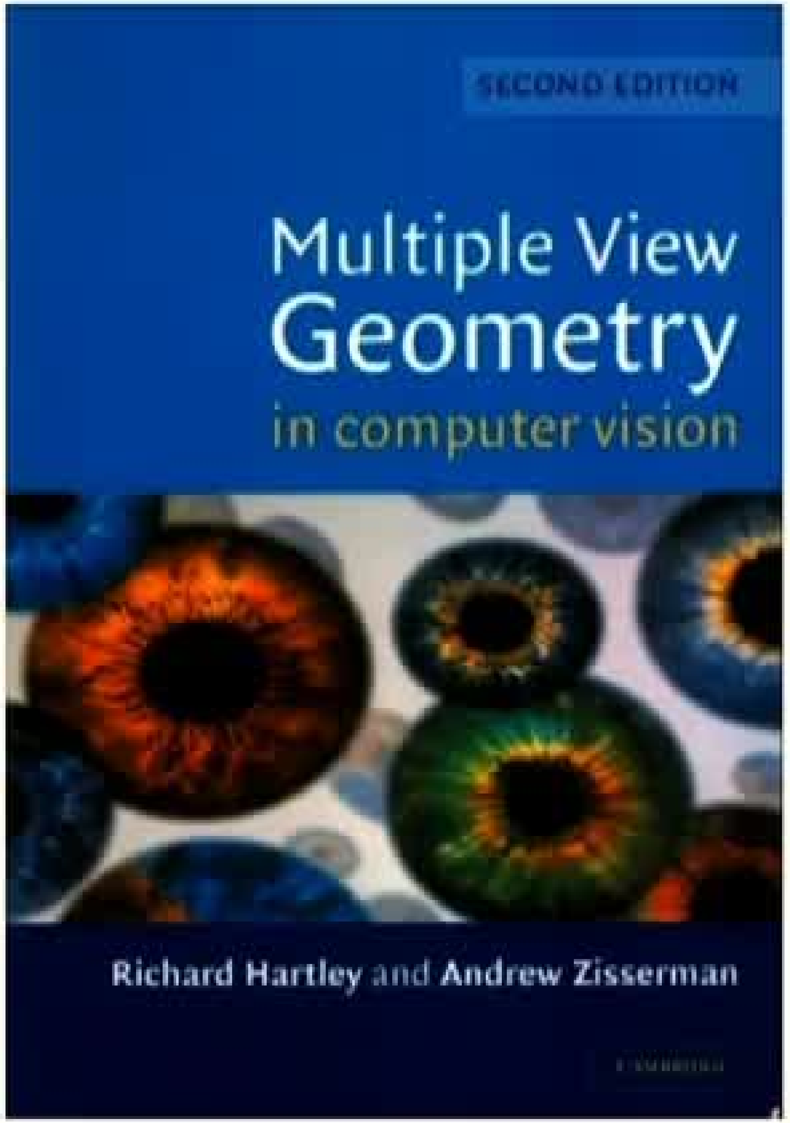
Algebraic Vision

Rekha R. Thomas
University of Washington



Multiple View Geometry in Computer Vision

Richard Hartley and Andrew Zisserman,
Cambridge University Press, 2000

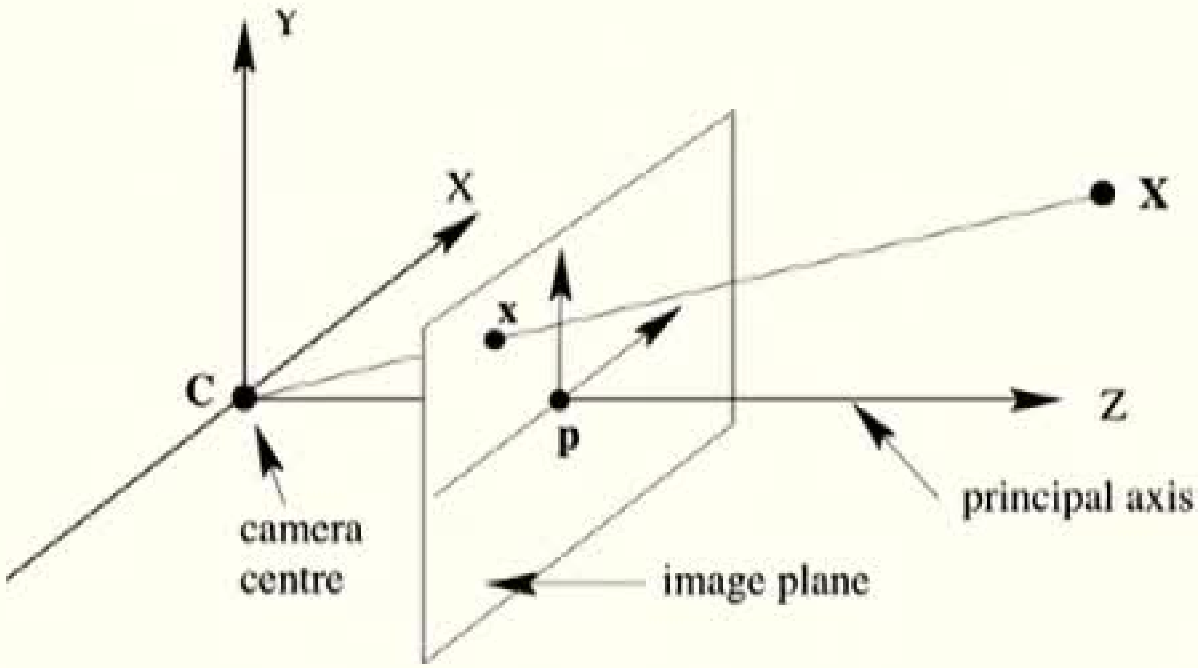
An Invitation to 3D Vision

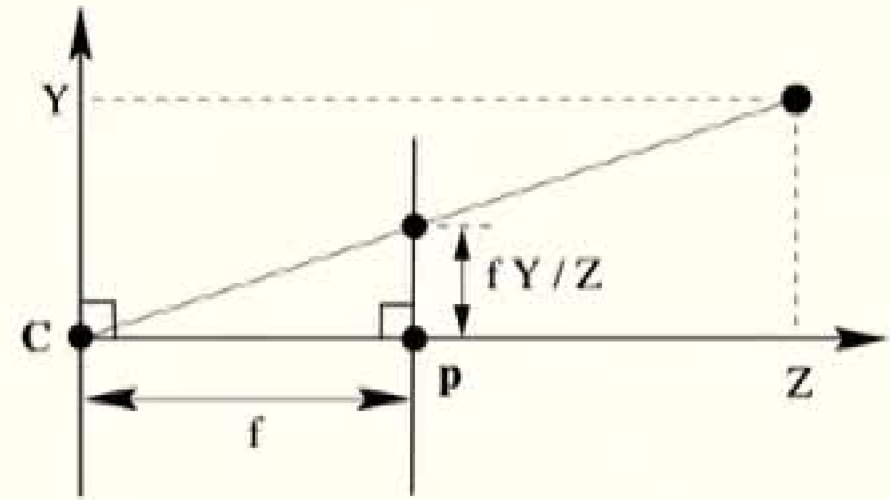
Yi Ma, Stefano Soatto, Jana Kosecka, Shankar Sastry,
Springer-Verlag 2004

Theory of Reconstruction from Image Motion

Stephen Maybank, Springer-Verlag 1993

The basic pinhole camera





$$\begin{array}{ccc} \mathbb{R}^3 & \rightarrow & \mathbb{R}^2 \\ (X, Y, Z) & \mapsto & \left(\frac{fX}{Z}, \frac{fY}{Z}\right) \\ \mathbf{X} & & \mathbf{x} \end{array}$$

$$\begin{array}{ccc} \mathbb{P}^3 & \rightarrow & \mathbb{P}^2 \\ (X, Y, Z, 1) & \mapsto & \underbrace{\left(\frac{fX}{Z}, \frac{fY}{Z}, 1\right)} \end{array}$$

central projection from
world to image coordinates

$$\begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

camera matrix

general projective camera

$$\mathbb{P}^3 \dashrightarrow \mathbb{P}^2$$
$$\mathbf{X} \mapsto A\mathbf{X} =: \mathbf{x}$$

$A = [B \mid b] \in \mathbb{R}^{3\times 4}$
rank $(A) = 3$

---

camera center $\quad c \in \mathbb{P}^3 \, : \, Ac = 0$

---

finite camera $\quad B$ nonsingular $\Rightarrow c = \left(-B^{-1}b, 1\right)$

A finite $\Rightarrow A = K[R \;\; t]$

K upper triangular w/ positive diagonal

R rotation matrix

$t \in \mathbb{R}^3$

K contains the internal parameters of the camera; focal length, skew etc

---

calibrated camera $\quad A = K[R \; t] \quad K$ known *linear model of image formation*

$$\mathbf{x} \sim A\mathbf{X}$$

image — *camera* — *world point*

Reconstruction Problems

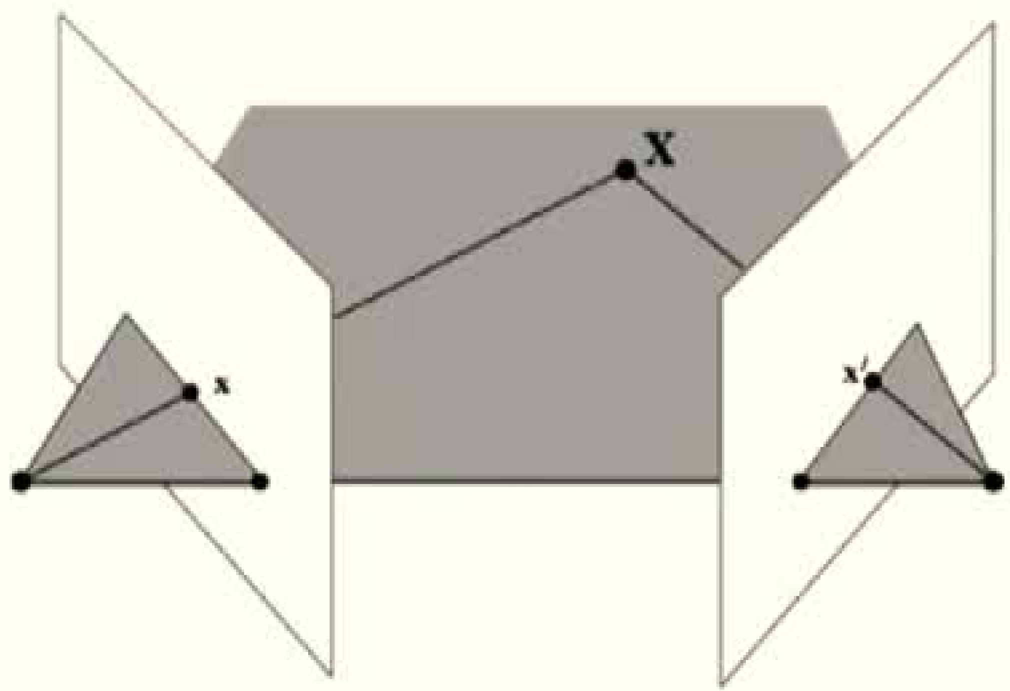
	Given	*Find*
Resectioning	$\mathbf{X}_i \leftrightarrow \mathbf{x}_i$	$A \,:\, \mathbf{x}_i \sim A\mathbf{X}_i$
Triangulation	$A_i,\ \mathbf{x}_i$	$\mathbf{X} \,:\, \mathbf{x}_i \sim A_i\mathbf{X}$
Reconstruction	$\mathbf{x}_i^1 \leftrightarrow \mathbf{x}_i^2 \cdots \leftrightarrow \mathbf{x}_i^n$	$A_j, \mathbf{X}_i \,:\, \mathbf{x}_i^j \sim A_j\mathbf{X}_i$

in practice, images are noisy and need to find MLE

Triangulation

Want to recover a 3D point $\mathbf{X}$ *from noisy images* $\hat{\mathbf{x}}_i$ *in n cameras*

$$\mathcal{A} = (A_1, \ldots, A_n)$$



Polynomial Optimization

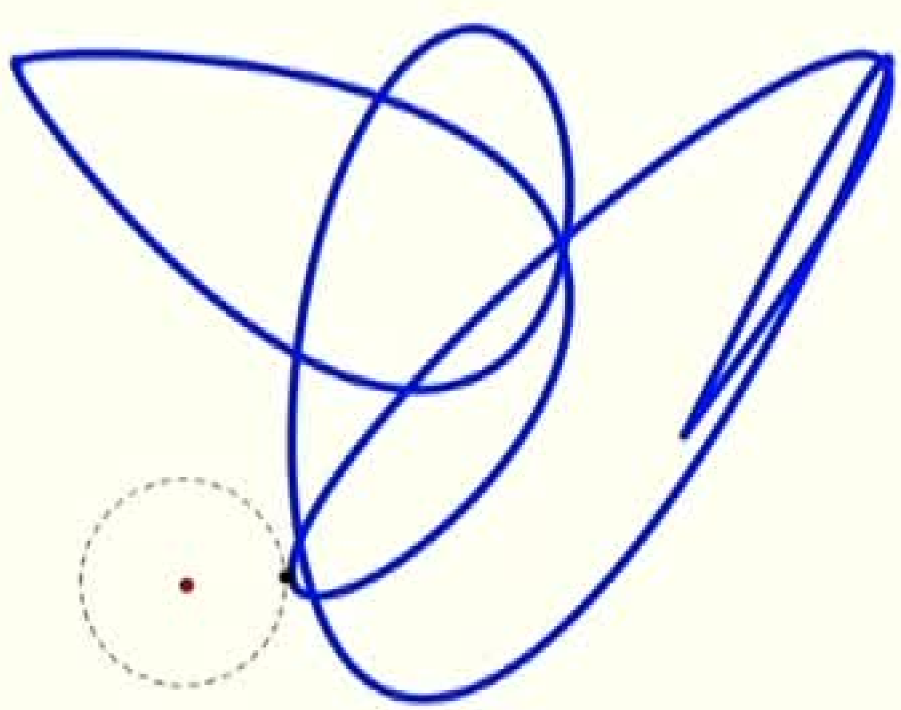
$$\begin{array}{ll} \min & \sum \|x_i - \hat{\mathbf{x}}_i\|^2 \\ \text{s.t.} & (x_1, \ldots, x_n) \in \overline{\varphi(\mathbb{P}^3)} \end{array}$$

Algebraic Geometry

$$\begin{array}{llll} \varphi : & \mathbb{P}^3 & \rightarrow & \mathbb{P}^2 \times \cdots \times \mathbb{P}^2 \\ & X & \mapsto & (A_1 X, \ldots, A_n X) \end{array}$$

$$\overline{\varphi(\mathbb{P}^3)} \subset (\mathbb{P}^2)^n$$

multiview variety of $\mathcal{A}$



$J_{\mathcal{A}}$ *multiview ideal* — *vanishing ideal of the multiview variety*

$$A_i \mathbf{X} = \lambda_i \mathbf{x}_i \;\; \forall i = 1, \ldots, n \;\; \Leftrightarrow \begin{bmatrix} A_1 & \mathbf{x}_1 & 0 & 0 & \cdots & 0 \\ A_2 & 0 & \mathbf{x}_2 & 0 & \cdots & 0 \\ \vdots & & & & & \\ A_n & 0 & 0 & 0 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{pmatrix} \mathbf{X} \\ \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0$$

so all maximal minors of this matrix lie in $J_{\mathcal{A}}$

b_{ij}	t_{ijk}	q_{ijkl}
bilinear/epipolar constraints	*trilinear/trifocal constraints*	*quadrilinear/quadrifocal constraints*

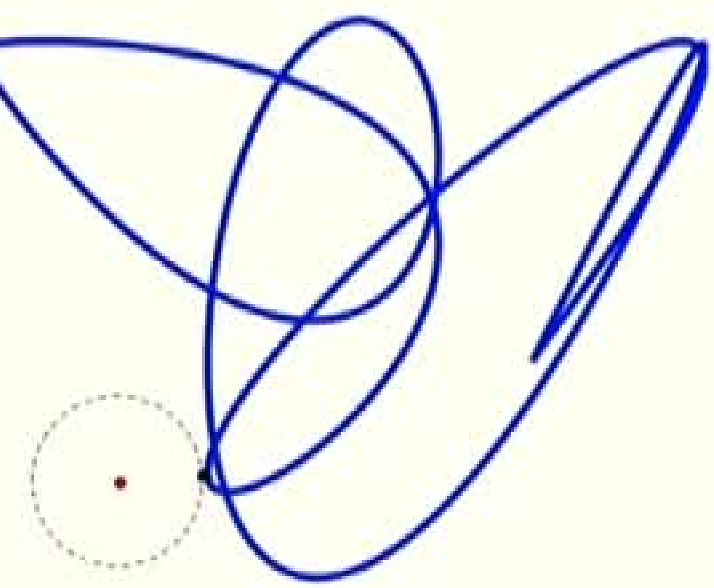
Aholt, Sturmfels, T. (2011) *strengthens* Heyden-Åström (1997)

$\{b_{ij}, t_{ijk}, q_{ijk}\}$ *form a universal Gröbner basis for* $J_{\mathcal{A}}$ *generically*

… *Hilbert schemes in vision* AST (2011), Lieblich-van Meter (2018)

Polynomial Optimization

$$\begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & x \in \overline{\varphi(\mathbb{P}^3)} \end{array} \quad = \quad \begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & b_{ij} = 0 \\ & t_{ijk} = 0 \end{array}$$

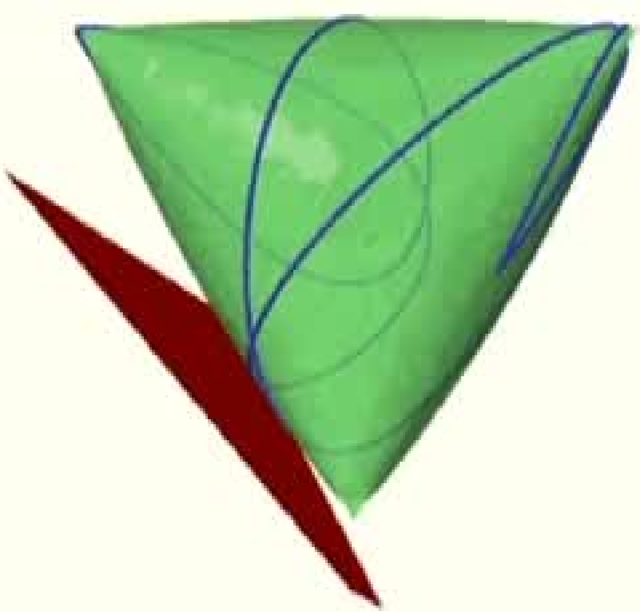


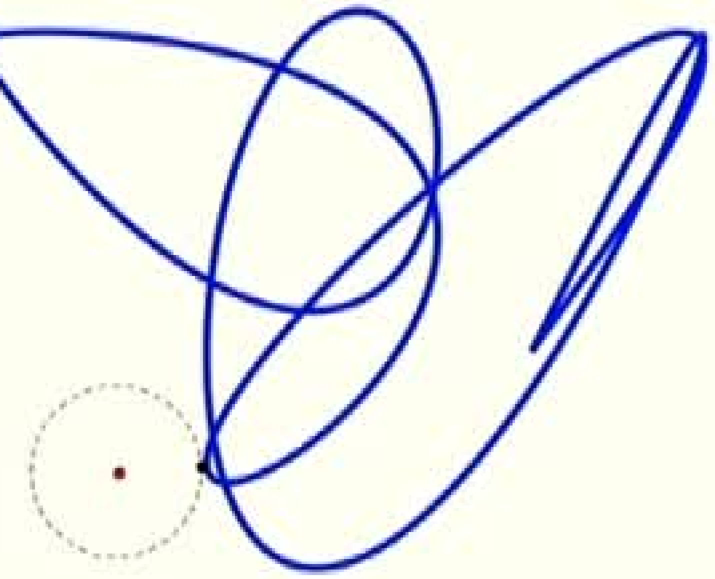
$$\begin{array}{ll} \min & \langle G, Y \rangle \\ \text{s.t.} & \langle B_{ij}, Y \rangle = 0 \\ & Y \succeq 0 \end{array}$$

semidefinite programming
(SDP)
relaxation

Polynomial Optimization

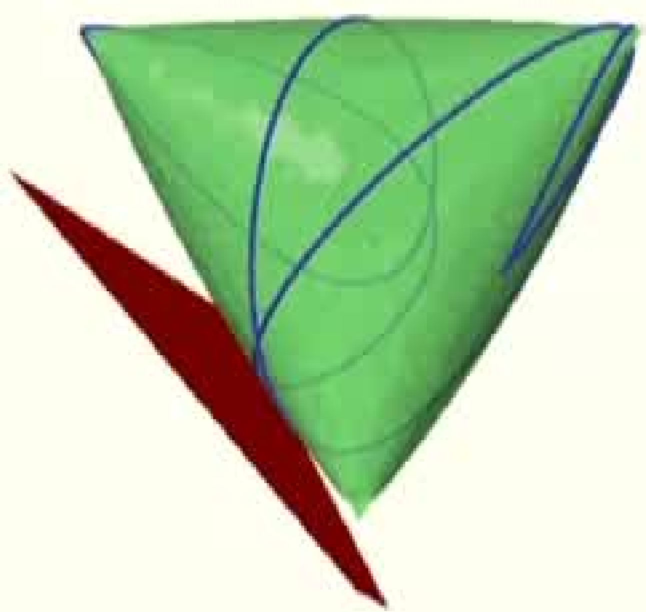
$$\begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & x \in \overline{\varphi(\mathbb{P}^3)} \end{array} \quad = \quad \begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & b_{ij} = 0 \\ & t_{ijk} = 0 \end{array}$$

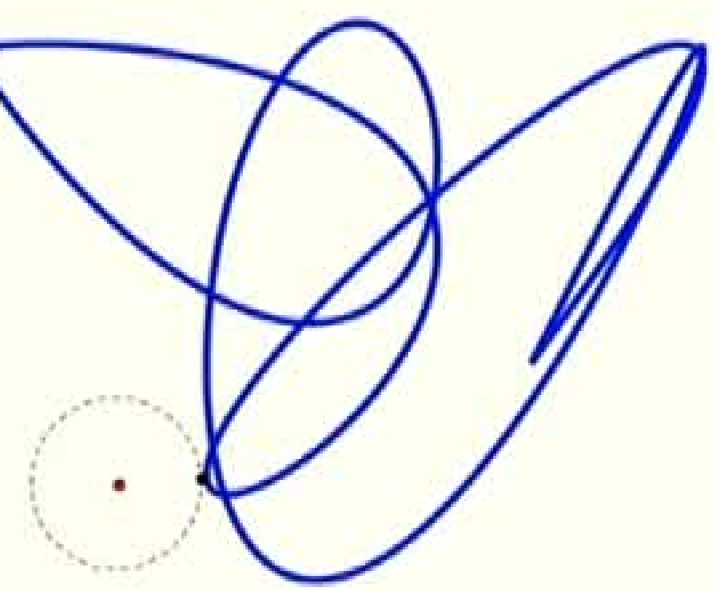




Polynomial Optimization

$$\begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & x \in \overline{\varphi(\mathbb{P}^3)} \end{array} \quad = \quad \begin{array}{ll} \min & \|x - \hat{\mathbf{x}}\|^2 \\ \text{s.t.} & b_{ij} = 0 \\ & t_{ijk} = 0 \end{array}$$

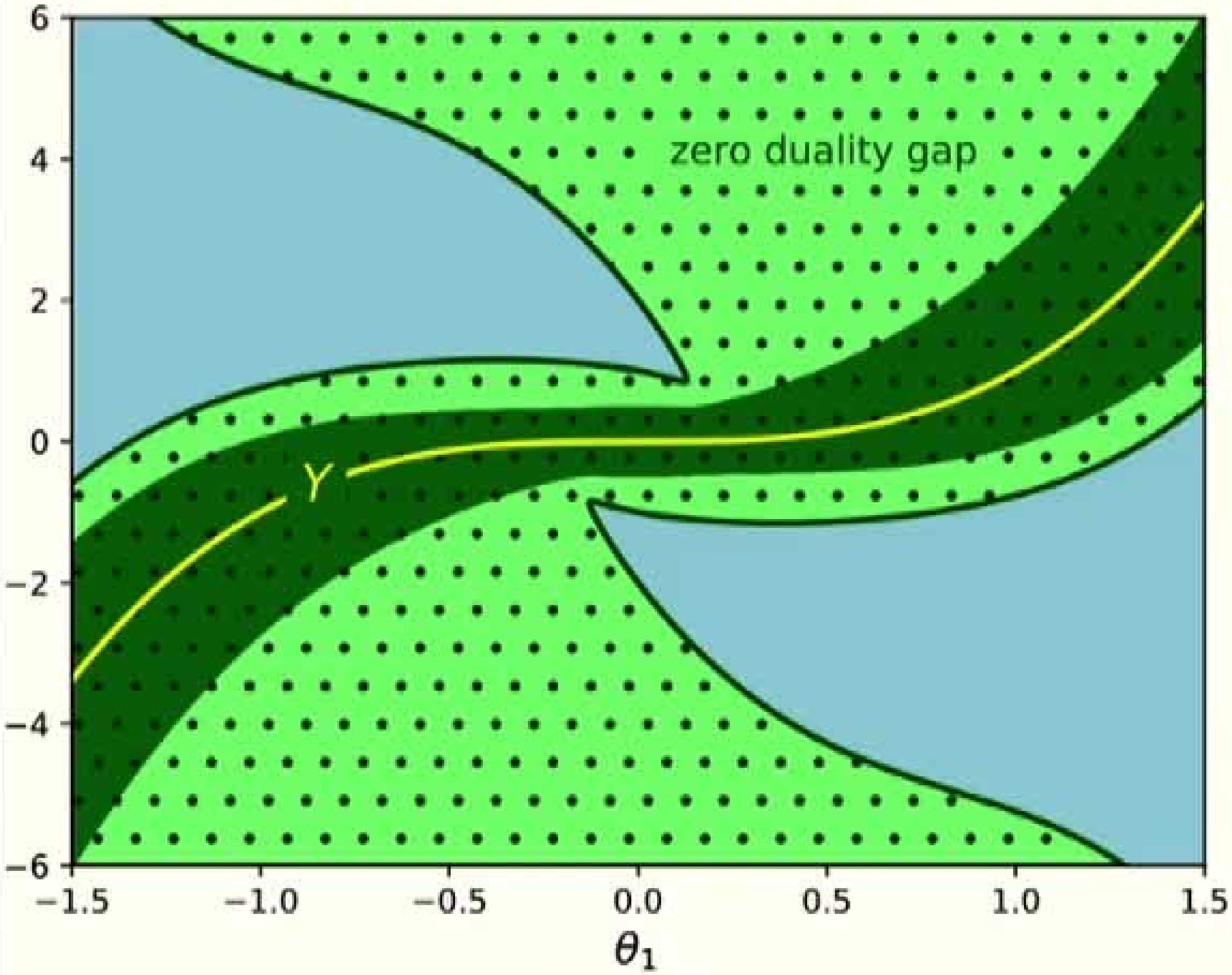




Theorem: **Aholt, Agarwal, T. (2012)** *This SDP relaxation solves the original triangulation problem exactly, under low noise.*

Local stability of SDP relaxations of QCQPs

Cifuentes, Agarwal, Parrilo, T. (2017)

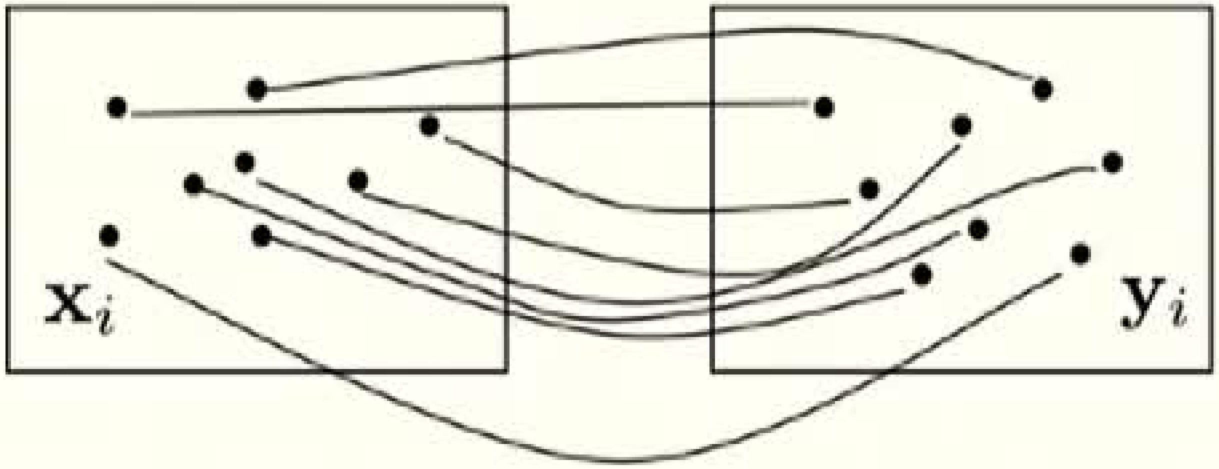


$$\begin{array}{ll} \min & \|y-\theta\|^2 \\ \text{s.t.} & y_2 = y_1^2,\ y_3 = y_1 y_2 \end{array}$$

A simple theorem: Suppose $\bar{\theta} \in Y$ (quadratic variety) is smooth. Then SDP relaxation solves the QCQP $\forall \theta$ sufficiently close to $\bar{\theta}$.

Two View Geometry

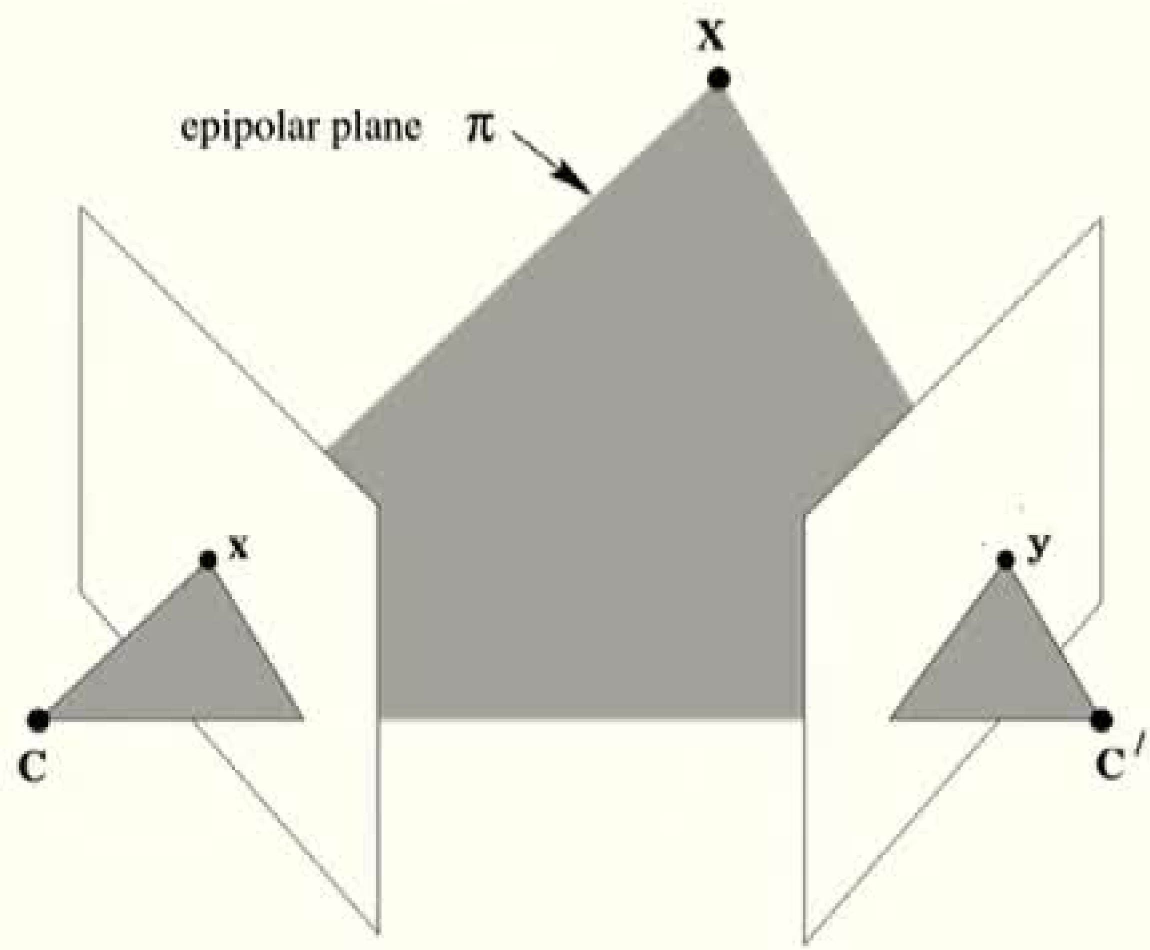
Given m correspondences: $\{(\mathbf{x}_i, \mathbf{y}_i) \in \mathbb{R}^2 \times \mathbb{R}^2,\ i = 1, \ldots, m\}$



Are these images of m world points in 2 cameras?
If yes, find the cameras and reconstruct the 3D point cloud.

Two View Geometry

$$\mathbf{x} \sim A\mathbf{X}, \;\; \mathbf{y} \sim A'\mathbf{X} \qquad \Rightarrow \qquad \mathbf{X}, \mathbf{x}, \mathbf{y}, C, C' \text{ all coplanar}$$

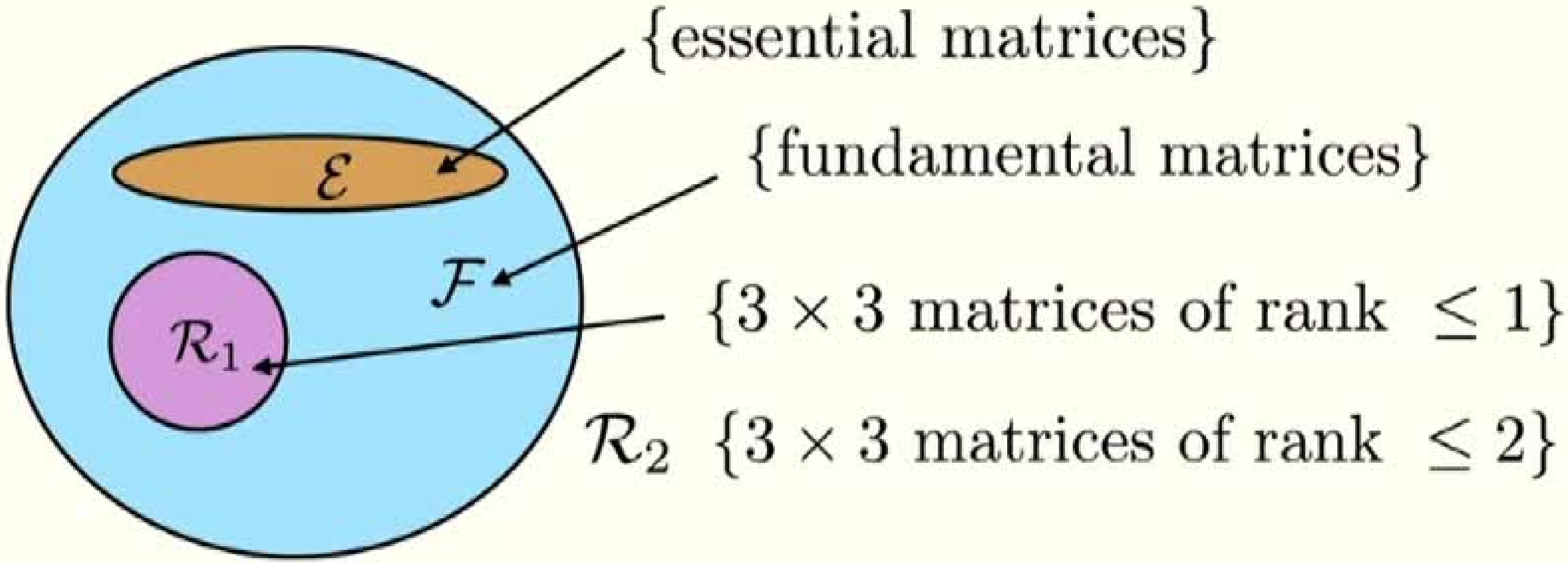


$$\Rightarrow \qquad \exists\, 3 \times 3 \text{ matrix } M, \; \text{rank}(M) = 2 \text{ s.t.}$$

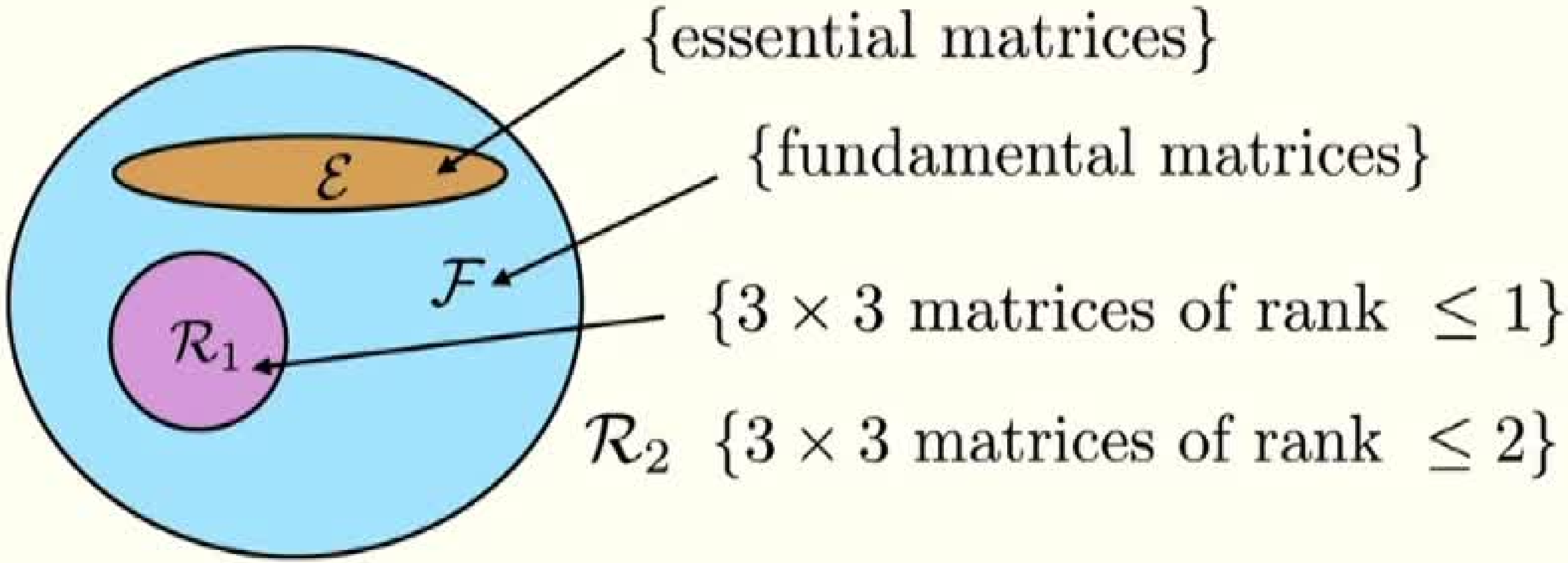
$$\underbrace{\mathbf{y}^\top M \mathbf{x} = 0}$$

epipolar equation

uncalibrated cameras	calibrated cameras
$M = F$	$M = E$
fundamental matrix	*essential matrix*
	$\sigma_1(E) = \sigma_2(E) > 0$



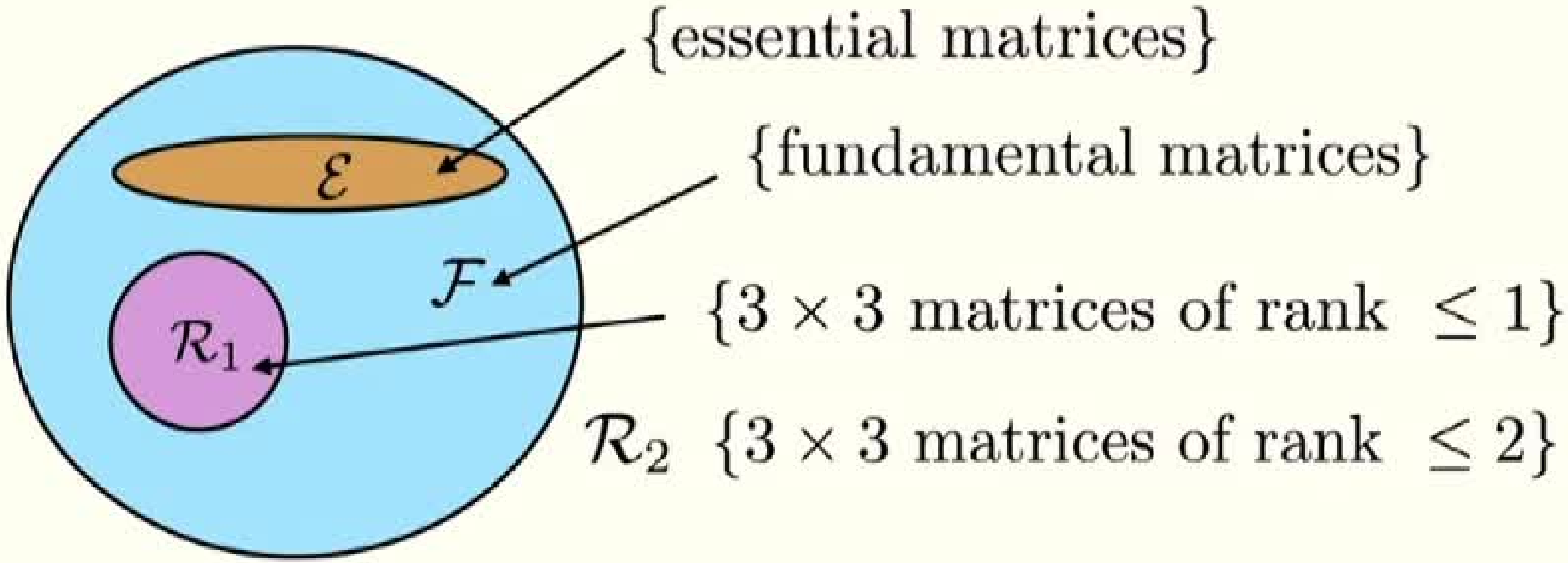
m point correspondences $\leftrightarrow$ subspace L of codim m



m point correspondences $\leftrightarrow$ subspace L of codim m

$\exists\ F \ \Leftrightarrow\ L \cap \mathcal{F} \neq \emptyset \qquad\qquad \exists\ E \ \Leftrightarrow\ L \cap \mathcal{E} \neq \emptyset$

$\mathcal{F} = \mathcal{R}_2 \backslash \mathcal{R}_1$ *matrices of rank exactly 2*	$\mathcal{E}$ *real algebraic variety cut out by* $2EE^\top E - \mathrm{trace}(EE^\top)E = 0,\ \det(E) = 0$ *(10 cubics)* **Demazure (1988)**



m point correspondences $\leftrightarrow$ subspace L of codim m

$\exists\ F \ \Leftrightarrow\ L \cap \mathcal{F} \neq \emptyset$ $\qquad$ $\exists\ E \ \Leftrightarrow\ L \cap \mathcal{E} \neq \emptyset$

Theorems:

Agarwal, Lee, Sturmfels, T. (2017)

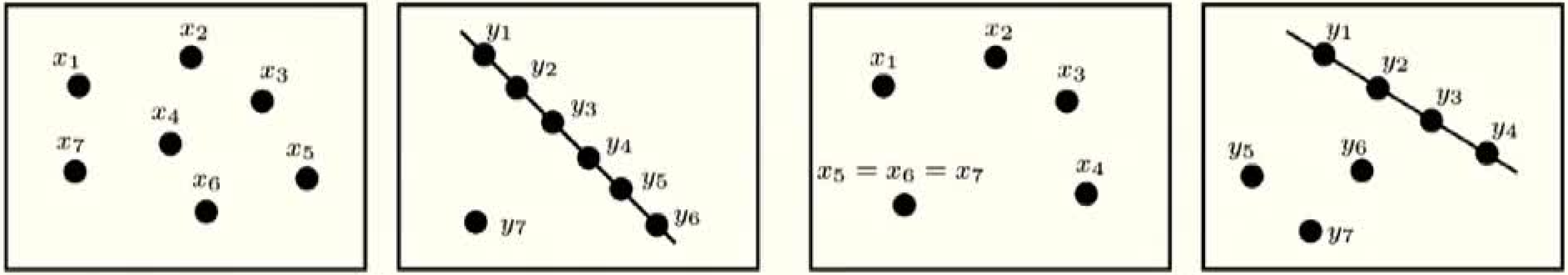
$m \leq 5 \ \Rightarrow\ L \cap \mathcal{F} \neq \emptyset$

$m \leq 4 \ \Rightarrow\ L \cap \mathcal{E} \neq \emptyset$

Fundamental Matrices

$$m = 6, 7, 8$$

There is always a real 3x3 matrix in the intersection but it may not have rank two.

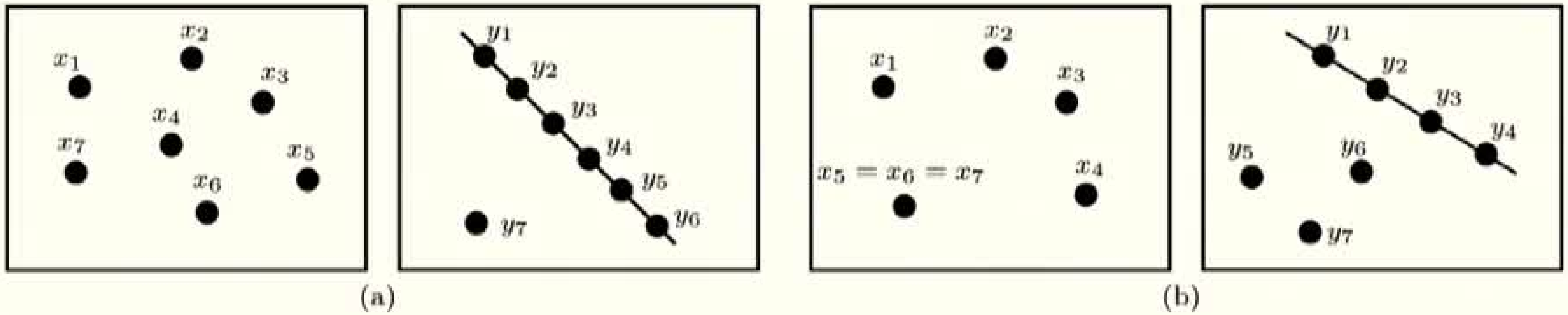


(a) (b)

Fundamental Matrices

$m = 6, 7, 8$

There is always a real 3x3 matrix in the intersection but it may not have rank two.



Essential Matrices

$m = 5, 6$ ***open*** *There maybe no real points in the intersection.*

$$X = \begin{bmatrix} 3 & 0 & 1 \\ 9 & 1 & 1 \\ 1 & 2 & 1 \\ 8 & 8 & 1 \\ 4 & 8 & 1 \end{bmatrix} \quad Y = \begin{bmatrix} 2 & 0 & 1 \\ 5 & 4 & 1 \\ 9 & 6 & 1 \\ 2 & 5 & 1 \\ 1 & 4 & 1 \end{bmatrix}$$

$$L \cap \mathcal{E} = \emptyset$$

$$L \cap \mathcal{E}_{\mathbb{C}} = 10$$

5-point algorithm for E Nistér (2004)

1. *Plug* $sA_1 + tA_2 + uA_3 + A_4 \in L$ *into the 10 Demazure cubics.*

2. *Gauss-Jordan elimination yields:*

$\mathcal{A}$	s^3	t^3	s^2t	st^2	s^2u	s^2	t^2u	t^2	stu	st	s	t	1
(a)	1										[2]	[2]	[3]
(b)		1									[2]	[2]	[3]
(c)			1								[2]	[2]	[3]
(d)				1							[2]	[2]	[3]
(e)					1						[2]	[2]	[3]
(f)						1					[2]	[2]	[3]
(g)							1				[2]	[2]	[3]
(h)								1			[2]	[2]	[3]
(i)									1		[2]	[2]	[3]
(j)										1	[2]	[2]	[3]

Kukelova's thesis('12)

~10.6 micro secs

many such solvers
Pajdla's group @ Czech TU

3. *Define:*

$$\begin{aligned}(k) &:= (e) - u\cdot(f)\\ (l) &:= (g) - u\cdot(h)\\ (m) &:= (i) - u\cdot(j)\end{aligned}$$

and arrange into a 3x3 matrix:

$\mathcal{B}$	s	t	1
(k)	[3]	[3]	[4]
(l)	[3]	[3]	[4]
(m)	[3]	[3]	[4]

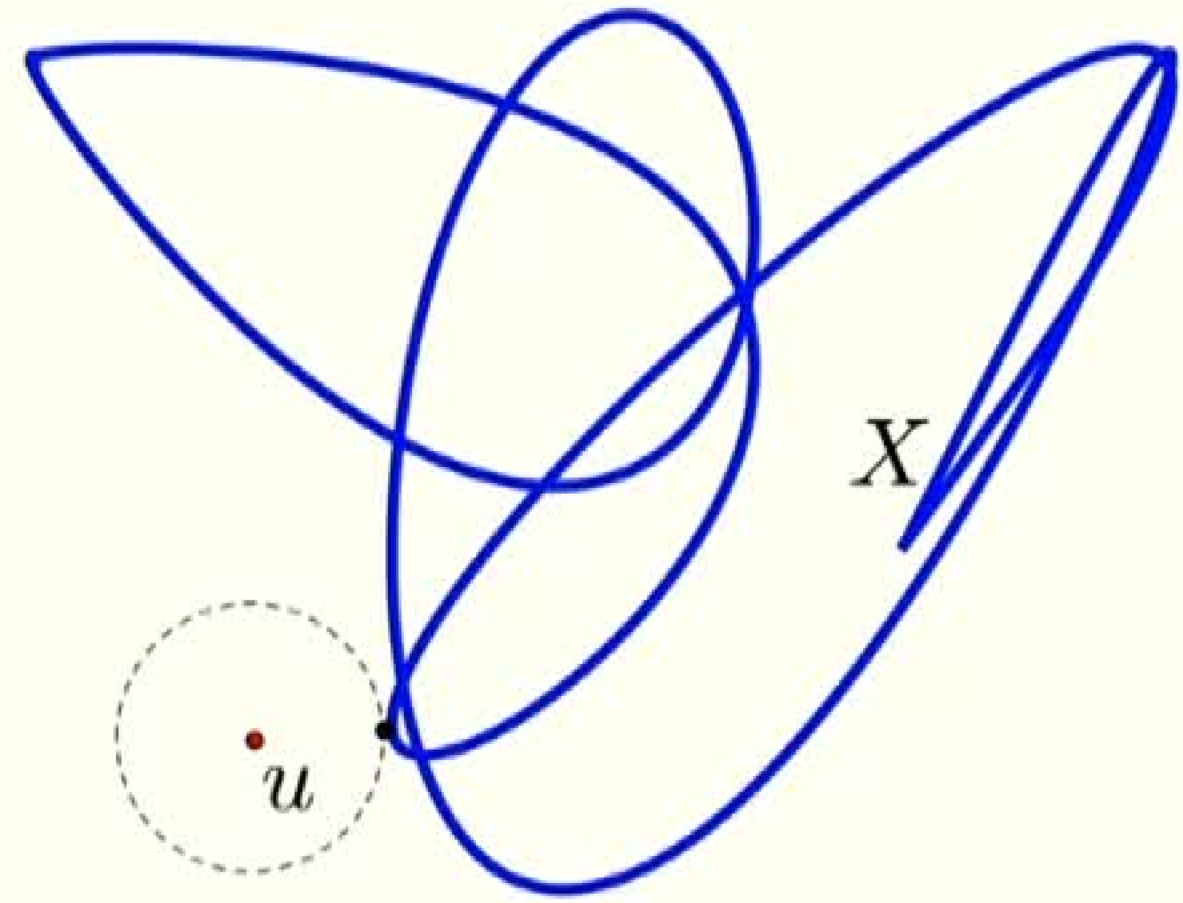
4. $\det(\mathcal{B})$ *is a univariate polynomial in u of degree 10.*
Solve using Sturm sequences/eigenvalue methods

Euclidean distance degree of an algebraic variety

$u \in \mathbb{R}^n$ *data point*

$X \subseteq \mathbb{R}^n$ *real algebraic variety*

$$\min_{x \in X} \|x - u\|^2$$

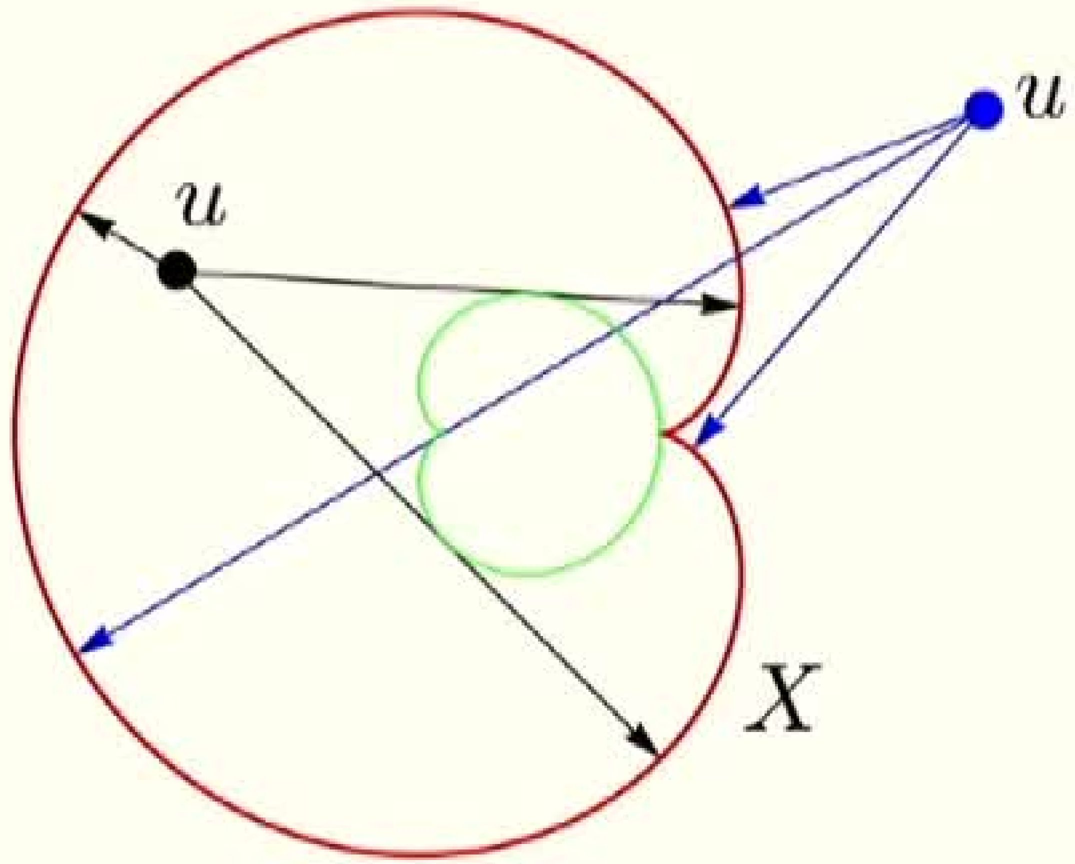


Euclidean distance degree of an algebraic variety

$u \in \mathbb{R}^n$ *data point*

$X \subseteq \mathbb{R}^n$ *real algebraic variety*

$$\min_{x \in X} \|x - u\|^2$$



Theorem: Draisma, Horobet, Ottaviani, Sturmfels, T. (2016)

smooth critical points of this problem is a constant for generic u

Euclidean distance degree of X =

smooth critical points of this problem

Euclidean distance degree of the essential variety

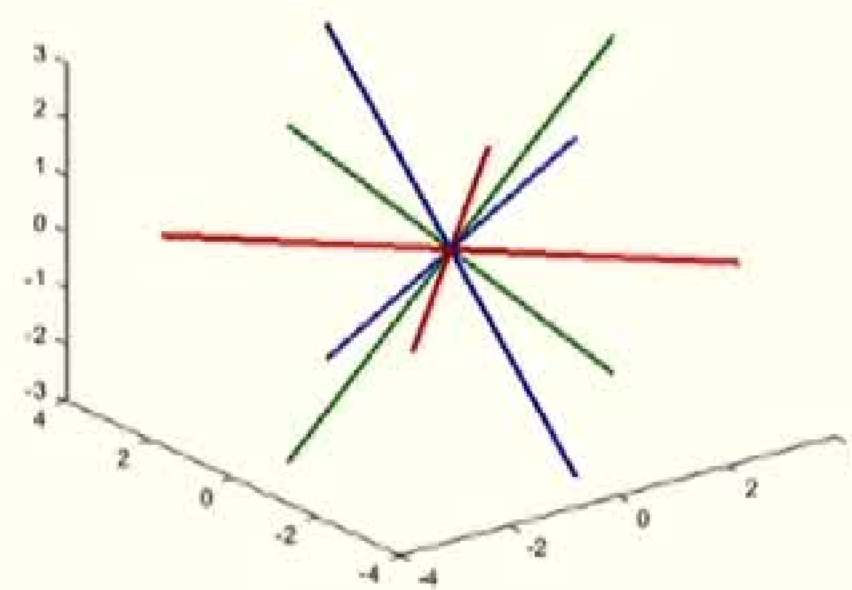
$$\mathcal{E} = \{E \in \mathbb{R}^{3\times 3} \, : \, \mathrm{rank}(E) = 2, \; \sigma_1(E) = \sigma_2(E) > 0\}$$

Theorem: Drusvyatskiy, Lee, Ottaviani, T. (2017)

$$\mathrm{EDdegree}(\mathcal{E}_{\mathbf{C}}) = 6$$

A transfer principle:

the EDdegree of orthogonally invariant matrix varieties equals the EDdegree of their singular values variety after symmetrization.



Euclidean distance degree of triangulation

X_n *multiview variety from n cameras*

Stewenius, Shaffalitzky, Nistér (2005):

n	2	3	4	5	6	7
$\text{EDdegree}(X_n)$	6	47	148	336	638	1081

Conjecture: $$\text{EDdegree}(X_n) = \frac{9}{2}n^3 - \frac{21}{2}n^2 + 8n - 4$$

A complex variant with a slightly different (cubic) formula was proven by **Harris & Lowengrub (2017)**