\mathrm{s}}(X) - 2.$$

Here, $g_{\mathrm{s}}(X)$ is the *sectional genus* of X, that is, the arithmetic genus of the curve $X \cap L$ for general $L \in \mathbb{G}(n-k+1, \mathbb{P}^n)$. Under the assumption $\operatorname{codim}_X X_{\mathrm{sing}} \geq 2$, such curves are smooth by Bertini, and so one can equivalently use the geometric genus.

Definition 4.1.11 (Hurwitz form)**.** The element $\mathrm{Hu}_X \in S(\mathbb{G}(n-k, \mathbb{P}^n))_{\deg \mathrm{Hu}_X}$ in the Plücker ring defining $\mathcal{Z}_1(X)$ is the *Hurwitz form*.

Example 4.1.12. If $C \subseteq \mathbb{P}^n$ is a curve, then $\mathcal{Z}_1(C) \subseteq \mathbb{G}(n-1, \mathbb{P}^n) = (\mathbb{P}^n)^\vee$ is the (closure of the) set of tangent hyperplanes. In other words, $\mathcal{Z}_1(C) = C^\vee$ is the dual

variety. In this case Theorem 4.1.10 gives the degree $\deg \mathrm{Hu}_C = 2\deg C + 2g(C) - 2$ for *smooth* C. For a smooth plane curve of degree d, this recovers the usual formula for the degree of the dual curve:

$$\deg C^\vee = \deg \mathrm{Hu}_C = 2d + 2\binom{d-1}{2} - 2 = d(d-1).$$

For singular plane curves this formula has to be modified using *Plücker's formula.* ◁

To obtain a finer invariant than the degree of $\mathrm{Hu}_X \in S(\mathbb{G}(n-k, \mathbb{P}^n))$, one can use the $\mathbb{Z}_+^{n+1}$-grading induced by the natural torus action of $(\mathbb{C}^\times)^{n+1}$ on $\mathrm{Gr}(n-k+1, \mathbb{C}^{n+1})$:

If $\Lambda \in \mathbb{G}(n-k, \mathbb{P}^n) \cong \mathrm{Gr}(k, (\mathbb{C}^{n+1})^\vee)$ is represented as the *kernel* of $L \in \mathbb{C}^{k\times(n+1)}$, with Plücker coordinates $p_I([\Lambda]) \coloneqq \det L_I$, $I \in \binom{[n+1]}{k}$, then

$$\deg p_{i_1,\dots,i_d} = \mathrm{e}_{i_1} + \cdots + \mathrm{e}_{i_d} \in \mathbb{Z}^{n+1}.$$

If $w \in \mathbb{Z}^{n+1}$ is a weight vector corresponding to a 1-parameter family of the torus action on $\mathbb{P}^n$, then it induces a 1-parameter family on $\mathbb{G}(n-k, \mathbb{P}^n)$ with corresponding weight vector $(w_I = \sum_{i\in I} w_i)_I$. The following theorem relates the initial form of the Hurwitz form to the initial ideal of $\mathrm{I}(X)$ in the square-free case.

A square-free monomial ideal M defines a reduced union of coordinate subspaces $\mathbb{V}(M) \subseteq \mathbb{C}^{n+1}$. For a subset $S \in \binom{[n+1]}{k}$ with $Z = \langle \mathrm{e}_S \rangle \coloneqq \langle \{ \mathrm{e}_s \mid s \in S \} \rangle \subset \mathbb{V}(M)$, let $\nu(S)$ be the number of irreducible components of $\mathbb{V}(M)$ containing Z.

Theorem 4.1.13 ([Stu17, Corollary 4.4]). *If $w \in \mathbb{Z}^{n+1}$ is a weight vector, and if $M = \mathrm{in}_w(\mathrm{I}(X))$ is a square-free monomial ideal, then*

$$\mathrm{in}_w(\mathrm{Hu}_X) = \prod_S p_S^{2(\nu(S)-1)},$$

where the product runs over $S \in \binom{[n+1]}{k}$ with $\langle \mathrm{e}_S \rangle \subseteq \mathbb{V}(M)$.

4.2 Logarithmic discriminants in likelihood geometry

In this section we briefly introduce the likelihood geometry setup from algebraic statistics, with a view towards the maximum likelihood degree. The ML degree of a smooth subvariety $X \subseteq (\mathbb{C}^\times)^n$ is the number of complex critical points of the log likelihood function $\mathcal{L}_u(x) = \sum_{i=0}^n u_i \log(x_i)$ on X for general data points $u \in \mathbb{C}^{n+1}$. We define the logarithmic discriminant of X as the discriminant of the maximum likelihood estimation problem, the closure of parameters u which introduce degenerate critical points. An introduction to likelihood geometry is given in [HS14].

4.2.1 Maxmimum likelihood degree

A discrete statistical model is a subset of probability distributions

$$X \subseteq \Delta_n \coloneqq \left\{ p \in \mathbb{R}^{n+1}_{>0} \;\middle|\; p_0 + \cdots + p_n = 1 \right\}$$

in the *probability simplex* Δ_n. Given a data point $u \in \mathbb{N}^{n+1}$, the *maximum likelihood estimation* (MLE) problem asks to find the model parameters $x \in X$ that minimize the *likelihood*

$$\underset{x \in X}{\text{minimize}}\, x_0^{u_0} \cdots x_n^{u_n}.$$

Assuming X is parametrized by smooth functions $\iota = (\iota_0, \ldots, \iota_n) \colon \Theta \to \Delta_n$, it is equivalent to minimize the *log-likelihood function*[1]

$$\mathcal{L}_u(\theta) = \log(\iota_0(\theta)^{u_0} \cdots \iota_n(\theta)^{u_n}) = u_0 \log \iota_0(\theta) + \cdots + u_n \log \iota_n(\theta).$$

In this formulation, one can find the minimizer by studying the *log-likelihood equations* $\nabla_\theta \mathcal{L}_u(\theta) = 0$.

In the context of *algebraic statistics*, the algebraic model is given by an algebraic variety $X \subseteq \mathbb{T}^{n+1} \coloneqq (\mathbb{C}^\times)^{n+1}$, where the underlying statistical model is then $X \cap \Delta_n$. Usually, one assumes $X \cap \mathbb{V}(p_0 + \cdots + p_n) = \emptyset$, in which case one can normalize to have $X \subset \mathbb{V}(p_0 + \cdots + p_n - 1)$, though this is not essential for our discussion. Closed subvarieties of the algebraic torus are called *very affine varieties*. They are (abstractly) characterized among affine varieties $\operatorname{Spec} A$ by the fact that A is generated as a $\mathbb{C}$-algebra by the units $A^\times$.

Pick local coordinates $\theta_1, \ldots, \theta_d$ around a regular point $x \in X_{\mathrm{reg}}$, then the likelihood equations are locally given by

$$\nabla_\theta \mathcal{L}_u(\theta) = \left(\sum_{j=0}^{n} u_j \frac{1}{\iota_j(\theta)} \frac{\partial \iota_j(\theta)}{\partial \theta_i} \right)_{i=1,\ldots,d} = 0.$$

In particular, they make sense for arbitrary $u \in \mathbb{P}^n$ and define an algebraic subset of $X_{\mathrm{reg}} \times \mathbb{P}^n_u$.

Theorem 4.2.1 ([HS14, Theorem 1.6])**.** *Let $X \subseteq \mathbb{T}^{n+1}$ be a very affine variety of dimension d. The incidence*

$$\mathcal{L}_X^\circ \coloneqq \{\, (x, u) \mid \nabla \mathcal{L}_u(x) = 0 \,\} \subseteq X_{\mathrm{reg}} \times \mathbb{P}^n$$

is a projective bundle of rank $n - \dim X$ over X_{reg}. The projection $\mathcal{L}_X^\circ \to \mathbb{P}^n$ is generically finite.

We will see a proof of this theorem in the special case of linear models in Lemma 4.3.3. This theorem shows that for general $u \in \mathbb{N}^{n+1}$ (or $u \in \mathbb{C}^{n+1}$), the set of complex solutions $\operatorname{Crit}_X(u) \subseteq X_{\mathrm{reg}}$ to the critical equations is finite and reduced.

Definition 4.2.2. The *likelihood correspondence* is the closure of the incidence variety $\mathcal{L}_X \coloneqq \overline{\mathcal{L}_X^\circ} \subseteq X \times \mathbb{P}^n$ from the previous theorem. The *maximum likelihood degree* $\operatorname{MLdeg}(X)$ of the variety $X \subseteq \mathbb{T}^{n+1}$ is the degree of the morphism $\mathcal{L}_X \to \mathbb{P}^n_u$, that is, the number of complex critical points $\#\operatorname{Crit}_X(u)$ for general $u \in \mathbb{P}^n$.

[1] We are inconsistent with the statistics literature here, where $\mathcal{L}$ denotes the likelihood and ℓ denotes the log-likelihood. However, we reserve the letter ℓ for linear forms in the upcoming sections.

Example 4.2.3. Maximum likelihood degree 1 occurs precisely when there is a rational map from $\mathbb{P}^n_u$ to X which gives the MLE. A famous example is the binomial distribution with parameter p

$$X = \left\{ \left(\binom{n}{i} p^i (1-p)^{n-i} \right)_i \;\middle|\; p \in \mathbb{C} \setminus \{0,1\} \right\} \subseteq \mathbb{T}^{n+1}.$$

The maximum likelihood estimator (for $u \in \mathbb{N}^{n+1}$) is

$$\mathrm{MLE}_X(u) = \frac{1}{n(u_0 + u_1 + \cdots + u_n)} \sum_{i=0}^{n} i u_i.$$

This is a rational function in the data u, so $\mathrm{MLdeg}(X) = 1$. ◁

The following result gives a surprising connection between the topology of a very affine variety and its ML degree.

Theorem 4.2.4 ([Huh13])**.** *If $X \subseteq \mathbb{T}^{n+1}$ is a* smooth *very-affine variety of dimension d, then* $\mathrm{MLdeg}(X) = (-1)^d \chi(X)$.

In particular, the ML degree is independent of the embedding $X \hookrightarrow \mathbb{T}^{n+1}$. For an exposition of the proof and other aspects of likelihood geometry in algebraic statistics, consider the article [HS14].

4.2.2 Logarithmic discriminants

In the view of the Ehresmann Theorem 4.1.4 and Theorem 4.2.4, over a dense open set of parameters $u \in \mathbb{P}^n_d$, the likelihood correspondence is a $|\chi(X)|$-sheeted covering. When passing from generic to non-generic data, this property can fail in one of the following ways:

(i) Two critical points collide to form a degenerate (isolated) critical point.

(ii) A positive-dimensional component of critical points appears.

(iii) The point escapes to infinity (the boundary of $\mathbb{T}^{n+1}$), or to the singular locus X_{sing}.

The (closure of the) union of these loci in $\mathbb{P}^n_u$ was coined the *data discriminant* by Rodriguez & Tang [RT15], who propose algorithms to compute the data discriminant. Conditions (i) and (ii) describe the branch locus of the morphism $\mathcal{L}^\circ_X \to \mathbb{P}^n$ of irreducible smooth varieties in the sense of Definition 4.1.1. The points in (iii) were studied by Sattelberger & van der Veer in [SV23] with tools from tropical geometry and Bernstein–Sato theory.

Definition 4.2.5. Let $X \subseteq \mathbb{T}^{n+1}$ be a very affine variety. The *logarithmic discriminant* $\nabla_{\log}(X) \subseteq \mathbb{P}^n_u$ is the branch locus of the morphism $\mathcal{L}^\circ_X \to \mathbb{P}^n_u$. If $\nabla_{\log}(X) \subset \mathbb{P}^n$ is a hypersurface, then its defining equation, unique up to scaling, is denoted by $\Delta_{\log}(X)$.

By Theorem 4.1.3, the logarithmic discriminant describes the set

$$\nabla_{\log}(X) \coloneqq \overline{\{\, u \in \mathbb{P}^n \mid \mathrm{Crit}_X(u) \text{ is infinite or non-reduced} \,\}}.$$

Unraveling the definition leads to a rather concrete description of the ramification locus.

Proposition 4.2.6. *The ramification locus of the projection $f\colon \mathcal{L}_X^\circ \to \mathbb{P}^n$ is locally defined by the Hessian determinant of $\mathcal{L}_u(x)$:*

$$\det H_x(\mathcal{L}_u) \coloneqq \det\left[\frac{\partial^2}{\partial x_j \partial x_k} \sum_{i=0}^{n} u_i \log \iota_i(x)\right]_{j,k=1}^{d} = 0.$$

Proof. The ramification locus is defined by the determinant of the Jacobian of $\nabla\mathcal{L}_u(x)$, which is precisely the Hessian determinant. □

Example 4.2.7. If $\mathrm{MLdeg}(X) = 1$ (such as in Example 4.2.3), then $\mathrm{Crit}_X(u)$ can never contain an isolated degenerate point by Zariski's Main Theorem [Sta25, Lemma 05K0]. Therefore, in that case $\nabla_{\log}(X)$ consists of the (closure of the) parameters u with positive-dimensional critical locus. By the fiber dimension theorem, $\dim \nabla_{\log}(X)$ will have codimension at least 2. In [HS14, Example 1.14], this is illustrated for the binary independence model $X = \{x_0x_3 - x_1x_2 = 0\} \subseteq \mathbb{T}^4$ which has $\mathrm{MLdeg}(X) = 1$. ◁

Perhaps the simplest interesting example of a non-trivial logarithmic discriminant is the following, already in the spirit of the following sections:

Example 4.2.8. Consider the smooth very-affine variety $X \subseteq \mathbb{T}^3$ parametrized by $x \mapsto (x, x-1, x-2)$. The image is a line in $\mathbb{C}^3$ intersected with the torus, isomorphic to $\mathbb{C}^1 \setminus \{0,1,2\}$. As such it has signed Euler characteristic $(-1)^1\chi(X) = 2$, so the logarithmic discriminant consists of parameters where the two critical points merge into a double point. The (single) critical equation reads

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(u_0 \log x + u_1 \log(x-1) + u_2 \log(x-2)\right) = \frac{u_0}{x} + \frac{u_1}{x-1} + \frac{u_2}{x-2} \overset{!}{=} 0.$$

Clearing denominators reveals a quadratic equation in x:

$$(u_0+u_1+u_2)x^2 + (-3u_0 - 2u_1 - u_2)x + 2u_0 \overset{!}{=} 0.$$

The discriminant of this monic polynomial reveals the logarithmic discriminant

$$\Delta_{\log}(X) = u_0^2 + 4u_0u_1 + 4u_1^2 - 2u_0u_2 + 4u_1u_2 + u_2^2 \in \mathbb{C}[u_0,u_1,u_2]_2.$$ ◁

Many more examples will be collected in Section 4.3.3.

4.3 Linear models and their discriminants

In this section we consider perhaps the simplest class of algebraic discrete statistical models: Linear models, represented by hyperplane arrangement complements. We review basic definitions of hyperplane arrangements and their combinatorics and geometry. We then describe how to compute the logarithmic discriminant of a hyperplane arrangement and study many concrete examples.

4.3.1 Hyperplane arrangements

We introduce basic notions on hyperplane arrangements that we need later. A standard reference is [OT92]. In this section, $\mathbb{K}$ is either $\mathbb{C}$ or $\mathbb{R}$.

Definition 4.3.1 (Hyperplane arrangement)**.** A hyperplane arrangement is a collection of affine hyperplanes $H_i = \{\ell_i = 0\} \subseteq \mathbb{K}^d$, that is,

$$\mathcal{A} = \bigcup_{i=0}^{n} H_i = \mathbb{V}(\ell_0 \cdots \ell_n) \subseteq \mathbb{K}^d, \qquad \ell_i \in \mathbb{K}[x_1, \ldots, x_d]_{\leq 1}.$$

By slight abuse of notation, write $\mathcal{A}$ both for the set of hyperplanes as well as for their union. We collect the coefficients of the affine-linear forms ℓ_i as

$$(\ell_0(x), \ldots, \ell_n(x))^\mathsf{T} = Ax + b, \qquad L^\mathsf{T} \coloneqq [b \mid A] \in \mathbb{K}^{(n+1)\times(d+1)}.$$

If $\mathcal{A}$ is a real arrangement, then we denote by $\mathcal{A}_\mathbb{C}$ its complexification, i. e. the complex arrangement defined by the same equations.

We denote $\operatorname{rank}\mathcal{A} \coloneqq \operatorname{rank} A$. $\mathcal{A}$ is *essential* if the normals of H_i span the dual space, equivalently, $\operatorname{rank}\mathcal{A} = d$.

An arrangement $\mathcal{A}$ is *central* if $\bigcap_i H_i \neq \emptyset$, and *non-central* otherwise.

Observe that an arrangement is essential and non-central if and only if $n + 1 \geq d + 1$ and $\operatorname{rank} L = d + 1$.

Example 4.3.2 (Bi-uniform arrangement)**.** An arrangement of affine hyperplanes in $\mathbb{K}^d$ is *in general position* [OT92, Definition 5.19] or *bi-uniform* if the following equivalent conditions are satisfied:

(a) Every intersection of $k \leq d+1$ of the hyperplanes is an affine subspace of codimension k (or empty in the case $k = d + 1$);

(b) Any collection of $k \leq d$ rows of A is linearly independent and any collection of $k \leq d + 1$ rows of L^T is linearly independent;

(c) The central arrangement $\overline{\mathcal{A}} \cup H_\infty \subseteq \mathbb{P}^d$ is uniform, in the sense that all $(d+1)$-minors of the matrix $[L \mid \mathrm{e}_0] \in \mathbb{K}^{(d+1)\times(n+2)}$ are non-zero.

A generic choice of linear forms produces such an arrangement. A bi-uniform arrangement of $n + 1$ hyperplanes has rank $\min\{n + 1, d\}$, hence it is essential if and only if $n + 1 \geq d$. It is non-central if and only if $n + 1 \geq d + 1$. An example of $n + 1 = 4$ lines in $\mathbb{R}^2$ in general position is displayed in Figure 4.1. ◁

If $X = \mathbb{C}^d \setminus \mathcal{A}$ is a complex hyperplane arrangement complement, then the linear forms give a morphism of varieties

$$\ell \colon X \to \mathbb{T}^{n+1}.$$

X is a very affine variety if and only if $\mathcal{A}$ is essential, in which case ℓ gives a closed embedding. The image is a linear subspace of $\mathbb{C}^{n+1}$ intersected with $\mathbb{T}^{n+1}$, for this reason X is called a *linear model* in algebraic statistics. In this case, the likelihood correspondence has a particularly simple description (compare also [HS14, Proposition 1.19]):

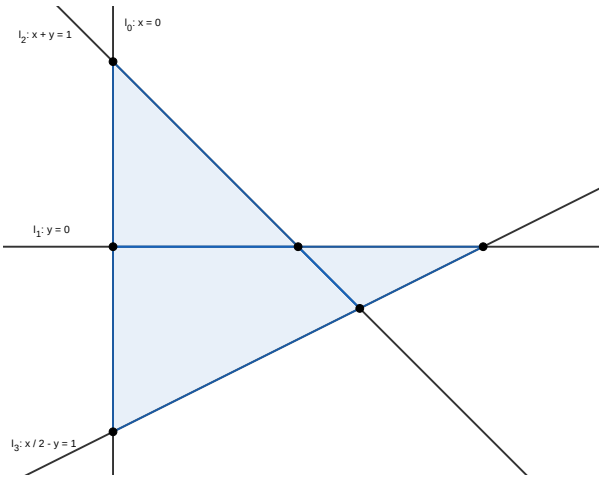


Figure 4.1: A bi-uniform arrangement of 4 lines in the real plane.

Lemma 4.3.3. *Let $X = \mathbb{C}^d \setminus \mathcal{A}$ be an essential hyperplane arrangement complement. The likelihood correspondence $\mathcal{L}_X \subseteq X \times \mathbb{P}^n_u$ is a projective bundle over X globally defined by the equations*

$$A^\mathsf{T} \cdot \operatorname{diag}(1/\ell_0(x), \dots, 1/\ell_n(x)) \cdot \begin{pmatrix} u_0 \\ \vdots \\ u_n \end{pmatrix} = 0. \tag{4.1}$$

Proof. For an affine linear form $\ell(x) = a^\mathsf{T} x + b$ we have $\nabla_x \log \ell(x) = a^\mathsf{T} \cdot \frac{1}{\ell(x)}$. Apply this to the log-likelihood function $\mathcal{L}_u(x) = \sum_i u_i \log \ell_i(x)$. □

4.3.2 Euler characteristics of hyperplane arrangements

An essential hyperplane arrangement complement X is a smooth very affine variety, hence its ML degree is given by the signed Euler characteristic $(-1)^d \chi(\mathbb{C}^d \setminus \mathcal{A})$. In fact, Theorem 4.2.4 was first conjectured for hyperplane arrangement complements by Varchenko [Var95] and proven for real arrangements, and the case of complex arrangements was settled by Orlik & Terao [OT95].

To give a combinatorial interpretation of this Euler characteristic, we briefly introduce the characteristic polynomial and its relationship to the topology of $\mathbb{K}^d \setminus \mathcal{A}$.

Definition 4.3.4 (Lattice of flats $\mathrm{L}(\mathcal{A})$, characteristc polynomial $\chi_\mathcal{A}$)**.** Let $\mathcal{A}$ be an arrangement. The *semilattice of flats* $\mathrm{L}(\mathcal{A})$ is the set of all non-empty intersections of hyperplanes (including $\bigcap_\emptyset = \mathbb{K}^d$), partially ordered by reverse inclusion.

For $\mathcal{B} \subseteq \mathcal{A}$, denote $\mathrm{f}(\mathcal{B}) := \dim_\mathbb{K} \bigcap_{H \in \mathcal{B}} H \geq -1$, then the *characteristic polynomial* of $\mathcal{A}$ is

$$\chi_\mathcal{A}(t) = \sum_{\substack{\mathcal{B} \subseteq \mathcal{A} \\ \mathrm{f}(\mathcal{B}) \geq 0}} (-1)^{|\mathcal{B}|} t^{\mathrm{f}(\mathcal{B})} = \sum_{F \in \mathrm{L}(\mathcal{A})} \mu(F)\, t^{\dim F}.$$

Here $\mu(-) = \mu_{\mathrm{L}(\mathcal{A})}(\mathbb{K}^d, -)$ is the *Möbius function* of the semilattice of flats, characterized by $\sum_{\mathbb{K}^n \leq G \leq F} \mu(G) = 0$ if $F \neq \mathbb{K}^d$ and $\mu(\mathbb{K}^d) = 1$.

The following theorem links the characteristic polynomial $\chi_{\mathcal{A}}$ to topological properties of the complement of $\mathcal{A}$. For a discussion of the result and the topology of hyperplane arrangements, see [OT92, Chapter 5, specifically Theorem 5.93].

Theorem 4.3.5. *Let $\mathcal{A}$ be an arrangement and consider the $\mathbb{K}$-manifold $X = \mathbb{K}^d \setminus \mathcal{A}$.*

($\mathbb{C}$) *Let $\mathcal{A}$ be a complex arrangement, then the cohomology groups $\mathrm{H}^i_{\mathrm{sing}}(X, \mathbb{Z})$ are free of finite rank $b_i(X)$. Let $P_X(t) \coloneqq \sum_{i \geq 0} b_i(X) t^i$, then*

$$P_X(t) = (-t)^d \chi_{\mathcal{A}}(-t^{-1}) \in \mathbb{Z}[t].$$

In particular, $\chi_{\mathcal{A}}(1) = \chi(X)$ is the Euler characteristic of X.

($\mathbb{R}$) *Let $\mathcal{A}$ be a real arrangement, then X has a finite number $\mathrm{ch}(\mathcal{A})$ of connected components. Let $\mathrm{bch}(\mathcal{A})$ be the number of bounded components (or* relatively bounded *if $\mathcal{A}$ is not essential). Then we have the relations*

$$\mathrm{ch}(\mathcal{A}) = (-1)^d \chi_{\mathcal{A}}(-1), \qquad \mathrm{bch}(\mathcal{A}) = (-1)^{\operatorname{rank} \mathcal{A}} \chi_{\mathcal{A}}(1).$$

In particular, the number of bounded components of the complement of a real essential arrangement coincides with the signed Euler characteristic of its complexification.

Example 4.3.6 (Generic arrangements)**.** The bi-uniform arrangement $\mathcal{U}$ displayed in Figure 4.1 has one flat of dimension 2 (all of $\mathbb{R}^2$), 4 flats of dimension 1 (lines), and 6 flats of dimension 0 (intersection points). Consequently, its characteristic polynomial and Euler characteristic are

$$\chi_{\mathcal{U}}(t) = t^2 - 4t + 6, \qquad \chi(\mathbb{C}^2 \setminus \mathcal{U}_{\mathbb{C}}) = \chi_{\mathcal{U}}(1) = 3.$$

We see that this agrees with the number of bounded chambers, moreover, the total number of chambers is indeed $(-1)^2 \chi_{\mathcal{U}}(-1) = 11$.

In general, a bi-uniform arrangement $\mathcal{A}$ of $n+1$ hyperplanes has $\binom{n+1}{k}$ flats of codimension k, $k = 0, \ldots, d$. From this we can give a closed expression for the characteristic polynomial and Euler characteristic:

$$\chi_{\mathcal{A}}(t) = \sum_{k=0}^{d} (-1)^k \binom{n+1}{k} t^{d-k}, \qquad \chi(\mathbb{C}^d \setminus \mathcal{A}_{\mathbb{C}}) = \sum_{k=0}^{d} (-1)^k \binom{n+1}{k} = (-1)^d \binom{n}{d}. \quad \triangleleft$$

Example 4.3.7 (The $\mathcal{M}_{0,m}$-arrangement)**.** In the context of positive geometry, the equations $\nabla_x \mathcal{L}_u = 0$ appear as the scattering equations for bi-adjoint scalar ϕ^3-theories in particle physics. In particular, the tree-level amplitude for such theories is a global residue over the complex critical points [ST21, Theorem 13]. This is called the CHY formula, after

the authors of [CHY14]. For an m-particle scattering process, the arrangement $\mathcal{A} \subset \mathbb{C}^{m-3}$ is given by the $\binom{m-1}{2} - 1$ non-constant 2-minors of the following $(2 \times m)$-matrix:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 0 \\ 0 & 1 & x_1 & x_2 & \cdots & x_{m-3} & 1 \end{bmatrix}. \tag{4.2}$$

The complement $\mathbb{C}^{m-3} \setminus \mathcal{A}$ models the moduli space $\mathcal{M}_{0,m}$ of smooth projective genus zero curves with m marked points. The function $\mathcal{L}_u$ is the scattering potential, and the u_i are Mandelstam invariants. For a mathematical treatment, see [Lam24, Lecture 3].

This arrangement has $\chi(\mathcal{M}_{0,m}) = (m-3)!$. To see this, note that the arrangement is essential and real, so we can use Theorem 4.3.5($\mathbb{R}$) to equivalently count the number of bounded regions of $\mathcal{M}_{0,m-3}(\mathbb{R}) = \mathbb{R}^{m-3} \setminus \mathcal{A}$. A point $x \in \mathcal{M}_{0,m-3}(\mathbb{R})$ corresponds to a tuple of m distinct points $(0, 1, x_1, \ldots, x_{m-3}, \infty)$ in $\mathbb{P}^1(\mathbb{R})$. x is in a bounded chamber if and only if $0 < x_i < 1$ for all i. To each x we can associate a permutation $\sigma \in \mathfrak{S}_{m-3}$ such that $x_{\sigma(1)} < \cdots < x_{\sigma(m-3)}$. Conversely, any two points in $\mathcal{M}_{0,m-3}(\mathbb{R})$ with the same associated permutation can be connected by a path, so

$$\chi(\mathcal{M}_{0,m}) = \operatorname{bch}(\mathbb{R}^{m-3} \setminus \mathcal{A}) = \#\mathfrak{S}_{m-3} = (m-3)!$$

This is illustrated for $\mathcal{M}_{0,5}(\mathbb{R})$ in Figure 4.2, which has two bounded components, corresponding to $0 < x_1 < x_2 < 1$ and $0 < x_2 < x_1 < 1$. ◁

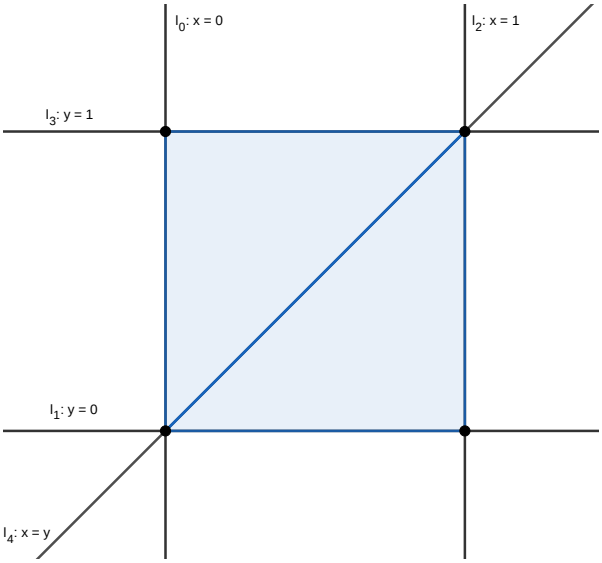


Figure 4.2: The arrangement of 5 lines associated to $\mathcal{M}_{0,5}$.

Theorem 4.3.5 shows that the ML degree of a real essential arrangement equals the number of bounded chambers. In this case, Varchenko observed a close relationship between critical points of the log-likelihood function and bounded cells:

Theorem 4.3.8 (Varchenko [Var95, Theorem 1.2.1]). *Let $\mathcal{A} \subseteq \mathbb{R}^d$ be a real essential arrangement and $u \in \mathbb{R}^{n+1}_{>0}$.*

(i) *The set of complex critical points* $\mathrm{Crit}_X(u) \subseteq \mathbb{C}^d \setminus \mathcal{A}_{\mathbb{C}}$ *of* $\mathcal{L}_u(x) = \sum_{i=0}^n u_i \log \ell_i(x)$ *is finite, non-degenerate, and all points are real.*

(ii) *The critical points are contained in the bounded components of* $\mathbb{R}^d \setminus \mathcal{A}$*, and each such component contains exactly one critical point.*

4.3.3 Examples and computations

In this section we collect examples of hyperplane arrangements and their logarithmic discriminants. The examples will guide towards the general theory developed in subsequent sections. Recall that we defined the logarithmic discriminant only for essential arrangements so that $X = \mathbb{C}^d \setminus \mathcal{A}_{\mathbb{C}}$ is very-affine.

Example 4.3.9 (The minimum number of hyperplanes)**.** In order to be essential, an arrangement needs to have at least d hyperplanes which can be normalized as $H_i = \mathbb{V}(x_i)$, $i = 1, \dots, d$. If there are exactly d hyperplanes, then the Euler characteristic is 0. If we add another hyperplane H_0, then the Euler characteristic increases to 1 if (and only if) the arrangement of $d+1$ hyperplanes is in linear general position. Up to the action of the affine group $\mathrm{Aff}(\mathbb{A}^d)$ we can assume $H_0 = \mathbb{V}(x_1 + \cdots + x_d + 1)$.

In this case the critical equations read

$$\frac{u_1}{x_1} + \frac{u_0}{x_1 + \cdots + x_d + 1} = \cdots = \frac{u_d}{x_d} + \frac{u_0}{x_1 + \cdots + x_d + 1} = 0.$$

Clearing denominators shows that the critical locus is an affine-linear subspace of $\mathbb{C}^d$, given by the linear equations $(x_1 + \cdots + x_d + 1)u_i + u_0 x_i = 0$. Furthermore, since the critical point *determines* the parameter u via $u_i/u_0 = x_i/(x_1 + \cdots + x_d + 1) \neq 0$, we have that $\mathrm{Crit}_X(u)$ is either empty or consists of a single reduced point. This shows that in this case the logarithmic discriminant is not only of codimension ≥ 2 (like in Example 4.2.7), but actually empty.

For other arrangements of maximum likelihood degree 1, the discriminant may be non-empty. For example, the square arrangement $\ell = (x_1, x_2, x_1 - 1, x_2 - 1)$ in $\mathbb{C}^2$ has discriminant $\mathbb{V}(u_0, u_2) \cup \mathbb{V}(u_1, u_3) \subseteq \mathbb{P}^3_u$. ◁

Example 4.3.10 (Points on the line)**.** A hyperplane arrangement in $\mathbb{C}^1$ or $\mathbb{R}^1$ is a collection of $n+1$ distinct points. While not all such configurations are equivalent under the action of $\mathrm{Aff}(\mathbb{A}^1)$, they are all bi-uniform. The number of critical points is $(-1)^1\chi(X) = n$, we have seen one particular instance in the smallest case $n = 2$ in Example 4.2.8. For three arbitrary distinct points $b_0, b_1, b_2 \in \mathbb{C}^1$, the logarithmic discriminant is

$$\Delta_{\log} = (u_0, u_1, u_2) \begin{bmatrix} -(b_2-b_1)^2 & (b_2-b_1)(b_0-b_2) & (b_2-b_1)(b_1-b_0) \\ (b_2-b_1)(b_0-b_2) & -(b_0-b_2)^2 & (b_1-b_0)(b_0-b_2) \\ (b_2-b_1)(b_1-b_0) & (b_1-b_0)(b_0-b_2) & -(b_1-b_0)^2 \end{bmatrix} \begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix}.$$

This quadric in u_0, u_1, u_2 has discriminant $4(b_1-b_0)^2(b_2-b_1)^2(b_0-b_2)^2$, hence is smooth and irreducible for any choice of distinct points.

The likelihood correspondence is a hypersurface in $\mathbb{P}^n \times \mathbb{C}^1 \setminus \{b_0, \dots, b_n\}$ defined by a single equation $g(\underline{u}; x) = 0$, so the logarithmic discriminant is related to the discriminant of g (after clearing denominators). We will prove in section Section 4.5.1 that these two discriminants indeed coincide. ◁

Smoothness of discriminants is usually a rare phenomenon — typically singularities appear where higher multiplicities appear in fibers (for example a triple point or two double points).

To compute more sophisticated examples, we now outline how logarithmic discriminants can be computed using computer algebra software.

Proposition 4.3.11. *The ramification locus* $\mathrm{Ram}(\mathcal{L}_X \to \mathbb{P}^n)$ *is defined by the likelihood equations (Lemma 4.3.3) and the Hessian determinant*

$$h \coloneqq \det\left(A^\mathsf{T} \cdot \operatorname{diag}\left(\frac{u_0}{\ell_0^2}, \dots, \frac{u_n}{\ell_n^2}\right) \cdot A\right) = \sum_{I \subseteq [n+1],\, |I|=d} |A_I|^2 \frac{u^I}{(\ell^I)^2} = 0.$$

Here, $u^I = \prod_{i \in I} u_i$ *and* $|A_I|$ *is the* $d \times d$*-minor of* A *indexed by* $I \subseteq [n+1] \coloneqq \{0, \dots, n\}$.

As such, the ramification locus is either empty or a global complete intersection given by this set of $d+1$ *equations.*

Proof. The ramification locus is given by these equations by Proposition 4.2.6. The determinant formula is derived from the Jacobian determinant of the likelihood equations from Lemma 4.3.3. The second equality is an expansion of the determinant by the Cauchy–Binet formula. The final statement follows as $\mathcal{L}_X \subseteq X \times \mathbb{P}^n_u$ is a complete intersection and $\mathrm{Ram}(\mathcal{L}_X \to \mathbb{P}^n)$ is a divisor. □

Using this description of the ramification locus, we can compute the logarithmic discriminant from the matrix $L^\mathsf{T} \in \mathbb{Q}^{(n+1)\times(d+1)}$ of the arrangement as follows:

1. Construct the likelihood equations $\nabla_x \mathcal{L}_u$ and the Hessian determinant h as described in Proposition 4.3.11. Using auxiliary variables y_i representing $1/\ell_i$, these $d+1$ equations lie in a polynomial ring $\mathbb{Q}[u_0, \dots, u_n, x_1, \dots, x_d, y_0, \dots, y_n]$.

2. Impose the condition $\prod_i \ell_i(x) \neq 0$. This can be accomplished either by clearing denominators and saturating by the product of the ℓ_i (in this case the extra variables y_i are not required), or by adding the equations $\{\, y_i \ell_i - 1 = 0 \mid i = 0, \dots, n \,\}$ and eliminating $y_0, \dots, y_n$.

3. Eliminate the variables $x_1, \dots, x_n$.

An implementation of this in `Macaulay2` [GS] can be found on the MathRepo page https://mathrepo.mis.mpg.de/LogDiscHyperplanes/. All of the following examples in this section have been computed in this way.

Example 4.3.12 (Continuing Example 4.3.7)**.** Five-particle scattering in bi-adjoint scalar ϕ^3-theories leads us to consider the variety $\mathcal{M}_{0,5}$. It is isomorphic to the complement of five lines in $\mathbb{C}^2$, given by the affine-linear functions

$$\ell_0 = x_1, \quad \ell_1 = x_2, \quad \ell_2 = x_1 - 1, \quad \ell_3 = x_2 - 1, \quad \ell_4 = x_2 - x_1. \tag{4.3}$$

This arrangement has $\mathrm{MLdeg}(\mathcal{M}_{0,5}) = 2$, and its logarithmic discriminant is defined by the following homogeneous polynomial

$$\Delta_{\log}(\mathcal{M}_{0,5}) = (u_0u_3 + u_0u_4 + u_1u_4 + u_1u_2 + u_2u_4 + u_3u_4 + u_4^2)^2 - 4u_0u_1u_2u_3. \tag{4.4}$$

A symmetric determinantal expression for $\Delta_{\log}$ is given in Remark 4.6.1. ◁

Passing from generic to specific arrangements, the degree of the discriminant can change:

Example 4.3.13 (From general to special)**.** Consider the bi-uniform arrangement from Figure 4.1 given by $\mathcal{A} = \mathbb{V}(x \cdot y \cdot (x+y-1) \cdot (\frac{1}{2}x - y - 1)) \subseteq \mathbb{R}^2$. Its discriminant is defined by the degree 4 form

$$\begin{aligned}\Delta_{\log}(\mathcal{A}) = {} & u_0^4+4u_0^3u_1-12u_0^2u_1^2-32u_0u_1^3+64u_1^4+10u_0^3u_2+60u_0^2u_1u_2+144u_0u_1^2u_2+256u_1^3u_2+37u_0^2u_2^2+200u_0u_1u_2^2+352u_1^2u_2^2 \\ & +60u_0u_2^3+192u_1u_2^3+36u_2^4-8u_0^3u_3+36u_0^2u_1u_3-120u_0u_1^2u_3+160u_1^3u_3-50u_0^2u_2u_3+80u_0u_1u_2u_3+400u_1^2u_2u_3-90u_0u_2^2u_3 \\ & +216u_1u_2^2u_3-36u_2^3u_3+22u_0^2u_3^2-100u_0u_1u_3^2+148u_1^2u_3^2+54u_0u_2u_3^2+180u_1u_2u_3^2-27u_2^2u_3^2-24u_0u_3^3+60u_1u_3^3+18u_2u_3^3+9u_3^4.\end{aligned}$$

This polynomial is irreducible and contains all $\binom{7}{3} = 35$ monomials of degree 4.

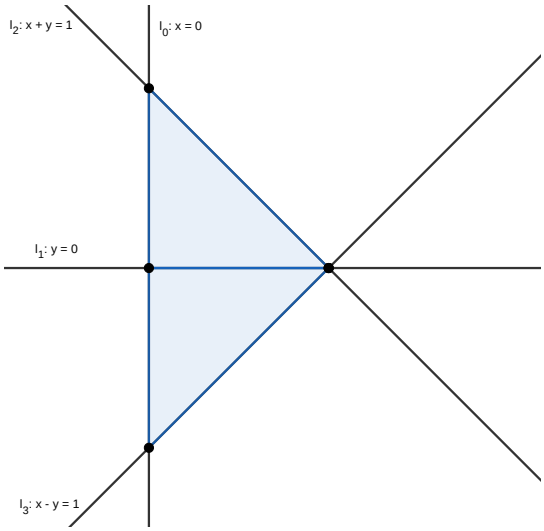


Figure 4.3: A special arrangement of 4 lines in the real plane.

On the other hand, the arrangement $\mathcal{A}' = \mathbb{V}(x \cdot y \cdot (x+y-1) \cdot (x-y-1)) \subseteq \mathbb{R}^2$ displayed in Figure 4.3 has only two bounded chambers, and its discriminant is the union of the two components

$$\nabla_{\log}(\mathcal{A}') = \mathbb{V}(4u_1^2 + 4u_1u_2 + u_2^2 + 4u_1u_3 - 2u_2u_3 + u_3^2) \cup \mathbb{V}(u_0, u_1 + u_2 + u_3).$$

The first component is a quadric surface, while the second component is a line contained in the coordinate hyperplane $u_0 = 0$. In fact, we can introduce a parameter ε and consider the family of arrangements $\mathcal{A}_\varepsilon = \mathbb{V}(x \cdot y \cdot (x+y-1) \cdot (\varepsilon x - y - 1))$, then $\mathcal{A} = \mathcal{A}_{1/2}$ and $\mathcal{A}' = \mathcal{A}_1$. Symbolically computing the discriminant (over $\mathbb{Q}(\varepsilon)$) gives a polynomial $\Delta_{\log}(\mathcal{A}_\varepsilon) \in \mathbb{C}[\varepsilon][u_0, \dots, u_3]$ of degree 4 in u, and specializing this polynomial to $\varepsilon = 1$ yields (up to constants)

$$\Delta_{\log}(\mathcal{A}_\varepsilon)|_{\varepsilon=1} = (4u_1^2 + 4u_1u_2 + u_2^2 + 4u_1u_3 - 2u_2u_3 + u_3^2) \cdot (u_1 + u_2 + u_3)^2. \qquad \triangleleft$$

We have seen in Example 4.3.13 that the logarithmic discriminant of an arrangement can be reducible with a component of codimension ≥ 1. The following example shows that even the codimension 1 part of the discriminant can split:

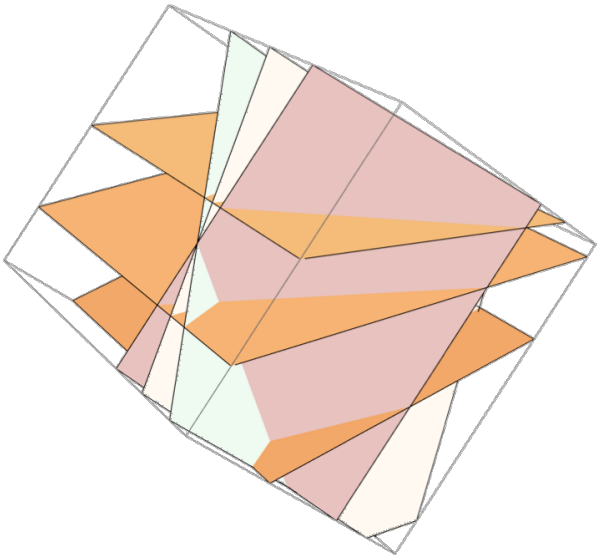


Figure 4.4: The six planes in space from Example 4.3.14. The three near-vertical planes intersect in a line and the same holds for the three near-horizontal planes.

Example 4.3.14. Consider the arrangement of six planes in $\mathbb{R}^3$ displayed in Figure 4.4, given by the matrix

$$L = [b \mid A]^\mathsf{T} = \left[\begin{array}{cccccc} 1 & 2 & 1 & 0 & 0 & 0 \\ \hline 1 & 1 & 2 & 1 & 0 & 1 \\ 1 & \frac{3}{2} & \frac{3}{2} & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right].$$

The first three planes intersect in the line $\mathbb{V}(x-1, y+2) = (1,-2,0) + \mathbb{R}(0,0,1)$, the last three planes intersect in the line $\mathbb{V}(x+z, y+z) = \mathbb{R}(1,1,-1)$, and there are no other dependencies among the planes. A computation shows that

$$\begin{aligned} \nabla_{\log}(X) = \; & \mathbb{V}(144u_0^2 + 120u_0u_1 + 168u_0u_2 + 25u_1^2 - 70u_1u_2 + 49u_2^2) \\ & \cup \mathbb{V}(u_3^2 - 2u_3u_4 + 4u_3u_5 + u_4^2 + 4u_4u_5 + 4u_5^2) \\ & \cup \mathbb{V}(u_0 + u_1 + u_2, u_3 + u_4 + u_5). \end{aligned}$$

The codimension 1 part of $\nabla_{\log}(X)$ is reducible and there is an additional linear codimension 2 component. In particular, $\nabla_{\log}(X)$ is not a hypersurface and therefore cannot agree with $\nabla_{\mathrm{Hu}}(X)$ (not even as sets). For general u in the codimension 2 component, $\mathrm{Crit}_X(u)$ has dimension 1. ◁

4.4 The Hurwitz discriminant

In this section we relate the likelihood correspondence and the logarithmic discriminant to the projective geometry of the variety $\mathcal{R}_L$ parametrized by the reciprocals $(1/\ell_i)_{i=0}^n$. The Hurwitz form of this variety is closely related to the logarithmic discriminant. As an application, we show that the logarithmic discriminant avoids the positive orthant under mild genericity conditions.

4.4.1 Reciprocal linear spaces

Reformulating the equation $A^\mathsf{T} \cdot \mathrm{diag}(1/\ell_0(x), \dots, 1/\ell_n(x)) \cdot u = 0$ from (4.1), we see that for $u \in \mathbb{P}^n$ and $x \in X = \mathbb{C}^d \setminus \mathcal{A}$,

$$x \in \mathrm{Crit}_X(u) \qquad \text{if and only if} \qquad (u_0/\ell_0(x), \dots, u_n/\ell_n(x))^\mathsf{T} \in \mathrm{Ker}(A^\mathsf{T}).$$

Let $\mathbb{T} = \mathbb{T}(\mathbb{P}^n) \coloneqq \mathbb{P}^n \setminus \mathbb{V}(u_0 \cdots u_n)$ and consider the locally closed embeddings

$$\begin{aligned} \gamma\colon \mathbb{C}^d \setminus \mathcal{A} &\to \mathbb{P}^n, & x &\mapsto (1/\ell_0(x) : \cdots : 1/\ell_n(x)) \\ \varphi\colon \mathbb{T}(\mathbb{P}^n) &\to \mathbb{G}(n-d, \mathbb{P}^n), & u &\mapsto \mathrm{Ker}(A^\mathsf{T} \,\mathrm{diag}(u_0, \dots, u_n)). \end{aligned}$$

With this notation, the critical equations become *linear equations* on $\mathrm{Im}(\gamma)$ imposed by the linear subspace $\varphi(u)$:

Lemma 4.4.1. (i) *$x \in \mathrm{Crit}_X(u)$ if and only if $\gamma(x) \in \varphi(u) \cap \mathrm{Im}(\gamma)$.*

(ii) *Let $f\colon \mathcal{L}_X \to \mathbb{P}^n$ be as in Theorem 4.2.1 and denote by $\mathcal{L}_X|_{\mathbb{T}}$ the restriction of $\mathcal{L}_X \cap (\mathbb{T} \times X) \subseteq \mathbb{P}^n \times X$. Consider the incidence of complementary linear spaces intersecting $\mathrm{Im}(\gamma)$*

$$\mathcal{I}^\circ = \{\, ([\Lambda], y) \mid y \in \Lambda \cap \mathrm{Im}(\gamma) \,\} \subseteq \mathbb{G}(n-d, \mathbb{P}^n) \times \mathrm{Im}(\gamma).$$

The following commutative diagram is cartesian:

$$\begin{array}{ccc} \mathcal{L}_X|_{\mathbb{T}} & \xrightarrow{\gamma \times \varphi} & \mathcal{I}^\circ \\ {\scriptstyle f}\downarrow & \lrcorner & \downarrow{\scriptstyle \mathrm{pr}_\Lambda} \\ \mathbb{T} & \xrightarrow{\varphi} & \mathbb{G}(n-d, \mathbb{P}^n). \end{array} \tag{4.5}$$

(iii) *If all $n+1$ $(n \times d)$-submatrices of A have full rank d, then φ is defined on $D \coloneqq \mathbb{P}^n \setminus \bigcup_{i<j} \mathbb{V}(u_i, u_j)$ and the diagram is also cartesian with D in place of $\mathbb{T}$.*

Proof. The discussion preceding the lemma shows (i). (ii) is a careful restatement of (i), stating that the incidence $\mathcal{I}^\circ$ pulls back to the likelihood equations. Imposing the rank condition on A, we see that $A^\mathsf{T}\operatorname{diag}(u_0,\dots,u_n)$ still has rank d, so φ is well-defined on $D \supseteq \mathbb{T}$. The argument about being cartesian is the same as before, proving (iii). □

This suggests that points in the logarithmic discriminant correspond to linear spaces intersecting $\operatorname{Im}(\gamma)$ non-transversally. The image closure $\mathcal{R}_L \coloneqq \overline{\operatorname{Im}(\gamma)}$ is the *reciprocal linear space* of the matrix $L = [\,b \mid A\,]^\mathsf{T}$.

Definition 4.4.2. Let $L \in \mathbb{C}^{(d+1)\times(n+1)}$ be a matrix of full rank $d+1$. The variety

$$\mathcal{R}_L = \overline{\left\{ [v_0^{-1} : \cdots : v_n^{-1}] \;\middle|\; v \in \operatorname{Rowspan}(L) \cap (\mathbb{C}^\times)^{n+1} \right\}} \subseteq \mathbb{P}^n$$

is the *reciprocal linear space* of L.

The branch locus of $\operatorname{pr}_\Lambda \colon \mathcal{I} \to \mathbb{G}(n-d,\mathbb{P}^n)$ is the first associated hypersurface $\mathcal{Z}_1(\mathcal{R}_L)$ from Section 4.1.3. This irreducible and reduced hypersurface is defined by the Hurwitz form in the Plücker ring $\operatorname{Hu}_{\mathcal{R}_L} \in S(\mathbb{G}(n-d,\mathbb{P}^n))$. We summarize the important properties of $\mathcal{R}_L$ for our investigations in the following theorem, for which we need some notation:

A *circuit* of $L = [l_0|\dots|l_n]$ is a subset $C \subseteq [n+1]$ such that $\{l_c\}_{c\in C}$ is a minimal dependent set. If $0 \neq a \in \mathbb{C}^C$ is its linear relation $(\sum_{c\in C} a_i l_i = 0)$, then define the polynomial (up to scaling)

$$f_C \coloneqq \sum_{c\in C} a_c y^{C\setminus\{c\}} \in \mathrm{I}(\mathcal{R}_L) \subseteq \mathbb{C}[y_0,\dots,y_n].$$

Denote by $\mathcal{C} \subseteq \mathbb{C}^{d+1}$ the central arrangement defined by the homogeneous forms $l_i^\mathsf{T} \cdot x$.

Theorem 4.4.3. *Let $L \in \mathbb{C}^{(d+1)\times(n+1)}$ be a matrix of full rank.*

(i) [PS06, Theorem 4] *$\mathcal{R}_L \subseteq \mathbb{P}^n$ is a projectively Cohen–Macaulay variety. A universal Gröbner basis of $\mathrm{I}(\mathcal{R}_L)$ is given by $\{\, f_C \mid C$ circuit of $L \,\}$.*

(ii) [PS06, Lemma 2] $\deg \mathcal{R}_L = (-1)^{d+1}\chi_{\mathcal{C}}(0)$.

(iii) [SSV13, Theorem 2] $\deg \operatorname{Hu}_{\mathcal{R}_L} = 2(-1)^{d+1}((d+1)\chi_{\mathcal{C}}(0) + \chi'_{\mathcal{C}}(0))$.

Example 4.4.4 (Uniform linear spaces)**.** If $L \in \mathbb{C}^{(d+1)\times(n+1)}$ is uniform, then the characteristic polynomial satisfies $\chi_{\mathcal{C}}(0) = (-1)^{d+1}\binom{n}{d}$ and $\chi'_{\mathcal{C}}(0) = (-1)^d\binom{n+1}{d}$, hence

$$\deg \mathcal{R}_L = \binom{n}{d}, \qquad \deg \operatorname{Hu}_{\mathcal{R}_L} = 2\left((d+1)\binom{n}{d} - \binom{n+1}{d}\right) = 2(n-d)\binom{n}{d-1}. \quad \triangleleft$$

Definition 4.4.5 (Hurwitz discriminant)**.** The closure of the pullback of $\mathcal{Z}_1(\mathcal{R}_L)$ along $\varphi \colon \mathbb{P}^n \dashrightarrow \mathbb{G}(n-d,\mathbb{P}^n)$ is the *Hurwitz discriminant* $\nabla_{\mathrm{Hu}}(X) \coloneqq \varphi^{-1}(\mathcal{Z}_1(\mathcal{R}_L))$.

This is well-defined, since the complement of the maximal domain of definition of φ has codimension at least 2, so the hypersurface $\nabla_{\mathrm{Hu}}(X) \subseteq \mathbb{P}^n$ is uniquely determined.

Remark 4.4.6. On $\mathbb{T}$, one obtains the defining equation for ∇_{Hu} from that of $\mathcal{Z}_1(\mathcal{R}_L)$ by substituting $p_I = \det(A_I^\perp) \cdot (u^I)^{-1}$. Here, for $I \in \binom{[n+1]}{n-d+1}$, the p_I are the Plücker coordinates on $\mathbb{G}(n-d, \mathbb{P}^n)$ and $A^\perp$ represents $\ker(A^\mathsf{T})$, i.e., $\mathrm{Im}(A^\perp) = \ker(A^\mathsf{T})$. The matrix $A_I^\perp$ consists of the rows of $A^\perp$ indexed by I. ◁

Whenever $\nabla_{\mathrm{Hu}}(X)$ is neither empty nor all of $\mathbb{P}^n$, it is necessarily a hypersurface. The same need not be true for $\nabla_{\log}(X)$, as Example 4.3.14 shows; in particular $\nabla_{\log} \neq \nabla_{\mathrm{Hu}}$ in that case. Nonetheless, Lemma 4.4.1 indicates that the two should be closely related, and indeed we always have the following:

Proposition 4.4.7. *We have the inclusion $\nabla_{\log}(X) \cap \mathbb{T} \subseteq \nabla_{\mathrm{Hu}}(X) \cap \mathbb{T}$ of schemes. If all $(n \times d)$ of A have full rank d, then this also holds on D (notation as in Lemma 4.4.1).*

Proof. Apply Lemma 4.1.5 to the cartesian diagrams from Lemma 4.4.1. □

We will use this in Section 4.5.3 below to show that for any bi-uniform arrangement, the Hurwitz discriminant coincides with the logarithmic discriminant.

4.4.2 Positivity of the logarithmic discriminant

As a first application of the Hurwitz discriminant construction, we show positivity of the logarithmic discriminant for real arrangements. Our construction relies on the following property of the projective closure of the arrangement.

Definition 4.4.8. Let $\mathcal{A} \subseteq \mathbb{K}^n$ be an arrangement and $\overline{\mathcal{A}} \subseteq \mathbb{P}^n(\mathbb{K})$ the arrangement of the projective closures of the H_i. The *flats at infinity* are the intersections of the projective hyperplanes in $\overline{\mathcal{A}}$ which are contained in the hyperplane at infinity $\mathbb{P}^n(\mathbb{K}) \setminus \mathbb{K}^n$.

Theorem 4.4.9. *If $\mathcal{A}$ has no flats at infinity and if $u \in \mathbb{T}$ is such that $\mathrm{Crit}_X(u)$ consists of $r = \mathrm{MLdeg}(X)$ reduced points, then $u \notin \nabla_{\log}(X)$.*

Our proof relies on the following result which relates the ML degree of $X = \mathbb{C}^d \setminus \mathcal{A}$ to the degree of the reciprocal linear space.

Proposition 4.4.10 ([BBT24])**.** $\mathrm{MLdeg}(X) \leq \deg \mathcal{R}_{\mathcal{L}}$*, with equality if and only if $\mathcal{A}$ has no flats at infinity.*

Proof idea. If $f\colon \mathcal{L}_X \to \mathbb{P}^n$ is the projection, then $\mathrm{MLdeg}(X) = \#f^{-1}(u)$ for general $u \in \mathbb{T}$. Let $[\Lambda] = \varphi(u) = \mathrm{Ker}(A^\mathsf{T} \operatorname{diag}(u_0, \ldots, u_n))$, then by Lemma 4.4.1

$$\gamma(f^{-1}(u)) = \mathrm{Im}(\gamma) \cap \Lambda \subseteq \mathcal{R}_L \cap \Lambda.$$

Applying Bézout's inequality (Corollary 0.3.3) shows $\mathrm{MLdeg}(X) \leq \deg \mathcal{R}_L$.

One can show that the intersection $\Lambda \cap \mathcal{R}_L$ is finite for general u [BBT24, Theorem 3.5]. Intersection points in the boundary $\mathcal{R}_L \setminus \mathrm{Im}(\gamma)$ appear if and only if $\mathcal{A}$ has flats at infinity. For details see [BBT24, Corollary 3.9]. □

Example 4.4.11. A bi-uniform arrangement has no flats at infinity. This can be verified either from the definition, or by observing that $\mathrm{MLdeg}(X) = \binom{n}{d} = \deg \mathcal{R}_L$ by comparing Examples 4.3.6 and 4.4.4. Note that there exist arrangements with non-uniform matrix A which still have no flats at infinity, for example the arrangement from Example 4.3.14. ◁

Proof of Theorem 4.4.9. It suffices to show that the set of $u \in \mathbb{C}^{n+1}$ with exactly $r := \mathrm{MLdeg}(X)$ non-degenerate critical points is open. By the implicit function theorem, perturbing u by a small $\varepsilon \in \mathbb{C}^{n+1}$ has the effect that

$$\mathrm{Crit}_X(u+\varepsilon) = \{\xi_1(\varepsilon), \dots, \xi_r(\varepsilon)\} \cup Z$$

where $\xi_1(\varepsilon), \dots, \xi_r(\varepsilon)$ are the distinct perturbed critical points and Z is a union of (potentially positive-dimensional) components not containing the ξ_i. As $\mathcal{A}$ has no flats at infinity, we have $r = \deg \mathcal{R}_L$, and hence we must have $\mathcal{R}_L \cap \varphi(u) = \gamma(\{\xi_1(\varepsilon), \dots, \xi_r(\varepsilon)\})$ by Bézout's Inequality 0.3.3. This shows $Z = \emptyset$. □

The argument shows in fact that $\nabla_{\mathrm{Hu}} \cap \mathbb{R}^{n+1}_+ = \emptyset$ (which implies the statement about the logarithmic discriminant by Proposition 4.4.7).

Corollary 4.4.12 (Positivity of the discriminant)**.** *If $\mathcal{A}$ is a real arrangement that has no flats at infinity, then $\nabla_{\log} \cap \mathbb{R}^{n+1}_+ = \emptyset$.*

Proof. Combine Theorem 4.4.9 with Varchenko's Theorem 4.3.8. □

The following example shows that the assumption of having no flats at infinity is not always necessary, and we conjecture that Corollary 4.4.12 holds for all real arrangement.

Example 4.4.13 (Continuing Example 4.3.12)**.** The arrangement $\mathcal{M}_{0,5}$ has MLdeg 2, but the corresponding reciprocal linear space has degree 4. Indeed, the parallel lines introduce two flats at infinity (compare Figure 4.2).

The defining polynomial $\Delta_{\log}(\mathcal{M}_{0,5})$ from Example 4.3.12 is positive on $\mathbb{R}^5_+$: The inequality of arithmetic and geometric mean shows that $(u_0u_3 + u_1u_2)^2 \geq 4u_1u_2u_3u_4$, so $\Delta_{\log}(\mathcal{M}_{0,5}) \geq (u_0+u_1+u_2+u_3+u_4)^2u_4^2 > 0$. It is, however, not positive on $\mathbb{R}^5$, for example $\Delta_{\log}(-\frac{1}{2}, 1, 2, -\frac{1}{2}, -1) = -\frac{7}{16}$. ◁

4.5 Bi-uniform arrangements

In this section we turn to studying the logarithmic discriminant of sufficiently uniform hyperplane arrangements. In the simplest case of $n+1$ points on the affine line (which are always bi-uniform), we can give a complete answer about degree, irreducibility and the connection to the Hurwitz form. We then give a general irreducibility criterion, depending on the arrangement containing a bi-uniform subarrangement of sufficient size. Afterwards, we study the Hurwitz discriminant for a bi-uniform arrangement, calculate its degree and show that it set-theoretically agrees with the logarithmic discriminant. This concludes the proof of Main Results 4.

4.5.1 Arrangements of points on the line

We start in the smallest interesting case $d = 1$, that is, arrangements of $n+1 \geq 3$ distinct points $\mathcal{A}$ on the affine line $\mathbb{C}^1$. Here, we are able to provide a complete description; similar statements in higher dimensions may not hold or are harder to prove. Here and later, $f\colon \mathcal{L}_X \to \mathbb{P}^n$ denotes the projection from the likelihood correspondence.

Theorem 4.5.1. *The discriminant $\nabla_{\log}(L) \subseteq \mathbb{P}^n$ of an arrangement of $n+1 \geq 3$ points in $\mathbb{C}^1$ is an irreducible reduced hypersurface of degree $2(n-1)$. The class of $\operatorname{Ram}(f)$ in $\mathrm{A}(\mathbb{P}^n \times \mathbb{P}^1) = \mathbb{Z}[\alpha,\beta]/\langle \alpha^{n+1}, \beta^2\rangle$ equals $\alpha^2 + 2(n-1)\alpha\beta$.*

Without loss of generality we can assume $\ell_i(x) = x + b_i$ (so $X = \mathbb{C}^1 \setminus \{-b_0, \dots, -b_n\}$). The ramification locus is defined in $\mathbb{P}^n \times X$ by the two equations

$$\begin{aligned} 0 = g_1(u;x) &= \frac{u_0}{x+b_0} + \cdots + \frac{u_n}{x+b_n}, \\ 0 = g_2(u;x) &= -\frac{\partial g_1}{\partial x} = \frac{u_0}{(x+b_0)^2} + \cdots + \frac{u_n}{(x+b_n)^2}. \end{aligned}$$

Lemma 4.5.2. *The ramification locus $\operatorname{Ram}(f)$ is a smooth $(n-1)$-dimensional irreducible variety. The projection $f\colon \operatorname{Ram}(f) \to \operatorname{Branch}(f)$ is birational. Hence, $\nabla_{\log} \subseteq \mathbb{P}^n$ is an irreducible and reduced hypersurface.*

Proof. The cone over $\operatorname{Ram}(f)$ in $\mathbb{C}^{n+1} \times (\mathbb{C}^1 \setminus \{-b_0, \dots, -b_n\})$ is defined by $\langle g_1, g_2\rangle \subseteq R[u_0, \dots, u_n; x]_{\ell_0 \cdots \ell_n}$. Substituting u_0 and u_1 using g_1 and g_2, we see

$$\frac{\mathbb{C}[u_0, \dots, u_n; x]_{\ell_0 \cdots \ell_n}}{\langle g_1, g_2\rangle} \cong \frac{\mathbb{C}[u_1, \dots, u_n; x]_{\ell_0 \cdots \ell_n}}{\langle \tilde{g}_2\rangle} \cong \mathbb{C}[u_2, \dots, u_n; x]_{\ell_0 \cdots \ell_n},$$

where $\tilde{g}_2 = g_2$ after substituting u_0. This shows smoothness and irreducibility of $\operatorname{Ram}(f)$.

Since $\dim X = 1$, the projection is generically finite. It remains to show generic injectivity. For this, it suffices to show that the variety

$$\mathcal{B}_{n+1} := \left\{ (u,x,y) \in \mathbb{P}^n \times X \times X \;\middle|\; \begin{array}{c} x \neq y, \\ g_1(u;x) = g_2(u;x) = g_1(u;y) = g_2(u;y) = 0 \end{array} \right\}$$

has dimension at most $n-2$, as it then can not map dominantly to $\nabla_{\log}$. In fact, $\mathcal{B}_{n+1}$ is a $\mathbb{P}^{n-4}$-bundle over the two-dimensional irreducible variety $\{\, (x,y) \in X \times X \mid x \neq y \,\}$ for $n \geq 4$. To see this, we need to show that the coefficient matrix B_{n+1} of $g_1(-;x) = g_2(-;x) = g_1(-;y) = g_2(-;y) = 0$ has full rank 4 always. The matrix B_4 is a submatrix of B_{n+1} for all $n \geq 3$, so it suffices to show that

$$B_4 := \begin{bmatrix} \frac{1}{x} & \frac{1}{x+1} & \frac{1}{x+b_2} & \frac{1}{x+b_3} \\ \frac{1}{x^2} & \frac{1}{(x+1)^2} & \frac{1}{(x+b_2)^2} & \frac{1}{(x+b_3)^2} \\ \frac{1}{y} & \frac{1}{y+1} & \frac{1}{y+b_2} & \frac{1}{y+b_3} \\ \frac{1}{y^2} & \frac{1}{(y+1)^2} & \frac{1}{(y+b_2)^2} & \frac{1}{(y+b_3)^2} \end{bmatrix}$$

has rank 4. Computing the determinant symbolically confirms this:

$$\det C_4 = -\frac{b_2\,(b_2-1)\,b_3\,(b_3-1)\,(b_2-b_3)\,(x-y)^4}{y^2\,(y+1)^2\,(y+b_3)^2\,(y+b_2)^2\,x^2\,(x+1)^2\,(x+b_3)^2\,(x+b_2)^2}.$$

This establishes $\dim \mathcal{B} = n-2$ (for $n = 2, 3$, $\mathcal{B}$ is empty). □

This lemma allows us to compute the cohomology class of the ramification divisor in $\mathrm{A}(\mathbb{P}^n \times \mathbb{P}^1)$ and its pushforward using standard intersection theory.

Proof of Theorem 4.5.1. We first study the intersection of g_1 and g_2 in $\mathbb{P}^n \times \mathbb{P}^1$. Clearing denominators yields the bihomogeneous equations

$$g_1^{\mathrm{h}}(u; x_0, x_1) = \sum_{i=0}^{n} u_i \prod_{k\neq i}(x_1 + b_k x_0), \qquad g_2^{\mathrm{h}}(u; x_0, x_1) = \sum_{i=0}^{n} u_i \prod_{k\neq i}(x_1 + b_k x_0)^2.$$

Let $W_i = \mathbb{V}(g_i^{\mathrm{h}}) \subseteq \mathbb{P}^n \times \mathbb{P}^1$. In the boundary $\mathbb{P}^n \times \{-b_0, \dots, -b_n, \infty\} \subseteq \mathbb{P}^n \times \mathbb{P}^1$ we have

$$\langle g_1^{\mathrm{h}}, g_2^{\mathrm{h}} \rangle = \begin{cases} \langle u_i, x_1 + b_i x_0 \rangle & \text{in } \mathcal{O}_{\mathbb{P}^n\times\mathbb{P}^1, \mathbb{V}(u_i)\times\{-b_i\}}, \\ \langle u_0 + \cdots + u_n, x_0 \rangle & \text{in } \mathcal{O}_{\mathbb{P}^n\times\mathbb{P}^1, \mathbb{V}(u_0+\cdots+u_n)\times\{\infty\}}. \end{cases}$$

This gives the decomposition into irreducible reduced components

$$W_1 \cap W_2 = \mathrm{Ram}(f) \cup \left(\bigcup_{i=0}^{n} \mathbb{V}(u_i) \times \{-b_i\} \right) \cup \Big(\mathbb{V}(u_0 + \cdots + u_n) \times \{\infty\} \Big).$$

Each of the $n+2$ terms on the right has class $\alpha\beta$. We also have

$$[W_1] = \alpha + n\beta, \quad [W_2] = \alpha + 2n\beta, \quad [W_1 \cap W_2] = \alpha^2 + 3n\alpha\beta.$$

Combining the above displays we find the cycle class of $\mathrm{Ram}(f)$:

$$[\mathrm{Ram}(f)] = \alpha^2 + (3n - (n+2))\alpha\beta = \alpha^2 + 2(n-1)\alpha\beta.$$

Pushing forward via $\mathbb{P}^n \times \mathbb{P}^1 \to \mathbb{P}^n$, we obtain

$$\deg \nabla_{\log} = \frac{\text{coeff. of } \alpha\beta}{\deg(f|_{\mathrm{Ram}(f)})} \overset{4.5.2}{=} 2(n-1).$$ □

Corollary 4.5.3. *The logarithmic discriminant defines a flat family of hypersurfaces parametrized by $\mathcal{M}_{0,n+2}$*

$$\left\{ (u, \mathcal{A}) \in \mathbb{P}^n \times \mathcal{M}_{0,n+2} \;\middle|\; u \in \nabla_{\log}(\mathbb{C}^1 \setminus \mathcal{A}) \right\} \to \mathcal{M}_{0,n+2}.$$

It can be computed using the formulae

$$\Delta_{\log}(X) = \mathrm{Disc}_x(g_1^{\mathrm{h}}) = \frac{1}{u_0 \cdots u_n (u_0 + \cdots + u_n)} \mathrm{Res}_x(g_1^{\mathrm{h}}, g_2^{\mathrm{h}}).$$

Proof. The fibers of this family are hypersurfaces of constant degree $2(n-1)$ by Theorem 4.5.1, hence it is a flat family. It follows from Definition Theorem 4.1.3 that $\nabla_{\log}(X)$ is contained in $\mathbb{V}(\mathrm{Disc}_x(g_1^{\mathrm{h}}))$ and $\mathbb{V}(\mathrm{Res}_x(g_1^{\mathrm{h}}, g_2^{\mathrm{h}}))$. On the other hand, the discriminant of a degree n polynomial has degree $2(n-1)$ in its coefficients, so $\Delta_{\log}(X)$ and $\mathrm{disc}_x(g_1^{\mathrm{h}})$ must agree. Similarly, the resultant of g_1^{h} and g_2^{h} has degree $3n$ and splits off the given $n+2$ factors, so it also coincides with the logarithmic discriminant. □

The discriminant in Example 4.3.10 is smooth for any point configuration. This fails for $n+1 > 3$, as a computation of the discriminant of $\Delta_{\log}(X)$ shows:

Proposition 4.5.4. *For any arrangement of four distinct points in $\mathbb{C}^1$ the logarithmic discriminant is a singular quartic surface in $\mathbb{P}^3$.*

Proof. Corollary 4.5.3 allows us to compute a closed form expression of $\Delta_{\log}(X)$ as a polynomial in $u_0, \dots, u_3$ with coefficients in $\mathbb{C}[b_0, \dots, b_3]$. A computation with `Macaulay2` shows that the discriminant of this quartic vanishes identically, so for any choice of b_i the logarithmic discriminant is singular. □

Proposition 4.5.5. *In the case $d = 1$, the Hurwitz discriminant and logarithmic discriminant coincide.*

Proof. In the notation from Section 4.4, the map $\varphi\colon \mathbb{T}(\mathbb{P}^n_u) \to \mathrm{Gr}(n-1, \mathbb{P}^n) \cong (\mathbb{P}^n)^\vee$ given by $u \mapsto \mathrm{Ker}((1, \dots, 1)\,\mathrm{diag}(u_0, \dots, u_0)) = [u^\perp]$ is an isomorphism. Therefore, the likelihood correspondence is isomorphic to the Hurwitz incidence in Lemma 4.4.1. □

This yields an alternative derivation of $\deg \nabla_{\log} = 2(n-1)$ via Example 4.4.4.

4.5.2 Irreducibility

In this section we prove a sufficient criterion for the logarithmic discriminant to be an irreducible variety.

Theorem 4.5.6. *Let $\mathcal{A} \subseteq \mathbb{C}^d$ be an arrangement of $n+1 \geq d+2$ hyperplanes containing a bi-uniform subarrangement of $d+2$ hyperplanes. Then $\mathrm{Ram}(\mathcal{L}_X^\circ \to \mathbb{P}^n)$ and $\nabla_{\log}(X)$ are reduced and irreducible.*

For arrangements of lines in the plane, the theorem says that $\nabla_{\log}$ is irreducible as soon as it contains a bi-uniform arrangement of four lines such as shown in Figure 4.1.

We will prove that the ramification locus is irreducible and reduced, defined by a single irreducible polynomial on $\mathcal{L}_X$. Without loss of generality, we assume that the submatrix of the first $(d+2)$ columns of L satisfies the hypotheses. For convenience we may also assume that the leftmost $d \times d$ submatrix of A^T is the identity matrix.

The defining equations of $\mathrm{Ram}(f) \subseteq \mathbb{P}^n \times X$ are given by the polynomials in Lemma 4.3.3 and Proposition 4.3.11. By making the reversible substitution $v_i := u_i/\ell_i(x)$, the ramification scheme is isomorphic to the subscheme defined by

$$A^\mathsf{T} v = 0, \qquad \det\left(A^\mathsf{T}\,\mathrm{diag}(v_0/\ell_0, \dots, v_n/\ell_n)\,A\right) = \sum_{I \subseteq [n+1],\, |I|=d} |A_I|^2 \frac{v^I}{\ell^I} = 0.$$

By assumption the linear equations $A^\mathsf{T} v = 0$ have the form $v_j = \sum_{i=d}^{n} -a_{ij} v_i$ for $j = 0, \dots, d-1$. Linearly substituting $v_0, \dots, v_{d-1}$ and clearing denominators by multiplying with $\ell_0 \cdots \ell_n$ yields the polynomial $\tilde{h}$ given by

$$\tilde{h} = \sum_{I \subseteq [n+1],\, |I| = d} |A_I|^2 \, \tilde{v}^I \ell^{[n+1]\setminus I}, \qquad \tilde{v} = (-\sum_{i=d}^{n} a_{i,0} v_i, \dots, -\sum_{i=d}^{n} a_{i,(d-1)} v_i, v_d, \dots, v_n). \tag{4.6}$$

The discussion so far shows that $\mathbb{V}(\tilde{h}) \subseteq \mathbb{P}^{n-d} \times X$ is isomorphic to $\mathrm{Ram}(f)$. Since $\mathbb{C}[v_d, \dots, v_n, x_1, \dots, x_d]_{\ell_0 \cdots \ell_n}$ is a UFD, Theorem 4.5.6 is equivalent to the following:

Proposition 4.5.7. *The polynomial $\tilde{h} \in \mathbb{C}[v_d, \dots, v_n; x_1, \dots, x_d]_{\ell_0 \cdots \ell_n}$ is irreducible.*

For $n+1 = d+2$ we are going to show the (slightly stronger) claim of irreducibility in $\mathbb{C}[v_d, v_{d+1}; x_1, \dots, x_d]$.

Proof. We first reduce to the case $n+1 = d+2$ as follows: $\tilde{h}$ is homogeneous in v, and hence if $\tilde{h} = p \cdot q$ factors, then these factors are necessarily homogeneous too. Thus if there is a factorization into non-units, then setting $v_{d+2}, \dots, v_n$ to zero still yields a non-trivial decomposition of $\tilde{h}' \coloneqq \tilde{h}(v_d, v_{d+1}, 0, \dots, 0; \underline{x})$. Inspecting (4.6), we see that $\tilde{h}'$ is the $\tilde{h}$ of the sub-arrangement of the first $d+2$ hyperplanes multiplied by the unit $\ell_{d+2} \cdots \ell_n$. Hence, it suffices to consider $n+1 = d+2$.

Let $\tilde{h}^\mathrm{h}(v_d, v_{d+1}; x_0, \dots, x_d)$ be the homogenization of $\tilde{h}$ with respect to the new variable x_0. We also have the "naive" homogenization $N(v_d, v_{d+1}; x_0, \dots, x_d)$ obtained from (4.6) in which each ℓ_i is replaced by its homogenization $\ell_i^\mathrm{h} = b_i x_0 + a_{i,1} x_1 + \cdots + a_{i,d} x_d$. N has x-degree $n+1-d = 2$ and must be of the form $N = \tilde{h}^\mathrm{h} \cdot x_0^e$ for some $e \geq 0$. On the other hand, substituting $x_0 = 0$ in N yields identically zero:

$$\begin{aligned} &A^\mathsf{T} \operatorname{diag}(v) \operatorname{diag}(\ell_0^\mathrm{h}(0, x), \dots, \ell_{d+1}^\mathrm{h}(0, x))^{-1} A x \\ &= A^\mathsf{T} \operatorname{diag}(v) (1, 1, \dots, 1)^\mathsf{T} \;=\; A^\mathsf{T} v \;=\; 0. \end{aligned} \tag{4.7}$$

This shows $e \geq 1$ and hence $\tilde{h}^\mathrm{h}$ has x-degree ≤ 1.

By the previous discussion $\tilde{h}^\mathrm{h}$ is bihomogeneous of bidegree $(d, 1)$ or $(d, 0)$. In both cases any non-trivial factorization must involve $q \in \mathbb{C}[v_{d+1}, v_{d+2}]$, which has a non-zero root $(\lambda_0, \lambda_1) \in \mathbb{C}^2 \setminus 0$. We show that this leads to a contradiction.

The polynomial $\tilde{h}^\mathrm{h}(\lambda_0, \lambda_1; x_0, \dots, x_d)$ vanishes identically. Since the matrix L is uniform, the set of all pairwise products $(\ell^\mathrm{h})^J$, $|J| = 2$, are linearly independent (Lemma 4.5.8 below). From this we conclude that all coefficients of $\tilde{h}^\mathrm{h}(\lambda_0, \lambda_1; x_0, \dots, x_d)$ in front of $(\ell^\mathrm{h})^J$ must vanish identically. From (4.6) we see that these coefficients are

$$\mathrm{coeff}[(\ell^\mathrm{h})^J](s) = (\det A_{[d+2]\setminus J})^2 \cdot w^{[d+2]\setminus J}, \qquad w \coloneqq \tilde{v}|_{(v_d, v_{d+1}) = (\lambda_0, \lambda_1)} \in \mathbb{C}^{d+2}.$$

Since A^T is uniform, we conclude that all products of $d-1$ entries of w vanish, so w has at least three zero entries. But then $w = 0$, because $Aw = 0$ and no $d-1$ columns of A^T are dependent. This is a contradiction to $(w_d, w_{d+1}) = \lambda \neq 0$. □

Lemma 4.5.8. *Let $S \coloneqq \mathbb{C}[x_0, \dots, x_d]$ and consider linear forms $\{\tilde{\ell}_0, \dots, \tilde{\ell}_{d+1}\} \subseteq S_1$. The set of pairwise products $P \coloneqq \{\, \tilde{\ell}_i \tilde{\ell}_j \mid i < j \,\} \subseteq S_2$ is linearly independent if all subsets of $d+1$ forms of $\{\tilde{\ell}_0, \dots, \tilde{\ell}_{d+1}\}$ are linearly independent.*

Proof. Up the action of $\mathrm{GL}_{d+1}(\mathbb{C})$ on S_1 we may assume $\tilde{\ell}_0 = x_0, \dots, \tilde{\ell}_d = x_d$. Then $\tilde{\ell}_{d+1} = c_0 x_0 + \cdots + c_d x_d$ with all $c_i \neq 0$. It is then straightforward to see that P spans S_2. Since $|P| = \binom{d+2}{2} = \dim_{\mathbb{C}} S_2$, P is linearly independent. □

The sufficient condition of Theorem 4.5.6 is not necessary for irreducibility: An example is the arrangement $\mathcal{M}_{0,5}$ of $n+1 = 5$ lines in $\mathbb{C}^2$ (Example 4.3.12). On the other hand, we have also seen that in degenerate cases $\nabla_{\log}(X)$ may be reducible, even with several components of codimension 1 (Example 4.3.14).

Remark 4.5.9. Under the same assumptions as in Theorem 4.5.6, the Hurwitz ramification scheme $\mathrm{Ram}(\mathrm{pr}_\Lambda \colon \mathcal{I}^\circ \to \mathbb{G}(n-d, \mathbb{P}^n))$ is also reduced and irreducible. Irreducibility follows from standard projective geometry techniques and only uses that $X \cong \mathrm{Im}(\gamma)$ is smooth. Reducedness, on the other hand, is a consequence of Theorem 4.5.6 because the ramification scheme of $f \colon \mathcal{L}_X^\circ \to D$ is the scheme-theoretic preimage of the Hurwitz ramification scheme, and both are locally defined by a single equation. If the equation of the latter was a pure power with exponent > 1, then so would be the equation of the former. ◁

4.5.3 Logarithmic discriminants from Hurwitz forms

This section is devoted to the doubly uniform case, hence for the remainder of the section **both A and L are uniform matrices**. We compute the degree of $\nabla_{\mathrm{Hu}}(X)$ and prove that $\nabla_{\log}(X)$ and $\nabla_{\mathrm{Hu}}(X)$ agree as sets. By the *Newton polytope* of a hypersurface we mean that of its defining equation, see Section 4.6.

Proposition 4.5.10. *Let $\mathcal{A}$ be bi-uniform. Then*

$$\deg \nabla_{\mathrm{Hu}}(X) = 2d \binom{n-1}{d},$$

and the Newton polytope of $\nabla_{\mathrm{Hu}}(X) \subseteq \mathbb{P}^n$ is the full dilated standard simplex.

Proof. Recall from Theorem 4.4.3 that a universal Gröbner basis of $\mathrm{I}(\mathcal{R}_L)$ is given by f_C, C circuit of L. Since L is uniform, *all* subsets $C \in \binom{[n+1]}{d+2}$ are circuits. Let $w \in \mathbb{Z}_{>0}^{n+1}$ be a weight vector with distinct integer entries. The w-initial ideal $M \coloneqq \mathrm{in}_w(\mathrm{I}(\mathcal{R}_L))$ is the square-free monomial ideal generated by the w-leading monomials of the generators f_C. Explicitly, we have $\mathrm{in}_w(f_C) = y^{C \setminus i_C}$, where $i_C = \mathrm{argmin}_{c \in C} w_c$.

Let $i_w \in [n+1]$ be such that $w_{i_w} = \min_i w_i$, then $\mathbb{V}(M) \subseteq \mathbb{A}^{n+1}$ consists of the coordinate subspaces of dimension $d+1$ containing e_{i_w}. For a given coordinate subspace Z of dimension d containing e_{i_w}, there are $\nu(Z) = n+1-d$ such coordinate subspaces containing Z. It follows from Theorem 4.1.13 that

$$\mathrm{in}_w(\mathrm{Hu}_{\mathcal{R}_L}) = \prod_{i_w \in S \in \binom{[n+1]}{d}} p_{[n+1] \setminus S}^{2(n-d)},$$

where $p_{[n+1]\setminus S}$ denotes the Plücker variable corresponding to the complement of S. Now, since A is uniform, this monomial does not vanish at $\operatorname{Ker}(A^\mathsf{T}\operatorname{diag}(u))$ if only u_{i_w} is zero. Therefore, the preimage of $\operatorname{Hu}_{\mathcal{R}_L}$ via φ does not contain $\mathbb{V}(u_{i_w})$.

To compute the preimage on $\mathbb{T}$, we proceed as in Remark 4.4.6 and substitute $p_{[n+1]\setminus S}$ by $\det(A^\perp_{[n+1]\setminus S})\cdot(u^{[n+1]\setminus S})^{-1}$. By uniformity of A, all determinants $\det(A^\perp_{[n+1]\setminus S})$ are nonzero. Hence, we obtain a non-zero scalar multiple of

$$\prod_{i_w\in S}(u^{[n+1]\setminus S})^{-2(n-d)} = (u_0\cdots\widehat{u_{i_w}}\cdots u_n)^{-2(n-d)\binom{n-1}{n-d}} = (u_0\cdots\widehat{u_{i_w}}\cdots u_n)^{-2d\binom{n-1}{d}}.$$

Therefore, multiplying by $(u_0\cdots u_n)^{2d\binom{n-1}{d}}$, we obtain $u_{i_w}^{2d\binom{n-1}{d}}$ as a term in the substituted equation. Since w and hence i_w was arbitary, we see that the equation does not split off any u_i and the Newton polytope is $2d\binom{n-1}{d}\operatorname{Conv}(e_0,\dots,e_n)$. □

Proposition 4.5.11. *Let both L and A be uniform. Then $\nabla_{\log}(X)$ and $\nabla_{\mathrm{Hu}}(X)$ agree set-theoretically. In particular, $\nabla_{\log}(X)$ is a hypersurface.*

To prove Proposition 4.5.11, we need some more notation and a lemma. Let $\mathcal{I}$ be the closure of $\mathcal{I}^\circ$ from Lemma 4.4.1:

$$\mathcal{I} \coloneqq \{\,(\Lambda, y) \mid y\in\Lambda\,\} \subseteq \mathbb{G}(n-d,\mathbb{P}^n)\times\mathcal{R}_L.$$

Furthermore, consider again $D=\mathbb{P}^n\setminus\bigcup_{i<j}\mathbb{V}(u_i,u_j)$ and let Y be the incidence

$$Y \coloneqq \left\{\,(u,y)\;\middle|\; A^\mathsf{T}\operatorname{diag}(y)u=0\,\right\}\subseteq D\times\mathcal{R}_L,$$

These fit into the following diagram, each square being cartesian (compare Lemma 4.4.1)

$$\begin{array}{ccc} Y & \longrightarrow & \mathcal{I} \\ \downarrow & \lrcorner & \downarrow \\ D\times\mathcal{R}_L & \xrightarrow{\varphi\times\mathrm{id}} & \mathbb{G}(n-d,\mathbb{P}^n)\times\mathcal{R}_L \\ \downarrow{\scriptstyle \mathrm{pr}_u} & \lrcorner & \downarrow{\scriptstyle \mathrm{pr}_\Lambda} \\ D & \xrightarrow{\varphi} & \mathbb{G}(n-d,\mathbb{P}^n) \end{array} \tag{4.8}$$

By slight abuse of notation, we will denote by $\mathbb{T}$ also the torus in $\mathbb{P}^n_y$.

Lemma 4.5.12. *$\mathcal{R}_L\setminus\mathbb{T}$ is the union of all coordinate subspaces of codimension $n+1-d$. The image of $\mathcal{I}\cap(\operatorname{Im}(\varphi)\times(\mathcal{R}_L\setminus\mathbb{T}))$ under the projection pr_Γ is $\{[e_0],\dots,[e_n]\}\subset\mathbb{P}^n$. The preimage $\mathrm{pr}_u^{-1}([e_i]) = V_D(u_i)\times\{[e_i]\}$.*

Proof. Since L is uniform, it follows from the explicit description of the generators of the ideal of $\mathcal{R}_L$ (Theorem 4.4.3) that the set-difference $\mathcal{R}_L\setminus\mathbb{T}$ is the union of all coordinate linear subspaces of $\mathbb{P}^n$ of codimension $n+1-d$.

If y is in such a coordinate subspace, at least $n+1-d$ entries of y are zero. Then $A^\mathsf{T}\operatorname{diag}(u)y=0$ implies $\operatorname{diag}(u)y=0$, as A is uniform. So y has at most one non-zero entry since $u\in D$ has at most one zero entry. □

Corollary 4.5.13. *Y is an n-dimensional irreducible reduced variety.*

Proof. Y is smooth of dimension n on the complement $Y \setminus (D \times \{[e_0], \dots, [e_n]\})$ because over $\mathcal{R}_L \cap \mathbb{T}$ it is an open subset of a projective bundle via $\mathrm{pr}_y \colon Y \to \mathcal{R}_L$. Moreover, $\mathrm{pr}_y^{-1}([e_i]) = \mathbb{V}(u_i) \times \{[e_i]\}$ which has dimension $n-1$. Therefore, these fibers cannot be irreducible components since Y is cut out by d equations in $D \times \mathcal{R}_L$ (of dimension $n+d$). We conclude that Y is irreducible and generically reduced. Since $\mathcal{R}_L$ is Cohen–Macaulay (Theorem 4.4.3) and D is smooth, Y is Cohen–Macaulay and thus reduced. □

The advantage of using diagram (4.8) over the "open" version (4.5) is that the vertical arrows here are *proper* maps.

Proof of Proposition 4.5.11. By Remark 4.1.6, we actually have set-theoretic equality

$$\mathrm{pr}_u((\varphi \times \mathrm{id})^{-1}(\mathfrak{R})) = \varphi^{-1}(\mathrm{pr}_\Lambda(\mathfrak{R})) =: \nabla_{\mathrm{Hu}},$$

where $\mathfrak{R} := \overline{\mathrm{Ram}(\mathrm{pr}_\Lambda \colon \mathcal{I}^\circ \to \mathbb{G}(n-d, \mathbb{P}^n))}$. $\mathfrak{R}$ is reduced and irreducible of codimension 1 in $\mathcal{I}$ by Remark 4.5.9. Let Z be the preimage of $\mathfrak{R}$ in Y. Note that $Z|_{\mathcal{L}_X} = \mathrm{Ram}(f \colon \mathcal{L}_X \to \mathbb{P}^n)$. Its projection $\mathrm{pr}_u(Z)$ is ∇_{Hu}, which is a hypersurface by Proposition 4.5.10. If Z has an irreducible component of codimension > 1, then its image in D under pr_u must be in the image of some codimension 1 component of Z. Hence, it suffices to prove that Z does not contain any codimension 1 component of the complement $Y \setminus \mathcal{L}_X^\circ$, which is

$$Y \setminus \mathcal{L}_X^\circ = \mathrm{pr}_y^{-1}([e_0]) \cup \cdots \cup \mathrm{pr}_y^{-1}([e_n]) \cup (Y \cap (D \times \mathcal{R}_{A^\mathsf{T}})).$$

Here, $\mathcal{R}_{A^\mathsf{T}} \subseteq \mathbb{P}_y^n$ is the $(d-1)$-dimensional reciprocal linear space corresponding to the matrix A^T, which appears as the boundary of $\mathrm{Im}(\gamma \colon \mathbb{C}^d \setminus \mathcal{A} \to \mathbb{P}_y^n) \subseteq \mathcal{R}_L$ when moving $x \in \mathbb{C}^n$ to the hyperplane at infinity.

The image of $\mathrm{pr}_y^{-1}([e_i])$ in D is $\mathbb{V}(u_i)$ by the proof of Lemma 4.5.12. On the other hand, by Proposition 4.5.10, $\mathbb{V}(u_i)$ is not a component of ∇_{Hu}.

It remains to consider $Y \cap (D \times \mathcal{R}_{A^\mathsf{T}})$. We may restrict Y and $\mathcal{I}$ to the smooth dense open subvarieties

$$Y' := Y \cap (D \times (\mathcal{R}_L \cap \mathbb{T})), \qquad \mathcal{I}' := \mathcal{I} \cap (\mathbb{G}(n-d, \mathbb{P}^n) \times (\mathcal{R}_L \cap \mathbb{T})).$$

Then $\mathfrak{R} \cap \mathcal{I}'$ is the ramification scheme of $\mathcal{I}' \to \mathbb{G}(n-d, \mathbb{P}^n)$, and its preimage $Z \cap Y'$ is the ramification scheme of $\mathrm{pr}_u \colon Y' \to D$. Hence, we need to prove that $Y' \to D$ is not ramified everywhere along $Y' \cap (D \times \mathcal{R}_{A^\mathsf{T}})$. The latter is an irreducible subvariety of Y' of codimension 1 because it is an open of a projective bundle via pr_y.

We now proceed in a similar fashion as in the proof of Theorem 4.5.6: First identify $\mathcal{R}_L \cap \mathbb{T} \cong \mathbb{P}^d \setminus \overline{\mathcal{A}}$ where $\overline{\mathcal{A}}$ is the closure of $\mathcal{A}$ in $\mathbb{P}^d$. Under this isomorphism, $\mathcal{R}_{A^\mathsf{T}} \cap \mathbb{T}$ corresponds to the hyperplane at infinity $\{x_0 = 0\}$. In the affine open chart where, for instance, $x_1 = 1$, the ramification scheme of $Y' \to D$ is defined by the determinant of the Jacobian matrix of $A^\mathsf{T} \operatorname{diag}(u)(\ell_0(x)^{-1}, \dots, \ell_n(x)^{-1})^\mathsf{T} = 0$ but now with respect to the variables $x_0, x_2, \dots, x_d$. Let A' be the $(n+1) \times d$ submatrix of $L^\mathsf{T} = [b \mid A]$ obtained

by deleting the column corresponding to x_1. After the substitution $v_i = u_i/\ell_i(x)$, the ramification locus is defined by

$$A^\mathsf{T} v = 0, \qquad \det(A^\mathsf{T} \operatorname{diag}(v) \operatorname{diag}(\ell_0(x), \dots, \ell_n(x))^{-1} A') = 0$$

in the chosen chart. We may assume that the first $d \times d$ submatrix of A is the identity matrix. The substituted Hessian $\tilde{h}$ is defined as in the proof of Theorem 4.5.6 and Proposition 4.5.7. It only uses the variables $v_d, v_{d+1}, \dots, v_n$. The difference with Theorem Theorem 4.5.6 is that all occurrences of $\det(A_I)^2$ are now replaced by $\det(A_I)\det(A'_I)$.

Our goal is to show that $\tilde{h}$ is not identically zero, as then $Y' \to D$ is not ramified everywhere along $Y' \cap (D \times \mathcal{R}_{A^\mathsf{T}})$. As in the proof of Proposition 4.5.7, we will write $N(v_d, v_{d+1}, \dots, v_n; x_0, x_1, \dots, x_d)$ for the "naive" homogenization of $\tilde{h}$. The coefficient of v_d^d in N is

$$(-1)^{d-1} \cdot \Big(\prod_{j=0}^{d-1} a_{d,j} \Big) \cdot \Big(\prod_{k=d+1}^{n} \ell_k^{\mathrm{h}}(x) \Big) \cdot \Big(\sum_{i=0}^{d} \det(A_{[d+1]\setminus i}) \det(A'_{[d+1]\setminus i}) \frac{\ell_i^{\mathrm{h}}(x)}{a_{d,i}} \Big),$$

where $a_{d,i}$ denotes the entry of A for $0 \le i \le d-1$ and we set $a_{d,d} \coloneqq -1$. The sum on the right is the only factor that can possibly be zero for $x \in \mathbb{P}^d \setminus \overline{\mathcal{A}}$. It is enough to prove that this sum does not vanish identically after setting $x_0 = 0$.

By Equation (4.7), the sum vanishes identically at $x_0 = 0$ if A' is replaced by A. Note that $\ell_0^{\mathrm{h}}(0, x_1, \dots, x_d), \dots, \ell_{d+1}^{\mathrm{h}}(0, x_1, \dots, x_d)$ form a *circuit* of the matrix A. Hence, the coefficients of a linear relation among them are unique up to scaling. Therefore, if also the above sum vanishes identically after setting $x_0 = 0$, then there is some $\lambda \in \mathbb{C}$ such that $\det(A'_{[d+1]\setminus i}) = \lambda \det(A_{[d+1]\setminus i})$ for all $i = 0, \dots, d$. But this implies that the inverse of the first $(d+1) \times (d+1)$ submatrix of L^T has two linearly dependent columns, which is absurd. □

4.6 The logarithmic discriminant of $\mathcal{M}_{0,m}$

In this final section we focus on the arrangements $\mathcal{A}$ whose complement models the moduli space $\mathbb{C}^{m+3} \setminus \mathcal{A} \cong \mathcal{M}_{0,m}$ of genus zero curves with m distinct marked points (introduced before in Example 4.3.7). $\mathcal{M}_{0,m}$ is represented by a $2 \times m$ matrix of the form (4.2) whose 2-minors are nonzero.

These arrangements are relevant for scattering amplitudes in bi-adjoint scalar ϕ^3-theories, as studied by Cachazo–He–Yuan [CHY14]. It is customary to denote the exponents u by s_{ij} in this context, in such a way that s_{ij} is the exponent standing with the determinant of columns i and j in (4.2). In physics, each column corresponds to a particle with momentum p_i and s_{ij} are *Mandelstam invariants*, obtained by taking the Minkowski inner product of the sum of momentum vectors $p_i + p_j$ with itself. For more information on the physicist's perspective, see [Lam24, Section 3.2].

We start with a complete analysis for the first nontrivial case $m = 5$. We have studied the case of $\mathcal{M}_{0,5}$ already in Examples 4.3.7, 4.3.12 and 4.4.13. The corresponding matrices

are

$$\mathcal{M}_{0,5} = \mathbb{C}^2 \setminus D_1\left(\begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & x_1 & x_2 & 1 \end{bmatrix}\right), \qquad L = \left[\begin{array}{ccccc} 0 & 0 & -1 & -1 & 0 \\ \hline 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 \end{array}\right].$$
$$\begin{matrix} p_1 & p_2 & p_3 & p_4 & p_5 \end{matrix} \qquad \begin{matrix} s_{13} & s_{14} & s_{23} & s_{24} & s_{34} \end{matrix}$$

Using Mandelstam variables $(u_0, u_1, u_2, u_3, u_4) \mapsto (s_{13}, s_{14}, s_{23}, s_{24}, s_{34})$, the polynomial $\Delta_{\log}$ in (4.4) reads

$$\Delta_{\log} = (s_{13}s_{24} + s_{13}s_{34} + s_{14}s_{34} + s_{14}s_{23} + s_{23}s_{34} + s_{24}s_{34} + s_{34}^2)^2 - 4s_{13}s_{14}s_{23}s_{24}.$$

Remark 4.6.1. It was pointed out to us by Sebastian Mizera that in this case, the discriminant has a nice symmetric determinantal expression:

$$\Delta_{\log} = \det \begin{bmatrix} 0 & -s_+ & s_{13} & s_{14} \\ -s_+ & 0 & s_{23} & s_{24} \\ s_{13} & s_{23} & 0 & s_{34} \\ s_{14} & s_{24} & s_{34} & 0 \end{bmatrix} \qquad s_+ \coloneqq s_{13} + s_{14} + s_{23} + s_{24} + s_{34}$$

In physics, this is a principal minor of the 5×5 *Gram matrix* associated to our particles:

$$G = \begin{bmatrix} 0 & s_{12} & s_{13} & s_{14} & s_{15} \\ s_{12} & 0 & s_{23} & s_{24} & s_{25} \\ s_{13} & s_{23} & 0 & s_{34} & s_{35} \\ s_{14} & s_{24} & s_{34} & 0 & s_{45} \\ s_{15} & s_{25} & s_{35} & s_{45} & 0 \end{bmatrix} \in R^{5\times 5}, \quad R = \frac{\mathbb{C}[s_{11}, \ldots, s_{55}]}{\langle s_{ii} = s_{ij} - s_{ji} = \sum_{k=1}^5 s_{ik} = 0 \mid i, j\rangle}$$

Using the given linear relations, all s_{ij} can be expressed in terms of our five exponents, for example $s_{12} = -s_+$. In this way, all five principal 4-minors of G evaluate to $\Delta_{\log}$. ◁

The polynomial $\Delta_{\log}(s_{12}, \ldots, s_{34})$ captures when the following equations have a singular solution

$$\frac{s_{13}}{x_1} + \frac{s_{23}}{x_1 - 1} - \frac{s_{34}}{x_2 - x_1} = \frac{s_{14}}{x_2} + \frac{s_{24}}{x_2 - 1} + \frac{s_{34}}{x_2 - x_1} = 0. \tag{4.9}$$

The discriminant can be obtained algebraically as follows: First, one isolates x_2 in the first equation of (4.9) and substitutes the resulting expression $x_2 = f(x_1)$ into the second equation. This gives a rational function with numerator

$$(s_{13}(x_1 - 1) + s_{23}x_1)\,(a(s)x_1^2 + b(s)x_1 + c(s)) = 0, \qquad a, b, c \in \mathbb{C}[\underline{s}]_2.$$

The factor $(s_{13}(x_1 - 1) + s_{23}x_1)$ is the denominator in the expression $f(x_1)$ found for x_2, so it cannot vanish at a critical point x of $\mathcal{L}_u(x)$. The quadratic discriminant of the second factor $b(s)^2 - 4a(s)c(s)$ equals $\Delta_{\log}$. For $s \in \nabla_{\log}$, the degenerate critical point is

$$x_1 = \frac{-b(s)}{2\,a(s)}, \qquad x_2 = \frac{-b(s)\,(2\,a(s)\,(s_{13} + s_{34}) + b(s)\,(s_{13} + s_{23} + s_{34}))}{2\,a(s)\,(2\,a(s)s_{13} + b(s)(s_{13} + s_{23}))}.$$

The characteristic polynomial of the central arrangement $\mathcal{C}$ defined by L is

$$\chi_{\mathcal{C}}(t) = t^3 - 5t^2 + 8t - 4.$$

According to Theorem 4.4.3, the reciprocal linear space $\mathcal{R}_L \subseteq \mathbb{P}^4$ is a subvariety of dimension 2 and degree 4. Its Hurwitz form is a hypersurface in $\mathbb{G}(2, \mathbb{P}^4)$, represented modulo the Plücker ideal by a homogeneous polynomial of degree $-2(3\chi_{\mathcal{C}}(0) + \chi'_{\mathcal{C}}(0)) = 8$ in $p_{012}, p_{013}, \ldots, p_{234}$. This polynomial has 2285 terms, and can be downloaded at [KKT]. The Hurwitz discriminant from Section 4.4 is obtained by substituting $p_{ijk} = \det(A^{\perp}_{ijk}) \cdot (u_i u_j u_k)^{-1}$, see the discussion in Remark 4.4.6. Performing this substitution, we obtain the reducible polynomial

$$\Delta_{\mathrm{Hu}}(\mathcal{M}_{0,5}) = (s_{13} + s_{23} + s_{34})^2 \cdot (s_{14} + s_{24} + s_{34})^2 \cdot \Delta_{\log}(\mathcal{M}_{0,5}).$$

As we have already observed for $m = 5$, the arrangement $\mathcal{A}$ associated to $\mathcal{M}_{0,m}$ is not a generic arrangement of $n+1 = \binom{m-1}{2} - 1$ hyperplanes in $\mathbb{C}^{m-3}$. For example, the Euler characteristic is *much* smaller: Comparing Examples 4.3.6 and 4.3.7, we see

$$\mathrm{MLdeg}(\mathcal{M}_{0,m}) = (m-3)! \ll \binom{\binom{m-1}{2} - 2}{m-3} = \mathrm{MLdeg}(\text{“generic } \mathcal{A}\text{”}).$$

For $m = 10$, these numbers are 5040 and 5379616, respectively.

Conjecture 4.6.2. For any $m \geq 5$, the logarithmic discriminant $\nabla_{\log}(\mathcal{M}_{0,m})$ is an irreducible and reduced hypersurface in $\mathbb{P}^{\frac{m(m-3)}{2}-1}$. For $m = 5, 6, 7, 8$ its degree equals $4, 30, 208$ and 1540, respectively. ⑦

The conjecture holds for $m = 5$ by the discussion above. It is supported by a numerical computation for $m = 6, 7, 8$. The code is available at [KKT].

Contrary to the case $m = 5$, such an explicit analysis for $\mathcal{M}_{0,6}$ seems to be out of reach. $\Delta_{\log}(\mathcal{M}_{0,6})$ is (conjectured to be) a polynomial of degree 30 in nine variables. We have not been able to compute all its coefficients. However, we computed partial information which showcases an interesting recursive structure.

4.6.1 Soft limits

Our strategy for obtaining such partial information is based on *degenerations* of the logarithmic discriminant variety $\nabla_{\log}$, which we assume to have codimension one. Concretely, for a weight vector $w \in \mathbb{R}^{n+1}$, let

$$\Delta^w_{\log}(u) = \mathrm{in}_w(\Delta_{\log}) = \lim_{\varepsilon \to 0} \varepsilon^{\max_{\alpha \in P} w \cdot \alpha} \cdot \Delta_{\log}(\varepsilon^{-w} \cdot u).$$

be the *w-initial form* of $\Delta_{\log}$ (see Section 0.2.2). Note that we use the max-convention in this thesis, hence weight vector signs are consistently flipped relative to [KKT25, Section 8]. Geometrically, we consider the limit of hypersurfaces

$$\nabla_{\log}(\varepsilon) = \{ u \in \mathbb{P}^n \mid \Delta_{\log}(\varepsilon^{-w} \cdot u) = 0 \} \xrightarrow{\varepsilon \to 0} \nabla^w_{\log} = \{ u \in \mathbb{P}^n \mid \Delta^w_{\log}(u) = 0 \}.$$

We interpret $\Delta_{\log}^w$ in terms of the *Newton polytope* of $\Delta_{\log}$. This will justify the hope that $\Delta_{\log}^w$ is easier to compute. Recall that the Newton polytope of $\Delta_{\log}$ is the convex polytope spanned by the exponents

$$P = \mathrm{Conv}(\{\, \alpha \mid x^\alpha \in \mathrm{supp}(\Delta_{\log}) \,\}) \subset \mathbb{R}^{n+1}$$

By homogeneity, it has dimension at most n. A *face* of P is a subset of the form $P^w = \{\, \alpha' \in P \mid w \cdot \alpha' = \max_{\alpha \in P} w \cdot \alpha \,\}$ for some $w \in \mathbb{R}^{n+1}$. The *f-vector* $(f_0, f_1, \ldots, f_{n-1})$ records the number of faces f_i of dimension i. Facets are faces of dimension $\dim P - 1$ and vertices are zero-dimensional faces.

The Newton polytope P of $\Delta_{\log}$ for $\mathcal{M}_{0,5}$ is four-dimensional with f-vector $(7, 17, 18, 8)$. Its eight facets P^w and the corresponding w-initial forms are displayed in Table 4.1.

$-w$	$\Delta_{\log}^w$
$(1,0,0,0,0)$	$(s_{14}s_{23} + s_{14}s_{34} + s_{23}s_{34} + s_{24}s_{34} + s_{34}^2)^2$
$(0,1,0,0,0)$	$(s_{13}s_{24} + s_{13}s_{34} + s_{23}s_{34} + s_{24}s_{34} + s_{34}^2)^2$
$(0,0,1,0,0)$	$(s_{13}s_{24} + s_{13}s_{34} + s_{14}s_{34} + s_{24}s_{34} + s_{34}^2)^2$
$(0,0,0,1,0)$	$(s_{14}s_{23} + s_{13}s_{34} + s_{14}s_{34} + s_{23}s_{34} + s_{34}^2)^2$
$(0,0,1,1,1)$	$(s_{14}s_{23} + s_{13}s_{24} + s_{13}s_{34} + s_{14}s_{34})^2 - 4s_{13}s_{14}s_{23}s_{24}$
$(0,1,0,1,1)$	$(s_{14}s_{23} + s_{13}s_{24} + s_{13}s_{34} + s_{23}s_{34})^2 - 4s_{13}s_{14}s_{23}s_{24}$
$(1,0,1,0,1)$	$(s_{14}s_{23} + s_{13}s_{24} + s_{14}s_{34} + s_{24}s_{34})^2 - 4s_{13}s_{14}s_{23}s_{24}$
$(1,1,0,0,1)$	$(s_{14}s_{23} + s_{13}s_{24} + s_{23}s_{34} + s_{24}s_{34})^2 - 4s_{13}s_{14}s_{23}s_{24}$

13, 14, 23, 24, 34

Table 4.1: The w-initial forms $\Delta_{\log}^w(\mathcal{M}_{0,5})$ for the facets of the Newton polytope.

The relation between initial forms of $\Delta_{\log}$ and faces of P is as follows: $\Delta_{\log}^w(u)$ is the sum of all terms in $\Delta_{\log}$ whose exponents lie on the face P^w. In particular, $\Delta_{\log}^w$ is a nonzero polynomial whose degree is that of $\Delta_{\log}$. The challenge is to compute $\Delta_{\log}^w$ without computing $\Delta_{\log}$ first. This can be done by degenerating the ramification locus, as illustrated in the next examples.

Example 4.6.3. For $\mathcal{M}_{0,5}$, we first consider the facet with normal $w = -(0,1,0,1,1)$. We will obtain the w-initial form from the likelihood equations $\nabla \mathcal{L}_{\varepsilon^{-w} \cdot u} = 0$:

$$\frac{s_{13}}{x_1} + \frac{s_{23}}{x_1 - 1} - \frac{\varepsilon s_{34}}{x_2 - x_1} = \frac{\varepsilon s_{14}}{x_2} + \frac{\varepsilon s_{24}}{x_2 - 1} + \frac{\varepsilon s_{34}}{x_2 - x_1} = 0.$$

We study these equations in the limit for $\varepsilon \to 0$. This choice of w leads to the *soft limit* for the particle labeled 4 in the physics literature [CUZ20]. This is because we replace s_{ij} by εs_{ij} each time $4 \in \{i, j\}$. As ε approaches 0, the two complex critical points converge in $\mathcal{M}_{0,5}$ [CUZ20, Section 2]. Their limits are the two solutions of

$$\frac{s_{13}}{x_1} + \frac{s_{23}}{x_1 - 1} = \frac{s_{14}}{x_2} + \frac{s_{24}}{x_2 - 1} + \frac{s_{34}}{x_2 - x_1} = 0. \tag{4.10}$$

The condition for the solutions of (4.10) to coincide is precisely $\Delta^w_{\log} = 0$. To check this, one eliminates x_1 and computes a quadratic discriminant as earlier this section. Thus, $\nabla^w_{\log}$ is the projection of the ramification locus in the soft limit. ◁

Example 4.6.4. In Example 4.6.3, the critical points converge in $\mathcal{M}_{0,5}$ in the soft limit. The situation for the facet $w = -(0,0,1,1,1)$ is different. We consider

$$\frac{s_{13}}{x_1} + \frac{\varepsilon s_{23}}{x_1 - 1} - \frac{\varepsilon s_{34}}{x_2 - x_1} = \frac{s_{14}}{x_2} + \frac{\varepsilon s_{24}}{x_2 - 1} + \frac{\varepsilon s_{34}}{x_2 - x_1} = 0.$$

This time, when $\varepsilon \to 0$, the critical points move to the boundary of $\mathcal{M}_{0,5}$. With the Ansatz $x_1 = 1 + \varepsilon \bar{x}_1 + O(\varepsilon^2)$ and $x_2 = 1 + \varepsilon \bar{x}_2 + O(\varepsilon^2)$, we find

$$s_{13} + \frac{s_{23}}{\bar{x}_1} - \frac{s_{34}}{\bar{x}_2 - \bar{x}_1} = s_{14} + \frac{s_{24}}{\bar{x}_2} + \frac{s_{34}}{\bar{x}_2 - \bar{x}_1} = 0.$$

We compute that the solutions for $(\bar{x}_1, \bar{x}_2)$ coincide if and only if $\Delta^w_{\log} = 0$. This happens when the critical points of $\mathcal{L}_{\varepsilon^w \cdot u}$ approach the point $(1,1)$ in the same tangent direction. Thus, the initial form $\Delta^w_{\log}$ detects when these critical points collide for $\varepsilon \to 0$ in the Deligne–Mumford compactification $\overline{\mathcal{M}}_{0,5}$. ◁

We suggest a generalization of the observation in Example 4.6.3 concerning the initial forms of $\Delta_{\log}(\mathcal{M}_{0,m})$ corresponding to soft limits. For $3 \leq k \leq m-1$, let $w_k \in \mathbb{R}^{n+1}$ be the weight vector corresponding to the soft limit for particle k. That is, w assigns weight 1 to s_{ij} if $k \in \{i,j\}$, and weight 0 otherwise. Example 4.6.3 uses $k = 4, m = 5$. Let $\mathcal{M}^{(k)}_{0,m-1}$ be the complement of the hyperplane arrangement in $\mathbb{C}^{m-4}$ given by all non-constant minors of the submatrix of (4.2) obtained by deleting the k-th column. If its logarithmic discriminant is a hypersurface, then $\Delta_{\log}(\mathcal{M}^{(k)}_{0,m-1})$ is a polynomial in the Mandelstam variables s_{ij} with $k \notin \{i,j\}$.

Based on numerical computations and nearly complete theoretical insights, we believe that $\Delta_{\log}(\mathcal{M}^{(k)}_{0,m-1})^{m-3}$ divides the soft limit initial form $\Delta^{w_k}_{\log}(\mathcal{M}_{0,m})$. We leave a proof of this claim for future work, and end with an example.

Example 4.6.5. The scattering equations for $\mathcal{M}_{0,6}$ with $u \to \varepsilon^{-w_5} \cdot u$ are

$$\begin{aligned}
f_1 &= \frac{s_{13}}{x_1} + \frac{s_{23}}{x_1 - 1} - \frac{s_{34}}{x_2 - x_1} - \frac{\varepsilon s_{35}}{x_3 - x_1} = 0 \\
f_2 &= \frac{s_{14}}{x_2} + \frac{s_{24}}{x_2 - 1} + \frac{s_{34}}{x_2 - x_1} - \frac{\varepsilon s_{45}}{x_3 - x_2} = 0 \\
f_3 &= \frac{\varepsilon s_{15}}{x_3} + \frac{\varepsilon s_{25}}{x_3 - 1} + \frac{\varepsilon s_{35}}{x_3 - x_1} + \frac{\varepsilon s_{45}}{x_3 - x_2} = 0.
\end{aligned}$$

The discriminant of $f_1 = f_2 = 0$ with $\varepsilon = 0$ is the logarithmic discriminant $\Delta_{\log}(\mathcal{M}^{(5)}_{0,5})$ of degree 4. It appears as a factor $\Delta_{\log}(\mathcal{M}^{(5)}_{0,5})^3$ in $\Delta^{w_5}_{\log}(\mathcal{M}_{0,6})$. The degree of $\Delta_{\log}(\mathcal{M}_{0,6})$ is (conjecturally) 30. The missing factor of degree 18 captures when (x_1, x_2) leads to a double root in x_3 for $f_3 = 0$. To compute this, first isolate x_2 in $f_1 = 0$ to express x_2 in terms of x_1 in f_2 and f_3. After that, up to some spurious factors, our degree 18 factor is the resultant $\operatorname{res}_{x_1}(g_2, \Delta_{x_3}(g_3))$, with g_i the numerator of f_i. The face P^{w_5} of the Newton polytope P of $\Delta_{\log}(\mathcal{M}_{0,6})$ is a facet. It contains 600413 lattice points and has f-vector $(237, 907, 1432, 1195, 564, 149, 20)$. ◁

Chapter 5

Counting Critical Points to Low-rank Hankel Matrix Approximation

Contents

> [Algebraic geometry] seems to have acquired the reputation of being esoteric, exclusive, and very abstract, with adherents who are secretly plotting to take over all the rest of mathematics. In one respect this last point is accurate.
>
> DAVID MUMFORD [Mum99]

In this final chapter, we consider the complexity of a certain polynomial optimization problem: Low-rank Hankel matrix approximation. This is a fundamental problem in engineering applications, and can be formulated both in a finite-dimensional and infinite-dimensional setting. In the former setting, it is an instance of an Euclidean distance minimization problem to the variety $\mathbb{P}\mathcal{X}_{d,r} = \sigma_r\nu_d(\mathbb{P}^1) \subseteq \mathbb{P}^d$ with respect to a (positive-definite) quadric. The latter setting can be seen as a limit when $d \to \infty$, and corresponds to a rational function approximation problem. We study necessary conditions for an approximant to be a critical point, with the goal to count the (complex) solutions to the critical equations (known as the Euclidean distance degree (EDD) in the finite-dimensional case). In both finite and infinite dimensions, the critical equations can be interpreted as a multiparameter eigenvalue problem, whose structure allows us to apply the Thom–Porteous formula for a solution count. Main contributions of this chapter are:

Main Results 5 (5.3.11, 5.3.12, 5.3.15, 5.4.5, 5.4.12, 5.5.12)**.** *The generic Euclidean distance degree of* $\mathcal{X}_{d,r}$ *is*

$$\mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}) = \sum_{j=0}^{r} \binom{d-2r+j}{j} 3^j.$$

A quadric Q *satisfies* $\mathrm{EDD}_Q(\mathcal{X}_{d,r}) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r})$ *if the associated* bad locus *is empty. We develop an algorithm to compute bounds on* $\mathrm{EDD}_Q(\mathcal{X}_{d,r})$ *for non-generic* Q*. Computational studies lead to new results on* $\mathrm{EDD}_\Theta(\mathcal{X}_{d,r})$ *for the Bombieri inner product.*

In the infinite-dimensional setting, the algebraic degree of the Model Order Reduction problem from order R *to* r *is*

$$\sum_{j=0}^{r} \binom{R-r-1+j}{j} 2^j.$$

We present an optimal parameter homotopy to find the critical points in both cases.

This chapter is based on (at the time of writing) unpublished joint work with Sibren Lagauw. A version of this material which focuses on system theoretic aspects also appears in his doctoral thesis [Lag26]. The chapter is structured as follows: In Section 5.1 we introduce the notion of the Euclidean distance degree (EDD) from metric algebraic

geometry. Our goal is to study ED degrees of the variety of low-rank Hankel matrices. These matrices and their connection to linear recurrences and dynamical systems in engineering are studied in detail in Section 5.2. In Section 5.3, we study critical equations for the ED minimization problem for $\mathcal{X}_{d,r}$ via a resolution of singularities. We then realize the critical locus as a degeneracy locus, which allows us to derive a new formula for the generic EDD. In Section 5.4, we then consider specific weight matrices, for which the "bad locus" can often explain the ED defect. Finally, in Section 5.5 we turn to the infinite-dimensional analogue, where we study critical equations (the Walsh system) and compute the algebraic degree of the optimization problem.

5.1 Background: Euclidean Distance Degree

We start by introducing the theory of critical points to a distance function to an affine or projective variety. The Euclidean distance degree (EDD) counts the number of complex critical points for general data. If the distance function is given by a positive definite weight matrix Λ, then this number depends on Λ. In this case, generic and specific ED degrees for application-specific Λ's are of interest.

5.1.1 Critical points to the distance function

Let $\emptyset \neq X(\mathbb{R}) \subseteq \mathbb{R}^n$ be a real algebraic variety and $y \in \mathbb{R}^n$ a data point. We can consider the following optimization problem:

$$\text{minimize } ||\hat{y} - y||_2 \qquad \text{subject to } \hat{y} \in X(\mathbb{R}).$$

To have an algebraic distance function, we instead consider the equivalent formulation with the squared distance $d_y^2(\hat{y}) := \|\hat{y} - y\|_2^2$, which is a polynomial function.

Lemma 5.1.1. (i) *The above optimization problem has a global minimizer $y^* \in X(\mathbb{R})$.*

(ii) *If $X(\mathbb{R})$ is smooth, then every local minimizer y^* is a critical point of the function $d_y^2(\hat{y})$ on $X(\mathbb{R})$, i. e. $\nabla d_y^2|_{y^*} = 0$.*

(iii) *In general, consider the stratification $X \supset X_{\text{sing}} \supset (X_{\text{sing}})_{\text{sing}} \supset \dots$ where every singular locus is considered as an affine variety. Then every local minimizer y^* is a critical point of d_y^2 on the stratum containing y^*.*

Proof. (i) Pick any $x \in X(\mathbb{R})$, then any minimum has to be in the closed ball $B_{||x-y||}(x)$. Since $X(\mathbb{R})$ is closed, $X(\mathbb{R}) \cap B$ is compact and the minimum exists.

(ii) This is a standard fact from multivariate calculus.

(iii) Each stratum is smooth away from the next one, hence (iii) follows from (ii). □

Lemma 5.1.1 indicates that to find the minimizer, one can try to compute the critical points of d_y on X_{reg} and pick the one with minimal distance to y. Critical points on X_{reg} are characterized by the following properties:

Lemma 5.1.2. *Let $\hat{y} \in X_{\text{reg}}(\mathbb{R})$. The following are equivalent:*

(a) *$\hat{y}$ is a critical point of d_y^2 on $X_{\text{reg}}(\mathbb{R})$;*

(b) *$\hat{y} - y \in N_{\hat{y}}X \subseteq \mathbb{R}^n$, where $N_{\hat{y}}X = (T_{\hat{y}})^\perp$ is the normal space to X at $\hat{y}$ with respect to the standard inner product;*

(c) *If $f_1, \dots, f_s \in \mathbb{R}[x_1, \dots, x_n]$ define X in a neighborhood of $\hat{y}$, then there exist $\ell \in \mathbb{R}^n$ such that the* Lagrangian function

$$\mathcal{L}_y(\hat{y}, \ell) \colon \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}, \qquad (\hat{y}, \ell) \mapsto \frac{1}{2} d_y^2(\hat{y}) + \sum_{i=1}^{s} \ell_i f_i(\hat{y})$$

has vanishing gradient $\nabla \mathcal{L}_y|_{(\hat{y},\ell)}$.

Proof. The equivalence of (a)⇔(c) is exactly the method of Lagrange multipliers. For the equivalence of (b) and (c), note that

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \hat{y}} &= (\hat{y} - y) + \sum_{i=1}^{s} \ell_i \nabla f_i(\hat{y}) = 0 \\ \frac{\partial \mathcal{L}}{\partial \ell} &= (f_1(\hat{y}), \dots, f_s(\hat{y})) = 0. \end{aligned}$$

The second set of equations describe the condition $\hat{y} \in X$. The first set of equations exactly say that $y - \hat{y} \in N_{\hat{y}}X$, since $\nabla f_i(\hat{y})$ span the normal space. □

Condition 5.1.2(b) gives rise to an incidence correspondence of pairs $(y, \hat{y})$ such that $\hat{y} \in X_{\text{reg}}$ is a critical point of d_y^2. We generalize this slightly by allowing the distance function to be a weighted norm squared $x \mapsto \frac{1}{2} x^\mathsf{T} \Lambda x$, where $\Lambda \in \text{Sym}^2(\mathbb{R}^n)$ is a positive definite matrix. Conditions (b) from the previous lemma then become

$$\Lambda(\hat{y} - y) \in \text{Rowspan}\, J_f(\hat{y}), \qquad J_f(\hat{y}) = [\nabla f_1(\hat{y}), \dots, \nabla f_s(\hat{y})]^\mathsf{T},$$

where $J_f(\hat{y})$ is the Jacobi matrix. Thus, the incidence of data and critical points $(y, \hat{y})$ is the real part of the degeneracy locus of the *augmented Jacobi matrix*

$$\mathcal{E}_X \coloneqq \overline{\left\{ (y, \hat{y}) \in \mathbb{C}^n \times X_{\text{reg}} \;\middle|\; \text{rank} \begin{bmatrix} \Lambda(\hat{y} - y) \\ J_f(\hat{y}) \end{bmatrix} \leq c \coloneqq \text{codim}\, X \right\}} \subseteq \mathbb{C}^n \times X.$$

Lemma 5.1.3 ([Dra+15, Theorem 4.1])**.** *The restriction $\mathcal{E}_X^\circ = \mathcal{E}_X \cap (\mathbb{C}^n \times X_{\text{reg}})$ is a smooth determinantal variety of dimension n, defined by the $(c+1)$-minors of the augmented Jacobi matrix. It is an affine bundle of rank c over X_{reg}.*

Proof. The incidence is an affine bundle via the projection $\text{pr}_{\hat{y}} \colon \mathcal{E}_X^\circ \to X_{\text{reg}}$, its fibers are c-dimensional affine spaces passing through $\hat{y}$. Thus $\mathcal{E}_X^\circ$ is of dimension n, and the underlying (reduced) variety is smooth. To argue reducedness, we can argue locally around $\hat{y} \in X_{\text{reg}}$, where X is given by c equations, hence we can apply Proposition 0.4.15 to the Lagrange system to conclude that it is reduced. □

Thus, to find the critical points of d_y on X_{reg} means to compute the fiber of $\mathrm{pr}_y \colon \mathcal{E}_X \to \mathbb{C}^n$ and discarding points in X_{sing} (for general y, there are no such points).

Lemma 5.1.4. *For general $y \in \mathbb{C}^n$, the number of critical points in X_{reg} is finite and independent of y. If there is a point $x \in X_{\mathrm{reg}}$ such that $T_x X \cap (T_x X)^{\perp_\Lambda} = 0$, then this number is positive. This is the case if Λ is positive definite and $X_{\mathrm{reg}}(\mathbb{R}) \neq \emptyset$.*

Note that since any non-empty Euclidean-open subset $U \subseteq \mathbb{R}^n$ is Zariski-dense in $\mathbb{C}^n$, we could also restrict to data points y in such a U.

Proof. The projection $\mathrm{pr}_y \colon \mathcal{E}_X^\circ \to \mathbb{C}^n$ is a map of irreducible varieties of the same dimension n, hence the first statement follows from the fiber dimension theorem.

To show that $\mathrm{pr}_y \colon \mathcal{E}_X \to \mathbb{C}^n$ is dominant if $T_x X \cap T_x X^{\perp_\Lambda} = \{0\}$ for some $x \in X_{\mathrm{reg}}$, it suffices to show that the differential

$$\mathrm{d}_{(x,x)} \mathrm{pr} \colon T_{(x,x)} \mathcal{E}_X \to T_x \mathbb{C}^n = \mathbb{C}^n$$

is surjective at that point. This can be established by noticing that the image of the tangent space $(\mathrm{d}_{(x,x)} \mathrm{pr}_y)(T_{(x,x)} \mathcal{E}_X) \subseteq \mathbb{C}^n$ contains both $T_x X$ and $(T_x X)^{\perp_\Lambda}$; for details see [Dra+15, Theorem 4.1].

For the last part note that if $x \in X_{\mathrm{reg}}(\mathbb{R})$, then since Λ is positive-definite, the real tangent space intersects its normal space trivially, and hence the same holds for its complexification. □

Definition 5.1.5. The *Euclidean distance degree* $\mathrm{EDD}_\Lambda(X)$ of a (complex) variety $X \subseteq \mathbb{C}^n$ is the degree of the projection $\mathcal{E}_X \to \mathbb{C}^n$. By extension, the ED degree of a projective variety $X \subseteq \mathbb{P}^{n-1}$ is defined as $\mathrm{EDD}_\Lambda(X) \coloneqq \mathrm{EDD}_\Lambda(\widehat{X})$.

Example 5.1.6. Let $X = \mathbb{V}(x_1^2 + x_2^2 - 1) \subseteq \mathbb{R}^2$ be a circle, then the real critical points to $y = (y_1, y_2)$ are given by $\frac{\pm 1}{\sqrt{y_1^2 + y_2^2}}(y_1, y_2)$ when $y \neq 0$, and the entire circle if $y = 0$. The augmented Jacobi matrix has determinant

$$\det \begin{pmatrix} y_1 - \hat{y}_1 & y_2 - \hat{y}_2 \\ 2\hat{y}_1 & 2\hat{y}_2 \end{pmatrix} = 2(y_1 \hat{y}_2 - y_2 \hat{y}_1).$$

Intersecting this line with the circle indeed gives the critical points. In particular, $\mathrm{EDD}(X) = 2$.

Next let $X = \mathbb{V}(x_1^2 - x_2)$ be a parabola, then

$$\det \begin{bmatrix} y_1 - \hat{y}_1 & y_2 - \hat{y}_2 \\ 2\hat{y}_1 & -1 \end{bmatrix} = \hat{y}_1 - y_1 - 2\hat{y}_1 y_2 + 2\hat{y}_1 \hat{y}_2.$$

Intersecting this with the parabola by substituting $\hat{y}_2 = \hat{y}_1^2$ yields

$$\hat{y}_1 - y_1 - 2\hat{y}_1 y_2 + 2\hat{y}_1^3 = 0,$$

so the Euclidean distance degree of the parabola is $\mathrm{EDD}(X) = 3$.

Lastly, if X is an ellipse that is not a circle, or a hyperbola, then no cancellation happens in the intersection of X and the determinant of the augmented Jacobi matrix, so these plane conics satisfy $\mathrm{EDD}(X) = 4$. ◁

5.1.2 Generic Euclidean Distance Degree

For any projective variety $X \subseteq \mathbb{P}^{n-1}$ and any weight matrix Λ, the intersection of X with the isotropic quadric $\mathcal{Q}_\Lambda \coloneqq \mathbb{V}(x^\mathsf{T}\Lambda x)$ will be non-empty. If X is smooth and this intersection is transverse, then the ED degree can be computed by invariants of X:

Theorem 5.1.7 ([Dra+15, Theorem 5.4 & 5.8]). *If $X \subseteq \mathbb{P}^{n-1}$ is smooth of dimension m and the intersection $X \cap \mathcal{Q}_\Lambda$ is transversal, then*

$$\operatorname{EDD}_\Lambda(X) = \sum_{i=1}^{n-1} \operatorname{mdeg}_{(i,n-i)} \mathcal{N}_X = \sum_{i=0}^{m} (-1)^i (2^{m+1-i} - 1) \deg c_i(\mathcal{T}X).$$

Here, $\mathcal{N}_X \subseteq \mathbb{P}^{n-1} \times (\mathbb{P}^{n-1})^\vee$ is the conormal variety *and $\deg c_i(\mathcal{T}X)$ is the degree of the i-th Chern class of the tangent bundle after the pushforward $\mathrm{A}_{m-i}(\mathbb{P}^{n-1}) \cong \mathbb{Z}$.*

This result extends to singular varieties when replacing Chern classes with Chern–Mather classes [Pie88, Théorème 3] and suitably refining the transversality assumption (see [Dra+15, Theorem 5.4]). Since almost all quadrics are transversal to a given variety, this justifies the following definition:

Definition 5.1.8. The *generic Euclidean distance degree* of a variety $\operatorname{EDD}_{\text{gen}}(X)$ is the ED degree of X with respect to a general set of weights Λ. The *ED defect* with respect to Λ is the difference $\operatorname{EDdef}_\Lambda(X) = \operatorname{EDD}_{\text{gen}}(X) - \operatorname{EDD}_\Lambda(X)$.

The name "defect" is justified by the following proposition:

Lemma 5.1.9. *For every full rank $\Lambda \in \operatorname{Sym}^2(\mathbb{C}^n)$, $\operatorname{EDD}_\Lambda(X) \leq \operatorname{EDD}_{\text{gen}}(X)$.*

Proof. Pick an open set $U \subseteq \mathbb{C}^n$ and $f_1, \dots, f_c$ ($c = \operatorname{codim}_{\mathbb{C}^n} X$) such that $X \cap U = \mathbb{V}(f_1, \dots, f_c) \cap U \subseteq X_{\text{reg}}$ is a dense open set of the regular locus. Consider the incidence

$$\mathcal{E}_{\text{univ}} \coloneqq \left\{ (\Lambda, y, \hat{y}, \ell) \;\middle|\; \begin{array}{c} \Lambda(\hat{y} - y) = \ell_1 \nabla f_1 + \dots + \ell_c \nabla f_c \\ f_1(\hat{y}) = \dots = f_c(\hat{y}) = 0 \end{array} \right\} \subseteq \operatorname{Sym}^2 \mathbb{C}^n \times \mathbb{C}^n \times U \times \mathbb{C}^c.$$

This is an affine bundle of rank $\binom{n}{2} - c$ over $(y, \hat{y}) \in \mathbb{C}^n \times (X \cap U)$ (for example by Lemma 5.4.2 below) and has dimension

$$\dim \mathcal{E}_{\text{univ}} = n + (n - c) + \left(\binom{n}{2} + c \right) = \binom{n+1}{2} + n = \dim(\operatorname{Sym}^2 \mathbb{C}^n \times \mathbb{C}^n).$$

Now consider the projection $p \colon \mathcal{E}_{\text{univ}} \to \operatorname{Sym}(n, \mathbb{C}) \times \mathbb{C}^n$, this is a "universal" variant of the ED correspondence. Given $\Lambda \in \operatorname{Sym}(n, \mathbb{C})$, by definition $\operatorname{EDD}_\Lambda(X)$ is the number of isolated reduced points in the fiber $p^{-1}(\Lambda, y)$ for general $y \in \mathbb{C}^n$. This quantity is always bounded above by the number of reduced points in the fiber for general points (Λ', y') in $\operatorname{Sym}^2 \mathbb{C}^n \times \mathbb{C}^n$ [SW05, Theorem 7.1.4], which establishes $\operatorname{EDD}_\Lambda(X) \leq \operatorname{EDD}_{\text{gen}}(X)$. □

In some cases, the ED defect can be computed from the singularities of the intersection $X \cap \mathcal{Q}_\Lambda$.

Theorem 5.1.10 ([MRW20, Corollary 1.8]). *If X is smooth and the singular locus $(X \cap \mathcal{Q}_\Lambda)_{\text{sing}}$ is finite, then*

$$\operatorname{EDdef}_\Lambda(X) = \sum_{x \in (X \cap \mathcal{Q}_\Lambda)_{\text{sing}}} \mu_x,$$

where $\mu_x > 0$ is the Milnor number of the isolated singularity $(X \cap \mathcal{Q}_\Lambda, x)$.

Example 5.1.11. Consider a smooth conic $X \subseteq \mathbb{P}^2$. We compute its generic ED degree using the Chern classes. Note that $X \cong \mathbb{P}^1$, so $\mathcal{T}X = \mathcal{O}_{\mathbb{P}^1}(2)$. Thus,

$$c_0(\mathcal{T}X) = 1 \cdot [\mathbb{P}^1] \in \mathrm{A}_1(\mathbb{P}^1), \qquad c_1(\mathcal{T}X) = 2 \cdot [pt] \in \mathrm{A}_0(\mathbb{P}^1).$$

To calculate the degree of these Chern classes, we need to consider them under the pushforward $\mathrm{A}_i(X) \to \mathrm{A}_i(\mathbb{P}^2)$ via the inclusion $X \hookrightarrow \mathbb{P}^2$. We see

$$\deg c_0(\mathcal{T}X) = \deg X = 2, \qquad \deg c_1(\mathcal{T}X) = \deg(2 \cdot [pt]) = 2$$

and the generic ED degree is

$$\operatorname{EDD}_{\text{gen}}(X) = 1 \cdot 3 \cdot 2 + (-1) \cdot 1 \cdot 2 = 4.$$

This matches the computation in Example 5.1.6, as "generic affine plane conics" are ellipses and hyperbolas.

However, if $\widehat{X} = \mathbb{V}(x_0^2 + x_1^2 - x_2^2) \subseteq \mathbb{C}^3$ is the "standard" cone (its level sets $\widehat{X} \cap \{x_2 = r\}$ are circles), then its ED degree drops to 2: The intersection of X with the isotropic quadric consists of two ordinary double points

$$X \cap \mathcal{Q}_{\mathrm{Id}_3} = \mathbb{V}\left((x_0 + \mathrm{i}x_1)(x_0 - \mathrm{i}x_1), x_2^2\right) = 2 \cdot [1 : \mathrm{i} : 0] \cup 2 \cdot [1 : -\mathrm{i} : 0].$$

A double point has Milnor number 1, hence $\operatorname{EDdef}_{\mathrm{Id}_3}(X) = 2$. ◁

5.2 Background: Linear recurrences and their solutions

In this section we briefly discuss the basic theory of linear recurrence relations and their relation to Hankel and Toeplitz matrices. Here $y = (y_0, y_1, \dots) \in \mathbb{K}^{d+1}$ will always be a sequence, $d \in \mathbb{N}_+ \cup \{\infty\}$ (where we define $\mathbb{K}^\infty := \prod_{k \in \mathbb{N}} \mathbb{K}$), and we study when it satisfies a homogeneous linear recurrence relation. First, we characterize what solutions look like, and then characterize y's satisfying low-order recurrences using Hankel matrices.

It is no more difficult to develop the theory over an arbitrary (algebraically closed) field $\mathbb{K}$, which we shall do here. However, there is no intuition lost in considering only real or complex sequences.

5.2.1 The characteristic polynomial

We start by studying sequences which satisfy a linear recurrence relation $0 \neq a \in \mathbb{K}^{r+1}$:

$$a_0 y_k + a_1 y_{k+1} + \cdots + a_r y_{k+r} = 0, \qquad 0 \leq k \leq d - r.$$

To allow for more compact notation, consider the *Toeplitz matrix*

$$T(a) = T_d(a) \coloneqq \begin{bmatrix} a_0 & \dots & a_r & 0 & \cdots & \\ 0 & a_0 & \dots & a_r & 0 & \cdots \\ & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \in \mathbb{K}^{(d-r+1)\times(d+1)}.$$

Then the linear recurrence relation can be compactly written as $T(a) \cdot y = 0$.

Definition 5.2.1. We say that y is of *finite rank* if it satisfies a linear recurrence relation $a \neq 0$. The *order* of a is the number r, the *rank* of y is the smallest order of a recurrence a that is satisfied by y. The *characteristic polynomial* of a is

$$a(z_0, z_1) = a_0 z_0^r + a_1 z_0^{r-1} z_1 + \cdots + a_r z_1^r \in \mathbb{K}[z_0, z_1]_r,$$

we shall identify a with its dehomogenization $a(z) = a(1, z) \in \mathbb{K}[z]_{\leq r}$.

In order to not lose information in the dehomogenization, we will record extra factors z_0^e as $(z - \infty)^e$, indicating a root at $\infty \in \mathbb{P}^1(\mathbb{K})$.

Remark 5.2.2. Note that a factor of $z_0^e \mathrel{\hat{=}} (z - \infty)^e$ indicates that the r-term recurrence does not involve the last e terms of y if $d < \infty$ ($y_{d-e+1}, \dots, y_d$ do not appear). On the other hand, if $d = \infty$, then y also obeys the shorter relation a/z_0^e, as y extends indefinitely to the right. For this reason, we mostly ignore roots at infinity and use the dehomogenization. ◁

The Toeplitz matrix $T(a)$ always has maximal rank $(d - r + 1)$, hence the space of solutions to the recurrence a is r-dimensional. A convenient basis is provided by the *Vandermonde vectors* and their *confluent* derivatives:

$$\operatorname{vand}(\omega) \coloneqq (\omega^k)_k, \qquad \operatorname{vand}^{(j)}(\omega) \coloneqq \left(\binom{k}{j}\omega^{k-j}\right)_k \overset{j!\in\mathbb{K}^\times}{=} \frac{1}{j!}\frac{\partial^j \operatorname{vand}(x)}{\partial x^j}\bigg|_{x=\omega},$$
$$\operatorname{vand}(\infty) \coloneqq (\delta_{k,d})_k = e_d, \qquad \operatorname{vand}^{(j)}(\infty) \coloneqq e_{d-j}.$$

Here $\omega \in \mathbb{K}$ and $\operatorname{vand}^{(j)}(\infty)$ is only defined for $d < \infty$; note that $\operatorname{vand}^{(0)}(\omega) = \operatorname{vand}(\omega)$.

Proposition 5.2.3. *Assume the characteristic polynomial factors as*

$$a(z) = a_r(z - \omega_1)^{e_1} \cdots (z - \omega_m)^{e_m}, \qquad \omega_i \in \mathbb{K} \cup \{\infty\},\ e_i \geq 1.$$

Then there exist unique $\lambda_{ij} \in \mathbb{K}$ *such that*

$$y = \sum_{i=1}^{m} \sum_{j=0}^{e_i - 1} \lambda_{ij} \cdot \operatorname{vand}^{(j)}(\omega_i) \in \mathbb{K}^{d+1}.$$

Proof. To argue existence and uniqueness, it suffices to consider finite d (as infinite solutions are composed of finite solutions of arbitrary length). Let $\omega \in \mathbb{K}$ be a root of $a(s)$, then the Vandermonde vector is a solution to the difference equation:

$$a_0\omega^k + \cdots + a_r\omega^{k+r} = \omega^k a(\omega) = 0.$$

If ω has multiplicity $e > 1$, then the confluent Vandermonde vectors also give solutions, since ω is also a root of the derivatives of $a(s)$ of order $< e$:

$$a_0 \operatorname{vand}^{(j)}(\omega)_k + \cdots + a_r \operatorname{vand}^{(j)}(\omega)_{k+r} = \omega^k \frac{1}{j!}\frac{\partial^j}{\partial s^j} a(\omega) = 0.$$

If $\operatorname{char}\mathbb{K} \leq j$, then one replaces $\frac{1}{j!}\frac{\partial^j}{\partial s^j}$ with the *Hasse derivatives* $\mathcal{D}^{(j)}(x^k) = \binom{k}{j} x^{k-j}$, which admit the same conclusion [LL25, Corollary 3.2]. The case of $\omega = \infty$ is clear from Remark 5.2.2.

We have found $\deg a(z) = r$ solution vectors $\operatorname{vand}^{(j)}(\omega_i)$, so it suffices to show that they are linearly independent. For this we can further restrict to $d = r$ and show that the square matrix

$$V^{(\underline{e})}(\underline{\omega}) \coloneqq \begin{bmatrix} | & \cdots & | & \cdots & | \\ \operatorname{vand}(\omega_1) & \cdots & \operatorname{vand}^{(e_1-1)}(\omega_1) & \cdots & \operatorname{vand}^{(e_m-1)}(\omega_m) \\ | & \cdots & | & \cdots & | \end{bmatrix}$$

is invertible. This *confluent Vandermonde matrix* has determinant

$$\det V^{(\underline{e})}(\underline{\omega}) = \prod_{i_1<i_2} (\omega_{i_2} - \omega_{i_1})^{e_{i_1}e_{i_2}} \neq 0$$

(see for example [LL25, Theorem 4.2]), which finishes the proof. □

Example 5.2.4. Consider the Fibonacci recurrence relation $f_k = f_{k-1} + f_{k-2}$. The characteristic polynomial is $a(s) = s^2 - s - 1 = (s - \frac{1+\sqrt{5}}{2})(s - \frac{1-\sqrt{5}}{2})$, hence all solutions are of the form

$$f_k = \lambda_1 \left(\frac{1+\sqrt{5}}{2}\right)^k + \lambda_2 \left(\frac{1-\sqrt{5}}{2}\right)^k, \qquad \lambda_1, \lambda_2 \in \mathbb{C}.$$

The desired initial conditions are

$$\begin{aligned} 0 &= f_0 = \lambda_1 + \lambda_2 \\ 1 &= f_1 = \lambda_1 \frac{1+\sqrt{5}}{2} + y_1 + \lambda_1 \frac{1-\sqrt{5}}{2}. \end{aligned}$$

Solving this linear system reveals the scalars $\lambda_1 = -\lambda_2 = \frac{1}{\sqrt{5}}$, and hence *Binet's formula*

$$f_k = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^k. \qquad \triangleleft$$

Corollary 5.2.5. *If $a(z)$ is a recurrence relation of minimal order of y and $a(z)$ factors over $\mathbb{K}$ as $a_r(z-\omega_1)^{e_1}\cdots(z-\omega_m)^{e_m}$, then the contributions of $\operatorname{vand}^{(e_i-1)}(\omega_i)$ in the presentation of Proposition 5.2.3 is non-zero, that is, $\lambda_{i,e_i-1} \neq 0$ for all i. In particular, if $a(z)$ has simple roots, then all $\lambda_{i,0} \neq 0$, meaning all terms $\operatorname{vand}(\omega_i)$ contribute.*

Proof. The case of $\omega = \infty$ is immediate from Remark 5.2.2. Assume that $\lambda_{i,e_i-1} = 0$, then from the presentation it is evident that y also satisfies the shorter relation $a(z)/(z-\omega_i)$, a contradiction to minimality of the order of a. □

From the explicit description of Proposition 5.2.3 we can also see how the sequence $(y_k)_{k\in\mathbb{N}}$ behaves in the limit $k \to \infty$.

Proposition 5.2.6. *Let $y \in \mathbb{C}^\infty$ be a solution to the recurrence relation a, and assume a is of minimal order. Then the sequence converges to 0 if and only if all roots ω of $a(z)$ have magnitude $|\omega| < 1$. The sequence is bounded if and only if all roots have magnitude $|\omega| \leq 1$ and those with $|\omega| = 1$ have multiplicity 1.*

The proof relies on a basic lemma about sums of exponentials:

Lemma 5.2.7. *Let $y_k \coloneqq \sum_{i=1}^r \lambda_i \omega_i^k$ with $\lambda_1, \ldots, \lambda_r, \omega_1, \ldots, \omega_r \in \mathbb{C}^\times$ and all ω_i distinct and of magnitude $|\omega_i| = 1$. Then $(y_k)_k$ is bounded but not a null sequence.*

Proof. By the triangle inequality $|y_k| \leq \sum_{i=1}^r |\lambda_i|$, so the sequence is bounded. Assume for a contradiction that $\lim_{k\to\infty} y_k = 0$. The same is true for $y_k \cdot \omega_r^{-k}$, so we may assume without loss of generality that $\omega_r = 1$. If y_k converges, then so does the Cesàro sum

$$s_n \coloneqq \frac{1}{n}\sum_{k=0}^{n-1} y_k = \frac{1}{n}\sum_{k=0}^{n-1}\sum_{i=1}^{r} \lambda_i \omega_i^k = \sum_{i=1}^{r} \lambda_i \frac{1}{n}\sum_{k=0}^{n-1} \omega_i^k = \lambda_r + \sum_{i=1}^{r-1} \lambda_i \frac{\omega_1^n - 1}{n(\omega_i - 1)}.$$

From the geometric sum we see that each summand on the right tends to zero, so $\lim_{n\to\infty} s_n = \lambda_r \neq 0$. This contradicts $\lim_{n\to\infty} s_n = \lim_{k\to\infty} y_k = 0$. □

Proof of Proposition 5.2.6. Let $\delta \coloneqq \max_i |\omega_i|$. If $\delta < 1$ then y_k is a null sequence. In the subsequent analysis we may thus ignore all terms with $|\omega| < 1$.

If $\delta \geq 1$, let e be the highest multiplicity among roots with $|\omega| = \delta$ and consider $y'_k \coloneqq y_k/(\delta^k k^{e-1})$. By Corollary 5.2.5, the confluent Vandermonde vector $\operatorname{vand}^{(e-1)}(\omega)$ contributes for every ω that achieves this maximum. Then y'_k is bounded and not a null sequence by Lemma 5.2.7, hence $|y_k| \to \infty$ if $\delta > 1$ or if $\delta = 1$ and $e > 1$. In the remaining case $\delta = 1$ and $e = 1$, the lemma also implies that y_k is bounded but not a null sequence. □

Proposition 5.2.6 motivates the following definition:

Definition 5.2.8 (Stable polynomial)**.** A univariate polynomial $a(z) \in \mathbb{C}[z]$ is *stable* or *Schur-stable* if all its roots lie in the open unit disc $\mathbb{D} = \{\omega \in \mathbb{C} \mid |\omega| < 1\}$.

5.2.2 Hankel matrices

In the previous section we described how to obtain solutions from linear recurrence relations. In this section we ask the converse question: Which linear recurrences does a given sequence $y \in \mathbb{K}^{d+1}$ obey? The answer is closely related to the linear algebra of Hankel matrices:

Definition 5.2.9. Let $y \in \mathbb{K}^{d+1}$ and $1 \leq r < d+1$. The *Hankel matrices* of y are

$$H_r(y) = [y_{i+j}]_{\substack{0 \leq i \leq d-r \\ 0 \leq j \leq r}} = \begin{bmatrix} y_0 & y_1 & y_2 & \cdots & y_r \\ y_1 & y_2 & y_3 & \cdots & y_{r+1} \\ \vdots & \iddots & \iddots & \iddots & \vdots \end{bmatrix} \in \mathbb{K}^{(d-r+1)\times(r+1)}.$$

Hankel matrices $H_r(y)$ detect linear recurrence relations of y of order r.

Lemma 5.2.10 (Hankel–Toeplitz exchange)**.** *For any $y \in \mathbb{K}^{d+1}$ and $a \in \mathbb{K}^{r+1}$ one has*

$$T(a) \cdot y = H_r(y) \cdot a.$$

In particular, y satisfies a linear recurrence of order r if and only if $\operatorname{rank} H_r(y) \leq r$.

Proof. The given equation is immediate from inspection:

$$\begin{bmatrix} a_0 & a_1 & \cdots & a_r & & \\ & a_0 & a_1 & \cdots & a_r & \\ & & \ddots & \ddots & \ddots & \ddots \end{bmatrix} \begin{pmatrix} y_0 \\ y_1 \\ \vdots \end{pmatrix} = \begin{bmatrix} y_0 & y_1 & \cdots & y_r \\ y_1 & y_2 & \cdots & y_{r+1} \\ \iddots & \iddots & \iddots & \iddots \end{bmatrix} \begin{pmatrix} a_0 \\ \vdots \\ a_r \end{pmatrix}$$

For the second statement, note that this equation tells us that linear recurrences of y can be detected as elements in the kernel of $H_r(y)$, hence they exist if and only if this matrix is not of rank $r+1$. □

Example 5.2.11. We return to the start of the Fibonacci sequence $f = (1, 1, 2, 3, 5)$. Its Hankel matrices are the following:

$$H_0(f) = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 3 \\ 5 \end{bmatrix}, \qquad H_1(f) = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 3 \\ 3 & 5 \end{bmatrix}, \qquad H_2(f) = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{bmatrix}, \qquad \dots$$

The first two matrices are of maximal rank, while the third has $a = (-1, -1, 1)^\mathsf{T}$ in its kernel. Thus the minimal recurrence relation is given by $f_k = f_{k-1} + f_{k-2}$. ◁

In fact, we have seen finite Hankel matrices before in our encounter with apolarity theory in section 1.2.1. This connection gives us useful insight into the structure of the space of all recurrence relations that y satisfies.

Lemma 5.2.12. *Let $d < \infty$ and set $F = \sum_{i=0}^{d} y_i X_0^{d-i} X_1^i \in \operatorname{Sym}^d \mathbb{K}^2$.*

(i) *The Catalecticant map*

$$F_{j,d-j}\colon \mathrm{S}^j(\mathbb{K}^2)^\vee \to \mathrm{Sym}^{d-j}\mathbb{K}^2, \qquad f \mapsto f \bullet F$$

is represented by the Hankel matrix $H_j(y)$ when using the bases of monomials and divided monomials respectively. In particular, $\operatorname{Ker} H_j(y) = (F^\perp)_j \subseteq \mathbb{K}[z_0, z_1]_j$.

(ii) *If the minimal recurrence relation a has order r, then every relation b of order $\leq d - r + 1$ is a multiple of the minimal one, i. e. $a(z) \mid b(z)$.*

Proof. The first part is a reformulation of Lemma 1.2.3(ii). The second assertion can be deduced from the structure of the apolar ideal $F^\perp = \langle g_1, g_2 \rangle$, as $\operatorname{Ker} H_j(y) = (F^\perp)_j$. Since the second generator is in degree $\deg g_2 = d + 2 - r$ (Example 1.2.9), any element in earlier degrees is a multiple of g_1. □

The algebraic geometry of Hankel matrices is manifested in the Hankel variety.

Definition 5.2.13. Fix $d \in \mathbb{N}$ and $0 \leq r \leq d/2$. The *Hankel variety* $\mathcal{X}_{d,r}$ is the variety of sequences $y \in \mathbb{K}^{d+1}$ of rank $\leq r$:

$$\mathcal{X}_{d,r} \coloneqq D_r(H_r) = \{\, y \mid \operatorname{rank} H_r(y) \leq r \,\} \subseteq \mathbb{A}^{d+1}.$$

Its main properties are summarized in the following result:

Theorem 5.2.14. *Let $0 \leq r \leq d/2$ and consider the Hankel variety $\mathcal{X}_{d,r} \subseteq \mathbb{A}^{d+1}$.*

(i) *$\mathcal{X}_{d,r}$ is irreducible, has dimension $2r$ and $(\mathcal{X}_{d,r})_{\mathrm{sing}} = \mathcal{X}_{d,r-1}$.*

(ii) *The prime ideal $\mathrm{I}(\mathcal{X}_{d,r}) \subseteq \mathbb{K}[y_0, \ldots, y_d]$ is generated by the $(r+1)$-minors of $H_r(y)$.*

(iii) *$\mathbb{P}(\mathcal{X}_{d,r})$ is the r-th secant variety of the rational normal curve $\sigma_r\nu_d(\mathbb{P}^1) \subseteq \mathbb{P}^d$.*

(iv) *The ideal of minors $I_{r+1}(H_j(y))$ in* (ii) *is independent of the choice $r \leq j \leq d - j$ (so that $H_j(y)$ has at least $r + 1$ rows and columns).*

Proof. We begin with statement (iv). We give a proof "up to radical", for a proof of the full statement see [Wat97, Proposition 5]. Let y satisfy a minimal recurrence relation of degree s, then from Lemma 5.2.12(ii) we see

$$\operatorname{rank} H_j(y) = \begin{cases} s & \text{if } s \leq j \leq \lceil d/2 \rceil \\ j+1 & \text{if } j < s. \end{cases}$$

Thus, $\operatorname{rank} H_r(y) \leq r$ if and only if $\operatorname{rank} H_j(y) \leq r$ for any $r \leq j \leq d/2$, and we can extend this to $d-r$ as $H_j(y) = H_{d-j}(y)^\mathsf{T}$. This shows that $D_r(H_j)^{\mathrm{red}} = D_r(H_r)^{\mathrm{red}} \subseteq \mathbb{A}^{d+1}$ as sets, so the two ideals of minors are equal up to radical.

Our proof of (i) and (ii) follows [IK99, Theorem 1.45]. Consider the incidence variety

$$\begin{array}{ccc} \mathcal{Z} \coloneqq \{\, (a,y) \mid T(a)\cdot y = 0 \,\} & \xrightarrow{\mathrm{pr}_y} & \mathcal{X}_{d,r} \subseteq \mathbb{A}^{d+1}_y \\ \downarrow{\scriptstyle \mathrm{pr}_a} & & \\ \mathbb{P}^r_a & & \end{array}$$

The projection pr_a gives $\mathcal{Z}$ the structure of a rank r vector bundle over $\mathbb{P}^r_a$ with fibers $\operatorname{Ker} T(a)$. In particular, $\mathcal{Z}$ is smooth and irreducible of dimension $2r$. Now $\mathcal{X}_{d,r} = \mathrm{pr}_y(\mathcal{Z})$, hence $\mathcal{X}_{d,r}$ is irreducible as well.

The restriction $\mathrm{pr}_{y|\mathcal{Z}} \colon \mathcal{Z} \to \mathcal{X}_{d,r}$ is a birational morphism and has the inverse

$$\mathcal{X}_{d,r} \setminus \mathcal{X}_{d,r-1} \to \mathcal{Z}, \qquad y \mapsto (y, [\operatorname{Ker} H_r(y)]),$$

using that $\mathcal{X}_{d,r} \setminus \mathcal{X}_{d,r-1} = \{\, y \mid \operatorname{rank} H_r(y) = r \,\}$ by (iv). This shows that $\mathcal{X}_{d,r}$ also has dimension $2r$ and is smooth outside of $\mathcal{X}_{d,r-1}$. On the other hand, if $y \in \mathcal{X}_{d,r-1}$, then $y + \varepsilon \operatorname{vand}(\omega) \in \mathcal{X}_{d,r}$ for arbitrary ω. As the $\operatorname{vand}(\omega)$ are linearly independent, this shows $\dim_{\mathbb{K}} T_y \mathcal{X}_{d,r} = d+1 > 2r$, so $y \in (\mathcal{X}_{d,r})_{\mathrm{sing}}$.

For (ii), note that the ideal of minors defines a scheme of the expected dimension, hence is Cohen–Macaulay by Theorem 0.1.25. Restricting to $\mathbb{A}^{d+1} \setminus \mathcal{X}_{d,r-1}$, this scheme is reduced by (i) and Proposition 0.4.15, as $\operatorname{rank} H_r(y)$ drops only by 1. By the Unmixedness Theorem 0.1.26, the ideal of minors is reduced everywhere.

To prove (iii), recall that the rational normal curve is the image of the embedding

$$\nu_d \colon \mathbb{P}^1 \hookrightarrow \mathbb{P}^d, \qquad \nu([z_0 : z_1]) = [z_0^d : z_0^{d-1} z_1 : \cdots : z_1^d] = [\operatorname{vand}(z_1/z_0)].$$

In other words, $\nu_d(\mathbb{P}^1)$ coincides with the set of Vandermonde vectors in $\mathbb{P}^d$. By Proposition 5.2.3, general points in $\mathcal{X}_{d,r}$ are of the form $\lambda_1 \operatorname{vand}(\omega_1) + \cdots + \lambda_r \operatorname{vand}(\omega_r)$, $\lambda_i \in \mathbb{K}$, $\omega_i \in \mathbb{P}^1(\mathbb{K})$. By the previous observation, these are exactly points on r-secants to $\nu_d(\mathbb{P}^1)$. □

We now turn to the case of infinite sequences $y \in \mathbb{K}^\infty$. Lemma 5.2.10 also applies to infinite Hankel matrices, hence $\operatorname{rank} H_r(y) \le r$ still characterizes solutions to recurrence relations. Another characterization is given by Kronecker's theorem on the rationality of generating series of finite-rank sequences y.

Theorem 5.2.15 (Kronecker [Kro81])**.** *The following are equivalent:*

(a) *y satisfies a linear recurrence relation of minimal order r.*

(b) *The formal power series $Y(z) \coloneqq \sum_{k=0}^\infty y_k z^{-k} \in \mathbb{K}[\![z^{-1}]\!]$ is a rational function: $Y(z) = z \cdot b(z)/a(z)$, $\gcd(a,b) = 1$, and $r = \deg a > \deg b$.*

The polynomial $a(z)$ in (b) *is the characteristic polynomial of the minimal order recurrence relation.*

Proof. Let $a(z)$ be the characteristic polynomial of the recurrence and observe

$$Y(z) \cdot a(z) = \sum_{k=-r}^{\infty} (a_0 y_k + \cdots + a_r y_{k+r})\, z^{-k} \in \mathbb{K}(\!(z^{-1})\!).$$

with the convention $y_k = 0$ for $k < 0$. Thus the condition that y satisfies the relation a is equivalent to the statement that $Y(z) \cdot a(z)$ has only terms of positive degree:

$$Y(z) \cdot a(z) = z \cdot b(z) \in z\mathbb{K}[z].$$

In particular, $Y(z) = z \cdot b(z)/a(z)$ is rational and $\deg b(z) \leq r - 1$ by construction.

Conversely, any such representation of $Y(z) = z \cdot b(z)/a(z)$ as a quotient of polynomials witnesses that y satisfies the recurrence relation a. This also shows that $a(z)$ and $b(z)$ must be coprime if a is the minimal relation. □

The rational function used in Kronecker's theorem is the $\mathcal{Z}$-transform of the sequence y, a useful tool to study linear difference equations.

Definition 5.2.16 ($\mathcal{Z}$-transform)**.** Let $y \in \ell^2(\mathbb{C})$, then the *$\mathcal{Z}$-transform* of y is the analytic function

$$Y(z) = \mathcal{Z}\{y\}(z) \coloneqq \sum_{k=0}^{\infty} y_k z^{-k}.$$

defined on a neighborhood of the unit circle.

5.2.3 Low-rank Hankel matrix approximation

In applications such as system identification, one is interested in approximating a given noisy input $y \in \mathbb{R}^{d+1}$ by a simpler vector $\hat{y} \in \mathbb{R}^{d+1}$, where simple means $\hat{y}$ obeys a linear recurrence relation of low order, that is, $\hat{y} \in \mathcal{X}_{d,r}$ for some desired r. This problem not only depends on the input y and the rank r, but also on a norm $\|\cdot\|$ on a space of input sequences in $\mathbb{R}^{d+1}$. A reference for (infinite) Hankel matrix approximation in the context of signal processing and control theory is the book by Regalia [Reg18].

We first consider the finite-dimensional case $d < \infty$. Let Q be a quadratic form represented by a symmetric positive-definite matrix $\Lambda \in \operatorname{Sym}^2 \mathbb{R}^{d+1}$. Given $y \in \mathbb{R}^{d+1}$, the task is

$$\underset{\hat{y} \in \mathbb{R}^{d+1}}{\text{minimize}}\, Q(\hat{y} - y) \qquad \text{subject to } \hat{y} \in \mathcal{X}_r(\mathbb{R}).$$

This is exactly the setting of Euclidean distance degree — if we exclude solutions in the singular locus $(\mathcal{X}_{d,r})_{\text{sing}} = \mathcal{X}_{d,r-1}$ (by Theorem 5.2.14(i)). Thus we arrive at the following problem:

Problem 5.2.17. Can lower rank solutions be optimal for this minimization problem? What is the Euclidean distance degree $\operatorname{EDD}_\Lambda(\mathcal{X}_{d,r})$ for interesting or general Λ? ⁇

The Euclidean distance degree of $\mathcal{X}_{d,r}$ has been considered by several authors [OSS14; Pan25; Koz+24], and is well-understood only if $r = 1$. Our approach will be more

elementary and in many cases gives an algorithmic way to compute $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$ for specific Λ.

We now return to the infinite-dimensional setting. A natural normed space of sequences are square-summable sequences $y \in \ell^2$. We can characterize finite-rank sequences in ℓ^2 as follows:

Lemma 5.2.18. *Let $y \in \mathbb{C}^\infty$ satisfy a linear recurrence relation $a(z)$ of minimal order. Then $y \in \ell^2$ if and only if $a(z)$ is stable.*

Proof. The ℓ^2-convergence of confluent Vandermonde vectors with $|\omega| < 1$ follows, for example, from the ratio test. This shows the reverse direction, for the forward direction apply Proposition 5.2.6. □

Combining this result with Kronecker's Theorem 5.2.15 shows that a natural para-meter space for finite-rank sequences $y \in \ell^2$ is

$$\{\, (a,b) \in \mathbb{C}[z]_{\le r} \times \mathbb{C}[z]_{<r} \mid a \text{ stable} \,\} \subseteq \mathbb{C}^{r+1} \times \mathbb{C}^r \qquad \leftrightarrow \qquad y \mathrel{\hat{=}} Y(z) = \frac{zb(z)}{a(z)}.$$

Given $y \in \ell^2(\mathbb{R})$, the optimization task is to minimize the ℓ^2-distance to a sequence $\hat{y}$ of lower rank:

$$\underset{\hat{y} \in \ell^2(\mathbb{R})}{\text{minimize}} \, \|\hat{y} - y\|_{\ell^2} \qquad \text{subject to } \hat{y} \text{ satisfying a recurrence of order } r.$$

The solution space can be modeled by pairs $\hat{y} \mathrel{\hat{=}} (a(z), b(z)) \in \mathbb{R}[z]_{\le r} \times \mathbb{R}[z]_{<r}$ (subject to the stability condition), hence in a finite-dimensional affine space. By fixing the order $R \ge r$ of the recurrence of y, we can parametrize the input sequences y by a finite-dimensional variety as well. Thus the following infinite-dimensional analogue of the ED degree question makes sense:

Problem 5.2.19. What are first order necessary conditions of optimality for $\hat{y}$? For general inputs y, is the set of critical points finite and non-degenerate? What is the number of (complex) critical points, depending on r and R? ⍰

Remark 5.2.20. Note that $\|y - \hat{y}\|_{\ell^2} = \|y[k] - \hat{y}[k]\|_{\ell^2}$, where $y[k]_i = y_{i+k}$ is the same sequence shifted by k. Shifting the sequence corresponds to multiplying the $\mathcal{Z}$-transform by z^{-k}. From this point of view, we might as well use b/a instead of $z \cdot b/a$ as the rational function representing y. However, the choice of having a factor of z will play a subtle role in the statement and proof of Theorem 5.5.3 below. Regardless of which convention one uses, one would arrive at the same polynomial system in Theorem 5.5.4. ◁

5.2.4 Applications in dynamical systems

In this section we briefly give an outlook on how low-rank Hankel approximation problem is of interest to mathematical engineers in the context of dynamical systems.

Dynamical systems

The linear recurrence relation described by $a(z)$ can be thought of as the characteristic equation describing an autonomous, single-output linear time-invariant (LTI) dynamical system or order r. Here,

- *Autonomous* means that the system is not influenced by external factors.
- *Linear of order r* indicates that the output signal y obeys a linear recurrence relation of order r.
- The system is said to be *time-invariant* because the system parameters — the coefficients of $a(z)$ — are independent of time.

An output signal y is said to be *model-compliant* with respect to the system described by a if y satisfies the recurrence relation a. The space of all such model-compliant signals is the so-called *behavior* of the system [Wil86], which in our notation corresponds to the r-dimensional kernel $\operatorname{Ker} T(a) \subseteq \mathbb{R}^{d+1}$.

A state-space description of such a single-output autonomous LTI system takes the form

$$x_{k+1} = Ax_k, \qquad y_k = C^\mathsf{T} x_k = C^\mathsf{T} A^k x_0,$$

with $x_0, x_1, \ldots \in \mathbb{R}^r$ the state vectors, $A \in \mathbb{R}^{r\times r}$ the state-transition matrix and $C \in \mathbb{R}^r$ the output vector. If A is diagonalizable, then we see that y is a sum of Vandermonde vectors of the eigenvalues of A, and similar for the general case (using confluent Vandermonde vectors). In particular, if the description is minimal (that is, r is minimal among state-space descriptions with the same system behavior), then $\operatorname{charpol}_A(z) = a(z)$.

To extend the model class to allow for an input $(u_k)_{k\in\mathbb{N}}$ at each time step k, one considers discrete-time single-input single-output (SISO) LTI models of finite order r. They have a state space description

$$x_{k+1} = Ax_k + B \cdot u_k, \qquad y_k = C^\mathsf{T} x_k + d \cdot u_k,$$

where $B \in \mathbb{R}^r$ and $d \in \mathbb{R}$. If $x_0 = 0$ and $u_k = \delta_{0,k}$, then the *impulse response* is

$$\operatorname{response}[\delta_0] = y_k = \begin{cases} d & k = 0 \\ C^\mathsf{T} A^{k-1} B & k > 0. \end{cases}$$

Applying the $\mathcal{Z}$-transform to the sequence y reveals that the dynamics of the SISO model can equivalently be described by the *transfer function*:

$$\mathcal{Z}\{y\}(z) = H(z) \cdot \mathcal{Z}\{u\}(z), \qquad H(z) \coloneqq C^\mathsf{T}(Iz - A)^{-1}B + d = b(z)/a(z).$$

where $a(z), b(z) \in \mathbb{R}[z]_{\leq r}$. The system response to an arbitrary input u is the convolution

$$\operatorname{response}[u] = \operatorname{response}[\delta_0] * u,$$

as the $\mathcal{Z}$-transform turns convolution into multiplication and $\mathcal{Z}(\delta_0) = 1$.

System identification

System identification infers mathematical models from observed data, which is often assumed to be generated by an underlying dynamical system. In reality, the observed data $y \in \mathbb{R}^{d+1}$ is almost always corrupted by external disturbances such as measurement noise, so that exact identification is in practice almost never possible. Another obstruction to exact identification is model mismatch, that is, the chosen model (order) does not match the true underlying model. One approach, the so-called misfit modeling setup, is to modify the observed data as little as possible by subtracting misfits $\tilde{y}$, so that the corrected data $\hat{y} = y - \tilde{y}$ become model-compliant with respect to an instance of the prescribed model class. The least squares criterion $\|\tilde{y}\|_2 = \|y - \hat{y}\|_2$ is a natural choice, as explored for example in [RH95].

The low-rank Hankel matrix approximation problem under consideration formulates such an identification task: Given a noisy observed sequence of data y, retrieve the closest sequence $\hat{y}$ of finite rank $\leq r$. The parameters a of the corresponding autonomous, single-output LTI model of order r generating $\hat{y}$ can then be found in the kernel of $H_r(y)$.

This problem appears in a wide variety of data driven engineering fields, such as direction-of-arrival estimation in array processing, shape-from-moments problems, and realization algorithms for converting impulse response samples into state-space models, see for example [Moo20, Section 10] and references therein. Engineers typically resort to various iterative approaches to obtain estimates of $\hat{y}$ and a. The first idea goes back to Prony in 1795 [Pro95], see for example [HDS90] for more information.

Model order reduction

The approximation of a given higher-order discrete-time SISO system by one of lower order that approximates its input-output behavior as accurately as possible is known as *SISO model order reduction* (MOR). Since the impulse response of a SISO model completely encodes its dynamics (arbitrary responses can be recovered via convolution as above), the higher-order model is often approximated via its impulse response. The ℓ_2-norm of the impulse response difference is then a natural approximation criterion. When expressed using the transfer functions of the models, this leads to a rational function approximation problem (Section 5.5.1).

Reducing the model order while maintaining accuracy (in relevant input and output regimes), one obtains a simpler model that is cheaper to simulate and analyze in practice. The main use cases of model reduction arise in engineering applications where first-principles modeling leads to very large state-space descriptions, such as fluid dynamics, circuit simulation, or structural mechanics, making direct computation too expensive. State-of-the-art techniques solve the MOR problem by iteratively computing a solution that satisfies the first-order necessary conditions for optimality. See [GAB08] for a mathematical treatment, presenting a framework unifying various approaches.

5.3 Euclidean distance degrees of Hankel matrices

In this section we study Problem 5.2.17 by transforming the critical equations of the optimization problem into a system of determinantal equations. Using Thom–Porteous theorem, we can calculate the generic Euclidean distance degree, and provide an algorithmic approach to calculate the ED defect in many cases. The derivation of the system of equations (for unit weights) has appeared in [LVD26], the main novel contribution is a mathematical analysis of finiteness and number of complex critical points, as well as the generalization to arbitrary weight matrices.

5.3.1 The critical equations

Motivated by Lemma 5.2.12, we use the following notation:

- Let $S_d = \mathbb{C}[z]_{\leq d}$ be the vector space of univariate polynomials of degree at most d, thought of as the dehomogenization of $\mathrm{S}^d \mathbb{C}^2 \cong \mathbb{C}[z_0, z_1]_d$. It is equipped with the standard basis of monomials.
- The given data y and the low-rank approximation $\hat{y}$ are in the dual space $S_d^\vee \cong \mathbb{A}^{d+1}$ (or the real points thereof). This can be thought of as a space of symmetric tensors, or of polynomials in the dual variables. In particular, $\mathcal{X}_{d,r} \subseteq S_d^\vee$, and the Hankel matrix $H_r(\hat{y})$ is a linear map $S_r \to S_{d-r}^\vee$ with $H_r(\hat{y})^\mathsf{T} = H_{d-r}(\hat{y})$.
- The recurrence relation a is an element of S_r. The Toeplitz matrix $T_d(a) \colon S_d^\vee \to S_{d-r}^\vee$ is *dual* to the multiplication map $T_d(a)^\mathsf{T} = (f \mapsto af) \colon S_{d-r} \to S_d$.
- The quadratic form $Q \in \mathrm{S}^2(S_d)$ is given by an invertible map $\Lambda \colon S_d^\vee \to S_d$ with $\Lambda^\mathsf{T} = \Lambda$, $Q(y) = \frac{1}{2} y^\mathsf{T} \Lambda y$. In the given bases, it is represented by a symmetric matrix.

We start with the observation that the incidence from the proof of Theorem 5.2.14

$$\mathcal{Z} = \{\, (a, \hat{y}) \mid H_r(\hat{y}) \cdot a = 0 \,\} \xrightarrow{\mathrm{pr}_{\hat{y}}} \mathcal{X}_{d,r} \subseteq S_d^\vee$$

is a smooth variety whose dense open set $\{\, (a, \hat{y}) \mid \operatorname{rank} H_r(\hat{y}) = r \,\} = \mathrm{pr}_{\hat{y}}^{-1}((\mathcal{X}_{d,r})_{\mathrm{reg}})$ is isomorphic to the regular locus of $\mathcal{X}_{d,r}$. This allows us to consider the equivalent constrained optimization problem over the real locus of the smooth variety $\mathcal{Z}$:

$$\underset{\substack{\hat{y} \in S_d^\vee(\mathbb{R}) \\ 0 \neq a \in S_r}}{\text{minimize}}\, Q(\hat{y} - y) \qquad \text{subject to } H_r(\hat{y}) \cdot a = 0.$$

Using Lagrange multiplies, we turn this problem into an unconstrained optimization problem. The Lagrange function is

$$\mathcal{L}_y(\hat{y}, a, \ell) = Q(\hat{y} - y) + \ell \cdot H_r(\hat{y}) \cdot a \overset{5.2.10}{=} Q(\hat{y} - y) + \ell \cdot T(a) \cdot \hat{y},$$

where the unknown vector of Lagrange multipliers ℓ naturally resides in S_{d-r}. Taking the gradient of $\mathcal{L}_y$, we obtain the critical equations

$$0 = \frac{\partial \mathcal{L}_y}{\partial \hat{y}} = \Lambda(\hat{y} - y) + T(a)^\mathsf{T} \cdot \ell \tag{5.1a}$$

$$0 = \frac{\partial \mathcal{L}_y}{\partial a} = H_r(\hat{y})^\mathsf{T} \cdot \ell = H_{d-r}(\hat{y}) \cdot \ell \overset{5.2.10}{=} T(\ell) \cdot \hat{y} \tag{5.1b}$$

$$0 = \frac{\partial \mathcal{L}_y}{\partial \ell} = T(a) \cdot \hat{y}. \tag{5.1c}$$

The projection $(\hat{y}, a) \mapsto \hat{y}$ identifies (complex) critical points of this problem of rank r with that of the EDD problem for $\mathcal{X}_{d,r}$. However, this formulation may (and will) have more than $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$ critical points, since this system of equations allows $\hat{y} \in (\mathcal{X}_{d,r})_{\mathrm{sing}}$. In fact, the incidence variety of $(y, \hat{y}, a, \ell)$ splits into $r+1$ components, corresponding to the different values of $\operatorname{rank} H_r(\hat{y})$, as described in the following theorem:

Theorem 5.3.1. *The scheme defined by the Lagrange equations*

$$\{\, (y, \hat{y}, a, \ell) \mid \text{eqs. (5.1a) to (5.1c) hold} \,\} \subseteq S_d^\vee \times S_d^\vee \times (S_r \setminus 0) \times S_{d-r}$$

is reduced and has $r+1$ irreducible components corresponding to $0 \leq \operatorname{rank} H_r(\hat{y}) \leq r$. It is smooth outside the intersection of two components and each component is a variety of dimension $d+2$.

Note that eq. (5.1a) allows us to immediately eliminate $y = \hat{y} + \Lambda^{-1} T(a)^\mathsf{T} \ell$ from the equations. Thus it suffices to study the following incidence $\mathcal{K}$ instead:

$$\begin{aligned} \mathcal{K} &:= \{\, (\hat{y}, a, \ell) \mid H_r(\hat{y})a = 0,\ H_{d-r}(\hat{y})\ell = 0 \,\} && \subseteq S_d^\vee \times (S_r \setminus 0) \times S_{d-r} \\ &\qquad\qquad \big\downarrow \mathrm{pr} \\ \widehat{\mathcal{Z}} &:= \{\, (\hat{y}, a) \mid H_r(\hat{y})a = 0 \,\} && \subseteq S_d^\vee \times (S_r \setminus 0) \end{aligned}$$

Proof. The proof will proceed in two steps: We first show that $\mathcal{K}$ is a local complete intersection of codimension $d+1$, and along the way obtain a description of the irreducible components. In the second step we characterize the smooth points of $\mathcal{K}$, which yields reducedness from the Unmixedness Theorem 0.1.26.

Consider the locally closed subsets $\mathcal{K}_{=s}, \widehat{\mathcal{Z}}_{=s}$ where $\operatorname{rank} H_r(\hat{y}) = s$. Then $\widehat{\mathcal{Z}}_{=s}$ is a rank s vector bundle over $S_r \setminus 0$ and $\mathcal{K}_{=s}$ is a bundle over $\widehat{\mathcal{Z}}_{=s}$ of rank $d - r + 1 - s$. From this we deduce that the locally closed set $\mathcal{K}_{=s}$ is irreducible of dimension

$$\dim \mathcal{K}_{=s} = \dim(S_r \setminus 0) + \operatorname{rank}_{S_r \setminus 0} \widehat{\mathcal{Z}}_{=s} + \operatorname{rank}_{\widehat{\mathcal{Z}}_{=s}} \mathcal{K}_{=s} = d + 2.$$

On the open set where $a_i \neq 0$, the i-th row of $H_{d-r}(\hat{y})$ is a linear combination of the others, so it can be omitted from the equation $H_{d-r}(\hat{y}) \cdot \ell = 0$. This shows that $\mathcal{K}$ is locally defined by $\operatorname{codim} \mathcal{K} = d+1$ equations, hence a local complete intersection. It also follows that the $\mathcal{K}_s := \overline{\mathcal{K}_{=s}}$ $(s = 0, \ldots, r)$ form the irreducible components of $\mathcal{K}$.

Let $s = \operatorname{rank} H_r(\hat{y})$, so $\hat{y}$ satisfies a minimal recurrence relation $b(z)$ of order $0 \leq s \leq r$. By Lemma 5.2.12, $b(z) \mid \gcd(\ell(z), a(z))$, we will show that $(\hat{y}, a, \ell) \in \mathcal{K}_{\text{reg}}$ if and only if $b(z) = \gcd(\ell(z), a(z))$. To see that this implies the smoothness claim in the theorem, note that if $(\hat{y}, a, \ell)$ is a singular point, then pick a confluent Vandermonde vector $v = \operatorname{vand}^{(j)}(\omega)$ corresponding to an additional common root of $a(z)$ and $\ell(z)$ (with multiplicity $j+1$). Then $(\hat{y} + \varepsilon v, a, \ell) \in \mathcal{K}_{=s+1}$, showing that $(\hat{y}, a, \ell) \in \mathcal{K}_{=s} \cap \overline{\mathcal{K}_{=s+1}}$.

To test smoothness, consider the Jacobi matrix at $(\hat{y}, a, \ell) \in \mathcal{K}$

$$J = J_{(H_r(\hat{y})a, H_{d-r}(\hat{y})\ell)}(\hat{y}, a, \ell) = \begin{array}{c|c|c} \hat{y} & a & \ell \\ \hline T(a) & H_r(\hat{y}) & 0 \\ \hline T(\ell) & 0 & H_{d-r}(\hat{y}) \end{array}$$

By the Jacobi criterion, $(\hat{y}, a, \ell)$ is a smooth point if and only if the matrix has rank $d+1$, equivalently $\dim_{\mathbb{C}} \operatorname{Ker} J^{\mathsf{T}}$. The condition for $(f, g) \in S_{d-r} \oplus S_r$ to satisfy $(f, g) \cdot J = 0$ is

$$f \cdot a + g \cdot \ell = 0, \qquad H_{d-r}(\hat{y}) f = 0, \qquad H_r(\hat{y}) g = 0.$$

By Lemma 5.2.12 the last two conditions are equivalent to $b(z) = \gcd(f(z), g(z))$; let $\tilde{a}, \tilde{\ell}, \tilde{f}, \tilde{g}$ be the respective quotients by $b(z)$ (e. g. $\tilde{a}(z) = a(z)/b(z)$). Then elements in the left kernel are parametrized by

$$\operatorname{Ker} J^{\mathsf{T}} = \left\{ (b \cdot \tilde{f}, b \cdot \tilde{g}) \;\middle|\; \tilde{f}\tilde{a} + \tilde{g}\tilde{\ell} = 0 \right\}.$$

By Bézout's lemma this has a unique solution up to scaling if and only if $\gcd(\tilde{a}, \tilde{\ell}) = 1$, that is, $\gcd(a(z), \ell(z)) = b(z)$. □

Given y, we can rearrange eq. (5.1a) to $\hat{y} = y - \Lambda^{-1}(a \cdot \ell)$, so $\hat{y}$ is determined by a and ℓ. In fact, under a mild assumption, we can express both $\hat{y}$ and ℓ in terms of a.

Lemma 5.3.2. *Assuming the matrix $C \coloneqq T(a)\Lambda^{-1}T(a)^{\mathsf{T}}$ is invertible, we have*

$$\begin{aligned} \ell &= C^{-1}\, T(a)\, y \\ \hat{y} &= y - \Lambda^{-1}\, T(a)^{\mathsf{T}}\, C^{-1}\, T(a)\, y. \end{aligned}$$

This assumption is satisfied for real nonzero a and Λ positive definite. In that case $\hat{y}$ is the Λ-orthogonal projection of y onto $\operatorname{Ker} T(a)$.

Proof. We can substitute eq. (5.1a) into eq. (5.1c) to obtain

$$0 \overset{(5.1c)}{=} T(a)\hat{y} = T(a)y + T(a)(\hat{y} - y) \overset{(5.1a)}{=} T(a)y - T(a)\Lambda^{-1}T(a)^{\mathsf{T}} \cdot \ell.$$

Multiplying with $C^{-1} = (T(a)\Lambda^{-1}T(a)^{\mathsf{T}})^{-1}$ yields the first formula. The second is obtained by substituting the first into eq. (5.1a).

If a is real, then $T(a)\Lambda^{-1}T(a)^{\mathsf{T}}$ is positive definite (hence invertible), since $T(a)^{\mathsf{T}}$ is injective for any $a \neq 0$. Lastly, the formula for $\hat{y}$ is exactly the formula for the Λ-orthogonal projection onto the kernel of $T(a)$. □

In principle, this allows us to eliminate $\hat{y}$ and ℓ from the critical equations. Before doing so, we need to address the following questions

(1) In the real setting, can lower-rank solutions be (local) minimizers of $Q(\hat{y} - y)$?

(2) In view of Theorem 5.3.1, how can we characterize lower-rank solutions in terms of a and ℓ?

(3) Can we do the elimination without relying on $T(a)\Lambda^{-1}T(a)^\mathsf{T}$ being invertible?

The answer to the first question is negative (Theorem 5.3.3), and our answer to (2) in Lemma 5.3.4 will lead us to an elegant solution to (3) in the form of Theorem 5.3.5.

5.3.2 Characterizing lower-rank solutions

Theorem 5.3.3. *Let $y \in S_d^\vee(\mathbb{R}) \setminus \mathcal{X}_{d,r-1}(\mathbb{R})$ be* any *real point outside $X_{d,r-1}$, let $\hat{y} \in \mathcal{X}_{d,r-1}(\mathbb{R})$, and assume that Λ is positive definite. Then $\hat{y}$ is* not *a local minimum of the function $x \mapsto Q(x - y)$ on $\mathcal{X}_{d,r}(\mathbb{R})$.*

This result has appeared in the case of unit weights in [LVD26, Lemma 3], we give a generalized variant of the proof.

Proof. By decreasing r if necessary, we may assume that $\operatorname{rank} H_r(\hat{y}) = r - 1$, and let $a \in S_{r-1}$ be a non-zero kernel vector of $H_{r-1}(\hat{y})$. Assume for a contradiction that $\hat{y}$ was a local minimizer.

Let $\omega \in \mathbb{R}$ with $a(\omega) \neq 0$, $a^*(z) := a(z) \cdot (z - \omega)$ and let $\operatorname{vand}(\omega)$ be the corresponding Vandermonde vector. Let $P_a \colon S_d^\vee \to \operatorname{Ker} T(a)$ be the Λ-orthogonal projection onto $\operatorname{Ker} T_d(a)$ and let

$$u := P_a(\operatorname{vand}(\omega)) - \operatorname{vand}(\omega) \in \operatorname{Ker} T_d(a^*), \qquad u \perp_\Lambda \operatorname{Ker} T(a)$$

be the component of $\operatorname{vand}(\omega)$ orthogonal to $\operatorname{Ker} T_d(a)$. Thus the orthogonal projection P_{a^*} onto $\operatorname{Ker} T_d(a^*)$ is the sum of the projections onto $\operatorname{Ker} T_d(a)$ and $\langle u \rangle_\mathbb{R}$.

Since $T_d(a^*)\hat{y} = 0$ and $\hat{y}$ is a local minimum on $\mathcal{X}_{d,r}$, $(\hat{y}, a^*)$ is a real solution to the critical equations. In particular, by Lemma 5.3.2 $\hat{y} = P_{a^*} y = P_a y$, so the orthogonal projection onto $\langle u \rangle_\mathbb{R}$ must be zero, that is, $u \perp_\Lambda y = 0$.

By eq. (5.1a), $\Lambda(\hat{y} - y) \in \operatorname{Im} T_d(a)^\mathsf{T} = (\operatorname{Ker} T_d(a))^\perp$, that is, $\hat{y} - y \perp_\Lambda \operatorname{Ker} T_d(a)$. Furthermore $\hat{y} \perp_\Lambda u$ by choice of u, so we get

$$\langle y - \hat{y}, \operatorname{vand}(\omega) \rangle_\Lambda = \langle y - \hat{y}, P_a \operatorname{vand}(\omega) - u \rangle_\Lambda = \langle y - \hat{y}, -u \rangle_\Lambda = 0.$$

This implies $y = \hat{y}$, as $\omega \in \mathbb{R} \setminus \{a(\omega) = 0\}$ was arbitrary and the set of Vandermonde vectors spans the ambient space $S_d^\vee(\mathbb{R})$. But this is impossible, as $\operatorname{rank} H_r(y) \geq r > \operatorname{rank} H_r(\hat{y})$. □

This shows that both from the real optimization as well as the EDD perspective, lower-rank solutions should be removed. They can be detected by the property of the Lagrange vector ℓ as follows:

Lemma 5.3.4. *Assume* $(\hat{y}, a, \ell)$ *satisfy eqs.* (5.1a) *to* (5.1c).

(i) *If* $\operatorname{rank} H_r(\hat{y}) = r$, *then* $\ell \in S_{d-r}$ *is a multiple of* $a \in S_r$.

(ii) *Conversely, if* ℓ *is a multiple of* a *and the data vector* $y \in S_d^\vee$ *is sufficiently general, then* $\operatorname{rank} H_r(\hat{y}) = r$.

Proof. Since $\hat{y}$ satisfies the recurrence $a(z)$, it also satisfies any $a^*(z) = a(z) \cdot g(z)$. In other words, $S_{d-2r}a \subseteq \operatorname{Ker} H_{d-r}(\hat{y})$. If $\operatorname{rank} H_r(\hat{y}) = r$, then

$$\dim_{\mathbb{C}} \operatorname{Ker} H_{d-r}(\hat{y}) = d - r + 1 - \operatorname{rank} H_{d-r}(\hat{y}) = d - 2r + 1 = \dim_{\mathbb{C}} S_{d-2r}a,$$

so equality holds. By eq. (5.1b) $\ell \in \operatorname{Ker} H_{d-r}(\hat{y})$, so it is a multiple of a.

For the converse we need some genericity of y. By the description of the irreducible components in Theorem 5.3.1, smooth points $(\hat{y}, a, \ell) \in \mathcal{K}_{=s}$ have the property that $\hat{y}$ satisfies a relation $b(z)$ of order s and $b(z) = \gcd(a(z), \ell(z))$. If $s < r$, then this means $a(z) \nmid \ell(z)$. Thus, adding $a(z) \mid \ell(z)$ to the Lagrange equations, all but the component generically of $\operatorname{rank} H_r(\hat{y})$ get cut down to dimension $\leq d + 1$. Since the solutions are invariant under scaling $(y, \hat{y}, \lambda a, \frac{1}{\lambda}\ell)$, the solution set for any given y are at least one-dimensional, so the projection

$$\mathrm{pr}_y \colon \{\, (y, \hat{y}, a, \ell) \mid \text{eqs. (5.1a) to (5.1c) hold, } \ell(z) \mid a(z) \text{ and } \operatorname{rank} H_r(\hat{y}) < r \,\} \to S_d^\vee$$

is not dominant by this dimension argument. □

Thus, to remove the lower-rank solutions, we can substitute $\ell = a \cdot g$ for $g \in S_{d-2r}$. This transforms the critical equations as follows:

$$0 = \Lambda(\hat{y} - y) + T_d(a)^\mathsf{T} T_{d-r}(a)^\mathsf{T} g \tag{5.2a}$$
$$0 = T_{d-r}(g) \cdot T_d(a) \cdot \hat{y} \tag{5.2b}$$
$$0 = T_d(a) \cdot \hat{y}. \tag{5.2c}$$

In this formulation, eq. (5.2b) is a direct consequence of eq. (5.2c), so we may omit it. We can rearrange eq. (5.2a) to $\hat{y} = y - \Lambda^{-1}(a^2 g)$ and substitute it into eq. (5.2c), to arrive at the following reformulation:

Theorem 5.3.5. *For general* y, *the complex critical points* $(\hat{y}, a, \ell)$ *with* $\operatorname{rank} H_r(\hat{y}) = r$ *are in bijection to the solutions* $(a, g) \in S_r \times S_{d-2r}$ *of the system*

$$T_d(a) y = T_d(a) \Lambda^{-1} T_d(a^2)^\mathsf{T} g. \tag{5.3}$$

The bijection is given by $\ell = a \cdot g$ *and* $\hat{y} = y - \Lambda^{-1}(a^2 \cdot g)$. *If* y *and* Λ *are real, then this bijection preserves real solutions.*

In particular, the number of complex solutions to eq. (5.3) *equals* $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$.

With this perspective, we can also explain the solution count to the Lagrange system including lower-rank solutions (modulo the rescaling action of $\mathbb{C}^\times$):

Theorem 5.3.6. *For general y, the number of solutions to eqs. (5.1a) to (5.1c) is*

$$\sum_{s=0}^{r} \binom{d-2s}{r-s} \mathrm{EDD}_\Lambda(\mathcal{X}_{d,s}).$$

Proof. Let $y \in S_d^\vee$ be general and $\mathcal{L}_{=s}(\mathcal{X}_{d,r})$ be the subset of the solutions to the Lagrange equations for which $\operatorname{rank} H_r(\hat{y}) = s$. By (the proof of) Theorem 5.3.1, every $(\hat{y}, a, \ell) \in \mathcal{L}_{=s}(\mathcal{X}_{d,r})$ yields a rank s solution to the critical equations for $\mathcal{X}_{d,s}$ via

$$\mathcal{L}_{=s}(\mathcal{X}_{d,r}) \ni (\hat{y}, a, \ell) \mapsto \left(\hat{y}, \gcd(a,\ell), \frac{a \cdot \ell}{\gcd(a,\ell)}\right) \in \mathcal{L}_{=s}(\mathcal{X}_{d,s}).$$

Conversely, if $(\hat{y}, a', \ell') \in \mathcal{L}_{=s}(\mathcal{X}_{d,s})$, then we can pick $r-s$ roots of $\ell'/\gcd(a',\ell')$ and transfer them from ℓ' to a' to obtain all preimages in $\mathcal{L}_{=s}(\mathcal{X}_{d,r})$.

By genericity of y, all roots are distinct, and since $\deg \ell'/\gcd(a',\ell') = d-2s$, the number of preimages is $\binom{d-2s}{r-s}$. The rank s component $\mathcal{L}_{=s}(\mathcal{X}_{d,s})$ contains $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,s})$ solutions, hence summing over all irreducible components yields the formula. □

5.3.3 The substituted ED correspondence

In the usual ED correspondence, one considers pairs of data y and critical points $\hat{y}$. In our approach we have replaced $\hat{y}$ with a, and in this section we describe the equations of the substituted ED correspondence under mild assumptions on the weight matrix Λ.

Definition 5.3.7 ($B_\Lambda(a)$, $M_\Lambda(a,y)$, good and bad locus)**.** Define the matrices

$$B_\Lambda(a) \coloneqq T(a)\Lambda^{-1}T(a^2)^\mathsf{T} \colon S_{d-2r} \to S_{d-r}^\vee$$
$$M_\Lambda(a,y) \coloneqq [T(a)y \mid B_\Lambda(a)] \colon \mathbb{C} \oplus S_{d-2r} \to S_{d-r}^\vee.$$

The *bad locus* is the scheme $\mathcal{B} = \mathcal{B}_\Lambda \subseteq S_r$ defined by the maximal minors $I_{d-2r+1}(B_\Lambda)$. The *good locus* is its complement $\mathcal{G} = S_r \setminus \mathcal{B}$.

The good locus is dense open if $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) > 0$ as the next lemma shows, and it is generically all of $S_r \setminus 0$ (Theorem 5.4.1 below). Complementarily, the bad locus is a strict subset $\mathbb{P}\mathcal{B}_\Lambda \subsetneq \mathbb{P}S_r$ if $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) > 0$, and generically empty.

With this notation, the system eq. (5.3) can be compactly written as a *multi-parameter eigenvalue problem*

$$M_\Lambda(a,y) \cdot \begin{pmatrix} -1 \\ g \end{pmatrix} = 0.$$

This perspective was explored in previous work by one of the authors [LVD26], we will come back to this in Section 5.3.4. To make the equations homogeneous linear in $g \mathrel{\hat{=}} g_0 + \cdots + g_{d-2r} z^{d-2r} \in S_{d-2r}$, we homogenize with the additional variable g_{-1}:

Lemma 5.3.8. *Whenever $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) > 0$, for example when Λ is real positive definite, we have $\mathcal{G} \neq \emptyset$. If $\mathcal{G} \neq \emptyset$, then the homogenized incidence*

$$\mathcal{Y}_\mathcal{G} = \{\, (y, a, (g_{-1} : g)) \mid T(a)y \cdot g_{-1} = B_\Lambda(a)g \,\} \subseteq S_d^\vee \times \mathcal{G} \times \mathbb{P}(\mathbb{C} \oplus S_{d-2r}).$$

is an irreducible and smooth complete intersection of dimension $d+2$.

Proof. We first show the claim that $\mathcal{G} \neq \emptyset$ if $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) > 0$ by contraposition. Let $\mathcal{G} = \emptyset$, and let $(y, a, g) \in S_d^\vee \times S_r \times S_{d-2r}$ be a solution to eq. (5.3). By assumption, $\operatorname{rank} B_\Lambda(a) \leq d - 2r$, so $\operatorname{Ker} B_\Lambda(a) \neq 0$. Thus the solution set for a given y is invariant under the free group action

$$(\lambda, g') \cdot (y, a, g) \coloneqq (y, \lambda a, \lambda^{-2} g + g'), \qquad (\lambda, g') \in \mathbb{C}^\times \ltimes \operatorname{Ker} B_\Lambda(a).$$

In particular, the solution set for any y is either empty or of dimension ≥ 2. On the other hand, the irreducible component of the Lagrange equations where $\operatorname{rank} H_r(\hat{y}) = r$ has dimension $d + 2$ by Theorem 5.3.1, so the projection to $y \in S_d^\vee \cong \mathbb{A}^{d+1}$ can not be dominant. This shows that for general y there are *no* solutions, i. e. $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) = 0$.

Now we prove the claimed properties of $\mathcal{Y}_\mathcal{G}$ by an affine bundle argument. The projection onto (a, g) yields the open set $AG \coloneqq \mathcal{G} \times (1, S_{d-2r}) \subseteq \mathcal{G} \times \mathbb{P}(\mathbb{C} \oplus S_{d-2r})$, since for $g_{-1} = 0$ there is no solution by the maximal rank assumption on $B_\Lambda(a)$ for $a \in \mathcal{G}$. This turns the incidence $\mathcal{Y}_\mathcal{G}$ into a rank r affine bundle over AG, as any choice $(a, g) \in AG$ determines y as a solution to the affine linear system $T(a)y = B_\Lambda(a)g$. Consequently, $\mathcal{Y}_\mathcal{G}$ is a smooth variety of dimension $(r + 1) + (d - 2r + 1) + r = d + 2$. □

Remark 5.3.9. It is an interesting question to ask whether $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) > 0$ is also *necessary* for $\mathcal{G} \neq \emptyset$. We prove this in Theorem 5.4.6 below in the case $r = 1$. In fact, we are not aware of any invertible weight matrix Λ with $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) = 0$ for $r \geq 2$. ◁

From this incidence we now eliminate $(g_{-1} : g)$. Set $k \coloneqq d - 2r + 1$, then the projection of the incidence onto $\mathcal{G} \times S_d^\vee$ yields the determinantal scheme

$$\mathcal{E}_{d,r} \coloneqq D_k(M_\Lambda(a, y)) = \{\, (a, y) \mid \operatorname{rank} M_\Lambda(a, y) \leq k \,\} \subseteq \mathbb{P}\mathcal{G} \times \mathbb{P}(S_d^\vee).$$

Theorem 5.3.10. *$\mathcal{E}_{d,r}$ is a smooth variety of dimension d and codimension r. It has the following two descriptions:*

(i) *On $\mathbb{P}\mathcal{G} \times \mathbb{P}S_d^\vee \subseteq \mathbb{P}^r \times \mathbb{P}^d$, $\mathcal{E}_{d,r}$ is the degeneracy locus of the vector bundle morphism*

$$\mathcal{O}^{d-r+1} \xrightarrow{M_\Lambda(a,y)^\mathsf{T}} \mathcal{O}(1,1) \oplus \mathcal{O}(3,0)^k.$$

(ii) *It is the projective subbundle $\mathcal{E}_{d,r} = \mathbb{P}\mathcal{F} \subseteq \mathbb{P}\mathcal{O}_{\mathbb{P}\mathcal{G}}^{d+1} = \mathbb{P}\mathcal{G} \times \mathbb{P}^d$ given by*

$$\mathcal{F} \coloneqq \underbrace{\operatorname{Ker}(\mathcal{O}^{d+1} \xrightarrow{T(a)} \mathcal{O}(1)^{k+r})}_{=:\mathcal{F}_1,\ \operatorname{rank} = r} \oplus \underbrace{\operatorname{Im}(\mathcal{O}(-2)^k \xrightarrow{\Lambda^{-1}T(a^2)} \mathcal{O}^{d+1})}_{=:\mathcal{F}_2,\ \operatorname{rank} = k}.$$

Proof. The first bundle description is the definition of $\mathcal{E}_{d,r}$. To see reducedness of the ideal of minors, apply Proposition 0.4.15 to the smooth incidence Lemma 5.3.8.

For the second condition, notice that the rank condition on $\mathcal{E}_{d,r}$ is equivalent to $T_d(a)y \in \operatorname{Im} B_\Lambda(a)$, which is equivalent to $y \in \operatorname{Im} \Lambda^{-1} T_d(a^2)^\mathsf{T}$ modulo the kernel of T^a. The sum of these two is direct for any $a \in \mathcal{G}$ by the assumption on $B_\Lambda(a)$. This description as a projective subbundle also shows that $\mathcal{E}_{d,r}$ is smooth. □

The determinantal scheme $\mathcal{E}_{d,r} \subseteq \mathbb{P}\mathcal{G} \times \mathbb{P}^d$ thus has the expected codimension, hence we can apply Thom–Porteous theorem to show:

Theorem 5.3.11. *Assume $\mathbb{P}\mathcal{G} = \mathbb{P}S_r$, then $\mathcal{E}_{d,r} \subseteq \mathbb{P}^r \times \mathbb{P}^d$ is a biprojective variety of bidegree*

$$[\mathcal{E}_{d,r}] = \left\{ \frac{1}{(1-(\alpha+\beta))(1-3\alpha)^k} \right\}^r = \sum_{j=0}^{r} \binom{k-1+j}{j} 3^j \cdot \alpha^j (\alpha+\beta)^{r-j}$$
$$= \sum_{j=0}^{r} \sum_{i=0}^{j} \binom{k+r}{j-i} \binom{k-1+i}{i} 2^i \alpha^j \beta^{r-j}$$

in the Chow ring $\mathrm{A}(\mathbb{P}^r \times \mathbb{P}^d) = \mathbb{Z}[\alpha,\beta]/\langle \alpha^{r+1}, \beta^{d+1} \rangle$.

Proof. The Thom–Porteous formula in corank 1 (Corollary 0.4.13) can be applied to the description from Theorem 5.3.10(i) as the degeneracy locus of

$$M_\Lambda(a,y) \colon \mathcal{O}(-1,-1) \oplus \mathcal{O}(-3,0)^k \longrightarrow \mathcal{O}^{d-r+1}.$$

The formula states that $[D_k(M(a,y))]$ is the Segre class

$$[\mathcal{E}_{d,r}] = s_r(\mathcal{O}(-1,-1) \oplus \mathcal{O}(-3,0)^k) = \left\{ \frac{1}{(1-(\alpha+\beta))(1-3\alpha)^k} \right\}^r$$
$$= \sum_{j=0}^{r} \binom{k-1+j}{j} 3^j \cdot \alpha^j (\alpha+\beta)^{r-j}.$$

On the other hand, we can also apply the formula for classes of projective subbundles [EH16, Prop. 9.13] to the second description from Theorem 5.3.10. Here (in the notation of [EH16]) $\mathcal{E} = \mathcal{O}_{\mathbb{P}^r}^{d+1}$, $c_1(\mathcal{O}_{\mathbb{P}\mathcal{E}}(1)) = \beta$ and, using Whitney's formula,

$$c(\mathcal{E}/\mathcal{F}) = s(\mathcal{F}) = s(\mathcal{F}_1) \cdot s(\mathcal{F}_2) = c(\mathcal{O}(1)^{k+r}) \cdot s(\mathcal{O}(-2)^k)$$
$$= \left(\sum_{i=0}^{k+r} \binom{k+r}{i} \alpha^i \right) \left(\sum_{i=0}^{\infty} \binom{k-1+i}{i} 2^i \alpha^i \right).$$

This gives the second formula

$$[\mathcal{E}_{d,r}] = \sum_{j=0}^{r} c_1(\mathcal{O}_{\mathbb{P}\mathcal{E}}(1))^{r-j} \cdot c_j(\mathcal{E}/\mathcal{F}) = \sum_{j=0}^{r} \left(\sum_{i=0}^{j} \binom{k+r}{j-i} \binom{k-1+i}{i} 2^i \right) \alpha^j \beta^{r-j}. \qquad \square$$

Specializing to a general y, we can use the previous calculation to obtain the generic ED degree.

Corollary 5.3.12. (i) *If $\mathbb{P}\mathcal{B}_\Lambda = \mathbb{P}S_r$, then $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) = 0$.*

(ii) *If $\mathbb{P}\mathcal{B}_\Lambda \subsetneq \mathbb{P}S_r$, then for general y the number of solutions in $\mathcal{E}_{d,r}|_y \subseteq \mathbb{P}\mathcal{G}$ is $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$, and $g, \ell, \hat{y}$ can be uniquely reconstructed from a.*

(iii) *If* $\mathbb{P}\mathcal{B}_\Lambda = \emptyset$, *then* $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}) = \sum_{j=0}^{r} \binom{d-2r+j}{j} 3^j$.

Proof. The first two claims are restatements/consequences of Lemma 5.3.8 and Theorem 5.3.5. Intersecting the expression in Theorem 5.3.11 with d general hyperplanes in $\mathbb{P}^d$ gives the formula

$$\sum_{j=0}^{r} \binom{k-1+j}{j} 3^j \cdot \alpha^j (\alpha+\beta)^{r-j} \cdot \beta^d = \sum_{j=0}^{r} \binom{k-1+j}{j} 3^j \cdot \alpha^r \beta^d.$$

The condition of $B(a)$ not dropping rank is satisfied for generic Λ (Theorem 5.4.1 below), and so this condition is sufficient for having generic ED degree. □

Remark 5.3.13. Evaluating the other formula in Theorem 5.3.11 gives the expression reported in [OSS14, Theorem 3.7]. ◁

Corollary 5.3.14. *For general* Λ, *the number of solutions up to rescaling to the Lagrange system eqs. (5.1a) to (5.1c) is*

$$\sum_{s=0}^{r} \binom{d-2s}{r-s} \sum_{j=0}^{s} \binom{d-2s+j}{j} 3^j.$$

Proof. Evaluate the formula from Theorem 5.3.6 in the generic EDD from Corollary 5.3.12. □

5.3.4 An optimal parameter homotopy

In this subsection we show how Homotopy continuation using a polyhedral start system can solve the system of equations derived in Theorem 5.3.5 efficiently.

We continue with the assumption that $\mathbb{P}\mathcal{B}_\Lambda = \emptyset$, so $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r})$. The equations $M_\Lambda(a,y) \cdot \binom{-1}{g} = 0$ form a system of $n := d-r+1$ polynomial equations in $(a,g) \in S_r \times S_{d-2r} \cong \mathbb{C}^{n+1}$. Dehomogenizing a (for example by $a_r = 1$), we obtain a square system in n variables with isolated solutions (for general y)

$$T(a_0,\dots,a_{r-1},1) \cdot y = B_\Lambda(a_0,\dots,a_{r-1},1) \cdot \begin{pmatrix} g_0 \\ \vdots \\ g_{d-2r} \end{pmatrix}. \tag{5.4}$$

This problem formulation is far better suited for (numerical) system solving than, say, the system of minors $I_{k+1}(M_\Lambda(a,y))$. It turns out that applying the homotopy continuation method to eq. (5.4) with the polyhedral start system tracks the optimal number of paths.

Theorem 5.3.15. *For general* y, *the Newton polytopes of eq. (5.4) have mixed volume* $\mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r})$ *and homotopy continuation from a polyhedral start system can locate all solutions.*

The proof relies on the multiparameter eigenvalue problem structure on the equations. We start by looking at the structure of the Newton polytope of one of the equations.

Lemma 5.3.16. *Consider $\mathbb{R}^{r+k}$ with standard basis $e_1, \dots, e_r, b_1, \dots, b_k$, and let $\Delta \coloneqq \operatorname{Conv}(0, e_1, \dots, e_r)$. For positive integers $d_0, \dots, d_k$, consider the polytope*

$$P = \operatorname{Conv}(d_0 \cdot \Delta, b_1 + d_1 \cdot \Delta, \dots, b_k + d_k \cdot \Delta).$$

The integer points of P are all located in the facets:

$$P \cap \mathbb{Z}^{r+k} = (d_0\Delta \cap \mathbb{Z}^r) \cup \bigcup_{i=1}^{k} b_i + (d_i\Delta \cap \mathbb{Z}^r).$$

Proof. Using the convenient notation $b_0 \coloneqq 0 \in \mathbb{R}^{r+k}$, then every point in P is of the form

$$p = \sum_{i=0}^{k} \lambda_i(b_i + d_i \cdot \delta_i), \qquad (\lambda_0, \dots, \lambda_k) \in \Delta_k, \quad \delta_0, \dots, \delta_n \in \Delta.$$

If $p \in \mathbb{Z}^{r+k} \cap P$, then necessarily $\lambda_i = \langle p, b_i \rangle \in \mathbb{Z} \cap [0,1] = \{0,1\}$ for $i = 1, \dots, k$. Since $\sum_i \lambda_i = 1$, all but one λ_i are zero, which gives the desired statement. □

Proof of Theorem 5.3.15. Let P be the polytope from the previous lemma (with $d_0 = 1, d_1 = \cdots = d_k = 3$) and consider general polynomials with Newton polytope P $f_1, \dots, f_n \in \mathbb{C}[a_0, \dots, a_{r-1}, g_0, \dots, g_{k-1}]$. By the lemma, each monomial has degree ≤ 1 in the g_i, hence we can write the polynomial system $f(a, g) = 0$ as

$$N(a) \cdot \begin{pmatrix} 1 \\ g_0 \\ \vdots \\ g_{k-1} \end{pmatrix} = 0, \qquad N(a) \in \mathbb{C}[a_0, \dots, a_{r-1}]^{n \times k+1}.$$

The entries of $N(a)$ in column i have degree $\leq d_i$; let $M^{\mathrm{h}}(a_0, \dots, a_r)$ be the homogenization to degree d_i in each column. By genericity, the finite vanishing set

$$\left\{ N^{\mathrm{h}}(a) \cdot \begin{pmatrix} g_{-1} \\ g \end{pmatrix} = 0 \right\} \subseteq (\mathbb{C}^{r+1} \setminus 0) \times \mathbb{P}(\mathbb{C} \oplus \mathbb{C}^k)$$

is reduced, its only solutions are in the torus $(\mathbb{C}^\times)^{r+1} \times (\{1\} \times (\mathbb{C}^\times)^k)$, and the rank of $N^{\mathrm{h}}(a)$ drops exactly by 1. In particular, the number of solutions is simultaneously counted by Thom–Porteous and by the BKK theorem via the volume of P:

$$\mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}) = \deg s_r(\mathcal{O}_{\mathbb{P}^r}(-1) \oplus \mathcal{O}_{\mathbb{P}^r}(-3)^{\oplus k}) = \#\, \mathbb{V}(f_1, \dots, f_{d-r+1}) = n!\, \mathrm{vol}_n(P).$$

Now return to our specific polynommial system $N^{\mathrm{h}} = M_\Lambda(a, y)$, and let $Q_1, \dots, Q_n \subseteq P$ be the Newton polytopes. Then then desired equality follows

$$\mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}) \leq n!\, \mathrm{mvol}_n(Q_1, \dots, Q_n) \leq n!\, \mathrm{vol}_n(P) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}).$$ □

Remark 5.3.17. This generalizes without difficulty to any multimarameter eigenvalue problem $M(a) \cdot g = 0$, even with non-constant degrees along the columns (as long as M represents a vector bundle morphism over $\mathbb{P}^r_a$). The only assumption is that the solution set is isolated and located in the torus of $\mathbb{P}^r_a$. ◁

5.4 Specific ED degrees

Having gained a good understanding of $\mathcal{E}_{d,r}$ and $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$ for *general* $\Lambda \in \mathrm{Sym}(d+1)$, we now turn our attention to *specific* weights of interest. Towards this goal, we start by taking a closer look at the bad locus $\mathcal{B}_\Lambda \subseteq S_r$. In the cases where $\mathbb{P}\mathcal{X}_{d,r}$ is a curve ($r=1$) or a hypersurface ($d=2r$), we realize the bad locus as the singular locus of certain forms. We also derive an algorithm to compute the ED defect for many specific weights, and apply this to weights previously considered in the literature [OSS14].

5.4.1 Structure of the bad locus

We start by confirming that the absence of the bad locus indeed holds generically:

Theorem 5.4.1. *Let $\mathcal{S} \coloneqq \{\, \Lambda \in \mathbb{P}\,\mathrm{Sym}(d+1)$ invertible$\,\}$. The subset of $\mathcal{S}$ such that $\mathbb{P}\mathcal{B}_\Lambda = \emptyset$, is a dense open set.*

We need a small linear algebra lemma:

Lemma 5.4.2. *Let $v \in \mathbb{P}(\mathbb{C}^n)$, and $W \subseteq \mathbb{C}^n$ be a linear subspace of dimension r. Then the set $\{\, \Lambda \in \mathrm{Sym}(n) \mid \Lambda v \in W \,\}$ is a linear subspace of dimension $\binom{n}{2} + r$.*

Proof. After applying a congruence transformation $\Lambda \mapsto B\Lambda B^\mathsf{T}$ that maps the given vector v to e_1, we may assume $v = e_1$. The condition $\Lambda v \in W$ means that the first column of Λ is a vector $w \in W$. The other columns can be chosen arbitrarily as long as the resulting matrix is symmetric:

$$\Lambda = \left[\begin{array}{c|ccc} w_0 & w_1 & \dots & w_d \\ \hline w_1 & & & \\ \vdots & & \Lambda' & \\ w_d & & & \end{array}\right]$$

The dimension of the solution set is thus $\dim W + \dim \mathrm{Sym}(n-1) = \binom{n}{2} + r$. □

Proof of Theorem 5.4.1. The set $\nabla \coloneqq \{\, \Lambda \in \mathcal{S} \mid \mathbb{P}\mathcal{B}_\Lambda \neq \emptyset \,\}$ is the projection of the following incidence to $\mathcal{S}$:

$$\left\{\, (a, \Lambda, g) \in \mathbb{P}S_r \times \mathcal{S} \times \mathbb{P}S_{d-2r} \;\middle|\; B_\Lambda(a)g = T_d(a)\Lambda^{-1}T_d(a^2)g = 0 \,\right\}.$$

The fibers over the projection to $(a, g) \in \mathbb{P}S_r \times \mathbb{P}S_{d-2r}$ are matrices $\Lambda^{-1} \in \mathrm{Sym}(d+1)$ which map the vector $a^2 g$ into the linear subspace $\mathrm{Ker}\, T_d(a)$. By the previous Lemma 5.4.2, these fibers have dimension $\binom{d+1}{2} + r - 1$, so the incidence has dimension

$$r + (d-2r) + \binom{d+1}{2} + r - 1 = \binom{d+2}{2} - 2 = \dim \mathcal{S} - 1.$$

Thus, the image $\mathrm{Im}(\mathrm{pr}_\mathcal{S})$ is a closed subset of codimension at least one. This shows that its complement $\{\, \Lambda \mid \mathbb{P}\mathcal{B}_\Lambda = \emptyset \,\}$ is non-empty and open. □

If the bad locus is non-empty but finite, then its degree gives a bound on $\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r})$.

Theorem 5.4.3 (Bounds on $\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r})$)**.** *Assume that $\mathbb{P}\mathcal{B}_\Lambda$ is a non-empty finite scheme, then $\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r}) > 0$. More precisely, for general y we have*

$$\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r}) = \sum_{a \in \mathbb{P}\mathcal{B}_\Lambda} \mathrm{mult}_a\, I_{k+1}(M_\Lambda(a,y)) \geq \deg \mathbb{P}\mathcal{B}_\Lambda.$$

Moreover, equality holds (for general y) if and only if it holds for some particular y for which the critical locus is finite.

Proof. If $\mathbb{P}\mathcal{B}_\Lambda$ is finite, then we can still apply Thom–Porteous to the ideal of minors $I \coloneqq I_{k+1}(M_\Lambda(a,y))$, whose zero set in $\mathbb{P}S_r$ is finite for general y. We can partition the degeneracy locus into two sets, the solutions in $\mathcal{G}$ and the solutions supported in the bad locus $\mathbb{P}\mathcal{B}_\Lambda$:

$$\mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,r}) = \deg I_{k+1}(M_\Lambda(a,y)) = \mathrm{EDD}_\Lambda(\mathcal{X}_{d,r}) + \sum_{a \in \mathbb{P}\mathcal{B}_\Lambda} \mathrm{mult}_a\, I.$$

This shows the first formula. Furthermore, as $M_\Lambda(a,y) = [T_d(a) \mid B_\Lambda(a)]$, it follows that $I \subseteq I_k(B_\Lambda(a))$. This gives the claimed inequality, and the last claim is a standard semicontinuity argument. □

Example 5.4.4. The given inequality is actually an equality in many cases (see Table 5.1), where $\deg \mathbb{P}\mathcal{B}_\Lambda$ "explains" the ED defect. The first case when this *fails* is the following situation of *Bombieri weights*, discussed in Section 5.4.3 below:

$$(d,r) = (7,2) \qquad \Theta \coloneqq \mathrm{diag}(1,7,21,35,35,21,7,1), \qquad \deg \mathbb{P}\mathcal{B}_\Theta = 58.$$

The bad locus is supported in two points $[1 : \pm 2\mathrm{i} : -1] \mathrel{\hat{=}} (z \pm \mathrm{i})^2 \in \mathbb{P}S_2$, each having multiplicity 29. However, numerical and symbolic computation with random $y \in \mathbb{Q}^8$ strongly suggests that $\mathrm{EDdef}_\Theta(\mathcal{X}_{7,2}) = 60 > 58$ and $\mathrm{EDD}_\Theta(\mathcal{X}_{7,2}) = 43$. This indicates that the scheme structure of the component of $D_4(M_\Lambda(a,y)) \subseteq \mathbb{P}S_2$ supported in $p \in \mathbb{P}\mathcal{B}_\Theta$ has a multiplicity of 30 each, and this non-reduced structure varies with y. ◁

From the theorem we can deduce a procedure to compute the ED defect of specific weights, which we now describe:

For a prime number p and a polynomial matrix A defined over $\mathbb{Q}$, denote by $\overline{A}$ the reduction modulo p, assuming denominators of entries of A are not divisible by p. Recall from Lemma 5.3.8 the solution set to the substituted ED correspondence in the good locus

$$\mathcal{Y}_\mathcal{G}|_y = \{\, (a,g) \mid T(a)y = B_\Lambda(a)g \,\} \subseteq \mathcal{G} \times S_{d-2r},$$

where we specialize to a particular $y \in S_d^\vee$. The idea of the following approach is to check that the degeneracy locus of $M_\Lambda(a,y)$ is the disjoint union of the subschemes $\mathbb{P}\mathcal{B}_\Lambda$ and $\mathcal{Y}_\mathcal{G}|_y$ (the latter considered as a subset of $\mathbb{P}S_r$ via the projection pr_a). This can be done efficiently using a Hilbert polynomial computation:

Theorem 5.4.5. *Let Λ be a rational weight matrix, $y \in \mathbb{Q}^{d+1}$, p a prime and define*

$$\overline{X} \coloneqq D_k(\overline{M_\Lambda(a,y)}) \subseteq \mathbb{P}^r_{\overline{\mathbb{F}_p}}, \qquad Y \coloneqq \mathcal{Y}_\mathcal{G}|_y \cap \{a_r = 1\} \subseteq S_r \times S_{d-2r}.$$

If γ is an integer such that the following conditions are satisfied:

(i) $\gamma \leq \operatorname{length}(Y) < \infty$;

(ii) $\operatorname{HP}_{\overline{X}}(t) = \operatorname{HP}_{\mathbb{P}\mathcal{B}_\Lambda}(t) + \gamma$;

(iii) $\dim \mathbb{P}\mathcal{B}_\Lambda \leq 0$ *or* $D_{k-2}(B_\Lambda) = \emptyset \subseteq \mathbb{P}^r$;

then $\operatorname{EDD}_\Lambda(\mathcal{X}_{d,r}) = \operatorname{length}(Y) = \gamma$.

Proof. Let $X \coloneqq D_k(M_\Lambda(a,y)) \subseteq \mathbb{P}^r_\mathbb{C}$ be the degeneracy locus over $\mathbb{C}$ (or $\mathbb{Q}$ for computational purposes). Since $\mathbb{P}\mathcal{B}_\Lambda \cap \operatorname{pr}_a(Z) = \emptyset$ and both are contained in X, we have the inequalities

$$\operatorname{HP}_{\mathbb{P}\mathcal{B}_\Lambda}(j) + \gamma \overset{\text{(i)}}{\leq} \operatorname{HP}_{\mathbb{P}\mathcal{B}_\Lambda}(j) + \operatorname{HP}_{\operatorname{pr}_a(Z)}(j) \leq \operatorname{HP}_X(j) \leq \operatorname{HP}_{\overline{X}}(j), \qquad j \in \mathbb{Z}.$$

By assumption (ii), the last term equals first term, hence equality holds for all j. For $j \gg 0$ this implies $X = \operatorname{pr}_a(Y) \cup \mathbb{P}\mathcal{B}_\Lambda$. In particular, for *general* $y \in S_d^\vee$, the same decomposition holds. Assumption (iii) now guarantees that γ is the cardinality of the solution set for *general* y as well: In the first case X has expected dimension 0 and we can apply Thom–Porteous Theorem 5.4.3. In the second case the assumption allows us to apply the *excess Porteous formula* [Ful98, Example 14.4.7]. □

This theorem provides an efficient algorithm to compute $\operatorname{EDD}_\Lambda(\mathcal{X}_{d,r})$ for specific Λ. We briefly comment on our computational approach to each step:

(i) This is easy to verify: One can check $\operatorname{length}(Y) < \infty$ over a finite field first and then numerically solve the polynomial system defining Y to obtain a certified lower bound on $\operatorname{length}(Y)$.

(ii) The computation of the Hilbert polynomials is computationally the most expensive step. However, since we can compute $\operatorname{HP}_{\overline{X}}(t)$ over a finite field, there is no coefficient swelling despite random y. Condition (ii) can be impossible to satisfy for special weights, as Example 5.4.4 indicates.

(iii) This is a technical condition, and we are not aware of any instance where this fails (and (ii) would hold). The property $\dim \mathbb{P}\mathcal{B}_\Lambda \leq 0$ can be read off from $\operatorname{HP}_{\mathbb{P}\mathcal{B}_\Lambda}$ which is required to compute in (ii) anyways, while emptiness of the lower-rank degeneracy locus $D_{k-2}(B_\Lambda)$ can again be verified over a finite field.

We will apply this and the previous theorem to compute many specific ED degrees in Section 5.4.3.

5.4.2 The bad locus as a singular locus

In this section we describe how the bad locus can be identified with the singular locus of a hypersurface in the cases where $\mathcal{X}_{d,r}$ is a curve ($r = 1$) or a hypersurface ($d = 2r$).

We start with the rational normal curve $\mathbb{P}\mathcal{X}_{d,1} \subseteq \mathbb{P}^d$. For a symmetric matrix $\Lambda \in \mathrm{Sym}(d+1)$, denote by f_Λ the binary form

$$f_\Lambda = \nu_d(x_0, x_1)^\mathsf{T} \cdot \Lambda \cdot \nu_d(x_0, x_1) \in \mathbb{C}[x_0, x_1]_{2d}, \qquad \nu_d(x_0, x_1) = (x_0^d, x_0^{d-1}x_1, \ldots, x_1^d)^\mathsf{T}.$$

Denote by $\mathrm{Sing}(f) = V_+(\gcd(\frac{\partial}{\partial x_1} f, \frac{\partial}{\partial x_2} f)) \subseteq \mathbb{P}^1_{(x_0:x_1)}$ the singular locus of f.

Theorem 5.4.6. *Let Λ be a complex invertible weight matrix.*

(i) *We have a linear isomorphism identifying $\mathcal{B}_\Lambda$ with the singular locus of f_Λ:*

$$\psi\colon \mathbb{P}^1_a \to \mathbb{P}^1_x,\ (a_0 : a_1) \mapsto (-a_1 : a_0), \qquad \psi(\mathbb{P}\mathcal{B}_\Lambda) = \mathrm{Sing}(f_\Lambda).$$

(ii) *The following are equivalent:*

(a) $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,1}) = 0$, *so* $\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r}) = 3d - 2$;

(b) $\mathbb{P}\mathcal{B}_\Lambda = \mathbb{P}S_1$, *so* $\mathcal{G} = \emptyset$;

(c) $f_\Lambda = 0$;

(d) $y^\mathsf{T}\Lambda y \in \mathrm{I}(\mathcal{X}_{d,1}) \subseteq \mathbb{C}[y_0, \ldots, y_d]$, *that is,* $\mathcal{Q}_\Lambda \supseteq \mathbb{P}\mathcal{X}_{d,1}$.

(iii) *If the preceding equivalent conditions are* not *satisfied, then*

$$\mathrm{EDdef}_\Lambda(\mathcal{X}_{d,r}) = \deg \mathcal{B}_\Lambda = \deg \mathrm{Sing}(f_\Lambda) \leq 2d - 1.$$

This gives an algebraic proof of Theorem 5.1.10 in the special case of $X = \mathcal{X}_{d,1}$.

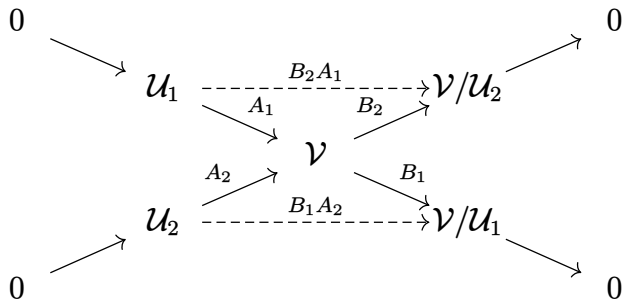
Example 5.4.7. In the case of the standard weights $\mathbf{1} \coloneqq \mathrm{Id}_{d+1}$, we have

$$f_\mathbf{1} = \sum_{i=0}^{d} x_0^{2d-2i} x_1^{2i} = \frac{x_1^{2d+2} - x_0^{2d+2}}{x_1^2 - x_0^2}.$$

This polynomial has $2d$ distinct roots, as it is a product of cyclotomic polynomials. By the theorem $\mathrm{EDdef}_\mathbf{1}(\mathcal{X}_{d,1}) = 0$ and $\mathrm{EDD}_\mathbf{1}(\mathcal{X}_{d,1}) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,1}) = 3d - 2$. ◁

We need the following lemma about degeneracy loci, to which we were pointed by Nathan Pflueger, compare section 3 in [Pfl25].

Lemma 5.4.8. *Consider vector bundles $\mathcal{U}_1, \mathcal{U}_2 \subseteq \mathcal{V}$ of $\mathrm{rank}\,\mathcal{U}_i = r_i$ on a scheme X. Assume that the $\mathcal{U}_i$ are locally direct summands of $\mathcal{V}$, so we have two short exact sequences of vector bundles*



Then for any $0 \leq k \leq \min\{r_1, r_2\}$ *the following degeneracy loci coincide:*

$$D_{r_1-k}(B_2 A_1) = D_{r_1+r_2-k}(A_1 \oplus A_2) = D_{r_2-k}(B_1 A_2) \subseteq X.$$

Proof. By symmetry, it suffices to show the first equality. Moreover, the statement is local on X, so we may assume that all bundles are trivial and that $\mathcal{U}_1$ is a direct summand of $\mathcal{V}$. Pick a basis of $\mathcal{U}_1$ and extend it to a basis of $\mathcal{V}$, then A_1 is the identity matrix and $A_1 \oplus A_2$ is represented by the matrix

$$A_1 \oplus A_2 \hat{=} \begin{array}{c} \\ \mathcal{U}_1 \\ \mathcal{V}/\mathcal{U}_1 \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{U}_1 & \mathcal{U}_2 \end{array} \\ \left[\begin{array}{c|c} \mathrm{id}_{r_1} & * \\ \hline 0 & B_1 A_2 \end{array}\right] \end{array}$$

From this block diagonal structure we see that the ideal of $(r_1 + r_2 - k + 1)$-minors of $A_1 \oplus A_2$ equals the ideal of $(r_2 - k + 1)$-minors of $B_1 A_2$. □

Proof of Theorem 5.4.6. (i) We show that ψ identifies $\mathcal{B}_\Lambda$ into $\mathrm{Sing}(f_\Lambda)$ by applying Lemma 5.4.8 to $D_{d-2}(B_\Lambda(a))$ on the standard affine charts of $\mathbb{P}^1$. By Proposition 5.2.3, the kernel of the Toeplitz matrices is parametrized by (confluent) Vandermonde vectors. On the chart $a_1 = -1$ we have (in the notation of the proof of 5.2.3)

$$\begin{aligned} \operatorname{Ker} T(a) &= \operatorname{Im} V^{(1)}(a_0) = \langle \{ v \coloneqq \mathrm{vand}(a_0) \} \rangle_{\mathbb{C}} \\ \operatorname{Ker} T(a^2) &= \operatorname{Im} V^{(2)}(a_0) = \langle \{ v, v' \coloneqq \mathrm{vand}^{(1)}(a_0) \} \rangle_{\mathbb{C}}. \end{aligned}$$

We apply Lemma 5.4.8 to the exact sequences on $\{a_1 = -1\} = \mathbb{A}^1$

$$0 \longrightarrow \mathcal{O}^{d-1} \xrightarrow{T(a^2)^\mathsf{T}} \mathcal{O}^{d+1} \xrightarrow{(V^{(2)})^\mathsf{T}} \mathcal{O}^2 \longrightarrow 0$$

$$0 \longrightarrow \mathcal{O}^{1} \xrightarrow{\Lambda V^{(1)}} \mathcal{O}^{d+1} \xrightarrow{T(a)\Lambda^{-1}} \mathcal{O}^{d} \longrightarrow 0$$

and the integer $k = 0$ to obtain the equality of ideals of minors

$$I_{d-1}(T(a)\Lambda^{-1} \cdot T(a^2)^\mathsf{T}) = I_1((V^{(2)})^\mathsf{T} \cdot \Lambda V^{(1)}) = \langle v^\mathsf{T} \Lambda v, (v')^\mathsf{T} \Lambda v \rangle_{\mathbb{C}[a_0]}.$$

Under the identification $\psi \colon (x_0 : x_1) \mapsto (-a_1 : a_0)$, these two polynomials are exactly $f_\Lambda(1, x_1)$ and its first derivative. A completely analogous calculation yields the result on the chart $a_0 \neq 0$. This establishes that $\psi(\mathcal{B}_\Lambda) = \mathrm{Sing}(f_\Lambda)$ as schemes.

(ii) By (i), $\mathbb{P}\mathcal{B}_\Lambda = \mathbb{P}S_1$ is equivalent to $\mathrm{Sing}(f_\Lambda) = \mathbb{P}^1$, which only happens for $f_\Lambda = 0$. From the definition of $f_\Lambda = \nu_d(x)^\mathsf{T} \cdot \Lambda \cdot \nu_d(x)$, we see that this happens if and only if

$$y^\mathsf{T} \Lambda y \in \operatorname{Ker}(\nu_d^* \colon \mathbb{C}[\underline{y}] \to \mathbb{C}[x_0, x_1]) = \mathrm{I}(\mathcal{X}_{d,1}).$$

For the remaining equivalence with (a), note that by Lemma 5.3.8, $\mathcal{G} = \emptyset$ implies $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,1}) = 0$. The converse follows from (iii), namely if $\mathbb{P}\mathcal{B}_\Lambda$ is finite, then

$$\mathrm{EDD}_\Lambda(\mathcal{X}_{d,1}) = \mathrm{EDD}_{\mathrm{gen}}(\mathcal{X}_{d,1}) - \mathrm{EDdef}(\mathcal{X}_{d,1}) \geq (3d-2) - (2d-1) > 0.$$

(iii) The ideal of minors of $M(a,y)$ is in the case $r=1$ a single determinant, which factors in $R=\mathbb{C}[a_0,a_1;y_0,\dots,y_d]$ as

$$\det M(a,y) = g(a,y)\cdot b(a), \qquad g(a,y)\in R_{e,1} \text{ irred.},\ b\in R_{3d-2-e,0}.$$

The polynomial $b(a)$ defines the base locus of $D_{d-1}(M(a,y))\subset\mathbb{P}^1$ as $y\in\mathbb{P}^d$ varies, which is exactly $\mathbb{P}\mathcal{B}_\Lambda$. For general y, the e roots of the polynomial $g(a,y)$ correspond to the $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,1})$-many critical points. This establishes

$$\mathrm{EDdef}(\mathcal{X}_{d,1}) = \deg b(a) = \deg \mathbb{P}\mathcal{B}_\Lambda \overset{\text{(i)}}{=} \deg \mathrm{Sing}(f_\Lambda).$$

For the inequality $\mathrm{EDdef}(\mathcal{X}_{d,1})\le 2d-1$, note that the singular locus of a binary form f of degree $2d$ with root multiplicities $e_1,\dots,e_m$ has degree

$$\deg \mathrm{Sing}\, f = \sum_{i=1}^m e_i - 1 = 2d-m.$$

This is maximized for $m=1$, in which case $\deg \mathrm{Sing}\, f = 2d-1$. □

In the case $d=2r$, where $\mathcal{X}_{2r,r}$ is the Catalecticant hypersurface, we also have a representation of $\mathcal{B}_\Lambda$ as a singular locus. The dual variety is the variety of squares

$$(\mathbb{P}\mathcal{X}_{2r,r})^\vee = \mathbb{P}\{\,\mathrm{sq}(a)=a^2 \mid a\in S_r\,\}\subseteq \mathbb{P}S_{2r},$$

see for example [LS16, Proposition 3.1].

Lemma 5.4.9. *Let $d=2r$. The bad locus equals the singular locus of the quartic $q=Q_{\Lambda^{-1}}\circ\mathrm{sq}$ which is the composition of the squaring map with the dual quadric:*

$$\mathbb{P}\mathcal{B}_\Lambda = \mathrm{Sing}(q)\subseteq\mathbb{P}S_r, \qquad q(a)=Q_{\Lambda^{-1}}(a^2).$$

Proof. Here the matrix $B_\Lambda(a)=T(a)^\mathsf{T}\Lambda^{-1}T(a^2)^\mathsf{T}$ is a single column, so $\mathbb{P}\mathcal{B}_\Lambda$ is defined by the vanishing of its entries. Note that the squaring map is identical to the map $a\mapsto T(a^2)^\mathsf{T}=T_{2r}(a)^\mathsf{T}T_r(a)^\mathsf{T}$ with Jacobi matrix $2T(a)^\mathsf{T}$. We compute the gradient of $q(a)$ to conclude:

$$q(a)=Q_{\Lambda^{-1}}(T(a^2)), \qquad \nabla q(a) = T(a^2)\Lambda^{-1}\cdot 2T(a)^\mathsf{T} = 2B_\Lambda(a)^\mathsf{T}.$$ □

5.4.3 Specific weights

With Theorems 5.4.3 and 5.4.5 we are able to compute many specific ED degrees of $\mathcal{X}_{d,r}$. As a demonstration of its usefulness, we revisit Table 4 from [OSS14] in Table 5.1. There the following weights matrices are considered (adapted to our notation)

$$\begin{aligned}
\text{Unit:}\quad & \mathbf{1} = \mathrm{diag}(1,\dots,1)\\
\text{Frobenius:}\quad & F = \mathrm{diag}((\min\{i+1,d-i+1\})_{i=0,\dots,d})\\
\text{Bombieri:}\quad & \Theta = \mathrm{diag}(\tbinom{d}{0},\tbinom{d}{1},\dots,\tbinom{d}{d})
\end{aligned}$$

	$\Lambda = \mathbf{1}$ Unit				$\Lambda = F$ Frobenius				$\Lambda = \Theta$ Bombieri			
$d \setminus r$	1	2	3	4	1	2	3	4	1	2	3	4
1	1				1				1			
2	4				2				2			
3	7	1			7	1			3	1		
4	10	13			6	9			4	7		
5	13	34	1		13	34	1		5	16	1	
6	16	64	40		10	38	34		6	28	20	
7	19	103	142	1	19	103	142	1	7	43..45	62..64	1
8	22	151	334	121	14	103	246	113	8	61..65	134..142	53
9	25	208	643	547	25	208	643	543	9	82..88	243..263	229

Table 5.1: The computed Euclidean distance degrees for $\Lambda \in \{\mathbf{1}, F, \Theta\}$. Green indicates that the defect is completely explained by Theorem 5.4.3, dark green being $\mathrm{EDD}_{\mathrm{gen}}$. Yellow indicates a range between the upper bound from 5.4.3 and the (very likely correct) lower bound from a computation with random y. Orange indicates that the bad locus is positive-dimensional.

Frobenius weights correspond to the pullback of the standard norm on $\mathbb{R}^{\lceil d/2 \rceil \times \lfloor d/2 \rfloor}$ via $y \mapsto H_{\lfloor d/2 \rfloor}(y)$. Bombieri weights correspond to the pullback of the standard norm in $\mathbb{R}^{2^d} \cong (\mathbb{R}^2)^{\otimes d}$ under the inclusion $y \in \mathrm{Sym}^d \mathbb{R}^2 \subseteq (\mathbb{R}^2)^{\otimes d}$; equivalently under the identification of $\mathrm{Sym}^d \mathbb{R}^2 \cong \mathrm{S}^d \mathbb{R}^2$ (Lemma 1.2.4).

Theorem 5.4.10. *Table 5.1 displays the correct values for* $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$, $\Lambda \in \{\mathbf{1}, F\}$.

Proof. In all these cases Theorem 5.4.3 applies and allows us to compute the exact value of $\mathrm{EDD}_\Lambda(\mathcal{X}_{d,r})$. This was checked using an implementation in `Macaulay2`, available at

`https://github.com/leokayser/hankel-edd`

The first table took 3 seconds to verify, the second around 65 seconds. □

These values indicate the following (conjectural) behavior:

- Unit weights have empty bad locus and hence exhibit generic EDdegree.
- Frobenius weights are defective for even d (but also for some odd d, see e. g. $\mathcal{X}_{9,4}$), and their defect can be explained by the bad locus.

The Bombieri weights are very special in that they exhibit very large EDD defect (conjecturally the largest [Koz+24]). It turns out that the degree of the bad locus can only explain part of this defect: For $d \geq 7$ the condition in Theorem 5.4.3 for equality is (very likely) not met and for $r \geq 4$ the bad locus is not finite. The results for $\Lambda = \Theta$ in Table 5.1 took 40 seconds to calculate the first three columns, and 275 seconds to do the two saturations for column 4 (over $\mathbb{F}_{32003}$).

r	$\mathrm{EDD}_\Theta(\mathcal{X}_{k+2r-1,r})$	$\mathrm{EDdef}_\Theta(\mathcal{X}_{k+2r-1,r})$
1	$1+k$	$2k$
2	$1+3k+3\binom{k+1}{2}$	$6\binom{k+1}{2}$
3	$1+3k+9\binom{k+1}{2}+7\binom{k+2}{3}$	$20\binom{k+2}{3}$
4	$1+3k+9\binom{k+1}{2}+25\binom{k+2}{3}+19\binom{k+3}{4}$	$2\binom{k+2}{3}+62\binom{k+3}{4}$
5	$1+3k+9\binom{k+1}{2}+27\binom{k+2}{3}+71\binom{k+3}{4}+51\binom{k+4}{5}$	$10\binom{k+3}{4}+192\binom{k+4}{5}$
6	$1+3k+9\binom{k+1}{2}+27\binom{k+2}{3}+81\binom{k+3}{4}+201\binom{k+4}{5}+141\binom{k+5}{6}$	$42\binom{k+4}{5}+588\binom{k+5}{6}$

$k \setminus r$	1	2	3	4	5	6
1	2	7	20	*53*	162	463
2	3	16	62	229	803	2 740
3	4	28	134	599	2 470	9 718
4	5	43	243	1 268	5 984	26 578
5	6	61	396	2 356	12 492	61 702
6	7	82	600	4 002	23 518	127 720

Table 5.2: Conjectural polynomials following $k \mapsto \mathrm{EDD}_\Theta(\mathcal{X}_{k+2r-1,r})$, and their values (with *exception* $\mathrm{EDD}_\Theta(\mathcal{X}_{8,4}) = 53$). Green cases covered by Theorem 5.4.12.

Extensive numerical computation leads us to make the following conjecture about the ED degrees for the Bombieri weights, refining the conjectures stated in [Pan25]. Let $C := \mathrm{Im}(\mathbb{P}S_1 \to \mathbb{P}S_r,\ \ell \mapsto \ell^r)$ be the rescaled rational normal curve (again under the isomorphism from Lemma 1.2.4).

Conjecture 5.4.11. The bad locus for $\Lambda = \Theta$ equals the $(\lceil r/3 \rceil - 1)$-th osculating space of C at $(z \pm \mathrm{i})^r$, that is,

$$(\mathbb{P}\mathcal{B}_\Theta)^{\mathrm{red}} = \mathbb{P}((z+\mathrm{i})^{r+1-\lceil r/3 \rceil} S_{\lceil r/3 \rceil - 1}) \cup \mathbb{P}((z-\mathrm{i})^{r+1-\lceil r/3 \rceil} S_{\lceil r/3 \rceil - 1}).$$

The lower osculating spaces (including these points themselves) are nested embedded points. For fixed r, the sequence $k \mapsto \mathrm{EDD}_\Theta(\mathcal{X}_{k+2r-1,r})$ is an integer polynomial of degree r, the first six (conjectural) polynomials are displayed in Table 5.2.

There is a known exception $\mathrm{EDD}_\Theta(\sigma_4\nu_8\mathbb{P}^1) = 53 < 57$, where $\mathbb{P}\mathcal{B}_\Theta \subseteq \mathbb{P}^4$ additionally contains a conic contained in the plane $\langle (z+\mathrm{i})^4, (z+\mathrm{i})^2(z-\mathrm{i})^2, (z-\mathrm{i})^4 \rangle_{\mathbb{P}}$. ⍰

This conjecture is supported not only by computational experiments, but also by the following results:

Theorem 5.4.12. (i) *The conjecture is true for $r = 1$, in particular $\mathrm{EDD}_\Theta(\mathcal{X}_{d,1}) = d$.*

(ii) *The following values of* $\mathrm{EDD}_\Theta(\mathcal{X}_{d,r})$ *are correct:*

(d,r)	$\mathrm{EDD}_\Theta(\mathcal{X}_{d,r})$	$\mathrm{EDdef}_\Theta(\mathcal{X}_{d,r})$	$(\mathbb{P}\mathcal{B}_\Theta)^{\mathrm{red}} \subseteq \mathbb{P}S_r$	$\mathrm{HP}_{\mathbb{P}\mathcal{B}_\Theta}(t)$
$(d,1)$	d	$2d-2$	$(z \pm \mathrm{i})^d$	$2\cdot(d-1)$
$(4,2)$	7	6	$(z \pm \mathrm{i})^2$	$2\cdot 3$
$(5,2)$	16	18	$(z \pm \mathrm{i})^2$	$2\cdot 9$
$(6,2)$	28	36	$(z \pm \mathrm{i})^2$	$2\cdot 18$
$(6,3)$	20	20	$(z \pm \mathrm{i})^3$	$2\cdot 10$
$(8,4)$	53	68	$S_1(z \pm \mathrm{i})^3 \cup Q$	$4(t+1)+31$
$(10,5)$	162	202	$S_1(z \pm \mathrm{i})^4$	$2\cdot(5(t+1)+44)$

Here $Q \subseteq \mathbb{P}S_4$ *is the exceptional plane conic mentioned in Conjecture 5.4.11.*

(iii) *If* $d = 2r$*, then the bad locus* $\mathbb{P}\mathcal{B}_\Theta$ *contains the two osculating spaces.*

Proof. (i) In the notation of Theorem 5.4.6, the binary form f_Θ is

$$f_\Theta = \sum_{k=0}^{d} \binom{d}{k} x_0^{2k} x_1^{2d-2k} = (x_0^2 + x_1^2)^d = (x_0 + \mathrm{i}x_1)^d (x_0 - \mathrm{i}x_1)^d.$$

Its singular locus thus consists of two points of multiplicity $(d-1)$, corresponding to $(1 : \pm\mathrm{i}) \hat{=} z \pm \mathrm{i}$ under the isomorphism ψ. We deduce $\mathrm{EDD}_\Theta(\mathcal{X}_{d,1}) = (3d-2) - 2(d-1) = d$.

(ii) The first case is a restatement of (i). The other cases are proven using a direct calculation with Theorem 5.4.3 as before. Note that in the cases $r \geq 4$ special care has to be applied, as $\dim \mathbb{P}\mathcal{B}_\Theta > 0$. There, we can successfully apply Theorem 5.4.5 instead.

(iii) This was shown in [Pan25, Theorem 1.2(ii)] for the singular locus of the quartic q from Lemma 5.4.9. By said lemma, this singular locus coincides with $\mathbb{P}\mathcal{B}_\Theta$. □

5.5 Infinite Hankel matrix approximation

In this final section we turn to the infinite-dimensional analogue of low-rank Hankel matrix approximation, that is, Problem 5.2.19. We will always denote by R and r the (finite!) ranks of the data sequence $y \in \ell^2$ and its lower-rank approximant $\hat{y}$, respectively. In other words, we are dealing with rational functions

$$\hat{y} \hat{=} \widehat{Y} = \frac{zb(z)}{a(z)} \qquad a \in \mathbb{C}[z]_{\leq r} \setminus 0,\ b \in \mathbb{C}[z]_{<r}$$
$$y \hat{=} Y = \frac{zd(z)}{c(z)} \qquad c \in \mathbb{C}[z]_{\leq R} \setminus 0,\ d \in \mathbb{C}[z]_{<R}.$$

We start by deriving the first order necessary conditions for $\hat{y}$ to be a critical point to $\|y - \hat{y}\|_{\ell^2}$ for given $y \in \ell^2$. We then show that, after restricting to approximants with simple poles, the incidence of critical pairs $(a, b; c, d)$ is an irreducible variety of the expected dimension. Similar to the finite-dimensional case, the critical equations exhibit the structure of an affine MEP. We then perform a similar analysis using Thom–Porteous to obtain the number of complex critical points for general data (c, d) depending on R, r.

5.5.1 Deriving the critical equations

In contrast to the finite-dimensional case, our target function

$$\hat{y} \mapsto \|\hat{y} - y\|_{\ell^2}^2 = \sum_{k=1}^{\infty} |\hat{y}_k - y_k|^2$$

does not seem to depend algebraically on the approximant. Our first goal is therefore to characterize critial points (a, b) in a way that depends evidently polynomially on the coefficients of (a, b).

A first step is to express the ℓ^2-norm of y in terms of the $\mathcal{Z}$-transform $Y(z) = \sum_{k=0}^{\infty} y_k z^{-k}$.

Lemma 5.5.1 (Parseval's identity)**.** *For any sequence $y \in \ell^2(\mathbb{C})$ with $\mathcal{Z}$-transform $Y(z)$ we have*

$$\|y\|_{\ell^2}^2 := \sum_{k=0}^{\infty} |y_k|^2 = \frac{1}{2\pi \mathrm{i}} \oint |Y(z)|^2 z^{-1} \, \mathrm{d}z = \int_0^1 |Y(e^{2\pi \mathrm{i} t})|^2 \, \mathrm{d}t.$$

Here $\oint F(z) \, \mathrm{d}z = \int_0^{2\pi} F(e^{\mathrm{i}t}) \mathrm{i} e^{\mathrm{i}t} \, \mathrm{d}t$ is the contour integral along the unit circle.

Proof. This is a special case of Parseval's identity, which expresses that the $\mathcal{Z}$-transform is *unitary*:

$$\langle x, y \rangle_{\ell^2} = \sum_{k=0}^{\infty} x_k \overline{y_k} = \frac{1}{2\pi \mathrm{i}} \oint \mathcal{Z}\{x\}(z) \, \overline{\mathcal{Z}\{y\}(z)} \, z^{-1} \, \mathrm{d}z.$$

By bilinearity and density of finite sequences in $\ell^2(\mathbb{C})$, it suffices to prove the statement for $x = (\delta_{i,k})_k, y = (\delta_{j,k})_k$. Then, using that $\overline{z} = z^{-1}$ on the unit circle,

$$\frac{1}{2\pi \mathrm{i}} \oint z^{-i} \overline{z^{-j}} z^{-1} \, \mathrm{d}z = \frac{1}{2\pi \mathrm{i}} \oint z^{j-i-1} \, \mathrm{d}z = \delta_{i,j} = \langle x, y \rangle_{\ell^2}. \qquad \square$$

Thus, our optimization problem is equivalent to the following:

$$\underset{(a,b) \in \mathbb{R}[z]_{\le r} \times \mathbb{R}[z]_{< r}}{\text{minimize}} \; J(a, b) := \frac{1}{2\pi \mathrm{i}} \oint \left| \frac{z b(z)}{a(z)} - \frac{z d(z)}{c(z)} \right|^2 z^{-1} \, \mathrm{d}z,$$

where a, c are stable to ensure that the integral converges. Using this expression of the norm, we can derive an elegant description of the critical points. First, we cite a result from [Reg18] analogous to Theorem 5.3.3 in the finite-dimensional case:

Proposition 5.5.2. *If $\hat{y}$ is a local minimizer of $J(\hat{y})$, then $\hat{y}$ has rank exactly r.*

Proof. This is [Reg18, Theorem 5.3]; the idea is similar to our proof of Theorem 5.3.3. $\square$

This proposition allows us to restrict our attention to the case where $a(z)$ is of degree r and coprime to $b(z)$, so $\hat{y}$ has rank exactly r. The following result is originally due to Walsh [Wal32, Section 5], we adapt the proof strategy of Meier & Luenberger [ML67] who consider the continuous-time case.

Theorem 5.5.3. *Let $\hat{y}$ be of rank exactly r, so $\widehat{Y}(z) = zb(z)/a(z)$ with $\deg a = r$ and $\gcd(a,b) = 1$. Then $\hat{y}$ is a critical point to $J(a,b) = \|\hat{y} - y\|_{\ell^2}^2$ if and only if*

$$Y^{(j)}(1/\omega) = \widehat{Y}^{(j)}(1/\omega), \qquad \text{for all poles } \omega \text{ of } \widehat{Y} \text{ and } j = 0, \dots, \operatorname{ord}_\omega a(z)$$

where $Y^{(j)}, \widehat{Y}^{(j)}$ denotes the j-th derivative.

To make sense of the expression, note that $Y, \widehat{Y}$ are rational functions with poles in the unit circle, so $Y(1/z), \widehat{Y}(1/z)$ are rational functions with poles outside the unit circle. Thus it makes sense to evaluate $Y, \widehat{Y}$ and their derivatives at the inverses of the poles of $\widehat{Y}$, including at ∞ if $\omega = 0$.

In the theorem "critical point" is understood in the following sense: After normalizing the leading coefficient $a_r = 1$, the approximant is parametrized by $\mathbb{R}^{2r}$, and $\hat{y}$ is a critical point of the smooth map $\mathbb{R}^{2r} \to \mathbb{R}$ given by the squared norm J.

Proof. Let $v \in \mathbb{R}^{2r}$ be a direction, then, using $\overline{z} = z^{-1}$ on the unit circle,

$$\begin{aligned}\frac{\partial J(\widehat{Y})}{\partial v} &= \frac{1}{2\pi \mathrm{i}} \oint \overline{(\widehat{Y} - Y)(z)} \cdot \frac{\partial \widehat{Y}(z)}{\partial v} z^{-1} \,\mathrm{d}z + \frac{1}{2\pi \mathrm{i}} \oint (\widehat{Y} - Y)(w) \cdot \overline{\frac{\partial \widehat{Y}(w)}{\partial v}} w^{-1} \,\mathrm{d}w \\ &\overset{w=z^{-1}}{=} \frac{2}{2\pi \mathrm{i}} \oint (\widehat{Y} - Y)(z^{-1}) \cdot \frac{\partial \widehat{Y}(z)}{\partial v} z^{-1} \,\mathrm{d}z.\end{aligned}$$

For a critical point these derivatives must be zero, in this way we obtain the conditions

$$\oint \widehat{Y}(z^{-1}) \cdot \frac{\partial \widehat{Y}(z)}{\partial v} z^{-1} \,\mathrm{d}z = \oint Y(z^{-1}) \cdot \frac{\partial \widehat{Y}(z)}{\partial v} z^{-1} \,\mathrm{d}z,$$

where v ranges over $2r$ independent directions. Let $\omega_1, \dots, \omega_m \in \mathbb{D}$ be all poles of $\widehat{Y}(z)/z$, $e_i \coloneqq \operatorname{ord}_{\omega_i} a(z)$, and consider the partial fraction decomposition

$$\widehat{Y}(z) = \frac{zb(z)}{a(z)} = \sum_{i=1}^{m} \sum_{j=1}^{e_i} \frac{r_{ij} z}{(z - \omega_i)^j} \qquad r_{ij} \in \mathbb{C}.$$

Then a basis of directions v in a neighborhood of $\widehat{Y}$ is given by perturbing each r_i or each ω_i (for complex poles, perturbations come in conjugate pairs to preserve reality). For notational simplicity, we only treat the case of real poles. The complex case can be treated analogously using *Wirtinger calculus* instead, for details see [LAD23; ML67; Wal32]. In the real case, we have, for a particular pole ω_i of order e_i, the partial derivatives

$$\frac{\partial \widehat{Y}}{\partial r_{ij}} = \frac{z}{(z - \omega_i)^j}, \qquad \frac{\partial \widehat{Y}}{\partial \omega_i} = \sum_{j=1}^{e_i} \frac{j r_{ij} z}{(z - \omega_i)^{j+1}}, \qquad \frac{\partial \widehat{Y}}{\partial \omega_i} - \sum_{j=1}^{e_i - 1} j r_{i,j} \frac{\partial \widehat{Y}}{\partial r_{i,j+1}} = \frac{e_i r_{i,e_i} z}{(z - \omega_i)^{e_i + 1}}.$$

By the assumption that $\hat{y}$ has order exactly r, the highest order coefficient r_{i,e_i} is non-zero (analogously to Corollary 5.2.5). Putting things together, a suitable set of directions, and hence conditions, is given by

$$\oint \frac{\widehat{Y}(z^{-1})}{(z - \omega_i)^j} \,\mathrm{d}z = \oint \frac{Y(z^{-1})}{(z - \omega_i)^j} \,\mathrm{d}z, \qquad i = 1, \dots, m, \ j = 1, \dots, e_i + 1.$$

Let F be a meromorphic function with poles only in the open unit disc, then by Cauchy's integral formula

$$F^{(j)}(1/\omega_i) = \frac{j!}{2\pi \mathrm{i}} \oint \frac{F(z^{-1})}{(z-\omega_i)^{j+1}} \,\mathrm{d}z.$$

Applying this to $F = \widehat{Y}, Y$ in the previous equation yields the theorem. □

From now on until the end of the chapter, we impose the restriction that **the approximant has simple poles only**. In this situation the optimality conditions of Theorem 5.5.3 lead to the following elegant description:

Theorem 5.5.4 (Walsh's theorem)**.** *Let $\widehat{Y}(z) = b(z)/a(z)$ have simple poles, then $\hat{y}$ is a critical point if and only if there exists a polynomial $g \in \mathbb{C}[z]_{\leq R-r-1}$ such that*

$$a \cdot d - b \cdot c = \tilde{a}^2 \cdot g, \qquad \tilde{a}(z) \coloneqq z^r a(1/z) \in \mathbb{C}[z]_{\leq r}.$$

Proof. By Theorem 5.5.3, the rational function $F(z) = zd(z)/c(z) - zb(z)/a(z)$ has zeros of order 2 at the inverse $1/\omega$ of each root ω of $a(z)$. In other words, $F = \tilde{a}^2 \cdot G$, where $G(z)$ is a rational function with no poles outside the unit disc. Multiply F by $a \cdot c$ to obtain

$$\underbrace{a \cdot c \cdot G}_{=: g'} \cdot \tilde{a}^2 = a \cdot c \cdot F = z(a \cdot d - b \cdot c).$$

The right hand side is a polynomial, hence $\tilde{a}^2 \cdot g'$ is a polynomial. g' is a rational function whose only (potential) poles are in the unit disc (among the poles of F). However, $\tilde{a}^2$ can only cancel poles outside the unit disc, hence g' must be a polynomial itself. From the equation we also see that $z \mid g'$, so we can divide $g'/z =: g$ to obtain $g \cdot \tilde{a}^2 = ad - bc$. Thus, if $zb(z)/a(z)$ is a critical point, then there exists a polynomial g with the desired properties, and this argument is reversible. □

5.5.2 Irreducibility of the Walsh variety

In this section we study the system of equations on (a, b, c, d, g) given by Walsh's Theorem 5.5.4. In the following analysis it will be advantageous to homogenize all polynomials to binary forms in $S \coloneqq \mathbb{C}[z_0, z_1]$. For example, the transformation $a \mapsto \tilde{a}$ is simply given by $a(z_0, z_1) \mapsto a(z_1, z_0)$, also the statements of Lemmas 5.5.7 and 5.5.10 are more cumbersome in the inhomogeneous notation. To clean up notation, we give each vector space of polynomials its own name:

$$a \in A \coloneqq S_r, \quad b \in B \coloneqq S_{r-1}, \quad c \in C \coloneqq S_R, \quad d \in D \coloneqq S_{R-1}, \quad g \in G \coloneqq S_{R-r-1}.$$

Additionally, denote $A^\circ \coloneqq A \setminus 0$, and we use, for example, $A^\circ BG$ to denote the product $(A \setminus 0) \times B \times G$.

Definition 5.5.5. The *Walsh variety* is the quasi-affine scheme given by

$$\mathcal{W} \coloneqq \left\{ (a, b, c, d, g) \;\middle|\; ad - bc = \tilde{a}^2 g \in S_{R+r-1} \right\} \subseteq A^\circ BCDG. \tag{5.5}$$

The defining equations 5.5 are of binary forms in S_{R+r-1}, hence they consist of $R+r$ equations in an ambient space of dimension $3R+r+2$.

Theorem 5.5.6. *The Walsh variety $\mathcal{W}$ is a reduced and irreducible ideal-theoretic complete intersection in $A^\circ BCDG$ of dimension $2R+2$.*

Before proving Theorem 5.5.6, we note a useful structure in the equations defining $\mathcal{W}$. Consider the projection

$$\pi\colon \mathcal{W} \to ABG,\ (a,b,c,d,g) \to (a,b,g).$$

For fixed (a,b,g), the right-hand side of the defining equation of $\mathcal{W}$ is constant, so the unknowns (c,d) are constrained by an affine system of linear equations:

$$\begin{bmatrix} T_{R+r}(a)^\mathsf{T} & -T_{R+r}(b)^\mathsf{T} \end{bmatrix} \cdot \begin{pmatrix} d \\ c \end{pmatrix} = \tilde{a}^2 g.$$

Lemma 5.5.7. *For given $(a,b,g) \in ABG$, $a \neq 0$, the affine-linear subspace $\pi^{-1}(a,b,g)$ is non-empty if and only if $\gcd(a,b) \mid \tilde{a}^2 g$. In that case we have*

$$\dim \pi^{-1}(a,b,g) = R - r + 1 + \deg \gcd(a,b).$$

Proof. Let $k \coloneqq \deg \gcd(a,b)$, $a' \coloneqq a/\gcd(a,b)$ and $b' \coloneqq b/\gcd(a,b)$. It is immediate from

$$\gcd(a,b) \cdot (a'd - b'c) = \tilde{a}^2 g$$

that $\gcd(a,b) \mid \tilde{a}^2 g$ is *necessary* for the fiber to be non-empty.

On the other hand, since a' and b' are coprime of degree $r-k$ and $r-k-1$, they generate all polynomials in degree $\geq 2r-2k-2$ by Bézout's Lemma 0.1.14. This means that, since $R+r-k-1 \geq 2r-2k-2$, we have

$$\{\, a'd - b'c \mid (c,d) \in CD \,\} = S_{R+r-k-1} \ni \tilde{a}^2 g/\gcd(a,b)$$

This shows that the divisibility condition is also *sufficient*, and the dimension of the solution space is equal to the dimension of the vector space of syzygies of a', b' in degree $R+r-k-1$. This dimension can be read of from Bézout's Lemma 0.1.14 as

$$\dim \pi^{-1}(a,b,g) = \dim_{\mathbb{C}} \{\, (c,d) \mid a'd - b'c = 0 \,\} = R - r + 1 + k. \qquad \square$$

Proof of Theorem 5.5.6. The proof will proceed in the following steps: First we stratify $\mathcal{W}$ by the degree of the greatest common divisor of a and b, this will allow us to determine the dimension of $\mathcal{W}$. Then we show that $\mathcal{W}$ is a set-theoretic complete intersection and irreducible. Lastly, reducedness will follow from a bundle argument, so $\mathcal{W}$ is an ideal-theoretic complete intersection.

For $0 \leq k \leq r-1$, consider the two locally-closed subsets of $A^\circ BG$

$$\begin{aligned} \mathcal{S}_{\geq k} &\coloneqq \{\, (a,b,g) \mid \deg \gcd(a,b) \geq k \,\}, \qquad \mathcal{S}_k \coloneqq \mathcal{S}_{\geq k} \setminus \mathcal{S}_{\geq k+1}, \\ \mathcal{T}_k &\coloneqq \left\{\, (a,b,g) \in \mathcal{S}_k \;\middle|\; \gcd(a,b) \text{ divides } \tilde{a}^2 g \,\right\}. \end{aligned}$$

The defining equations of $\mathcal{S}_{\geq k}$ are given by the vanishing of $(k+1)$-minors of the Sylvester matrix $\mathrm{Syl}(a;b)$. Within $\mathcal{S}_k$, the divisibility constraint of $\mathcal{T}_k$ is given by $(k+1)$-minors of $\mathrm{Syl}(a;b)$ concatenated with the coefficient vector of $\widetilde{a}^2 g$.

The variety $\mathcal{S}_{\geq k}$ is irreducible as the image of the polynomial map

$$(S_k \setminus 0) \times (S_{r-k} \setminus 0) \times S_{r-k-1} \times G \to A^\circ BG, \qquad (f,u,v,g) \mapsto (fu, fv, g).$$

Given $(a,b,g) \in \mathcal{S}_{\geq k}$, there are only finitely many factorizations $a = fu$, $b = fv$ up to rescaling f by $\lambda \in \mathbb{C}^\times$, hence the fibers of this map are 1-dimensional. From this we conclude that $\mathcal{S}_{\geq k}$ is an irreducible variety of dimension $R + r + 1 - k$.

If $k = 0$, then the divisibility constraint is trivial and $\mathcal{T}_0 = \mathcal{S}_0$. For $k \geq 1$, $\mathcal{T}_k \subseteq \mathcal{S}_k$ is a strict subvariety, since there are tuples $(a,b,g) \in \mathcal{S}_k$ that do *not* satisfy the divisibility constraint, for example

$$a = (z-1)^r, \quad b = (z-1)^k (z-2)^{r-1-k}, \quad g = (z-3)^{R-r-1}.$$

Therefore $\dim \mathcal{T}_0 = R + r + 1$ and $\dim \mathcal{T}_k < R + r + 1 - k$ for $k \geq 1$.

We return to the variety $\mathcal{W}$. Stratifying by $\gcd(a,b)$ then gives us this locally closed decomposition

$$\mathcal{W} = \bigcup_{k=0}^{n-1} \mathcal{W}_k, \qquad \mathcal{W}_k \coloneqq \{\, (a,b,c,d,g) \in \mathcal{W} \mid \deg \gcd(a,b) = k \,\}.$$

By Lemma 5.5.7, $\mathcal{W}_k$ is an affine bundle over $\mathcal{T}_k$ of rank $R - r + 1 + k$. From this we can read off information about the dimensions of the $\mathcal{W}_k$ and $\mathcal{W}$ as

$$\dim \mathcal{W}_k = \dim \mathcal{T}_k + \operatorname{rank}_{\mathcal{T}_k} \mathcal{W}_k \begin{cases} = 2R+2 & k = 0 \\ < 2R+2 & k \geq 1 \end{cases} \qquad \dim \mathcal{W} = \max_k \dim \mathcal{W}_k = 2R+2.$$

Comparing this with the number of equations $R+r$ and the ambient dimension $3R+r+2$ shows that $\mathcal{W}$ is a (set-theoretic) complete intersection. In particular, *every* irreducible component of $\mathcal{W}$ has dimension exactly $2N+2$.

On the other hand, $\mathcal{W} = \bigcup_{k=0}^{n-1} \overline{\mathcal{W}_k}$ is a union of closed subsets with $\dim \overline{\mathcal{W}_k} < 2R+2$ for $k \geq 1$, so $\mathcal{W} = \overline{\mathcal{W}_0}$. We have seen that $\mathcal{W}_0$ is a bundle over $\mathcal{T}_0 = \mathcal{S}_0$ which is smooth and irreducible, so we conclude that $\mathcal{W}$ is irreducible. Finally, we see that $\mathcal{W}_0$ is a dense open subset of $\mathcal{W}$ which is reduced (even smooth), and by the unmixedness Theorem 0.1.26 $\mathcal{W}$ is reduced everywhere. □

Remark 5.5.8. We excluded the case $a = 0$, as it is not meaningful in $\widehat{Y}(z) = b(z)/a(z)$. If we set $a = 0$ in eq. (5.5), then the equation becomes $b \cdot c = 0$, and the solution set is the union of two linear spaces

$$\underbrace{(0 \times B \times 0 \times D \times G)}_{= \mathbb{V}(a,c) \cong \mathbb{C}^{2R}} \cup \underbrace{(0 \times 0 \times C \times D \times G)}_{= \mathbb{V}(a,b) \cong \mathbb{C}^{3R-r+1}}.$$

Combining the cases $a \neq 0$ and $a = 0$ shows that the solution set of eq. (5.5) in the "full" ambient space $ABCDG$ is given by $\overline{\mathcal{W}} \cup \mathbb{V}(a,b) \cup \mathbb{V}(a,c)$. These three varieties

are irreducible, but $\dim \mathbb{V}(a,c) < 2R+2$ is too small for $\mathbb{V}(a,c)$ to be an irreducible component. Hence $\mathbb{V}(a,c) \subseteq \overline{\mathcal{W}}$ and the two irreducible components are given by $\overline{\mathcal{W}}$ and $\mathbb{V}(a,b)$. ◁

Corollary 5.5.9. *Consider the projection $\tau\colon \mathcal{W} \to CD$, $(a,b,c,d,g) \to (c,d)$. For general $(c,d) \in CD$, every component of $\tau^{-1}(c,d)$ is reduced, of dimension 1, and consists of tuples (a,b,c,d,g) such that $a(z)$ has r distinct roots and $\gcd(a,b) = 1$.*

Note that this corollary does *not* claim the general fiber to be *non-empty*; though this will follow from Theorem 5.5.12 below. The one-dimensionality is due to the fact that eq. (5.5) is weighted-homogeneous in (a,b,g), hence solutions can be rescaled.

Proof. If τ is not dominant, then the theorem is vacuously true. Otherwise, by the fiber dimension theorem we have that general fibers are reduced of dimension

$$\dim \tau^{-1}(c,d) = \dim \mathcal{W} - \dim CD = 1.$$

The (negation of the) additional properties claimed in the theorem are given by polynomial equations in the coefficients of (a,b,g), such as the discriminant $\Delta(a)$ and the resultant $\operatorname{res}(a,b)$. Since $\mathcal{W}$ is irreducible and none of these conditions is trivial on $\mathcal{W}$, they form a smaller-dimensional subset. As such, they cannot map dominantly to CD, and so most fibers do not contain any such points. □

5.5.3 Multiparameter eigenvalue problem

In the previous section we have already noticed and exploited the affine-linear structure of the Walsh system. Since we treat (c,d) as the input data and (a,b,g) as the unknowns, the following reformulation of eq. (5.5) is a useful step towards eliminating b and g from the system:

$$M(a,c,d) \cdot \begin{pmatrix} -1 \\ b \\ g \end{pmatrix} = 0, \quad M(a,c,d) \coloneqq \begin{bmatrix} T(a \cdot d)^\mathsf{T} & T(c)^\mathsf{T} & T(\tilde{a}^2)^\mathsf{T} \end{bmatrix} \in \mathbb{C}[\underline{a}]^{(R+r)\times(R+1)}.$$

(all Toeplitz matrices have index $R+r-1$ here). The matrix $M(a,c,d)$ corresponds to the multiparameter eigenvalue problem formulation proposed in [LAD23].

Thus we are in a very similar position to Section 5.3.3, and we could introduce the bad locus as the subset of a such that the submatrix $B(a,c) \coloneqq [T(c)^\mathsf{T} \mid T(\tilde{a}^2)^\mathsf{T}]$ is rank-deficient. Luckily for us, in the infinite-dimensional case there is no bad locus for general data!

Lemma 5.5.10. *For given $(a,c) \in A^\circ C$, we have*

$$\dim_{\mathbb{C}} \operatorname{Ker} B(a,c) = \max\{0, \deg \gcd(c, \tilde{a}^2) - r\}.$$

If either c has distinct roots or if both a and c are stable, then $B(a,c)$ has maximal rank.

Proof. The kernel of B consists of syzygies $(u, v) \in S_{r-1} \times S_{N-n-1}$ such that $cu + \tilde{a}^2 v = 0$. One verifies exactly as in Lemma 5.5.7 that this space of syzygies has dimension equal to $\deg \gcd(c, \tilde{a}^2) - r$ if this number is non-negative, and 0 otherwise.

This formula shows that $B(a, c)$ has maximal rank if and only if $\deg \gcd(c, \tilde{a}^2) > r$. If c has distinct roots, then c and $\tilde{a}^2$ can have at most $\deg \tilde{a} = r$ roots in common. If both a and c are stable, then all roots of $c(z)$ are in the unit circle and all roots of $\tilde{a}$ are outside, so $\gcd(c, \tilde{a}^2) = 1$. □

Corollary 5.5.11. *For given (c, d) with c having distinct roots, the projection*

$$\mathrm{pr}_a \colon \{\, (a, b, g) \in A^\circ BG \mid (a, b, c, d, g) \in \mathcal{W} \,\} \to A^\circ$$

is an isomorphism onto the degeneracy locus of $M(a, c, d)$. If (c, d) is general, then the projective degeneracy locus $D_R(M(a, c, d)) \subseteq \mathbb{P}A$ is a finite set of reduced points.

Proof. By assumption and Lemma 5.5.10, the matrix $M(a, c, d)$ has rank R or $R + 1$ for any $a \neq 0$. Now apply Proposition 0.4.15 and, for the final part, Corollary 5.5.9. □

To obtain a count of the number of solutions, it remains to apply Thom–Porteous formula to the degeneracy locus $D_{c,d} \coloneqq D_R(M(a, c, d)) \subseteq \mathbb{P}A$.

Theorem 5.5.12. *If (c, d) is such that the set $D_{c,d}$ is finite, then*

$$\#\{\,\text{sol'ns to (5.5)}\,\} \leq \deg D_{c,d} = \sum_{j=0}^{r} \binom{R - r - 1 + j}{j} 2^j.$$

Moreover, if (c, d) are general, then equality holds.

Proof. Under the assumption that $D_{c,d}$ has dimension 0, we can apply the Thom–Porteous formula 0.4.13 in corank 1. The matrix represents a vector bundle morphism

$$M(a, c, d) \colon \mathcal{O}(-1) \oplus \mathcal{O}^r \oplus \mathcal{O}(-2)^{R-r} \longrightarrow \mathcal{O}^{R+r}$$

The corresponding Segre class is

$$\begin{aligned}[D_{c,d}] = s_r(\mathcal{O}(-1) \oplus \mathcal{O}(-2)^{R-r}) &= \left\{ \frac{1}{(1 - T)(1 - 2T)^{R-r}} \right\}^r \\ &= \sum_{j=0}^{r} \binom{R - r - 1 + j}{j} 2^j. \end{aligned}$$

□

Several values of the formula in Theorem 5.5.12 are depicted in Table 5.3. Notice for example that if we reduce the order of the given model by one, i. e., $R = r + 1$, then the formula specializes to $\sum_{j=0}^{r} 2^j = 2^{r+1} - 1$.

Example 5.5.13 (The four disk system)**.** In [LAD23, Example 2], the authors considered the critical equations for $(R, r) = (6, 2)$ and data point

$$\frac{d(z)}{c(z)} = \frac{0.0448z^5 + 0.2368z^4 + 0.0013z^3 + 0.0211z^2 + 0.2250z + 0.0219}{z^6 - 1.2024z^5 + 2.3675z^4 - 2.0039z^3 + 2.2337z^2 - 1.0420z + 0.8513}.$$

$R \setminus r$	1	2	3	4	5	6	7	8
1	1							
2	3	1						
3	5	7	1					
4	7	17	15	1				
5	9	31	49	31	1			
6	11	49	111	129	63	1		
7	13	71	209	351	321	127	1	
8	15	97	351	769	1 023	769	255	1
9	17	127	545	1 471	2 561	2 815	1 793	511

Table 5.3: Evaluation of the formula from Theorem 5.5.12 for various problem sizes (R, r).

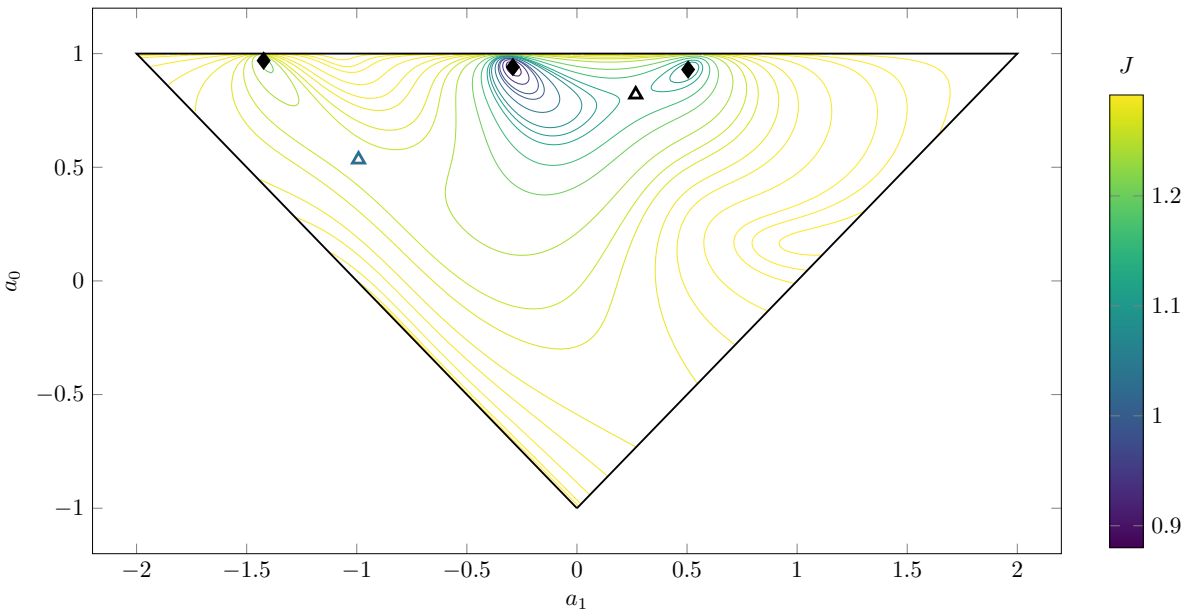


Figure 5.1: Contour plot of the objective function $J(b/a)$ from Example 5.5.13 for stable $a(z) = z^2 + a_1 z + a_0 \in \mathbb{R}[z]_{\leq 2}$ and optimal $b(z)$. There are five real stable critical points: Three local minima and two saddle points.

They report the system from Walsh's theorem to have 49 complex solutions, exactly matching the value predicted by Theorem 5.5.12. For this instance, 11 of the solutions are real-valued, 5 of which corresponded to stable models. We can visualize the optimization landscape in this case as follows:

Given a (real, stable) denominator $a(z)$, the optimal choice $b(z)$ that minimizes $J(b/a)$ is uniquely determined as an orthogonal projection. Therefore we can map a stable $a(z)$ to the corresponding value of $J(b/a)$, creating a contour plot. This is depicted in Figure 5.1 for the particular example, highlighting the five stable critical points. ◁

Completely analogously to the results in Section 5.3.4, we also obtain a result on the optimality of the polyhedral homotopy. In order to turn the system eq. (5.5) from Walsh's theorem into a zero-dimensional system, we again normalize for example by $a_r = 1$.

Theorem 5.5.14. *For general (c, d), the Newton polytopes of eq. (5.5) have mixed volume $\sum_{j=0}^{r} \binom{R-r-1+j}{j} 2^j$ and homotopy continuation from a polyhedral start system can locate all solutions.*

We can also give an estimate for certain special data (c, d):

Proposition 5.5.15. *If (c, d) is such that $D_{c,d}$ is finite but c has exactly one double root, then the number of critical points is bounded by*

$$\#\{\text{ sol'ns to (5.5)}\} \leq \sum_{j=0}^{r} \binom{R-r-1+j}{j} 2^j - \binom{R-2}{r-1}.$$

Proof. Write $c = (s-p_1)^2 \prod_{i=2}^{N-1}(s-p_i)$ with $p_1, \ldots, p_{R-1}$ distinct. The set $D_{c,d}$ can be split into those points that correspond to (5.5) and those corresponding to the bad locus where $B(a, c)$ drops rank. From Lemma 5.5.10 we deduce that the latter happens if and only if $\tilde{a} = (s-p_1)\prod_{i\in I}(s-p_i)$, where $I \subseteq \{2, \ldots, R-1\}$ is a selection of $r-1$ roots of c. Thus the bad locus consists of exactly $\binom{N-2}{n-1}$ points, potentially with multiplicity. We conclude using the first part of Theorem 5.5.12

$$\begin{aligned}\#\{\text{ sol'ns to (5.5)}\} &\leq \deg D_{c,d} - \#\{\, a \mid \operatorname{rank} B(a,c) < R\,\} \\ &\leq \sum_{j=0}^{r} \binom{R-r-1+j}{j} 2^j - \binom{R-2}{r-1}.\end{aligned}$$

□

Conclusions

> If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is.
>
> JOHN VON NEUMANN [Jor]

In this thesis we have solved various computational challenges, ranging from Mukai lifting to symmetric tensor decomposition and to Euclidean distance minimization. We have seen that logarithmic discriminants interact with the combinatorics of reciprocal linear spaces in interesting ways, and that infinitely generated initial algebras can sometimes still be tamed combinatorially. In this final section, we explore future research directions related to the topics of this thesis.

Chapter 1 has laid the foundation for further investigation on the Hilbert functions of chopped ideals of points, yet there remains a lot to research. Possible future directions include proving the Expected Syzygy Conjecture 1.3.10 in other families, extending the results to positive characteristics and to the multigraded setting.

In Chapter 2 we have successfully tackled the Mukai lifting problem using the skew normal form homotopy. As mentioned in Section 2.4, Petrakiev's map f (Theorem 2.2.15) is conjecturally birational modulo SL(6). Future work on the (numerical) orbit membership problem could complete the steps outlined in Section 2.4.2 to test this conjecture.

Studying the computational complexity of polynomial subalgebras in Chapter 3 has revealed many parallels but also differences to the ideal world. A major open challenge is to algorithmically decide $\text{SAGBIFINITE}_{\mathbb{K}}$, that is, finiteness of the initial algebra. Perhaps a positive or negative answer to the Semilinearity Conjecture 3.5.12 could reveal some insight into the inherent complexity of infinitely generated initial algebras.

In Chapter 4 we have studied the logarithmic discriminant of linear models from algebraic statistics, focusing mostly on the bi-uniform case. Possible next steps include a deeper investigation of special arrangements such as $\mathcal{M}_{0,m}$ or graphical arrangements. Another possibility is to look at hypersurface arrangements of higher degree or other relevant models from algebraic statistics such as scaled toric varieties.

Finally, in Chapter 5 we pushed the state-of-the-art about what is known about specific ED degrees of $\mathcal{X}_{d,r}$. It would be very desirable to investigate Conjecture 5.4.11 further and prove it in the cases at least when the bad locus is (conjecturally) finite ($r \leq 3$). In the infinite Hankel matrix case, Theorem 5.5.3 naturally leads to the question about critical points with multiple poles, and the question whether global minimizers can be found on these other components of the optimization landscape.

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 pp. 95, 96, 99).

[ame25] silviana amethyst. *Computer programs*. Accessed: 2025-08-05. 2025. URL: https://silviana.org/computer_programs/ (cit. on p. 72).

[And13] Dave Anderson. "Okounkov bodies and toric degenerations". In: *Mathematische Annalen* 356.3 (2013), pp. 1183–1202. ISSN: 1432-1807. DOI: 10.1007/s00208-012-0880-3 (cit. on pp. 22, 84).

[And16] Jenny Anderson. *A scientist calculated the cost of not being a straight man, and she wants a tax cut*. 2016. URL: https://qz.com/631455/a-scientist-cacluated-the-cost-of-not-being-a-straight-man-and-she-wants-a-tax-cut (visited on 09/03/2025) (cit. on p. 120).

[Ati02] Michael Atiyah. "Mathematics in the 20th century". In: *Bulletin of the London Mathematical Society* 34.1 (Jan. 2002), pp. 1–15. ISSN: 1469-2120. DOI: 10.1112/s0024609301008566 (cit. on p. 8).

[Bal87] E. Ballico. "Generators for the homogeneous ideal of s general points in $\mathbb{P}^3$". In: *J. Algebra* 106.1 (1987), pp. 46–52. DOI: 10.1016/0021-8693(87)90020-2 (cit. on p. 47).

[BBT24] Barbara Betti, Viktoriia Borovik, and Simon Telen. "Proudfoot-Speyer degenerations of scattering equations". In: *arXiv:2410.03614* (2024) (cit. on p. 140).

[BC13] Peter Bürgisser and Felipe Cucker. *Condition: The Geometry of Numerical Algorithms*. Springer Berlin Heidelberg, 2013. ISBN: 9783642388965. DOI: 10.1007/978-3-642-38896-5 (cit. on p. 2).

[BCP83] A. Borodin, S. Cook, and N. Pippenger. "Parallel computation for well-endowed rings and space-bounded probabilistic machines". In: *Information and Control* 58.1 (1983), pp. 113–136. ISSN: 0019-9958. DOI: 10.1016/S0019-9958(83)80060-6 (cit. on pp. 103, 104).

[Ber+17] A. Bernardi, N. S. Daleo, J. D. Hauenstein, and B. Mourrain. "Tensor decomposition and homotopy continuation". In: *Diff. Geom. Appl.* 55 (2017), pp. 78–105. DOI: 10.1016/j.difgeo.2017.07.009 (cit. on p. 43).

[Ber+18] A. Bernardi, E. Carlini, M. V. Catalisano, A. Gimigliano, and A. Oneto. "The hitchhiker guide to: Secant varieties and tensor decomposition". In: *Mathematics* 6.12 (2018), p. 314. DOI: 10.3390/math6120314 (cit. on pp. 24, 40).

[Ber79] D. N. Bernshtein. "The number of roots of a system of equations". In: *Functional Analysis and Its Applications* 9.3 (1979), pp. 183–185. ISSN: 1573-8485. DOI: 10.1007/bf01075595 (cit. on p. 73).

[Bet25] Barbara Betti. "Khovanskii and Gröbner Degenerations in Computational Commutative Algebra". PhD thesis. Leipzig, Germany: Fakultät für Mathematik und Informatik, Universität Leipzig, 2025 (cit. on p. ix).

[Bez+12] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. "Julia: A Fast Dynamic Language for Technical Computing". In: (2012). arXiv: 1209.5145 [cs.PL] (cit. on p. 70).

[BK25] Barbara Betti and Leonie Kayser. "Mukai Lifting of Self-Dual Points in $\mathbb{P}^6$". In: *Experimental Mathematics* (June 2025), pp. 1–13. ISSN: 1944-950X. DOI: 10.1080/10586458.2025.2513603 (cit. on pp. ix, 70).

[BL13] J. Buczyński and J. M. Landsberg. "Ranks of tensors and a generalization of secant varieties". In: *Lin. Alg. Appl.* 438.2 (2013), pp. 668–689. DOI: 10.1016/j.laa.2012.05.001 (cit. on p. 24).

[BM93] Dave Bayer and David Mumford. *What can be computed in algebraic geometry?* 1993. DOI: 10.48550/ARXIV.ALG-GEOM/9304003 (cit. on p. 2).

[BPT25] Barbara Betti, Marta Panizzut, and Simon Telen. "Solving equations using Khovanskii bases". In: *Journal of Symbolic Computation* 126 (2025), p. 102340. ISSN: 0747-7171. DOI: https://doi.org/10.1016/j.jsc.2024.102340 (cit. on p. 83).

[BS88] David Bayer and Michael Stillman. "On the complexity of computing syzygies". In: *Journal of Symbolic Computation* 6.2 (1988), pp. 135–147. ISSN: 0747-7171. DOI: 10.1016/S0747-7171(88)80039-7 (cit. on pp. 111, 112).

[BSW23] M. Burr, F. Sottile, and E. Walker. "Numerical homotopies from Khovanskii bases". In: *Mathematics of Computation* 92.343 (Mar. 2023), pp. 2333–2353. ISSN: 1088-6842. DOI: 10.1090/mcom/3689 (cit. on p. 73).

[BT18] Paul Breiding and Sascha Timme. "HomotopyContinuation.jl: A Package for Homotopy Continuation in Julia". In: *Mathematical Software – ICMS 2018*. Ed. by James H. Davenport, Manuel Kauers, George Labahn, and Josef Urban. Cham: Springer International Publishing, 2018, pp. 458–465 (cit. on p. 72).

[BT25] Matías Bender and Simon Telen. "Yet another eigenvalue algorithm for solving polynomial systems". In: *Varieties, Polyhedra, Computation*. EMS Press, Dec. 2025, pp. 355–393. ISBN: 9783985475964. DOI: 10.4171/ecr/22/12 (cit. on p. 36).

[Buc06] Bruno Buchberger. "Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal". In: *Journal of Symbolic Computation* 41.3 (2006), pp. 475–511. ISSN: 0747-7171. DOI: 10.1016/j.jsc.2005.09.007 (cit. on p. 2).

[Bur+24] Michael Burr, Oliver Clarke, Timothy Duff, Jackson Leaman, Nathan Nichols, and Elise Walker. "SubalgebraBases in Macaulay2". In: *Journal of Software for Algebra and Geometry* 14.1 (May 2024), pp. 97–109. ISSN: 1948-7916. DOI: 10.2140/jsag.2024.14.97 (cit. on pp. 4, 94, 107).

[Bür98] Peter Bürgisser. "On the Parallel Complexity of the Polynomial Ideal Membership Problem". In: *Journal of Complexity* 14.2 (1998), pp. 176–189. ISSN: 0885-064X. DOI: 10.1006/jcom.1998.0472 (cit. on p. 23).

[BV88] Winfried Bruns and Udo Vetter. *Determinantal Rings*. Springer Berlin Heidelberg, 1988. ISBN: 9783540392743. DOI: 10.1007/bfb0080378 (cit. on p. 15).

[Cay46] Arthur Cayley. "Sur quelques propriétés des déterminants gauches." French. In: *Journal für die reine und angewandte Mathematik* 32 (1846), pp. 119–123. URL: http://eudml.org/doc/147336 (cit. on p. 77).

[Cha07] Marc Chardin. "Some results and questions on Castelnuovo-Mumford regularity". In: *Lecture Notes in Pure and Applied Mathematics* 254 (2007), p. 1 (cit. on p. 16).

[Chi03] Yasuko Chikuse. *Statistics on Special Manifolds*. en. 2003rd ed. Lecture notes in statistics. New York, NY: Springer, Feb. 2003. DOI: 10.1007/978-0-387-21540-2 (cit. on p. 85).

[Cho92] Timothy Y. Chow. *Reconciling alternative definitions of graded Cohen-Macaulay rings*. https://timothychow.net/CM.pdf. 1992 (cit. on p. 15).

[CHY14] Freddy Cachazo, Song He, and Ellis Ye Yuan. "Scattering equations and Kawai-Lewellen-Tye orthogonality". In: *Physical Review D* 90.6 (2014), p. 065001 (cit. on pp. 133, 149).

[CLO16] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer International Publishing, Oct. 2016. DOI: 10.1007/978-3-662-41154-43 (cit. on pp. 18–20).

[CUZ20] Freddy Cachazo, Bruno Umbert, and Yong Zhang. "Singular solutions in soft limits". In: *Journal of High Energy Physics* 2020.5 (2020), pp. 1–33 (cit. on p. 152).

[DE05] Alicia Dickenstein and Ioannis Z. Emiris, eds. *Solving Polynomial Equations*. Springer-Verlag, 2005. ISBN: 3540243267. DOI: 10.1007/b138957 (cit. on p. 33).

[Dic+91] Alicia Dickenstein, Noaï Fitchas, Marc Giusti, and Carmen Sessa. "The membership problem for unmixed polynomial ideals is solvable in single exponential time". In: *Discrete Applied Mathematics* 33.1 (1991), pp. 73–94. ISSN: 0166-218X. DOI: 10.1016/0166-218X(91)90109-A (cit. on p. 105).

[Dij86] E. W. Dijkstra. "On a cultural gap". In: *The Mathematical Intelligencer* 8.1 (Mar. 1986), pp. 48–52. ISSN: 0343-6993. DOI: 10.1007/bf03023921 (cit. on p. 94).

[Dil90] Annie Dillard. *The writing life*. New York, NY: HarperPerennial, Sept. 1990 (cit. on p. 32).

[DK02] Harm Derksen and Gregor Kemper. *Computational Invariant Theory*. Springer Berlin Heidelberg, 2002. ISBN: 9783662049587. DOI: 10.1007/978-3-662-04958-7 (cit. on p. 89).

[DO88] Igor Dolgachev and David Ortland. *Point sets in projective spaces and theta functions*. en. Astérisque 165. Société mathématique de France, 1988. URL: http://www.numdam.org/item/AST_1988__165__1_0/ (cit. on pp. 78, 79).

[DR88] John. De Pillis and Nicholas J. Rose. *Mathematical maxims and minims*. eng. Raleigh N.C: Rome Press, 1988 (cit. on p. 58).

[Dra+15] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R. Thomas. "The Euclidean Distance Degree of an Algebraic Variety". In: *Foundations of Computational Mathematics* 16.1 (Jan. 2015), pp. 99–149. ISSN: 1615-3383. DOI: 10.1007/s10208-014-9240-x (cit. on pp. 158–160).

[DS95] John Dalbec and Bernd Sturmfels. "Introduction to Chow Forms". In: *Invariant Methods in Discrete and Computational Geometry*. Springer Netherlands, 1995, pp. 37–58. ISBN: 9789401584029. DOI: 10.1007/978-94-015-8402-9_2 (cit. on p. 125).

[Dub90] Thomas W. Dubé. "The Structure of Polynomial Ideals and Gröbner Bases". In: *SIAM Journal on Computing* 19.4 (Aug. 1990), pp. 750–773. DOI: 10.1137/0219053 (cit. on p. 105).

[EH06] David Eisenbud and Joe Harris. *The geometry of schemes*. Vol. 197. Springer Science & Business Media, 2006 (cit. on p. 121).

[EH16] D. Eisenbud and J. Harris. *3264 and All That - A Second Course in Algebraic Geometry*. Cambridge: Cambridge University Press, 2016 (cit. on pp. 8, 25, 27, 29, 60, 179).

[EHL22] Lawrence Ein, Huy Tài Hà, and Robert Lazarsfeld. "Saturation bounds for smooth varieties". In: *Algebra & Number Theory* 16.6 (Sept. 2022), pp. 1531–1546. ISSN: 1937-0652. DOI: 10.2140/ant.2022.16.1531 (cit. on p. 55).

[Ehr51] Charles Ehresmann. "Les connexions infinitésimales dans un espace fibré différentiable". In: *Colloque de topologie (espaces fibrés), Bruxelles, 1950*. Georges Thone, Liège, 1951, pp. 29–55 (cit. on p. 122).

[Eis+02] D. Eisenbud, S. Popescu, F.-O. Schreyer, and C. Walter. "Exterior algebra methods for the minimal resolution conjecture". In: *Duke Math. J.* 112.2 (2002), pp. 379–395. DOI: 10.1215/S0012-9074-02-11226-5 (cit. on p. 47).

[Eis05] D. Eisenbud. *The Geometry of Syzygies: A Second Course in Commutative Algebra and Algebraic Geometry*. Springer New York, 2005. DOI: 10.1007/0-387-26456-6_3 (cit. on pp. 8, 12, 16).

[Eis95] D. Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*. Vol. 150. Graduate Texts in Mathematics. New York: Springer-Verlag, 1995 (cit. on pp. 8, 16, 39, 47, 65).

[EJT12] Michael Elberfeld, Andreas Jakoby, and Till Tantau. "Algorithmic Meta Theorems for Circuit Classes of Constant and Logarithmic Depth". In: *STACS 2012*. Vol. 14. LIPIcs. Dagstuhl, Germany, 2012, pp. 66–77. ISBN: 978-3-939897-35-4. DOI: 10.4230/LIPIcs.STACS.2012.66 (cit. on p. 114).

[EP00] David Eisenbud and Sorin Popescu. "The Projective Geometry of the Gale Transform". In: *Journal of Algebra* 230.1 (2000), pp. 127–173. ISSN: 0021-8693. DOI: 10.1006/jabr.1999.7940 (cit. on pp. 76, 80).

[EP96] D. Eisenbud and S. Popescu. "Gale duality and free resolutions of ideals of points". In: *Inventiones mathematicae* 136 (1996), pp. 419–449. DOI: 10.1007/s002220050315 (cit. on p. 47).

[ES94] David Eisenbud and Bernd Sturmfels. "Binomial Ideals". In: *Duke Math. J.* 84 (Feb. 1994). DOI: 10.1215/S0012-7094-96-08401-X (cit. on pp. 21, 22).

[Frö85] R. Fröberg. "An inequality for Hilbert series of graded algebras". In: *Mathematica Scandinavica* 56.2 (1985), pp. 117–144. URL: https://www.jstor.org/stable/24491560 (cit. on pp. 48, 49, 67).

[Ful98] William Fulton. *Intersection Theory*. Springer New York, 1998. ISBN: 9781461217008. DOI: 10.1007/978-1-4612-1700-8 (cit. on pp. 8, 23, 25, 27, 29, 184).

[GAB08] S. Gugercin, A.C. Antoulas, and C. Beattie. "$\mathcal{H}_2$ Model Reduction for Large-Scale Linear Dynamical Systems". In: *SIAM Journal on Matrix Analysis and Applications* 30.2 (2008), pp. 609–638 (cit. on p. 171).

[Gei+23] Alheydis Geiger, Sachi Hahimoto, Bernd Sturmfels, and Raluca Vlad. "Self-Dual Matroids from Canonical Curves". In: *Experimental Mathematics* (2023), pp. 1–22. DOI: 10.1080/10586458.2023.2239282 (cit. on pp. 4, 70).

[GGR86] A. V. Geramita, D. Gregory, and L. Roberts. "Monomial ideals and points in projective space". In: *J. Pure Appl. Algebra* 40 (1986), pp. 33–62. DOI: 10.1016/0022-4049(86)90029-0 (cit. on pp. 47, 52).

[GJ20] Andre Galligo and Zbigniew Jelonek. "Elimination ideals and Bézout relations". In: *Journal of Algebra* 562 (2020), pp. 621–626. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2020.06.022 (cit. on p. 108).

[GKT25] Fulvio Gesmundo, Leonie Kayser, and Simon Telen. "Hilbert Functions of Chopped Ideals". In: *Journal of Algebra* 666 (2025), pp. 415–445. ISSN: 0021-8693. DOI: https://doi.org/10.1016/j.jalgebra.2024.11.017 (cit. on pp. ix, 24, 25, 32).

[GKZ08] Israel M Gelfand, Mikhail Kapranov, and Andrei Zelevinsky. *Discriminants, resultants, and multidimensional determinants*. Springer Science & Business Media, 2008 (cit. on pp. 123–125).

[GM84] A. V. Geramita and P. Maroscia. "The ideal of forms vanishing at a finite set of points in $\mathbb{P}^n$". In: *J. Algebra* 90.2 (1984), pp. 528–555. DOI: 10.1016/0021-8693(84)90188-1 (cit. on p. 47).

[GO82] A.V Geramita and F Orecchia. "Minimally generating ideals defining certain tangent cones". In: *Journal of Algebra* 78.1 (Sept. 1982), pp. 36–57. ISSN: 0021-8693. DOI: 10.1016/0021-8693(82)90101-6 (cit. on p. 46).

[GS] D. R. Grayson and M. E. Stillman. *Macaulay2, a software system for research in algebraic geometry*. Available at http://www.math.uiuc.edu/Macaulay2/. (version 1.24.11) (cit. on pp. 1, 44, 135).

[GS66] Seymour Ginsburg and Edwin H. Spanier. "Semigroups, Presburger formulas, and languages." In: *Pacific Journal of Mathematics* 16.2 (1966), pp. 285–296 (cit. on pp. 115, 117).

[Har92] J. Harris. *Algebraic geometry. A first course*. Vol. 133. Graduate Texts in Mathematics. New York: Springer-Verlag, 1992. ISBN: 0-387-97716-3 (cit. on p. 8).

[HDS90] J. F. Hauer, C. J. Demeure, and L. L. Scharf. "Initial results in Prony analysis of power system response signals". In: *IEEE Transactions on Power Systems* 5.1 (1990), pp. 80–89 (cit. on p. 171).

[Her26] Grete Hermann. "Die Frage der endlich vielen Schritte in der Theorie der Polynomideale". In: *Mathematische Annalen* 95 (1926), pp. 736–788 (cit. on p. 105).

[HH99] J. Herzog and T. Hibi. "Componentwise linear ideals". In: *Nagoya Mathematical Journal* 153.none (1999), pp. 141–153 (cit. on p. 44).

[Hil02] David Hilbert. "Mathematical problems". In: *Bulletin of the American Mathematical Society* 8.10 (1902), pp. 437–479 (cit. on p. 1).

[Hof99] Douglas R Hofstadter. *Godel, Escher, Bach*. London, England: Basic Books, Jan. 1999 (cit. on p. iii).

[HS14] June Huh and Bernd Sturmfels. "Likelihood Geometry". In: *Combinatorial Algebraic Geometry*. Springer International Publishing, 2014, pp. 63–117. ISBN: 9783319048703. DOI: 10.1007/978-3-319-04870-3_3 (cit. on pp. 5, 126–130).

[HS95] Birkett Huber and Bernd Sturmfels. "A polyhedral method for solving sparse polynomial systems". In: *Mathematics of Computation* 64.212 (1995), pp. 1541–1555. ISSN: 1088-6842. DOI: 10.1090/s0025-5718-1995-1297471-4 (cit. on p. 74).

[HS96] A. Hirschowitz and C. Simpson. "La résolution minimale de l'idéal d'un arrangement général d'un grand nombre de points dans $\mathbb{P}^n$". In: *Inventiones mathematicae* 126 (1996), pp. 467–503. DOI: 10.1007/s002220050107 (cit. on p. 47).

[HU79] John E Hopcroft and Jeffrey D Ullman. *An introduction to automata theory, languages, and computation*. Addison-Wesley series in computer science. Pearson, Jan. 1979. ISBN: 978-0201029888 (cit. on p. 95).

[Huh13] June Huh. "The maximum likelihood degree of a very affine variety". In: *Compositio Mathematica* 149.8 (2013), pp. 1245–1266. DOI: 10.1112/S0010437X13007057 (cit. on p. 128).

[IK99] A. Iarrobino and V. Kanev. *Power sums, Gorenstein algebras, and determinantal loci*. Vol. 1721. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 1999 (cit. on pp. 3, 39, 40, 42, 45, 167).

[Joa17] Jürgen Gerhard Joachim von zur Gathen. *Modern Computer Algebra*. Cambridge University Press, Mar. 2017. 812 pp. ISBN: 1107039037. DOI: 10.1017/CBO9781139856065 (cit. on pp. 18, 19, 95).

[Jor] Palle Jorgensen. URL: https://homepage.math.uiowa.edu/~jorgen/vonneumannquotesource.html (visited on 01/15/2026) (cit. on p. 201).

[Kan17] Daniel M. Kane. *Unary Subset-Sum is in Logspace*. 2017. arXiv: 1012.1336 [cs.CC] (cit. on pp. 102, 114).

[Kay22] Leonie Kayser. "Gröbner Bases and Their Complexity". MA thesis. Oct. 2022 (cit. on p. 111).

[Kay25] Elisabeth Leonie Kayser. "Computational Complexity of Polynomial Subalgebras". In: *Proceedings of the 2025 International Symposium on Symbolic and Algebraic Computation*. ISSAC '25. ACM, July 2025, pp. 337–344. DOI: 10.1145/3747199.3747578 (cit. on pp. ix, 94).

[KKT] Leonie Kayser, Andreas Kretschmer, and Simon Telen. MathRepo page LogDiscHyperplanes https://mathrepo.mis.mpg.de/LogDiscHyperplanes, 2024 (cit. on pp. 45, 151).

[KKT25] L. Kayser, A. Kretschmer, and S. Telen. "Logarithmic discriminants of hyperplane arrangements". In: *Le Matematiche* 80.1 (2025). ISSN: 03733505, 20375298. DOI: 10.4418/2025.80.1.13 (cit. on pp. ix, 120, 151).

[KM19] Kiumars Kaveh and Christopher Manon. "Khovanskii Bases, Higher Rank Valuations, and Tropical Geometry". In: *SIAM Journal on Applied Algebra and Geometry* 3.2 (2019), pp. 292–336. DOI: 10.1137/17M1160148 (cit. on pp. 83, 84).

[KM96] Klaus Kühnle and Ernst W. Mayr. "Exponential Space Computation of Gröbner Bases". In: *Proceedings of the 1996 International Symposium on Symbolic and Algebraic Computation*. ISSAC '96. Zurich, Switzerland: Association for Computing Machinery, 1996, pp. 63–71. ISBN: 0897917960. DOI: 10.1145/236869.236900 (cit. on pp. 106–108).

[Koz+24] Khazhgali Kozhasov, Alan Muniz, Yang Qi, and Luca Sodomaco. *On the minimal algebraic complexity of the rank-one approximation problem for general inner products*. 2024. arXiv: 2309.15105 [math.AG] (cit. on pp. 168, 188).

[KR00] Martin Kreuzer and Lorenzo Robbiano. *Computational Commutative Algebra 1*. Springer Berlin Heidelberg, 2000. DOI: 10.1007/978-3-540-70628-1 (cit. on pp. 19, 21).

[Kra03] Marcus Kracht. *The mathematics of language*. en. Studies in Generative Grammar [SGG]. Berlin, Germany: De Gruyter Mouton, Aug. 2003 (cit. on p. 117).

[Kro81] L. Kronecker. "Zur Theorie der Elimination einer Variablen aus zwei algebraischen Gleichungen". In: *Monatsberichte der Königlich Preussischen Akademie der Wissenschaften (Berlin)* (1881). Reprinted as pp. 113–192 in Mathematische Werke, vol. 2, B.G. Teubner, Leipzig, 1897., pp. 535–600 (cit. on p. 167).

[Kur02] Shigeru Kuroda. "The infiniteness of the SAGBI bases for certain invariant rings". In: *Osaka Journal of Mathematics* 39.3 (2002), pp. 665–680 (cit. on p. 116).

[LAD23] S. Lagauw, O. M. Agudelo, and B. De Moor. "Globally Optimal SISO H_2-Norm Model Reduction Using Walsh's Theorem". In: *IEEE Control Systems Letters* 7 (2023), pp. 1670–1675 (cit. on pp. 192, 196, 197).

[Lag26] S. Lagauw. "Global Solution to Problems in System Identification and Model Order Reduction". PhD thesis. Leuven, Belgium: Faculty of Engineering, KU Leuven, 2026 (cit. on pp. ix, 156).

[Lam24] Thomas Lam. "Moduli spaces in positive geometry". In: *arXiv:2405.17332* (2024) (cit. on pp. 133, 149).

[Lan12] J. M. Landsberg. *Tensors: Geometry and Applications*. Vol. 128. Graduate Studies in Mathematics. Providence, RI: American Mathematical Society, 2012, pp. xx+439. ISBN: 978-0-8218-6907-9 (cit. on pp. 40, 53).

[LL25] Chi-Kwong Li and Jephian C. -H. Lin. *Confluent Vandermonde matrix and related topics*. 2025. arXiv: 2403.01474 [math.CO]. URL: https://arxiv.org/abs/2403.01474 (cit. on p. 163).

[Lor93] A. Lorenzini. "The minimal resolution conjecture". In: *J. Algebra* 156.1 (1993), pp. 5–35. DOI: 10.1006/jabr.1993.1060 (cit. on p. 46).

[LS16] Hwangrae Lee and Bernd Sturmfels. "Duality of multiple root loci". In: *Journal of Algebra* 446 (Jan. 2016), pp. 499–526. ISSN: 0021-8693. DOI: 10.1016/j.jalgebra.2015.08.029 (cit. on p. 187).

[LVD26] S. Lagauw, L. Vanpoucke, and B. De Moor. "Exact Solution to the Least Squares Realization Problem as a Multiparameter Eigenvalue Problem". In: *Automatica* 185 (2026), p. 112792 (cit. on pp. 172, 175, 177).

[Mat70] Yu. V. Matiyasevich. "Enumerable sets are diophantine". English. In: *Sov. Math., Dokl.* 11 (1970), pp. 354–358. ISSN: 0197-6788 (cit. on p. 96).

[Mat87] H. Matsumura. *Commutative Ring Theory*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1987. DOI: 10.1017/CBO9781139171762 (cit. on p. 8).

[May89] Ernst W. Mayr. "Membership in polynomial ideals over Q is exponential space complete". In: *STACS 89*. Ed. by B. Monien and R. Cori. Berlin, Heidelberg: Springer Berlin Heidelberg, 1989, pp. 400–406. ISBN: 978-3-540-46098-5 (cit. on pp. 94, 104).

[May97] Ernst W. Mayr. "Some Complexity Results for Polynomial Ideals". In: *Journal of Complexity* 13.3 (1997), pp. 303–325. ISSN: 0885-064X. DOI: 10.1006/jcom.1997.0447 (cit. on pp. 105, 109, 112).

[MFK94] David Mumford, John Fogarty, and Frances Kirwan. *Geometric Invariant Theory*. Ergebnisse der Mathematik und Ihrer Grenzgebiete, 3 Folge/A Series of Modern Surveys in Mathematics Series. Springer Berlin Heidelberg, 1994. ISBN: 9783540569633 (cit. on p. 78).

[Mig98] J. Migliore. *Introduction to liaison theory and deficiency modules*. Vol. 165. Springer Science & Business Media, 1998 (cit. on p. 65).

[ML67] L. Meier and D. Luenberger. "Approximation of linear constant systems". In: *IEEE Transactions on Automatic Control* 12.5 (1967), pp. 585–588. DOI: 10.1109/TAC.1967.1098680 (cit. on pp. 191, 192).

[MM82] Ernst W. Mayr and Albert R. Meyer. "The complexity of the word problems for commutative semigroups and polynomial ideals". In: *Advances in Mathematics* 46.3 (Dec. 1982), pp. 305–329. DOI: 10.1016/0001-8708(82)90048-2 (cit. on pp. 94, 104, 105, 109, 111, 112).

[MN02] J. Migliore and U. Nagel. "Liaison and related topics: Notes from the Torino Workshop/School". In: *arXiv:0205161* (2002) (cit. on p. 65).

[Mon21] Pinaki Mondal. *How Many Zeroes?: Counting Solutions of Systems of Polynomials via Toric Geometry at Infinity*. Springer International Publishing, 2021. ISBN: 9783030751746. DOI: 10.1007/978-3-030-75174-6 (cit. on pp. 73, 74).

[Moo20] B. de Moor. "Least squares optimal realisation of autonomous LTI systems is an eigenvalue problem". In: *Communications in Information and Systems* 20.2 (2020), pp. 163–207 (cit. on p. 171).

[MR13] Ernst W. Mayr and Stephan Ritscher. "Dimension-dependent bounds for Gröbner bases of polynomial ideals". In: *Journal of Symbolic Computation* 49 (2013). The International Symposium on Symbolic and Algebraic Computation, pp. 78–94. ISSN: 0747-7171. DOI: 10.1016/j.jsc.2011.12.018 (cit. on p. 105).

[MR17] Abraham Martín del Campo and Jose Israel Rodriguez. "Critical points via monodromy and local methods". In: *Journal of Symbolic Computation* 79 (Mar. 2017), pp. 559–574. ISSN: 0747-7171. DOI: 10.1016/j.jsc.2016.07.019 (cit. on p. 88).

[MRW20] Laurentiu G. Maxim, Jose Israel Rodriguez, and Botong Wang. "Defect of Euclidean distance degree". In: *Advances in Applied Mathematics* 121 (Oct. 2020), p. 102101. ISSN: 0196-8858. DOI: 10.1016/j.aam.2020.102101 (cit. on p. 161).

[MS21] Mateusz Michalek and Bernd Sturmfels. *Invitation to nonlinear algebra*. Graduate studies in mathematics 211. Providence, Rhode Island: American Mathematical Society, 2021. ISBN: 978-1-4704-5367-1. URL: https://bookstore.ams.org/gsm-211 (cit. on p. 82).

[MT17] Ernst W. Mayr and Stefan Toman. "Complexity of Membership Problems of Different Types of Polynomial Ideals". In: *Algorithmic and Experimental Methods in Algebra, Geometry, and Number Theory*. Ed. by Gebhard Böckle, Wolfram Decker, and Gunter Malle. Cham: Springer International Publishing, 2017, pp. 481–493. ISBN: 978-3-319-70566-8. DOI: 10.1007/978-3-319-70566-8_20 (cit. on pp. 4, 94, 102).

[Muk88] Shigeru Mukai. "Curves, K3 Surfaces and Fano 3-folds of Genus ≤ 10". In: *Algebraic Geometry and Commutative Algebra*. Academic Press, 1988, pp. 357–377. ISBN: 978-0-12-348031-6. DOI: 10.1016/B978-0-12-348031-6.50026-7 (cit. on pp. 4, 70, 74, 80).

[Muk92] Shigeru Mukai. "Curves and Grassmannians". In: *Algebraic geometry and related topics. Proceedings of the international symposium, held in Incheon, Republic of Korea, February 11-13, 1992*. Cambridge, MA: International Press, 1992, pp. 19–40. ISBN: 1-57146-013-6 (cit. on pp. 4, 70, 80).

[Mum99] David B. Mumford. "Appendix: Curves and Their Jacobians". In: *The Red Book of Varieties and Schemes*. Springer Berlin Heidelberg, 1999, pp. 225–291. ISBN: 9783540460213. DOI: 10.1007/978-3-540-46021-3_4 (cit. on p. 156).

[OM17] Ben Obiero and Damian Maingi. "On the minimal resolution conjecture for the ideal of general points in P^4". In: *Pure Mathematical Sciences* 6 (2017), pp. 67–85. ISSN: 1314-7560. DOI: 10.12988/pms.2017.6916 (cit. on p. 47).

[OSC24] The OSCAR Team. *OSCAR – Open Source Computer Algebra Research system, Version 1.0.2*. 2024. URL: https://www.oscar-system.org (cit. on p. 1).

[OSS14] G. Ottaviani, P.-J. Spaenlehauer, and B. Sturmfels. "Exact Solutions in Structured Low-Rank Approximation". In: *SIAM Journal on Matrix Analysis and Applications* 35.4 (2014), pp. 1521–1542 (cit. on pp. 168, 180, 182, 187).

[OT92] Peter Orlik and Hiroaki Terao. *Arrangements of Hyperplanes*. Springer Berlin Heidelberg, 1992. ISBN: 9783662027721. DOI: 10.1007/978-3-662-02772-1 (cit. on pp. 130, 132).

[OT95] Peter Orlik and Hiroaki Terao. "The number of critical points of a product of powers of linear functions". In: *Inventiones mathematicae* 120 (1995), pp. 1–14 (cit. on p. 131).

[Pan25] Belal Sefat Panah. "The computation of the Euclidean distance degree for the middle Catalacticant for the binary forms". In: *Communications in Algebra* 0.0 (2025), pp. 1–9. DOI: 10.1080/00927872.2025.2532175 (cit. on pp. 168, 189, 190).

[Pet09] Ivan Petrakiev. "On Self-Associated Sets of Points in Small Projective Spaces". In: *Communications in Algebra* 37.2 (2009), pp. 397–405. DOI: 10.1080/00927870802451407 (cit. on pp. 4, 70, 80, 81, 88).

[Pfl25] Nathan Pflueger. "Versality of Brill–Noether flags and degeneracy loci of twice-marked curves". In: *Algebraic Geometry* (July 2025), pp. 497–518. ISSN: 2214-2584. DOI: 10.14231/ag-2025-014 (cit. on p. 185).

[Pie88] Ragni Piene. "Cycles polaires et classes de Chern pour les variétés projectives singulières. (On polar cycles and Chern classes of singular projective varieties)". In: *Travaux en Cours.* 37 (Jan. 1988) (cit. on p. 160).

[Pro95] Gaspard de Prony. "Essai experimental et analytique sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de l'alkool à différentes températures". In: *Journal de l'École Polytechnique de Paris* 1 (1795), pp. 24–76 (cit. on p. 171).

[PS06] Nicholas Proudfoot and David Speyer. "A broken circuit ring". In: *Beiträge Algebra Geom.* 47.1 (2006), pp. 161–166. ISSN: 0138-4821 (cit. on p. 139).

[Reg18] Phillip A. Regalia. *Adaptive IIR Filtering in Signal Processing and Control*. Routledge, Apr. 2018. ISBN: 9781315136653. DOI: 10.1201/9781315136653 (cit. on pp. 168, 191).

[Rei03] Zinovy Reichstein. "SAGBI bases in rings of multiplicative invariants". In: *Commentarii Mathematici Helvetici* 78.1 (Mar. 2003), pp. 185–202. ISSN: 1420-8946. DOI: 10.1007/s000140300008 (cit. on pp. 4, 94).

[RH95] B. Roorda and C. Heij. "Global Total Least Squares Modeling of Multivariable Time Series". In: *IEEE Transactions on Automatic Control* 40.1 (Jan. 1995), pp. 50–63 (cit. on p. 171).

[Rit12] Stephan Ritscher. "Degree Bounds and Complexity of Gröbner Bases of Important Classes of Polynomial Ideals". PhD thesis. Technische Universität München, 2012 (cit. on pp. 103, 106).

[RS06] Lorenzo Robbiano and Moss Sweedler. "Subalgebra bases". In: *Commutative Algebra: Proceedings of a Workshop held in Salvador, Brazil, Aug. 8–17, 1988*. Springer. 2006, pp. 61–87. DOI: 10.1007/BFb0085537 (cit. on p. 83).

[RS90] Lorenzo Robbiano and Moss Sweedler. "Subalgebra bases". In: *Commutative Algebra*. Ed. by Winfried Bruns and Aron Simis. Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 61–87. ISBN: 978-3-540-47136-3 (cit. on pp. 18, 113–115).

[RT15] Jose Israel Rodriguez and Xiaoxian Tang. "Data-discriminants of likelihood equations". In: *Proceedings of the 2015 ACM on international symposium on symbolic and algebraic computation*. 2015, pp. 307–314 (cit. on p. 128).

[Rus16] F. Russo. *On the Geometry of Some Special Projective Varieties*. Springer International Publishing, 2016. ISBN: 978-3-319-26765-4. DOI: 10.1007/978-3-319-26765-4 (cit. on pp. 8, 24).

[Sau85] T. Sauer. "The number of equations defining points in general position." In: *Pacific J. Math.* 120.1 (1985), pp. 199–213. DOI: 10.2140/pjm.1985.120.199 (cit. on pp. 47, 52, 62).

[Sch14] Arnold Schwarzenegger. *Arnold Schwarzenegger's post*. 2014. URL: https://www.facebook.com/arnold/posts/10152403566761760 (visited on 09/03/2025) (cit. on p. 70).

[Sch78] Thomas J. Schaefer. "The complexity of satisfiability problems". In: *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*. STOC '78. San Diego, California, USA: Association for Computing Machinery, 1978, pp. 216–226. ISBN: 9781450374378. DOI: 10.1145/800133.804350 (cit. on pp. 96, 100).

[Spe77] David Spear. "A constructive approach to commutative ring theory". In: 1977. URL: https://api.semanticscholar.org/CorpusID:117037310 (cit. on p. 106).

[SS88] David Shannon and Moss Sweedler. "Using Gröbner bases to determine algebra membership, split surjective algebra homomorphisms determine birational equivalence". In: *Journal of Symbolic Computation* 6.2 (1988), pp. 267–273. ISSN: 0747-7171. DOI: 10.1016/S0747-7171(88)80047-6 (cit. on p. 106).

[SSV13] Raman Sanyal, Bernd Sturmfels, and Cynthia Vinzant. "The entropic discriminant". In: *Advances in Mathematics* 244 (Sept. 2013), pp. 678–707. ISSN: 0001-8708. DOI: 10.1016/j.aim.2013.05.019 (cit. on p. 139).

[ST21] Bernd Sturmfels and Simon Telen. "Likelihood equations and scattering amplitudes". In: *Algebraic Statistics* 12.2 (2021), pp. 167–186 (cit. on p. 132).

[Sta25] The Stacks project authors. *The Stacks project*. https://stacks.math.columbia.edu. 2025 (cit. on pp. 17, 122, 123, 129).

[Sti+] M. Stillman, G. G. Smith, S. A. Strømme, D. Eisenbud, F. Galetto, and J. W. Skelton. *Points: sets of points. Version 3.0.* A `Macaulay2` package available at https://github.com/Macaulay2/M2/tree/master/M2/Macaulay2/packages (cit. on pp. 44, 52).

[Stu17] Bernd Sturmfels. "The Hurwitz form of a projective variety". In: *J. Symbolic Comput.* 79 (2017), pp. 186–196. ISSN: 0747-7171,1095-855X. DOI: 10.1016/j.jsc.2016.08.012 (cit. on pp. 125, 126).

[Stu96] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes.* Memoirs of the American Mathematical Society. American Mathematical Society, 1996. ISBN: 9780821882672 (cit. on p. 83).

[SV23] Anna-Laura Sattelberger and Robin van der Veer. "Maximum likelihood estimation from a tropical and a Bernstein–Sato perspective". In: *International Mathematics Research Notices* 2023.6 (2023), pp. 5263–5292 (cit. on p. 128).

[SW05] Andrew J Sommese and Charles W Wampler. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science.* World Scientific, Mar. 2005. ISBN: 9789812567727. DOI: 10.1142/5763 (cit. on pp. 71, 72, 160).

[Tel20] S. Telen. "Solving Systems of Polynomial Equations". Available at https://simontelen.webnode.page/publications/. PhD thesis. KU Leuven, Leuven, Belgium, 2020 (cit. on pp. 33, 35, 36).

[Tim21] S. Timme. "Numerical nonlinear algebra". Available at https://sascha.timme.xyz/. PhD thesis. Technische Universität Berlin (Germany), 2021 (cit. on p. 71).

[TMB18] Simon Telen, Bernard Mourrain, and Marc Van Barel. "Solving Polynomial Systems via Truncated Normal Forms". In: *SIAM Journal on Matrix Analysis and Applications* 39.3 (Jan. 2018), pp. 1421–1447. ISSN: 1095-7162. DOI: 10.1137/17m1162433 (cit. on p. 34).

[TV22] S. Telen and N. Vannieuwenhoven. "A normal form algorithm for tensor rank decomposition". In: *ACM Trans. on Math. Soft.* 48.4 (2022), pp. 1–35. DOI: 10.1145/3555369 (cit. on pp. 4, 43).

[Var95] Alexander Varchenko. "Critical points of the product of powers of linear functions and families of bases of singular vectors". In: *Compositio Mathematica* 97.3 (1995), pp. 385–401 (cit. on pp. 131, 133).

[Ver25] Christof Vermeersch. *Home | MacaulayLab — christofvermeersch.github.io.* https://macaulaylab.net. 2025 (cit. on p. 36).

[Vol99] Heribert Vollmer. *Introduction to Circuit Complexity: A Uniform Approach.* Springer Berlin Heidelberg, 1999. ISBN: 9783662039274. DOI: 10.1007/978-3-662-03927-4 (cit. on p. 101).

[Wal32] J. L. Walsh. "On interpolation and approximation by rational functions with preassigned poles". In: *Transactions of the American Mathematical Society* 34.1 (1932), pp. 22–74. ISSN: 1088-6850. DOI: 10.1090/s0002-9947-1932-1501629-1 (cit. on pp. 191, 192).

[Wal95] C. H. Walter. "The minimal free resolution of the homogeneous ideal of s general points in $\mathbb{P}^4$". In: *Math. Zeit.* 219 (1995), pp. 231–234. DOI: 10.1007/BF02572362 (cit. on p. 47).

[Wat97] Junzo Watanabe. "Hankel matrices and Hankel ideals". In: *Proceedings of the School of Science, Tokai University* 32 (1997), pp. 11–21 (cit. on p. 166).

[Wil86] Jan C. Willems. "From Time Series to Linear System - Part I. Finite Dimensional Linear Time Invariant Systems". In: *Automatica* 22.5 (1986), pp. 561–580 (cit. on p. 170).

Index

Daten zur Autorin

Name:	Elisabeth Leonie Kayser
Geburtsdatum:	20. Februar 1998
Geburtsort:	Hannover, Deutschland

Wissenschaftlicher Werdegang

01/2023-06/2026	**Doktorstudium der Mathematik** Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig
04/2021-11/2022	**M.Sc. Informatik**, Thesis: *Gröbner Bases and Their Complexity* Leibniz Universtität Hannover
10/2019-01/2022	**M.Sc. Mathematik** Thesis: *The Waring problem for polynomials: Geometry and applications* Leibniz Universtität Hannover

Selbstständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialien oder erbrachten Dienstleistungen als solche gekennzeichnet.

(Ort, Datum)	(Unterschrift)

Bibliographische Daten

Computation and Complexity in Algebraic Geometry
(Zu Berechnungen und Komplexität in der Algebraischen Geometrie)
Kayser, Elisabeth Leonie
Universität Leipzig, Dissertation, 2026
245 Seiten, 16 Abbildungen, 6 Tabellen, 162 Referenzen