Hey! You got your algebraic geometry in my computer vision!

Jessie Loucks-Tavitas

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

Doctor of Philosophy

University of Washington

2024

Reading Committee:

Max Lieblich, Chair

Rekha Thomas

Jarod Alper

Program Authorized to Offer Degree:

Mathematics

University of Washington

Abstract

Hey! You got your algebraic geometry in my computer vision!

Jessie Loucks-Tavitas

Chair of the Supervisory Committee:

Professor Max Lieblich

Mathematics

We consider problems in the intersection of algebraic geometry and computer vision. In particular, we study algebraic varieties associated with the camera resectioning problem. We characterize these resectioning varieties' multigraded vanishing ideals using Gröbner basis techniques. As an application, we derive and re-interpret celebrated results in geometric computer vision related to camera-point duality. We also clarify some relationships between the classical problems of optimal resectioning and triangulation, state a conjectural formula for the Euclidean distance degree of the resectioning variety, and discuss how this conjecture relates to the recently-resolved multiview conjecture. Additionally, we work toward a compactification of the moduli of camera pairs that allows for a continuous geometric interpolation between concentric and non-concentric pairs. Furthermore, we give a new construction of the space of essential matrices from first principles. This construction enables us to re-prove the fundamental results of Demazure and to re-prove the recent description of the essential variety due to Kileel–Fløystad–Ottaviani. We also describe a new five-point solver for generic data based upon generic Gröbner bases for symmetric bilinear forms. Finally, we perform some intersection-theoretic calculations to provide basic geometric invariants of the classical multiview variety as well as the universal imaging variety.

TABLE OF CONTENTS

Page

List of Figures iii

Dedication v

The Menu vi

Chapter 1: Introduction 1
- 1.1 History and motivation 1
- 1.2 Basics of algebraic vision 3
- 1.3 Organization 5

Chapter 2: Camera resectioning and Carlsson-Weinshall duality 7
- 2.1 Introduction 7
- 2.2 Resectioning vs triangulation 11
- 2.3 Carlsson-Weinshall duality revisited 17
- 2.4 Proof of Theorem 2.2.6 28
- 2.5 Optimal single-camera resectioning 35
- 2.6 Conclusion 37

Chapter 3: Towards a compactification of the moduli of camera pairs 39
- 3.1 The fundamental variety 39
- 3.2 Attempts at compactification 41

Chapter 4: The essential variety revisited 46
- 4.1 Camera calibration and the essential variety 46
- 4.2 A modular compactification of the classical double cover 49
- 4.3 Where do we go from here? 59

Chapter 5: Intersection theory in algebraic vision 60
5.1 The story of this project 60
5.2 Towards the total Chern class of the classical multiview variety 60
5.3 Some questions regarding the univariate multiview variety 65
5.4 The Picard group of the universal imaging variety 67

Bibliography 70

LIST OF FIGURES

1.1 The "Cellini bell" (left) and a 3-D reconstruction (right). Images © Factum Foundation. 1

1.2 A pinhole camera. 2

1.3 A point correspondence, $(A_1(q), A_2(q)) \in \mathbf{P}^2 \times \mathbf{P}^2$ between two pinhole cameras, A_1 and A_2. 2

1.4 A triangulation problem (# cameras = 2, # world points = 1). 4

1.5 A resectioning problem (# cameras = 1, # world points = 2). 4

2.1 Two reduced cameras viewing three 3D points (left) are Carlsson-Weinshall dual to three reduced cameras viewing two 3D points (right). See Section 2.3.1 for details. 10

2.2 Four fixed points $E_1, \ldots, E_4 \in \mathbf{P}^3$ determine a reference tetrahedron. They project through the camera center $c \in \mathbf{P}^3$ to the four points $e_1, \ldots e_4 \in \mathbf{P}^2$. The Cremona transformation $\mathcal{C}$ maps the line through c and q to the unique twisted cubic passing through $\mathcal{C}(q), \mathcal{C}(c), E_1, \ldots, E_4$. 20

2.3 An atlas for the reduced pinhole camera, cf. [Aga+23, Figure 1]. 25

2.4 Schematic outline of the proof of Theorem 2.2.6. 29

3.1 A panorama of the interior of St. Peter's Basilica, made in Google Photos from several consecutive photographs. 42

4.1 Projective ambiguity of 3D reconstructions, c/o [HZ03, Fig. 1.4] 46

4.2 The two directions d_1 and d_2 are conjugate with respect to C. 48

4.3 Two cones intersecting in a smooth quartic, two smooth conics, and a smooth conic and doubled line, respectively. 51

ACKNOWLEDGMENTS

Firstly, I would like to express my deepest gratitude to my advisor Max Lieblich. His advice, kindness, empathy, and general presence during my time here have been irreplaceable, and I could not have asked for a better guide through the proverbial forest. This dissertation would not exist without the guidance and collaboration of Rekha Thomas and Sameer Agarwal, and even more so the contributions of my coauthors Timothy Duff and Erin Connelly. I am very thankful for their time and input. I am also grateful for the guidance I have received from my professors and mentors from graduate school and other programs such as EDGE and the SDSU REU. I am indebted to my undergraduate professors (and future colleagues!)—particularly Corey Shanbrom, Matthew Krauel, and Jay Cummings—for telling me to go to grad school, and for their mentorship even beyond my time at Sacramento State. My fellow graduate students have been a continual source of camaraderie and joy while we wander through our respective forests. I am especially thankful for the friendship of Martin Bishop, Paige Helms, and Alex Wang, as well as the Cool Book Club, AG Lunch (whose budget partially supported this dissertation), and the STEMinists. I extend a special thanks to Alex [redacted] and Kevin [redacted]. I thank my cats, Timothie and Ottoman, for helpfully ensuring that I left the house at least twice a week in search of a more peaceful work environment. Without their interventions I would doubtless have become a recluse while writing this dissertation. I also thank Hope Midgett, Jaimie Cooper, the Elessia group chat, my Tuesday morning discipleship group, the Central District community group, and the people of Downtown Cornerstone Church for being my spiritual family and support system. I am immensely thankful to my parents, Stuart and Jennifer Loucks, and my siblings and their families, Jason and Alexis, Jared and London, Joanna, Jamie, and Jordan Loucks, for their unconditional love from afar. I thank my incredible husband Ryan Tavitas, the best man I know, for his unmatched love, patience, selflessness, faithfulness, and partnership. These characteristics define not only his support for my career but also every other part of our life together. Finally, I owe not only the entirety of this dissertation, but my whole life and being to the Creator of my soul (and mathematics).

DEDICATION

to my grandparents:
KC and Sally Loo,
Jim and Jean Loucks, and
Eb Wilkinson and Sue Reynolds.

The Menu

Soup

(chapter 1)

computer vision consommé

refined with a thread of algebraic geometry

Salad

(chapters 3 to 5)

chopped heirloom moduli spaces

with a rank-deficient matrix emulsion

Entrée

(chapter 2)

algebraic variety served two ways

crowned with a Gröbner basis foam

Dessert

(section 2.3)

deconstructed Carlsson-Weinshall duality

with a reduced birational geometry glaze

Chapter 1

INTRODUCTION

1.1 History and motivation

Computer vision finds its origin in *photogrammetry*, dating back to the late 1800s. Photogrammetry's original goal was to recover data from photographs in order to model coastlines and other geographic areas that were otherwise inaccessible by standard cartographic methods. To this day, computer vision algorithms use techniques from photogrammetry to model three-dimensional objects such as coastlines, architecture, and historical articafts (see Figure 1.1).

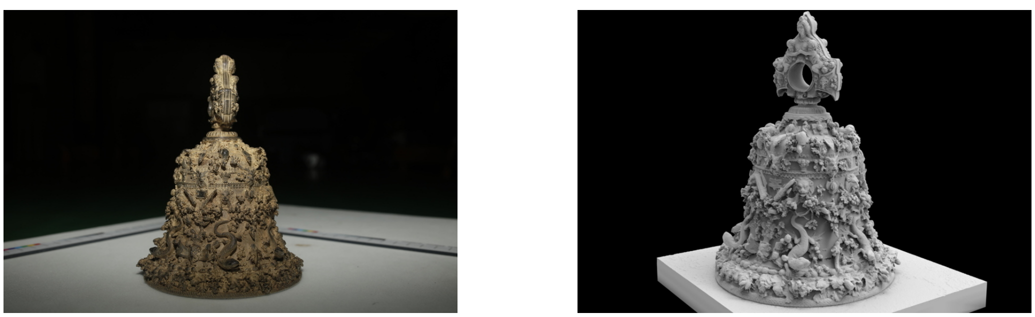


Figure 1.1: The "Cellini bell" (left) and a 3-D reconstruction (right).
Images © Factum Foundation.

Two classical questions arising in multi-view geometry are those of *triangulation* and *resectioning*.

Question 1.1.1 (Triangulation)**.** Given a three-dimensional object or scene and multiple images of it, can we determine the (relative) positions of the cameras in the world?

Question 1.1.2 (Resectioning)**.** Given multiple images as well as (relative) camera locations, can we reconstruct the scene or object being photographed?

Questions in computer vision have long-since been modeled by projective geometry, which led to the development of *multi-view geometry*. One can even argue that projective geometry was invented with the purpose of solving problems in computer vision, since projective spaces were created to model linear perspective.

A classical projective pinhole camera is modeled by choosing a *camera center* $c \in \mathbf{P}^3$ and an *image plane* $P \subset \mathbf{P}^3$. The image of another point in the world, say $q \in \mathbf{R}^3$, is the point of intersection, $\overline{qc} \cap P$ (see Figure 1.2). As we will see in section 1.2, this model can be mathematically formalized as a linear projection $\mathbf{A} : \mathbf{P}^3 \dashrightarrow \mathbf{P}^2$. Every A is undefined at exactly one point which we call the *camera center*. We call points in $\mathbf{P}^3$ *world points* and points in $\text{im}(A)$ *image points*.

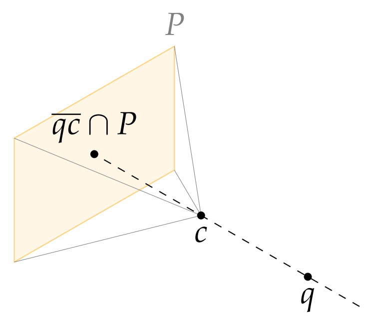


Figure 1.2: A pinhole camera.

In this framework, answering the questions of triangulation and resectioning boils down to characterizing the *point correspondences* of a collection of cameras or world points. For example, locating a world point q given two cameras A_1, A_2 is equivalent to locating its image points $A_1(q), A_2(q)$ and intersecting the lines $\overline{A_1(q)c}$ and $\overline{A_2(q)c}$ (see Figure 1.3).

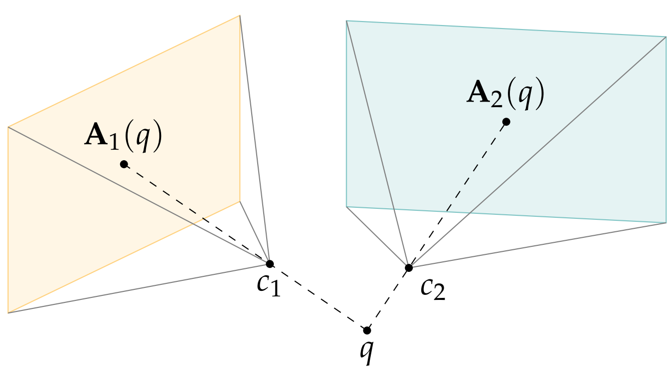


Figure 1.3: A point correspondence, $(A_1(q), A_2(q)) \in \mathbf{P}^2 \times \mathbf{P}^2$ between two pinhole cameras, A_1 and A_2.

1.2 Basics of algebraic vision

1.2.1 Definitions

We now introduce the basic definitions used to model multi-view geometry questions algebro-geometrically.

Note 1.2.1. One can define pinhole cameras and multiview configurations relatively over any given scheme S, which is necessary for defining the functor $\mathrm{Cam}_n^\circ : \mathbf{Sch}/S \to \mathbf{Set}$ in [LV20]. We omit the details of relativization here, and, for intuition, we work over $\mathbf{R}$ or $\mathbf{C}$ throughout unless otherwise noted.

Definition 1.2.2 (Projective pinhole cameras)**.** A *(projective) pinhole camera* is a linear projection

$$A : \mathbf{P}^3 \dashrightarrow \mathbf{P}^2$$

given by a full-rank 3×4 matrix in $\mathbf{P}(M_{3\times 4})$ (or, equivalently, three linearly independent sections of $\mathcal{O}_{\mathbf{P}^3}(1)$). The *center* of A is the unique point $c \in \mathbf{P}^3$ at which A is undefined.

We say a camera is *finite* if c does not lie on the plane at infinity $x_3 = 0$.

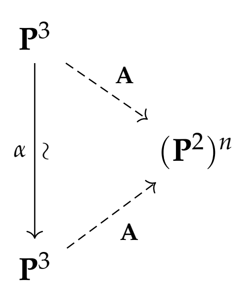
Definition 1.2.3 (Multiview configurations)**.** A *multiview configuration* (or *camera configuration*) *of length n* is a tuple

$$\mathbf{A} = (A_1, \dots, A_n) : \mathbf{P}^3 \dashrightarrow (\mathbf{P}^2)^n.$$

We often denote the center of A_i as c_i for each i.

Note 1.2.4. In two views, we call the line $\ell = \overline{c_1 c_2}$ the *epipolar line*. The *epipoles* e_1, e_2 are the images of each center in the other camera's image plane; i.e. the intersections of ℓ with the image plane.

Definition 1.2.5 (Pinhole camera isomorphisms)**.** An *isomorphism* between multiview configurations $\mathbf{A}$ and $\mathbf{A}'$ of length n is an automorphism $\alpha \in \mathrm{PGL}_4$ such that $A_i = A'_i \circ \alpha$ for all $1 \leq i \leq n$. That is, α fits into a commutative diagram



Definition 1.2.6 (Multiview varieties)**.** The *multiview variety*, or *joint image*, of a camera configuration of length n is $MV_n(\mathbf{A}) := \overline{\mathrm{im}(\mathbf{A})} \subset (\mathbf{P}^2)^n$ (where $\overline{X}$ denotes the Zariski closure of X in its ambient space). A point in $MV_n(\mathbf{A})$ is called a *point correspondence*.

Notation 1.2.7. For consistency with [Aga+23], we often denote $MV_n(\mathbf{A})$ as $\Gamma^{m,1}_{\bar{\mathbf{A}},\mathbf{p}}$.

To conclude this section, we rephrase Questions 1.1.1 and 1.1.1 in this new language.

Question 1.2.8 (Triangulation rephrased)**.** Given a (known) multiview configuration $\bar{\mathbf{A}} = (A_1, \ldots, A_m)$ and their corresponding image points $\mathbf{p} = (p_{11}, \ldots, p_{mn})$ of a(n unknown) tuple of world points $\mathbf{q} = (q_1, \ldots, q_n)$, can we determine the world point positions?

Question 1.2.9 (Resectioning rephrased)**.** Given a (known) tuple of world points $\bar{\mathbf{q}} = (q_1, \ldots, q_n)$ in the world and their corresponding image points $\mathbf{p} = (p_{11}, \ldots, p_{mn})$ under a(n unknown) multiview configuration $\mathbf{A} = (A_1, \ldots, A_m)$, can we determine the relative positions of the cameras?

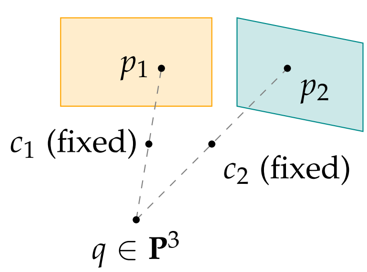


Figure 1.4: A triangulation problem (# cameras = 2, # world points = 1).

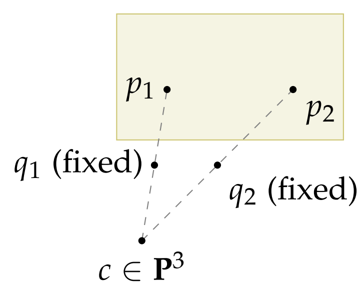


Figure 1.5: A resectioning problem (# cameras = 1, # world points = 2).

1.2.2 Relevant results

In this section we include various results from [LV20], which give one the hope that other results in vision can be proven with similar functorial methods.

Definition 1.2.10 (The moduli space of camera configurations, [LV20])**.** The *functor of length-n camera configurations*, denoted Cam°_n, has value over a scheme S the set of isomorphism classes of general relative multiview configurations of length n.

Remark 1.2.11. See [LV20, Def. 2.25, Def. 2.27] for the technical definition of a *general relative multiview configuration of length n.* For our purposes, the reader can imagine it to be a multiview configuration of length n, as defined above, whose centers are pairwise distinct.

Proposition 1.2.12 ([LV20]). *Two multiview configurations* $\mathbf{A}$ *and* $\mathbf{A}'$ *are isomorphic if and only if their multiview varieties are equal as subschemes of* $(\mathbf{P}^2)^n$.

The following is the main theorem of [LV20], which expands on the main theorem of [AST13].

Theorem 1.2.13 ([LV20]). *There are smooth irreducible varieties* Cam_n° *and* CalCam_n° *parametrizing general n-view camera configurations and n-view calibrated camera configurations, respectively.*

1. *The variety* Cam_n° *has dimension* $11n - 15$*. For all* $n > 1$*, sending a configuration to its joint image defines a locally closed embedding*

$$\mathrm{Cam}_n^\circ \hookrightarrow \mathrm{Hilb}_{(\mathbf{P}^2)^n}.$$

   If $n > 2$ *then this morphsim is an open immersion, so that* Cam_n° *is identified with an open subscheme of the smooth locus of* $\mathrm{Hilb}_{(\mathbf{P}^2)^n}$*.*

2. *The variety* CalCam_n° *has dimension* $6n - 7$*. For all* $n > 1$*, there is a natural locally closed embedding*

$$\mathrm{CalCam}_n^\circ \hookrightarrow \mathrm{Hilb}_{C_1 \times \cdots \times C_n \subset (\mathbf{P}^2)^n}.$$

   (where the latter is a diagram Hilbert scheme). If $n > 2$ *then this morphism is an open immersion.*

3. *The natural decalibration morphism* $\nu_n : \mathrm{CalCam}_n^\circ \to \mathrm{Cam}_n$ *is finite, proper, and unramified. The morphism* ν_2 *is an étale cover of its image with general fiber of order* 2*. For* $n > 2$ *the morphism* ν_n *is generally injective but not injective.*

1.3 Organization

I have found my time studying algebraic vision to be hugely edifying to my mathematical maturity and abilities. In particular, the framework of studying problems specific to computer vision gave me the confidence to learn more about topics in algebraic geometry that were previously intimidating. For instance, the mathematics discussed in this thesis ranges from standard Gröbner basis methods and ideal-theoretic results to discussion of moduli spaces and quotient stacks. One way to view this document is as a tasting platter of the algebraic geometry one can employ to explore, conceptualize, and prove theorems related to problems in computer vision.

The main chapters are organized by topic, as follows: Chapter 2 is a (very slightly edited) version of recent joint work with Erin Connelly and Timothy Duff in [CDL23]. Chapter 3 details the insights I have gained in a long-term effort to provide a geometrically intuitive compactification of the moduli space Cam_2°. In Chapter 4, I review and restructure the results and techniques regarding the essential variety. This chapter is based in large part on work of Benjamin Antieau, Max Lieblich, Bianca Viray, and Lucas Van Meter, which can be found in the thesis of Van Meter [Van19]. Chapter 5 provides some results regarding the intersection theory of various mathematical objects arising in classical algebraic vision questions.

Chapter 2

CAMERA RESECTIONING AND CARLSSON-WEINSHALL DUALITY

The work in this chapter is joint with Timothy Duff and Erin Connelly, and is almost entirely identical to the contents of [CDL23].

2.1 *Introduction*

The *dramatis personae* of the classical pinhole camera model are a full-rank 3×4 matrix A representing a camera, a 4×1 matrix q representing a world point, and a 3×1 matrix p representing its projection into an image. Image formation may be understood via the projective-linear map

$$\begin{aligned} A : \mathbf{P}^3 &\dashrightarrow \mathbf{P}^2 \\ q &\mapsto Aq, \end{aligned} \tag{2.1}$$

and we write $Aq \sim p$ if these two vectors represent the same point in $\mathbf{P}^2$. The *center* of the camera A is the unique point where the map (2.1) is undefined.

The pinhole camera, despite its simplicity, remains a good model of physical cameras. This explains its importance in modern computer vision applications such as structure-from-motion (SfM) and Simultaneous Localization and Mapping (SLAM). On the other hand, classical problems associated with 3D reconstruction have been studied long before the advent of computers, and the role played by algebraic methods in their solution has long been apparent. For instance, Hesse in 1863 formulated the problem of constructing two homographic configurations of 7 lines in space, each prescribed to pass through a configuration of 7 points in a plane [Hes63]. Hesse's reduction of this problem to computing the roots of a cubic equation may be understood as an early instance of the so-called 7 point algorithm. Similarly, Grunert's 1841 "3D Pothenot problem" [Gru41] is known nowadays as the perspective 3-point (P3P) problem, and his general strategy reducing the problem to a quartic equation remains in use today.

In recent years, the name *algebraic vision* [KK22] has been coined to describe a body of interdisciplinary research in which notions from algebra and vision flow freely. To date, algebraic vision has largely focused on problems which we refer to as the *full reconstruction problem* and *triangulation*.

In the full reconstruction problem, we are given a collection of image points $\tilde{p}_{11}, \dots, \tilde{p}_{mn}$, and our task is to recover a set of cameras $A_1, \dots, A_m$ and world points $q_1, \dots, q_n$ that is consistent with these observations. Hesse's solution treats the “minimal" case $(m, n) = (2, 7)$. Today, there are many works which solve analogous minimal problems which can be used effectively in SfM pipelines (see eg. [Lar+18; Duf+19b; Duf+20; KBP08; LAO17].)

In triangulation, we are given not only image points, but also the cameras that produced them, $\bar{A}_1, \dots, \bar{A}_m$. We need only recover one or more unknown world points. Already for $m = 2$, an exact solution to this problem will typically not exist, due to the fact that the lines in $\mathbf{P}^3$ projecting to generic image points under $\bar{A}_1, \bar{A}_2$ will be skew. Nevertheless, algebraic methods have led to a wealth of knowledge about the triangulation problem. For example, the *multiview ideal* associated to $\bar{A}_1, \dots, \bar{A}_m$ gives rise to a complete set of algebraic constraints on *any* m-tuple of image points they produce. There is a considerable literature related to multiview ideals [APT21; AST13; FM95; HÅ97]. Theorem 2.2.3 collects some important previous results.

Often regarded as being “dual" to triangulation is the problem of camera *resectioning*. Here, we assume n image points are given along with the configuration of world points $\bar{\mathbf{q}} = (\bar{q}_1, \dots, \bar{q}_n) \in (\mathbf{P}^2)^n$ from which they were produced by a single unknown camera A. Grunert's 1841 paper gives a minimal solution for $(m, n) = (1, 3)$ under the assumption that A is *Euclidean*. Without this assumption, A is a general 3×4 matrix, and we need $n \geq 6$.

2.1.1 Results and organization

In this chapter, we aim to bring the general resectioning problem up-to-speed with the latest developments in algebraic vision. In Section 2.2, after recalling some previous results about multiview varieties, we state our first main result, Theorem 2.2.6. This characterizes a complete set of algebraic constraints for the resectioning problem, under the genericity assumption that no four of the given world points are coplanar. These constraints are given by k-linear polynomials for $6 \leq k \leq 12$ which generate the *resectioning ideal* $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$ (Definition 2.2.1). Our work is a natural continuation of recent work by Agarwal et al. [Aga+23], and we resolve three of its open questions. For instance, Theorem 2.2.6 resolves [Aga+23, §8.1, Q4] for generic $\bar{\mathbf{q}}$ by determining a universal Gröbner basis for $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$.

We note that resectioning ideals have several pleasant properties from the point of view of commutative algebra: namely, for generic $\bar{\mathbf{q}} \in (\mathbf{P}^3)^n$,

1. For fixed m and n, resectioning ideals are homogeneous with respect to a natural $\mathbf{Z}^{mn}$-grading, and have the same $\mathbf{Z}^{mn}$-graded Hilbert function as long as no four points are coplanar. Proposition 2.2.7 implies that this Hilbert function may be obtained by specializing a combinatorial formula of Li [Li18, Theorem 1.1], based on the inclusion-exclusion rule. Our ideal-theoretic result also considerably strengthens Li's set-theoretic description, and reduces the degrees of the equations that are needed.

2. The multidegrees of resectioning ideals are always equal to 1. A geometric explanation of this phenomenon follows along the lines explained in [Bre+23, §4]. See also [CCM23, Theorem 4.2] for an explanation using multigraded Rees algebras.

3. For any monomial order $<$, the initial ideal $\mathrm{in}_<(I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}))$ and the multigraded generic initial ideal $\mathrm{gin}_<(I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}))$, although not equal as in the case of multiview ideals [AST13], are both radical. In particular, $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$ belongs to the class of *Cartwright-Sturmfels ideals*, recently surveyed by Conca, De Negri, and Gorla [CDG22].

Our first basic insight is that the projection of a point $q \in \mathbf{P}^3$ under a pinhole camera $A : \mathbf{P}^3 \dashrightarrow \mathbf{P}^2$ may be viewed as the projection of a point $\mathrm{vec}(A) \in \mathbf{P}^{11}$ under what we call a "hypercamera" $Q : \mathbf{P}^{11} \dashrightarrow \mathbf{P}^2$. This is reminiscent, and in fact a generalization, of a well-studied principle in computer vision known as *Carlsson-Weinshall duality* [CW98]. This is the subject of Section 2.3. Our Theorem 2.3.5 develops a reduced analogue of the "atlas" of algebraic varieties proposed in [Aga+23]. This addresses [Aga+23, §8.2, Q2]. The *reduced joint image* and its dual, recently studied by Trager, Ponce, and Hebert, are two members of this atlas. Carlsson-Weinshall duality amounts to a simple linear isomorphism between these two varieties. In Example 2.2.8 and Section 2.3.2, we explain how our perspective unifies previous approaches to resectioning constraints in the computer vision literature [CW98; THP19; Qua95; Sch+00], which can all be obtained from the ideal $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$ by specialization.

Theorem 2.3.6 in Section 2.3.2 shows that reduced resectioning varieties for generic point configurations are scheme-theoretically cut out by bilinear forms. This stands in stark contrast to the high degree polynomials in Theorem 2.2.6, whose proof we complete

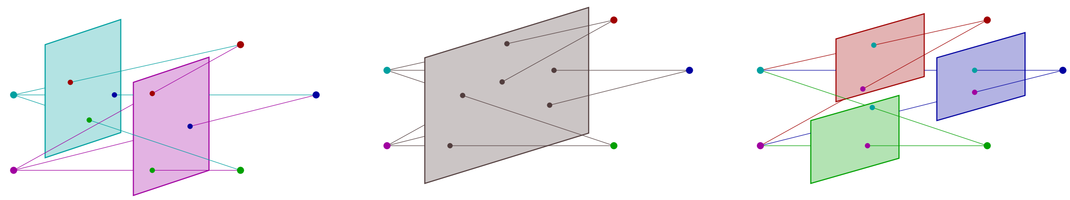


Figure 2.1: Two reduced cameras viewing three 3D points (left) are Carlsson-Weinshall dual to three reduced cameras viewing two 3D points (right). See Section 2.3.1 for details.

in Section 2.4. Finally, in Section 2.5, we address [Aga+23, §8.1, Q6] by investigating the *Euclidean distance degree* of the resectioning variety in affine pixel coordinates. This is a number that quantifies the algebraic complexity of a natural Euclidean distance optimization formulation of the camera resectioning problem. Our main contribution, based on evidence supplied by computational experiments, is Conjecture 2.5.1, giving a formula for this quantity as a cubic polynomial in n. The statement is analogous to, and inspired by, the *multiview conjecture*, recently resolved by Maxim, Rodriguez, and Wang [MRW20]. We conclude with a short discussion in Section 2.6.

2.1.2 Notation and conventions

Our notation largely follows that established in [Aga+23]. Our basic algebro-geometric objects are affine and projective varieties over the field of complex numbers $\mathbf{C}$. The symbol $\mathbf{P}^n$ denotes complex n-dimensional projective space, which we may also identify with the projectivization $\mathbf{P}(V)$ of any $(n+1)$-dimensional complex vector space V. As in the introduction, known quantities will usually be designated with a bar $\bar{\bullet}$. This bar is also used to denote the *Zariski closure* of a set: its usage will be clear from the context. If we wish to emphasize that given quantities in certain scenarios may be "noisy" due to deviations from the pinhole model or erroneous measurements, we instead use $\widetilde{\bullet}$.

2.2 *Resectioning vs triangulation*

Let us recall a "universal" version of the imaging map (2.1). This is the map which sends m cameras $A_1, \ldots, A_m \in \mathbf{P}\left(\mathrm{Hom}_{\mathbf{C}}(\mathbf{C}^4, \mathbf{C}^3)\right) \cong \mathbf{P}^{11}$ and n points $q_1, \ldots, q_n \in \mathbf{P}^3$ to mn points in $\mathbf{P}^2$. The graph of this rational map is an incidence correspondence, dubbed the *image formation correspondence* in [Aga+23],

$$\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} = \overline{\{(\mathbf{A}, \mathbf{q}, \mathbf{p}) \in (\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n \times (\mathbf{P}^2)^{mn} \mid A_i q_j \sim p_{ij} \quad \forall i \in [m], j \in [n]\}}. \tag{2.2}$$

Given a generic camera arrangement $\bar{\mathbf{A}} = (\bar{A}_1, \ldots, \bar{A}_m) \in \left(\mathbf{P}^{11}\right)^m$, one may also consider the associated *multiview variety*. In the notation of [Aga+23], this may be defined as

$$\Gamma^{m,n}_{\bar{\mathbf{A}},\mathbf{p}} = \{\mathbf{p} \in (\mathbf{P}^2)^{mn} \mid (\bar{\mathbf{A}}, \mathbf{q}, \mathbf{p}) \in \Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} \text{ for some } \mathbf{q} \in \left(\mathbf{P}^3\right)^n\}. \tag{2.3}$$

Multiview varieties and their vanishing ideals are well-understood objects. Our present study of camera resectioning is based on the following definition, which parallels (2.3) in that the role of cameras and 3D points are switched.

Definition 2.2.1 (Resectioning variety/ideal)**.** The m-camera *resectioning variety* associated to a given point arrangement $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$ is the multiprojective variety

$$\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}} = \left\{\mathbf{p} \in \left(\mathbf{P}^2\right)^{mn} \mid (\mathbf{A}, \bar{\mathbf{q}}, \mathbf{p}) \in \Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} \text{ for some } \mathbf{A} \in \left(\mathbf{P}^{11}\right)^m\right\}. \tag{2.4}$$

The vanishing ideal $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$ is the *resectioning ideal* of $\bar{\mathbf{q}}$.

Remark 2.2.2. It turns out that $\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}} = (\mathbf{P}^2)^{mn}$ if and only if $n < 6$, assuming $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$ is sufficiently generic. Thus we assume $n \geq 6$ throughout this section.

To better explain the analogy between resectioning and triangulation, we collect several previous results about the multiview ideals $I(\Gamma^{m,n}_{\bar{\mathbf{A}},\mathbf{p}})$ in Theorem 2.2.3 below. Our first main result, Theorem 2.2.6, involves certain multilinear *focal polynomials* which belong to the resectioning ideal $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$. These are structurally very similar to the classically-known focal polynomials belonging to $I(\Gamma^{m,n}_{\bar{\mathbf{A}},\mathbf{p}})$. We briefly recall a derivation of these constraints. Suppose we are given a camera arrangement $\bar{\mathbf{A}} \in \left(\mathbf{P}^{11}\right)^n$. Consider a generic point

$$(\bar{\mathbf{A}}, \mathbf{q}, \mathbf{p}) = (\bar{A}_1, \ldots, \bar{A}_m, q_1, \ldots, q_n, p_{11}, \ldots, p_{mn}) \in \Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}.$$

Fixing representatives for this point in homogeneous coordinates, there exist nonzero scalars $\lambda_{11}, \ldots, \lambda_{mn} \in \mathbf{C}$ which satisfy the equations

$$\bar{A}_i q_j = \lambda_{ij} p_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n. \tag{2.5}$$

From these conditions, one may obtain certain multilinear polynomials in $\bar{A}_i, p_{ij}$ alone, known in various sources as k-focals or k-multilinearities. Specifically, for each $j = 1, \ldots, n$ and any subset $\sigma = \{\sigma_1, \ldots, \sigma_k\} \subset [m]$ of size ≥ 2, the matrix

$$\begin{bmatrix} \bar{A}_{\sigma_1} & p_{\sigma_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{A}_{\sigma_k} & 0 & \cdots & p_{\sigma_k} \end{bmatrix} \tag{2.6}$$

must be rank-deficient. The maximal $(4+k) \times (4+k)$ minors of these matrices are the *k-focals* associated with the camera arrangement $\bar{\mathbf{A}}$.

In Theorem 2.2.3, we collect several previous results which make the relationship between $\Gamma^{m,n}_{\bar{\mathbf{A}},\mathbf{q},\mathbf{p}}$ and the k-focals more precise. These results impose progressively stronger genericity assumptions on the camera arrangement $\bar{\mathbf{A}}$.

Theorem 2.2.3. *Let $\bar{\mathbf{A}} = (\bar{A}_1, \ldots, \bar{A}_m)$, for $m \geq 2$, be a fixed camera arrangement.*

1. *[AST13, Theorem 2.1] If all maximal 4×4 minors of the matrix $\left[\bar{A}_1^T \mid \cdots \mid \bar{A}_m^T \right]$ are nonzero, then the k-focals for $k \in \{2, 3, 4\}$ form a universal Gröbner basis for $I(\Gamma^{m,n}_{\bar{\mathbf{A}},\mathbf{p}})$.*
2. *[APT21, Theorem 3.7] If $\bar{\mathbf{A}}$ is such that the camera centers are distinct, then the k-focals for $k \in \{2, 3\}$ generate the vanishing ideal $I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}})$.*
3. *[APT21, Theorem 5.6] If $\bar{\mathbf{A}}$ is such that the camera centers are distinct and do not lie in a common plane, then the 2-focals determine $\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$ as a subscheme of $\left(\mathbf{P}^2\right)^{mn}$.*

Turning now to camera resectioning, suppose we are instead given $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$. Similar to (2.5), we wish to obtain conditions involving only $\bar{q}_j$ and p_{ij} from

$$A_i \bar{q}_j = \lambda_{ij} p_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n. \tag{2.7}$$

To obtain these conditions, we may apply a well-known identity involving the matrix Kronecker product, denoted $\otimes$, and the vectorization operator $\text{vec}(\bullet)$, which stacks the columns of a matrix vertically.

Proposition 2.2.4 (See eg. [HJ94, p252, Exercise 22])**.** *For any $M \in \mathbf{C}^{q \times r}$, $N \in \mathbf{C}^{r \times s}$,*

$$\text{vec}\,(MN) = (I_{s \times s} \otimes M)\, \text{vec}\,(N)\,, \tag{2.8}$$

where $I_{s \times s} \in \mathbf{C}^{s \times s}$ is the identity matrix.

We apply this identity with $M = \bar{q}_j^\top$ and $N = A_i^\top$. For the 3×12 matrix $I_{3\times 3} \otimes \bar{q}_j^\top$, we introduce the notation

$$\bar{Q}_j := I_{3\times 3} \otimes \bar{q}_j^\top = \begin{bmatrix} \bar{q}_j^\top & 0 & 0 \\ 0 & \bar{q}_j^\top & 0 \\ 0 & 0 & \bar{q}_j^\top \end{bmatrix}. \tag{2.9}$$

Combining (2.7) and Proposition 2.2.4, we deduce that

$$\bar{Q}_j \operatorname{vec}\left(A_i^\top\right) = \lambda_{ij} p_{ij}, \quad 1 \le i \le m, 1 \le j \le n.$$

Equivalently, for each $i = 1, \dots, m$ we have

$$\begin{bmatrix} \bar{Q}_1 & p_{i1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \\ \bar{Q}_n & 0 & \cdots & p_{in} \end{bmatrix} \begin{bmatrix} \operatorname{vec}(A_i^\top) \\ -\lambda_{i1} \\ \vdots \\ -\lambda_{in} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, if $\mathbf{p} \in \Gamma_{\bar{\mathbf{q}},\mathbf{p}}^{m,n}$, then we have the rank constraints

$$\operatorname{rank} \begin{bmatrix} \bar{Q}_1 & p_{i1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{Q}_n & 0 & \cdots & p_{in} \end{bmatrix} < 12 + n. \tag{2.10}$$

We observe that this rank constraint is equivalent to the vanishing of all maximal $(12 + n) \times (12 + n)$ minors. These minors are homogeneous polynomials in the entries of each $\bar{Q}_i$ and p_{ij}; indeed, for any nonzero scalars $c_1, \dots, c_n, c_1', \dots, c_n'$,

$$\begin{aligned} \operatorname{rank} \begin{bmatrix} c_1\bar{Q}_1 & c_1' p_{i1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ c_n\bar{Q}_n & 0 & \cdots & c_n' p_{in} \end{bmatrix} &= \operatorname{rank} \begin{bmatrix} \bar{Q}_1 & c_1^{-1} p_{i1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{Q}_n & 0 & \cdots & c_n^{-1} p_{in} \end{bmatrix} \\ &= \operatorname{rank} \begin{bmatrix} \bar{Q}_1 & p_{i1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{Q}_n & 0 & \cdots & p_{in} \end{bmatrix}. \end{aligned} \tag{2.11}$$

One may of course consider such rank constraints not only for 3×12 matrices of the form (2.9), but for any given arrangement of surjective linear maps,

$$\bar{B}_j : \mathbf{P}^{11} \dashrightarrow \mathbf{P}^2, \quad j = 1, \dots, n,$$

represented by generic 3×12 matrices. To prevent confusion with cameras A_i, we refer to each $\bar{B}_j$ as a *hypercamera*. We denote a general arrangement of hypercameras by $\bar{\mathbf{B}} = (\bar{B}_1, \ldots, \bar{B}_n) \in \left(\mathbf{P}^{35}\right)^n$. However, we instead write $\bar{\mathbf{Q}}$ to denote the special hypercamera arrangement associated to a point arrangement $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$ by the rule (2.9).

Let us also note that rank constraints analogous to (2.10) hold for any subset of at least 6 world points and their corresponding images. This motivates the following definition, as well as the statement of our first result.

Definition 2.2.5 (*k*-focal, focal ideal)**.** Fix a hypercamera arrangement $\bar{\mathbf{B}} = (\bar{B}_1, \ldots, \bar{B}_n) \in \left(\mathbf{P}^{35}\right)^n$. For any set $\{\sigma_1, \ldots, \sigma_k\} \subset [n]$ of size $k \geq 6$ and an index $i \in [m]$, a *k-focal* polynomial is any maximal $(12+k) \times (12+k)$ minor of the $3k \times (12+k)$ matrix

$$\begin{bmatrix} \bar{B}_{\sigma_1} & p_{i\sigma_1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \bar{B}_{\sigma_k} & 0 & \cdots & p_{i\sigma_k} \end{bmatrix}. \tag{2.12}$$

From context, it will be clear whether "focals" refers to the polynomials in Definition 2.2.5 or their triangulation counterparts. The ideal in $\mathbf{C}[\mathbf{p}]$ generated by all k-focals, $6 \leq k \leq m$, is the m-camera *focal ideal* $I_m(\bar{\mathbf{B}})$. For a given point arrangement $\bar{\mathbf{q}}$, we define its focal ideal $I_m(\bar{\mathbf{q}})$ to be the focal ideal $I_m(\bar{\mathbf{Q}})$ for the associated hypercamera arrangement $\bar{\mathbf{Q}}$.

Theorem 2.2.6. *Let $m, n \geq 1$ be integers. For any point arrangement $\bar{\mathbf{q}} \in (\mathbf{P}^3)^n$ such that no four points are coplanar, we have*

$$I(\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}) = I_m(\bar{\mathbf{q}}),$$

and the set of all k-focals for $6 \leq k \leq 12$ forms a universal Gröbner basis for this ideal.

Theorem 2.2.6 is the resectioning analogue of Theorem 2.2.3 part (1). Directly adapting the proof of this result is not straightforward. This is because $\bar{\mathbf{Q}}$ is a very special hypercamera arrangement. Nevertheless, the noncoplanarity hypothesis in Theorem 2.2.6 ensures that $\bar{\mathbf{Q}}$ is generic enough for Gröbner basis arguments to be applied.

In the setting of triangulation, we note that the range of interesting focals $2 \leq k \leq 4$ is much smaller than in Theorem 2.2.6, and in this setting the k-focals correspond to well-understood objects in multiview geometry—namely, fundamental matrices, trifocal tensors, and quadrifocal tensors [HZ04, cf. Ch. 17]. It would seem that the k-focals for resectioning are less well-understood. Nevertheless, in Example 2.2.8, Section 2.3.2, we observe that they do specialize to "dual" multiview constraints appearing in the literature.

As a warm-up, we establish a set-theoretic variant of Theorem 2.2.6. By analogy with (2.3), let us define for any hypercamera arrangement $\bar{\mathbf{B}} \in \left(\mathbf{P}^{35}\right)^n$ the variety $\Gamma^{n,m}_{\bar{\mathbf{B}},\mathbf{p}}$ to be the closed image of the associated imaging map $\left(\mathbf{P}^{11}\right)^m \dashrightarrow \left(\mathbf{P}^2\right)^{mn}$. In other words, $\Gamma^{n,m}_{\bar{\mathbf{B}},\mathbf{p}}$ is a hypercamera version of the multiview variety. When $\bar{\mathbf{B}} = \bar{\mathbf{Q}}$, we have the following result.

Proposition 2.2.7. *Fix $\bar{\mathbf{q}} = (\bar{q}_1, \ldots, \bar{q}_n) \in (\mathbf{P}^3)^n$ with no four $\bar{q}_i$ coplanar. Then*

$$\Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}} = \Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}} = \mathrm{V}(I_m(\bar{\mathbf{q}})).$$

Proof. It is relatively straightforward to prove the inclusions

$$\Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}} \subset \Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}} \subset \mathrm{V}(I_m(\bar{\mathbf{q}})),$$

so we focus on the harder inclusion $\mathrm{V}(I_m(\bar{\mathbf{q}})) \subset \Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}}$. This is also where we need the noncoplanarity assumption. Consider any point

$$\mathbf{p} \in \mathrm{V}(I_m(\bar{\mathbf{q}})).$$

We will construct a sequence of points $(\mathbf{p}^{(k)}) \in \Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}}$ converging to $\mathbf{p}$. To simplify notation in what follows, we consider the case $m = 1$. When $m > 1$, the same construction applies component-wise. We write p_i in place of p_{1i}, so that the kernel of the matrix

$$\begin{bmatrix} \bar{Q}_1 & p_1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ \bar{Q}_n & 0 & \cdots & p_n \end{bmatrix}$$

contains a point $v = [v_1 : \cdots : v_{12+n}] \in \mathbf{P}^{11+n}$. Let us fix homogeneous coordinates for $p_1, \ldots, p_n, \bar{q}_1, \ldots, \bar{q}_n, v$. We define

$$A = \begin{bmatrix} v_1 & \cdots & v_4 \\ \vdots & \ddots & \vdots \\ v_9 & \cdots & v_{12} \end{bmatrix}, \quad \lambda_j = -v_{12+j}.$$

Let us first observe that the matrix A is nonzero, for otherwise we would have

$$\lambda_j p_j = A\bar{q}_j = \bar{Q}_j \operatorname{vec}(A^\top) = 0 \quad \Rightarrow \quad \lambda_j = 0$$

for all j, contradicting the fact that $v \neq 0$. Next, observe that at most three of the λ_j can be zero: otherwise, four of the points $\bar{q}_j$ would lie in some plane containing the kernel of A, contradicting our hypothesis that $\bar{\mathbf{q}}$ is noncoplanar. It follows that we can find a nonzero 3×4 matrix A' with $A'\bar{q}_j = p_j$ for each j with $\lambda_j = 0$. Fix such a matrix A'. We now construct the desired sequence $(\mathbf{p}^{(k)})_{k \geq 1}$. Set

$$\begin{aligned} A^{(k)} &= A + (1/k)A', \\ p_j^{(k)} &= A^{(k)}\bar{q}_j, \quad j = 1, \ldots, n. \end{aligned}$$

To show convergence, note that when $\lambda_j = 0$ we have

$$p_j^{(k)} = (1/k)A'\bar{q}_j \sim A'\bar{q}_j = p_j.$$

For $\lambda_j \neq 0$ we attain p_j in the limit as $k \to \infty$, since

$$p_j^{(k)} = A\bar{q}_j + (1/k)A'\bar{q}_j \to A\bar{q}_j \sim p_j.$$

In summary, we have found for any point $\mathbf{p} \in \mathrm{V}(I_m(\bar{\mathbf{q}}))$ a sequence of points $(\mathbf{p}^{(k)}) \in \Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}}$ converging to $\mathbf{p}$. Since $\Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}}$ is closed in the Euclidean topology, we deduce the needed inclusion: $\mathrm{V}(I_m(\bar{\mathbf{q}})) \subset \Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}}$. □

We conclude this section with the simplest interesting example of a resectioning variety.

Example 2.2.8. For $m = 1$ camera and a generic point arrangement $\bar{\mathbf{q}} = (\bar{q}_1, \ldots, \bar{q}_6) \in (\mathbf{P}^3)^n$, the resectioning variety $\Gamma^{1,6}_{\bar{\mathbf{q}},\mathbf{p}}$ is a hypersurface in $(\mathbf{P}^2)^6$. Applying a suitable permutation to the rows of the 6-focal matrix, this hypersurface has an 18×18 determinantal representation,

$$\det \begin{bmatrix} \bar{q}_1^T & & & p_1[1] & & \\ \vdots & & & & \ddots & \\ \bar{q}_6^T & & & & & p_6[1] \\ & \bar{q}_1^T & & p_1[2] & & \\ & \vdots & & & \ddots & \\ & \bar{q}_6^T & & & & p_6[2] \\ & & \bar{q}_1^T & p_1[3] & & \\ & & \vdots & & \ddots & \\ & & \bar{q}_6^T & & & p_6[3] \end{bmatrix} = 0. \tag{2.13}$$

Out of the $3^6 = 729$ possible terms of a sextilinear form on $(\mathbf{P}^2)^6$, the special structure of the 6-focal determinant dictates that only $\binom{6}{2}\binom{4}{2} = 90$ can be nonzero. On the other hand, Lemma 2.4.3 below shows that applying a general linear change of coordinates to (2.13) has the effect that all 729 possible terms become nonzero. This highlights an important distinction between resectioning and multiview ideals—the initial ideal for generic data $\bar{\mathbf{q}}$ is not the same as the $\mathbf{Z}^6$-graded Borel-fixed generic initial ideal (cf. [AST13, §3].) Letting $<$ denote the lexicographic order with $p_6[3] < p_6[2] < p_6[1] < p_5[3] < \cdots < p_1[1]$, we have

$$\begin{aligned} \text{in}_<(I(\Gamma^{1,6}_{\bar{\mathbf{q}},\mathbf{p}}) &= \langle p_1[1]p_2[1]p_3[2]p_4[2]p_5[3]p_6[3]\rangle, \\ \text{gin}_<(I(\Gamma^{1,6}_{\bar{\mathbf{q}},\mathbf{p}}) &= \langle p_1[1]p_2[1]p_3[1]p_4[1]p_5[1]p_6[1]\rangle. \end{aligned} \tag{2.14}$$

Interestingly, (2.13) also has several smaller determinantal representations. Many of these may be obtained from (2.13) using Schur complements. For example, we have the 12×12 determinantal representation

$$(p_1[3]\cdots p_6[3])^{-1}\det\begin{bmatrix} p_1[3]\bar{q}_1^T & & -p_1[1]\bar{q}_1^T \\ \vdots & & \vdots \\ p_6[3]\bar{q}_6^T & & -p_6[1]\bar{q}_6^T \\ & p_1[3]\bar{q}_1^T & -p_1[2]\bar{q}_1^T \\ & \vdots & \vdots \\ & p_6[3]\bar{q}_6^T & -p_6[2]\bar{q}_6^T \end{bmatrix} = 0. \tag{2.15}$$

The 6-focal determinant also has a 6×6 determinantal representation, which specializes to (2.26) below after fixing $\bar{q}_1, \ldots, \bar{q}_4, p_1, \ldots, p_4$. This is the classical form of the constraint appearing in works such as [Qua95; CW98]. Finally, we note that Schaffilitzky et al. [Sch+00] derive a 3×3 determinantal constraint relating 3D points and their 2D projections that is linear in a distinguished image point p_6. Theorem 2.2.6 implies that their determinant is a multiple of the 6-focal determinant. Notably, Schaffilitzky et al. use their constraint to solve the minimal problem of reconstructing 6 points from 3 views. Earlier works, eg. [CW98], had already observed that this problem is equivalent to the classical 7 point problem in 2 views. This equivalence follows from the principle of Carlsson-Weinshall duality, which we revisit in the next section.

2.3 *Carlsson-Weinshall duality revisited*

Recall the image formation variety $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ from (2.2). Previous work of Agarwal et al. [Aga+23] explains how the problems of reconstruction, triangulation, and resectioning may all be understood in terms of slicing and projection operations on this variety. The relationships between the varieties produced by these operations are summarized in a diagram designated as an *atlas* for the pinhole camera. One striking feature of the atlas's appearance is the apparently symmetric roles of cameras in $\mathbf{P}^{11}$ and world points in $\mathbf{P}^3$. A simple explanation for this phenomenon is as follows: for a given camera center $c \in \mathbf{P}^3$, world point $q \in \mathbf{P}^3$, and image plane $L \in \mathrm{Gr}(\mathbf{P}^2, \mathbf{P}^3)$, we obtain the same projected point on L whether we project c through q or project q through c. If we want to express this symmetry in terms of camera matrices instead of camera centers, one approach is to introduce coordinates on the image plane. Indeed, there are an additional $\dim \mathrm{PGL}_3 = 8 = 11 - 3$ degrees of freedom in choosing projective coordinates on L. A particular choice of coordinates leads directly to the framework of *Carlsson-Weinshall (CW) duality* from the multiview geometry literature.

In this section, we point out that several world-to-image point constraints which were previously discovered using CW duality arise naturally as specializations of our focal constraints. We also show in Theorem 2.3.5 that Carlsson-Weinshall duality gives rise to a rational quotient of the image formation correspondence, and develop a *reduced* version of the atlas that better explains the symmetry between cameras and world points—see Figure 2.3.

A direct application of the focal constraints described in Section 2.2 arises naturally in the setting of Carlsson-Weinshall (CW) duality. In the eponymous authors' celebrated work, CW duality is described as the notion that "*problems of [resectioning] and [triangulation] from image data are... dual in the sense that they can be solved with the same algorithm depending on the number of [world] points and cameras*" [CW98].

In this section, we develop CW duality in the context of a *reduced atlas*, analogous to that of [Aga+23], which makes the symmetry between cameras and 3D points evident. Theorems 2.3.5 and 2.3.6 explain how nodes in this atlas arise as rational quotients of their non-reduced counterparts. The latter result also includes an analogue of Theorem 2.2.6: the reduced resectioning variety is cut out scheme-theoretically by bilinear forms for a sufficiently generic point configuration.

Remark 2.3.1. Recent work by Trager, Hebert, and Ponce [THP19] demonstrates that the exact coordinates of the camera centers and world points are not essential features of CW duality, contrary to the original setup. For simplicity, we state the main results of this section with respect to the conventional projective frame defined in (2.20).

2.3.1 Geometric formulation

For m cameras and n world points, we define the *reduced image formation correspondence* to be the variety

$$\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} = \overline{\{(\mathbf{a},\mathbf{q},\mathbf{p}) \in (\mathbf{P}^3)^m \times (\mathbf{P}^3)^n \times (\mathbf{P}^2)^{mn} \mid A(a_i)\cdot q_j \sim p_{ij} \quad \forall i \in [m],\, j \in [n]\}}, \tag{2.16}$$

where for $a_i = [a_{i1} : a_{i2} : a_{i3} : a_{i4}] \in \mathbf{P}^3$ we define

$$A(a_i) = \begin{bmatrix} a_{i1} & 0 & 0 & a_{i4} \\ 0 & a_{i2} & 0 & a_{i4} \\ 0 & 0 & a_{i3} & a_{i4} \end{bmatrix}. \tag{2.17}$$

When $A(a_i)$ is of full rank, we call it the *reduced camera matrix* associated to the point a_i. The center of a reduced camera matrix $A(a_i)$ is $\mathcal{C}(a_i)$, where $\mathcal{C}$ is the quadratic Cremona involution

$$\begin{aligned} \mathcal{C} : \mathbf{P}^3 &\dashrightarrow \mathbf{P}^3 \\ [a_1 : a_2 : a_3 : a_4] &\mapsto [1/a_1 : 1/a_2 : 1/a_3 : -1/a_4]. \end{aligned} \tag{2.18}$$

Note that $\mathcal{C}(a_i)$ is defined exactly when at most one a_{ij} is zero, or equivalently, when $A(a_i)$ is a full-rank camera matrix with a well-defined center.

The key observation of Carlsson-Weinshall duality is expressed by the symmetric roles of a 3D point q_j and a reduced camera $A(a_i)$ in image formation:

$$A(a_i)q_j = A(q_j)a_i \quad \forall i = 1,\ldots,m,\ j = 1,\ldots,n. \tag{2.19}$$

The special form of the reduced camera matrix arises from fixing a projective basis in each image and a partial projective basis in the world. We adopt the notation of [HZ04, Ch. 16]:

$$\begin{aligned} E_1 &= [1:0:0:0], \quad e_1 = [1:0:0], \\ E_2 &= [0:1:0:0], \quad e_2 = [0:1:0], \\ E_3 &= [0:0:1:0], \quad e_3 = [0:0:1], \\ E_4 &= [0:0:0:1], \quad e_4 = [1:1:1], \\ E_5 &= [1:1:1:1]. \end{aligned} \tag{2.20}$$

Each set of four points $E_1,\ldots E_4 \in \mathbf{P}^3$, $e_1,\ldots,e_4 \in \mathbf{P}^3$ is said to span a *reference tetrahedron* in $\mathbf{P}^3$. The geometry relating these points and the Cremona transformation $\mathcal{C}$ can be appreciated in Figure 2.2.

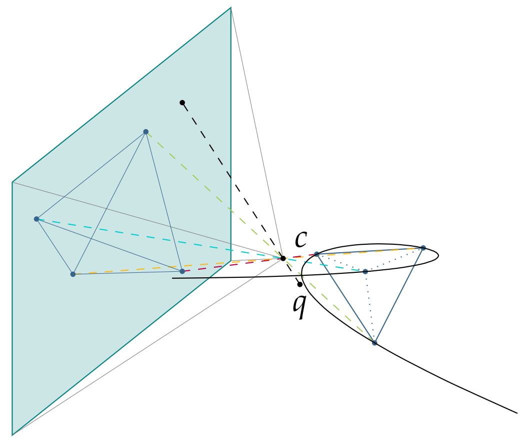


Figure 2.2: Four fixed points $E_1, \ldots, E_4 \in \mathbf{P}^3$ determine a reference tetrahedron. They project through the camera center $c \in \mathbf{P}^3$ to the four points $e_1, \ldots e_4 \in \mathbf{P}^2$. The Cremona transformation $\mathcal{C}$ maps the line through c and q to the unique twisted cubic passing through $\mathcal{C}(q), \mathcal{C}(c), E_1, \ldots, E_4$.

Given a camera of the form $A = A(a), a \in \mathbf{P}^3$, we have that $A(a)E_i = e_i$ for $i = 1, \ldots, 4$. The converse is true as well; that is, a camera matrix takes the reduced form (2.17) if and only if it sends E_i to e_i for each $i = 1, \ldots, 4$. As we will soon demonstrate, there is a rational group action by an algebraic group $\mathcal{G}_m$ on $\Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ for which each $\mathcal{G}_m$-orbit, where defined, contains a unique element of $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$. That is, $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ can be thought of as a kind of quotient of the general image formation correspondence. Theorem 2.3.5 makes this precise using the notion of a *rational quotient.* In this reduced setting, the roles of camera centers and world points are manifestly symmetric: a point $(\mathbf{a}, \mathbf{q}, \mathbf{p}) \in \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ is also a point in $\mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ after swapping the $\mathbf{a}$ and $\mathbf{q}$ factors. By this observation, we then get the isomorphism $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \simeq \mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}}$.

Just as a point in $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ can be thought of as a configuration of cameras and points, a point in $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ can be thought of as such a configuration up to certain coordinate changes. More precisely, points in $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ correspond to orbits in $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ under the action of a group $\mathcal{G}_m$ consisting of coordinate changes in the world and each of the m images. Up to this group action, we may assume the image planes $L_1, \ldots, L_m \in \mathrm{Gr}(\mathbf{P}^2, \mathbf{P}^3)$ are all equal, ie. $L_1 = \ldots = L_m$. This explains the center image in Figure 2.1.

We now transition into a formal treatment of the notions described above. Define $\mathcal{G}_m = (\mathrm{PGL}_3)^m \times \mathrm{Stab}_{\,\mathrm{PGL}_4}(E_5)$, an algebraic group of dimension $8m + 12$ which acts rationally on $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ as follows:

$$\begin{aligned} \mathcal{G}_m \times \Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} &\dashrightarrow \Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} \\ (T_1, \ldots, T_m, S) \cdot (A_1, \ldots, A_m, q_1, \ldots q_n, p_{11}, \ldots, p_{mn}) & \\ = (T_1 A_1 S^{-1}, \ldots, T_m A_m S^{-1}, &S q_1, \ldots, S q_n, T_1 p_{11}, \ldots, T_m p_{mn}). \end{aligned} \tag{2.21}$$

To formalize the intuition that $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ is a quotient of $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ by $\mathcal{G}_m$, we recall the definition of a *rational quotient* as follows.

Definition 2.3.2 (rational quotient)**.** (cf. [Dol03, §6.2].) Let X and Y be irreducible algebraic varieties and G an algebraic group acting rationally on X. We say Y is a *rational quotient* of X by G, and write $X/G \simeq_{\mathbf{Bir}} Y$, if Y is a model for the field of G-invariant rational functions on X: that is, if there exists an isomorphism $\mathbf{C}(Y) \cong \mathbf{C}(X)^G$.

A classical result due to Rosenlicht states that rational quotients always exist over any algebraically closed field (cf. [Dol03, Theorem 6.2].) The following simple lemma provides sufficient conditions for recognizing a particular class of rational quotients in which the action yields a birational equivalence of X with $G \times Y$.

Lemma 2.3.3. *Let G be an algebraic group acting rationally on an irreducible variety X. For an irreducible subvariety $Y \subset X$, we have $X/G \simeq_{\mathbf{Bir}} Y$ if there exists a rational map*

$$\begin{aligned} \mu_G : X &\dashrightarrow G \\ x &\mapsto \mu_G(x) \end{aligned}$$

such that

1. *for all x in some dense open subset $U \subset X$ such that $U \cap Y \neq \emptyset$, we have that $\mu_G(x) = id$ if and only if $x \in Y$, and*
2. *$\mu_G(g \cdot x) = \mu_G(x)\, g^{-1}$ for all (g, x) in dense open subset of $G \times X$.*

Moreover, these assumptions imply that

$$\begin{aligned} X &\dashrightarrow G \times Y \\ x &\mapsto (\mu_G(x), \mu_G(x) \cdot x) \end{aligned}$$

is a birational equivalence, with a rational inverse given by

$$\begin{aligned} G \times Y &\dashrightarrow X \\ (g, y) &\mapsto g^{-1} \cdot y. \end{aligned}$$

Remark 2.3.4. The assumptions of Lemma 2.3.3 imply the map $X \dashrightarrow G \times Y$ is well-defined:

$$\mu_G(\mu_G(x) \cdot x) = \mu_G(x)\mu_G(x)^{-1} = \mathrm{id} \quad \Rightarrow \quad \mu_G(x) \cdot x \in Y.$$

Proof. A function $f \in \mathbf{C}(Y)$ pulls back to a function $h \in \mathbf{C}(X)$ defined by $h(x) = f(\mu_G(x) \cdot x)$ on X. Our assumptions imply h is G-invariant, since

$$h(g \cdot x) = f(\mu_G(g \cdot x) \cdot (g \cdot x)) = f\left((\mu_G(x)g^{-1}) \cdot (g \cdot x)\right) = h(x).$$

Let us write $\varphi^* : \mathbf{C}(Y) \to \mathbf{C}(X)^G$. Since $Y \subset X$, we also have the induced map $\iota^* : \mathbf{C}(X)^G \to \mathbf{C}(Y)$. We show that φ^* and ι^* are mutual inverses. Taking any function $f \in \mathbf{C}(Y)$ and $y \in Y$ in its domain of definition, we calculate

$$\iota^* \varphi^* f(y) = f(\mu_G(y) \cdot y) = f(y).$$

Similarly, for any fixed $f \in \mathbf{C}(X)^G$, the values $f(x)$ and $f(\mu_G(x) \cdot x)$ are defined for a dense open subset of $x \in X$, for which we compute

$$\varphi^* \iota^* f(x) = f(\mu_G(x) \cdot x) = f(x).$$

This proves $X/G \simeq_{\mathbf{Bir}} Y$. The birational equivalence of $G \times Y$ and X follows similarly. □

Theorem 2.3.5 (CW duality). *For any $m, n \geq 0$, we have a birational equivalence of varieties*

$$\Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} \cong_{Bir} \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \times \mathcal{G}_m,$$

which yields the following commutative diagram (in which each arrow labeled $\sim$ is a birational or biregular isomorphism).

$$\begin{array}{ccc}
\Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} \times (\mathrm{PGL}_3)^n & \overset{\sim}{\dashrightarrow} & \Gamma^{n,m+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} \times (\mathrm{PGL}_3)^m \\
\wr \downarrow & & \wr \downarrow \\
(\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \times \mathcal{G}_m) \times (\mathrm{PGL}_3)^n & \overset{\sim}{\dashrightarrow} & (\mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}} \times \mathcal{G}_n) \times (\mathrm{PGL}_3)^m \\
\downarrow & & \downarrow \\
\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} & \overset{\sim}{\longrightarrow} & \mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}}
\end{array}$$

This diagram has the following additional properties:

1. *If $\nu_{m,n}$ denotes any of the horizontal maps, we have $\nu_{n,m} \circ \nu_{m,n} = \mathrm{id}$ wherever both maps are defined.*

2. *The vertical maps express the reduced image formation variety as a rational quotient of the image formation correspondence,*

$$\Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} / \mathcal{G}_m \simeq_{\mathit{Bir}} \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}. \tag{2.22}$$

3. *The duality between the problems of exact resectioning and triangulation may be expressed in terms of this commutative diagram and certain projections: eg., for the bottom row, if $\pi'_{\mathbf{a}}, \pi'_{\mathbf{q}}$ denote the projections from $\mathrm{P}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ that forget the $\mathbf{a}$ and $\mathbf{q}$ factors, then the diagram below commutes.*

$$\begin{array}{ccccc} & & (\mathbf{P}^3)^n \times (\mathbf{P}^2)^{mn} & & \\ & \nearrow^{\pi'_{\mathbf{a}}} & & \nwarrow^{\pi'_{\mathbf{q}}} & \\ \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} & & \xrightarrow{\quad\sim\quad} & & \mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}} \\ & \searrow_{\pi'_{\mathbf{q}}} & & \swarrow_{\pi'_{\mathbf{a}}} & \\ & & (\mathbf{P}^3)^m \times (\mathbf{P}^2)^{mn} & & \end{array}$$

Proof. We begin by constructing the maps that yield the rational quotient (2.22). This part follows by applying Lemma 2.3.3 with $X = \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}}$, $Y = \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$, and $G = \mathcal{G}_m$. To obtain the inclusion $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \subset \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ we define

$$\begin{aligned} \iota : \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} &\to \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} \\ (a_1, \ldots, a_m, q_1, \ldots, q_n, p_{11}, \ldots, p_{mn}) &\mapsto \\ (A(a_1), \ldots, A(a_m), E_1, E_2, E_3, E_4, &q_1, \ldots, q_n, e_1, e_2, e_3, e_4, p_{11}, \ldots, p_{mn}). \end{aligned}$$

To construct the map $\mu_{\mathcal{G}_m} : \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} \dashrightarrow \mathcal{G}_m$, consider first the map

$$\begin{aligned} S : (\mathbf{P}^3)^4 &\dashrightarrow \mathrm{Stab}_{\mathrm{PGL}_4}(E_5) \\ (q_1, \ldots, q_4) &\mapsto \left(\begin{bmatrix} q_1 \mid q_2 \mid q_3 \mid q_4 \end{bmatrix} \cdot \mathrm{diag}\left([5234]_{\mathbf{q}}, [1534]_{\mathbf{q}}, [1254]_{\mathbf{q}}, [1235]_{\mathbf{q}}\right) \right)^{-1}, \end{aligned}$$

where each $[5234]_{\mathbf{q}}, \ldots, [1235]_{\mathbf{q}}$ is the determinant of a matrix obtained by replacing $q_1, \ldots, q_4$ with E_5 in the 4×4 matrix whose columns are $q_1, \ldots, q_4$. We verify that $S(q_1, \ldots, q_4)$ is well-defined and contained in $\mathrm{Stab}_{\mathrm{PGL}_4}(E_5)$ using linear algebra. To ease notation, we write $Q = \left[\, q_1 \mid q_2 \mid q_3 \mid q_4 \,\right]$ and $D = \operatorname{diag}\left([5234]_{\mathbf{q}}, [1534]_{\mathbf{q}}, [1254]_{\mathbf{q}}, [1235]_{\mathbf{q}}\right)$. Rescaling any of the $q_1, \ldots, q_4$ then rescales the matrix product QD. Using Cramer's rule, we calculate that

$$\begin{aligned} S(q_1, \ldots, q_4)E_5 &= D^{-1} \cdot Q^{-1} E_5 \\ &= D^{-1}[[5234]_{\mathbf{q}} : [1534]_{\mathbf{q}} : [1254]_{\mathbf{q}} : [1235]_{\mathbf{q}}] \\ &= E_5. \end{aligned}$$

An analogous calculation verifies that for any $S_0 \in \mathrm{PGL}_4$ we have

$$S(S_0 q_1, \ldots, S_0 q_4) = \left(S_0 Q \cdot (\det(S_0))^3\, D\right)^{-1} = S(q_1, \ldots, q_4)\, S_0^{-1}. \tag{2.23}$$

Note additionally, that if $S_0 \in \mathrm{Stab}_{\mathrm{PGL}_4}(E_5)$, then so also must $S(S_0 q_1, \ldots, S_0 q_4) \in \mathrm{Stab}_{\mathrm{PGL}_4}(E_5)$. Similar to our definition of S above, we may define a map

$$\begin{aligned} T : (\mathbf{P}^2)^4 &\dashrightarrow \mathrm{PGL}_3 \\ (p_1, \ldots, p_4) &\mapsto \left(\left[\, p_1 \mid p_2 \mid p_3 \,\right] \cdot \operatorname{diag}\left([423]_{\mathbf{p}}, [143]_{\mathbf{p}}, [124]_{\mathbf{p}}\right) \right)^{-1}, \end{aligned}$$

but we replace $p_1, \ldots, p_3$ by p_4 (rather than e_4) when forming the expressions $[423]_{\mathbf{p}}, \ldots, [124]_{\mathbf{p}}$. Once again, for $T_0 \in \mathrm{PGL}_3$ we have

$$T(T_0 p_1, \ldots, T_0 p_4) = T(p_1, \ldots, p_4)\, T_0^{-1}. \tag{2.24}$$

Finally, we define

$$\begin{aligned} \mu_{\mathcal{G}_m} : \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}} &\dashrightarrow \mathcal{G}_n \\ (A_1, \ldots, A_m, q_1, \ldots, q_{n+4}, p_{11}, \ldots, p_{mn}) &\mapsto \\ &(T(p_{11}, \ldots, p_{14}), \ldots, T(p_{11}, \ldots, p_{m4}), S(q)). \end{aligned}$$

We check that the two assumptions of Lemma 2.3.3 are satisfied. The map $\mu_{\mathcal{G}_m}$ fixes $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ pointwise since $T(e_1, e_2, e_3, e_4)$ and $S(E_1, E_2, E_3, E_4)$ both act as the identity. Similarly, for sufficiently generic $g \in \mathcal{G}_m$ and $x \in \Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ the assumption that $\mu_{\mathcal{G}_m}(g \cdot x) = \mu_{\mathcal{G}_m}(x)\, g^{-1}$ follows from (2.23) and (2.24).

Thus, we may conclude from Lemma 2.3.3 that we have the rational quotient (2.22), giving property 2 in the statement of the theorem. Moreover, the lemma implies that $\Gamma^{m,n+4}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ is birationally equivalent to $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \times \mathcal{G}_m$, which allows us to define the vertical maps in the main diagram. To complete the diagram, it suffices to define the bottommost map, which is

$$\nu_{m,n} : \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \to \mathrm{P}^{n,m}_{\mathbf{a},\mathbf{q},\mathbf{p}}$$
$$(a_1, \dots, a_m, q_1, \dots, q_n, p_{11}, \dots p_{mn}) \mapsto (q_1, \dots, q_n, a_1, \dots, a_m, p_{11}, \dots p_{nm}).$$

Now, to show that $\nu_{m,n}$ is an isomorphism, we use the symmetric equations (2.19). The remaining parts of the theorem now follow easily. □

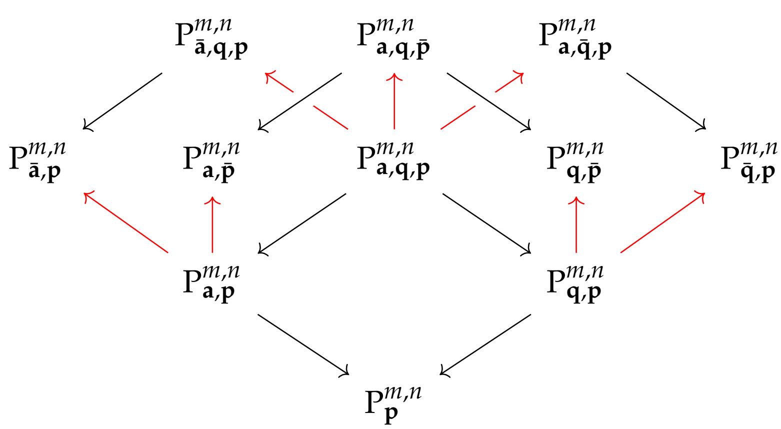


Figure 2.3: An atlas for the reduced pinhole camera, cf. [Aga+23, Figure 1].

The reduced image formation correspondence $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ sits at the center of the *reduced atlas* depicted in Figure 2.3. Following [Aga+23], we may define the remaining entities in this figure using slices and projections of the reduced image formation correspondence. For instance, the varieties $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{p}}$ and $\mathrm{P}^{m,n}_{\mathbf{q},\mathbf{p}}$ are defined, respectively, as the image under the coordinate projections $\pi'_{\mathbf{q}} : \mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}} \to \left(\mathbf{P}^3\right)^m \times \left(\mathbf{P}^2\right)^{mn}$, $\pi'_{\mathbf{a}} : \left(\mathbf{P}^3\right)^n \times \left(\mathbf{P}^2\right)^{mn}$ appearing in Theorem 2.3.5. Slicing the variety $\mathrm{P}^{m,n}_{\mathbf{q},\mathbf{p}}$ with the coordinate planes defined by $\mathbf{q} = \bar{\mathbf{q}}$, we obtain the *reduced resectioning variety* $\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$. Up to Zariski closure, this is the *dual reduced joint image* introduced in [THP19]. Theorem 2.3.5 above and Theorem 2.3.6 below illustrate the correspondence between these varieties and their non-reduced counterparts in [Aga+23], and provide an explanation for the vertical symmetry present in both atlases.

2.3.2 Algebraic consequences

Readers familiar with multiview geometry will no doubt wonder how the focal constraints of Theorem 2.2.6 relate to various "dual multiview constraints", derived by Carlsson, Weinshall, and others. All of these previously-studied constraints may be interpreted as polynomials vanishing on $\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$. Specializing the 6-focal constraints to $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$, we obtain

$$\det \begin{bmatrix} E_1^\top \otimes I & e_1 & & & & & \\ E_2^\top \otimes I & & e_2 & & & & \\ E_3^\top \otimes I & & & e_3 & & & \\ E_4^\top \otimes I & & & & e_4 & & \\ \bar{Q}_{j_1} & & & & & p_{ij_1} & \\ \bar{Q}_{j_2} & & & & & & p_{ij_2} \end{bmatrix} = 0, \quad \forall i = 1, \dots, m,\ 1 \le j_1 < j_2 \le n. \tag{2.25}$$

Permuting the rows, (2.25) implies that

$$\det \left[\begin{array}{c|cccccc} & E_1 & & & E_4 & & \\ I_{12\times 12} & & E_2 & & E_4 & & \\ & & & E_3 & E_4 & & \\ \hline \bar{q}_{j_1}^\vee & & & & & p_{ij_1} & \\ \bar{q}_{j_2}^\vee & & & & & & p_{ij_2} \end{array} \right] = 0,$$

and taking the Schur complement, we find

$$\det \left(\begin{bmatrix} 0_{3\times 4} & p_{ij_1} & \\ 0_{3\times 4} & & p_{ij_2} \end{bmatrix} - \begin{bmatrix} \bar{q}_{j_1}^\vee \\ \bar{q}_{j_2}^\vee \end{bmatrix} \begin{bmatrix} E_1 & & & E_4 \\ & E_2 & & E_4 \\ & & E_3 & E_4 \end{bmatrix} \right) =$$

$$\det \begin{bmatrix} A(\bar{q}_{j_1}) & p_{ij_1} & \\ A(\bar{q}_{j_2}) & & p_{ij_2} \end{bmatrix} = 0. \tag{2.26}$$

Equation (2.26) is a bilinear form in p_{ij_1} and p_{ij_2}, which may be represented by Carlsson and Weinshall's 3×3 *dual fundamental matrix* (cf. [CW98, eq. 18]),

$$\begin{bmatrix} 0 & \bar{q}_{j_1}[2]\bar{q}_{j_2}[1] \det \begin{bmatrix} \bar{q}_{j_1}[4] & \bar{q}_{j_1}[3] \\ \bar{q}_{j_2}[4] & \bar{q}_{j_2}[3] \end{bmatrix} & \bar{q}_{j_1}[3]\bar{q}_{j_2}[1] \det \begin{bmatrix} \bar{q}_{j_1}[2] & \bar{q}_{j_1}[4] \\ \bar{q}_{j_2}[2] & \bar{q}_{j_2}[4] \end{bmatrix} \\ \bar{q}_{j_1}[1]\bar{q}_{j_2}[2] \det \begin{bmatrix} \bar{q}_{j_1}[3] & \bar{q}_{j_1}[4] \\ \bar{q}_{j_2}[3] & \bar{q}_{j_2}[4] \end{bmatrix} & 0 & \bar{q}_{j_1}[3]\bar{q}_{j_2}[2] \det \begin{bmatrix} \bar{q}_{j_1}[4] & \bar{q}_{j_1}[1] \\ \bar{q}_{j_2}[4] & \bar{q}_{j_2}[1] \end{bmatrix} \\ \bar{q}_{j_1}[1]\bar{q}_{j_2}[3] \det \begin{bmatrix} \bar{q}_{j_1}[4] & \bar{q}_{j_1}[2] \\ \bar{q}_{j_2}[4] & \bar{q}_{j_2}[2] \end{bmatrix} & \bar{q}_{j_1}[2]\bar{q}_{j_2}[3] \det \begin{bmatrix} \bar{q}_{j_1}[1] & \bar{q}_{j_1}[4] \\ \bar{q}_{j_2}[1] & \bar{q}_{j_2}[4] \end{bmatrix} & 0 \end{bmatrix}. \tag{2.27}$$

This construction yields a total of $m\binom{n}{2}$ bilinear equations vanishing on the reduced joint image $\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$. A similar application of this Schur complement trick to suitably-chosen 7- and 8-focals leads to the *dual trifocal and quadrifocal tensors* (cf. [CW98, §6.3–6.4], [THP19, §3, 4]).

For a sufficiently generic point configuration $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$, it turns out that the reduced 2-focals (2.26) determine $\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$ as a subscheme of $\left(\mathbf{P}^2\right)^{mn}$. Theorem 2.3.6 states precise genericity conditions such that this occurs. Thus, while equations needed to cut out $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ in a strong sense have very high degree, only bilinear equations are needed to cut out its quotient $\mathrm{P}^{m,n}_{\mathbf{a},\mathbf{q},\mathbf{p}}$ in a weaker sense. The essential insight is, via Carlsson-Weinshall duality, that $\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$ is simply the direct product of "ordinary" multiview varieties,

$$\mathrm{P}^{m,n}_{\bar{\mathbf{q}},\mathbf{p}} = \Gamma^{n,m}_{A(\bar{\mathbf{q}}),\mathbf{p}} \cong \Gamma^{n,1}_{A(\bar{\mathbf{q}}),\mathbf{p}} \times \cdots \times \Gamma^{n,1}_{A(\bar{\mathbf{q}}),\mathbf{p}} \quad \text{where } A(\bar{\mathbf{q}}) = (A(\bar{q}_1), \ldots, A(\bar{q}_n)) . \tag{2.28}$$

A previous result, namely part (3) of Theorem 2.2.3, states that the multiview variety of a sufficiently generic camera arrangement is cut out by the bilinear forms in its vanishing ideal. The Cremona transformation $\mathcal{C}$ allows us to translate these genericity conditions on the cameras $A(\bar{\mathbf{q}})$ into conditions on the point arrangement $\bar{\mathbf{q}} \in \left(\mathbf{P}^2\right)$. A 4-nodal cubic surface in $\mathbf{P}^3$ containing the points $E_1, \ldots E_4$ is given by an equation of the form

$$a_1 x_2 x_3 x_4 + a_2 x_1 x_3 x_4 + a_3 x_1 x_2 x_4 + a_4 x_1 x_2 x_3 = 0, \quad [a_1 : a_2 : a_3 : a_4] \in \mathbf{P}^3. \tag{2.29}$$

If all a_i are nonzero, then such a surface is projectively equivalent to *Cayley's nodal cubic surface*, for which $a_1 = a_2 = a_3 = a_4 = 1$ and the points $E_1, \ldots, E_4$ comprise the singular locus. We also allow degenerate cases where one or more $a_i = 0$ in (2.29), in which case the surface degenerates to the union of a plane and a quadric, or the union of three planes.

Theorem 2.3.6. *Fix n distinct points* $\bar{q}_1, \ldots, \bar{q}_n \in \mathbf{P}^3 \setminus \{E_1, E_2, E_3, E_4\}$, $n \geq 2$, *such that no four* $\bar{q}_j$ *lie on a common 4-nodal cubic surface through* $E_1, \ldots E_4$. *Write*

$$\bar{\mathbf{q}} = (\bar{q}_1, \ldots, \bar{q}_n, E_1, \ldots, E_4),$$

and $\bar{\mathbf{q}}'$ *for the sub-arrangement of* $\bar{\mathbf{q}}$ *obtained by deleting the* $E_1, \ldots, E_4$. *We have a birational equivalence of varieties*

$$\Gamma^{m,n+4}_{\bar{\mathbf{q}},\mathbf{p}} \simeq_{\mathbf{Bir}} \mathrm{P}^{m,n}_{\bar{\mathbf{q}}',\mathbf{p}} \times (\mathrm{PGL}_3)^m,$$

which realizes the reduced resectioning variety $\mathrm{P}^{m,n}_{\bar{\mathbf{q}}',\mathbf{p}}$ *as a rational quotient of* $\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$ *by* $(\mathrm{PGL}_3)^m$. *Additionally,* $\mathrm{P}^{m,n}_{\bar{\mathbf{q}}',\mathbf{p}}$ *is cut out scheme-theoretically by the* $m\binom{n}{2}$ *bilinear equations* (2.26).

Proof. The statements involving rational quotients follow similarly as in Theorem 2.3.5. Under the isomorphism (2.28), the bilinear constraints in the theorem statement are the usual 2-focals vanishing on the multiview variety. We recall from part (3) of Theorem 2.2.3 that the 2-focals cut out the multiview variety scheme-theoretically whenever the camera centers are distinct and do not lie on a common plane. Now, since $\bar{q}_i$ is not in the span of any three E_j, the center of the camera $A(\bar{q}_i)$ is given by the Cremona transformation $\mathcal{C}(\bar{q}_i)$. Since $\mathcal{C}$ maps any plane in $\mathbf{P}^3$ to a 4-nodal cubic surface, and vice-versa, we are done. □

From the practitioner's point of view, the genericity assumptions of Theorem 2.3.6, as well as the implicit assumption that we can fix four fiducial 3D points and their images to the standard positions (2.20), may be quite reasonable. This is supported by the experiments of [THP19, §5], suggesting some potential uses of Carlsson-Weinshall duality in SfM settings.

2.4 *Proof of Theorem 2.2.6*

Our proof of Theorem 2.2.6 follows the general strategy used in the proof of [Aga+23, Theorem 3.2], but requires some nontrivial modifications.

Remark 2.4.1. Unlike triangulation, resectioning is an interesting problem even for $m = 1$ camera. In fact, most of the work needed to prove Theorem 2.2.6 involves the special case $m = 1$. As in the proof of Proposition 2.2.7, we fix $m = 1$ and write p_i in place of p_{1i}. We also write $I(\bullet)$ in place of $I_1(\bullet)$, and $\Gamma_{\bar{\mathbf{q}},\mathbf{p}}$ instead of $\Gamma^{1,n}_{\bar{\mathbf{q}},\mathbf{p}}$. Finally, let us recall the variety $\Gamma^{n,m}_{\bar{\mathbf{Q}},\mathbf{p}} \subset \left(\mathbf{P}^2\right)^{mn}$ introduced in Proposition 2.2.7. In place of $\Gamma^{n,1}_{\bar{\mathbf{Q}},\mathbf{p}}$, we simply write $\Gamma_{\bar{\mathbf{Q}},\mathbf{p}}$.

2.4.1 *Proof outline and preliminary facts*

To begin, we describe our proof strategy at a high level. The main steps of our proof can be understood via the diagram in Figure 2.4, with each of the steps (1)–(4) explained below.

(1) **Coordinate change to obtain a generic hypercamera arrangement from the structured one.** For $\bar{q}_1, \ldots, \bar{q}_n \in \mathbf{P}^3$ with no four coplanar, we establish in Lemmas 2.4.3 and 2.4.4 that there exist 3×3 invertible matrices $H_1, \ldots, H_n$ such that the trans-

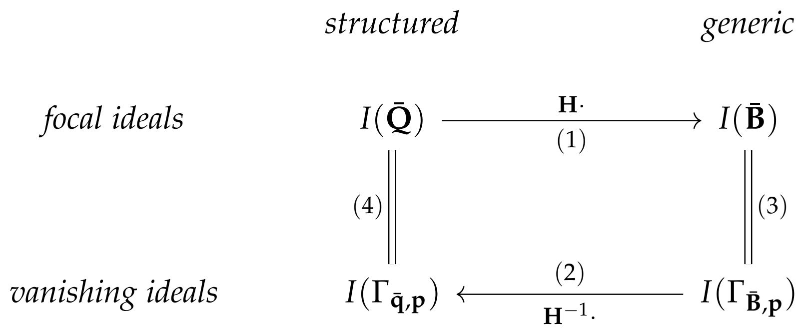


Figure 2.4: Schematic outline of the proof of Theorem 2.2.6.

formed hypercamera arrangement

$$\bar{\mathbf{B}} := (H_1\bar{Q}_1, \ldots, H_n\bar{Q}_n)$$

is *minor-generic* in the sense of Definition 2.4.2. Applying the coordinate change $\mathbf{H} = (H_1, \ldots, H_n)$ to $(\mathbf{P}^2)^n$ reduces the study of $I(\bar{\mathbf{q}})$ for the structured arrangement $\bar{\mathbf{Q}}$ to that of $I(\bar{\mathbf{B}})$ for the generic $\bar{\mathbf{B}}$.

(2) If we apply the inverse of the coordinate change $\mathbf{H}^{-1} = (H_1^{-1}, \ldots, H_n^{-1})$ from step (1) to $(\mathbf{P}^2)^n$, we can specialize from $I(\Gamma_{\bar{\mathbf{B}},\mathbf{p}})$ to $I(\Gamma_{\bar{\mathbf{q}},\mathbf{p}})$:

$$\bar{B}_j(A) = p_j \quad \Longleftrightarrow \quad H_j^{-1}\bar{B}_j(A) = H_j^{-1}p_j \quad \Longleftrightarrow \quad \bar{Q}_j(A) = H_j^{-1}p_j.$$

(3) **Show equality of focal and vanishing ideals in the generic case.** By the previous two steps, it is sufficient to show $I_{\text{foc}}(\bar{\mathbf{B}}) = I(\Gamma_{\bar{\mathbf{B}},\mathbf{p}})$. We establish this using Gröbner bases, as described in Section 2.4.2.

(4) **Show equality of focal and vanishing ideals in the structured case.** Combine steps (1)–(3).

For the first step in the proof outline, we need the following definition.

Definition 2.4.2 (minor-generic hypercamera arrangement)**.** We say the hypercamera arrangement $\bar{\mathbf{B}} = (\bar{B}_1, \ldots, \bar{B}_n) \in \left(\mathbf{P}^{35}\right)^n$ is *minor-generic* if all 12×12 minors of the $12 \times 3n$ matrix $\left(\bar{B}_1^\top \mid \cdots \mid \bar{B}_n^\top\right)$ are nonzero.

This is a direct analogue of the genericity condition in Theorem 2.2.3, part (1). We also need the following result. Let $\mathbf{F}$ be a field, and consider s matrices $A_1, \ldots, A_s \in \mathbf{F}^{M \times N}$.

We say $A_1, \dots, A_s$ are *rowspan-uniform* if, for any subset $S \subset [s]$ of size at least M/N, we have

$$\sum_{i \in S} \operatorname{rowspan}(A_i) = \mathbf{F}^N. \tag{2.30}$$

Lemma 2.4.3. *If $A_1, \dots, A_s \in \mathbf{F}^{M \times N}$ are rowspan-uniform, then there exists a dense Zariski-open set of matrices $(H_1, \dots, H_s) \in \mathrm{GL}(\mathbf{F}^M)^s$ such that the maximal $n \times n$ minors of the the $sM \times N$ matrix*

$$\begin{pmatrix} H_1 A_1 \\ \hline \vdots \\ \hline H_s A_s \end{pmatrix} \tag{2.31}$$

are all nonzero.

We leave the proof of this result to [CDL23, Appendix A]. This result is a direct generalization of [APT21, Lemma 3.6], in the setting of triangulation. In our setting of resectioning, we take $(M, N) = (3, 12)$, and deduce that we can transform the arrangement $\bar{\mathbf{Q}}$ for suitably generic $\bar{\mathbf{q}}$ to a minor-generic arrangement $\bar{\mathbf{B}}$ using the following result.

Lemma 2.4.4. *Suppose that $\bar{\mathbf{q}} \in (\mathbf{P}^3)^n$ is a point arrangement such that no four points are coplanar. Then $\bar{\mathbf{Q}}$ is rowspan-uniform.*

Proof. For any subset $S \subset [n]$ of size at least 4, we must show

$$\sum_{j \in S} \operatorname{rowspan}(\bar{Q}_j) = \mathbf{C}^{12}.$$

Noting the compatible direct-sum decompositions

$$\mathbf{C}^{12} \simeq \mathbf{C}^4 \oplus \mathbf{C}^4 \oplus \mathbf{C}^4,$$
$$\operatorname{rowspan}(\bar{Q}_j) \simeq \operatorname{rowspan}(\bar{q}_j^\top) \oplus \operatorname{rowspan}(\bar{q}_j^\top) \oplus \operatorname{rowspan}(\bar{q}_j^\top),$$

it suffices to observe that any set of four elements from the set $\{\bar{q}_j\}_{j \in S}$ span $\mathbf{P}^3$, from our assumption that such a set is noncoplanar. □

Lemma 2.4.4 gives us a geometric interpretation of when $\bar{\mathbf{Q}}$ is minor-generic. Algebraically, this condition is precisely what we need to obtain the Gröbner basis of Theorem 2.2.6 via a standard specialization argument. This is the focus of the next subsection.

2.4.2 Gröbner basis tools

To realize $I(\bar{\mathbf{q}})$ as the specialization of an ideal that is independent of $\bar{\mathbf{q}}$, we could replace the arrangement $\bar{\mathbf{Q}}$ with $\mathbf{B} = (B_1, \ldots, B_n)$, where

$$B_i = \begin{bmatrix} B_i[1,1] & B_i[1,2] & B_i[1,3] & B_i[1,4] & B_i[1,5] & B_i[1,6] & B_i[1,7] & B_i[1,8] & B_i[1,9] & B_i[1,10] & B_i[1,11] & B_i[1,12] \\ B_i[2,1] & B_i[2,2] & B_i[2,3] & B_i[2,4] & B_i[2,5] & B_i[2,6] & B_i[2,7] & B_i[2,8] & B_i[2,9] & B_i[2,10] & B_i[2,11] & B_i[2,12] \\ B_i[3,1] & B_i[3,2] & B_i[3,3] & B_i[3,4] & B_i[3,5] & B_i[3,6] & B_i[3,7] & B_i[3,8] & B_i[3,9] & B_i[3,10] & B_i[3,11] & B_i[3,12] \end{bmatrix}, \tag{2.32}$$

thereby introducing $36n$ new indeterminates. Alternatively, we could replace $\bar{\mathbf{Q}}$ with the symbolic arrangement $\mathbf{B}^\star = (B_1^\star, \ldots, B_n^\star)$, where

$$B_i^\star = \begin{bmatrix} B_i^\star[1,1] & B_i^\star[1,2] & B_i^\star[1,3] & B_i^\star[1,4] & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & B_i^\star[2,1] & B_i^\star[2,2] & B_i^\star[2,3] & B_i^\star[2,4] & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & B_i[3,1]^\star & B_i^\star[3,2] & B_i^\star[3,3] & B_i^\star[3,4] \end{bmatrix}, \tag{2.33}$$

for a total of $12n$ new indeterminates. For either $\mathbf{B}$ or $\mathbf{B}^\star$ and for each k with $6 \le k \le m$, we are interested in the determinants of the matrices

$$(\mathbf{B} \mid \mathbf{p})[\mathbf{r}]_\sigma = \begin{bmatrix} B_{\sigma_1}[\mathbf{r}_1, :] & p_{\sigma_1}[\mathbf{r}_1] & 0 & \ldots & 0 \\ B_{\sigma_2}[\mathbf{r}_2, :] & 0 & p_{\sigma_2}[\mathbf{r}_2] & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B_{\sigma_k}[\mathbf{r}_k, :] & 0 & \ldots & 0 & p_{\sigma_k}[\mathbf{r}_k] \end{bmatrix}, \tag{2.34}$$

$$(\mathbf{B}^\star \mid \mathbf{p})[\mathbf{r}]_\sigma = \begin{bmatrix} B^\star_{\sigma_1}[\mathbf{r}_1, :] & p_{\sigma_1}[\mathbf{r}_1] & 0 & \ldots & 0 \\ B^\star_{\sigma_2}[\mathbf{r}_2, :] & 0 & p_{\sigma_2}[\mathbf{r}_2] & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ B^\star_{\sigma_k}[\mathbf{r}_k, :] & 0 & \ldots & 0 & p_{\sigma_k}[\mathbf{r}_k] \end{bmatrix}, \tag{2.35}$$

where $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_n) \subset [3]^n$, with

$$\begin{aligned} \mathbf{r}_1, \ldots, \mathbf{r}_k &\neq \emptyset, \\ \#\mathbf{r}_1 + \cdots + \#\mathbf{r}_k &= 12+k, \\ \sigma = \{\sigma_1, \ldots, \sigma_k\} &\subset [n]. \end{aligned}$$

Upon specializing $\mathbf{B} \to \bar{\mathbf{Q}}$, or $\mathbf{B}^\star \to \bar{\mathbf{Q}}$, the respective determinants of (2.34) or (2.35) specialize to the same k-focal.

Proposition 2.4.5. *Equip either ring $\mathbf{C}[\mathbf{B}, \mathbf{p}]$ or $\mathbf{C}[\mathbf{B}^\star, \mathbf{p}]$ with the $\mathbf{Z}^{2n}$-grading defined on generators by $\deg(B_i[j,k]) = \deg(B_i^\star[j,k]) = e_i$ and $\deg(p_i[j]) = e_{n+i}$. The polynomial $\det(\mathbf{B} \mid \mathbf{p})[\mathbf{r}]_\sigma$ is homogeneous of multidegree*

$$\sum_{i=1}^{k} (\#\mathbf{r}_i - 1) e_{\sigma_i} + e_{n+\sigma_i}. \tag{2.36}$$

The same is true for $\det(\mathbf{B}^\star \mid \mathbf{p})[\mathbf{r}]_\sigma$, *provided it is nonzero. When* $k > 12$, *we have* $\mathbf{r}_j = \{l\}$ *for some* $j \in [k]$, $l \in [3]$, *and hence*

$$\begin{aligned} \det(\mathbf{B} \mid \mathbf{p})[\mathbf{r}]_\sigma &= p_{\sigma_j}[l] \cdot \det(\mathbf{B} \mid \mathbf{p})[\mathbf{r}_1, \ldots, \widehat{\mathbf{r}_j}, \ldots, \mathbf{r}_k]_{\sigma \setminus \{j\}}, \\ \det(\mathbf{B}^\star \mid \mathbf{p})[\mathbf{r}]_\sigma &= p_{\sigma_j}[l] \cdot \det(\mathbf{B}^\star \mid \mathbf{p})[\mathbf{r}_1, \ldots, \widehat{\mathbf{r}_j}, \ldots, \mathbf{r}_k]_{\sigma \setminus \{j\}}. \end{aligned} \tag{2.37}$$

Proof. The argument is nearly identical to [Aga+23, Proposition 3.3]. The multidegree formula (2.36) follows from a calculation analagous to that already given in (2.11). Since $12 = \sum_{i=1}^{k} (\#\mathbf{r}_i - 1)$, it follows that at most 12 of the sets $\mathbf{r}_i$ can contain more than one element. If some $\mathbf{r}_j$ is a singleton, the factorization (2.37) follows by Laplace expansion. □

We now define four auxiliary ideals.

Definition 2.4.6 (*Q*-focal ideals)**.** The ideals $I_{6..12}(\mathbf{B}), I_m(\mathbf{B}) \subset \mathbf{C}[\mathbf{B}, \mathbf{p}]$ are those which are generated by all determinants of (2.34) for all k, respectively, in the ranges $6 \le k \le 12$, and $k = m$. Similarly, $I_{6..12}(\mathbf{B}^\star), I_m(\mathbf{B}^\star) \subset \mathbf{C}[\mathbf{B}^\star, \mathbf{p}]$ are generated by all determinants of (2.35).

Each of the four auxiliary ideals in Definition 2.4.6 is useful for different reasons. For example, $I_m(\mathbf{B})$ and $I_m(\mathbf{B}^\star)$ are the ideals of maximal minors of a *sparse generic matrix* whose nonzero entries are distinct indeterminates. Thus, $I_m(\mathbf{B})$ and $I_m(\mathbf{B}^\star)$ belong to the class of *sparse determinantal ideals*, whose structure has been analyzed in several previous works [GM82; Boo12]. Most relevant to our work is the result of [Boo12] which directly implies that the m-focals form a universal Gröbner basis for either of these ideals.

For the other two ideals $I_{6..12}(\mathbf{B})$, $I_{6..12}(\mathbf{B}^\star)$, we do not know whether or not the focals form universal Gröbner bases. However, Proposition 2.4.7 shows that they *do* form Gröbner bases for a class of product orders that allows us to make the necessary specialization argument.

We recall that a *product order* on $\mathbf{C}[\mathbf{B}, \mathbf{p}]$ with $\mathbf{B} < \mathbf{p}$ is a monomial order defined by comparing monomials first with some fixed monomial order in $\mathbf{p}$, then breaking any ties with some other monomial order in $\mathbf{B}$.

Proposition 2.4.7.

1. *The set* $G = \{\det(\mathbf{B} \mid \mathbf{p})[\mathbf{r}]_\sigma \mid 6 \le \#\sigma \le 12\}$ *forms a Gröbner basis for the ideal* $I_{6..12}(\mathbf{B})$ *for any product order with* $\mathbf{B} < \mathbf{p}$.

2. *The set* $G^\star = \{\det(\mathbf{B}^\star \mid \mathbf{p})[\mathbf{r}]_\sigma \mid 6 \le \#\sigma \le 12\}$ *forms a Gröbner basis for the ideal* $I_{6..12}(\mathbf{B}^\star)$ *for any product order with* $\mathbf{B}^\star < \mathbf{p}$.

We provide a proof in [CDL23, Appendix A]. A specialization argument applied to the two parts of Proposition 2.4.7 gives, respectively, the two parts of Lemma 2.4.8 below.

Lemma 2.4.8. *For any monomial order on* $\mathbf{C}[\mathbf{p}]$, *the following hold:*

1. *Let* $\bar{\mathbf{B}}$ *be a specialization of* $\mathbf{B}$ *such that the* $\bar{\mathbf{B}}$ *is minor generic. Then the specialized* $6-12$ *focals form a Gröbner basis for the ideal they generate.*

2. *Let* $\bar{\mathbf{Q}}$ *be a specialization of* $\mathbf{B}^\star$ *derived from a point arrangement* $\bar{\mathbf{q}} \in (\mathbf{P}^3)$ *with no four points coplanar. Then the specialized* $6-12$ *focals form a Gröbner basis.*

Proof. For part (1), let $<$ be any product order with $\mathbf{B} < \mathbf{p}$. Then

$$\operatorname{in}_< \left(\sum_{p^{\alpha_1} < \ldots < p^{\alpha_k}} g_{\alpha_i}(\mathbf{B}) \mathbf{p}^{\alpha_i} \right) = \operatorname{in}_<(g_{\alpha_k}(\mathbf{B})) p^{\alpha_k}. \tag{2.38}$$

Standard specialization results for Gröbner bases with respect to product orders [CLO15, Theorem 2, §4.7] imply that the $\bar{\mathbf{B}}$ specialized $6-12$ focals form a Gröbner basis if each coefficient $g_{\alpha_k}(\bar{\mathbf{B}})$ is nonzero. Each of these coefficients is a 12×12 minor of $(\bar{B}_1^\top \mid \cdots \mid \bar{B}_n^\top)$. By minor-genericity, none of these coefficients vanish.

Similarly, for part (2), consider any product order $<$ with $\mathbf{B}^* < \mathbf{p}$. In this case, the nonzero coefficients $g_{\alpha_k}(\mathbf{B}^\star)$ are always products of three 4×4 determinants,

$$\prod_{i=1}^{3} \det \begin{bmatrix} B^*_{j_{i,1}}[i,1] & B^*_{j_{i,1}}[i,2] & B^*_{j_{i,1}}[i,3] & B^*_{j_{i,1}}[i,4] \\ B^*_{j_{i,2}}[i,1] & B^*_{j_{i,2}}[i,2] & B^*_{j_{i,2}}[i,3] & B^*_{j_{i,2}}[i,4] \\ B^*_{j_{i,3}}[i,1] & B^*_{j_{i,3}}[i,2] & B^*_{j_{i,3}}[i,3] & B^*_{j_{i,3}}[i,4] \\ B^*_{j_{i,4}}[i,1] & B^*_{j_{i,4}}[i,2] & B^*_{j_{i,4}}[i,3] & B^*_{j_{i,4}}[i,4] \end{bmatrix}. \tag{2.39}$$

Our noncoplanarity assumption implies that the $\mathbf{B}^\star \to \bar{\mathbf{Q}}$ specialization of (2.39) is nonzero. □

Finally, we have the following result on the vanishing ideal of $\Gamma_{\bar{\mathbf{B}},\mathbf{p}}$ when $\bar{\mathbf{B}}$ is a minor-generic hypercamera arrangement. See [CDL23, Appendix A] for the proof.

Proposition 2.4.9. *For a minor-generic hypercamera arrangement* $\bar{\mathbf{B}}$, *we have that*

$$I(\Gamma_{\bar{\mathbf{B}},\mathbf{p}}) = I(\bar{\mathbf{B}}).$$

2.4.3 *Completing the proof*

Using the results of the previous sections, we may complete the proof of Theorem 2.2.6, following the overall structure presented in Figure 2.4.

We first prove the statement for $m = 1$ camera. Let $\bar{\mathbf{q}} \in (\mathbf{P}^3)^m$ be a point arrangement with no four points coplanar. Lemma 2.4.4 then implies that the hypercamera arrangement $\bar{\mathbf{Q}}$ is rowspan-uniform, and thus Lemma 2.4.3 implies there exist coordinate changes in the images, $\mathbf{H} = (H_1, \ldots, H_m) \in (\mathrm{PGL}_3)^m$, such that the arrangement $\bar{\mathbf{B}} = (H_1\bar{Q}_1, \ldots, H_m\bar{Q}_m)$ is minor-generic. Noting

$$\begin{pmatrix} \bar{Q}_1 & p_1 & & \\ \vdots & & \ddots & \\ \bar{Q}_n & & & p_n \end{pmatrix} = \begin{pmatrix} H_1^{-1} & & \\ & \ddots & \\ & & H_n^{-1} \end{pmatrix} \begin{pmatrix} \bar{B}_1 & H_1 p_1 & & \\ \vdots & & \ddots & \\ \bar{B}_n & & & H_n p_n \end{pmatrix}, \tag{2.40}$$

we define the isomorphism of multigraded rings

$$\begin{aligned} L_{\mathbf{H}} : \mathbf{C}[\mathbf{p}] &\to \mathbf{C}[\mathbf{p}] \\ p_i &\to H_i p_i, \end{aligned}$$

and observe that

$$I(\bar{\mathbf{q}}) = L_{\mathbf{H}}(I(\bar{\mathbf{B}})).$$

To see this, take any focal $f \in I(\bar{\mathbf{q}})$. Just as in the proof of Lemma 2.4.3, corresponding minor of the focal matrix on the left of Equation (2.40) may written as a $\mathbf{C}$-linear combination of focals for the arrangement $\bar{\mathbf{B}}$. Hence the inclusion $I(\bar{\mathbf{q}}) \subset L_{\mathbf{H}}(I(\bar{\mathbf{B}}))$ holds, and the reverse follows similarly. Thus, we have

$$\begin{aligned} I(\bar{\mathbf{q}}) &= L_{\mathbf{H}}(I(\bar{\mathbf{B}})) \\ &= L_{\mathbf{H}}(I(\Gamma_{\bar{\mathbf{B}},\mathbf{p}})) && \text{(Proposition 2.4.9)} \\ &= I(\Gamma_{\bar{\mathbf{Q}},\mathbf{p}}) \\ &= I(\Gamma_{\bar{\mathbf{q}},\mathbf{p}}). && \text{(Proposition 2.2.7)} \end{aligned}$$

Thus the focals generate $I(\Gamma_{\bar{\mathbf{q}},\mathbf{p}})$. Moreover, Lemma 2.4.8 part (2) implies that they form a universal Gröbner basis, which completes the proof when $m = 1$.

Finally, if $m > 1$, it suffices to observe that $\Gamma^{m,n}_{\bar{\mathbf{q}},\mathbf{p}}$ is the direct product of varieties $\Gamma_{\bar{\mathbf{q}},\mathbf{p}}$, and hence the vanishing ideals sum. Moreover, two k-focals corresponding to different factors have disjoint support in $\mathbf{C}[\mathbf{p}]$, so their S-polynomials reduce to zero for any term order, and we may conclude that the focals form a universal Gröbner basis for any number of cameras.

2.5 *Optimal single-camera resectioning*

The results of Section 2.3 express a duality principle for the exact versions of the camera resectioning and triangulation problems. A consequence of this duality is that, in a certain sense, resectioning and triangulation are equivalent problems. However, we should be mindful that this equivalence holds in an idealized setting which assumes that the pinhole camera is exact and there is no measurement noise. In practice, neither of these assumptions hold.

In this section, we fix a generic point arrangement $\bar{\mathbf{q}}$ and consider $\Gamma^{1,n}_{\bar{\mathbf{q}},\mathbf{p}} \subset (\mathbf{P}^2)^n$ intersected with the affine chart where $p_i[3] \neq 0$ for all $1 \leq i \leq n$. We denote this affine variety by $X_{\bar{\mathbf{q}},n}$. In other words, for a point arrangement $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^n$ such that no four points are coplanar, the affine variety $X_{\bar{\mathbf{q}},n}$ is, by Proposition 2.2.7, equal to the closed image of the rational map

$$\begin{aligned}\psi_{\bar{\mathbf{q}},n} &: \mathbf{P}^{11} \dashrightarrow \mathbf{C}^{2n} \\ A &\mapsto \left(\frac{A[1,:]\bar{q}_1}{A[3,:]\bar{q}_1}, \frac{A[2,:]\bar{q}_1}{A[3,:]\bar{q}_1}, \ldots, \frac{A[1,:]\bar{q}_n}{A[3,:]\bar{q}_n}, \frac{A[2,:]\bar{q}_n}{A[3,:]\bar{q}_n}\right). \end{aligned} \tag{2.41}$$

In the resectioning problem, we are given world points $\bar{\mathbf{q}} = (\bar{q}_1, \ldots, \bar{q}_n) \in (\mathbf{P}^3)^n$ and pixel values of n corresponding image points, $(\tilde{u}_i, \tilde{v}_i)$ for $i = 1, \ldots, n$. We denote the vector of image measurement data by $\tilde{d}_{uv} = (\tilde{u}_1, \ldots, \tilde{v}_n) \in \mathbf{C}^{2n}$. In practice, $\tilde{d}_{uv}$ and $\bar{\mathbf{q}}$ are both defined over the real numbers. Our task is to recover a camera A such that

$$\psi_{\bar{\mathbf{q}},n}(A) = \tilde{d}_{uv}. \tag{2.42}$$

In an idealized setting, the pinhole model is exact and there is no measurement noise. Hence, we can recover A by computing the kernel of the n-focal matrix, and we expect a unique solution as soon as $n \geq 6$. This is the basis of the so-called “5.5-point" minimal solver.

In practice, the pinhole model is *not* exact and there *is* measurement noise. Thus, for $n \geq 6$, we should expect $\tilde{d}_{uv} \notin X_{\bar{\mathbf{q}},n}$, meaning that no solution to (2.42) can exist. However, we can still consider the following optimization problem:

$$L_{\tilde{d}_{uv}}(u_1, v_1, \ldots, u_n, v_n) = \sum_{i=1}^{n} (u_i - \tilde{u}_i)^2 + (v_i - \tilde{v}_i)^2 \quad \text{s.t.} \quad (u_1, \ldots, v_n) \in X_{\bar{\mathbf{q}},n}. \tag{2.43}$$

This is essentially the formulation of the optimal resectioning problem that is used in Hartley and Zisserman’s classic text [HZ04][§7.2]. The only minor difference, implicit in their formulation, is that our feasible set $X_{\bar{\mathbf{q}},n}$ differs from theirs by a set of measure zero picked up through Zariski closures. Similar formulations, which make the camera matrix explicit, appear in other sources, eg. in work of Cifuentes [Cif21, Example 6.5] who studied sums-of-squares relaxations of this problem. Hartley and Zisserman refer to the squared Euclidean loss function $L_{\tilde{d}_{uv}}$ as the *geometric error*, and suggest using local methods like Levenberg-Marquardt to optimize it. Here, we address the complexity of computing the *global minimum* of (2.43).

We recall the notion of the *Euclidean distance degree* of an affine variety, [Dra+16, §2]. For $X_{\bar{\mathbf{q}},n}$, we denote this quantity by $\mathrm{ED}(X_{\bar{\mathbf{q}},n})$. Given a generic data point $\tilde{d}_{uv}$, this is the number of critical points of the squared Euclidean loss $L_{\tilde{d}_{uv}}$ restricted to the smooth locus of $X_{\bar{\mathbf{q}},n}$.

Conjecture 2.5.1. *For all $n \geq 6$ and generic $\bar{\mathbf{q}} \in (\mathbf{P}^3)^n$, we have*

$$\mathrm{ED}(X_{\bar{\mathbf{q}},n}) = (80/3)n^3 - 368n^2 + (5068/3)n - 2580. \tag{2.44}$$

We return to Example 2.2.8, to verify the simplest case of this conjecture.

Example 2.5.2. Consider the resectioning hypersurface $H(u_1, \ldots, v_6) = 0$, ie. (2.13) in the chart

$$p_1[3] = p_2[3] = p_3[3] = p_4[3] = p_5[3] = p_6[3] = 1. \tag{2.45}$$

The affine variety $X_{\bar{\mathbf{q}},6} \subset \mathbf{C}^{12}$ is in fact the cone over a projective variety in $\mathbf{P}^{11}$. This can be seen from the determinantal representation of H in (2.15). It follows that $X_{\bar{\mathbf{q}},n}$ is singular. More precisely, the singular locus of $X_{\bar{\mathbf{q}},n}$ has dimension 9.

Working over the finite field $\mathbf{F} = \mathbf{Z}_{32003}$, we may verify Conjecture 2.5.1 with symbolic computation using the computer algebra system Macaulay2 [GS]. To do so, we draw a $\mathbf{F}$-valued point configuration $\bar{\mathbf{q}} \in \left(\mathbf{P}^3\right)^6$ and data vector $\tilde{d}_{uv} \in \mathbf{F}^{12}$ uniformly at random. The critical points of (2.43) correspond to points $(u_1, \ldots, v_6) \in X_{\bar{\mathbf{q}},6}$ such that

$$\operatorname{rank} \begin{bmatrix} u_1 - \tilde{u}_1 & \cdots & v_6 - \tilde{v}_6 \\ \frac{\partial H}{\partial u_1} & \cdots & \frac{\partial H}{\partial v_6} \end{bmatrix} \leq 1. \tag{2.46}$$

To remove the singular points on $X_{\bar{\mathbf{q}},6}$ which cause rank-deficiency in (2.46), it is sufficient take the ideal generated by the 2×2 minors of this matrix and $H(u_1, \ldots, v_6)$ and compute its ideal quotient with respect to the ideal $\langle \frac{\partial H}{\partial u_1}, \frac{\partial H}{\partial v_1} \rangle$. The result of this operation is a zero-dimensional ideal of degree 68. Moreover, we may compute that the vanishing locus of this ideal consists of 68 distinct, nonsingular points on $X_{\bar{\mathbf{q}},6}$. The number 68 may be seen as quantifying the intrinsic algebraic difficulty of solving the constrained optimization problem (2.43). This is further reinforced by heuristically computing the Galois/monodromy group of this problem, as in [Duf+22], which reveals the full symmetric group S_{68}.

Our conjectural formula (2.44) is reminiscent of recent results characterizing the Euclidean distance degree of the *affine multiview variety* $X_{\bar{\mathbf{A}},m}$. This can be defined by taking analagous affine charts on the multiview variety $\Gamma^{m,1}_{\bar{\mathbf{A}},\mathbf{p}}$. Using a topological formula for the ED-degree of a smooth variety, Maxim, Rodriguez, and Wang [MRW20] proved

$$\mathrm{ED}(X_{\bar{\mathbf{A}},m}) = (9/2)m^3 - (21/2)m^2 + 8m - 4. \tag{2.47}$$

We compare this formula with ours in Table 2.1. We confirmed the entries of this table using numerical monodromy heuristics [Duf+19a], using both the implementations provided in Macaulay2 [GS] and Julia [BT18]. For these computations, it is advantageous to use the rational parametrization (2.41) instead of the implicit focal constraints in Theorem 2.2.6.

A surprising aspect of Conjecture 2.5.1 is that $\mathrm{ED}(X_{\bar{\mathbf{q}},n})$ is a polynomial of degree 3 in n. On the other hand, if we were to apply the methods of [MRW20] to computing the affine ED-degree of the variety $\Gamma^{n,1}_{\bar{\mathbf{B}},\mathbf{q}}$ associated to a *generic* hypercamera arrangement $\bar{\mathbf{B}} \in \left(\mathbf{P}^{11}\right)$, this would give instead a polynomial of degree 11. This highlights some special properties of the hypercamera arrangement $\bar{\mathbf{Q}}$, and provides contrast with the results of previous sections. One explanation for this contrast is the fact that the projective coordinate changes used in Theorems 2.3.5 and 2.3.6 do not preserve the Euclidean distance. For similar reasons, the affine ED degree of the reduced resectioning variety, which is the same as $\mathrm{ED}(X_{\bar{\mathbf{A}},m})$, appears to be unrelated to that of the general resectioning variety.

We close this section by noting one immediate obstacle to proving Conjecture 2.5.1. As already seen in Example 2.5.2, the variety $X_{\bar{\mathbf{q}},n}$ for generic data $\bar{\mathbf{q}}$ is *not smooth* for any $n \geq 6$. This contrasts with the case of $X_{\bar{\mathbf{A}},m}$, which is smooth for a sufficiently generic arrangement of $m \geq 3$ cameras $\bar{\mathbf{A}}$. Thus, to prove (2.47) with similar techniques, the basic Euler characteristic formulas valid in the smooth case would need to be replaced by their singular counterparts involving Euler obstruction functions, eg. [MRW21, Theorem 1.3].

2.6 Conclusion

In summary, our work takes several first steps in studying the resectioning problem for general projective cameras from the algebro-geometric perspective, with a focus on

m / n	$\mathrm{ED}(X_{\bar{\mathbf{A}},m})$	$\mathrm{ED}(X_{\bar{\mathbf{q}},n})$
2	6	—
3	47	—
4	148	—
5	336	—
6	638	68
7	1081	360
8	1692	1036
9	2498	2256
10	3526	4180
11	4803	6968
12	6356	10780
13	8212	15776
14	10398	22116
15	12941	29960

Table 2.1: Euclidean distance degrees for optimal triangulation from m generic cameras (middle column) and optimal resectioning from n 3D points (right.)

Gröbner bases, Carlsson-Weinshall duality, and Euclidean distance optimization. Our discoveries provide many parallels with the already well-studied multiview ideals associated with the triangulation problem. Still, many open questions remain.

In this paper, we considered resectioning in the setting of general projective cameras. Returning to the classical P3P problem [Gru41], it would be worthwhile to carry out a parallel study in the setting of *Euclidean cameras*, as proposed in [Aga+23, §8.3, Q1]. In view of Theorem 2.2.3 parts (2)–(3), it is natural to ask: are all k-focals for $6 \leq k \leq 12$ are needed to generate $I_m(\bar{\mathbf{q}})$ under the noncoplanarity assumption of Theorem 2.2.6? What can we say about $I_m(\bar{\mathbf{q}})$ if this noncoplanarity assumption is relaxed? Using the reduced atlas developed Section 2.3 to answer more of the open questions in [Aga+23, §8] is yet another interesting avenue to pursue. Our focus on resectioning for linear maps $\mathbf{P}^3 \dashrightarrow \mathbf{P}^2$ was motivated by computer vision. However, it would make just as much sense to study resectioning varieties in the context of general projections $\mathbf{P}^N \dashrightarrow \mathbf{P}^M$, [Li18], or even matrix multiplication maps as in [Aga+23, §8.3, Q3]. Finally, we offer Conjecture 2.5.1 as a challenge in Euclidean distance degree computation.

Chapter 3

TOWARDS A COMPACTIFICATION OF THE MODULI OF CAMERA PAIRS

In this chapter we consider possible compactifications of Cam_2, the moduli space of camera pairs.

Definition 3.0.1 (The moduli of camera pairs)**.** The functor of camera pairs, denoted Cam_2, has value over a scheme S the set of isomorphism classes of (not necessarily general) relative multiview configurations of length 2.

The moduli space Cam_2 contains a well-studied open locus, denoted Cam_2° (denoted by Cam_2 in [LV20]), which comprises of pairs $\mathbf{A}$ whose centers are distinct.

Viewed as the quotient stack $[(U \times U)/\operatorname{PGL}_4]$ (where U is the rank-3 locus inside $\mathrm{M}_{3\times 4}$), Cam_2 is not even Deligne-Mumford. Indeed, the automorphism group of any concentric pair is five-dimensional. In this chapter we work towards a modular representation of camera pairs containing Cam_2° whose boundary points correspond to concentric cameras, but whose geometry is "nicer" (i.e., more reduced) than that of Cam_2 with its current definition.

3.1 The fundamental variety

It is well-known (e.g., [HZ03]) that a nonconcentric pair of cameras $\mathbf{A} = (A_1, A_2)$ is entirely determined by a corresponding 3×3 matrix called the *fundamental matrix* of $\mathbf{A}$ which is derived from the data given by A_1 and A_2. In turn, the space of fundamental matrices corresponds to the rank-2 locus of $\mathbf{P}(M_{3\times 3})$, denoted $\mathbf{F}^\circ$. Taking the Zariski closure of $\mathbf{F}^\circ$ in $\mathbf{P}(M_{3\times 3})$, we get the *fundamental variety* $\mathbf{F}$ which is actually the locus $(\det M = 0) \subset \mathbf{P}(M_{3\times 3})$. However, the boundary points of $\mathbf{F}$ (that is, points in $\mathbf{F} \setminus \mathbf{F}^\circ$) do not coincide with concentric camera pairs in $\mathrm{Cam}_2 \setminus \mathrm{Cam}_2^\circ$. It is of interest, then, to find a geometrically intuitive presentation of Cam_2 by a scheme or perhaps even a (nice) stack, whose boundary points are also in correspondence with the boundary points of Cam_2.

3.1.1 Preliminaries

Recall that a length-n camera configuration $\mathbf{A}$ is modeled by an n-tuple of 3×4 matrices $A_i = [R_i \mid \mathbf{t_i}]$, which form an n-linear rational map $\mathbf{A} : \mathbf{P}^3 \dashrightarrow (\mathbf{P}^2)^n$ (where each $A_i = \pi_i \circ \mathbf{A}$ is a linear projection).

Lemma 3.1.1. *A pair of cameras* $([R_1 \mid \mathbf{t_1}], [R_2 \mid \mathbf{t_2}]) \in \mathrm{Cam}_2$ *can be represented by a pair of the form* $([I \mid \mathbf{0}], [R \mid \mathbf{t}])$.

Proof. If R_1 is not invertible, multiply $[R_1 \mid \mathbf{t_1}]$ by the elementary matrix that swaps the dependent column of R_1 with $\mathbf{t_1}$ (this is always possible because $[R_1 \mid \mathbf{t_1}]$ is full-rank). Thus we can assume that R_1^{-1} is invertible. Multiply both cameras by $\begin{bmatrix} R_1^{-1} & -R_1^{-1}\mathbf{t_1} \\ 0 & 1 \end{bmatrix} \in \mathrm{PGL}_4$ to get $[I \mid \mathbf{0}]$ in the first factor, $[R \mid \mathbf{t}]$ in the second. □

Proposition 3.1.2 ([HZ03])**.** *Let* $\mathbf{A} = (A_1, A_2)$ *be a pair of pinhole cameras with distinct centers. Then there exists a unique (up to isomorphism class of camera pairs) rank-2 matrix F such that F satisfies*

$$p_2^\intercal\, F\, p_1 = 0 \quad \iff \quad A_1(q) = p_1, \quad A_2(q) = p_2 \quad \textit{for some } q \in \mathbf{P}^3. \tag{3.1}$$

Proof. Omitted. □

Definition 3.1.3 (Fundamental matrix)**.** Let $\mathbf{A} = ([I \mid \mathbf{0}], [R \mid \mathbf{t}])$ be a camera pair in Cam_2 with distinct centers (i.e., $\mathbf{t} \neq \mathbf{0}$). The rank-2 matrix F from Proposition 3.1.2 corresponding to $\mathbf{A}$ is called the *fundamental matrix* of $\mathbf{A}$.

Corollary 3.1.4. *The fundamental matrix of a non-concentric camera pair* $([I \mid \mathbf{0}], [R \mid \mathbf{t}])$ *is of the form* $[\mathbf{t}]_\times R$*, where*

$$[\mathbf{t}]_\times := \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}. \tag{3.2}$$

Proof. One can check directly that $[\mathbf{t}]_\times R$ fulfills the equation $p_2^\intercal([\mathbf{t}]_\times R)p_1 = 0$ if and only if $[I \mid \mathbf{0}]q = p_1, [R \mid \mathbf{t}]q = p_2$. □

Definition 3.1.5 (Fundamental variety)**.** The *fundamental variety* $\mathbf{F}$ is the Zariski closure of the image $\mathbf{F}^\circ$ of the fundamental matrix map

$$\begin{aligned} \varphi : \mathrm{Cam}_2 &\dashrightarrow \mathbf{P}(M_{3\times 3}) \\ ([I \mid \mathbf{0}], [R \mid \mathbf{t}]) &\longmapsto [\mathbf{t}]_\times R. \end{aligned} \tag{3.3}$$

Proposition 3.1.6. *The fundamental matrix map φ of eq.* (3.3) *is regular exactly on* Cam_2°*, and moreover φ induces an isomorphism*

$$\mathrm{Cam}_2^\circ \xrightarrow{\sim} \mathbf{F}^\circ. \tag{3.4}$$

That is, a pair of non-concentric cameras is defined by its fundamental matrix, and conversely a fundamental matrix allows us to recover the (isomorphism class of the) pair of cameras.

Proof. Omitted. See, e.g., [HD00; LV20]. □

3.2 *Attempts at compactification*

One can represent Cam_2 as the quotient stack $[(U \times U)/\,\mathrm{PGL}_4]$, where $U \subset \mathbf{P}(M_{3\times 4})$ is the rank-3 locus. However, the unstable points of this quotient—points in $U \times U$ with nontrivial PGL_4 stabilizers—are exactly the matrix pairs with coincident kernels.

Lemma 3.1.1 tells us that a point $([R_1 \mid \mathbf{t_1}], [R_2 \mid \mathbf{t_2}])$ in this quotient can always be represented by the pair $([I \mid \mathbf{0}], [R \mid \mathbf{t}])$, so it suffices to look at the quotient space $[U/\mathbf{G}]$, where $\mathbf{G} < \mathrm{PGL}_4$ is the subgroup fixing $[I \mid \mathbf{0}]$. To begin the compactification process, it is natural to look for a resolution of the rational map φ obtained from assigning to a camera pair its fundamental matrix.

3.2.1 *Blow-ups, vector bundles, and tangent spaces*

In the quotient stack representation $[U/\mathbf{G}]$ of Cam_2, we see that Cam_2° corresponds to $[U^\circ/G]$ where

$$U^\circ = \{[R \mid \mathbf{t}] \in U : \mathbf{t} \neq \mathbf{0}\}. \tag{3.5}$$

It is easily computed that the points $[R \mid \mathbf{t}] \in U$ with non-trivial $\mathbf{G}$-stabilizers are exactly those with $\mathbf{t} = \mathbf{0}$. The stabilizer of such points is G itself.

Note 3.2.1. The map $\varphi : [R \mid \mathbf{t}] \mapsto [\mathbf{t}]_\times R$ reflects the fact that there is no notion of a fundamental matrix for concentric cameras. We see that φ depends on t being nonzero: If $t = 0$, then $[\mathbf{t}]_\times R = 0$ as well, which is not defined in $\mathbf{P}^8$ (and thus not in $\mathbf{F}$). So then the only information we have for concentric cameras Cam_2 is the homography $R : \mathbf{P}^2 \to \mathbf{P}^2$ which transforms the first image plane into the second. Since concentric cameras come up in the real world it would be useful to recover more information about the point correspondences in this case.

Example 3.2.2. One real-world example of a concentric camera situation is the construction of a panoramic photograph. These are often pieced together from multiple images taken from the same camera rotated around an axis (e.g. a tripod).

Figure 3.1: A panorama of the interior of St. Peter's Basilica, made in Google Photos from several consecutive photographs.

A seemingly natural way to resolve this indeterminacy along $Z = V(\mathbf{t})$ would be to look at the blow up of Cam_2 at Z. Blowing up will give a generic extension $\tilde{\varphi} : \widetilde{\mathrm{Cam}_2} \to \mathbf{F}$, where $\widetilde{\mathrm{Cam}_2} = \mathrm{Bl}_Z \mathrm{Cam}_2 \subset \mathrm{Cam}_2 \times \mathbf{P}^2$.

Note 3.2.3. The indeterminacy locus of φ and the non-separated locus of Cam_2 coincide here. One hope is to resolve φ in a way that also shrinks the non-separated locus. A guaranteed benefit of blowing up here is the resolution of φ, but we will see that the non-separated locus of Cam_2 does not get any smaller when we pass to $\widetilde{\mathrm{Cam}_2}$.

We will compute explicitly the blow-up $\widetilde{\mathrm{Cam}_2}$ and resolution $\tilde{\varphi}$ which give the diagram

$$\begin{array}{ccc} \widetilde{\mathrm{Cam}_2} & & \\ \downarrow \pi & \searrow \tilde{\varphi} & \\ \mathrm{Cam}_2 & \overset{\varphi}{\dashrightarrow} & \mathbf{F} \subset \mathbf{P}^8. \end{array} \tag{3.6}$$

The blow-up $\widetilde{\mathrm{Cam}_2}$ is given by the equations

$$T_1t_2 = T_2t_1, \qquad T_1t_3 = T_3t_1, \qquad T_2t_3 = T_3t_2. \tag{3.7}$$

We can directly compute the fiber of a point $[R \mid \mathbf{t}] \in \mathrm{Cam}_2$ away from Z, assuming $t_1 = 1$. Any point lying above $[R \mid \mathbf{t}]$ will be of the form $([R \mid \mathbf{t}], (T_1 : T_2 : T_3))$, but 3.7 requires that, when $T_1 \neq 0$,

$$(T_1 : T_2 : T_3) = (T_1 : T_1t_2 : T_1t_3) = (1 : t_2 : t_3) = t. \tag{3.8}$$

This can be checked for $t_2 = 1$ and $t_3 = 1$ as well. Patching everything together we see that $\widetilde{\mathrm{Cam}_2}$ is isomorphic to Cam_2 away from Z.

The fiber $\pi^{-1}([R \mid \mathbf{0}])$ of a point in Z, however, allows for all $T \in \mathbf{P}^2$. This means that a point in $\pi^{-1}([R \mid \mathbf{0}])$ is $[R \mid \mathbf{0}]$ along with some $T \in \mathbf{P}^2$. The tangent vector T is describing infinitesimal motion of the second camera. We can think of T as the direction from which the second camera "came" to become concentric with the first. Allowing tangent vectors on the boundary points of Cam_2 thus gives us more information about the relationship between the two cameras.

We can compute a resolution of φ on the affine patches of $\mathbf{P}^2_T$. If $T_2 = 1$, then we have

$$t_1 = T_1t_2 \qquad \text{and} \qquad t_3 = T_3t_2 \tag{3.9}$$

so, generically in $\mathbf{P}^8$,

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -T_3t_2 & t_2 \\ T_3t_2 & 0 & -T_1t_2 \\ -t_2 & T_1t_2 & 0 \end{bmatrix} = t_2 \cdot \begin{bmatrix} 0 & -T_3 & 1 \\ T_3 & 0 & -T_1 \\ 1 & T_1 & 0 \end{bmatrix} = t_2 \cdot [T]_\times = [T]_\times. \tag{3.10}$$

This can be checked on the patches $T_i = 1$ for $i = 2, 3$ as well, so the resolution of φ is

$$\begin{aligned} \tilde{\varphi} : \widetilde{\mathrm{Cam}_2} &\to F \\ ([R \mid \mathbf{t}], T) &\mapsto [T]_\times R \end{aligned} \tag{3.11}$$

As demonstrated, the resolution $\tilde{\varphi}$ gives rise to a "fundamental matrix" for concentric cameras: A point $([R \mid \mathbf{0}], T)$ in the exceptional fiber $E = \pi^{-1}(Z)$ now has image $[T]_\times R \in \mathbf{F}^\circ$. Clearly this is no longer an isomorphism of spaces, but now we have a map that gives us a fundamental matrix even for the boundary points of Cam_2.

Unfortunately this is about as good as it gets. The issue we started with is that the non-separated part of Cam_2 includes all concentric cameras. However, the action of G on $\mathrm{Bl}_Z U$ does not take T into account at all:

$$([R \mid \mathbf{t}], T) \cdot M = ([R \mid \mathbf{t}] \cdot M, T) \tag{3.12}$$

since otherwise the identity matrix would not fix $([R \mid \mathbf{0}], T)$, nor would it be an extension of the action of G on U. So the points in $\mathrm{Bl}_Z U$ with non-trivial stabilizer are still those with $t = 0$, and the stabilizer is still the entirety of G. Thus the non-separated locus has not shrunk; in fact, now it is the entire exceptional divisor.

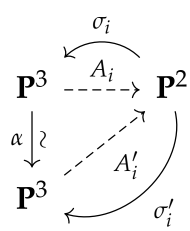
3.2.2 *Reformulating* Cam_2

In the 3×4 matrix case, we are able to ignore the physical image planes completely. However, bringing them into the picture may provide insight into an appropriate compactification of the moduli of nonconcentric camera pairs.

Consider the reformulation of a camera pair (or more generally, a length-n camera configuration), where the restriction σ to the i-th $\mathbf{P}^2$ factor is a section σ_i of the camera A_i.

$$\mathbf{P}^3 \overset{\mathbf{A}}{\dashrightarrow} \mathbf{P}^2 \times \mathbf{P}^2 \overset{\sigma}{\hookrightarrow} \mathbf{P}^3 \times \mathbf{P}^3 \tag{3.13}$$

These are considered up to automorphism $\alpha \in \mathrm{PGL}_4$ such that the following diagram commutes for $i = 1, 2$:



(3.14)

Requiring the additional data of the σ_i assigns image planes $P_1 \times P_2 = \mathrm{im}(\sigma) \subset \mathbf{P}^3 \times \mathbf{P}^3$.

In linear algebra terms, we can view σ_i as the right psuedo-inverse of A_i. Also, with this definition of isomorphism classes, the 4×4 composition $\sigma \circ \mathbf{A}$ is unique up to conjugation by PGL_4. Let $\mathbf{A} = (A_1, A_2) = ([I \mid \mathbf{0}], [R \mid \mathbf{t}])$ be a camera pair, with sections $\sigma = (\sigma_1, \sigma_2) = ([I \mid \mathbf{u}]^\intercal, [S \mid \mathbf{v}]^\intercal)$. For simplicity, we assume that R, S are invertible. Then the 4×4 camera-plane pairs are

$$\sigma_1 A_1 = \begin{pmatrix} I & \mathbf{0} \\ \mathbf{u}^\intercal & 0 \end{pmatrix}, \qquad \sigma_2 A_2 = \begin{pmatrix} S^\intercal R & S^\intercal \mathbf{t} \\ (S^{-1}\mathbf{v})^\intercal S^\intercal R & (S^{-1}\mathbf{v})^\intercal S^\intercal \mathbf{t} \end{pmatrix}. \tag{3.15}$$

Without loss of generality, we can assume that $\mathbf{u} = \mathbf{0}$ because $\sigma_1 A_1$ is diagonalizable with characteristic polynomial $\chi(\sigma_1 A_1) = \lambda(1-\lambda)^3$. This gives us a presentation of the pair $(\mathbf{A}, \sigma)$ by

$$\mathbf{A} = ([I \mid \mathbf{0}], [R \mid \mathbf{t}]), \qquad \sigma = ([I \mid \mathbf{0}]^\intercal, [S \mid \mathbf{v}]^\intercal). \tag{3.16}$$

Letting $\mathbf{q} = (q_0 : q_1 : q_2 : q_3)$, the image planes are then given by

$$P_1 : \mathbf{q}^\intercal \begin{pmatrix} \mathbf{0} \\ -1 \end{pmatrix} = 0, \qquad P_2 : \mathbf{q}^\intercal \begin{pmatrix} S^{-1}\mathbf{v} \\ -1 \end{pmatrix} = 0. \tag{3.17}$$

Using this formulation of a pinhole camera pair, one can then apply point-plane duality to get the "dual" pair

$$(\Sigma, a) = (([I \mid \mathbf{0}], [S \mid \mathbf{v}]), ([I \mid \mathbf{0}]^\intercal, [R \mid \mathbf{t}]^\intercal)).$$

Note that, so long as $\mathbf{v} \neq \mathbf{0}$, even if $\mathbf{t} = \mathbf{0}$ we can still identify (Σ, a) with the fundamental matrix of Σ, which is $[\mathbf{v}]_\times S$. In the $\mathbf{t} = \mathbf{0}$ case we get $S = R^{-1}$, identifying this new fundamental matrix with an "inverse" or "dual" fundamental matrix. The necessary computations to work out the exact underlying structure here is beyond the scope of this document, but this setup provides a way to view camera pairs both algebraically (as maps $\mathbf{P}^3 \dashrightarrow \mathbf{P}^2$) and geometrically (as a tuple of point-plane pairs). From here, it may be possible to obtain a compactification of Cam_2° that provides continuity between the non-concentric and concentric camera parameter spaces.

Chapter 4

THE ESSENTIAL VARIETY REVISITED

This chapter can be seen as a rewrite of the work of Lucas Van Meter, Max Lieblich, Bianca Viray, and Benjamin Antieau on the essential variety, which can be found in the dissertation of Van Meter [Van19].

4.1 Camera calibration and the essential variety

In the problem of triangulation, we note that scene reconstruction from a multiview configuration is only unique up to *projective automorphism*. To obtain geometrically accurate

Figure 4.1: Projective ambiguity of 3D reconstructions, c/o [HZ03, Fig. 1.4]

reconstructions from multiview configurations, it is necessary to introduce additional data which ensures that the Euclidean structure of the world is preserved under the camera transformation. This data comes in the form of a *calibrating conic* and its image, which tell us whether a given camera preserve orthogonality in the world.

Definition 4.1.1 (Image of the absolute conic)**.** Let $D = V(x_0^2 + x_1^2 + x_2^2) \subset \mathbf{P}^2$ be the canonical conic curve at infinity in $\mathbf{P}^2$. We call D the *image of the absolute conic (IAC)*.

Definition 4.1.2 (Relative calibration datum [LV20])**.** A *relative calibration datum with respect to D* for a pinhole camera A is a degree 2 plane curve $C \subset \mathbf{P}^3$ such that the restriction $A|_C : C \dashrightarrow \mathbf{P}^2$ factors through the inclusion $D \hookrightarrow \mathbf{P}^2$. If C is smooth, the calibration datum is called *smooth* or *non-degenerate*; otherwise it is called *degenerate*.

Remark 4.1.3. The plane curve C in Definition 4.1.2 is called "relative" because we have specified a "calibrated plane" $(\mathbf{P}^2, D)$, whereas in [LV20] a (non-relative) calibration datum allows the additional parameter of choosing a different smooth conic $D' \subset \mathbf{P}^2$ for which $A|_C : C \xrightarrow{\sim} D'$ is an isomorphism.

Definition 4.1.4 (Calibrated camera, calibrating curve)**.** A *calibrated camera* is a pair (A, C) where A is a pinhole camera and C is a relative calibration datum for A with respect to D. We call C the *calibrating curve* or *absolute conic* of A.

Definition 4.1.5 (Simultaneously calibrated multiview configuration)**.** A *(simultaneously) calibrated multiview configuration* is a pair $(\mathbf{A}, C)$ where $\mathbf{A}$ is a multiview configuration and C is a calibrating curve for each A_i.

Remark 4.1.6. In order for a multiview configuration to be simultaneously calibrated with a non-degenerate calibrating curve, the intersection of the cones $C_i := A_i^{-1}(D)$ must contain a conic curve, a fact that comes in handy in Section 4.2.1.

Definition 4.1.7 (Isomorphism of calibrated multiview configurations)**.** Two simultaneously calibrated multiview configurations $(\mathbf{A}, C), (\mathbf{A}', C')$ of common length n are *isomorphic* if there exists an isomorphism $\alpha : \mathbf{A} \xrightarrow{\sim} \mathbf{A}'$ of multiview configurations such that $\alpha(C) = C'$.

Definition 4.1.8 (Essential matrix)**.** If a pair of cameras (A_1, A_2) is simultaneously calibrated, then their fundamental matrix is called an *essential matrix*.

Definition 4.1.9 (The essential variety/essential divisors)**.** The *essential variety* $\mathbf{E}$ is the Zariski closure of the subvariety of the fundamental variety $\mathbf{F}$ containing all essential matrices. We say that members $J \in E$ are called *essential divisors*.

Remark 4.1.10. The condition of preserving the absolute conic is an algebraic way of saying that A must preserve the Euclidean structure of the world up to scaling. Although projective space has no innate Euclidean structure, we can let the affine patches $\mathbf{A}^3 \simeq D_{x_3}$ and $\mathbf{A}^2 \simeq D_{x_2}$ represent the "real" (visible) world and image plane, respectively. In practice, these affine patches are endowed with Euclidean structure. Thus in the following paragraphs we do not concern ourselves with the fact that "angles" are ill-defined in $\mathbf{P}^3$ and $\mathbf{P}^2$.

4.1.1 *Intuition behind the absolute conic*

Given a pinhole camera A, why does the preservation of the absolute conic ensure that the image of A preserves the Euclidean structure of the world? In this section we provide intuition for this idea by setting the absolute conic to be $C := V(x_0^2 + x_1^2 + x_2^2) \cap \pi_\infty \subset \mathbf{P}^3$, where $\pi_\infty = V(x_3)$ is the standard plane at infinity.

Note that every line $\ell \subset \mathbf{P}^3$ not lying on π_∞ has one intersection point $d = \ell \cap \pi_\infty$ which we can think of as its "direction." In a Euclidean sense, the intersection is exactly the direction vector of the line (assuming that the "true" Euclidean world is $D_{x_3} \subset \mathbf{P}^3$ as mentioned before).

Definition 4.1.11 (Polar lines, conjugate points)**.** For a point $q \in \pi_\infty$, we can define the *polar line of* q with respect to C as the line through the two points on C whose tangent lines contain q. Two points $d_1, d_2 \in \pi_\infty$ are *conjugate* with respect to C if they lie on each other's polar lines (see Figure 4.2).

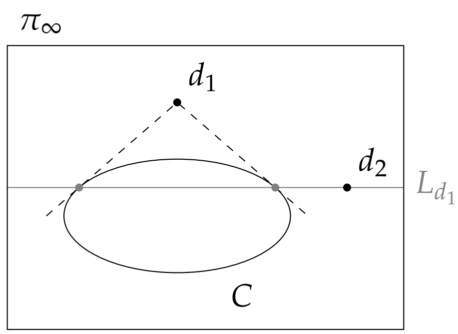


Figure 4.2: The two directions d_1 and d_2 are conjugate with respect to C.

Setting $C = V(x_0^2 + x_1^2 + x_2^2) \cap \pi_\infty$, we compute that the polar line L_{d_1} of C with respect to the point $d_1 = (y_0 : y_1 : y_2 : 0)$ is

$$L_{d_1} = V(2x_0y_0 + 2x_1y_1 + 2x_2y_2) = V(x_0y_0 + x_1y_1 + x_2y_2). \tag{4.1}$$

Thus $d_2 \in L_{d_1}$ if and only if $d_2^\intercal d_1 = 0$. Note that this condition corresponds exactly to d_1 and d_2 being orthogonal in Euclidean space. So two lines $\ell_1, \ell_2 \subset \mathbf{P}^3$ are orthogonal within the standard Euclidean structure if and only if their directions are conjugate with respect to the absolute conic. The same condition for orthogonality holds in $\mathbf{P}^2$ as well; that is, if A maps C to D isomorphically, then the orthogonality of lines is preserved.

Remark 4.1.12. Let $\mathbf{A} = (A_1, A_2)$ be a simultaneously calibrated pair of cameras, with arbitrary calibrating curve C. Then we can choose a representation of $\mathbf{A}$ such that C is the plane conic $V(x_0^2 + x_1^2 + x_2^2) \cap \pi_\infty \subset \mathbf{P}^3$. In this case, $\mathbf{A}$ can be written as $\mathbf{A} = ([I \mid \mathbf{0}], [R \mid \mathbf{t}])$, where $R \in \mathrm{SO}_3(\mathbf{C})$ and $t \in \mathbf{P}^2$. Descending to $\mathbf{R}$ and applying the "de-calibration" map we get the classical twisted pair cover

$$\mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) \to \mathbf{E}(\mathbf{R}). \tag{4.2}$$

4.2 *A modular compactification of the classical double cover*

We summarize and rephrase the main results of [Van19, Chapter 5]. The main theorem can be expressed as follows.

Theorem 4.2.1 ([Van19]). *The classical twisted pair cover*

$$\mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) \to \mathbf{E}(\mathbf{R}) \tag{4.3}$$

admits a unique extension to a finite étale cover $\pi : \widetilde{\mathbf{E}} \to \mathbf{E}$ by a projective scheme. Moreover, the scheme $\pi^{-1}(\mathbf{E}^{\mathrm{sm}}) \subset \widetilde{\mathbf{E}}$ is naturally the moduli space of calibrated pairs of cameras. Furthermore,

(i) The covering map π can be realized as a restriction to hyperplanes of the de-ordering map $\mathbf{P}^3 \times \mathbf{P}^3 \to \mathrm{Sym}^2 \mathbf{P}^3$.

$$\begin{array}{ccc} \mathbf{P}^3 \times \mathbf{P}^3 & \xrightarrow{\quad\pi\quad} & \mathrm{Sym}^2 \mathbf{P}^3 \\ \cup & & \cup \\ \widetilde{\mathbf{E}} & \xrightarrow{\quad\pi|_{\widetilde{\mathbf{E}}}\quad} & \mathbf{E} \\ \downarrow & & \downarrow \\ \mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) & \longrightarrow & \mathbf{E}(\mathbf{R}) \end{array} \tag{4.4}$$

In particular, $\mathbf{E}$ is naturally a hyperplane section of $\mathrm{Sym}^2 \mathbf{P}^3$.

(ii) The classical degree 10 embedding $\mathbf{E} \hookrightarrow \mathbf{P}^8$ is canonical, and arises from the ample generator for $\mathrm{Pic}(\mathbf{E})$, which is isomorphic to $\mathbf{Z}$. Moreover, the ideal of $\mathbf{E}$ in $\mathbf{P}^8$ via this embedding is defined by exactly 10 cubic forms, which are the classical Demazure equations.

We prove part *(i)* in Section 4.2.2, which builds off the contents of Section 4.2.1. It is worth noting that part *(ii)* is a re-statement of the results of [FKO18]. Unlike their work, however, the method of proof used in [Van19] does not rely on the existence of the Demazure equations but rather re-derives them. This observation, as well as the proof of part *(ii)* is discussed in detail in Section 4.2.3.

4.2.1 *A reducibility criterion for simultaneous calibration*

We describe a reducibility criterion for simultaneous calibration. The key result is the following theorem.

Theorem 4.2.2 (A reducibility criterion for simultaneous calibration, [Van19])**.** *Let* $\mathbf{A} : \mathbf{P}^3 \dashrightarrow \mathbf{P}^2 \times \mathbf{P}^2$ *be a non-concentric pair of cameras. Let* $J = \overline{\operatorname{im} \mathbf{A}} \subset \mathbf{P}^2 \times \mathbf{P}^2$ *denote the joint image, or* $(1,1)$*-divisor of* $\mathbf{P}^2 \times \mathbf{P}^2$*, that is cut out by the fundamental matrix* F *of* $\mathbf{A}$*.*

Then $\mathbf{A}$ *is a calibrated configuration if and only if the restriction* $J \cap (D \times D)$ *of* J *to the calibrating conics in the image planes is reducible as a* $(2,2)$*-divisor of* $\mathbf{P}^1 \times \mathbf{P}^1 \simeq D \times D$ *in one of the following ways:*

(i) $J \cap (D \times D) = E + F \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(2,2)|$ *where* $E, F \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)|$ *are both smooth, or*

(ii) $J \cap (D \times D) = E + F \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(2,2)|$ *where* $E \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)|$ *is smooth and contains the singular point of* J*, and* F *is the union of a* $(1,0)$ *curve and a* $(0,1)$ *curve that both pass through the singular point.*

The background work for the theorem begins with the following observation.

Proposition 4.2.3 (Reducibility criterion for simultaneous calibration, [Van19])**.** *Let* $\mathbf{A}$ *be the multiview configuration* $\mathbf{A} = (A_1, \ldots, A_n)$*. Let* $H_i = A_i^{-1}(D)$ *be the cone of* A_i *over* D *for each* i*. Then* A_i *is calibrated for all* i *if and only if there is a smooth conic curve* $C \subset H_1 \cap \cdots \cap H_n$*.*

Proof. It is clear that if such a conic C exists then each A_i is calibrated with calibrating conic C, and hence $\mathbf{A}$ is. On the other hand, suppose $A_1, \ldots, A_n$ is calibrated for each i with calibrating curve C. But then $C \subset H_i$ for each i, so it must lie within the intersection $H_1 \cap \cdots \cap H_n$. □

The key ingredient for the proof of Theorem 4.2.2 is the following classification of cone intersections.

Proposition 4.2.4 (E.g., [Van19])**.** *Two cones contained in* $\mathbf{P}^3$ *with distinct cone points intersect in:*

(i) a smooth quartic,

(ii) the union of two smooth conics, or

(iii) the union of a smooth conic and a doubled line.

Proof. Omitted. □

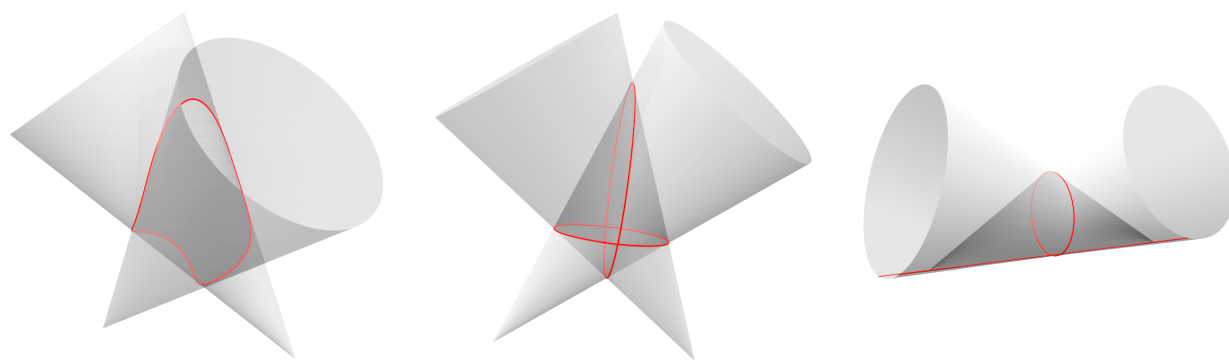


Figure 4.3: Two cones intersecting in a smooth quartic, two smooth conics, and a smooth conic and doubled line, respectively.

From this classification we immediately get the following corollary of Proposition 4.2.3.

Corollary 4.2.5 ([Van19])**.** *Two cameras $A_1, A_2 : \mathbf{P}^3 \dashrightarrow \mathbf{P}^2$ with distinct centers are simultaneously calibrated if and only if the intersection curve $H_1 \cap H_2 \subset \mathbf{P}^3$ is reducible.*

Proof. By Proposition 4.2.3 $H_1 \cap H_2$ must contain a conic curve, which, by Proposition 4.2.4, is possible exactly when $H_1 \cap H_2$ is reducible. □

Remark 4.2.6. Let $\mathbf{A} = (A_1, A_2)$ be a nonconcentric pair of cameras with joint image $J \subset \mathbf{P}^2 \times \mathbf{P}^2$. By Corollary 4.2.5, J is an essential divisor if and only if the intersection of cones $A_1^{-1}(D) \cap A_2^{-1}(D) = H_1 \cap H_2$ is reducible.

Since a pair of cameras is uniquely determined by its joint image, which is a divisor of $\mathbf{P}^2 \times \mathbf{P}^2$, it is natural to seek a geometric or divisor-theoretic condition on J to detect the "reducedness" of the intersection of the cones over D.

To build intuition, consider the case when $H_1 \cap H_2$ is the union of two smooth conics. Here $H_1 \cap H_2$ contains neither of the camera centers, so $J \cap (D \times D) = \mathbf{A}(H_1 \cap H_2)$ does not contain the singular point of J. Then the divisor J is rationally equivalent to the union $(\ell_1 \times \mathbf{P}^2) \cup (\mathbf{P}^2 \times \ell_2)$ for lines $\ell_1, \ell_2 \subset \mathbf{P}^2$ intersecting D transversely. It follows that the intersection $J \cap (D \times D)$ is rationally equivalent to

$$(\{p_1, p_2\} \times D) \cup (D \times \{q_1, q_2\}). \tag{4.5}$$

Moreover, since D is a smooth conic isomorphic to $\mathbf{P}^1$, $J \cap (D \times D)$ has class $(2,2)$ in $\mathrm{Pic}(D \times D) \simeq \mathbf{Z} \times \mathbf{Z}$. This is true in general for any joint image J and in fact the intersection correspondence $J \mapsto J \cap (D \times D)$ gives a bijection of linear systems.

Lemma 4.2.7 ([Van19])**.** *The map*

$$\begin{aligned} |\mathcal{O}_{\mathbf{P}^2 \times \mathbf{P}^2}(1,1)| &\to |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(2,2)| \\ J &\mapsto J \cap (D \times D) \end{aligned}$$

is an isomorphism of linear systems.

Proof. Omitted. □

By Lemma 4.2.7, a camera pair $\mathbf{A} = (A_1, A_2)$ is determined exactly by $\overline{\mathbf{A}(H_1 \cap H_2)} = J \cap (D \times D)$. We know that $\mathbf{A}$ is injective apart from the camera centers, and its "calibratedness" relies entirely on the reducibility of $H_1 \cap H_2$, so the same reducibility criterion can be applied to $\mathbf{A}(H_1 \cap H_2)$.

We now prove Theorem 4.2.2, which follows neatly from the above results.

Proof of Theorem 4.2.2. We show that one of these two criteria is fulfilled if and only if the conditions of Corollary 4.2.5 are. Notice that both criteria can only be fulfilled if the intersection curve $H_1 \cap H_2$ is reducible; otherwise $J \cap (D \times D)$ will not be reducible.

If the intersection curve $H_1 \cap H_2$ is reducible, it contains a smooth conic not containing either camera center, say C_1. Then $\mathbf{A}(C_1) \subset J \cap (D \times D)$ is a smooth curve E.

The residual curve $J \cap (D \times D) - E$ is the image of the other degree 2 curve in $H_1 \cap H_2$, say C_2. If C_2 is another smooth conic, also not containing the camera centers, then indeed $\mathbf{A}(C_2) = J \cap (D \times D) - E$ is smooth. Otherwise, C_2 is a doubled line containing the epipolar line. In this case, E indeed must contain the singular point of J as C_1 passes through this line (see Figure 4.3 for visual intuition). Furthermore, one can compute directly that $\mathbf{A}(C_2) = D \times A_2(c_1) \cup A_1(c_2) \times D$, which is exactly the union of a $(1,0)$- and $(0,1)$-curve containing the singular point $(A_1(c_2), A_2(c_1))$. □

Corollary 4.2.8 ([Van19])**.** *Suppose $J \subset \mathbf{P}^2 \times \mathbf{P}^2$ is a fundamental divisor. If the singular point (d_1, d_2) of J is contained in $D \times D$, then $J \cap (D \times D)$ contains $\{d_1\} \times D + D \times \{d_2\}$, which has class $(1,1)$ in* $\mathrm{Pic}(\mathbf{P}^1 \times \mathbf{P}^1)$.

Proof. If $J \cap (D \times D)$ contains the singular point, then the intersection curve $H_1 \cap H_2$ contains the epipolar line. This corresponds exactly to case *(ii)* in Theorem 4.2.2, in whose proof we saw that $J \cap (D \times D)$ contains the desired union of curves. □

The reducibility criterion of Theorem 4.2.2 gives a new description of the essential variety $\mathbf{E}$ in terms of divisors.

Corollary 4.2.9 ([Van19]). *The essential variety $\mathbf{E} \subset |\mathcal{O}_{\mathbf{P}^2 \times \mathbf{P}^2}(1,1)|$ is the closure of the locus of divisors J such that*

(a) J lies in the determinantal hypersurface $\mathbf{F} \subset |\mathcal{O}_{\mathbf{P}^2 \times \mathbf{P}^2}(1,1)|$;

(b) The image of J under the isomorphism

$$\begin{aligned} |\mathcal{O}_{\mathbf{P}^2 \times \mathbf{P}^2}(1,1)| &\to |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(2,2)| \\ J &\mapsto J \cap (D \times D) \end{aligned}$$

lies in the locus of divisors of the form $E + F$ where $E \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)|$ is smooth, and $F \in |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)|$ is either smooth or equal to $\{d_1\} \times D + D \times \{d_2\}$. In particular, $E = \mathbf{A}(C)$ where C is the calibrating conic of $\mathbf{A}$.

Proof. This follows directly from Theorem 4.2.2. Either F is the smooth residual curve, or Corollary 4.2.8 ensures that F is $\{d_1\} \times D + D \times \{d_2\}$. □

4.2.2 Double covering and compactification of the essential variety

As Theorem 4.2.2 makes clear, the essential variety lives inside the locus of $|\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(2,2)|$ consisting of reducible divisors containing a smooth member of $|\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)|$. In particular, the smooth member is the image of the calibrating conic. A brief analysis of this locus produces a projective scheme $\widetilde{\mathbf{E}}$ that is naturally a compactification of the moduli space of calibrated camera pairs.

Suppose $\mathbf{A} = (A_1, A_2)$ is a calibrated pair of nonconcentric cameras with calibrating conic C and joint image J. Either E or F is the image of the calibrating conic C under $\mathbf{A}$. In fact, when we restrict to $\mathbf{R}$ (where the doubled line case can never happen) the act of swapping the roles of F and E is exactly the swap occurring in the classical twisted pair covering $\mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) \to \mathbf{E}(\mathbf{R})$. The goal of this section is to formalize this idea by proving the following proposition.

Proposition 4.2.10 (A compactified double cover of the essential variety, [Van19])**.** *The classical twisted pair cover* $\mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) \to \mathbf{E}(\mathbf{R})$ *admits a unique extension to a generically 2-to-1 étale cover* $\pi : \widetilde{\mathbf{E}} \to \mathbf{E}$*, where the cover* π *can be seen as a restriction to hyperplanes of the natural "unordering" map* $\mathbf{P}^3 \times \mathbf{P}^3 \to \mathrm{Sym}^2\,\mathbf{P}^3$.

$$\begin{array}{ccc} \mathbf{P}^3 \times \mathbf{P}^3 & \xrightarrow{\quad \pi \quad} & \mathrm{Sym}^2\,\mathbf{P}^3 \\ \cup & & \cup \\ \widetilde{\mathbf{E}} & \xrightarrow{\quad \pi|_{\widetilde{\mathbf{E}}} \quad} & \mathbf{E} \\ \downarrow & & \downarrow \\ \mathrm{SO}_3(\mathbf{R}) \times \mathbf{P}^2(\mathbf{R}) & \longrightarrow & \mathbf{E}(\mathbf{R}) \end{array} \tag{4.6}$$

Moreover, the scheme $\widetilde{\mathbf{E}}^\circ := \pi^{-1}(\mathbf{E}^{\mathrm{sm}}) \subset \widetilde{\mathbf{E}}$ *is naturally the moduli space* CalCam_2 *of calibrated pairs of cameras.*

As Theorem 4.2.2 makes clear, J is essential only if $\alpha(J)$ is reducible and contains a smooth member of $|\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)|$. Any such divisor in $|\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(2,2)|$ is in the image of one of the following divisor addition maps:

$$\begin{aligned} \alpha &: |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)| \longrightarrow |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(2,2)| \\ \beta_1 &: |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,0)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(0,1)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)| \longrightarrow |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(2,2)| \\ \beta_2 &: |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,0)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(0,1)| \longrightarrow |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(2,2)| \end{aligned} \tag{4.7}$$

Remark 4.2.11. Notice that, for each i, β_i factors through α via addition of the $(1,0)$- and $(0,1)$-divisors, denoted by γ_i. This divisor addition is given by the map $\gamma : |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,0)| \times |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(0,1)| \to |\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(1,1)|$, which is equivalent to the Segre embedding $\sigma : \mathbf{P}^1 \times \mathbf{P}^1 \to \mathbf{P}^3$. So $\gamma_1 = \sigma \times \mathrm{id}$ and $\gamma_2 = \mathrm{id} \times \sigma$.

Projectivizing each line bundle, we obtain the following diagram

$$\begin{array}{l} \mathbf{P}^1 \times \mathbf{P}^1 \times \mathbf{P}^3 \\ \qquad\qquad \searrow^{\gamma_1} \\ \qquad\qquad\qquad \mathbf{P}^3 \times \mathbf{P}^3 \xrightarrow{\alpha} \mathbf{P}^8 \supset \mathbf{F} \supset \mathbf{E} \\ \qquad\qquad \nearrow_{\gamma_2} \\ \mathbf{P}^3 \times \mathbf{P}^1 \times \mathbf{P}^1 \end{array} \tag{4.8}$$

where, slightly abusing notation, $\alpha : \mathbf{P}^3 \times \mathbf{P}^3 \to \mathbf{P}^8$ is the composition of the divisor addition map as defined above and the isomorphism $|\mathcal{O}_{\mathbf{P}^1\times\mathbf{P}^1}(2,2)| \xrightarrow{\sim} |\mathcal{O}_{\mathbf{P}^2\times\mathbf{P}^2}(1,1)|$.

Lemma 4.2.12 ([Van19]). *The preimage $\widetilde{\mathbf{E}} = \alpha^{-1}(\mathbf{E})$ is a smooth bilinear divisor of $\mathbf{P}^3 \times \mathbf{P}^3$. Moreover, $\widetilde{\mathbf{E}}$ is invariant under the swapping action of $\mathbf{Z}/2\mathbf{Z}$, so $\alpha|_{\widetilde{\mathbf{E}}}$ factors through the quotient map $\pi : \mathbf{P}^3 \times \mathbf{P}^3 \to \mathrm{Sym}^2 \mathbf{P}^3$. In particular, $\mathbf{E} \simeq \pi(\widetilde{\mathbf{E}})$ is expressed as a hyperplane section of* $\mathrm{Sym}^2 \mathbf{P}^3$.

$$\begin{array}{ccc} \widetilde{\mathbf{E}} = \alpha^{-1}(\mathbf{E}) & \xrightarrow{\alpha|_{\widetilde{\mathbf{E}}}} & \mathbf{E} \\ \cap & & \cap \\ \mathbf{P}^3 \times \mathbf{P}^3 & \xrightarrow{\alpha} & \mathbf{P}^8 \\ \downarrow \pi & \nearrow & \\ \mathrm{Sym}^2 \mathbf{P}^3 & & \end{array} \tag{4.9}$$

Proof. Members of $\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,1)$ are bilinear forms on $\mathbf{P}^1 \times \mathbf{P}^1$, so each divisor is represented by a 2×2 matrix. Moreover, we can see members of $\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(0,1)$ as column vectors $\mathbf{v}$ and those of $\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1}(1,0)$ as row vectors $\mathbf{u}^\intercal$. The sum of two such divisors corresponds to the product $\mathbf{v}\mathbf{u}^\intercal$, which is rank-deficient and thus not smooth. Let

$$a = \begin{pmatrix} a_0 & a_1 \\ a_2 & a_3 \end{pmatrix}, \qquad b = \begin{pmatrix} b_0 & b_1 \\ b_2 & b_3 \end{pmatrix}$$

be coordinates on the first and second $\mathbf{P}^3$ factors, respectively. Notice that, with this formulation, we have $\mathrm{im}(\gamma_1) = V(a_0a_3 - a_1a_2)$, and similarly $\mathrm{im}(\gamma_2) = V(b_0b_3 - b_1b_2)$.

By construction, α sends

$$(a,b) \mapsto \begin{pmatrix} a_0b_0 & a_1b_0 + a_0b_1 & a_1b_1 \\ a_2b_0 + a_0b_2 & a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0 & a_3b_1 + a_1b_3 \\ a_2b_2 & a_3b_2 + a_2b_3 & a_3b_3 \end{pmatrix}, \tag{4.10}$$

which has determinant $(a_0a_3 - a_1a_2)(b_0b_3 - b_1b_2)(a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0)$.

Let $\widetilde{\mathbf{F}} = \alpha^{-1}(\mathbf{F}) = \alpha^{-1}(\det = 0)$. We see that $\widetilde{\mathbf{F}}$ decomposes as the sum of irreducible divisors

$$\widetilde{\mathbf{F}} = \mathrm{im}(\gamma_1) + \mathrm{im}(\gamma_2) + \widetilde{\mathbf{E}}$$

where $\widetilde{\mathbf{E}} := V(a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0)$.

The residual divisor $\widetilde{\mathbf{E}}$ must be the closure of the locus of divisor pairs (E,F) where at least one of E,F is smooth. That is, $\widetilde{\mathbf{E}} = \alpha^{-1}(\mathbf{E})$. Moreover, the polynomial cutting out $\widetilde{\mathbf{E}}$ is symmetric in a,b, so $\widetilde{\mathbf{E}}$ is indeed stable under the $\mathbf{Z}/2\mathbf{Z}$ swapping action. Noting that

$\pi(\widetilde{\mathbf{E}}) = V(a_0b_3 - a_1b_2 - a_2b_1 + a_3b_0) \subset \operatorname{Sym}^2 \mathbf{P}^3$ is a hyperplane and isomorphic to $\mathbf{E}$, we get the last statement. □

We see that $\alpha|_{\widetilde{\mathbf{E}}} : \widetilde{\mathbf{E}} \to \mathbf{E}$ gives a 2-to-1 cover of $\mathbf{E}$. Moreover, removing the rank-one locus from $\widetilde{\mathbf{E}}$ results in a smooth scheme $\widetilde{\mathbf{E}}^\circ$ that parametrizes all calibrated, nondegenerate, non-concentric camera pairs. We formalize this notion below.

Lemma 4.2.13 ([Van19])**.** *The scheme $\widetilde{\mathbf{E}}^\circ$ represents the functor* CalCam_2° *of [LV20] whose T-valued points are tuples (P, A_1, A_2, C) where $P \to T$ is a Zariski form of $\mathbf{P}^3_T$, $A_i : P \dashrightarrow \mathbf{P}^2_T$ $(i = 1, 2)$ are surjective linear projections with distinct kernels, and $C \subset P$ is a T-flat family of curves of degree 2 such that A_i is defined on C and the restriction $\mathbf{A}_i|_C$ factors through the closed immersion $D_T \subset \mathbf{P}^2_T$.*

Proof. By construction, there is a morphism $\alpha|_{\widetilde{\mathbf{E}}} : \widetilde{\mathbf{E}} \to \mathbf{F}$. Restricting to the rank-two locus $\widetilde{\mathbf{E}}^\circ := \alpha^{-1}(\mathbf{E}^{\mathrm{sm}})$, we further get a morphism $\widetilde{\mathbf{E}}^\circ \to \mathrm{Cam}_2^\circ$. Fix an S-scheme T and consider the point $f \in h_{\widetilde{\mathbf{E}}^\circ}(T)$, and let $g = \alpha \circ f$ be the resulting T-point in Cam_2.

$$\begin{array}{ccc} T & & \\ \downarrow f & \searrow^{g=\alpha\circ f} & \\ \widetilde{\mathbf{E}}^\circ & \xrightarrow{\alpha} & \mathrm{Cam}_2 \end{array}$$

A point $(E, F) \in \widetilde{\mathbf{E}}^\circ$ represents a calibrated camera pair $\mathbf{A}$ with $\mathbf{A}^{-1}(E) = C_0$ for some smooth conic $C_0 \subset \mathbf{P}^3$. Consequently, the image $\mathrm{im}(f)$ gives a family of smooth conics

$$\mathcal{C}_0 = \{C_0 := \mathbf{A}^{-1}(E) : (E, F) \in \mathrm{im}(f)\}$$

Pulling back $\mathcal{C}_0$ to P gives us the additional data of a T-flat family C of degree-2 curves that satisfies the desired conditions.

Conversely, a T-point gives a morphism $g : T \to \mathrm{Cam}_2$ which, given the additional data of C, factors through $\widetilde{\mathbf{E}}^\circ$. □

Combining Lemmas 4.2.12 and 4.2.13 and seeing $\mathbf{E}$ as a subvariety of $\operatorname{Sym}^2 \mathbf{P}^3$ gives us Proposition 4.2.10.

4.2.3 Connections to previous work

To wrap up the contents of Theorem 4.2.1, we consider the expression of $\mathbf{E}$ as a subscheme of $\operatorname{Sym}^2 \mathbf{P}^3$. This section also compares Theorem 4.2.1 to the results of [FKO18], in which $\mathbf{E}$ is described as a hyperplane section of $X_4^{\leq 2}$, the scheme of symmetric 4×4 matrices of rank at most 2.

Also, to recover the Demazure equations, we construct a presentation of $\operatorname{Sym}^2 \mathbf{P}^3$ as a subscheme of $\mathbf{P}^9$ to get $\mathbf{E} \simeq \operatorname{Sym}^2 \mathbf{P}^3 \cap \mathbf{P}^8$ (where $\mathbf{P}^8$ is a hyperplane in $\mathbf{P}^9$). By analyzing $\operatorname{Sym}^2 \mathbf{P}^3 \cap \mathbf{P}^8 \subset \mathbf{P}^8$, we see that $\mathbf{E}$ can be realized as a subscheme generated by 10 cubic forms in $\mathbf{P}^8$.

For ease of reading (and writing), we omit most of the proofs in this section.

Proposition 4.2.14 ([Van19])**.** *Let $X_4^{\leq 2}$ denote the scheme of symmetric 4×4 matrices of rank at most 2. This is the cone in the affine space of symmetric 4×4 matrices whose ideal is generated by the 3×3 minors. The map $\gamma : (x, y) \mapsto xy^\intercal + yx^\intercal$ defines an isomorphism*

$$\operatorname{Sym}^2 \mathbf{P}^3 \to \mathbf{P}(X_4^{\leq 2}) \tag{4.11}$$

Proof. Omitted. □

We saw in Lemma 4.2.12 that $\mathbf{E}$ is a hyperplane section of $\operatorname{Sym}^2 \mathbf{P}^3$. Therefore Proposition 4.2.14 tells us that Proposition 4.2.10 is actually a reformulation of the aforementioned result of Fløystad, Kileel, and Ottaviani.

4.2.4 *A presentation of* $\operatorname{Sym}^2 \mathbf{P}^3$ *in* $\mathbf{P}^9$

Here we give the desired presentation of $\operatorname{Sym}^2 \mathbf{P}^3$ as a subscheme of $\mathbf{P}^9$. The results in this section are proved in [Van19] by way of certain results in sheaf cohomology.

Proposition 4.2.15 ([Van19])**.** *There is a canonical embedding $\operatorname{Sym}^2 \mathbf{P}^3 \hookrightarrow \mathbf{P}^9$, unique up to automorphism of $\mathbf{P}^9$, such that any hyperplane section has dimension 5 and degree 10.*

Proof. Omitted. □

Lemma 4.2.16. *The map* $s : \mathrm{Sym}^2 \mathbf{P}^3 \to \mathbf{P}^9$ *given in homogeneous coordinates by*

$$\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} \mapsto \begin{pmatrix} a_0 b_0 \\ a_1 b_0 + a_0 b_1 \\ a_1 b_1 \\ a_2 b_0 + a_0 b_2 \\ a_0 b_3 + a_3 b_0 \\ a_1 b_2 + a_2 b_1 \\ a_3 b_1 + a_1 b_3 \\ a_2 b_2 \\ a_3 b_2 + a_2 b_3 \\ a_3 b_3 \end{pmatrix} =: \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \end{pmatrix} \tag{4.12}$$

is an embedding of schemes. Moreover, the ideal of $\mathrm{im}(s) \simeq \mathrm{Sym}^2 \mathbf{P}^3$ *is generated by 10 cubic forms.*

Proof. The well-definedness of s follows from the fact that each coordinate polynomial is symmetric in a, b. Injectivity is verified directly.

Using Macaulay2, we find that the ideal of $\mathrm{im}(s) \subset \mathbf{P}^9$ is generated by the 10 cubic polynomials

$$\begin{aligned}
x_6^2 x_7 - x_5 x_6 x_8 + x_2 x_8^2 + x_5^2 x_9 - 4 x_2 x_7 x_9 \\
2x_4 x_6 x_7 - x_4 x_5 x_8 - x_3 x_6 x_8 + x_1 x_8^2 + 2x_3 x_5 x_9 - 4x_1 x_7 x_9 \\
x_4^2 x_7 - x_3 x_4 x_8 + x_0 x_8^2 + x_3^2 x_9 - 4x_0 x_7 x_9 \\
x_4 x_5 x_6 - x_3 x_6^2 - 2x_2 x_4 x_8 + x_1 x_6 x_8 + 4x_2 x_3 x_9 - 2x_1 x_5 x_9 \\
x_4 x_5^2 - x_3 x_5 x_6 - 4x_2 x_4 x_7 + 2x_1 x_6 x_7 + 2x_2 x_3 x_8 - x_1 x_5 x_8 \\
x_4^2 x_5 - x_3 x_4 x_6 - x_1 x_4 x_8 + 2x_0 x_6 x_8 + 2x_1 x_3 x_9 - 4x_0 x_5 x_9 \\
x_3 x_4 x_5 - x_3^2 x_6 - 2x_1 x_4 x_7 + 4x_0 x_6 x_7 + x_1 x_3 x_8 - 2x_0 x_5 x_8 \\
x_2 x_4^2 - x_1 x_4 x_6 + x_0 x_6^2 + x_1^2 x_9 - 4x_0 x_2 x_9 \\
2x_2 x_3 x_4 - x_1 x_4 x_5 - x_1 x_3 x_6 + 2x_0 x_5 x_6 + x_1^2 x_8 - 4x_0 x_2 x_8 \\
x_2 x_3^2 - x_1 x_3 x_5 + x_0 x_5^2 + x_1^2 x_7 - 4x_0 x_2 x_7.
\end{aligned}$$

□

Since $\mathbf{E}$ is a hyperplane in $\mathrm{Sym}^2 \mathbf{P}^3$, we can now write $\mathbf{E} = \mathrm{Sym}^2 \mathbf{P}^3 \cap \mathbf{P}^8$ where $\mathbf{P}^8$ is a hyperplane. This results in the following proposition.

Proposition 4.2.17 ([Van19]). *Let $\mathbf{P}^8 \subset \mathbf{P}^9$ be a hyperplane and $\operatorname{Sym}^2 \mathbf{P}^3 \hookrightarrow \mathbf{P}^9$ an embedding. Then the embedding $E = \operatorname{Sym}^2 \mathbf{P}^3 \cap \mathbf{P}^8 \subset \mathbf{P}^8$ of the hyperplane section is the complete linear system on $\mathcal{O}_E(1)$, and the homogeneous ideal of E in $\mathbf{P}^8$ is generated by precisely 10 cubic forms.*

Proof. Details omitted. We outline a sketch:

(i) Let E be the vanishing of a hyperplane section of $\operatorname{Sym}^2 \mathbf{P}^3$, and let $B = \bigoplus_{n \geq 0} B_i$ be the graded ring for which $E = \operatorname{Proj}(B)$.

(ii) Letting $Q = \operatorname{Sym}^* B_1$ be the symmetric algebra of B_1, we see that $E = \operatorname{Proj}(B) \subset \operatorname{Proj}(Q) \simeq \mathbf{P}^8$.

(iii) Then, using a cohomology argument, we see that the ideal of $\operatorname{Sym}^2 \mathbf{P}^3 \subset \mathbf{P}^9$ descends to the ideal $J = \ker Q \to B$, which is the ideal of $E \subset \mathbf{P}^8$.

(iv) Thus, by Lemma 4.2.16, the ideal of E is generated by 10 cubic forms in $\mathbf{P}^8$.

□

4.3 Where do we go from here?

Recall that the divisor addition morphism

$$\alpha : |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1(1,1)}| \times |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1(1,1)}| \to |\mathcal{O}_{\mathbf{P}^1 \times \mathbf{P}^1(2,2)}| \tag{4.13}$$

gives an equivalent morphism $\mathbf{P}^3 \times \mathbf{P}^3 \to \mathbf{P}^8$.

A point correspondence is a linear constraint on $\mathbf{P}^8$, so determining a calibrated camera from five correspondences is equivalent to intersecting five hyperplanes with the essential variety. However, we could pull back the point-correspondence hyperplanes to bilinear forms on $\mathbf{P}^3 \times \mathbf{P}^3$ and intersect them there, too. This would give us each member of the possible twisted pairs. Intersecting five hyperplanes with the essential variety is basically intersecting five linear equations with the ten cubics (the Demazure equations) defining the essential variety in $\mathbf{P}^9$.In contrast, intersecting five bilinear forms with the essential cover only requires intersecting with one additional bilinear form on $\mathbf{P}^3 \times \mathbf{P}^3$. A brief computational experiment run by my colleague Duff gives us the suspicion that this could lend itself to a practical and efficient algorithm for reconstructing calibrated cameras. However, my own limited computational background prevents me providing an exact description of its efficiency and usefulness. It is a future research goal, though.

Chapter 5

INTERSECTION THEORY IN ALGEBRAIC VISION

This chapter contains a discussion and various computations related to the intersection theory of various algebraic varieties arising in algebraic vision. Similar to work in [HL18], these computations may be of use in the proof of Conjecture 2.5.1.

5.1 The story of this project

Like I mentioned in Section 1.3, studying problems in computer vision with an algebro-geometric lens has given me the opportunity to learn about “scary” math topics with a much narrower scope, which—for me—takes the scare away. Since taking Jake Levinson’s intersection theory class in Spring 2020, I was very intersted in becoming more proficient with the tools used in this area, but I was intimidated by the broad scope. A few years later, my collaborator and mentor Tim Duff showed me a paper by Harris and Lowengrub [HL18] that uses intersection-theoretic techniques to compute an upper bound for the Euclidean distance (ED) degree of the multiview variety. Duff and I, along with Erin Connelly, were working at the time on [CDL23] which includes at the end a conjecture regarding the ED degree of the *resectioning* variety. I took this as an opportunity to learn more about intersection theory with the end-goal of utilizing the methods to prove Conjecture 2.5.1. My time spent on the contents of this chapter were very rewarding and enlightening for me, although I do not yet know the larger narrative in which these results might belong.

5.2 Towards the total Chern class of the classical multiview variety

Knowing the total Chern class of the (tangent bundle of the) classical multiview variety may prove useful for verifying its ED degree, which then could help in proving Conjecture 2.5.1. It should be noted, though, that the relationship between $c(\mathcal{T}_{\Gamma^{1,n}_{\bar{\mathbf{A}},\mathbf{p}}})$ and $\mathrm{ED}(\Gamma^{1,n}_{\bar{\mathbf{A}},\mathbf{p}})$ is not so clear as in [HL18], since in their work Harris and Lowengrub rely on the fact that their formulation of the multiview variety is the *projective closure of an affine cone*, which is not the case for the classical construction.

The goal of this section is to prove the following proposition.

Proposition 5.2.1. *Let $\widetilde{\phi} : \widetilde{\mathbf{P}}^3 \hookrightarrow (\mathbf{P}^2)^n$ be the closed immersion whose image is the multiview variety $\Gamma^{1,n}_{\tilde{\mathbf{A}},\mathbf{p}}$. Then the pushforward of the total Chern class $c(\mathcal{T}_{\widetilde{\mathbf{P}}^3})$ under $\widetilde{\phi}$ is*

$$\widetilde{\phi}_* c(\mathcal{T}_{\widetilde{\mathbf{P}}^3}) = \sum_{\substack{\sum a_i = 2n-3 \\ a_i \le 2}} y^a + 4 \sum_{\substack{\sum a_i = 2n-2 \\ a_i \le 2}} y^a + 6 \sum_{\substack{\sum a_i = 2n-1 \\ a_i \le 2}} y^a + (4+2n) \prod_{i=1}^n y_i - 2 \sum_{i=1}^n \prod_{j \ne i} y_j^2.$$

Notation 5.2.2. We often use the notation $c(X)$ to denote the total Chern class of the tangent bundle of X, that is, $c(\mathcal{T}_X)$.

The immersion $\widetilde{\phi}$ is defined as follows.

Proposition 5.2.3 (E.g., [LV20])**.** *Let $\phi : \mathbf{P}^3 \dashrightarrow (\mathbf{P}^2)^n$ be a multiview configuration with pairwise distinct camera centers. Then ϕ admits a resolution $\widetilde{\phi} : \widetilde{\mathbf{P}}^3 \to (\mathbf{P}^2)^n$ that is an isomorphism onto its image. Moreover, $\operatorname{im}(\widetilde{\phi}) = \overline{\operatorname{im}(\phi)}$.*

Proposition 5.2.4. *The Chow ring of the multivew variety $\widetilde{\mathbf{P}}^3 = \operatorname{Bl}_{c_i} \mathbf{P}^3$ is*

$$A^\bullet(\widetilde{\mathbf{P}}^3) = \mathbf{Z}[x_0, \ldots, x_n] / (\{x_i \cdot x_j\}_{i \ne j}, \{x_i^3 - x_0^3\}_{i=1}^n)$$

where x_0 is the class of the strict transform of any hyperplane in $\mathbf{P}^3$, and, for $i > 0$, x_i is the class of the exceptional divisor $E_i = \pi^{-1}(c_i)$.

Proof. This follows directly from the classification of the Chow rings of blow-ups of projective space at finitely many distinct points [Por23, Thm. 3.6]. We note that, since none of the c_i are proximal (i.e., all of the points are disjoint in the ground scheme $\mathbf{P}^3$), the change-of-basis matrices from the strict transforms to the total transforms (and vice-versa) as defined in [Por23, Eq. 1] are just the identity.

One can also invoke [EH16, Cor. 9.6] and apply a change of variables to get the same result. □

Lemma 5.2.5. *The Chow ring of $(\mathbf{P}^2)^n$ is*

$$A^\bullet((\mathbf{P}^2)^n) = A^\bullet(\mathbf{P}^2) \otimes \cdots \otimes A^\bullet(\mathbf{P}^2) = \frac{\mathbf{Z}[y_1, \ldots, y_n]}{(y_1^3, \ldots, y_n^3)}$$

where y_i represents the class of the i-th hyperplane (i.e. $[(\mathbf{P}^2)^{i-1} \times \ell_i \times (\mathbf{P}^2)^{n-i}]$).

Proof. This follows directly from the fact that the Chow ring of the product of projective spaces is the tensor product of their Chow rings. □

The following proposition comes into use when computing the total Chern class of $\widetilde{\mathbf{P}}^3$.

Proposition 5.2.6 ([Ful84, Example 15.4.2])**.** *Let Y be a smooth scheme and $X \subset Y$ be a closed smooth subscheme of codimension d. Consider the following blowup diagram.*

$$\begin{array}{ccc} \widetilde{X} & \xrightarrow{j} & \widetilde{Y} \\ \downarrow{\scriptstyle g} & & \downarrow{\scriptstyle f} \\ X & \xrightarrow[i]{} & Y. \end{array} \tag{5.1}$$

Suppose that $c_k(N_{X/Y}) = i^* c_k$ for some $c_k \in A^k(Y)$, and that $c(X) = i^*\alpha$ for some $\alpha \in A^\bullet(Y)$. Let $\eta = c_1(\mathcal{O}_{\widetilde{Y}}(\widetilde{X})$. Then,

$$c(\widetilde{Y}) - f^* c(Y) = f^*(\alpha) \cdot \beta \tag{5.2}$$

where

$$\beta = (1+\eta)\sum_{i=0}^{d}(1-\eta)^i f^* c_{d-i} - \sum_{i=0}^{d} f^* c_{d-i}. \tag{5.3}$$

Proposition 5.2.7. *The total Chern class of $\mathcal{T}_{\widetilde{\mathbf{P}}^3}$ is*

$$c(\widetilde{\mathbf{P}}^3) = (1+x_0)^4 + \sum_{i=1}^{n}\Big((1+x_i)(1-x_i)^3 - 1\Big) \tag{5.4}$$

$$= 1 + 4x_0 + 6x_0^2 + (4+2n)x_0^3 - 2\sum_{i=1}^{n} x_i. \tag{5.5}$$

Proof. The proof can essentially be extracted from the proof of [HL18, Prop. 3], being careful to note where our notation differs. In particular, applying Proposition 5.2.6 to $Y = \mathbf{P}^3$ and X being the i-th camera center (here we avoid the usual notation c_i to avoid confusion between the i-th camera center and the i-th Chern class), we see that the contribution of blowing up each camera center is given by $\alpha = 1, c_0 = 1, c_k = 0$ $(k > 0)$. The first Chern class of the exceptional divisor over the i-th camera center is x_i, so we get $\beta = (1+x_i)(1-x_i)^3 - 1$.

Moreover, $f^*c(Y) = (1+x_0)^4$ in $\widetilde{\mathbf{P}}^3$, so we have

$$c(\widetilde{\mathbf{P}}^3) - (1+x_0)^4 = \sum_{i=1}^{n}\Big((1+x_i)(1-x_i)^3 - 1\Big) \tag{5.6}$$

as desired. Expanding eq. (5.6) with respect to the relations in Proposition 5.2.4, we get

$$\begin{aligned}
c(\widetilde{\mathbf{P}}^3) &= (1+x_0)^4 + \sum_{i=1}^{n} \Big((1+x_i)(1-x_i)^3 - 1) \Big) \\
&= (1+x_0)^4 + \sum_{i=1}^{n} (2x_i^3 - 2x_i) \\
&= 1 + 4x_0 + 6x_0^2 + 4x_0^3 + 2\sum_{i=1}^{n} x_0^3 - x_i \\
&= 1 + 4x_0 + 6x_0^2 + (4+2n)x_0^3 - 2\sum_{i=1}^{n} x_i.
\end{aligned}$$

□

Proposition 5.2.8. *The class of $\widetilde{\mathbf{P}}^3$ as a subvariety of $(\mathbf{P}^2)^n$ is*

$$[\widetilde{\phi}(\widetilde{\mathbf{P}}^3)] = \sum_{\substack{\sum a_i = 2n-3 \\ a_i \le 2}} y^a.$$

Proof. We use the method of undetermined coefficients as described in [EH16, Sec. 2.1]. The dimension of $\widetilde{\phi}(\widetilde{\mathbf{P}}^3)$ is 3, so its class is

$$\delta = [\widetilde{\phi}(\widetilde{\mathbf{P}}^3)] = \sum_{\substack{\sum a_i = 2n-3 \\ a_i \le 2}} c_a y^a$$

where $a = (a_1, \ldots, a_n)$ and $y^a = y_1^{a_1} \cdots y_n^{a_n}$. We determine the coefficients c_a by multiplying both sides with classes of complementary codimension. Then we get $c_a = \deg(\delta \cdot y^{a'})$ where $a'_i = 2 - a_i$. There are only two possible types of $y^{\alpha'}$, either $y_i^2 y_j$ or $y_i y_j y_k$ ($i \neq j \neq k$). Notice that $\widetilde{\phi}^*(y_i^2 y_j)$ is the intersection of a line through c_i and a plane through c_j, which is a point. Pushing forward, we have

$$y_i^2 y_j \cdot \delta = \widetilde{\phi}_* \widetilde{\phi}^* (y_i^2 y_j) = [\mathrm{pt}] = \prod_{k=1}^{n} y_k$$

so $c_a = 1$ for $y^a = \frac{\prod_{k=1}^n y_k}{y_i^2 y_j}$ for all $i \neq j$. Similarly, $\widetilde{\phi}^*(y_i y_j y_k)$ is the intersection of three planes, which is again a point, so we get

$$y_i y_j y_k \cdot \delta = \widetilde{\phi}_* \widetilde{\phi}^* (y_i y_j y_k) = [\mathrm{pt}] = \prod_{k=1}^{n} y_k.$$

That is, the coefficients are all 1, so we have

$$\delta = [\widetilde{\phi}(\widetilde{\mathbf{P}}^3)] = \sum_{\substack{\sum a_i = 2n-3 \\ a_i \leq 2}} y^a.$$

□

Finally, we compute the pushforward $\widetilde{\phi}_* c(\widetilde{\mathbf{P}}^3)$ as stated in Proposition 5.2.1.

Proof of Proposition 5.2.1. Recall that $c(\widetilde{\mathbf{P}}^3) = 1 + 4x_0 + 6x_0^2 + (4+2n)x_0^3 - 2\sum_{i=1}^n x_i$. So to determine the pushforward, it suffices to compute the classes of x_0, x_0^2, x_0^3, and x_i for $i = 1, \ldots, n$.

We immediately see that $\widetilde{\phi}_*(x_0^3) = [\mathrm{pt}] = \prod_{i=1}^n y_i$. Again implementing the method of undetermined coefficients and the push-pull formula, we get (for all i, j not necessarily distinct)

$$y_i y_j \cdot \widetilde{\phi}_*(x_0) = \widetilde{\phi}_*(\widetilde{\phi}^*(y_i y_j) \cdot x_0) = [\mathrm{pt}] \implies \widetilde{\phi}_*(x_0) = \sum_{\substack{\sum a_i = 2n-2 \\ a_i \leq 2}} y^a \tag{5.7}$$

$$y_i \cdot \widetilde{\phi}_*(x_0^2) = \widetilde{\phi}_*(\widetilde{\phi}^*(y_i) \cdot x_0^2) = [\mathrm{pt}] \implies \widetilde{\phi}_*(x_0^2) = \sum_{\substack{\sum a_i = 2n-1 \\ a_i \leq 2}} y^a \tag{5.8}$$

For x_i, we have

$$y_j y_k \cdot \widetilde{\phi}_*(x_i) = \widetilde{\phi}_*(\widetilde{\phi}^*(y_j y_k) \cdot x_i) = \begin{cases} \widetilde{\phi}_*(x_0^2 \cdot x_i) = 0 & i, j, k \text{ not all equal} \\ \widetilde{\phi}_*(\underbrace{(x_0^2 - x_i^2) \cdot x_i}_{= -x_i^3 = x_0^3}) = [\mathrm{pt}] & i = j = k \end{cases} \tag{5.9}$$

so $\widetilde{\phi}_*(x_i) = \prod_{j \neq i} y_j^2$. Combining terms, we get

$$\widetilde{\phi}_* c(\widetilde{\mathbf{P}}^3) = \sum_{\substack{\sum a_i = 2n-3 \\ a_i \leq 2}} y^a + 4 \sum_{\substack{\sum a_i = 2n-2 \\ a_i \leq 2}} y^a + 6 \sum_{\substack{\sum a_i = 2n-1 \\ a_i \leq 2}} y^a + (4+2n)\prod_{i=1}^n y_i - 2\sum_{i=1}^n \prod_{j \neq i} y_j^2.$$

□

We note that $c(\widetilde{\phi}_* \mathcal{T}_{\widetilde{\mathbf{P}}^3})$ is *not* equal to $A_* c(\mathcal{T}_{\widetilde{\mathbf{P}}^3})$, but it is certainly the case that they are related. For instance, Grothendieck-Riemann-Roch states the following.

Theorem 5.2.9 (Grothendieck-Riemann-Roch, e.g., [Ful84, Ch. 15])**.** *Let $j : X \to Y$ be a proper map of nonsingular varieties, and let $\mathcal{E}$ be a vector bundle on X. Then*

$$\operatorname{ch}(j_*\mathcal{E})\operatorname{td}(\mathcal{T}_Y) = j_*(\operatorname{ch}(\mathcal{E})\operatorname{td}(\mathcal{T}_X))$$

(where $\operatorname{ch}(\mathcal{E})$ *is the Chern character of* $\mathcal{E}$*, and* $\operatorname{td}(\mathcal{E})$ *is its Todd character).*

Letting $X = \widetilde{\mathbf{P}}^3, Y = (\mathbf{P}^2)^n, j = \widetilde{\phi}, \mathcal{E} = \mathcal{T}_{\widetilde{\mathbf{P}}^3}$, we get

$$\operatorname{ch}(\widetilde{\phi}_*\mathcal{T}_{\widetilde{\mathbf{P}}^3})\operatorname{td}(\mathcal{T}_{(\mathbf{P}^2)^n}) = \widetilde{\phi}_*(\operatorname{ch}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})\operatorname{td}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})).$$

The computations of $\operatorname{ch}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})$ and $\operatorname{td}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})$ are beyond the scope of this thesis, but my expectation is that they are not so complicated, or perhaps there is a nice formula for their product. Additionally, we have $\operatorname{td}(\mathcal{T}_{(\mathbf{P}^2)^n}) = \prod_{i=1}^n \left(\frac{y_i}{1-e^{-y_i}}\right)^3$. It follows that

$$\operatorname{ch}(\widetilde{\phi}_*\mathcal{T}_{\widetilde{\mathbf{P}}^3}) = \widetilde{\phi}_*(\operatorname{ch}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})\operatorname{td}(\mathcal{T}_{\widetilde{\mathbf{P}}^3})) \cdot \prod_{i=1}^{n}\left(\frac{1-e^{-y_i}}{y_i}\right)^3.$$

From here it would be straightforward to extract the total Chern class $c(\widetilde{\phi}_*\mathcal{T}_{\widetilde{\mathbf{P}}^3})$, which is the goal.

5.3 *Some questions regarding the univariate multiview variety*

Here we lay out some questions that, when answered, may shed light on some key differences between the "univariate" and "multivariate" multiview varieties.

Definition 5.3.1 (The univariate multiview variety, [HL18])**.** Let P_i be a 3×4 matrix, considering each row ℓ_{ij} as the affine function sending $q = (x, y, z)$ to $\ell_{ij} \cdot (x, y, z, 1)$. So we can see each P_i as a rational map from $\mathbf{C}^3 \to \mathbf{C}^2$ by $P_i(v) = (\ell_{i1}(q)/\ell_{i3}(q), \ell_{i2}(q)/\ell_{i3}(q))$. The ℓ_{ij} naturally extend to projective space, where $\ell_{ij}(q) = \ell_{ij} \cdot q$.

The *univariate multiview variety* MV_n' is the Zariski closure of $\operatorname{im}(\phi)$, where

$$\phi : \mathbf{P}^3 \dashrightarrow \mathbf{P}^{2n} \tag{5.10}$$

$$q \longmapsto ((\ell_{11} \cdot \prod_{i\neq 1} \ell_{i3})(q) :: (\ell_{12} \cdot \prod_{i \neq 1} \ell_{i3})(q) :: \cdots : (\ell_{n2} \cdot \prod_{i\neq n} \ell_{i3})(q) : \prod_{i=1}^{n} \ell_{i3}(q)). \tag{5.11}$$

Question 5.3.2. In the multivariate case, we know that the joint images for a length-2 multiview configuration are $(1,1)$-divisors of $\mathbf{P}^2 \times \mathbf{P}^2$. In the univariate case, what kind of divisors are the joint images in $\mathbf{P}^4$?

Question 5.3.3. What does the essential variety look like in the univariate setting?

Question 5.3.4. The multivariate multiview variety corresponding to an n-camera arrangement is nonsingular for $n > 2$, whereas the univariate multiview variety is singular for all n. However, computing a fairly tight upper bound for the univariate multiview variety is straightforward using the Chern-Mather class (see, e.g. [Mac74]). Does this method translate to the resectioning case, which is both higher-dimensional and which is singular even in the multivariate case?

Question 5.3.5. Two multivariate camera configurations are isomorphic if and only if their multiview varieties are equal as subschemes of $(\mathbf{P}^2)^n$. Is this still true in the univariate case?

Below we answer Question 5.3.5, whose proofs follow exactly as in [LV20].

Lemma 5.3.6. *Suppose $\mathbf{A}_1, \mathbf{A}_2 : \mathbf{P}^3 \dashrightarrow \mathbf{P}^{2n}$ are camera configurations and $\alpha : \mathbf{P}^3 \dashrightarrow \mathbf{P}^3$ is a birational automorphism such that $\mathbf{A}_2 = \mathbf{A}_1 \circ \alpha$. If α and $\mathbf{A}_1 \circ \alpha$ are both regular on an open subset $U \subset \mathbf{P}^3$ whose complement has codimension at least 2 then α extends to a unique regular automorphism $\tilde{\alpha} : \mathbf{P}^3 \to \mathbf{P}^3$.*

Proof. We can remove the indeterminacy locus of $\mathbf{A}_1$ from U if necessary, which is a union of lines; so U will still have complement of codimension at least 2, and now $\mathbf{A}_1, \alpha, \mathbf{A}_2$ are all regular on U. By assumption, $\mathbf{A}_i^* \mathcal{O}(1) = \mathcal{O}_U(1)$, so $\alpha^* \mathcal{O}(1) = \mathcal{O}(1)$. Then since $\Gamma(U, \mathcal{O}(1)) = \Gamma(\mathbf{P}^3, \mathcal{O}(1))$, by the universal property of projective space the morphism $U \to \mathbf{P}^3$ must extend to a unique $\widetilde{\alpha} : \mathbf{P}^3 \to \mathbf{P}^3$; and since α was birational, this must be an automorphism as desired. □

Proposition 5.3.7. *Two camera configurations $\mathbf{A}_1$ and $\mathbf{A}_2$ of length n are isomorphic if and only if their associated multiview varieties in $\mathbf{P}^{2n}$ are equal as closed subschemes.*

Proof. It is straightforward from the definitions that two isomorphic camera configurations have equal multiview varieties.

Conversely, since $\mathbf{A}_i$ is birational onto its image ($i = 1, 2$), we see that if $\overline{\mathrm{im}(\mathbf{A}_1)} = \overline{\mathrm{im}(\mathbf{A}_2)}$ then there is a birational automorphism $\alpha : \mathbf{P}^3 \dashrightarrow \mathbf{P}^3$ such that $\mathbf{A}_2 = \mathbf{A}_1 \circ \alpha$. Moreover, $\alpha, \mathbf{A}_1, \mathbf{A}_2$ are all regular on the open subscheme $U \subset \mathbf{P}^3$ which is the complement of the union of lines on which $\mathbf{A}_1, \mathbf{A}_2$ are not defined. The union of lines has codimension 2, so we can apply Lemma 5.3.6 to get $\mathbf{A}_1 \simeq \mathbf{A}_2$ as camera configurations. □

Finally, we include a conjecture regarding the degree of the univariate multiview variety.

Conjecture 5.3.8. *The degree of the univariate multiview variety is* $\frac{n(n-1)(n+4)}{6}$.

5.4 The Picard group of the universal imaging variety

The goal of this section is to prove the following:

Theorem 5.4.1. *The Picard group* $\mathrm{Pic}(\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}})$ *of the universal imaging variety is isomorphic to*

$$\mathrm{Pic}(\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}) \simeq \mathbf{Z}^{mn+1}.$$

The generators are x_{ij}, y $(1 \leq i, j \leq m, n)$*, where* x_{ij} *is the class of the pullback of the generator of* $\mathrm{Pic}(\mathbf{P}^2)$ *in the* (i, j)*-th factor, and* y *is the class of the image of the exceptional divisor of the blow-up* $\mathrm{Bl}_Z(\mathbf{P}^3)$ *at the locus of indeterminacy of the universal imaging map* $\Phi : (\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n \dashrightarrow (\mathbf{P}^2)^{mn}$.

Recall that the universal imaging variety $\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}$ is defined as the Zariski closure of the graph of the universal imaging map

$$\begin{aligned} \Phi : (\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n &\dashrightarrow (\mathbf{P}^2)^{mn} \\ (A_i, q_j) &\longmapsto A_i q_j. \end{aligned}$$

Proposition 5.4.2. *Let* $f : X \dashrightarrow Y$ *be a rational map of projective schemes over* $\mathbf{C}$ *whose base locus* Z *can be decomposed as a union of smooth closed subschemes* $Z = \bigcup_{i=1}^n Z_i$ *such that* Z_i, Z_j *intersect transversely for all* $i \neq j$*. Then*

(i) $\widetilde{f} : \mathrm{Bl}_Z X \to Y$ *is a morphism, and*

(ii) $\mathrm{Bl}_Z X \simeq \overline{\Gamma(f)}$.

Sketch of proof. By [Li09, Theorem 1.3], we have

$$\mathrm{Bl}_Z X = \mathrm{Bl}_{\widetilde{Z}_{i_n}} \cdots \mathrm{Bl}_{\widetilde{Z}_{i_2}} \mathrm{Bl}_{Z_{i_1}} X$$

(for some appropriate ordering of the Z_i) and this iterated blow-up is exactly the one that makes $\tilde{f}$ regular. Furthermore, $\mathrm{im}(\widetilde{f}) \subset X \times Y$ contains $\mathrm{im}(f|_{X \setminus Z})$, and is a reduced and irreducible subscheme of $X \times Y$, so it must coincide with $\overline{\Gamma(f)}$ □

Corollary 5.4.3. *Let Z denote the base locus of the universal imaging map Φ. Then*

$$\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}} \simeq \mathrm{Bl}_Z\left((\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n\right).$$

Proof. By Proposition 5.4.2, it remains to show that Z is a union of smooth closed subschemes whose intersections are transverse.

Rewrite Z as $Z = \bigcup_{i,j=(1,1)}^{m,n} Z_{ij}$, where

$$Z_{ij} := \{(\mathbf{A}, \mathbf{q}) \in (\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n : A_i q_j = \mathbf{0}\}.$$

Clearly each Z_{ij} is smooth.

Let $\pi_j : Z_{ij} \to \mathbf{P}^3$ be the j-th projection, and notice that the fibers of π_j are dimension 8, since $Aq = 0$ is cut out by three linear equations in A. Thus we can conclude that $\dim Z_{ij} = 11(m-1) + 3n + 8$; that is, $\operatorname{codim} Z_{ij} = 3$.

Consider the intersection $Z_{ab} \cap Z_{cd}$. If $a \neq c$ and $b \neq d$, the dimension of their intersection is clearly $11(m-2) + 3n + 16$; that is, codimension 6, as expected. If $b = d$, then we have $A_a q_b = A_c q_b = 0$. This again is of the expected codimension, since A_a and A_c are independent.

Finally, consider the case $a = c$. Generically, if a given $A_a q_b = A_a q_d = 0$, it must be that $q_b = q_d = \ker A$. This also gives us a codimension 6 locus.

Thus in all cases, $\operatorname{codim} Z_{ij} \cap Z_{k\ell} = 6$ as desired. ☐

Proof of Theorem 5.4.1. We have the blow-up diagram

$$\begin{array}{ccc} E & \xhookrightarrow{\ j\ } & W = \mathrm{Bl}_Z(X) \\ \Big\downarrow{\scriptstyle \pi_E} & & \Big\downarrow{\scriptstyle \pi} \\ Z & \xhookrightarrow{\ i\ } & X := (\mathbf{P}^{11})^m \times (\mathbf{P}^3)^n \end{array}$$

which induces the exact sequence of Chow groups

$$A^0(E) \xrightarrow{j_*} A^1(W) \xrightarrow{\pi_*} A^1(X) \longrightarrow 0.$$

Let $e = j_*[E]$ denote the class of E in $A^1(W)$. By [EH16, Theorem 13.12], we know that $j^*(e) = j^* j_*[E] = -\zeta$, where $\zeta = c_1(\mathcal{O}_E(1))$. Also, by [EH16, Theorem 9.6], we see that ζ

generates (additively) a subgroup of

$$A(E) \simeq \bigoplus_{i=0}^{\dim E} \zeta^i \cdot A(X) \simeq \bigoplus_{i=0}^{\dim E} \zeta^i \cdot \frac{\mathbf{Z}[x_1, \ldots, x_m, y_1, \ldots, y_n]}{(x_i^{12}, y_j^4)_{(i,j)=1}^{m,n}}.$$

In particular, ζ is a non-torsion element of $A(E)$.

If j_* were not injective, there would be some $n \in \mathbf{Z}$ for which $j_*(n[E]) = n \cdot e = 0$. But then $0 = j^*(n \cdot e) = -n\zeta$, a contradiction as we just saw that ζ is non-torsion.

Therefore the sequence must be short exact, and moreover it must split as $A^1(X) = \mathbf{Z}^{mn}$ is an injective $\mathbf{Z}$-module. That is,

$$A^1(W) \simeq A^0(E) \oplus \mathbf{Z}^{mn} \simeq \mathbf{Z} \times \mathbf{Z}^{mn} = \mathbf{Z}^{mn+1}.$$

Moreover, by Corollary 5.4.3, we have $\mathrm{Pic}(\Gamma^{m,n}_{\mathbf{A},\mathbf{q},\mathbf{p}}) \simeq A^1(W)$, completing the proof. $\square$

BIBLIOGRAPHY

[APT21] S. Agarwal, A. Pryhuber, and R. R. Thomas. "Ideals of the Multiview Variety". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 43.04 (Apr. 2021), pp. 1279–1292. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2019.2950631.

[Aga+23] Sameer Agarwal, Timothy Duff, Max Lieblich, and Rekha R Thomas. "An Atlas for the Pinhole Camera". In: *Foundations of Computational Mathematics* (2023), pp. 1–51.

[AST13] Chris Aholt, Bernd Sturmfels, and Rekha Thomas. "A Hilbert scheme in computer vision". In: *Canad. J. Math.* 65.5 (2013), pp. 961–988. ISSN: 0008-414X. DOI: 10.4153/CJM-2012-023-2. URL: https://doi-org.offcampus.lib.washington.edu/10.4153/CJM-2012-023-2.

[Boo12] Adam Boocher. "Free resolutions and sparse determinantal ideals". In: *Math. Res. Lett.* 19.4 (2012), pp. 805–821. ISSN: 1073-2780. DOI: 10.4310/MRL.2012.v19.n4.a6. URL: https://doi-org.offcampus.lib.washington.edu/10.4310/MRL.2012.v19.n4.a6.

[Bre+23] Paul Breiding, Felix Rydell, Elima Shehu, and Angélica Torres. "Line multiview varieties". In: *SIAM Journal on Applied Algebra and Geometry* 7.2 (2023), pp. 470–504.

[BT18] Paul Breiding and Sascha Timme. "HomotopyContinuation.jl: A package for homotopy continuation in Julia". In: *International Congress on Mathematical Software*. Springer. 2018, pp. 458–465.

[CW98] Stefan Carlsson and Daphna Weinshall. "Dual Computation of Projective Shape and Camera Positions from Multiple Images". In: *International Journal of Computer Vision* 27.3 (1998), pp. 227–241. DOI: 10.1023/A:1007961913417. URL: https://doi.org/10.1023/A:1007961913417.

[CCM23] Yairon Cid-Ruiz, Oliver Clarke, and Fatemeh Mohammadi. "A study of nonlinear multiview varieties". In: *Journal of Algebra* 620 (2023), pp. 363–391. ISSN: 0021-8693.

[Cif21] Diego Cifuentes. "A convex relaxation to compute the nearest structured rank deficient matrix". In: *SIAM Journal on Matrix Analysis and Applications* 42.2 (2021), pp. 708–729.

[CDG22] Aldo Conca, Emanuela De Negri, and Elisa Gorla. "Radical generic initial ideals". In: *Vietnam Journal of Mathematics* 50.3 (2022), pp. 807–827. ISSN: 2305-221X.

[CDL23] Erin Connelly, Timothy Duff, and Jessie Loucks-Tavitas. "Algebra and Geometry of Camera Resectioning". In: *arXiv preprint arXiv:2309.04028* (2023). Submitted to Journal of Algebra.

[CLO15] David A. Cox, John Little, and Donal O'Shea. *Ideals, varieties, and algorithms*. Fourth. Undergraduate Texts in Mathematics. An introduction to computational algebraic geometry and commutative algebra. Springer, Cham, 2015, pp. xvi+646. ISBN: 978-3-319-16720-6.

[Dol03] Igor Dolgachev. *Lectures on invariant theory*. Vol. 296. London Mathematical Society Lecture Note Series. Cambridge University Press, Cambridge, 2003, pp. xvi+220. ISBN: 0-521-52548-9.

[Dra+16] Jan Draisma, Emil Horobeţ, Giorgio Ottaviani, Bernd Sturmfels, and Rekha R Thomas. "The Euclidean distance degree of an algebraic variety". In: *Foundations of Computational Mathematics* 16.1 (2016), pp. 99–149.

[Duf+19a] Timothy Duff, Cvetelina Hill, Anders Jensen, Kisun Lee, Anton Leykin, and Jeff Sommars. "Solving polynomial systems via homotopy continuation and monodromy". In: *IMA Journal of Numerical Analysis* 39.3 (2019), pp. 1421–1446.

[Duf+20] Timothy Duff, Kathlén Kohn, Anton Leykin, and Tomas Pajdla. "PL1P—Point-line Minimal Problems under Partial Visibility in Three Views". In: *European Conference on Computer Vision* (2020).

[Duf+19b] Timothy Duff, Kathlén Kohn, Anton Leykin, and Tomas Pajdla. "PLMP-Point-Line Minimal Problems in Complete Multi-View Visibility". In: *IEEE International Conference on Computer Vision*. 2019, pp. 1675–1684.

[Duf+22] Timothy Duff, Viktor Korotynskiy, Tomas Pajdla, and Margaret H. Regan. "Galois/monodromy groups for decomposing minimal problems in 3D reconstruction". In: *SIAM Journal on Applied Algebra and Geometry* 6.4 (2022), pp. 740–772. ISSN: 2470-6566. DOI: 10.1137/21M1422872.

[EH16] David Eisenbud and Joe Harris. *3264 and All That: A Second Course in Algebraic Geometry*. Cambridge University Press, 2016. DOI: 10.1017/CBO9781139062046.

[FM95] Olivier D. Faugeras and Bernard Mourrain. "On the Geometry and Algebra of the Point and Line Correspondences Between N Images". In: *IEEE International Conference on Computer Vision*. 1995, pp. 951–956. DOI: 10.1109/ICCV.1995.466832. URL: https://doi.org/10.1109/ICCV.1995.466832.

[FKO18] Gunnar Fløystad, Joe Kileel, and Giorgio Ottaviani. "The Chow form of the essential variety in computer vision". In: *Journal of Symbolic Computation* 86 (2018), pp. 97–119.

[Ful84] W. Fulton. *Intersection Theory*. Ergebnisse der Mathematik und ihrer Grenzgebiete : a series of modern surveys in mathematics. Folge 3. Springer-Verlag, 1984. ISBN: 9783540121763. URL: https://books.google.com/books?id=cmoPAQAAMAAJ.

[GM82] M. Giusti and M. Merle. "Singularités isolées et sections planes de variétés déterminantielles. II. Sections de variétés déterminantielles par les plans de coordonnées". In: Lecture Notes in Math. 961 (1982), pp. 103–118. DOI: 10.1007/BFb0071278. URL: https://doi-org.offcampus.lib.washington.edu/10.1007/BFb0071278.

[GS] Daniel R. Grayson and Michael E. Stillman. *Macaulay2, a software system for research in algebraic geometry*. Available at http://www2.macaulay2.com.

[Gru41] J. A. Grunert. "Das Pothenotische Problem in erweiterter Gestalt nebst Über seine Anwendungen in der Geodäsie". In: *Grunerts Archiv für Mathematik und Physik* 1 (1841), pp. 238–248.

[HL18] Corey Harris and Daniel Lowengrub. "The Chern-Mather class of the multiview variety". In: *Communications in Algebra* 46.6 (2018), pp. 2488–2499. ISSN: 0092-7872. DOI: 10.1080/00927872.2017.1392545. eprint: https://doi.org/10.1080/00927872.2017.1392545. URL: https://doi.org/10.1080/00927872.2017.1392545.

[HZ04] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004. ISBN: 9780511811685. DOI: 10.1017/cbo9780511811685. URL: https://doi.org/10.1017/cbo9780511811685.

[HZ03] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Second. With a foreword by Olivier Faugeras. Cambridge University Press, Cambridge, 2003, pp. xvi+655. ISBN: 0-521-54051-8.

[HD00] Richard I Hartley and Nicolas Y Dano. "Reconstruction from six-point sequences". In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*. Vol. 2. IEEE. 2000, pp. 480–486.

[Hes63] O. Hesse. "Die cubische Gleichung, von welcher die Lösung des Problems der Homographie von M. Chasles abhängt". In: *Journal für die Reine und Angewandte Mathematik.* 62 (1863), pp. 188–192. ISSN: 0075-4102.

[HÅ97] Anders Heyden and Kalle Åström. "Algebraic properties of multilinear constraints". In: *Math. Methods Appl. Sci.* 20.13 (1997), pp. 1135–1162. ISSN: 0170-4214. DOI: 10.1002/(SICI)1099-1476(19970910)20:13<1135::AID-MMA908>3.0.CO;2-9. URL: https://doi-org.offcampus.lib.washington.edu/10.1002/(SICI)1099-1476(19970910)20:13%3C1135::AID-MMA908%3E3.0.CO;2-9.

[HJ94] Roger A. Horn and Charles R. Johnson. *Topics in matrix analysis*. Corrected reprint of the 1991 original. Cambridge University Press, Cambridge, 1994, pp. viii+607. ISBN: 0-521-46713-6.

[KK22] Joe Kileel and Kathlén Kohn. "Snapshot of algebraic vision". In: *arXiv preprint arXiv:2210.11443* (2022). DOI: 10.48550/ARXIV.2210.11443. URL: https://arxiv.org/abs/2210.11443.

[KBP08] Zuzana Kukelova, Martin Bujnak, and Tomas Pajdla. "Automatic generator of minimal problem solvers". In: *Computer Vision–ECCV 2008* (2008), pp. 302–315.

[LAO17] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. "Efficient solvers for minimal problems by syzygy-based reduction". In: *IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 820–829.

[Lar+18] Viktor Larsson, Magnus Oskarsson, Kalle Åström, Alge Wallis, Zuzana Kukelova, and Tomás Pajdla. "Beyond Gröbner Bases: Basis Selection for Minimal Solvers". In: *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. Computer Vision Foundation / IEEE Computer Society, 2018, pp. 3945–3954. DOI: 10.1109/CVPR.2018.00415. URL: http://openaccess.thecvf.com/content%5C_cvpr%5C_2018/html/Larsson%5C_Beyond%5C_Grobner%5C_Bases%5C_CVPR%5C_2018%5C_paper.html.

[Li18] Binglin Li. "Images of rational maps of projective spaces". In: *Int. Math. Res. Not. IMRN* 13 (2018), pp. 4190–4228. ISSN: 1073-7928. DOI: 10.1093/imrn/rnx003. URL: https://doi.org/10.1093/imrn/rnx003.

[Li09] Li Li. "Wonderful compactification of an arrangement of subvarieties". In: *Michigan mathematical journal* 58.2 (2009), pp. 535–563.

[LV20] Max Lieblich and Lucas Van Meter. "Two Hilbert Schemes in Computer Vision". In: *SIAM Journal on Applied Algebra and Geometry* 4.2 (2020), pp. 297–321. DOI: 10.1137/18M1200117. eprint: https://doi.org/10.1137/18M1200117. URL: https://doi.org/10.1137/18M1200117.

[Mac74] Robert D MacPherson. "Chern classes for singular algebraic varieties". In: *Annals of Mathematics* 100.2 (1974), pp. 423–432.

[MRW20] Laurentiu G Maxim, Jose I Rodriguez, and Botong Wang. "Euclidean distance degree of the multiview variety". In: *SIAM Journal on Applied Algebra and Geometry* 4.1 (2020), pp. 28–48. DOI: 10.1137/18M1233406. URL: https://doi.org/10.1137/18M1233406.

[MRW21] Laurentiu G. Maxim, Jose Israel Rodriguez, and Botong Wang. "Euclidean distance degree of projective varieties". In: *International Mathematics Research Notices.* 20 (2021), pp. 15788–15802. ISSN: 1073-7928,1687-0247. DOI: 10.1093/imrn/rnz266. URL: https://doi.org/10.1093/imrn/rnz266.

[Por23] Daniel Camazón Portela. "The Chow ring of a sequence of point blow-ups". In: *arXiv preprint arXiv:2301.06737* (2023).

[Qua95] Long Quan. "Invariants of six points and projective reconstruction from three uncalibrated images". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17.1 (1995), pp. 34–46. DOI: 10.1109/34.368154.

[Sch+00] Frederik Schaffalitzky, Andrew Zisserman, Richard I. Hartley, and Philip H. S. Torr. "A Six Point Solution for Structure and Motion". In: *European Conference on Computer Vision*. Springer, 2000, pp. 632–648. DOI: 10.1007/3-540-45054-8_41. URL: https://doi.org/10.1007/3-540-45054-8%5C_41.

[THP19] Matthew Trager, Martial Hebert, and Jean Ponce. "Coordinate-Free Carlsson-Weinshall Duality and Relative Multi-View Geometry". In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 2019, pp. 225–233. DOI: 10.1109/CVPR.2019.00031. URL: http://openaccess.thecvf.com/content%5C_CVPR%5C_2019/html/Trager%5C_Coordinate-Free%5C_Carlsson-Weinshall%5C_Duality%5C_and%5C_Relative%5C_Multi-View%5C_Geometry%5C_CVPR%5C_2019%5C_paper.html.

[Van19] Lucas Van Meter. "A Functorial Approach to Algebraic Vision". PhD thesis. University of Washington, 2019.