

Ideals of the Multiview Variety

Sameer Agarwal¹, Andrew Pryhuber, and Rekha R. Thomas

Abstract—The multiview variety of an arrangement of cameras is the Zariski closure of the images of world points in the cameras. The prime vanishing ideal of this complex projective variety is called the multiview ideal. We show that the bifocal and trifocal polynomials from the cameras generate the multiview ideal when the foci are distinct. In the computer vision literature, many sets of (determinantal) polynomials have been proposed to describe the multiview variety. We establish precise algebraic relationships between the multiview ideal and these various ideals. When the camera foci are noncoplanar, we prove that the ideal of bifocal polynomials saturate to give the multiview ideal. Finally, we prove that all the ideals we consider coincide when dehomogenized, to cut out the space of finite images.

Index Terms—Projective geometry, structure-from-motion, nonlinear algebra

1 INTRODUCTION

A general projective camera is a rank three matrix in $\mathbb{R}^{3 \times 4}$. Given a camera arrangement $\mathcal{A} = (A_1, \dots, A_n)$, the image formation map

$$\varphi_{\mathcal{A}} : \mathbb{P}_{\mathbb{R}}^3 \dashrightarrow (\mathbb{P}_{\mathbb{R}}^2)^n,$$

sends a homogenized world point $\mathbf{q} \in \mathbb{P}_{\mathbb{R}}^3$ to its images $(\mathbf{p}_1 = A_1 \mathbf{q}, \dots, \mathbf{p}_n = A_n \mathbf{q}) \in (\mathbb{P}_{\mathbb{R}}^2)^n$. The i th copy of $\mathbb{P}_{\mathbb{R}}^2$ in the codomain of $\varphi_{\mathcal{A}}$ is the homogenized image plane of camera i . The unique point $\mathbf{c}_i \in \mathbb{P}_{\mathbb{R}}^3$ in the kernel of A_i is the focal point of camera i . The map $\varphi_{\mathcal{A}}$ is defined at all points in $\mathbb{P}_{\mathbb{R}}^3$ except at the foci $\mathbf{c}_1, \dots, \mathbf{c}_n$. Triggs called $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ the *joint image* [1] and Heyden-Åström call it the *natural descriptor* [2]. We are interested in studying the complete set of polynomials that vanish on $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$.

Definition 1.1. Given a set $S \subseteq \mathbb{P}_{\mathbb{C}}^{d-1}$, the collection of all polynomials in $\mathbb{C}[x_1, \dots, x_d]$ that vanish on S is a homogeneous ideal, known as the vanishing ideal of S , and denoted as $\mathbf{I}(S)$. The variety $\mathbf{V}(\mathbf{I}(S))$ is the smallest complex projective variety that contains S , known as the Zariski closure of S .

We refer the reader to [3] for the basics on ideals and varieties. In this paper we will be interested in the vanishing ideal of the joint image $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$.

Definition 1.2. The multiview ideal of $\mathcal{A}$, denoted $M_{\mathcal{A}}$, is the vanishing ideal of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ in $\mathbb{C}[p_1, \dots, p_n]$ where $p_i = (x_i, y_i, z_i)$ are the coordinates on the i th copy of $\mathbb{P}_{\mathbb{C}}^2$. The Zariski closure of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ in $(\mathbb{P}_{\mathbb{C}}^2)^n$ is the complex projective variety $\mathbf{V}(M_{\mathcal{A}})$, which we call the multiview variety of $\mathcal{A}$.

- S. Agarwal is with Google Inc., Mountain View, CA 94043 USA. E-mail: sameeragarwal@google.com.
- A. Pryhuber and R.R. Thomas are with the Department of Mathematics, University of Washington, Seattle, WA 98105 USA. E-mail: {pryhuber, rrrthomas}@u.w.edu.

Manuscript received 9 Apr. 2019; revised 12 Sept. 2019; accepted 22 Oct. 2019. Date of publication 31 Oct. 2019; date of current version 4 Mar. 2021.

(Corresponding author: Sameer Agarwal.)

Recommended for acceptance by H. Li.

Digital Object Identifier no. 10.1109/TPAMI.2019.2950631

The terminology *multiview ideal* and *multiview variety* comes from [4]. Following Triggs [1], Trager et al. refer to the multiview variety as the *joint image variety*.

Starting with the seminal work of Longuet-Higgins [5], researchers have studied various systems of polynomials that vanish on $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$. In the computer vision literature these equations are known as *multiview constraints* [2], [6], [7], [8], [9]. Obviously, the ideals generated by these systems of polynomials are contained in $M_{\mathcal{A}}$. However, there hasn't been much discussion of whether these polynomials generate $M_{\mathcal{A}}$ since the focus of all these papers has been on the multiview variety and not its vanishing ideal. The aim of this paper is to provide a complete description of the multiview ideal and study its relationship to the above sets of polynomials.

It can be difficult to determine the vanishing ideal of a variety. However, there are various advantages to knowing it. To be able to do any computations with a variety or to study its structure using algebra, we need a description in terms of polynomials and the vanishing ideal is the optimal algebraic description. This manifests itself in a number of ways.

The set of all polynomial functions on X is precisely $\mathbb{C}[x_1, \dots, x_d]/\mathbf{I}(X)$, known as the coordinate ring of X . In particular, a polynomial g vanishes on X if and only if g belongs to $\mathbf{I}(X)$. Knowledge of a generating set $\{g_1, \dots, g_k\}$ of $\mathbf{I}(X)$ also informs us about the local structure of X , since a point $x \in X$ is smooth if and only if the Jacobian matrix $(\frac{\partial g_i}{\partial x_j})$ has rank equal to the codimension of X . More generally, if $X \subset \mathbb{P}_{\mathbb{C}}^{d-1}$ is a projective variety then $\mathbf{I}(X)$ carries all the geometric information about X allowing algebra (and algebraic algorithms) to infer geometric properties of X . For example, the dimension and degree of X can be read off from the *Hilbert polynomial* of $\mathbf{I}(X)$ which also carries many more sophisticated invariants of X . See [3] for all the above.

In multiview geometry, many estimation problems can be phrased as polynomial optimization problems over varieties [4], [10]. In particular, the triangulation problem under Gaussian noise amounts to projecting a point onto the multiview variety [11].

In general, polynomial optimization on a variety $X \subseteq \mathbb{R}^n$ boils down to certifying the non-negativity of a polynomial

f on X by expressing it as a sum-of-squares (sos) modulo an ideal J vanishing on X [12]. This means finding a sos polynomial $s = \sum p_i^2$ such that $f - s$ lies in J . This expressibility is maximized, and the algorithms terminate in the lowest possible degree, when $J = \mathbf{I}(X)$. We illustrate this on a very small example.

Example 1.3. The polynomial $x + 1$ is non-negative on $X = \{0\} \subset \mathbb{R}$. The ideal $\langle x^2 \rangle$ cuts out X but $\mathbf{I}(X) = \langle x \rangle$. Now $(x + 1) - 1 \in \langle x \rangle$ allowing $s = 1$ as the sos certificate. On the other hand, if $x + 1 - s \in \langle x^2 \rangle$ then s has to have degree at least 2; for instance $(x + 1) - (1 + \frac{1}{2}x)^2 \in \langle x^2 \rangle$.

The above phenomenon can have a major impact on the number of rounds of convex relaxations needed to solve a polynomial optimization problem such as the well-known *Lasserre/sos hierarchies* [13], [14], where each round looks for sos certificates of a fixed degree with degrees increasing monotonically with rounds. In each round the semidefinite program being solved is of size $O(n^d)$, where n is the number of variables and d is degree in that round. As a result, in many cases only the first round maybe computationally feasible and having access to $\mathbf{I}(X)$ can make the difference between the problem being tractable or not.

The rest of the paper is structured as follows. After a brief discussion of the notation used in this paper we begin in Section 2 by introducing a family of ideals associated with every camera arrangement $\mathcal{A}$ which we call the k -focal ideals. We describe how these ideals behave under change of coordinates, and dispel the popular myth that, under a change of image coordinates, k -focal polynomials go to k -focal polynomials. In Section 3, we prove our first main theorem (Theorem 3.7), that the well-known bifocal (epipolar constraints) and trifocal polynomials generate $M_{\mathcal{A}}$ when the camera foci in $\mathcal{A}$ are distinct. Next, in Section 4, we consider three different types of determinantal polynomials proposed to cut out the multiview variety by Heyden-Åström [2], Faugeras et al. [7] and Ma et al. [9]. We show that while the ideals they generate are all contained in $M_{\mathcal{A}}$, none of them actually coincide with $M_{\mathcal{A}}$. We establish their precise algebraic relationship with $M_{\mathcal{A}}$. In Section 5, we consider the relationship of the multiview ideal to bifocal polynomials and prove the algebraic analog of the statement that the bifocal polynomials cut out the multiview variety when the camera foci are noncoplanar. In Section 6, we study how the various ideals relate to each other when we restrict our attention to finite images, i.e., exclude points at infinity. We conclude in Section 7 with a summary.

Many results in this paper require explicit computation. We recommend the reader have a copy of Macaulay2 [15] (or equivalent symbolic algebra software) handy. The Macaulay2 codes for our computations can be found at https://sites.math.washington.edu/~thomas/papers/Multiview_Ideal.zip

1.1 Notation

In the rest of the paper, we will use $\mathbb{P}$ to denote $\mathbb{P}_{\mathbb{C}}$. The ideal generated by the polynomials $f_1, \dots, f_s$ will be denoted as $\langle f_1, \dots, f_s \rangle$.

We will use A for cameras and G for matrices in GL_n . $\mathcal{A}$ and $\mathcal{G}$ will denote arrangements of corresponding matrices. Bold, lower-case roman letters will be used to indicate

vectors, and lower-case greek letters will be used for functions. Given a partial symbolic matrix M , $\text{minors}(k, M)$ will denote the ideal generated by all $k \times k$ minors of the matrix M . The symbol $[n]$ denotes the set $\{1, \dots, n\}$ and $\binom{[n]}{m}$ denotes the set of all size m subsets of $[n]$.

2 THE k -FOCAL IDEALS OF A CAMERA ARRANGEMENT

Let p_i be the tuple of variables (x_i, y_i, z_i) denoting the coordinates associated to the projective plane $\mathbb{P}_{\mathbb{R}}^2$ corresponding to the i th camera image. Write $p = (p_1, \dots, p_n)$, and consider the partially symbolic matrix

$$\mathcal{A}(p) := \begin{bmatrix} A_1 & p_1 & & & \\ A_2 & & p_2 & & \\ \vdots & & & \ddots & \\ A_n & & & & p_n \end{bmatrix}. \quad (1)$$

Let $\mathcal{A}(p)$ denote the evaluation of $\mathcal{A}(p)$ at $p = \mathbf{p}$. If $\mathbf{p} := (p_1, \dots, p_n) \in \varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ then there exists some $\mathbf{q} \in \mathbb{P}_{\mathbb{R}}^3$ and scalars $\lambda_i \in \mathbb{R}$ such that $A_i \mathbf{q} = \lambda_i p_i$ for all $i = 1, \dots, n$. Therefore, $\mathcal{A}(p)$ has a non-trivial kernel since it contains the point $(\mathbf{q}, -\lambda_1, \dots, -\lambda_n)$, and hence the maximal minors of $\mathcal{A}(p)$, which are polynomials in $p_1, \dots, p_n$, vanish on $\mathbf{p}$. Since $\mathbf{p}$ was arbitrary, these maximal minors vanish on all of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ and on the multiview variety. Therefore

$$\text{minors}(4 + n, \mathcal{A}(p)) \subseteq M_{\mathcal{A}}.$$

In this section, we describe further minors of $\mathcal{A}(p)$ and the ideals they generate, which will play an important role in the description of $M_{\mathcal{A}}$.

Definition 2.1. For a subset $\sigma = \{\sigma_1, \dots, \sigma_k\} \subseteq [n]$ where $k \geq 2$, consider the partially symbolic matrix

$$\mathcal{A}_{\sigma}(p) = \begin{pmatrix} A_{\sigma_1} & p_{\sigma_1} & 0 & \dots & 0 \\ A_{\sigma_2} & 0 & p_{\sigma_2} & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{\sigma_k} & 0 & \dots & 0 & p_{\sigma_k} \end{pmatrix}, \quad (2)$$

of size $3k \times (4 + k)$. A maximal $(4 + k) \times (4 + k)$ minor of $\mathcal{A}_{\sigma}(p)$ is called a k -focal polynomial of $\mathcal{A}$. The k -focal ideal of $\mathcal{A}$, $H_{\mathcal{A}}^k$, is the ideal sum

$$H_{\mathcal{A}}^k = \sum_{\sigma \in \binom{[n]}{k}} \text{minors}(4 + k, \mathcal{A}_{\sigma}(p)).$$

Trager et al. also study the k -focal polynomials and refer to them as k -linearities [16], [17]. Note that every k -focal polynomial is multilinear and of total degree k . Such a minor involves choosing $4 + k$ rows of $\mathcal{A}_{\sigma}(p)$, and by a pigeonhole argument, at most four cameras may contribute more than one row to the minor when $k > 4$. Indeed, if more than four cameras contributed at least two rows each, then at least 10 rows are accounted for, which leaves at most $k - 6$ rows to take from the remaining $k - 5$ cameras. So at least one camera will be left out entirely which means

that the submatrix of that $4 + k$ minor has a zero column and the minor is zero.

A useful fact for us will be that for two positive integers $l > k \geq 2$, there is a simple way to “bump up” a k -focal polynomial to an l -focal polynomial by multiplying the k -focal polynomial with a monomial.

Lemma 2.2. *Suppose f is a k -focal polynomial from cameras $\sigma = \{\sigma_1, \dots, \sigma_k\} \subset [n]$ where $k \geq 2$. For any $l > k$ cameras $\tau = \{\sigma_1, \dots, \sigma_k, \tau_1, \dots, \tau_{l-k}\}$, there is a l -focal polynomial g such that $(\prod_{i=1}^{l-k} w_{\tau_i})f = g$ for any choice of variables $w_{\tau_i} \in \{x_{\tau_i}, y_{\tau_i}, z_{\tau_i}\}$, one for each camera.*

Proof. Add the row and column associated to coordinate w_{τ_i} to $\mathcal{A}_\sigma(p)$ for $\tau_1, \dots, \tau_{l-k}$ as follows:

$$\begin{pmatrix} A_{\sigma_1} & p_{\sigma_1} & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{\sigma_k} & 0 & \dots & p_{\sigma_k} & 0 & \dots & 0 \\ (A_{\tau_1})_{w_{\tau_1}} & 0 & \dots & 0 & w_{\tau_1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (A_{\tau_{l-k}})_{w_{\tau_{l-k}}} & 0 & \dots & 0 & 0 & \dots & w_{\tau_{l-k}} \end{pmatrix}.$$

Taking the determinant of this matrix yields the l -focal polynomial $g = (\prod_{i=1}^{l-k} w_{\tau_i})f$. $\square$

Combining the above facts we get that any l -focal polynomial for $l > 4$ is of the form $(\prod_{i=1}^{l-k} w_{\tau_i})f$ where f is a $k \leq 4$ focal polynomial. This is a generalization of Proposition 2 in [16] that showed that every n -focal polynomial is a monomial multiple of a k -focal polynomial for $k \leq 4$. As a result, we will primarily focus on the ideals $H_{\mathcal{A}}^2$, $H_{\mathcal{A}}^3$, and $H_{\mathcal{A}}^4$, called the *bifocal*, *trifocal*, and *quadrifocal* ideals of $\mathcal{A}$.

A closer look at $H_{\mathcal{A}}^2$ reveals that it is the ideal generated by the $\binom{n}{2}$ epipolar constraints, since $\mathcal{A}_{\{i,j\}}$ is a 6×6 matrix, whose determinant is the epipolar constraint between images i and j . By Lemma 2.2, $H_{\mathcal{A}}^3$ contains the bumped up version of $H_{\mathcal{A}}^2$ and for every triplet of images $\{i, j, k\}$, the 27 trifocals implied by the three trifocal tensors relating them. And finally, $H_{\mathcal{A}}^4$ contains the bumped up versions of $H_{\mathcal{A}}^2$ and $H_{\mathcal{A}}^3$ and the 81 quadrifocals implied by the quadrifocal tensor. The fact that we only need to study $H_{\mathcal{A}}^2$, $H_{\mathcal{A}}^3$, and $H_{\mathcal{A}}^4$ lines up with the well known fact in multiview geometry that when studying n -view constraints, one only needs to study the epipolar matrix, the trifocal tensor and the quadrifocal tensor. See Chapter 17 in the book by Hartley & Zisserman [8] for explicit computations of the generators of $H_{\mathcal{A}}^2$, $H_{\mathcal{A}}^3$, and $H_{\mathcal{A}}^4$ and their history.

In the remainder of this section, we will investigate how k -focal ideals transform under certain linear transformations on cameras. It is widely known that, from image data, the geometry of a camera arrangement can only be determined up to an arbitrary choice of $\mathbb{P}^3$ coordinates. This is reflected in the following lemma.

Lemma 2.3 (Projective Ambiguity). *Suppose $G \in \text{GL}_4$. Then for any k , $H_{\mathcal{A}}^k = H_{\mathcal{A}G}^k$ where $\mathcal{A}G = (A_1G, A_2G, \dots, A_kG)$.*

Proof. This follows since $(\mathcal{A}G)_\sigma(p) = \mathcal{A}_\sigma(p)\text{diag}(G, I_k)$ for any k -element subset $\sigma \subset [n]$ which implies that any k -focal of $\mathcal{A}G$ differs from the same k -focal of $\mathcal{A}$ by a factor of $\det(G) \neq 0$. $\square$

From the proof of Lemma 2.3, we see that a $\mathbb{P}^3$ coordinate change that sends $\mathbf{q} \mapsto G\mathbf{q}$ maps k -focals to k -focals, picking up only a scalar factor $\det G \neq 0$. We will now see that change of coordinates on the image planes $\mathbb{P}^2$ affect the k -focals in a more subtle way.

Let $\mathcal{G} = (G_1, \dots, G_n) \in (\text{GL}_3)^n$ be a sequence of invertible matrices and consider the camera arrangement $\mathcal{G}\mathcal{A} := (G_1A_1, \dots, G_nA_n)$ obtained from a given arrangement $\mathcal{A}$ by left-multiplying A_i with G_i . Note that the focal point of the camera A_i is the same as the focal point of the camera G_iA_i . Since $p_i = (x_i, y_i, z_i)$, we denote the ring $\mathbb{C}[x_1, y_1, z_1, \dots, x_n, y_n, z_n]$ by $\mathbb{C}[p_1, \dots, p_n]$ and a polynomial in it by $f(p_1, \dots, p_n)$. The sequence $\mathcal{G}$ induces a camera-wise linear change of coordinates $\chi_{\mathcal{G}}$ on $\mathbb{C}[p_1, \dots, p_n]$ by sending

$$\chi_{\mathcal{G}} : \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} \mapsto G_i^{-1} \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}. \quad (3)$$

Note that this amounts to a change of coordinates in the image planes $\mathbb{P}^2$ of the cameras in $\mathcal{A}$. Let $G^{-1}p$ denote $\chi_{\mathcal{G}}(p) = (G_1^{-1}p_1, \dots, G_n^{-1}p_n)$. In what follows we will also need the notation $\mathcal{G}^{-1} := (G_1^{-1}, \dots, G_n^{-1})$, $\mathcal{G}^{-1}\mathcal{A} := (G_1^{-1}A_1, \dots, G_n^{-1}A_n)$ and $\chi_{\mathcal{G}^{-1}}(p_i) = G_i p_i$.

To analyze the effect of $\chi_{\mathcal{G}}$ on k -focal ideals, we recall the classical Cauchy-Binet formula, a proof of which can be found in [18].

Lemma 2.4 (Cauchy-Binet). *If A and B are rectangular matrices of size $m \times n$ and $n \times m$, respectively, where $m \leq n$, then the determinant of the square matrix AB is*

$$\det(AB) = \sum_{\sigma \in \binom{[n]}{m}} \det(A_{[:,\sigma]}) \det(B_{[\sigma,:]}),$$

where $:$ indicates that all rows/columns are taken.

Lemma 2.5. *For the k -focal ideal $H_{\mathcal{A}}^k$, $\chi_{\mathcal{G}}(H_{\mathcal{A}}^k) = H_{\mathcal{G}\mathcal{A}}^k$. Similarly, $\chi_{\mathcal{G}^{-1}}(H_{\mathcal{G}\mathcal{A}}^k) = H_{\mathcal{A}}^k$.*

Proof. We prove the first statement and the other follows similarly. We will show that the k -focal ideal of $\mathcal{A}_{[k]}$ is sent to the k -focal ideal of $(\mathcal{G}\mathcal{A})_{[k]}$. The result then follows for the full k -focal ideal $H_{\mathcal{A}}^k$ by summing the k -focal ideals of all $\mathcal{A}_\sigma$ as σ varies over all k -subsets of $[n]$.

Recall that a k -focal polynomial of $\mathcal{A}_{[k]} := (A_1, \dots, A_k)$ is a maximal minor of

$$\mathcal{A}_{[k]}(p) = \begin{bmatrix} A_1 & p_1 & & & \\ A_2 & & p_2 & & \\ \vdots & & & \ddots & \\ A_k & & & & p_k \end{bmatrix}.$$

Applying $\chi_{\mathcal{G}}$ to this maximal minor is the same as taking the same maximal minor of

$$\mathcal{A}_{[k]}(\chi_{\mathcal{G}}(p)) = \begin{bmatrix} A_1 & G_1^{-1}p_1 & & & \\ A_2 & & G_2^{-1}p_2 & & \\ \vdots & & & \ddots & \\ A_k & & & & G_k^{-1}p_k \end{bmatrix}.$$

The corresponding k -focal polynomial of $\mathcal{GA}$ is the same maximal minor of

$$(\mathcal{GA})_{[k]}(p) = \text{diag}(G_1, \dots, G_k) \mathcal{A}_{[k]}(p). \quad (4)$$

The ideal $\chi_{\mathcal{G}}(H_{\mathcal{A}_{[k]}}^k)$ is generated by the maximal minors of $\mathcal{A}_{[k]}(\chi_{\mathcal{G}}(p))$, namely

$$\left\{ \det(\mathcal{A}_{[k]}(G^{-1}p)_{[\sigma, \cdot]}) : \sigma \in \binom{[3k]}{4+k} \right\},$$

while $H_{(\mathcal{GA})_{[k]}}^k$ is generated by the maximal minors of $(\mathcal{GA})_{[k]}(p)$. We need to show that these ideals coincide.

Let G denote the block diagonal matrix with blocks $G_1, \dots, G_n$. A $(4+k)$ -minor of $(\mathcal{GA})_{[k]}(p)$ is the determinant of a submatrix with $4+k$ rows indexed by some $\tau \in \binom{[3k]}{4+k}$. Such a submatrix has the form $G_{\tau} \mathcal{A}_{[k]}(G^{-1}p)$ where G_{τ} is the submatrix of G consisting of the rows of G indexed by τ . By the Cauchy-Binet formula

$$\begin{aligned} \det(G_{\tau} \mathcal{A}_{[k]}(G^{-1}p)) \\ = \sum_{\sigma \in \binom{[3k]}{4+k}} \det((G_{\tau})_{[\cdot, \sigma]}) \det(\mathcal{A}_{[k]}(G^{-1}p)_{[\sigma, \cdot]}). \end{aligned}$$

This implies that $\det(G_{\tau} \mathcal{A}_{[k]}(G^{-1}p))$ lies in the ideal $\chi_{\mathcal{G}}(H_{\mathcal{A}_{[k]}}^k)$, and hence, $H_{(\mathcal{GA})_{[k]}}^k \subseteq \chi_{\mathcal{G}}(H_{\mathcal{A}_{[k]}}^k)$.

The reverse containment follows by applying the same argument to $\mathcal{A}_{[k]}(p) = G^{-1}G \mathcal{A}_{[k]}(p)$ and $G \mathcal{A}_{[k]}(p)$ where G^{-1} is the block diagonal matrix with blocks $G_1^{-1}, \dots, G_n^{-1}$.

Summing over all k camera subsets, the result follows:

$$\begin{aligned} \chi_{\mathcal{G}}(H_{\mathcal{A}}^k) &= \chi_{\mathcal{G}} \left(\sum_{\sigma \in \binom{[n]}{k}} H_{\mathcal{A}_{\sigma}}^k \right) = \sum_{\sigma \in \binom{[n]}{k}} \chi_{\mathcal{G}}(H_{\mathcal{A}_{\sigma}}^k) \\ &= \sum_{\sigma \in \binom{[n]}{k}} H_{(\mathcal{GA})_{\sigma}}^k = H_{\mathcal{GA}}^k. \end{aligned}$$

□

This proof shows that, contrary to popular belief, it is not true that k -focal polynomials go to k -focal polynomials under the change of coordinates given by $\chi_{\mathcal{G}}$, but the ideals do as in Lemma 2.5.

3 THE MULTIVIEW IDEAL

Recall from Definition 1.2 that the multiview ideal $M_{\mathcal{A}}$ of the camera arrangement $\mathcal{A}$ is the vanishing ideal of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$, meaning that it is the set of all polynomials in $\mathbb{C}[p_1, \dots, p_n]$ that vanish on $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$. Since $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ is a subset of $(\mathbb{P}_{\mathbb{R}}^2)^n$, $M_{\mathcal{A}}$ is, in fact, generated by polynomials with real coefficients.¹

1. Let $h(x) = f(x) + ig(x)$ be a complex polynomial, where $f(x)$ and $g(x)$ are real polynomials. Then if $h(x)$ vanish on a set of real points, then so must $f(x)$ and $g(x)$.

The complex projective variety $V(M_{\mathcal{A}}) \subset (\mathbb{P}^2)^n$, which is the complex Zariski closure of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$, is the multiview variety of $\mathcal{A}$. One might wonder if it is better to study the real Zariski closure of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$ and its vanishing ideal since complex points in the multiview variety do not have any physical meaning, and hence no relevance to multiview geometry. However, observe that if the real Zariski closure was strictly smaller than the set of real points in $V(M_{\mathcal{A}})$, then there would be a polynomial not in $M_{\mathcal{A}}$ that vanishes on $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$, which would contradict that $M_{\mathcal{A}}$ is the vanishing ideal of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$. Therefore, $M_{\mathcal{A}}$ is also the vanishing ideal of the real Zariski closure of $\varphi_{\mathcal{A}}(\mathbb{P}_{\mathbb{R}}^3)$, and hence a *real radical ideal* [19, Section 12.5].

Further, since $\varphi_{\mathcal{A}}$ is a polynomial map and $\mathbb{P}_{\mathbb{R}}^3$ is irreducible, $V(M_{\mathcal{A}})$ is an irreducible three-dimensional variety in $(\mathbb{P}^2)^n$. Hence $M_{\mathcal{A}}$ is a *prime* (homogeneous) ideal, meaning that if $fg \in M_{\mathcal{A}}$ then either f or g is in $M_{\mathcal{A}}$.

It was shown in [4] that the bifocals, trifocals and quadrifocals of $\mathcal{A}$ form a *universal Gröbner basis* of $M_{\mathcal{A}}$ under a certain genericity assumption on the cameras. This means that this collection of polynomials form a *Gröbner basis* for $M_{\mathcal{A}}$ with respect to any term order [3]. We will use this result to establish a generating set for $M_{\mathcal{A}}$ when the camera foci are distinct.

We first note what happens to $M_{\mathcal{A}}$ under the change of coordinates $\chi_{\mathcal{G}}$ defined in the previous section. Recall that $\chi_{\mathcal{G}}$ sends a polynomial $f(p_1, \dots, p_n) \in \mathbb{C}[p_1, \dots, p_n]$ to $f(G_1^{-1}p_1, \dots, G_n^{-1}p_n)$.

Lemma 3.1. *The image of the multiview ideal $M_{\mathcal{A}}$ under the map $\chi_{\mathcal{G}}$ is $M_{\mathcal{GA}}$, the multiview ideal of $\mathcal{GA}$. i.e., $\chi_{\mathcal{G}}(M_{\mathcal{A}}) = M_{\mathcal{GA}}$. Similarly, $\chi_{\mathcal{G}^{-1}}(M_{\mathcal{GA}}) = M_{\mathcal{A}}$.*

Proof. Again, we will prove that $\chi_{\mathcal{G}}(M_{\mathcal{A}}) = M_{\mathcal{GA}}$. The proof that $\chi_{\mathcal{G}^{-1}}(M_{\mathcal{GA}}) = M_{\mathcal{A}}$ is similar.

From the definition we see that a polynomial $f(p_1, \dots, p_n)$ vanishes on the multiview variety $V(M_{\mathcal{A}})$ if and only if $f(A_1\mathbf{q}, \dots, A_n\mathbf{q}) = 0$ for all $\mathbf{q} \in \mathbb{P}^3 \setminus \{c_1, \dots, c_n\}$, equivalently, if and only if

$$f(G_1^{-1}(G_1A_1\mathbf{q}), \dots, G_n^{-1}(G_nA_n\mathbf{q})) = 0,$$

for all $\mathbf{q} \in \mathbb{P}^3 \setminus \{c_1, \dots, c_n\}$. The multiview variety of $\mathcal{GA}$ is the Zariski closure of the points $(G_1A_1\mathbf{q}, \dots, G_nA_n\mathbf{q})$ as $\mathbf{q}$ varies over $\mathbb{P}^3 \setminus \{c_1, \dots, c_n\}$. Therefore, f vanishes on $V(M_{\mathcal{A}})$ if and only if $\chi_{\mathcal{G}}(f)$ vanishes on $V(M_{\mathcal{GA}})$. This proves that $\chi_{\mathcal{G}}(M_{\mathcal{A}}) \subseteq M_{\mathcal{GA}}$.

To finish the proof we need to argue that if $g(p_1, \dots, p_n) \in M_{\mathcal{GA}}$ then $g = \chi_{\mathcal{G}}(f)$ for some $f \in M_{\mathcal{A}}$. A polynomial $g \in M_{\mathcal{GA}}$ if and only if $g(G_1A_1\mathbf{q}, \dots, G_nA_n\mathbf{q}) = 0$ for all $\mathbf{q} \in \mathbb{P}^3 \setminus \{c_1, \dots, c_n\}$ if and only if $g(G_1\mathbf{p}_1, \dots, G_n\mathbf{p}_n) = 0$ for all $(\mathbf{p}_1, \dots, \mathbf{p}_n) \in V(M_{\mathcal{A}})$. Define $g(G_1p_1, \dots, G_np_n) = : f \in M_{\mathcal{A}}$. Then $\chi_{\mathcal{G}}(f) = g(p_1, \dots, p_n)$. □

We will use the results obtained so far to give an elementary proof that the bifocals and trifocals generate the multiview ideal $M_{\mathcal{A}}$ for any arrangement $\mathcal{A}$ of cameras with pairwise distinct foci. An important tool will be *translational* cameras.

Definition 3.2. *A camera T is said to be translational if its left 3×3 block is the identity matrix, i.e., $T = [I \ \mathbf{t}]$ for some $\mathbf{t} \in \mathbb{R}^3$.*

Lemma 3.3. *If $\mathcal{T}$ is an arrangement of translational cameras, then $H_{\mathcal{T}}^4 \subseteq H_{\mathcal{T}}^3$.*

Proof. Using Macaulay2, this statement can be checked for $n = 4$ translational cameras with foci represented symbolically as $(t_{i1}, t_{i2}, t_{i3}, -1)$. For $n \geq 4$, since $H_{\mathcal{T}}^4 = \sum_{\sigma \in \binom{[n]}{4}} H_{\mathcal{T}\sigma}^4$ and $H_{\mathcal{T}}^3 = \sum_{\sigma \in \binom{[n]}{3}} H_{\mathcal{T}\sigma}^3$, the statement follows. $\square$

We now use translational cameras to show that the quad-rifocals are not needed in a generating set of $M_{\mathcal{A}}$. This is done by extending the result for translational cameras to finite cameras. Recall that a *finite* camera is a camera whose left 3×3 block is invertible, or equivalently a camera whose focal point is not a point at infinity. Observe that any finite camera can be obtained by multiplying some translational camera on the left by an invertible 3×3 matrix.

Corollary 3.4. *If $\mathcal{A}$ is any arrangement of cameras, then $H_{\mathcal{A}}^4 \subseteq H_{\mathcal{A}}^3$.*

Proof. If $\mathcal{A}$ is an arrangement of finite cameras, then $A_i = G_i[I \ t_i]$ for some $G_i \in GL_3$. Therefore $\mathcal{A} = \mathcal{GT}$ where $\mathcal{T}$ is an arrangement of translational cameras. By Lemma 3.3, $H_{\mathcal{T}}^4 \subseteq H_{\mathcal{T}}^3$. Hence, Lemma 2.5 implies

$$H_{\mathcal{A}}^4 = H_{\mathcal{GT}}^4 = \chi_{\mathcal{G}}(H_{\mathcal{T}}^4) \subseteq \chi_{\mathcal{G}}(H_{\mathcal{T}}^3) = H_{\mathcal{GT}}^3 = H_{\mathcal{A}}^3.$$

For any four cameras indexed by $\sigma \in \binom{[n]}{4}$, there exists some $G \in GL_4$ which takes the foci of $\mathcal{A}_{\sigma}$ off of the plane at infinity, i.e., so that $\mathcal{A}_{\sigma}G$ is an arrangement of finite cameras. Inverting this $\mathbb{P}^3$ -coordinate change does not change ideal containment by Lemma 2.3. The general result follows since $H_{\mathcal{A}}^4 = \sum_{\sigma \in \binom{[n]}{4}} H_{\mathcal{A}\sigma}^4 \subseteq \sum_{\sigma \in \binom{[n]}{3}} H_{\mathcal{A}\sigma}^3 = H_{\mathcal{A}}^3$. $\square$

To get to our main result, we will need a result from [4] about camera arrangements $\mathcal{A}$ that are generic in the sense that all 4×4 minors of $[A_1^T \ A_2^T \ \cdots \ A_n^T]$ are non-zero. We call such an $\mathcal{A}$ *minor-generic*.

Corollary 3.5. *Suppose $\mathcal{A}$ is minor-generic. Then $M_{\mathcal{A}} = H_{\mathcal{A}}^2 + H_{\mathcal{A}}^3$.*

Proof. Theorem 2.1 in [4] says that if $\mathcal{A}$ is minor-generic, then the bifocals, trifocals and quad-rifocals form a universal Gröbner basis of $M_{\mathcal{A}}$. In particular, this implies that $M_{\mathcal{A}} = H_{\mathcal{A}}^2 + H_{\mathcal{A}}^3 + H_{\mathcal{A}}^4$. The statement is then immediate from Corollary 3.4. $\square$

Minor-genericity is a purely algebraic condition on camera arrangements. The following statement, which appears as a brief comment in [4] without proof, gives a geometric reinterpretation of this condition.

Lemma 3.6. *If $\mathcal{A}$ is minor-generic, then the foci of the cameras in $\mathcal{A}$ are pairwise distinct. Conversely, if the cameras in $\mathcal{A}$ have pairwise distinct foci, then there exist $G_i \in GL_3$ such that $\mathcal{GA}$ is minor-generic.*

Proof. Let $L_i \subset \mathbb{C}^4$ denote the three-dimensional row span of A_i . If A_i and A_j have the same focal point then $L_i = L_j$ and hence any four of the six rows of A_i and A_j are linearly dependent and $\mathcal{A}$ is not minor-generic. This proves the first statement.

Now suppose the foci of cameras in $\mathcal{A}$ are pairwise distinct. This means that the planes L_i are pairwise distinct. For any $G_i \in GL_3$, the rows of $G_i A_i$ form a basis of L_i . By choosing G_i appropriately, the three rows of A_i can be sent to any choice of three linearly independent vectors in L_i . We need to show that there is a choice of G_i such that no four rows from the matrices $G_i A_i$ are linearly dependent.

Consider the $3n \times 4$ matrix obtained by vertically stacking the cameras in $\mathcal{A}$, as a point in $(\mathbb{C}^4)^{3n}$, with coordinates x_{kl}^i representing the (k, l) -entry of the i th camera. We will identify this point in $(\mathbb{C}^4)^{3n}$ with the corresponding $3n \times 4$ matrix, and stack of n cameras, and call all of them $\mathcal{A}$. Let $\mathcal{A}(x)$ denote the symbolic $3n \times 4$ matrix with entries x_{kl}^i . For $\sigma \in \binom{[3n]}{4}$, let d_{σ} denote the determinant of the 4×4 submatrix of $\mathcal{A}(x)$ with rows indexed by σ . These cut out $\binom{3n}{4}$ quartic hypersurfaces $V(d_{\sigma})$ in $(\mathbb{C}^4)^{3n}$. Let v_i denote the normal of the hyperplane $L_i \subset \mathbb{C}^4$. Impose linear conditions saying that the rows of $\mathcal{A}(x)$, numbered $3i, 3i+1, 3i+2$, dot to zero with v_i . These $3n$ equations determine a subspace L in $(\mathbb{C}^4)^{3n}$ of dimension at least $9n = 12n - 3n$. The given point $\mathcal{A}$ lies in L . We need to show that there is a choice of $\mathcal{G} \in (GL_3)^n$ such that $\mathcal{GA}$ (which again lies in L) avoids the determinantal surfaces. This is equivalent to picking a basis for each L_i that stack together to a $\mathcal{B} \in L \setminus \bigcup_{\sigma} V(d_{\sigma})$.

We first show that L is not contained in any $V(d_{\sigma})$ by exhibiting a point in $L \setminus V(d_{\sigma})$ for each σ . Since at most four cameras can be involved in any d_{σ} , we may assume without loss of generality that σ involves only rows of the first four cameras. There are four cases to consider depending on how many rows these four cameras contribute to σ — the possibilities being $(3,1,0,0)$, $(2,2,0,0)$, $(2,1,1,0)$, and $(1,1,1,1)$. In each case we will produce a $\mathcal{B} \in L \setminus V(d_{\sigma})$. A key observation is that A_i and A_j having distinct foci implies $L_i \cap L_j$ is a proper subspace of both L_i and L_j for all i, j . Our starting point in each case below is $\mathcal{A} \in L$ which we modify to the needed $\mathcal{B}$ by replacing the bases of L_i that provide the rows of A_i .

Case 1. $(3,1,0,0)$: Modify $\mathcal{A}$ to $\mathcal{B}$ by choosing a basis for L_2 to be the three rows of B_2 so that no element in this basis lies in $L_1 \cap L_2$. Then $\mathcal{B}$ does not vanish on d_{σ} .

Case 2. $(2,2,0,0)$: Choose a basis for L_1 such that the two rows v_1, v_2 contributing to σ from the first camera are chosen from $L_1 \setminus L_2$. Then $L_2 \cap \text{Span}\{v_1, v_2\}$ is a proper subspace of L_2 of dimension at most one. Therefore taking two linearly independent vectors v_3, v_4 outside of this subspace as the two rows from L_2 creates a $\mathcal{B}$ that does not vanish on d_{σ} .

Case 3. $(2,1,1,0)$: Choose a basis for L_1 such that the two contributing rows v_1, v_2 from the first camera lie in $L_1 \setminus (L_2 \cup L_3)$. Choose the row v_3 from L_2 such that $v_3 \in L_2 \setminus (\text{Span}\{v_1, v_2\} \cup L_3)$, which forces $L_3 \cap \text{Span}\{v_1, v_2, v_3\}$ to be a proper subspace of L_3 . Taking v_4 outside this subspace, we get a point $\mathcal{B} \in L$ at which d_{σ} does not vanish.

Case 4. (1,1,1,1): Choose $v_1 \in L_1 \setminus (L_2 \cup L_3 \cup L_4)$, $v_2 \in L_2 \setminus (\text{Span}\{v_1\} \cup L_3 \cup L_4)$, $v_3 \in L_3 \setminus (\text{Span}\{v_1, v_2\} \cup L_4)$, and $v_4 \in L_4 \setminus (\text{Span}\{v_1, v_2, v_3\})$. By construction, we get a point in L at which d_σ does not vanish.

Therefore, $L \cap V(d_\sigma)$ is a proper subvariety of L for each σ , and a generic choice of $\mathcal{G}$ will put $\mathcal{G}_A \in L \setminus \bigcup_\sigma V(d_\sigma)$. $\square$

We note that $\mathcal{A}$ having distinct foci does not imply that $\mathcal{A}$ is minor-generic. A simple example would be an arrangement of four translational cameras; the submatrix consisting of the four first rows in each camera has zero determinant. However, having distinct foci allows the camera arrangement to be made minor-generic by the action of a tuple $\mathcal{G}$. We are now ready to prove the main theorem of this section.

Theorem 3.7. *Let $\mathcal{A}$ be an arrangement of cameras with distinct foci. Then $M_A = H_A^2 + H_A^3$.*

Proof. By Lemma 3.6, there exists $\mathcal{G} \in (GL_3)^n$ such that $\mathcal{G}_A$ is minor-generic. Then, by Corollary 3.5, $M_{\mathcal{G}_A} = H_{\mathcal{G}_A}^2 + H_{\mathcal{G}_A}^3$. Therefore, by Lemmas 3.1 and 2.5, we get

$$M_A = \chi_{\mathcal{G}^{-1}}(M_{\mathcal{G}_A}) = \chi_{\mathcal{G}^{-1}}(H_{\mathcal{G}_A}^2) + \chi_{\mathcal{G}^{-1}}(H_{\mathcal{G}_A}^3) = H_A^2 + H_A^3. \quad \square$$

Proposition 5(1) in [16] says that the H_A^2 and H_A^3 together cut out the multiview variety which implies that $H_A^2 + H_A^3 \subseteq M_A$. Theorem 3.7 shows that these polynomials also generate the multiview ideal providing the analogous ideal-theoretic statement.

Theorem 3.7 improves on Corollary 2.7 in [4] which states that when the foci of the cameras A_i are in linearly general position, then M_A is generated by the bifocals and trifocals. Theorem 3.7 requires no sophisticated condition on the cameras beyond the foci being pairwise distinct.

Conca et al. [20] and Li [21] also consider the vanishing ideal of the image of linear map from a projective space to a product of projective spaces. It is shown in [20] that this ideal is Cartwright-Sturmfels, meaning that its initial ideal is radical after a generic change of coordinates. Both of these works allow for projective spaces of arbitrary dimension. Specializing to our situation, Li's results show that $M_A = \sum_{k=2}^n H_A^k$ while we prove that $M_A = H_A^2 + H_A^3$.

Just like in [16] where the results automatically generalized from projective cameras to euclidean cameras, Theorem 3.7 also generalizes to euclidean cameras. Recall that a camera A_i is euclidean if it is of the form $A_i = [R_i \ t_i]$ where $R_i \in \text{SO}_3$.

Corollary 3.8. *Let $\mathcal{A}$ be an arrangement of euclidean cameras with pairwise distinct foci. Then $M_A = H_A^2 + H_A^3$.*

We state one more consequence of Theorem 3.7 which will be needed in the next section.

Corollary 3.9. *Let $\mathcal{A}$ be a camera arrangement with pairwise distinct foci. Then for any $\mathbf{p}_i \in \mathbb{P}^2$, the points $(A_1\mathbf{c}_i, A_2\mathbf{c}_i, \dots, \mathbf{p}_i, \dots, A_n\mathbf{c}_i)$ lie in $V(M_A)$ where $\mathbf{c}_i$ is the focal point of A_i .*

Proof. By Theorem 3.7, it suffices to show that for any i , the bifocals and trifocals vanish on the points $(A_1\mathbf{c}_i, A_2\mathbf{c}_i, \dots, \mathbf{p}_i, \dots, A_n\mathbf{c}_i)$. For any pair of cameras $\{i, j\}$, observe that $(\mathbf{c}_i, 0, -1)$ is a nonzero element of $\ker \mathcal{A}_{\{i,j\}}$

$(\mathbf{p}_i, A_j\mathbf{c}_i)$. For any pair $\{j, k\}$ not containing camera i , $(\mathbf{c}_i, -1, -1)$ is a nonzero element of $\ker \mathcal{A}_{\{j,k\}}(A_j\mathbf{c}_i, A_k\mathbf{c}_i)$. Hence all polynomials of H_A^2 vanish on $(A_1\mathbf{c}_i, A_2\mathbf{c}_i, \dots, \mathbf{p}_i, \dots, A_n\mathbf{c}_i)$. A similar argument applies to any triples of cameras, from which it follows that all polynomials in H_A^3 vanish on $(A_1\mathbf{c}_i, A_2\mathbf{c}_i, \dots, \mathbf{p}_i, \dots, A_n\mathbf{c}_i)$. $\square$

The image of focal point i in image j , i.e., $A_j\mathbf{c}_i$, is called the *epipole in image j relative to image i* . Corollary 3.9 shows that while the product of an arbitrary point in image i with all epipoles relative to image i does not appear in the image of φ_A , these points appear in the multiview variety after taking Zariski closure. See also Proposition 1 in [16].

We conclude this section by showing that the hypothesis in Theorem 3.7 cannot be relaxed, namely if a pair of foci of cameras in $\mathcal{A}$ coincide, then the multiview ideal is strictly larger than the ideal generated by bifocals and trifocals.

Example 3.10. Consider the four translational camera arrangement $\mathcal{A}$ where $\mathbf{t}_1, \mathbf{t}_2 = (0, 0, 0)$, $\mathbf{t}_3 = (1, 1, 1)$, $\mathbf{t}_4 = (-1, -1, -1)$. Eliminating the variables q and λ_i from the ideal $\langle A_i q - \lambda_i p_i : i = 1, \dots, n \rangle$, we can directly obtain M_A . Computing a primary decomposition of $H_A^2 + H_A^3$, we find that

$$H_A^2 + H_A^3 = M_A \cap \langle y_4 - z_4, y_3 - z_3, x_4 - z_4, x_3 - z_3 \rangle.$$

The extra component $\langle y_4 - z_4, y_3 - z_3, x_4 - z_4, x_3 - z_3 \rangle$ cuts out the points $(\mathbf{p}_1, \mathbf{p}_2, A_3\mathbf{c}_1, A_4\mathbf{c}_1)$, and from the primary decomposition we see that the projective variety they form is not contained in $V(M_A)$.

4 MORE IDEALS FOR THE MULTIVIEW VARIETY

In the computer vision literature, there are several sets of polynomials that have been shown to vanish on the space of images $\varphi_A(\mathbb{P}_{\mathbb{R}}^3)$, and hence they also vanish on the multiview variety. We now consider three such sets of polynomials and the ideals they generate, and compare them to the multiview ideal M_A .

4.1 Heyden and Åström [2]

Heyden and Åström were the first to do an algebraic study of the multiview variety, by studying the n -focal ideal H_A^n [2]. The variety of this ideal is indeed the multiview variety.

Lemma 4.1. *For any camera arrangement $\mathcal{A}$ with pairwise distinct foci, $V(M_A) = V(H_A^n)$.*

Proof. Recall from the image formation equations, $A_i\mathbf{q} = \lambda_i\mathbf{p}_i$ for all $i = 1, \dots, n$, that if $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_n)$ lies in the image of φ_A then the matrix $\mathcal{A}(\mathbf{p})$ has a non-trivial kernel. This means that all maximal minors of $\mathcal{A}(\mathbf{p})$ vanish on the image of φ_A , and therefore also on its Zariski closure, which is the multiview variety. Therefore, $V(M_A) \subseteq V(H_A^n)$.

To see the reverse inclusion, suppose $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_n) \in V(H_A^n)$ which means that $\mathcal{A}(\mathbf{p})$ is rank deficient and there is a nonzero vector of the form $(\mathbf{q}, -\lambda_1, \dots, -\lambda_n)$ in the kernel of $\mathcal{A}(\mathbf{p})$. If $\mathbf{q} = 0$, then we will get that $\lambda_i\mathbf{p}_i = 0$ for all i . However, since $\mathbf{p}_i \neq 0$, it must be that $\lambda_i = 0$ for all i and hence the vector in the kernel is the zero vector which is a contradiction. Therefore, there is a nonzero vector $\mathbf{q}$ such that $A_i\mathbf{q} = \lambda_i\mathbf{p}_i$ for some

λ_i . If $\mathbf{q}$ is not the focal point of any camera, then $\mathbf{p}$ lies in $\varphi_A(\mathbb{P}_C^3)$. Since φ_A is continuous, $\varphi_A(\overline{\mathbb{P}_C^3}) \subseteq \overline{\varphi_A(\mathbb{P}_C^3)}$. It follows that $\varphi_A(\mathbb{P}_C^3) \subseteq V(M_A)$ because $\overline{\mathbb{P}_C^3} = \mathbb{P}_C^3$ and so $\mathbf{p} \in V(M_A)$. On the other hand, if $\mathbf{q}$ is the focal point $\mathbf{c}_i$ of camera i , then $\mathbf{p}_j = A_j \mathbf{c}_i$ for all $j \neq i$, and by Corollary 3.9, $\mathbf{p} \in V(M_A)$. Thus we get that $V(M_A) \supseteq V(H_A^n)$. $\square$

Example 3.10 shows that the assumption of distinct foci is necessary for Lemma 4.1. In this example, $n = 4$ and $H_A^4 = H_A^2 + H_A^3$ by Corollary 3.4. We see that $V(H_A^4)$ has a component other than $V(M_A)$.

4.2 Faugeras et al. [22]

The second set of polynomials we will study were constructed by Faugeras & Mourrain while proving that the multiview variety is cut out by epipolar/bifocal and trifocal polynomials, and that the quadrifocal constraints corresponding to the quadrifocal tensor were not needed [7], [22].

Observe that $A_i \mathbf{q} = \lambda_i \mathbf{p}_i$ implies $A_i \mathbf{q} \times \mathbf{p}_i = 0$, for each i , or equivalently, $[p_i]_{\times} A_i \mathbf{q} = 0$, where

$$[p_i]_{\times} = \begin{pmatrix} 0 & -z_i & y_i \\ z_i & 0 & -x_i \\ -y_i & x_i & 0 \end{pmatrix}, \quad (5)$$

represents taking cross product with p_i , i.e., $[p_i]_{\times} v = p_i \times v$. Stacking all 3×4 matrices $[p_i]_{\times} A_i$, we get the $3n \times 4$ partially symbolic matrix

$$\mathcal{A}^F(p) := \begin{pmatrix} [p_1]_{\times} A_1 \\ [p_2]_{\times} A_2 \\ \vdots \\ [p_n]_{\times} A_n \end{pmatrix}. \quad (6)$$

If there is a world point $\mathbf{q}$ satisfying $A_i \mathbf{q} \times \mathbf{p}_i = 0$, then this matrix is rank deficient and all maximal minors of $\mathcal{A}^F(p)$ vanishes on the multiview variety.

Definition 4.2. The ideal of all maximal 4×4 minors of $\mathcal{A}^F(p)$, denoted by F_A , will be called the Faugeras ideal of the arrangement $\mathcal{A}$. We denote the subideals of F_A generated by minors involving only two and three cameras by F_A^2 and F_A^3 , respectively.

We now describe a sequence of matrix transformations that allow us to obtain $\mathcal{A}^F(p)$ from $\mathcal{A}(p)$. Let $P(p) := \text{diag}([p_1]_{\times}, \dots, [p_n]_{\times})$ be the symbolic block diagonal matrix of size $3n \times 3n$. Multiplying $\mathcal{A}(p)$ on the left by the block diagonal matrix $P(p)$ and dropping the rightmost n columns of the resulting matrix, we obtain $\mathcal{A}^F(p)$

$$\mathcal{A}^F(p) = P(p) \mathcal{A}(p) \begin{bmatrix} I_4 \\ 0_{n \times 4} \end{bmatrix} = P(p) \mathcal{A}, \quad (7)$$

where as before, we abuse notation to let $\mathcal{A}$ also represents the $3n \times 4$ matrix $[A_1; \dots; A_n]$ obtained by stacking the cameras vertically. From the matrix constructions of H_A^n and F_A , we observe that their projective vanishing sets in $(\mathbb{P}^2)^n$ coincide.

Lemma 4.3. For any camera arrangement $\mathcal{A}$ with pairwise distinct foci, $V(M_A) = V(F_A)$.

Proof. The proof will follow from Lemma 4.1 if we can show that $V(F_A) = V(H_A^n)$. If $\mathbf{p} \in (\mathbb{P}^2)^n$ is such that $\mathcal{A}^F(\mathbf{p})$ drops rank, then there exists a nonzero $\mathbf{q} \in \ker(\mathcal{A}^F(\mathbf{p}))$ so that $A_i \mathbf{q} \times \mathbf{p}_i = 0$ for all i . This means there exist nonzero scale factors λ_i such that $A_i \mathbf{q} = \lambda_i \mathbf{p}_i$. The vector $(\mathbf{q}, -\lambda_1, \dots, -\lambda_n)$ is a nontrivial element in $\ker(\mathcal{A}(\mathbf{p}))$, so $\mathcal{A}(\mathbf{p})$ is rank deficient. Therefore $V(F_A) \subseteq V(H_A^n)$.

For the other inclusion, if there is a nontrivial $(\mathbf{q}, -\lambda_1, \dots, -\lambda_n) \in \ker(\mathcal{A}(\mathbf{p}))$ for some $\mathbf{p} \in (\mathbb{P}^2)^n$, then as in the proof of Lemma 4.1, $\mathbf{q}$ must be nonzero, and so $\mathbf{q}$ is a nontrivial element of $\ker(\mathcal{A}^F(\mathbf{p}))$. This shows that $V(F_A) \supseteq V(H_A^n)$, hence $V(F_A) = V(H_A^n) = V(M_A)$. $\square$

4.3 Ma et al. [9]

The third and final set of polynomials we will study are the so called *multiview rank constraints* which were proposed by Ma and collaborators [9] as an alternative to the multilinear constraints studied for example in Hartley & Zisserman [8].

Suppose $A_1 = [I \ 0]$ and $A_i = [B_i \ t_i]$ for $i \geq 2$. Starting with $\mathcal{A}(p)$, a series of matrix operations are described in Chapter 8 in [9] to arrive at a new set of determinantal polynomials, arising as maximal minors of

$$\mathcal{A}^Y(p) := \begin{bmatrix} p_1 \times (I p_1) & p_1 \times 0 \\ p_2 \times (B_2 p_1) & p_2 \times t_2 \\ \vdots & \vdots \\ \vdots & \vdots \\ p_n \times (B_n p_1) & p_n \times t_n \end{bmatrix}. \quad (8)$$

Definition 4.4. The ideal of all maximal 2×2 minors of $\mathcal{A}^Y(p)$, denoted by Y_A , will be called the Ma ideal of the arrangement $\mathcal{A}$.

We observe that $\mathcal{A}^Y(p)$ can be obtained from $\mathcal{A}^F(p)$ by multiplying by a single matrix on the right

$$\mathcal{A}^Y(p) = \mathcal{A}^F(p) \begin{bmatrix} p_1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (9)$$

From this we observe that Y_A has the same projective vanishing set as F_A , and hence H_A^n and M_A .

Lemma 4.5. For any camera arrangement $\mathcal{A}$ with pairwise distinct foci and $A_1 = [I \ 0]$, $V(M_A) = V(Y_A)$.

Proof. If $\mathbf{p} \in (\mathbb{P}^2)^n$ is such that $\mathcal{A}^Y(\mathbf{p})$ drops rank, then there exists a nontrivial $(v_1, v_2) \in \ker(\mathcal{A}^Y(\mathbf{p}))$. Therefore, $\mathbf{q} = (v_1 \mathbf{p}_1, v_2) \in \ker(\mathcal{A}^F(\mathbf{p}))$ is nontrivial. Note that it is necessary that we assume $A_1 = [I \ 0]$ so that $[p_1]_{\times} A_1 (v_1 \mathbf{p}_1, v_2) = v_1 [p_1]_{\times} \mathbf{p}_1 = 0$. This shows that $V(Y_A) \subseteq V(F_A)$.

For the other inclusion, if $0 \neq \mathbf{q} \in \ker(\mathcal{A}^F(\mathbf{p}))$ for some $\mathbf{p} \in (\mathbb{P}^2)^n$, then since $\mathbf{p}_1 \times [I \ 0] \mathbf{q} = 0$, there exists a scalar v_1 such that $v_1 \mathbf{p}_1 = (\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3)$. This means that $(v_1, \mathbf{q}_4) \in \ker(\mathcal{A}^Y(\mathbf{p}))$, which is nontrivial because if $v_1 = 0$, then $(\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) = 0$, so $\mathbf{q}_4 \neq 0$. This shows $V(Y_A) \supseteq V(F_A)$, and the desired result follows from Lemma 4.3. $\square$

Observe that Y_A is generated by polynomials of total degree 3. This fact has an interesting consequence. As we

mentioned earlier, Y_A has been proposed as an alternate algebraic foundation for multi-view geometry. From Lemma 4.5, we know that it cuts out the multiview variety. Since M_A is the vanishing ideal of the multiview variety, we get that $Y_A \subseteq M_A$. However, from Theorem 3.7 we know that $M_A = H_A^2 + H_A^3$, i.e., it is generated by polynomials of degree two and three, which means that in general $Y_A \neq M_A$ and instead $Y_A \subset M_A$ or equivalently $Y_A \subset H_A^2 + H_A^3$. This means that the bifocals and trifocals imply the multiview rank constraints, but not the other way around. Similarly, H_A^n and F_A , which are generated by polynomials of total degree n and four respectively, are properly contained in M_A . We see this in Example 4.6 below.

4.4 Relationships to the Multiview Ideal

We now compute the three ideals on an example, foreshadowing their structural properties, which we examine next.

Example 4.6. Consider the translational arrangement $\mathcal{A}$ where $\mathbf{t}_1 = (0, 0, 0)$, $\mathbf{t}_2 = (1, 0, 0)$, $\mathbf{t}_3 = (0, 1, 0)$ whose multiview ideal is

$$M_A = \langle y_1 z_2 - y_2 z_1, x_2 z_3 - x_3 z_2 + y_2 z_3 - y_3 z_2, \\ x_1 z_3 - x_3 z_1, x_1 x_3 y_2 + x_1 y_2 y_3 - x_2 x_3 y_1 - x_3 y_1 y_2 \rangle.$$

The primary decompositions of H_A^n , F_A , and Y_A are

$$H_A^n = M_A \cap \langle z_1, y_1, x_1 \rangle \cap \langle z_2, y_2, x_2 \rangle \cap \langle z_3, y_3, x_3 \rangle, \\ Y_A = M_A \cap \langle z_1, y_1, x_1 \rangle \cap \langle y_3, y_2, x_3, x_2, z_3^2, z_2 z_3, z_2^2 \rangle \\ \cap \langle z_1, y_2, x_3, x_2, x_1, z_3^2, z_2 z_3, z_2^2 \rangle \\ \cap \langle z_1, y_3, y_2, y_1, x_3, z_3^2, z_2 z_3, z_2^2 \rangle,$$

$$F_A = M_A \cap \langle y_2, y_1, x_2, x_1, z_2^2, z_1 z_2, z_1^2 \rangle \\ \cap \langle y_3, y_1, x_3, x_1, z_3^2, z_1 z_3, z_1^2 \rangle \\ \cap \langle z_2, z_1, y_3, y_2, y_1, x_3, z_3^2 \rangle \\ \cap \langle z_3, z_2, y_1, x_3 + y_3, x_2 + y_2, x_1, z_1^2 \rangle \\ \cap \langle y_2, x_2, z_3^2, z_2 z_3, z_2^2, y_3 z_3, y_3 z_2, y_3^2, x_3 z_3, \\ x_3 z_2, x_3 y_3, x_3^2, x_1 x_3 + x_1 y_3 - x_3 y_1 \rangle \\ \cap \langle z_3, y_2, x_2, z_2^2, z_1 z_2, z_1^2, x_3 z_2, x_3 z_1, x_3^2, \\ x_1 z_2, x_1 z_1, x_1 y_3 - x_3 y_1, x_1 x_3, x_1^2 \rangle \cap C,$$

where C is a component minimally generated by 133 polynomials of total degree up to eight.

While each of H_A^n , F_A , and Y_A notably contains M_A as a component, the nature of their other components is worth further investigation.

To analyze the extra components, we rely on several notions from commutative algebra, which we define next. The first notion is that of a *multigraded* ring. Consider the ring $\mathbb{C}[p_1, \dots, p_n]$ endowed with the $\mathbb{Z}^n$ -grading $\circ(w_i) = \mathbf{e}_i$ where $w_i \in \{x_i, y_i, z_i\}$ and $\mathbf{e}_i$ is the i th standard basis vector in $\mathbb{R}^n$. We say a polynomial in this ring is homogeneous if each of its terms have the same multidegree.

The *irrelevant ideal* in this grading, which we denote by $\mathfrak{m}$, is the intersection of the ideals $\mathfrak{m}_i := \langle x_i, y_i, z_i \rangle$.

$$\mathfrak{m} := \bigcap_{i=1}^n \mathfrak{m}_i = \bigcap_{i=1}^n \langle x_i, y_i, z_i \rangle. \quad (10)$$

Observe that $\mathfrak{m}$ is generated by all multilinear monomials of multidegree $(1, 1, \dots, 1)$ and total degree n . It is the maximal ideal in the ring $\mathbb{C}[p_1, \dots, p_n]$ generated by homogeneous elements of strictly positive multidegree.

The *radical* of an ideal I is the ideal $\sqrt{I} := \{f : f^k \in I \text{ for some } k \in \mathbb{N}\}$. If I is a homogeneous ideal then so is its radical, and $I \subseteq \sqrt{I}$. The *colon* of an ideal I with the ideal J , denoted as $(I : J)$ is the set of all polynomials f such that $fg \in I$ for all $g \in J$, i.e., $I : J = \{f : fJ \subseteq I\}$.

Recall that the projective varieties of the ideals H_A^n , F_A , and Y_A all agree and equal the multiview variety $V(M_A)$. We can now state a first relationship among the ideals that follows easily from the projective Nullstellensatz in our multigraded setting, whose statement and proof will appear in Appendix A.

Theorem 4.7. For any $\mathcal{A}$ with pairwise distinct foci,

- (a) $\sqrt{H_A^n} : \mathfrak{m} = M_A$.
- (b) $\sqrt{F_A} : \mathfrak{m} = M_A$.
- (c) $\sqrt{Y_A} : \mathfrak{m} = M_A$ when $A_1 = [I \ 0]$.

Proof. See Appendix A. $\square$

In the language of algebraic geometry what this says is that $\sqrt{H_A^n}$, $\sqrt{F_A}$ and $\sqrt{Y_A}$ all cut out the multiview variety *scheme-theoretically*. They are not equal as ideals but they agree in high enough multidegree with M_A , see [23, pp 50].

We now strengthen Theorem 4.7 (a) and (b) to show that the operation of taking the radical is not needed, i.e., $H_A^n : \mathfrak{m} = M_A$ and $F_A : \mathfrak{m} = M_A$. This means that H_A^n and F_A already cut out the multiview variety *scheme-theoretically*. Experimental evidence suggests that when $A_1 = [I \ 0]$, such a result is also true for Y_A , but an explicit proof is made difficult by the convoluted structure of the 2×2 minors of $\mathcal{A}^V(p)$.

We first show that the simple structure of the primary decomposition of H_A^n observed in Example 4.6 holds in general.

Lemma 4.8. For any camera arrangement $\mathcal{A}$ with pairwise distinct foci, $H_A^n = M_A \cap \mathfrak{m}$. In particular, H_A^n is a radical ideal with prime decomposition $M_A \cap \mathfrak{m}_1 \cap \mathfrak{m}_2 \cap \dots \cap \mathfrak{m}_n$.

Proof. Suppose f is a generator of $\in H_A^n$, i.e., a maximal minor of $\mathcal{A}(p)$. Then $f \in \mathfrak{m}$. Also, since f vanishes on $V(M_A)$, $f \in M_A$. Therefore, $H_A^n \subseteq M_A \cap \mathfrak{m}$.

Now suppose $f \in M_A \cap \mathfrak{m}$. Since M_A is generated by bifocals and trifocals $f = \sum \lambda_i r_i b_i + \sum \mu_j s_j t_j$ where b_i 's are bifocals, t_j 's are trifocals, r_i, s_j are monomials, and λ_i, μ_j are scalars. Further, since $f \in \mathfrak{m}$, every term in f is divisible by some generator $\prod_{i=1}^n w_i$ of $\mathfrak{m}$ where $w_i \in \{x_i, y_i, z_i\}$. Now consider $r_i b_i$. Since b_i involves only two cameras, it must be that r_i contains a variable w_i from each of the other $n - 2$ cameras so that each term of $r_i b_i$ lies in $\mathfrak{m}$. This makes $r_i b_i$ a monomial multiple of a n -focal by Lemma 2.2. The same argument holds for $s_j t_j$. Thus, $f \in H_A^n$. $\square$

Proposition b3 in [17] proves that when $\mathcal{A}$ is minor-generic, $H_{\mathcal{A}}^n$ is a radical ideal. Lemma 4.8 shows that $H_{\mathcal{A}}^n$ is always a radical ideal under the weaker assumption of distinct foci.

Theorem 4.9. *For any camera arrangement $\mathcal{A}$ with pairwise distinct foci, $H_{\mathcal{A}}^n : \mathfrak{m} = M_{\mathcal{A}}$.*

Proof. We first note that $M_{\mathcal{A}} : \mathfrak{m} = M_{\mathcal{A}}$. Suppose $f \in M_{\mathcal{A}} : \mathfrak{m}$. Then $fu \in M_{\mathcal{A}}$ for any monomial generator u of $\mathfrak{m}$. Since $M_{\mathcal{A}}$ is prime and does not contain any monomials, $f \in M_{\mathcal{A}}$. Since $H_{\mathcal{A}}^n = M_{\mathcal{A}} \cap \mathfrak{m}$ by Lemma 4.8, $H_{\mathcal{A}}^n : \mathfrak{m} = M_{\mathcal{A}} : \mathfrak{m} = M_{\mathcal{A}}$. $\square$

We now consider the Faugeras ideal $F_{\mathcal{A}}$ and prove that $F_{\mathcal{A}} : \mathfrak{m} = M_{\mathcal{A}}$. The nontrivial part is to argue that $M_{\mathcal{A}}$ is contained in $F_{\mathcal{A}} : \mathfrak{m}$. This fact relies on the following technical lemma, similar in flavor to Lemma 2.2, which shows that bifocals and trifocals can both be multiplied by any generator of $\mathfrak{m}$ to fall into $F_{\mathcal{A}}$.

Lemma 4.10.

- a) For $n = 2$ cameras, and any monomial $p_{1j}p_{2k}$, there exists a 4×4 minor f of $\mathcal{A}^F(p)$ such that $f = (-1)^{j+k} p_{1j}p_{2k} \det \mathcal{A}(p)$.
- b) Let $n = 3$ and i_1, i_2, i_3 be pairwise distinct. Then for any trifocal $\det \mathcal{A}(p)_{\{p_{i_1j_1}p_{i_2j_2}\}}$ and any coordinate p_{i_3k} , there exists a 4×4 minor f of $\mathcal{A}^F(p)$ such that $f = (-1)^k p_{i_3k} \det \mathcal{A}(p)_{\{p_{i_1j_1}p_{i_2j_2}\}}$.

Proof. See Appendix B. $\square$

Theorem 4.11. *For any camera arrangement $\mathcal{A}$ with pairwise distinct foci, $F_{\mathcal{A}} : \mathfrak{m} = M_{\mathcal{A}}$.*

Proof. The containment $F_{\mathcal{A}} : \mathfrak{m} \subseteq M_{\mathcal{A}}$ follows as in Theorem 4.9 because $F_{\mathcal{A}} \subseteq M_{\mathcal{A}}$ and hence, $F_{\mathcal{A}} : \mathfrak{m} \subseteq M_{\mathcal{A}} : \mathfrak{m} = M_{\mathcal{A}}$. The other containment will follow by showing $H_{\mathcal{A}}^2, H_{\mathcal{A}}^3 \subseteq F_{\mathcal{A}} : \mathfrak{m}$. For general camera arrangements with n cameras, recall that $F_{\mathcal{A}}^2$ (resp. $F_{\mathcal{A}}^3$) is the ideal generated by all 4×4 minors of $\mathcal{A}^F(p)$ that involve only two (resp. three) cameras. By Lemma 4.10(a), for any multilinear monomial $(\prod_{m=1}^n w_m)$ and any bifocal b_{ij} , $(\prod w_m)b_{ij} \in (f)$ for some Faugeras minor $f \in F_{\mathcal{A}}^2$, hence $H_{\mathcal{A}}^2 \subseteq F_{\mathcal{A}} : \mathfrak{m}$. We address the trifocals in two cases. First consider the case when the two rows eliminated from $\mathcal{A}_{\{i,j,k\}}(p)$ to form a trifocal $t \in H_{\mathcal{A}}^3_{\{i,j,k\}}$ come from the same camera, say without loss of generality, from camera i . In this case, $t = w_i b_{jk}$ for some w_i , and Lemma 4.10(a) again implies $t \in F_{\mathcal{A}} : \mathfrak{m}$. For the case when the two rows from $\mathcal{A}_{\{i,j,k\}}(p)$ to form $t \in H_{\mathcal{A}}^3_{\{i,j,k\}}$ come from different cameras, Lemma 4.10(b) implies that, for any $(\prod w_m)$, $(\prod w_m)t \in (f)$ for some $f \in F_{\mathcal{A}}^3$. We conclude that $H_{\mathcal{A}}^3 \subseteq F_{\mathcal{A}} : \mathfrak{m}$, as desired. $\square$

5 THE BIFOCAL IDEAL

We saw in Theorem 3.7 that the bifocals and trifocals together generate the multiview ideal when the camera foci are pairwise distinct. In this section, we investigate

how imposing further conditions on the cameras can lead to an even simpler description of the multiview ideal. Heyden and Åström [2] and Trager et al. [16] show that when the camera foci are not all on a plane, the bifocals are necessary and sufficient to cut out the multiview variety. There has also been work to further reduce this description by considering the minimal number of bifocals needed ([2], [24]), though we will not address this question here. In this section, we focus on the ideal-theoretic relationship between the bifocal ideal $H_{\mathcal{A}}^2$ and the multiview ideal $M_{\mathcal{A}}$ when the camera foci are noncoplanar.

To motivate our investigation, we start with some examples. We say that a camera arrangement $\mathcal{A}$ is coplanar, noncoplanar or collinear if their foci have the corresponding property.

Example 5.1. Consider the four noncoplanar translational camera arrangement $\mathcal{A}_1$ where $\mathbf{t}_1 = (0, 0, 0)$, $\mathbf{t}_2 = (1, 0, 0)$, $\mathbf{t}_3 = (0, 1, 0)$, $\mathbf{t}_4 = (0, 0, 1)$. Eliminating the variables q and λ_i from the ideal $\langle A_i q - \lambda_i p_i : i = 1, \dots, n \rangle$, we observe $M_{\mathcal{A}_1}$ occurs as a component in $H_{\mathcal{A}_1}^2$

$$\begin{aligned} H_{\mathcal{A}_1}^2 &= M_{\mathcal{A}_1} \cap \langle x_2, y_2, z_2, x_1, x_3, x_4 \rangle \\ &\cap \langle x_1, y_1, z_1, x_2 + y_2 + z_2, x_3 + y_3 + z_3, x_4 + y_4 + z_4 \rangle \\ &\cap \langle x_3, y_3, z_3, y_1, y_2, y_4 \rangle \cap \langle x_4, y_4, z_4, z_1, z_2, z_3 \rangle. \end{aligned}$$

Example 5.2. Consider the four coplanar translational camera arrangement $\mathcal{A}_2$ where $\mathbf{t}_1 = (1, 0, 0)$, $\mathbf{t}_2 = (0, 1, 0)$, $\mathbf{t}_3 = (0, 0, 1)$, $\mathbf{t}_4 = (1/3, 1/3, 1/3)$. We observe that $H_{\mathcal{A}_2}^2 = M_{\mathcal{A}_2} \cap C$ where

$$C = \langle x_4 + y_4 + z_4, x_3 + y_3 + z_3, x_2 + y_2 + z_2, x_1 + y_1 + z_1 \rangle.$$

In Example 5.1, each extra component of $H_{\mathcal{A}_1}^2$ contains an irrelevant ideal $\mathfrak{m}_i$ and hence does not contribute to $V(H_{\mathcal{A}_1}^2)$. Saturating the bifocal ideal $H_{\mathcal{A}_1}^2$ with respect to the full irrelevant ideal $\mathfrak{m}$ removes these components. We will prove that this is always true when camera foci are noncoplanar. We begin by proving a series of three lemmas.

Lemma 5.3. *Suppose $\mathcal{A}$ is an arrangement of $n \geq 4$ cameras with pairwise distinct foci. Then $\mathcal{A}$ is noncoplanar $\Rightarrow H_{\mathcal{A}}^n \subseteq H_{\mathcal{A}}^2$.*

Proof. $n = 4, 5, 6$. If $\mathcal{A}$ is noncoplanar, then there is some subset of four cameras that is noncoplanar. Order the cameras in $\mathcal{A}$ so that these are the cameras $A_1, \dots, A_4$. By a change of coordinates on $\mathbb{P}^3$, we can send the foci of the cameras $A_1, \dots, A_4$ to the foci of the cameras in $\mathcal{A}_1$ from Example 5.1. Then, by Lemma 2.5, applying $\mathbb{P}^2$ coordinate changes using some $\mathcal{G} \in (GL_3)^n$, we can assume that $\mathcal{A}$ is an arrangement of translational cameras. These transformations fix the first four cameras, and we think of the cameras A_i for $i \geq 5$ as variable, represented symbolically by their translations, and the implication can be confirmed by direct calculation in Macaulay2.

$n = 7$. In this case, the full computation is too expensive. To make the computation feasible, we split the proof into two cases, depending on whether the arrangement has five collinear cameras or not.

Case I. If a noncoplanar arrangement of seven cameras has at most four collinear cameras, then every four camera subset can be augmented with two additional cameras to get a noncoplanar arrangement of six cameras. Thus every 7-focal of such an arrangement, which looks like $w_i w_j w_k q$ for some quadrifocal q , has the form of a 6-focal from a noncoplanar arrangement, say $w_i w_j q$, multiplied by a coordinate w_k . The $n = 6$ case shows that $w_i w_j q$ is generated by 2-focals, hence $w_i w_j w_k q$ is generated by 2-focals.

Case II. We now consider the case of noncoplanar seven camera arrangements in which five cameras are collinear. In this case, by a proper choice of camera ordering and $\mathbb{P}^3$ coordinate change, we can assume the translations of A_5, A_6, A_7 are of the form $t_5 = (\lambda_5, 0, 0)^\top$, $t_6 = (\lambda_6, 0, 0)^\top$, $t_7 = (\lambda_7, 0, 0)^\top$ where the λ_i are symbolic. This makes A_1, A_2, A_5, A_6, A_7 collinear. The choice to take the line that the cameras lie on to be the x axis is arbitrary, but can be made without loss of generality. This arrangement is now described by few enough variables to enable a direct computation showing that $H_A^7 \subseteq H_A^2$.

$n \geq 8$. Now suppose $n \geq 8$ and f is an n -focal of $\mathcal{A}$. Recall that f involves all n cameras but at most four cameras can contribute two rows to the matrix whose determinant is f . At one extreme, these four cameras maybe $A_1, \dots, A_4$ and at the other extreme they might be four cameras different from the first four, which we call $A_5, \dots, A_8$. Thus the n -focal $f \in H_A^n$ is a monomial multiple of a 8-focal $g = mq$ of $\{A_1, \dots, A_4, A_5, \dots, A_8\}$ where q is a quadrifocal and m is a monomial.

If the four cameras contributing to q involve $A_1, \dots, A_4$, then g is a multiple of a 7-focal from noncoplanar cameras. On the other hand, if $q \in H_{A_5, \dots, A_8}^4$, then q can be generated by the trifocals of $A_5, \dots, A_8$ by Lemma 3.3:

$$g = m \left(\sum_{t_i \in H_{A_5, \dots, A_8}^3} h_i t_i \right) = \sum_{t_i \in H_{A_5, \dots, A_8}^3} h_i (m t_i).$$

In particular, this shows that g can be generated from 7-focals, $m t_i$. These come from noncoplanar seven camera arrangements because $A_1, \dots, A_4$ are noncoplanar. In either case, we know that such 7-focals can be generated by 2-focals, hence $g \in H_A^2$. It follows that $f \in H_A^2$, as desired. $\square$

Lemma 5.4. Suppose $\mathcal{A}$ is an arrangement of $n \geq 4$ cameras with pairwise distinct foci. Then $H_A^n \subseteq H_A^2 \Rightarrow M_A = H_A^2 : m$.

Proof. If $f \in H_A^2 : m$, then $f(\prod z_i) \in H_A^2 \subseteq M_A$, vanishes on $V(M_A)$. Since M_A is prime and does not contain any monomials, $f \in M_A$. Therefore, $H_A^2 : m \subseteq M_A$. For the other containment, by Theorem 3.7, it suffices to show that H_A^2 and H_A^3 are contained in $H_A^2 : m$. It is clear that $H_A^2 \subseteq H_A^2 : m$. By Lemma 2.2, multiplying any $t \in H_A^3$ by a generator $\prod w_i$ of m yields a monomial multiple of an n -focal. By assumption, this n -focal lies in H_A^2 . Thus, $t \in H_A^2 : m$ and $M_A \subseteq H_A^2 : m$. $\square$

Lemma 5.5. Suppose $\mathcal{A}$ is an arrangement of $n \geq 4$ cameras with pairwise distinct foci. Then $M_A = H_A^2 : m \Rightarrow \mathcal{A}$ is noncoplanar.

Proof. We prove the contrapositive, namely that if $\mathcal{A}$ is coplanar then $M_A \neq H_A^2 : m$. We will construct a point $p \in V(H_A^2 : m) \setminus V(M_A)$, from which the result will follow.

Let $n \in \mathbb{P}^3$ be the normal vector of a plane containing the foci of the cameras in $\mathcal{A}$. If the foci are not collinear then n is unique, otherwise we choose any plane containing the foci and its normal n . Let $l_i \subseteq \mathbb{P}^2$ denote the image of the plane $n^\perp$ in camera i , and let $e_{i,j}$ denote the image of the focal point of camera j in image i . Then $e_{i,j} \in l_i$ since the focal point of camera j lies in $n^\perp$. Choose $p_1 \in l_1 \setminus \{e_{1,2}, e_{1,3}\}$ and $p_2 \in l_2 \setminus \{e_{2,1}, e_{2,3}\}$. Then there is a unique world point q on $n^\perp$ whose images in cameras 1 and 2 are p_1 and p_2 . Let $\tilde{p}_3 \in l_3$ be the (unique) image of q in camera 3. Then $p_1, p_2, \tilde{p}_3$ satisfy trifocal constraints. Choose $p_3 \in l_3 \setminus \{\tilde{p}_3\}$ and some $p_i \in l_i$ for $i \geq 4$. By construction, $p \notin V(M_A)$. Since the cameras are coplanar, the epipolar plane given by q and any two cameras i and j is $n^\perp$ for any pair i, j . By choosing $p_i \in l_i$ for all i , we force every bifocal polynomial to vanish on p . Therefore by construction, $p \in V(H_A^2) \setminus V(M_A)$, but since $V(H_A^2) = V(H_A^2 : m)$, we conclude that $H_A^2 : m \neq M_A$. $\square$

Together, Lemmas 5.3, 5.4, 5.5 imply the following theorem.

Theorem 5.6. Suppose $\mathcal{A}$ is an arrangement of $n \geq 4$ cameras with pairwise distinct foci. Then the following are equivalent.

- (a) $\mathcal{A}$ is noncoplanar.
- (b) $H_A^n \subseteq H_A^2$.
- (c) $M_A = H_A^2 : m$.

We now make some observations about Theorem 5.6.

Theorem 6.1 in [2] observes that $V(H_A^2) = V(M_A)$ for noncoplanar $\mathcal{A}$ while Proposition 5 (2) in [16] further shows that $V(H_A^2) = V(M_A)$ is equivalent to the foci of $\mathcal{A}$ being noncoplanar. Our Theorem 5.6 proves the analogous ideal statement, namely that noncoplanarity of foci is equivalent to $M_A = H_A^2 : m$.

Example 5.2 shows how Theorem 5.6 fails when $\mathcal{A}$ is coplanar. The bifocal ideal $H_{\mathcal{A}_2}^2$ contains the component $\langle x_1 + y_1 + z_1, x_2 + y_2 + z_2, x_3 + y_3 + z_3, x_4 + y_4 + z_4 \rangle$, which cannot be removed by saturating with respect to m . Its variety cuts out the projections of the plane containing the foci of $\mathcal{A}_2$ in each camera image. This plane in $\mathbb{P}^3$ has normal vector $(1, 1, 1, -1)$. The following example shows that further degeneracy occurs when camera foci are collinear.

Example 5.7. Consider the four collinear translational camera arrangement $\mathcal{A}_3$ where $t_1 = (0, 0, 0)$, $t_2 = (1, 0, 0)$, $t_3 = (2, 0, 0)$, $t_4 = (3, 0, 0)$. Here, $H_{\mathcal{A}_3}^2 \subseteq M_{\mathcal{A}_3}$, but both ideals are prime, so $M_{\mathcal{A}_3}$ cannot occur as a component of $H_{\mathcal{A}_3}^2$. In addition, the dimension of $H_{\mathcal{A}_3}^2$ is one larger than that of $M_{\mathcal{A}_3}$. This is explained by the fact that there is an entire one-dimensional family of planes that contains the camera centers of $\mathcal{A}_3$.

As seen in the above examples and discussion, the relation between H_A^2 and M_A can be complicated when camera centers are coplanar or collinear. Determining the exact

relationship between ideals in these degenerate settings would be an interesting problem for the future.

In Theorem 5.6 we showed that when cameras are noncoplanar, the n -focal ideal becomes a subset of the 2-focal ideal. We now give an example to show that this containment need not hold for $H_{\mathcal{A}}^k$ where $n > k > 2$. The construction relies on having three of five cameras being collinear.

Example 5.8. Consider the five translational camera arrangement $\mathcal{B}$ with $\mathbf{t}_1 = (0, 0, 0)$, $\mathbf{t}_2 = (0, 0, 1)$, $\mathbf{t}_3 = (0, 0, 2)$, $\mathbf{t}_4 = (0, 1, 0)$, $\mathbf{t}_5 = (0, 0, 1)$. Theorem 5.6 shows that $H_{\mathcal{B}}^5 \subseteq H_{\mathcal{B}}^2$ since $\mathcal{B}$ is noncoplanar. However the following trifocal from B_1, B_2, B_3

$$t = -x_1y_2y_3 + 2x_2y_1y_3 - x_3y_1y_2,$$

is not in $H_{\mathcal{B}}^2$. Similarly, the quadrifocal

$$q = x_4t = -x_1x_4y_2y_3 + 2x_2x_4y_1y_3 - x_3x_4y_1y_2,$$

from cameras B_1, B_2, B_3, B_4 is not in $H_{\mathcal{B}}^2$.

6 FINITE IMAGES

The results of the previous sections have important practical consequences when we restrict attention to the set of all finite images, that is to all $(\mathbf{p}_1, \dots, \mathbf{p}_n) \in V(M_{\mathcal{A}})$ with $z_i \neq 0$ for all i . The vanishing ideal of this affine patch is obtained by dehomogenizing $M_{\mathcal{A}}$ with respect to the variables z_i from each image plane. We call this the *affine multiview ideal* of $\mathcal{A}$ and denote it $\pi(M_{\mathcal{A}})$, where $\pi : \mathbb{C}[x_i, y_i, z_i] \rightarrow \mathbb{C}[x_i, y_i]$ is the map setting each z_i to 1. From Theorem 3.7, we see that $\pi(M_{\mathcal{A}})$ is generated by dehomogenized bifocals and dehomogenized trifocals when the foci of $\mathcal{A}$ are pairwise distinct.

Corollary 6.1. *If $\mathcal{A}$ is a camera arrangement with pairwise distinct foci, then $\pi(M_{\mathcal{A}}) = \pi(H_{\mathcal{A}}^2) + \pi(H_{\mathcal{A}}^3)$.*

Using the following fact about dehomogenizing colon ideals, the results of Section 4 yield a nice relation among $\pi(H_{\mathcal{A}}^n)$, $\pi(F_{\mathcal{A}})$, $\pi(Y_{\mathcal{A}})$, and the affine multiview ideal, $\pi(M_{\mathcal{A}})$.

Lemma 6.2. *For ideals $I, J \subset \mathbb{C}[x_i, y_i, z_i]$, $\pi(I : J) = \pi(I) : \pi(J)$.*

Proof. If $f \in \pi(I : J)$, then $f = \pi(g)$ for some g which satisfies $gh \in I$ for all $h \in J$. Therefore $f\pi(h) = \pi(g)\pi(h) = \pi(gh) \in \pi(I)$ for any $h \in J$, proving $f \in \pi(I) : \pi(J)$. If $f \in \pi(I) : \pi(J)$, then for any $h \in J$, $f\pi(h) \in \pi(I)$, i.e., there exists $g \in I$ such that $f\pi(h) = \pi(g)$. Denote the homogenization of f with respect to $z_1, \dots, z_n$ by $\tilde{f}$. We claim that $\tilde{f} \in I : J$. Indeed for any $h \in J$, $\pi(\tilde{f}h) = \pi(\tilde{f})\pi(h) = f\pi(h) = \pi(g)$ for some $g \in I$. Homogenizing both sides, we get $\tilde{f}h = g \in I$, and we conclude that $\pi(I) : \pi(J) \subseteq \pi(I : J)$. $\square$

Corollary 6.3. *If $\mathcal{A}$ is a camera arrangement with pairwise distinct foci, then $\pi(M_{\mathcal{A}}) = \pi(H_{\mathcal{A}}^n) = \pi(F_{\mathcal{A}}) = \pi(\sqrt{Y_{\mathcal{A}}})$.*

Proof. Lemma 6.2 implies that $\pi(I : \mathfrak{m}) = \pi(I) : (1) = \pi(I)$ for any ideal I . Dehomogenizing Theorems 4.9, 4.11, and 4.7, each equality follows. $\square$

Observe that the last equality in Corollary 6.3 requires $A_1 = [I \ 0]$. Geometrically, Corollary 6.3 shows that while

the homogenous ideals $H_{\mathcal{A}}^n$, $F_{\mathcal{A}}$, $Y_{\mathcal{A}}$, and $M_{\mathcal{A}}$ do not coincide, they are the same away from the origin in each image plane. In particular, this is the case on the affine patch $\{\mathbf{p} \in \mathbb{P}^{2n} : z_1 = \dots = z_n = 1\}$ corresponding to finite image data.

Using Theorem 5.6 we see that, when $\mathcal{A}$ is noncoplanar, the dehomogenized bifocals alone suffice to generate the affine multiview ideal $\pi(M_{\mathcal{A}})$.

Corollary 6.4. *Suppose $\mathcal{A}$ is a noncoplanar camera arrangement with pairwise distinct foci. Then*

$$\pi(M_{\mathcal{A}}) = \pi(H_{\mathcal{A}}^2).$$

Proof. Dehomogenizing the result of Theorem 5.6, we get $\pi(M_{\mathcal{A}}) = \pi(H_{\mathcal{A}}^2 : \mathfrak{m}) = \pi(H_{\mathcal{A}}^2) : \pi(\mathfrak{m}) = \pi(H_{\mathcal{A}}^2)$. $\square$

Corollary 6.4 shows that $\pi(M_{\mathcal{A}})$ is generated by quadratics whenever $\mathcal{A}$ satisfies the noncoplanarity assumption. This observation was used in [11] to create a semidefinite programming relaxation of the triangulation problem which is can be seen as minimizing euclidean distance from an observed noisy data point to the affine multiview variety. It was shown that when the noise is small, the semidefinite relaxation solves triangulation. Of course, Corollary 6.3 needs the foci of the cameras to be noncoplanar and indeed, the experiments in [11] show that the quality of the semidefinite programming solution deteriorates as the foci become coplanar and then collinear.

Geometrically, we can understand how the quality of the relaxation deteriorates because the bifocal ideal cuts out more than the multiview variety for coplanar arrangements. In the coplanar case, the bifocal ideal cuts out the image of the plane that contains the camera centers. These points are not the images of true 3D points. It is therefore possible that the nearest point problem yields a spurious solution on this extra component. Similarly, in the collinear case, the bifocal ideal cuts out a strictly larger variety than just the multiview variety. In this case, the dimension of the vanishing set of the bifocal ideal is one larger than the multiview variety.

7 SUMMARY

The multiview variety is a foundational geometric object in multiview geometry and understanding its vanishing ideal $M_{\mathcal{A}}$ precisely is important for any algebraic algorithm that solves problems on this variety. There have been many partial results about the algebraic structure of the multiview variety. The aim of our paper is to put them all into a unified algebraic setting and give a complete description of $M_{\mathcal{A}}$.

Our main result is that when the foci of the cameras are pairwise distinct, $M_{\mathcal{A}}$ is generated by the bifocal and trifocal polynomials of $\mathcal{A}$ (Theorem 3.7). The proof requires an understanding of the behavior of coordinate changes on k -focal ideals (Lemma 2.5), and translational cameras (Lemma 3.3). The main result holds for euclidean cameras as well (Corollary 3.8). We also give an example to illustrate that the assumption of distinct foci cannot be relaxed for this result to hold (Example 3.10).

Next we study three sets of polynomials that have been proposed to cut out the multiview variety, by Heyden-

Åström, Faugeras and Ma et. al. respectively. We show that the ideals generated by these polynomials are all properly contained in M_A . We establish the exact algebraic relationships between the above ideals and M_A (Theorems 4.7, 4.9 and 4.11).

We then prove that if the camera foci are assumed to be noncoplanar, then in fact M_A is the saturation of the bifocal ideal by the irrelevant ideal (Theorem 5.6). In this situation the n -focal ideal is a subset of the bifocal ideal.

Finally we prove that the dehomogenization of the ideals by Heyden-Åström, Faugeras and Ma et. al. all agree with the dehomogenization of M_A (Corollary 6.3). Similarly, under noncoplanarity of foci, the bifocal ideal also has the same dehomogenization (Corollary 6.4). This means that all of these ideals cut out the space of finite images.

TECHNICAL PROOFS

A.1 Multigraded Projective Nullstellensatz

In this section, we state and prove the projective Nullstellensatz in our multigraded setting, which we use to prove Theorem 4.7 in Section 4. Let $I \subseteq \mathbb{C}[p_1, \dots, p_n]$ be homogeneous with respect to the $\mathbb{Z}^n$ -grading $\deg w_i = \mathbf{e}_i$. To be clear about projective versus affine varieties, we define $V_P(I) := V(I) = \{\mathbf{p} \in (\mathbb{P}^2)^n : f(\mathbf{p}) = 0 \text{ for all } f \in I\}$, and for a set $S \subseteq (\mathbb{P}^2)^n$, we define

$$I_P(S) = \{f \in \mathfrak{m} : f(\mathbf{p}) = 0 \text{ for all } \mathbf{p} \in S\}.$$

We say that $V_P(I)$ is the projective vanishing set of I in $(\mathbb{P}^2)^n$ and $I_P(S)$ is the largest homogeneous ideal vanishing on S contained in $\mathfrak{m}$. While we force $I_P(S) \subseteq \mathfrak{m}$, it also makes sense to consider the largest homogeneous ideal vanishing on S without intersecting with $\mathfrak{m}$. As before we denote this ideal by $I(S)$, and notice that $I_P(S) = I(S) \cap \mathfrak{m}$. In the usual grading on $\mathbb{C}[p_1, \dots, p_n]$, a vanishing ideal $I(S)$ is homogeneous in the usual sense which means that it is contained in the usual irrelevant ideal $\langle x_1, y_1, z_1, \dots, x_n, y_n, z_n \rangle$. Under the multi-grading, $I_P(S)$ is required to be in the corresponding irrelevant ideal $\mathfrak{m}$. We will use the following variant of the Nullstellensatz.

Lemma A.1. *For any homogeneous ideal $I \subseteq \mathbb{C}[p_1, \dots, p_n]$ such that $I \subseteq \mathfrak{m}$, $I_P(V_P(I)) = \sqrt{I}$.*

Proof. Define the affine operations

$$\begin{aligned} V_A(I) &= \{\mathbf{p} \in (\mathbb{A}^3)^n : f(\mathbf{p}) = 0 \text{ for all } f \in I\} \\ I_A(S) &= \{f \in \mathbb{C}[p_1, \dots, p_n] : f(\mathbf{p}) = 0 \text{ for all } \mathbf{p} \in S\} \end{aligned}$$

where we treat S as a subset of $(\mathbb{A}^3)^n$. We will use the affine version of the Nullstellensatz on the cone over $V := V_P(I)$, i.e., the set $C_V = V_A(I) \subseteq (\mathbb{A}^3)^n$. We claim that

$$I_A(C_V) = I_P(V). \quad (11)$$

First suppose $f \in I_A(C_V)$. Given $\mathbf{p} = (p_1, \dots, p_n) \in V$, all homogeneous coordinates of $\mathbf{p}$, represented by scalings $(\lambda_1 p_1, \dots, \lambda_n p_n)$, lie in C_V , so f vanishes for all homogeneous coordinates of $\mathbf{p}$. This means that the homogeneous components $f_{i_1, \dots, i_n}$ of f , consisting of all terms with

multidegree $(i_1, \dots, i_n)$, vanish at $\mathbf{p}$, so $f \in I(V)$, hence $I_A(C_V) \subseteq I(V)$. By the Nullstellensatz in $(\mathbb{A}^3)^n$, $I_A(C_V) = I_A(V_A(I)) = \sqrt{I}$, and by the assumption that $I \subseteq \mathfrak{m}$, $\sqrt{I} \subseteq \sqrt{\mathfrak{m}} = \mathfrak{m}$. This shows that $I_A(C_V) \subseteq I(V) \cap \mathfrak{m} = I_P(V)$.

Conversely, suppose $f \in I_P(V)$. Since any point $\mathbf{p}$ of C_V such that $p_i \neq 0$ for all i gives homogeneous coordinates for a point in V , it follows that f vanishes on $C_V \setminus \bigcup_{i=1}^n \mathbb{A}^3 \times \dots \times \{0\}_i \times \dots \times \mathbb{A}^3$. We need to show that f vanishes on each of the sets $\mathbb{A}^3 \times \dots \times \{0\}_i \times \dots \times \mathbb{A}^3$. Since $f \subseteq \mathfrak{m}$, it has strictly positive multidegree, and every monomial in f contains at least one coordinate from each copy of $\mathbb{A}^3$. Setting all 3 coordinates to zero in any $\mathbb{A}^3$ forces f to be zero, so we conclude that $f \in I_A(C_V)$. Finally, from (11), we conclude

$$\sqrt{I} = I_A(V_A(I)) = I_A(C_V) = I_P(V) = I_P(V_P(I)). \quad \square$$

Corollary A.2. *For any homogeneous ideal $I \subseteq \mathbb{C}[p_1, \dots, p_n]$, $I_P(V_P(I)) = \sqrt{I} \cap \mathfrak{m}$.*

Proof. Observe that

$$V_P(I \cap \mathfrak{m}) = V_P(I) \cup V_P(\mathfrak{m}) = V_P(I)$$

and

$$\sqrt{I \cap \mathfrak{m}} = \sqrt{I} \cap \sqrt{\mathfrak{m}} = \sqrt{I} \cap \mathfrak{m} \subseteq \mathfrak{m}$$

Therefore by Lemma A.1, $I_P(V_P(I)) = \sqrt{I} \cap \mathfrak{m}$. $\square$

Corollary A.3. *For any A with pairwise distinct foci,*

$$M_A \cap \mathfrak{m} = \sqrt{H_A^n} \cap \mathfrak{m} = \sqrt{F_A} \cap \mathfrak{m} = \sqrt{Y_A} \cap \mathfrak{m}.$$

Proof. We have already shown in Section 4 that $V_P(H_A^n) = V_P(F_A) = V_P(Y_A) = V_P(M_A)$. Since M_A is radical, the result follows by Corollary A.2. $\square$

We can now prove Theorem 4.7, restated here, from the main body of the paper.

Theorem 4.7. *For any A with pairwise distinct foci,*

- (1) $\sqrt{H_A^n} : \mathfrak{m} = M_A$
- (2) $\sqrt{F_A} : \mathfrak{m} = M_A$
- (3) $\sqrt{Y_A} : \mathfrak{m} = M_A$ when $A_1 = [I \mid 0]$

Proof. Taking colon ideal with $\mathfrak{m}$, the desired result follows from Corollary A.3 and the fact that $M_A : \mathfrak{m} = M_A$, which was proven in Theorem 4.9. $\square$

A.2 Proof of Lemma 4.10

In this section, we elaborate on the technical details used to prove Theorem 4.11. Recall that the nontrivial statement there was that bifocals and trifocals can be multiplied by any generator of $\mathfrak{m}$ to fall into F_A . This requires understanding the 4×4 minors of $\mathcal{A}^F(p)$ for which we once again invoke the Cauchy-Binet formula and the observation that $\mathcal{A}^F(p) = P(p)A$ from (7).

First we characterize certain 4×4 minors of $P(p)$. Let p_{ij} denote the j th coordinate of p_i , i.e., $p_{i1} = x_i$, $p_{i2} = y_i$, and $p_{i3} = z_i$. Having the subscript (resp. superscript) p_{ij} on $P(p)$ indicates eliminating from $P(p)$ the unique row (resp. column)

of $[p_i]_\times$ that does not contain p_{ij} . On the other hand, having the subscript p_{ij} on the matrices $\mathcal{A}$ and $\mathcal{A}(p)$ will stand for eliminating the unique row of the matrix containing p_{ij} .

We will only need to consider the 4×4 minors of $P(p)$ when $n = 2$ and $n = 3$. Let $R_i, C_i \subseteq \{p_{i1}, p_{i2}, p_{i3}\}$ denote collections of coordinates, and write $R = \bigcup_{i=1}^n R_i, C = \bigcup_{i=1}^n C_i$. When $n = 2$, a 4×4 minor of $P(p)$ is $\det(P(p)_R^C)$ for some R, C of size $|R| = |C| = 2$, and when $n = 3$, $|R| = |C| = 5$. Observe that if $|R_i| \neq |C_i|$ for any i , then the submatrix $P(p)_R^C$ has at least two linearly dependent rows or columns, yielding a zero minor. When $|R_i| = |C_i|$ for all i , $P(p)_R^C$ is block diagonal, so $\det(P(p)_R^C) = \prod_{i=1}^n \det([p_i]_\times)_{R_i}^{C_i}$.

Lemma A.2. *Let $n = 2$. The nonzero 4×4 minors of $P(p)$ are determined by collections of coordinates R, C with $|R_1| = |C_1| = |R_2| = |C_2| = 1$. For $R = \{p_{1j}, p_{2k}\}$ and $C = \{p_{1l}, p_{2m}\}$, the 4×4 minor $\det(P(p)_R^C)$ is the monomial*

$$\det(P(p)_R^C) = (-1)^{j+k+l+m} p_{1j} p_{2k} p_{1l} p_{2m}.$$

Proof. As noted above, if $|R_i| \neq |C_i|$ for either i , then $\det(P(p)_R^C) = 0$, whereas if $|R_i| = |C_i| = 2$ for either i , then $P(p)_R^C$ has a rank 2 block on its diagonal, hence $\det(P(p)_R^C) = 0$, proving the first statement. For $R = \{p_{1j}, p_{2k}\}$ and $C = \{p_{1l}, p_{2m}\}$, the 4×4 minor $\det(P(p)_R^C)$ is

$$\begin{aligned} \det(P(p)_R^C) &= \det([p_1]_\times)_{R_1}^{C_1} \det([p_2]_\times)_{R_2}^{C_2} \\ &= ((-1)^{j+l} p_{1j} p_{1l}) ((-1)^{k+m} p_{2k} p_{2m}) \\ &= (-1)^{j+k+l+m} p_{1j} p_{2k} p_{1l} p_{2m}. \end{aligned}$$

□

Lemma A.3. *Let $n = 3$. Suppose $|R_3| = |C_3| = 1$, and $|R_1| = |C_1| = |R_2| = |C_2| = 2$. For $R_3 = \{p_{3j}\}, C_3 = \{p_{3k}\}$, the 4×4 minor $\det(P(p)_R^C)$ is the monomial*

$$\begin{cases} (-1)^{j+k+l+m} p_{3j} p_{3k} p_{1l} p_{2m} & \text{if } R_1 \neq C_1, R_2 \neq C_2 \\ 0 & \text{otherwise.} \end{cases}$$

where p_{1l} is the coordinate common to R_1 and C_1 and p_{2m} is the coordinate common to R_2 and C_2 .

Proof. When $R_i = C_i$ as sets for $i = 1$ or $i = 2$, then $([p_i]_\times)_{R_i}^{C_i} = 0$, hence $\det(P(p)_R^C) = \prod_{i=1}^n \det([p_i]_\times)_{R_i}^{C_i} = 0$. On the other hand, when $R_1 \neq C_1$, $\det([p_1]_\times)_{R_1}^{C_1} = (-1)^l p_{1l}$ where $p_{1l} = R_1 \cap C_1$. Similarly $\det([p_2]_\times)_{R_2}^{C_2} = (-1)^m p_{2m}$ where $p_{2m} = R_2 \cap C_2$ when $R_2 \neq C_2$. □

We now show that bifocals and trifocals can both be multiplied by any generator of $\mathfrak{m}$ to fall into $F_{\mathcal{A}}$.

Lemma A.4.

- (1) For $n = 2$ cameras, and any monomial $p_{1j} p_{2k}$, there exists a 4×4 minor f of $\mathcal{A}^F(p)$ such that $f = (-1)^{j+k} p_{1j} p_{2k} \det(\mathcal{A}(p))$.
- (2) Let $n = 3$ and i_1, i_2, i_3 be pairwise distinct. Then for any trifocal $\det(\mathcal{A}(p)_{\{p_{i_1 j_1}, p_{i_2 j_2}\}})$ and any coordinate $p_{i_3 k}$, there exists a 4×4 minor f of $\mathcal{A}^F(p)$ such that $f = (-1)^k p_{i_3 k} \det(\mathcal{A}(p)_{\{p_{i_1 j_1}, p_{i_2 j_2}\}})$.

Proof. (a) Fix some $p_{1j} p_{2k}$. Since $n = 2$, $P(p)\mathcal{A}$ is a 6×4 matrix and we need to delete two rows to get a 4×4

minor. Using Lemma A.2 and Cauchy-Binet, the result follows from the computation below:

$$\begin{aligned} f &= \det(P(p)_{\{p_{1j}, p_{2k}\}} \mathcal{A}) \\ &= \sum_{|C|=2} \det(P(p)_{\{p_{1j}, p_{2k}\}}^C) \det(\mathcal{A}_C) \\ &= \sum_{|C_1|=|C_2|=1} \det(P(p)_{\{p_{1j}, p_{2k}\}}^C) \det(\mathcal{A}_C) \\ &= \sum_{1 \leq l, m \leq 3} \det([p_1]_\times)_{\{p_{1j}\}}^{C_1} \det([p_2]_\times)_{\{p_{2k}\}}^{C_2} \\ &\quad \times \det(\mathcal{A}_{\{p_{1l}, p_{2m}\}}) \\ &= \sum_{1 \leq l, m \leq 3} (-1)^{j+k+l+m} p_{1j} p_{2k} p_{1l} p_{2m} \det(\mathcal{A}_{\{p_{1l}, p_{2m}\}}) \\ &= (-1)^{j+k} p_{1j} p_{2k} \sum_{1 \leq l, m \leq 3} (-1)^{l+m} p_{1l} p_{2m} \det(\mathcal{A}_{\{p_{1l}, p_{2m}\}}) \\ &= (-1)^{j+k} p_{1j} p_{2k} \det(\mathcal{A}(p)). \end{aligned}$$

where the last equality follows from expanding the determinant of $\mathcal{A}(p)$ along the last two columns.

(b) Without loss of generality, let $i_1 = 1, i_2 = 2, i_3 = 3$ and let p_{3k} be arbitrary. For simplicity, suppose $j_1 = j_2 = 1$. Therefore, we consider the trifocal $\det(\mathcal{A}(p)_{\{p_{11}, p_{21}\}})$. Using Lemma A.3 and Cauchy-Binet, we expand $f = \det(P(p)_R^C \mathcal{A})$ where $R_1 = \{p_{12}, p_{13}\}, R_2 = \{p_{22}, p_{23}\}, R_3 = \{p_{3k}\}$ as follows:

$$\begin{aligned} f &= \det(P(p)_R^C \mathcal{A}) \\ &= \sum_{C: |C_1|=|C_2|=2, |C_3|=1} \det(P(p)_{\{p_{12}, p_{13}, p_{22}, p_{23}, p_{3k}\}}^C) \det(\mathcal{A}_C) \\ &= \sum_{|C_3|=1} \det([p_3]_\times)_{R_3}^{C_3} \times \sum_{|C_1|=|C_2|=2} \left(\det([p_1]_\times)_{R_1}^{C_1} \right. \\ &\quad \times \det([p_2]_\times)_{R_2}^{C_2} \det(\mathcal{A}_C) \Big) \\ &= \sum_{i=1}^3 (-1)^{i+k} p_{3k} p_{3i} \times \sum_{\substack{|C_1|=|C_2|=2 \\ C_1 \neq R_1, C_2 \neq R_2}} \left(\det([p_1]_\times)_{R_1}^{C_1} \right. \\ &\quad \times \det([p_2]_\times)_{R_2}^{C_2} \det(\mathcal{A}_{\{C_1, C_2, p_{3i}\}}) \Big) \\ &= (-1)^k p_{3k} \sum_{i=1}^3 (-1)^i p_{3i} \times \sum_{2 \leq l, m \leq 3} \left(\det([p_1]_\times)_{\{p_{11}, p_{1l}\}}^{C_1} \right. \\ &\quad \times \det([p_2]_\times)_{\{p_{21}, p_{2m}\}}^{C_2} \det(\mathcal{A}_{\{p_{11}, p_{1l}, p_{21}, p_{2m}, p_{3i}\}}) \Big) \\ &= (-1)^k p_{3k} \sum_{i=1}^3 (-1)^i p_{3i} \\ &\quad \times \sum_{2 \leq l, m \leq 3} (-1)^{l+m} p_{1l} p_{2m} \det(\mathcal{A}_{\{p_{11}, p_{1l}, p_{21}, p_{2m}, p_{3i}\}}) \\ &= (-1)^k p_{3k} \det(\mathcal{A}(p)_{\{p_{11}, p_{21}\}}) \end{aligned}$$

Observe that the final equality follows from expanding the determinant of $\mathcal{A}(p)_{\{p_{11}, p_{21}\}}$ on the p_3 column.

For general j_1, j_2 , performing the same computation with $R_1 = \{p_{11}, p_{12}, p_{13}\} \setminus \{p_{1j_1}\}$, $R_2 = \{p_{21}, p_{22}, p_{23}\} \setminus \{p_{2j_2}\}$ and $R_3 = \{p_{3k}\}$ yields $\det(P(p)_{R\mathcal{A}}) = (-1)^k p_{3k} \det(\mathcal{A}(p)_{\{j_1, j_2\}})$. $\square$

ACKNOWLEDGMENTS

The authors would like to thank the referees of this article for their careful reading and suggestions. In particular, their comments helped fill a gap in the proof of the main theorem of Section 5. Andrew Pryhuber and Rekha R. Thomas acknowledge support from the U.S. National Science Foundation through the grant DMS-1719538.

REFERENCES

- [1] B. Triggs, "Matching constraints and the joint image," in *Proc. IEEE Int. Conf. Comput. Vision*, 1995, pp. 338–343.
- [2] A. Heyden and K. Åström, "Algebraic properties of multilinear constraints," *Math. Methods Appl. Sci.*, vol. 20, pp. 1135–1162, Sep. 1997.
- [3] D. A. Cox, J. Little, and D. O'Shea, *Ideals, Varieties, and Algorithms*, 4th ed., Berlin, Germany: Springer, 2015.
- [4] C. Aholt, B. Sturmfels, and R. Thomas, "A Hilbert scheme in computer vision," *Can. J. Math.*, vol. 65, no. 5, pp. 961–988, 2013.
- [5] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, no. 5828, 1981, Art. no. 133.
- [6] S. Maybank, *Theory of Reconstruction from Image Motion*. Berlin, Germany: Springer-Verlag, 1993.
- [7] O. Faugeras, Q.-T. Luong, and T. Papadopolou, *The Geometry of Multiple Images: The Laws that Govern the Formation of Images of a Scene and Some of Their Applications*. Cambridge, MA, USA: MIT Press, 2001.
- [8] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [9] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*. Berlin, Germany: Springer, 2012.
- [10] F. Kahl and D. Henrion, "Globally optimal estimates for geometric reconstruction problems," *Int. J. Comput. Vision*, vol. 74, no. 1, pp. 3–15, 2007.
- [11] C. Aholt, S. Agarwal, and R. Thomas, "A QCQP approach to triangulation," in *Proc. Eur. Conf. Comput. Vision*, 2012, pp. 654–667.
- [12] G. Blekherman, P. A. Parrilo, and R. R. Thomas, *Semidefinite Optimization and Convex Algebraic Geometry*. Philadelphia, PA, USA: SIAM, 2012.
- [13] J. B. Lasserre, "Global optimization with polynomials and the problem of moments," *SIAM J. Optim.*, vol. 11, no. 3, pp. 796–817, 2000.
- [14] P. A. Parrilo, "Semidefinite programming relaxations for semialgebraic problems," *Algebr. Geometric Methods Discrete Optim.*, vol. 96, no. 2, pp. 293–320, 2003. [Online]. Available: <https://doi.org/offcampus.lib.washington.edu/10.1007/s10107-003-0387-5>
- [15] D. R. Grayson and M. E. Stillman, "Macaulay 2, a software system for research in algebraic geometry," 2002. [Online]. Available: <https://faculty.math.illinois.edu/Macaulay2/>
- [16] M. Trager, M. Hebert, and J. Ponce, "The joint image handbook," in *Proc. IEEE Int. Conf. Comput. Vision*, 2015, pp. 909–917.
- [17] M. Trager, M. Hebert, and J. Ponce, "The joint image handbook supplementary material," in *Proc. IEEE Int. Conf. Comput. Vision*, Dec. 2015, pp. 909–917.
- [18] J. G. Broida and S. G. Williamson, *A Comprehensive Introduction to Linear Algebra*. Reading, MA, USA: Addison-Wesley, 1989.
- [19] M. Marshall, *Positive Polynomials and Sums of Squares*. Providence, RI, USA: American Mathematical Society, 2008.
- [20] A. Conca, E. De Negri, and E. Gorla, "Cartwright-Sturmfels ideals associated to graphs and linear spaces," *J. Comb. Algebra*, vol. 2, no. 3, pp. 231–257, 2018.
- [21] B. Li, "Images of rational maps of projective spaces," *Int. Math. Res. Notices*, vol. 2018, no. 13, pp. 4190–4228, 2018.
- [22] O. Faugeras and B. Mourrain, "On the Geometry and Algebra of the point and line correspondences between n images," in *Proc. IEEE Int. Conf. Comput. Vision*, 1995, pp. 951–956.
- [23] J. Harris, *Algebraic Geometry: A First Course*. Berlin, Germany: Springer, 1995.
- [24] M. Trager, B. Osserman, and J. Ponce, "On the solvability of viewing graphs," in *Proc. Eur. Conf. Comput. Vision*, 2018.



Sameer Agarwal received the PhD degree in computer science from the University of California, San Diego, in 2006. He is currently an engineer with Google and an affiliate professor with the University of Washington. His research interests include computer vision, optimization, and algebra.



Andrew Pryhuber received the BA degree in mathematics from Bowdoin College, Brunswick, Maine. He is currently working toward the doctoral degree in the Department of Mathematics, University of Washington, Seattle, WA. His research interests include computational algebraic geometry, computer vision, and optimization.



Rekha R. Thomas received the PhD degree in operations research from Cornell University, NY, US, in 1994. She is currently a professor of mathematics with the University of Washington, Seattle, WA. Her research interests include optimization and applied algebraic geometry.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/csdl.