

A Glimpse at Algebraic Vision: An Introduction to the Multiview Variety

Hans Elliott

June 2026

Contents

1	Introduction	2
2	Background	2
2.1	Pinhole Camera Model	2
2.2	Real Projective Space	3
2.3	Projective Algebraic Geometry	3
2.4	Projective Pinhole Camera	4
3	Triangulation with Two Images	4
3.1	Epipolar Geometry and Triangulation	4
3.2	Algebraic Varieties in Epipolar Geometry	6
3.2.1	Epipolar Constraint Revisited	6
3.2.2	Normalized Coordinate System	6
3.2.3	The Two-View Variety	7
3.3	Implications for Triangulation	8
4	References	9
5	Appendix	10

1 Introduction

Computer vision, the science of teaching computers to see, pulls from a wide range of mathematics to solve its problems, including algebraic geometry. After covering some background material, we will explore a particular problem in computer vision, known as “triangulation,” where three-dimensional structure is reconstructed from two or more images of a scene. An important step in triangulation is identifying corresponding points in the images, such as a particular feature that appears in both images (see appendix for examples). In theory, triangulation is simple to perform once these point correspondences are obtained, but in practice, the measuring process is subject to error and the points need to be “de-noised.” We will see that algebraic geometry can be applied to study the mathematical objects that arise from this problem.

2 Background

2.1 Pinhole Camera Model

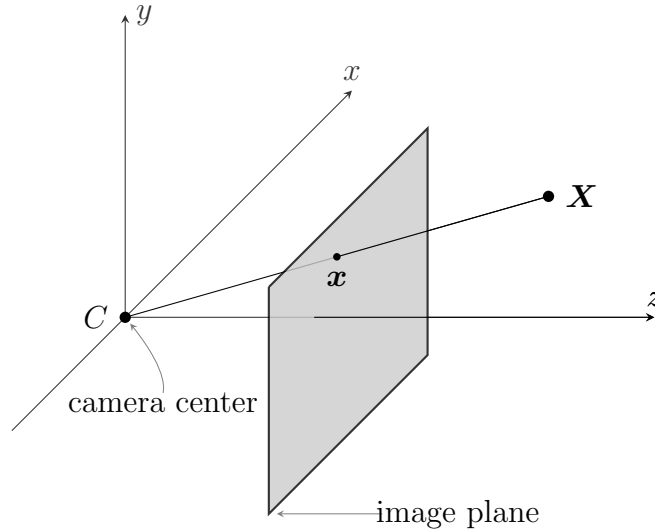


Figure 1: A pinhole camera with camera center C at the origin of $\mathbb{R}^3$. The world point $\mathbf{X}$ is projected centrally to the image point $\mathbf{x}$ on the image plane.

A basic model of a camera is given by the “pinhole model,” which is a central projection of the three-dimensional world space onto a two-dimensional image. Consider a world point $\mathbf{X} = (X, Y, Z) \in \mathbb{R}^3$. The **camera center** is assumed to be at the origin of $\mathbb{R}^3$ and the camera’s **image plane** is assumed to be the $z = 1$ plane. Unless the point $\mathbf{X}$ lies in the (x, y) -plane (so that $Z = 0$), the line passing through both the camera center and $\mathbf{X}$ will intersect the image plane at a unique point. Letting ℓ be this line and parameterizing it as $\ell(t) = (tX, tY, tZ)$ for $t \in \mathbb{R}$, we see that $\ell(t)$ intersects the $z = 1$ plane precisely when $t = 1/Z$, so that the point in the image plane corresponding to $\mathbf{X}$ is given by $(X/Z, Y/Z, 1)$.

Thus, we can think of the camera as a map from $\{(x, y, z) \in \mathbb{R}^3 \mid z \neq 0\}$ to $\mathbb{R}^2$ given by $(X, Y, Z) \mapsto (X/Z, Y/Z)$. See Hartley and Zisserman [HZ04] §6.1 for more details.

2.2 Real Projective Space

A given line in $\mathbb{R}^2$ is defined by the equation $ax + by + c = 0$ for varying $(x, y) \in \mathbb{R}^2$ and thus can be represented by the vector (a, b, c) . Since any points satisfying this equation will also satisfy $\lambda(ax + by + c) = 0$ for any $\lambda \in \mathbb{R} \setminus \{0\}$, the vectors (a, b, c) and $(\lambda a, \lambda b, \lambda c)$ represent the same line. An equivalence class of vectors $[(a, b, c)]$ under the equivalence relation that $(a, b, c) \sim (d, e, f)$ if and only if $(a, b, c) = \lambda(d, e, f)$ for some $\lambda \in \mathbb{R} \setminus \{0\}$, is known as a **homogeneous vector**. The set of equivalence classes of vectors in $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ is called the **projective space**, $\mathbb{P}^2$, where $(0, 0, 0)$ is excluded since it does not correspond to any line. A point $(x, y) \in \mathbb{R}^2$ lies on a line (a, b, c) if it satisfies $ax + by + c = 0$, which can be written in terms of the inner product $(x, y, 1)(a, b, c)^T = 0$, so that (x, y) is represented as a three-vector by $(x, y, 1)$. The vector $(\lambda x, \lambda y, \lambda)$ will also satisfy this equation and is thus an equivalent representation of the point. Generally, points in $\mathbb{P}^2$ are given by homogeneous three-vectors (x, y, z) which represent points $(x/z, y/z) \in \mathbb{R}^2$. However, if $z = 0$ the homogeneous point can not be simply converted to a representation in $\mathbb{R}^2$. Points with a third coordinate of 0 are called “points at ∞ ” or “ideal points” and, while an important aspect of projective space, are not directly relevant here. The key fact we will use about $\mathbb{P}^2$ is that points (and lines) are given by three-vectors (or “homogeneous coordinates”) that are defined only up to scale. Similarly, we can define $\mathbb{P}^3$ as the set of equivalence classes of vectors in $\mathbb{R}^4 \setminus \{(0, 0, 0, 0)\}$ in which points are specified by homogeneous four-vectors. More details can be found in [HZ04] §2.2 and [CLO15] §8.

2.3 Projective Algebraic Geometry

We can generalize the study of affine varieties to projective space, but since points in $\mathbb{P}^n$ are defined by homogeneous vectors, we must take care in how we define projective varieties. A polynomial is **homogeneous** if each of its terms has the same total degree. Then for homogeneous polynomials $f_1, \dots, f_s$, we say $\mathbf{V}(f_1, \dots, f_s)$ is the **projective variety** defined by $f_1, \dots, f_s$. In [CLO15] §8.2, it is shown that if a homogeneous polynomial vanishes at one set of homogeneous coordinates for a point in $\mathbb{P}^n$, then it vanishes at all homogeneous coordinates, and thus projective varieties are well-defined. For example, $f = xyz - z^3$ is homogeneous of total degree 3 but $g = xyz - z$ is not homogeneous. Note that the point $(1, 1, 1) \in \mathbb{P}^2$ satisfies $g(1, 1, 1) = 0$, but the equivalent homogeneous coordinates $(2, 2, 2)$ yield $g(2, 2, 2) = 6 \neq 0$. Meanwhile, $f(1, 1, 1) = f(2, 2, 2) = 0$. We can also define **homogeneous ideals**, which can be thought of as ideals generated by homogeneous polynomials ([CLO15] §8.3). Then, the map $\mathbf{I}$ takes projective varieties to homogeneous ideals and the map $\mathbf{V}$ takes homogeneous ideals to projective varieties. Importantly, $\mathbf{V}(\mathbf{I}(V)) = V$ holds for any projective variety V , just as it holds for affine varieties.

2.4 Projective Pinhole Camera

It is useful to extend the pinhole camera model to projective space, so that cameras can be represented by matrices. Let $\mathbf{X}$ be a world point, which we will model as a point in $\mathbb{P}^3$ by writing $\mathbf{X} = (X, Y, Z, 1)$. Then a projective camera is a map $A : \mathbb{P}^3 \rightarrow \mathbb{P}^2$ represented by a 3×4 **camera matrix**, A , of rank 3. The process of taking an image is modeled by the equation $A\mathbf{X} = \lambda\mathbf{x}$, where the world point is mapped to an image point $\mathbf{x} \in \mathbb{P}^2$, with $\lambda \in \mathbb{R}$ a scalar that is introduced due to the homogeneity of the coordinates. The elements of the camera matrix encode the properties of the camera, such as its orientation and focal length. As an example, we can extend the pinhole camera described above. The camera center lies at $(0, 0, 0, 1) \in \mathbb{P}^3$ and the central projection of $\mathbf{X}$ is given by

$$A\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}.$$

One important note about the projective camera map is that it is undefined at the camera center. If $c \in \mathbb{P}^3$ is the camera center, then A will always map c to $(0, 0, 0)$, which is not an element of $\mathbb{P}^2$. This is shown in [HZ04] §6.2.

3 Triangulation with Two Images

We now consider the triangulation problem and its relation to algebraic geometry. Suppose we have two pictures of the world point $\mathbf{X} = (X, Y, Z, 1) \in \mathbb{P}^3$ from projective pinhole cameras at two distinct positions, so that $\mathbf{X}$ appears projected onto each image. Letting A and A' be the known 3×4 projective camera matrices, we have the corresponding image points $\mathbf{x}, \mathbf{x}' \in \mathbb{P}^2$ satisfying $\lambda\mathbf{x} = A\mathbf{X}$ and $\lambda'\mathbf{x}' = A'\mathbf{X}$ for some $\lambda, \lambda' \neq 0$. Our goal is to determine the coordinates of $\mathbf{X}$ from the information in $\mathbf{x}, \mathbf{x}', A, A'$. By repeating this process for many world points, we could create a three-dimensional model of the scene from the two images. Given the cameras and perfectly measured image points, the process of triangulation is theoretically straightforward. However, in practice, the measuring of point correspondences between two images is subject to error, and points need to be “de-noised”. To understand why, we consider the geometry of the two views of $\mathbf{X}$, which is known as epipolar geometry.

3.1 Epipolar Geometry and Triangulation

Figure 2 below models our setup. We have two cameras which have imaged the same world point, $\mathbf{X}$. Let c and c' be their camera centers and H and H' be their image planes. The *baseline* is the line in world space which connects the camera centers. The **epipole** e is the point of intersection of the baseline with the image plane of the first camera, H , and e' is defined similarly for the second image. Note that since the epipole e lies on the line passing through the other camera’s center, c' , which intersects H at e , it follows that e is the image

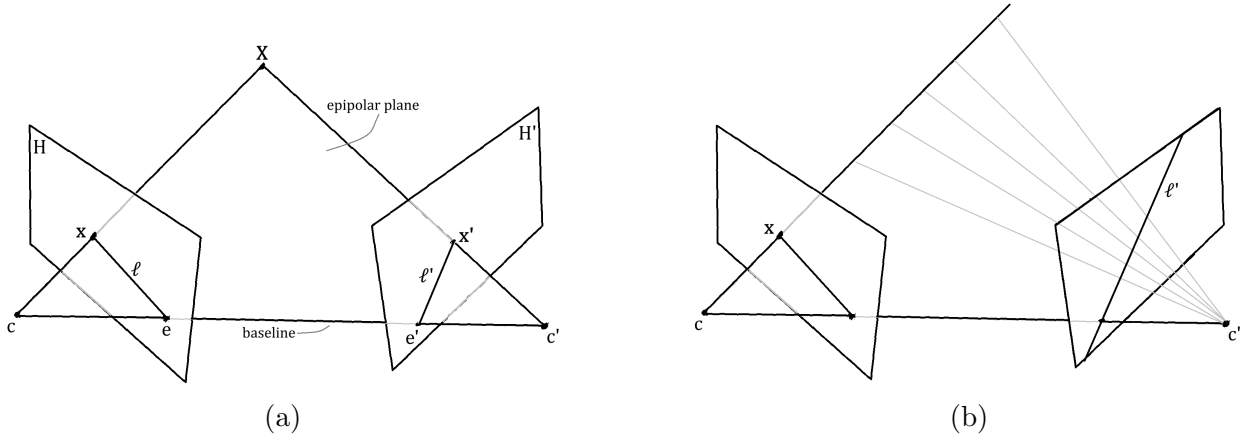


Figure 2: The geometry of corresponding points in two images.

of c' under the first camera. This implies that $Ac' = e$, and similarly, $A'c = e'$. The world point $\mathbf{X}$, image points $\mathbf{x}$ and $\mathbf{x}'$, and camera centers c and c' all lie in a common plane, which we call the *epipolar plane*. The **epipolar line** ℓ is the intersection of the epipolar plane with the first image plane, and ℓ' is defined similarly for the second image.

Examining Figure 2 (b), suppose we know the positions of the camera centers and image planes, as well as the image point $\mathbf{x}$, but do not know the position of $\mathbf{X}$ or $\mathbf{x}'$. The world point $\mathbf{X}$ must lie on the line $\overline{c\mathbf{x}}$ that passes through c and $\mathbf{x}$ by the definition of a pinhole camera, and thus $\mathbf{x}'$ must lie on a line that intersects with $\overline{c\mathbf{x}}$ and the other camera center, c' . Any such line must intersect the epipolar line ℓ' , so $\mathbf{x}'$ must lie on ℓ' . Similarly, $\mathbf{x}$ is constrained to lie on ℓ . Then, there is a map from $\mathbf{x}$ to ℓ' and similarly from $\mathbf{x}'$ to ℓ , which takes points in one image to lines in the other image. This map is represented by the **fundamental matrix**, denoted by F , which satisfies $\ell' = F^T \mathbf{x}$ and $\ell = F \mathbf{x}'$. Since F maps points in $\mathbb{P}^2$ to lines in $\mathbb{P}^2$, which are both represented by 3-vectors, we can see that F is a 3×3 matrix. It can also be shown that F must have rank 2 (a construction of F is given in [HZ04] §9.1). Note that since $\mathbf{x}$ lies on ℓ , it must be that $\mathbf{x}^T \ell = \mathbf{x}^T F \mathbf{x}' = 0$, which is known as the **epipolar constraint**. In fact, for any two images acquired by cameras at distinct centers, the fundamental matrix is the unique 3×3 , rank 2, homogeneous matrix F that satisfies $\mathbf{x}^T F \mathbf{x}' = 0$ for any pair of corresponding points $\mathbf{x}, \mathbf{x}'$ in the two images. Given camera matrices, one can compute F , and given F , the camera matrices can be retrieved. These topics are discussed in greater detail in [HZ04] §9.

In theory, triangulating $\mathbf{X}$ is simple since it sits at the intersection of two lines. From the equations $\mathbf{x} = A\mathbf{X}$ and $\mathbf{x}' = A'\mathbf{X}$, one can arrange and solve a system of linear equations which will give the coordinates of $\mathbf{X}$, as discussed in [HZ04] §12.2. However, our first step of measuring corresponding points in the two images is subject to error, so that the measured points may be a noisy representation of the true points. Suppose we have measured a pair of corresponding points $\mathbf{x}, \mathbf{x}'$ in the two images via some noisy measurement process (e.g., by hand or using a computer for automatic detection). Our cameras are known, so we can compute the fundamental matrix F and use it to test that the measured points do not exactly satisfy the epipolar constraint. This motivates the optimization problem of finding points

$\hat{\mathbf{x}}, \hat{\mathbf{x}}'$ which are close to $\mathbf{x}, \mathbf{x}'$ but which do satisfy the epipolar constraint. In [HZ04] §12, algorithms are given for solving this optimization problem. Here, we consider this problem from the perspective of algebraic geometry.

3.2 Algebraic Varieties in Epipolar Geometry

In [HÅ97], Heyden and Åström examine the two image triangulation problem from the perspective of algebraic geometry. Since the maps $A, A' : \mathbb{P}^3 \rightarrow \mathbb{P}^2$ are undefined at the camera centers, c and c' , as we saw in the section on projective cameras, they define $\dot{\mathbb{P}}^3 = \mathbb{P}^3 \setminus \{c, c'\}$. Then they inspect the joint map $\Phi = (A, A')$ where

$$\Phi : \dot{\mathbb{P}}^3 \rightarrow \mathbb{P}^2 \times \mathbb{P}^2$$

maps $\mathbf{X} \in \mathbb{P}^3$ to $(A\mathbf{X}, A'\mathbf{X}) \in \mathbb{P}^2 \times \mathbb{P}^2$. It can be shown that $\mathbb{P}^2 \times \mathbb{P}^2$ is a projective variety, and we are interested in describing the **joint image**, $\Phi(\dot{\mathbb{P}}^3) \subseteq \mathbb{P}^2 \times \mathbb{P}^2$.

3.2.1 Epipolar Constraint Revisited

From the camera model, we have the system of equations $A\mathbf{X} = \lambda\mathbf{x}$ and $A'\mathbf{X} = \lambda'\mathbf{x}'$. These can be rewritten as

$$Mu = \begin{pmatrix} A & \mathbf{x} & 0 \\ A' & 0 & \mathbf{x}' \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ -\lambda \\ -\lambda' \end{pmatrix} = \mathbf{0}$$

where M is a 6×6 matrix, u is a 6×1 vector, and $\mathbf{0}$ represents the zero vector. If the determinant of M is nonzero so that M is invertible, we find $M^{-1}Mu = M^{-1}\mathbf{0}$ so that $u = \mathbf{0}$, which implies $\mathbf{X} = (0, 0, 0, 0)$ and $\lambda = \lambda' = 0$. However, we are not interested in this case (we assume $\mathbf{X} \in \mathbb{P}^3$ and $(0, 0, 0, 0)$ is not even an element of $\mathbb{P}^3$). Thus, we require that $\det(M) = 0$, which gives rise to a polynomial constraint. In particular, letting $\mathbf{x} = (x_1, y_1, z_1)$ and $\mathbf{x}' = (x_2, y_2, z_2)$, since the elements of A and A' are known parameters that specify the properties of the camera, then $\det(M)$ gives a polynomial in variables $x_1, y_1, z_1, x_2, y_2, z_2$. The constraint that $\det(M(\mathbf{x}, \mathbf{x}')) = 0$ for a given $\mathbf{x}, \mathbf{x}'$ correspondence is equivalent to the epipolar constraint given above that $\mathbf{x}^T F \mathbf{x}' = 0$ (see [HZ04] §17.1). Since the determinant is a polynomial in the coordinates of the corresponding points, any pair of points which satisfies this constraint will lie on a variety.

3.2.2 Normalized Coordinate System

Since Heyden and Åström are interested in algebraic relationships, they simplify computations by choosing specific coordinate systems for $\mathbb{P}^3$ and for each copy of $\mathbb{P}^2$ in the image planes. Specifically, they propose to pick coordinate systems so that the camera centers are given by $c = (0, 0, 0, -1)$ and $c' = (1, 0, 0, -1)$, and so that camera matrices are given by

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, A' = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

which satisfy $Ac = 0$ and $A'c' = 0$ as required. Note this also defines the epipoles $e = Ac' = (1, 0, 0)$ and $e' = A'c = (-1, 0, 0)$. Picking coordinates in this way is possible since the camera centers are assumed to be distinct, and we can apply a projective transformation to the camera matrices (see [HZ04] §9.6).

Letting $\mathbf{x} = (x_1, y_1, z_1)$ and $\mathbf{x}' = (x_2, y_2, z_2)$, we can write out the matrix M from above as

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & x_1 & 0 \\ 0 & 1 & 0 & 0 & y_1 & 0 \\ 0 & 0 & 1 & 0 & z_1 & 0 \\ 1 & 0 & 0 & 1 & 0 & x_2 \\ 0 & 1 & 0 & 0 & 0 & y_2 \\ 0 & 0 & 1 & 0 & 0 & z_2 \end{pmatrix}.$$

Computing (see appendix) the $\det(M)$, we obtain the polynomial

$$b = z_1 y_2 - y_1 z_2,$$

so that the epipolar constraint is summarized by the equation $b = 0$.

3.2.3 The Two-View Variety

Since $b = z_1 y_2 - y_1 z_2$ is a homogeneous polynomial (referred to as “bihomogeneous”), we may consider the projective variety $\mathbf{V}(b) \subseteq \mathbb{P}^2 \times \mathbb{P}^2$. We will see that this variety is related to the joint image $\Phi(\dot{\mathbb{P}}^3)$, and will show that $\Phi(\dot{\mathbb{P}}^3)$ is not itself a variety.

Theorem (Theorem 4.2 from [HÅ97]). *We have $\Phi(\dot{\mathbb{P}}^3) \subsetneq \mathbf{V}(b)$.*

This is because the image $\Phi(\dot{\mathbb{P}}^3)$ does not contain the points in $\mathbb{P}^2 \times \mathbb{P}^2$ where one point is an epipole and the other is not. That is, $\Phi(\dot{\mathbb{P}}^3)$ is missing $(e, x) \in \mathbb{P}^2 \times \mathbb{P}^2$ for $x \neq e'$ and $(x', e') \in \mathbb{P}^2 \times \mathbb{P}^2$ for $x' \neq e$. For example, consider the point (e, x) for some $x = (x_2, y_2, z_2) \in \mathbb{P}^2 \setminus \{e'\}$. Then $(e, x) \notin \Phi(\dot{\mathbb{P}}^3)$ since any point that Φ maps to e in the first image will map to e' in the second image, which is clear from Figure 2 (a) since if a world point maps to the epipole e in the first image, it must lie on the baseline. However, since $e = (1, 0, 0)$, we have $b(1, 0, 0, x_2, y_2, z_2) = 0 \cdot y_2 - 0 \cdot z_2 = 0$, showing that $(e, x) \in \mathbf{V}(b)$. Thus, $\Phi(\dot{\mathbb{P}}^3) \subsetneq \mathbf{V}(b)$.

Lemma (Part 1 of Theorem 4.3 from [HÅ97]). *The ideal $\mathbf{I}(\Phi(\dot{\mathbb{P}}^3)) \subseteq \mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2]$ generated by all polynomials that vanish on the joint image is equal to the ideal generated by b . That is, $\mathbf{I}(\Phi(\dot{\mathbb{P}}^3)) = \langle b \rangle$.*

Proof. Note that $\Phi(\dot{\mathbb{P}}^3)$ consists of all points $(\mathbf{x}, \mathbf{x}') \in \mathbb{P}^2 \times \mathbb{P}^2$ satisfying equations $A\mathbf{X} = \lambda\mathbf{x}$ and $A'\mathbf{X} = \lambda'\mathbf{x}'$ for some given world coordinates $\mathbf{X} = (X, Y, Z, W) \in \mathbb{P}^3$ and scalars $\lambda, \lambda' \in \mathbb{R}$. Letting $\mathbf{x} = (x_1, y_1, z_1)$ and $\mathbf{x}' = (x_2, y_2, z_2)$ and writing these equations fully as

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} = \lambda' \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix},$$

we obtain the system of 6 equations given by equating the below polynomials to zero:

$$\begin{aligned} X - \lambda x_1, \quad Y - \lambda y_1, \quad Z - \lambda z_1, \\ (X + W) - \lambda' x_2, \quad Y - \lambda' y_2, \quad Z - \lambda' z_2. \end{aligned} \tag{1}$$

Allowing (X, Y, Z, W) and λ, λ' to vary, these equations are polynomials and generate an ideal $J \subseteq \mathbb{R}[X, Y, Z, W, \lambda, \lambda', x_1, \dots, z_2]$. However, we are interested in characterizing the ideal $I = \mathbf{I}(\Phi(\dot{\mathbb{P}}^3)) \subseteq \mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2]$ of polynomials in image coordinates which vanish on $\Phi(\dot{\mathbb{P}}^3)$. Thus our strategy will be to compute a Gröbner basis for J using an elimination monomial order and to take its intersection with $\mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2]$. By the elimination theorem ([CLO15] §3.1) this will provide a Gröbner basis for I . However, if we compute a Gröbner basis for J with, for example, the lexicographic order to eliminate the world coordinates and scale parameters, we will find that the scale parameters can not be eliminated (see appendix). In particular, the polynomial $\lambda'(y_1 z_2 - z_1 y_2)$ is one of the generators in the computed Gröbner basis. This shows that any point in $\mathbf{V}(J)$ must satisfy either $\lambda' = 0$ or $y_1 z_2 - z_1 y_2 = 0$. The second equation is equivalent to $b = 0$, which is not surprising since this is the epipolar constraint, and any point where the equations (1) vanish must satisfy the epipolar constraint. However, we are not interested in cases where $\lambda = 0$ or $\lambda' = 0$, since these parameters scale the homogeneous coordinates of the image points, which are never equal to $(0, 0, 0)$. This suggests we need to add a constraint to our system of equations, which can be achieved by introducing a new variable, μ , and adding a seventh equation to (1). That equation is $\mu\lambda\lambda' - 1 = 0$, which ensures that $\lambda, \lambda' \neq 0$. Then we compute a Gröbner basis for the ideal generated by these 7 polynomials, using lexicographic order with $\mu > X > Y > Z > W > \lambda > \lambda' > x_1 > \dots > z_2$, and take the intersection of this basis with $\mathbb{R}[x_1, y_1, z_1, x_2, y_2, z_2]$, and find (see appendix) that $\{y_1 z_2 - z_1 y_2\}$ is a Gröbner basis for I . Since this generator is $-b$, it follows that $\mathbf{I}(\Phi(\dot{\mathbb{P}}^3)) = \langle b \rangle$. $\square$

Having shown that $\Phi(\dot{\mathbb{P}}^3) \subsetneq \mathbf{V}(b)$ and that $\mathbf{I}(\Phi(\dot{\mathbb{P}}^3)) = \langle b \rangle$, the next theorem follows simply.

Theorem (Part 2 of Theorem 4.3 from [HÅ97]). *The joint image $\Phi(\dot{\mathbb{P}}^3)$ is not a variety, but its Zariski closure $\overline{\Phi(\dot{\mathbb{P}}^3)} = \mathbf{V}(b)$.*

Proof. By definition, $\overline{\Phi(\dot{\mathbb{P}}^3)} = \mathbf{V}(\mathbf{I}(\Phi(\dot{\mathbb{P}}^3))) = \mathbf{V}(b)$, where the second equality uses the above lemma. Since any variety V satisfies $V = \mathbf{V}(\mathbf{I}(V))$ but $\mathbf{V}(\mathbf{I}(\Phi(\dot{\mathbb{P}}^3))) = \mathbf{V}(b) \supsetneq \Phi(\dot{\mathbb{P}}^3)$ by the first theorem, it follows that $\Phi(\dot{\mathbb{P}}^3)$ is not a variety. $\square$

The variety $\mathbf{V}(b)$ is known as the “multiview variety”, and it can be generalized to more than two views.

3.3 Implications for Triangulation

Returning to the triangulation problem, we see that de-noising corresponding image points prior to triangulation can be re-phrased geometrically. Our measured point correspondences $\mathbf{x}, \mathbf{x}'$ are typically close to the multiview variety but do not lie exactly on it due to noise from

the measurement process. Since the measured points do not lie on the multiview variety, they can not lie in the joint image, which is a subset of the variety (that is, the noisy points can't simultaneously satisfy $A\mathbf{X} = \lambda\mathbf{x}$, $A'\mathbf{X} = \lambda'\mathbf{x}'$ for a given world point). Therefore, we can't hope to recover the world point captured by the cameras until we find some nearby points $\hat{\mathbf{x}}, \hat{\mathbf{x}}'$ that do lie on the multiview variety. In practice, de-noising may be accomplished by finding the points $\hat{\mathbf{x}}, \hat{\mathbf{x}}'$ which are minimally distant from the original points (using some measure of distance) subject to the epipolar constraint that $\hat{\mathbf{x}}^T F \hat{\mathbf{x}}' = 0$ (see [HZ04] §12.5). This agrees with our intuition, since the epipolar constraint generates the multiview variety. Thus, geometrically, de-noising is the process of projecting the measured image points onto the multiview variety.

The previous section also shows that the joint image of a scene under two fixed cameras is not a variety, since it is missing some of the points which satisfy the epipolar constraint. However, these missing points are exactly the pairs in $\mathbb{P}^2 \times \mathbb{P}^2$ where one point is an epipole and the other point is not an epipole. These points cannot appear in the joint image due to the epipolar geometry arising from the two views, but they do satisfy the epipolar constraint. Thus, de-noised points will always be valid image points in $\Phi(\mathbb{P}^3)$ unless one of the de-noised points is an epipole and the other is not. In a triangulation program, these points could be discarded since we know they could never appear in the joint image, and thus may be due to error.

4 References

- [HÅ97] Anders Heyden and Kalle Åström. “Algebraic Properties of Multilinear Constraints”. In: *Mathematical Methods in the Applied Sciences* 20.13 (1997). DOI: [https://doi.org/10.1002/\(SICI\)1099-1476\(19970910\)20:13<1135::AID-MMA908>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1099-1476(19970910)20:13<1135::AID-MMA908>3.0.CO;2-9).
- [HZ04] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2004. ISBN: 978-0-511-18618-9.
- [CLO15] David A. Cox, John Little, and Donal O’Shea. *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra*. Springer, 2015. ISBN: 978-3-319-16721-3.

5 Appendix

The SageMath computer algebra system (www.sagemath.org) was used to perform several computations to support the above text.

First, we use Sage to verify the calculation in §3.2.2 that $\det(M) = z_1y_2 - y_1z_2$.

```
## Define matrix M
var("x1 y1 z1 x2 y2 z2")
M = matrix(SR, [
    [1,0,0,0,x1,0],
    [0,1,0,0,y1,0],
    [0,0,1,0,z1,0],
    [1,0,0,1,0,x2],
    [0,1,0,0,0,y2],
    [0,0,1,0,0,z2]
])

# Compute the determinant
det(M)
```

(Output):

y2*z1 - y1*z2

Next, we can verify the Gröbner basis computations from §3.2.3.

```
R.<X,Y,Z,W,lambda1,lambda2,x1,y1,z1,x2,y2,z2> = PolynomialRing(QQ, order="lex")

camera_eqns = [
    # equations from camera 1
    X - lambda1 * x1,
    Y - lambda1 * y1,
    Z - lambda1 * z1,
    # equations from camera 2
    X + W - lambda2 * x2,
    Y - lambda2 * y2,
    Z - lambda2 * z2
]

J = Ideal(camera_eqns)
print("Groebner basis for J")
print(J.groebner_basis())

# Constrain so lambda, lambda2 \ne 0
R.<u,X,Y,Z,W,lambda1,lambda2,x1,y1,z1,x2,y2,z2> = PolynomialRing(QQ, order="lex")
```

```

camera_eqns_constrained = camera_eqns + [ u*lambda1*lambda2 - 1 ]

J_tilde = Ideal(camera_eqns_constrained)
G_tilde = J_tilde.groebner_basis()

# Extract generators in elimination ideal
elim_gens = [
    g for g in G_tilde
        if not any(v in g.variables() for v in [u,X,Y,Z,W,lambda1,lambda2])
]
print("\nGroebner basis for I")
print(elim_gens)

```

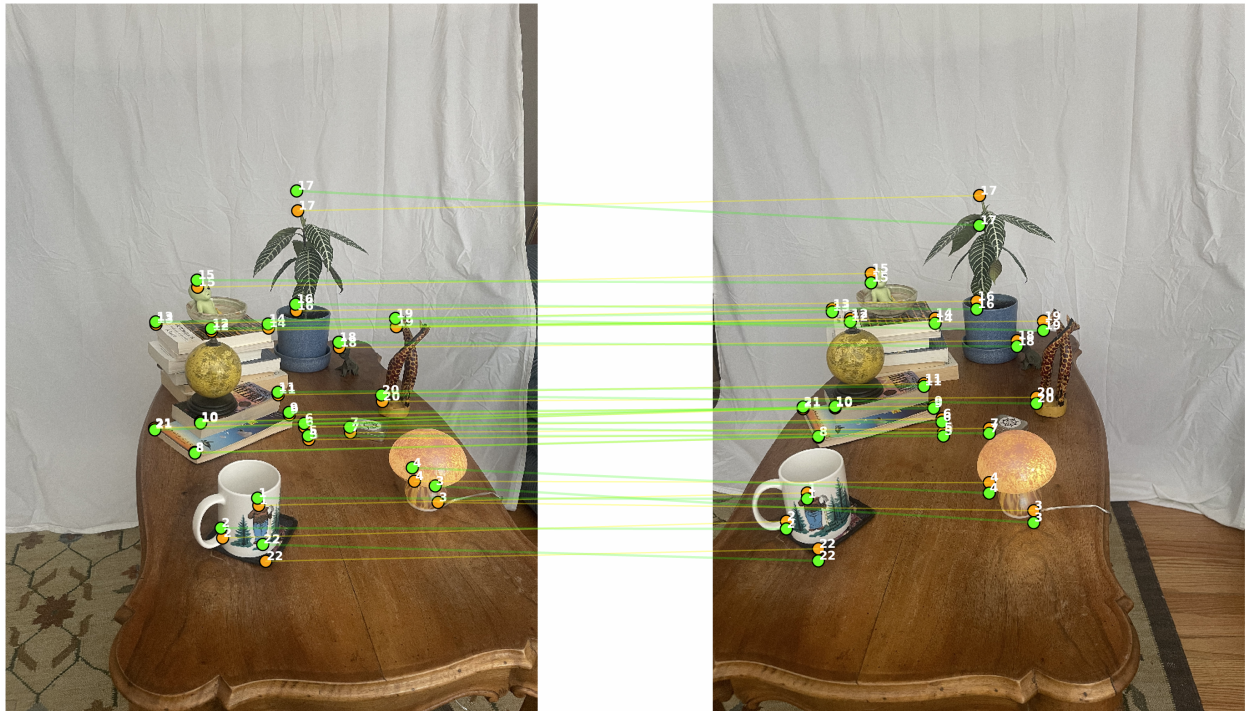
(Output)

```

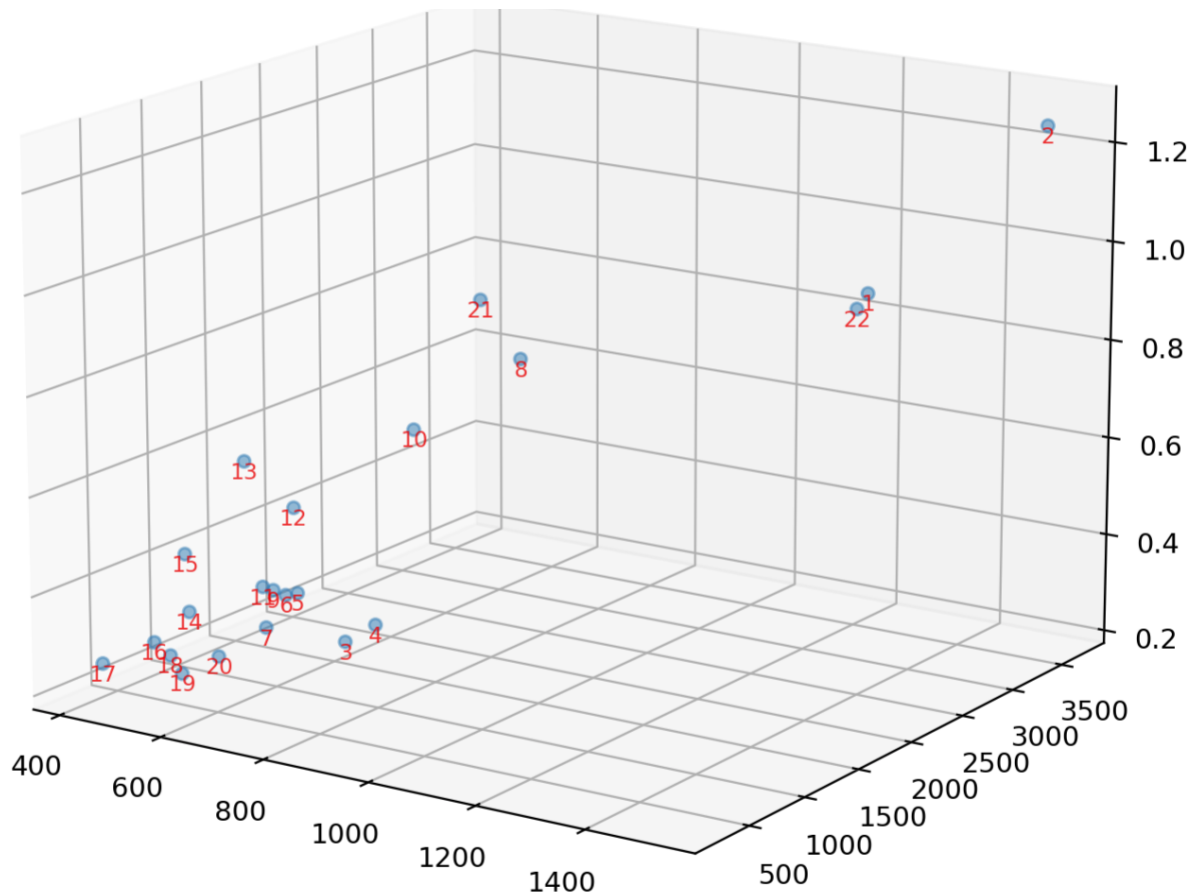
Groebner basis for J
[X - lambda1*x1, Y - lambda2*y2, Z - lambda2*z2, W + lambda1*x1 - lambda2*x2,
 lambda1*y1 - lambda2*y2, lambda1*z1 - lambda2*z2, lambda2*y1*z2 -
 lambda2*z1*y2]

Groebner basis for I
[y1*z2 - z1*y2]

```



- Two images of the same scene taken from different positions.
- Point correspondences were hand selected, and are shown by the orange points.
- Algorithms from [HZ04] were used to compute corrected points, shown in green. The “normalized 8-point algorithm” was used to estimate F , and then the “optimal triangulation algorithm” was used to project the points onto the multiview variety. Information about the cameras was not incorporated (the “canonical camera matrices” were used).
- Code: https://github.com/hans-elliott99/3d_triangulation



- Triangulated 3-D points produced from the corrected point correspondences using the linear triangulation algorithm. Since “uncalibrated” cameras were used, the 3-D reconstruction is only accurate up to “projective ambiguity”. That is, some projective map needs to be applied to the triangulated points to recover a sensible 3-D reconstruction. Identifying this map requires additional information about the cameras and the scene.