CSE 152: Computer Vision

Hao Su

Lecture 15: Fundamental Matrix

Agenda

- Why is stereo useful?
- Epipolar constraints
- **Fundamental matrix**

Cross Product as Matrix Multiplication

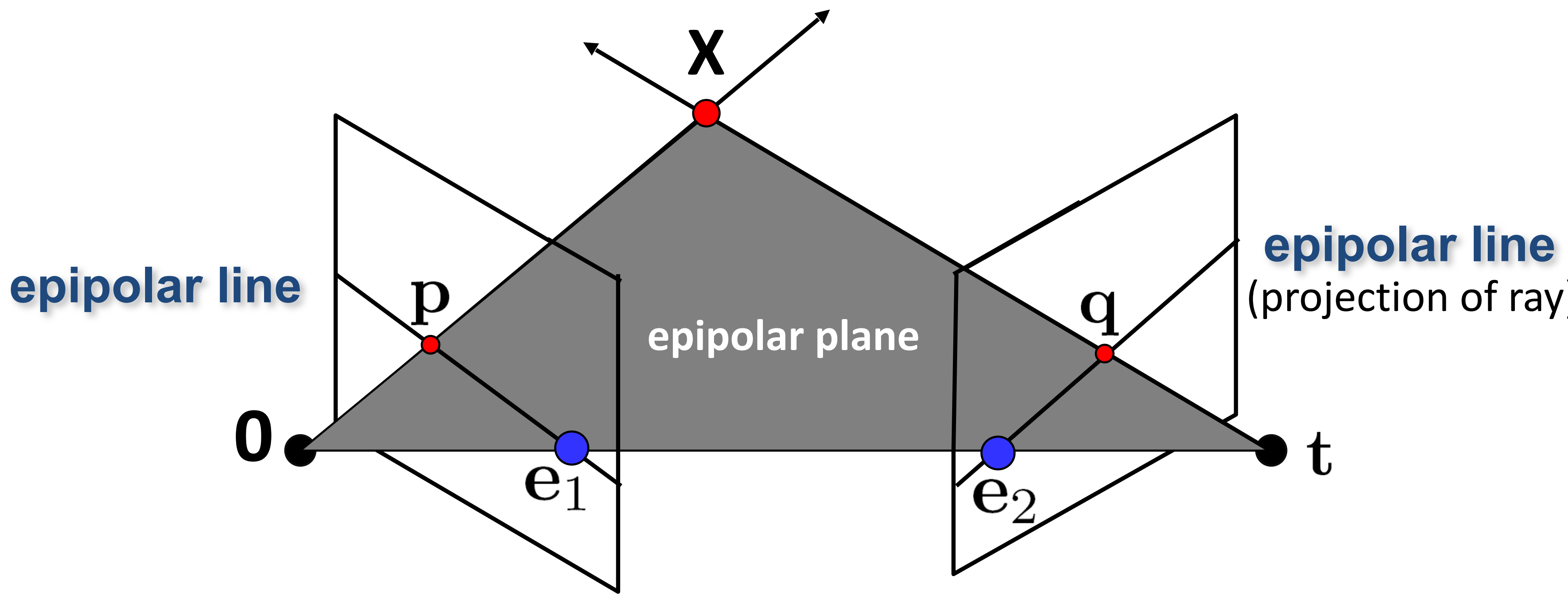
$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{bmatrix} \begin{bmatrix} b_x \\ b_y \\ b_z \end{bmatrix} = [\mathbf{a}_\times]\mathbf{b}$$

$$[a_\times] = -[a_\times]^T$$

“skew-symmetric matrix”

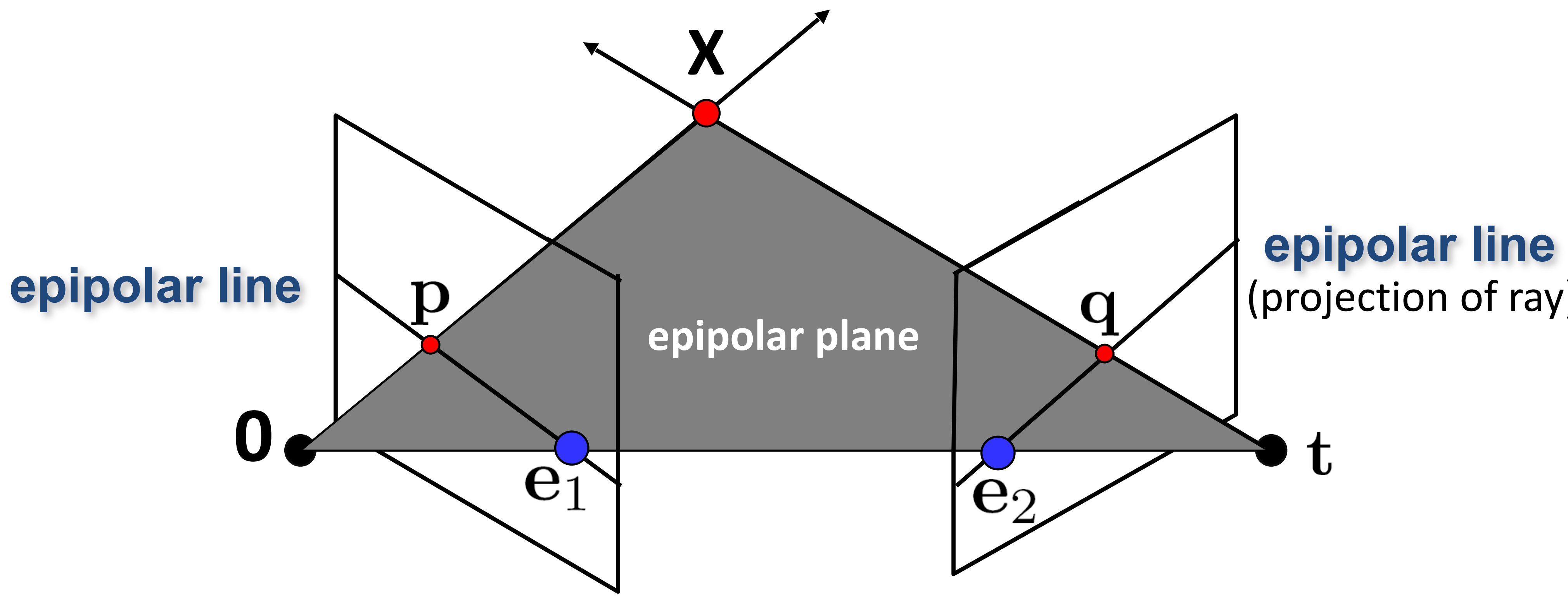
rank 2

Essential Matrix



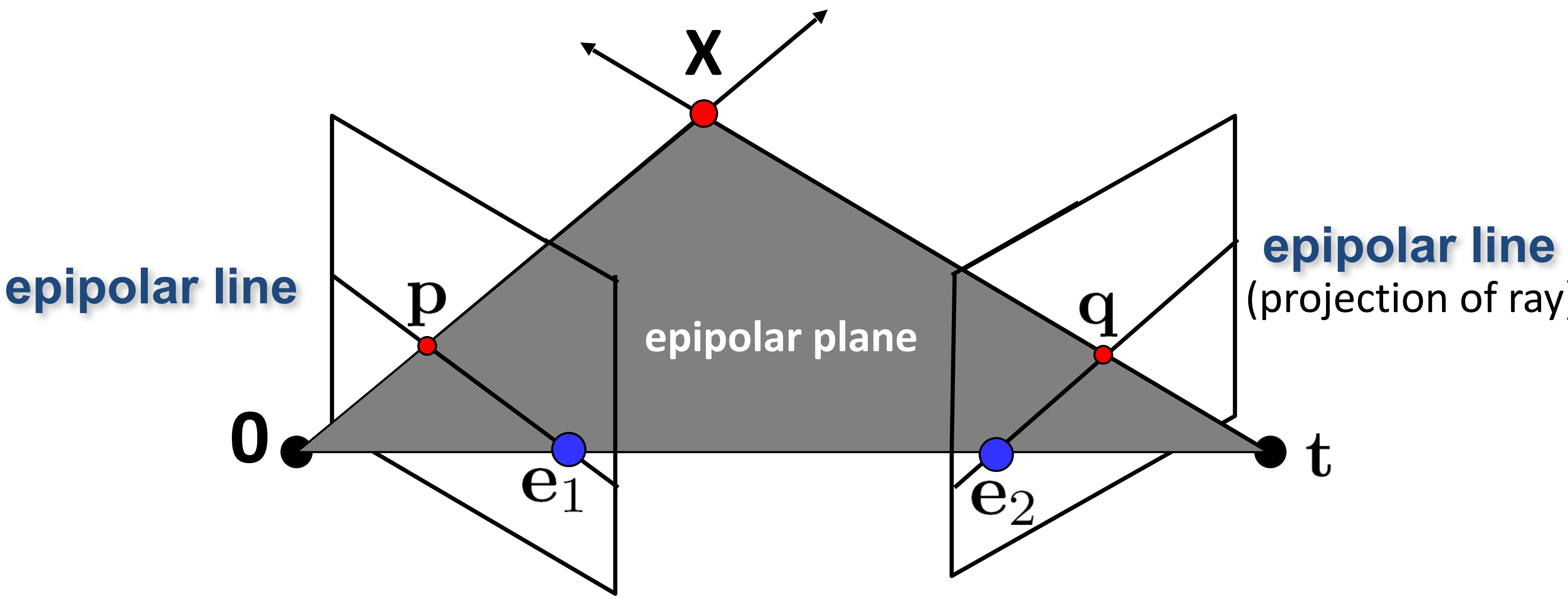
- Assume p and q in $\mathbb{R}^3$ are two points on the (virtual) image plane of two cameras
- Denoted by the pinhole frame coordinate in the corresponding cameras

Essential Matrix



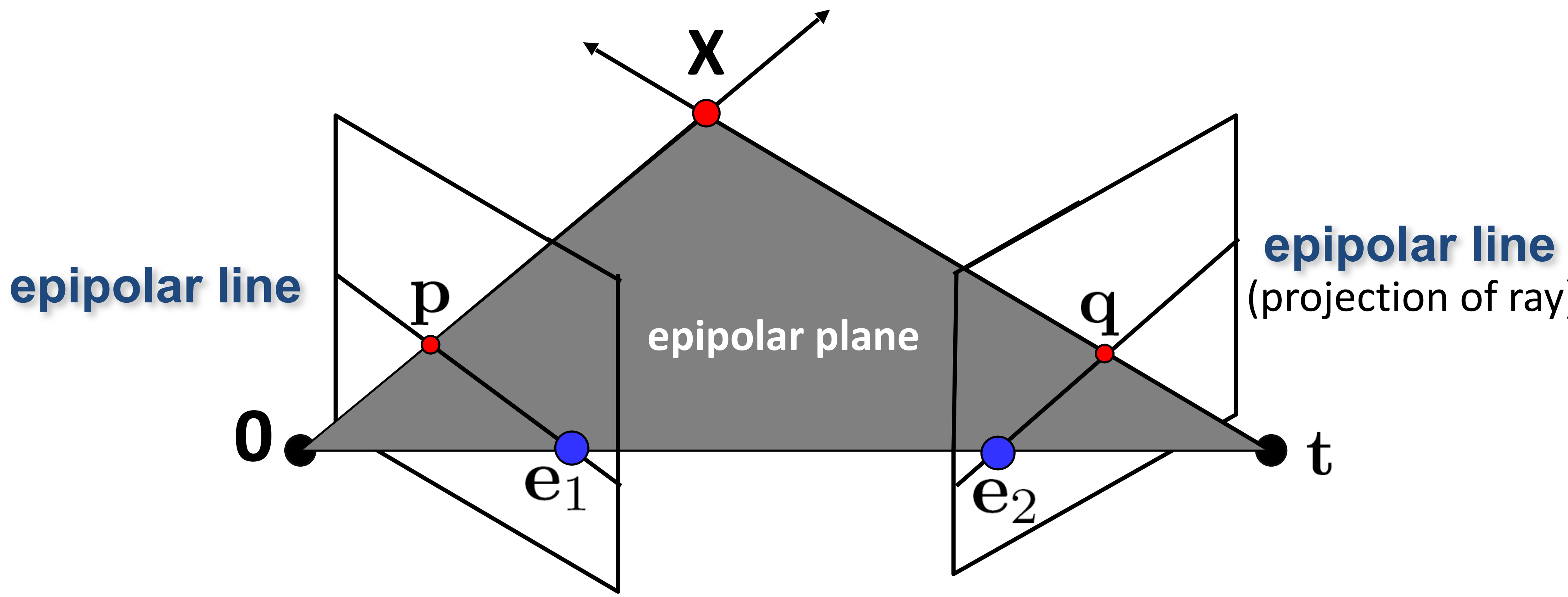
- Let camera 1 be [**I**, **0**] and camera 2 be [**R**, **t**].

Essential Matrix



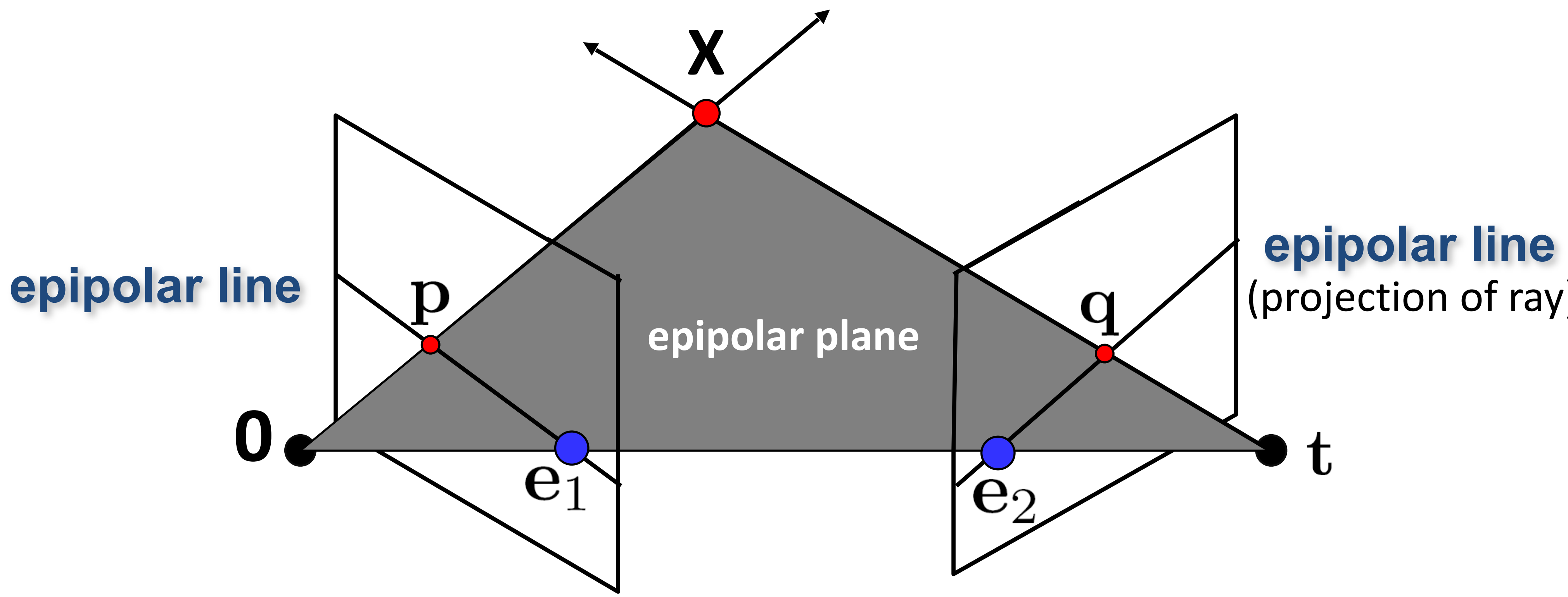
- Let camera 1 be [**I**, **0**] and camera 2 be [**R**, **t**].
- In camera 1 pinhole frame, 3D point **X** is given by $\mathbf{X}_1 = \lambda_1 \mathbf{p}$

Essential Matrix



- Let camera 1 be [**I**, **0**] and camera 2 be [**R**, **t**].
- In camera 1 pinhole frame, 3D point **X** is given by $\mathbf{X}_1 = \lambda_1 \mathbf{p}$
- In camera 2 pinhole frame, 3D point **X** is given by $\mathbf{X}_2 = \lambda_2 \mathbf{q}$

Essential Matrix

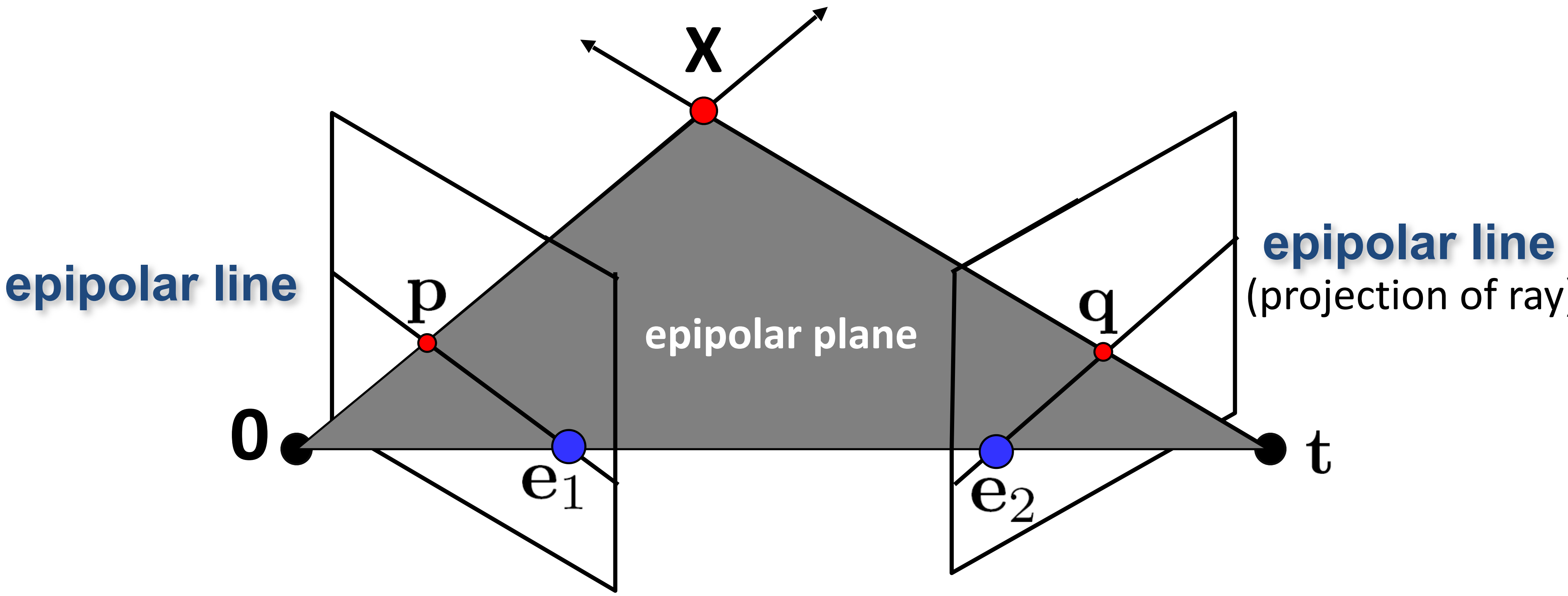


- Let camera 1 be [**I**, **0**] and camera 2 be [**R**, **t**].
- In camera 1 pinhole frame, 3D point **X** is given by $\mathbf{X}_1 = \lambda_1 \mathbf{p}$
- In camera 2 pinhole frame, 3D point **X** is given by $\mathbf{X}_2 = \lambda_2 \mathbf{q}$
- Since camera 2 is related to camera 1 by rigid-body motion [**R**, **t**]

$$X_1 = RX_2 + t$$
$$\lambda_1 p = \lambda_2 Rq + t$$

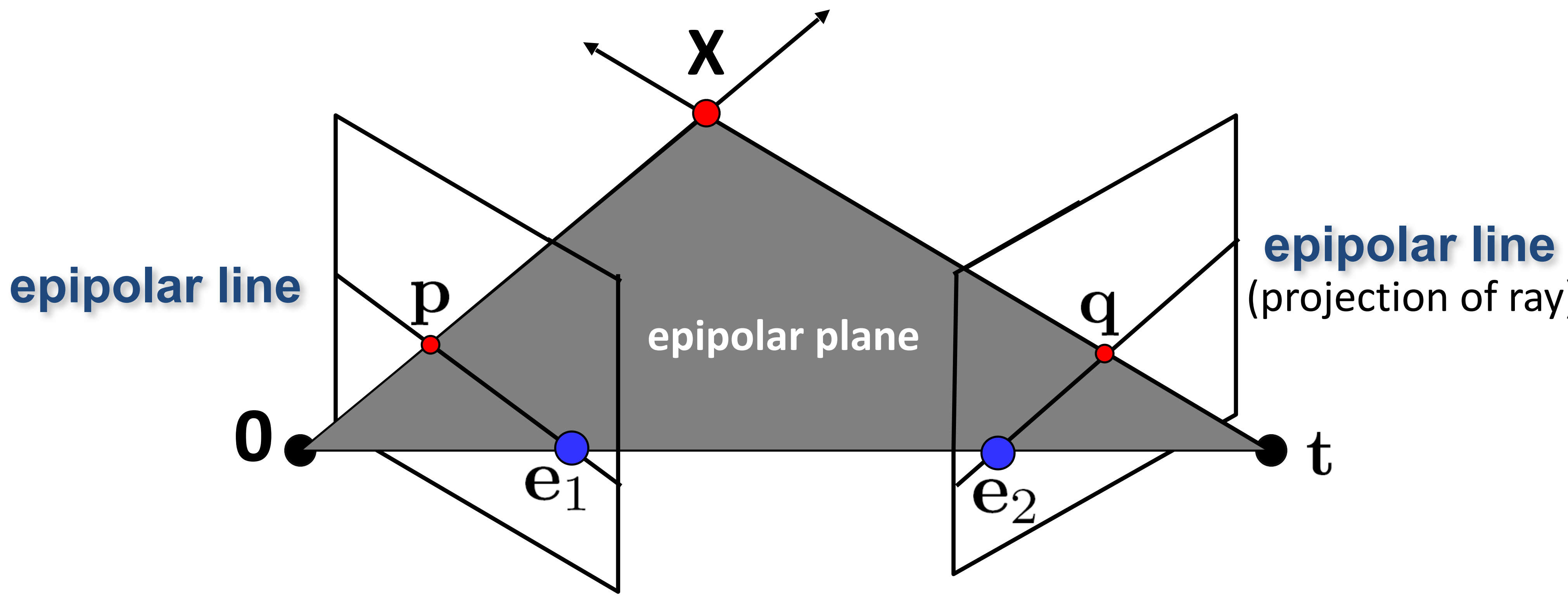
(change of coordinate system)

Essential Matrix



- We have: $\lambda_1 p = \lambda_2 R q + t$

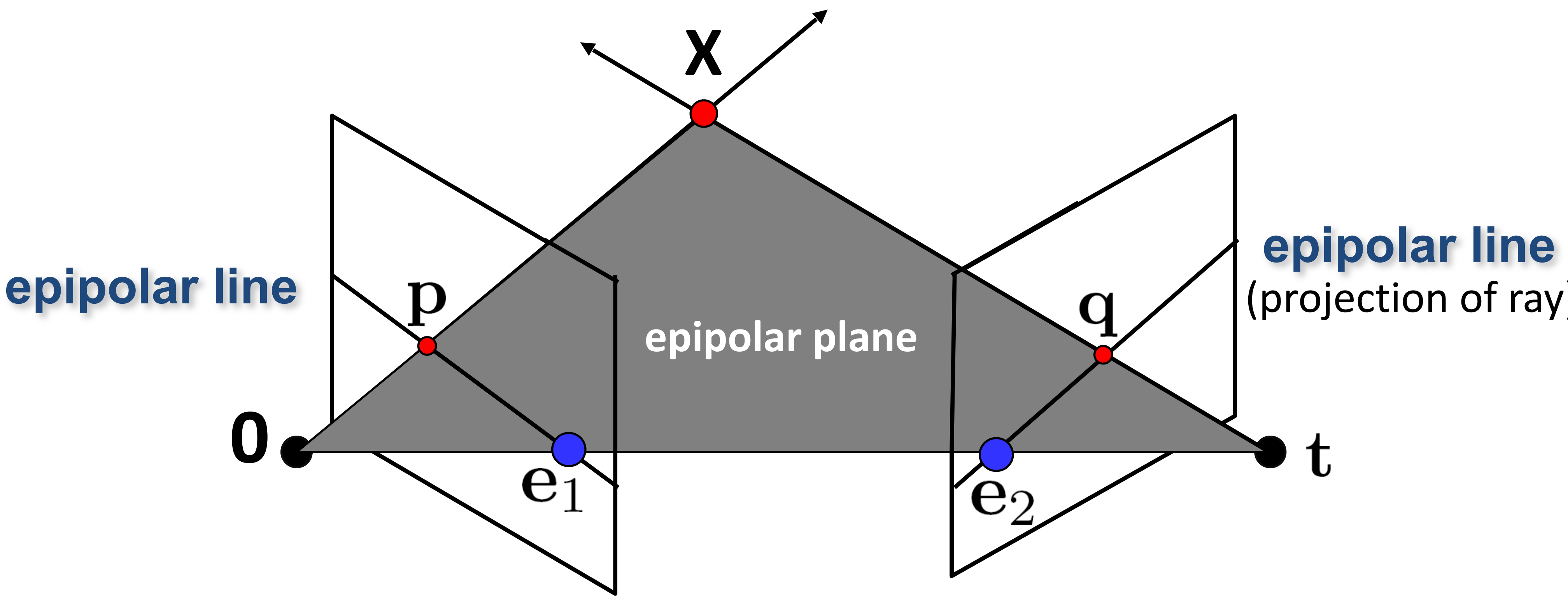
Essential Matrix



- We have: $\lambda_1 p = \lambda_2 Rq + t$
- Take cross-product with respect to **t**:

$$\lambda_1 [t]_\times p = \lambda_2 [t]_\times (Rq + t)$$

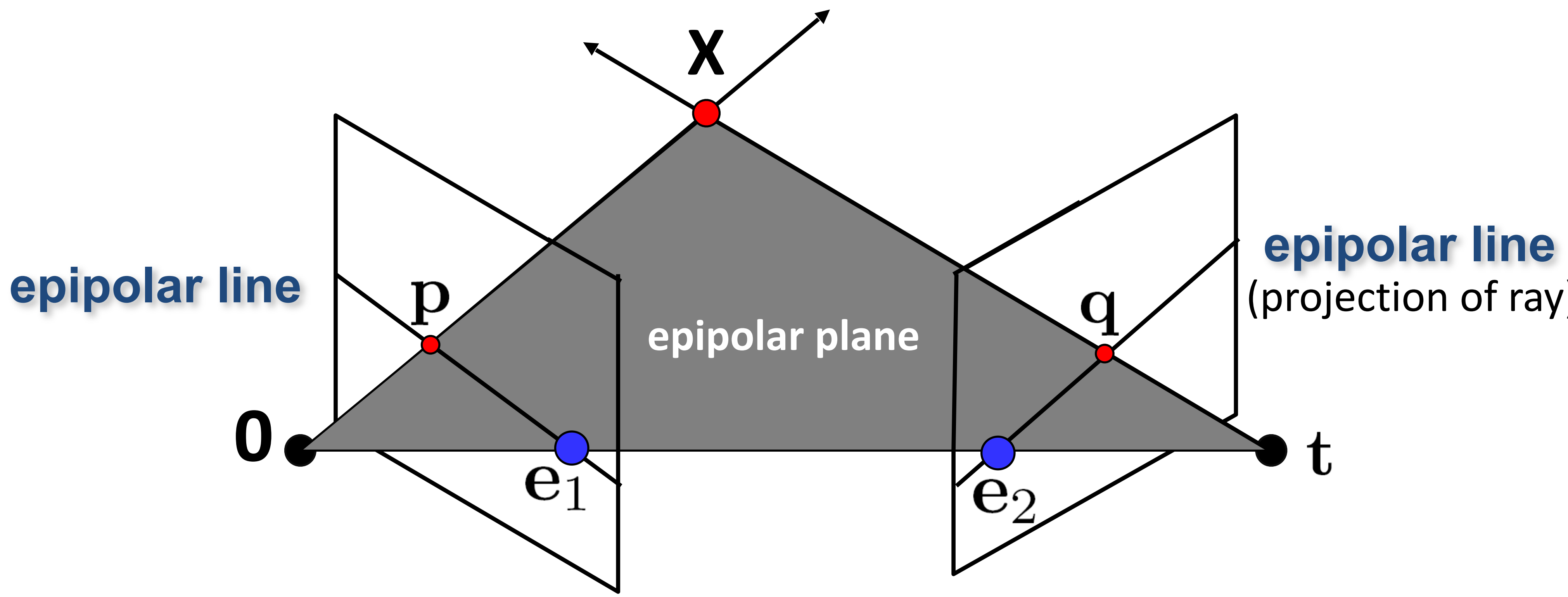
Essential Matrix



- We have: $\lambda_1 p = \lambda_2 R q + t$
- Take cross-product with respect to **t**:

$$\lambda_1 [t]_\times p = \lambda_2 [t]_\times R q$$

Essential Matrix



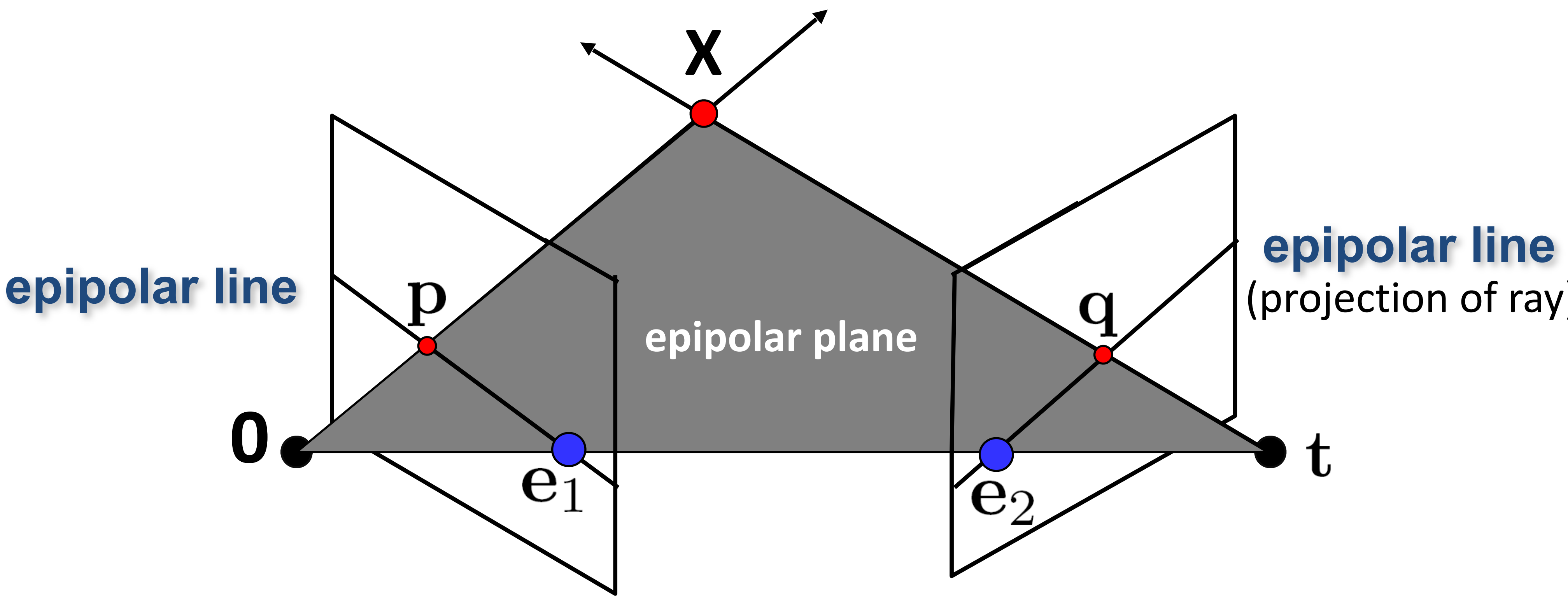
- We have: $\lambda_1 p = \lambda_2 R q + t$
- Take cross-product with respect to **t**:

$$\lambda_1 [t]_\times p = \lambda_2 [t]_\times R q$$

- Take dot-product with respect to **p**:

$$0 = \lambda_2 p^T [t]_\times R q$$

Essential Matrix



- We have: $p^T[t]_\times Rq = 0$
- Define: $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$

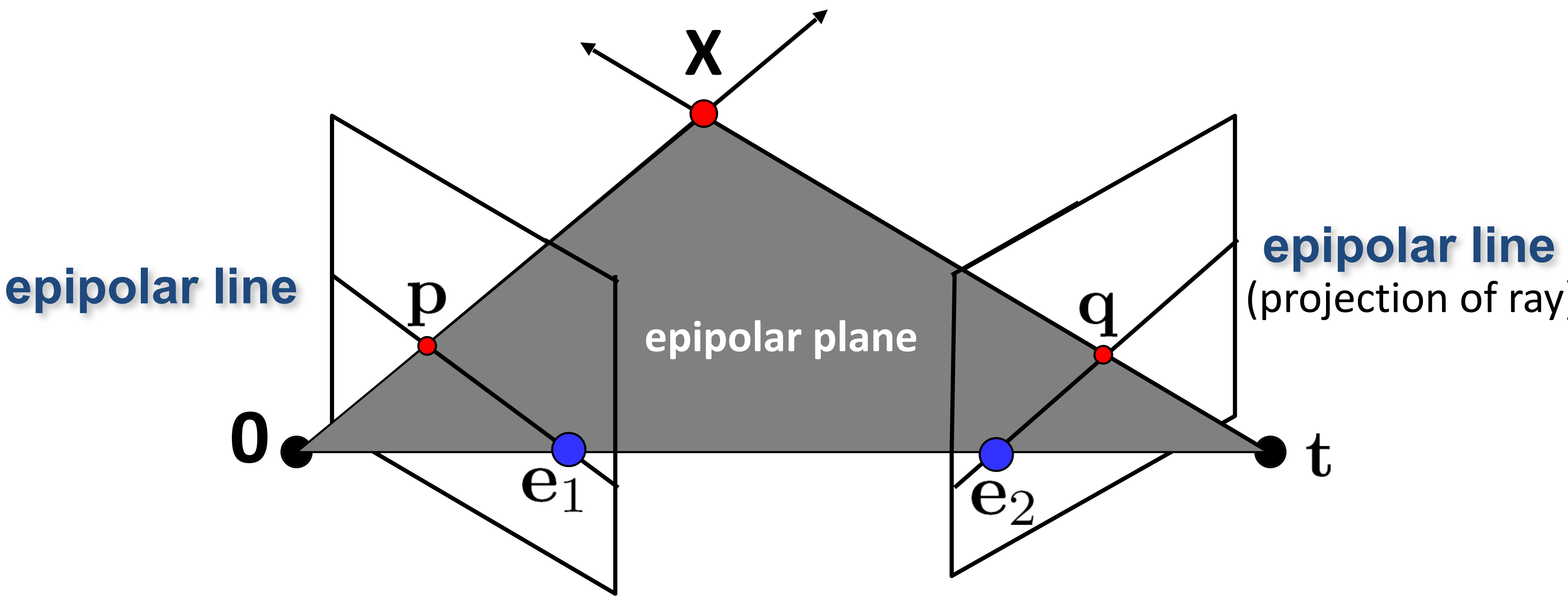
Essential matrix

rank(E)=2

- Then, we have:

$$\boxed{p^T E q = 0}$$

Fundamental Matrix



- Consider intrinsic camera matrices
- Then, **p** and **q** are in the pinhole frame and pixel counterparts are:

$$\mathbf{p}' = \mathbf{K}_1\mathbf{p} \qquad \mathbf{q}' = \mathbf{K}_2\mathbf{q}$$

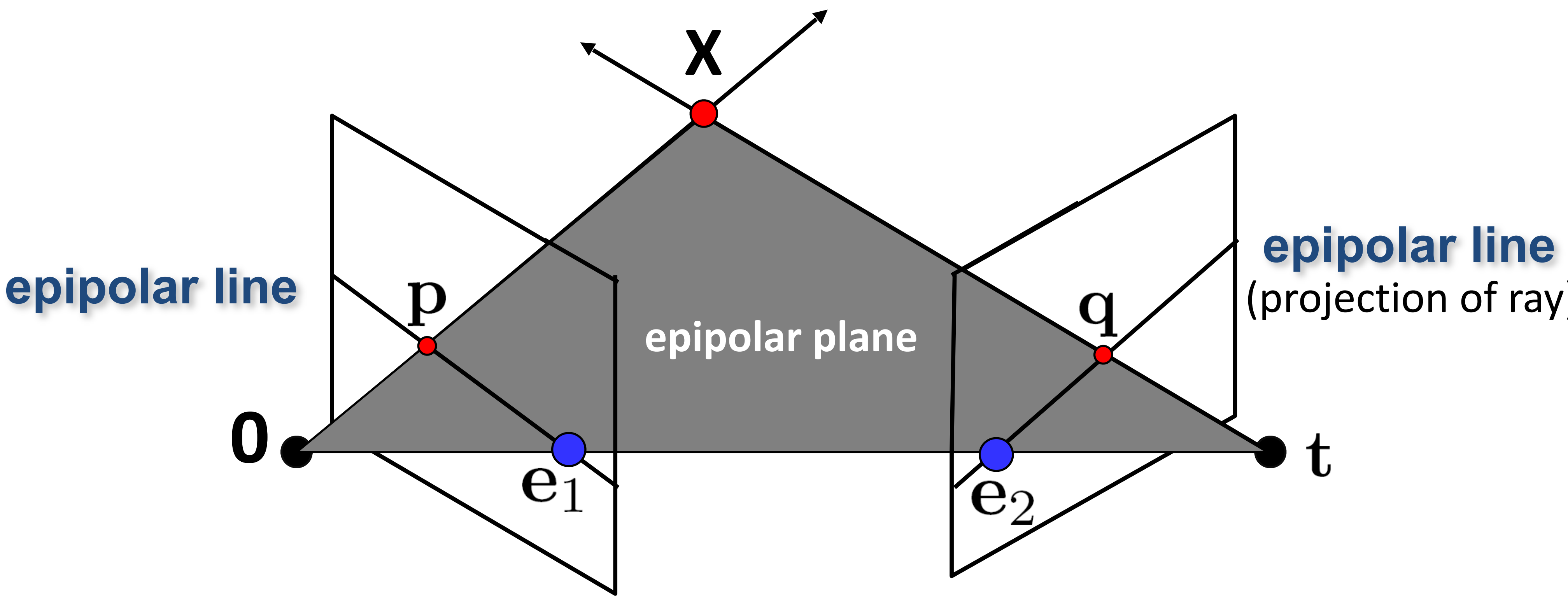
- Recall essential matrix constraint:

$$\boxed{p^T E q = 0}$$

- Substituting, we have:

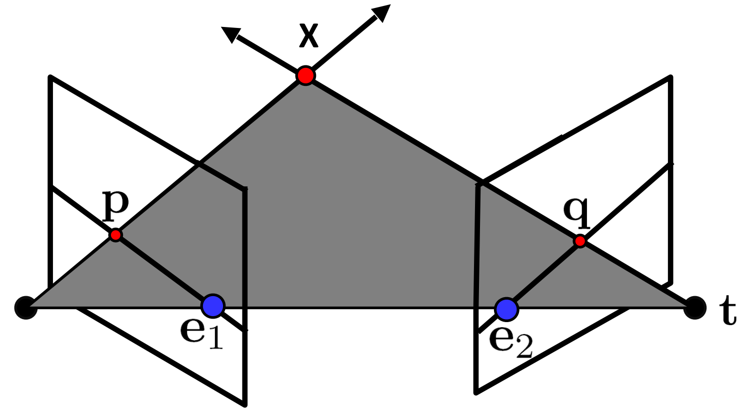
$$(K_1^{-1}p')^T E (K_2^{-1}q') = 0$$

Fundamental Matrix



- Essential matrix constraint in pixel space: $(K_1^{-1}p')^T E(K_2^{-1}q') = 0$.
- Rearranging: $p'^T K_1^{-T} E K_2^{-1} q' = 0$
- Define: $F = K_1^{-T} E K_2^{-1}$ ← Fundamental matrix rank(F)=2
- Then, we have: $\boxed{p'^T F q' = 0}$

Fundamental Matrix



$$\boxed{p'^T F q' = 0}$$

$$[x_p, y_p, z_p] \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x_q \\ y_q \\ z_q \end{bmatrix} = 0$$

$\cdot /z_p$ $\qquad$ $\cdot /z_q$

$$[x_p/z_p, y_p/z_p, 1] \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x_q/z_q \\ y_q/z_q \\ 1 \end{bmatrix} = 0$$

pixel coordinates

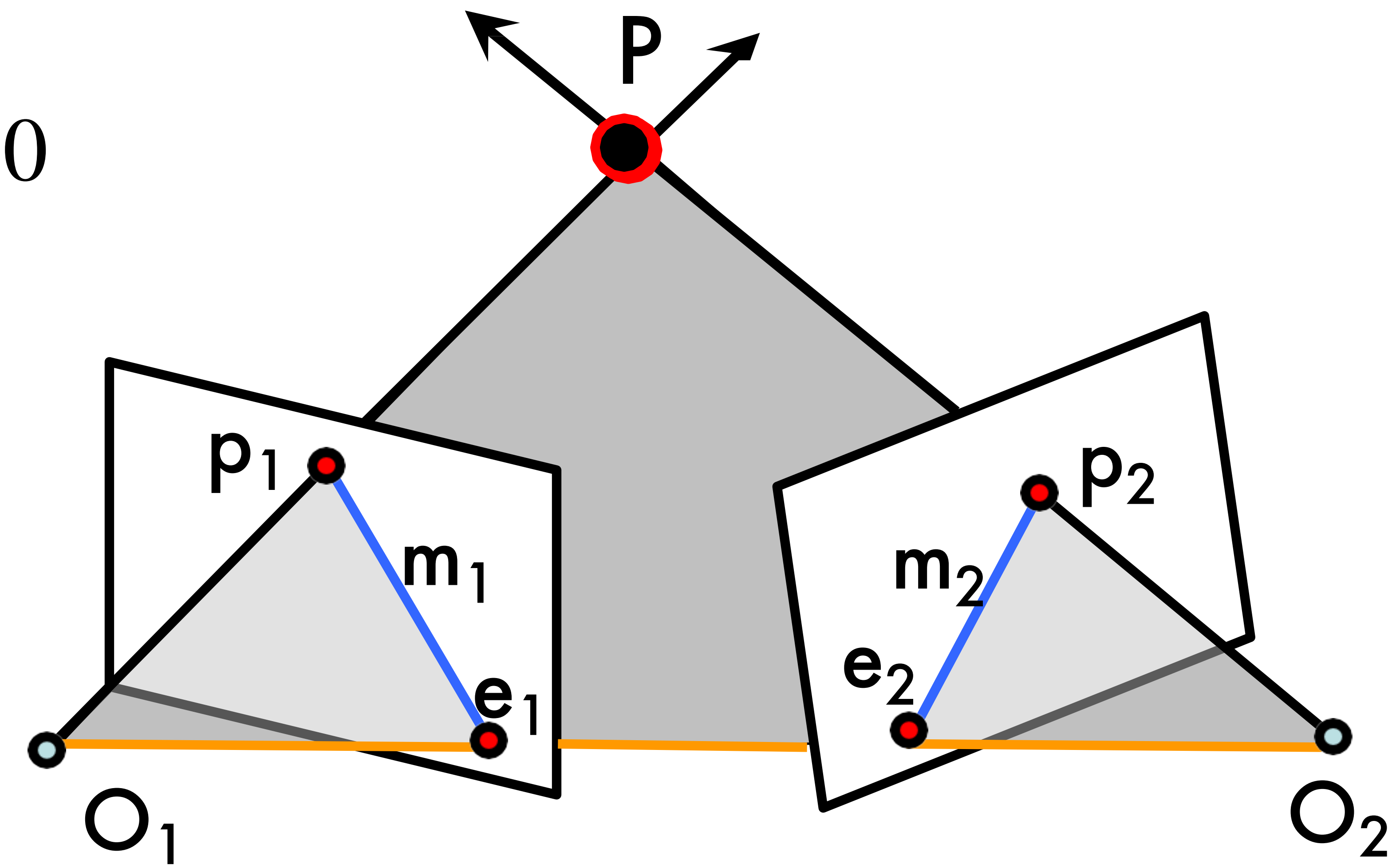
Epipolar Constraint

$$p_1^T \cdot Fp_2 = 0$$

- $w_1 = Fp_2$ defines an equation $w_1^T p_1 = 0$
- Note that, p_1 is the corresponding point of p_2 by the derivation of F
- So $w_1^T p_1 = 0$ is the line constraint that the corresponding point of p_2 has to satisfy
- But the line that corresponds to p_2 is the epipolar line, by the definition of epipolar line
- So, $w_1 = Fp_2$ defines the epipolar line of p_2

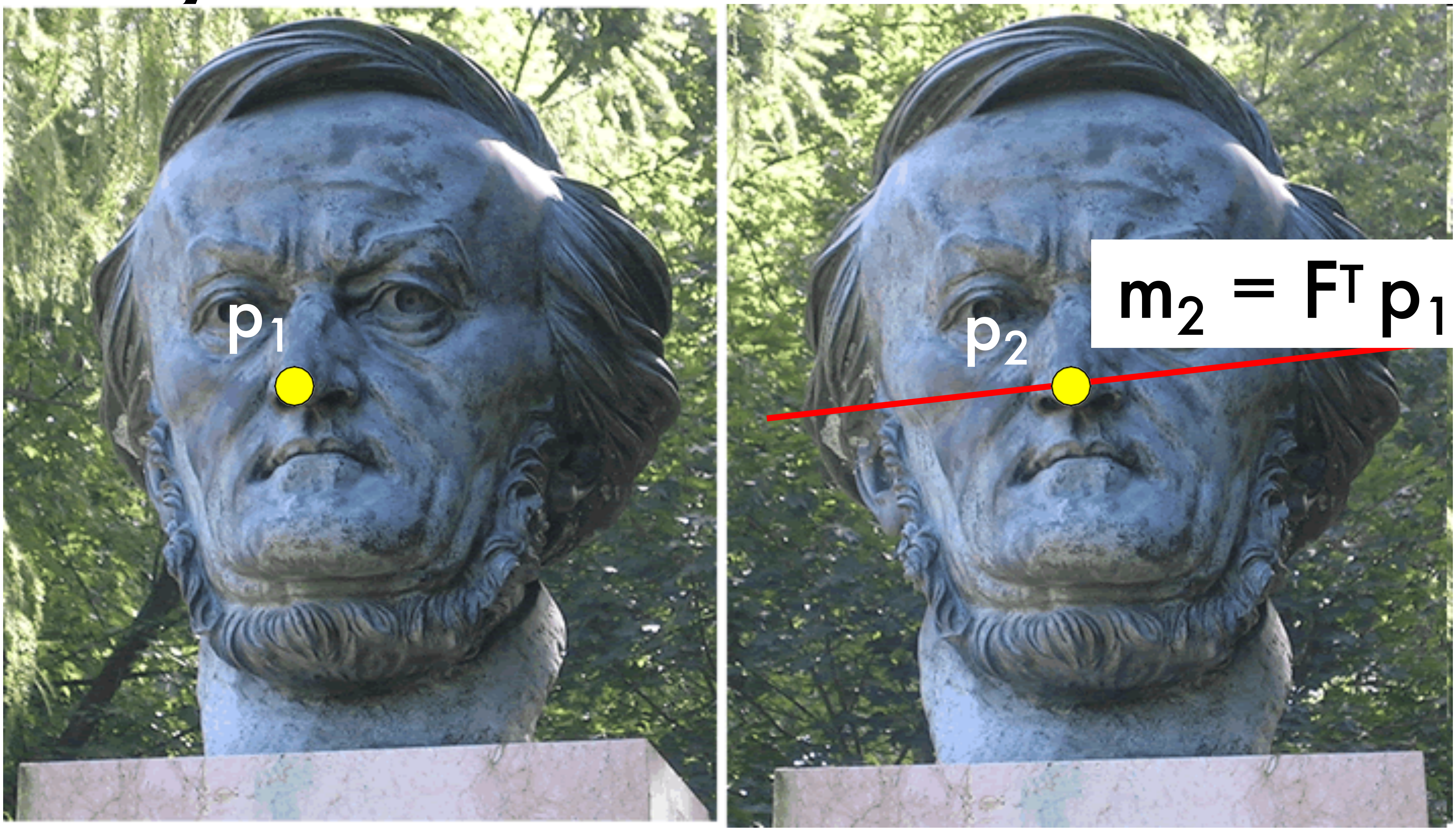
Epipolar Constraint

$$p_1^T \cdot F p_2 = 0$$



- w_1 = F p_2 defines an equation $w_1^T p_1 = 0$, the epipolar line m_1 of p_2
- $w_2 = F^T p_1$ defines an equation $w_2^T p_2 = 0$, the epipolar line m_2 of p_1
- F is singular (rank two)
- F e_2 = 0 and $F^T e_1 = 0$

Why F is useful?



- Suppose F is known
- No additional information about the scene and camera is given
- Given a point on left image, we can compute the corresponding epipolar line in the second image

Why F is useful?

- F captures information about the epipolar geometry of 2 views + camera parameters

- MORE IMPORTANTLY: F gives constraints on how the scene changes under view point transformation (without reconstructing the scene!)

- Powerful tool in:
  - 3D reconstruction
  - Multi-view object/scene matching