Omnidirectional vision geometry and signal processing

Kostas Daniilidis
with Christopher Geyer (now at UC Berkeley)
and Ameesh Makadia
GRASP Laboratory
Computer and Information Science Department
University of Pennsylvania

Demystifying catadioptric cameras

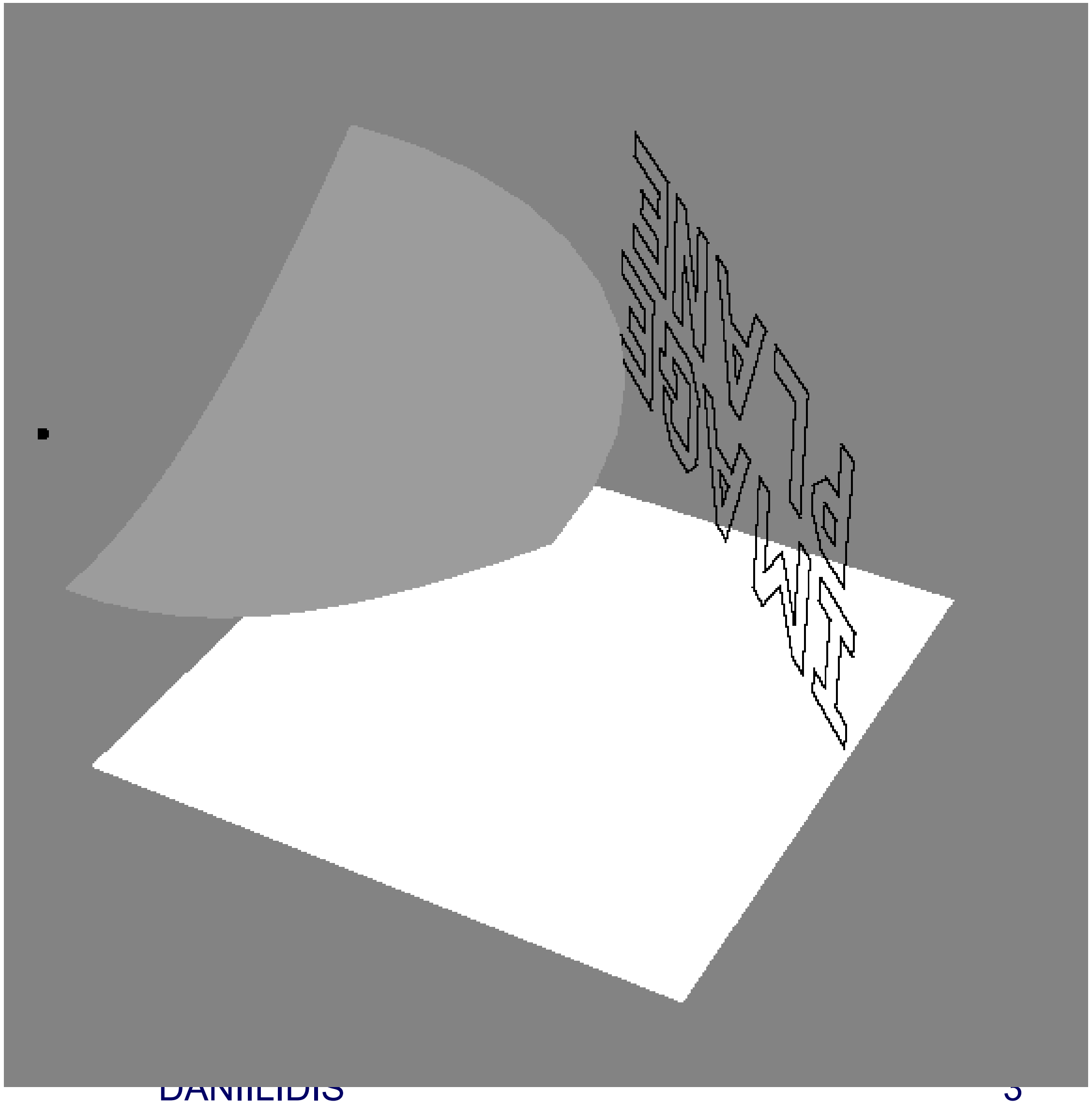
Simplify:
Catadioptric projections can be described by simple, intuitive models

Revelations:
Modeling catadioptric projections gives us insight into perspective cameras

Motion: To give a framework for studying structure-from-motion in parabolic mirror cameras

Signals: How to deal with the intensities on a sphere.

Central Catadioptric Projection



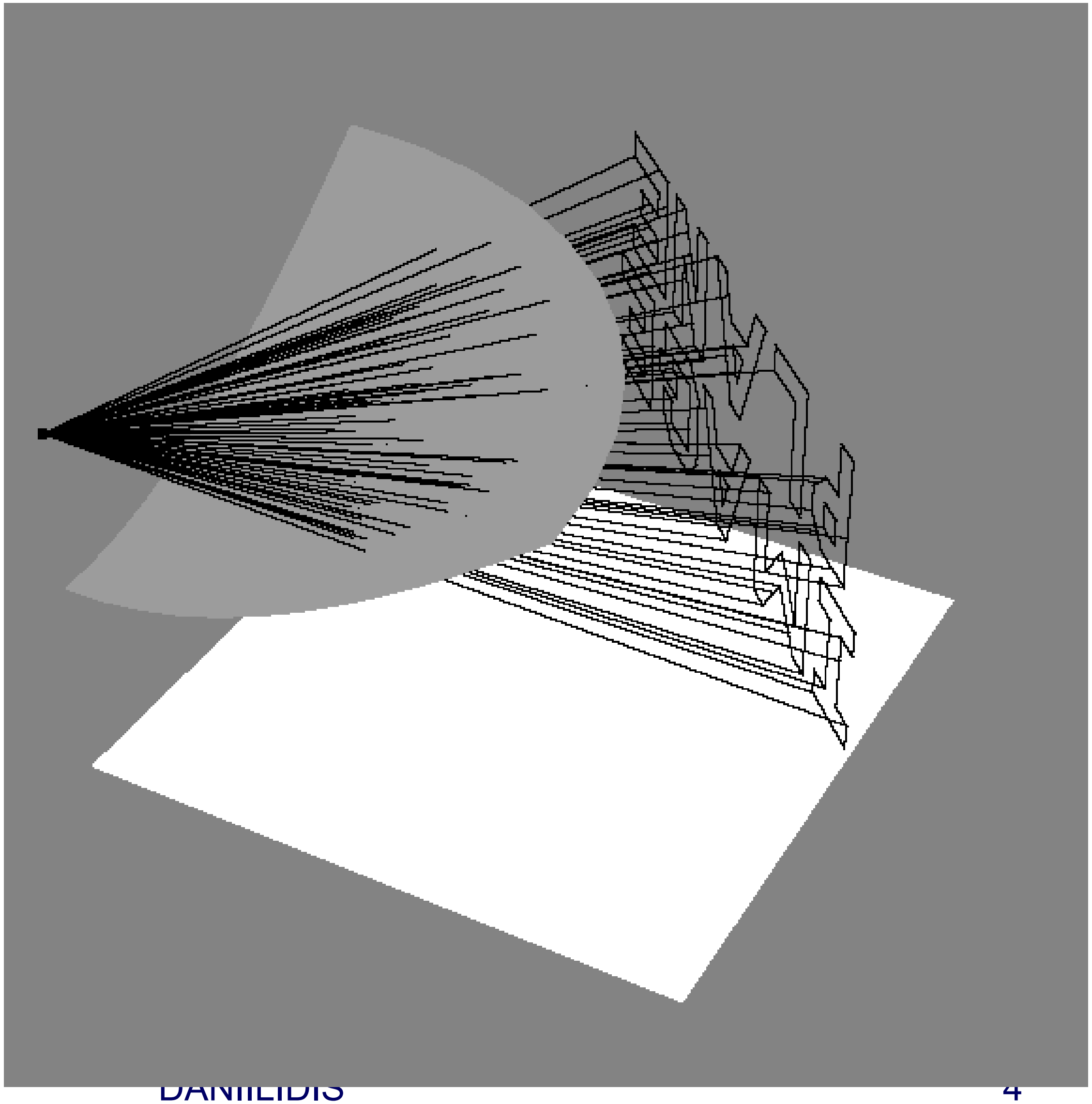
Setup.

An object in space

Hyperbolic mirror

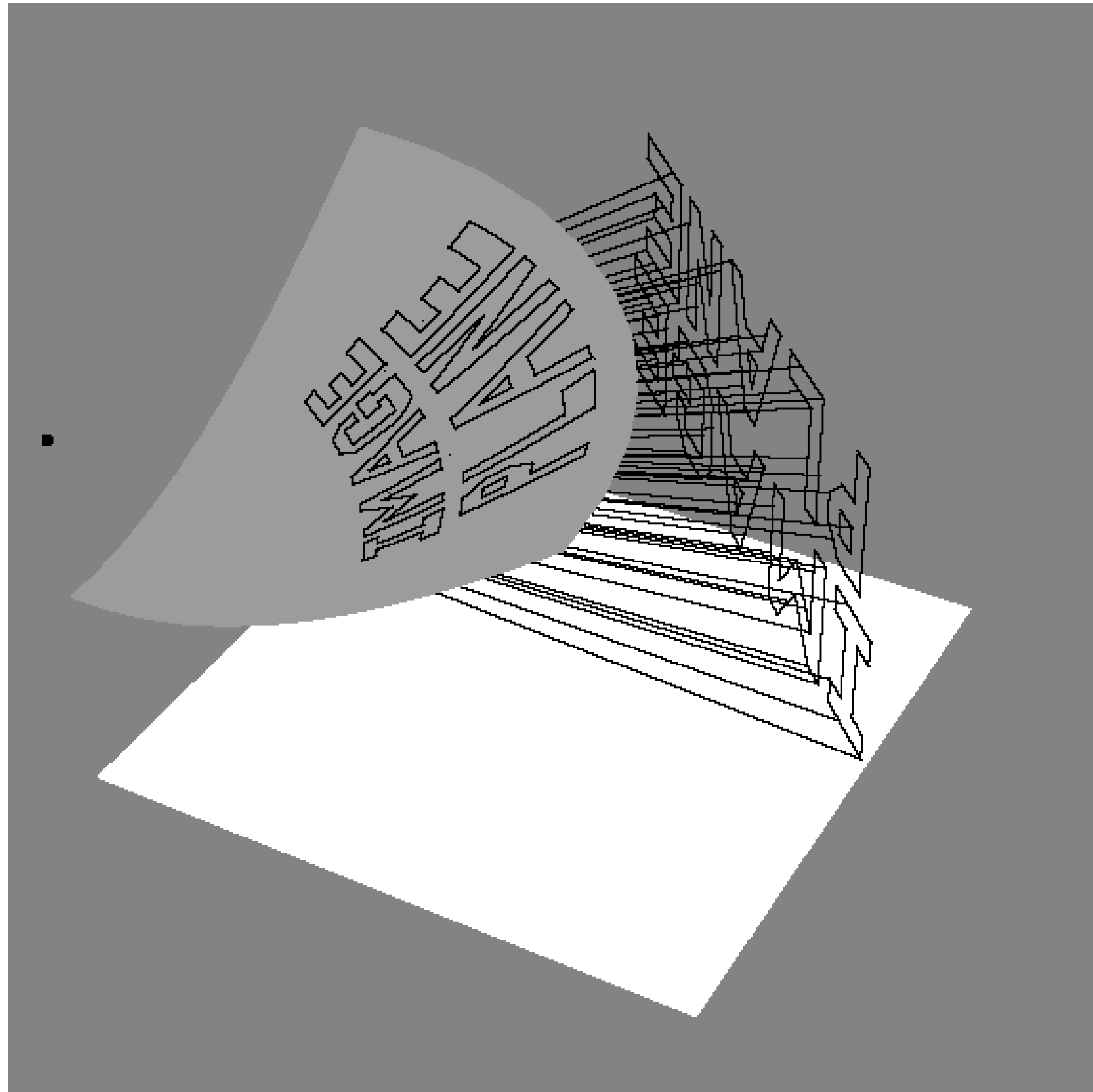
Image plane

Central Catadioptric Projection



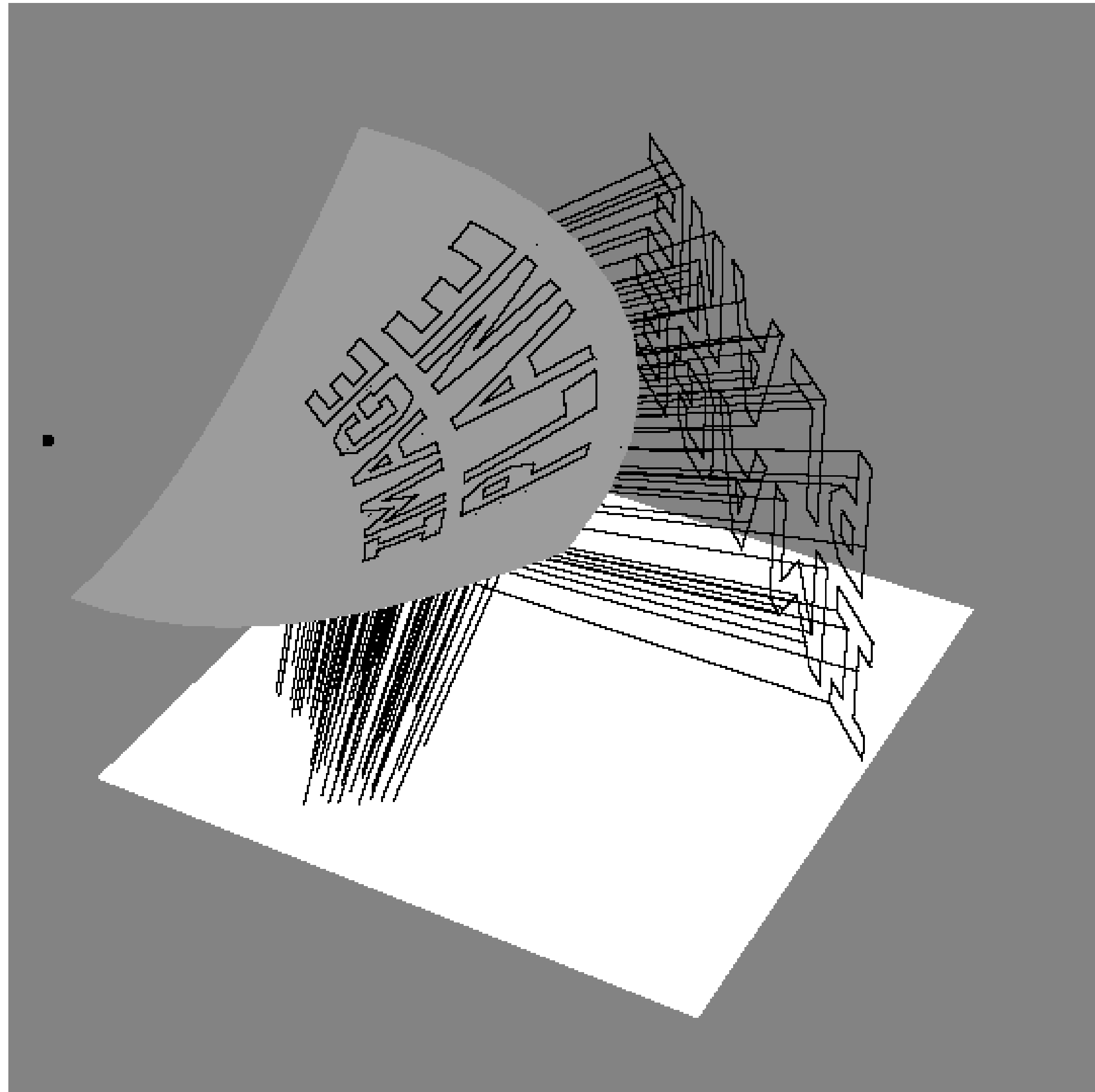
Rays through focus

Central Catadioptric Projection



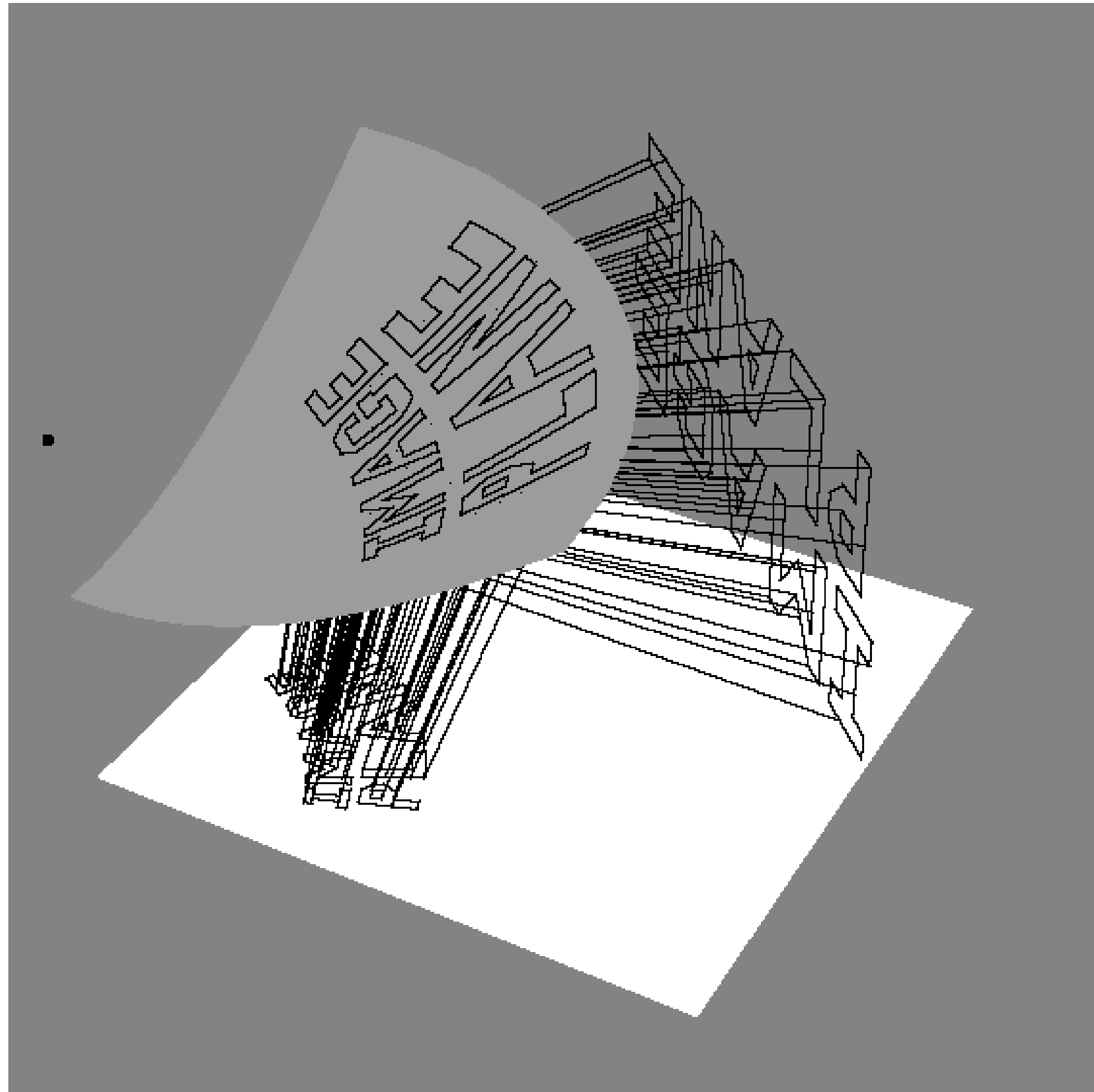
Intersected with hyperbola

Central Catadioptric Projection



Reflected rays incident with second focus

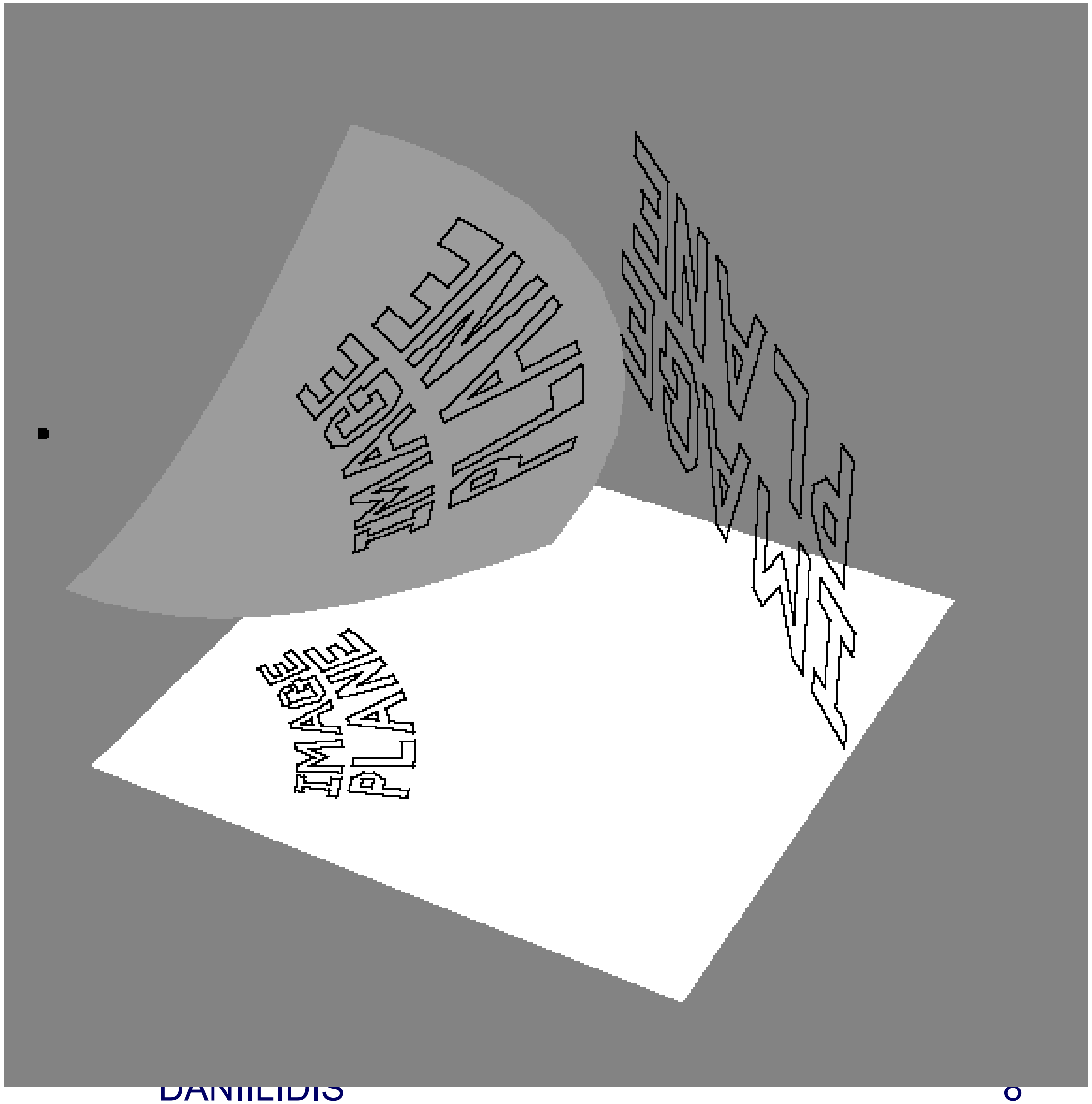
Central Catadioptric Projection



Intersected
with the image plane

Central Catadioptric Projection

is a double projection:
First on the mirror, then on the image plane.

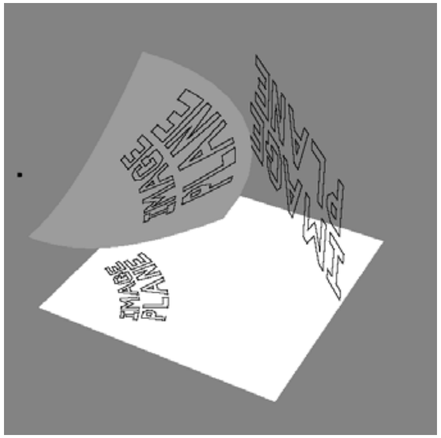


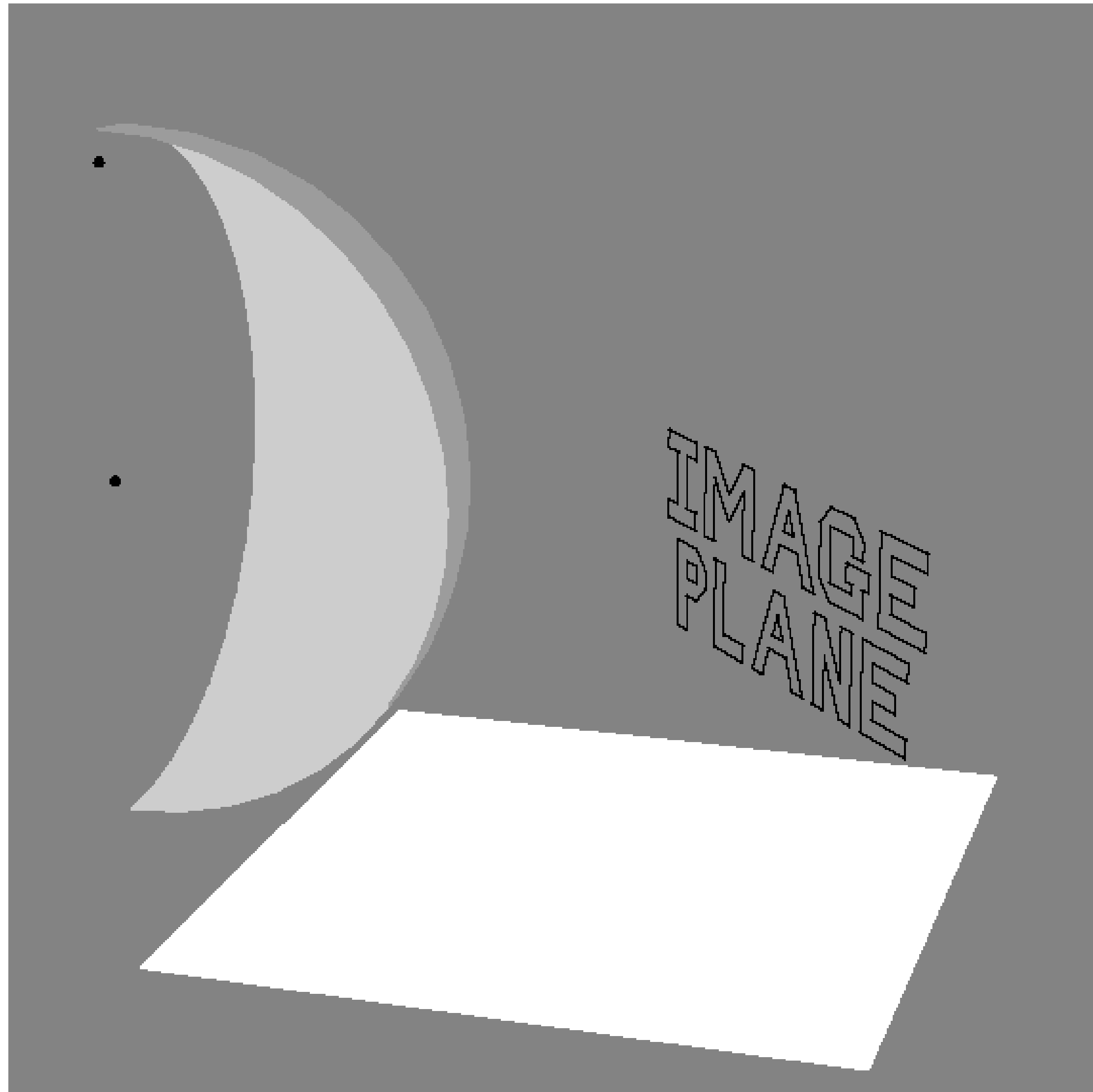
Unifying Theorem:

All central catadioptric projections are equivalent to double projection through the sphere.

Corollary: Conventional cameras are just a singularity.

Equivalence with the sphere





Setup.

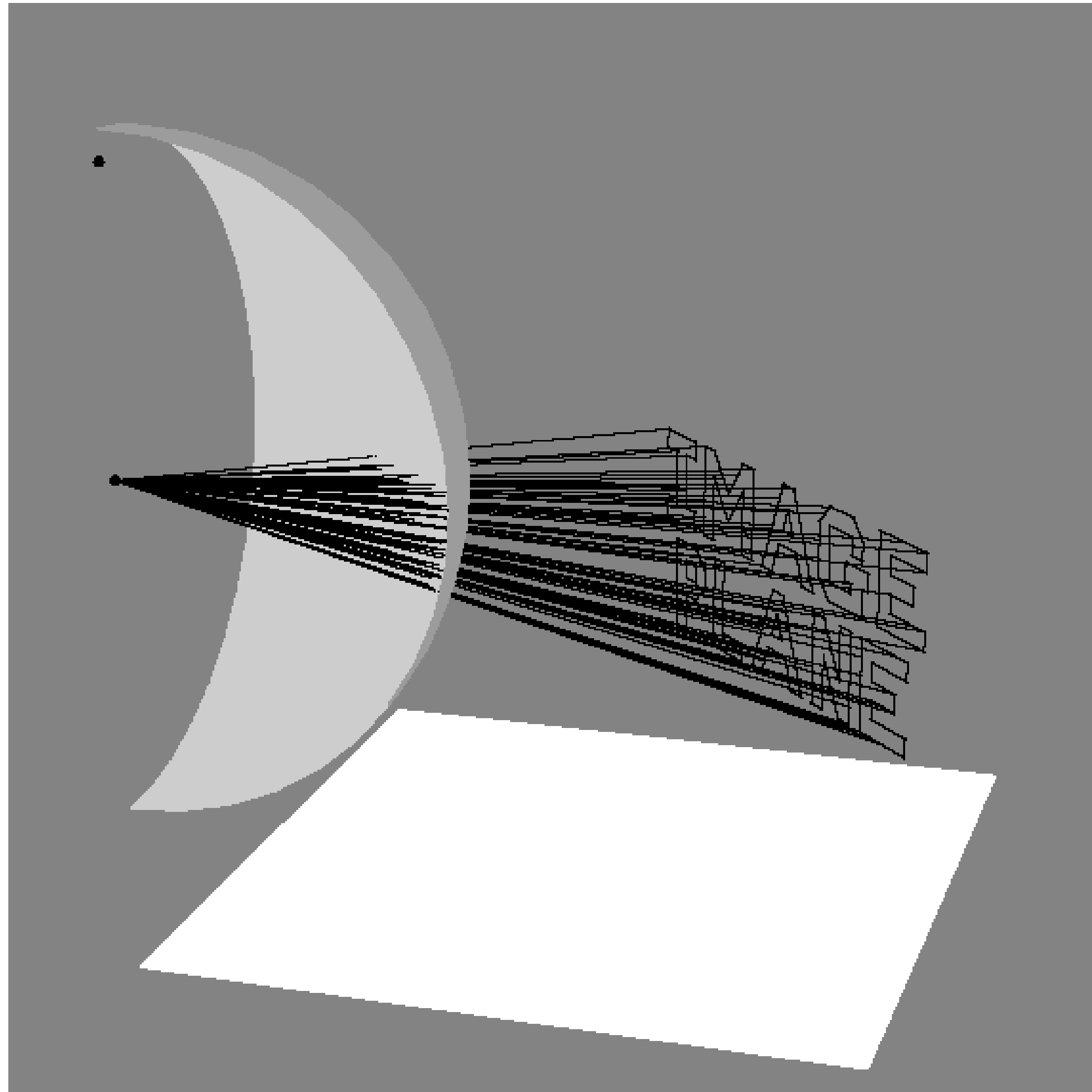
Object.

Sphere. Point on its axis.

Image plane.

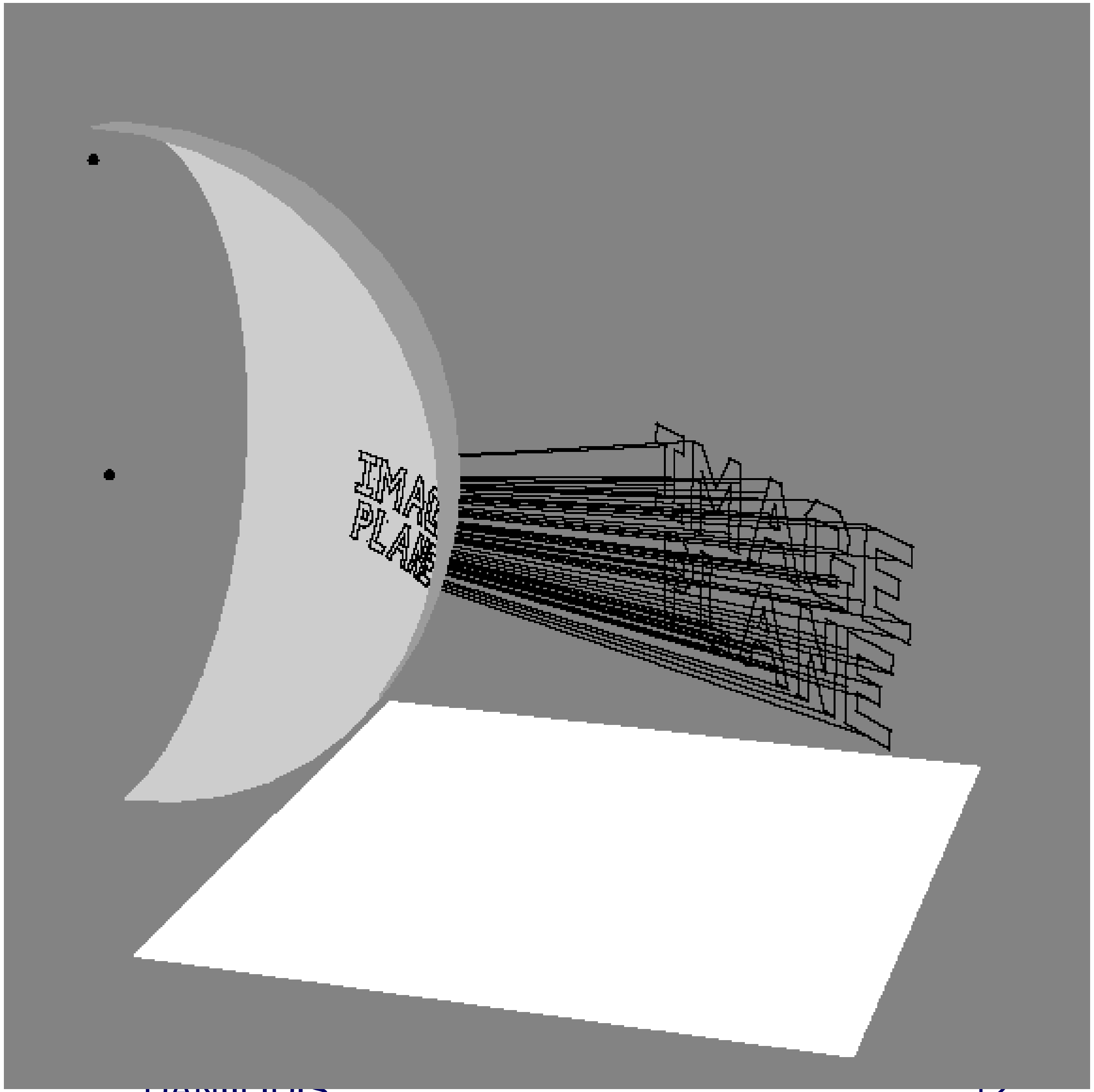
Equivalence with the sphere

Rays through sphere center



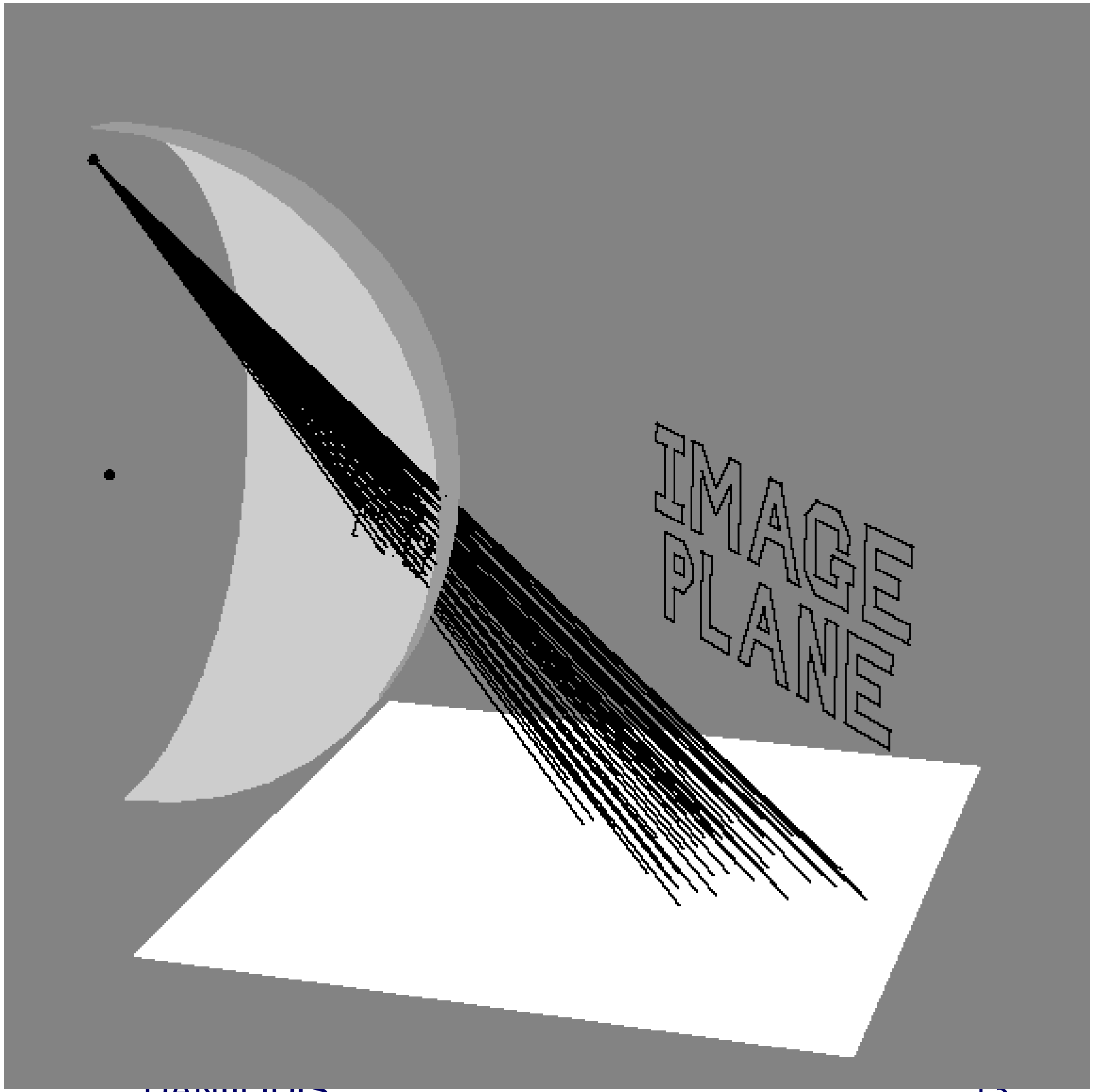
Equivalence with the sphere

Rays intersected with sphere



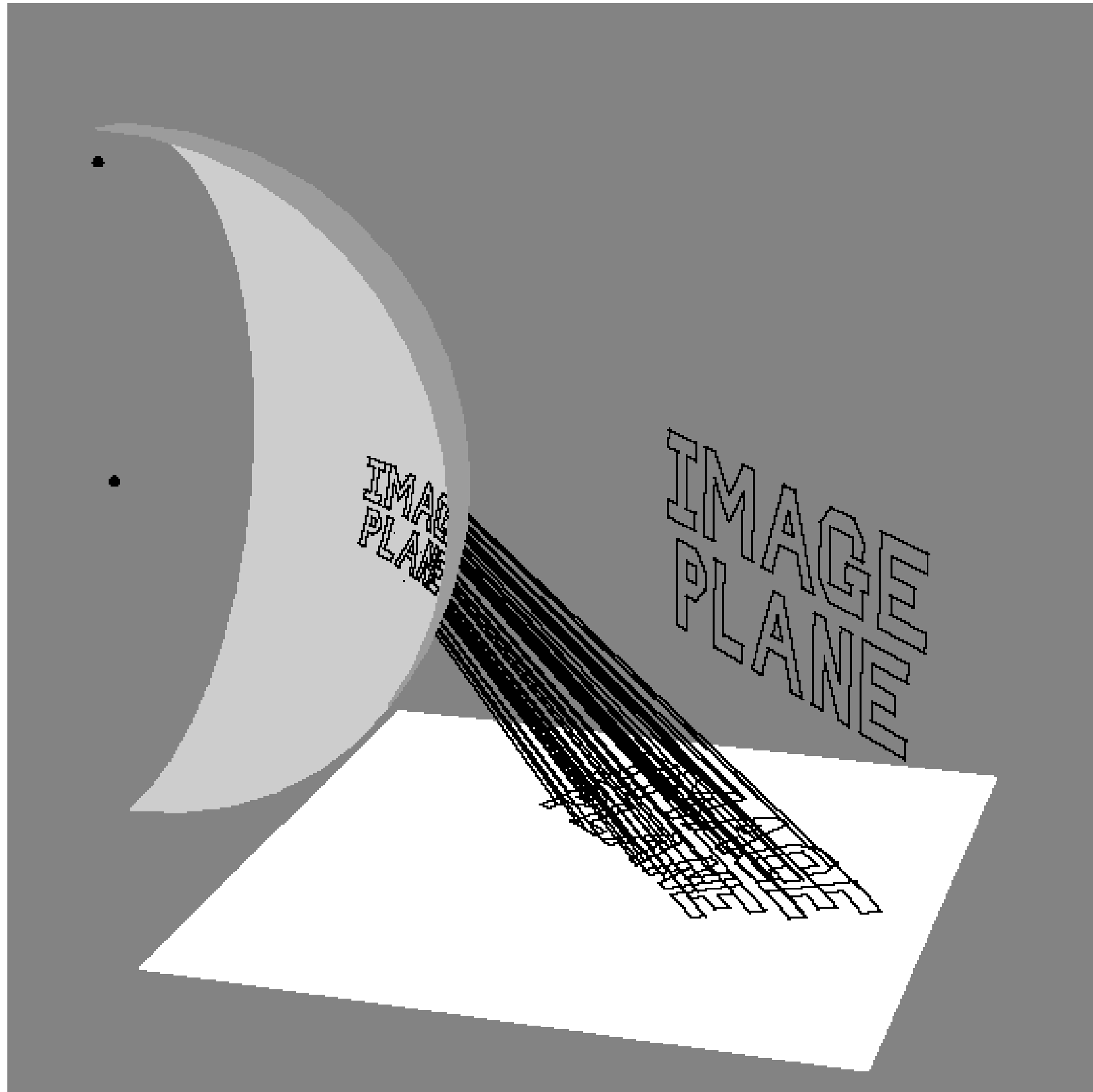
Equivalence with the sphere

Rays through point on axis



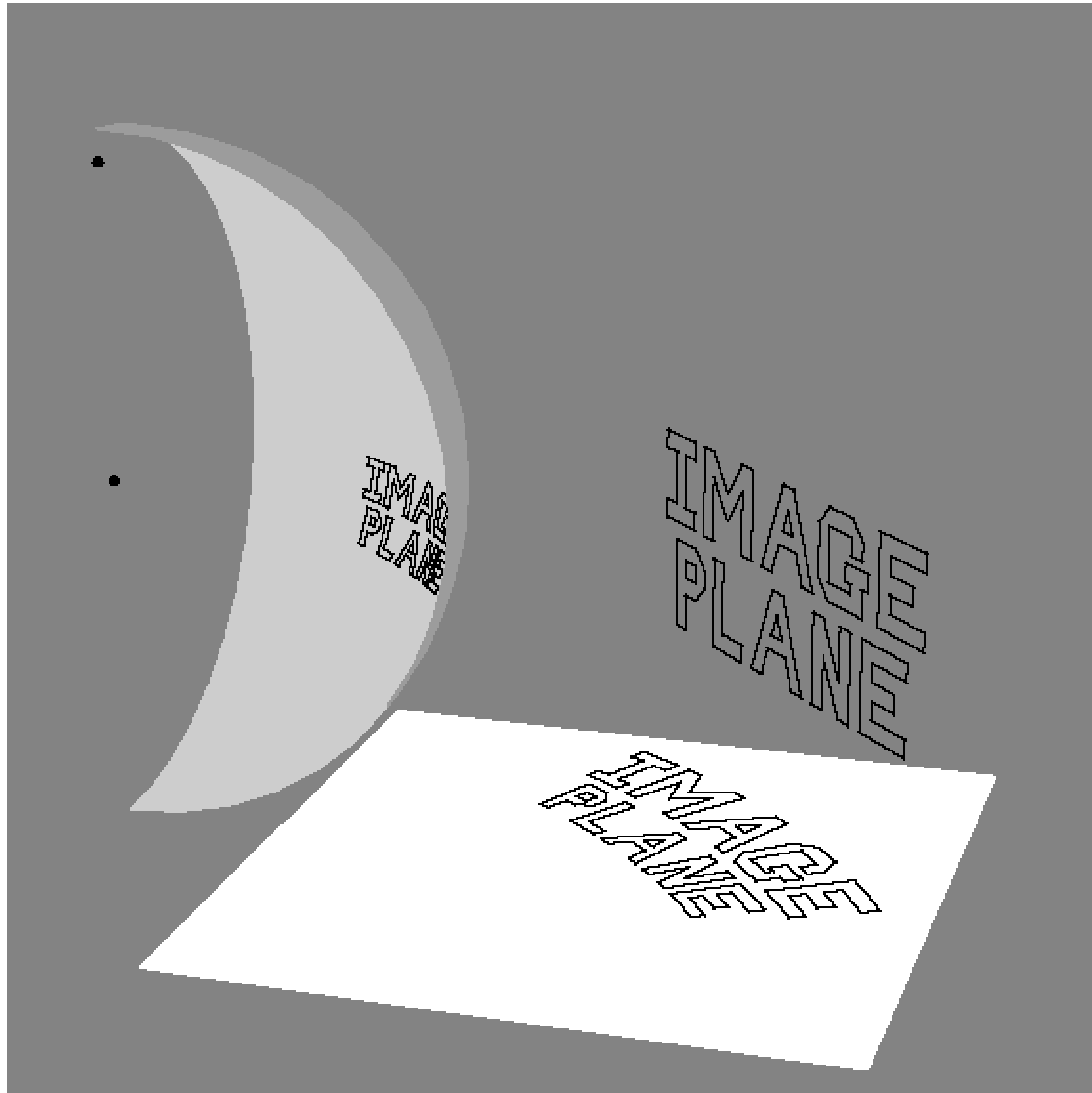
Equivalence with the sphere

Intersected with image plane



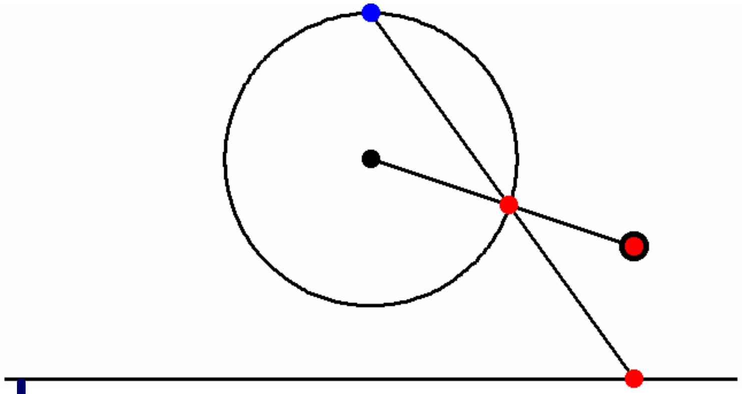
Equivalence with the sphere

Image of object obtained on image plane identical to catadioptric projection

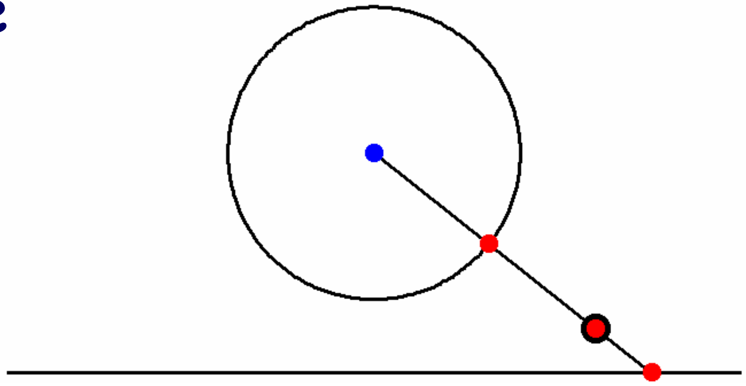


Two facts:

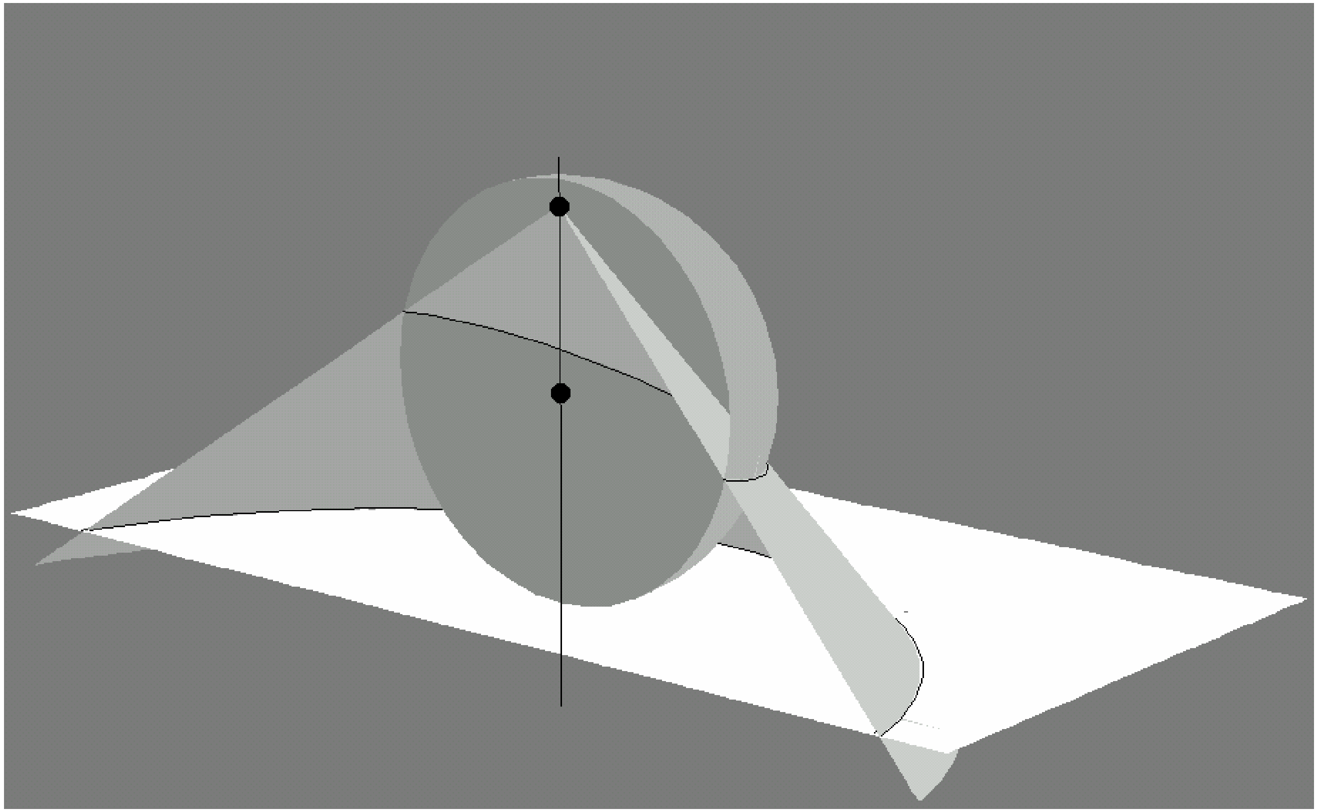
1. Parabolic projection = central projection to the sphere then stereo-graphic projection to a plane



2. Perspective projection = central projection to the sphere followed by central projection to a plane from the same center ! Our model covers all conventional perspective cameras!!



The projection of a line in space is a conic section and in parabolic mirrors it is a circle.

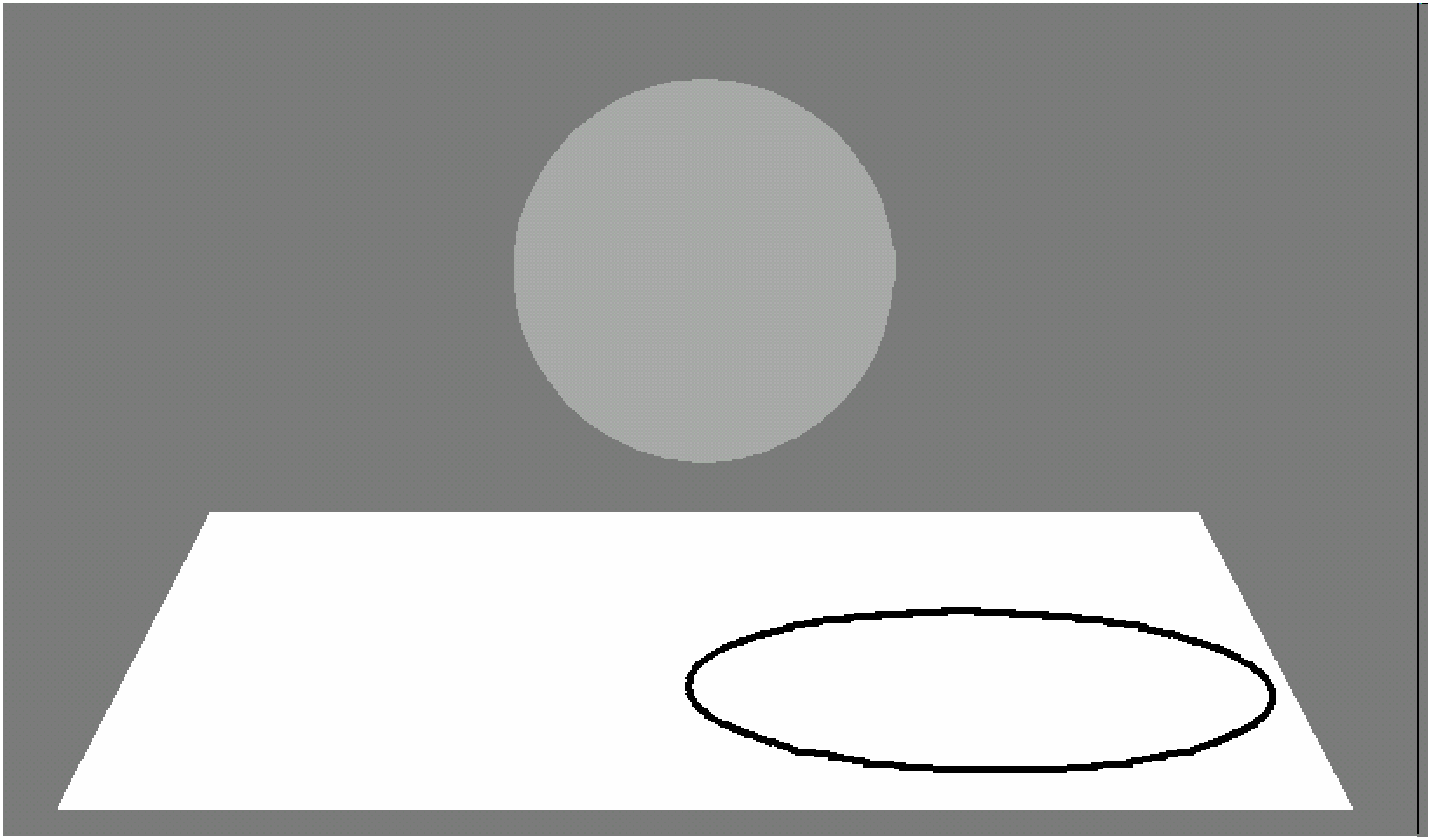


A new representation of image features

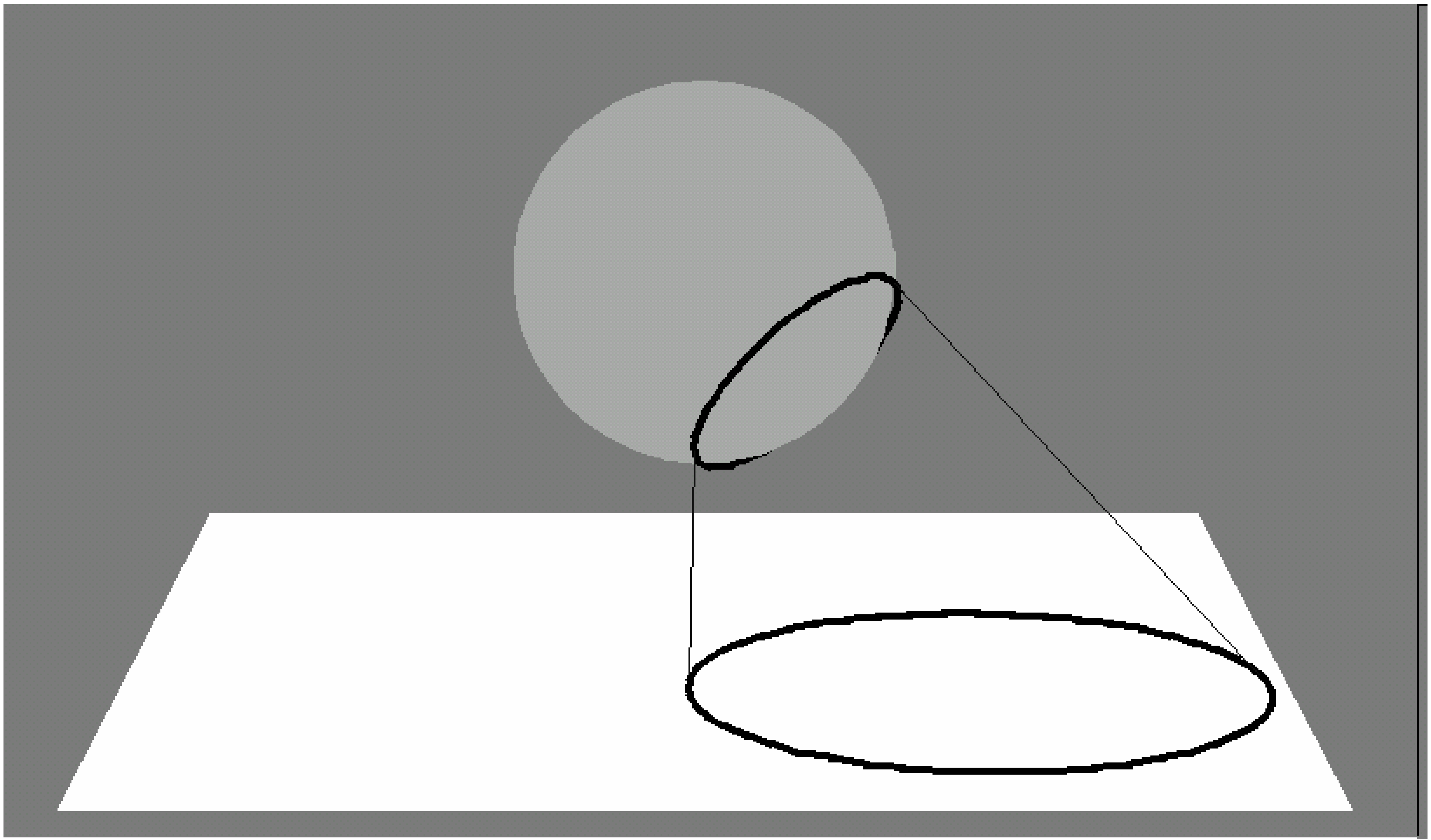
While the projective plane captures both points and lines, we do not have a space suitable for points and circles. We need a

CIRCLE SPACE!

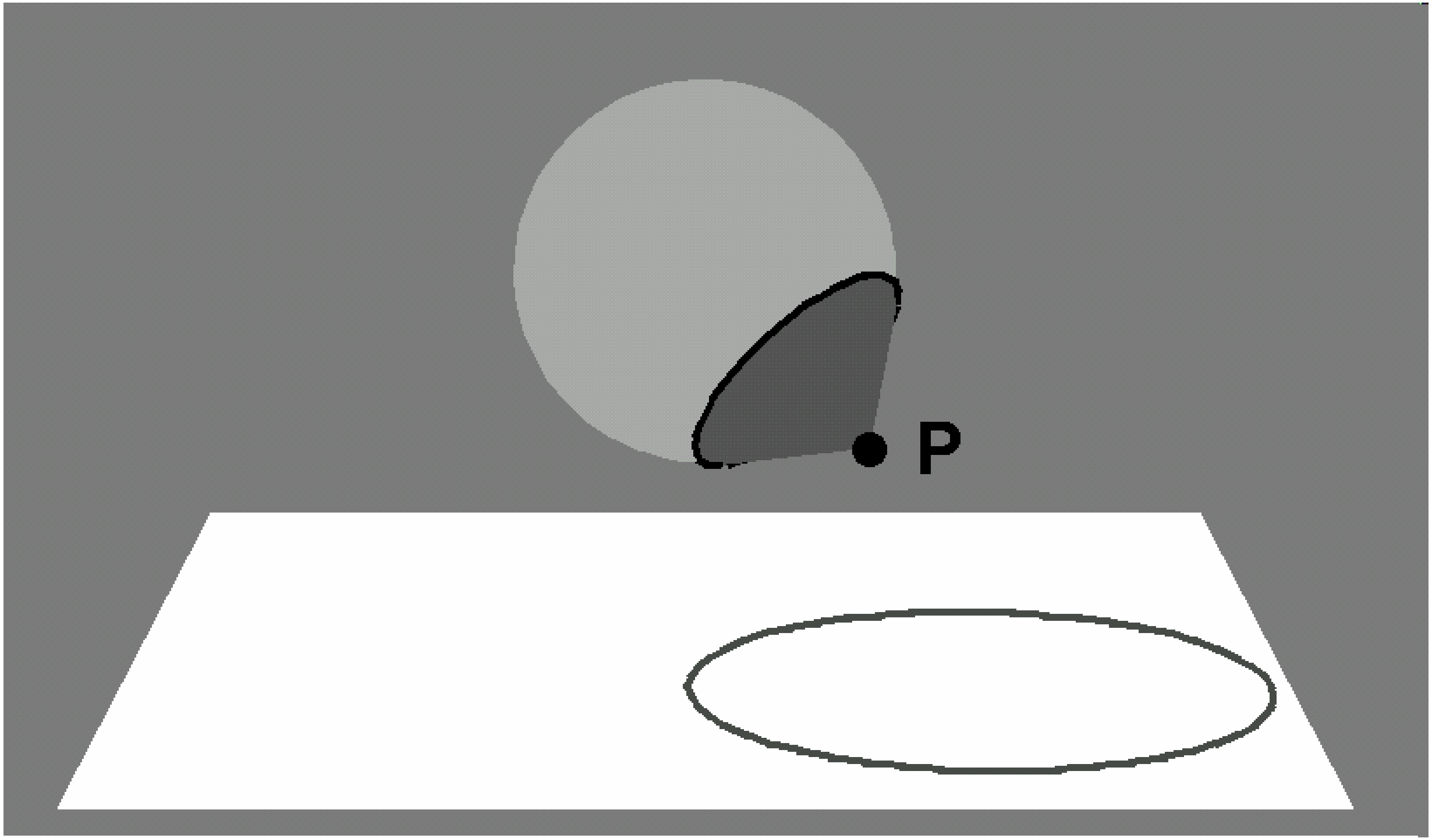
Lift a circle (line projection in parabolic omnicameras)



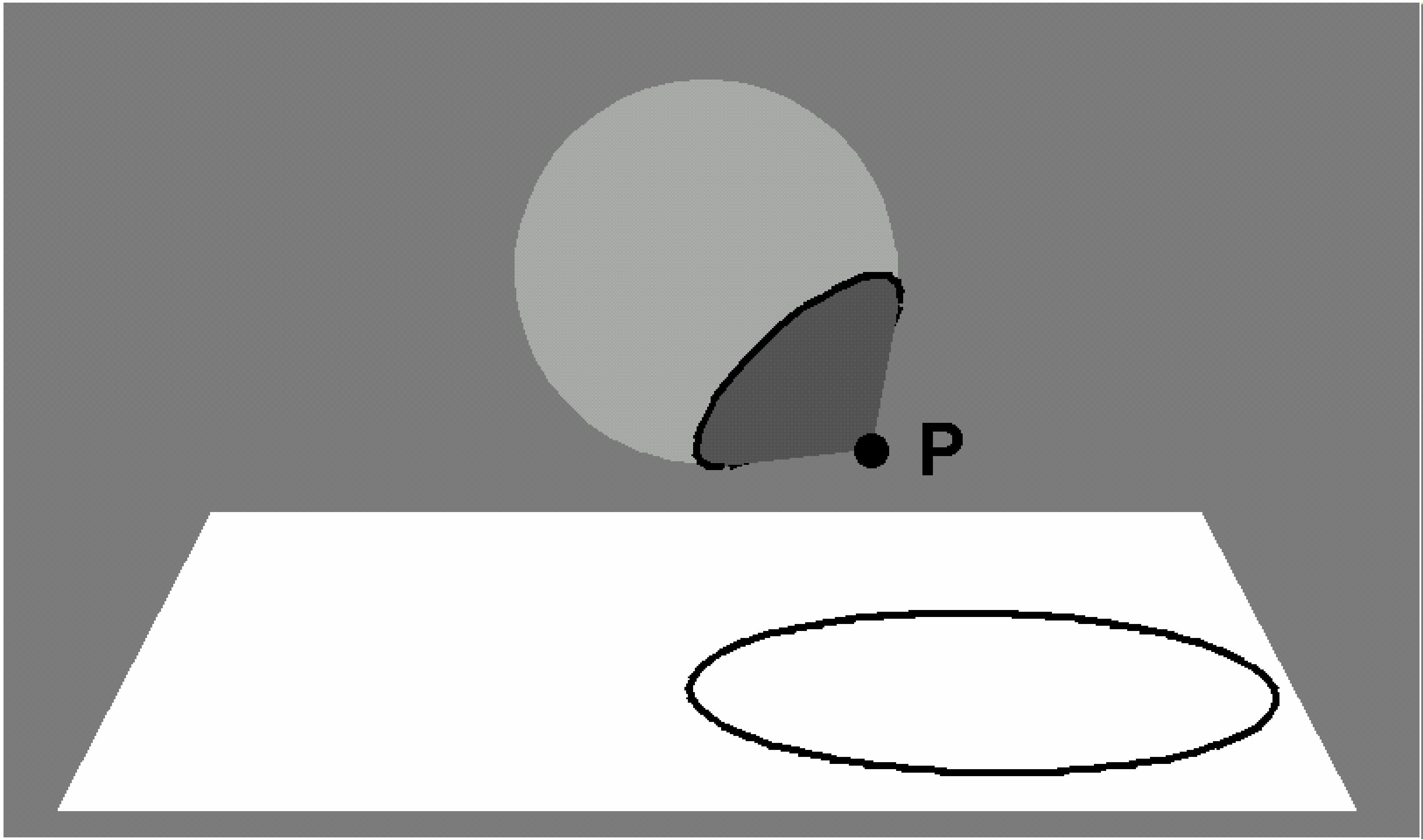
Take inverse stereographic image



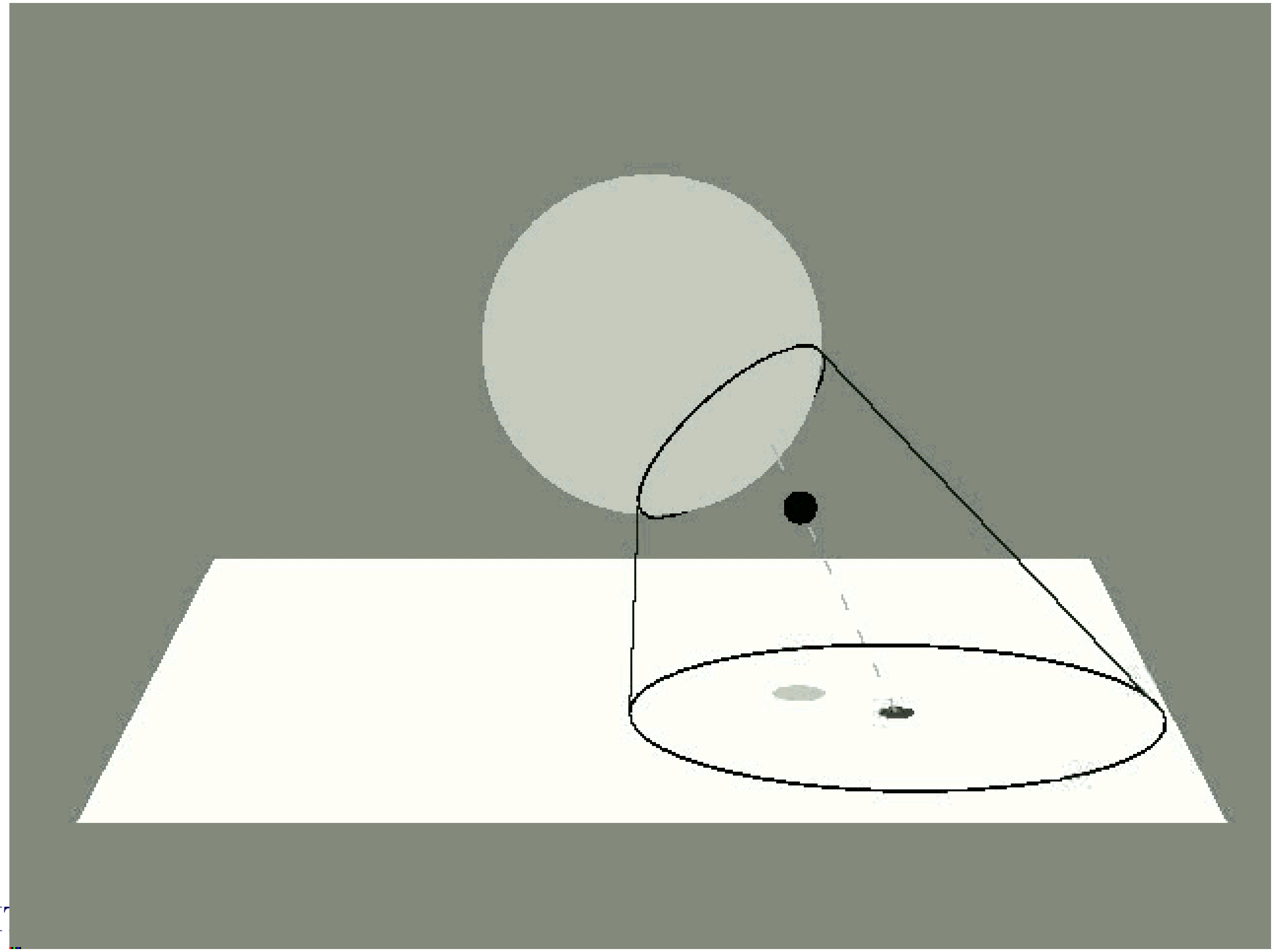
Construct cone tangent to locus

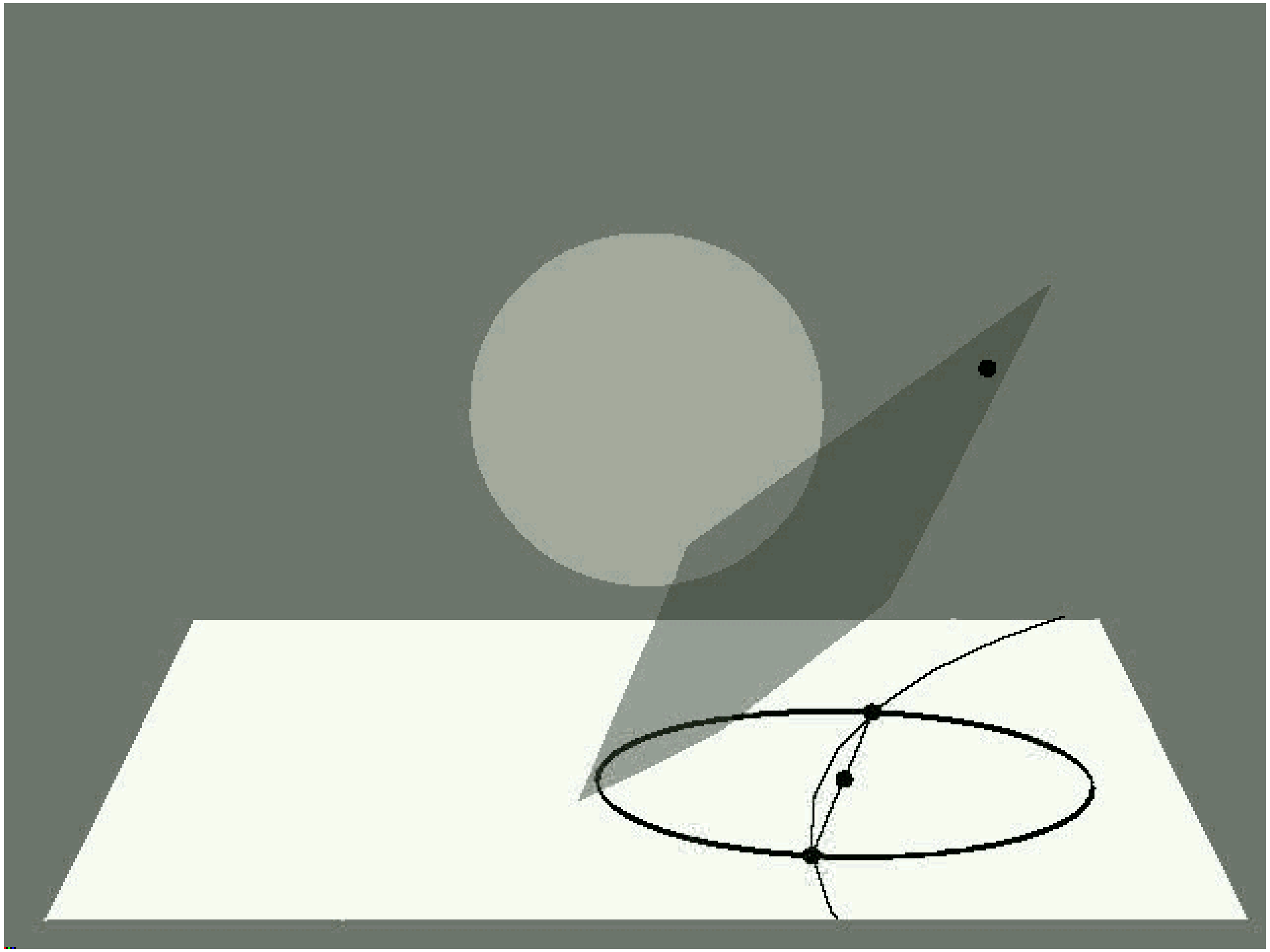


P is the representation of the circle



By varying the radius we model points, circles, and imaginary circles!





Not every circle is a line projection (it has to be projection of a great circle). All these feasible lines lie on a plane in circle space.

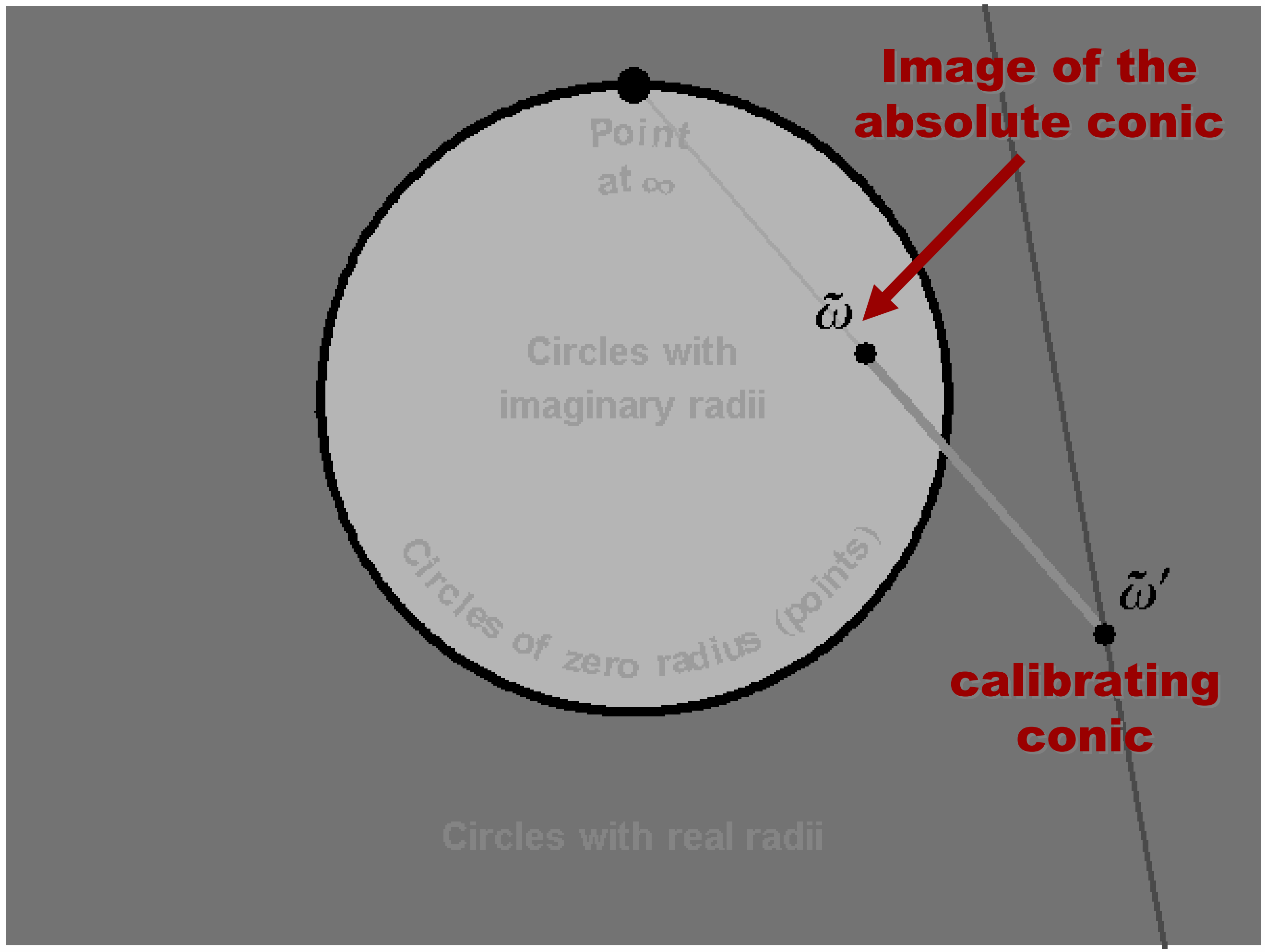


Image of the
absolute conic
Point
at ∞
$\tilde{\omega}$
Circles with
imaginary radii
Circles of zero radius (points)
$\tilde{\omega}'$
calibrating
conic
Circles with real radii

Transformations of circle space

Motivation: In the perspective case the group of transformations is the set of collineations, i.e. non-singular matrices in PGL(3)

Goal: find the natural transformation group of circle space.

A translation in the plane….

If the sphere has projective quadratic form

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

Then for A to preserve the sphere we must have

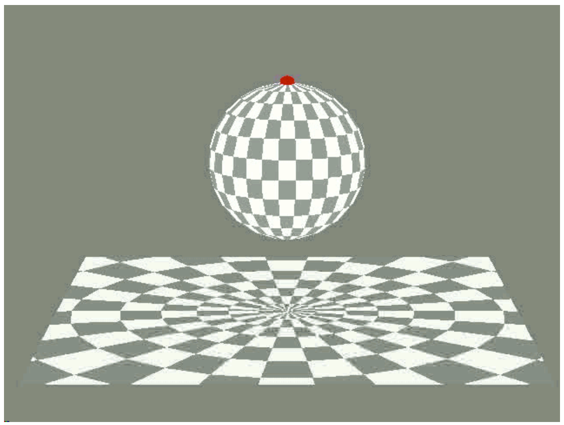
$$A^T Q A = Q$$

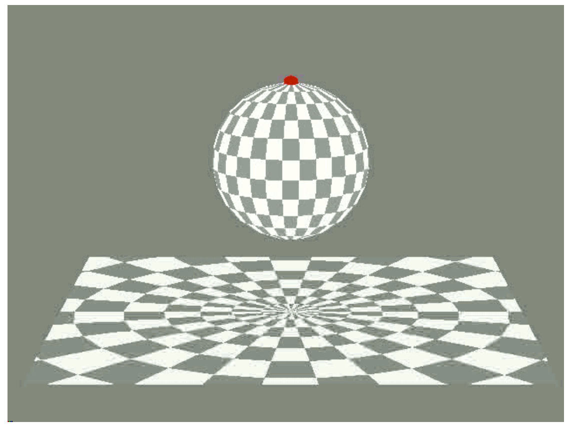
(Note similarity with $R^T R = I$)

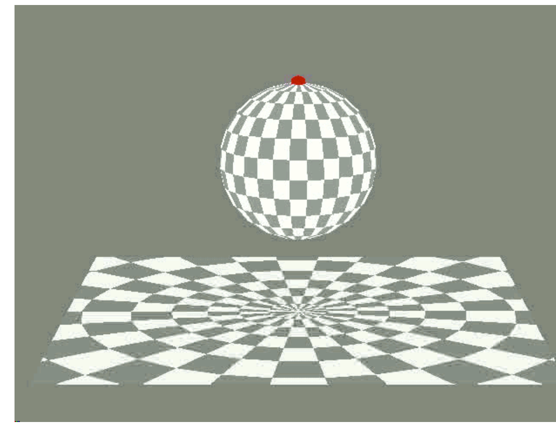
The Lorentz group O(3,1)

<u>**It is a six dimensional Lie group**</u>

Infinitessimal generators of the Lorentz group





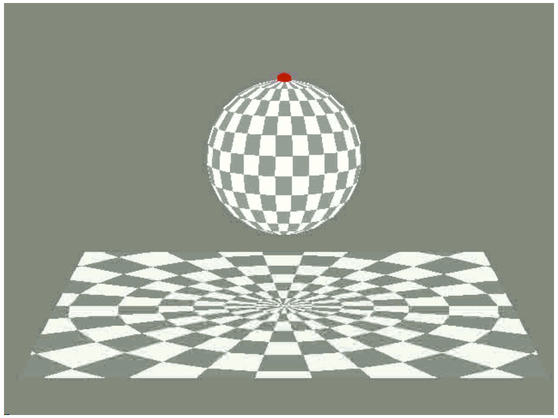


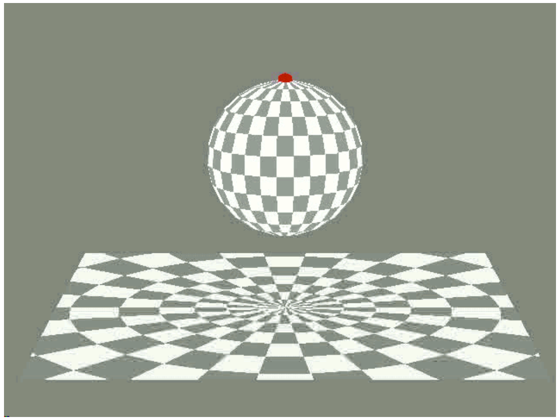
$$\exp \delta\theta \left(\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

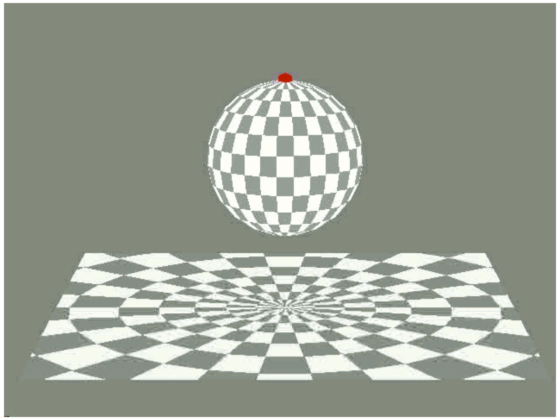
$$\exp \delta\theta \left(\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

$$\exp \delta\theta \left(\begin{array}{ccc|c} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right)$$

Infinitessimal generators of the Lorentz group





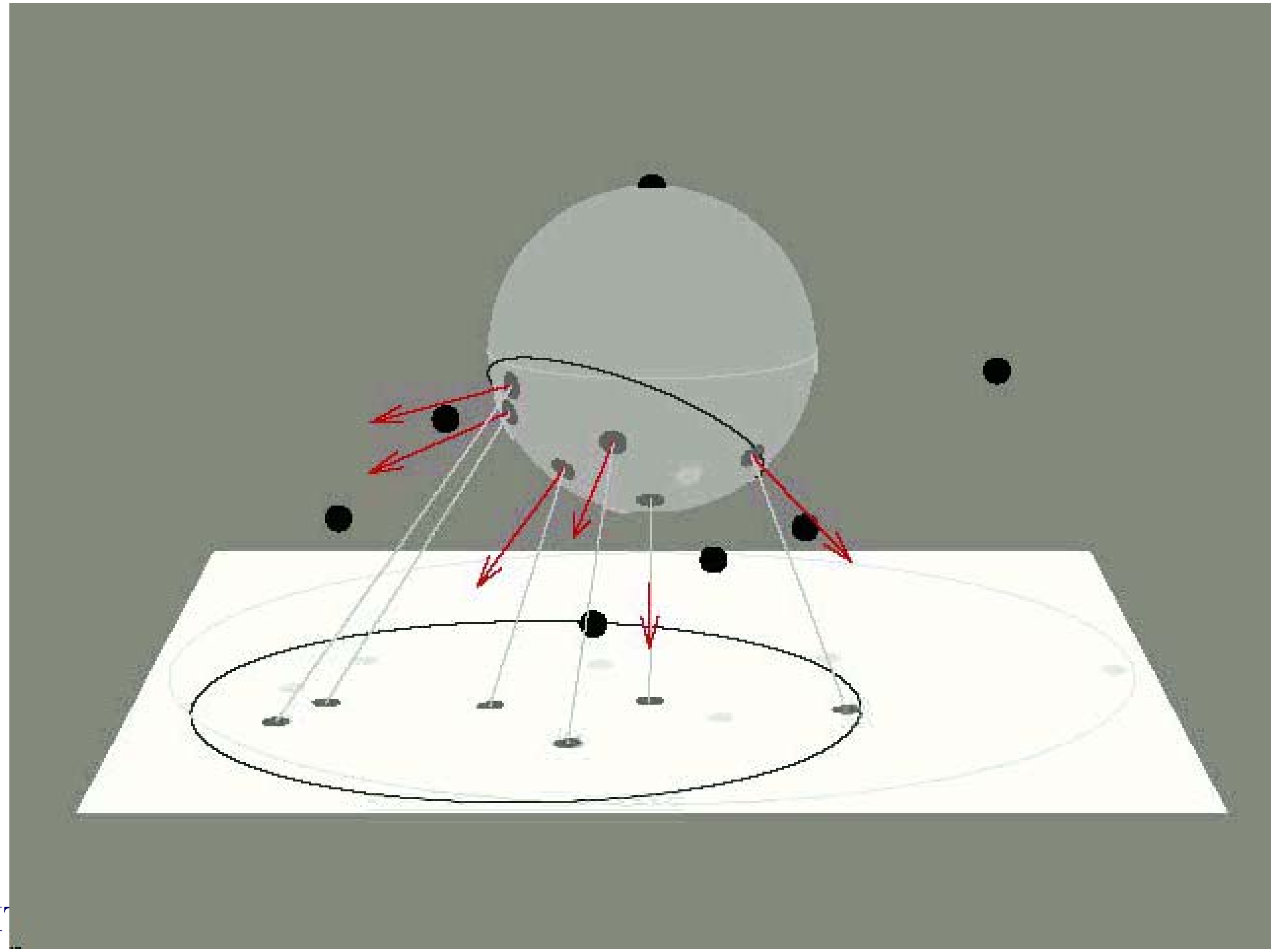


$$\exp \delta\theta \begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

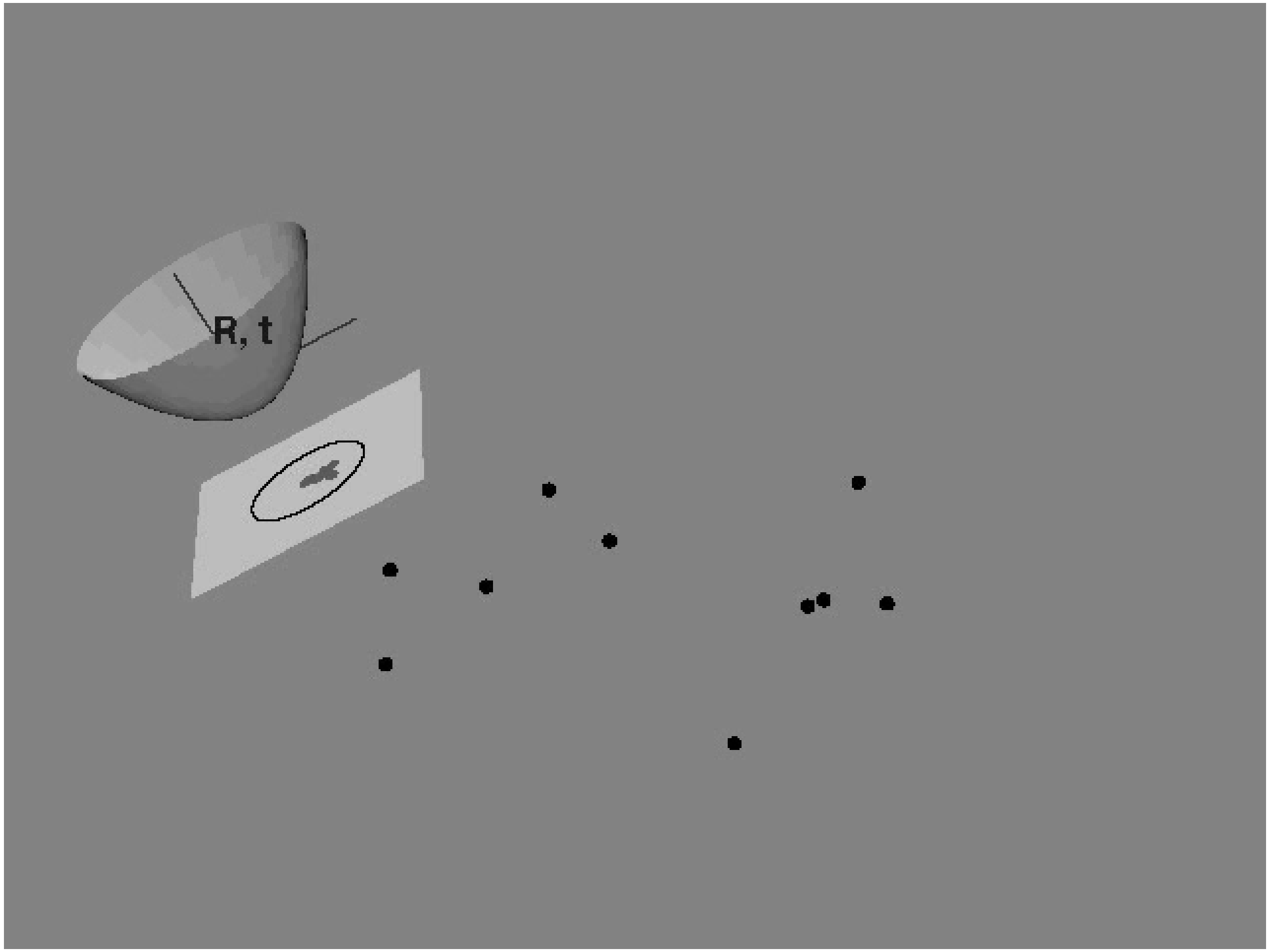
$$\exp \delta\theta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$\exp \delta\theta \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

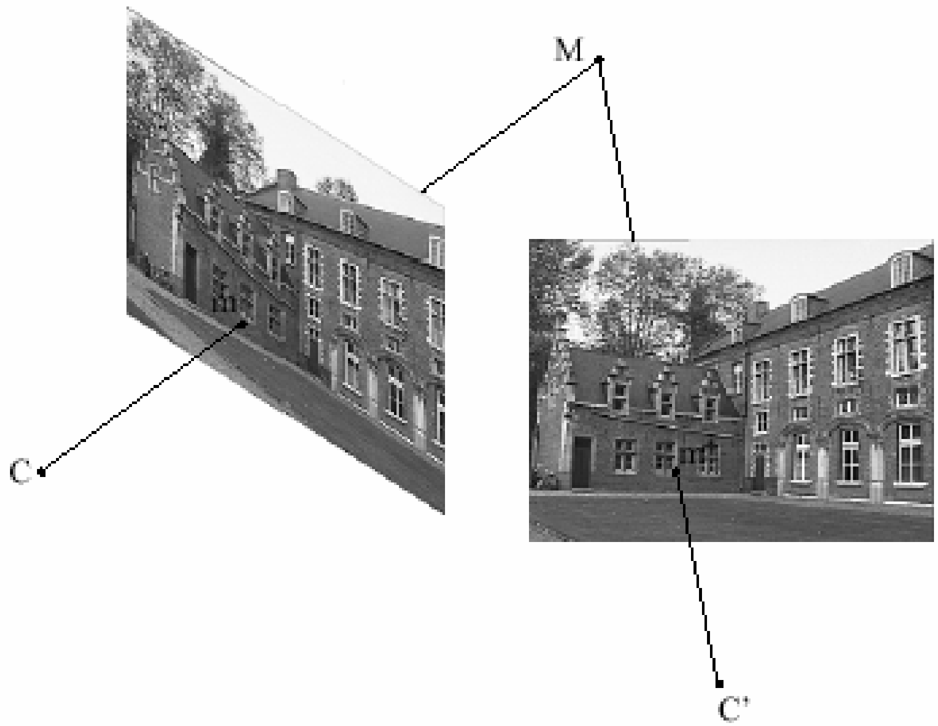
•We need a linear transformation from uncalibrated pixels to calibrated rays.
Such a linear transformation exists and its kernel contains the parameters of this mapping.



Motion estimation



Perspective image pair: Epipolar constraint describes coplanarity between two projection centers and image point

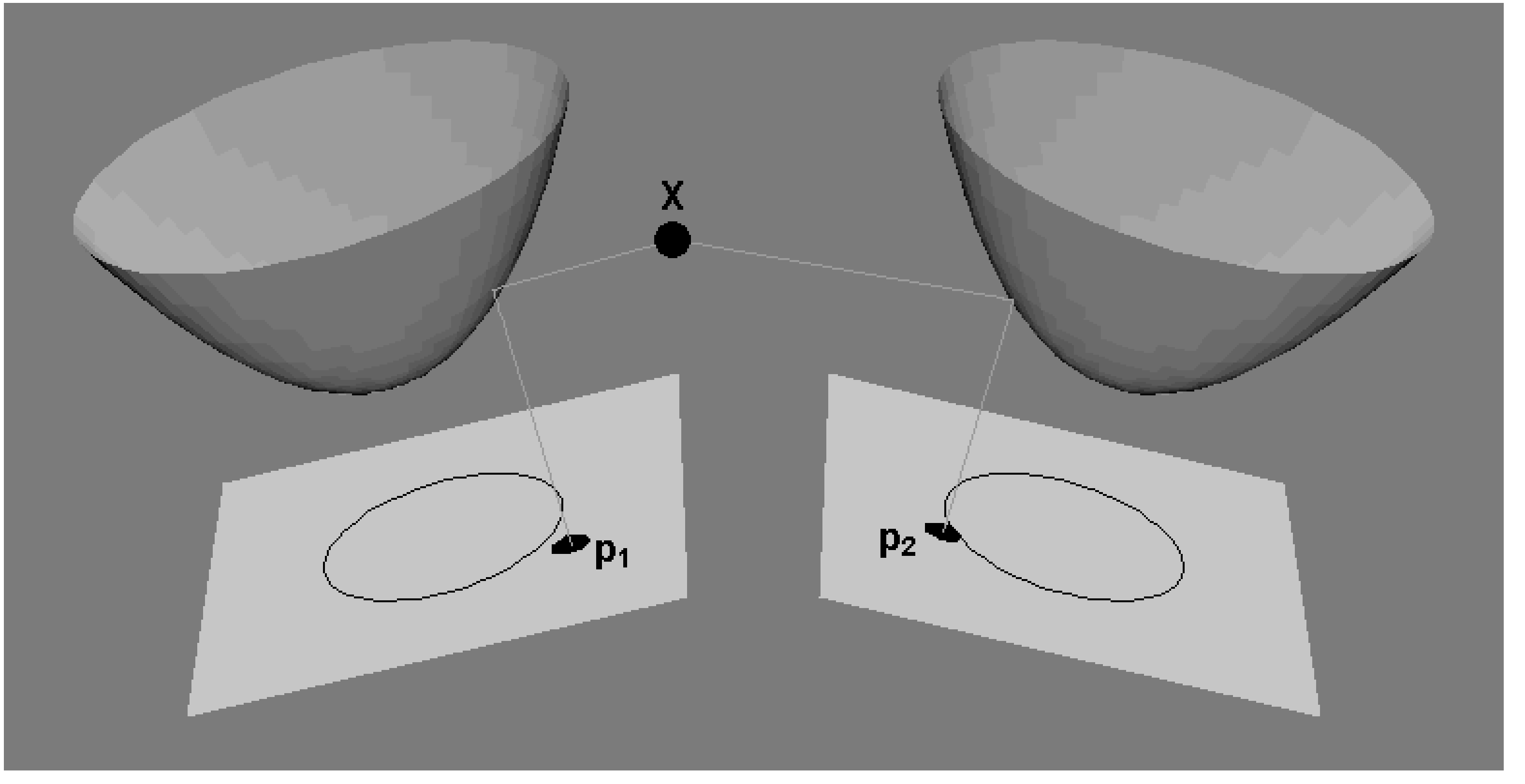


Two view perspective: the essential matrix

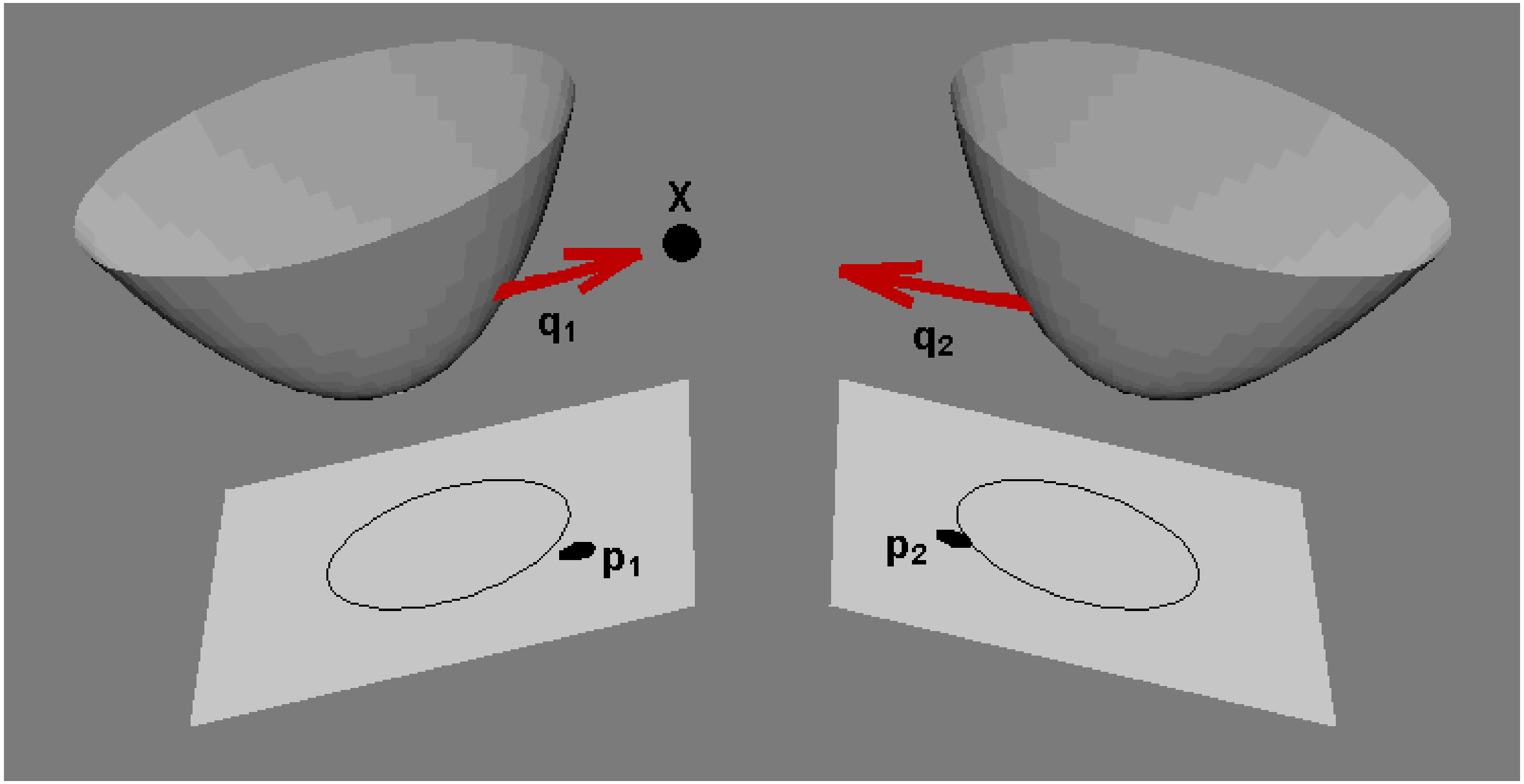
Recall that two images p_1, p_2 of the same space point X satisfy the bilinear constraint

$$p_1^T E p_2 = 0$$

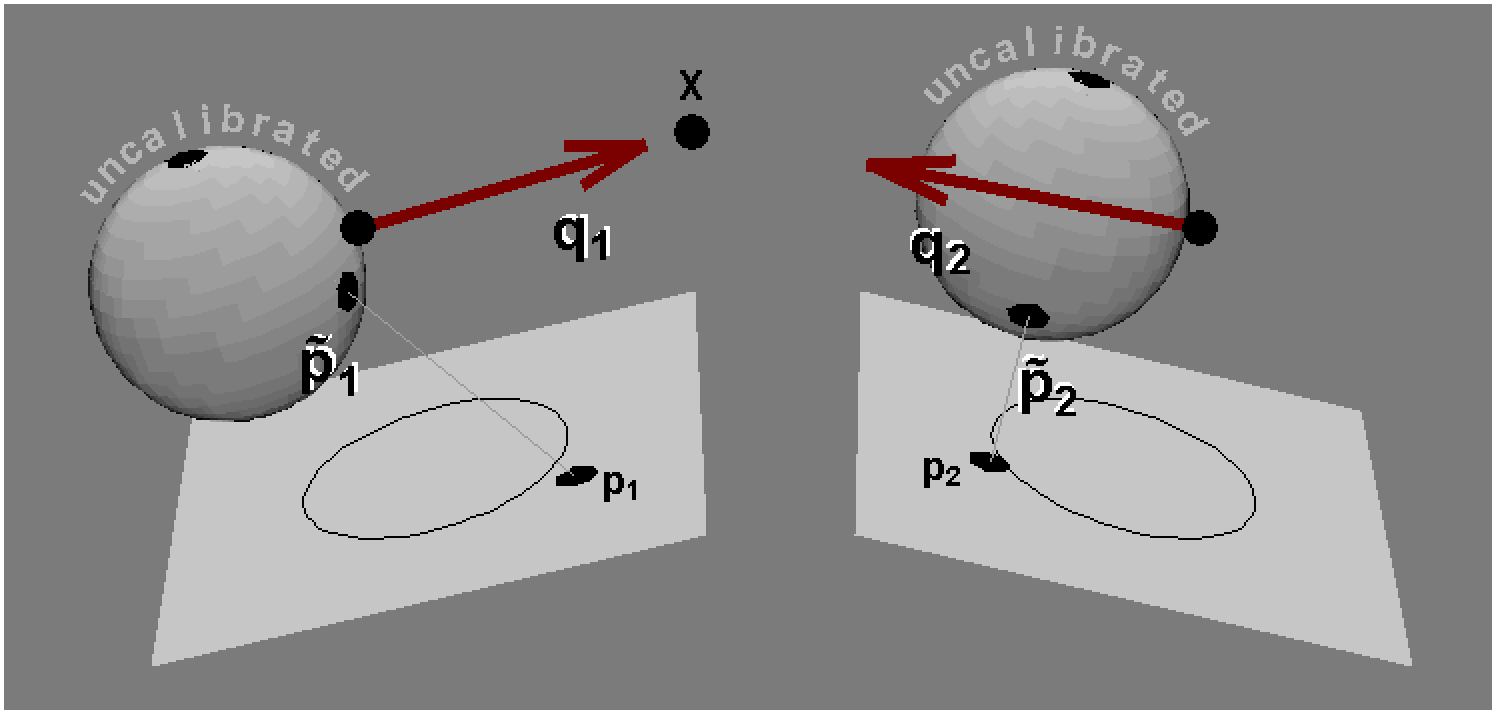
where E is a 3×3 rank 2 matrix independent of X,



Assume p_1 and p_2 are the catadioptric projections of X



$$q_2^T E q_1 = 0$$



However there exist Lorentz group elements K_1 & K_2 such that

$$q_1 = K_1 \tilde{p}_1 \quad \textbf{and} \quad q_2 = K_2 \tilde{p}_2$$

Catadioptric fundamental matrix

$$\breve{p}_2^T \underbrace{K_2^T E K_1}_{F} \breve{p}_1 = 0$$

i.e. the <u>lifted</u> image points satisfy a *bilinear epipolar constraint*!!!

<u>F is the 4×4 catadioptric fundamental matrix</u>

The kernel of F is the kernel of K.

Reconstruction algorithm much simpler than in perspective !

1. When intrinsics constant recover camera parameters with kernel computation and intersect
$$\ker F \cap \ker F^T = \{\lambda \tilde{\omega}\}\,.$$
2. Recover rotation and translation
3. Reconstruct environment or produce novel views.

A characterization of parabolic fundamental matrices

Recall that a 3×3 matrix *E* is an essential matrix if and only if

$$\mathbf{E} = \mathbf{U}\begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}\mathbf{V}^{\mathbf{T}}$$

for some *U*, *V* in *SO*(3)

Claim: A 4×4 matrix *F* is a parabolic fundamental matrix if and only if

$$\mathbf{E} = \mathbf{U}\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 0 & \\ & & & 0 \end{pmatrix}\mathbf{V}^{\mathbf{T}}$$

for some *U*, *V* in *SO*(3,1)

What are the properties of F?

- Rank 2 (4 constraints)
- Two Lorentzian Singular Values are equal (? Constraints)

How many degrees of freedom are in F?

- 5 for motion
- 3+3 for intrinsics left and right
- =11

However algebraically:

- 4 constraints already for rank of a 4x4 to be 2
- ? Constraints for Lorentzian Singular Values to be equal

The set of all *gx* in X for any g in G is called the **orbit** of x. If the group possesses an orbit, that means for any a,b in X, ga=b for a g in G, then the group action is called **transitive**. For example, there is always a rotation mapping one point on the sphere to another.

If a subgroup H of G fixes a point x in X then H is called the **isotropy group**. A typical example of an isotropy group is the subgroup SO(2) of SO(3) acting on the north-pole of a sphere.

A space X with a transitive Lie group action G is called **homogeneous space**.

If the isotropy group is H, it is denoted with **G/H**.

Group theoretic analysis of bilinear constraints

- Let's examine the LSVD characterization of parabolic fundamental matrices:

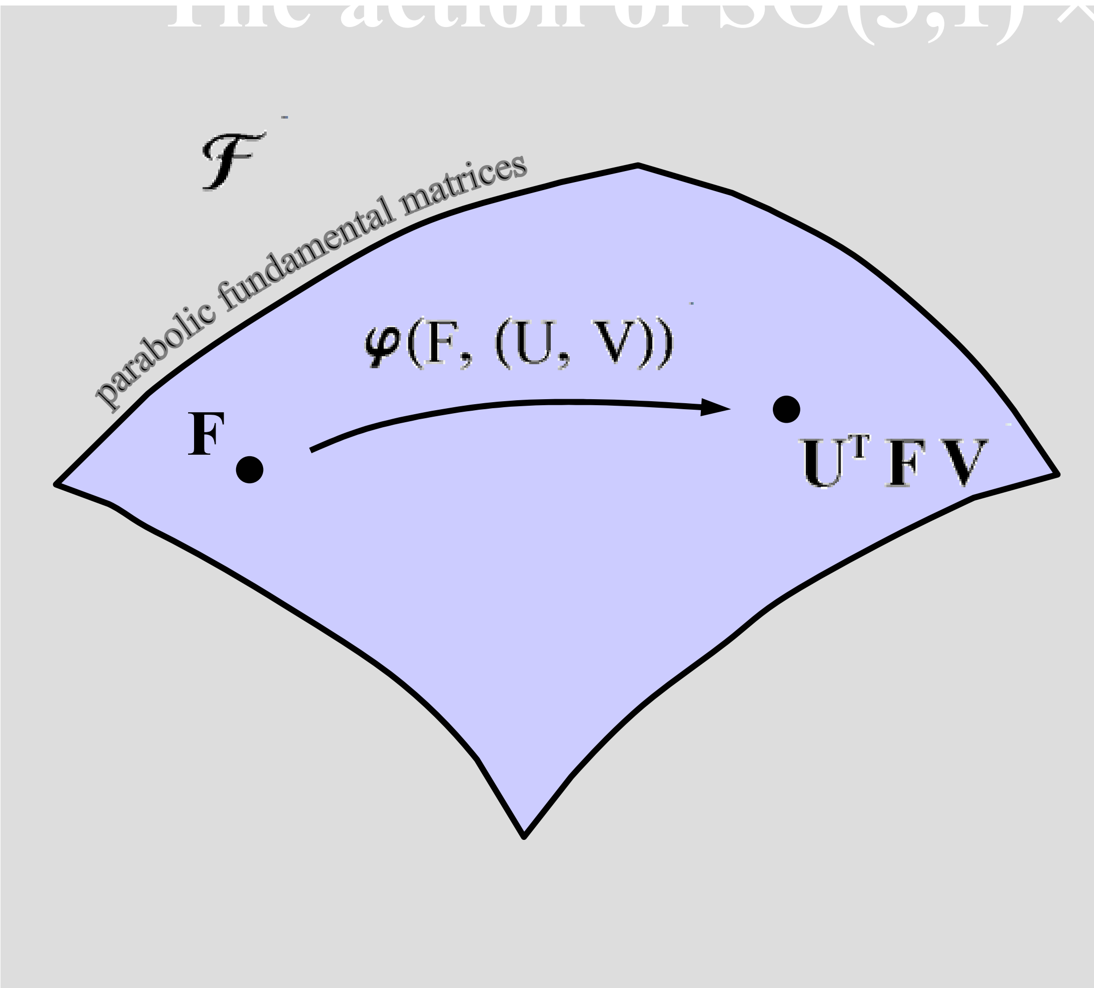
$$F = U \operatorname{diag}(1, 1, 0, 0) V^T \qquad U, V \in SO(3,1)$$

implies fundamental matrices are closed under left or right multiplication by Lorentz transformations, i.e.

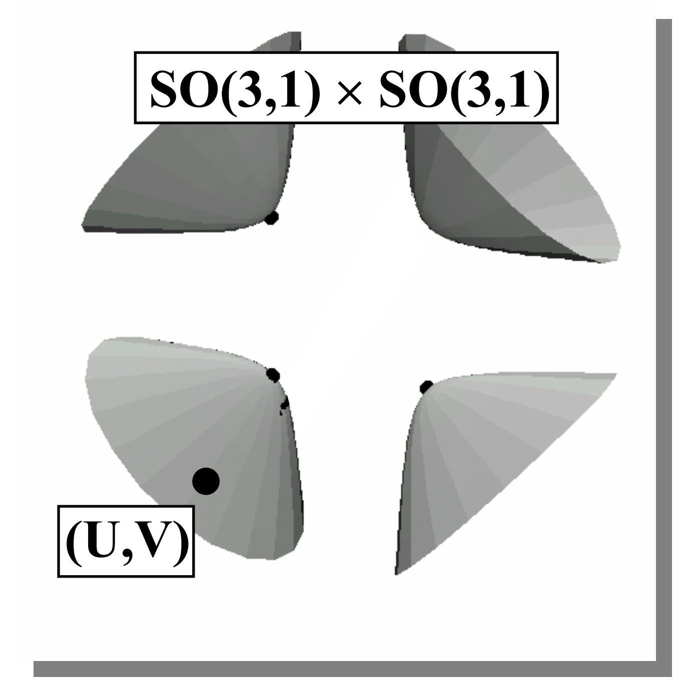
$$U'^T F V' \qquad U', V' \in SO(3,1)$$

is also a parabolic fundamental matrix.

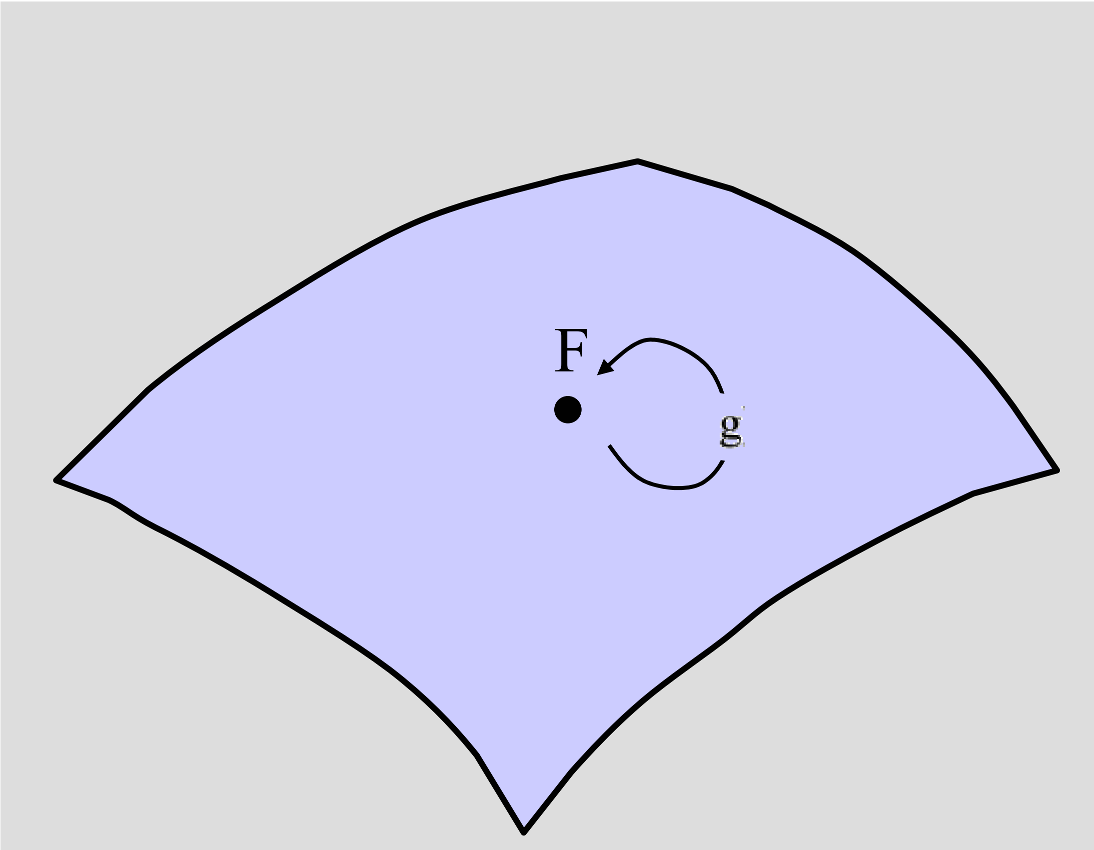
Note: the same reasoning applies to essential matrices.



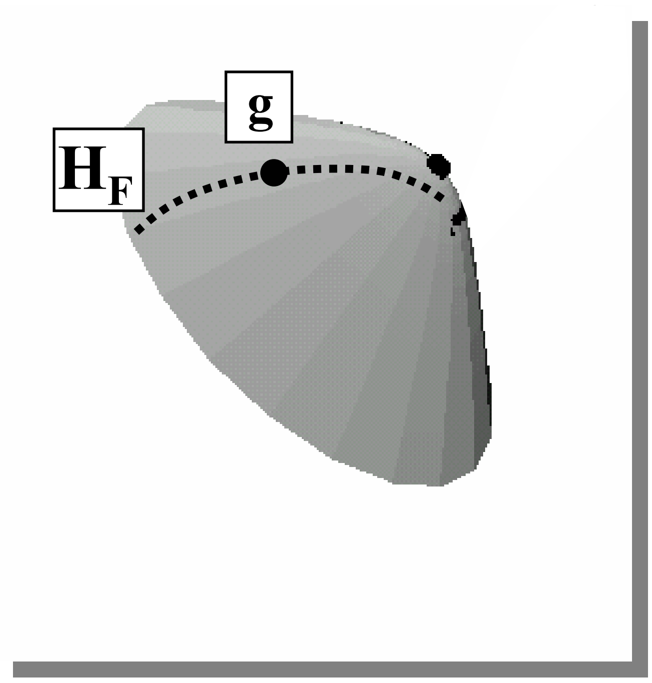
Thus SO(3,1) × SO(3,1) *acts* upon the set of fundamental matrices



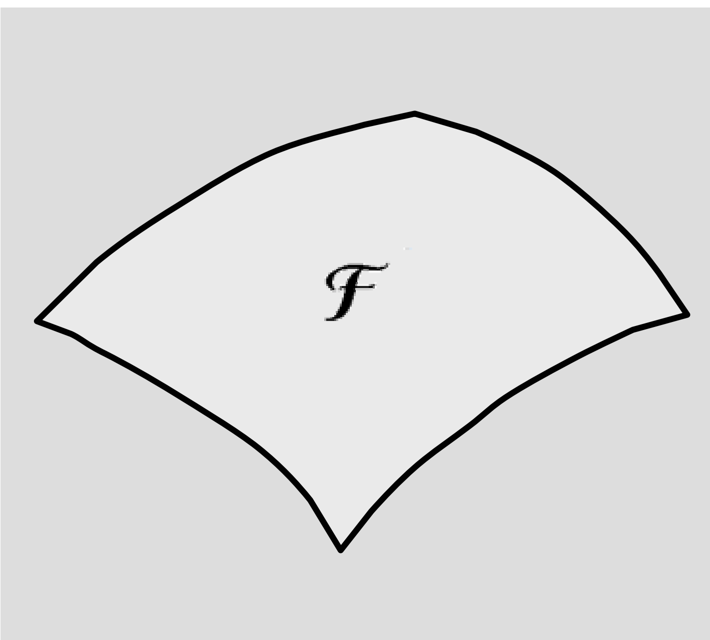
The isotropy group H_F: all g's leaving F invariant

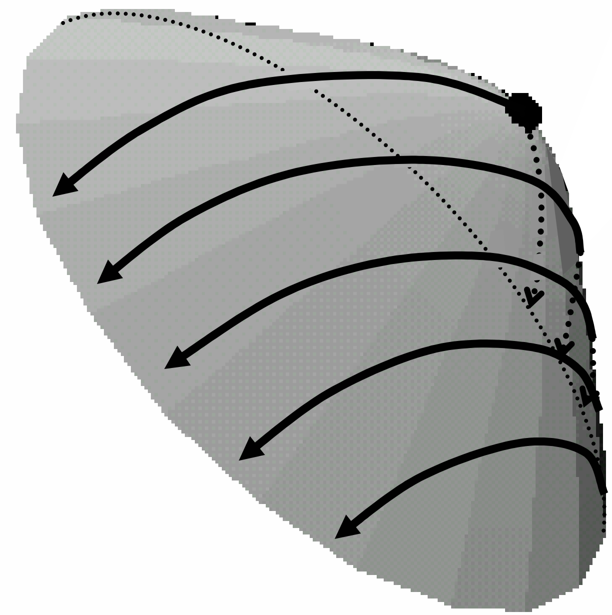


$$H_F = \{g : \varphi(g, F) = F\}$$



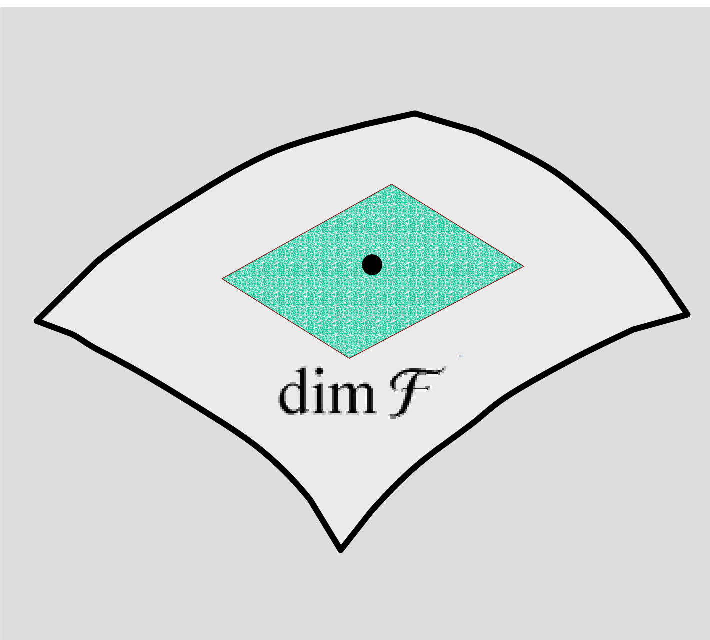
The set of fundamental matrices form a quotient space

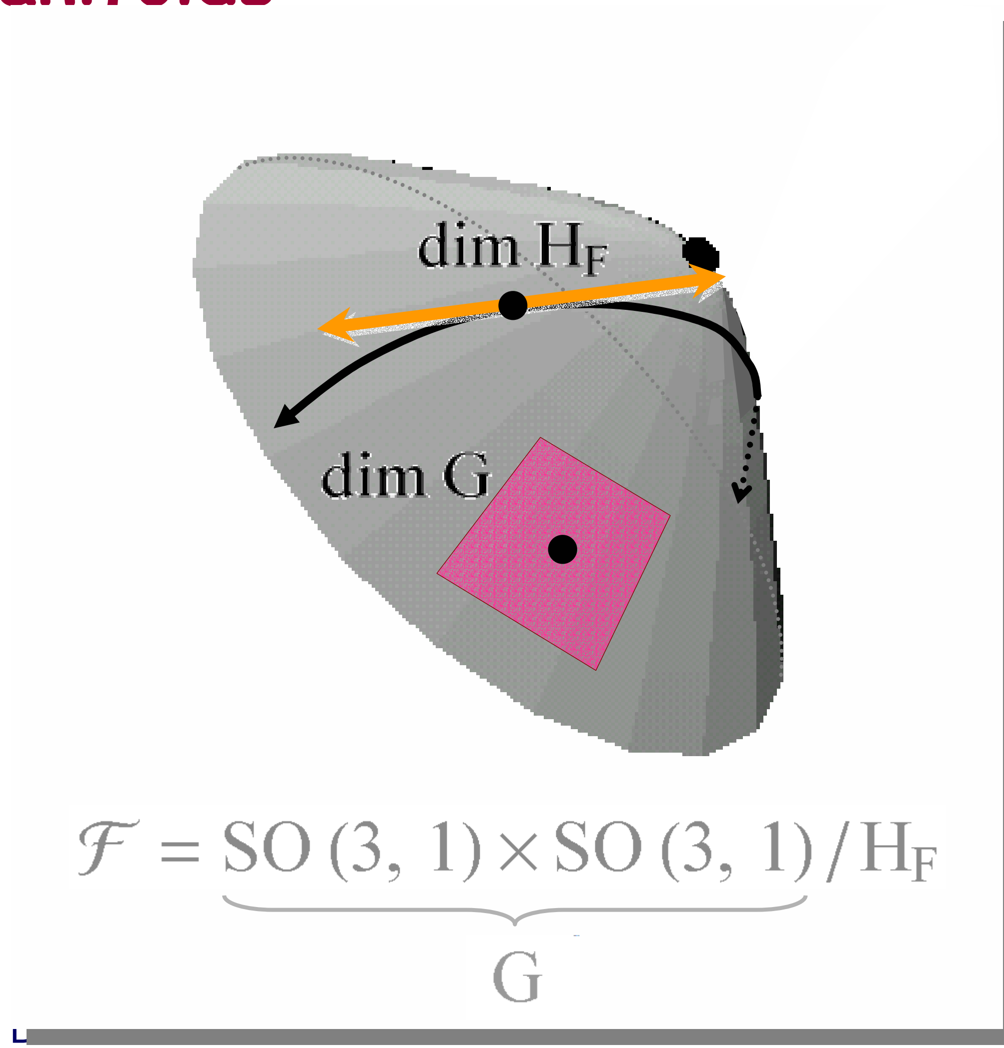




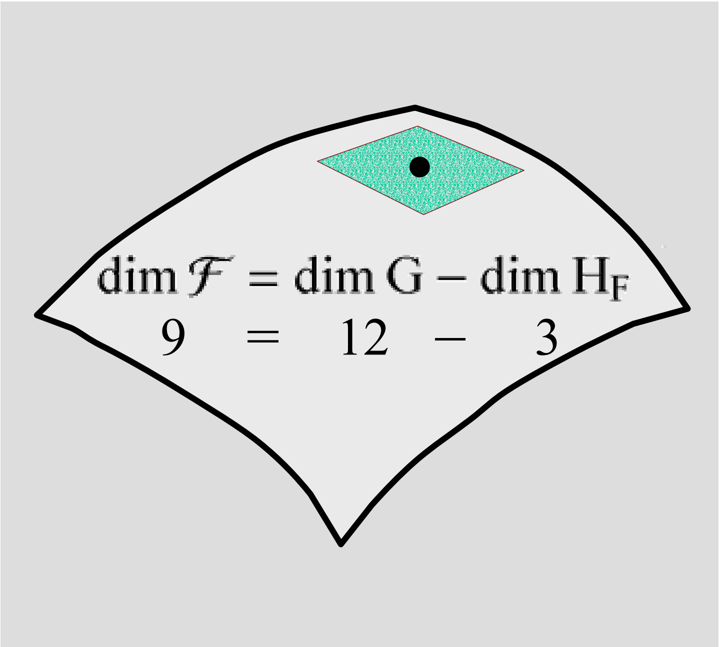
$$\mathcal{F} = SO(3,1) \times SO(3,1) / H_F$$

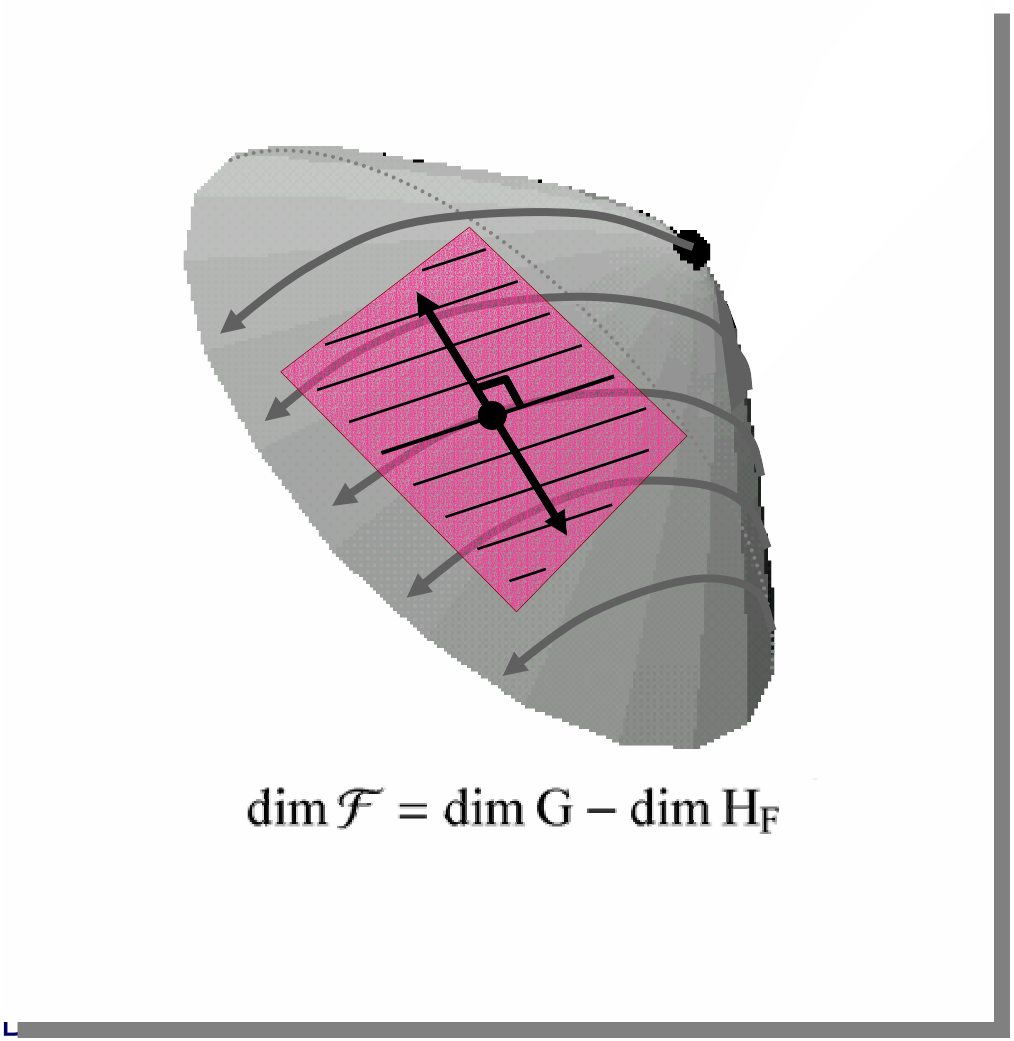
Quotient of Lie groups are automatically manifolds



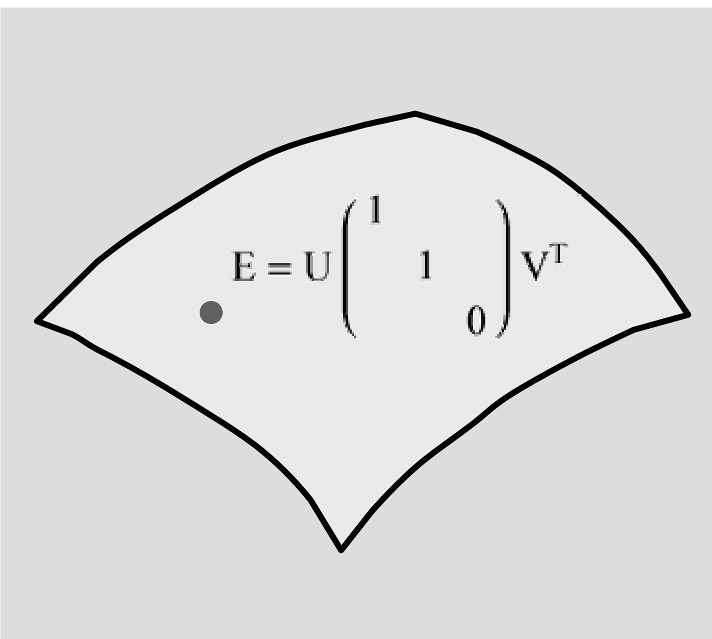


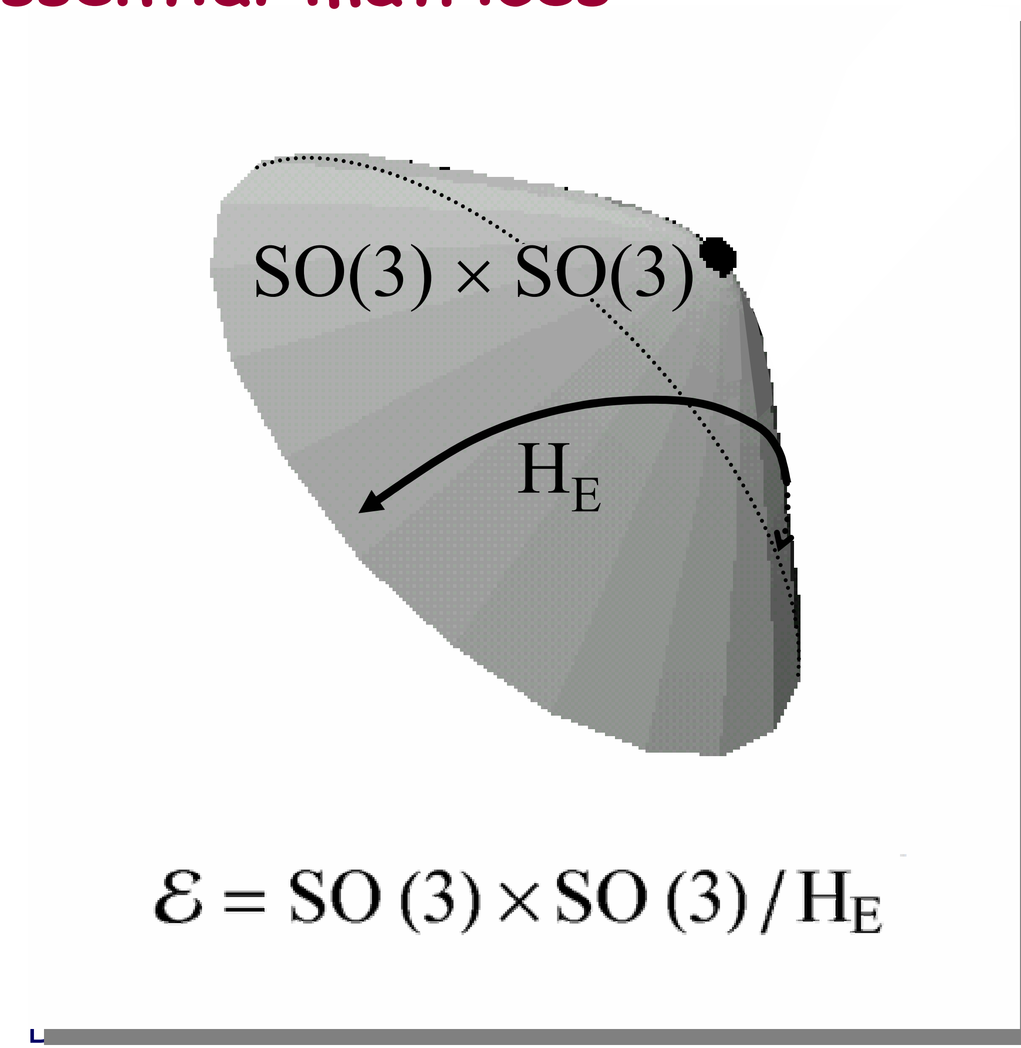
Quotient of Lie groups are automatically manifolds





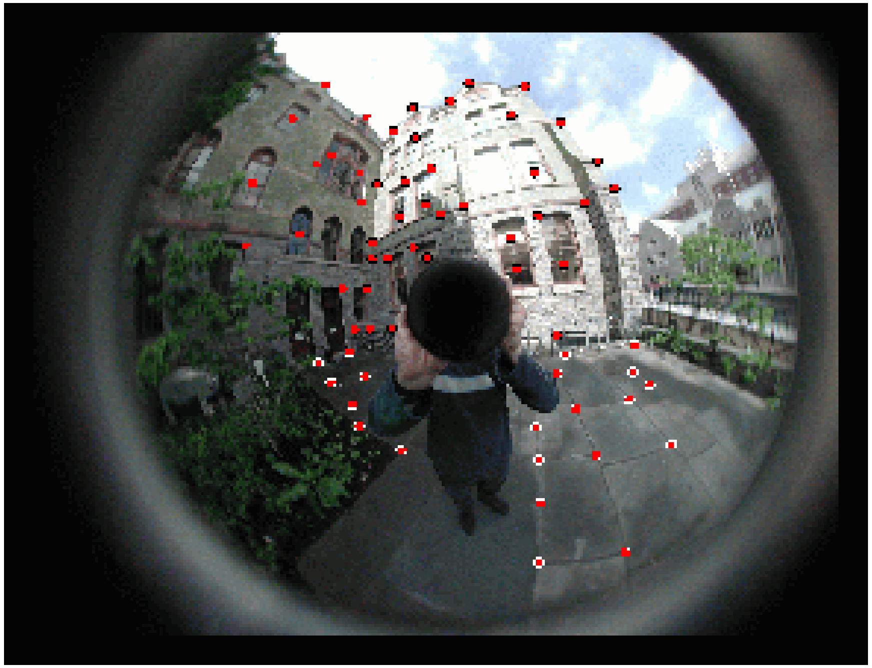
All of these results also apply to essential matrices

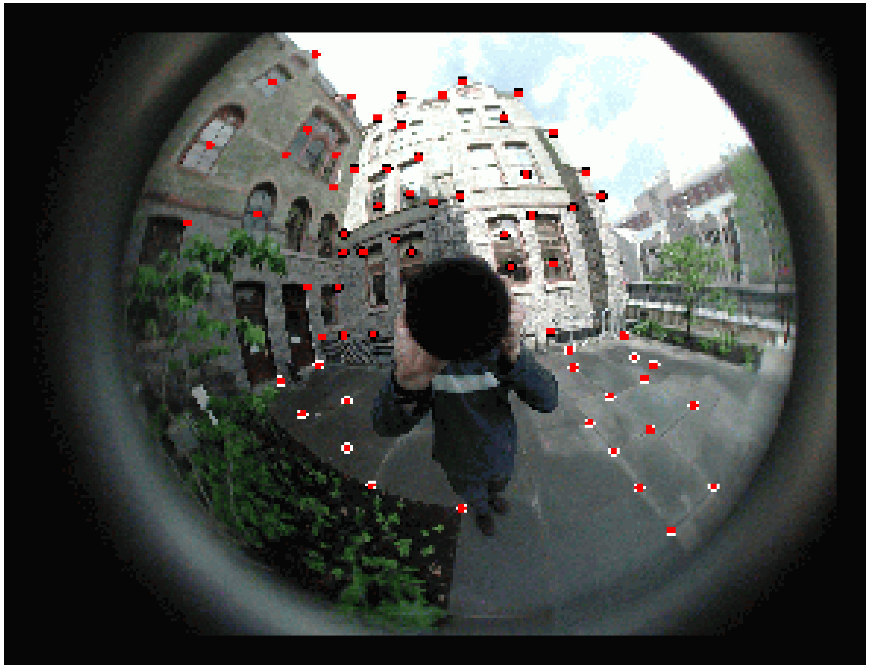




Two view example

Given these two views with corresponding points estimate the parabolic fundamental matrix





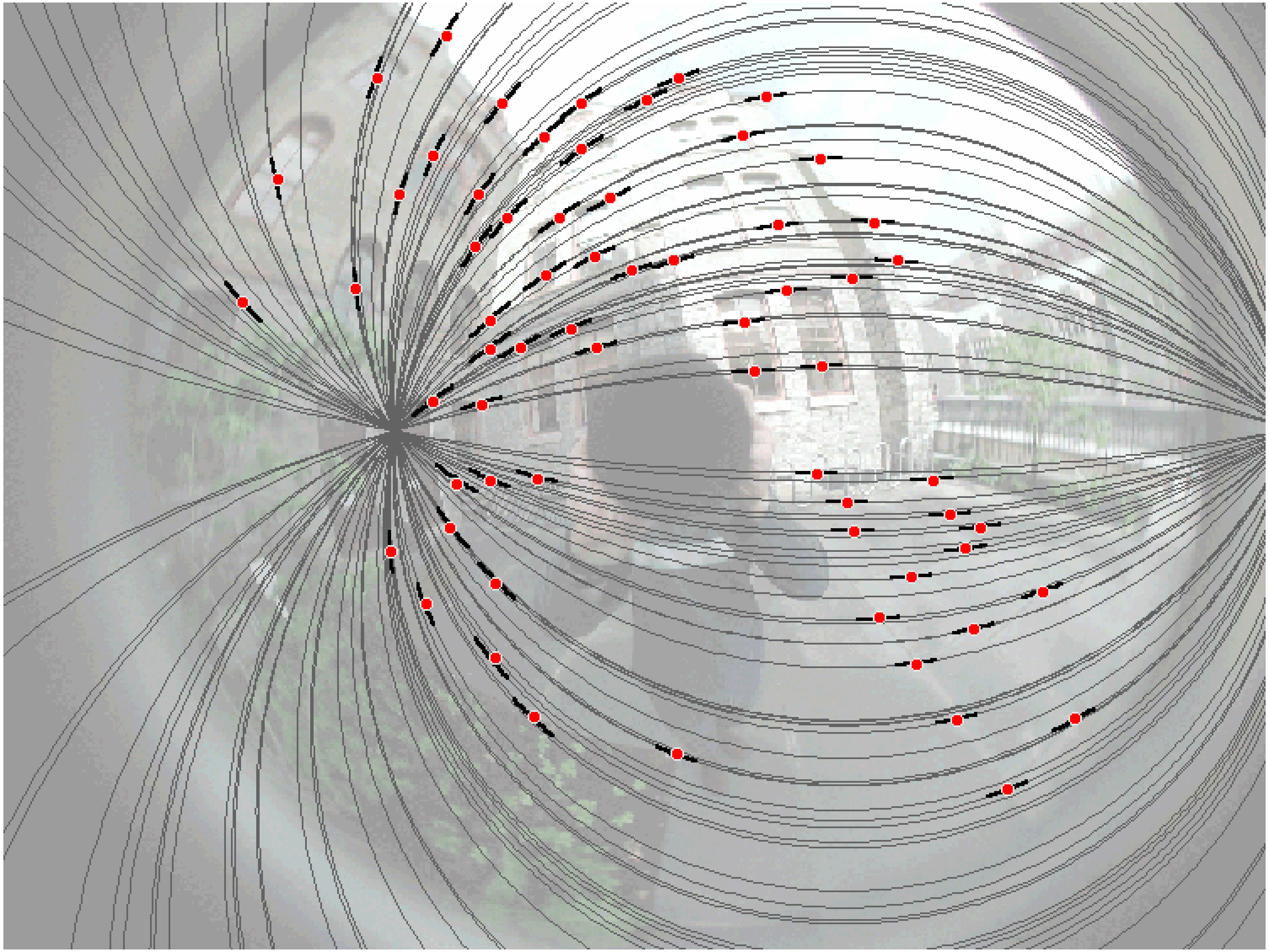
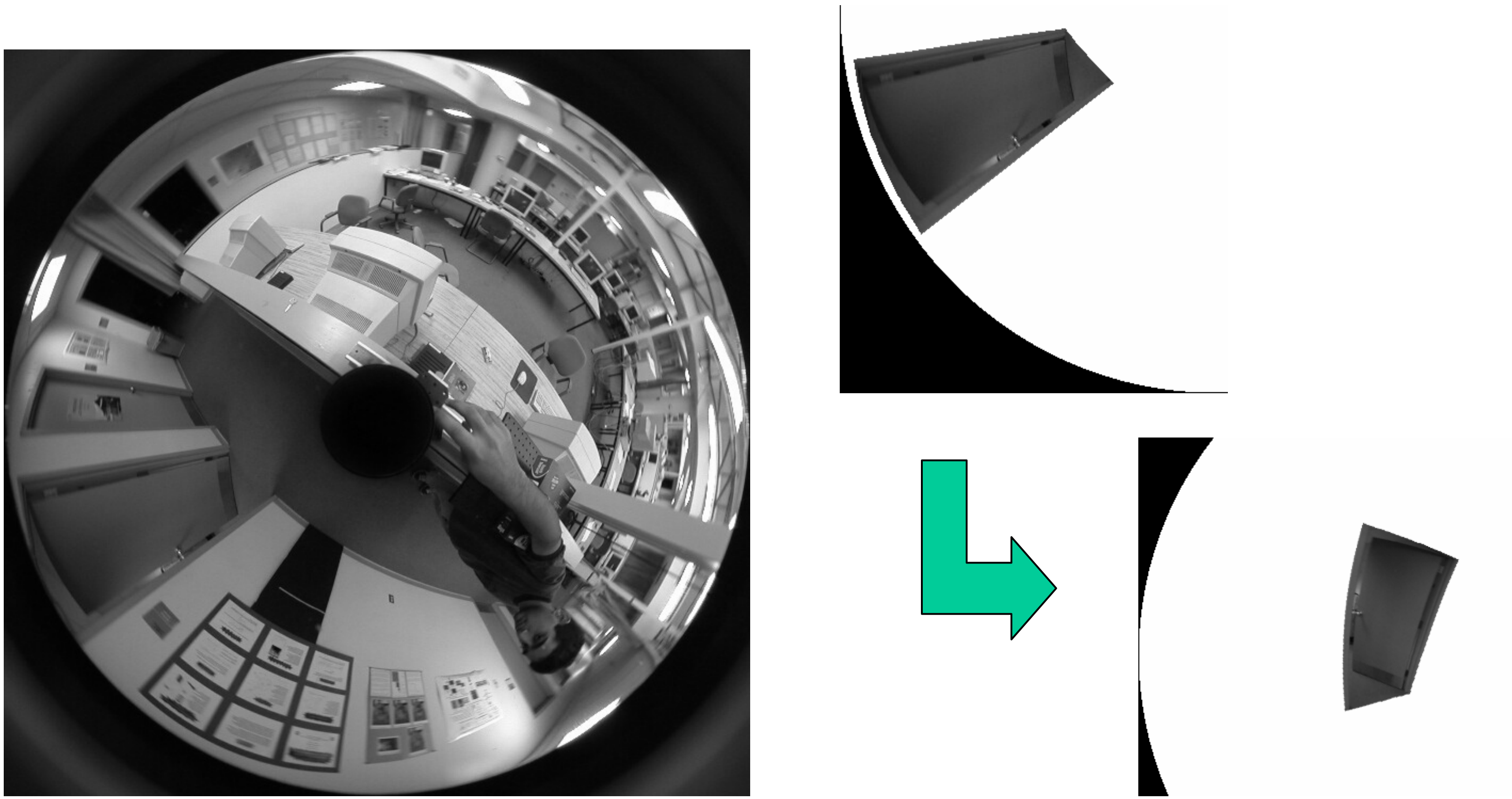


Image processing in perspective images

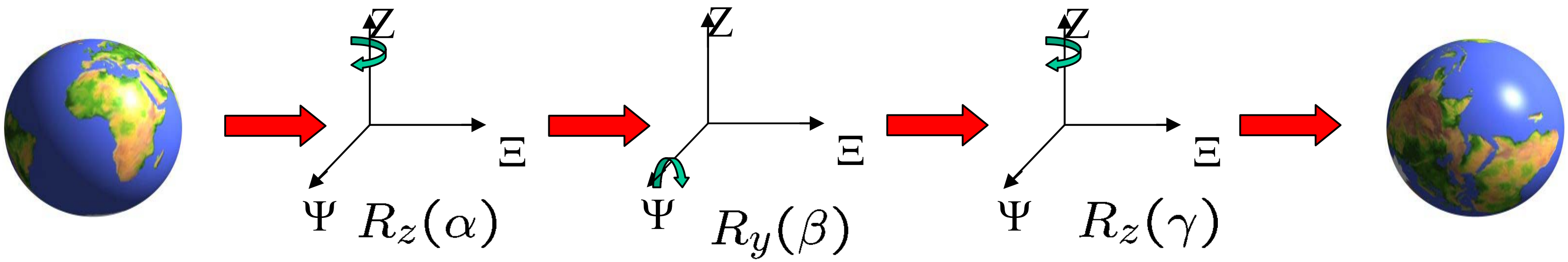
- Images obtained through perspective projection undergo local mappings:
  - Translations
  - Similitude
  - Affine
  - Projective (Collineations).

Template deformation in an omni-image is not covered by any of these mappings



Original image → calibration → Spherical image

Definitions



The rotation of a function $f(\eta)$ by an element $g \in SO(3)$ is defined with the operator Λ_g as $\Lambda_g f(\eta) = f(g^{-1}\eta)$

The integration of a function $f(\eta) \in L^2(S^2)$ is defined as

$$\int_{\eta \in S^2} f(\eta) d\eta = \int_0^{2\pi} \int_0^{\pi} f(\theta, \phi) sin(\theta) d\theta d\phi$$

How does convolution look like on the sphere?

- What is the "shift" in the convolution?
- It is a 3D-rotation acting as an operator:

$$(f * h)(\eta) = \int_{g \in SO(3)} f(gn)h(g^{-1}\eta)\, dg, \quad \eta \in S^2$$

North pole

$$\eta := (\cos(\varphi)\sin(\vartheta),\ \sin(\varphi)\sin(\vartheta),\ \cos(\vartheta)),$$

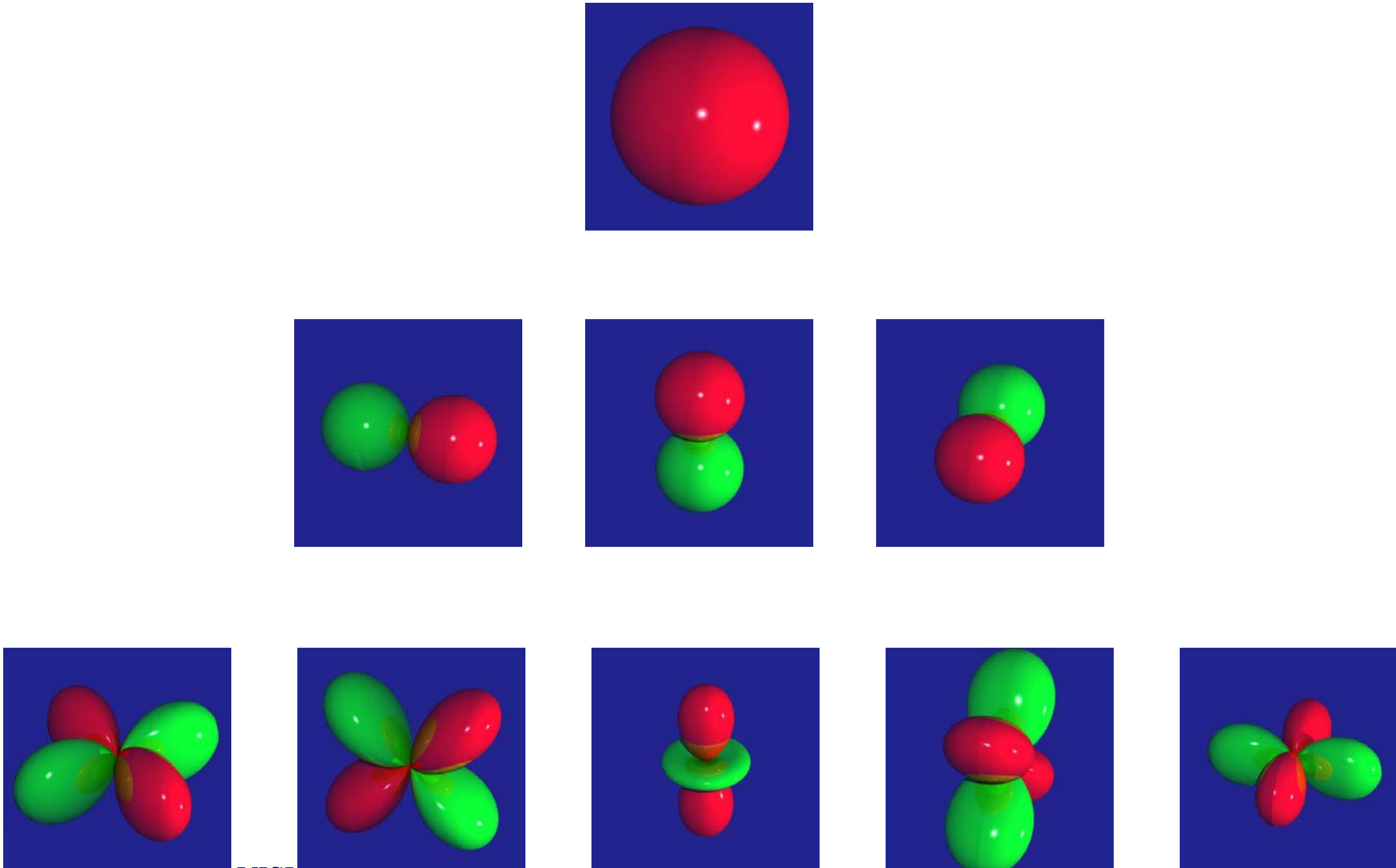
What about a Fourier transform on the sphere?

Look for a decomposition of functions on the sphere into subspaces invariant under SO(3): Eigenfunctions of the Laplace equation, the spherical harmonics

$$Y_m^l(\theta,\phi) = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_m^l(\cos\theta) e^{im\phi}$$

$$P_m^l(x) = \frac{(1-x^2)^{\frac{m}{2}}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l$$

Spherical Harmonics



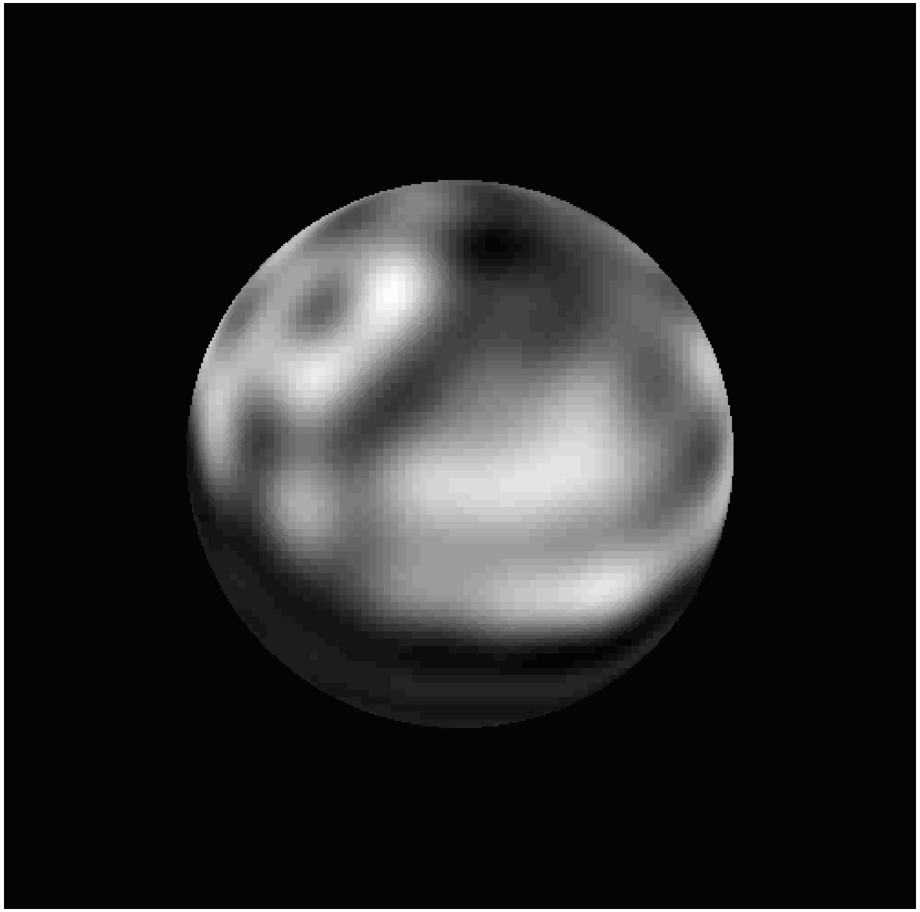
Spherical Harmonic Transform

$$f(\theta,\phi) = \sum_{l\in\mathbb{N}} \sum_{|m|\leq l} \hat{f}_{lm} Y_m^l(\theta,\phi)$$

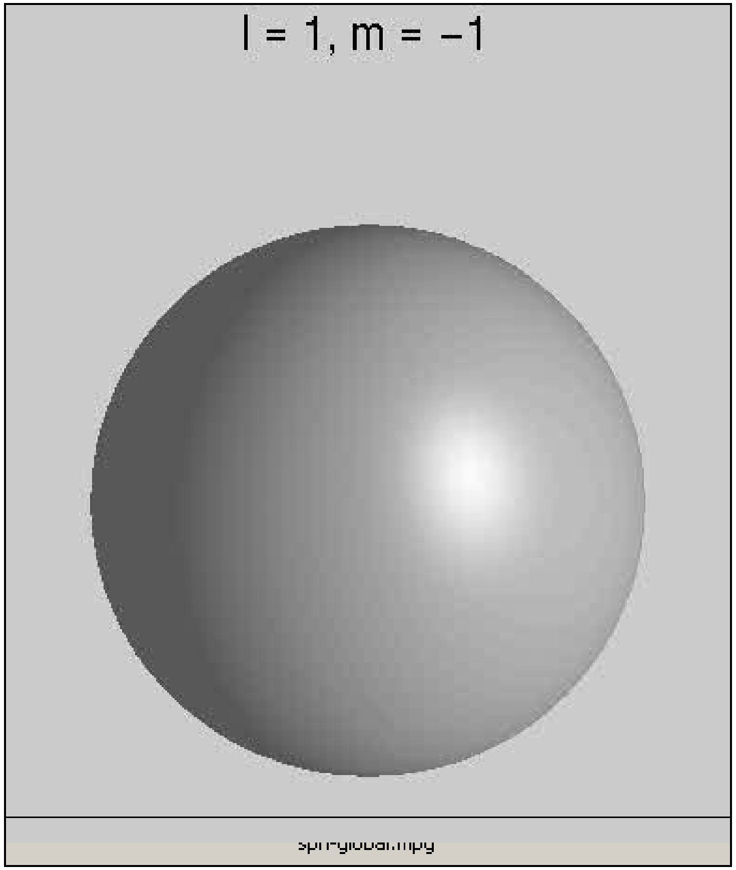
$$\hat{f}_{lm} = \int_{\eta\in S^2} f(\eta)\overline{Y_m^l}(\eta)d\eta$$

The $(2l+1)$ $\hat{f}_{lm}$ are the spherical harmonic coefficients of degree l.

Reconstruction with Spherical Harmonics



Spherical range images



Let us put it in a more general framework….

- **Images are functions on homogeneous spaces (group quotients) and mappings are groups acting on them.**

A **representation** of a group G is a homomorphism $T : G \rightarrow GL(V)$.

A representation is **unitary** if for all $g \in G$

$$(T(g)v, T(g)w) = (v, w) \quad \forall v, w \in V.$$

A representation T is **reducible** if there is a proper subspace W of V which is invariant under T. Otherwise, T is irreducible.

Let T be representation of group G in vector space V. Then T is **irreducible** if and only if the only $A : V \rightarrow V'$ satisfying $T(g)A = AT(g)$, $\forall g \in G$ are $A = \lambda I$.

If the acting group is unimodular and locally compact, the group has an irreducible representation $U(g, p)$.

The Fourier transform of a function on the homogeneous space G/H exists:

$$F(p) = \int_{\eta \in G/H} f(\eta) U(g^{-1}, proj(p)) d\eta.$$

In the case of the sphere $S^2 = SO(3)/SO(1)$

$$\hat{f}^l_m = \int_{\eta \in G/H} f(\eta) U^l_{m0}(\eta) d\eta$$

where U^l_{mn} the irreducible unitary representation of $SO(3)$.

SO(3) irreducible unitary representation

$$U^l_{mn}(g(\gamma,\beta,\alpha)) = e^{-im\gamma} P^l_{mn}(\cos(\beta)) e^{-in\alpha}$$

Framework for image processing in various domains

- Identify the domain of definition of the signal as a homogeneous space and the group acting on it.
- Check whether an irreducible unitary representation exists for the acting group. Compute the generalized Fourier transform of the image.
- Compute the transformation (group action) from a generalized shift theorem. Compute invariants from the magnitude of the Fourier coefficients.

Problem 1: Rotation estimation

Sphere SO(3)/SO(2)

Problem: Compute the rotation of a spherical image directly from its spherical harmonic coefficients (no correspondence).

Current methods: Iterative closest point gradient decent minimization or hierarchical flow algorithms.

Shift Theorem

$$\widehat{f}_{lm}^{g} = \sum_{|p| \leq l} U_{pm}^{l} \widehat{f}_{lp}$$

$$\widehat{f}_{l}^{g} \equiv \Lambda_g \widehat{f}_l = U^l(g)^T \widehat{f}_l$$

Image Invariants

$$\Lambda_g \hat{f}_l = U^l(g)^T \hat{f}_l$$

U^l is a unitary matrix,

$$K_l(f(\eta)) = \sum_{|m| \le l} \overline{\hat{f}_{lm}} \hat{f}_{lm}$$

Approach

Problem: Determine if two spherical images A and B are related by a rotation $g(\alpha, \beta, \gamma)$, and if so, what are α, β, and γ.

We can use our invariant function $K_l(f(\eta))$ to determine if two images are identical up to a rotation.

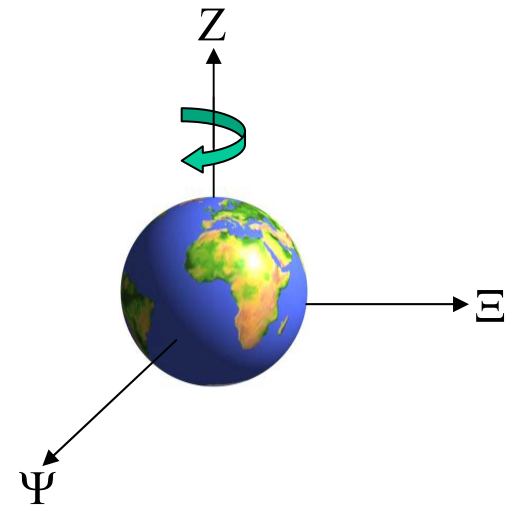
We extract Euler angles of rotation from the Shift Theorem.

$$U^l \text{ gives } U^l(g_1 g_2) = U^l(g_1)U^l(g_2)$$

$$g(\alpha, \beta, \gamma) = g_1(\alpha + \tfrac{\pi}{2}, \tfrac{\pi}{2}, 0) g_2(\beta + \pi, \tfrac{\pi}{2}, \gamma + \tfrac{\pi}{2})$$

$$\Lambda_{g_2 g_1} \hat{f}_l = (U^l(g_1))^T (U^l(g_2))^T \hat{f}_l$$

$$\hat{f}^g_{lm} = e^{-im(\gamma+\frac{\pi}{2})} \sum_{|p|\leq l} e^{-ip(\alpha+\frac{\pi}{2})} \hat{f}_{lp} \sum_{|k|\leq l} P^l_{pk}(0) P^l_{km}(0) e^{-ik(\beta+\pi)}$$

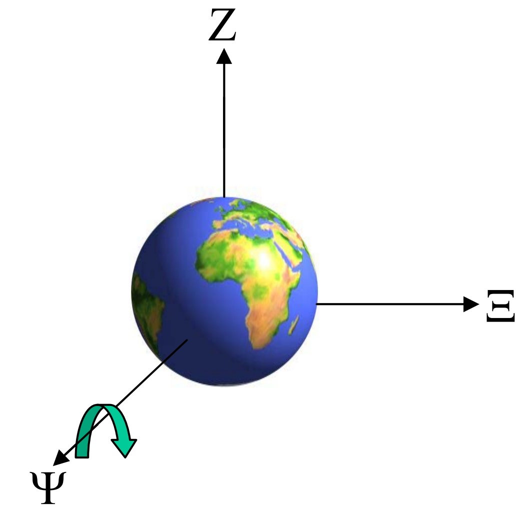


Re-examining the shift property we see that the angle alpha appears only once.

$$\hat{f}^{g}_{lm} = e^{-im(\gamma+\frac{\pi}{2})} \sum_{|p|\leq l} e^{-ip(\alpha+\frac{\pi}{2})} \hat{f}_{lp} \sum_{|k|\leq l} P^{l}_{pk}(0) P^{l}_{km}(0) e^{-ik(\beta+\pi)}$$

We can generate an over-constrained system using multiple coefficients with m>0

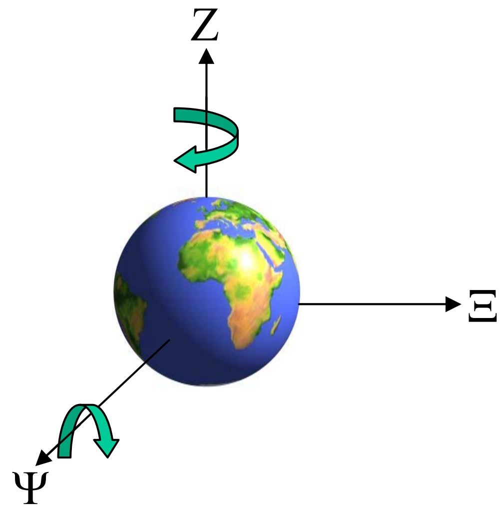
Estimating Rotation Around Y-axis



Without loss of generality, we assume that only beta is nonzero (apply known alpha and gamma rotations to images prior to estimation).

Rewriting the shift property we get

$$
\begin{aligned}
\Lambda_g \hat{f}_{lm} &= \sum_{|p|\leq l} e^{-ip\beta} C^l_{mp} \\
C^l_{mp} &= e^{-im(\frac{\pi}{2})} \Big(\sum_{|l|\leq l} e^{-ik(\frac{\pi}{2})} \hat{f}_{lk} P^l_{kp}(0) P^l_{pm}(0) e^{-ik\pi} \Big)
\end{aligned}
$$



Estimation is done in two steps

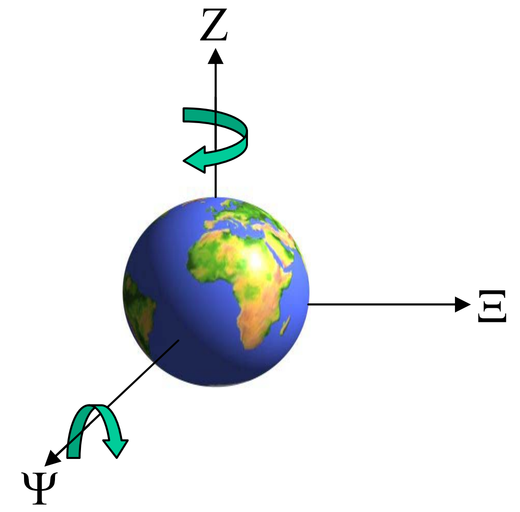
- Generate estimates for beta and gamma.

- Use beta and gamma as input to solving for alpha, which we already know how to calculate.

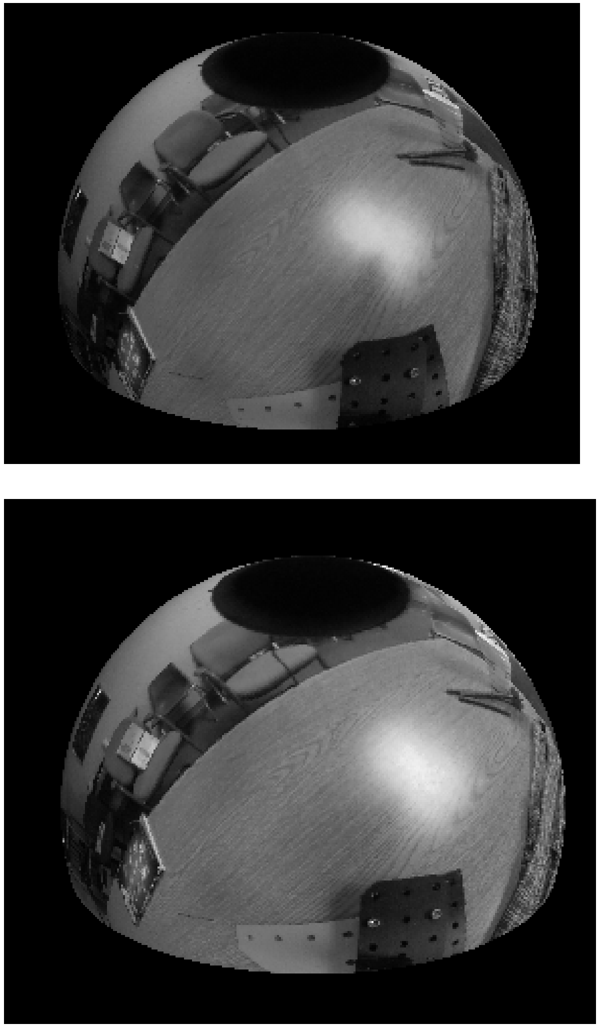
Estimating beta and gamma

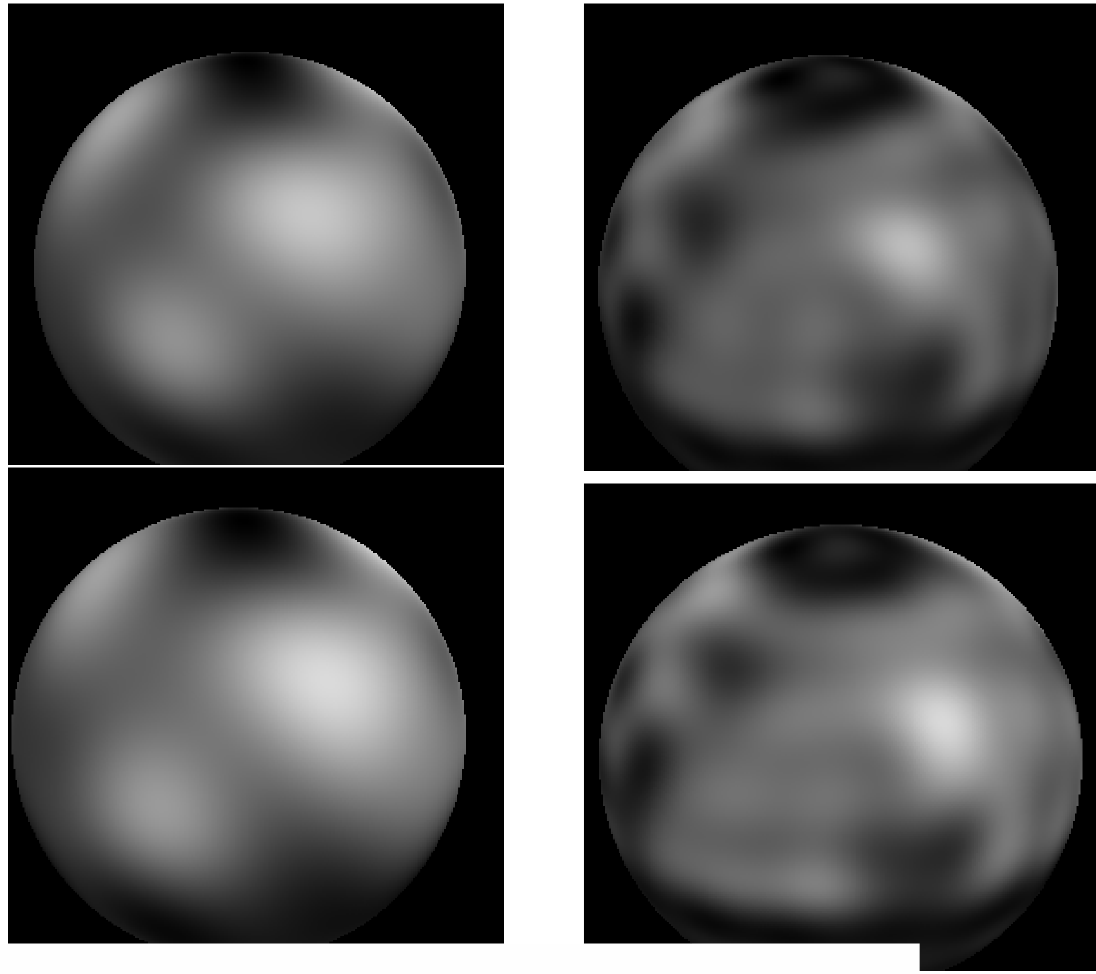
The first rotation of alpha is not reflected in the coefficients f_l0

Using only the equations for the coefficients f_l0, we get an over-constrained system for the two unknowns beta and gamma



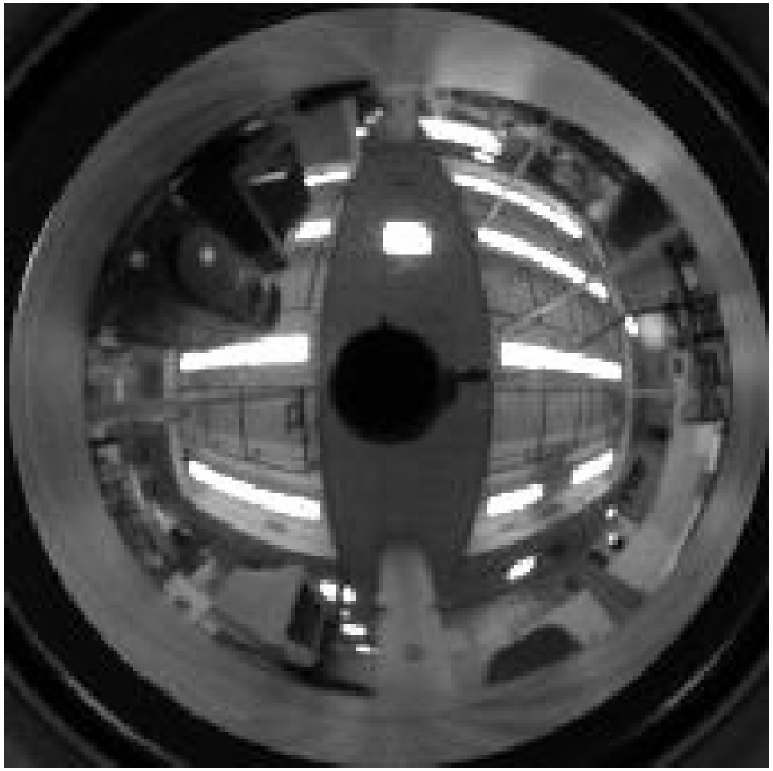
Estimation from very few coefficients!

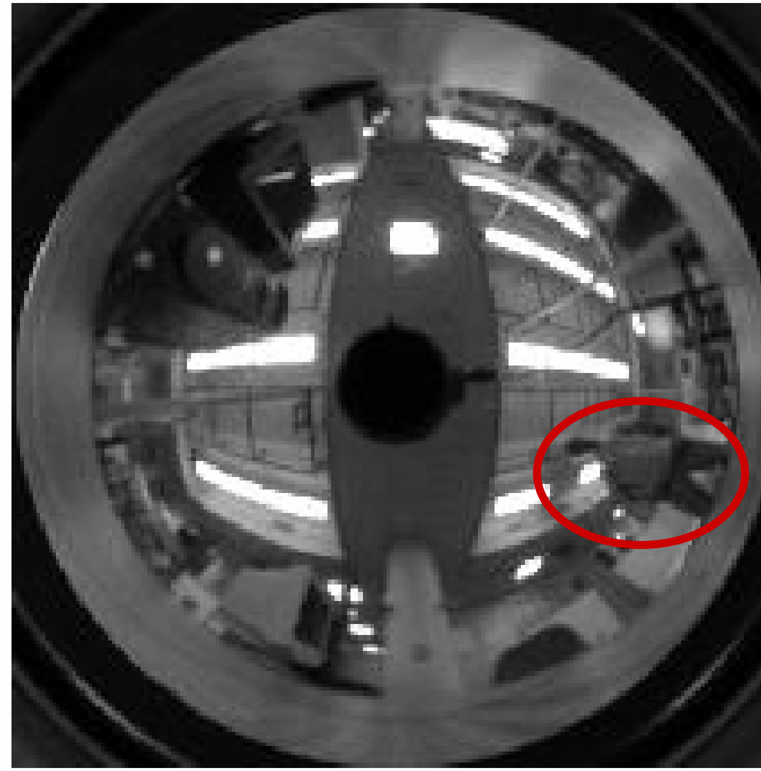




Angle	$l \leq 5$	$l \leq 8$	$l \leq 16$	Flow
$\alpha = -2.4^\circ$	-2.2°	-2.3°	-2.2°	-2.7°
$\beta = 9.3^\circ$	7.2°	7.6°	7.3°	8.1°
$\gamma = 2.2^\circ$	2.5°	2.4°	2.7°	1.9°

Resistant to clutter





Angle	$l \leq 8$	$l \leq 16$	Flow	$l \leq 8$	$l \leq 16$	Flow	$l \leq 8$	$l \leq 16$	Flow
$\alpha = 15°$	14.96°	14.96°	14.88°	14.57°	14.83°	14.76°	14.19°	14.19°	14.45°
$\beta = 13.8°$	13.87°	14.03°	13.88	13.87°	13.81°	13.90	13.96°	13.96°	13.98
$\gamma = 12.8°$	13.01°	12.89°	12.94°	13.11°	13.11°	13.41°	13.74°	13.68	13.50

6%, 10%, and 13% clutter

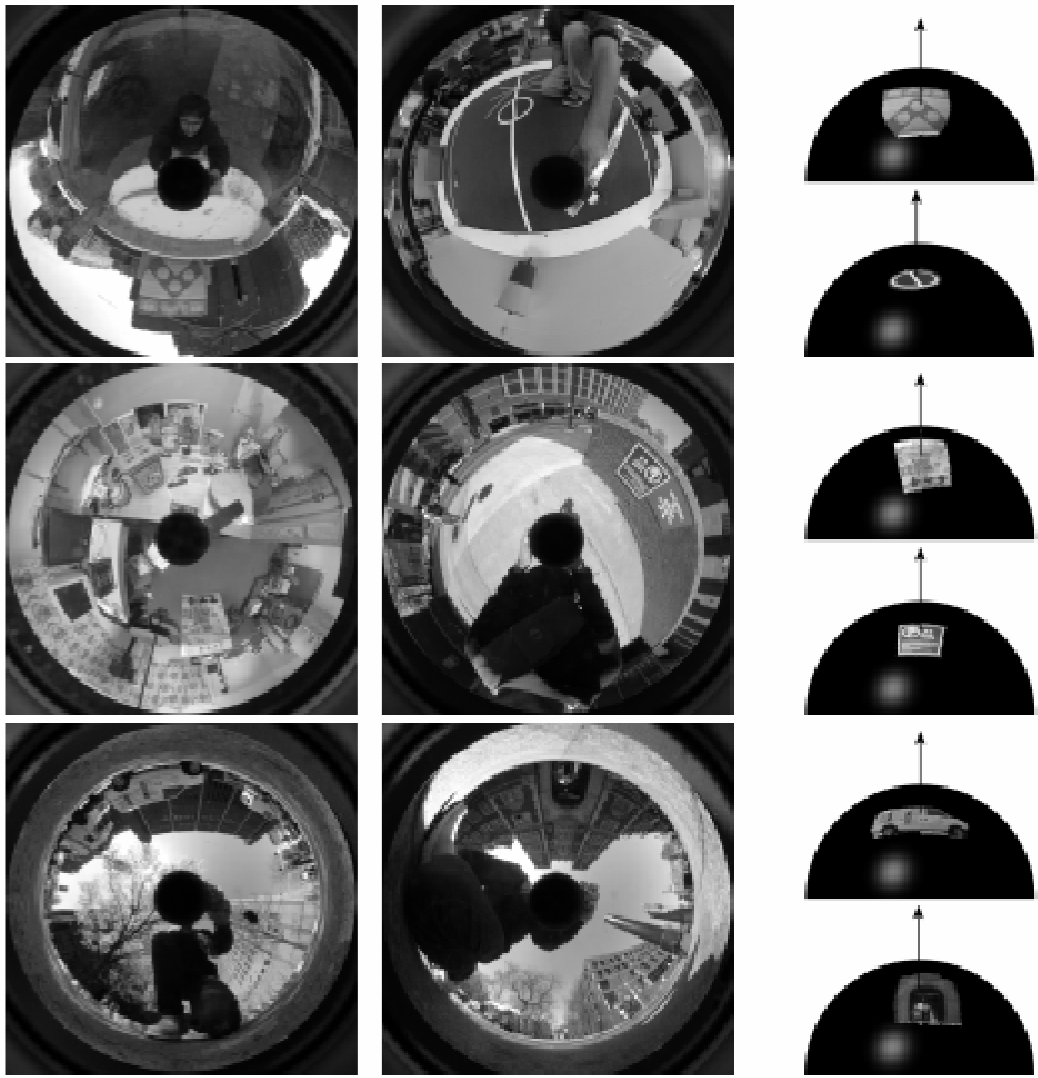
Problem 2: Template matching

Given $f(\eta), h(\eta) \in L^2(S^2)$, the correlation between $f(\eta)$ and $h(\eta)$ is defined as

$$g(\alpha, \beta, \gamma) = \int_{S^2} f(\eta)\Lambda(R)h(\eta)\ d\eta$$

Correlation can be obtained from the spherical harmonics $\hat{f}^l_m$ and $\hat{h}^l_m$ via the 3-D Inverse Discrete Fourier Transform as

$$g(\alpha, \beta, \gamma) = IDFT\{\sum_l \hat{f}^l_m \overline{\hat{h}^l_k} U^l_{m,h}(\pi/2) U^l_{h,k}(\pi/2)\}.$$



Harmonic analysis

- Global shape descriptors (moment, Fourier-descriptors) of the 60's-80's have been abandoned because of occlusions.
- Omnidirectional images give you large closed areas persistent in images (many appearance based techniques)
- Classical Fourier can not be applied anyway due to the new deformations.
- Let us re-think Fourier-transforms!