Polynomials in Multiview Geometry

Christopher Aholt

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Rekha Thomas, Chair

Sameer Agarwal, Chair

Isabella Novik

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

Polynomials in Multiview Geometry

Christopher Aholt

Co-Chairs of the Supervisory Committee:
Professor Rekha Thomas
Mathematics

Doctor Sameer Agarwal
Computer Science and Engineering

We study multiview geometry and some of its applications through the use of polynomials. A three-dimensional world point gives rise to $n \geq 2$ two-dimensional projections in n given cameras. The object of focus in this thesis is the multiview variety, the space of all possible n-tuples of such projections. By applying tools and techniques from algebraic geometry, representation theory, optimization, and others, we are able to provide a more complete picture of the multiview variety than has existed before. We apply this understanding to solving triangulation, the problem of reconstructing a world point from noisy projections.

ACKNOWLEDGMENTS

I would like to thank Rekha Thomas and Sameer Agarwal for advising me in more than just my thesis and graduate studies. Thank you both for helping me accomplish my dreams.

DEDICATION

To Nathalie

TABLE OF CONTENTS

Page

Chapter 1

INTRODUCTION

Multiview geometry is the study of $n \geq 2$ two-dimensional images of a three-dimensional scene. In this thesis we focus on describing the space of all possible images, called the *multiview variety*, and on understanding its geometry. From this we tackle the fundamental problem of reconstructing 3D geometry directly from known images.

Applications are exciting and varied. One example is a movie camera moving along a path in a scene. At n different time steps, the camera has recorded n different images of the same scene from different locations. Knowledge of the 3D scene directly from these 2D images allows filmmakers to add in convincing animation and graphics to the scene after everything is already shot. Another application is in robotics: creating a virtual 3D space from multiple cameras mounted on an autonomous vehicle, for instance. A few more examples are photo tourism and immersive maps [1, 50]. An end goal of our description of the geometry of the multiview variety is an overall understanding of where the real difficulties of 3D reconstruction lie, and conversely where simplicity can be exploited.

This thesis studies the multiview variety and applications via polynomials. Our study is divided into three separately published articles, each forming the entirety of one of the remaining chapters beyond this introduction. Each chapter is committed to the analysis of a different facet of the multiview variety, altogether giving a more complete picture of the underlying algebra and geometry than has existed in previous work.

The remaining parts of this chapter are organized as follows. In Section 1.1 we describe the mathematical model of a pinhole camera. Section 1.2 provides a precise description of the multiview variety, along with an outline of the main results of each of the three remaining chapters, including details of our contributions.

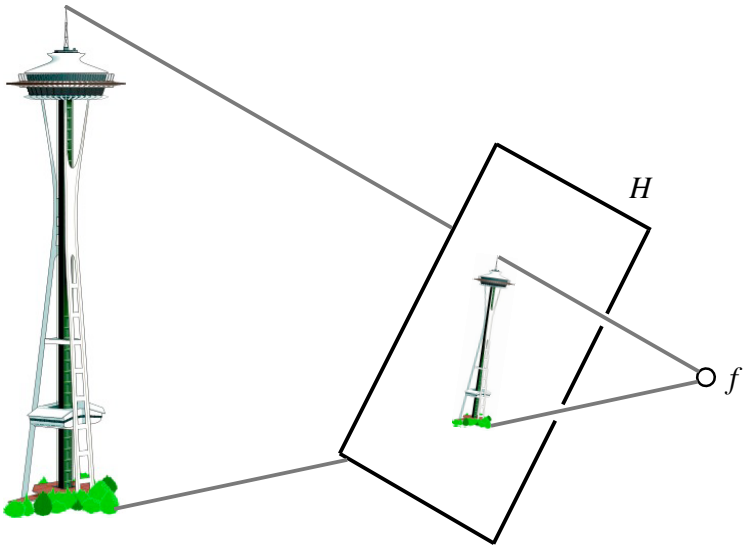


Figure 1.1: The projection of a scene under a pinhole camera with center of projection f and viewing plane H.

1.1 The Pinhole Camera

We first review the fundamentals of multiview geometry and computer vision for the unfamiliar reader. A fantastic source for more details is [22]. In the simplest form, a camera is an assignment of three-dimensional *world points* to two-dimensional *image points*. Let us explore the specifics of the *pinhole camera model*.

A pinhole camera consists of two pieces of information: a *center of projection* $f \in \mathbb{R}^3$, and a *viewing plane* $H \subseteq \mathbb{R}^3$, where $\psi : H \xrightarrow{\simeq} \mathbb{R}^2$ is a coordinatization of H. For $X \in \mathbb{R}^3$, let $\Pi(X) \in \mathbb{R}^2$ denote the image of X in this camera. This two-dimensional image is obtained by first constructing the line ℓ between X and f. If this line ℓ intersects H at $x \in H$, then $\Pi(X) := \psi(x)$ by definition. See Figure 1.1 for the associated picture.

It is important to realize that every world point on ℓ, except the center of projection, is actually projected to the exact same image point. Also, for world points on the plane

through f parallel to H, the line ℓ will not intersect H. For all points except the center of projection f, this can be resolved by extending the image points to include points *at infinity*. In other words, we will extend the image to $\mathbb{P}^2$ and the world to $\mathbb{P}^3$. Here we are using the notation that m-dimensional projective space is

$$\mathbb{P}^m = \left(\mathbb{R}^{m+1} \setminus \{\mathbf{0}\}\right) / \sim,$$

where $v \sim w$ if and only if $v = \lambda w$ for some $0 \neq \lambda \in \mathbb{R}$.

By extending our definitions to projective space, it becomes clear that a pinhole camera is actually a fundamental concept from projective geometry. Indeed, the explicit math behind the pinhole camera becomes much easier to write down. Here, a pinhole camera is modeled as a full rank matrix $P \in \mathbb{R}^{3\times 4}$. We call any such matrix a **camera matrix**. The camera matrix determines a rational map $\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$ given by $v \mapsto Pv$. This rational map is defined except at $f \in \mathbb{P}^3$, where f is any generator for the (1-dimensional) kernel of P. We call f the **focal point** of camera P.

To see that this model behaves exactly as we expect from Figure 1.1, let's take a moment to familiarize ourselves with how a camera matrix acts on the affine patch $\mathbb{R}^3 \subseteq \mathbb{P}^3$ consisting of all vectors with last coordinate equal to 1. We will also normalize the coordinates of the projections to lie in the similarly coordinatized $\mathbb{R}^2 \subseteq \mathbb{P}^2$. For $X \in \mathbb{R}^3$, recall that we denote the image of X in $\mathbb{R}^2$ under camera P by $\Pi(X)$.

A more precise definition is in order. First, for any vector $v \in \mathbb{R}^m$, let $\widetilde{v} \in \mathbb{R}^{m+1}$ denote its *homogenization*: $\widetilde{v} = (v, 1)$. Now, we say that $x = \Pi(X) \in \mathbb{R}^2$ is the **image** or **projection** of $X \in \mathbb{R}^3$ in camera P if there exists some $0 \neq \lambda \in \mathbb{R}$ such that

$$P\widetilde{X} = \lambda\widetilde{x}. \tag{1.1}$$

The scalar λ is called the *depth* of X in camera P. A more realistic model of a pinhole camera would require $\lambda > 0$, but we ignore this semialgebraic constraint. Under this definition it can be the case that some X does not have an image in camera P. This could happen, for instance, if $\widetilde{X} \in \ker(P)$ (in which case we also call X the *affine focal point* or

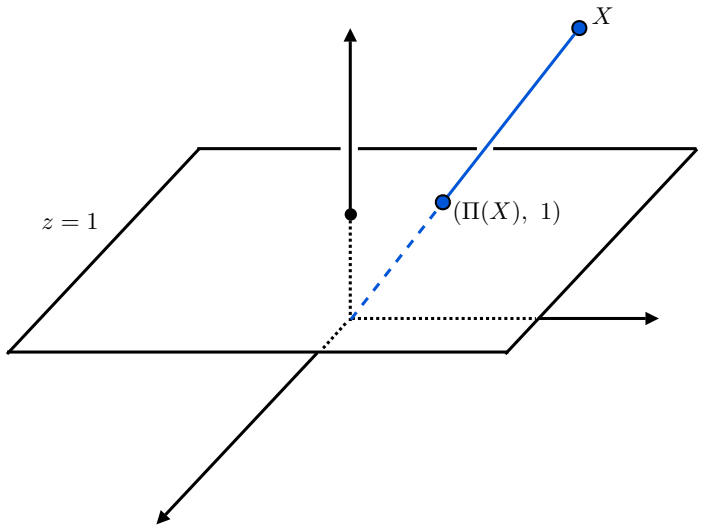


Figure 1.2: The projection of $X \in \mathbb{R}^3$ under the camera matrix $P = [I \mid \mathbf{0}]$ is given by $\Pi(X) \in \mathbb{R}^2$.

center of projection). The following two examples provide a geometric understanding of this definition.

Example 1.1 $\mathbf{P} = [\mathbf{I} \mid \mathbf{0}]$. *Following the defining equation* (1.1)*, if* $X = (x, y, z)$ *and* $z \neq 0$*, then* $\Pi(X) = (x/z,\ y/z)$*. The coordinates of* $\Pi(X)$ *are linear rational functions in the coordinates of the world point* X*. The center of projection is the origin* $\mathbf{0}$*, and the viewing plane is the plane* $z = 1$ *(See Figure 1.2). Any nonzero scalar multiple of* X *will give the exact same image in this camera.*

Points on the xy-plane do not have projections in this camera under the definition (1.1)*. Indeed, in this camera matrix* P*, such points are projected to points in* $\mathbb{P}^2$ *with last coordinate equal to* 0*, i.e. points at infinity.*

Example 1.2 $\mathbf{P} = [\mathbf{R} \mid \mathbf{t}]$. *Here* R *is an orthogonal matrix, and* $t \in \mathbb{R}^3$*. The camera* P *in this case is nothing more than a rotation and translation of* $\mathbb{R}^3$*, followed by the camera from Example 1.1 above. The coordinates of* $\Pi(X)$ *are again easily seen to be rational in* X*. The*

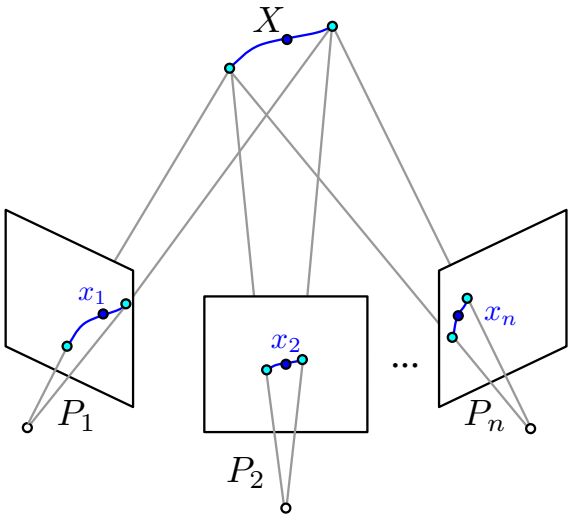


Figure 1.3: The multiview variety V is the variety described by the space of all $(x_1, \ldots, x_n) \in (\mathbb{P}^2)^n$ such that each x_i is the image of the same world point $X \in \mathbb{P}^3$.

picture here is exactly the same as Figure 1.2, but with a rotated and translated plane and center of projection. The center of projection for this camera is $-R^\top t$.

1.2 Results and Contributions

The main object of study in this thesis is the space of all possible image projections of a scene under $n \geq 2$ given camera matrices $P_1, \ldots, P_n \in \mathbb{R}^{3\times 4}$. More explicitly, let $V \subseteq (\mathbb{P}^2)^n$ be the variety defined by the image of the rational map $\mathbb{P}^3 \dashrightarrow (\mathbb{P}^2)^n$ given by

$$v \mapsto (P_1 v, \ldots, P_n v). \tag{1.2}$$

We call this variety the **multiview variety** associated with the given tuple of camera matrices. Figure 1.3 depicts V as the space cut out in the n-fold product of two-dimensional image planes as X moves freely in $\mathbb{P}^3$.

1.2.1 "A Hilbert Scheme in Computer Vision"

The content of Chapter 2 is taken from a paper coauthored with Bernd Sturmfels and Rekha Thomas, which is to appear in the *Canadian Journal of Mathematics* [4]. This chapter studies the algebraic geometry underlying the multiview variety $V \subseteq (\mathbb{P}^2)^n$ defined by the rational map (1.2). The ideal of polynomials which vanish on this irreducible variety in $(\mathbb{P}^2)^n$ is called the *multiview ideal* associated with the given n-tuple of cameras. Certain determinantal polynomials were previously known in the computer vision community to generate the variety for a generic tuple of cameras [24]. We extend this result by determining a universal Gröbner basis of a generic multiview ideal consisting of $\binom{n}{2}$ quadratic, $3\binom{n}{3}$ cubic, and $\binom{n}{4}$ quartic determinantal polynomials.

From this Gröbner basis we discover the generic multigraded initial ideal of any multiview ideal, from which we can successfully read off each ideal's multigraded Hilbert function. The multigraded Hilbert function is a way of encapsulating combinatorial data of a given multiview ideal into a single object. Surprisingly, we show that as long as the cameras have distinct focal points, all multiview ideals have the same multigraded Hilbert function.

Armed with knowledge of the multigraded Hilbert function, we embark on a combinatorial study of all multiview ideals. For $n = 3$ and 4, we determine specific cameras which yield multiview ideals which are toric. Study of these toric ideals and the corresponding initial ideals through symbolic computational algebra allows us to completely classify all mixed subdivisions of a truncated tetrahedron (a tetrahedron with its corners sliced off). We identify all regular and non-regular triangulations, placing them into 55 symmetry classes.

We also construct a multigraded Hilbert scheme $\mathcal{H}_n$ which facilitates a rigorous study of the space of all multiview ideals. General multigraded Hilbert schemes are well-studied tools in algebraic geometry for parameterizing families of ideals with the same multigraded Hilbert function [19]. All multiview ideals live on $\mathcal{H}_n$, but there are many more ideals on $\mathcal{H}_n$ which do not arise as the multiview ideal for any tuple of cameras. We show that $\mathcal{H}_n$ is connected, and that all ideals on it are radical and Cohen-Macaulay.

The main theorem of Chapter 2 is that for $n \geq 3$, the collection of all multiview ideals fill

out a full component of dimension $11n - 15$ inside of $\mathcal{H}_n$. This is a surprising result, saying that these ideals which come from real-world applications in computer vision somehow sit quite nicely inside of the multigraded Hilbert scheme. The proof is technical and non-trivial: it requires the discovery of a sequence of degenerations of multiview ideals arising from collinear cameras, followed by the identification of a basis for the tangent space of a particular monomial ideal on $\mathcal{H}_n$.

1.2.2 "The Ideal of the Trifocal Variety"

The material in Chapter 3 comes from a paper with the same title, written with Luke Oeding. This paper is currently under review for *Mathematics of Computation* [3]. This chapter extends Chapter 2 in the case of $n = 3$ cameras by taking a focused look at a certain class of polynomials from the latter. Specifically, we look at the cubics from the universal Gröbner basis of the multiview ideal.

These multihomogeneous cubics can be interpreted as tensors in $\mathbb{P}(\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3) = \mathbb{P}^{26}$, called *trifocal tensors*. Notice that this chapter uses $\mathbb{C}$ instead of $\mathbb{R}$ as the underlying field. For the geometric picture of this tensor, suppose the three cameras have affine centers of projection f_1, f_2, f_3 and viewing planes $H_1, H_2, H_3 \simeq \mathbb{C}^2$. The trifocal tensor can be viewed as a bilinear map $\mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}^3$, and we will view it as a map $\mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2$, where the ith copy of $\mathbb{P}^2$ is the set of all lines in the viewing plane H_i.

Indeed, consider lines $\ell_i \subseteq H_i$ for $i = 1, 2$. The planes $\langle f_1, \ell_1 \rangle$ and $\langle f_2, \ell_2 \rangle$ generically intersect in a line $\ell_{1,2} \subseteq \mathbb{P}^3$. Moreover, the plane $\langle f_3, \ell_{1,2} \rangle$ generically intersects the viewing plane H_3 in a line ℓ_3. The desired map $\mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2$ is given by $(\ell_1, \ell_2) \mapsto \ell_3$. Figure 1.4 depicts this map.

The collection of all trifocal tensors forms the *trifocal variety* in $\mathbb{P}^{26}$, and the collection of all homogeneous polynomials which vanish on this variety is the prime *trifocal ideal*. We set out on a computational crusade to determine the minimal generators of the trifocal ideal. The trifocal variety and the trifocal ideal are invariant with respect to a certain $G = \mathrm{SL}(3)^{\times 3}$ action. In fact, the variety is the orbit of a single tensor under this group

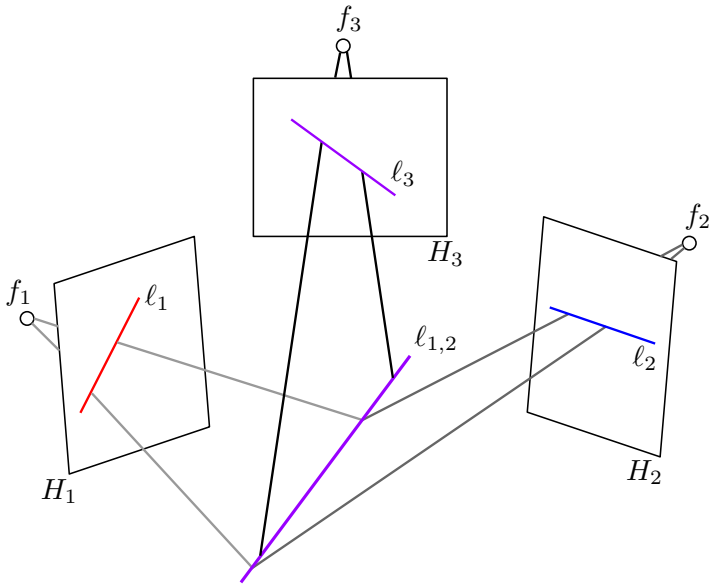


Figure 1.4: The trifocal tensor as a map $\mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2$.

action. This observation has been exploited previously by [5], and we use it as the workhorse behind an algorithm which builds the trifocal ideal from the ground up. By breaking up the polynomial ring over $\mathbb{P}^{26}$ into irreducible pieces under this G-action, we are able to check each piece for membership in the ideal via representation theory. The machinery behind this technique was introduced in [32].

The main theorem of Chapter 3 is that the trifocal ideal is minimally generated by 10 cubics, 81 qunitics, and 1980 polynomials of degree six. We first find all these generators through representation theory. Then we use numerical algebraic geometry and symbolic computational algebra to help us know that we have found the whole ideal.

The proof of the main result is in many steps. First we find the $10+81+1980$ generators using the representation theory techniques described above. Then the numerical algebraic geometry package `Bertini` [7] tells us a numerical primary decomposition of the smaller ideal generated by the 10 cubic generators. Previous work classifies all orbits of $\mathbb{P}^{26}$ under

the G-action [44]. This classification and the `Bertini` calculation allows us to determine the numerical degree of the trifocal variety. We confirm that this calculation agrees with the degree of the ideal generated by the calculated polynomial generators through degree 9 by using the symbolic computational algebra package `Macaulay2` [17]. An exercise in commutative algebra allows us to then finish the proof of the main theorem.

1.2.3 "A QCQP Approach to Triangulation"

Chapter 4 is coauthored with Sameer Agarwal and Rekha Thomas and was accepted as an oral presentation at the 2012 European Conference in Computer Vision [2]. The problem of triangulating/reconstructing a world point from its n images is the application which first started our study of the multiview variety V. In this chapter, we apply our understanding of the multiview variety to an algorithm for solving the triangulation problem to optimality.

Given only the projection of a world point in a single camera, it is impossible to determine which world point gave the projection. Indeed, there is a single dimension's worth of possible pre-images X, as in Figure 1.2. With two or more cameras, however, it is possible to resolve this ambiguity. In this case, it is a simple exercise in linear algebra to reconstruct the correct world point. Figure 1.3 depicts how this is nothing more than finding the intersection of n lines which have been *back-projected* from the known images and centers of projection. Notice that with $n \geq 3$ cameras, the resulting system of equations is overconstrained.

The problem of reconstructing the three-dimensional X from noisy two-dimensional observations is called **triangulation**. In any application, the triangulation problem is not as simple as the noiseless situation outlined above. With noisy projections in Figure 1.3, the corresponding lines would not intersect, and the choice of "correct" world point becomes a much more interesting question.

Let's take a moment to state the triangulation problem more precisely. We assume given n camera matrices $P_1, P_2, \ldots, P_n \in \mathbb{R}^{3\times 4}$ and n noisy observations $\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n \in \mathbb{R}^2$ of some $X \in \mathbb{R}^3$. In other words, we know that for each $i = 1, 2, \ldots, n$,

$$\widehat{x}_i = \Pi_i(X) + \eta_i,$$

where $\Pi_i(X)$ is the true projection of X in camera i, and η_i is some two-dimensional noise. We assume $\eta_i \sim \mathcal{N}(0, \sigma I)$, so the triangulation problem is to find the maximum likelihood estimate of X:

$$\arg\min_X \sum_i^n \|\Pi_i(X) - \widehat{x}_i\|^2. \tag{1.3}$$

This is an unconstrained rational optimization problem over $\mathbb{R}^3$. A non-linear least squares optimization algorithm such as Levenberg-Marquardt [42] can be applied to this problem in hopes of finding an optimum. Unfortunately, such algorithms only search for local optima, and there is no guarantee on the quality of the returned solution.

We take a novel viewpoint to solving the triangulation problem (1.3) by first recognizing triangulation as a constrained polynomial optimization problem over $(\mathbb{R}^2)^n = \mathbb{R}^{2n}$. The constraints are affine versions of the $\binom{n}{2}$ quadratic and $\binom{n}{3}$ cubic polynomials from the Gröbner basis described in Chapter 2. The objective function in this reformulation is the convex quadratic function $\|x - \widehat{x}\|^2$, where $\widehat{x} = (\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_n)$ is an n-tuple of the noisy two-dimensional observations. We then relax this polynomial formulation of triangulation by removing the cubic constraints, giving us a quadratically-constrained quadratic program (QCQP) with equality constraints. We know from previous work that this QCQP is equivalent to the triangulation problem as long as all the camera centers of projection are not coplanar [24].

Solving QCQPs in general is a difficult task, so we apply a standard relaxation to the triangulation QCQP to obtain a semidefinite program (SDP) which can be set up and solved in time polynomial in n, the number of cameras. This SDP relaxation is nothing more than the first relaxation in a much more general hierarchy of relaxations that can be applied to any polynomial optimization problem [34]. The first relaxation is the Lagrangian dual of the QCQP. In practice, we had observed that solving just the first relaxation in this hierarchy would often seem to solve the original triangulation problem for generic synthetic instances. In Chapter 4 our theorems explain this phenomenon, and we justify the theory with numerical experiments.

In our main theorem, we provide necessary and sufficient conditions for the optimal

value of the QCQP to be equal to the optimal value of the SDP in the case of triangulation. Moreover, there is a standard way of extracting a candidate solution for any QCQP from an optimal solution to its SDP relaxation. We also determine an easily checkable sufficient condition for this extracted solution to be equal to the desired solution to the triangulation problem. Both these results can be stated more generally for *any* QCQP with equality constraints, with nothing more than a change in notation.

Our algorithm for solving triangulation is to solve the relaxed SDP, extract the candidate solution, and then check using our condition whether or not we can guarantee optimality of that solution for triangulation. We provide more detailed analysis of our easily checkable sufficient condition in the case that the objective function is squared Euclidean distance, as it is in the triangulation problem. We find that when the true optimal value of the QCQP is small, the first SDP relaxation can be guaranteed to solve the QCQP. For triangulation this means that for all problem instances in which the noisy observations are all not too far perturbed from true projections of some world point, the triangulation problem can be solved in time polynomial in n by means of the first SDP relaxation.

We provide extensive experiments to confirm this theoretical result and to help quantify what we mean by *not too far perturbed*. In synthetic experiments, for small n we randomly generate world points and perturb their images in n cameras by increasing levels of noise. We apply our algorithm to solving the triangulation problem on each of the noisy observations and measure the fraction of times in which our theorem can guarantee global optimality of the candidate solution. For generic cameras, we find our algorithm to perform fantastically, even with large noise levels. For more degenerate configurations of cameras (coplanar and collinear), we see more rapid degredation with larger noise. Our algorithm also shows a very large fraction of certifiable global optima from four different real-world test cases.

Notation

A word of warning on notation: it can vary slightly between each chapter. This is partially due to the order in which the articles were written, and mostly due to the fact that each

chapter uses notation which best serves the purposes therein. One example of notation changing between chapters is using A_i for the cameras in Chapters 2 and 3, but using P_i in Chapter 4. Another example is that Chapter 2 uses the triple (x, y, z) to denote homogeneous coordinates on $\mathbb{P}^2$, whereas Chapter 4 uses x to denote a two-dimensional coordinate in $\mathbb{R}^2$. The notation is clearly defined in each chapter, so it should not cause any ambiguity or confusion.

Chapter 2

A HILBERT SCHEME IN COMPUTER VISION [4]

2.1 Introduction

Computer vision is based on mathematical foundations known as *multiview geometry* [14, 18] or *epipolar geometry* [22, §9]. In that subject one studies the space of pictures of three-dimensional objects seen from $n \geq 2$ cameras. Each camera is represented by a 3×4-matrix A_i of rank 3. The matrix specifies a linear projection from $\mathbb{P}^3$ to $\mathbb{P}^2$, which is well-defined on $\mathbb{P}^3 \backslash \{f_i\}$, where the focal point f_i is represented by a generator of the kernel of A_i.

The space of pictures from the n cameras is the image of the rational map

$$\phi_A \,:\, \mathbb{P}^3 \;\dashrightarrow\; (\mathbb{P}^2)^n, \quad \mathbf{x} \mapsto (A_1\mathbf{x}, A_2\mathbf{x}, \ldots, A_n\mathbf{x}). \tag{2.1}$$

The closure of this image is an algebraic variety, denoted V_A and called the *multiview variety* of the given n-tuple of 3×4-matrices $A = (A_1, A_2, \ldots, A_n)$. In geometric language, the multiview variety V_A is the blow-up of $\mathbb{P}^3$ at the cameras $f_1, \ldots, f_n$, and we study this threefold as a subvariety of $(\mathbb{P}^2)^n$.

The *multiview ideal* J_A is the prime ideal of all polynomials that vanish on the multiview variety V_A. It lives in a polynomial ring $K[x, y, z]$ in $3n$ unknowns (x_i, y_i, z_i), $i = 1, 2, \ldots, n$, that serve as coordinates on $(\mathbb{P}^2)^n$. In Section 2 we give a determinantal representation of J_A for generic A, and identify a universal Gröbner basis consisting of multilinear polynomials of degree 2, 3 and 4. This extends previous results of Heyden and Åström [24].

The multiview ideal J_A has a distinguished initial monomial ideal M_n that is independent of A, provided the configuration A is generic. Section 3 gives an explicit description of M_n and shows that it is the unique Borel-fixed ideal with its $\mathbb{Z}^n$-graded Hilbert function. Following [10], we introduce the multigraded Hilbert scheme $\mathcal{H}_n$ which parametrizes $\mathbb{Z}^n$-homogeneous ideals in $K[x, y, z]$ with the same Hilbert function as M_n. We show in Section 6

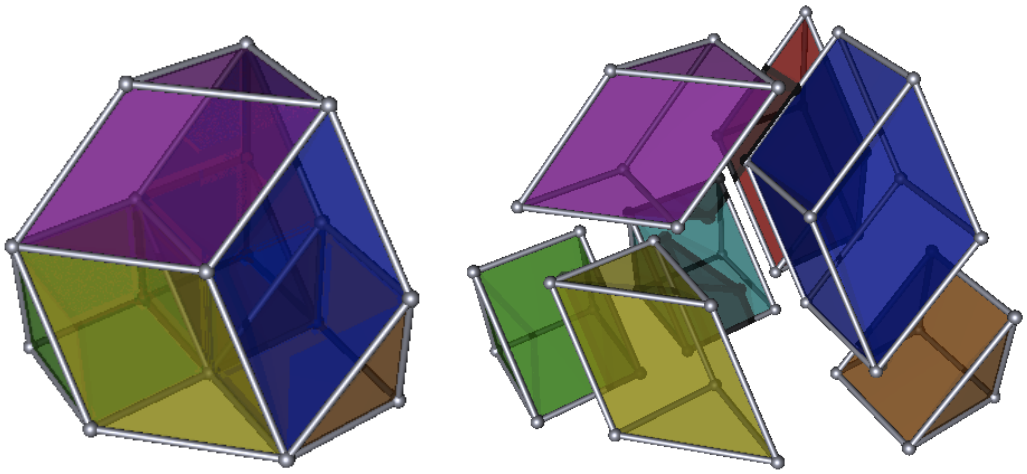


Figure 2.1: A multiview variety V_A for $n = 3$ cameras degenerates into six copies of $\mathbb{P}^1 \times \mathbb{P}^2$ and one copy of $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$.

that, for $n \geq 3$, $\mathcal{H}_n$ has a distinguished component of dimension $11n-15$ which compactifies the space of camera positions studied in computer vision. For two cameras, that space is an irreducible cubic hypersurface in $\mathcal{H}_2 \simeq \mathbb{P}^8$.

Section 4 concerns the case when $n \leq 4$ and the focal points f_i are among the coordinate points $(1{:}0{:}0{:}0), \ldots, (0{:}0{:}0{:}1)$. Here the multiview variety V_A is a toric threefold, and its degenerations are parametrized by a certain toric Hilbert scheme inside $\mathcal{H}_n$. Each initial monomial ideal of the toric ideal J_A corresponds to a three-dimensional mixed subdivision as seen in Figure 2.1. A classification of such mixed subdivisions for $n = 4$ is given in Theorem 2.19.

In Section 5 we place our n cameras on a line in $\mathbb{P}^3$. Moving them very close to each other on that line induces a two-step degeneration of the form

$$\text{trinomial ideal} \longrightarrow \text{binomial ideal} \longrightarrow \text{monomial ideal}. \tag{2.2}$$

We present an in-depth combinatorial study of this curve of multiview ideals.

In Section 6 we finally define the Hilbert scheme $\mathcal{H}_n$, and we construct the space of

camera positions as a GIT quotient of a Grassmannian. Our main result (Theorem 2.26) states that the latter is an irreducible component of $\mathcal{H}_n$. As a key step in the proof, the tangent space of $\mathcal{H}_n$ at the monomial ideal in (2.2) is computed and shown to have the correct dimension $11n - 15$. Thus, the curve (2.2) consists of smooth points on the distinguished component of $\mathcal{H}_n$. For $n \geq 3$, our Hilbert scheme has multiple components. This is seen from our classification of monomial ideals on $\mathcal{H}_3$, which relates closely to [10, §5].

The triangulation problem in computer vision is the problem of determining the point $\mathbf{x} \in \mathbb{P}^3$ as in (2.1) from a measured point $p = (p_1, p_2, \ldots, p_n)$ in the multiview variety V_A. As stated, this reconstruction is a simple exercise in linear algebra, and so the more accurate problem is to consider triangulation when the point p is a noisy measurement and hence does not lie on V_A. Choosing affine coordinates, this can be formulated as a maximum likelihood optimization problem which is constrained over the multiview variety V_A. The equations defining V_A and in particular, a degree lexicographic Gröbner basis of the multiview ideal J_A, are necessary to initiate certain convex optimization schemes to solve this maximum likelihood problem. This was one of our motivations for embarking on a thorough study of the multiview variety and its ideal. The results obtained here go well beyond this initial goal and expose the rich combinatorial, algebraic and geometric properties of these ideals and varieties that arise naturally in computer vision.

2.2 A Universal Gröbner Basis

Let K be any algebraically closed field, $n \geq 2$, and consider the map ϕ_A defined as in (2.1) by a tuple $A = (A_1, A_2, \ldots, A_n)$ of 3×4-matrices of rank 3 with entries in K. The subvariety $V_A = \overline{\text{image}(\phi_A)}$ of $(\mathbb{P}^2)^n$ is the *multiview variety*, and its ideal $J_A \subset K[x, y, z]$ is the *multiview ideal*. Note that J_A is prime because its variety V_A is the image under ϕ_A of an irreducible variety.

We say that the camera configuration A is *generic* if all 4×4-minors of the $(4 \times 3n)$-matrix $\begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}$ are non-zero. In particular, if A is generic then the focal points of the n cameras are pairwise distinct in $\mathbb{P}^3$. For any subset $\sigma = \{\sigma_1, \ldots, \sigma_s\} \subseteq [n]$ we

consider the $3s \times (s+4)$-matrix

$$A_\sigma \;:=\; \begin{bmatrix} A_{\sigma_1} & p_{\sigma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ A_{\sigma_2} & \mathbf{0} & p_{\sigma_2} & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_{\sigma_s} & \mathbf{0} & \cdots & \mathbf{0} & p_{\sigma_s} \end{bmatrix},$$

where $p_i := \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}^T$ for $i \in [n]$. Assuming $s \geq 2$, each maximal minor of A_σ is a homogeneous polynomial of degree $s = |\sigma|$ that is linear in p_i for $i \in \sigma$. Thus for $s = 2, 3, \ldots$ these polynomials are bilinear, trilinear, etc. The matrix A_σ and its maximal minors are considered frequently in multiview geometry [22, 24]. Recall that a *universal Gröbner basis* of an ideal is a subset that is a Gröbner basis of the ideal under all term orders. The following is the main result in this section.

Theorem 2.1 *If A is generic then the maximal minors of the matrices A_σ for $2 \leq |\sigma| \leq 4$ form a universal Gröbner basis of the multiview ideal J_A.*

The proof rests on a sequence of lemmas. Here is the most basic one.

Lemma 2.2 *The maximal minors of A_σ for $|\sigma| \geq 2$ lie in the prime ideal J_A.*

Proof: If $(p_1, \ldots, p_n) \in (K^3)^n$ represents a point in image(ϕ_A) then there exists a non-zero vector $q \in K^4$ and non-zero scalars $c_1, \ldots, c_n \in K$ such that $A_i q = c_i p_i$ for $i = 1, 2, \ldots, n$. This means that the columns of A_σ are linearly dependent. Since A_σ has at least as many rows as columns, the maximal minors of A_σ must vanish at every point $p \in V_A$. $\square$

Later we shall see that when A is generic, J_A has only one initial monomial ideal up to symmetry. We now identify that ideal. Let M_n denote the ideal in $K[x, y, z]$ generated by the $\binom{n}{2}$ quadrics $x_i x_j$, the $3\binom{n}{3}$ cubics $x_i y_j y_k$, and the $\binom{n}{4}$ quartics $y_i y_j y_k y_l$, where i, j, k, l runs over distinct indices in $[n]$.

We fix the lexicographic term order $\prec$ on $K[x, y, z]$ which is specified by

$$x_1 \succ \cdots \succ x_n \succ y_1 \succ \cdots \succ y_n \succ z_1 \succ \cdots \succ z_n.$$

Our goal is to prove that the initial monomial ideal $\mathrm{in}_\prec(J_A)$ is equal to M_n. We begin with the easier inclusion.

Lemma 2.3 *If A is generic then $M_n \subseteq \mathrm{in}_\prec(J_A)$.*

Proof: The generators of M_n are the quadrics $x_i x_j$, the cubics $x_i y_j y_k$, and the quartics $y_i y_j y_k y_l$. By Lemma 2.2, it suffices to show that these are the initial monomials of maximal minors of $A_{\{ij\}}$, $A_{\{ijk\}}$ and $A_{\{ijkl\}}$ respectively.

For the quadrics this is easy. The matrix $A_{\{ij\}}$ is square and we have

$$\det(A_{\{ij\}}) \;=\; \det \begin{bmatrix} A_i^1 & x_i & 0 \\ A_i^2 & y_i & 0 \\ A_i^3 & z_i & 0 \\ A_j^1 & 0 & x_j \\ A_j^2 & 0 & y_j \\ A_j^3 & 0 & z_j \end{bmatrix} \;=\; \det \begin{bmatrix} A_i^2 \\ A_i^3 \\ A_j^2 \\ A_j^3 \end{bmatrix} x_i x_j \,+\, \text{lex. lower terms.} \tag{2.3}$$

where A_t^r is the rth row of A_t. The coefficient of $x_i x_j$ is non-zero because A was assumed to be generic. For the cubics, we consider the 9×7-matrix

$$A_{\{ijk\}} \quad = \quad \begin{bmatrix} A_i & p_i & 0 & 0 \\ A_j & 0 & p_j & 0 \\ A_k & 0 & 0 & p_k \end{bmatrix}. \tag{2.4}$$

Now, $x_i y_j y_k$ is the lexicographic initial monomial of the 7×7-determinant formed by removing the fourth and seventh rows of $A_{\{ijk\}}$. Here we are using that, by genericity, the vectors $A_i^2, A_i^3, A_j^3, A_k^3$ are linearly independent.

Finally, for the quartic monomial $y_i y_j y_k y_l$ we consider the 12×8 matrix

$$A_{\{ijkl\}} \quad = \quad \begin{bmatrix} A_i & p_i & 0 & 0 & 0 \\ A_j & 0 & p_j & 0 & 0 \\ A_k & 0 & 0 & p_k & 0 \\ A_l & 0 & 0 & 0 & p_l \end{bmatrix}. \tag{2.5}$$

Removing the first row from each of the four blocks, we obtain an 8×8-matrix whose determinant has $y_i y_j y_k y_l$ as its lex. initial monomial. $\square$

The next step towards our proof of Theorem 2.1 is to express the multiview variety V_A as a projection of a diagonal embedding of $\mathbb{P}^3$. This will put us in a position to utilize the results of Cartwright and Sturmfels in [10].

We extend each camera matrix A_i to an invertible 4×4-matrix $B_i = \begin{bmatrix} b_i \\ A_i \end{bmatrix}$ by adding a row b_i at the top. Our diagonal embedding of $\mathbb{P}^3$ is the map

$$\psi_B : \mathbb{P}^3 \,\rightarrow\, (\mathbb{P}^3)^n, \quad \mathbf{x} \mapsto (B_1\mathbf{x}, B_2\mathbf{x}, \ldots, B_n\mathbf{x}). \tag{2.6}$$

Let $V^B := \text{image}(\psi_B) \subset (\mathbb{P}^3)^n$ and $J^B \subset K[w,x,y,z]$ its prime ideal. Here $(w_i : x_i : y_i : z_i)$ are coordinates on the ith copy of $\mathbb{P}^3$ and (w,x,y,z) are coordinates on $(\mathbb{P}^3)^n$. The ideal J^B is generated by the 2×2-minors of the $4 \times n$ matrix

$$\begin{bmatrix} & \begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix} & \begin{bmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix} & & \begin{bmatrix} w_n \\ x_n \\ y_n \\ z_n \end{bmatrix} \\ B_1^{-1} & B_2^{-1} & \cdots & B_n^{-1} & \end{bmatrix}. \tag{2.7}$$

These equations can be seen to generate the prime ideal J^B by noting that V^B is nothing more than the image of the diagonal embedding of $\mathbb{P}^3$ in $(\mathbb{P}^3)^n$ after a linear change of coordinates. Now consider the coordinate projection

$$\pi \,:\, (\mathbb{P}^3)^n \dashrightarrow (\mathbb{P}^2)^n, \;\; (w_i : x_i : y_i : z_i) \mapsto (x_i : y_i : z_i) \text{ for } i = 1, \ldots, n.$$

The composition $\pi \circ \psi_B$ is a rational map, and it coincides with ϕ_A on its domain of definition $\mathbb{P}^3 \backslash \{f_1, \ldots, f_n\}$. Therefore, $V_A = \overline{\pi(V^B)}$ and

$$J_A \;=\; J^B \cap K[x,y,z]. \tag{2.8}$$

The polynomial ring $K[w,x,y,z]$ admits the natural $\mathbb{Z}^n$-grading $\deg(w_i) = \deg(x_i) = \deg(y_i) = \deg(z_i) = e_i$ where e_i is the standard unit vector in $\mathbb{R}^n$. Under this grading,

$K[w,x,y,z]/J^B$ has the multigraded Hilbert function

$$\mathbb{N}^n \to \mathbb{N},\ \ (u_1,\ldots,u_n) \mapsto \begin{pmatrix} u_1+\cdots+u_n+3 \\ 3 \end{pmatrix}.$$

The multigraded Hilbert scheme $H_{4,n}$ which parametrizes $\mathbb{Z}^n$-homogeneous ideals in $K[w,x,y,z]$ with that Hilbert function was studied in [10]. More generally, the multigraded Hilbert scheme $H_{d,n}$ represents degenerations of the diagonal $\mathbb{P}^{d-1}$ in $(\mathbb{P}^{d-1})^n$ for any d and n. For the general definition of multigraded Hilbert schemes see [19]. It was shown in [10] that $H_{d,n}$ has a unique Borel-fixed ideal $Z_{d,n}$. Here *Borel-fixed* means that $Z_{d,n}$ is stable under the action of $\mathcal{B}^n$ where $\mathcal{B}$ is the group of lower triangular matrices in $\mathrm{PGL}(d,K)$. Here is what we shall need about the monomial ideal $Z_{4,n}$.

Lemma 2.4 (Cartwright-Sturmfels [10, §2] and Conca [11, §5])

1. *The unique Borel-fixed monomial ideal $Z_{4,n}$ on $H_{4,n}$ is generated by the following monomials where i,j,k,l are distinct indices in $[n]$:*

$$w_iw_j,\ w_ix_j,\ w_iy_j,\ x_ix_j,\ \ x_iy_jy_k,\ \ y_iy_jy_ky_l.$$

2. *This ideal $Z_{4,n}$ is the lexicographic initial ideal of J^B when B is sufficiently generic. The lexicographic order here is $w \succ x \succ y \succ z$ with each block ordered lexicographically in increasing order of indices.*

Using these results, it was deduced in [10] that all ideals on $H_{4,n}$ are radical and Cohen-Macaulay, and that $H_{4,n}$ is connected. We now use this distinguished Borel-fixed ideal $Z_{4,n}$ to prove the equality in Lemma 2.3.

Lemma 2.5 *If A is generic then $M_n = \mathrm{in}_\prec(J_A)$.*

Proof: We fix the lexicographic term order $\prec$ on $K[w,x,y,z]$ and its restriction to $K[x,y,z]$. Lemma 2.4 (1) shows that $M_n = Z_{4,n} \cap K[x,y,z]$. Lemma 2.4 (2) states that

$Z_{4,n} = \mathrm{in}_\prec(J^B)$ when B is generic. The lexicographic order has the important property that it allows the operations of taking initial ideals and intersections to commute [12, Chapter 3]. Therefore,

$$\begin{aligned}\mathrm{in}_\prec(J_A) &= \mathrm{in}_\prec(J^B \cap K[x,y,z]) \\ &= \mathrm{in}_\prec(J^B) \cap K[x,y,z] \\ &= Z_{4,n} \cap K[x,y,z] \;=\; M_n.\end{aligned}$$

This identity is valid whenever the conclusion of Lemma 2.4 (2) is true. We claim that, for this to hold, the appropriate genericity notion for B is that all 4×4-minors of the $(4 \times 4n)$-matrix $\begin{bmatrix} B_1^T & B_2^T & \cdots & B_n^T \end{bmatrix}$ are non-zero. Indeed, under this hypothesis, the maximal minors of the $4s \times (s+4)$-matrix

$$B_\sigma \;:=\; \begin{bmatrix} B_{\sigma_1} & \tilde{p}_{\sigma_1} & \mathbf{0} & \cdots & \mathbf{0} \\ B_{\sigma_2} & \mathbf{0} & \tilde{p}_{\sigma_2} & \ddots & \mathbf{0} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B_{\sigma_s} & \mathbf{0} & \cdots & \mathbf{0} & \tilde{p}_{\sigma_s} \end{bmatrix}, \text{ where } \tilde{p}_i := \begin{bmatrix} w_i & x_i & y_i & z_i \end{bmatrix}^T \text{ for } i \in [n],$$

have non-vanishing leading coefficients. We see that $Z_{4,n} \subseteq \mathrm{in}_\prec(J^B)$ by reasoning akin to that in the proof of Lemma 2.3. The equality $Z_{4,n} = \mathrm{in}_\prec(J^B)$ is then immediate since $Z_{4,n}$ is the multigraded generic initial ideal of J^B; see Lemma 2.4 (2). Hence, for any generic camera positions A, we can add a row to A_i and get B_i that are "sufficiently generic" for Lemma 2.4 (2). This completes the proof. □

Proof of Theorem 2.1: Lemma 2.5 and the proof of Lemma 2.3 show that the maximal minors of the matrices A_σ for $2 \leq |\sigma| \leq 4$ are a Gröbner basis of J_A for the lexicographic term order. Each polynomial in that Gröbner basis is multilinear, thus the initial monomials remain the same for any term order satisfying $x_i \succ y_i \succ z_i$ for $i = 1, 2, \ldots, n$. So, the minors form a Gröbner basis for that term order. The set of minors is invariant under permuting $\{x_i, y_i, z_i\}$ for each i. Moreover, the genericity of A implies that every monomial which can possibly appear in the support of a minor does so. Hence, these minors form a universal Gröbner basis of J_A. □

Remark 2.6 *Computer vision experts have known for a long time that multiview varieties V_A are defined set-theoretically by the above multilinear constraints of degree at most 4. We refer to work of Heyden and Åström [24, 23]. What is new here is that these constraints define V_A in the strongest possible sense: they form a universal Gröbner basis for the prime ideal J_A.*

The n cameras are in *linearly general position* if no four focal points are coplanar and no three are collinear. While the number of multilinear polynomials in our lex Gröbner basis of J_A is $\binom{n}{2} + 3\binom{n}{3} + \binom{n}{4}$, far fewer suffice to generate the ideal J_A when A is in linearly general position.

Corollary 2.7 *If A is in linearly general position then the ideal J_A is minimally generated by $\binom{n}{2}$ bilinear and $\binom{n}{3}$ trilinear polynomials.*

Proof: This can be shown for $n \leq 4$ by a direct calculation. Alternatively, these small cases are covered by transforming to the toric ideals in Section 4. First map the focal points of the cameras to the torus fixed focal points of the toric case, followed by multiplying each A_i by a suitable $g_i \in \mathrm{PGL}(3, K)$.

Now let $n \geq 5$. For any three cameras i, j, k, the maximal minors of (2.4) are generated by only one such maximal minor modulo the three bilinear polynomials (2.3). Likewise, for any four cameras i, j, k and l, the maximal minors of (2.5) are generated by the trilinear and bilinear polynomials. This implies that the resulting $\binom{n}{2} + \binom{n}{3}$ polynomials generate J_A, and, by restricting to two or three cameras, we see that they minimally generate. □

2.3 The Generic Initial Ideal

We now focus on combinatorial properties of our special monomial ideal

$$M_n \quad = \quad \bigl\langle\, x_i x_j,\, x_i y_j y_k,\, y_i y_j y_k y_l \;:\; \forall\, i,j,k,l \in [n] \text{ distinct} \bigr\rangle.$$

We refer to M_n as the *generic initial ideal* in multiview geometry because it is the lex initial ideal of any multiview ideal J_A after a generic coordinate change via the group G^n where

$G = \mathrm{PGL}(3, K)$. Indeed, consider **any** rank 3 matrices $A_1, A_2, \ldots, A_n \in K^{3\times 4}$ with pairwise distinct kernels $K\{f_i\}$. If $g = (g_1, g_2, \ldots, g_n)$ is generic in G^n then $g \circ A$ is generic in the sense that all 4×4-minors of the matrix $\left[(g_1A_1)^T\ (g_2A_2)^T \cdots (g_nA_n)^T\right]$ are non-zero. Thus, by the results of Section 2, M_n is the initial ideal of $J_{g\circ A}$, or, using standard commutative algebra lingo, M_n is the generic initial ideal of J_A.

The careful reader will notice that the term "generic initial ideal" is often associated with the action of the full general linear group, whereas here we are using it in reference to our G^n action. For our purposes, the G^n action more naturally captures the structure of the problem at hand. The term *multigraded generic initial ideal* was used for this construction in [10, 11].

Since M_n is a squarefree monomial ideal, it is radical. Hence M_n is the intersection of its minimal primes, which are generated by subsets of the variables x_i and y_j. We begin by computing this prime decomposition.

Proposition 2.8 *The generic initial ideal M_n is the irredundant intersection of $\binom{n}{3} + 2\binom{n}{2}$ monomial primes. These are the monomial primes P_{ijk} and $Q_{ij} \subseteq K[x, y, z]$ defined below for any distinct indices $i, j, k \in [n]$:*

- *P_{ijk} is generated by $x_1, \ldots, x_n$ and all y_l with $l \notin \{i, j, k\}$,*
- *Q_{ij} is generated by all x_l for $l \neq i$ and y_l for $l \notin \{i, j\}$.*

Proof: Let L denote the intersection of all P_{ijk} and Q_{ij}. Each monomial generator of M_n lies in P_{ijk} and in Q_{ij}, so $M_n \subseteq L$. For the reverse inclusion, we will show that $V(M_n)$ is contained in $V(L) = (\cup V(P_{ijk})) \cup (\cup V(Q_{ij}))$.

Let $(\tilde{x}, \tilde{y}, \tilde{z})$ be any point in the variety $V(M_n)$. First suppose $\tilde{x}_i = 0$ for all $i \in [n]$. Since $\tilde{y}_i\tilde{y}_j\tilde{y}_k\tilde{y}_l = 0$ for distinct indices, there are at most three indices i, j, k such that $\tilde{y}_i$, $\tilde{y}_j$ and $\tilde{y}_k$ are nonzero. Hence $(\tilde{x}, \tilde{y}, \tilde{z}) \in V(P_{ijk})$.

Next suppose $\tilde{x}_i \neq 0$. The index i is unique because $x_ix_j \in M_n$ for all $j \neq i$. Since $\tilde{x}_i\tilde{y}_j\tilde{y}_k = 0$ for all $j, k \neq i$, we have $\tilde{y}_j \neq 0$ for at most one index $j \neq i$. These properties

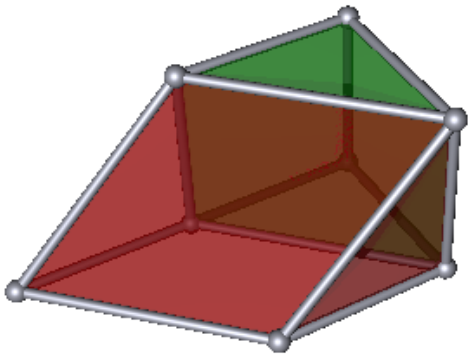


Figure 2.2: The variety of the generic initial ideal M_2 seen as two adjacent facets of the 4-dimensional polytope $\Delta_2 \times \Delta_2$.

imply $(\tilde{x}, \tilde{y}, \tilde{z}) \in V(Q_{ij})$. □

We regard the monomial variety $V(M_n)$ as a threefold inside the product of projective planes $(\mathbb{P}^2)^n$. If the focal points are distinct, V_A has a Gröbner degeneration to the reducible threefold $V(M_n)$. The irreducible components of $V(M_n)$ are

$$V(P_{ijk}) \simeq \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1 \quad \text{and} \quad V(Q_{ij}) \simeq \mathbb{P}^2 \times \mathbb{P}^1. \tag{2.9}$$

We find it convenient to regard $(\mathbb{P}^2)^n$ as a toric variety, so as to identify it with its polytope $(\Delta_2)^n$, a direct product of triangles. The components in (2.9) are 3-dimensional boundary strata of $(\mathbb{P}^2)^n$, and we identify them with faces of $(\Delta_2)^n$. The corresponding 3-dimensional polytopes are the 3-*cube* and the *triangular prism*. The following three examples illustrate this view.

Example 2.9 [Two cameras $(n = 2)$] *The variety of $M_2 = \langle x_1 \rangle \cap \langle x_2 \rangle$ is a hypersurface in $\mathbb{P}^2 \times \mathbb{P}^2$. The two components are triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$, which are glued along a common square $\mathbb{P}^1 \times \mathbb{P}^1$, as shown in Figure 2.2.* □

Example 2.10 [Three cameras $(n = 3)$] *The variety of M_3 is a threefold in $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$.*

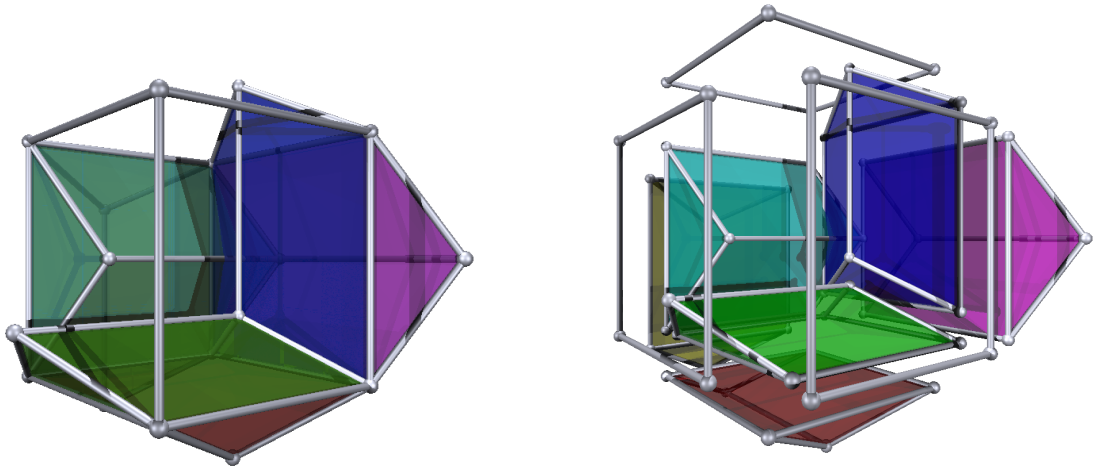


Figure 2.3: The monomial variety $V(M_3)$ as a subcomplex of $(\Delta_2)^3$.

Its seven components are given by the prime decomposition

$$\begin{aligned} M_3 \quad = \quad & \langle x_1, x_2, y_1 \rangle \cap \langle x_1, x_2, y_2 \rangle \cap \langle x_1, x_3, y_1 \rangle \\ & \cap \langle x_1, x_3, y_3 \rangle \cap \langle x_2, x_3, y_2 \rangle \cap \langle x_2, x_3, y_3 \rangle \cap \langle x_1, x_2, x_3 \rangle. \end{aligned}$$

The last component is a cube $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$, and the other six components are triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. These are glued in pairs along three of the six faces of the cube. For instance, the two triangular prisms $V(x_1, x_2, y_1)$ and $V(x_1, x_3, y_1)$ intersect the cube $V(x_1, x_2, x_3)$ in the common square face $V(x_1, x_2, x_3, y_1) \simeq \mathbb{P}^1 \times \mathbb{P}^1$. This polyhedral complex lives in the boundary of $(\Delta_2)^3$, and it shown in Figure 2.3. Compare this picture with Figure 2.1. □

Example 2.11 [Four cameras ($n = 4$)] *The variety $V(M_4)$ is a threefold in $(\mathbb{P}^2)^4$, regarded as a 3-dimensional subcomplex in the boundary of the 8-dimensional polytope $(\Delta_2)^4$. It consists of four cubes and twelve triangular prisms. The cubes share a common vertex, any two cubes intersect in a square, and each of the six squares is adjacent to two triangular prisms.* □

From the prime decomposition in Proposition 2.8 we can read off the *multidegree* [39, §8.5] of the ideal M_n. Here and in what follows, we use the natural $\mathbb{Z}^n$-grading on $K[x, y, z]$

given by $\deg(x_i) = \deg(y_i) = \deg(z_i) = e_i$. Each multiview ideal J_A is homogeneous with respect to this $\mathbb{Z}^n$-grading.

Corollary 2.12 *The multidegree of the generic initial ideal M_n is equal to*

$$\mathcal{C}\big(K[x,y,z]/M_n;\mathbf{t})\big) \;=\; t_1^2t_2^2\cdots t_n^2\cdot\left(\sum_{1\le i<j<k\le n}\frac{1}{t_it_jt_k}+\sum_{1\le i,j\le n}\frac{1}{t_i^2t_j}\right) \tag{2.10}$$

A more refined analysis also yields the Hilbert function in the $\mathbb{Z}^n$-grading.

Theorem 2.13 *The multigraded Hilbert function of $K[x,y,z]/M_n$ equals*

$$\mathbb{N}^n\,\to\,\mathbb{N},\;(u_1,\ldots,u_n)\,\mapsto\,\binom{u_1+\cdots+u_n+3}{3}-\sum_{i=1}^{n}\binom{u_i+2}{3}. \tag{2.11}$$

Proof: Fix $u\in\mathbb{N}^n$. A K-basis $\mathfrak{B}_u$ for $(K[x,y,z]/M_n)_u$ is given by all monomials $x^ay^bz^c\notin M_n$ such that $a+b+c=u$. Therefore, either (i) $a=0$ and at most three components of b are non-zero; or (ii) $a\neq 0$, in which case only one a_i can be non-zero and $b_j\neq 0$ for at most one $j\in[n]\backslash\{i\}$.

We shall count the monomials in $\mathfrak{B}_u$. Monomials of type (i) look like y^bz^c, with at most three nonzero entries in b. Also, b determines c since $c_i=u_i-b_i$ for all $i\in[n]$, and so we count the number of possibilities for y^b. There are u_i choices for $b_i\neq 0$, and thus $U:=u_1+\cdots+u_n$ many monomials in the set $\mathcal{Y}:=\{y_i^{b_i}\,:\,1\le b_i\le u_i,\,i=1,\ldots,n\}$. The factor y^b in y^bz^c is the product of 0, 1, 2 or 3 monomials from $\mathcal{Y}$ with distinct subscripts.

To resolve over-counting, consider a fixed index i. There are $\binom{u_i}{2}$ ways of choosing two monomials from $\mathcal{Y}$ with subscript i and $\binom{u_i}{3}$ ways of choosing three monomials from $\mathcal{Y}$ with subscript i. Also, there are $\binom{u_i}{2}(U-u_i)$ ways of choosing two monomials from $\mathcal{Y}$ with subscript i and a third monomial with a different subscript. Hence, the number of choices for y^b in y^bz^c is

$$\binom{U}{0}+\binom{U}{1}+\left[\binom{U}{2}-\sum_{i=1}^{n}\binom{u_i}{2}\right]+\left[\binom{U}{3}-\sum_{i=1}^{n}\binom{u_i}{3}-U\sum_{i=1}^{n}\binom{u_i}{2}+\sum_{i=1}^{n}u_i\binom{u_i}{2}\right].$$

For case (ii) we count all monomials $x^a y^b z^c \in \mathfrak{B}_u$ with $a_i \neq 0$ and all other $a_j = 0$. It suffices to count the choices for the factor $x^a y^b$. For fixed i, there are $\binom{u_i+1}{2}$ monomials of the form $x_i^{a_i} y_i^{b_i}$ with $a_i + b_i \leq u_i$ and $a_i \geq 1$. Such a monomial may be multiplied with $y_j^{b_j}$ such that $j \neq i$ and $0 \leq b_j \leq u_j$. This amounts to choosing zero or one monomial from $\mathcal{Y} \backslash \{y_i, y_i^2, \ldots, y_i^{u_i}\}$ for which there are $1 + U - u_i$ choices. Hence, there are

$$[1+U] \sum_{i=1}^{n} \binom{u_i+1}{2} - \sum_{i=1}^{n} u_i \binom{u_i+1}{2}$$

monomials in $\mathfrak{B}_u$ of type (ii). Adding the two expressions, we get

$$\begin{aligned} |\mathfrak{B}_u| &= 1 + U + \binom{U}{2} + \binom{U}{3} + (1+U) \sum_{i=1}^{n} \binom{u_i}{1} - \sum_{i=1}^{n} u_i \binom{u_i}{1} - \sum_{i=1}^{n} \binom{u_i}{3} \\ &= 1 + U + \binom{U}{2} + \binom{U}{3} + (1+U)U - \sum_{i=1}^{n} \binom{u_i+2}{3} \\ &= \binom{U+3}{3} - \sum_{i=1}^{n} \binom{u_i+2}{3}. \end{aligned}$$

□

Our analysis of M_n has the following implication for the multiview ideals J_A. Note that these are $\mathbb{Z}^n$-homogeneous for any camera configuration A.

Theorem 2.14 *For an n-tuple of camera matrices $A = (A_1, \ldots, A_n)$ with $\operatorname{rank}(A_i) = 3$ for each i, the multiview ideal J_A has the Hilbert function (2.11) if and only if the focal points of the n cameras are pairwise distinct.*

Proof: The if-direction follows from the argument in the first paragraph of this section. If the n camera positions $f_i = \ker(A_i)$ are distinct in $\mathbb{P}^3$ then M_n is the generic initial ideal of J_A, and hence both ideals have the same $\mathbb{Z}^n$-graded Hilbert function. For the only-if-direction we shall use:

$$\text{If } Q \in \mathrm{PGL}(4, K) \text{ and } AQ := (A_1 Q, \ldots, A_n Q), \text{ then } J_A = J_{AQ}. \tag{2.12}$$

This holds because Q defines an isomorphism on $\mathbb{P}^3$ and hence ϕ_A as in (2.1) has the same image in $(\mathbb{P}^2)^n$ as ϕ_{AQ}.

Suppose first that $n = 2$ and A_1 and A_2 have the same focal point and hence the same (three-dimensional) rowspace W. We can map W to the hyperplane $\{x_1 = 0\}$ by some $Q \in \mathrm{PGL}(4, K)$, and (2.12) ensures that $J_A = J_{AQ}$. Thus we may assume that $A_1 = \begin{bmatrix} \mathbf{0} & C_1 \end{bmatrix}$ and $A_2 = \begin{bmatrix} \mathbf{0} & C_2 \end{bmatrix}$ where C_1 and C_2 are invertible matrices and $\mathbf{0}$ is a column of zeros. Choosing $f_1 = f_2 = (1, 0, 0, 0)$ as the top row of B_1 and B_2 (as in Section 2), we have

$$B_1^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C_1^{-1} \end{bmatrix}, \; B_2^{-1} = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & C_2^{-1} \end{bmatrix}.$$

The ideal J^B is generated by the 2×2 minors of the matrix (2.7) which is

$$D = \begin{bmatrix} w_1 & w_2 \\ p_1(x_1, y_1, z_1) & q_1(x_2, y_2, z_2) \\ p_2(x_1, y_1, z_1) & q_2(x_2, y_2, z_2) \\ p_3(x_1, y_1, z_1) & q_3(x_2, y_2, z_2) \end{bmatrix}$$

where the p_i's and q_i's are linear polynomials. The ideal I generated by the 2×2 minors of the submatrix of D obtained by deleting the top row lies on the Hilbert scheme $H_{3,2}$ from [10] and hence $K[x, y, z]/I$ has Hilbert function

$$\mathbb{N}^2 \to \mathbb{N}, \;\; (u_1, u_2) \mapsto \begin{pmatrix} u_1 + u_2 + 2 \\ 2 \end{pmatrix}.$$

For $(u_1, u_2) = (1, 1)$, this has value 6. Since $I \subseteq J_A = J^B \cap K[x, y, z]$, the Hilbert function of $K[x, y, z]/J_A$ has value ≤ 6, while (2.11) evaluates to 8.

If $n > 2$, we may assume without loss of generality that A_1 and A_2 have the same rowspace. The argument for $n = 2$ shows that $J_A = J^B \cap K[x, y, z] \supseteq I$. The Hilbert function value of $K[x, y, z]/J_A$ in degree $e_1 + e_2$ is again 8, while the Hilbert function value of $K[x, y, z]/I$ in degree $e_1 + e_2$ coincides with the value 6 for $K[x_1, y_1, z_1, x_2, y_2, z_2]/I$. So we again conclude that $K[x, y, z]/J_A$ does not have Hilbert function (2.11). $\square$

For $G = \mathrm{PGL}(3, K)$, the product G^n acts on $K[x, y, z]$ by left-multiplication

$$(g_1, \ldots, g_n) \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \;=\; g_i \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}.$$

An ideal I in $K[x, y, z]$ is said to be *Borel-fixed* if it is fixed under the induced action of $\mathcal{B}^n$ where $\mathcal{B}$ is the subgroup of lower triangular matrices in G.

Proposition 2.15 *The generic initial ideal M_n is the unique ideal in $K[x, y, z]$ that is Borel-fixed and has the Hilbert function (2.11) in the $\mathbb{Z}^n$-grading.*

Proof: The proof is analagous to that of [10, Theorem 2.1], where $Z_{d,n}$ plays the role of M_n. The ideal M_n is Borel-fixed because it is a generic initial ideal. The same approach as in [13, §15.9.2] can be used to prove this fact. One could also use the generators of M_n to prove Borel-fixedness directly.

The multidegree of any $\mathbb{Z}^n$-graded ideal is determined by its Hilbert series [39, Claim 8.54]. Thus any ideal I with Hilbert function (2.11) has multidegree (2.10). Let I be such a Borel-fixed ideal. This is a monomial ideal.

Each maximum-dimensional associated prime P of I has multidegree either $t_1^2 t_2^2 \cdots t_n^2/(t_i t_j t_k)$ or $t_1^2 t_2^2 \cdots t_n^2/(t_i^2 t_j)$, by [39, Theorem 8.53]. In the first case P is generated by $2n-3$ indeterminates, one associated with each of the three cameras i, j, k and two each from the other $n-3$ cameras. Borel-fixedness of I tells us that the generators indexed by each camera must be the most expensive variables with respect to the order $\prec$. Hence $P = P_{ijk}$. Similarly, $P = Q_{ij}$ in the case when P has multidegree $t_1^2 t_2^2 \cdots t_n^2/(t_i^2 t_j)$.

Every prime component of M_n is among the minimal associated primes of I. This yields the containments $I \subseteq \sqrt{I} \subseteq M_n$. Since I and M_n have the same $\mathbb{Z}^n$-graded Hilbert function, the equality $I = M_n$ holds. □

The *Stanley-Reisner complex* of a squarefree monomial ideal M in a polynomial ring $K[t_1, \ldots, t_s]$ is the simplicial complex on $\{1, \ldots, s\}$ whose facets are the sets $[s] \backslash \sigma$ where

$P_\sigma := \{t_i : i \in \sigma\}$ is a minimal prime of M. A *shelling* of a simplicial complex is an ordering $F_1, F_2, \ldots, F_q$ of its facets such that, for each $1 < j \leq q$, there exists a unique minimal face of F_j (with respect to inclusion) among the faces of F_j that are not faces of some earlier facet F_i, $i < j$; see [52, Definition 2.1]. If the Stanley-Reisner complex of M is shellable, then $K[t_1, \ldots, t_s]/M$ is Cohen-Macaulay [52, Theorem 2.5].

Proposition 2.16 *The Stanley-Reisner complex of the generic initial ideal M_n is shellable. Hence the quotient ring $K[x, y, z]/M_n$ is Cohen-Macaulay.*

Proof: This proof is similar to that for $Z_{d,n}$ given in [10, Corollary 2.6]. Let Δ_n denote the Stanley-Reisner complex of the ideal M_n. By Proposition 2.8, there are two types of minimal primes for M_n, namely P_{ijk} and Q_{ij}, which we describe uniformly as follows. Let $P = (p_{ij})$ be the $3 \times n$ matrix whose ith column is $[x_i\ y_i\ z_i]^T$. For $u \in \{0, 1, 2\}^n$ define $P_u := \langle p_{ij} : i \leq u_j, 1 \leq j \leq n \rangle$. Then the minimal primes P_{ijk} of M_n are precisely the primes P_u as u varies over all vectors with three coordinates equal to one and the rest equal to two, and the minimal primes Q_{ij} are those P_u where u has one coordinate equal to zero, one coordinate equal to one and the rest equal to two. The facet of Δ_n corresponding to the minimal prime P_u is then $F_u := \{p_{ij} : u_j < i \leq 3,\ 1 \leq j \leq n\}$. We claim that the ordering of the facets F_u induced by ordering the u's lexicographically starting with $(0, 1, 2, 2, \ldots, 2)$ and ending with $(2, 2, \ldots, 2, 1, 0)$ is a shelling of Δ_n.

Consider the face $\eta_u := \{p_{ij} : j > 1, i = u_j + 1 \leq 2\}$ of the facet F_u. We will prove that η_u is the unique minimal one among the faces of F_u that have not appeared in a facet $F_{u'}$ for $u' < u$. Suppose G is a face of F_u that does not contain η_u. Pick an element $p_{u_j+1,j} \in \eta_u \backslash G$. Then $j > 1$, $u_j \leq 1$ and so if F_u is not the first facet in the ordering, then there exists $i < j$ such that $u_i > 0$ because $u > (0, 1, 2, 2, \ldots, 2)$ and of the form described above. Pick i such that $i < j$ and $u_i > 0$ and consider $F_{u+e_j-e_i} = F_u \backslash \{p_{u_j+1,j}\} \cup \{p_{u_i,i}\}$. Then $u + e_j - e_i < u$ and G is a face of $F_{u+e_j-e_i}$. Conversely, suppose G is a face of F_u that is also a face of $F_{u'}$ where $u' < u$. Since $\sum u'_j = \sum u_j$, there exists some $j > 1$ such that $u'_j > u_j$. Therefore, G does not contain $p_{u_j+1,j}$ which belongs to η_u. Therefore, η_u is not contained in G. $\square$

2.4 A Toric Perspective

In this section we examine multiview ideals J_A that are toric. For an introduction to toric ideals we refer the reader to [55]. We now assume that, for each camera i, each of the four torus fixed points in $\mathbb{P}^3$ either is the camera position or is mapped to a torus fixed point in $\mathbb{P}^2$. This implies $n \leq 4$. We fix $n = 4$ and $f_i = e_i$ for $i = 1, 2, 3, 4$. Up to permuting and rescaling columns, our assumption implies that the configuration A equals

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; A_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \; A_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

For this camera configuration, the multiview ideal J_A is indeed a toric ideal:

Proposition 2.17 *The ideal J_A is obtained by eliminating the diagonal unknowns w_1, w_2, w_3 and w_4 from the ideal of 2×2-minors of the 4×4-matrix*

$$\begin{pmatrix} w_1 & x_2 & x_3 & x_4 \\ x_1 & w_2 & y_3 & y_4 \\ y_1 & y_2 & w_3 & z_4 \\ z_1 & z_2 & z_3 & w_4 \end{pmatrix}. \tag{2.13}$$

This toric ideal is minimally generated by six quadrics and four cubics:

$$\begin{aligned} J_A \; = \; \langle & y_1y_4 - x_1z_4, y_3x_4 - x_3y_4, y_2x_4 - x_2z_4, z_1y_3 - x_1z_3, z_2x_3 - x_2z_3, z_1y_2 - y_1z_2, \\ & y_2z_3y_4 - z_2y_3z_4, \; y_1z_3x_4 - z_1x_3z_4, \; x_1z_2x_4 - z_1x_2y_4, \; x_1y_2x_3 - y_1x_2y_3 \rangle \end{aligned}$$

Proof: We extend A_i to a 4×4-matrix B_i as in Section 2 by adding the row $b_i = e_i^T$. The B_i's are then all permutation matrices, and the matrix in (2.7) equals the matrix in (2.13). The ideal J^B is generated by the 2×2 minors of that matrix of unknowns. The multiview ideal is $J_A = J^B \cap K[x, y, z]$. We find the listed binomial generators by performing the elimination with a computer algebra package such as `Macaulay2`. Toric ideals are precisely those prime ideals generated by binomials and hence J_A is a toric ideal. □

Remark 2.18 *The* normalized coordinate system in multiview geometry *proposed by Heyden and Åström [24] is different from ours and does not lead to toric varieties. Indeed, if one uses the camera matrices in [24, §2.3], then J_A is also generated by six quadrics and four cubics, but seven of the ten generators are not binomials. One of the cubic generators has six terms.* □

In commutative algebra, it is customary to represent toric ideals by integer matrices. Given $\mathcal{A} \in \mathbb{N}^{p\times q}$ with columns $a_1, \ldots, a_q$, the *toric ideal* of $\mathcal{A}$ is

$$I_{\mathcal{A}} \;:=\; \langle t^u - t^v \,:\, \mathcal{A}u \;=\; \mathcal{A}v,\, u, v \in \mathbb{N}^q \rangle \;\subset\; K[t] \,:=\, K[t_1, \ldots, t_q],$$

where t^u represents the monomial $t_1^{u_1} t_2^{u_2} \cdots t_q^{u_q}$. If $\mathcal{A}'$ is the submatrix of $\mathcal{A}$ obtained by deleting the columns indexed by $j_1, \ldots, j_s$ for some $s < q$, then the toric ideal $I_{\mathcal{A}'}$ equals the elimination ideal $I_{\mathcal{A}} \cap K[t_j \,:\, j \notin \{j_1, \ldots, j_s\}]$; see [55, Prop. 4.13 (a)]. The integer matrix $\mathcal{A}$ for our toric multiview ideal J_A in Proposition 2.17 is the following *Cayley matrix* of format 8×12:

$$\mathcal{A} \;=\; \begin{bmatrix} A_1^T & A_2^T & A_3^T & A_4^T \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}$$

where $\mathbf{1} = [1\,1\,1]$ and $\mathbf{0} = [0\,0\,0]$. This matrix $\mathcal{A}$ is obtained from the following 8×16 matrix by deleting columns $1, 6, 11$ and 16:

$$\begin{bmatrix} I_4 & I_4 & I_4 & I_4 \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \tag{2.14}$$

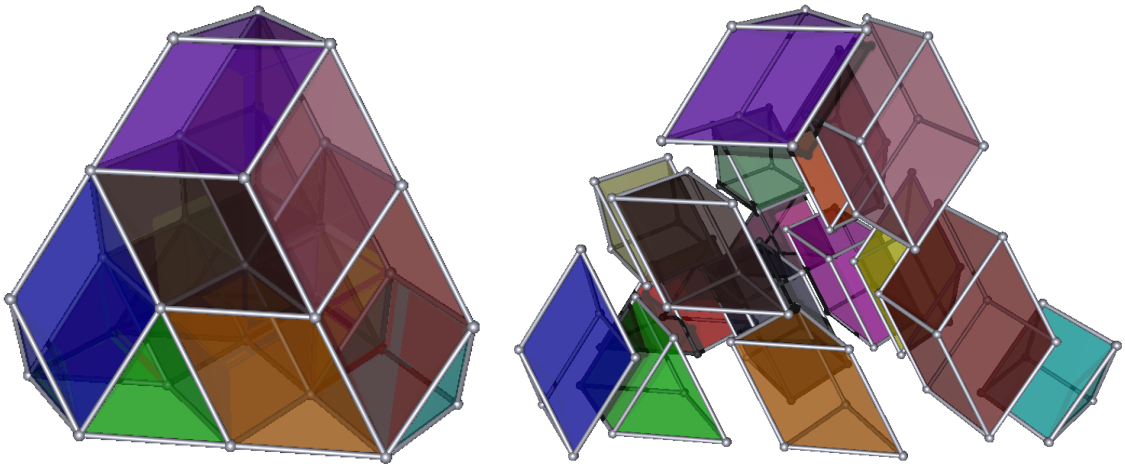


Figure 2.4: Initial monomial ideals of the toric multiview variety correspond to mixed subdivisions of the truncated tetrahedron P. These have 4 cubes and 12 triangular prisms.

The vectors $\mathbf{1}$ and $\mathbf{0}$ now have length four, I_4 is the 4×4 identity matrix and we assume that the columns of (2.14) are indexed by

$$w_1, x_1, y_1, z_1, x_2, w_2, y_2, z_2, x_3, y_3, w_3, z_3, x_4, y_4, z_4, w_4.$$

The matrix (2.14) represents the direct product of two tetrahedra, and its toric ideal is known (by [55, Prop. 5.4]) to be generated by the 2×2 minors of (2.13). Its elimination ideal in the ring $K[x, y, z]$ is $I_{\mathcal{A}}$, and hence $J_A = I_{\mathcal{A}}$.

The matrix $\mathcal{A}$ has rank 7 and its columns determine a 6-dimensional polytope $\mathrm{conv}(\mathcal{A})$ with 12 vertices. The normalized volume of $\mathrm{conv}(\mathcal{A})$ equals 16, and this is the degree of the 6-dimensional projective toric variety in $\mathbb{P}^{11}$ defined by J_A. In our context, we don't care for the 6-dimensional variety in $\mathbb{P}^{11}$ but we are interested in the threefold in $\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2$ cut out by J_A. To study this combinatorially, we apply the *Cayley trick*. This means we replace the 6-dimensional polytope $\mathrm{conv}(\mathcal{A})$ by the 3-dimensional polytope

$$P \;=\; \mathrm{conv}(A_1^T) + \mathrm{conv}(A_2^T) + \mathrm{conv}(A_3^T) + \mathrm{conv}(A_4^T).$$

This is the Minkowski sum of the four triangles that form the facets of the standard tetrahedron. Equivalently, P is the scaled tetrahedron $4\Delta_3$ with its vertices sliced off. Triangulations of $\mathcal{A}$ correspond to mixed subdivisions of P. Each 6-simplex in $\mathcal{A}$ becomes a cube or a triangular prism in P. Each mixed subdivision has four cubes $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ and twelve triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. Such a mixed subdivision of P is shown in Figure 2.4. Note the similarities and differences relative to the complex $V(M_4)$ in Example 2.11.

We worked out a complete classification of all mixed subdivisions of P:

Theorem 2.19 *The truncated tetrahedron P has* 1068 *mixed subdivisions, one for each triangulation of the Cayley polytope* $\mathrm{conv}(\mathcal{A})$. *Precisely* 1002 *of the* 1068 *triangulations are regular. The regular triangulations form* 48 *symmetry classes, and the non-regular triangulations form* 7 *symmetry classes.*

We offer a brief discussion of this result and how it was obtained. Using the software `Gfan` [26], we found that $I_{\mathcal{A}}$ has 1002 distinct monomial initial ideals. These ideals fall into 48 symmetry classes under the natural action of $(S_3)^4 \rtimes S_4$ on $K[x, y, z]$ where the i-th copy of S_3 permutes the variables x_i, y_i, z_i, and S_4 permutes the labels of the cameras. The matrix $\mathcal{A}$ being unimodular, each initial ideal of $I_{\mathcal{A}}$ is squarefree and each triangulation of $\mathcal{A}$ is unimodular. To calculate all non-regular triangulations, we used the bijection between triangulations and $\mathcal{A}$-graded monomial ideals in [55, Lemma 10.14]. Namely, we ran a second computation using the software package `CaTS` [25] that lists all $\mathcal{A}$-graded monomials ideals, and we found their number to be 1068, and hence $\mathcal{A}$ has 66 non-regular triangulations.

The 48 distinct initial monomial ideals of the toric multiview ideal J_A can be distinguished by various invariants. First, their numbers of generators range from 12 to 15. There is precisely one initial ideal with 12 generators:

$$\begin{aligned} Y_1 \;=\; \langle\, & y_1z_2, z_1y_3, x_1z_4, z_2x_3, y_2x_4, x_3y_4, \\ & x_1y_2x_3, z_1y_2x_3, x_1z_2x_4, z_1x_3z_4, z_2y_3x_4, z_2y_3z_4 \,\rangle. \end{aligned}$$

At the other extreme, there are two classes of initial ideals with 15 generators. These are the only classes having quartic generators, as all ideals with ≤ 14 generators require only

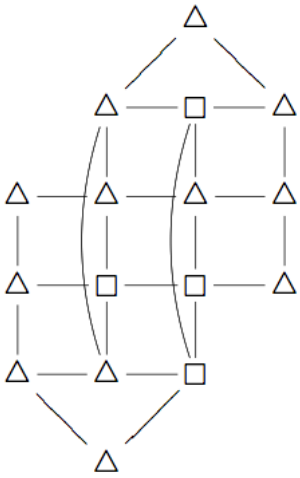


Figure 2.5: The dual graph of the mixed subdivision given by Y_1.

quadrics and cubics. A representative is

$$\begin{aligned} Y_2 \;=\; \langle\, & z_1y_2, x_1z_3, x_1z_4, x_2z_3, y_2x_4, y_3x_4,\; y_1z_2x_3y_4, \\ & x_1y_2x_3, x_1z_2x_3, x_1z_2x_4, x_4z_2y_1, y_1z_3x_4, y_1z_3y_4, y_2x_3y_4, y_2z_3y_4 \,\rangle. \end{aligned}$$

All non-regular $\mathcal{A}$-graded monomial ideal have 14 generators. One of them is

$$\begin{aligned} Y_3 \;=\; \langle\, & z_1y_2, z_1y_3, x_1z_4, x_2z_3, x_2z_4, y_3x_4,\; x_1y_2z_3, y_1x_2y_3, \\ & x_1y_2x_4, x_1z_2x_4, x_1z_3x_4, y_1z_3x_4, y_2z_3x_4, y_2z_3y_4 \,\rangle. \end{aligned}$$

A more refined combinatorial invariant of the 55 types is the dual graph of the mixed subdivision of P. The 16 vertices of this graph are labeled with squares and triangles to denote cubes and triangular prisms respectively, and edges represent common facets. The graph for Y_1 is shown in Figure 2.5.

For complete information on the classification in Theorem 2.19 see the website

`www.math.washington.edu/~aholtc/HilbertScheme`.

That website also contains the same information for the toric multiview variey in the easier case of $n = 3$ cameras. Taking A_1, A_2 and A_3 as camera matrices, the corresponding Cayley matrix has format 7×9 and rank 6:

$$\mathcal{A} \;=\; \begin{bmatrix} A_1^T & A_2^T & A_3^T \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix} \;=\; \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}$$

This is the transpose of the matrix $A_{\{123\}}$ in (2.4) when evaluated at $x_1 = y_1 = \cdots = z_3 = 1$. The corresponding 6-dimensional Cayley polytope $\mathrm{conv}(\mathcal{A})$ has 9 vertices and normalized volume 7, and the toric multiview ideal equals

$$J_A \;=\; \langle z_1y_3 - x_1z_3, z_2x_3 - x_2z_3, z_1y_2 - y_1z_2, x_1y_2x_3 - y_1x_2y_3 \rangle. \tag{2.15}$$

We note that the quadrics cut out V_A plus an extra component $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$:

$$\langle z_1y_3 - x_1z_3, z_2x_3 - x_2z_3, z_1y_2 - y_1z_2 \rangle \;=\; J_A \cap \langle z_1, z_2, z_3 \rangle \tag{2.16}$$

This equation is precisely [24, Theorem 5.6] but written in toric coordinates.

The toric ideal J_A has precisely 20 initial monomial ideals, in three symmetry classes, one for each mixed subdivision of the 3-dimensional polytope

$$P \;=\; \mathrm{conv}(A_1^T) + \mathrm{conv}(A_2^T) + \mathrm{conv}(A_3^T).$$

Thus P is the Minkowski sum of three of the four triangular facets of the regular tetrahedron. Each mixed subdivision of P uses one cube $\mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ and six triangular prisms $\mathbb{P}^2 \times \mathbb{P}^1$. A picture of one of them is seen in Figure 2.1.

Remark 2.20 *Our toric study in this section is universal in the sense that* **every** *multiview variety V_A for $n \leq 4$ cameras in linearly general position in $\mathbb{P}^3$ is isomorphic to the toric*

multiview variety under a change of coordinates in $(\mathbb{P}^2)^n$. *This fact can be proved using the coordinate systems for the Grassmannian* $\mathrm{Gr}(4, 3n)$ *furnished by the construction in [56, §4]. Here is how it works for* $n = 4$. *The coordinate change via* $\mathrm{PGL}(3, K)^4$ *gives*

$$\begin{bmatrix} A_1^T & A_2^T & A_3^T & A_4^T \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & * & * & * & * & * & * & * & * & * \\ * & * & * & 0 & 0 & 0 & * & * & * & * & * & * \\ * & * & * & * & * & * & 0 & 0 & 0 & * & * & * \\ * & * & * & * & * & * & * & * & * & 0 & 0 & 0 \end{bmatrix}, \tag{2.17}$$

where the 3×3*-matrices indicated by the stars in the four blocks are invertible. Now, the* 4×12*-matrix (2.17) gives a* support set Σ *that satisfies the conditions in [56, Proposition 3.1]. The corresponding Zariski open set* $\mathcal{U}_\Sigma$ *of the Grassmannian* $\mathrm{Gr}(4, 12)$ *is non-empty. In fact, by [56, Remark 4.9(a)], the set* $\mathcal{U}_\Sigma$ *represents configurations whose cameras* f_1, f_2, f_3, f_4 *are not coplanar. Now, Theorem 4.6 in [56] completes our proof because (the universal Gröbner basis of) the ideal* J_A *depends only on the point in* $\mathcal{U}_\Sigma \subset \mathrm{Gr}(4, 12)$ *represented by (2.17) and not on the specific camera matrices* $A_1, \ldots, A_4$. □

2.5 Degeneration of Collinear Cameras

In this section we consider a family of collinear camera positions. The degeneration of the associated multiview variety will play a key role in proving our main results in Section 6, but they may be of independent interest. Collinear cameras have been studied in computer vision, for example in [22].

Let ε be a parameter and fix the configuration $A(\varepsilon) := (A_1, \ldots, A_n)$ where

$$A_i := \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \varepsilon^{n-i} & 0 & 0 & 1 \end{bmatrix}$$

The focal point of camera i is $f_i = (-1 : 1 : 1 : \varepsilon^{n-i})$ and hence the n cameras given by $A(\varepsilon)$ are collinear in $\mathbb{P}^3$. Note that these camera matrices stand in sharp contrast to those for which A is generic which was the focus of Sections 2 and 3. They also differ from the toric situation in Section 4.

We consider the multiview ideal $J_{A(\varepsilon)}$ in the polynomial ring $K(\varepsilon)[x,y,z]$, where $K(\varepsilon)$ is the field of rational functions in ε with coefficients in K. Then $J_{A(\varepsilon)}$ has the Hilbert function (2.11), by Theorem 2.14. Let $\mathcal{G}_n$ be the set of polynomials in $K(\varepsilon)[x,y,z]$ consisting of the $\binom{n}{2}$ quadratic polynomials

$$x_i y_j - x_j y_i \qquad \text{for } \ 1 \leq i < j \leq n \tag{2.18}$$

and the $3\binom{n}{3}$ cubic polynomials below for all choices of $1 \leq i < j < k \leq n$:

$$\begin{aligned}
&(\varepsilon^{n-k} - \varepsilon^{n-i}) x_i z_j x_k + (\varepsilon^{n-j} - \varepsilon^{n-k}) z_i x_j x_k + (\varepsilon^{n-i} - \varepsilon^{n-j}) x_i x_j z_k \\
&(\varepsilon^{n-k} - \varepsilon^{n-i}) y_i z_j y_k + (\varepsilon^{n-j} - \varepsilon^{n-k}) z_i y_j y_k + (\varepsilon^{n-i} - \varepsilon^{n-j}) y_i y_j z_k \\
&(\varepsilon^{n-k} - \varepsilon^{n-i}) y_i z_j x_k + (\varepsilon^{n-j} - \varepsilon^{n-k}) z_i y_j x_k + (\varepsilon^{n-i} - \varepsilon^{n-j}) y_i x_j z_k
\end{aligned} \tag{2.19}$$

Let L_n be the ideal generated by (2.18) and the following binomials from the first two terms in (2.19):

$$L_n \,:=\, \langle x_i y_j - x_j y_i \,:\, 1 \leq i < j \leq n \rangle + \left\langle \begin{array}{l} x_i z_j x_k - z_i x_j x_k, \\ y_i z_j y_k - z_i y_j y_k, \\ y_i z_j x_k - z_i y_j x_k \end{array} \,:\, 1 \leq i < j < k \leq n \right\rangle.$$

Let N_n be the ideal generated by the leading monomials in (2.18) and (2.19):

$$N_n \,:=\, \langle x_i y_j \,:\, 1 \leq i < j \leq n \rangle \,+\, \langle x_i z_j x_k,\, y_i z_j y_k,\, y_i z_j x_k \,:\, 1 \leq i < j < k \leq n \rangle.$$

The main result in this section is the following construction of a two-step flat degeneration $J_{A(\varepsilon)} \rightarrow L_n \rightarrow N_n$. This gives an explicit realization of (2.2). We note that $V_{A(\varepsilon)}$ can be seen as a variant of the *Mustafin varieties* in [9].

Theorem 2.21 *The three ideals $J_{A(\varepsilon)}$, L_n and N_n satisfy the following:*

(a) The multiview ideal $J_{A(\varepsilon)}$ is generated by the set $\mathcal{G}_n$.

(b) The binomial ideal L_n equals the special fiber of $J_{A(\varepsilon)}$ for $\varepsilon = 0$.

(c) The monomial ideal N_n is the initial ideal of L_n, in the Gröbner basis sense, with respect to the lexicographic term order with $x \succ y \succ z$.

The rest of this section is devoted to explaining and proving these results. Let us begin by showing that $\mathcal{G}_n$ is a subset of $J_{A(\varepsilon)}$. The determinant of

$$A(\varepsilon)_{\{ij\}} \;=\; \begin{bmatrix} A_i & p_i & \mathbf{0} \\ A_j & \mathbf{0} & p_j \end{bmatrix}$$

equals $(\varepsilon^{n-j}-\varepsilon^{n-i})(x_iy_j-x_jy_i)$. Hence $J_{A(\varepsilon)}$ contains (2.18), by the argument in Lemma 2.2. Similarly, for any $1\leq i<j<k\leq n$, consider the 9×7 matrix

$$A(n)_{\{ijk\}} \;=\; \begin{bmatrix} 1 & 1 & 0 & 0 & x_i & 0 & 0 \\ 1 & 0 & 1 & 0 & y_i & 0 & 0 \\ \varepsilon^{n-i} & 0 & 0 & 1 & z_i & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & x_j & 0 \\ 1 & 0 & 1 & 0 & 0 & y_j & 0 \\ \varepsilon^{n-j} & 0 & 0 & 1 & 0 & z_j & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & x_k \\ 1 & 0 & 1 & 0 & 0 & 0 & y_k \\ \varepsilon^{n-k} & 0 & 0 & 1 & 0 & 0 & z_k \end{bmatrix}.$$

The three cubics (2.19), in this order and up to sign, are the determinants of the 7×7 submatrices of $A(\varepsilon)_{\{ijk\}}$ obtained by deleting the rows corresponding to y_j and y_k, the rows corresponding to x_j and x_k, and the rows corresponding to x_i and y_k respectively. We conclude that $\mathcal{G}_n$ lies in $J_{A(\varepsilon)}$.

We next discuss part (b) of Theorem 2.21. Every rational function $c(\varepsilon) \in K(\varepsilon)$ has a unique expansion as a Laurent series $c_1\varepsilon^{a_1} + c_2\varepsilon^{a_2} + \cdots$ where $c_i \in K$ and $a_1 < a_2 < \cdots$ are integers. The function $\mathrm{val} : K(\varepsilon) \to \mathbb{Z}$ given by $c(\varepsilon) \mapsto a_1$ is then a valuation on $K(\varepsilon)$, and $K[\![\varepsilon]\!] = \{c \in K(\varepsilon) \,:\, \mathrm{val}(c) \geq 0\}$ is its valuation ring. The unique maximal ideal in $K[\![\varepsilon]\!]$ is $m = \langle c \in K(\varepsilon) \,:\, \mathrm{val}(c) > 0\rangle$. The residue field $K[\![\varepsilon]\!]/m$ is isomorphic to K, so

there is a natural map $K[\![\varepsilon]\!] \to K$ that represents the evaluation at $\varepsilon = 0$. The *special fiber* of an ideal $I \subset K(\varepsilon)[x,y,z]$ is the image of $I \cap K[\![\varepsilon]\!][x,y,z]$ under the induced map $K[\![\varepsilon]\!][[x,y,z] \to K[x,y,z]$. The special fiber is denoted $\mathrm{in}(I)$. It can be computed from I by a variant of Gröbner bases (cf. [38, §2.4]).

What we are claiming in Theorem 2.21 (b) is the following identify

$$\mathrm{in}(J_{A(\varepsilon)}) \;=\; L_n \qquad \text{in } K[x,y,z].$$

It is easy to see that the left hand side contains the right hand side: indeed, by multiplying the trinomials in (2.19) by ε^{k-n} and then evaluating at $\varepsilon = 0$, we obtain the binomial cubics among the generators of L_n.

Finally, what is claimed in Theorem 2.21 (c) is the following identity

$$\mathrm{in}_{\prec}(L_n) \;=\; N_n \qquad \text{in } K[x,y,z].$$

Here, $\mathrm{in}_{\prec}(L_n)$ is the lexicographic initial ideal of L_n, in the usual Gröbner basis sense. Again, the left hand side contains the right hand side because the initial monomials of the binomial generators of L_n generate N_n.

Note that N_n is distinct from the generic initial ideal M_n. Even though M_n played a prominent role in Sections 2 and 3, the ideal N_n will be more useful in Section 6. The reason is that M_n is the most singular point on the Hilbert scheme $\mathcal{H}_n$ while, as we shall see, N_n is a smooth point on $\mathcal{H}_n$.

In summary, what we have shown thus far is the following inclusion:

$$N_n \;\subseteq\; \mathrm{in}_{\prec}\big(\mathrm{in}(J_{A(\varepsilon)})\big) \tag{2.20}$$

We seek to show that equality holds. Our proof rests on the following lemma.

Lemma 2.22 *The monomial ideal N_n has the $\mathbb{Z}^n$-graded Hilbert function* (2.11).

Proof: Let $u = (u_1, \ldots, u_n) \in \mathbb{N}^n$, and let $\mathfrak{B}_u$ be the set of all monomials of multidegree u in $K[x,y,z]$ which are not in N_n. We need to show that

$$|\mathfrak{B}_u| \;=\; \binom{u_1 + \cdots + u_n + 3}{3} - \sum_{i=1}^{n} \binom{u_i + 2}{3}.$$

It can be seen from the generators of N_n that the monomials in $\mathfrak{B}_u$ are of the form $z^a y^b x^c z^d$ for $a, b, c, d \in \mathbb{N}^n$ such that $u = a + b + c + d$ and

$$\begin{aligned} a &= (a_1, \ldots, a_i, 0, \ldots, 0) \\ b &= (0, \ldots, 0, b_i, \ldots, b_j, 0, \ldots, 0) \\ c &= (0, \ldots, 0, c_j, \ldots, c_k, 0, \ldots, 0) \\ d &= (0, \ldots, 0, d_k, \ldots, d_n) \end{aligned}$$

for some triple i, j, k with $1 \leq i \leq j \leq k \leq n$.

We count the monomials in $\mathfrak{B}_u$ using a combinatorial "stars and bars" argument. Each monomial can be formed in the following way. Suppose there are $u_1 + \cdots + u_n + 3$ blank spaces laid left to right. Fill exactly three spaces with bars. This leaves $u_1 + \cdots + u_n$ open blanks to fill in, which is the total degree of a monomial in $\mathfrak{B}_u$. The three bars separate the blanks into four compartments, some possibly empty. From these compartments we greedily form a, b, c, and d to make $z^a y^b x^c z^d$ as described below.

In what follows, $\star$ is used as a placeholder symbol. Fill the first u_1 blanks with the symbol $\star_1$, the next u_2 blanks with $\star_2$, and continue to fill up until the last u_n blanks are filled with $\star_n$. Now we pass once more through these symbols and replace each $\star_i$ with either x_i, y_i, or z_i such that all variables in the first compartment are z's, those in the second are y's, then x's and in the fourth compartment z's. Removing the bars gives $z^a y^b x^c z^d$ in $\mathfrak{B}_u$.

There are $\binom{u_1 + \cdots + u_n + 3}{3}$ ways of choosing the three bars. The monomials in $\mathfrak{B}_u$ are overcounted only when $i = j = k$ if z_i appears in both the first and fourth compartments. Indeed, in such cases if we require $a_i = 0$, the monomial is uniquely represented, so we are overcounting by the $\binom{u_i + 2}{3}$ choices when $a_i \neq 0$. □

We are now prepared to derive the main result of this section.

Proof of Theorem 2.21: Lemma 2.22 and Theorem 2.14 tell us that N_n and $J_{A(\varepsilon)}$ have the same $\mathbb{Z}^n$-graded Hilbert function (2.11). We also know from [38, §2.4] that $\mathrm{in}(J_{A(\varepsilon)})$ has the same Hilbert function, just as passing to an initial monomial ideal for a term order

preserves Hilbert function. Hence the equality $N_n \subseteq \text{in}_\prec\big(\text{in}(J_{A(\varepsilon)})\big)$ holds in (2.20). This proves parts (b) and (c). We have shown that $\mathcal{G}_n$ is a Gröbner basis for the homogeneous ideal $J_{A(\varepsilon)}$ in the valuative sense of [38, §2.4]. This implies that $\mathcal{G}_n$ generates $J_{A(\varepsilon)}$. □

Remark 2.23 *The polyhedral subcomplexes of $(\Delta_2)^n$ defined by the binomial ideal L_n and the monomial ideal N_n are combinatorially interesting. For instance, L_n has prime decomposition $I_3 \cap I_4 \cap \cdots \cap I_n \cap I_{n+1}$, where*

$$\begin{aligned} I_t \;:=\; & \langle\, x_i, y_i : i = t, t+1, \ldots, n \,\rangle \,+ \\ & \langle\, x_i y_j - x_j y_i : 1 \leq i < j < t \,\rangle \,+ \\ & \langle\, x_i z_j - x_j z_i,\, y_i z_j - y_j z_i : 1 \leq i < j < t-1 \,\rangle. \end{aligned}$$

The monomial ideal N_n is the intersection of $\text{in}_\prec(I_t)$ for $t = 3, \ldots, n+1$. □

2.6 The Hilbert Scheme

We define $\mathcal{H}_n$ to be the multigraded Hilbert scheme which parametrizes all $\mathbb{Z}^n$-homogeneous ideals in $K[x, y, z]$ with the Hilbert function in (2.11). According to the general construction given in [19], $\mathcal{H}_n$ is a projective scheme. The ideals J_A and $\text{in}_\prec(J_A)$ for n distinct camera positions, as well as the combinatorial ideals M_n, L_n and N_n all correspond to closed points on $\mathcal{H}_n$.

Our Hilbert scheme $\mathcal{H}_n$ is closely related to the Hilbert scheme $H_{4,n}$ which was studied in [10]. We already utilized results from that paper in our proof of Theorem 2.1. Note that $H_{4,n}$ parametrizes degenerations of the diagonal $\mathbb{P}^3$ in $(\mathbb{P}^3)^n$ while $\mathcal{H}_n$ parametrizes blown-up images of that $\mathbb{P}^3$ in $(\mathbb{P}^2)^n$.

Let $G = \text{PGL}(3, K)$ and $\mathcal{B} \subset G$ the Borel subgroup of lower-triangular 3×3 matrices modulo scaling. The group G^n acts on $K[x, y, z]$ and this induces an action on the Hilbert scheme $\mathcal{H}_n$. Our results concerning the ideal M_n in Section 3 imply the following corollary, which summarizes the statements analogous to Theorem 2.1 and Corollaries 2.4 and 2.6 in [10].

Corollary 2.24 *The multigraded Hilbert scheme $\mathcal{H}_n$ is connected. The point representing the generic initial ideal M_n lies on each irreducible component of $\mathcal{H}_n$. All ideals that lie on $\mathcal{H}_n$ are radical and Cohen-Macaulay.*

In particular, every monomial ideal in $\mathcal{H}_n$ is squarefree and can hence be identified with its variety in $(\mathbb{P}^2)^n$, or, equivalently, with a subcomplex in the product of triangles $(\Delta_2)^n$. One of the first questions one asks about any multigraded Hilbert scheme, including $\mathcal{H}_n$, is to list its monomial ideals.

This task is easy for the first case, $n = 2$. The Hilbert scheme $\mathcal{H}_2$ parametrizes $\mathbb{Z}^2$-homogeneous ideals in $K[x, y, z]$ having Hilbert function

$$h_2 : \mathbb{N}^2 \to \mathbb{N},\ (u_1, u_2) \mapsto \binom{u_1 + u_2 + 3}{3} - \binom{u_1 + 2}{3} - \binom{u_2 + 2}{3}.$$

There are exactly nine monomial ideals on $\mathcal{H}_2$, namely

$$\langle x_1x_2\rangle,\ \langle x_1y_2\rangle,\ \langle x_1z_2\rangle,\ \langle y_1x_2\rangle,\ \langle y_1y_2\rangle,\ \langle y_1z_2\rangle,\ \langle z_1x_2\rangle,\ \langle z_1y_2\rangle,\ \langle z_1z_2\rangle.$$

In fact, the ideals on $\mathcal{H}_2$ are precisely the principal ideals generated by bilinear forms, and $\mathcal{H}_2$ is isomorphic to an 8-dimensional projective space

$$\mathcal{H}_2 \ = \ \{\langle c_0x_1x_2 + c_1x_1y_2 + \cdots + c_8z_1z_2\rangle \ : (c_0 : c_1 : \cdots : c_8) \in \mathbb{P}^8\}.$$

The principal ideals J_A which actually arise from two cameras form a cubic hypersurface in this $\mathcal{H}_2 \simeq \mathbb{P}^8$. To see this, we write A_i^j for the j-th row of the i-th camera matrix and $[A_{i_1}^{j_1}A_{i_2}^{j_2}A_{i_3}^{j_3}A_{i_4}^{j_4}]$ for the 4×4-determinant formed by four such row vectors. The bilinear form can be written as

$$\mathbf{x}_2^T F \mathbf{x}_1 \ = \ \begin{bmatrix} x_2 & y_2 & z_2 \end{bmatrix} \begin{bmatrix} c_0 & c_3 & c_6 \\ c_1 & c_4 & c_7 \\ c_2 & c_5 & c_8 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix},$$

where F is the *fundamental matrix* [22]. In terms of the camera matrices,

$$F \ = \ \begin{bmatrix} [A_1^2A_1^3A_2^2A_2^3] & -[A_1^1A_1^3A_2^2A_2^3] & [A_1^1A_1^2A_2^2A_2^3] \\ -[A_1^2A_1^3A_2^1A_2^3] & [A_1^1A_1^3A_2^1A_2^3] & -[A_1^1A_1^2A_2^1A_2^3] \\ [A_1^2A_1^3A_2^1A_2^2] & -[A_1^1A_1^3A_2^1A_2^2] & [A_1^1A_1^2A_2^1A_2^2] \end{bmatrix}. \tag{2.21}$$

This matrix has rank ≤ 2, and every 3×3-matrix of rank ≤ 2 can be written in this form for suitable camera matrices A_1 and A_2 of size 3×4.

The formula in (2.21) defines a map $(A_1, A_2) \mapsto F$ from pairs of camera matrices with distinct focal points into the Hilbert scheme $\mathcal{H}_2$. The closure of its image is a compactification of the space of camera positions. We now precisely define the corresponding map for arbitrary $n \geq 2$. The construction is inspired by the construction due to Thaddeus discussed in [10, Example 7].

Let $\mathrm{Gr}(4, 3n)$ denote the Grassmannian of 4-dimensional linear subspaces of K^{3n}. The n-dimensional algebraic torus $(K^*)^n$ acts on this Grassmannian by scaling the coordinates on K^{3n}, where the ith factor K^* scales the coordinates indexed by $3i-2, 3i-1$ and $3i$. Thus, if we represent each point in $\mathrm{Gr}(4, 3n)$ as the row space of a $(4 \times 3n)$-matrix $\begin{bmatrix} A_1^T & A_2^T & \cdots & A_n^T \end{bmatrix}$, then $\lambda = (\lambda_1, \ldots, \lambda_n) \in (K^*)^n$ sends this matrix to $\begin{bmatrix} \lambda_1 A_1^T & \lambda_2 A_2^T & \cdots & \lambda_n A_n^T \end{bmatrix}$. The multi-view ideal J_A is invariant under this action by $(K^*)^n$. In symbols, $J_{\lambda \circ A} = J_A$. In the next lemma, GIT stands for *geometric invariant theory*.

Lemma 2.25 *The assignment $A \mapsto J_A$ defines an injective rational map γ from a GIT quotient $\mathrm{Gr}(4, 3n)//(K^*)^n$ to the multigraded Hilbert scheme $\mathcal{H}_n$.*

Proof: For the proof it suffices to check that $J_A \neq J_{A'}$ whenever A and A' are generic camera configurations that are not in the same $(K^*)^n$-orbit. □

We call γ the camera map. Since we need γ only as a rational map, the choice of linearization does not matter when we form the GIT quotient. The closure of its image in $\mathcal{H}_n$ is well-defined and independent of that choice of linearization. We define the *compactified camera space*, for n cameras, to be

$$\Gamma_n \;:=\; \overline{\gamma(\mathrm{Gr}(4, 3n)//(K^*)^n)} \;\subseteq\; \mathcal{H}_n.$$

The projective variety Γ_n is a natural compactification of the parameter space studied by Heyden in [23]. Since the torus $(K^*)^n$ acts on $\mathrm{Gr}(4, 3n)$ with a one-dimensional stabilizer,

Lemma 2.25 implies that the compactified space of n cameras has the dimension we expect from [23], namely,

$$\dim(\Gamma_n) \;=\; \dim(\mathrm{Gr}(4,3n)) - (n-1) \;=\; 4(3n-4) - (n-1) \;=\; 11n - 15.$$

We regard the following theorem as the main result in this paper.

Theorem 2.26 *For $n \geq 3$, the compactified camera space Γ_n appears as a distinguished irreducible component in the multigraded Hilbert scheme $\mathcal{H}_n$.*

Note that the same statement if false for $n = 2$: Γ_2 is not a component of $\mathcal{H}_3 \simeq \mathbb{P}^8$. It is the hypersurface consisting of the fundamental matrices (2.21).

Proof: By definition, the compactified camera space Γ_n is a closed subscheme of $\mathcal{H}_n$. The discussion above shows that the dimension of any irreducible component of $\mathcal{H}_n$ that contains Γ_n is no smaller than $11n-15$. We shall now prove the same $11n-15$ as an upper bound for the dimension. This is done by exhibiting a point in Γ_n whose tangent space in the Hilbert scheme $\mathcal{H}_n$ has dimension $11n-15$. This will imply the assertion.

For any ideal $I \in \mathcal{H}_n$, the tangent space to the Hilbert scheme $\mathcal{H}_n$ at I is the space of $K[x,y,z]$-module homomorphisms $I \to K[x,y,z]/I$ of degree $\mathbf{0}$. In symbols, this space is $\mathrm{Hom}(I, K[x,y,z]/I)_{\mathbf{0}}$. The K-dimension of the tangent space provides an upper bound for the dimension of any component on which I lies. It remains to specifically identify a point on Γ_n that is smooth on $\mathcal{H}_n$, an ideal which has tangent space dimension exactly $11n-15$.

It turns out that the monomial ideal N_n described in the previous section has this desired property. Lemmas 2.27 and 2.28 below give the details. □

Lemma 2.27 *The ideals L_n and N_n from the previous section lie in Γ_n.*

Proof: The image of γ in $\mathcal{H}_n$ consists of all multiview ideals J_A, where A runs over configurations of n distinct cameras, by Theorem 2.14. Let $A(\varepsilon)$ denote the collinear configuration in Section 5, and consider any specialization of ε to a non-zero scalar in K. The resulting ideal $J_{A(\varepsilon)}$ is a K-valued point of Γ_n, for any $\varepsilon \in K \backslash \{0\}$. The special fiber $J_{A(0)} = L_n$ is in

the Zariski closure of these points, because, locally, any regular function vanishing on the coordinates of $J_{A(\varepsilon)}$ for all $\varepsilon \neq 0$ will vanish for $\varepsilon = 0$. We conclude that L_n is a K-valued point in the projective variety Γ_n. Likewise, since $N_n = \text{in}_{\prec}(L_n)$ is an initial monomial ideal of L_n, it also lies on Γ_n. $\square$

Lemma 2.28 *The tangent space of the multigraded Hilbert scheme $\mathcal{H}_n$ at the point represented by the monomial ideal N_n has dimension $11n - 15$.*

Proof: The tangent space at N_n equals $\text{Hom}(N_n, K[x,y,z]/N_n)_{\mathbf{0}}$. We shall present a basis for this space that is broken into three distinct classes: those homomorphisms that act nontrivially only on the quadratic generators, those that act nontrivially only on the cubics, and those with a mix of both.

Each $K[x,y,z]$-module homomorphism $\varphi : N_n \to K[x,y,z]/N_n$ below is described by its action on the minimal generators of N_n. Any generator not explicitly mentioned is mapped to 0 under φ. One checks that each is in fact a well-defined $K[x,y,z]$-module homomorphism from N_n to $K[x,y,z]/N_n$.

Class I: For each $1 \leq i < n$, we define the following maps

- $\alpha_i : x_i y_k \mapsto y_i y_k$ for all $i < k \leq n$,
- $\beta_i : x_i y_{i+1} \mapsto x_{i+1} y_i$.

For each $1 < k \leq n$, we define the following map

- $\gamma_k : x_i y_k \mapsto x_i x_k$ for all $1 \leq i < k$.

We define two specific homomorphisms

- $\delta_1 : x_1 y_2 \mapsto y_1 z_2$,
- $\delta_2 : x_{n-1} y_n \mapsto z_{n-1} x_n$.

Class II: For each $1 < j < n$, we define the following maps. Each homomorphism is defined on every pair (i, k) such that $1 \leq i < j < k \leq n$.

- $\rho_j : x_i z_j x_k \mapsto x_i x_j x_k$ and $y_i z_j x_k \mapsto y_i x_j x_k$,
- $\sigma_j : x_i z_j x_k \mapsto x_i x_j z_k$ and $y_i z_j x_k \mapsto y_i x_j z_k$,
- $\tau_j : x_i z_j x_k \mapsto x_i z_j z_k$ and $y_i z_j x_k \mapsto y_i z_j z_k$,
- $\nu_j : y_i z_j x_k \mapsto y_i y_j x_k$ and $y_i z_j y_k \mapsto y_i y_j y_k$,
- $\mu_j : y_i z_j x_k \mapsto z_i y_j x_k$ and $y_i z_j y_k \mapsto z_i y_j y_k$,
- $\pi_j : y_i z_j x_k \mapsto z_i z_j x_k$ and $y_i z_j y_k \mapsto z_i z_j y_k$.

Class III: For each $1 \leq i < n$, we define the map

- $\epsilon_i : x_i y_k \mapsto z_i y_k$ and $x_i z_j x_k \mapsto z_i z_j x_k$ for $i < k \leq n$ and $i < j < k$.

For each $1 < k \leq n$, we define the map

- $\zeta_k : x_i y_k \mapsto x_i z_k$ and $y_i z_j y_k \mapsto y_i z_j z_k$ for $1 \leq i < k$ and $i < j < k$.

All these maps are linearly independent over the field K. There are $n-1$ maps each of type α_i, β_i, γ_k, ϵ_i, and ζ_k, for a total of $5(n-1)$ different homomorphisms. Each subclass of maps in class II has $n-2$ members, adding $6(n-2)$ more homomorphisms. Finally adding δ_1 and δ_2, we arrive at the total count of $5(n-1)+6(n-2)+2 = 11n-15$ homomorphisms.

We claim that any $K[x, y, z]$-module homomorphism $N_n \to K[x, y, z]/N_n$ can be recognized as a K-linear combination of those from the three classes described above. To prove this, suppose that $\varphi : N_n \to K[x, y, z]/N_n$ is a module homomorphism. For $1 \leq i < k \leq n$, we can write $\varphi(x_i y_k)$ as a linear combination of monomials of multidegree $e_i + e_k$ which are not in N_n. By subtracting appropriate multiples of α_i, ϵ_i, γ_k, and ζ_k, we can assume that

$$\varphi(x_i y_k) = a\, y_i x_k + b\, y_i z_k + c\, z_i x_k + d\, z_i z_k$$

for some scalars $a, b, c, d \in K$. We show that this can be written as a linear combination of the maps described above by considering a few cases.

In the first case we assume $i + 1 < k$. We use $K[x, y, z]$-linearity to infer

$$\varphi(x_i y_{i+1} y_k) = a\, y_i y_{i+1} x_k + b\, y_i y_{i+1} z_k + c\, z_i y_{i+1} x_k + d\, z_i y_{i+1} z_k = y_k\, \varphi(x_i y_{i+1}).$$

Specifically, y_k divides the middle polynomial. But none of the four monomials are zero in the quotient $K[x, y, z]/N_n$. Hence, $0 = a = b = c = d$.

For the subsequent cases we assume $k = i + 1$. This allows us to further assume that $a = 0$, since we can subtract off $a\, \beta_i(x_i y_{i+1})$. Now suppose that we have strict inequality $k < n$. As before, the $K[x, y, z]$-linearity of φ gives

$$\varphi(x_i y_k y_n) = d\, z_i z_k y_n = y_k\, \varphi(x_i y_n).$$

Specifically, y_k divides the middle term. Hence, $d = 0$. Similarly, $c = 0$:

$$\varphi(x_i y_k z_k x_n) = c\, z_i x_k z_k x_n = y_k\, \varphi(x_i z_k x_n).$$

Suppose we further have the strict inequality $1 < i$. Then necessarily $b = 0$:

$$\varphi(y_1 z_i x_i y_k) = b\, y_1 z_i y_i z_k = x_i\, \varphi(y_1 z_i y_k).$$

However, if $i = 1$ and $k = 2$, we have that $\varphi(x_1 y_2) = b\, \delta_1(x_1 y_2)$.

The only case that remains is $k = n$ and $i = n - 1$. Here, we can also assume that $c = 0$ by subtracting $c\, \delta_2(x_{n-1} y_n)$. We will show that $d = 0 = b$ by once more appealing to the fact that φ is a module homomorphism:

$$\varphi(x_1 x_{n-1} y_n) = d\, x_1 z_{n-1} z_n = x_{n-1}\, \varphi(x_1 y_n),$$

which gives $d = 0$. This subsequently implies the desired $b = 0$, because

$$\varphi(y_1 x_i z_i y_n) = b\, y_1 y_i z_i z_n = x_i\, \varphi(y_1 z_i y_n).$$

This has finally put us in a position where we can assume that $\varphi(x_i y_k) = 0$ for all $1 \leq i < k \leq n$. To finish the proof that φ is a linear combination of the $11n - 15$ classes described

above, we need to examine what happens with the cubics. Suppose $1 \leq i < j < k \leq n$, and consider $\varphi(y_i z_j x_k)$. This can be written as a linear sum of the 17 standard monomials of multidegree $e_i + e_j + e_k$ which are not in N_n. Explicitly, these standard monomials are:

$$\begin{array}{llll} x_i x_j x_k, & x_i x_j z_k, & x_i z_j z_k, & y_i x_j x_k, \quad y_i x_j z_k \\ y_i y_j x_k, & y_i y_j y_k, & y_i y_j z_k, & y_i z_j z_k, \\ z_i x_j x_k, & z_i x_j z_k, & z_i y_j x_k, & z_i y_j y_k, \\ z_i y_j z_k, & z_i z_j x_k, & z_i z_j y_k, & z_i z_j z_k. \end{array}$$

By subtracting off multiples of the maps ρ_j, σ_j, τ_j, ν_j, μ_j, and π_j, we can assume that this is a sum of the 11 monomials remaining after removing $y_i x_j x_k$, $y_i x_j z_k$, $y_i z_j z_k$, $y_i y_j x_k$, $z_i y_j x_k$, and $z_i z_j x_k$. However, now note that

$$\varphi(x_i y_i z_j x_k) = x_i \, \varphi(y_i z_j x_k) = y_i \, \varphi(x_i z_j x_k).$$

This means that for every one of the 11 monomials m appearing in the sum, either $x_i m = 0$ or y_i divides m. Similarly,

$$\varphi(y_i z_j x_k y_k) = y_k \, \varphi(y_i z_j x_k) = x_k \, \varphi(y_i z_j y_k),$$

and so either $y_k m = 0$ or x_k divides m. Taking these both into consideration actually kills every one of the 11 possible standard monomials (we spare the reader the explicit check), and hence we can assume that $\varphi(y_i z_j x_k) = 0$.

Now consider what happens with $\varphi(x_i z_j x_k)$. Indeed,

$$0 = x_i \, \varphi(y_i z_j x_k) = \varphi(x_i y_i z_j x_k) = y_i \, \varphi(x_i z_j x_k).$$

So for every one of the 17 standard monomials m which possibly appears in the support of $\varphi(x_i z_j x_k)$ we must have that $y_i m = 0$ in $K[x, y, z]/N_n$. This actually leaves us with only two possible such standard monomials – namely $z_i z_j x_k$ and $z_i z_j y_k$. We write $\varphi(x_i z_j x_k) = a \, z_i z_j x_k + b \, z_i z_j y_k$.

The fact that we assume $\varphi(x_i y_k) = 0$ implies $a = 0 = b$. This is because

$$0 = z_j x_k \, \varphi(x_i y_k) = \varphi(x_i z_j x_k y_k) = y_k \, \varphi(x_i z_j x_k).$$

To sum up, we have shown that, under our assumptions, if $\varphi(y_i z_j x_k) = 0$ holds then it also must be the case that $\varphi(x_i z_j x_k) = 0$. We can prove in a similar manner that $\varphi(y_i z_j y_k) = 0$, and this finishes the proof that φ can be written as a K-linear sum of the $11n - 15$ classes of maps described. $\square$

We reiterate that Theorem 2.26 fails for $n = 2$, since $\mathcal{H}_2 \simeq \mathbb{P}^8$, and Γ_2 is a cubic hypersurface cutting through $\mathcal{H}_2$. We offer a short report for $n = 3$.

Remark 2.29 *The Hilbert scheme $\mathcal{H}_3$ contains* $13,824$ *monomial ideals. These come in* 16 *symmetry classes under the action of $(S_3)^3 \rtimes S_3$. A detailed analysis of these symmetry classes and how we found the* $13,824$ *ideals appears on the website*

www.math.washington.edu/∼aholtc/HilbertScheme.

For seven of the symmetry classes, the tangent space dimension is less than $\dim(\Gamma_3) = 18$. *From this we infer that $\mathcal{H}_3$ has components other than Γ_3.*

We note that the number $13,824$ *is exactly the number of monomial ideals on $H_{3,3}$ as described in [10]. Moreover, the monomial ideals on $H_{3,3}$ also fall into* 16 *distinct symmetry classes. We do not yet fully understand the relationship between $\mathcal{H}_n$ and $H_{3,n}$ suggested by this observation.*

Moreover, it would be desirable to coordinatize the inclusion $\Gamma_3 \subset \mathcal{H}_3$ and to relate it to the equations defining trifocal tensors*, as seen in [5, 23]. It is our intention to investigate this topic in a subsequent publication.*

Our study was restricted to cameras that take 2-dimensional pictures of 3-dimensional scenes. Yet, residents of *flatland* might be more interested in taking 1-dimensional pictures of 2-dimensional scenes. From a mathematical perspective, generalizing to arbitrary dimensions makes sense: given n matrices of format $r \times s$ we get a map from $\mathbb{P}^{s-1}$ into $(\mathbb{P}^{r-1})^n$, and one could study the Hilbert scheme parametrizing the resulting varieties. Our focus on $r = 3$ and $s = 4$ was motivated by the context of computer vision.

Chapter 3

THE IDEAL OF THE TRIFOCAL VARIETY [3]

3.1 Introduction

In the field of multiview geometry one studies $n \geq 2$ planar images of points in space. Given n full rank 3×4 matrices $A_1, \ldots, A_n$ over $\mathbb{C}$, these *camera matrices* determine a rational map

$$\phi : \mathbb{P}^3 \dashrightarrow (\mathbb{P}^2)^n \qquad x \mapsto (A_1 x, \cdots, A_n x)$$

from projective 3-space into the n-fold product of projective planes. For any given tuple $(A_1, \ldots, A_n)$ the image of this map determines a variety $\overline{\phi(\mathbb{P}^3)} \subseteq (\mathbb{P}^2)^n$ called the *multiview variety associated to* $(A_1, \ldots, A_n)$.

In [4] the authors determine the prime ideal defining the multiview variety for a generic fixed tuple of cameras such that the camera matrices $A_1, \ldots, A_n$ have pairwise distinct kernels. In this paper we focus on a different, but related variety in the special case of $n = 3$ cameras: the variety of all trifocal tensors [22, Ch. 15]. The essential difference is that for the multiview variety the camera matrices are fixed and this determines a map from the world to a set of images, but in the trifocal setup we consider the set of tensors determined by all possible general configurations of triples of cameras. Algebraically, the collection of trifocal tensors is parameterized by the 4×4 minors of the 4×9 matrix $(A_1^T \mid A_2^T \mid A_3^T)$ which involve one column each from the first two blocks, and two columns from the third block [23]. Geometrically, a trifocal tensor arises from a bilinear map describing the geometry of a given configuration of cameras. We give a complete description of the ideal describing this subvariety of tensors.

We further describe this geometric map. Each camera matrix A_i determines a *focal point* $f_i \in \mathbb{P}^3$ and a *viewing plane* $\pi_i \simeq \mathbb{P}^2 \subseteq \mathbb{P}^3$. The image in camera i of a point $x \in \mathbb{P}^3$ is

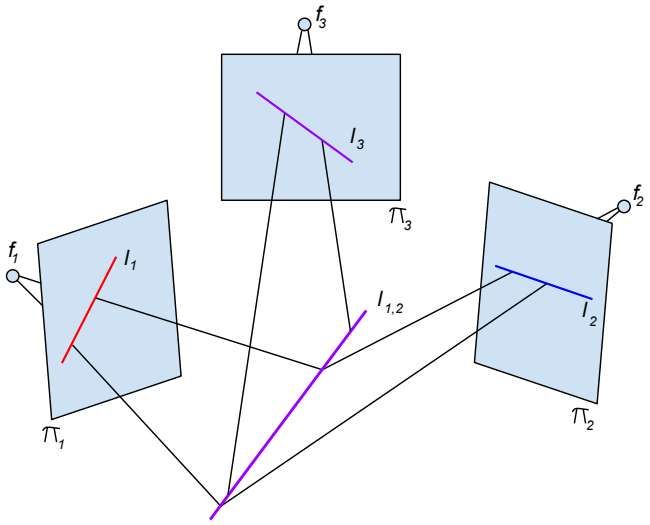


Figure 3.1: The trifocal tensor as a map $\mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2$.

determined by intersecting the line $\langle f_i, x\rangle$ with the plane π_i. Now consider lines $l_i \subseteq \pi_i$ for $i = 1, 2$. The planes $\langle f_1, l_1\rangle$ and $\langle f_2, l_2\rangle$ generically intersect in a line $l_{1,2} \subseteq \mathbb{P}^3$, and the plane $\langle f_3, l_{1,2}\rangle$ generically intersects π_3 in a line l_3. See Figure 3.1.

We have described for a sufficiently general camera configuration, a map

$$\mathbb{P}^2 \times \mathbb{P}^2 \to \mathbb{P}^2,$$

given by $(l_1, l_2) \mapsto l_3$. This map must come from a bilinear map

$$\mathbb{C}^3 \times \mathbb{C}^3 \to \mathbb{C}^3.$$

To help avoid ambiguity, fix $A, B, C \simeq \mathbb{C}^3$ so that this map is now $A \times B \to C$. This bilinear map is equivalently a tensor $T \in A^* \otimes B^* \otimes C$, called a *trifocal tensor* because of its derivation. For more details, see [22, Chapter 15],[5], [23], [46].

One way to connect these two seemingly different algebraic and geometric constructions is via the following construction, which shows (in an invariant way) how the parametrization

using special minors of a 4×9 matrix gives rise to a tensor parametrization. This also relates to the compactified camera space considered in [4].

The row space of $\left(A_1^T \mid A_2^T \mid A_3^T\right)$ determines a point in the Grassmannian $\mathrm{Gr}(4,9)$. Set U_1, U_2, U_3 respectively as the 3-dimensional column spaces of the blocked matrix $\left(A_1^T \mid A_2^T \mid A_3^T\right)$. The direct sum $W = U_1 \oplus U_2 \oplus U_3$ is 9-dimensional, and we can view the matrix $\left(A_1^T \mid A_2^T \mid A_3^T\right)$ as describing a point in the Grassmannian

$$\mathrm{Gr}(4, W) \subset \mathbb{P}(\textstyle\bigwedge^4 W).$$

Consider the group $G = \mathrm{SL}(U_1) \times \mathrm{SL}(U_2) \times SL(U_3) \subset SL(W)$, which can be thought of as the group of (unit determinant) 3×3 blocks on the diagonal of a 9×9 matrix. Now $\bigwedge^4 W$ decomposes as a G-module as follows

$$\bigwedge^4(U_1 \oplus U_2 \oplus U_3) = \left(\bigoplus_{i\neq j} U_i \otimes \bigwedge^3 U_j\right) \oplus \left(\bigoplus_{i\neq j} \bigwedge^2 U_i \otimes \bigwedge^2 U_j\right) \oplus \left(\bigoplus_{i,j,k \text{ distinct}} U_i \otimes U_j \otimes \bigwedge^2 U_k\right).$$

If we take $A^* = U_1, B^* = U_2$ and $C^* = U_3$, we see that the factor $U_1 \otimes U_2 \otimes \bigwedge^2 U_3$ is isomorphic to $A^* \otimes B^* \otimes C$ and corresponds to the space of maximal minors of a 4×9 matrix using 1 column from the first two 4×3 blocks and 2 columns from the last 4×3 block. So we get a G-equivariant projection

$$\pi : \mathrm{Gr}(4,9) \dashrightarrow \mathbb{P}\left(U_1 \otimes U_2 \otimes \bigwedge^2 U_3\right) \cong \mathbb{P}(A^* \otimes B^* \otimes C),$$

and the closure of the image of this projection is the trifocal variety X. Because the projection is G-equivariant, the image X is automatically G-invariant. The generic fiber of the projection π is 3 dimensional. However, there is a $(\mathbb{C}^*)^3$ action by scaling each A_i that leaves the set of trifocal tensors invariant. The GIT quotient $\mathrm{Gr}(4,9)//(\mathbb{C}^*)^3$ has dimension 18, which is the dimension of X, [4, 5]. So if one works with the GIT quotient, the map

$$\pi : \mathrm{Gr}(4,9)//(\mathbb{C}^*)^3 \dashrightarrow \mathbb{P}\left(U_1 \otimes U_2 \otimes \bigwedge^2 U_3\right),$$

becomes birational to X.

One would like to know when a given tensor in $V = A^* \otimes B^* \otimes C$ arose as a trifocal tensor. The Zariski closure of the set of all such trifocal tensors defines an irreducible algebraic variety, called the *trifocal variety* and hereafter denoted by $X \subset \mathbb{P}V$. Let $I(X)$ denote the ideal of polynomial functions vanishing on X, hereafter called the *trifocal ideal*. Since a tensor T is a trifocal tensor (or a limit of such) if and only if T is a zero of every polynomial in the ideal $I(X)$, the question of identifying trifocal tensors can be answered (at least for general tensors in an open set of X) by determining the minimal generators of $I(X)$.

In [5] the authors determine a set of polynomials that cut out X as a set. However, their set of polynomials does not generate the ideal $I(X)$. We note that [46] and [48] also found some equations vanishing on X, but neither described the entire trifocal ideal. The focus of this article is to determine the minimal generators of $I(X)$.

Choosing bases $\{a_1, a_2, a_3\}$, $\{b_1, b_2, b_3\}$, and $\{c_1, c_2, c_3\}$ of A^*, B^* and C respectively, any tensor $T \in V$ can be realized as

$$T = \sum_{i,j,k=1}^{3} T_{i,j,k} a_i \otimes b_j \otimes c_k$$

via the 27 variables $T_{i,j,k}$ for $1 \leq i, j, k \leq 3$ Therefore, the trifocal ideal lives in the polynomial ring $k[T_{ijk}]$.

The cubic polynomials in the ideal are the 10 coefficients of

$$\det(x_1 T_{ij1} + x_2 T_{ij2} + x_3 T_{ij3}).$$

One component of the zero set of these polynomials is our variety. To remove the other components we add polynomials of degree 5 and 6.

To simplify matters, we will take the ground field to be $k = \mathbb{C}$; however, we note that in practice, one works over $\mathbb{R}$. A tensor with real entries is on the complex trifocal variety if and only if it is a zero of all polynomials in $I(X)$. And indeed, all of the generating polynomials in $I(X)$ can be taken with rational coefficients, and thus are in the ideal of X when considered as a variety over $\mathbb{R}$.

Our result is the following.

Theorem 3.1 *Let X denote the trifocal variety. The prime ideal $I(X)$ is minimally generated by 10 polynomials in degree 3, 81 polynomials in degree 5, and 1980 polynomials in degree 6.*

There are noticeably more generators here than in [5], which showed that 10 equations of degree 3, 20 of degree 9, and 6 of degree 12 cut out X set-theoretically. On the other hand, the degrees of our equations are lower and we know that they are the minimal degree polynomials that generate the ideal.

3.2 Outline

To prove Theorem 3.1 and determine the minimal generators of the trifocal ideal, we use a mixture of several different computational and theoretical tools that we now outline. In short, our strategy is to first find equations in the ideal in the lowest degrees, next show that the equations we found cut out the variety set theoretically and thus define an ideal that agrees up to radical with the ideal we want, and then we show that the two ideals are actually equal.

Because the trifocal construction is unchanged by changes of coordinates in each camera plane, we have a large group G that acts on X. We describe this symmetry and various representations for points on X in Section 3.3. Then we describe the geometry of related G-varieties in $\mathbb{P}V = \mathbb{P}(A^* \otimes B^* \otimes C)$ in Section 3.4.

We let $I(X)_d$ denote the degree d piece of $I(X)$, and denote by $M_d(X)$, (or M_d when the context is clear), the (vector space of) minimal generators of $I(X)$ occurring in degree d. The group G acts on X and thus on $I(X)$. This facilitates the search for all polynomials in $I(X)_d$ in low degree (for $d \leq 9$) via computations using classical representation theory. In Section 3.5 we describe our computations and identify which modules of polynomials are minimal generators assisted by symbolic computations in Maple and Macaulay2. In particular, we find that the only minimal generators of $I(X)$ for $d \leq 9$ occur in degrees 3,

5 and 6. Next, we compute a Gröbner basis of the ideal $J = \langle M_3 + M_5 + M_6 \rangle$, and find (again in Macaulay2) that the degree of J is 297.

Another valuable tool for understanding the zero-set of a collection of polynomials is a numerical primary decomposition via numerical algebraic geometry. For this we used Bertini [7] for experiments and computations; see also [51]. In Section 3.6 we consider only M_3, the lowest degree (degree 3) part of $I(X)$, which has a basis of 10 polynomials in the 27 variables. Here we obtain a numerical primary decomposition of $V(M_3)$ using Bertini. In particular, we find that up to the numerical accuracy of Bertini, $V(M_3)$ has 4 components, and we are even given their degrees. This numerical result provides us with tangible data from which we are able to conjecture (and eventually prove) the true structure of $V(M_3)$. In particular, we find that up to the numerical accuracy of Bertini, X has degree 297.

In Section 3.7 we use geometric considerations and resort to Nurmiev's classification of orbits and their closures ([44, 43]) to geometrically identify all the components found by the Bertini computation. This geometric understanding allows us to conclude in Proposition 3.10 that the zero-set $\mathcal{V}(J)$ is equal to X (as sets), so $\sqrt{J} = I(X)$. In Section 3.8 we again use the classification of orbits and the orbit poset structure to show in Theorem 3.15 that J is prime and thus $J = I(X)$.

3.3 The Trifocal Variety as an Orbit Closure

Consider $V = A^* \otimes B^* \otimes C$ and the natural left action of $G = \mathrm{SL}(A) \times \mathrm{SL}(B) \times \mathrm{SL}(C) \simeq \mathrm{SL}(3)^{\times 3}$ on V. There is also a natural action of the symmetric group $\mathfrak{S}_3$ permuting the three factors in the tensor product, and it is easy to see that X is invariant under the action of the $\mathfrak{S}_2$ permuting A^* and B^*. However, this finite invariance does not provide much computational advantage.

Remark 3.2 *Since we are working over* $\mathbb{C}$ *we consider general changes of coordinates by* $\mathrm{SL}(3, \mathbb{C})$. *However, were we to work over* $\mathbb{R}$, *we would want to change our analysis to consider rotations in the three planes, and the group action would be by* $\mathrm{SO}(3, \mathbb{R})^{\times 3}$.

Since the trifocal variety $X \subset \mathbb{P}V$ is invariant under changes of coordinates in the camera planes, we say that X is a G-variety. Moreover, [5] shows that X is actually the closure of a single G-orbit in $\mathbb{P}V$.

Because $G \simeq \mathrm{SL}(3)^{\times 3}$ is 24 dimensional acting on $V \simeq \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$, which is 27 dimensional, there must be infinitely many G-orbits in V. Even so, the orbits happen to have been classified, apparently independently, by several authors. Since elements of V can be interpreted in a number of ways (as triples of 3×3 matrices or 3×3 matrices with linear entries depending on 3 variables, as $3-3-3$ trilinear forms or ternary trilinear forms, as cuboids or elements of a triple tensor product, or as a G-submodule of $\bigwedge^3 \mathbb{C}^9$), the various classifications occurred in different settings — see [57, 41, 44].

We prefer to use Nurmiev's version of the classification, which follows Vinberg's conventions and uses the results and techniques of [59]. One main reason for this choice is that Nurmiev also computed the closures of all the nilpotent orbits in a note [43], in the same language as the previous paper. There are 4 continuous families of orbits called semi-simple orbits, and one finite family of nilpotent orbits. To every orbit $\mathcal{O}$ is associated a *normal form*, which is a representative $v \in V$ such that $G.v = \mathcal{O}$. Though obviously not unique, we will typically choose a normal form that is as simple as possible, or that clearly reveals some of the structure of the orbit.

To use the Nurmiev classification, we first identify the trifocal variety as one of the orbits on Nurmiev's list. Indeed, Alzati and Tortora give a normal form for the orbit of trifocal tensors that we now recall. A general trifocal tensor T may be, after a possible change of coordinates by G, identified as a tensor whose slices in the C-direction are

$$T^1 = \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} T^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix} T^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

Choose bases $A^* = \mathrm{span}\{a_1, a_2, a_3\}$, $B^* = \mathrm{span}\{b_1, b_2, b_3\}$, and $C = \mathrm{span}\{c_1, c_2, c_3\}$. Then we can write

$$T = (-a_1) \otimes b_2 \otimes c_1 + (-a_3) \otimes (-b_1) \otimes c_1 + (-a_2) \otimes b_2 \otimes c_2$$

$$+a_3 \otimes b_2 \otimes c_2 - a_3 \otimes b_2 \otimes c_3 + a_3 \otimes b_3 \otimes c_3,$$

which by changing coordinates (via G) may be written as

$$T = a_1 \otimes b_2 \otimes c_1 + a_3 \otimes b_1 \otimes c_1 + a_2 \otimes b_2 \otimes c_2 + a_3 \otimes b_3 \otimes c_3.$$

It is also useful to express a tensor T via matrices with linear entries. For this, one considers a pure tensor $a_i \otimes b_j \otimes c_k$ as a matrix with an a_i in the j, k position of the matrix. Then do this for all pure tensors in an expression for T and add the matrices. In fact, this describes the projections for the P-Rank varieties defined in Section 3.4. A normal form for the trifocal variety has matrices of linear forms

$$T(A) = \begin{pmatrix} a_3 & 0 & 0 \\ a_1 & a_2 & 0 \\ 0 & 0 & a_3 \end{pmatrix}, \; T(B) = \begin{pmatrix} b_2 & 0 & 0 \\ 0 & b_2 & 0 \\ b_1 & 0 & b_3 \end{pmatrix}, \; T(C) = \begin{pmatrix} 0 & c_1 & 0 \\ 0 & c_2 & 0 \\ c_1 & 0 & c_3 \end{pmatrix}.$$

The difference here is that $T(A)$, $T(B)$, $T(C)$ each individually represent T, but the entire set $\{T^1, T^2, T^3\}$ also represents T. One advantage to considering a tensor as a matrix in linear forms is that it is now clear that P-Rank$(T) = (3, 3, 2)$, so $X \subset \text{P-Rank}^{3,3,2}$. In particular, $T(C)$ has rank 2, and thus must satisfy the equations implied by $\det(T(C)) \equiv 0$, while $T(A)$ and $T(B)$ do not satisfy this relation.

Remark 3.3 *The construction of the matrix $T(A)$ from the tensor T shows that the G action on T corresponds to an action on the matrix $T(A)$ by left and right multiplication by elements of* $\mathrm{SL}(B)$ *and* $\mathrm{SL}(C)$*, and by linear changes of variables on the entries of $T(A)$, with similar descriptions for the action on $T(B)$ and $T(C)$.*

Nurmiev lists the G-orbits in V as a list of integers. To a triple of integers ijk Nurmiev associates the tensor $e_i \otimes e_j \otimes e_k$, with $1 \leq i, j-3, k-6 \leq 3$. The spaces in each expression corresponds to summation.

For example, consider orbit 11 on Nurmiev's list: 149 167 248 357. We choose bases of A^*, B^* and C so that $a_i = e_i$, $b_{j-3} = e_j$, and $c_{k-6} = e_k$. So orbit 11 corresponds to the

tensor

$$a_1 \otimes b_1 \otimes c_3 + a_1 \otimes b_3 \otimes c_1 + a_2 \otimes b_1 \otimes c_2 + a_3 \otimes b_2 \otimes c_1,$$

which corresponds to the matrix of linear forms

$$\begin{pmatrix} 0 & a_2 & a_1 \\ a_3 & 0 & 0 \\ a_1 & 0 & 0 \end{pmatrix}.$$

Finally, notice that $T(C)$, for instance, can be moved by a change of coordinates to

$$T(C) = \begin{pmatrix} 0 & c_1 & 0 \\ 0 & c_2 & 0 \\ c_1 & 0 & c_3 \end{pmatrix} \cong \begin{pmatrix} 0 & 0 & c_1 \\ 0 & 0 & c_2 \\ c_1 & c_3 & 0 \end{pmatrix} \cong \begin{pmatrix} 0 & c_1 & c_3 \\ c_1 & 0 & 0 \\ c_2 & 0 & 0 \end{pmatrix} \cong \begin{pmatrix} 0 & c_3 & c_1 \\ c_2 & 0 & 0 \\ c_1 & 0 & 0 \end{pmatrix}.$$

Then swapping the roles of c_2 and c_3 we obtain

$$T(C) \cong \begin{pmatrix} 0 & c_2 & c_1 \\ c_3 & 0 & 0 \\ c_1 & 0 & 0 \end{pmatrix}.$$

This shows that the normal form of T is congruent to orbit $11''$ on Nurmiev's list (where representative $11''$ is obtained from representative 11 by performing the permutation $a \to b \to c \to a$ twice).

Nurmiev's list also contains the dimension of the stabilizer of this orbit. This confirms for us that the codimension of the trifocal variety X is 8 (an already well-established fact).

One of the utilities of having a group action is the following. If a group $G \subset GL(V)$ preserves a variety $X \subset \mathbb{P}V$ (i.e. $G.X = X$), we may consider $I(X)$ as a G-module and in particular as a G-submodule of $S^\bullet V^*$, the space of symmetric tensors on V. Recall that $S^\bullet V^*$ is isomorphic to the space of homogeneous polynomials on V. The representation theory of $G = \mathrm{SL}(A) \times \mathrm{SL}(B) \times \mathrm{SL}(C)$-modules is well known; however, the reader may wish to consult [31] or [16] for reference. One fact we will use is if $V = A^* \otimes B^* \otimes C$, then irreducible G-modules in $S^\bullet V^*$ are all of the form $S_\lambda A \otimes S_\mu B \otimes S_\nu C^*$, where λ, μ and ν are all partitions of the same nonnegative integer.

3.4 F-Rank and P-Rank Varieties

In the previous section we saw that the matrices T^1, T^2, T^3 representing the slices of a trifocal tensor do not have full rank. This condition depends on the choice of coordinates. On the other hand, the 3×9 flattening matrix $(T^1 \mid T^2 \mid T^3)$ does have full rank, and this condition is not dependent on the choice of coordinates. Of course slicing in a different direction may yield a different result, but it is easy to check that trifocal tensors have full rank flattenings for all slices. We refer to this condition as *flattening rank (*F-Rank*)*, and note that the general trifocal tensor has F-Rank$(T) = (3, 3, 3)$.

The matrix $T(C)$ with linear forms in C does not have full rank, while the matrices $T(A)$ and $T(B)$ do have full rank. The construction of $T(C)$ describes a projection $A^* \otimes B^* \otimes C \to A^* \otimes B^*$, so it is natural to refer to the tuple of ranks of the various projections as *projection rank* (P-Rank). A general trifocal tensor T has P-Rank$(T) = (3, 3, 2)$.

These two considerations lead to the study of subspace varieties (the former) and rank varieties (the latter). Understanding algebraic and geometric properties of these varieties will help us find equations for the trifocal variety. In what follows we highlight some of these properties, which are specific cases of much more general constructions. For more details, see [31, Chapter 7]).

3.4.1 Subspace varieties

For $p \leq 3$, $q \leq 3$, $r \leq 3$, the *subspace variety* $\mathrm{Sub}_{p,q,r} \subset \mathbb{P}V$ is the projectivization of the set of tensors that have F-Rank at most (p, q, r):

$$\mathrm{Sub}_{p,q,r} = \mathbb{P}\{T \in V \mid \text{F-Rank}(T) \leq (p, q, r)\},$$

where we write $(a, b, c) \leq (p, q, r)$ if $a \leq p$ and $b \leq q$ and $c \leq r$. Subspace varieties are irreducible, and their ideals are defined by minors of flattenings (see [33, Theorem 3.1]). For the sake of a reader unfamiliar with these concepts, we recall the construction of these equations.

A tensor $T \in V$ is realized via 27 variables $T_{i,j,k}$ for $1 \leq i, j, k \leq 3$

$$T = \sum_{i,j,k=1}^{3} T_{i,j,k} a_i \otimes b_j \otimes c_k,$$

choosing bases $\{a_1, a_2, a_3\}$, $\{b_1, b_2, b_3\}$, and $\{c_1, c_2, c_3\}$ of A^*, B^* and C respectively. There are three directions in which we may slice T to get triples of matrices. Let $W_i = (T_{i,j,k})_{j,k}$, $Y_j = (T_{i,j,k})_{i,k}$, $Z_k = (T_{i,j,k})_{i,j}$ denote these slices. Then the matrices $W = (W_1 \mid W_2 \mid W_3)$ — respectively $Y = (Y_1 \mid Y_2 \mid Y_3)$, and $Z = (Z_1 \mid Z_2 \mid Z_3)$ — are the three flattenings with respect to the three slicings of the tensor T.

A special case of [33, Theorem 3.1] is that the ideals generated by the 3-minors of flattenings are the ideals of subspace varieties. Namely,

$$I(\mathrm{Sub}_{2,3,3}) = \langle minors(3, W) \rangle,$$

$$I(\mathrm{Sub}_{3,2,3}) = \langle minors(3, Y) \rangle,$$

$$I(\mathrm{Sub}_{3,3,2}) = \langle minors(3, Z) \rangle.$$

Moreover, the intersection of two subspace varieties yields another, and this holds ideal theoretically as well. $\mathrm{Sub}_{2,2,3} = \mathrm{Sub}_{2,3,3} \cap \mathrm{Sub}_{3,2,3}$ and

$$I(\mathrm{Sub}_{2,2,3}) = \langle minors(3, W) \rangle + \langle minors(3, Y) \rangle,$$

and similarly for permutations. Likewise

$$I(\mathrm{Sub}_{2,2,2}) = \langle minors(3, W) \rangle + \langle minors(3, Y) \rangle + \langle minors(3, Z) \rangle.$$

It is also easy to check the dimensions of subspace varieties. A convenient tool is to use the Kempf-Weyman desingularization via vector bundles. Let $\mathcal{S}_i$ denote the canonical (subspace) rank i vector bundle over the Grassmannian $\mathrm{Gr}(i, n)$. The desingularization is

$$\mathbb{P}(\mathcal{S}_p \otimes \mathcal{S}_q \otimes \mathcal{S}_r) \times \mathrm{Gr}(p, 3) \times \mathrm{Gr}(q, 3) \times \mathrm{Gr}(r, 3) \dashrightarrow \mathrm{Sub}_{p,q,r}\,.$$

In particular, we have

$$\dim(\mathrm{Sub}_{p,q,r}) = pqr - 1 + p(3-p) + q(3-q) + r(3-r).$$

We computed the degrees of each subspace variety using Macaulay 2:

variety	$\mathrm{Sub}_{2,3,3}$	$\mathrm{Sub}_{2,2,3}$	$\mathrm{Sub}_{2,2,2}$
dimension	19	15	13
codimension	7	11	13
degree	36	306	783

Another description of $I(\mathrm{Sub}_{p,q,r})$ in Representation Theoretic language will allow us to compare the equations here with any other G-invariant sets of equations, no matter how they are presented. Another way to state [33, Theorem 3.1] is that for each integer d, $I(\mathrm{Sub}_{p,q,r})_d$ consists of all representations $S_\lambda A \otimes S_\mu B \otimes S_\nu C^*$ with partitions $\lambda, \mu, \nu \vdash d$ such that either $|\lambda| > p$, $|\mu| > q$ or $|\nu| > r$.

In fact, the ideals of subspace varieties are generated in the minimal degree d possible. To save space, we write $S_\lambda S_\mu S_\nu$ for the representation $S_\lambda A \otimes S_\mu B \otimes S_\nu C^*$, and $\bigwedge^3$ in place of S_{111}

$$I(\mathrm{Sub}_{2,3,3}) = \langle \textstyle\bigwedge^3 \bigwedge^3 S^3 \oplus \bigwedge^3 S_{21} S_{21} \oplus \bigwedge^3 S^3 \bigwedge^3 \rangle$$

$$I(\mathrm{Sub}_{223}) = I(\mathrm{Sub}_{233}) + \langle S_{21} \textstyle\bigwedge^3 S_{21} \oplus S^3 \bigwedge^3 \bigwedge^3 \rangle$$

$$I(\mathrm{Sub}_{2,2,2}) = I(\mathrm{Sub}_{2,2,3}) + \langle S_{21} S_{21} \textstyle\bigwedge^3 \rangle$$

Finally, comparing to Nurmiev's list [43], the variety $\mathrm{Sub}_{2,3,3}$ corresponds to nilpotent orbit 9 (and $9'$ and $9''$ correspond to permutations of $\mathrm{Sub}_{2,3,3}$). The variety $\mathrm{Sub}_{2,2,3}$ corresponds to nilpotent orbit number 17 (and $17'$ and $17''$ for permutations). $\mathrm{Sub}_{2,2,2}$ is also equal to the secant variety of a Segre product, $\sigma_2(\mathrm{Seg}(\mathbb{P}^2 \times \mathbb{P}^2 \times \mathbb{P}^2))$ and corresponds to nilpotent orbit number 20 on Nurmiev's list.

3.4.2 P-Rank varieties

P-Rank varieties are defined by considering the three images of the projections of a tensor onto two of the factors and restricting the rank of points in the image. In particular (see [31, §7.2.2]) Rank^r_A is the projectivization of the set

$$\{T \in V \mid \mathrm{rank}(T(A)) \leq r\}.$$

Here recall that $T(A) \in B^* \otimes C$ is the projection of T from V, canonical after a choice of coordinates. Indeed, for any choice of coordinates, if the slices of T in the A^*-direction are W_1, W_2, and W_3, then $T(A)$ is identified with the matrix $W_1 + W_2 + W_3$. We see that if $\operatorname{rank}(T(A)) \leq r$, then $W_1 + W_2 + W_3$ has rank $\leq r$ for *all* choices of coordinates for T. Also, if we write the matrix representing $T(A)$ as a matrix depending linearly on the entries of A (the parameters), then this matrix has rank $\leq r$ for all choices of parameters.

The rank varieties Rank_B^r and Rank_C^r are defined similarly. Let $\text{P-Rank}^{p,q,r}$ denote the projectivization of the set of tensors T with $\text{P-Rank}(T) \leq (p, q, r)$. Equivalently,

$$\text{P-Rank}^{p,q,r} = \operatorname{Rank}_A^p \cap \operatorname{Rank}_B^q \cap \operatorname{Rank}_C^r .$$

It is easy to check that $\text{P-Rank}^{p,q,r}$ is $\mathrm{SL}(A) \times \mathrm{SL}(B) \times \mathrm{SL}(C)$-invariant.

While P-Rank varieties can be considered in arbitrary dimensions, we restrict to the case that A, B and C are 3-dimensional.

Landsberg points out that Rank_A^r is usually far from irreducible. In particular, there are at least two subspace varieties in Rank_A^2

$$\operatorname{Sub}_{3,3,2} \cup \operatorname{Sub}_{3,2,3} \subset \operatorname{Rank}_A^2 = \text{P-Rank}^{2,3,3} .$$

In fact, we will see later that there is yet another component. Similarly

$$\operatorname{Sub}_{3,3,2} \cup \operatorname{Sub}_{3,2,3} \subset \text{P-Rank}^{2,3,3},$$

$$\operatorname{Sub}_{3,3,2} \cup \operatorname{Sub}_{2,3,3} \subset \text{P-Rank}^{3,2,3}$$

imply a third containment $\operatorname{Sub}_{3,3,2} \cup \operatorname{Sub}_{2,2,3} \subset \text{P-Rank}^{2,2,3}$.

Moreover the 3-way intersection certainly contains the following

$$\operatorname{Sub}_{3,2,2} \cup \operatorname{Sub}_{2,3,2} \cup \operatorname{Sub}_{2,2,3} \subset \text{P-Rank}^{2,2,2} .$$

But, in fact, all of the inclusions above are strict containments.

In Section 3.7 we consider the poset of orbit closures in $\text{P-Rank}^{3,3,2}$. This will allow us to show that $\text{P-Rank}^{2,2,2}$ is irreducible and corresponds to the orbit closure consisting of tensors

whose projections $T(A)$, $T(B)$ and $T(C)$ are skew-symmetrizable 3×3 matrices. One can also show that the irreducible components of P-Rank^{3,2,2} are $\mathrm{Sub}_{3,3,2}$ and P-Rank^{2,2,2} (which contains $\mathrm{Sub}_{2,2,3}$). Further, we'll see that P-Rank^{3,3,2} consists of four distinct components, namely the two subspace varieties $\mathrm{Sub}_{2,3,3} \cup \mathrm{Sub}_{3,2,3}$, the trifocal variety X, and P-Rank^{2,2,2}.

3.5 Symbolic Computations Using Representation Theory

In this section we compute the trifocal ideal $I(X)$ up to degree 9, and then find the minimal generators among those polynomials.

A systematic algorithm to compute all polynomials in low degree in the ideal of a G-variety is described in [32]. We carry out this algorithm for the trifocal variety. In a word, the test is to decompose the ambient coordinate ring as a G-module and check every module of equations in low degree for membership via representation theory.

Recall that the trifocal variety X is the closure of a G-orbit in $\mathbb{P}V$, with $G = \mathrm{SL}(A) \times \mathrm{SL}(B) \times \mathrm{SL}(C)$. This allows us to use tools from representation theory to compute and understand the ideal of X.

The coordinate ring of $\mathbb{P}V$ has a G-module decomposition in each degree as

$$S^d V^* = \bigoplus_{\lambda,\mu,\nu \vdash d} (S_\lambda A \otimes S_\mu B \otimes S_\nu C^*) \otimes \mathbb{C}^{m_{\lambda,\mu,\nu}},$$

where $S_\lambda A \otimes S_\mu B \otimes S_\nu C^*$ is an isotypic module associated to partitions λ, μ, ν of d, and $m_{\lambda,\mu,\nu}$ is the multiplicity of that isotypic component, [32, Proposition 4.1]. One advantage to considering polynomials in modules is that we can compare different sets of polynomials no matter in what basis they are presented and determine if they are the same.

Even more useful is the following. Because we have a reductive group acting, we have the following splitting as G-modules

$$k[V] = I(X) \oplus k[V]/I(X).$$

To determine $I(X)$, then we just need to determine which irreducible G-modules are in $I(X)$.

Here is a short synopsis of the algorithm in [32]. Suppose M is an irreducible G-module in $S^\bullet V^*$. To determine whether $M \subset I(X)$ or $M \subset k[V]/I(X)$, it suffices to check whether a random point on X vanishes on the highest weight vector of M. Random points of X may be constructed by acting on a normal form by random elements of G. If M is an isotypic component of $S^\bullet V^*$ that occurs with multiplicity m, we first construct a basis of the highest weight space $\mathbb{C}^m$ of M, by a straightforward construction involving Young symmetrizers. Then we select m random points from X and construct an $m \times m$ matrix whose i, j-entry is the ith point evaluated on jth basis vector. The kernel of this matrix tells the linear subspace of M that is in the ideal of X.

Of course because we use random points, we should then re-verify all vanishing results symbolically to rule out false positives (there are no false negatives because non-vanishing at random points of X implies non-vanishing on X.) We did these extensive computations in Maple, and we have provided a sample of our code in the ancillary files accompanying the arXiv version of this article.

We can further determine which among the polynomial modules we find are actually minimal generators. Suppose the degree d piece of the ideal, $I(X)_d$, is known and has been input into Macaulay2 as `I`. Then in the next step of the Landsberg-Manivel algorithm, for each new highest weight vector f (a polynomial of degree $d+1$) check if $f \in \langle I(X)_d \rangle$ by quickly computing `f%I`. The module $\{G.f\}$ associated to the highest weight vector f is in $\langle I(X)_d \rangle$ if and only if `f%I` is zero.

We tabulate the results of this test applied to the trifocal variety below. Again, to save space we write $S_\lambda S_\mu S_\nu$ in place of $S_\lambda A \otimes S_\mu B \otimes S_\nu C^*$. The trifocal variety has an $\mathfrak{S}_2$ symmetry permuting the A^* and B^* factors. To further save space, where appropriate, we will write $\mathfrak{S}_2.(S_\lambda S_\mu)S_\nu$ to indicate the sum $S_\lambda S_\mu S_\nu + S_\mu S_\lambda S_\nu$.

Proposition 3.4 *Let X denote the trifocal variety in $\mathbb{P}V$ and let M_d denote the space of minimal generators in degree d of $I(X)$. There are 10 minimal generators in degree 3, 81 in degree 5, and 1980 in degree 6. The G-module structure of the minimal generators is as*

follows.

$$M_3 = \bigwedge^3 \bigwedge^3 S^3,$$

$$M_5 = (S_{221}S_{221})(S_{311} \oplus S_{221})$$

$$M_6 = (\mathfrak{S}_2.(S_{222}S_{33})(S_{33} \oplus S_{411})) \oplus \mathfrak{S}_2.(S_{33}S_{321})S_{321}$$

$$\oplus (\mathfrak{S}_2.(S_{33}S_{411}) \oplus S_{33}S_{33})\, S_{222},$$

and there are no other minimal generators in degree ≤ 9.

By recording the dimension of all modules that occur, this computes the first nine values of the Hilbert function of $k[V]/I(X)$:

$$27, 378, 3644, 27135, 166050, 865860, 3942162, 15966072, 58409126.$$

During our tests in Maple, we computed a basis of each module and provide these equations in the ancillary files accompanying the arXiv version of this paper.

We relate some of the polynomials we found to the known polynomials in [5]. Landsberg proves that Rank^r_A is the zero-set of $S^{r+1}A \otimes \bigwedge^{r+1} B \otimes \bigwedge^{r+1} C^*$ [31, Proposition 7.2.2.2]. One can also phrase the condition that $T \in \mathrm{Rank}^r_A$ as the requirement that the matrix $T(A)$ of linear forms from A has rank not exceeding r. If A is m-dimensional, a basis of the module $S^{r+1}A \otimes \bigwedge^{r+1} B \otimes \bigwedge^{r+1} C$ is given as follows. Consider the slices $T^1, \ldots, T^m$ of T in the A-direction, and dummy variables $x_1, \ldots, x_m$. The condition that $\mathrm{rank}(\sum_{i=1}^m x_i T^i) \leq r$ is the condition that all coefficients (on the x_i) of the $(r+1) \times (r+1)$ minors of the matrix $\sum_{i=1}^m x_i T^i$ vanish. So a basis of $S^{r+1}A \otimes \bigwedge^{r+1} B \otimes \bigwedge^{r+1} C$ is given by the (polynomial) coefficients of these minors.

Recall that a normal form for a point T on the trifocal variety has

$$T(C) \cong \begin{pmatrix} 0 & c_2 & c_1 \\ c_3 & 0 & 0 \\ c_1 & 0 & 0 \end{pmatrix}.$$

And this matrix clearly has rank ≤ 2.

The above discussion implies that $X \subset \mathrm{Rank}^2_C$, and the module

$$M_3 := \bigwedge^3 A \otimes \bigwedge^3 B \otimes S^3 C^*$$

is in the trifocal ideal. These equations were also identified in [5].

We now describe the two modules in M_5. The highest weight vectors are polynomials with (respectively) 104 and 244 monomials and multi degrees [(2,2,1),(2,2,1),(3,1,1)] and [(2,2,1),(2,2,1),(2,2,1)], in the ring

$$k[a_{11}, \ldots, a_{33}, b_{11}, \ldots, b_{33}, c_{11}, \ldots, c_{33}].$$

Here we are using a_{ij} to denote T_{ij1}, $b_{ij} = T_{ij2}$, and $c_{ij} = T_{ij3}$. Typical terms of the highest weight vectors are

$$\cdots - b_{31}c_{22}a_{13}b_{12}a_{11} + b_{31}a_{12}^2 b_{23}c_{11} - b_{31}c_{21}a_{12}^2 b_{13} \\ + b_{22}c_{31}a_{13}a_{12}b_{11} - b_{22}c_{31}a_{11}b_{13}a_{12} - a_{32}c_{22}b_{11}^2 a_{13} \ldots$$

and

$$\cdots - a_{21}^2 b_{12}b_{33}c_{12} - 2a_{12}a_{33}b_{11}b_{22}c_{21} - a_{21}^2 b_{12}b_{13}c_{32} \\ + a_{12}a_{23}b_{11}b_{21}c_{32} + a_{21}^2 b_{12}^2 c_{33} \ldots,$$

respectively. The basis of $S_{221}S_{221}S_{221}$ consists of 27 polynomials which are all equal after a change of indices. All the coefficients come from the set $\{-5, -2, -1, 1, 2, 4\}$. The basis of $S_{221}S_{221}S_{311}$ consists of 54 polynomials which are of two different types having either 104 or 64 monomials and coefficients in the set $\{-1, 1\}$. The polynomials themselves can be downloaded from the web as mentioned above.

For M_6 we can give a similar description. $S_{222}S_{33}S_{33}$ and $S_{222}S_{33}S_{411}$ and $S_{33}S_{222}S_{411}$ are all 100-dimensional, each with a basis consisting of polynomials that have between 66 and 666 monomials and small (absolute value no greater than 4) integer coefficients. Similarly $S_{321}S_{33}S_{321}$ is 640-dimensional with a basis consisting of polynomials that have between 60

and 732 monomials and small integer coefficients. The full set of polynomials is available with the ancillary files on the arXiv version of the paper.

After computing $I(X)_d$ for small d, we computed a Gröbner basis of $J = \langle M_3 + M_5 + M_6 \rangle$ in Macaulay2. Surprisingly, this computation finished in a few minutes — it actually took longer to load the polynomials into M2 than it took to compute the Gröbner basis. We were also able to compute a Gröbner basis of $\mathfrak{S}_3.M_3 = \left(\bigwedge^3 \bigwedge^3 S^3\right) \oplus \left(\bigwedge^3 S^3 \bigwedge^3\right) \oplus \left(S^3 \bigwedge^3 \bigwedge^3\right)$, the 30 cubic equations defining the rank variety P-Rank^{2,2,2}. We record the results of these computations:

Proposition 3.5 *Let X denote the trifocal variety and let M_d denote the space of minimal generators in $I(X)_d$. The following computations done over $\mathbb{Q}$ hold:*

$$\deg(\mathcal{V}(\mathfrak{S}_3.M_3)) = 1035 \text{ and } \operatorname{codim}(\mathcal{V}(\mathfrak{S}_3.M_3)) = 10.$$

$$\deg(\mathcal{V}(M_3 + M_5 + M_6)) = 297 \text{ and } \operatorname{codim}(\mathcal{V}(M_3 + M_5 + M_6)) = 8.$$

Proof: The proof is by computations in Macaulay2 [17] that we provide with the ancillary files in the arXiv version of the paper. □

Though we don't need it for our proof, we were also able to compute the Hilbert polynomial of $J = \langle M_3 + M_5 + M_6 \rangle$:

$$\begin{aligned} 69P_5 - 423P_6 + 882P_7 - 204P_8 - 2565P_9 & \\ + 5751P_{10} - 6129P_{11} + 3402P_{12} - 783P_{13} + 100P_{14} & \\ - 525P_{15} + 1038P_{16} - 909P_{17} + 297P_{18}, & \end{aligned}$$

where we use the variables P_i following the standard normalization used in Macaulay2 to describe the Hilbert Polynomial.

3.6 Numerical Algebraic Geometry: Bertini

In Numerical algebraic geometry, and specifically using the program Bertini, one can compute numerical primary decompositions of ideals if the number of equations and the degrees

of those equations are relatively small. In contrast to Gröbner basis computations where typically more equations is better, in numerical algebraic geometry it is better to start with the lowest degree and lowest number of equations that one can understand. Then one can try to compute a numerical primary decomposition and attempt to work by other means to obtain a geometric description of the components indicated by Bertini.

Following this philosophy, we started with the 10 equations in degree 3 given by the complete vanishing of

$$\det(x_1Z_1 + x_2Z_2 + x_3Z_3),$$

which define Rank_C^2. Recall that $(Z_1 \mid Z_2 \mid Z_3)$ is the flattening of a tensor in the C-direction. These are specifically the 10 equations defining the module M_3; that is, a basis for M_3.

After about 6.5 hours of computational time on 2 processors (and some help from J. Hauenstein getting the initial parameters right), or just under 10 minutes on Hauenstein's cluster, Bertini succeeded to compute the following numerical decomposition.

Computation 3.6 *Let M_3 denote the 10 coefficients (in x_1, x_2, x_3) of the cubic $\det(x_1Z_1 + x_2Z_2 + x_3Z_3)$. Up to the numerical precision of Bertini, the zero set of M_3 has precisely 4 components:*

In codimension 7 there are 2 components, each of degree 36.

In codimension 8 there is 1 component of degree 297.

In codimension 10 there is 1 component of degree 1035.

It is not too hard to show that the two components in codim 7 are the subspace varieties

$$\text{Sub}_{3,2,3} \cup \text{Sub}_{2,3,3}$$

This is because they have the correct dimension, M_3 is in both ideals, and their ideals are generated (respectively) by

$$(\textstyle\bigwedge^3\bigwedge^3 S^3) \oplus (S^3\bigwedge^3\bigwedge^3) \qquad \text{and} \qquad (\bigwedge^3\bigwedge^3 S^3) \oplus (\bigwedge^3 S^3\bigwedge^3).$$

We know that the trifocal tensor variety is in the zero set and has codimension 8. It is not contained in either subspace variety, so we may conclude that X corresponds to the codimension 8 component in the numerical decomposition. We also learn that X has degree 297.

The variety P-Rank^{2,2,2} must correspond to the codimension 10 component, which we prove in Proposition 3.8 below. In addition, this Bertini computation also tells that P-Rank^{2,2,2} has degree 1035.

3.7 Nurmiev's Classification and Proof of the Main Theorem

The orbits of $\mathrm{SL}(3)^{\times 3}$ acting on $\mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$ have been classified by Nurmiev [44], who also computed the closure of most orbits.

Using Nurmiev's list of normal forms, we can quickly check which orbits are contained in $\mathcal{V}(M_3)$ by taking a parameterized representative for each orbit (normal form) and evaluating the polynomials in M_3 on that representative. This can be carried out in a straightforward manner in any computer algebra system.

The following orbits from Nurmiev's list of nilpotent orbits [43, Table 4] are in $\mathcal{V}(M_3)$: 9, 9′, 11″, 12, 12′, 13, 13′, 14, 15, 15′, 16, 16′, 17, 17′, 17″, 18, 18′, 18″, 19, 19′, 19″, 20 (= 20′ = 20″), 21, 21′, 21″, 22, 22′, 23, 23′, 23″, 24 (= 24′ = 24″), 25 (= ∅), see [44] for an explanation of the notation used.

After considering the nilpotent orbits, we must also consider the semi-simple orbits together with their nilpotent parts. Among these orbits, our tests found that only one semi-simple orbit, namely the one corresponding to Nurmiev's fourth family, is in $\mathcal{V}(M_3)$. In our notation we may represent this normal form as

$$F = \lambda(a_1 \otimes b_2 \otimes c_3 + a_2 \otimes b_3 \otimes c_1 + a_3 \otimes b_1 \otimes c_2 \\ - a_1 \otimes b_3 \otimes c_2 - a_2 \otimes b_1 \otimes c_3 - a_3 \otimes b_2 \otimes c_1),$$

for any scalar $\lambda \neq 0$, but over the complex numbers this scalar may be absorbed.

Remark 3.7 *As a matrix with linear entries either in A^*, B^* or C a normal form for F is always of the form*

$$\begin{pmatrix} 0 & x & -y \\ -x & 0 & z \\ y & -z & 0 \end{pmatrix}.$$

Namely this orbit corresponds to the skew symmetric matrices. Moreover, since this matrix is skew-symmetric, it always has even rank, and thus we find that $F \in \text{P-Rank}^{2,2,2}$ with no computation necessary.

Next we consider the closures of all the nilpotent orbits in [43, Table 4]. Our inclusion poset diagram in Figure 3.7 is enlightening. For all arrows except for those emanating from F, the diagram is a restatement of results in [43]. Namely, if orbit Q is in the closure of orbit P (as indicated by Nurmiev's table) and there isn't already a directed path from P to Q we draw an arrow from P to Q.

We then consider all orbits that are in the zero set of M_3 for all permutations of A,B and C. We write this zero set in shorthand as $\mathcal{V}(\mathfrak{S}_3.M_3)$. Since F is a zero of $\mathfrak{S}_3.M_3$, every orbit in its closure must also be in this zero set. A straightforward computation shows that these orbits in $\mathcal{V}(\mathfrak{S}_3.M_3)$ are numbers 17-21, 23, 24 (and all of their primed versions). Of course this does not imply that these orbits are actually in the closure of F. However, it is enough to check that orbits 17, 18 (and their primed versions) are in the closure of F.

Proposition 3.8 *The zero set $\mathcal{V}(\mathfrak{S}_3.M_3)$ (which equals $\text{P-Rank}^{2,2,2}$ by definition) is irreducible and is the closure of the orbit F above.*

Moreover, the orbit associated to the normal form F is not contained in any of the other orbit closures in Figure 3.7

Proof: This proof is entirely computational, but since we did not find it in the literature, we provide the computations here.

Orbit closures consist of one orbit of the top dimension along with other orbits of lower dimension. So we conclude by counting dimensions that none of the orbits 9, $9'$, $11''$, 12,

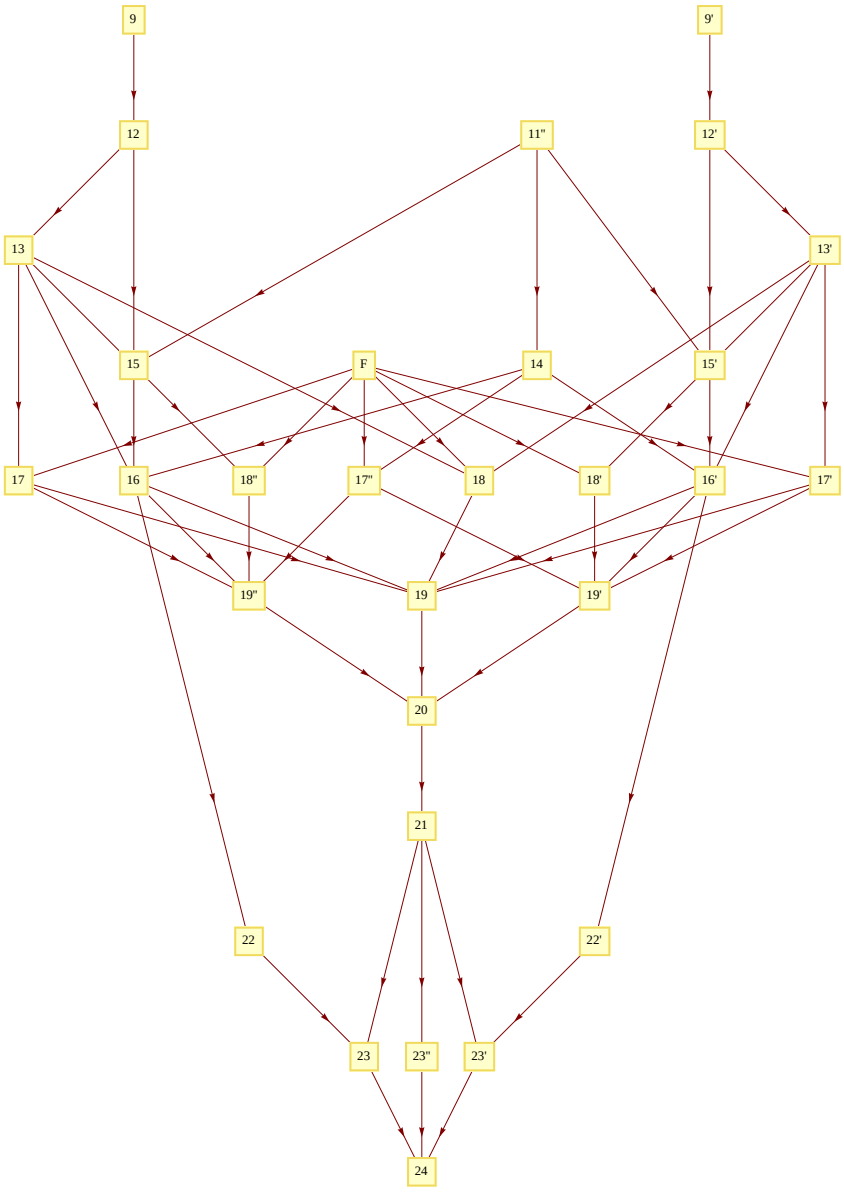


Figure 3.2: A poset diagram for orbit closures in Rank^2_C.

$12'$, 13, $13'$, 14, 15 or $15'$ are in the closure of F. Later we will show that 17, 18 and all their primed versions are in the closure of F.

We claim that none of the orbits of higher dimension (9, $9'$, $11''$, 12, $12'$, 13, $13'$) contain F in their closure. Consider the normal form of F

$$\begin{pmatrix} 0 & x & -y \\ -x & 0 & z \\ y & -z & 0 \end{pmatrix},$$

and flatten it to the matrix

$$\left(\begin{array}{ccc|ccc|ccc} 0 & x & 0 & 0 & 0 & -y & 0 & 0 & 0 \\ -x & 0 & 0 & 0 & 0 & 0 & 0 & 0 & z \\ 0 & 0 & 0 & y & 0 & 0 & 0 & -z & 0 \end{array}\right),$$

which, for general choices of x, y, z has full rank. The other slices have similar format, and this shows that F is not contained in either of the subspace varieties $\mathrm{Sub}_{2,3,3}$ or $\mathrm{Sub}_{3,2,3}$ (the closures of orbits 9 and $9''$). This also implies that F is not in the closure of any orbit contained in the closure of 9 or $9''$.

To show that F is not contained in X, we could demonstrate a polynomial in $I(X)$ that does not vanish on F. We already noted that the skew-symmetric matrices in F has rank 2 or less and thus vanishes on all polynomials in M_3. The ideal $I(X)$ has no minimal generators in degree 4, so we must start to consider polynomials in degree 5 or higher. Here we notice by direct computation that neither of the modules in M_5 vanish on F, separating F from X.

Another way to conclude $\mathrm{Rank}_C^2 \not\subset \mathcal{V}(M_5)$, without computation, is to consider the degree 5 piece of the ideal generated by $M_3 = \bigwedge^3 \bigwedge^3 S_3$. The Pieri formula gives that every module of $I(M_3)$ in degree 5 must have a partition in the first position whose first part is at least 3. On the other hand, the module $S_{221}S_{221}S_{221}$ in M_5 fails this property. So it cannot be in the ideal $I(M_3)$ of Rank_C^2, and thus $\mathrm{Rank}_C^2 \not\subset X$.

Any orbit in $\mathcal{V}(M_3)$ is either in the closure of F or in the closure of 9, $9'$ or $11''$. Moreover, because F is not in the closure of 9, $9'$ or $11''$, its closure must be an irreducible component

of $\mathcal{V}(M_3)$.

Since $\mathfrak{S}_3.M_3$ manifestly has $\mathfrak{S}_3$ symmetry, so does its zero-set. Thus to prove that $\mathcal{V}(\mathfrak{S}_3.M_3)$ is irreducible, we need to show that orbits 17, 18 and their primed versions are contained in the closure of F. This will imply irreducibility of $\mathcal{V}(\mathfrak{S}_3.M_3)$ because every orbit contained there is in the closure of a single orbit.

Since F is symmetric with respect to permutation by $\mathfrak{S}_3$, it suffices to prove that orbits 17 and 18 are contained in the closure of F. To do this we exhibit a sequence of points in the orbit of F that converge to each orbit. We prefer to work with the normal forms both as tensors and as matrices of linear forms. The group operations allowed are row and column operations as well as general linear changes of coordinates on each linear form appearing.

Nurmiev's orbit 17 has normal form given by the matrix

$$\begin{pmatrix} 0 & a_1 & a_2 \\ a_1 & 0 & 0 \\ a_2 & 0 & 0 \end{pmatrix},$$

or the tensor

$$a_1 \otimes (b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes (b_1 \otimes c_3 + b_3 \otimes c_1).$$

Replacing b_1 with $-b_1$ we obtain the tensor

$$a_1 \otimes (-b_1 \otimes c_2 + b_2 \otimes c_1) + a_2 \otimes (-b_1 \otimes c_3 + b_3 \otimes c_1)$$

$$= -a_1 \otimes b_1 \otimes c_2 + a_1 \otimes b_2 \otimes c_1) + -a_2 \otimes b_1 \otimes c_3 + a_2 \otimes b_3 \otimes c_1),$$

which corresponds to the (skew symmetric) matrix

$$\begin{pmatrix} 0 & -a_1 & -a_2 \\ a_1 & 0 & 0 \\ a_2 & 0 & 0 \end{pmatrix}.$$

This matrix is a limit of the following matrices in the orbit of F

$$\begin{pmatrix} 0 & -a_1 & -a_2 \\ a_1 & 0 & z_n \\ a_2 & -z_n & 0 \end{pmatrix},$$

where $z_n \to 0$.

For the orbit 18, we consider the following representative as a tensor

$$a_1 \otimes (b_1 \otimes c_1 + b_2 \otimes c_2) + a_2 \otimes (b_1 \otimes c_2 + b_2 \otimes c_3)$$

which corresponds to the matrix

$$\begin{pmatrix} a_1 & a_2 & 0 \\ 0 & a_1 & a_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

By cycling rows 1, 2 and 3 we at least have a matrix with 0 diagonal,

$$\begin{pmatrix} 0 & a_1 & a_2 \\ 0 & 0 & 0 \\ a_1 & a_2 & 0 \end{pmatrix}.$$

Now we work with the following matrix normal form for F:

$$\begin{pmatrix} 0 & x & -y \\ -x & 0 & z \\ y & -z & 0, \end{pmatrix}$$

which corresponds to the tensor

$$F = z \otimes b_2 \otimes c_3 + y \otimes b_3 \otimes c_1 + x \otimes b_1 \otimes c_2 - z \otimes b_3 \otimes c_2 - y \otimes b_1 \otimes c_3 - x \otimes b_2 \otimes c_1.$$

First we set $x = a_1$ and $-y = a_2$. Considering a limit of such tensors gives a tensor in the closure with $b_2 = 0$ corresponding to the matrix

$$\begin{pmatrix} 0 & a_1 & a_2 \\ 0 & 0 & 0 \\ -a_2 & -z & 0 \end{pmatrix}.$$

Multiply the 3rd row by $\frac{-a_1}{a_2}$ to get another matrix in the same orbit

$$\begin{pmatrix} 0 & a_1 & a_2 \\ 0 & 0 & 0 \\ a_1 & -z\frac{-a_1}{a_2} & 0 \end{pmatrix}.$$

Finally we set $z = \frac{a_2^2}{a_1}$ to yield a matrix in the orbit 18. $\square$

The above discussion provides the following effective test for a given tensor T to be a trifocal tensor. Namely the orbit of trifocal tensors is precisely the G-invariant set of tensors with P-Rank$(T) = (3, 3, 2)$ or some permutation thereof, and F-Rank$(T) = (3, 3, 3)$. These two conditions contain both vanishing and non-vanishing conditions. We phrase this as an algorithm to determine whether a given tensor is a trifocal tensor.

Algorithm 1

Input: *A tensor* $T \in \mathbb{C}^3 \otimes \mathbb{C}^3 \otimes \mathbb{C}^3$.

- *Replace* T *by a change of coordinates (either arbitrary or random) from* $GL(A) \times GL(B) \times GL(C)$ *applied to* T.

- *Compute the projections* $T(A), T(B), T(C)$. *Is* P-Rank$(T) = (3, 3, 2)$ *(or some permutation) and no less?*

  NO: stop, T *is not a trifocal tensor.*

  YES: continue

- *Compute all 3 flattenings. Is* F-Rank$(T) = (3, 3, 3)$ *and no less?*

  NO: stop, T *is not a trifocal tensor.*

  YES: T *is a trifocal tensor.*

If one uses arbitrary changes of coordinates (with parameters) the conclusions of Algorithm 1 hold without modification. However, it may be difficult to perform the tests. If one uses random changes of coordinates, Algorithm 1 will go quickly, and the negative conclusions are sure, but the positive conclusions will hold only with high probability.

This test is effective because it involves computing the ranks of three 3×3 matrices and the ranks of three 3×9 matrices. This test is similar in spirit to the results in [5, Section 4].

Proposition 3.8 yields the following geometric statement.

Proposition 3.9 *Let X denote the trifocal variety (the closure of orbit $11''$). Then the irreducible decomposition of $\mathcal{V}(M_3) = \text{P-Rank}^{3,3,2}$ is*

$$\mathcal{V}(M_3) = \text{Sub}_{2,3,3} \cup \text{Sub}_{3,2,3} \cup X \cup \text{P-Rank}^{2,2,2}\,.$$

Proof: To see that $\mathcal{V}(M_3)$ contains the four listed components, just construct a normal form for each and notice that the associated matrix in linear forms has rank < 3. To see that these are the only components, look at the orbit closure diagram in Figure 3.7, which displays all orbits in $\mathcal{V}(M_3)$ and is justified by [43] and Proposition 3.8. Notice that there are 4 sources in the directed graph representing the poset and these correspond to the only irreducible components in the decomposition. □

Nurmiev's table includes the dimension of the stabilizer of each orbit, which tells the codimension of each of the components: $\text{codim}(\text{Sub}_{2,3,3}) = 7$, $\text{codim}(X) = 8$, $\text{codim}(\text{P-Rank}^{2,2,2}) = 10$. Nurmiev's computation is confirmed by the computation done in Bertini; however, Bertini tells us a bit more, namely the degree of each component.

Proposition 3.10 *The zero set $\mathcal{V}(M_3 + M_5 + M_6)$ is irreducible and agrees with X set-theoretically.*

Proof: We need to show that when we intersect $\mathcal{V}(M_3)$ with $\mathcal{V}(M_5 + M_6)$ that all of the orbits that remain are actually in the trifocal variety. It suffices to show that orbits 17, $17'$, $18'$, $18''$ are not in $\mathcal{V}(M_5 + M_6)$. This is because by considering the orbit closure poset diagram in Figure 3.7, these orbits are contained in all other orbits in $\mathcal{V}(M_3)$ that are not in the trifocal variety X, so if they are not in $\mathcal{V}(M_5 + M_6)$, then no other G-orbit in $\mathcal{V}(M_3)$ outside of X is in $\mathcal{V}(M_3 + M_5 + M_6)$.

By direct computation, we find that the module $S_{33}S_{222}S_{411}$ does not vanish on orbit 17, the module $S_{222}S_{33}S_{411}$ does not vanish on orbit 17', the module $S_{33}S_{33}S_{222}$ does not vanish on 18', and $S_{222}S_{33}S_{33}$ does not vanish on 18". On the other hand, each of these modules are in M_6. □

3.8 The Ideal J is Prime

Let $J = \langle M_3 + M_5 + M_6 \rangle$, where M_d are the minimal generators of $I(X)$ in degree d. The trifocal variety X is irreducible because it is a parameterized variety. The fact that the zero set $V(J)$ is irreducible and equals X is the content of Proposition 3.10. So $I(X)$ and J agree up to radical. It remains to check that there are no embedded components.

The classification of G-orbits in V also yields a classification of minimal G-invariant prime ideals. To every orbit is associated the prime ideal of its orbit closure.

Remark 3.11 *Here we also use the fact that if G is a connected group, and J is a G-stable ideal, then the minimal primes in any primary decomposition of J are G-stable. This essentially follows from the fact that if $J = \cap_i Q_i$ is a primary decomposition with primary ideals Q_i associated to primes P_i, then $gJ = J = \cap_i gQ_i$ for any $g \in G$ and this action must permute the P_i by the uniqueness of minimal primes. But since G is connected, this permutation must be trivial.*

The poset in Figure 3.7 shows that the minimal prime ideals that contain $I(X)$ are those corresponding to orbits 14, 15, and 15′. Let P_{14}, P_{15} and $P_{15'}$ denote the corresponding prime ideals. Then we must have $J \subset P_{14} \cap P_{15} \cap P_{15'}$. On the other hand, we know that $\sqrt{J}$ is prime and equals $I(X)$. So a primary decomposition of J is of the form $J = I(X) \cap Q_{14} \cap Q_{15} \cap Q_{15}$, for some primary ideals Q_i associated to the primes P_i. We will show that the multiplicity of each Q_i with respect to P_i is zero.

If we show this, we don't have to consider possible embedded components coming from the other orbits in the closure of X because these ideals contain P_{14}, P_{15} and $P_{15'}$. Moreover, since X and J have an $\mathfrak{S}_2$ symmetry, if we show that the P_{15} does not occur in the primary decomposition, then neither does $P_{15'}$.

We will use a basic fact from commutative algebra. We found [8, Theorem 12.1] a useful formulation for understanding this type of test.

Proposition 3.12 *[6, Proposition 4.7] Let $\mathfrak{a}$ be a decomposable ideal in a ring A, let $\mathfrak{a} = \cap_{i-1}^{n} \mathfrak{q}_i$ be a minimal primary decomposition and let $\mathfrak{p}_i$ be the prime ideal associated to the*

primary ideal $\mathfrak{q}_i$. *Then*

$$\cup_{i=1}^{n} \mathfrak{p}_i = \{x \in A \mid (\mathfrak{a} : x) \neq \mathfrak{a}\}.$$

In particular, if the zero ideal is decomposable, the set D *of zero-divisors of* A *is the union of the prime ideals belonging to* 0.

We also have the following well-known fact (see for example [13]).

Proposition 3.13 *Let* $R = k[x_0, \ldots x_n]$, *let* J *be an ideal in* R *and suppose* $f \in R$ *has degree* d *and is not a zero-divisor in* R/J. *Then we have the following identity of Hilbert series:*

$$(1 - t^d)H_{R/J}(t) = H_{R/(J+f)}(t).$$

Proof: This is completely standard, but we recall the proof here for the reader's convenience and because it elucidates the ideas we will use later.

If f is not a zero divisor, the following sequence is exact

$$0 \longrightarrow (R/J)(-d) \xrightarrow{f} R/J \longrightarrow R/(J+f) \longrightarrow 0. \tag{3.1}$$

Since $H_{(R/J)(-d)}(t) = t^d H_{(R/J)}(t)$, the result follows from the additivity of Hilbert series. □

Remark 3.14 *If* f *is actually a zero-divisor of* R/J, *then in some degree* t', *the graded version of the sequence* (3.1) *will have a kernel* K *larger than expected. This will force the inequality in*

$$t'^d H_{(R/J)}(t') = H_{(R/J)(-d)}(t') < H_K(t').$$

In this case we will have

$$t'^d H_{(R/J)}(t') - H_{R/J}(t') < H_K(t') - H_{R/J}(t') = H_{R/(J+f)}(t'),$$

which implies that

$$(1 - t'^d)H_{R/J}(t') < H_{R/(J+f)}(t').$$

The previous results allow for the following test. Since the zero-divisors of J correspond to the union of prime ideals P_i that contain J, we can select one $f \in P_i$ of degree d which vanishes on the subvariety $\mathcal{V}(P_i) \subset \mathcal{V}(J)$ but does not vanish on $\mathcal{V}(J)$. If we show that $(1-t^d)H_{R/J}(t) = H_{R/(J+f)}(t)$, then f is not a zero-divisor of R/J. This would show that P_i could not have been a prime ideal associated to J.

We provide the results of this computational test with the prime ideals P_{15} and P_{14}. We wanted to check if a map had a kernel, so we worked over characteristic 101. Non-vanishing modulo a prime p implies non-vanishing in characteristic 0.

In Macaulay 2 we computed a Gröbner basis of $J = \langle M_3 + M_5 + M_6 \rangle$ in about 30 seconds. The Poincare polynomial P_J of R/J is

$$\begin{aligned} P_J = 1 - 10T^3 - 81T^5 - 1605T^6 + 18117T^7 - 77517T^8 + 192794T^9 - 315792T^{10} \\ + 350676T^{11} - 243572T^{12} + 48438T^{13} + 116883T^{14} - 175239T^{15} + 140238T^{16} \\ - 75330T^{17} + 27954T^{18} - 6912T^{19} + 1026T^{20} - 69T^{21}. \end{aligned}$$

For the prime ideal P_{15} we constructed slices in the B-direction Y_1, Y_2, Y_3, computed $\det(x_1Y_1 + x_2Y_2 + x_3Y_3) \equiv 0$ and selected the polynomial f as the coefficient of x_1^3. Precisely,

$$f = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ b_{11} & b_{12} & b_{13} \\ c_{11} & c_{12} & c_{13} \end{pmatrix}.$$

This polynomial f vanishes on P-Rank^{3,2,3} and $\mathcal{V}(P_{15})$ but not on P-Rank^{3,3,2}, and thus not on X.

Computing a Gröbner basis of $J + \langle f \rangle$ took about 10 hours on a server that allowed us to use 16GB of RAM and up to 8 Intel(R) Xeon(R) CPU X5460 3.16GHz processors. The

Poincare polynomial P_f of $R/(J+f)$ is

$$\begin{aligned}P_f = 1 - 11t^3 - 81t^5 - 1595t^6 + 18117t^7 - 77436t^8 + 194399t^9 \\ - 333909t^{10} + 428193t^{11} - 436366t^{12} + 364230t^{13} - 233793t^{14} \\ + 68333t^{15} + 91800t^{16} - 192213t^{17} + 203193t^{18} - 147150t^{19} \\ + 76356t^{20} - 28023t^{21} + 6912t^{22} - 1026t^{23} + 69t^{24}.\end{aligned}$$

Now it is easy to check that $(1-t^3)P_J = P_f$, which implies

$$(1-t^3)H_{R/J}(t) = H_{R/(J+f)}(t),$$

and thus f is not a zero-divisor of R/J. The prime P_{15} is thus not an embedded prime of J. By the $\mathfrak{S}_2$ symmetry of J, we conclude that $P_{15'}$ is also not an embedded prime of J.

For the prime ideal P_{14}, the module $S_{22}S_{211}S_{211}$ vanishes on $\mathcal{V}(P_{14})$ but not on X. We select the highest weight polynomial g for our test:

$$\begin{aligned}g = a_{13}a_{21}b_{12}c_{21} - a_{13}a_{21}c_{12}b_{21} + 3a_{23}a_{12}b_{11}c_{21} + c_{13}a_{21}a_{12}b_{21} - \\ b_{13}a_{21}a_{12}c_{21} - 3a_{23}a_{12}c_{11}b_{21} + a_{23}a_{11}b_{22}c_{11} - a_{23}a_{11}c_{22}b_{11} + \\ 2c_{22}a_{21}a_{13}b_{11} - c_{22}a_{13}a_{11}b_{21} - 2b_{22}a_{21}a_{13}c_{11} + b_{22}a_{21}c_{13}a_{11} + \\ b_{22}a_{13}a_{11}c_{21} - c_{22}a_{21}b_{13}a_{11} + c_{23}a_{11}a_{22}b_{11} + 2b_{23}a_{21}a_{12}c_{11} - \\ b_{23}a_{21}c_{12}a_{11} - b_{23}a_{12}a_{11}c_{21} - c_{13}b_{12}a_{21}^2 + b_{13}c_{12}a_{21}^2 - \\ c_{23}b_{22}a_{11}^2 + b_{23}c_{22}a_{11}^2 - a_{22}a_{21}b_{13}c_{11} - 2c_{23}a_{21}a_{12}b_{11} + \\ c_{23}a_{21}b_{12}a_{11} + c_{23}a_{12}a_{11}b_{21} - b_{23}a_{11}a_{22}c_{11} + a_{23}a_{21}b_{12}c_{11} - \\ a_{23}a_{21}c_{12}b_{11} + a_{22}a_{21}c_{13}b_{11} - 3a_{22}a_{13}b_{11}c_{21} + 3a_{22}a_{13}c_{11}b_{21} - \\ 2c_{21}a_{23}b_{12}a_{11} + 2c_{21}a_{22}b_{13}a_{11} + 2b_{21}a_{23}c_{12}a_{11} - 2b_{21}a_{22}c_{13}a_{11}.\end{aligned}$$

Computing a Gröbner basis of $J + \langle g \rangle$ took about 45 hours to finish on a server that allowed us to use 16GB of RAM and up to 8 processors. The Poincare polynomial P_g of

$R/(J+g)$ is

$$\begin{aligned} P_g = 1 - 10t^3 - t^4 - 81t^5 - 1605t^6 + 18127t^7 - 77517t^8 + 192875t^9 \\ - 314187t^{10} + 332559t^{11} - 166055t^{12} - 144356t^{13} + 432675t^{14} \\ - 525915t^{15} + 383810t^{16} - 123768t^{17} - 88929t^{18} + 168327t^{19} \\ - 139212t^{20} + 75261t^{21} - 27954t^{22} + 6912t^{23} - 1026t^{24} + 69t^{25}. \end{aligned}$$

It is again a simple check that $(1-t^4)P_J = P_g$, which implies

$$(1-t^4)H_{R/J}(t) = H_{R/(J+g)}(t).$$

As before, g is not a zero-divisor of R/J, and the prime P_{14} is not an embedded prime of J.

We have shown the following

Theorem 3.15 *The ideal $J = \langle M_3 + M_5 + M_6 \rangle$ is prime.*

Proof: By Proposition 3.10 we know that $\sqrt{J} = I(X)$. By Proposition 3.5, we know that the degree of the top dimensional component of J is 297, counted with multiplicity. By Computation 3.6, we know that the degree of X is 297. So we know that in a primary decomposition of J, $I(X)$ occurs with multiplicity 1. It only remains to rule out embedded primes. By the above discussion, if we have a primary decomposition of the form $J = I(X) \cap Q_{14} \cap Q_{15} \cap Q_{15'}$, where the Q_i are primary ideals associated to the prime ideals P_i, then we showed that their multiplicity must be zero. So $J = I(X)$, and in particular J is prime. □

This completes the proof of Theorem 3.1. We conjecture that a similar calculation will work to show that the ideal of the orbit closure associated to F is minimally generated by $\mathfrak{S}_3.M_3$.

Chapter 4

A QCQP APPROACH TO TRIANGULATION [2]

4.1 Introduction

We consider the problem of triangulating a point $X \in \mathbb{R}^3$ from $n \geq 2$ noisy image projections. This is a fundamental problem in multi-view geometry and is a crucial subroutine in all structure-from-motion systems [22].

Formally, let the point $X \in \mathbb{R}^3$ be visible in $n \geq 2$ images. Also let $P_i \in \mathbb{R}^{3\times 4}$ be a projective camera and $x_i \in \mathbb{R}^2$ be the projection of X in image i, i.e,

$$x_i = \Pi P_i \tilde{X}, \ \forall \ i = 1, \ldots, n, \tag{4.1}$$

where, using `MATLAB` notation, $\tilde{X} = \Big[X; 1\Big]$ and $\Pi \Big[u; v; w\Big] = \Big[u/w; v/w\Big]$.

Given the set $\{x_i\}$ of noise free projections, it is easy to determine X using a linear algorithm based on singular value decomposition (SVD) [22]. However, in practice we are given $\widehat{x}_i = x_i + \eta_i$, where η_i is noise, and there are no guarantees on the quality of the solution returned by the linear algorithm.

For simplicity, we assume that $\eta_i \sim \mathcal{N}(0, \sigma I)$. Then the *triangulation problem* is to find the maximum likelihood estimate of X given the noisy observations $\{\widehat{x}_i\}$. Assuming such a point X always exists, this is equivalent to solving:

$$\arg\min_X \sum_i^n \|\Pi P_i \tilde{X} - \widehat{x}_i\|^2. \tag{4.2}$$

Here and in the rest of the paper we will ignore the constraint that the point X has positive depth in each image. The above optimization problem is an instance of fractional programming which is in general hard [15]. An efficient and optimal solution of (4.2) is the subject of this paper.

For $n = 2$, Hartley & Sturm showed that (4.2) can be solved optimally in polynomial time [21]. For $n = 3$, a Gröbner basis based algorithm for (4.2) was proposed in [53]. This algorithm relies on the observation that generically, the optimal solution to (4.2) is one among the 47 solutions to a certain system of polynomial equations. This Gröbner basis method is not usefully extendable to higher n and efficient optimal triangulation for $n \geq 4$ views has remained an unsolved problem. Other approaches either use the linear SVD based algorithm as initialization followed by non-linear refinement which lead to locally optimal solutions with no guarantees on the run time complexity [22], or optimal algorithms whose worst case complexity is exponential in n [28, 29, 37].

We present a new triangulation algorithm for $n \geq 2$ views. Based on semidefinite programming, the algorithm in polynomial time either determines a globally optimal solution to (4.2) or informs the user that it is unable to do so. Theoretically, the operating range (in terms of image noise) of the algorithm is limited and depends on the particular configuration of the cameras. In practice our method computes the optimal solution in the vast majority of test cases. In the rare case that optimality cannot be certified, the algorithm returns a solution which can be used as an initializer for nonlinear least squares iteration.

The paper is organized as follows. In Section 4.2 we formulate triangulation as a constrained quadratic optimization problem. We present semidefinite relaxations to this problem in Section 4.3. We propose our triangulation algorithm in Section 4.4 and analyze its performance. We also provide theoretical explanation for why the algorithm works. Section 4.5 presents experiments on synthetic and real data and we conclude in Section 4.6 with a discussion.

Notation. We will use `MATLAB` notation to manipulate matrices and vectors, e.g., $A[1:2, 2:3]$ refers to a 2×2 submatrix of A. $P = \{P_1, \ldots, P_n\}$ denotes the set of cameras. $x = [x_1; \ldots; x_n]$ denotes a vector of image points, one in each camera, and $\widehat{x} = [\widehat{x}_1; \ldots; \widehat{x}_n]$ denotes the vector of image observations. Both x and $\widehat{x}$ lie in $\mathbb{R}^{2n}$. If $y \in \mathbb{R}^m$ is a vector, then $\tilde{y} = \begin{bmatrix} y; 1 \end{bmatrix}$ is the homogenized version of y. The inner product space of $k \times k$ real symmetric matrices is denoted $\mathcal{S}^k$ with the inner product $\langle A, B \rangle = \sum_{1 \leq i,j \leq k} A_{ij} B_{ij}$. The

set $\mathcal{S}_+^k \subseteq \mathcal{S}^k$ denotes the closed convex cone of positive semidefinite matrices. We write $A \succ 0$ (resp. $A \succeq 0$) to mean that A is positive definite (resp. positive semidefinite).

4.2 Triangulation As Polynomial Optimization

With an eye towards the future, let us re-state the triangulation problem (4.2) as the constrained optimization problem

$$\underset{x_1,\ldots,x_n,X}{\arg\min} \sum_i \|x_i - \hat{x}_i\|^2, \quad \text{s.t. } x_i = \Pi P_i \tilde{X}, \quad \forall i = 1,\ldots,n. \tag{4.3}$$

In this formulation, the constraints state that each x_i is the projection of X in image i. Let us now denote the feasible region for this optimization problem by

$$V_P = \left\{ x \in \mathbb{R}^{2n} \mid \exists X \in \mathbb{R}^3 \text{ s.t. } x_i = \Pi P_i \tilde{X},\ \forall i = 1,\ldots,n \right\}. \tag{4.4}$$

For any $x \in V_P$, we can recover the corresponding X using the SVD based algorithm [22]. Now we can re-state (4.3) purely in terms of x

$$\arg\min_x \|x - \hat{x}\|^2, \quad \text{s.t. } x \in V_P. \tag{4.5}$$

Let $\overline{V}_P \supseteq V_P$ be the closure of V_P, meaning that $\overline{V}_P$ contains all the limit points of V_P. Then consider the following optimization problem:

$$\arg\min_x \|x - \hat{x}\|^2, \quad \text{s.t. } x \in \overline{V}_P. \tag{4.6}$$

The objective function in these two optimization problems is the squared distance from $\widehat{x}$ to the sets V_P and $\overline{V}_P$ respectively. Since $\overline{V}_P$ is the topological closure of V_P it can be shown that any solution x^* to (4.6) which is not in V_P is arbitrarily close to a point in V_P, and the optimal objective function values for (4.5) and (4.6) are the same. Thus, solving (4.6) is essentially equivalent to solving (4.5).

The set V_P is a quasi-projective variety. A *variety* is the zero set of a finite set of polynomials, and a *quasi-projective variety* is the set difference of two varieties. Therefore,

$\overline{V}_P$ is also a variety [49]. Heyden & Åström [24] show that

$$\overline{V}_P = \left\{ x \in \mathbb{R}^{2n} \;\middle|\; \begin{array}{l} f_{ij}(x) = 0,\ 1 \le i < j \le n \\ t_{i,j,k}(x) = 0,\ 1 \le i < j < k \le n \end{array} \right\}. \tag{4.7}$$

Here $f_{ij}(x) = \tilde{x}_i^\top F_{ij} \tilde{x}_j = 0$ are the bilinear/quadratic epipolar constraints, where $F_{ij} \in \mathbb{R}^{3\times 3}$ is the fundamental matrix for images i and j. The second set of constraints $t_{ijk}(x_i, x_j, x_k) = 0$ are the the trilinear/cubic constraints defined by the trifocal tensor on images i, j and k.

At the risk of a mild abuse of notation, we will also use F_{ij} to denote a $(2n+1)\times(2n+1)$ matrix such that $f_{ij}(x) = \tilde{x}^\top F_{ij} \tilde{x}$. The construction of this matrix involves embedding two copies of the 3×3 fundamental matrix (with suitable reordering of the entries) in an all zero $(2n+1) \times (2n+1)$ matrix.

Now, let W_P be the quadratic variety

$$W_P = \left\{ x \in \mathbb{R}^{2n} \mid \tilde{x}^\top F_{ij} \tilde{x} = 0,\ 1 \le i < j \le n \right\}. \tag{4.8}$$

For $n = 2$, since there are no trilinear constraints $W_P = \overline{V}_P$ but for $n \ge 3$, in general $W_P \supseteq \overline{V}_P$. For $n \ge 4$, Heyden & Åström show that if the camera centers are not co-planar then $W_P = \overline{V}_P$ [24]. Note that for $n = 3$, the camera centers are always co-planar. Therefore, for $n = 2$, and for $n \ge 4$ when the camera centers are non-co-planar, we can just optimize over the quadratic variety W_P:

$$\arg\min_x \|x - \hat{x}\|^2 \quad \text{s.t. } x \in W_P. \tag{4.9}$$

However, we cannot just ignore the co-planar case as a degeneracy since it is a common enough occurrence, e.g., an object rotating on a turntable in front of a fixed camera. If all the camera centers lie on a plane π_P, then solving (4.9) instead of (4.6) can result in spurious solutions, i.e. a projection vector x^* for which there is no single point $X^* \in \mathbb{R}^3$ that projects to x_i^* for each image i. This can happen if each x_i^* lies on the image of the plane π_P in image i. It is easy to reject such a spurious x^* by checking if it satisfies the trilinear constraints.

From here on, we will focus our attention on solving (4.9) and in Section 4.5 we will show that solving (4.9) instead of (4.6) is not a significant source of failures.

Let us now define the polynomial

$$g(x) = \|x - \widehat{x}\|^2 = \tilde{x}^\top \begin{bmatrix} I & -\widehat{x} \\ -\widehat{x}^\top & \|\widehat{x}\|^2 \end{bmatrix} \tilde{x} = \tilde{x}^\top G \tilde{x}, \tag{4.10}$$

where I is the $2n \times 2n$ identity matrix. Observe that $G \in \mathcal{S}_+^{2n+1}$, and the Hessian of g is $\nabla^2 g = 2I$. Similarly, let $F_{ij} = \begin{bmatrix} H_{ij} & b_{ij} \\ b_{ij}^\top & \beta_{ij} \end{bmatrix}$, where $H_{ij} \in \mathcal{S}^{2n}$, $b_{ij} \in \mathbb{R}^{2n}$, and $\beta_{ij} \in \mathbb{R}$. Then $\nabla^2 f_{ij} = 2H_{ij}$. We re-write (4.9) as the quadratically constrained quadratic program (QCQP)

$$\arg\min_{x \in \mathbb{R}^{2n}} \quad \tilde{x}^\top G \tilde{x} \quad \text{s.t.} \quad \tilde{x}^\top F_{ij} \tilde{x} = 0,\ 1 \le i < j \le n. \tag{4.11}$$

4.3 Semidefinite Relaxation

We re-write (4.11) as the following rank constrained semidefinite program (SDP).

$$\begin{aligned} \arg\min_Y \quad & \langle G, Y \rangle \\ \text{s.t.} \quad & \langle F_{ij}, Y \rangle = 0,\ 1 \le i < j \le n, \\ & \langle E, Y \rangle = 1, \\ & Y \in \mathcal{S}_+^{2n+1}, \\ & \operatorname{rank}(Y) = 1. \end{aligned} \tag{4.12}$$

Here $E \in \mathcal{S}^{2n+1}$ is an all zero matrix except for its bottom right entry which equals one. The problems (4.11) and (4.12) are equivalent: x is feasible (optimal) for (4.11) if and only if $Y = \tilde{x}\tilde{x}^\top$ is feasible (optimal) for (4.12).

Solving rank constrained semidefinite programs is NP-hard [58]. Dropping the rank constraint gives the primal semidefinite program

$$\begin{aligned} \arg\min_Y \quad & \langle G, Y \rangle \\ \text{s.t.} \quad & \langle F_{ij}, Y \rangle = 0,\ 1 \le i < j \le n, \\ & \langle E, Y \rangle = 1, \\ & Y \in \mathcal{S}_+^{2n+1}. \end{aligned} \tag{4.13}$$

The dual of this primal semidefinite program is

$$\begin{aligned} \underset{\lambda_{ij},\rho}{\arg\max} \quad & \rho \\ \text{subject to} \quad & G + \textstyle\sum \lambda_{ij} F_{ij} - \rho E \succeq 0, \\ & \lambda_{ij}, \rho \in \mathbb{R},\ 1 \leq i < j \leq n. \end{aligned} \tag{4.14}$$

The primal SDP (4.13) is also known as the first *moment* relaxation and its dual SDP (4.14) is known as the first *sum of squares* relaxation. They are instances of a general hierarchy of semidefinite relaxations for polynomial optimization problems [34]. Problem (4.14) is also the Lagrangian dual of (4.11).

The remainder of this paper is dedicated to the possibility that solving the triangulation problem is equivalent to solving these semidefinite relaxations. Let us denote by $g^*, g^{\text{quad}}, g^{\text{mom}}, g^{\text{sos}}$, the optimal solutions to the optimization problems (4.6), (4.11), (4.13), (4.14) respectively. Then the following lemmas hold.

Lemma 4.1 *For all n, $g^* \geq g^{\text{quad}}$. For $n = 2$, or $n \geq 4$ with non-co-planar cameras, $g^* = g^{\text{quad}}$.*

Proof: The claim follows from the discussion in Section 4.2 after the definition of W_P. □

Lemma 4.2 *$g^{\text{quad}} \geq g^{\text{mom}}$.*

Proof: This is true because (4.13) is a relaxation of (4.11). □

Lemma 4.3 *For all n, $g^{\text{mom}} = g^{\text{sos}}$. Moreover, there exist optimal Y^*, λ_{ij}^* and ρ^* that achieve these values.*

Proof: The inequality $g^{\text{mom}} \geq g^{\text{sos}}$ follows from weak duality. Equality, and the existence of Y^*, λ_{ij}^* and ρ^* which attain the optimal values follow if we can show that the feasible regions of both the primal and dual problems have nonempty interiors [58, Theorem 3.1] (also known as Slater's constraint qualification).

For the primal problem, let $x \in \mathbb{R}^{2n}$ be any feasible point for the triangulation problem (4.9) (such a feasible point always exists) and let $D = \text{diag}(1, \ldots, 1, 0) \in \mathcal{S}^{2n+1}$. It is easy to show that $Y = \tilde{x}\tilde{x}^\top + D$ is positive definite and primal feasible. For the dual problem, take $\lambda_{ij} = 0$ and $\rho = -1$ and verify $G + E \succ 0$. □

4.4 The Algorithm and its Analysis

We propose Algorithm 1 as a method for triangulation.

Algorithm 1 Triangulation

Require: Image observation vector $\widehat{x} \in \mathbb{R}^{2n}$ and the set of cameras $P = \{P_1, \ldots, P_n\}$.

Solve the primal (4.13) and dual (4.14) SDPs to optimal Y^*, λ_{ij}^* and ρ^*.
$x = Y^*[1 : 2n, 2n + 1]$ (i.e., x is the last column of Y^* without its last entry)
Use the SVD algorithm to determine a world point $X \in \mathbb{R}^3$ from x.
if $I + \sum \lambda_{ij}^* H_{ij} \succ 0$ **then**
  if $n = 2$, or $n \geq 4$ and the cameras P are non-co-planar, **then**
    Return (OPTIMAL, X).
  end if
  if $x_i = \Pi P_i \tilde{X}$ $\forall i = 1, \ldots, n$ **then**
    Return (OPTIMAL, X)
  end if
end if
Return (SUBOPTIMAL, X).

4.4.1 Correctness

Theorem 4.4 *Algorithm 1 terminates in time polynomial in n.*

Proof: The proof is based on three facts. One, the primal (4.13) and dual (4.14) have descriptions of size polynomial in n, the number of images. Two, SDPs can be solved to

arbitrary precision in time which is polynomial in the size of their descriptions [58]. Three, the eigenvalue decomposition of a matrix can be computed in polynomial time. □

Before moving forward, we need the following definition.

Definition 4.5 *A triangulation problem is* SDP-EXACT *if* $g^{\text{quad}} = g^{\text{sos}} = g^{\text{mom}}$*, i.e., the relaxations are tight.*

We will first describe the conditions under which a triangulation problem is SDP-EXACT. We will then show that if Algorithm 1 returns OPTIMAL then triangulation is SDP-EXACT, and further, the solution X returned by the algorithm is indeed optimal for triangulation.

Theorem 4.6 *Let x^* be an optimal solution to the quadratic program* (4.11)*. The triangulation problem is* SDP-EXACT *if and only if there exist $\lambda_{ij} \in \mathbb{R}$ such that*

$$\text{(i) } \nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0 \qquad \textit{and} \qquad \text{(ii) } I + \sum \lambda_{ij} H_{ij} \succeq 0. \tag{4.15}$$

Before proving this theorem, we observe that it is not immediately useful from a computational perspective. Indeed, *a priori* verifying condition (i) requires knowledge of the optimal solution x^*. However, the theorem will help us understand *why* the triangulation problem is so often SDP-EXACT in Section 4.4.3.

Let $L(x, \lambda_{ij}, \rho) = g(x) + \sum_{ij} \lambda_{ij} f_{ij}(x) - \rho = \tilde{x}^\top \left(G + \sum \lambda_{ij} F_{ij} - \rho E\right) \tilde{x}$. Observe that $\nabla_x L(x, \lambda_{ij}, \rho) = \nabla g(x) + \sum \lambda_{ij} \nabla f_{ij}(x)$ and $\nabla_x^2 L(x, \lambda_{ij}, \rho) = 2(I + \sum \lambda_{ij} H_{ij})$. We require that the following two simple lemmas, the proofs of which can be found in the Appendix.

Lemma 4.7 *If x^* is the optimal solution to* (4.11) *and λ_{ij} satisfy condition (i), then $L(x, \lambda_{ij}, g(x^*)) = (x - x^*)^\top \left(I + \sum \lambda_{ij} H_{ij}\right)(x - x^*)$. Further, if condition (ii) is satisfied as well, then $L(x, \lambda_{ij}, g(x^*)) \geq 0,\ \forall x \in \mathbb{R}^{2n}$.*

Lemma 4.8 *If $x^\top A x + 2b^\top x + c \geq 0,\ \forall x$, then* $\begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \succeq 0.$

Proof: [Theorem 4.6] For the *if* direction, let λ_{ij} satisfy conditions (i) and (ii). Then from Lemma 4.7 we have $L(x, \lambda_{ij}, g(x^*)) \geq 0\ \forall x \in \mathbb{R}^{2n}$, which combined with Lemma 4.8 gives

$G + \sum \lambda_{ij} F_{ij} - g(x^*)E \succeq 0$. Therefore λ_{ij} and $\rho = g(x^*)$ are dual feasible, which in turn means that $g^{\text{sos}} \geq \rho = g(x^*) = g^{\text{quad}}$. Lemmas 4.2 and 4.3 give the reverse inequality, thus $g^{\text{sos}} = g^{\text{mom}} = g^{\text{quad}} = g(x^*)$.

For the *only if* direction, let ρ^* and λ_{ij}^* be the optimal solution to the dual (4.14). If the problem is SDP-EXACT we have the equality $\rho^* = g(x^*) = g^{\text{quad}}$ and from the dual feasibility of λ_{ij}^* and ρ^* we have that $G + \sum \lambda_{ij}^* F_{ij} - \rho^* E \succeq 0$. Taken together, these two facts imply that $L(x, \lambda_{ij}^*, g(x^*)) \geq 0, \ \forall x \in \mathbb{R}^{2n}$.

Since $L(x^*, \lambda_{ij}^*, g(x^*)) = 0$, $L(x, \lambda_{ij}^*, g(x^*))$ is a non-negative quadratic polynomial that vanishes at x^*. Non-negativity implies condition (ii) (that the Hessian of the polynomial is positive semidefinite) and the fact that zero is the minimum possible value of a non-negative polynomial implies that its gradient vanishes at x^* which is exactly condition (i). $\square$

Condition (ii) of Theorem 4.6 is automatically satisfied by *any* feasible λ_{ij} for the dual (4.14). Hence, verifying SDP exactness using the dual optimal λ_{ij}^* reduces to checking condition (i) of Theorem 4.6. Checking this condition is not computationally practical, since it requires knowledge of the optimum x^*. By slightly tightening condition (ii) we can bypass condition (i).

Theorem 4.9 *If* $\{Y^*, \ \lambda_{ij}^*, \ \rho^*\}$ *are primal-dual optimal and* $I + \sum \lambda_{ij}^* H_{ij} \succ 0$, *then* $\text{rank}(Y^*) = 1$, *and triangulation is* SDP-EXACT.

Proof: Notice that $I + \sum \lambda_{ij}^* H_{ij}$ is the top left $(2n) \times (2n)$ block in the larger $(2n+1) \times (2n+1)$ positive semidefinite matrix $G + \sum \lambda_{ij}^* F_{ij} - \rho^* E$. By hypothesis, $I + \sum \lambda_{ij}^* H_{ij}$ is nonsingular and thus has full rank equal to $2n$, which implies

$$\text{rank}\left(G + \sum \lambda_{ij}^* F_{ij} - \rho^* E\right) \geq 2n. \tag{4.16}$$

The dual and the primal SDP solutions satisfy *complementary slackness*, which means that $\left\langle G + \sum \lambda_{ij}^* F_{ij} - \rho^* E, Y^* \right\rangle = 0$. In particular it implies that

$$\text{rank}(G + \sum \lambda_{ij}^* F_{ij} - \rho^* E) + \text{rank}(Y^*) \leq 2n + 1, \tag{4.17}$$

where we use the standard fact that whenever $\langle A, B \rangle = 0$ for $A, B \in \mathcal{S}^N_+$, then $\text{rank}(A) + \text{rank}(B) \leq N$. From (4.16) and (4.17) we have $\text{rank}(Y^*) \leq 1$. Since $\langle E, Y \rangle = 1$, we have $Y^* \neq 0$ and hence $\text{rank}(Y^*) = 1$. □

Line 4 of Algorithm 1 uses Theorem 4.9 to establish that we have solved (4.11). Lines 5–10 of the algorithm are then devoted to making sure that the solution actually lies in $\overline{V}_P$. Thus, Algorithm 1 is correct.

4.4.2 Implications

Theorem 4.10 *If the image observations are noise free, i.e. there exists $X^* \in \mathbb{R}^3$ such that $\widehat{x}_i = \Pi P_i X^*, \ \forall i = 1, \ldots, n$, then Algorithm 1 returns* OPTIMAL

Proof: Setting $\lambda_{ij} = 0$ satisfies the hypothesis for Theorems 4.6 and 4.9. □

Theorem 4.11 *Two view triangulation is* SDP-EXACT.

Proof: For $n = 2$ the triangulation problem (4.11) involves minimizing a quadratic objective over a single quadratic equality constraint $f_{12}(x) = 0$. The conditions of Theorem 4.6 in this case reduce to finding $\lambda \in \mathbb{R}$ satisfying

$$\nabla g(x^*) + \lambda \nabla f_{12}(x^*) = 0 \qquad \text{and} \qquad I + \lambda H_{12} \succeq 0. \tag{4.18}$$

The existence of such a λ follows directly from [40, Theorem 3.2]. □

Theorems 4.9 and 4.11 nearly imply that two-view triangulation can be solved in polynomial time. We say *nearly* because, despite Theorem 4.11, it is possible that the matrix $I + \lambda^* H_{12}$ is singular for the dual optimal λ^* (see Appendix for an example). This is not a contradiction, since Theorem 4.9 is only a sufficient condition for optimality. Despite such pathologies, we shall see in Section 4.5 that in practice Algorithm 1 usually returns OPTIMAL for two-view triangulation.

4.4.3 Geometry of the Algorithm

Recall that the optimization problem (4.11) can be interpreted as determining the closest point x^* to $\widehat{x}$ in the variety W_P. This viewpoint gives geometric intuition for why Algorithm 1 can be expected to perform well in practice.

Lemma 4.12 *Given $\widehat{x}$, and assuming appropriate regularity conditions at the optimal solution x^* to* (4.11)*, there exist $\lambda_{ij} \in \mathbb{R}$ such that*

$$\widehat{x} = x^* + \sum \frac{\lambda_{ij}}{2} \nabla f_{ij}(x^*) \tag{4.19}$$

Proof: It follows from Lagrange multiplier theory [42] that there exist Lagrange multipliers $\lambda_{ij} \in \mathbb{R}$ such that $\nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0$. Observe that for (4.11) $\nabla g(x^*) = 2(x^* - \widehat{x})$, which finishes the proof. □

If $\|x^* - \widehat{x}\|$ is small, i.e. $\widehat{x}$ is close to the variety W_P, then there must exist some λ_{ij} satisfying (4.19) such that $\|\lambda\|$ is small and hence $I + \sum \lambda_{ij} H_{ij}$ is a small perturbation of the positive definite identity matrix I. Since I lies in the interior of the positive semidefinite cone, these small perturbations also lie in the interior, that is $I + \sum \lambda_{ij} H_{ij} \succ 0$.

This, coupled with the fact that λ_{ij} are Lagrange multipliers at x^*, yields the sufficient conditions in Theorem 4.6 for a triangulation problem to be SDP-EXACT. Thus, if the amount of noise in the observations is small, Algorithm 1 can be expected to recover the optimal solution to the triangulation problem. Since H_{ij} depends only on the cameras P_i and not on $\widehat{x}$, the amount of noise which Algorithm 1 will tolerate depends only on P_i. We summarize this formally in the following theorem.

Theorem 4.13 *Let $N(x^*) = \{x^* + \sum \frac{\lambda_{ij}}{2} \nabla f_{ij}(x^*) \mid I + \sum \lambda_{ij} H_{ij} \succ 0\}$. For any $\widehat{x} \in \mathbb{R}^{2n}$, if Algorithm 1 returns* OPTIMAL *and x^* is the optimal image projection vector then $\widehat{x} \in N(x^*)$. Conversely, if $\widehat{x} \in N(x^*)$ and x^* is the closest point in W_P to $\widehat{x}$, then Algorithm 1 will return* OPTIMAL*.*

Proof: The proof is a straightforward application of Lemma 4.12 and Theorem 4.9. □

4.5 Experiments

Algorithm 1 was implemented using `YALMIP` [36], `SeDuMi` [54] and `MATLAB`. These tools allow for easy implementation and the timings below are for completeness and should not be used to judge the runtime performance of the algorithm.

Fundamental matrices and epipolar constraints are specified only up to scale and badly scaled fundamental matrices lead to poorly conditioned SDPs. This was easily fixed by dividing each F_{ij} by its largest singular value.

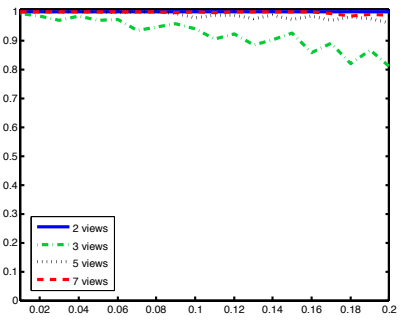
Algorithm 1 verifies optimality by checking the positive definiteness of $I + \sum \lambda_{ij}^* H_{ij} \succ 0$, which requires that its smallest eigenvalue be greater than some $\delta > 0$. In all our experiments we set $\delta = 0.05$.

Either Algorithm 1 returns OPTIMAL and the solution is guaranteed to be optimal, or it returns SUBOPTIMAL in which case we cannot say anything about the quality of the solution, even though it could still be optimal or at least serve as a good guess for non-linear iteration. Thus, we run Algorithm 1 on a number of synthetic and real-world datasets and report the fraction of cases in which the algorithm certifies optimality.

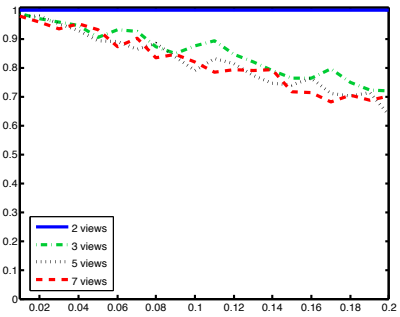
4.5.1 Synthetic Data

To test the performance of the algorithm as a function of camera configuration and image noise we first tested Algorithm 1 on three synthetic data sets. Following [45], we created instances of the triangulation problem by randomly generating points inside the unit cube in $\mathbb{R}^3$.

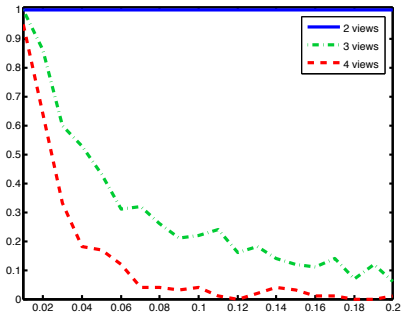
For the first experiment, a varying number of cameras (2, 3, 5 and 7) were placed uniformly at random on a sphere of radius 2. In the second experiment, the same number of cameras were uniformly distributed on a circle of radius 2 in the xy-plane. In the third experiment they were restricted to the x-axis and were placed at a distance of 3, 5, 7, and 9 units. These setups result in image measurements with an approximate square length of 2 units. Gaussian noise of varying standard deviation was added to the image measurements.



(a) Cameras on a sphere.



(b) Cameras on a circle.



(c) Cameras on a line.

Figure 4.1: Fraction of synthetic experiments in which Algorithm 1 returned OPTIMAL versus the standard deviation of the noise level added to the images. (a) Cameras placed randomly on the sphere of radius 2. (b) Three cameras all placed on the xy-plane. (c) Three cameras placed along the x-axis.

The maximum standard deviation was 0.2, which corresponds to about 10% of the image size. For each noise level we ran the experiment 375 times. Figures 4.5.1(a), (b) and (c) show the fraction of test cases for which Algorithm 1 returned OPTIMAL as a function of the the standard deviation of perturbation noise.

We first note that for $n = 2$ cameras, in all three cases, independent of camera geometry and noise, we are able to solve the triangulation problem optimally 100% of the time. This experimentally validates Theorem 4.11 and provides strong evidence that for $n = 2$ there is no gap between Theorems 4.6 and 4.9 in practice.

Observe that for cameras on a sphere (Figure 4.5.1(a)), the algorithm performs very well, and while the performance drops as the noise is increased, the drop is not significant for practical noise levels. Another interesting feature of this graph is that finding and certifying optimality becomes easier with increasing number of cameras. The reason for this increase in robustness is not clear and we plan to further pursue this intriguing observation in future research.

Table 4.1: Performance of Algorithm 1 on real data. The column OPTIMAL reports the fraction of triangulation problems which were certified optimal.

Data set	# images	# points	OPTIMAL	Time (sec)
Model House	10	672	1.000	143
Corridor	11	737	0.999	193
Dinosaur	36	4983	1.000	960
Notre Dame	48	16,288	0.984	7200

Cameras on a circle (Figure 4.5.1(b)) is one of the degenerate cases of our algorithm where $\overline{V}_P \neq W_P$. We expect a higher rate of failure here and indeed this is the case. It is however worth noting that the algorithm does not immediately breakdown and shows a steady decline in performance as a function of noise like the previous experiment. Unlike cameras on a sphere, increasing the number of cameras does not increase the performance of the algorithm, which points to the non-trivial gap between $\overline{V}_P$ and W_P for co-planar cameras.

Finally let us consider the hard problem of triangulating a point when the camera is moving on a line. This case is hard geometrically and algorithmically as the cameras are co-planar, and this difficulty is reflected in the performance of the algorithm. In contrast to the previous experiments, we observe rapid degradation with increasing noise.

4.5.2 Real Data

We tested Algorithm 1 on four real-world data sets: the Model House, Corridor, and Dinosaur from Oxford University and the Notre Dame data set [50]. Our results are summarized in Table 4.1.

Model house is a simple data set where the camera moves laterally in front of the model house. Global optimality was achieved in all cases.

Corridor is a geometrically hard sequence where most of the camera motion is forward, straight down a corridor. This is similar to the synthetic experiment where the cameras were all on the x-axis. The algorithm returned OPTIMAL in all but one case.

Dinosaur consists of images of a (fake) dinosaur rotating on a turntable in front of a fixed camera. Even though the camera configuration is hard for our algorithm (cameras are co-planar), global optimality was achieved in all cases.

The three Oxford datasets are custom captures from the same camera. Notre Dame consists of images of the Notre Dame cathedral downloaded from Flickr. The data set comes with estimates of the radial distortion for each camera, which we ignore as we are only considering projective cameras in this paper. Algorithm 1 returned OPTIMAL in 98.36% of cases.

It is worth noting here that the synthetic datasets were designed to test the limits of the algorithm as a function of noise. A standard deviation equal to 0.2 translates to image noise of approximately 10% of the image size. In practice such high levels of noise rarely occur, and when they do, they typically correspond to outliers. The results in Table 4.1 indicate that the algorithm has excellent performance in practice.

4.6 Discussion

We have presented a semidefinite programming algorithm for triangulation. In practice it usually returns a global optimum and a certificate of optimality in polynomial time. Of course there are downsides which must be taken into consideration. Solving SDPs is not nearly as fast as gradient descent methods. Moreover, what happens in the rare cases that global optimality is not certifiable? Regardless, the lure of a polynomial time algorithm and a better understanding of the geometry of the triangulation problem is far too great to allow these hiccups to close the door on SDPs.

Hartley and Sturm [21] show that two-view triangulation can be solved by finding the roots of a degree 6 univariate polynomial. For $n = 3$, [53] gives a Gröbner basis based algorithm for triangulation by finding all solutions to a polynomial system with 47 roots.

These methods do not extend in a computationally useful manner to $n \geq 4$. Our algorithm solves the triangulation problem for all values of $n \geq 2$ under the conditions of Theorem 4.9 when the camera centers are not co-planar.

Similar to our setup, Kanatani et al. also frame triangulation as finding the closest point on a variety from noisy observations [30]. Unlike our work, they use both epipolar and trifocal constraints, which makes the constraint set much more complicated. Further, they do not prove finite convergence or the optimality of their procedure. Our work answers their question of analyzing the shape of the variety to obtain a *noise threshold* for optimality guarantees.

Semidefinite relaxations for optimization problems in multi-view geometry were first studied by Kahl & Henrion in [29]. They formulate triangulation as an optimization problem over quartic inequalities and study its moment relaxations. Kahl & Henrion observe that good solutions can be obtained from the first few relaxations. In our method the very first relaxation already yields high quality solutions. Further, the quadratic equality based formulation has nice theoretical properties that explain the empirical performance of our algorithm.

Hartley & Seo proposed a method for verifying global optimality of a given solution to the triangulation problem [20]. The relation between their test and the definiteness test of Theorem 4.9 is a fascinating direction for future research.

The study of QCQPs has a long history. There is a wide body of work devoted to verifying optimality of a given solution for various classes of QCQPs and using semidefinite programming to solve them. Thus it is natural that statements similar to Theorems 4.6 and 4.9 for various subclasses of QCQPs have appeared several times in the past e.g. [27, 35, 47].

The astute reader will notice that even though Theorems 4.6 and 4.9 are presented for triangulation, the proofs apply to any QCQP with equality constraints, and can be interpreted in terms of the Karush-Kuhn-Tucker (KKT) conditions for (4.11). More recently, and independently from us, Zheng et al. have proved versions of Theorems 4.6 and 4.9 for the more general case of inequality-constrained QCQPs [60]. Our formulation of triangulation

as minimizing distance to a variety allows for a geometric interpretation of these optimality conditions, which in turn explains the performance of our algorithm.

Finally, we use the general purpose SDP solver `SeDuMi` for our experiments. Much improvement can be expected from a custom SDP solver which makes use of the explicit symmetry and sparsity of the triangulation problem (e.g. for all n the linear matrix constraints in the primal semidefinite program (4.13) are sparse and have rank at most five).

Appendix

Failure of Theorem 4.9 for two-view triangulation

Consider the cameras

$$A_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & a \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & b \end{bmatrix}, \tag{4.20}$$

where $0 < a < b$. This models the situation in which two cameras are placed on the x-axis at $(a, 0, 0)$ and $(b, 0, 0)$, respectively, both facing the origin (i.e. viewing down the x-axis).

We can show that, given $\varepsilon > 0$ and $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) = \left(\begin{bmatrix} 0; \varepsilon \end{bmatrix}, \begin{bmatrix} \varepsilon; 0 \end{bmatrix} \right)$, the minimum value of the QCQP (4.11) is ε^2 and all points of the form

$$x_1^* = \begin{bmatrix} \sqrt{\mu(\varepsilon - \mu)}; & \mu \end{bmatrix} \tag{4.21}$$

$$x_2^* = \begin{bmatrix} \varepsilon - \mu; & \sqrt{\mu(\varepsilon - \mu)} \end{bmatrix} \tag{4.22}$$

where $0 \leq \mu \leq \varepsilon$ are optimal.

The non-uniqueness of the minimizer implies that the verification matrix $I + \lambda H_{12}$ cannot be positive definite. So in this case, the hypothesis of Theorem 4.9 is not satisfied, although Theorem 4.11 still holds.

The proof of these facts is a straight-forward application of the KKT conditions. The minimum distance ε^2 can be attained by setting one of the image points equal to the epipole, but not both. However, the minimum distance is also attained at points which correspond to neither image being the epipole.

Proof of Lemma 4.7

Recall conditions (i) $\nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0$, and (ii) $I + \sum \lambda_{ij} H_{ij} \succeq 0$. Let $H = \sum \lambda_{ij} H_{ij}$, $b = \sum \lambda_{ij} b_{ij}$, and $\beta = \sum \lambda_{ij} \beta_{ij}$. Then

$$L(x, \lambda_{ij}, g(x^*)) = g(x) + \sum \lambda_{ij} f_{ij}(x) - g(x^*) \tag{4.23}$$

$$= x^\top (I + H)\, x + 2\,(b - \widehat{x})^\top x + \widehat{x}^\top \widehat{x} + \beta - g(x^*). \tag{4.24}$$

We wish to show that (4.24) is equal to $(x - x^*)^\top (I + H)\,(x - x^*)$. The fact that $L(x, \lambda_{ij}, g(x^*)) \geq 0$ is then immediate because $I + H \succeq 0$. Looking at the quadratic, linear, and constant terms separately, it suffices to prove

$$\text{(a) } (I + H)x^* = \widehat{x} - b \qquad \text{and} \qquad \text{(b) } {x^*}^\top (I + H)\, x^* = \widehat{x}^\top \widehat{x} + \beta - g(x^*). \tag{4.25}$$

Indeed, (a) is just a restatement of condition (i), since $\nabla g(x) = 2(x - \widehat{x})$ and $\nabla f_{ij}(x) = 2(H_{ij}x + b_{ij})$. To prove (b), first recall that since x^* is feasible for problem (4.11), $f_{ij}(x^*) = 0$ for all i, j. In particular, $\beta = -2b^\top x^* - {x^*}^\top H x^*$. Using this and part (a), one can verify (b) through straightforward manipulations. □

Proof of Lemma 4.8

Suppose the matrix $M = \left[A, b\,; b^\top, c\right]$ is not positive semidefinite, i.e. there exists y such that $y^\top M y < 0$. Write $y = [y'; \gamma]$ for some $\gamma \in \mathbb{R}$. If $\gamma = 0$, then $0 > y^\top M y = {y'}^\top A y'$. But then we arrive at a contradiction by considering $x = \mu y'$ for a scalar $\mu \in \mathbb{R}$, since for μ large enough, $x^\top A x + 2b^\top x + c < 0$.

Now if $\gamma \neq 0$, setting $x = y'/\gamma$ gives

$$x^\top A x + 2b^\top x + c = \frac{1}{\gamma^2} \left(y^\top M y \right) < 0, \tag{4.26}$$

which again contradicts the hypothesis. □

BIBLIOGRAPHY

[1] S. Agarwal, N. Snavely, I. Simon, S.M. Seitz, and R. Szeliski. Building rome in a day. In *ICCV*, pages 72–79, 2009.

[2] C. Aholt, S. Agarwal, and R. Thomas. A QCQP approach to triangulation. In *ECCV*, 2012.

[3] C. Aholt and L. Oeding. The ideal of the trifocal variety. *arXiv preprint arXiv:1205.3776*, 2012.

[4] C. Aholt, B. Sturmfels, and R. Thomas. A Hilbert scheme in computer vision. *Canadian Journal of Mathematics*, published electronically in 2012, doi:10.4153/CJM-2012-023-2.

[5] A. Alzati and A. Tortora. A geometric approach to the trifocal tensor. *J. Math. Imaging Vision*, 38(3):159–170, 2010.

[6] M. F. Atiyah and I. G. MacDonald. *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969.

[7] D.J. Bates, J.D. Hauenstein, A.J. Sommese, and C.W. Wampler. Bertini: Software for numerical algebraic geometry. Available at http://www.nd.edu/∼sommese/bertini, 2010.

[8] W. Bruns and U. Vetter. *Determinantal rings*, volume 1327 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988.

[9] D. Cartwright, M. Häbich, B. Sturmfels, and A. Werner. Mustafin varieties. *Selecta Mathematica, New Series*, pages 1–37, 2010.

[10] D. Cartwright and B. Sturmfels. The Hilbert scheme of the diagonal in a product of projective spaces. *International Mathematics Research Notices*, 2010(9):1741–1771, 2010.

[11] A. Conca. Linear spaces, transversal polymatroids and asl domains. *Journal of Algebraic Combinatorics*, 25(1):25–41, 2007.

[12] D.A. Cox, J.B. Little, and D. O'Shea. *Ideals, Varieties and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*, volume 10. Springer, 2007.

[13] D. Eisenbud. *Commutative Algebra : with a View Toward Algebraic Geometry (Graduate Texts in Mathematics)*. Springer, 1999.

[14] O. Faugeras and Q.T. Luong. *The geometry of multiple images: the laws that govern the formation of multiple images of a scene and some of their applications*. MIT press, 2004.

[15] Roland W. Freund and Florian Jarre. Solving the sum-of-ratios problem by an interior-point method. *J. Glob. Opt.*, 19(1):83–102, 2001.

[16] W. Fulton and J. Harris. *Representation theory: a first course*, volume 129. Springer, 1991.

[17] D.R. Grayson and M.E. Stillman. Macaulay 2, a software system for research in algebraic geometry, 2002.

[18] F.D. Grosshans. On the equations relating a three-dimensional object and its two-dimensional images. *Advances in Applied Mathematics*, 34(2):366–392, 2005.

[19] M. Haiman and B. Sturmfels. Multigraded Hilbert schemes. *Journal of Algebraic Geometry*, 13(4):725–769, 2004.

[20] Richard Hartley and Yongduek Seo. Verifying global minima for l_2 minimization problems. In *CVPR*, 2008.

[21] Richard Hartley and Peter Sturm. Triangulation. *CVIU*, 68(2):146–157, 1997.

[22] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, second edition, 2003.

[23] A. Heyden. Tensorial properties of multiple view constraints. *Math. Methods Appl. Sci.*, 23(2):169–202, 2000.

[24] Anders Heyden and Kalle Åström. Algebraic properties of multilinear constraints. *Math. Methods Appl. Sci.*, 20(13):1135–1162, 1997.

[25] A. Jensen. Cats, a software system for toric state polytopes, 2003.

[26] A.N. Jensen. Gfan, a software system for gröbner fans and tropical varieties, 2009.

[27] V. Jeyakumar, AM Rubinov, and ZY Wu. Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. *Mathematical Programming*, 110(3):521–541, 2007.

[28] Fredrik Kahl, Sameer Agarwal, Manmohan Krishna Chandraker, David J. Kriegman, and Serge Belongie. Practical global optimization for multiview geometry. *IJCV*, 79(3):271–284, 2008.

[29] Fredrik Kahl and Didier Henrion. Globally optimal estimates for geometric reconstruction problems. *IJCV*, 74(1):3–15, 2007.

[30] K. Kanatani, H. Niitsuma, and Y. Sugaya. Optimization without search: Constraint satisfaction by orthogonal projection with applications to multiview triangulation. *Memoirs of the Faculty of Engineering*, 44:32–41, 2010.

[31] J. M. Landsberg. *Tensors: geometry and applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[32] J. M. Landsberg and L. Manivel. On the ideals of secant varieties of Segre varieties. *Found. Comput. Math.*, 4(4):397–422, 2004.

[33] J. M. Landsberg and J. Weyman. On the ideals and singularities of secant varieties of Segre varieties. *Bull. Lond. Math. Soc.*, 39(4):685–697, 2007.

[34] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, 2009.

[35] G. Li. Global quadratic minimization over bivalent constraints: Necessary and sufficient global optimality condition. *Journal of Optimization Theory and Applications*, pages 1–17, 2012.

[36] J. Lofberg. YALMIP: A toolbox for modeling and optimization in matlab. In *Int. Symp. on Computer Aided Control Systems Design*, pages 284–289, 2004.

[37] Fangfang Lu and Richard I. Hartley. A fast optimal algorithm for l_2 triangulation. In *ACCV*, volume 2, pages 279–288, 2007.

[38] D. Maclagan and B. Sturmfels. Introduction to tropical geometry. *Book in preparation*, 2009.

[39] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer, 2004.

[40] J. Moré. Generalizations of the trust region problem. *Optimization methods and Software*, 2(3-4):189–209, 1993.

[41] K. O. Ng. The classification of $(3, 3, 3)$ trilinear forms. *J. Reine Angew. Math.*, 468:49–75, 1995.

[42] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Verlag, 1999.

[43] A. G. Nurmiev. Closures of nilpotent orbits of cubic matrices of order three. *Uspekhi Mat. Nauk*, 55(2(332)):143–144, 2000.

[44] A. G. Nurmiev. Orbits and invariants of third-order matrices. *Mat. Sb.*, 191(5):101–108, 2000.

[45] C. Olsson, F. Kahl, and R. Hartley. Projective least-squares: Global solutions with local optimization. In *CVPR*, pages 1216–1223, 2009.

[46] T. Papadopoulo and O. Faugeras. A new characterization of the trifocal tensor. In Hans Burkhardt and Bernd Neumann, editors, *Proceedings of the 5th European Conference on Computer Vision*, volume 1406–1407 of *Lecture Notes in Computer Science*, Freiburg, Germany, June 1998. Springer–Verlag.

[47] MC Pinar. Sufficient global optimality conditions for bivalent quadratic optimization. *Journal of optimization theory and applications*, 122(2):433–440, 2004.

[48] C. Ressl. *Geometry, Constraints, and Computation of the Trifocal Tensor*. PhD thesis, Universität Bonn, 2003.

[49] I. Shafarevich. *Basic Algebraic Geometry I: Varieties in Projective Space*. Springer-Verlag, 1998.

[50] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Photo tourism: Exploring photo collections in 3d. In *SIGGRAPH*, pages 835–846, 2006.

[51] A.J. Sommese and C.W. Wampler. *Numerical solution of polynomial systems arising in engineering and science*. World Scientific, Singapore, 2005.

[52] R.P. Stanley. *Combinatorics and commutative algebra*. Birkhäuser Boston, 2004.

[53] H. Stewenius, F. Schaffalitzky, and D. Nister. How hard is 3-view triangulation really? In *ICCV*, pages 686–693, 2005.

[54] J.F. Sturm. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Opt. Meth. and Soft.*, 11-12:625–653, 1999.

[55] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8. Amer Mathematical Society, 1996.

[56] B. Sturmfels and A. Zelevinsky. Maximal minors and their leading terms. *Advances in mathematics*, 98(1):65–112, 1993.

[57] R. M. Thrall and J. H. Chanler. Ternary trilinear forms in the field of complex numbers. *Duke Math. J.*, 4(4):678–690, 1938.

[58] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[59] È. B. Vinberg and A. G. Èlašvili. A classification of the three-vectors of nine-dimensional space. *Trudy Sem. Vektor. Tenzor. Anal.*, 18:197–233, 1978.

[60] X.J. Zheng, X.L. Sun, D. Li, and Y.F. Xu. On zero duality gap in nonconvex quadratic programming problems. *Journal of Global Optimization*, 52(2):229–242, 2012.