Computer Vision and Computer Algebra

vorgelegt von

M. Sc. André Wagner

geb. in Landstuhl

von der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation

Promotionsausschuss:

Vorsitzende: Prof. Dr. Gitta Kutyniok

Gutachter: Prof. Dr. Michael Joswig

Gutachter: Directeur de recherche Didier Henrion

Tag der wissenschaftlichen Aussprache: 23. November 2017

Berlin 2017

„Orpheus could not aspire to charm the wild beasts with his music until the end of time. However, one could have hoped that Orpheus himself would not become a wild beast.“

Julien Benda

Zusammenfassung

André Wagner

Computer Vision and Computer Algebra

In Multiview-Geometrie, einem Teilgebiet der Computer-Vision, werden Bilder mit Kameras von 3D-Objekten aus verschiedenen Perspektiven aufgenommen. Diese Arbeit studiert Multiview-Geometrie mit Methoden aus der Computer-Algebra. Hauptprobleme des Maschinellen Sehens sind die Objektrekonstruktion und die Bestimmung der Kameraparameter. In diesen beiden fundamentalen Fragestellungen erzielen wir neue Erkenntnisse unter der Voraussetzung, dass zusätzliche Informationen über die ursprünglichen dreidimensionalen Objekte zur Verfügung stehen.

In Multiview-Geometrie beschreibt die Multiview-Varietät die Beziehungen der Bildpunkte in verschiedenen Bildern. Wir entwickeln zwei Verallgemeinerungen der Multiview-Varietät, die Rigid Multiview-Varietät und die Unlabeled Multiview-Varietät. Beide Varietäten sind durch konkrete Anwendungsbeispiele im Maschinellen Sehen motiviert. Sie können benutzt werden, um Qualität und Geschwindigkeit der Identifikation von Markern auf 3D-Objekten zu verbessern. Die Rigid Multiview-Varietät wird mengentheoretisch beschrieben und ein Triangulierungs-Algorithmus für die Unlabeled Multiview-Varietät konzipiert.

Der 8-Punkt Algorithmus ist einer der wichtigsten Algorithmen in Multiview-Geometrie und einer der meist genutzten Algorithmen, um die Fundamentalmatrix zu bestimmen. Es ist bekannt, dass er bei Bildern vom Einheitswürfel unabhängig von der Lage der Kameras scheitert. Wir erweitern diese Aussage auf Bilder von allgemeinen kombinatorischen Würfeln. Für diesen Fall beschreiben wir einen neuen Algorithmus, der die Qualität der Rekonstruktion der Fundamentalmatrix im Vergleich zum 7- und 8-Punkt Algorithmus drastisch verbessert.

Abschließend bestimmen wir Teilschnitte der primären Zerlegung der mengentheoretischen Gleichungen der Veronese-Varietät. Da die Veronese-Varietät ein Binomial-Ideal ist, können wir diese wir mit Hilfe von Kombinatorik beschreiben.

Abstract

André Wagner

Computer Vision and Computer Algebra

In multiview geometry, a field of computer vision, images of a three-dimensional scene are taken by several cameras from various perspectives. We dedicate ourselves to studying multiview geometry by means of computer algebra. Three-dimensional scene and camera parameter reconstruction is at the core of computer vision. This thesis obtains novel results about these two fundamental topics in certain cases if additional information about the original three-dimensional scene is available.

The multiview variety encodes the space of three-dimensional points seen through various views. We extend the knowledge about the multiview variety to two generalizations of it, the rigid multiview variety and the unlabeled multiview variety. These two varieties are inspired by specific applications in computer vision. They can be used to improve and speed-up the identification of unlabeled marker configurations. We give a set-theoretical description of the rigid multiview variety and design a triangulation algorithm for the unlabeled multiview variety.

The 8-point algorithm is one of the most important algorithms in multiview geometry, and the most commonly used algorithm for fundamental matrix estimation. It is known that the unit cube defeats the 8-point algorithm. We extend this result to all combinatorial cubes. Two perspective projections of a combinatorial cube defeat the 8-point algorithm independent of the position of the cameras. In this case we describe a new algorithm that drastically improves the quality of reconstruction of the fundamental matrix compared to the 7- and 8-point algorithm.

Finally we determine subintersections with omitted single intersectands of the primary decomposition of the set-theoretic equations of the Veronese variety. As the Veronese variety is an binomial ideal, these can actually be described by combinatorial means.

Acknowledgements

This thesis would not have been possible without the support of many people. First of all I would like to thank my advisor, Michael Joswig, for continuous motivation and many helpful discussions. Thanks goes as well to Didier Henrion for being my second assessor.

Many mathematicians inspired and helped me by discussions and answering my questions. Especially, I would like to thank my co-authors, Michael Joswig, Thomas Kahle, Joe Kileel and Bernd Sturmfels. Furthermore I would like to thank my colleagues of the Discrete Mathematics/Geometry Group and the Einstein Visiting Group, Carlos Amendola, Benjamin Assarf, Simon Hampe, Lars Karstner, Kathleen Kohn, Robert Loewe, Georg Loho, Benjamin Lorenz, Marta Panizzut, Manuel Radons, Benjamin Schröter and Kristin Shaw for great discussions and uncountably many lunches at *Who's That Burger*.

I am very grateful to Robert Loewe, Georg Loho, Thorsten Kud, Lars Karstner, Manuel Radons, Chris Redlich, and my father for proof reading. Additionally I'd like to thank Georg Hasselberg, Aaron de Wolff, Kim de Wolff for our climbing- and Thorsten for our crypto currency-sessions.

Finally, I thank Ivalu for her inexplicable strong love.

Contents

List of Figures

List of Tables

List of Abbreviations

$\mathbb{P}^n$	n-dimensional Projective space
$GL(\cdot)$	General linear group
$PGL(\cdot)$	Projective general linear group
$S(\cdot)$	Symmetric group
$O(\cdot)$	Orthogonal group
$\mathrm{Sym}_2(\mathbb{R}^n)$	$n \times n$ symmetric matrix with entries in $\mathbb{R}$
$\mathrm{adj}(\cdot)$	Classical adjoint matrix
$SE(\cdot)$	Special Euclidean
$V(\cdot)$	Algebraic variety cut out by an ideal
V_A	Multiview variety
J_A	Rigid multiview ideal
$\mathrm{Sym}_m(V_A)$	Unlabeled multiview variety of m unlabeled points
$\cdot : \cdot$	Colon ideal
f_i	Focal point
e_{ij}	Epipole in the i-th picture on the line of f_i, f_j
β_{ij}	Baseline of the focal points f_i, f_j
F	Fundamental matrix
I_n	Veronese ideal of degree 2 in n indeterminates
$\mathbb{N}M$	Monoid $\{Mu : u \in \mathbb{Z}^n\}$
$M^{-1}[b]$	Fiber $\{\alpha \in \mathbb{N}^n : Mx = b\}$
m_+	Positive part of a vector
m_-	Negative part of a vector
ρ	Partial character

to my fiancée Ivalu and Fnuß ...

Chapter 1

Introduction

This thesis is concerned with computer algebra approaches to computer vision. Multiview geometry is at the core of computer vision. In multiview geometry images of a three-dimensional scene are taken by several cameras from various perspectives. It has a vast number of applications, like real-time scene analysis and autonomously driving cars. When multiple cameras view a three-dimensional scene, then numerous geometric relations between the two-dimensional and three-dimensional scenes relate to algebraic constraints. In particular, the case of multiview geometry seems to be manageable for computer algebra due to the low dimension of the varieties related to it. In recent years, computer algebra systems [GS; Dec+16] have made great progress. Problems that used to be out of reach computationally are now attackable. The research in this thesis is strongly influenced by methods from computer algebra. The first chapters are in direct relation to multiview geometry. In the last chapter our knowledge about algebraic vision and computer algebra systems is used to solve problems in classical commutative algebra, as of that we have implemented several algorithms in software systems like `polymake` [GJ00], `Macaulay 2` [GS], `Singular` [Dec+16] and `matlab` [MAT16].

We are only concerned with *pinhole cameras*. A pinhole camera can be modeled as a 3×4 matrix with real entries. Such a camera maps a *world point* in the three-dimensional projective space $\mathbb{P}^3$ down to an *image point* in the projective plane $\mathbb{P}^2$. This camera model is the most commonly used model and is algebraically nice to handle. Among multiple views the image points must satisfy the *multilinear constraints*. For all pairs of images these are the *bilinear constraints* given by the *fundamental matrix* $F \in \mathbb{R}^{3\times3}$, for all triplets of images these are the *trilinear constraints* given by the *trifocal tensor* $T \in \mathbb{R}^{3\times3\times3}$, and so on. These relations among images are crucial to three-dimensional scene and camera position reconstruction.

The first approaches to the field of *algebraic vision* have already been made in [Hr97]. Heyden and Åström were one of the first to phrase multiview geometry in the language of algebraic geometry and already used Gröbner basis approaches. The *multiview variety* V_A is the algebraic variety of the relations between multiple images, e. g. the algebraic variety cut out by the multilinear constraints. Heyden and Åström prove that the relations between all pairs of pictures are sufficient to cut out V_A set-theoretically. Later in [AST13] a generating system for the prime ideal of the multiview variety and a universal Gröbner basis was determined. The prime ideal is generated by all bilinear and trilinear constraints. Further a universal Gröbner basis is given by all bilinear, trilinear and quadrilinear constraints. Up to isomorphism the multiview variety of four cameras itself can be constructed from the 2-minors of a generic 4×4 matrix. Thus the multiview variety can be studied via binomial ideals. Many computer vision algorithms are based on the descriptions above.

FIGURE 1.1: Upper left: Training VRmagic's Eyesi indirect ophthalmoscope. Lower left: Eyesi optical tracking system. Right: Eyesi device and multiview sketch. Pictures courtesy of VRmagic.

Anyhow this dissertation is not purely driven by algebraic interests, but it investigates various application motivated scenarios in multiview geometry with algebraic methods. The individual topics have been worked out in discussions with *VRmagic*, a company for ophthalmic surgery simulation using augmented reality. In their medical simulator *Eyesi*, a trainee surgeon is performing a simulated ophthalmic surgery. The surgeon wears a virtual reality headset with four cameras mounted at the front. Active or passive markers are attached to simulator instruments. In Figure 1.1 the black screen (upper left) and the lens (right) depict such a simulator device with marker configurations attached. The markers can be freely placed on each device. Different problems arise in this context:

- How to improve the reconstruction with a prior knowledge of the relative Euclidean distance of the markers in the three-dimensional scene?
- What are "good or bad" choices of marker configurations for robust reconstruction of camera positions?
- How to identify the markers in the two-dimensional pictures, if they are unlabeled?

In Chapters 2-5 the mathematical foundations providing answers to such questions are laid. These results also yield improvements to ophthalmic surgery simulations and similar, even more general problems, in computer vision. Novel results about the image relations, three-dimensional scene and camera position reconstruction are obtained for the case that additional information about the original three-dimensional scene is available. In these cases we construct relations between multiple views and design algorithms to reconstruct the structure defining objects of the multiview geometry. There are essentially two such crucial objects that need to be restored in the majority of applications, e.g. world points and cameras. Both cases are covered for special scenarios. We also apply our knowledge in algebraic vision to purely theoretical results in commutative algebra, namely understanding the set-theoretical description of the Veronese variety and its primary decomposition.

In Chapter 2 a brief introduction to the mathematical basics in this dissertation is given.

In Chapter 3 we study a generalization of the multiview variety. The multiview variety itself is not capable of addressing additional constraints in the three-dimensional scene. Our aim is to make use of additional constraints among multiple world points. Think of taking pictures of a three-dimensional marker configuration. The relative Euclidean distance between marker pairs is fixed. Obviously, two world points seen from one camera can map to the same image points independently on their Euclidean distance. However, when multiple cameras view two world points, then their Euclidean distance gives algebraic relations between the two-dimensional pictures. The multiview variety is not able to address these constraints directly. The *rigid multiview variety* $V(J_A)$ is the product of two multiview varieties, where the image points are projections of all pairs of world points with a fixed Euclidean distance of one. Despite the complexity of the rigid multiview variety in Theorem 3.7, we determine a complete set-theoretical description of this variety. Our proof is somewhat more general, such that we extend it in Theorem 3.10 to broad classes of related varieties even with multiple world points present.

In Chapter 4 we make use of mathematical optimization software and apply it to the results of Chapter 3. Our aim is to improve the accuracy of the *triangulation problem* when rigid data is available. Especially, we are interested in global optimization solvers. However, our computational experiments exhibit difficulties for standard global polynomial and non-linear solvers to deal with the rigid approach to triangulation.

Chapter 5 is about a common problem of multiview geometry, namely that the image point correspondences are unknown. This means that the image data is unlabeled. We try to determine whether there is a labeling of the image points such that they could have been seen by the given camera configuration. Two equivalent descriptions of the so called *unlabeled multiview variety* $\mathrm{Sym}_m(V_A)$ are used. This is the variety that takes all possible labelings of the image points into account. It can either be formulated by using the Chow variety to give a description in terms of symmetric tensors, or by taking unions of multiview varieties with interchanged image point correspondences (Proposition 5.4). We reconstruct the unlabeled world point configuration and design the *unlabeled triangulation algorithm* (Algorithm 2) that solves the *triangulation problem* [HZ03, §9.1] in the unlabeled case. For large numbers of pictures, this is much faster then going through all possible permutations of image point correspondences. Subsequently, the ambiguities of the unlabeled triangulation problem are analyzed.

In Chapter 6 we are concerned with reconstructing the *fundamental matrix* for a family of world point configurations. The most commonly used method to reconstruct the fundamental matrix is the *8-point algorithm* [HZ03, §11.2]. We say that an algorithm to reconstruct the fundamental matrix is *defeated by a world point configuration* in $(\mathbb{P}^3)^n$ if the algorithm fails to produce a unique fundamental matrix from the projections in $(\mathbb{P}^2)^n$ of that world point configuration, independent on the choice of cameras. It is known that the unit cube defeats the 8-point algorithm. We extend this result to all combinatorial cubes in $\mathbb{R}^3$ by using the Turnbull-Young invariant [TY26]. To understand the interplay between combinatorics and the 8-point algorithm we implemented several algorithms of numerical linear algebra in the open source software system `polymake` [GJ00]. Further we analyze when a reconstruction of the fundamental matrix is possible for two perspective projections of the vertices of a combinatorial cube. We then deduce a new algorithm (Algorithm 4) for this pathological situation. Algorithm 4 is based on projections onto rank deficient matrices by a singular value decomposition. Algorithms implemented for this chapter are being used in `polymake` [GJ00] in polyhedral computations.

Finally, in Chapter 7 we aim at understanding a complete intersection within the Veronese ideal. The second Veronese ideal I_n is generated by the 2-minors of a generic-symmetric $n \times n$ matrix. It contains, as a natural complete intersection, the ideal J_n which is generated by the principal 2-minors of a generic-symmetric $n \times n$ matrix. This ideal cuts out the second Veronese variety set-theoretically, and it was studied classically, e.g. by Gröbner [Grö65]. Both ideals I_n and J_n

are generated purely by binomials. In Theorem 7.11 sub-intersections of the primary decomposition of J_n, where one intersectand is omitted, are determined. Since the Veronese ideal itself is a prime component of J_n, computing these sub-intersections is a generalization of taking the colon ideal $I_n : J_n$. This is part of a bigger scheme. In Theorem 7.23 we determine the sub-intersection of the primary decomposition of general *Laurent binomial ideals* where one intersectand is omitted. Methods used in this chapter are at the interplay between combinatorics and algebra. In general binomial ideals have been used to construct various worst case scenarios of computer algebra. Only a few ideal-theoretic constructions actually preserve binomiality. Our combinatorial approach yields insight about this fact for ideal quotients of binomial ideals.

Declaration of Authorship

I, André WAGNER, declare that this thesis titled, "Computer Vision and Computer Algebra" and the work presented in it are my own or based on joint work of my own with co-authors. Parts of this thesis are prepublished.

- Chapter 3 is based on the joint paper *Rigid Multiview Varieties* [Jos+16] with Michael Joswig (TU Berlin), Joe Kileel (UC Berkeley) and Bernd Sturmfels (MPI Leipzig and UC Berkeley). Electronic version of an article published as *International Journal for Algebra and Computations, 26, Issue 04, 2016, 775,* doi:10.1142/S021819671650034X © World Scientific Publishing Company, http://www.worldscientific.com/worldscinet/ijac.

- Chapter 5 is based on the paper *Algebraic Relations and Triangulation of Unlabeled Image Points* [Wag17a]. A preprint is available at arXiv:1707.08722.

- Chapter 6 is based on the paper *Pictures of Combinatorial Cubes* [Wag17b]. A preprint is available at arXiv:1707.06563.

- Chapter 7 is based on the joint paper *Veronesean Almost Binomial Almost Complete Intersections* [KW16] with Thomas Kahle (OvGU Magdeburg). A preprint is available at arXiv:1608.03499.

To avoid redundancy we removed the general introductions to multiview geometry from the papers [Jos+16; Wag17a; Wag17b] in the Chapters 3, 5, 6, as these are covered in Chapter 2 of this thesis.

Chapter 2

Preliminaries

We use this chapter as a brief introduction to the mathematical concepts of this thesis. Further we aim to a establish a common notation that is usually kept throughout the dissertation. Two different mathematical fields are mainly covered, computer vision and binomial ideals. Their essential definitions and results for the following chapters are pithy introduced in the sections below. The statements presented here are well known in their communities.

2.1 Introduction to Multiview Geometry

In multiview geometry images of a three-dimensional scene are taken by several cameras from various perspectives. Algebraic vision is an emerging field of multiview geometry concerned with interactions between computer vision and algebraic geometry. A central role in this endeavor is played by projective varieties that arise in multiview geometry [HZ03]. We give a brief introduction to multiview geometry and usually stick to the notation established here throughout the dissertation for multiview geometry.

The set-up is as follows: A *camera* is a linear map from the three-dimensional projective space $\mathbb{P}^3$ to the projective plane $\mathbb{P}^2$, both over $\mathbb{R}$. We represent n

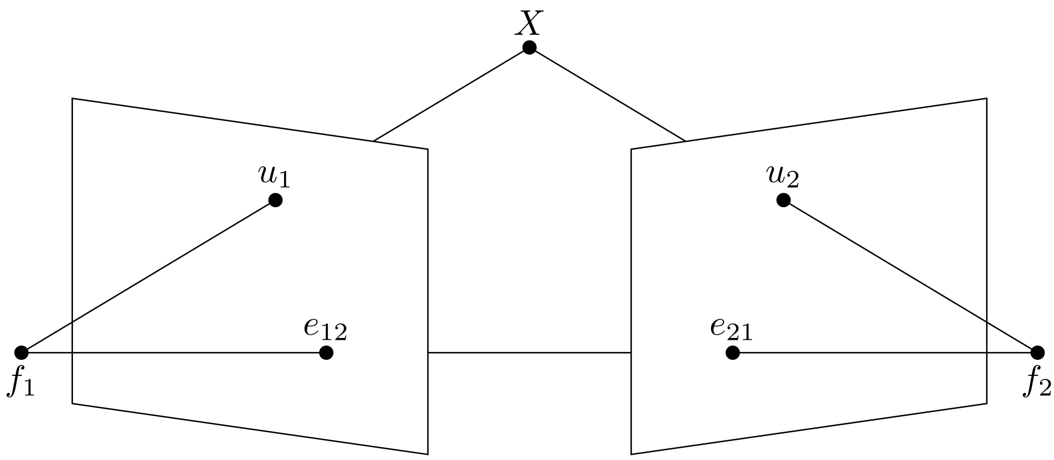


FIGURE 2.1: Two-view geometry

cameras by matrices $A_1, A_2, \ldots, A_n \in \mathbb{R}^{3\times 4}$ of rank 3. A *world point* $X \in \mathbb{P}^3$ is mapped by the *perspective relation* of the camera $A_j \in \mathbb{R}^{3\times 4}$

$$A_j X = \lambda_j u_j,\ \lambda_j \in \mathbb{R} \setminus \{0\}$$

to the *image point* $u_j \in \mathbb{P}^2$. The kernel of A_j is the *focal point* $f_j \in \mathbb{P}^3$. Each image point $u_j = (u_{j0}, u_{j1}, u_{j2}) \in \mathbb{P}^2$ of camera A_j has a line through f_j as its fiber in $\mathbb{P}^3$. This is the *back-projected line.* On that back-projected line lies the world point $X \in \mathbb{P}^3$. A camera is called *normalized* if it is of the form $[I \mid c]$ for some vector $c \in \mathbb{R}^3$, with $I \in \mathbb{R}^{3\times 3}$ being the identity matrix.

We assume throughout the thesis that the focal points of the n cameras are in *general position*, i.e. all distinct, no three on a line, and no four on a plane. Let β_{jk} denote the line in $\mathbb{P}^3$ spanned by the focal points f_j and f_k. This is the *baseline* of the camera pair A_j, A_k. The image of the focal point f_j in the image plane $\mathbb{P}^2$ of the camera A_k is the *epipole* e_{kj}. Note that the baseline β_{jk} is the back-projected line of e_{kj} with respect to A_j and also the back-projected line of e_{jk} with respect to A_k. See Figure 2.1 for a sketch.

For three cameras A_j, A_k, A_l the plane spanned by their focal points is called *trifocal plane.*

Fix a point X in $\mathbb{P}^3$ which is not on the baseline β_{jk}, and let u_j and u_k be the images of X under A_j and A_k. Since X is not on the baseline, neither image point is the epipole for the other camera. The two back-projected lines of u_j and u_k meet in a unique point, which is X. This process of reconstructing X from two images u_j and u_k is called *triangulation* [HZ03, §9.1].

The triangulation procedure amounts to solving the linear equations

$$B^{jk} \begin{bmatrix} X \\ -\lambda_j \\ -\lambda_k \end{bmatrix} = 0 \qquad \text{where} \quad B^{jk} = \begin{bmatrix} A_j & u_j & 0 \\ A_k & 0 & u_k \end{bmatrix} \in \mathbb{R}^{6\times 6}. \tag{2.1}$$

For general data we have $\operatorname{rank}(B^{jk}) = \operatorname{rank}(B_1^{jk}) = \cdots = \operatorname{rank}(B_6^{jk}) = 5$, where B_i^{jk} is obtained from B^{jk} by deleting the i-th row. *Cramer's Rule* can be used to recover X. Let $\wedge_5 B_i^{jk} \in \mathbb{R}^6$ be the column vector formed by the signed maximal minors of B_i^{jk}. Write $\widetilde{\wedge}_5 B_i^{jk} \in \mathbb{R}^4$ for the first four coordinates of $\wedge_5 B_i^{jk}$. These are bilinear functions of u_j and u_k. They yield

$$X = \widetilde{\wedge}_5 B_1^{jk} = \widetilde{\wedge}_5 B_2^{jk} = \cdots = \widetilde{\wedge}_5 B_6^{jk}. \tag{2.2}$$

There is a bilinear relation between image points if a reconstruction X is possible. We can derive this relation from the equation system of Equation 2.1. This equations system is solvable if a reconstruction X from u_j and u_k is possible. Hence $\det(B^{jk}) = 0$ must be satisfied. The equation $\det(B^{jk}) = 0$ is bilinear in the indeterminates u_j and u_k.

$$\det(B^{jk}) = 0 \Leftrightarrow u_k^T F u_j = 0, \text{ with } F \in \mathbb{R}^{3\times 3}$$

The matrix F is called *fundamental matrix.* F is only defined up to scale and its entries $F_{rs} \in \mathbb{R}$ can be computed from the camera matrices A_j, A_k as

$$F_{rs} = (-1)^{r+s} \det \begin{bmatrix} (A_j)_r \\ (A_k)_s \end{bmatrix} \in \mathbb{R},\ r, s \in [3],$$

where $(\cdot)_n$ is obtained by deleting the n-th row. Thus $(A_j)_n$ is a 2×4 matrix.

The triangulation can also be based on multiple views. Then B denotes the matrix constructed from all views

$$B \begin{bmatrix} X \\ -\lambda_1 \\ \vdots \\ -\lambda_n \end{bmatrix} = 0 \qquad \text{where} \quad B = \begin{bmatrix} A_1 & u_1 & 0 & \dots & 0 \\ A_2 & 0 & u_2 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A_n & 0 & \dots & 0 & u_n \end{bmatrix} \in \mathbb{R}^{3n\times(4+n)}. \tag{2.3}$$

In practice this equation system is surely not solved with Cramer's rule, however by a singular value decomposition and similar approaches [HS97]. We note that in most practical applications, the data $u_1, \dots, u_n$ will be noisy, in which case triangulation requires techniques from optimization, see [AAT12; KH05] to only state a few.

Based on multiple views we can derive multilinear relations, which the image points must satisfy in order for a reconstruction X being possible. Unlike in the two-view case these relations are not unique and are obtained from requiring that the maximal minors of B vanish. Let $\sigma \subseteq [n]$ with $|\sigma| = k$ and $A_\sigma = (A_{\sigma_1}^T, \dots, A_{\sigma_k}^T)$ then define the matrix B_σ as

$$B_\sigma = \begin{bmatrix} A_{\sigma_1} & u_{\sigma_1} & & \\ \vdots & & \ddots & \\ A_{\sigma_k} & & & u_{\sigma_k} \end{bmatrix}. \tag{2.4}$$

In that notation the matrix $B_{\{jk\}}$ aligns with the previously defined matrix B^{jk}. The multilinear equations are formed as maximal minors of the matrices B_σ. For more than four views the maximal minors of B_σ, $|\sigma| \geq 5$ are monomial multiples of the maximal minors of B_σ, $2 \leq |\sigma| \leq 4$ obtained from two, three and four views. Thus they do not contain any new information and it suffices to compute the maximal minors of the matrices B_σ with $2 \leq |\sigma| \leq 4$ to obtain all relevant multilinear relations.

The *multiview variety* V_A of the camera configuration $A = (A_1, \ldots, A_n)$ was defined in [AST13] as the closure of the image of the rational map

$$\begin{array}{rccc} \phi_A : & \mathbb{P}^3 & \dashrightarrow & \mathbb{P}^2 \times \mathbb{P}^2 \times \cdots \times \mathbb{P}^2, \\ & X & \mapsto & (A_1X, A_2X, \ldots, A_nX). \end{array} \tag{2.5}$$

The points $(u_1, u_2, \ldots, u_n) \in V_A$ are the consistent views in n cameras. The prime ideal I_A of V_A was determined in [AST13, Corollary 2.7]. It is generated by the $\binom{n}{2}$ bilinear polynomials $\det(B^{jk})$ plus $\binom{n}{3}$ further trilinear polynomials. A universal Gröbner basis of I_A is given by the maximal minors of the matrices B_σ with $2 \leq |\sigma| \leq 4$ [AST13, Theorem 2.1]. See [Li17] for the natural generalization of this variety to higher dimensions. As the multilinear relations can be expressed via maximal minors of matrices with certain entries being indeterminates, the multiview variety is a determinantal variety.

2.2 Introduction to Binomial Ideals

Chapter 7 is concerned with a special class of polynomial ideals called *binomial ideals.* These contain a rich combinatorial structure and faster algorithms are known to compute their generating sets and Gröbner bases. Many worst case scenarios in commutative algebra have been constructed using binomial ideals, e.g. the double exponential complexity for Gröbner bases [MM82] and the upper bound in the effective Nullstellensatz [Kol88]. Even the multiview variety of four and less views can be written up to a change of coordinate system as a binomial (toric) ideal. Binomial ideals have been extensively studied in [ES96].

By a *binomial* in a polynomial ring $\mathbb{K}[\mathbb{N}^n] := \mathbb{K}[x_1, \cdots, x_n]$ over the field $\mathbb{K}$ we mean a polynomial with at most two terms, say $ax^\alpha + bx^\beta$, where $a, b \in \mathbb{K}$ and $\alpha, \beta \in \mathbb{Z}^n_+$. A *binomial ideal* is defined as an ideal solely generated by binomials.

Example 2.1. *Consider the four camera matrices*

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

the multiview ideal is a binomial ideal and is generated by the binomials

$$\begin{aligned} I_A = \langle \quad & u_{11}u_{41} - u_{10}u_{42},\ u_{31}u_{40} - u_{30}u_{41},\ u_{21}u_{40} - u_{20}u_{42}, \\ & u_{12}u_{31} - u_{10}u_{32},\ u_{22}u_{30} - u_{20}u_{32},\ u_{12}u_{21} - u_{11}u_{22}, \\ & u_{21}u_{32}u_{41} - u_{22}u_{31}u_{42}, \quad u_{11}u_{32}u_{40} - u_{12}u_{30}u_{42}, \\ & u_{10}u_{22}u_{40} - u_{12}u_{20}u_{41}, \quad u_{10}u_{21}u_{30} - u_{11}u_{20}u_{31} \quad \rangle. \end{aligned}$$

Any multiview ideal for $n \leq 4$ cameras in linearly general position is isomorphic to the ideal of this example under a change of coordinate system in the image planes.

Many other famous varieties can be expressed via binomial ideals, so are the Segre, Veronese ideals determinantal ideals of two minors from a generic matrix, respectively generic-symmetric matrix.

Example 2.2. *The Veronese ideal* $I_{2,4} \subset \mathbb{Q}[a,b,c,d,e,f]$ *of degree two in four variables is generated by the* 2-minors *of the* generic-symmetric matrix

$$\begin{bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{bmatrix}.$$

There are only a few ideal-theoretic constructions that actually preserve binomiality, e.g. if I is a binomial ideal and p a binomial then the colon ideal $I : p$ usually is not a binomial ideal. In that context it almost seems miraculous that binomial ideals possess a minimal primary decomposition into binomial ideals [ES96, Theorem 2.1, Corollary 2.2].

It is easier first to study *Laurent binomial ideals.* Consider the ring $\mathbb{K}[\mathbb{Z}^n] = \mathbb{K}[x_1, \cdots, x_n, x_1^{-1}, \cdots, x_n^{-1}]$ of Laurent polynomials with coefficients over the field $\mathbb{K}$. Again a *Laurent binomial ideal* is an ideal solely generated by binomials in $\mathbb{K}[\mathbb{Z}^n]$. Any binomial $ax^\alpha + bx^\beta$ in the Laurent ring can be written as $x^m - c_m$ for some $m \in \mathbb{Z}^n$ and $c_m \in \mathbb{K}^*$. It is known that any Laurent binomial ideal can be constructed from a *partial character* ρ and a sublattice L_ρ of $\mathbb{Z}^n$. By partial character we mean a homomorphism ρ from the sublattice L_ρ of $\mathbb{Z}^n$ to the multiplicative group $\mathbb{K}^*$. The Laurent binomial ideal then is defined as

$$I(\rho) := \langle x^m - \rho(m) : m \in L_\rho \rangle.$$

If L is a sublattice of $\mathbb{Z}^n$, then the *saturation* of L is the lattice

$$\mathrm{sat}(L) := \{m \in \mathbb{Z}^n | dm \in L \text{ for some } d \in \mathbb{Z}\}.$$

A Laurent binomial ideal is prime if its lattice is *saturated.* Chapter 7 strongly relies on the following Theorem due to Eisenbud and Sturmfels [ES96, Theorem 2.1,d), Corrolary 2.2], which enables us to write down the prime decomposition of Laurent binomial ideals.

Theorem 2.3 (Eisenbud, Sturmfels 1996)**.** *Let* $\mathbb{K}[\mathbb{Z}^n]$ *be a Laurent ring over an algebraically closed field* $\mathbb{K}$ *of characteristic zero. Let* ρ *be a partial character on* $L_\rho \subseteq \mathbb{Z}^n$*. Write* g *for the order of* $\mathrm{sat}(L_\rho)/L_\rho$*. There are* g *distinct partial characters* $\rho_1, \ldots, \rho_g$ *of* $\mathrm{sat}(L_\rho)$ *extending* ρ*. The minimal prime decomposition of* $I(\rho) \subseteq \mathbb{K}[\mathbb{Z}^n]$ *is*

$$I(\rho) = \bigcap_{i=1}^{g} I(\rho_i).$$

With a little care and more formality Theorem 2.3 can be extended to more general fields. The result above can be transferred to binomial ideals. Let $m_+, m_- \in \mathbb{Z}^n_+$ denote the *positive part* and the *negative part* of a vector $m \in \mathbb{Z}^n$. Given a partial character ρ on $\mathbb{Z}^n$ we define the polynomial ideal

$$I_+(\rho) := \langle x^{m_+} - \rho(m)x^{m_-} : m \in L_\rho \rangle \in \mathbb{K}[\mathbb{N}^n].$$

Now if $I \in \mathbb{K}[\mathbb{N}^n]$ is a binomial ideal not containing any monomial, then there is a unique partial character ρ on the lattice $\mathbb{Z}^n$ such that the colon ideal $I : (x_1, \ldots, x_n)^\infty = I_+(\rho)$. The statement of Theorem 2.3 continues to hold if we replace $I(-)$ by $I_+(-)$ and prime by primary.

The most well studied class of binomial ideals are toric ideals. Due to their strong connection to regular triangulations of point configuration by the Sturmfels correspondence, they are of high interest to commutative algebra as to discrete geometry.

Consider the matrix $M \in \mathbb{R}^{m \times n}$, then the *toric ideal* I associated to the sublattice $L = \{m \in \ker(M), m \in \mathbb{Z}^n\}$ of $\mathbb{Z}^n$ is the binomial ideal

$$I := \langle x^{m_+} - x^{m_-} : m \in \ker(M), m \in \mathbb{Z}^n \rangle \in \mathbb{K}[\mathbb{N}^n].$$

The multiview ideal of four views can actually be written not just as a binomial ideal, but even as a toric ideal.

Example 2.1 (Continued). *Consider the matrix*

$$M = \begin{bmatrix} A_1^T & A_2^T & A_3^T & A_4^T \\ \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix},$$

where $\mathbf{1} = [1, 1, 1]$ *and* $\mathbf{0} = [0, 0, 0]$. *The multiview ideal of Example 2.1 is a toric ideal associated to the matrix* M.

Since a toric ideal I is generated by the integer points $m \in \mathbb{Z}^n$ in the kernel of M, it is frequently of interest to look at the *fibers*. For $b \in \mathbb{N}M$ the *fiber* of b with respect to M is defined as

$$M^{-1}[b] := \{\alpha \in \mathbb{N}^n : Mx = b\}.$$

Now if $x^\alpha - x^\beta \in I$, then $\alpha - \beta \in \ker(M)$ and $M\alpha = b$, $M\beta = b$. Therefore α and β are in the same fiber $M^{-1}[b]$. Consider the *monoid* $\mathbb{N}M := \{Mu : u \in \mathbb{Z}^n\}$. Then every binomial in I is homogeneous with respect to the multigrading $\mathbb{N}M$. The multidegree of the binomial $x^\alpha - x^\beta$ is $M\alpha = M\beta$.

Example 2.1 (Continued). *Consider the bilinear polynomial* $u_{12}u_{21} - u_{11}u_{22}$ *between camera one* A_1 *and two* A_2*. Its exponent vectors* m_+, m_- *and multidegree* b *are*

$$m_+ = (0,0,1,0,1,0,0,0,0,0,0,0), \quad m_- = (0,1,0,0,0,1,0,0,0,0,0,0),$$
$$b = (0,1,1,0,1,1,0,0) \qquad .$$

Many more interesting features of binomial ideals are known and we recommend reading [Stu96, §4] and [ES96] for a more detailed introduction.

Chapter 3

Rigid Multiview Variety

3.1 Introduction

The analysis of the multiview variety in [AST13] was restricted to a single world point $X \in \mathbb{P}^3$. In this chapter we study the case of two world points $X, Y \in \mathbb{P}^3$ that are linked by a distance constraint. Consider the hypersurface $V(Q)$ in $\mathbb{P}^3 \times \mathbb{P}^3$ defined by

$$Q = (X_0Y_3 - Y_0X_3)^2 + (X_1Y_3 - Y_1X_3)^2 + (X_2Y_3 - Y_2X_3)^2 - X_3^2Y_3^2. \quad (3.1)$$

The affine variety $V_{\mathbb{R}}(Q) \cap \{X_3{=}Y_3{=}1\}$ in $\mathbb{R}^3 \times \mathbb{R}^3$ consists of pairs of points whose Euclidean distance is 1.

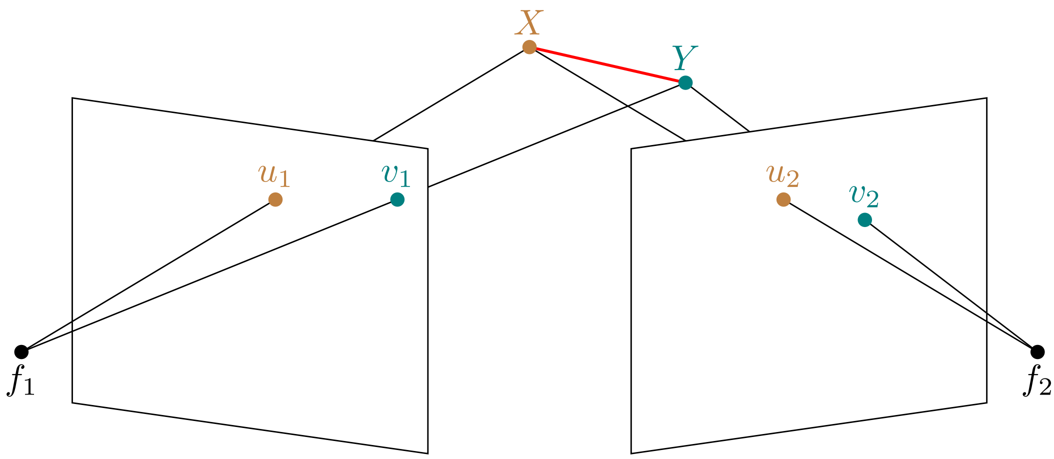


FIGURE 3.1: Two-view geometry of two rigid points

The *rigid multiview map* is the rational map

$$\begin{array}{rccccc} \psi_A : & V(Q) & \hookrightarrow & \mathbb{P}^3 \times \mathbb{P}^3 & \dashrightarrow & (\mathbb{P}^2)^n \times (\mathbb{P}^2)^n, \\ & (X, Y) & & & \mapsto & \big((A_1X, \ldots A_nX), (A_1Y, \ldots A_nY)\big). \end{array} \quad (3.2)$$

The *rigid multiview variety* is the image of this map. This is a 5-dimensional subvariety of $(\mathbb{P}^2)^{2n}$. Its multi-homogeneous prime ideal J_A lives in the polynomial ring $\mathbb{R}[u, v] = \mathbb{R}[u_{i0}, u_{i1}, u_{i2}, v_{i0}, v_{i1}, v_{i2} : i = 1, \ldots, n]$, where $(u_{i0}{:}u_{i1}{:}u_{i2})$

and $(v_{i0}{:}v_{i1}{:}v_{i2})$ are coordinates for the ith factor $\mathbb{P}^2$ on the left respectively right in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$. Our aim is to determine the ideal J_A. Knowing generators of J_A has the potential of being useful for designing optimization tools as in [AAT12] for triangulation in the presence of distance constraints.

The choice of world and image coordinates for the camera configuration $A = (A_1, \ldots, A_n)$ gives our problem the following group symmetries. Let N be an element of the *Euclidean group of motions* $\mathrm{SE}(3, \mathbb{R})$, which is generated by rotations and translations. We may multiply the camera configuration on the right by N to obtain $AN = (A_1N, \ldots, A_nN)$. Then $J_A = J_{AN}$ since $V(Q)$ is invariant under $\mathrm{SE}(3, \mathbb{R})$. For $M_1, \ldots, M_n \in \mathrm{GL}(3, \mathbb{R})$, we may multiply A on the left to obtain $A' = (M_1A, \ldots, M_nA)$. Then $J_{A'} = (M_1 \otimes \ldots \otimes M_n)J_A$.

The chapter is organized as follows. In Section 2 we present the explicit computation of the rigid multiview ideal for $n = 2, 3, 4$. Our main result, to be stated and proved in Section 3, is a system of equations that cuts out the rigid multiview variety $V(J_A)$ for any n. Section 4 is devoted to generalizations. The general idea is to replace $V(Q)$ by arbitrary subvarieties of $(\mathbb{P}^3)^m$ that represent polynomial constraints on $m \geq 2$ world points. We focus on scenarios that are of interest in applications to computer vision.

Our results in Propositions 3.1, 3.3, 3.4 and Corollary 3.2 are proved by computations with `Macaulay2` [GS]; for details see Section 3.5. Following standard practice in computational algebraic geometry, we carry out the computation on many samples in a Zariski dense set of parameters, and then conclude that it holds generically.

3.2 Two, Three and Four Cameras

In this section we offer a detailed case study of the rigid multiview variety when the number n of cameras is small. We begin with the case $n = 2$. The prime ideal J_A lives in the polynomial ring $\mathbb{R}[u, v]$ in 12 variables. This is the homogeneous coordinate ring of $(\mathbb{P}^2)^4$, so it is naturally $\mathbb{Z}^4$-graded. The variables u_{10}, u_{11}, u_{12} have degree $(1, 0, 0, 0)$, the variables u_{20}, u_{21}, u_{22} have degree $(0, 1, 0, 0)$, the variables v_{10}, v_{11}, v_{12} have degree $(0, 0, 1, 0)$, and the variables v_{20}, v_{21}, v_{22} have degree $(0, 0, 0, 1)$. Our ideal J_A is $\mathbb{Z}^4$-homogeneous.

Throughout this section we shall assume that the camera configuration A is *generic* in the sense of algebraic geometry. This means that A lies in the complement of a certain (unknown) proper algebraic subvariety in the affine space of all n-tuples of 3×4-matrices. All our results in Section 2 were obtained by symbolic computations with sufficiently many random choices of A (see Section 3.5 for details). Such choices of camera matrices are generic. They will be attained with with probability 1.

Proposition 3.1. *For $n = 2$, the rigid multiview ideal J_A is minimally generated by eleven $\mathbb{Z}^4$-homogeneous polynomials in twelve variables, one of degree* $(1, 1, 0, 0)$*, one of degree* $(0, 0, 1, 1)$*, and nine of degree* $(2, 2, 2, 2)$.

We prove this result by sufficiently many random computations with the computer algebra system `Macaulay2` [GS]. A slightly simplified version of the code is shown in Listing 3.1 in Section 3.5.

Let us look at the result in more detail. The first two bilinear generators are the familiar 6×6-determinants

$$\det \begin{bmatrix} A_1 & u_1 & 0 \\ A_2 & 0 & u_2 \end{bmatrix} \qquad \text{and} \qquad \det \begin{bmatrix} A_1 & v_1 & 0 \\ A_2 & 0 & v_2 \end{bmatrix}. \tag{3.3}$$

These cut out two copies of the multiview threefold $V_A \subset (\mathbb{P}^2)^2$, in separate variables, for $X \mapsto u = (u_1, u_2)$ and $Y \mapsto v = (v_1, v_2)$. If we write the two bilinear forms in (3.3) as $u_1^\top F u_2$ and $v_1^\top F v_2$ then F is a real 3×3-matrix of rank 2, known as the *fundamental matrix* [HZ03, §9] of the camera pair (A_1, A_2).

The rigid multiview variety $V(J_A)$ is a divisor in $V_A \times V_A \subset (\mathbb{P}^2)^2 \times (\mathbb{P}^2)^2$. The nine octics that cut out this divisor can be understood as follows. We write B and C for the 6×6-matrices in (3.3), and B_i and C_i for the matrices obtained by deleting their ith rows. The kernels of these 5×6-matrices are represented, via Cramer's Rule, by $\wedge_5 B_i$ and $\wedge_5 C_i$. We write $\widetilde{\wedge}_5 B_i$ and $\widetilde{\wedge}_5 C_i$ for the vectors given by their first four entries. As in (2.2), these represent the two world points X and Y in $\mathbb{P}^3$. Their coordinates are bilinear forms in (u_1, u_2) or (v_1, v_2), where each coefficient is a 3×3-minor of $\left[A_1^T, A_2^T\right]$. For instance, writing a_i^{jk} for the (j, k) entry of A_i, the first coordinate of $\widetilde{\wedge}_5 B_1$ is

$$\begin{aligned}
&-(a_1^{32}a_2^{23}a_2^{34} - a_1^{32}a_2^{24}a_2^{33} - a_1^{33}a_2^{22}a_2^{34} + a_1^{33}a_2^{24}a_2^{32} + a_1^{34}a_2^{22}a_2^{33} - a_1^{34}a_2^{23}a_2^{32})\, u_{11}u_{20} \\
&+(a_1^{32}a_2^{13}a_2^{34} - a_1^{32}a_2^{14}a_2^{33} - a_1^{33}a_2^{12}a_2^{34} + a_1^{33}a_2^{14}a_2^{32} + a_1^{34}a_2^{12}a_2^{33} - a_1^{34}a_2^{13}a_2^{32})\, u_{11}u_{21} \\
&-(a_1^{32}a_2^{13}a_2^{24} - a_1^{32}a_2^{14}a_2^{23} - a_1^{33}a_2^{12}a_2^{24} + a_1^{33}a_2^{14}a_2^{22} + a_1^{34}a_2^{12}a_2^{23} - a_1^{34}a_2^{13}a_2^{22})\, u_{11}u_{22} \\
&+(a_1^{22}a_2^{23}a_2^{34} - a_1^{22}a_2^{24}a_2^{33} - a_1^{23}a_2^{22}a_2^{34} + a_1^{23}a_2^{24}a_2^{32} + a_1^{24}a_2^{22}a_2^{33} - a_1^{24}a_2^{23}a_2^{32})\, u_{12}u_{20} \\
&-(a_1^{22}a_2^{13}a_2^{34} - a_1^{22}a_2^{14}a_2^{33} - a_1^{23}a_2^{12}a_2^{34} + a_1^{23}a_2^{14}a_2^{32} + a_1^{24}a_2^{12}a_2^{33} - a_1^{24}a_2^{13}a_2^{32})\, u_{12}u_{21} \\
&+(a_1^{22}a_2^{13}a_2^{24} - a_1^{22}a_2^{14}a_2^{23} - a_1^{23}a_2^{12}a_2^{24} + a_1^{23}a_2^{14}a_2^{22} + a_1^{24}a_2^{12}a_2^{23} - a_1^{24}a_2^{13}a_2^{22})\, u_{12}u_{22}.
\end{aligned}$$

Recall that the two world points in $\mathbb{P}^3$ are linked by a distance constraint (3.1), expressed as a biquadratic polynomial Q. We set $Q(X, Y) = T(X, X, Y, Y)$, where $T(\bullet, \bullet, \bullet, \bullet)$ is a quadrilinear form. We regard T as a tensor of order 4. It lives in the subspace $\mathrm{Sym}_2(\mathbb{R}^4) \otimes \mathrm{Sym}_2(\mathbb{R}^4) \simeq \mathbb{R}^{100}$ of $(\mathbb{R}^4)^{\otimes 4} \simeq \mathbb{R}^{256}$. Here $\mathrm{Sym}_k(\,\cdot\,)$ denotes the space of symmetric tensors of order k.

We now substitute our Cramer's Rule formulas for X and Y into the quadrilinear form T. For any choice of indices $1 \le i \le j \le 6$ and $1 \le k \le l \le 6$,

$$T\big(\,\widetilde{\wedge}_5 B_i\,,\ \widetilde{\wedge}_5 B_j\,,\ \widetilde{\wedge}_5 C_k\,,\ \widetilde{\wedge}_5 C_l\,\big) \tag{3.4}$$

is a multi-homogeneous polynomial in (u_1, u_2, v_1, v_2) of degree $(2,2,2,2)$. This polynomial lies in J_A but not in the ideal $I_A(u) + I_A(v)$ of $V_A \times V_A$, so it can serve as one of the nine minimal generators described in Proposition 3.1.

The number of distinct polynomials appearing in (3.4) equals $\binom{7}{2}^2 = 441$. A computation verifies that these polynomials span a real vector space of dimension 126. The image of that vector space modulo the degree $(2,2,2,2)$ component of the ideal $I_A(u) + I_A(v)$ has dimension 9.

We record three more features of the rigid multiview with $n = 2$ cameras. The first is the *multidegree* [MS05, §8.5], or, equivalently, the cohomology class of $V(J_A)$ in $H^*\big((\mathbb{P}^2)^4, \mathbb{Z}\big) = \mathbb{Z}[u_1, u_2, v_1, v_2]/\langle u_1^3, u_2^3, v_1^3, v_2^3\rangle$. It equals

$$2u_1^2v_1 + 2u_1u_2v_1 + 2u_2^2v_1 + 2u_1^2v_2 + 2u_1u_2v_2 + 2u_2^2v_2$$
$$+2u_1v_1^2 + 2u_1v_1v_2 + 2u_1v_2^2 + 2u_2v_1^2 + 2u_2v_1v_2 + 2u_2v_2^2.$$

This is found with the built-in command `multidegree` in `Macaulay2` [GS].

The second is the table of the Betti numbers of the minimal free resolution of J_A in the format of `Macaulay2` [GS]. In that format, the columns correspond to the syzygy modules, while rows denote the degrees. For $n = 2$ we obtain

```
        0  1  2  3 4 5
total:  1 11 25 22 8 1
    0:  1  .  .  . . .
    1:  .  2  .  . . .
    2:  .  .  1  . . .
    7:  .  9 24 22 8 1
```

The column labeled `1` lists the minimal generators from Proposition 3.1. Since the codimension of $V(J_A)$ is 3, the table shows that J_A is not Cohen-Macaulay. The unique 5th syzygy has degree $(3,3,3,3)$ in the $\mathbb{Z}^4$-grading.

The third point is an explicit choice for the nine generators of degree $(2,2,2,2)$ in Proposition 3.1. Namely, we take $i = j \leq 3$ and $k = l \leq 3$ in (3.4). The following corollary is also found by computation:

Corollary 3.2. *The rigid multiview ideal J_A for $n = 2$ is generated by $I_A(u) + I_A(v)$ together with the nine polynomials $Q\big(\widetilde{\wedge}_5 B_i, \widetilde{\wedge}_5 C_k\big)$ for $1 \leq i, k \leq 3$.*

We next come to the case of three cameras:

Proposition 3.3. *For $n = 3$, the rigid multiview ideal J_A is minimally generated by 177 polynomials in 18 variables. Its Betti table is given in Table 3.1.*

Proposition 3.3 is proved by computation. The 177 generators occur in eight symmetry classes of multidegrees. Their numbers in these classes are

(110000) : 1	(220111) : 3	(220220) : 9	(211211) : 1
(111000) : 1	(211111) : 1	(220211) : 3	(111111) : 1

	0	1	2	3	4	5	6	7	8	9	10	11
total:	1	177	1432	5128	10584	13951	12315	7410	3018	801	126	9
0:	1	.	.	.	.	.	.	.	.	.	.	.
1:	.	6	.	.	.	.	.	.	.	.	.	.
2:	.	2	21	6	.	.	.	.	.	.	.	.
3:	.	.	6	36	18	.	.	.	.	.	.	.
4:	.	.	1	12	42	36	9	.	.	.	.	.
5:	.	1	.	.	.	.	.	.	.	.	.	.
6:	.	24	108	166	120	42	6	.	.	.	.	.
7:	.	144	1296	4908	10404	13873	12300	7410	3018	801	126	9

TABLE 3.1: Betti numbers for the rigid multiview ideal with $n = 3$.

For instance, there are nine generators in degree $(2, 2, 0, 2, 2, 0)$, arising from Proposition 3.1 for the first two cameras. Using various pairs among the three cameras when forming the matrices B_i, B_j, C_k and C_l in (3.4), we can construct the generators of degree classes $(2, 2, 0, 2, 1, 1)$ and $(2, 1, 1, 2, 1, 1)$.

Table 3.1 shows the Betti table for J_A in `Macaulay2` format. The first two entries (`6` and `2`) in the `1`-column refer to the eight minimal generators of $I_A(u) + I_A(v)$. These are six bilinear forms, representing the three fundamental matrices, and two trilinear forms, representing the *trifocal tensor* of the three cameras (cf. [AO14], [HZ03, §15]). The entry `1` in row `5` of column `1` marks the unique sextic generator of J_A, which has $\mathbb{Z}^6$-degree $(1, 1, 1, 1, 1, 1)$.

For the case of four cameras we obtain the following result.

Proposition 3.4. *For $n = 4$, the rigid multiview ideal J_A is minimally generated by 1176 polynomials in 24 variables. All of them are induced from $n = 3$. Up to symmetry, the degrees of the generators in the $\mathbb{Z}^8$-grading are*

$$\begin{array}{llll} (11000000) : 1 & (22001110) : 3 & (22002200) : 9 & (21102110) : 1 \\ (11100000) : 1 & (21101110) : 1 & (22002110) : 3 & (11101110) : 1 \end{array}$$

We next give a brief explanation of how the rigid multiview ideals J_A were computed with `Macaulay2` [GS]. For the purpose of efficiency, we introduce projective coordinates for the image points and affine coordinates for the world points. We work in the corresponding polynomial ring

$$\mathbb{Q}[u, v][X_0, X_1, X_2, Y_0, Y_1, Y_2].$$

The rigid multiview map ψ_A is thus restricted to $\mathbb{R}^3 \times \mathbb{R}^3$. The prime ideal of its graph is generated by the following two classes of polynomials:

1. the 2×2 minors of the 3×2 matrices

$$\left[\ A_i \cdot (X_0, X_1, X_2, 1)^\top \ \middle|\ u_i\ \right], \quad \left[\ A_i \cdot (Y_0, Y_1, Y_2, 1)^\top \ \middle|\ v_i\ \right], \tag{3.5}$$

2. the dehomogenized distance constraint

$$Q\big((X_0, X_1, X_2, 1)^\top, (Y_0, Y_1, Y_2, 1)^\top\big).$$

From this ideal we eliminate the six world coordinates $\{X_0, X_1, X_2, Y_0, Y_1, Y_2\}$.

For a speed up, we exploit the group actions described in Section 3.1. We replace $A = (A_1, ..., A_n)$ and $Q = Q(X, Y)$ by $A' = (M_1A_1N, ..., M_nA_nN)$ and $Q' = Q(N^{-1}X, N^{-1}Y)$. Here $M_i \in \mathrm{GL}_3(\mathbb{R})$ and $N \in \mathrm{GL}_4(\mathbb{R})$ are chosen so that A' is sparse. The modification to Q is needed since we generally use $N \notin \mathrm{SE}(3, \mathbb{R})$. The elimination above now computes the ideal $(M_1 \otimes \ldots \otimes M_n)J_A$, and it terminates much faster. For example, for $n = 4$, the computation took two minutes for sparse A' and more than one hour for non-sparse A. For $n = 5$, `Macaulay2` ran out of memory after 18 hours of CPU time for non-sparse A. The complete code used in this chapter can be accessed via http://www3.math.tu-berlin.de/combi/dmg/data/rigidMulti/.

One last question is whether the Gröbner basis property in [AST13, §2] extends to the rigid case. This does not seem to be the case in general. Only in Proposition 3.1 can we choose minimal generators that form a Gröbner basis.

Remark 3.5. Let $n = 2$. The reduced Gröbner basis of J_A in the reverse lexicographic term order is a minimal generating set. For a generic choice of cameras the initial ideal equals

$$\begin{aligned} \mathrm{in}(J_A) \;=\; \langle\, & u_{10}u_{20},\ v_{10}v_{20},\ u_{10}^2u_{21}^2v_{10}^2v_{21}^2,\ u_{10}^2u_{21}^2v_{11}^2v_{20}v_{21},\ u_{10}^2u_{21}^2v_{11}^2v_{20}^2, \\ & u_{11}^2u_{20}^2v_{10}^2v_{21}^2,\ u_{11}^2u_{20}u_{21}v_{10}^2v_{21}^2,\ u_{11}^2u_{20}^2v_{11}^2v_{20}v_{21}, \\ & u_{11}^2u_{20}^2v_{11}^2v_{20}^2,\ u_{11}^2u_{20}u_{21}v_{11}^2v_{20}v_{21},\ u_{11}^2u_{20}u_{21}v_{11}^2v_{20}^2\,\rangle. \end{aligned}$$

For special cameras the exact form of the initial ideal may change. However, up to symmetry the degrees of the generators in the $\mathbb{Z}^4$-grading stay the same. In general, a universal Gröbner basis for the rigid multiview ideal J_A consists of octics of degree $(2, 2, 2, 2)$ plus the two quadrics (3.3). This was verified using the `Gfan`[Jen] package in `Macaulay2` [GS]. Analogous statements do not hold for $n \geq 3$.

3.3 Equations for the Rigid Multiview Variety

The computations presented in Section 2 suggest the following conjecture.

n\degree	2	3	6	7	8	total	timing (s)
2	2				9	11	< 1
3	6	2	1	24	144	177	14
4	12	8	16	240	900	1176	130
5	20	20	100	1200	3600	4940	24064

TABLE 3.2: The known minimal generators of the rigid multiview ideals, listed by total degree, for up to five cameras. There are no minimal generators of degrees 4 or 5. Average timings (in seconds), using the speed up described above, are in the last column.

Conjecture 3.6. *The rigid multiview ideal J_A is minimally generated by $\frac{4}{9}n^6 - \frac{2}{3}n^5 + \frac{1}{36}n^4 + \frac{1}{2}n^3 + \frac{1}{36}n^2 - \frac{1}{3}n$ polynomials. These polynomials come from two triples of cameras, and their number per class of degrees is*

$$
\begin{array}{llll}
(110..000..): & 1 \cdot 2\binom{n}{2} & (220..111..): & 3 \cdot 2\binom{n}{2}\binom{n}{3} \\
(220..220..): & 9 \cdot \binom{n}{2}^2 & (211..211..): & 1 \cdot n^2\binom{n-1}{2}^2 \\
(111..000..): & 1 \cdot 2\binom{n}{3} & (211..111..): & 1{\cdot}2n\binom{n-1}{2}\binom{n}{3} \\
(220..211..): & 3{\cdot}2n\binom{n}{2}\binom{n-1}{2} & (111..111..): & 1 \cdot \binom{n}{3}^2
\end{array}
$$

At the moment we have a computational proof only up to $n = 5$. Table 3.2 offers a summary of the corresponding numbers of generators.

Conjecture 3.6 implies that $V(J_A)$ is set-theoretically defined by the equations coming from triples of cameras. It turns out that, for the set-theoretic description, pairs of cameras suffice. The following is our main result:

Theorem 3.7. *Suppose that the n focal points of A are in general position in $\mathbb{P}^3$. The rigid multiview variety $V(J_A)$ is cut out as a subset of $V_A \times V_A$ by the $9\binom{n}{2}^2$ octic generators of degree class (220..220..). In other words, equations coming from any two pairs of cameras suffice set-theoretically.*

With notation as in the introduction, the relevant octic polynomials are

$$T\big(\,\widetilde{\wedge}_5 B_{i_1}^{j_1k_1}\,,\,\widetilde{\wedge}_5 B_{i_2}^{j_1k_1}, \widetilde{\wedge}_5 C_{i_3}^{j_2k_2}\,,\,\widetilde{\wedge}_5 C_{i_4}^{j_2k_2}\,\big),$$

for all possible choices of indices. Let H_A denote the ideal generated by these polynomials in $\mathbb{R}[u, v]$, the polynomial ring in $6n$ variables. As before, we write $I_A(u) + I_A(v)$ for the prime ideal that defines the 6-dimensional variety $V_A \times V_A$ in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$. It is generated by $2\binom{n}{2}$ bilinear forms and $2\binom{n}{3}$ trilinear forms, corresponding to fundamental matrices and trifocal tensors. In light of Hilbert's Nullstellensatz, Theorem 3.7 states that the radical of $H_A + I_A(u) + I_A(v)$ is equal to J_A. To prove this, we need a lemma.

A point u in the multiview variety $V_A \subset (\mathbb{P}^2)^n$ is *triangulable* if there exists a pair of indices (j, k) such that the matrix B^{jk} has rank 5. Equivalently, there exists a pair of cameras for which the unique world point X can be found by triangulation. Algebraically, this means $X = \widetilde{\wedge}_5 B_i^{jk}$ for some i.

Lemma 3.8. *All points in V_A are triangulable except for the pair of epipoles, (e_{12}, e_{21}), in the case where $n = 2$. Here, the rigid multiview variety $V(J_A)$ contains the threefolds $V_A(u) \times (e_{12}, e_{21})$ and $(e_{12}, e_{21}) \times V_A(v)$.*

Proof. Let us first consider the case of $n = 2$ cameras. The first claim holds because the back-projected lines of the two camera images u_1 and u_2 always span a plane in $\mathbb{P}^3$ except when $u_1 = e_{12}$ and $u_2 = e_{21}$. In that case both back-projected lines agree with the common baseline β_{12}. Alternatively, we can check algebraically that the variety defined by the 5×5-minors of the matrix B consists of the single point (e_{12}, e_{21}).

For the second claim, fix a generic point X in $\mathbb{P}^3$ and consider the surface

$$X^Q \;=\; \big\{ Y \in \mathbb{P}^3 \,:\, Q(X, Y) = 0 \big\}. \tag{3.6}$$

Working over $\mathbb{C}$, the baseline β_{12} is either tangent to X^Q, or it meets that quadric in exactly two points. Our assumption on the genericity of X implies that no point in the intersection $\beta_{12} \cap X^Q$ is a focal point. This gives

$$(A_1X, A_2X, A_1Y_X, A_2Y_X) \;=\; (A_1X, A_2X, e_{12}, e_{21}). \tag{3.7}$$

The point (A_1X, A_2X) lies in the multiview variety $V_A(u)$. Each generic point in $V_A(u)$ has this form for some X. Hence (3.7) proves the desired inclusion $V_A(u) \times (e_{12}, e_{21}) \subset V(J_A)$. The other inclusion $(e_{12}, e_{21}) \times V_A(v) \subset V(J_A)$ follows by switching the roles of u and v.

If there are more than two cameras, then for each world point X, due to general position of the cameras, there is a pair of cameras such that X avoids the pair's baseline. This shows that each point is triangulable if $n \geq 3$. □

Proof of Theorem 3.7. It follows immediately from the definition of the ideals in question that the following inclusion of varieties holds in $(\mathbb{P}^2)^n \times (\mathbb{P}^2)^n$:

$$V(J_A) \;\subseteq\; V\big(I_A(u) + I_A(v) + H_A\big).$$

We prove the reverse inclusion. Let (u, v) be a point in the right hand side.

Suppose that u and v are both triangulable. Then u has a unique preimage X in $\mathbb{P}^3$, determined by a single camera pair $\{A_{j_1}, A_{k_1}\}$. Likewise, v has a unique preimage Y in $\mathbb{P}^3$, also determined by a single camera pair $\{A_{j_2}, A_{k_2}\}$. There exist indices $i_1, i_2 \in \{1,2,3,4,5,6\}$ such that

$$X = \widetilde{\wedge}_5 B_{i_1}^{j_1 k_1} \quad \text{and} \quad Y = \widetilde{\wedge}_5 C_{i_2}^{j_2 k_2}.$$

Suppose that (u,v) is not in $V(J_A)$. Then $Q(X,Y) \neq 0$. This implies

$$Q(X,Y) \;=\; T(X,X,Y,Y) \;=\; T\big(\widetilde{\wedge}_5 B_{i_1}^{j_1 k_1}, \widetilde{\wedge}_5 B_{i_1}^{j_1 k_1}, \widetilde{\wedge}_5 C_{i_2}^{j_2 k_2}, \widetilde{\wedge}_5 C_{i_2}^{j_2 k_2}\big) \;\neq\; 0,$$

and hence $(u,v) \notin V(H_A)$. This is a contradiction to our choice of (u,v).

It remains to consider the case where v is not triangulable. By Lemma 3.8, we have $n=2$, as well as $v = (e_{12}, e_{21})$ and $(u,v) \in V(J_A)$. The case where u is not triangulable is symmetric, and this proves the theorem. □

The equations in Theorem 3.7 are fairly robust, in the sense that they work as well for many special position scenarios. However, when the cameras $A_1, A_2, \ldots, A_n$ are generic then the number $9\binom{n}{2}^2$ of octics that cut out the divisor $V(J_A)$ inside $V_A \times V_A$ can be reduced dramatically, namely to 16.

Corollary 3.9. *As a subset of the* 6*-dimensional ambient space $V_A \times V_A$, the* 5*-dimensional rigid multiview variety $V(J_A)$ is cut out by* 16 *polynomials of degree class* (220..220..)*. One choice of such polynomials is given by*

$$\begin{matrix} Q\big(\widetilde{\wedge}_5 B_i^{12}, \widetilde{\wedge}_5 C_k^{12}\big), & Q\big(\widetilde{\wedge}_5 B_i^{12}, \widetilde{\wedge}_5 C_k^{13}\big) \\ Q\big(\widetilde{\wedge}_5 B_i^{13}, \widetilde{\wedge}_5 C_k^{12}\big), & Q\big(\widetilde{\wedge}_5 B_i^{13}, \widetilde{\wedge}_5 C_k^{13}\big) \end{matrix} \qquad \textit{for all } 1 \leq i,k \leq 2.$$

Proof. First we claim that for each triangulable point u at least one of the matrices B^{12} or B^{13} has rank 5, and the same for v with C^{12} or C^{13}. We prove this by contradiction. By symmetry between u and v, we can assume that $rk(B^{12}) = rk(B^{13}) = 4$. Then $u_3 = e_{31}$, $u_2 = e_{21}$, and $u_1 = e_{12} = e_{13}$. However, this last equality of the two epipoles is a contradiction to the hypothesis that the focal points of the cameras A_1, A_2, A_3 are not collinear.

Next we claim that if B^{12} has rank 5, then at least one of the submatrices B_1^{12} or B_2^{12} has rank 5, and the same for B^{13}, C^{12} and C^{13}. Note that the bottom 4×6 submatrix of B^{12} has rank 4, since the first four columns are linearly independent, by genericity of A_1 and A_2. The claim follows. □

3.4 Other Constraints and More Points

In this section we discuss several extensions of our results. A first observation is that there was nothing special about the constraint Q in (3.1). For instance, fix positive integers d and e, and let $Q(X,Y)$ be any irreducible polynomial that is bihomogeneous of degree (d,e). Its variety $V(Q)$ is a hypersurface of degree

(d, e) in $\mathbb{P}^3 \times \mathbb{P}^3$. The following analogue to Theorem 3.7 holds, if we define the map ψ_A as in (3.2).

Theorem 3.10. *The closure of the image of the map ψ_A is cut out in $V_A \times V_A$ by $9\binom{n}{2}^2$ polynomials of degree class $(d, d, 0, \ldots, e, e, 0, \ldots)$. In other words, the equations coming from any two pairs of cameras suffice set-theoretically.*

Proof. The tensor T that represents Q now lives in $\mathrm{Sym}_d(\mathbb{R}^4) \otimes \mathrm{Sym}_e(\mathbb{R}^4)$. The polynomial (3.4) vanishes on the image of ψ_A and has degree (d, d, e, e). The proof of Theorem 3.7 remains valid. The surface X^Q in (3.6) is irreducible of degree e in $\mathbb{P}^3$. These polynomials cut out that image inside $V_A \times V_A$. □

Remark 3.11. In the generic case, we can replace $9\binom{n}{2}^2$ by 16, as in Corollary 3.9.

Another natural generalization is to consider m world points $X_1, \ldots, X_m$ that are linked by one or several constraints in $(\mathbb{P}^3)^m$. Taking images with n cameras, we obtain a variety $V(J_A)$ which lives in $(\mathbb{P}^2)^{mn}$. For instance, if $m = 4$ and X_1, X_2, X_3, X_4 are constrained to lie on a plane in $\mathbb{P}^3$, then $Q = \det(X_1, X_2, X_3, X_4)$ and $V(J_A)$ is a variety of dimension 11 in $(\mathbb{P}^2)^{4n}$. Taking 6×6-matrices B, C, D, E as in (2.1) for the four points, we then form

$$\det\big(\widetilde{\wedge}_5 B_i, \widetilde{\wedge}_5 C_j, \widetilde{\wedge}_5 D_k, \widetilde{\wedge}_5 E_l\big) \qquad \text{for all } 1 \leq i, j, k, l \leq 6. \tag{3.8}$$

For $n = 2$ we verified with `Macaulay2` that the prime ideal J_A is generated by 16 of these determinants, along with the four bilinear forms for ${V_A}^4$.

Proposition 3.12. *The variety $V(J_A)$ is cut out in ${V_A}^4$ by the $16\binom{n}{2}^4$ polynomials from (3.8). In other words, the equations coming from any two pairs of cameras suffice set-theoretically.*

Proof. Each polynomial (3.8) is in J_A. The proof of Theorem 3.7 remains valid. The planes $(X_i, X_j, X_k)^Q$ intersect the baseline β_{12} in one point each. □

To continue the theme of rigidity, we may impose distance constraints on pairs of points. Fixing a nonzero distance d_{ij} between points i and j gives

$$Q_{ij} = (X_{i0}X_{j3} - X_{j0}X_{i3})^2 + (X_{i1}X_{j3} - X_{j1}X_{i3})^2 + (X_{i2}X_{j3} - X_{j2}X_{i3})^2 - d_{ij}^2 X_{i3}^2 X_{j3}^2.$$

We are interested in the image of the variety $\mathcal{V} = V(Q_{ij} : 1 \leq i < j \leq m)$ under the multiview map ψ_A that takes $(\mathbb{P}^3)^m$ to $(\mathbb{P}^2)^{mn}$. For instance, for $m = 3$, we consider the variety $\mathcal{V} = V(Q_{12}, Q_{13}, Q_{23})$ in $(\mathbb{P}^3)^3$, and we seek the equations for its image under the multiview map ψ_A into $(\mathbb{P}^2)^{3n}$. Note that $\mathcal{V}$ has dimension 6, unless we are in the collinear case. Algebraically,

$$(d_{12} + d_{13} + d_{23})(d_{12} + d_{13} - d_{23})(d_{12} - d_{13} + d_{23})(-d_{12} + d_{13} + d_{23}) \;=\; 0. \tag{3.9}$$

If this holds then $\dim(\mathcal{V}) = 5$. The same argument as in Theorem 3.7 yields:

Corollary 3.13. *The rigid multiview variety $\overline{\psi_A(\mathcal{V})}$ has dimension six, unless (3.9) holds, in which case the dimension is five. It has real points if and only if d_{12}, d_{13}, d_{23} satisfy the triangle inequality. It is cut out in ${V_A}^3$ by $27\binom{n}{2}^2$ biquadratic equations, coming from the $9\binom{n}{2}^2$ equations for any two of the three points.*

3.5 Computations

We performed several random experiments in this Chapter. Our hardware was a cluster with Intel Xeon X2630v2 Hexa-Cores (2.8 GHz) and 64GB main memory per node. The software was `Macaulay2` [GS], version 1.8.2.1. All computations were single-threaded.

The tests were repeated several times with random input. The exact running times vary, even with identical input; the Table 3.2 lists the average values. It is not surprising that increasing n, the number of cameras, increases the running times considerably. Therefore we adapted the number of experiments according to n.

For all the statements in Section 3.3 regarding two cameras, (i.e. $n = 2$) we performed at least 1000 computations, with one exception. The statement regarding the universal Gröbner basis in Remark 3.5 is based on 20 experiments. Regarding three and four cameras (i.e., $n \in \{3, 4\}$) we performed at least 100 computations each. For $n = 5$ we performed at least 20 computations each.

In Listing 3.1 we show `Macaulay2` code which can be employed to establish Proposition 3.1. The complete code for all our results can be accessed via http://www3.math.tu-berlin.de/combi/dmg/data/rigidMulti/.

Lines 1–4 define the rings in which the computations take place. Lines 6–7 produce random camera matrices. Here the code shown differs slightly from the code used. Lines 7 – 13 produces random generic cameras by the definition of generic given in [AST13, §2]. However, our experiments suggest that it suffices to check that the focal points of the cameras are in linearly general position. The multiview map ϕ_A from (2.5) is encoded in lines 17–21. Line 13 is the rigid constraint (3.1). The actual computation is the elimination in line 23. The rigid multiview ideal J_A is defined in lines 25–26, and the final output are the multidegrees of J_A.

In these computations the world coordinates are dehomogenized by setting the last coordinate to 1, as explained at the end of Section 3.2. Notice that the code below line 7 does not need to be modified if we increase n.

```
R1 = QQ[u_(1,0)..u_(1,2)] ** QQ[u_(2,0)..u_(2,2)] **
     QQ[v_(1,0)..v_(1,2)] ** QQ[v_(2,0)..v_(2,2)];
R2 = QQ[X_0..X_2] ** QQ[Y_0..Y_2];
S  = R1 ** R2;

n  = 2;
AList=0
while (numgens minors(4,transpose matrix AList)=!=binomial(n*3,4))do(
   AList={};
   for i from 1 to n do(
      A_i=random(ZZ^3,ZZ^4,Height=>20);
      AList=AList| entries A_i; )
   );

I = ideal();
for j from 1 to n do (
   I = I + minors(2,A_j * (genericMatrix(S,X_0,3,1)||matrix{{1}})|
       genericMatrix(S,u_(j,0),3,1));
   I = I + minors(2,A_j * (genericMatrix(S,Y_0,3,1)||matrix{{1}})|
       genericMatrix(S,v_(j,0),3,1)); );
I = I + ideal((X_0-Y_0)^2 + (X_1-Y_1)^2 + (X_2-Y_2)^2-1);
I = eliminate({X_0,X_1,X_2,Y_0,Y_1,Y_2},I);

F = map(R1,S);
J = F(I);

degrees(J)
```

Listing 3.1: Compute J_A for two cameras

Chapter 4

Triangulation via the Rigid Multiview Variety

4.1 Introduction

There are many different optimization formulations known to solve the triangulation problem with presence of noise, these are essential to computer vision. Most of the algorithms use local optimization to reconstruct the world point. However a few of them also give global optimality certificates by using semidefinite programming [AAT12; KH05]. The rigid multiview variety can be understood as a generalization of the multiview variety. A natural question in that context is whether one can use the rigid multiview variety $V(J_A)$ to formulate an algorithm to improve the quality of the reconstruction by additional enforcing the rigidity constraints.

Let $\hat{u} \in (\mathbb{R}^2)^n$ be a vector of given noisy data. This is one world point $X \in \mathbb{R}^3$ seen through n cameras. We introduce a vector of indeterminates $u \in (\mathbb{R}^2)^n$ of the same length $2n$. We write $u_i = (u_i^{(1)}, u_i^{(2)})$ corresponding to the two coordinates in the i-th copy of $\mathbb{R}^2$, and we write $\tilde{u}$ for the homogenization which describes a point in the space $(\mathbb{P}^2)^n$. More precisely, we choose a chart such that $\tilde{u}_i = (u_i, 1)$ for all i. Similarly, we have $\hat{v}$, v, v_i and $\tilde{v}$ for the same kind of information derived from a second world point $Y \in \mathbb{R}^3$. Finally, let $\widetilde{X}, \widetilde{Y}$ be the homogenizations of X, Y and δ be the Euclidean distance between X, Y.

For noisefree points on the rigid multiview variety, we can use the classical triangulation approach to reconstruct the world points. For two views this can be done by solving Equation 2.1. In practice with multiple views this is usually done by a singular value decomposition of the matrix B from Equation 2.3. For noisy points the triangulation problem aims to find the maximum likelihood estimate of X for given noisy data. We try to extend this problem to the rigid multiview variety. Then the *rigid triangulation problem* is to find the maximum likelihood estimate of u, v given the distance δ between X, Y and noisy data $\hat{u}, \hat{v}$. We propose the rigid triangulation problem:

$$\min_{(u,v)} \sum_i \left(\|u_i - \hat{u}_i\|^2 + \|v_i - \hat{v}_i\|^2\right) : A_i\widetilde{X} = \lambda_i \tilde{u}_i,\ A_i\widetilde{Y} = \mu_i \tilde{v}_i\,, \|X - Y\|^2 = \delta^2. \quad (4.1)$$

The optimization problem aims to be as close as possible in its formulation to the one of the classical triangulation problem in [AAT12], however it is able to account additional information.

4.2 Projecting Onto the Rigid Multiview Variety

We can use the rigid multiview variety $V(J_A)$ to eliminate the variables X and Y from Equation 4.1.

$$\arg\min_{(u,v)} \sum_i \left(\|u_i - \hat{u}_i\|^2 + \|v_i - \hat{v}_i\|^2\right) : (\tilde{u}, \tilde{v}) \in V(J_A)\ . \quad (4.2)$$

A solution of Equation 4.2 is a point $(\tilde{u}, \tilde{v})$ on the rigid multiview variety $V(J_A)$ which is closest to given data $(\hat{u}, \hat{v}) \in (\mathbb{R}^2)^n \times (\mathbb{R}^2)^n$. This means that we are orthogonally projecting the points of $(\mathbb{P}^2)^n$ onto the rigid multiview variety. After projecting onto $V(J_A)$ one can simply use a singular value decomposition to reconstruct X, Y [HS97], because the rigid multiview variety $V(J_A)$ is a subset of the multiview variety V_A.

A set-theoretical description of W_A is given in Theorem 3.7. There are two types of constraints. First the bilinear ones and second the octics of Equation 3.4. Hence the optimization problem (4.2) can be rewritten as

$$\arg\min_{(u,v)} \sum_i \left(\|u_i - \hat{u}_i\|^2 + \|v_i - \hat{v}_i\|^2\right) : \quad \begin{array}{c} \tilde{u}_j F_{ij} \tilde{u}_i = 0,\ \forall i, j \\ \tilde{v}_j F_{ij} \tilde{v}_i = 0,\ \forall i, j \\ Q(\widetilde{\wedge}_5 B_i^{jk}, \widetilde{\wedge}_5 C_i^{jk}) = \delta,\ \forall i, j, k \end{array}$$

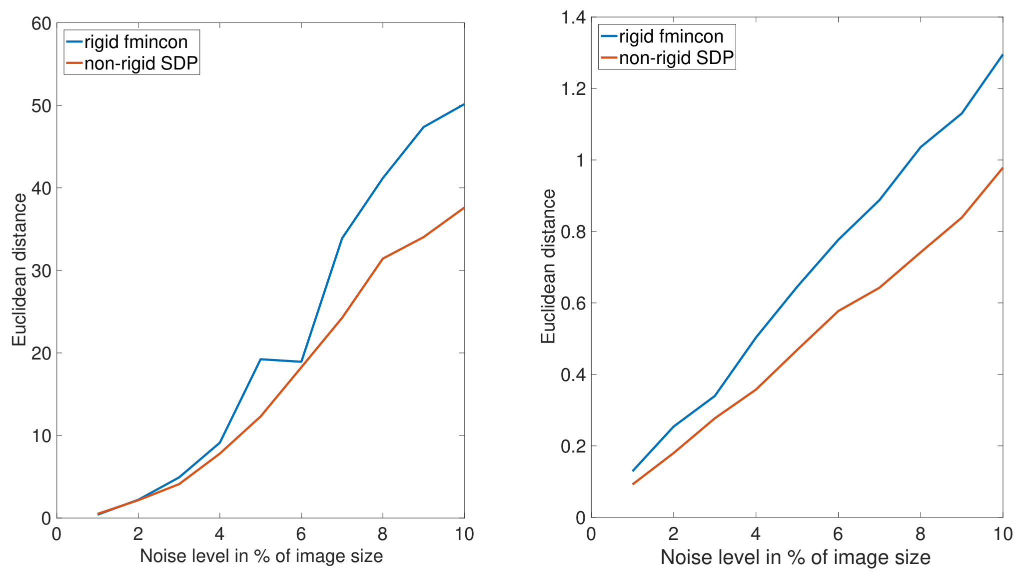
Not all octics are needed for a set-theoretical description, by Corollary 3.9 in total only 16 octics suffice. For two views even four octics are sufficient.

4.3 Computational Experiments

We used the formulations of Equation 4.1 and Equation 4.2 as input to different mathematical optimization solvers and software, like `SeDuMi` [Stu99], `Mosek` [ApS15], `fmincon` [MAT16], `SCIP` [Gam+16], `GloptiPoly 3` [HLL09], `YALMIP` [Lof05]. We aimed to employ these solvers only with out of the box methods.

The data was taken from the `Model House` data set. It consists of ten different camera positions and 672 world points. We applied additional Gaussian noise onto the images of the data set with standard deviation between $0\% - 10\%$

of the image sizes and zero mean. Since the algorithms used are sensitive to scaling of the fundamental matrix, we divide the fundamental matrices by the largest singular value. Using the local nonlinear constraint solver `fmincon` of `MATLAB`, for $n = 2$ cameras the solver mostly converges and gives compatible, yet slightly inferior results to the SDP approach of [AAT12]. However, adding the rigidity constraints seems not to improve the quality of the reconstruction and often fails for instances whenever $n \geq 3$.



(A) Image data approximation. (B) World data approximation.

FIGURE 4.1: Comparison between our rigid triangulation approach with `fmincon` and non-rigid triangulation `SDP` approach of [AAT12]. y-axis measures the distance between data and approximation.

For global optimization we worked with `YALMIP` and `GloptiPoly 3` accessing the solvers `SeDuMi` and `Mosek`. There we used the `solvemoment` and `solvesos` methods of `YALMIP`. Both of them failed: `solvemoment` got stuck in `YALMIP`'s preprocessing routine, and `SeDuMi` did not even get started. The method `solvesos` usually did not converge or got stuck in `YALMIP`'s preprocessing routine. The method `msol` of `GloptiPoly 3` using the fourth moment relaxation did not converge within a day of computing time.

We also employed `SCIP` with the global nonlinear constraint solver `IPOPT`. The rigid multiview variety $V(J_A)$ is the topological closure of the set

$$S_A := \{(u,v) \in (\mathbb{P}^2 \times \mathbb{P}^2)^n : A_i X = \lambda_i u_i,\ A_i Y = \mu_i v_i,\ \|X - Y\|^2 = \delta^2\}, \quad (4.3)$$

where $X, Y \in \mathbb{P}^3$ and $\lambda_i, \mu_i \in \mathbb{R}\backslash\{0\}$. By replacing $V(J_A)$ with S_A in Equation 4.2 noisy image points are projected onto S_A instead onto $V(J_A)$. We obtain the optimization problem

$$\min_{(u,v)} \sum_i \left(\|u_i - \hat{u}_i\|^2 + \|v_i - \hat{v}_i\|^2\right) : A_i\tilde{X} = \lambda_i \tilde{u}_i,\ A_i\tilde{Y} = \mu_i \tilde{v}_i\,,\ \|X - Y\|^2 = \delta^2. \quad (4.4)$$

This formulation has the advantage that we indirectly avoid the octics, whose high degree leads to severe numerical instabilities. We can eliminate λ_i, μ_i from these equations. Since $\tilde{X}$ and $\tilde{Y}$ are embedded into $\mathbb{P}^3$ at height one, the perspective relations $A_i\tilde{X} = \lambda_i\tilde{u}_i$ is equivalent to $A_i\tilde{X}||\tilde{u}_i$. Hence we obtain $A_i\tilde{X} \times \tilde{u}_i = 0$. This is the same approach described in Equation 3.5. `SCIP` could find within 30 seconds an optimality gap of 100%, however it failed to converge. `GloptiPoly 3` could not certifiy global optimality for computable moment relaxations.

4.4 Conclusion

The way we constructed the set-theoretical equations of the multiview variety strongly depends on Cramer's rule to triangulate the original world point coordinates. We then insert these into the Euclidean distance constraint. In practice Cramer's rule is a bad choice to solve linear equation system and is numerically instable. Other methods like singular value decomposition are used to obtain solutions. However, we are bound to using Cramer's rule to obtain an implicit description of the rigid multiview variety $V(J_A)$. Thus the octics of Equation 3.4 inherit this numerical instability and even amplify it. This results in an instable optimization problem, which is hard to handle. Clearly the high degree of the constraints makes it extremely difficult for solvers to obtain good results. One could use different methods instead of sums of squares to certify non-negativity. A possible choice would be circuit polynomials in combination with geometric programming [DIW16]. These are conjectured not to depend that much on the total degree of the involved polynomials. We are currently working on such an approach. Anyhow, unluckily the rigid multiview variety $V(J_A)$ does not inherit many of the algebraic properties of the multiview variety V_A. The rigid constraints actually destroy many of the nice features of the multiview variety V_A. For example $V(J_A)$ is not generated by multilinear constraints nor is it Cohen-Macaulay. At first sight it might seem surprising that using first moment relaxations to project onto the multiview variety V_A works as done in [AAT12], but probably this goes back to the multiview variety being very well behaved.

4.5 Code

The code in Listings 4.1 describes our approach to project onto the rigid multiview variety $V(J_A)$. We used `Matlab 2016a` and `GloptiPoly 3.8`. The code usually terminated. However, the SDP relaxation used in the computation where not tight and thus a solution could not be extracted due to an incomplete monomial basis.

The code gets the data input $\hat{u}_i \to$ `u1mes`, $A_i \to$ `A1`, $\delta \to$ `dist` as noted in lines 1-3. In lines 6-9 the variables $u_i \to$ `ui` of the optimization problem and their embeddings $\tilde{u} \to$ `uihom` into projective space $\mathbb{P}^2$ are defined. Lines 12-16 construct the set-theoretical constraints of $V(J_A)$. Only four of the 36 octics are chosen as one can read of from lines 29 and 31. Line 19 defines the objective function. Finally, lines 22-24 set up the SDP. Appended in lines 26-53 are the custom user functions used in this code snippet.

```
% Input: Image data:          u1mes, u2mes, v1mes, v2mes
%        cameras:             A1, A2
%        distance (delta^2):  dist

% variables
mpol u1 2; u1hom=[u1;1];
mpol u2 2; u2hom=[u2;1];
mpol v1 2; v1hom=[v1;1];
mpol v2 2; v2hom=[v2;1];

% the two bilinear equations
BilinU=u2hom'*F*u1hom;
BilinV=v2hom'*F*v1hom;

% rigid constraints
rigid=rigidConstraint(u1,u2,v1,v2,A1,A2,dist);

% distance between variables and data
edist=eucdist([u1;u2;v1;v2]',[u1mes;u2mes;v1mes;v2mes]');

% Gloptipoly 3 input
K=[rigid==0, BilinU==0, BilinV==0];
P = msdp(min(edist), K);
[status,obj] = msol(P);

%%%%%%%%%%%%%%%%%%%%%
function [ceq] = rigidConstraint(u1,u2,v1,v2,A1,A2,dist)
ceq=[];
for i=1:2 % only 4 octics are chosen
    X=triangulationViaCramersRule(u1,u2,A1,A2,i);
    for j=1:2
        Y=triangulationViaCramersRule(v1,v2,A1,A2,j);
        ceq=[ceq rigidPolynomial(X,Y,dist)];
    end
end
%%%%%%%%%%%%%%%%%%%
function X = triangulationViaCramersRule(x,y,A1,A2,i)
xzeros=[x(1);x(2);1;0;0;0];
yzeros=[0;0;0;y(1);y(2);1];
B=[vertcat(A1,A2) xzeros yzeros];
X=[-det(B([1:i-1,i+1:end],2:6)),
    det(B([1:i-1,i+1:end],[1 3 4 5 6])),
   -det(B([1:i-1,i+1:end],[1 2 4 5 6])),
    det(B([1:i-1,i+1:end],[1 2 3 5 6]))];
%%%%%%%%%%%%%%%%%%%
function Q = rigidPolynomial(X,Y,dist)
Q= (X(1)*Y(4)-Y(1)*X(4))^2+(X(2)*Y(4)-Y(2)*X(4))^2+
   (X(3)*Y(4)-Y(3)*X(4))^2-dist^2*X(4)^2*Y(4)^2;
%%%%%%%%%%%%%%%%%%%
function f = eucdist(x,xMeasurement)
x=reshape(x,2,[]);
xMeasurement=reshape(xMeasurement,2,[]);
diffMat=x-xMeasurement;
diffMat=reshape(diffMat,1,[]);
f=sum(sum(diffMat.^2));
```

LISTING 4.1: Projecting onto $V(J_A)$.

Chapter 5

Relations and Triangulation of Unlabeled Image Points

5.1 Introduction

In many computer vision applications, the correspondences among views are unknown. Hence the m world points in $\mathbb{P}^3$ and their images points in $\mathbb{P}^2$ will be unlabeled. To study such questions, we propose to work with the *unlabeled multiview variety.* This is the variety of products of *multiview varieties* [AST13] with unknown correspondences. An unlabeled point configuration in $(\mathbb{P}^2)^m$ is a point in the *Chow variety* $\mathrm{Sym}_m(\mathbb{P}^2)$ [Lan, §8.6]. Algebraically the *unlabeled multiview variety* is the image of the multiview variety under the quotient map $\big((\mathbb{P}^2)^m\big)^n \to \big(\mathrm{Sym}_m(\mathbb{P}^2)\big)^n$ for the symmetric group $S(m)$ action. Our focus is mostly on two unlabeled points. We design Algorithm 2 to triangulate two unlabeled points.

While labeled world and image configurations are points in $(\mathbb{P}^2)^m$ and $(\mathbb{P}^2)^m$, unlabeled image configurations are points in the *Chow varieties* which live as subvarieties in $\mathrm{Sym}_m(\mathbb{P}^2)$ and $\mathrm{Sym}_m(\mathbb{P}^3)$. This is the variety of ternary forms that are products of m linear forms (cf. [Lan, §8.6]), respectively quaternary forms that are products of m linear forms. It is embedded in the space $\mathbb{P}^{\binom{m+2}{m}-1}$ of all ternary forms of degree m.

We define the unlabeled multiview variety to be the closure of the image of the rational map ζ_A

$$\zeta_A : \mathrm{Sym}_m(\mathbb{P}^3) \dashrightarrow \big(\mathrm{Sym}_m(\mathbb{P}^2)\big)^n.$$

Then ζ_A maps an order m symmetric $4 \times \ldots \times 4$ tensor to n order m symmetric $3 \times \ldots \times 3$ tensors. For each world point these tensors have one axis. This gives rise to the *unlabeled multiview variety* $\mathrm{Sym}_m(V_A)$ in $\big(\mathbb{P}^{\binom{m+2}{m}-1}\big)^n$, where V_A is the multiview variety. Let $X, Y \in \mathbb{P}^3$ be two labeled world points. We denote their images in the i-th picture as u_i, v_i,

$$A_iX = \lambda_i u_i,\ A_iY = \mu_i v_i,\ \lambda_i, \mu_i \in \mathbb{R} \setminus \{0\}.$$

Then $u_i = (u_{i0}, u_{i1}, u_{i2})$ and $v_i = (v_{i0}, v_{i1}, v_{i2})$ are the coordinates of the image points.

Example 5.1. *Let $m = n = 2$. The Chow variety $\mathrm{Sym}_2(\mathbb{P}^2)$ is the hypersurface in $\mathbb{P}^5$ defined by the determinant of a symmetric 3×3-matrix $N = (N^{ij})_{i,j\in[3]}$. From the two image points u_i, v_i we can construct an unlabeled image point on the Chow variety $\mathrm{Sym}_m(\mathbb{P}^2)$, this is the symmetric matrix $N = u_i v_i^T + v_i u_i^T$. Up to the symmetry of coordinates of N, this reads as*

$$\begin{array}{ccc} N_1^{00} = 2u_{10}v_{10}, & N_1^{11} = 2u_{11}v_{11}, & N_1^{22} = 2u_{12}v_{12}, \\ N_1^{01} = u_{11}v_{10} + u_{10}v_{11}, & N_1^{02} = u_{12}v_{10} + u_{10}v_{12}, & N_1^{12} = u_{12}v_{11} + u_{11}v_{12}. \end{array}$$

The quotient map $(\mathbb{P}^2)^2 \to \mathrm{Sym}_2(\mathbb{P}^2) \subset \mathbb{P}^5$ is given by the formulas above. Similarly, for the two unlabeled images of the second camera we use

$$\begin{array}{ccc} N_2^{00} = 2u_{20}v_{20}, & N_2^{11} = 2u_{21}v_{21}, & N_2^{22} = 2u_{22}v_{22}, \\ N_2^{01} = u_{21}v_{20} + u_{20}v_{21}, & N_2^{02} = u_{22}v_{20} + u_{20}v_{22}, & N_2^{12} = u_{22}v_{21} + u_{21}v_{22}. \end{array}$$

We compute the image of $V_A \times V_A$ in $\mathbb{P}^5 \times \mathbb{P}^5$, denoted $\mathrm{Sym}_2(V_A)$. Its ideal has seven minimal generators, three of degree $(1,1)$, and one each in degrees $(3,0), (2,1), (1,2), (0,3)$. The generators in degrees $(3,0)$ and $(0,3)$ are $\det(N_1)$ and $\det(N_2)$. The five others depend on the cameras A_1, A_2.

If $m = 2$ we can use symmetric matrices in $\mathrm{Sym}_2(\mathbb{R}^4)$ and $\mathrm{Sym}_2(\mathbb{R}^3)$ to describe the unlabeled multiview variety, even for n larger than two. This is computationally easier to handle. Then the unlabeled world point configuration of the two labeled world points $X, Y \in \mathbb{P}^3$ can be represented by a symmetric 4×4 rank two matrix $M = XY^T + YX^T$. The unlabeled image point configuration of two labeled image points in the i-th picture $u_i, v_i \in \mathbb{P}^2$ can be represented by the symmetric 3×3 rank two matrix $N_i = u_i v_i^T + v_i u_i^T$.

Remark 5.2. Since X, Y, u_i, v_i are points in projective spaces the symmetric rank two matrices M, N_i are *only defined up to scale.*

For a symmetric 4×4 rank two matrix M that is suitably generic, ξ_A maps the matrix M to n symmetric 3×3 of rank two matrices. Thus the unlabeled analogue of the perspective relation $AX = \lambda u$ for a pinhole camera $A \in \mathbb{R}^{3\times 4}$ reads as

$$A(\underbrace{XY^T + YX^T}_{=M})A^T = \lambda(\underbrace{uv^T + vu^T}_{=N}),\ \lambda \in \mathbb{R}\setminus\{0\} \tag{5.1}$$

The unlabeled multiview map of two points is

$$\begin{array}{ccccc} \xi_A : & \mathrm{Sym}_2(\mathbb{P}^3) & \hookrightarrow & \mathrm{Sym}_2(\mathbb{R}^4) & \dashrightarrow & \big(\mathrm{Sym}_2(\mathbb{R}^3)\big)^n, \\ & & & M & \mapsto & (A_1 M A_1^T, \dots, A_n M A_n^T). \end{array} \tag{5.2}$$

If $m = 2$ the unlabeled multiview variety $\mathrm{Sym}_2(V_A)$ is the closure of the image of the rational map ξ_A.

We will pay special attention to the algebraic variety $V(\xi_A)$ of the closure of the image of ξ_A in Section 5.3.

5.2 Relabeling the Unlabeled

The Chow variety is a tool to easier handle unlabeled points, in order to understand the geometry it is sometimes more convenient not to think about unlabeled points as symmetric tensors. While in the section above we viewed two unlabeled points as one point on the Chow variety $\mathrm{Sym}_m(\mathbb{P}^2)$, in this section we will stick to unlabeled points as points in $\mathbb{P}^2$ and $\mathbb{P}^3$. Let ω be the map that takes a labeled point configuration in $(\mathbb{P}^2)^m$ to its unlabeled configuration represented as a symmetric tensor, e.g. if $n = 2$, $m = 2$ then ω takes (u, v) and (v, u), both in $\mathbb{P}^2 \times \mathbb{P}^2$, to the same symmetric matrix $uv^T + vu^T$. By construction of the Chow variety the closure of the preimage of the unlabeled multiview variety under the map ω is the union of labeled multiview varieties with the labeling of their image points interchanged.

Example 5.3. *The image points $u_1, v_1, u_2, v_2 \in \mathbb{P}^2$ are on the product of two two-view varieties if they satisfy*

$$u_2^T F u_1 = 0 \text{ and } v_2^T F v_1 = 0.$$

The image points $u_1, v_1, u_2, v_2 \in \mathbb{P}^2$ are on the closure of the preimage under ω of the unlabeled two-view variety of two points if they satisfy (s. Figure 5.2)

$$\big(u_2^T F u_1 = 0 \text{ and } v_2^T F v_1 = 0\big) \text{ or } \big(v_2^T F u_1 = 0 \text{ and } u_2^T F v_1 = 0\big). \tag{5.3}$$

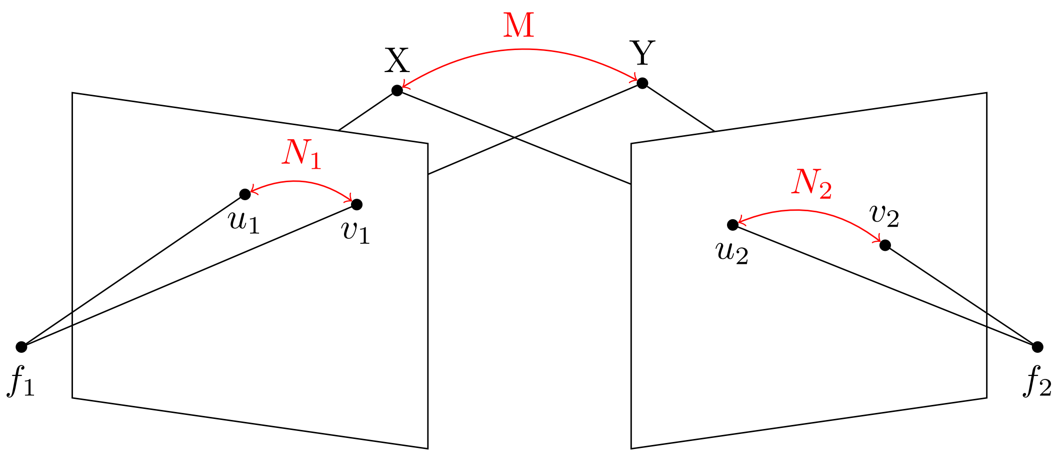
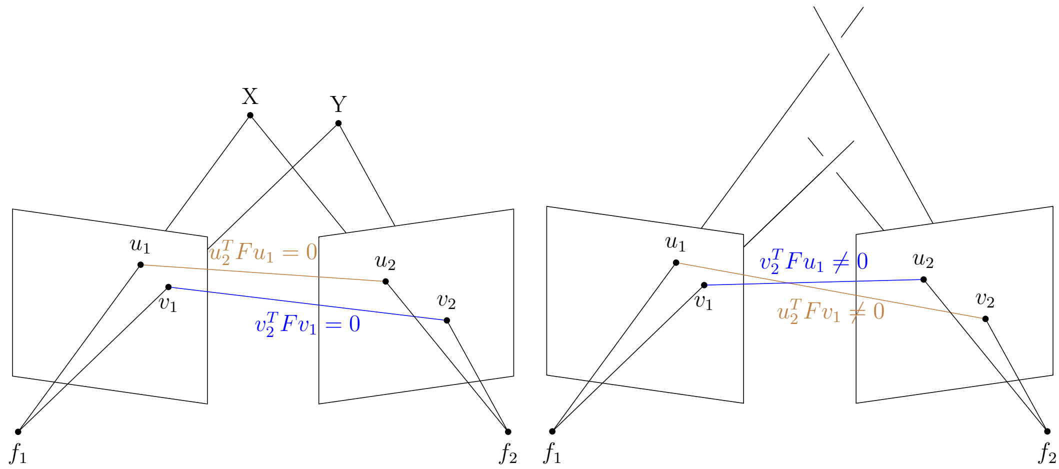


FIGURE 5.1: Unlabeled two-view geometry.



(A) $u_1 \sim u_2$, $v_1 \sim v_2$.

(B) $v_1 \sim u_2$, $u_1 \sim v_2$.

FIGURE 5.2: The two multiview varieties with permuted image point correspondence.

This variety is cut out set-theoretically by the ideal

$$\big\langle (u_2^T F u_1)(v_2^T F u_1), (u_2^T F u_1)(u_2^T F v_1), (v_2^T F v_1)(v_2^T F u_1), (v_2^T F v_1)(u_2^T F v_1) \big\rangle.$$

The approach of Example 5.3 extends to more pictures and more unlabeled world points. The idea here is to permute the labeling of the image points in each picture. Each permutation of image points gives a new product of multiview varieties with permuted coordinates. The union over all these product varieties then is the unlabeled multiview variety.

We construct a graph with all the image points as vertices, see Figure 5.3. Let G be the n-partite complete graph on $n \cdot m$ vertices, where each partition of G has cardinality m. Let $K \subset G$ be a perfect n-dimensional matching and by $\mathcal{K}$ we denote the set of all such perfect n-dimensional matchings on G. For a fixed K we denote $P_i(j)$, that is an image point u_j^* in the j-th picture, as the vertex in the j-th partition of G in the i-th path P_i of K, with $i \in [n]$, $j \in [m]$.

In this section we have not been working with unlabeled points in $\mathrm{Sym}_m(\mathbb{P}^2)$ but with labeled points $(\mathbb{P}^2)^m$ and their orbits under group action on the image points of the symmetric group. The following theorem illustrates the relation between these two approaches.

Proposition 5.4. *The closure of the preimage under ω of the unlabeled multiview variety is the union of products of multiview varieties, with permuted image points correspondences*

$$\bigcup_{K \in \mathcal{K}} \Big(V_A\big(P_1(1), P_2(1), \ldots, P_n(1)\big) \times \ldots \times V_A\big(P_1(m), P_2(m), \ldots, P_n(m)\big) \Big).$$

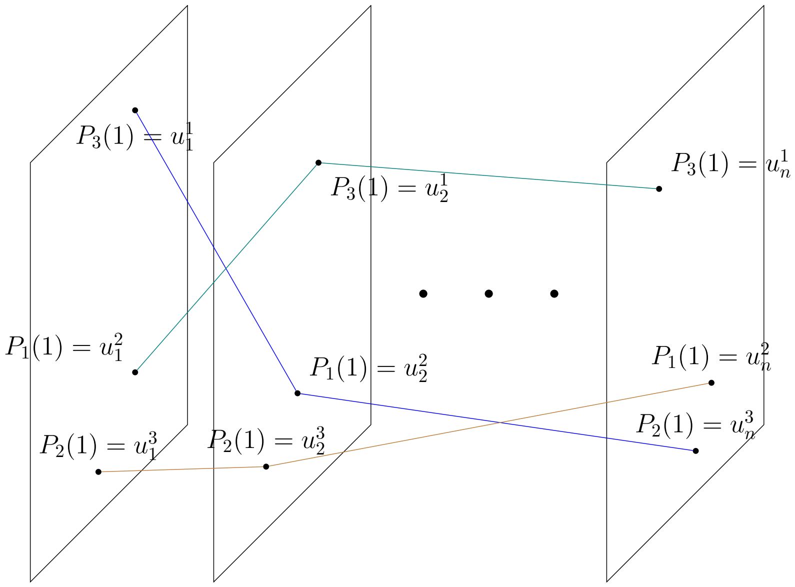


FIGURE 5.3: Permuted image point correspondence graph.

Proof. The usual multiview variety knows the labeling of the images points, however the unlabeled does not. Thus the preimage under ω of the unlabeled variety is nothing but the union of all multiview varieties with permuted image point correspondences. A permuted image point correspondence is an n-dimensional matching through the n images, as depicted in Figure 5.3. Taking the union over all such possible correspondences yields the preimage of the unlabeled multiview variety. □

Remark 5.5. The multiview variety is cut out set-theoretically by the bilinear constraints, if the cameras are in linearly general position [Hr97]. Thus the unlabeled multiview variety is cut out by the products of bilinear constraints.

The usual two-view triangulation does result in either zero solutions (u_1, u_2 are not on the two-view variety), one solution (u_1, u_2 are on the two-view variety) or a one-dimensional linear subspace of solutions (u_1, u_2 are the epipoles). The situation is a little different if we forget the labeling of the image points. Anyway, the triangulation problem still amounts in intersecting the back-projected lines. The reconstruction of the original world points is called *unlabeled triangulation*. In the following its ambiguities are studied. The unlabeled triangulation is *ambiguous* if there are two or more unlabeled world point configurations that project down to the same unlabeled image point configuration.

Proposition 5.6. *Let n=m=2. The unlabeled triangulation is not unique, if and only if the image points are on the unlabeled multiview variety and satisfy the two conditions*

$$|e_{12}, u_1, v_1| = 0, \quad |e_{21}, u_2, v_2| = 0 \ .$$

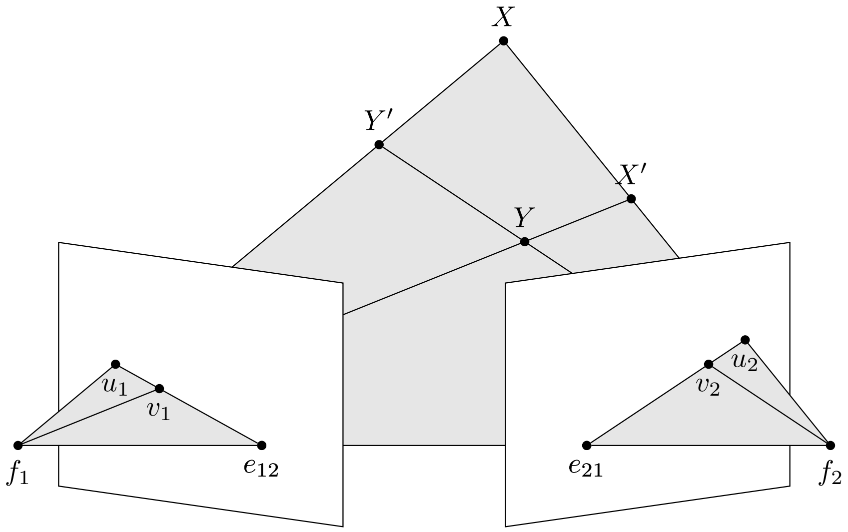


FIGURE 5.4: Ambiguous unlabeled two-view triangulation.

If the image points u_1, u_2, v_1, v_2 are distinct from the epipoles e_1, e_2, then there are exactly two possible unlabeled world point pairs (X, Y) and (X', Y') that are possible reconstructions.

Proof. By construction the arrangement of the four back-projected lines of u_1, u_2, v_1, v_2 have the focal points f_1 and f_2 as intersection points. Such that the triangulation is ambiguous the four back-projected lines must have at least four additional intersections. Thus the back-projected lines need to be coplanar by the Veblen-Young axiom. Because u_1, u_2, v_1, v_2 are on the unlabeled multiview variety this is equivalent to e_{ij}, u_1, v_1 being collinear, as depicted in Figure 5.4. Now if e_{12}, u_1, v_1 are pairwise distinct and e_{21}, u_2, v_2 are pairwise distinct, then the back-projected lines have in total six intersections of which four are not the focal points. □

Remark 5.7. Let n=m=2. In world coordinates the ambiguity of the unlabeled triangulation revolves around the world points X, Y and the baseline β_{12} through the focal points f_1, f_2 being coplanar. Hence the pencil of planes with the baseline β_{12} as its axis encodes the ambiguities of the two view unlabeled triangulation.

The ambiguity of the unlabeled triangulation of two views and two unlabeled points extends to the case $n = 2, m \geq 2$.

Corollary 5.8. *If a subset of the world point configuration and the baseline β_{ij} are coplanar, then the unlabeled triangulation is ambiguous.*

Proof. Follows immediately from Proposition 5.6. □

When a subset of k unlabeled world points and the baseline are on one plane, then all their back-projected lines intersect each other pairwise. There are in total k^2 such intersection points. To construct an alternative world point configurations, each back-projected line has to be used exactly one time. Otherwise the resulting image point configuration does not align with the original image point configuration.

Remark 5.9. Consider k world points that are coplanar with the baseline, then there are $k!$ different solutions to the unlabeled triangulation from two views.

For generic points on the unlabeled multiview variety the back-projected lines intersect in only two points and the reconstructing of the unlabeled world point configuration is unique.

When $m > 2$ the back-projected lines represent a line arrangement in $\mathbb{P}^3$, where through each focal point m lines pass. Let the *intersection degree* of a point in $\mathbb{P}^3$ denote the number of back-projected lines that intersect in that point. The unlabeled triangulation amounts to find points in $\mathbb{P}^3$ with intersection degree equal to n that are distinct from the focal points. A configuration of world points is a valid triangulation of the image points if it covers all back-projected lines.

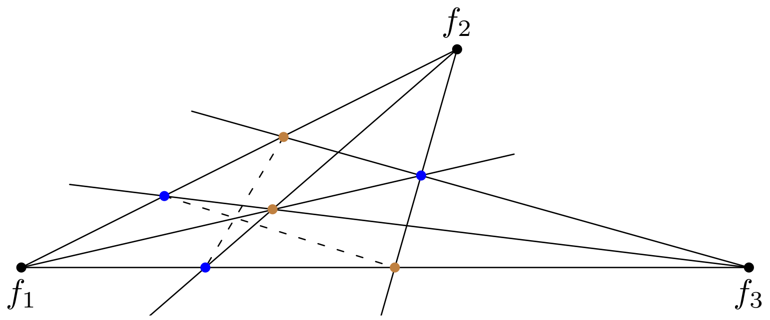


FIGURE 5.5: The blue and brown vertices describe different world point configurations that project to the same unlabeled image point configuration. The dashed lines denote the missing incidences in Pappus's hexagon theorem.

In general classifying the ambiguities of the unlabeled triangulation is a complex problem.

For three views $n = 3$ and three unlabeled points one can employ Pappus's hexagon theorem to construct an ambiguous unlabeled triangulation. In this case a configuration that possesses an ambiguous unlabeled triangulation must lie on the trifocal plane. Figure 5.5 depicts such an situation on the trifocal plane.

5.3 Two Unlabeled Points

In this section we will only consider the case of two (m=2) unlabeled world points. The complicated part of the unlabeled multiview map of Equation 5.2 is the rank two constraint on the symmetric matrix M. However if M has rank two we can be sure that its projections N_i have rank two.

Lemma 5.10. *The matrices N_i have rank at most two if M has rank two. If $M = XY^T + YX^T$, then N can be written as $N_i = uv^T + vu^T$ for some $u, v \in \mathbb{P}^2$.*

Proof. If M has rank two, then $rk(N_i) = rk(A_i M A_i^T) \leq 2$. The second statement is obtained from the fact

$$A_i(XY^T + YX^T)A_i^T = A_i X (YA_i)^T + A_i Y (XA_i)^T = \lambda_i uv^T + vu^T = \lambda N_i.$$

□

Remark 5.11. The reverse of the second statement of Lemma 5.10 does not hold $N_i = uv^T + vu^T \, \forall \, i \nRightarrow M = XY^T + YX^T$

We first analyze $\mathrm{Sym}_2(V_A)$ by dropping the rank two constraints in ξ_A. This gives us the rational map

$$\begin{array}{llll} \theta_A : & \mathrm{Sym}_2(\mathbb{R}^4) & \dashrightarrow & \left(\mathrm{Sym}_2(\mathbb{R}^4)\right)^n, \\ & M & \mapsto & (A_1 M A_1^T, \ldots, A_n M A_n^T). \end{array}$$

We will denote the ideal and the variety of the closure of the image of as I_θ and V_θ. This variety is a relaxation of the unlabeled multiview variety $\mathrm{Sym}_2(V_A)$. The map θ_A is linear in the entries of the symmetric matrix M, thus by *row-wise vectorizing* the upper triangular entries of $M \in \mathrm{Sym}_2(\mathbb{R}^4)$ to $\overrightarrow{M} \in \mathbb{R}^{10}$ the map θ_A can be rewritten in its standard form of linear equations, with the coefficient matrices $\widetilde{A}_i$. In the coordinates of $A_i = (a_{jk})_{j \in [3], k \in [4]}$ the coefficient matrix $\widetilde{A}_i$ then reads as

$$\widetilde{A}_i = \begin{pmatrix} a_{11}^2 & a_{11}a_{21} & a_{11}a_{31} & a_{21}^2 & a_{21}a_{31} & a_{31}^2 \\ 2a_{11}a_{12} & a_{12}a_{21} + a_{11}a_{22} & a_{12}a_{31} + a_{11}a_{32} & 2a_{21}a_{22} & a_{22}a_{31} + a_{21}a_{32} & 2a_{31}a_{32} \\ 2a_{11}a_{13} & a_{13}a_{21} + a_{11}a_{23} & a_{13}a_{31} + a_{11}a_{33} & 2a_{21}a_{23} & a_{23}a_{31} + a_{21}a_{33} & 2a_{31}a_{33} \\ 2a_{11}a_{14} & a_{14}a_{21} + a_{11}a_{24} & a_{14}a_{31} + a_{11}a_{34} & 2a_{21}a_{24} & a_{24}a_{31} + a_{21}a_{34} & 2a_{31}a_{34} \\ a_{12}^2 & a_{12}a_{22} & a_{12}a_{32} & a_{22}^2 & a_{22}a_{32} & a_{32}^2 \\ 2a_{12}a_{13} & a_{13}a_{22} + a_{12}a_{23} & a_{13}a_{32} + a_{12}a_{33} & 2a_{22}a_{23} & a_{23}a_{32} + a_{22}a_{33} & 2a_{32}a_{33} \\ 2a_{12}a_{14} & a_{14}a_{22} + a_{12}a_{24} & a_{14}a_{32} + a_{12}a_{34} & 2a_{22}a_{24} & a_{24}a_{32} + a_{22}a_{34} & 2a_{32}a_{34} \\ a_{13}^2 & a_{13}a_{23} & a_{13}a_{33} & a_{23}^2 & a_{23}a_{33} & a_{33}^2 \\ 2a_{13}a_{14} & a_{14}a_{23} + a_{13}a_{24} & a_{14}a_{33} + a_{13}a_{34} & 2a_{23}a_{24} & a_{24}a_{33} + a_{23}a_{34} & 2a_{33}a_{34} \\ a_{14}^2 & a_{14}a_{24} & a_{14}a_{34} & a_{24}^2 & a_{24}a_{34} & a_{34}^2 \end{pmatrix}$$

and the according rational map is

$$\begin{array}{llll} \widetilde{\theta}_A : & \mathbb{P}^9 & \dashrightarrow & (\mathbb{P}^5)^n, \\ & \overrightarrow{M} & \mapsto & (\widetilde{A}_1 \overrightarrow{M}, \ldots, \widetilde{A}_n \overrightarrow{M}), \end{array} \tag{5.4}$$

This map strongly resembles the conventional multiview map of Equation 2.5 and can be understood as a higher dimensional analog of it. We use the techniques developed in [Li17] for a more general setup to describe the closure of the image of $\widetilde{\theta}_A$ in Corollary 5.16. But we first study the unlabeled triangulation. The results we obtain to understand the unlabeled triangulation will come in handy for Corollary 5.16.

Let $\sigma \subseteq [n]$ with $|\sigma| = k$ and $\widetilde{A}_\sigma = (\widetilde{A}^T_{\sigma_1}, \dots, \widetilde{A}^T_{\sigma_k})$, then define B_σ as

$$B_\sigma = \begin{bmatrix} \widetilde{A}_{\sigma_1} & \vec{N}_{\sigma_1} & & \\ \vdots & & \ddots & \\ \widetilde{A}_{\sigma_k} & & & \vec{N}_{\sigma_k} \end{bmatrix}. \tag{5.5}$$

Let the *unlabeled focal point* f_{ij} of the cameras A_i and A_j be $f_{ij} := f_i f_j^T + f_j f_i^T$.

Proposition 5.12. *Let $n = m = 2$. Let N_1 and N_2 be two generic points on the unlabeled two-view variety and denote their unlabeled triangulation $M_\Delta = XY^T + YX^T$. Then $rk(B_{\{1,2\}}) = 10$ and there exists λ_1, $\lambda_2 \in \mathbb{R}$, such that*

$$\operatorname{span}\big((\overrightarrow{f_{12}}, 0, 0)^T, (\overrightarrow{M}_\Delta, \lambda_1, \lambda_2)^T\big) = \ker(B_{\{1,2\}}).$$

Proof. Since N_1 and N_2 are generic points on the unlabeled multiview variety the triangulation is unique and $M_\Delta = XY^T + YX^T$ is its solution. Hence there are λ_1, λ_2, such that $(\overrightarrow{M}_\Delta, \lambda_1, \lambda_2)$ is in $\ker(B_{\{1,2\}})$. On the other hand $\widetilde{A}_i \overrightarrow{M} = \lambda_i \vec{N}_i$ is equivalent to $A_i M A_i^T = \lambda_i N_i$, thus inserting $f_1 f_2^T + f_2 f_1^T$ gives $A_1 \left(f_1 f_2^T + f_2 f_1^T\right) A_1^T = 0$ and inserting gives $A_2 \left(f_2 f_1^T + f_1 f_2^T\right) A_2^T = 0$. Hence for $\lambda_1 = \lambda_2 = 0$ the vector $(\vec{f}_{12}, 0, 0)$ is in $\ker(B_{\{1,2\}})$. □

Remark 5.13. By the proof of Proposition 5.12 $\ker(\widetilde{A}_{\{i,j\}})$ is spanned by $\vec{f}_{ij}$.

Proposition 5.14. *Let $n \geq 3$. For a generic choice of cameras $A_1, \dots, A_n$ the matrix $\widetilde{A}_\sigma$ has full rank.*

Proof. The unlabeled focal point f_{ij} is the kernel of the submatrix $\widetilde{A}_{\{i,j\}}$. Since f_{ij} does not depend on the other cameras it is not in the kernel of any of the other $\widetilde{A}_k$ for a generic choice of cameras. □

We can use the matrix B_σ to design a triangulation algorithm for the unlabeled multiview variety. We call this algorithm the *unlabeled triangulation algorithm*. The Proposition 5.12 indicates that the unlabeled triangulation is more complicated than the usual labeled triangulation. By Propositions 5.6 and Corollary 5.8 it can have multiple solutions. Especially the case of two views is more elaborate. Nonetheless it has a unique solution for generic points on the unlabeled multiview variety. We need a lemma before we can state the unlabeled triangulation algorithm.

Lemma 5.15. *Consider a two-dimensional linear subspace $V \subset \operatorname{Sym}_2(\mathbb{R}^4)$ spanned by two symmetric matrices. Suppose that V contains two distinct rank two symmetric matrices $M_1, M_2 \in \operatorname{Sym}_2(\mathbb{R}^4)$, such that $rk(M_1 - M_2) = 4$. Then the symmetric matrices M_1, M_2 are the only rank two symmetric matrices in V. Also the subspace V contains no rank three symmetric matrix.*

Proof. We can choose M_1 and M_2 as basis of V. The rank two matrices in V can be computed as some roots of $\det(\alpha M_1 + (1-\alpha)M_2)$. Clearly zero and one are roots of this univariate polynomial. However the roots of $\det(\alpha M_1 + (1-\alpha)M_2)$ are the generalized eigenvalues to the matrix equation $M_1 x = \alpha(M_2 - M_1)x$, $x \in \mathbb{R}^4$. But since M_1 has rank two the generalized eigenvalue $\alpha = 0$ has double algebraic and geometric multiplicity. Then switching the roles of M_1 and M_2 shows that $\alpha = 1$ also has double algebraic and geometric multiplicity. □

The unlabeled triangulation of two views results in finding the rank two matrices in linear two-dimensional subspaces. That subspace is constructed from $\ker(B_\sigma)$. We can find these matrices by finding the roots of a univariate quartic polynomial, as their determinant needs to vanish. If we only consider generic points on the unlabeled multiview variety, this polynomial has exactly two real solutions, the unlabeled focal point f_{ij} and the unlabeled triangulation M_Δ of the unlabeled image points.

Input: N_1, N_2 generic points on the unlabeled multiview variety;
Output: M_Δ unlabeled triangulation of N_1, N_2;
begin

1. Compute an element m of $\ker B_\sigma$, that is not a multiple of $\vec{f}_{12}$;
2. The first ten entries of m represent a symmetric 4×4 matrix M;
3. Compute the quartic polynomial $\det(\alpha M + (1-\alpha)f_{12})$;
4. It factors to $\alpha^2(\alpha - c)^2$, $c \in \mathbb{R}$;
5. Compute c; $M_\Delta := cM + (1-c)f_{12}$;

end

Algorithm 1: Unlabeled triangulation problem for two views.

Proof. Since M and f_{12} are symmetric with real entries $\det(\alpha M + (1-\alpha)f_{12})$ only has real roots. By construction $\alpha = 0$ is a root. Because $(N_1, N_2) \in \mathrm{Sym}_2(V_A)$ there is a real value $\alpha = c$ such that $M_\Delta := cM + (1-c)f_{12}$. Now by Lemma 5.15 these two roots are the only roots of $\det(\alpha M + (1-\alpha)f_{12})$. □

Since the roots of the univariate quartic polynomial $\det(\alpha M + (1-\alpha)f_{12})$ have double multiplicity for generic points, we can find the roots by using the *quadratic formula*.

If $n \geq 3$, then the procedure is simple and amounts to computing $\ker B_\sigma$, because the kernel is one-dimensional for generic data by Proposition 5.14. The complete triangulation algorithm can be stated as follows:

Input: $N_1, \ldots, N_n$ generic points on the unlabeled multiview variety;
Output: M_Δ unlabeled triangulation of $N_1, \ldots, N_n$;
if n=2 **then**
| use Algorithm 1;
else

1. Compute the generator m of $\ker B_\sigma$;
2. The first ten entries of m represent the symmetric matrix M_Δ;

end

Algorithm 2: Unlabeled triangulation problem.

Proof. Correctness of Algorithm 2 follows from correctness of Algorithm 1. □

In some cases one might not want to keep on working with symmetric matrices, but actually reconstruct the original world points. This amounts to find a decomposition of M^Δ into $XY^T + YX^T$. Finding such a decomposition is equivalent to splitting a degenerate quadrics $\mathbb{P}^3$ into two planes. To do this, the approach studied in [RG11] for quadrics in $\mathbb{P}^2$ generalizes to our case, if one replaces the adjoint matrix $\mathrm{adj}(M^\Delta)$ by the 6×6 matrix of signed 2-minors of M^Δ.

We go back to studying the relaxed ideal I_θ. With the help of Proposition 5.12 we are able to determine the generators of the prime ideal I_θ of the image of θ_A.

Corollary 5.16. *Let $\sigma \subseteq [n]$. Then the ideal I_θ of the image of θ_A is generated by the maximal minors of B_σ (for all $3 \leq |\sigma| \leq 10$), and the 11-minors of B_σ (for all $|\sigma| = 2$).*

Proof. By [Li17, §2.2] the ideal is generated by the $rk(\widetilde{A}_\sigma) + |\sigma|$ minors of B_σ. If $|\sigma| = 2$ then by the proof of Proposition 5.12 the matrix $\widetilde{A}_\sigma$ has rank nine. If $3 \leq |\sigma| \leq 10$ then $\widetilde{A}_\sigma$ has full rank by Proposition 5.14. Further maximal minors of matrices B_σ with $|\sigma| \geq 11$ are monomial multiples of the other maximal minors. We obtain a generating set by just computing the minors of correct size. □

Remark 5.17. Let n=2. Consider the ideal generated by the 11-minors and the ideal generated by the three degree $(1, 1)$ polynomials of Example 5.1. These two ideals are equal.

Clearly the relaxed ideal I_θ is not a good relaxation of $\mathrm{Sym}_2(V_A)$, but it can easily be tightened by enforcing the singularity condition $\det(N_i) = 0$. We define the ideal

$$I'_\theta := I_\theta + \langle det(N_1), \ldots, det(N_n) \rangle.$$

Still it is possible to construct rank two matrices on the variety $V(I'_\theta)$ that are not on $\mathrm{Sym}_2(V_A)$. Let be $N_1 = e_{12}v^T + ve_{12}^T, v \in \mathbb{P}^2$. Then $rk(B) \leq 10$ independent of N_2, thus requiring that the 11×11 minors of B vanish (s. Corollary 5.16) does not impose any constraints on N_2. Remember that Equation 5.3 gives a description of the preimage of the unlabeled two-view variety of two points. However

$$\big(V(e_{12}^T F u_2) \times V(v_1^T F v_2)\big) \cup \big(V(e_{12}^T F v_2) \times V(v_1^T F u_2)\big)$$

gives constraints on v_1, v_2, u_2, namely ($v_2^T F u_1 = 0$ or $v_2^T F v_1 = 0$). Thus not all points of

$$\big\{(e_{12}v_1^T + v_1 e_{12}^T, N_2) : v \in \mathbb{P}^2, N_2 \in \mathrm{Sym}_2(\mathbb{R}^4)\big\} \tag{5.6}$$

are on the unlabeled multiview variety $\mathrm{Sym}_2(V_A)$, but they are on $V(I'_\theta)$.

Proposition 5.18. *For n=2, the unlabeled multiview variety is minimally generated by 7 polynomials in 12 variables. It is Cohen-Macaulay and its Betti table is given in Table 5.1.*

Proof. By Remark 4.4 of [AST13] the toric set-up of Example 2.1 is universal in the sense that every multiview variety with cameras in linearly general position is isomorphic to the multiview variety of Example 2.1. Since the unlabeled multiview variety with cameras in linearly general position is the union of multiview varieties with correspondences interchanges, any unlabeled multiview variety is isomorphic to the unlabeled multiview variety with the camera matrices of Example 2.1. For two views we only chose the camera matrices A_1 and A_2 of Example 2.1 and then ran the code of Listing 5.1 in `Macaulay2` [GS] with these cameras as input. □

Remark 5.19. For cameras in linearly general position we have the freedom of choice of a world coordinate system in $\mathbb{P}^3$ and coordinate systems in the images in $\mathbb{P}^2$ without changing the unlabeled multiview variety up to isomorphism. In that sense we can freely choose five cameras with focal points in linearly general position to compute an isomorphic description of the unlabeled multiview variety.

As we see in Example 5.1 the generators of the unlabeled two-view variety of two points ($n = 2$, $m = 2$) fall into three different classes according to their

```
            0 1  2 3 4
     total: 1 7 11 8 3
         0: 1 .  . . .
         1: . 3  2 . .
         2: . 4  6 . .
         3: . .  . 2 .
         4: . .  3 6 3
```

Table 5.1: Betti numbers for the unlabeled multiview ideal with $n = 2$.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13
total:	1	60	468	1580	3071	4765	5715	4741	2808	1257	428	102	15	1
0:	1	.	.	.	.	.	.	.	.	.	.	.	.	.
1:	.	9	6	.	.	.	.	.	.	.	.	.	.	.
2:	.	20	102	159	145	66	12	.	.	.	.	.	.	.
3:	.	24	273	932	1242	468	60	.	.	.	.	.	.	.
4:	.	.	12	123	609	2116	2709	1800	657	123	12	.	.	.
5:	.	7	75	366	1075	2115	2934	2941	2151	1134	416	102	15	1

TABLE 5.2: Betti numbers for the unlabeled multiview ideal with $n = 3$.

multidegree, namely the ones with multidegree $(1, 1)$, the ones with degree type $(2, 1)$ and the ones with degree type $(3, 0)$. The ones of degree type $(3, 0)$ are the determinants of N_1 and N_2.

Proposition 5.20. *The ideal of three degree* $(1, 1)$ *equations of the unlabeled two-view variety is a subset of the ideal generated by the entries of* N_2FN_1*.*

Proof. The statement has been checked for sufficiently many random choices of camera pairs with `Macaulay2` [GS], see Section 5.5 for details. □

We currently have no way to construct the generators with degree type $(2, 1)$. However we do believe that they relate to the variety of Equation 5.6.

The following statements have been derived with `Macaulay2` [GS]. They are concerned with the prime ideal of the unlabeled multiview variety of three, four and five views.

Proposition 5.21. *For n=3, the unlabeled multiview variety is minimally generated by 60 polynomials in 18 variables and its Betti table is given in Table 5.2.*

Proof. Similarly to the proof of Proposition 5.18 we choose special cameras to compute the unlabeled multiview variety. Here the cameras A_1, A_2 and A_3 of Example 2.1 were chosen for the computation with `Macaulay2` [GS]. □

Proposition 5.22. *For n=4, the unlabeled multiview variety is minimally generated by 215 polynomials up to degree six in 24 variables. For n=5, the unlabeled multiview variety is minimally generated by 620 polynomials up to degree six in 30 variables.*

Proof. Similarly to the proof of Proposition 5.18 we choose special cameras to compute the unlabeled multiview variety. For $n = 4$, the cameras of Example 2.1 were chosen for the computation with `Macaulay2` [GS]. For $n = 5$, we chose the four cameras of Example 2.1 and additionally the normalized camera with focal point $e_1 = (1, 0, 0, 1)$ for the computation with `Macaulay2` [GS]. □

The *unlabeled rigid multiview variety* is the image of the rigid multiview variety $V(J_A) \subset V_A \times V_A$ under the quotient map that takes two copies of $(\mathbb{P}^2)^2$ to two copies of $\mathrm{Sym}_2(\mathbb{P})^2 \subset \mathbb{P}^5$. This quotient map is given by $(u_1, v_1) \mapsto N_1$, $(u_2, v_2) \mapsto N_2$.

Example 5.1 (Continued). *We can get equations for the* unlabeled rigid multiview variety*, if we intersect the ideal J_A with the subring $\mathbb{R}[N_1, N_2]$ of bisymmetric homogeneous polynomials in $\mathbb{R}[u, v]$. This results in nine new generators which represent the distance constraint. One of them is a quartic of degree $(2, 2)$ in (N_1, N_2). The other eight are quintics, four of degree $(2, 3)$ and four of degree $(3, 2)$.*

5.4 More than Two Unlabeled Points

The study of more than two unlabeled points is quite more elaborate compared to the case of two points as done in Section 5.3. In general, equations describing the Chow variety are unknown. Also when more than two unlabeled points are present we work with symmetric tensors instead of symmetric matrices. In Section 5.3 we strongly relied on tools from matrix algebra, these are not available or far more complicated for symmetric tensors. We can still extend some results to the case of $m \geq 3$.

One can construct a point M on the Chow variety $\mathrm{Sym}_m(\mathbb{P}^3)$ from their unlabeled world configuration $X_1, \ldots, X_m$ by taking the *sum over all permuted tensors products*

$$M = \sum_{\pi \in S(m)} \left(X_{\pi(1)} \otimes \ldots \otimes X_{\pi(m)} \right) \in \mathbb{R}^{4 \times \ldots \times 4},$$

where $S(\cdot)$ denotes the symmetric group. The tensor M is symmetric and only defined up to scale. Thus it can be embedded in $\mathbb{P}^{\binom{m+3}{m}-1}$. Similarly, the unlabeled image point configurations are represented by symmetric tensors $N_i \in \mathbb{R}^{3 \times \ldots \times 3}$ defined up to scale, embedded in $\mathbb{P}^{\binom{m+2}{m}-1}$. The perspective relation of Equation 5.1 for a pinhole camera $A \in \mathbb{R}^{3 \times 4}$ generalizes to

$$M(\underbrace{A, \ldots, A}_{m \text{ times}}) = \lambda N.$$

We can vectorize this equation above as done in the Section 5.3. This yields the linear equation

$$\widetilde{A}\overrightarrow{M} = \lambda \overrightarrow{N}$$

for some coefficient matrix $\widetilde{A}$, with $\binom{m+2}{m}$ rows and $\binom{m+3}{m}$ columns. Let $\sigma \subseteq [n]$ with $|\sigma| = k$ and $\widetilde{A}_\sigma = (\widetilde{A}^T_{\sigma_1}, \ldots, \widetilde{A}^T_{\sigma_k})$ then define (analogously to the two point case of Equation 5.5) the matrix B_σ as

$$B_\sigma = \begin{bmatrix} \widetilde{A}_{\sigma_1} & \overrightarrow{N}_{\sigma_1} & & \\ \vdots & & \ddots & \\ \widetilde{A}_{\sigma_k} & & & \overrightarrow{N}_{\sigma_k} \end{bmatrix}.$$

This matrix has $k\binom{m+2}{m}$ rows and columns $\binom{m+3}{m}+k$. The *unlabeled triangulation* M_Δ can be reconstructed from the matrix B_σ. Similar to the two point case one can construct the *unlabeled focal point* of m views, this is

$$f_{[m]} := \sum_{\pi\in S(m)} \left(f_{\pi(1)} \otimes \ldots \otimes f_{\pi(m)}\right) \in \mathbb{R}^{3\times\ldots\times 3}.$$

For exactly m views the unlabeled focal point concatenated with m zeros $(\overrightarrow{f}_{[m]}, 0, \ldots, 0)$ is in the kernel of $B_{[m]}$. For $m+1$ views the kernel of $B_{[m+1]}$ is one-dimensional and M_Δ can be computed by linear algebra.

It is of interest to characterize the pictures of m unlabeled points using n cameras and their ambiguities. Further it would be desirable to know the prime ideal of $\mathrm{Sym}_m(V_A)$ for any n and m. In Propositions 5.18, 5.21 and 5.22 generators of degree at most six appear, thus we believe that the generators of $\mathrm{Sym}_m(V_A)$ can be constructed from the information obtained by two and three views.

5.5 Computations

We performed several random experiments. Our hardware was a cluster with Intel Xeon X2630v2 Hexa-Cores (2.8 GHz) and 64GB main memory per node. The software was `Macaulay2` [GS], version 1.9.2 [GS]. All computations were single-threaded.

Our result in Proposition 5.20 was proved by computations with `Macaulay2` [GS]. Following standard practice in computational algebraic geometry, we carried out the computation on many samples in a Zariski dense set of parameters, and then conclude that it holds generically. Further instead of using special cameras as we did in the proofs of Propositions 5.18, 5.21, 5.22, we were also able to compute the unlabeled multiview variety with random cameras as input.

The computations were repeated several times with random input. It is not surprising that increasing n, the number of cameras, increases the running times considerably. In particular using the toric setup of Example 2.1 is much faster than choosing dense camera matrices A_i.

For Proposition 5.20 we performed at least 1000 computations to verify its correctness. Example 5.1 was checked with 50 choices of random cameras.

In Listing 5.1 we show `Macaulay2` code which can be employed to establish Proposition 5.18.

Lines 1–4 define the rings in which the computations take place. Lines 6–14 produce random camera matrices. However, our experiments suggest that it suffices to check that the focal points of the cameras are in linearly general position. The unlabeled multiview map θ_A from Equation 5.3 is encoded in lines 17–21. Lines 13–14 are unlabeled perspective relations (Equation 5.1) and line 26 are the determinantal constraints on the matrices. The actual computation is the elimination in line 28. The unlabeled multiview variety is defined in lines 30–31.

```
ImageRing=QQ[a00,a01,a02,a11,a12,a22]**QQ[b00,b01,b02,b11,b12,b22]
WorldRing=QQ[m00,m01,m02,m03,m11,m12,m13,m22,m23,m33]
MultipleRing=QQ[l,k]
S=WorldRing**ImageRing** MultipleRing

--generate random camera matrices
n  = 2;
AList=0;
while (numgens minors(4,transpose matrix AList)=!=binomial(n*3,4))do(
   AList={};
   for i from 0 to n-1 do(
      A_i=random(ZZ^3,ZZ^4,Height=>20);
      AList=AList| entries A_i; )
   );

--create matrices corresponding to unlabeled points
M=genericSymmetricMatrix(S,m00,4);
N_0=genericSymmetricMatrix(S,a00,3);
N_1=genericSymmetricMatrix(S,b00,3) ;

--unlabeled multiview map
I=ideal(
    A_0*M*transpose A_0-l*N_0,
    A_1*M*transpose A_1-k*N_1,
    l*k-1 ,
    det(N_0),det(N_1))+minors(3,M);

time I = eliminate({m00,m01,m02,m03,m11,m12,m13,m22,m23,m33,l,k},I)

F = map(ImageRing,S);
J = F(I);
```

Listing 5.1: Compute $\mathrm{Sym}_2(V_A)$ for two cameras.

Chapter 6

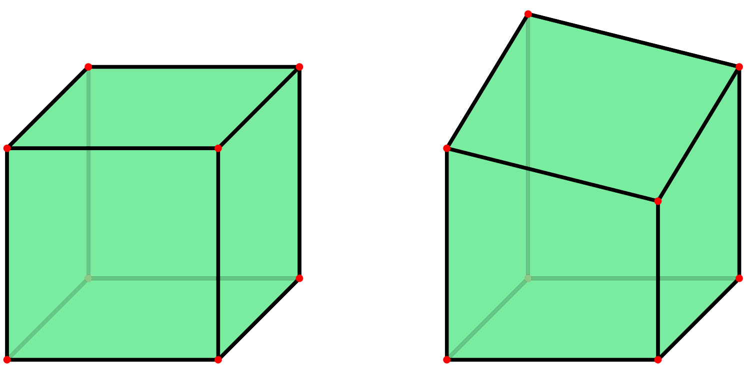
Pictures of Cubes

6.1 Introduction

The *8-point algorithm* [HZ03, Algorithm 11.1] is one of the key algorithms in epipolar geometry. It is successfully used in a vast number of applications to compute the fundamental matrix between two views. The 8-point algorithm purely relies on methods from linear algebra and is extremely fast. At the same time it only gives slightly inferior results compared to more involved algorithms that rely on nonlinear optimization.

An algorithm to reconstruct the fundamental matrix is *defeated by a world point configuration* in $(\mathbb{P}^3)^n$ if the algorithm fails to produce a unique fundamental matrix from the projections in $(\mathbb{P}^2)^n$ of that world point configuration independent on the choice of cameras. The 8-point algorithm is *defeated* by the eight vertices of a unit cube. We extend this result to arbitrary convex 3-polytopes bounded by six quadrilateral faces, whose vertex-facet incidence is the same as that of a cube. We will call such a polytope in $\mathbb{R}^3$ throughout this chapter *combinatorial cube*. The vertex-facet incidence of a polytope is the undirected bipartite graph formed by the containment of vertices within the facets [JT13, §3.5].

In the last section of this chapter we suggest how to handle the situation



(A) a standard cube. (B) a combinatorial cube.

FIGURE 6.1: Two projective non equivalent combinatorial cubes.

when the 8-point algorithm is defeated by the vertices of a combinatorial cube by using a modified version of the 7-point algorithm in this case.

There are multiple papers in multiview geometry which are concerned with critical configurations [May12; HK07]. Critical configurations in two-view geometry are point configurations consisting of n world points together with the two focal points $f_i \in \mathbb{P}^3$ of the cameras. For these configurations there exist two ambiguous fundamental matrices. They are a feature of the geometry itself and not of the choice of algorithm used to reconstruct the fundamental matrix $F \in \mathbb{R}^{3\times 3}$. Hartley and Kahl [HK07] give the complete description of critical configurations in multiview geometry. In the two-view case a necessary condition for a focal and world point configuration to be critical, is that the n world points $P_i \in \mathbb{P}^3$ and the two camera centers f_i lie on a *ruled quadric* [Kra41]. Thus it is not possible to reconstruct a unique fundamental matrix, independent of the chosen algorithm. Here we consider world point configurations where the 8-point algorithm always fails to reconstruct the fundamental matrix independent of the camera centers. These configurations are related to critical configurations but they do not align. There are only a few results about these configurations and usually rely on dimensional degeneracies, e.g. too many points on a plane or on a line. In this case the quadric running through the n points and the two cameras centers is degenerate. For example in [Phi98] it is shown, that all points but one on a plane defeat the 8-point algorithm. If we add the two focal points to this configuration then the point off the plane and the focal points span a plane. Hence we can fit a ruled quadric of two intersecting planes through the $n + 2$ points.

By $P_i \in \mathbb{P}^k$ we denote the i-th point in a point configuration $P = (P_1, \dots, P_n)$ of n points in $\mathbb{P}^k$. Pinhole cameras are represented as 3×4 matrices A_j with real entries. If two image points X_i, Y_i in two different pictures are perspective projections of the same world point P_i they must satisfy the perspective relation $A_1 P_i = \lambda_i X_i$, $A_2 P_i = \mu_i Y_i$ with $\lambda_i, \mu_i \in \mathbb{R} \setminus \{0\}$. As these equations must be satisfied at the same time one can deduce a bilinear relation which both points must satisfy.

$$Y_i^T F X_i = 0,$$

where F is the *fundamental matrix*. To reconstruct the fundamental matrix from the image point configurations $X, Y \in (\mathbb{P}^2)^n$ different algorithms are available [HZ03, §11]. The most commonly used algorithm to reconstruct the fundamental matrix is the 8-point algorithm [HZ03, Algorithm 11.1]. However standard implementations of the 8-point algorithm assume that the image point configurations are suitably generic, such that the fundamental matrix F can be determined as the solution of a system of linear equations. For exact data the 8-point algorithm then is as follows. Via vectorizing $F \in \mathbb{R}^{3\times 3}$ denoted by

Input: Two image point configurations $X, Y \in (\mathbb{P}^2)^n$;
Output: The fundamental matrix F;
begin

1. Compute $Z \in \mathbb{R}^{8\times 9}$ from X, Y;
2. Compute the kernel $\vec{F}$ of Z;
3. The nine coordinates of $\vec{F}$ form the fundamental matrix F;

end

Algorithm 3: 8-point algorithm (without noise)

$\vec{F} \in \mathbb{R}^9$ it can be computed as the kernel of the matrix

$$Z = \begin{bmatrix} X_1 \otimes Y_1 \\ \vdots \\ X_n \otimes Y_n \end{bmatrix} \in \mathbb{R}^{n\times 9},$$

where Z is the row wise tensor product of the image point configurations $X, Y \in (\mathbb{P}^2)^n$. The kernel of Z is one-dimensional and $\vec{F} = \ker(Z)$. This yields the fundamental matrix $F \in \mathbb{R}^{3\times 3}$. One necessary condition for the 8-point algorithm to successfully compute a fundamental matrix is

$$\dim(\ker(Z)) = 1. \tag{6.1}$$

Algorithm 3 states the 8-point algorithm in the absence of noise.

However, even if condition 6.1 is violated and the 8-point algorithm fails to construct a fundamental matrix, it is sometimes possible to retrieve a unique fundamental matrix using a different algorithm by additionally enforcing the rank two constraint of the fundamental matrix F, namely if there exists a unique rank two matrix in the kernel of Z.

Example 6.1. *Let* $P = \operatorname{conv}(\pm e_1 \pm e_2 \pm e_3)$ *be a standard cube and*

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

then

$$X = \begin{bmatrix} 1 & 3 & 1 & 3 & 1 & 3 & 1 & 3 \\ 2 & 2 & 4 & 4 & 2 & 2 & 4 & 4 \\ 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 & 3 & 1 & 3 & 1 & 3 & 1 & 3 \\ 2 & 2 & 4 & 4 & 2 & 2 & 4 & 4 \\ 0 & 0 & 0 & 0 & 2 & 2 & 2 & 2 \end{bmatrix}$$

and

$$Z = \begin{bmatrix} 1 & 2 & 1 & 2 & 4 & 2 & 0 & 0 & 0 \\ 9 & 6 & 3 & 6 & 4 & 2 & 0 & 0 & 0 \\ 1 & 4 & 1 & 4 & 16 & 4 & 0 & 0 & 0 \\ 9 & 12 & 3 & 12 & 16 & 4 & 0 & 0 & 0 \\ 1 & 2 & 3 & 2 & 4 & 6 & 2 & 4 & 6 \\ 9 & 6 & 9 & 6 & 4 & 6 & 6 & 4 & 6 \\ 1 & 4 & 3 & 4 & 16 & 12 & 2 & 8 & 6 \\ 9 & 12 & 9 & 12 & 16 & 12 & 6 & 8 & 6 \end{bmatrix}$$

The kernel of Z is two-dimensional and the 8-point algorithm is defeated by the vertices of P. However, there is only one matrix in $\ker(Z)$ *that is of rank two*

$$F = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and a unique reconstruction of F is possible.

6.2 The 8-Point Algorithm and a Cube

A quadric in $\mathbb{P}^3$ is defined by the algebraic equation $p^T Q p = 0$, where $Q \in \mathbb{R}^{4\times 4}$ and $p \in \mathbb{P}^3$. Since this is a quadratic equation in the indeterminate p we can choose the matrix Q to be symmetric. Clearly Q and any multiple of it λQ, $\lambda \in \mathbb{R}$ define the same quadric. By a slight abuse of notation we will refer to both the quadric and the matrix defining the quadric as Q.

One can try to fit a quadric through a point configuration $P \in (\mathbb{P}^3)^n$, then every point in the configuration gives a linear equation on the ten entries of the symmetric matrix Q. This results in a linear equation system with indeterminate vector $\vec{Q} = [Q_{00}, \ldots, Q_{04}, Q_{11}, \ldots, Q_{14}, Q_{22}, Q_{23}, Q_{24}, Q_{33}, Q_{34}, Q_{44}]$. Its coefficient matrix can be constructed with the Veronese map. The *Veronese map* $\nu_{2,4}$ in degree two and four indeterminates is the map from the four indeterminates to all monomials of degree two in these indeterminates.

$$\begin{array}{cccc} \nu_{2,4}: & \mathbb{P}^3 & \to & \mathbb{P}^9 \\ & [x_1, x_2, x_3, x_4] & \mapsto & [x_1^2, x_1x_2, x_1x_3, x_1x_4, x_2^2, x_2x_3, x_2x_4, x_3^2, x_3x_4, x_4^2]. \end{array}$$

For convenience of notation we apply the map $\nu_{2,4}$ to each point in a configuration separately. The map $\nu_{2,4}$ applied to $P \in (\mathbb{P}^3)^n$ gives a matrix $\nu_{2,4}(P) \in (\mathbb{P}^9)^n$. Therefore, if there exists a quadric Q through the points in a configuration $P \in (\mathbb{P}^3)^n$, it can be computed via the linear equation system $\nu_{2,4}(P)\vec{Q} = 0$.

The rank of Z is very essential to the 8-point algorithm and there is a relation to $\nu_{2,4}(P)$.

Lemma 6.2. *Let A_1, A_2 be two cameras and $P \in (\mathbb{P}^3)^n$ be a world point configuration. Let $A_1 P_i = \lambda_i X \in (\mathbb{P}^2)^n$ and $A_2 P_i = \mu_i Y \in (\mathbb{P}^2)^n$, then the rank of Z is bounded by the rank of $\nu_{2,4}(P)$.*

Proof. The statement can be rewritten. The number of linear independent equations in the system of linear equations $y_i^T F x_i = 0$ is bounded by the number of linear independent equations in the system of linear equations $P_i^T Q P_i = 0$, where Q is a 4×4 generic symmetric matrix of indeterminates. Without loss of generality we assume that $\nu_{2,4}(P)$ is of k and the first k equations $P_1^T Q P_1, \dots, P_k^T Q P_k$ are independent. Setting $Q = A_2^T F A_1$ we can write every point in the span of Z as $y^T F x = \sum_n \kappa_i y_i^T F x_i = \sum_n \kappa_i P_i^T A_2^T F A_1 P_i = \sum_n \kappa_i P_i^T Q P_i$ with $\kappa_i \in \mathbb{R}$ and by the independence of the first k equations $P_1^T Q P_1, \dots, P_k^T Q P_k$ we get

$$y^T F x = \sum_k \kappa_i P_i^T Q P_i = \sum_k \kappa_i y_i^T F x_i.$$

□

Let $P \in (\mathbb{P}^3)^{10}$ represent a configuration of ten points. The question whether the ten points of P are inscribable to a quaternary quadric has been studied in classical algebraic geometry. However, no geometric interpretation is known up till now and it probably would be too complicated to be of any use. In algebraic terms this condition can easily be phrased as $\det(\nu_{2,4}(P)) = 0$. Turnbull and Young give a description of the $PGL(3)$ invariant $\det(\nu_{2,4}(P)) = 0$ in the bracket algebra (Turnbull-Young invariant) [TY26]. Later the Turnbull-Young invariant has been straightened to a bracket polynomial of degree 5 with 138 monomials [Whi88]. Using the vertex labeling of Figure 6.2 the Turnbull-Young invariant given in [Whi88, p. 8-9] reduces to the bracket polynomial

$$\begin{aligned}&[\mathbf{0135}]\,[\mathbf{0247}]\,[\mathbf{1268}]\,[\mathbf{3469}]\,[\mathbf{5789}] - [\mathbf{0134}]\,[\mathbf{0257}]\,[\mathbf{1268}]\,[\mathbf{3569}]\,[\mathbf{4789}] + \\ &[\mathbf{0125}]\,[\mathbf{0346}]\,[\mathbf{1378}]\,[\mathbf{2479}]\,[\mathbf{5689}] - [\mathbf{0124}]\,[\mathbf{0356}]\,[\mathbf{1378}]\,[\mathbf{2579}]\,[\mathbf{4689}]\,,\end{aligned} \tag{6.2}$$

where $\mathbf{0}, \mathbf{1}, \cdots, \mathbf{9}$ denote the points in the configuration $[P, f_1, f_2] \in (\mathbb{P}^3)^{10}$.

We use this invariant to show that $\nu_{2,4}(P)$ is not of full rank if the points in P are the vertices of a combinatorial cube.

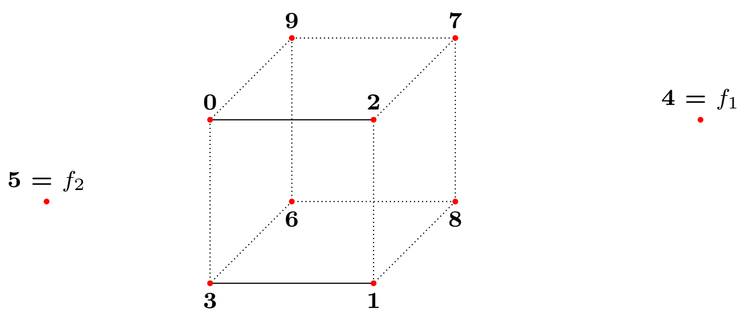


FIGURE 6.2: The chosen labeling of world points and camera centers.

Proposition 6.3. *Let P be the vertices of a combinatorial cube in $\mathbb{R}^3$, then the rank of $\nu_{2,4}(P)$ is at most seven.*

Proof. Consider the equation system $\nu_{2,4}([P, f_1, f_2])$. These are the equations $P_i^T Q P_i$ concatenated with the two equations $f_i^T Q f_i$ of two arbitrary points f_1, $f_2 \in \mathbb{P}^3$. If the rank of $\nu_{2,4}(P)$ is at most seven, then the rank of equation system $\nu_{2,4}([P, f_1, f_2]^T)$ is at most nine. Thus the ten points of the configuration $[P, f_1, f_2] \in (\mathbb{P}^3)^{10}$ are in special position and are inscribable to a quarternary quadric. This is equivalent to satisfying the Turnbull-Young invariant. We checked with `Macaulay2` [GS] that the polynomial of Equation 6.2 vanishes for all combinatorial cubes with the eight vertices $P = [\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{6}, \mathbf{7}, \mathbf{8}, \mathbf{9}]$ independent of the choice of the two points $f_1 = [\mathbf{4}]$ and $f_2 = [\mathbf{5}]$. □

Without using some speed-ups and simplifications solving Equation 6.2 is computationally out of reach. Thus we performed the computation in `Macaulay2` [GS] as follows. Since we are only interested in points in $\mathbb{R}^3$ we fixed the fourth coordinate of every point to one. Further we have the freedom of choice of a coordinate system in $\mathbb{R}^3$. We choose the point $\mathbf{0}$ as the origin and the points $\mathbf{3}$, $\mathbf{2}$, $\mathbf{9}$ as the three unit vectors. This implies that $\mathbf{1}$ is on the xy-plane, $\mathbf{5}$ is on the xz-plane and $\mathbf{7}$ is on the yz-plane. The `Macaulay2` [GS] code used to check Proposition 6.3 can be found in Listing 6.1.

Remark 6.4. Miquel's Theorem [Bob+08, p. 18] states: *Consider a combinatorial cube in $\mathbb{R}^3$ with planar faces. Assume that the vertices of three neighboring faces of the cube are circular, then the three circles determined by the triples of vertices corresponding to three remaining faces of the cube all intersect in one point.*

The point of intersection of the three circles is a vertex of the combinatorial cube. A combinatorial cube whose vertices of three neighboring faces are circular is sometimes called Miquel cube. Seven vertices are sufficient to define a Miquel cube. It is easy to see that all vertices of Miquel's cube lie on a sphere.

Proposition 6.5. *Let P be the vertices of a combinatorial cube in $\mathbb{R}^3$. If the quadric going through P, f_1, f_2 is a sphere, then Proposition 6.3 and Miquel's Theorem are equivalent.*

Proof. Consider the subset $P' = P \setminus P_i$ of the vertices of the combinatorial cube with the vertex P_i omitted. Since $\nu_{2,4}(P)$ is of rank seven at most the equation $P_i^T Q P_i = 0$ is linear dependent from the equations obtained from the points in P'. Hence if the vertices of P' are on a sphere, so is P_i on that sphere. Thus all vertices of the six faces of P are circular and Miquel's Theorem holds. On the other hand since all vertices of Miquel's cube lie on a sphere and the points vertices of P' are sufficient to define a Miquel cube, the matrix $\nu_{2,4}(P)$ is of rank at most seven. □

Due to Proposition 6.3 we understand the behavior of the 8-point algorithm if its input configurations X, Y are images of the vertices of a combinatorial cube.

Theorem 6.6. *Let A_1, A_2 be two arbitrary cameras and let $P \in (\mathbb{P}^3)^8$ be the vertices of a combinatorial cube. Then the 8-point algorithm with input $A_1 P, A_2 P$ fails to compute a fundamental matrix F. It is defeated by the vertices of P.*

Proof. By Proposition 6.3 the matrix $\nu_{2,4}(P)$ has at most rank seven, thus by Lemma 6.2 the matrix Z has at most rank seven. Hence the assumption in the 8-point algorithm that Z has at least rank eight is not satisfied. □

Theorem 6.6 states that if we take two pictures from the vertices of a combinatorial cube, then the matrix Z has at most rank seven and thus the 8-point algorithm is not able to compute the fundamental matrix.

Remark 6.7. Unlike to the conditions on critical configurations Theorem 6.6 does not impose any constraints on the camera centers.

6.3 Reconstruction of F From Cubes

Even if $\dim(\ker(Z)) = 2$ it is sometimes possible to reconstruct the fundamental matrix F. Since the matrix Z is of rank at most seven one can try to reconstruct the fundamental matrix by additionally enforcing the singularity condition of the fundamental matrix. Solving for F means finding the real roots of a univariate cubic polynomial. However for certain regions in $\mathbb{R}^3 \times \mathbb{R}^3$ of the two focal points, this polynomial has more than one solution and a unique reconstruction of F is not possible. These regions are semi-algebraic sets. We study these by first studying the simplest case, when P are the vertices of the *unit cube* $C_u = 1/2 \cdot \mathrm{conv}(\pm e_1 \pm e_2 \pm e_3)$.

For the three-dimensional unit cube C_u the matrix $Q_u \in \mathrm{Sym}_4(\mathbb{R})$ defining the family of quadrics through its vertices diagonalizes to $Q_u = \mathrm{diag}(\alpha, \beta, \gamma, \delta)$, $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, such that $\alpha + \beta + \gamma + \delta = 0$. This results in a two parameter family of quadrics running through the eight vertices of C_u. If we include the two camera centers then there is exactly one quadric Q running through all ten points and it is given as the solution of the linear equation system

$$\begin{aligned} \mathrm{tr}(Q) &= 0 \\ c_1^T Q c_1 &= 0 \\ c_2^T Q c_2 &= 0 \end{aligned} \quad \Leftrightarrow \quad \underbrace{\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1^2 & x_2^2 & x_3^2 & x_4^2 \\ y_1^2 & y_2^2 & y_3^2 & y_4^2 \end{bmatrix}}_{:=M} \begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} = 0 \tag{6.3}$$

where $f_1 = [x_1, x_2, x_3, x_4]$ and $f_2 = [y_1, y_2, y_3, y_4]$. By Cramer's rule we construct the solution of this linear equation system, as the vector of the four signed maximal minors of the matrix M,

$$\left[\begin{vmatrix} 1 & 1 & 1 \\ x_2^2 & x_3^2 & x_4^2 \\ y_2^2 & y_3^2 & y_4^2 \end{vmatrix}, \; -\begin{vmatrix} 1 & 1 & 1 \\ x_1^2 & x_3^2 & x_4^2 \\ y_1^2 & y_3^2 & y_4^2 \end{vmatrix}, \; \begin{vmatrix} 1 & 1 & 1 \\ x_1^2 & x_2^2 & x_4^2 \\ y_1^2 & y_2^2 & y_4^2 \end{vmatrix}, \; -\begin{vmatrix} 1 & 1 & 1 \\ x_1^2 & x_2^2 & x_3^2 \\ y_1^2 & y_2^2 & y_3^2 \end{vmatrix}\right]. \tag{6.4}$$

We denote by $f^{*2} \in \mathbb{P}^3$ the coordinate wise square of the vector $f \in \mathbb{P}^3$. Then Equation 6.4 can be written as the cross product of three vectors $(1,1,1,1)^T \times f_1^{*2} \times f_2^{*2}$.

Let $Q \in \mathrm{Sym}_4(\mathbb{R})$ be the quadric running through the vertices of a combinatorial cube C and the two camera centers f_1, f_2 . Further let $Q_u \in \mathrm{Sym}_4(\mathbb{R})$ be the quadric running through the vertices of the unit cube C_u and the cameras centers g_1, g_2.

Proposition 6.8. *If there is a projective transformation $T \in PGL(3)$ from C_u to C, then $(1,1,1,1)^T \times f_1^{*2} \times f_2^{*2}$ has the same sign pattern as $(1,1,1,1) \times Tg_1^{*2} \times Tg_2^{*2}$ and Q, Q_u have the same sign pattern.*

Proof. Let $P \in (\mathbb{P}^3)^{10}$ be the vertices of C together with f_1, f_2 and P_u be the vertices of C_u together with Tg_1, Tg_2.

$$(P_u^T Q_u P_u)_{ii} = 0 \,\forall\, i \Leftrightarrow (P^T T^T Q_u T P)_{ii} = 0 \,\forall\, i \Rightarrow T^T Q_u T = Q.$$

Now by Sylvester's law of inertia Q and Q_u have the same sign pattern. □

From Proposition 6.8 we are able to compute the type of the quadric Q by finding a projective transformation that maps the vertices of C to the vertices of C_u. In particular Q is ruled if $T^T Q_u T = Q$ is ruled.

Remark 6.9. If we interpret the vector of diagonal entries of Q as a point in $\mathbb{P}^3$ the condition $(1,1,1,1)^T \times f_1^{*2} \times f_2^{*2}$ is equivalent to $[\alpha, \beta, \gamma, \delta]$ being on the intersection of the three planes with normal vectors $[1,1,1,1]$, f_1^{*2}, $f_2^{*2} \in \mathbb{P}^3$.

The quadric is ruled if $[\alpha, \beta, \gamma, \delta]$ has a sign pattern of the following types $[-,-,+,+]$, $[+,-,+,-]$, or $[-,+,+,-]$. The boundaries of the components of the semi-algebraic set (where the signature of the quadric changes) are given as the vanishing set of the determinants α , β, γ and δ of Equation 6.4.

In some cases the signature changes, but still the quadric stays non-ruled. Therefore, to get a more explicit answer for this example it is useful to break up the symmetry of Equation 6.4. Since the quadric is independent on a scaling factor of Q we can set the last diagonal entry without loss of generality to $\delta = 1$. Thus the equation system of Equation 6.3 degenerates to an equation system of three equations in three variables and we can solve it explicitly. Then $Q = \mathrm{diag}(\alpha, \beta, -\alpha - \beta - 1, 1)$. There are two distinct cases when the quadric is ruled:

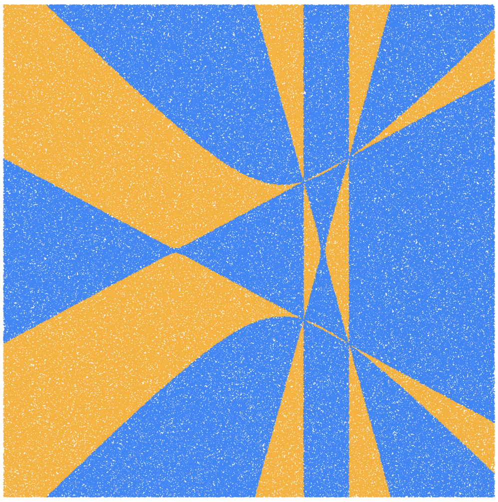


FIGURE 6.3: Region of failure (orange) of the 8-point algorithm (enforcing the singularity constraint) with a unit cube as input. We fixed the focal point of the first camera and randomly sampled the focal point of the second camera in a chosen plane.

1. If $\alpha, \beta \leq 0$ and $\alpha + \beta \leq -1$.

2. If α, β have different signs and $\alpha + \beta \geq 1$.

The vector of diagonal entries of Q then is given up to scale by

$$\begin{bmatrix} \begin{vmatrix} x_2^2 - x_3^2 & x_3^2 - 1 \\ y_2^2 - y_3^2 & y_3^2 - 1 \end{vmatrix} \\ \begin{vmatrix} x_1^2 - x_2^2 & x_3^2 - 1 \\ y_1^2 - y_2^2 & y_3^2 - 1 \end{vmatrix} \\ -\begin{vmatrix} x_2^2 - x_3^2 & x_3^2 - 1 \\ y_2^2 - y_3^2 & y_3^2 - 1 \end{vmatrix} - \begin{vmatrix} x_1^2 - x_2^2 & x_3^2 - 1 \\ y_1^2 - y_2^2 & y_3^2 - 1 \end{vmatrix} - \begin{vmatrix} x_1^2 - x_2^2 & x_2^2 - x_3^2 \\ y_1^2 - y_2^2 & y_2^2 - y_3^2 \end{vmatrix} \\ \begin{vmatrix} x_1^2 - x_2^2 & x_2^2 - x_3^2 \\ y_1^2 - y_2^2 & y_2^2 - y_3^2 \end{vmatrix} \end{bmatrix} \in \mathbb{P}^3$$

6.4 How to Handle Pictures of Cubes

In the sections above we did not consider noisy pictures. As seen in Theorem 6.6 the rank of Z drops by at least one if we take pictures of combinatorial cubes. However, in the presence of noise the matrix $Z \in \mathbb{R}^{8\times 9}$ again is of rank eight, but it is close to being singular. This results in a very bad performance of the algorithm in practice and we strongly advise against using it. It is simply the wrong choice of algorithm, since it is incapable of dealing with this set-up.

Input: Two image point configurations $X, Y \in (\mathbb{P}^2)^n$;
Output: The fundamental matrix F;
begin

1. Normalize X, Y
2. Compute $Z \in \mathbb{R}^{8\times 9}$ from X, Y;
3. Compute Z' that minimizes the Frobenius norm of $|Z - Z'|$.
   $Z' := U \operatorname{diag}(\sigma_1, \dots \sigma_7, 0, 0)V^T$.
4. Compute the two generators f_1 and f_2 of $\ker(Z')$ and solve
   $\det(\alpha F_1 + (1-\alpha)F_2) = 0$

if $\det(\alpha F_1 + (1-\alpha)F_2)$ has multiple real roots **then**
  | Choose the solution that minimizes the residual error on X, Y.
end

end

Algorithm 4: Cube-8-point algorithm.

As discussed in Section 6.2 the matrix Z has at most rank seven in the noisefree case, hence a standard implementation of the 7-point algorithm run on a 7-element subset of vertices can be used to retrieve the fundamental matrix. Thus a natural fix for the flaw of the 8-point algorithm is to use a modified version of the 7-point algorithm, that allows eight points as inputs. We use singular value decomposition on Z to obtain a matrix $Z' \in \mathbb{R}^{8\times 9}$ that is of rank seven and minimizes the Frobenius norm of $|Z - Z'|$. Let $Z = UDV^T$ be the singular value decomposition of Z, then by the Eckart-Young-Mirsky theorem $Z' = U \operatorname{diag}(\sigma_1, \dots \sigma_7, 0, 0)V^T$. Now we use the 7-point algorithm to obtain (one to three) possible solutions for the fundamental matrix. If there are multiple solutions, we chose the one that minimizes the residual error on the input points. The 8-point algorithm for cubes is given in Algorithm 4.

6.5 Numerical Experiments

We performed random tests on synthetic data to compare the performance of different algorithms. To do so, we sampled random cubes within the box $[-1, 1]^3$. The cameras were chosen with focal points roughly on a sphere with radius six. Gaussian noise was applied onto the images with standard deviation between $0\% - 10\%$ of the image sizes and zero mean. For each noise level we chose 2000 random samples and respectively computed 2000 approximations of fundamental matrices. As a measure to analyze the results of the algorithm we used the metric on the Grassmanian between two linear subspaces, namely the angle between the vectorization of the true fundamental matrix and the approximated fundamental matrix.

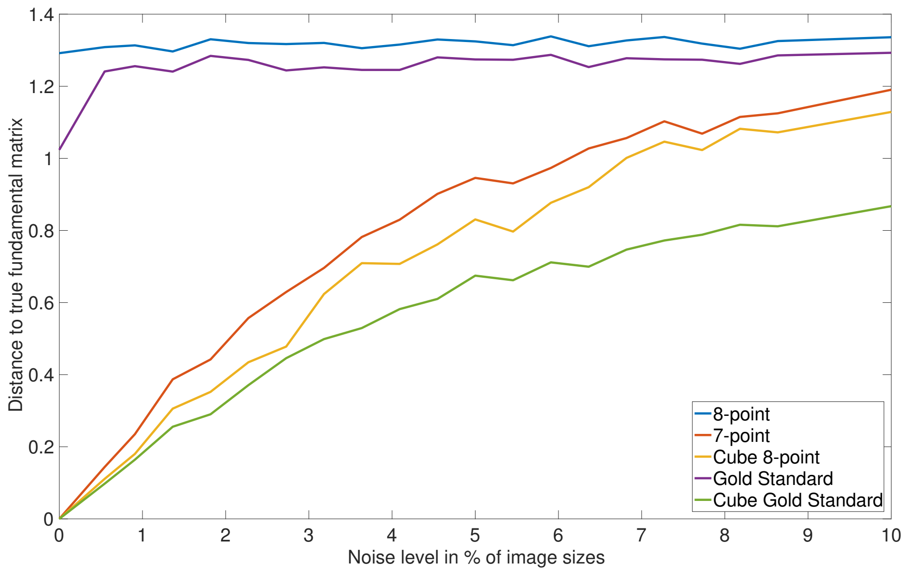


FIGURE 6.4: Comparison between algorithms to reconstruct the fundamental matrix. Plotted is the distance as angle (rad) between two one-dimensional subspaces versus the noise level.

The modified version of the 8-point algorithm for cubes (Algorithm 4) gives good results. Its running time is almost the same as of the usual 7-point algorithm, but unlike the unmodified version it is noise correcting. It also gives better results than algorithms that rely on non-linear optimization to estimate the fundamental matrix, like the *Gold Standard algorithm* [HZ03, Algorithm 11.3]. These algorithms usually use an initial guess of the fundamental matrix computed via the 8-point algorithm. But the results of the 8-point algorithm are so far off from the true fundamental matrix, such that the non-linear solvers get stuck in a local optimum far off from the global one. However by using estimates of the fundamental matrix computed with the modified 8-point algorithm for cubes (Algorithm 4) the Gold Standard algorithm can be improved. For example using an estimate of the fundamental matrix computed with Algorithm 4 as initial input, instead of a fundamental matrix computed with the 8-point algorithm [HZ03, Algorithm 11.1], improves the Gold Standard algorithm [HZ03, Algorithm 11.3]. In Figure 6.4 this version of the Gold Standard algorithm is denoted by *Cube Gold Standard*.

Note that there are also global solvers to find fundamental matrices based on semidefinite programming [Bug+15]. Noteworthy the situation in our case is different from the one depicted [Bug+15]. Usually there are more than 8 correspondent image point pairs available. In [Bug+15] ten points and more are considered. Thus due to noisy data the matrix Z is of rank nine. For the pathological case of only eight points and a rank drop in Z, Algorithm 1 in [Bug+15] has not been able to certify global optimality solutions based on

`GloptiPoly 3` [HLL09].

If one has the freedom of choice to place the eight points in $\mathbb{P}^3$ we suggest using the skew octagon for a more robust reconstruction of the fundamental matrix. The skew octagon is computational the optimal solution to various sphere placement problems, e.g. the Thomson problem see [Bro17].

6.6 Computations

Below you can find the code we used to check Proposition 6.3. Lines 2-12 define the vertices of a cube. Lines 16-19 define the facets of the cube. In lines 24-36 the reduced Turnbull-Young invariant of Equation 6.2 is defined. This invariant vanishes in the quotient ring $S = R/J$ of line 21.

```
R=QQ[x_1..x_15]
Cube=matrix{
    {0,0,0,1},              --0
    {x_1,x_2,0,1},          --1
    {0,1,0,1},              --2
    {1,0,0,1},              --3
    {x_10,x_11,x_12,1},     --4 f_1
    {x_13,x_14,x_15,1},     --5 f_2
    {x_5,0,x_6,1},          --6
    {0,x_3,x_4,1},          --7
    {x_7,x_8,x_9,1},        --8
    {0,0,1,1}}              --9

-- the six facets of the cube are coplanar
-- their determinants vanish
J=ideal(
det submatrix(Cube,{0,1,2,3},),det submatrix(Cube,{6,7,8,9},),
det submatrix(Cube,{0,3,6,9},),det submatrix(Cube,{1,2,7,8},),
det submatrix(Cube,{0,2,7,9},),det submatrix(Cube,{1,3,6,8},))

S=R/J
F=map(S,R)
-- the remaining 4 bracket monomials of the Turnbull-Young invariant
p=F(
det submatrix(Cube,{0,1,3,5},)*det submatrix(Cube,{0,2,4,7},)*
det submatrix(Cube,{1,2,6,8},)*det submatrix(Cube,{3,4,6,9},)*
det submatrix(Cube,{5,7,8,9},)-
det submatrix(Cube,{0,1,3,4},)*det submatrix(Cube,{0,2,5,7},)*
det submatrix(Cube,{1,2,6,8},)*det submatrix(Cube,{3,5,6,9},)*
det submatrix(Cube,{4,7,8,9},)+
det submatrix(Cube,{0,1,2,5},)*det submatrix(Cube,{0,3,4,6},)*
det submatrix(Cube,{1,3,7,8},)*det submatrix(Cube,{2,4,7,9},)*
det submatrix(Cube,{5,6,8,9},)-
det submatrix(Cube,{0,1,2,4},)*det submatrix(Cube,{0,3,5,6},)*
det submatrix(Cube,{1,3,7,8},)*det submatrix(Cube,{2,5,7,9},)*
det submatrix(Cube,{4,6,8,9},))
```

LISTING 6.1: Vanishing of Turnbull-Young Invariant.

Chapter 7

Veronesean Almost Binomial Almost Complete Intersections

7.1 Introduction

Ideals generated by minors of matrices are a mainstay of commutative algebra. Here we are concerned with ideals generated by 2-minors of generic-symmetric matrices. Ideals generated by arbitrary minors of generic-symmetric matrices have been studied by Kutz [Kut74] who proved, in the context of invariant theory, that the quotient rings are Cohen–Macaulay. Results of Goto show that the quotient ring is normal with divisor class group $\mathbb{Z}_2$ and Gorenstein if the format of the generic-symmetric matrix has the same parity as the size of the minors [Got77; Got79]. Conca extended these results to more general sets of minors of generic-symmetric matrices [Con94a] and determined Gröbner bases and multiplicity [Con94b].

Here we are concerned only with the binomial ideal I_n generated by the 2-minors of a *generic-symmetric* $(n \times n)$-matrix, which is a matrix with indeterminates as entries. This ideal cuts out the second Veronese variety and was studied classically, for example by Gröbner [Grö65]. It contains a complete intersection J_n generated by the principal 2-minors (Definition 7.3). Coming from liaison theory one may ask for the ideal $K_n = J_n : I_n$ on the other side of the complete intersection link via J_n. In this chapter we determine K_n.

Example 7.1. *Consider the ideal* $J_3 = \langle ad-b^2, af-c^2, df-e^2\rangle \subset \mathbb{Q}[a,b,c,d,e,f]$ *generated by the principal* 2*-minors of the generic-symmetric matrix* $\left(\begin{smallmatrix} a & b & c \\ b & d & e \\ c & e & f \end{smallmatrix}\right)$. *It is easy to check, for example with* `Macaulay2` *[GS], that* J_3 *is a complete intersection and has a prime decomposition* $J_3 = I_3 \cap K_3$ *where*

$$I_3 = J_3 + \langle ae - bc, cd - be, ce - bf\rangle$$

is the second Veronese ideal, generated by all 2*-minors, and*

$$K_3 = J_3 + \langle ae + bc, cd + be, ce + bf\rangle$$

is the image of I_3 under the automorphism of $\mathbb{Q}[a, \ldots, f]$ that maps $b, c,$ and e to their negatives and the remaining indeterminates to themselves. As a very special case of Theorem 7.11 we find that the generator $ae + bc$ is the generating function of the fiber

$$\left\{ u \in \mathbb{N}^6 : \begin{pmatrix} 2 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{pmatrix} \cdot u = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right\}$$

of the $\mathbb{Z}$-linear map V_3 that defines the fine grading of $\mathbb{Q}[a, \ldots, f]/I_3$. For $n \geq 4$ the extra generators are not binomials anymore and K_n is an intersection of ideals obtained from I_n by twisting automorphisms (Definition 7.7). In Example 7.12, for $n = 4$, we find $K_4 = J_4 + \langle p \rangle$ for one quartic polynomial p with eight terms.

General results on liaison theory of ideals of minors of generic-symmetric matrices have been obtained by Gorla [Gor07; Gor10]. Our methods rely on the combinatorics of binomial ideals and since K_n is not binomial we cannot explore the linkage class with the present method. Instead we are motivated by general questions about binomial ideals and their intersections. For example, [KM14, Problem 17.1] asks when the intersection of binomial ideals is binomial. From the primary (in fact, prime) decomposition of J_n we remove I_n and intersect the remaining binomial prime ideals. The result is not binomial. If n is even, $K_n = J_n + \langle p \rangle$ for one additional polynomial p. In the terminology of [Ber63], K_n is thus an *almost complete intersection*. It is also *almost binomial*, as it is principal modulo its binomial part [JKK16, Definition 2.1]. If n is odd, then there are n additional polynomials (Theorem 7.11). While these numbers can be predicted from general liaison theory, our explicit formulas reveal interesting structures at the boundary of binomiality and are thus a first step towards [KM14, Problem 17.1]

We determine K_n with methods from combinatorial commutative algebra, multigradings in particular (see [MS05, Chapters 7 and 8]). The principal observation that drives the proofs in Section 7.2 is that the Veronese-graded Hilbert function of the quotient $\mathbb{K}[\mathbf{x}]/J_n$ becomes eventually constant (Remark 7.14). The eventual value of the Hilbert function bounds the number of terms that a graded polynomial can have. The extra generators of K_n are the lowest degree polynomials that realize the bound. We envision that this structure could be explored independently and brought to unification with the theory of toral modules from [DMM10]. Our results also have possible extensions to higher Veronese ideals as we outline in Section 7.3.

Denote by $c_n := \binom{n}{2}$ the entries of the second diagonal in Pascal's triangle. Throughout, let $[n] := \{1, \ldots, n\}$ be the set of the first n integers. The second Veronese ideal lives in the polynomial ring $\mathbb{K}[\mathbb{N}^{c_{n+1}}]$ in c_{n+1} indeterminates over a field $\mathbb{K}$. For polynomial rings and quotients modulo binomial ideals we use monoid algebra notation (see, for instance, [KM14, Definition 2.15]). We make no apriori assumptions on $\mathbb{K}$ regarding its characteristic or algebraic closure, although care is necessary in characteristic two. The variables of $\mathbb{K}[\mathbb{N}^{c_{n+1}}]$ are denoted x_{ij}, for $i, j \in [n]$ with the implicit convention that $x_{ij} = x_{ji}$. For brevity we avoid a comma between i and j. We usually think about upper triangular matrices, that is $i \leq j$. The Veronese ideal I_n is the toric ideal of the *Veronese multigrading* $\mathbb{N}V_n$, defined by the $(n \times c_{n+1})$-matrix V_n with entries

$$(V_n)_{i,jk} := \begin{cases} 2 & \text{if } i = j = k, \\ 1 & \text{if } i = j, \text{ or } i = k, \text{ but not both}, \\ 0 & \text{otherwise.} \end{cases}$$

That is, the columns of V_n are the non-negative integer vectors of length n and weight two. For $\mathbf{b} \in \mathbb{N}V_n$, the *fiber* is $V_n^{-1}[\mathbf{b}] = \{u \in \mathbb{N}^{c_{n+1}} : V_n u = \mathbf{b}\}$. Computing the V_n-degree of a monomial is easy: just count how often each row or column index appears in the monomial. For example, $\deg(x_{12}x_{nn}) = (1, 1, 0, \ldots, 0, 2)$. We do not distinguish row and column vectors notationally, in particular we write columns as rows when convenient. Gröbner bases for a large class of toric ideals including I_n have been determined by Sturmfels [Stu96, Theorem 14.2]. The *Veronese lattice* $L_n \subset \mathbb{Z}^{c_{n+1}}$ is the kernel of V_n. The rank of L_n is c_n since the rank of V_n is n and $c_{n+1} - n = c_n$. Lemma 7.2 gives a lattice basis. With $\{e_{ij}, i \leq j \in [n]\}$ a standard basis of $\mathbb{Z}^{c_{n+1}}$, we use the following notation

$$[ij|kl] := e_{ik} + e_{jl} - e_{il} - e_{jk} \in \mathbb{Z}^{c_{n+1}}.$$

Then $[ij|kl]$ is the exponent vector of the minor $x_{ik}x_{jl} - x_{il}x_{jk}$.

7.2 Decomposing and Recomposing

Lemma 7.2. *The set*

$$\mathcal{B} = \{[in|jn] : i, j \in [n-1]\}$$

is a lattice basis of the Veronese lattice L_n.

Proof. Write the elements of $\mathcal{B}$ as the columns of a $(c_{n+1} \times c_n)$-matrix B. Deleting the rows corresponding to indices (i, n) for $i \in [n]$ yields the identity matrix I_{c_n}. Thus $\mathcal{B}$ spans a lattice of the correct rank and that lattice is saturated. Indeed, the Smith normal form of B must equal the identity matrix I_{c_n} concatenated with a zero matrix. Thus the quotient by the lattice spanned by $\mathcal{B}$ is free. $\square$

The Veronese ideal contains a codimension c_n complete intersection J_n generated by the principal 2-minors.

Definition 7.3. *The* principal minor ideal J_n *is generated by all principal* 2-*minors* $x_{ii}x_{jj} - x_{ij}^2$ *of a generic-symmetric matrix. The* principal minor lattice L'_n *is the lattice generated by the corresponding exponent vectors* $[ij|ij]$, $i, j \in [n]$.

It can be seen that the principal minor lattice is minimally generated by $[ij|ij]$. It is an unsaturated lattice meaning that it cannot be written as the kernel of an integer matrix, or equivalently, that the quotient $\mathbb{Z}^{c_{n+1}}/L'_n$ has torsion. Since there are no non-trivial coefficients on the binomials in J_n, Proposition 7.5 below says that it is a lattice ideal with lattice L'_n. The twisted group algebra in [KM14, Definition 10.4] is just a group algebra. Its torsion subgroup is given in the following proposition.

Proposition 7.4. *The principal minor lattice is minimally generated by*

$$\mathcal{B}' = \{2[in|jn] : i \neq j \in [n-1]\} \cup \{[in|in] : i \in [n-1]\}.$$

Furthermore the group L_n/L'_n *is (isomorphic to)* $(\mathbb{Z}/2\mathbb{Z})^{c_{n-1}}$.

Proof. It holds that $2[in|jn] = [in|in] + [jn|jn] - [ij|ij]$ and the map which includes the span of the elements $[ij|ij]$ into L'_n is unimodular. A presentation of the group can be read off the Smith normal form of the matrix whose columns are a lattice basis. Since $\mathcal{B}'$ is a basis of L'_n, the columns and rows can be arranged so that the diagonal matrix $\operatorname{diag}(2, \ldots, 2, 1, \ldots, 1)$ with c_{n-1} entries 2 is the top $(c_n \times c_n)$-matrix of the Smith normal form. Any entry below a two is divisible by two and thus row operations can be used to zero out the bottom part of the matrix. This yields the Smith normal form. □

The difference between the basis in Definition 7.3 and $\mathcal{B}'$ is that the transition matrix from $\mathcal{B}$ to $\mathcal{B}'$ is diagonal. Thus it is easy to describe the *half open fundamental parallelepiped* of $\mathcal{B}'$.

If $\operatorname{char}(\mathbb{K}) = 2$, then J_n is primary over I_n. In all other characteristics one can see that the Veronese ideal I_n is a minimal prime and in fact a primary component of J_n. These statements follow from [ES96] and are summarized in Proposition 7.10 below. Towards this observation, the next proposition says that J_n is a mesoprime ideal.

Proposition 7.5. J_n *is a mesoprime binomial ideal and its associated lattice is* L'_n.

Proof. According to [KM14, Definition 10.4] we show that $J_n = \langle x^{u^+} - x^{u^-} : u \in L'_n \rangle$, since the quotient by this ideal is the group algebra $\mathbb{K}[\mathbb{Z}^{c_{n+1}}/L'_n]$. By the correspondence between non-negative lattice walks and binomial ideals [DES98, Theorem 1.1] we prove that for any $u = u^+ - u^- \in L'_n$, the parts $u^+, u^- \in \mathbb{N}^{c_{n+1}}$ can be connected using *moves* $[ij|ij]$ without leaving $\mathbb{N}^{c_{n+1}}$.

The vectors u^+, u^- can be represented by upper triangular non-negative integer matrices. From Definition 7.3 it is obvious that all off-diagonal entries of $u^+ - u^-$ are divisible by two. Since

$$\langle x^{u^+} - x^{u^-} : u \in L'_n \rangle : x_{ij} = \langle x^{u^+} - x^{u^-} : u \in L'_n \rangle$$

for any variable x_{ij}, we can assume that u^+ and u^- have disjoint supports and thus individually have off-diagonal entries divisible by two. Consequently the moves $[ij|ij]$ allow to reduce all off-diagonal entries to zero, while increasing the diagonal entries. As visible from its basis, the lattice L'_n contains no nonzero diagonal matrices and thus u^+ and u^- have been connected to the same diagonal matrix. □

Remark 7.6. From Proposition 7.4 it follows immediately that the group algebra $\mathbb{K}[\mathbb{Z}^{c_{n+1}}]/J_n\mathbb{K}[\mathbb{Z}^{c_{n+1}}]$ is isomorphic to $\mathbb{K}[\mathbb{Z}^n \oplus (\mathbb{Z}/2\mathbb{Z})^{c_{n-1}}]$. In particular $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ is finely graded by the monoid $\mathbb{N}V_n \oplus (\mathbb{Z}/2\mathbb{Z})^{c_{n-1}}$.

Definition 7.7. *A* $\mathbb{Z}_2$-twisting *is a ring automorphism of a (Laurent) polynomial ring that maps the indeterminates either to themselves or to their negatives.*

The lattice points in a fundamental parallelepiped of L'_n play an important role in the following developments. The most succinct way to encode them is using their generating function (which in this case is simply a polynomial in the Laurent ring $\mathbb{K}[\mathbb{Z}^{c_{n+1}}]$). The explicit form, of course, depends on the coordinates chosen. The next lemma is immediate from the definition of $\mathcal{B}'$.

Lemma 7.8. *Let* $M = \{[in|jn] : i \neq j \in [n-1]\}$*. The generating function of the fundamental parallelepiped of* $\mathcal{B}'$ *is*

$$p_n = \prod_{m \in M} (x^m + 1) = \prod_{m \in M} (x^{m^+} + x^{m^-})$$

It is useful for the further development to pick the second representation of p_n in Lemma 7.8 as a representative of p_n in polynomial ring $\mathbb{K}[\mathbb{N}^{c_{n+1}}]$. Its image in the quotient $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ also has a natural representation. The terms of p_n can be identified with upper triangular integer matrices which arise as sums of positive and negative parts of elements of M. A positive part of $[in|jn] \in M$ has entries 1 at positions (i,j) and (n,n) while a negative part has two entries 1 in the last column, but not at (n,n). Modulo the moves $\mathcal{B}'$, any exponent matrix of a monomial of p_n can be reduced to have only entries 0 or 1 in its off-diagonal positions.

Remark 7.9. A simple count yields that p_n has V_n-degree $(n-2,\dots,n-2,2c_{n-1})$. In the natural representation of monomials of p_n as integer matrices with entries 0/1 off the diagonal, there is a lower bound for the value of the (n,n) entry. To achieve the lowest value, one would fill the last column with entries 1 using negative parts of elements of M, and then use positive parts (which increase (n,n)). For example, if n is even, there is one term of p_n whose last column arises from the negative parts of $[1n|2n],[3n|4n],\dots,[(n-3)n|(n-2)n]$ and then positive parts of the remaining elements of M. If n is odd, then there is one term of p_n, whose n-th column is $(1,\dots,1,\sigma_{n-1})$ for some value σ_{n-1}. In fact, since $|M| = c_{n-1}$, the lowest possible value of the (n,n) entry is $\sigma_{n-1} = c_{n-1} - \lfloor \frac{n-1}{2} \rfloor$.

The primary decomposition of J_n is given by [ES96, Theorem 2.1 and Corollary 2.2].

Proposition 7.10. *If* $\operatorname{char}(\mathbb{K}) = 2$*, the* J_n *is primary over* I_n*. In all other characteristics, there exists* $2^{c_{n-1}}$ $\mathbb{Z}_2$*-twisting* ϕ_i *with* $i \in [c_{n-1}]$*, such that the complete intersection* J_n *has prime decomposition*

$$J_n = \bigcap_i \phi_i(I_n). \tag{7.1}$$

Theorem 7.11. *If* n *is odd intersecting all but one of the components in* (7.1) *yields*

$$\bigcap_{i\neq l} \phi_i(I_n) = J_n + \langle \phi_l(p^+_{n,i}) : i \in [n] \rangle,$$

where $p^+_{n,i} \in \mathbb{K}[\mathbb{N}^{c_{n+1}}]$ *are homogeneous polynomials of degree* $\frac{(n-1)^2}{2}$ *that are given as generating functions of the fibers* $V_n^{-1}[(n-2,\dots,n-2)+e_i]$*. If* n *is even, then the same holds for a single polynomial* p^+_n *of degree* $\frac{n(n-2)}{2}$*, given as the generating function of* $V_n^{-1}[(n-2,\dots,n-2)]$*.*

The proof of Theorem 7.11 occupies the remainder of the section after the following example.

Example 7.12. *The complete intersection* J_4 *is a lattice ideal for the lattice* L'_4*. In the basis* $\mathcal{B}'$*, it is generated by the six elements*

$$\{2[i4|j4] : i < j \in [3]\} \cup \{[i4|i4] : i \in [3]\}.$$

Three of the six elements correspond to principal minors:

$$x_{11}x_{44} - x_{14}^2, x_{22}x_{44} - x_{24}^2, x_{33}x_{44} - x_{34}^2$$

The other elements give the binomials

$$x_{12}^2x_{44}^2 - x_{14}^2x_{24}^2, x_{13}^2x_{44}^2 - x_{14}^2x_{34}^2, x_{23}^2x_{44}^2 - x_{24}^2x_{34}^2$$

These six binomials do not generate J_4, but J_4 equals the saturation with respect to the product of the variables [MS05, Lemma 7.6]. The $2^3 = 8$ minimal prime components of J_n are obtained by all possible twist combinations of the monomials $\pm x_{14}x_{24}$, $\pm x_{14}x_{34}$, $\pm x_{24}x_{34}$. Consider the polynomial

$$p_4 = (x_{12}x_{44} + x_{14}x_{24})(x_{13}x_{44} + x_{14}x_{34})(x_{23}x_{44} + x_{24}x_{34}),$$

which is the generating function of the fundamental parallelepiped of L'_4 in the basis $\mathcal{B}'$ and of V_4-degree $(2, 2, 2, 6)$. In the Laurent ring $\mathbb{K}[\mathbb{Z}^{10}]$, *the desired ideal $J_4 : I_4$ equals $J_4 + \langle p_4 \rangle$. To do the computation in the polynomial ring, we need to saturate with respect to $\prod_{ij} x_{ij}$. If n is even, this saturation generates one polynomial, if n is odd, it generates n polynomials. Here, where $n = 4$, the ideal $J_4 : I_4$ is generated by J_4 and the single polynomial*

$$\begin{aligned} p_4^+ &= x_{11}x_{22}x_{33}x_{44} + x_{11}x_{23}x_{24}x_{34} + x_{13}x_{14}x_{22}x_{34} + x_{12}x_{14}x_{24}x_{33} \\ &\quad + x_{13}x_{14}x_{23}x_{24} + x_{12}x_{14}x_{23}x_{34} + x_{12}x_{13}x_{24}x_{34} + x_{12}x_{13}x_{23}x_{44}. \end{aligned}$$

Modulo the binomials in J_4, the polynomial p_4^+ equals p_4/x_{44}^2 (Lemma 7.18).

As a first step towards the proof of Theorem 7.11 we compute the monoid Q under which $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ is finely graded, meaning that its Q-graded Hilbert function takes values only zero or one. That is, we make Remark 7.6 explicit.

Lemma 7.13. *Fix $\mathbf{b} \in cone(V_n)$ for some n. The equivalence classes of lattice points in the fiber $V_n^{-1}[\mathbf{b}]$, modulo the moves $\mathcal{B}'$, are in bijection with set of 0/1 matrices $u \in \{0, 1\}^{c_{n+1}}$ of the following form*

- *$u_{ii} = 0$, for all $i \in [n]$*
- *$u_{in} = 0$, for all $i \in [n]$*
- *$\mathbf{b} - V_n u \in \mathbb{N}^n$.*

Proof. Each equivalence class of upper triangular matrices has a representative whose off-diagonal entries are all either zero or one. The bijection maps such an equivalence class to the c_{n-1} entries that are off-diagonal and off the last column. To prove that this is a bijection it suffices to construct the inverse map. To this end, let u satisfy the properties in the statement. In each row $i = 1, \ldots, n$, there are two values unspecified: the diagonal entry and the entry in the last column. Given $\mathbf{b}_i$, and using the choice of representative modulo $\mathcal{B}'$ whose last column entries is either 0 or 1, fixes the diagonal entry too by linearity. Therefore the map is a bijection. □

Remark 7.14. If $\mathbf{b}_i \geq (n - 2)$ for all $i \in [n]$, then any 0/1 upper triangular $(n - 2)$-matrix is a possible choice for the off-diagonal off-last column entries of

u in Lemma 7.13. An upper triangular $(n-2)$-matrix has c_{n-1} entries. Thus all those fibers have equivalence classes modulo $\mathcal{B}'$ that are in bijection with $\{0,1\}^{c_{n-1}}$. In particular, each of those fibers, has the same number of equivalence classes.

Remark 7.15. Remark 7.14 implies that in the V_n-grading, $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ is *toral* as in [DMM10, Definition 4.3]: its V_n-graded Hilbert function is globally bounded by $2^{c_{n-1}}$.

If n is odd, then $(n-2,\dots,n-2) \notin \mathbb{N}V_n$. Therefore the minimal (with respect to addition in the semigroup $cone(V_n)$) fibers that satisfy Remark 7.14 are $(n-1,n-2,\dots,n-2),\dots,(n-2,\dots,n-2,n-1)$. If n is even, there is only one minimal fiber.

For the proof of Theorem 7.11 it is convenient to work in the quotient ring $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$. Since $I_n \supseteq J_n$ and I_n is finely graded by $\mathbb{N}V_n$, each equivalence class is contained in a single fiber $V_n^{-1}[\mathbf{b}]$ and each fiber breaks into equivalence classes. The following definition sums the monomials corresponding to these classes for specific fibers.

Definition 7.16. *The* minimal saturated fibers *are the minimal fibers that satisfy Remark 7.14. The generating function of the equivalence class in a minimal saturated fiber is denoted by $p^+_{n,i}$. That is*

$$p^+_{n,i} = \sum_{\mathbf{a}\in V_n^{-1}[\mathbf{b}_i]/L'_n} \mathbf{x}^{\mathbf{a}} \in \mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n.$$

where $\mathbf{b}_i := (n-2,\dots,n-2)+e_i$ if n is odd and $\mathbf{b}_i = (n-2,\dots,n-2)$ if n is even.

If n is even, Definition 7.16 postulates only one polynomial which is simply denoted p^+_n when convenient. Sometimes, however, it can be convenient to keep the indices.

Remark 7.17. The construction of a generating function of equivalence classes of elements of the fiber in Definition 7.16 can be carried out for any fiber of V_n. For the fiber $(n-2,\dots,n-2,2c_{n-1})$ we get the polynomial p_n from Lemma 7.8.

The quantity $\sigma_{n-1} = c_{n-1} - \lfloor\frac{n-1}{2}\rfloor$ (that is $c_{n-1} - \frac{n-2}{2} = \frac{(n-2)^2}{2}$ for even n, and $c_{n-1} - \frac{n-1}{2} = \frac{(n-1)(n-3)}{2}$ for odd n) appeared in Remark 7.9 and shows up again in the next lemma: it almost gives the saturation exponent when passing from the Laurent ring to the polynomial ring.

Lemma 7.18. *As elements of $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$, if n is even then, $x_{nn}^{\sigma_{n-1}}p^+_{n,i} = p_n$, and if n is odd, then, $x_{nn}^{\sigma_{n-1}+1}p^+_{n,i} = x_{in}p_n$.*

Proof. If n is even, the product $x_{nn}^{\sigma_{n-1}}p^+_{n,i}$ has V_n-degree $(n-2,\dots,n-2,2c_{n-1})$. If n is odd, the degree of $x_{nn}^{\sigma_{n-1}+1}p^+_{n,i}$ equals $(n-2,\dots,n-2,2c_{n-1}+1)+e_i$.

Now these products equal p_n if n is even and $x_{in}p_n$ if n is odd by Remarks 7.14 and 7.17. □

Lemma 7.19. *If n is odd, then for any triple of distinct indices $i, j, k \in [n]$, in $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ we have $x_{ij}p^+_{n,k} = x_{jk}p^+_{n,i}$.*

Proof. By Proposition 7.5, the variables are nonzerodivisors on $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$. The multidegree of $p^+_{n,k}$ satisfies the conditions of Remark 7.14, thus there are bijections between the monomials of $x_{ij}p^+_{n,k}$ and $x_{jk}p^+_{n,i}$. Since all relations in J_n are equalities of monomials, multiplication with a variable does not touch coefficients. □

The following lemma captures an essential feature of our situation. Since the V_n-graded Hilbert function of $\mathbb{K}[\mathbb{N}^{c_{n+1}}]$ is globally bounded, there is a notion of *longest homogeneous polynomial* as one that uses all monomials in a given V_n-degree. If such a polynomial is multiplied by a term, it remains a longest polynomial.

Lemma 7.20. *The V_n-graded Hilbert functions of the $\mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$-modules, $\langle p_n \rangle$ and $\langle p^+_{n,i} \rangle$, $i = 1, \ldots, n$ take only zero and one as their values.*

Proof. We only prove the statement for $\langle p_n \rangle$ since the same argument applies also to $\langle p^+_{n,i} \rangle$. The claim is equivalent to the statement that any $f \in \langle p_n \rangle$ is a term (that is, a monomial times a scalar) times p_n. Let $f = gp_n$ with a V_n-homogeneous g. Let $t_1, \ldots, t_s$ be the terms of g. Since p_n is the sum of all monomials of degree $\deg(p_n)$, and multiplication by a term does not produce any cancellation, the number of terms of t_ip_n equals that of p_n. By Remark 7.14, the monomials in degree $\deg(t_ip_n)$ are in bijection with the monomials in degree $\deg(p_n)$, and therefore all t_ip_n are scalar multiples of the generating function of the fiber for $\deg(t_ip_n)$. □

Lemma 7.21. *For any $i \in [n]$, $\langle p^+_{n,i} \rangle : \left(\prod_{ij} x_{ij}\right)^\infty = \langle p^+_{n,k} : k \in [n] \rangle$.*

Proof. If n is odd, the containment of $p^+_{n,k}$ in the left hand side follows immediately from Lemma 7.19. If n is even, it is trivial. For the other containment, let f be a V_n-homogeneous polynomial that satisfies $mf \in \langle p^+_{n,i} \rangle$ for some monomial m. We want $f \in \langle p^+_{n,k} : k \in [n] \rangle$. By Lemma 7.20, $mf = tp^+_{n,i}$ for some term t. Since mf has the same number of terms as f and also the same number of terms as $tp^+_{n,i}$, this number must be $2^{c_{n-1}}$. By Remark 7.14, the only V_n-homogeneous polynomials with $2^{c_{n-1}}$ terms are monomial multiples of the $p^+_{n,k}$ for $k \in [n]$. □

Proposition 7.22. $(J_n + \langle p_n \rangle) : \left(\prod_{ij} x_{ij}\right)^\infty = J_n + \langle p^+_{n,j}, j \in [n] \rangle$.

Proof. Throughout we work in the quotient ring $S := \mathbb{K}[\mathbb{N}^{c_{n+1}}]/J_n$ and want to show

$$\langle p_n \rangle : \left(\prod_{ij} x_{ij} \right)^{\infty} = \langle p_{n,j}^+, j \in [n] \rangle.$$

Lemma 7.18 gives the inclusion $\supseteq$, since it shows that, modulo J_n, a monomial multiple of $p_{n,i}^+$ is equal to either p_n or $x_{in}p_n$ and thus lies in $\langle p_n \rangle$. For the other containment let

$$f \in \langle p_n \rangle : \left(\prod_{ij} x_{ij} \right)^{\infty},$$

that is $mf \in \langle p_n \rangle$ for some monomial m in S. This implies $mf = gp_n$ for some polynomial $g \in S$. By Lemma 7.18, $x_{in}mf = g' p_{n,i}^+$ for some $g' \in S$. So, $x_{in}mf \in \langle p_{n,i}^+ \rangle$ and thus $f \in \langle p_{n,i}^+ \rangle : x_{in}m$. Lemma 7.21 shows that $f \in \langle p_{n,k} : k \in [n] \rangle$. □

Having identified the minimal saturated fibers, the longest polynomials, and computed the saturation with respect to the variables x_{ij}, we are now ready to prove Theorem 7.11.

Proof of Theorem 7.11. After a potential renumbering, assume ϕ_1 is the identity. It suffices to prove the theorem for the omission of the Veronese ideal $i = 1$ from the intersection. The remaining cases follow by application of ϕ_l to the ambient ring.

Consider the extensions $J_n\mathbb{K}[\mathbb{Z}^{c_{n+1}}]$ and $I_n\mathbb{K}[\mathbb{Z}^{c_{n+1}}]$ to the Laurent polynomial ring. By the general Theorem 7.23

$$\bigcap_{i \neq 1} \phi_i(I_n\mathbb{K}[\mathbb{Z}^{c_{n+1}}]) = J_n\mathbb{K}[\mathbb{Z}^{c_{n+1}}] + \langle p_n \rangle.$$

Pulling back to the polynomial ring, we have

$$\bigcap_{i \neq 1} \phi_i(I_n) = (J_n + \langle p_n \rangle) : (\prod_{x_{ij}} x_{ij})^{\infty}.$$

Contingent on Theorem 7.23, the result now follows from Proposition 7.22. □

We have reduced the proof of Theorem 7.11 to a general result on intersection in the Laurent ring. It is a variation of [ES96, Theorem 2.1]. According to [ES96, Section 2], any binomial ideal in the Laurent polynomial ring $\mathbb{K}[\mathbb{Z}]$ is defined by its lattice $L \subset \mathbb{Z}^n$ of exponents and a partial character $\rho : L \to \mathbb{K}^*$. Such an ideal is denoted $I(\rho)$ where the lattice L is part of the definition of ρ.

Theorem 7.23. *Let* $\mathbb{K}$ *be a field such that* $\operatorname{char}(\mathbb{K})$ *is either zero or does not divide the order of the torsion part of* $\mathbb{Z}^n/L$ *and* $I(\rho) \subset \mathbb{K}[\mathbb{Z}^n]$ *be binomial. Let*

$I(\rho) = I(\rho_1') \cap \ldots \cap I(\rho_k')$ be a primary decomposition of $I(\rho)$ over the algebraic closure $\overline{\mathbb{K}}$ of $\mathbb{K}$. Omitting one component $I(\rho_{i^*}')$ yields

$$\bigcap_{i \neq i^*} I(\rho_i') = I(\rho) + \rho_{i^*}'(p_L)$$

where p_L is the generating function of a fundamental parallelepiped of the lattice L.

Proof. A linear change of coordinates in $\mathbb{Z}^n$ corresponds to a multiplicative change of coordinates in $\mathbb{K}[\mathbb{Z}^n]$. Since the inclusion of $L \subset \mathbb{Z}^n$ can be diagonalized using the Smith normal form, one can reduce to the case that $I(\rho)$ is generated by binomials $x_i^{q_i} - 1$. This case follows by multiplication of the results in the univariate case. The univariate case, in turn, is up to scaling given by the polynomials $(x^n - 1)/(x - 1)$. □

The assumption on $\operatorname{char}(\mathbb{K})$ in Theorem 7.23 can be relaxed at the cost of a case distinction similar to that in [ES96, Theorem 2.1].

The explicit form of p_L depends on a choice of lattice basis. Because the notions lattice basis ideal and lattice ideal are not the same in the polynomial ring (they are in the Laurent polynomial ring), one needs to pull back using colon ideals to get a result in the polynomial ring. Even if in the Laurent ring the subintersection in Theorem 7.23 is principal modulo $I(\rho)$, it need not be principal in the polynomial ring (as visible in Theorem 7.11). It would be very nice to find more effective methods for binomial subintersections in the polynomial ring, but at the moment the following remark is all we have.

Remark 7.24. Under the field assumptions in Theorem 7.23, let $I \subset \mathbb{K}[\mathbb{N}^n]$ be a lattice ideal in a polynomial ring with indeterminates $x_1, \ldots, x_n$. There exists a partial character $\rho : L \to \mathbb{K}^*$ such that $I = I(\rho) \cap \mathbb{K}[\mathbb{N}^n]$. The intersection of all but one minimal primary component of I is

$$(I(\rho) + \rho(p_L)) \cap \mathbb{K}[\mathbb{N}^n] = (I + \rho(p)m) : (\prod_{i=1}^{n} x_i)^\infty.$$

where p_L is the generating function of a fundamental parallelepiped of L, and m is any monomial such that $\rho(p_L)m \in \mathbb{K}[\mathbb{N}^n]$.

7.3 Extensions

The broadest possible generalization of the results in Section 7.2 may start from an arbitrary toric ideal $I \subset \mathbb{K}[\mathbb{N}^n]$, corresponding to a grading matrix $V \in \mathbb{N}^{d \times n}$, and a subideal $J \subset I$, for example a lattice basis ideal. One can then ask when the quotient $\mathbb{K}[\mathbb{N}^n]/J$ is toral in the grading V. The techniques in Section 7.2

depend heavily on this property and the very controllable stabilization of the Hilbert function. One can get the feeling that this happens if $J \subset I$ is a lattice ideal for some lattice that is of finite index in the saturated lattice $\ker_{\mathbb{Z}}(V)$. However, such a J cannot always be found: by a result of Cattani, Curran, and Dickenstein, there exist toric ideals that do not contain a complete intersection of the same dimension [CCD07].

A more direct generalization of the results of Section 7.2 was suggested to us by Aldo Conca. The d-th Veronese grading $V_{d,n}$ has as its columns all vectors of length n and weight d. The corresponding toric ideal is the d-th Veronese ideal $I_{d,n} \subset S = \mathbb{K}[\mathbb{N}^N]$ and it contains a natural complete intersection $J_{d,n}$ defined as follows. The set of columns of $V_{d,n}$ includes the multiples of the unit vectors $D := \{de_i, i = 1, \ldots, n\}$. For any column $v \notin D$, let $f_v = x_v^d - \prod_i x_{de_i}^{v_i}$. Then $J = \langle f_v : v \notin D \rangle \subset I_{d,n}$ is a complete intersection with $\operatorname{codim}(J_{d,n}) = \operatorname{codim}(I_{d,n})$. It is natural to conjecture that a statement similar to Proposition 7.5 is true. In this case, however, the group L/L' (cf. Proposition 7.4) has higher torsion. This implies that the binomial primary decomposition of J exists only if $\mathbb{K}$ has corresponding roots of unity. By results of Goto and Watanabe [GW78, Chapter 3] on the canonical module (cf. [BH98, Exercise 3.6.21]) the ring S/I is Gorenstein if and only if $d|n$, so that $J : I$ is equal to $J + (p)$ for some polynomial p exactly in this situation.

In Section 7.2, the notation can be kept in check because there is a nice representation of monomials as upper triangular matrices (Proposition 7.5, Lemma 7.13, etc.). To manage the generalization, it will be an important task to find a similarly nice representation. It is entirely possible that something akin to the string notation of [Stu96, Section 14] does the job. Additionally, experimentation with `Macaulay2` will be hard. For example, for $d = 3, n = 4$, the group L/L' from Proposition 7.4 is isomorphic to $(\mathbb{Z}/3\mathbb{Z})^{13}$ which means that a prime decomposition of $J_{3,4}$ has 1 594 323 components. Computing subintersections of it is out of reach. It may be possible to compute a colon ideal like $(J_{3,4} : I_{3,4})$ directly, but off-the-shelf methods failed for us.

Chapter 8

Discussion

Computer vision is a rich field with many challenges for computer algebra and algebraic geometry. In this thesis we prosperously employed many algebraic methods to different problems of multiview geometry.

The multiview variety V_A is a principle object in multiview geometry. Here we look into two natural generalizations of it, the rigid multiview variety $V(J_A)$ and unlabeled multiview variety $\mathrm{Sym}_m(V_A)$. Both varieties appear naturally in the context of computer vision. The behavior of both varieties has been analyzed successfully. For the rigid multiview variety $V(J_A)$ a set-theoretical description is given in Theorem 3.7. Moreover, a triangulation algorithm (Algorithm 2) in the unlabeled case is constructed. Nonetheless many things remain to be solved. For $V(J_A)$ and $\mathrm{Sym}_m(V_A)$ we still do not have any methods to construct their prime ideals. Similarly to the multiview variety V_A in both cases it seems the prime ideal can be obtained from the information that the bilinear and trilinear relations possess. Table 3.2 and Proposition 5.22 point towards this conjecture. Additionally the ambiguities of the unlabeled triangulation are not completely characterized.

Anyhow one can use both varieties to formulate optimization problems that project onto these, minimizing the distance with respect to some chosen norm. This becomes relevant when real world data is used and the image points are given noisy. one could then use triangulation to reconstruct the original world points, if a point on the varieties is found by solving such an optimization problem. In Chapter 4 we aimed to use mathematical optimization software to project onto the rigid multiview variety. However the rigid multiview variety seems to be too complicated for a straightforward approach. We would be highly interested in formulating an optimization problem based on the $V(J_A)$. To obtain globally optimal solutions this would probably require to exploit the structure of the equations of $V(J_A)$.

For the varieties V_A, $V(J_A)$, $\mathrm{Sym}_m(V_A)$ and triangulation algorithms the cameras are assumed to be known. However in many real world use cases this information is not available. Thus often the fundamental matrix needs to be computed from the image data at first. The main algorithm used in that pursuit is the 8-point algorithm. Chapter 6 describes an ill-conditioned set of this algorithm. A promising direction for future work could aim to completely classify the world point configurations that defeat the 8-point algorithm. If the world points are in convex position, we can construct a polytope with the points of that configuration as vertices. One could at first ask if those configurations that defeat the 8-point algorithm can be classified by the combinatorial type of the polytope associated to them. A combinatorial cube is a convex 3-polytope with eight vertices. In total there are 257 different combinatorial types of 3-polytopes with eight vertices. Enumerating one representative per equivalence class and checking if it defeats the 8-point algorithm could be an initial step in that direction.

We completely characterize the colon ideal $\phi(J_n) : I_n$ with Theorem 7.11, for some partial character on the Veronese lattice L_n. Theorem 7.24 states that this description could be an artifact of a more general fact. We conjecture that Theorem 7.24 holds true not just for Laurent binomial ideals but even for binomial ideals I_+ in the polynomial ring. This is based on substantial computational evidence. Even though most of the statements developed from the special case of $\phi(J_n) : I_n$ hold true in general, the whole proof seems to be deeply involved.

Bibliography

[AAT12] Chris Aholt, Sameer Agarwal, and Rekha Thomas. “A qcqp approach to triangulation”. In: *Computer Vision–ECCV 2012*. Springer Berlin Heidelberg, 2012, pp. 654–667.

[AO14] Chris Aholt and Luke Oeding. “The ideal of the trifocal variety”. In: *Mathematics of Computation* 83.289 (2014), pp. 2553–2574.

[ApS15] MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28)*. 2015. URL: http://docs.mosek.com/7.1/toolbox/index.html.

[AST13] Chris Aholt, Bernd Sturmfels, and Rekha Thomas. “A Hilbert scheme in computer vision”. In: *Canad. J. Math. Vol.* 65 (5) (2013), 961–988.

[Ber63] Robert Berger. “Differentialmoduln eindimensionaler lokaler Ringe”. In: *Mathematische Zeitschrift* 81.4 (1963), pp. 326–354. ISSN: 1432-1823. DOI: 10.1007/BF01111579. URL: http://dx.doi.org/10.1007/BF01111579.

[BH98] Winfried Bruns and Jürgen Herzog. *Cohen–Macaulay rings*. Cambridge University Press, 1998.

[Bob+08] Alexander I. Bobenko et al. *Discrete differential geometry*. Springer, 2008.

[Bro17] Kevin Brown. *Skew octagon solving the Thomson problem of eight points*. http://mathpages.com/home/kmath005/elec/elec8.htm. 2017.

[Bug+15] Florian Bugarin et al. “Rank-constrained fundamental matrix estimation by polynomial global optimization versus the eight-point algorithm”. In: *Journal of Mathematical Imaging and Vision* 53.1 (2015), pp. 42–60.

[BV88] Winfried Bruns and Udo Vetter. *Determinantal Rings*. Vol. 1327. LNM. Springer Verlag, Berlin, 1988.

[CCD07] Eduardo Cattani, Raymond Curran, and Alicia Dickenstein. “Complete intersections in toric ideals”. In: *Proceedings of the American Mathematical Society* 135.2 (2007), pp. 329–335.

[Con94a] Aldo Conca. "Divisor class group and canonical class of determinantal rings defined by ideals of minors of a symmetric matrix". In: *Archiv der Mathematik* 63.3 (1994), pp. 216–224.

[Con94b] Aldo Conca. "Gröbner bases of ideals of minors of a symmetric matrix". In: *Journal of Algebra* 166.2 (1994), pp. 406–421.

[Dec+16] Wolfram Decker et al. SINGULAR *4-1-0 — A computer algebra system for polynomial computations*. http://www.singular.uni-kl.de. 2016.

[DES98] Persi Diaconis, David Eisenbud, and Bernd Sturmfels. "Lattice walks and primary decomposition". In: *Mathematical Essays in Honor of Gian-Carlo Rota*. Ed. by B. Sagan and R. Stanley. Vol. 161. Progress in Mathematics. Birkhauser, Boston, 1998, pp. 173–193.

[DIW16] Mareike Dressler, Sadik Iliman, and Timo de Wolff. "An Approach to Constrained Polynomial Optimization via Nonnegative Circuit Polynomials and Geometric Programming". In: *preprint arXiv: 1602.06180* (2016).

[DMM10] Alicia Dickenstein, Laura Matusevich, and Ezra Miller. "Combinatorics of Binomial Primary Decomposition". In: *Mathematische Zeitschrift* 264.4 (2010), pp. 745–763. DOI: 10.1007/s00209-009-0487-x.

[ES96] David Eisenbud and Bernd Sturmfels. "Binomial Ideals". In: *Duke Mathematical Journal* 84.1 (1996), pp. 1–45.

[Gam+16] Gerald Gamrath et al. *The SCIP Optimization Suite 3.2*. eng. Tech. rep. 15-60. Takustr.7, 14195 Berlin: ZIB, 2016.

[GJ00] E. Gawrilow and M. Joswig. "`polymake`: a framework for analyzing convex polytopes". In: *Polytopes—combinatorics and computation*. Vol. 29. Basel: Birkhäuser, 2000, pp. 43–47.

[Gor07] Elisa Gorla. "The G-biliaison class of symmetric determinantal schemes". In: *Journal of Algebra* 310.2 (2007), pp. 880–902.

[Gor10] Elisa Gorla. "Symmetric ladders and G-biliaison". In: *Liaison, Schottky Problem and Invariant Theory*. Springer, 2010, pp. 49–62.

[Got77] Shiro Goto. "The divisor class group of a certain Krull domain". In: *J. Math. Kyoto Univ.* 17 (1977), pp. 47–50.

[Got79] Shiro Goto. "On the Gorensteinness of determinantal loci". In: *Journal of Mathematics of Kyoto University* 19.2 (1979), pp. 371–374.

[Grö65] Wolfgang Gröbner. "Über Veronesesche Varietäten und deren Projektionen". In: *Archiv der Mathematik* 16.1 (1965), pp. 257–264.

[GS] Daniel R. Grayson and Michael E. Stillman. *Macaulay2, a software system for research in algebraic geometry.* Available at http://www.math.uiuc.edu/Macaulay2/.

[GW78] Shiro Goto and Keiichi Watanabe. "On graded rings I". In: *Journal of the Mathematical Society of Japan* 30.2 (1978), pp. 179–213.

[HK07] Richard Hartley and Fredrik Kahl. "Critical configurations for projective reconstruction from multiple views". In: *International Journal of Computer Vision* 71.1 (2007), pp. 5–47.

[HLL09] Didier Henrion, Jean-Bernard Lasserre, and Johan Löfberg. "GloptiPoly 3: moments, optimization and semidefinite programming". In: *Optimization Methods & Software* 24.4-5 (2009), pp. 761–779.

[Hr97] Anders Heyden and Karl Åström. "Algebraic properties of multilinear constraints". In: *Mathematical Methods in the Applied Scinces* 20.13 (1997), pp. 1135–1162.

[HS97] Richard Hartley and Peter Sturm. "Triangulation". In: *Computer vision and image understanding* 68.2 (1997), pp. 146–157.

[HZ03] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2003. ISBN: 9780521540513.

[Jen] Anders N. Jensen. *Gfan, a software system for Gröbner fans and tropical varieties.* Available at http://home.imf.au.dk/jensen/software/gfan/gfan.html.

[JKK16] Anders Jensen, Thomas Kahle, and Lukas Katthän. "Finding binomials in polynomial ideals". In: *preprint, arXiv:1607:02135* (2016).

[Jos+16] Michael Joswig et al. "Rigid multiview varieties". In: *International Journal of Algebra and Computation* 26.04 (2016), pp. 775–788.

[JT13] Michael Joswig and Thorsten Theobald. *Polyhedral and algebraic methods in computational geometry.* Springer Science & Business Media, 2013.

[KH05] Fredrik Kahl and Didier Henrion. "Globally optimal estimates for geometric reconstruction problems". In: *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on.* Vol. 2. IEEE. 2005, pp. 978–985.

[KM14] Thomas Kahle and Ezra Miller. "Decompositions of commutative monoid congruences and binomial ideals". In: *Algebra & Number Theory* 8.6 (2014), pp. 1297–1364.

[Kol88] János Kollár. "Sharp effective nullstellensatz". In: *Journal of the American Mathematical Society* 1.4 (1988), pp. 963–975.

[Kra41] Josef Krames. “Zur Ermittlung eines Objektes aus zwei Perspektiven.(Ein Beitrag zur Theorie der “gefährlichen Örter”.)” In: *Monatshefte für Mathematik und Physik* 49.1 (1941), pp. 327–354.

[Kut74] Ronald E Kutz. “Cohen–Macaulay rings and ideal theory in rings of invariants of algebraic groups”. In: *Transactions of the American Mathematical Society* 194 (1974), pp. 115–129.

[KW16] Thomas Kahle and André Wagner. “Veronesean almost binomial almost complete intersections”. In: *arXiv preprint arXiv:1608.03499* (2016).

[Lan] J.M. Landsberg. *Tensors: Geometry and Applications*. Graduate studies in mathematics. American Mathematical Soc. URL: https://books.google.de/books?id=JTjv3DTvxZIC.

[Li17] Binglin Li. “Images of rational maps of projective spaces”. In: *International Mathematics Research Notices* (2017), rnx003.

[Lof05] Johan Lofberg. “YALMIP: A toolbox for modeling and optimization in MATLAB”. In: *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*. IEEE. 2005, pp. 284–289.

[MAT16] MATLAB. *version R2016a*. Natick, Massachusetts: The MathWorks Inc., 2016.

[May12] Stephen Maybank. *Theory of reconstruction from image motion*. Vol. 28. Springer Science & Business Media, 2012.

[MM82] Ernst W. Mayr and Albert A. Meyer. “The Complexity of the Word Problems for Commutative Semigroups and Polynomial Ideals”. In: *Advances in Mathematics* 46.3 (1982), pp. 305–329.

[MS05] Ezra Miller and Bernd Sturmfels. *Combinatorial Commutative Algebra*. Vol. 227. GTM. Springer, Berlin, 2005.

[Phi98] Johan Philip. “Critical point configurations of the 5-, 6-, 7-, and 8-point algorithms for relative orientation”. In: *Technical report TRITA-MAT-MA* (1998).

[RG11] Jürgen Richter-Gebert. *Perspectives on projective geometry: A guided tour through real and complex geometry*. Springer Science & Business Media, 2011.

[Stu96] Bernd Sturmfels. *Gröbner Bases and Convex Polytopes*. Vol. 8. University Lecture Series. American Mathematical Society, Providence, RI, 1996.

[Stu99] Jos F. Sturm. “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones”. In: *Optimization methods and software* 11.1-4 (1999), pp. 625–653.

[TY26] Herbert Turnbull and Alfred Young. “The linear invariants of ten quaternary quadrics”. In: *Trans. Camb. Phil. Soc* 23 (1926), pp. 265–301.

[Wag17a] André Wagner. “Algebraic Relations and Triangulation of Unlabeled Image Points”. In: *arXiv preprint arXiv:1707.08722* (2017).

[Wag17b] André Wagner. “Pictures of Combinatorial Cubes”. In: *arXiv preprint arXiv:1707.06563* (2017).

[Whi88] Neil White. “Implementation of the straightening algorithm of classical invariant theory”. In: *Proceedings, Invariant Theory and Tableaux, IMA* (1988).