Geometry and Algebra of Computer Vision

Autumn 97 - spring 98 (4 points prel.)

Fridays 13:15 - 15:00

BB2 Liljanshuset

LECTURE NOTES

.

GEOMETRY AND ALGEBRA OF PROJECTIVE VIEWS

Lecture Notes Nov. 1997

Stefan Carlsson
Computational Vision and Active Perception Laboratory
NADA-KTH, Stockholm, Sweden
stefanc@bion.kth.se

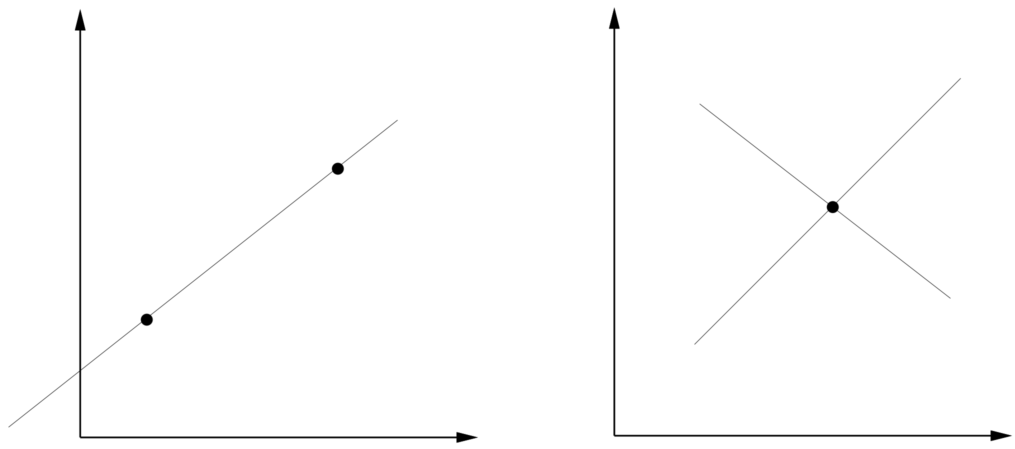

Figure 1: *The problems of finding the line joining two points and the point of intersection of two lines are dual*

1 Points Lines and Planes - Duality

1.1 Join and Intersection

If we have a coordinate system in 2-D, a point can be assigned coordinates $(\bar{x}, \bar{y})$. A line in 2-D consists of all points satisfying:

$$\bar{a}\bar{x} + \bar{b}\bar{y} + 1 = 0 \tag{1}$$

Let two points have coordinates $(\bar{x}_1, \bar{y}_1)$a nd $(\bar{x}_2, \bar{y}_2)$ respectively. What is the equation of the line connecting the two points ?

Since the line passes through the two points, it must satisfy the equations:

$$\begin{aligned} \bar{a}\bar{x}_1 + \bar{b}\bar{y}_1 + 1 &= 0 \\ \bar{a}\bar{x}_2 + \bar{b}\bar{y}_2 + 1 &= 0 \end{aligned} \tag{2}$$

This system can be solved for a and b and we get:

$$\bar{a} = \frac{\bar{y}_1 - \bar{y}_2}{\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1} \qquad \bar{b} = \frac{\bar{x}_2 - \bar{x}_1}{\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1} \tag{3}$$

Suppose now that we have two lines and want to find their point of intersection, $(\bar{x}, \bar{y})$

The two lines must therefore satisfy:

$$\begin{aligned} \bar{a}_1\bar{x} + \bar{b}_1\bar{y} + 1 &= 0 \\ \bar{a}_2\bar{x} + \bar{b}_2\bar{y} + 1 &= 0 \end{aligned} \tag{4}$$

We can solve for the point $(\bar{x}, \bar{y})$ and we get:

$$\bar{x} = \frac{\bar{b}_1 - \bar{b}_2}{\bar{a}_1 \bar{b}_2 - \bar{a}_2 \bar{b}_1} \qquad \bar{y} = \frac{\bar{a}_2 - \bar{a}_1}{\bar{a}_1 \bar{b}_2 - \bar{a}_2 \bar{b}_1} \tag{5}$$

There is a very nice symmetry or *duality* between these two problems. The coordinates for points and lines play exactly the same role in the two cases. We will see that this duality extends to other combinations of geometric objects.

1.2 Homogeneous coordinates

There are certain problems associated with special cases of these pairs of points and lines. Suppose that the two points are scalar multiples of each other, $\bar{x}_2 = k\bar{x}_1$ and $\bar{y}_2 = k\bar{y}_1$. This means that $\bar{x}_1\bar{y}_2 - \bar{x}_2\bar{y}_1 = 0$ and the parameters for the joining line cannot be determined. The joining line in this case will pass through the origin $(0, 0)$ of the coordinate system and this will create problems as we can see already in the equation for the line 1 which obviously cannot be used to represent lines through the origin. The same problem occurs when we want to find the intersection point for two parallel lines, i.e for the case $\bar{a}_2 = k\bar{a}_1$ and $\bar{b}_2 = k\bar{b}_1$. In order to get around this problem we introduce *homogeneous coordinates* for points and lines as:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = w \begin{pmatrix} \bar{x} \\ \bar{y} \\ 1 \end{pmatrix} \qquad \begin{pmatrix} a \\ b \\ c \end{pmatrix} = c \begin{pmatrix} \bar{a} \\ \bar{b} \\ 1 \end{pmatrix} \tag{6}$$

Where w and c are arbitrary scale factors. It is easy to see that the *cartesian coordinates* $\bar{x}, \bar{y}$ can be expressed in terms of the homogeneous coordinates as:

$$\bar{x} = \frac{x}{w} \qquad \bar{y} = \frac{y}{w} \tag{7}$$

Since the scale factor w is arbitrary, the homogeneous coordinates actually represents a line through the origin in euclidean 3-space, E^3 which is a one-dimensional linear subspace. This collection of lines through the origin of E^3 will be denoted as *projective 2-space*, or P^2.

The equation for the line can now be written as:

$$ax + by + cw = 0 \tag{8}$$

Lines through the origin can now be represented. They simply have $c = 0$. Likewise the intersection point of two parallel lines, the point at infinity, has $w = 0$.

The use of homogeneous coordinates will also simplify the analysis of join and intersection. Using homogeneous coordinates, the equations 2 becomes:

$$\begin{aligned} ax_1 + by_1 + cw_1 &= 0 \\ ax_2 + by_2 + cw_2 &= 0 \end{aligned} \tag{9}$$

similarly, the equations 4 can be written as:

$$a_1x + b_1y + c_1w = 0$$

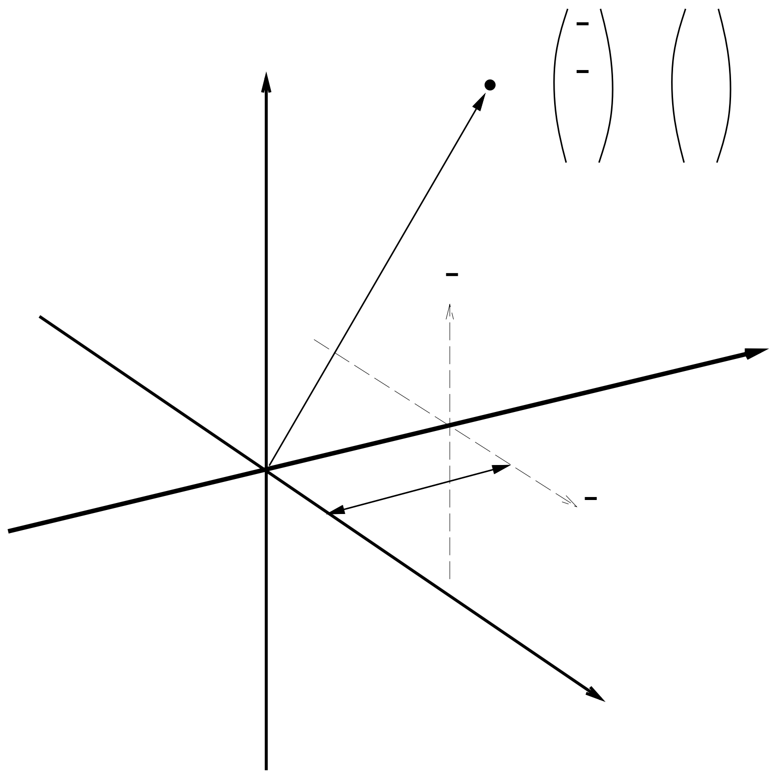

Figure 2: *Cartesian* $(\bar{x}, \bar{y})$ *and homogeneous* (x, y, w) *coordinates*

$$a_1 x \ + \ b_1 y \ + \ c_1 w \ = \ 0 \tag{10}$$

Consider the rectangular matrices:

$$\begin{pmatrix} x_1 & y_1 & w_1 \\ x_2 & y_2 & w_2 \end{pmatrix} \qquad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \tag{11}$$

From equations 3 and 5 we see that the line joining two points and the point of intersection of two lines can be expressed using sub-determinants of these rectangular matrices:

$$\bar{a} = \frac{\begin{bmatrix} y_1 & w_1 \\ y_2 & w_2 \end{bmatrix}}{\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}} \qquad \bar{b} = -\frac{\begin{bmatrix} x_1 & w_1 \\ x_2 & w_2 \end{bmatrix}}{\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}} \tag{12}$$

$$\bar{x} = \frac{\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}}{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}} \qquad \bar{y} = -\frac{\begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix}}{\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}} \tag{13}$$

But this means that the homogeneous line coordinates for the joining line and point of intersection are respectively

$$\begin{bmatrix} y_1 & w_1 \\ y_2 & w_2 \end{bmatrix}, \quad -\begin{bmatrix} x_1 & w_1 \\ x_2 & w_2 \end{bmatrix}, \quad \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \tag{14}$$

$$\begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}, \quad -\begin{bmatrix} a_1 & c_1 \\ a_2 & c_2 \end{bmatrix}, \quad \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \tag{15}$$

We see that two arbitrary points on a line can be used to express the line coordinates, by computing the subdeterminants in eq. 14. These subdeterminants are known as the *Grassmann coordinates* of the line. Another way to view a line in 2-D is as a two dimensional linear subspace in E^3, analogous to the way a point was a one-dimensional linear subspace. Two separate but otherwise arbitrary points on a line can be used as a basis and we consider the points:

$$\alpha \begin{pmatrix} x_1 \\ y_1 \\ w_1 \end{pmatrix} + \beta \begin{pmatrix} x_2 \\ y_2 \\ w_2 \end{pmatrix} \quad \alpha, \beta \in R \tag{16}$$

For arbitrary α and β this obviously represents a point on the line, which can be seen by insertion in eq. 8.

Using Grassmann coordinates for the line, we can express the condition that three points are on the same line. If the first two points are used to construct the line coordinates, we get the condition for the third point to be on the line as:

$$\begin{bmatrix} y_1 & w_1 \\ y_2 & w_2 \end{bmatrix} x_3 - \begin{bmatrix} x_1 & w_1 \\ x_2 & w_2 \end{bmatrix} y_3 + \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} w_3 = 0 \tag{17}$$

The left side is actually just the determinant formed from the homogeneous coordinates of the three points, and we have:

$$\begin{bmatrix} x_1 & y_1 & w_1 \\ x_2 & y_2 & w_2 \\ x_3 & y_3 & w_3 \end{bmatrix} = 0 \tag{18}$$

In the very same way we can consider three lines that intersect in a common point. This leads to the condition on the line coordinates.

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = 0 \tag{19}$$

1.3 Conics

Homogeneous coordinates can also be used to describe algebraic curves in general. A specially interesting class of curves are the conic sections which are second order polynomials in cartesian image coordinates $\bar{x}$ and $\bar{y}$.

$$a\bar{x}^2 + b\bar{y}^2 + 2c\bar{x}\bar{y} + 2d\bar{x} + 2e\bar{y} + f = 0 \tag{20}$$

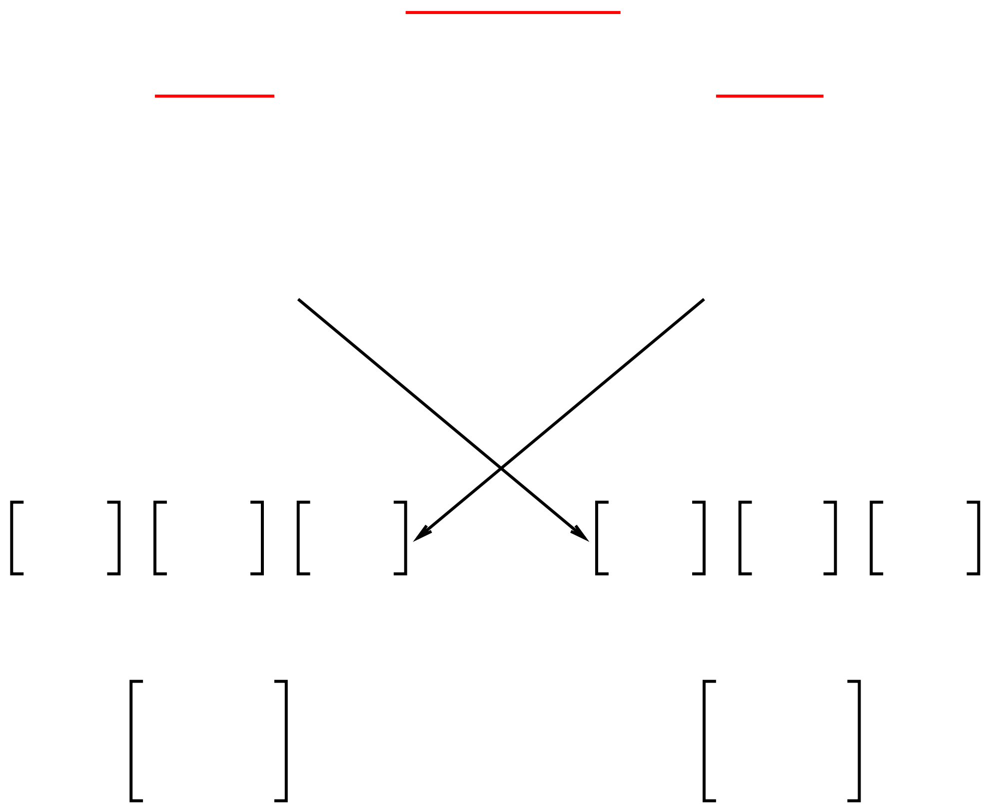

Figure 3: *Point line duality*

In homogeneous coordinates this equation becomes:

$$ax^2 \;+\; by^2 \;+\; 2cxy \;+\; 2dxw + \;2eyw + \;fw^2 \;=\; 0 \tag{21}$$

Which can be expressed as:

$$(x, y, w) \begin{pmatrix} a & c & d \\ c & b & e \\ d & e & f \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \;=\; 0 \tag{22}$$

We write this more compactly as:

$$p^T \mathbf{P} p \;=\; 0 \tag{23}$$

If p' is a point on the conic, The line with homogeneous line coordinates:

$$l' \;=\; \mathbf{P}p' \tag{24}$$

obviously has the point p' in common with the conic. It also has the same direction as the conic at that point. It is therefore the tangent line of the conic at point p'. For every point p on the conic we get the tangent line:

$$l \;=\; \mathbf{P}p \tag{25}$$

If $\mathbf{P}$ is non-singular, we have:

$$p = \mathbf{P}^{-1} l \tag{26}$$

Which can be inserted in the equation of the conic, 23.

$$p^T \mathbf{P} p = (\mathbf{P}^{-1} l)^T \mathbf{P} \mathbf{P}^{-1} l = l^T \mathbf{P}^{-1} \mathbf{P} \mathbf{P}^{-1} l = l^T \mathbf{P}^{-1} l = 0 \tag{27}$$

Where we have used the fact that $\mathbf{P}^{-T} = \mathbf{P}^{-1}$ due to symmetry.

The last identity is a quadratic constraint on the line coordinates of the tangent lines of the conic. It represents the conic equally well as the eq. 23, with point coordinates substituted for line coordinates.

1.4 Points and Planes in 3-D

If we move to 3-D space we will find a corresponding duality between points and planes. The equation for a plane in 3-D can be written using homogeneous point x, y, z, w and plane coordinates a, b, c, d

$$ax + by + cz + dw = 0 \tag{28}$$

In an analogous way as in 2-D, we can consider the plane passing through three points in space.

$$\begin{aligned} ax_1 + by_1 + cz_1 + dw_1 &= 0 \\ ax_2 + by_2 + cz_2 + dw_2 &= 0 \\ ax_3 + by_3 + cz_3 + dw_3 &= 0 \end{aligned} \tag{29}$$

If we define the vector $P_i^T = (x_i,\ y_i,\ z_i,\ w_i)$ this problem amounts to finding the vector (a, b, c, d) orthogonal to the P_i:s

$$P_{i,1} a + P_{i,2} b + P_{i,3} c + P_{i,4} d = 0 \qquad i = 1, 2, 3 \tag{30}$$

Since a determinant of a matrix is zero whenever two rows or columns are equal, The vectors P_i have the property:

$$[P_i\ P_1\ P_2\ P_3] = 0 \qquad i = 1, 2, 3 \tag{31}$$

but

$$[P_i\ P_1\ P_2\ P_3] =$$

$$P_{i,1}\ (P_1\ P_2\ P_3)_{234} + P_{i,2}\ (P_1\ P_2\ P_3)_{134} + P_{i,3}\ (P_1\ P_2\ P_3)_{124} + P_{i,4}\ (P_1\ P_2\ P_3)_{123}$$

where $(P_1\ P_2\ P_3)_{klm}$ is the minor formed by rows klm of the rectangular 4×3 matrix $(P_1\ P_2\ P_3)$

This means that we can take:

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} (P_1\ P_2\ P_3)_{234} \\ (P_1\ P_2\ P_3)_{134} \\ (P_1\ P_2\ P_3)_{124} \\ (P_1\ P_2\ P_3)_{123} \end{pmatrix} \tag{32}$$

as our solution vector. We see that the plane coordinates are formed from the three defining points in space in the same way as the line coordinates were formed from the two defining points in the plane, by considering the minors of the rectangular matrices formed by the homogeneous coordinate vectors of the points.

$$(P_1\ P_2\ P_3) = \begin{pmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \tag{33}$$

For a fourth point lying on the plane we get:

$$ax_4 + by_4 + cz_4 + dw_4 = 0 \tag{34}$$

Using the minors to express the plane coordinates, we have:

$$(P_1\ P_2\ P_3)_{234}\ x_4 + (P_1\ P_2\ P_3)_{134}\ y_4 + (P_1\ P_2\ P_3)_{124}\ z_4 + (P_1\ P_2\ P_3)_{123}\ w_4 = 0 \tag{35}$$

which is just an expansion of the determinant:

$$\begin{bmatrix} x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \\ x_3 & y_3 & z_3 & w_3 \\ x_4 & y_4 & z_4 & w_4 \end{bmatrix} = 0 \tag{36}$$

If four points are coplanar in space, the determinant computed from the matrix of the homogeneous coordinates of the points will vanish, analogously to the vanishing of the determinant of three collinear points in the plane.

In a dual way, we can consider the point of intersection of three planes in space. The homogeneous coordinates of this intersection point can be computed from the minors of the 4×3 matrix formed by the plane coordinates of the three planes.

Analogously as for three lines intersecting in a common point, eq. 19 the condition for four planes to intersect a common point, concurrency, can be written in terms of the determinant of their coordinates vanishing.

1.5 Lines in 3-D

What about lines in 3-D ?. Can they be characterized using homogeneous coordinates in the same way as points , planes and lines in 2-D ? The answer to this is partly negative. There is no single equation like 28 for as line in 3-D. What we are looking for is a unique representation of the points on the line, i.e. for as specific line we want to fined some parameters that can be used to compute these points. For that purpose we consider two distinct points with cartesian coordinate vectors P_1 and P_2 respectively. The points on the line passing through these points are given in cartesian coordinates by:

$$P = P_1 + \lambda(P_2 - P_1) \quad \lambda \in R \tag{37}$$

and in homogeneous coordinates as:

$$w \begin{pmatrix} P \\ - \\ 1 \end{pmatrix} = w(1-\lambda) \begin{pmatrix} P_1 \\ - \\ 1 \end{pmatrix} + w\lambda \begin{pmatrix} P_2 \\ - \\ 1 \end{pmatrix} =$$

$$= \frac{w(1-\lambda)}{w_1}\begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ w_1 \end{pmatrix} + \frac{w\lambda}{w_2}\begin{pmatrix} x_2 \\ y_2 \\ z_2 \\ w_2 \end{pmatrix} \tag{38}$$

This is the subspace in E^4 spanned by the homogeneous coordinate vectors the points P_1 and P_2 respectively. Obviously, two arbitrary points on a line in 3-D can be used to represent the line. This is very much the same situation as previously. Can we find a way to characterize this subspace similar to the line coordinates for lines in 2-D ?. Following the previous reasoning we can look at the minors of the rectangular matrix:

$$\begin{pmatrix} x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \end{pmatrix} \tag{39}$$

There are $\binom{4}{2} = 6$ different minors, i.e. 2×2 sub-determinants of this matrix.

$$\begin{bmatrix} y_1 & z_1 \\ y_2 & z_2 \end{bmatrix} \quad \begin{bmatrix} z_1 & x_1 \\ z_2 & x_2 \end{bmatrix} \quad \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \quad \begin{bmatrix} x_1 & w_1 \\ x_2 & w_2 \end{bmatrix} \quad \begin{bmatrix} y_1 & w_1 \\ y_2 & w_2 \end{bmatrix} \quad \begin{bmatrix} z_1 & w_1 \\ z_2 & w_2 \end{bmatrix} \tag{40}$$

The first question we can ask is if these minors are unique for a specific line, i.e. given two new points P_1', P_2' on the same line

$$\begin{pmatrix} x_1' & y_1' & z_1' & w_1' \\ x_2' & y_2' & z_2' & w_2' \end{pmatrix} \tag{41}$$

will the minors be the same up to scale as those for the original points 39

Since the new points are in the subspace spanned by points P_1 and P_2 we have:

$$\begin{pmatrix} x_1' & y_1' & z_1' & w_1' \\ x_2' & y_2' & z_2' & w_2' \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \end{pmatrix} \tag{42}$$

All the sub-matrices on the left side can be written as matrix products with the sub-matrices on the right:

$$\begin{pmatrix} x_1' & y_1' \\ x_2' & y_2' \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \tag{43}$$

$$\vdots$$

$$\begin{pmatrix} z_1' & w_1' \\ z_2' & w_2' \end{pmatrix} = \begin{pmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{pmatrix} \begin{pmatrix} z_1 & w_1 \\ z_2 & w_2 \end{pmatrix} \tag{44}$$

Using the product rule for taking determinants in matrix products, we get:

$$\begin{bmatrix} x_1' & y_1' \\ x_2' & y_2' \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \tag{45}$$

$$\vdots$$

$$\begin{bmatrix} z_1' & w_1' \\ z_2' & w_2' \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \begin{bmatrix} z_1 & w_1 \\ z_2 & w_2 \end{bmatrix} \tag{46}$$

we see that the new minors are scaled versions of the old, with the same scale factor. The representation using minors is therefore unique up to scale for a certain line. The minors in this representation are known as the *Plücker coordinates* of the line.

The Plücker coordinates is a six-dimensional homogeneous vector. The effective number of parameters in this vector is therefore five. The five parameters are however not independent, which can be seen from the following: The matrix :

$$\begin{pmatrix} x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \\ x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \end{pmatrix} \tag{47}$$

is obviously singular, i.e. its determinant is zero. If we use the Laplace-expansion in 2×2 minors of this matrix we get:

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \begin{bmatrix} z_1 & w_1 \\ z_2 & w_2 \end{bmatrix} - \begin{bmatrix} x_1 & z_1 \\ x_2 & z_2 \end{bmatrix} \begin{bmatrix} y_1 & w_1 \\ y_2 & w_2 \end{bmatrix} - \begin{bmatrix} x_1 & w_1 \\ x_2 & w_2 \end{bmatrix} \begin{bmatrix} y_1 & z_1 \\ y_2 & z_2 \end{bmatrix} = 0 \tag{48}$$

This relation means that there are only four independent parameters describing a line.

The interpretation of a line in 3-D as a join of two points has its dual interpretation as the intersection of two planes.

$$\begin{array}{l} a_1 x \ + \ b_1 y \ + \ c_1 z \ + d_1 w = \ 0 \\ a_2 x \ + \ b_2 y \ + \ c_2 z \ + d_2 w = \ 0 \end{array} \tag{49}$$

It can be shown that the line in 3-D will have Plücker coordinates as the minors of the rectangular matrix:

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{pmatrix} \tag{50}$$

Fig. 4 is an overall view of duality in 3-D space.

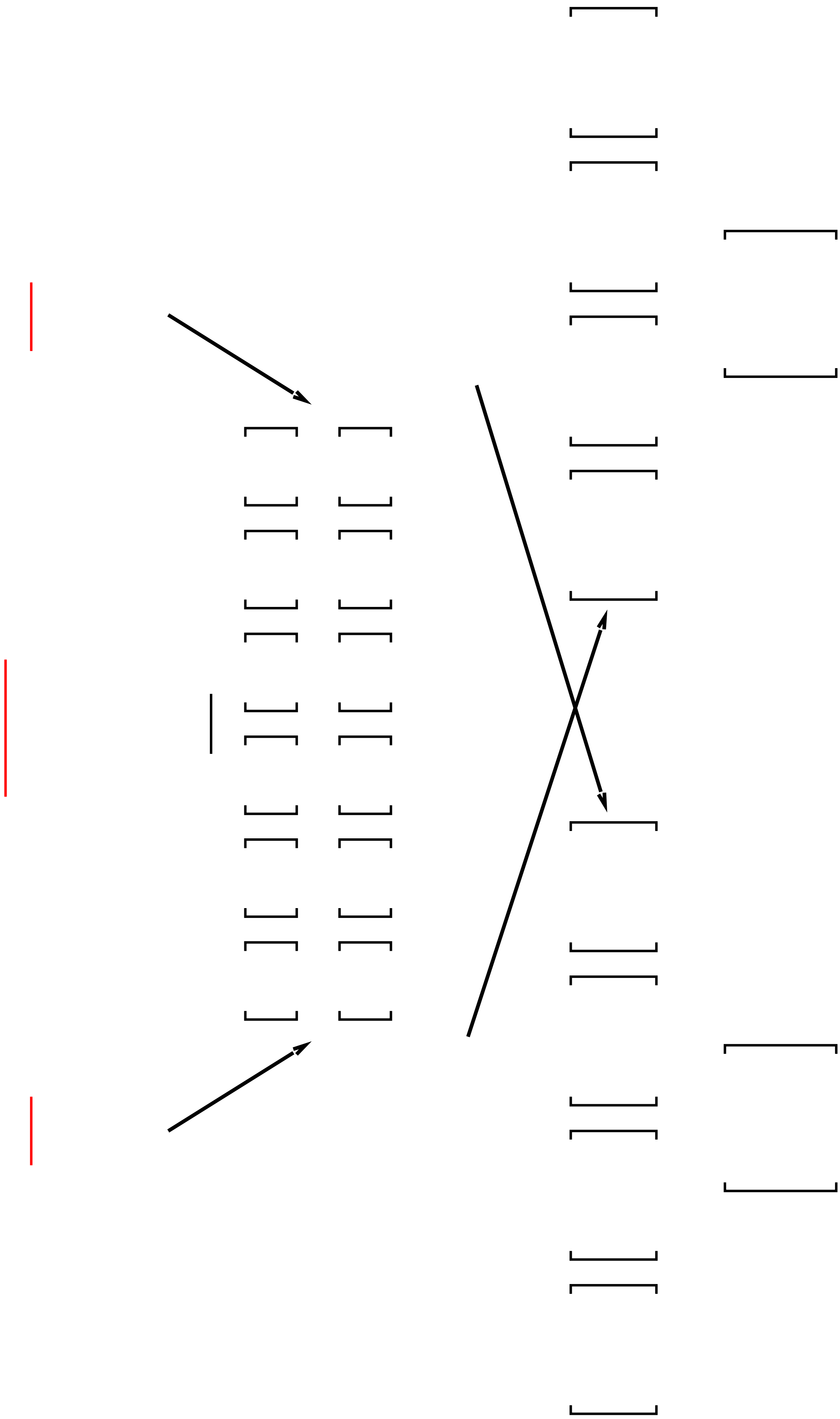

Figure 4: *Duality 3D*

2 Linear Transformations

2.1 Transformation Groups

We will have numerous reasons to study linear transformations of coordinates:

$$p' = \mathbf{T}p \tag{51}$$

The linear transformations can be of different kinds, ranging from the general case where the elements of the matrix $\mathbf{T}$ take on arbitrary values, to special cases where they are constrained.

A set of matrices is a *transformation group* if each member of the set depend on a set of parameters $\alpha_1 \dots \alpha_k$ in such a way that :

Multiplication of two matrices of the set results in a new member of the set:

$$\mathbf{T}(\alpha_1 \dots \alpha_k)\ \mathbf{T}(\beta_1 \dots \beta_k) = \mathbf{T}(\gamma_1 \dots \gamma_k) \tag{52}$$

The unit matrix $\mathbf{I}$ is a member of the set and to every member there is another such that:

$$\mathbf{T}(\alpha_1 \dots \alpha_k)\ \mathbf{T}(\alpha'_1 \dots \alpha'_k) = \mathbf{I} \tag{53}$$

The set of all non-singular transformations $\mathbf{T}$ form a group called the general linear group. All other linear transformation groups will be special cases of the general linear groups and are therefore called subgroups.

In this section we will present the transformation matrices associated two of the most important subgroups to the general linear group, the euclidean and affine groups. Later we will return to the concept of group when we study invariants of their transformations.

2.2 Rigid Motions - Euclidean Transformations

An important subgroup of the general linear transformation group is that corresponding to rigid motions. This means that an object is rotated and translated in space. In cartesian coordinates in 2-D, this can be written as.

$$\begin{pmatrix} \bar{x}' \\ \bar{y}' \end{pmatrix} = \mathbf{R} \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} + \begin{pmatrix} \bar{t}_x \\ \bar{t}_y \end{pmatrix} \tag{54}$$

with

$$\mathbf{R} = \begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix} \tag{55}$$

where θ is the rotation angle.

Using homogeneous coordinates, this transformation can be written as:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} \mathbf{R} & | & t \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{56}$$

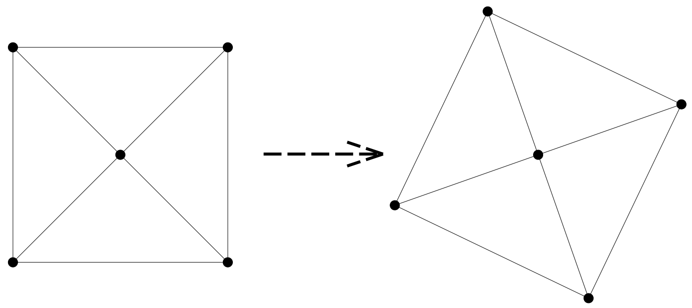

Figure 5: *Action of the euclidean group on five points*

where

$$t \;=\; \begin{pmatrix} \bar{t}_x \\ \bar{t}_y \end{pmatrix} \qquad 0^T \;=\; (0,0) \tag{57}$$

Note that identity is only up to scale. The use of homogeneous coordinates implies that rigid motion can be expressed as a linear transformation. Not all linear transformations will be rigid motion. The class of linear transformations whose matrices has the structure of eq. 56, is called euclidean transformations. Note that two successive rigid motions can be represented as a single motion:

$$\begin{pmatrix} \mathbf{R}_1 & | & t_1 \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \begin{pmatrix} \mathbf{R}_2 & | & t_2 \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \;=\; \begin{pmatrix} \mathbf{R}_1\mathbf{R}_2 & | & \mathbf{R}_1 t_2 + t_1 \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \tag{58}$$

The combination of two successive rotations $\mathbf{R_1}$ and $\mathbf{R_2}$ is the rotation $\mathbf{R_1 R_2}$ By choosing $\mathbf{R}_2 = \mathbf{R}_1{}^{-1}$ and $t_2 = -\mathbf{R}_1{}^{-1} t_1$ the rigid motion is reversed, i.e. there is a matrix that takes the coordinates back to the original position These properties characterize a group and the class of matrices with the structure of that of eq.56 is known as the *euclidean transformation group.* The euclidean transformations are all special cases of the general linear transformation, where the matrix elements are arbitrary. These general non-singular matrices also have the properties of a group which is called the *general linear transformation group* or *projective group.* The euclidean transformation group is then called a sub-group of the general linear group.

The rotation matrices $\mathbf{R}$ form a subgroup of themselves, the orthogonal group. Note that they have the important property that $\mathbf{R}^T\mathbf{R} = \mathbf{I}$

Various sub-groups can be caracterized in an interesting way by the geometric properties that are left invariant after the action of the group. The euclidean group has the property that angles and distances are left invariant. This follows intuitively from the fact that the euclidean group represents rigid motions, fig. 5 but can be seen formally by noting that angles and distances can all be expressed as inner products of differences of cartesian coordinates. Under the euclidean action these are transformed as:

$$< \bar{p}'_i - \bar{p}'_j \; \bar{p}'_k - \bar{p}'_l > \;=\; (\mathbf{R}\bar{p}_i - \mathbf{R}\bar{p}_j)^T \; (\mathbf{R}\bar{p}_k - \mathbf{R}\bar{p}_l) \;=$$

$$= \; (\bar{p}_i - \bar{p}_j)^T \mathbf{R}^T \mathbf{R} \; (\bar{p}_k - \bar{p}_l) \;=\; (\bar{p}_i - \bar{p}_j)^T \; (\bar{p}_k - \bar{p}_l) \;=$$

$$= < \bar{p}_i - \bar{p}_j \; \bar{p}_k - \bar{p}_l > \tag{59}$$

where the last equality follows from the orthogonality of the rotation matrix $\mathbf{R}$

2.3 Affine transformations

If the rotation matrix $\mathbf{R}$ in eq. 56 is replaced by a general non-singular transformation matrix $\mathbf{A}$ we get the *affine transformation*:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} \mathbf{A} & | & b \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{60}$$

These transformations form a subgroup of the general linear (projective) group in the same way as euclidean transformations. Note that euclidean transformations are special cases i.e. subgroups of the affine transformation groups. These groups therefore form a hierarchy in the sense that:

$$euclidean \subset affine \subset projective \tag{61}$$

Later we will see how various camera models and degrees of calibration will permit the computation of geometric information up to an arbitrary transformation of the kind above. We will speak of euclidean, affine and projective structure respectively.

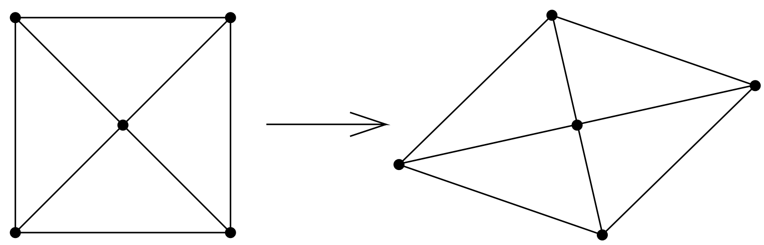

Figure 6: *Action of the affine group on five points*

Affine transformations will in general deform the shape of an object in addition to arbitrary rigid motions. They have the property however that parallell lines are preserved. This follows simply from the fact that points at infinity are mapped to points at infinity under affine transformations:

$$\begin{pmatrix} \mathbf{A} & | & b \\ -- & -- & -- \\ 0^T & | & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} x' \\ y' \\ 0 \end{pmatrix} \tag{62}$$

where:

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix} \tag{63}$$

Parallelism of two lines implies that they intersect in a point at infinity. Since points of intersection are mapped to points of intersection under general linar transformations this means that affinely transformed parallel lines will remain parallel.

2.4 General Linear Transformations

2.4.1 Properties

For general non-singular linear transformations $\mathbf{T}$

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ w \end{pmatrix} \tag{64}$$

we get the *general linear* or *projective* group of transformations. Projective transformations no longer preserve parallelism of lines which is a consequence of the fact that points at infinity can be mapped to arbitrary points and vice versa. Properties that are preserved however are e.g. collinearity of points. For three points this can be expressed

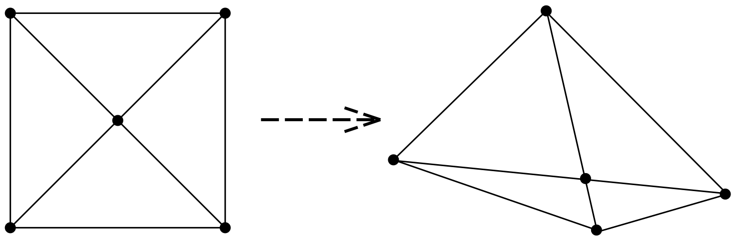

Figure 7: *Action of the projective group on five points*

as the vanishing of the bracket:

$$[p_1\ p_2\ p_3] = 0 \tag{65}$$

If we write:

$$p' = \mathbf{T}p \tag{66}$$

for the projective transformation of points p in the plane we get:

$$[p'_1\ p'_2\ p'_3] = [\mathbf{T}p_1\ \mathbf{T}p_2\ \mathbf{T}p_3] = \mathbf{T}[p_1\ p_2\ p_3] = 0 \tag{67}$$

2.4.2 Transformations of lines and conics

Just as coordinates of points are transformed, lines and conics will be transformed by linear transformations. Suppose the homogeneous coordinates p of a point are subjected to a linear transformation: $p' = \mathbf{T}p$. For a line with homogeneous line coordinates q:

$$q^T p = 0 \tag{68}$$

we get:

$$q^T p = q^T \mathbf{T}^{-1} p' = {q'}^T p' = 0 \tag{69}$$

I.e. the line coordinates q are transformed to :

$$q'^T = q^T \mathbf{T}^{-1} \tag{70}$$

or

$$q' = \mathbf{T}^{-1^T} q \tag{71}$$

A conic with equation:

$$p^T \mathbf{Q} p = 0 \tag{72}$$

will be transformed as:

$$p^T \mathbf{Q} p = (\mathbf{T}^{-1} p')^T \mathbf{Q} \mathbf{T}^{-1} p' = p'^T \mathbf{T}^{-1^T} \mathbf{Q} \mathbf{T}^{-1} p' = p'^T \mathbf{Q}' p' = 0 \tag{73}$$

i.e. the matrix representing the conic is transformed as:

$$\mathbf{Q}' = \mathbf{T}^{-1^T} \mathbf{Q} \mathbf{T}^{-1} \tag{74}$$

Later on we will give a more general view of the computation of *invaraints* of different groups and see how these can be used in the geometric description of objects.

3 Projection From 3-D to 2-D

3.1 The Projection Equation

The standard pin-hole camera model assumes that an image point is the projection of a point P in 3-D, with projection line passing through the projection point P_0. (fig. 8) By fixing a basis for the image plane in 3-D, we can assign image coordinates x, y, w to the image point a in 3-D. If P_1, P_2 and P_3 are the homogeneous 3-D coordinates of three points on the image plane in 3-D space, any linear combination $xP_1 + yP_2 + wP_3$ will be the homogeneous coordinates of a point on the image plane. Since the $P_i : s$ are homogeneous however, they can be scaled arbitrarily. and the p_i do not specify a unique point. If we fix the scale of the $P_i : s$ and use P_i^* for the corresponding specific four-vector, we can use P_1^*, P_2^*, P_3^* as a basis for the subspace containing the image points in 3-D.

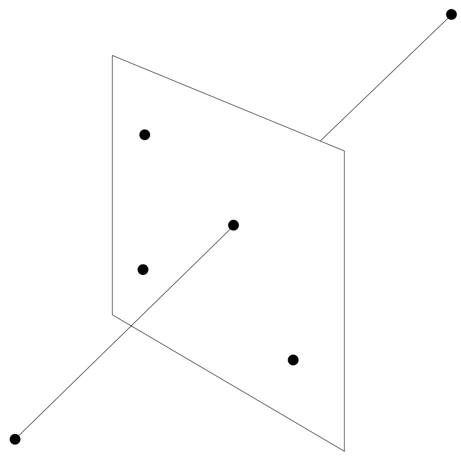

Figure 8: *Pin-hole camera model*

A certain linear combination:

$$P \ = \ xP_1^* \ + \ yP_2^* + \ wP_3^* \tag{75}$$

will represent the homogeneous 3-D coordinates of a unique point on the image plane, and x, y, w can be defined as its homogeneous image coordinates in the image plane spanned by the points P_1^*, P_2^*, P_3^*

The interesting question is of course how these image coordinates are related to the coordinates of the point U in 3-D and the coordinates P_0, P_1^*, P_2^*, P_3^* defining the camera. The points P_0, a and U are all on the same sight line. These three points and any arbitrary 4:th point will therefore always be coplanar. If we take points P_1 and P_2 and use the four point coplanarity condition of eq. 36 we get:

$$[P_0 \ P \ U \ P_1^*] \ = \ 0$$

$$[P_0 \ P \ U \ P_2^*] \ = \ 0 \tag{76}$$

Using eq. 75 for a we have:

$$y[P_0 \ P_2^* \ U \ P_1^*] \ + \ w[P_0 \ P_3^* \ U \ P_1^*] = \ 0$$

$$x[P_0\ P_1^*\ U\ P_2^*]\ +\ w[P_0\ P_3^*\ U\ P_2^*] =\ 0 \tag{77}$$

We can solve this for x, y, w, up to scale, and after some permutations in the determinants we can write:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} [P_0\ P_3^*\ P_2^*\ U] \\ \\ [P_0\ P_1^*\ P_3^*\ U] \\ \\ [P_0\ P_2^*\ P_1^*\ U] \end{pmatrix} \tag{78}$$

If the point U has homogeneous coordinates:

$$U = \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} \tag{79}$$

we can develop each determinant on the right side in eq. 78 linearly in these:

$$[P_0\ P_i^*\ P_j^*\ U] = X(P_0\ P_i^*\ P_j^*)_{234} - Y(P_0\ P_i^*\ P_j^*)_{134} + Z(P_0\ P_i^*\ P_j^*)_{124} - W(P_0\ P_i^*\ P_j^*)_{123}$$

where $(P_0\ P_i^*\ P_j^*)_{klm}$ denotes the minor formed from the rows klm in the 4×3 matrix $(P_0\ P_i^*\ P_j^*)$.

The important thing to note here is that the image coordinates x, y, w actually depend linearly on the space coordinates X, Y, Z, W, i.e. eq. 78 can be written as:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} [P_0\ P_2^*\ P_3^*\ U] \\ \\ [P_0\ P_1^*\ P_3^*\ U] \\ \\ [P_0\ P_2^*\ P_1^*\ U] \end{pmatrix} = \mathbf{P}\ U = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ W \end{pmatrix} \tag{80}$$

where $\mathbf{P}$ is a 3×4 matrix, whose elements are determined by the camera and image plane positions in space. This linear mapping from 3-D to 2-D is an example of a *projective transformation*

3.2 Internal parameters and external orientation

The matrix $\mathbf{P}$ in the projection equation 80 is 3×4 and is therefore described by 11 parameters since they can be scaled arbitrarily. The process of *camera calibration* consists of finding these parameters. Knowing the projection matrix means that we know the image coordinates of every space point. The projection matrix $\mathbf{P}$ can be structured in different ways in order to reveal how it depends on various camera parameters. Let $\bar{U}$, $\bar{P}_0$, $\bar{P}_1$, $\bar{P}_2$ and $\bar{P}_3$ be the cartesian 3-D coordinates of points U P_0 etc. I.e. we have:

$$U = W \begin{pmatrix} \bar{U} \\ -- \\ 1 \end{pmatrix} \qquad P_0 = W_0 \begin{pmatrix} \bar{P}_0 \\ -- \\ 1 \end{pmatrix} \qquad P_i^* = W_i \begin{pmatrix} \bar{P}_i \\ -- \\ 1 \end{pmatrix} \tag{81}$$

where W and W_0 are arbitrary. The image coordinates 80 can then be expresed as

$$[P_0\ P_i^*\ P_j^*\ U] = W_0\ W\ W_i\ W_j \begin{bmatrix} \bar{P}_0 & \bar{P}_i & \bar{P}_j & \bar{U} \\ -- & -- & -- & -- \\ 1 & 1 & 1 & 1 \end{bmatrix} =$$

$$= W_0 \; W \; W_i \; W_j \; [\bar{P}_i - \bar{P}_0 \; \bar{P}_j - \bar{P}_0 \; \bar{U} - \bar{P}_0] \tag{82}$$

Using this, eq. 80 can be written as:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} W_2 W_3 \; [\bar{P}_2 - \bar{P}_0 \quad \bar{P}_3 - \bar{P}_0 \quad \bar{U} - \bar{P}_0] \\ \\ W_1 W_3 \; [\bar{P}_1 - \bar{P}_0 \quad \bar{P}_3 - \bar{P}_0 \quad \bar{U} - \bar{P}_0] \\ \\ W_2 W_1 \; [\bar{P}_2 - \bar{P}_0 \quad \bar{P}_1 - \bar{P}_0 \quad \bar{U} - \bar{P}_0] \end{pmatrix} \tag{83}$$

where the common factor $W_0 W$ has been dropped since we have a homogeneous vector. The determinants on the right side can be developed linearly in the components of $\bar{U} - \bar{P}_0$.

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \mathbf{T} \;\; (\bar{U} - \bar{P}_0) \; = \; \mathbf{T} \, (\, \mathbf{I} \, | - \bar{P}_0 \,) \; U \tag{84}$$

where $\mathbf{T}$ is a general 3×3 matrix with elements taken from the minors of the rectangular matrices $W_i W_j (\bar{P}_i - \bar{P}_0 \; \bar{P}_j - \bar{P}_0)$

We can now use the fact that an arbitrary matrix can be factorized into a *upper triangular* $\mathbf{A}$ and an *orthogonal* $\mathbf{R}$ matrix:

$$\mathbf{T} \; = \; \mathbf{AR} \tag{85}$$

where

$$\mathbf{A} \; = \; \lambda \begin{pmatrix} \alpha_x & \gamma & \bar{x}_0 \\ 0 & \alpha_y & \bar{y}_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{86}$$

where λ is an arbitrary scale factor and the other parameters have geometrical interpretations relating to the choice of the image coordinate system.

The projection equation 84 then becomes:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \mathbf{A} \; \mathbf{R} \; (\, \mathbf{I} \, | - \bar{P}_0 \,) U \; = \; \mathbf{A} \; \mathbf{R} \begin{pmatrix} 1 & 0 & 0 & -\bar{X}_0 \\ 0 & 1 & 0 & -\bar{Y}_0 \\ 0 & 0 & 1 & -\bar{Z}_0 \end{pmatrix} U \tag{87}$$

3.3 Camera Calibration

The projection equation has been split up into various geometrical and physical components:

- $\mathbf{A}$ Internal camera parameters
- $\mathbf{R}$ Camera rotation
- $\bar{P}_0$ Camera position (projection point)

It is helpful to consider the coordinates of all points in a camera centered coordinate system, with origin in the projection point P_0 and rotation relative the world system described by a 3×3 matrix $\mathbf{R}$ whose elements are functions of the three rotation angles.

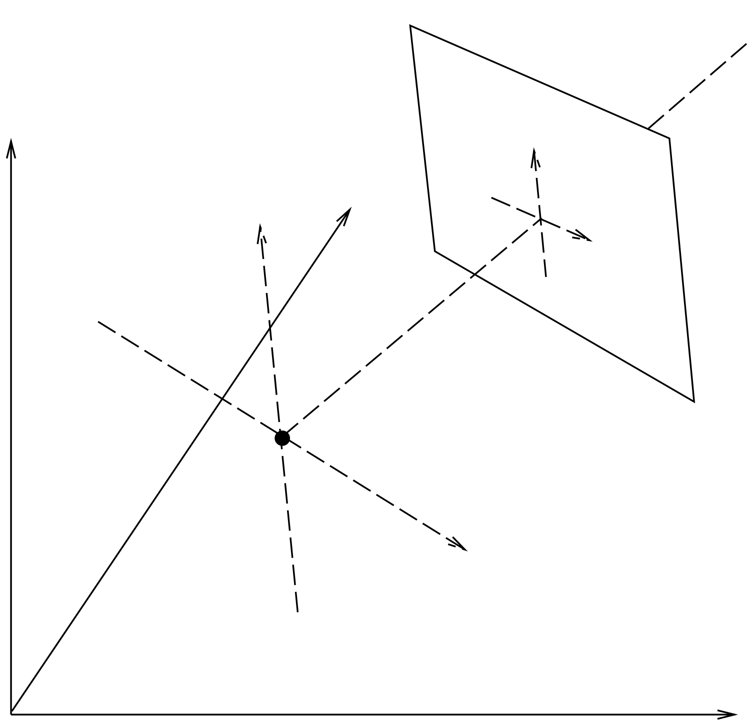

Figure 9: *Camera and world coordinate systems*

The process of *camera calibration* consists of determining these parameters using a set of known points P_i and their corresponding image points p_i. This problem is in general non-linear. It is instructive to note that the parameters of the matrix $\mathbf{A}$ depend only on the orientation of the image plane relative the camera coordinate system. These are therefore known as *internal orientation* parameters. The parameters $\mathbf{R}$ and P_0 depend on the position of the camera relative to the world coordinate system. They are therefore called *external orientation parameters.* Note that the effect of a camera rotation on the image coordinates is just a linear transformation which is independent of the point U in space. This is in contrast to the camera position P_0. A change of P_0 results in a transformation of the image coordinates that is highly dependent on the coordinate of the space point U.

Once the internal parameters, i.e. the matrix $\mathbf{A}$ have been determined, we can change the image coordinate system by a linear transformation:

$$\begin{pmatrix} x' \\ y' \\ w' \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} x \\ y \\ w \end{pmatrix} = \mathbf{R} \; (\mathbf{I}| - \bar{P}_0) \; U \tag{88}$$

These will be called *normalized image coordinates* fig 9 and they can be seen as the image coordinates in a system with othogonal axis and origin at the principal point of the image plane, i.e. the point where a line from P_0 cuts the image plane orthogonally.

In the camera system, the point U will have cartesian coordinates:

$$\bar{U}' = \mathbf{R} \; (\mathbf{I}| - \bar{P}_0) \; U = \mathbf{R}(\bar{U} - \bar{P}_0) \tag{89}$$

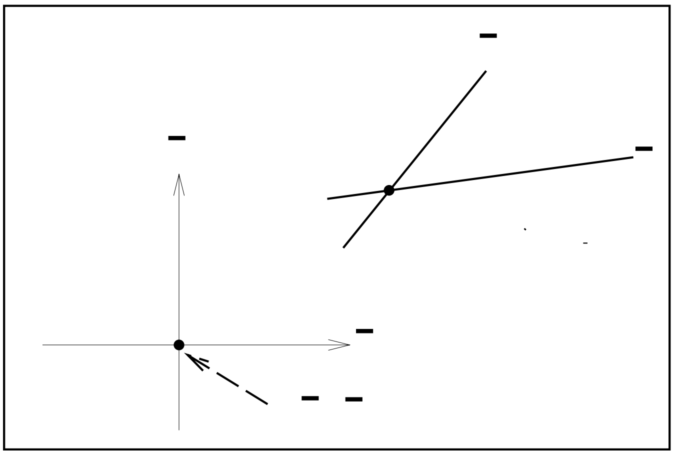

Figure 10: *Normalized coordinate system and original image system*

with components: $\bar{X}', \bar{Y}', \bar{Z}'$ Using these, we can write the projection as:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \lambda \begin{pmatrix} \alpha_x & \gamma & \bar{x}_0 \\ 0 & \alpha_y & \bar{y}_0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{X}' \\ \bar{Y}' \\ \bar{Z}' \end{pmatrix} \tag{90}$$

For points in the plane $\bar{Z}' = 0$ we get $w = 0$. I.e. all points in this plane, known as the *focal plane* project to points at infinity in the image plane. The focal plane is therefore parallel to the image plane and the Z'-axis is orthogonal to it. This is known as the *optical axis* of the camera, and its intersection with the image plane is known as *the principal point* Note that this interpretation is not possible unless we have defined an affine image coordinate system, since otherwise the points at infinity $(x, y, 0$ of the image coordinate system will be finite points in the image plane, contained in the line joining P_1 and P_2. This would have the implication that the focal plane would intersect the image plane.

Defining the cartesian normalized image coordinates:

$$\bar{x}' = \frac{\bar{X}'}{\bar{Z}'} \qquad \bar{y}' = \frac{\bar{Y}'}{\bar{Z}'} \tag{91}$$

we get for the cartesian image coordinates in the original system:

$$\begin{aligned} \bar{x} &= \alpha_x \bar{x}' + \gamma \bar{y}' + \bar{x}_0 \\ \bar{y} &= \alpha_y \bar{y}' + \bar{y}_0 \end{aligned} \tag{92}$$

From this we see that the point x_0, y_0 is just the coordinates the point $0, 0$ in the normalized image system, i.e. the point of intersection of the optical axis, i.e. Z'-axis with the image plane. α_x, and α_y are scale factors associated with the transformation from the normalized image coordinates to the original and γ measures the deviation from orthogonality of the original coordinate axis in the normalized system.

The internal and external parameters of the camera can now be determined in the following way:

- Find the 11 parameters of the projection matrix **P** using at least 6 known points in space and their image coordinates.
- Find the unique factorization 85. This gives the internal parameters of the matrix **A** and the external parameters of the camera rotation and position

Alternatively the parameters can be determined directly without computing the projection matrix first. This will require the use of constrained optimization since the parameters of the rotation matrix are nor arbitrary.

3.4 Projection of Lines

Given a pin hole camera model, a line in 3-D space will project to a line in the image. Using the projection equation 80 we can relate the the Plücker coordinates of the line in 3-D to the line coordinates in the image. Suppose that U_1 and U_2 are two points on the line in space. They project to image points with coordinates (x_1, y_1, w_1) and (x_2, y_2, w_2) respectively Form the projection equation 80 we then have:

$$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ w_1 & w_2 \end{pmatrix} = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{pmatrix} \begin{pmatrix} X_1 & X_2 \\ Y_1 & Y_2 \\ Z_1 & Z_2 \\ W_1 & W_2 \end{pmatrix} \tag{93}$$

Using the factorised form 87 of the projection matrix we can write this as:

$$\begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ w_1 & w_2 \end{pmatrix} = \mathbf{A}\,\mathbf{R} \begin{pmatrix} 1 & 0 & 0 & -\bar{X}_0 \\ 0 & 1 & 0 & -\bar{Y}_0 \\ 0 & 0 & 1 & -\bar{Z}_0 \end{pmatrix} \begin{pmatrix} X_1 & X_2 \\ Y_1 & Y_2 \\ Z_1 & Z_2 \\ W_1 & W_2 \end{pmatrix} \tag{94}$$

The image line coordinates are the minors of the rectangular left hand matrix. Using the rule for the determinant of a product of rectangular matrices and the linear transformation of minors, (see appendix), we get:

$$\begin{pmatrix} \begin{bmatrix} y_1 & y_2 \\ w_1 & w_2 \end{bmatrix} \\ \begin{bmatrix} w_1 & w_2 \\ x_1 & x_2 \end{bmatrix} \\ \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} \end{pmatrix} = (\mathbf{AR})^{-T}\,[\mathbf{AR}] \begin{pmatrix} 1 & 0 & 0 & 0 & -\bar{Z}_0 & \bar{Y}_0 \\ 0 & 1 & 0 & \bar{Z}_0 & 0 & -\bar{X}_0 \\ 0 & 0 & 1 & -\bar{Y}_0 & \bar{X}_0 & 0 \end{pmatrix} \begin{pmatrix} \begin{bmatrix} Y_1 & Y_2 \\ Z_1 & Z_2 \end{bmatrix} \\ \begin{bmatrix} Z_1 & Z_2 \\ X_1 & X_2 \end{bmatrix} \\ \begin{bmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{bmatrix} \\ \begin{bmatrix} X_1 & X_2 \\ W_1 & W_2 \end{bmatrix} \\ \begin{bmatrix} Y_1 & Y_2 \\ W_1 & W_2 \end{bmatrix} \\ \begin{bmatrix} Z_1 & Z_2 \\ W_1 & W_2 \end{bmatrix} \end{pmatrix} \tag{95}$$

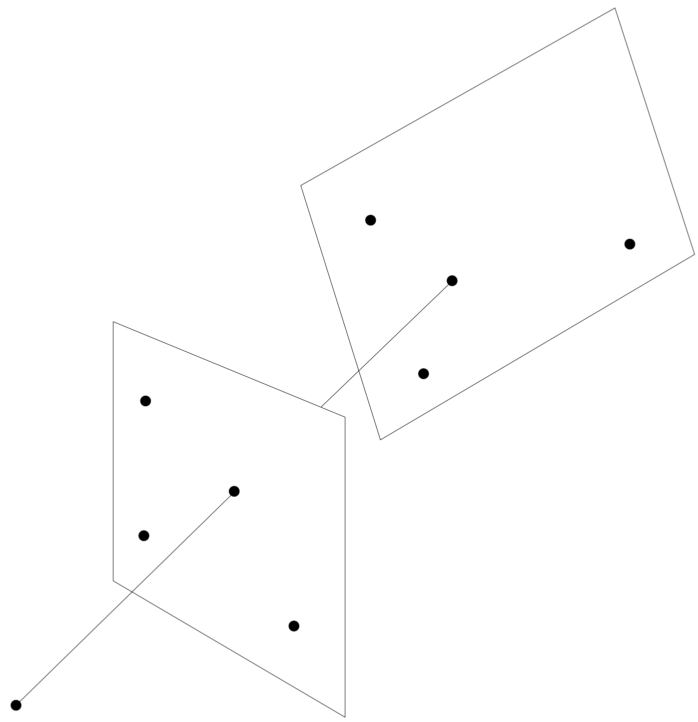

Figure 11: *Perspective mapping from a plane in space*

Given that we know a set of lines in space and their corresponding lines in the image, we can use this equation to find the interior and exterior orientation parameters of the camera. In the same way as for points, this will in general lead to a non linear problem.

3.5 Mapping From a planar surface

Suppose that the point U is confined to a plane in 3-D. The projection equation 80 then maps points from one plane to another. If we assign coordinates to the plane in space in the same way as we assigned image coordinates to the image plane, using points U_1^*, U_2^*, U_3^* as a basis for the subspace representing the plane in space, we can express the point U as:

$$U \;=\; x_u U_1^* \;+\; y_u U_2^* + \; w_u U_3^* \tag{96}$$

If we denote image coordinates with x_a, y_a, w_a, we have using eq. 80:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} \;=\; \mathbf{P}\, U \;=\; \mathbf{P}\, (U_1^*, U_2^*, U_3^*) \;\begin{pmatrix} x_u \\ y_u \\ w_u \end{pmatrix} \tag{97}$$

Since $\mathbf{P}$ is a 4×3 and (U_1^*, U_2^*, U_3^*) is a 3×4 matrix, their product will be 3×3. We can therefore write:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} \;=\; \mathbf{T}_{au} \begin{pmatrix} x_u \\ y_u \\ w_u \end{pmatrix} \tag{98}$$

where T is a 3×3 matrix. This is an example of a *projective transformation.* It will in general be invertible

3.6 Mapping from the plane at infinity

A special case of mapping from a planar surface is when the plane is at infinity. In that case we can take the basis points as:

$$U_1^* = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad U_2^* = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad U_3^* = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \tag{99}$$

These vectors can be used as a basis for all points with homogeneous coordinates $(X, Y, Z, 0)$, i.e. points that are "infinitely" far away. The 3-vector (X, Y, Z) then gives the direction to the point in the cartesian coordinate system.

And the linear transformation mapping coordinates in this basis to image coordinates becomes:

$$\mathbf{T}_\infty = \mathbf{A}\ \mathbf{R} \begin{pmatrix} 1 & 0 & 0 & -\bar{X}_0 \\ 0 & 1 & 0 & -\bar{Y}_0 \\ 0 & 0 & 1 & -\bar{Z}_0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \mathbf{A}\ \mathbf{R} \tag{100}$$

Due to the "large" distance to the points the mapping is independent of the position of the camera. The image coordinates will depend on the direction $X_\infty, Y_\infty, Z_\infty$ to the point:

$$U = X_\infty U_1^* + Y_\infty U_2^* + Z_\infty U_3^* \tag{101}$$

as:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} = \mathbf{A}\ \mathbf{R} \begin{pmatrix} X_\infty \\ Y_\infty \\ Z_\infty \end{pmatrix} \tag{102}$$

3.7 Affine Image Coordinates

For specific selections of the basis points in the image plane the internal parameters in the matrix $\mathbf{A}$ can be given a geometric interpretation.

For a certain image plane the choice of the three points P_1, P_2, P_3 is not unique. Using the expansion:

$$P = xP_1^* + yP_2^* + wP_3^* \tag{103}$$

we see that points P_1, P_2, P_3 will be represented by homogeneous image coordinates:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{104}$$

respectively. I.e. points P_1 and P_2 will be points at infinity in the image coordinate system. This is not in accordance with how an image coordinate system is constructed in an image in general. It is conventional to choose an affine image coordinate system by selecting a certain point $\bar{Q}_3$ as the origin and using two points $\bar{Q}_1$ and $\bar{Q}_2$ to define two coordinate axis. A point $\bar{P}$ in the image can then be expressed as:

$$\bar{P} = \bar{x}(\bar{Q}_1 - \bar{Q}_3) + \bar{y}(\bar{Q}_2 - \bar{Q}_3) + \bar{Q}_3 \tag{105}$$

using homogeneous coordinates, this can be written as:

$$P = W \begin{pmatrix} \bar{P} \\ -- \\ 1 \end{pmatrix} = x \begin{pmatrix} \bar{Q}_1 - \bar{Q}_3 \\ -- \\ 0 \end{pmatrix} + y \begin{pmatrix} \bar{Q}_2 - \bar{Q}_3 \\ -- \\ 0 \end{pmatrix} + w \begin{pmatrix} \bar{Q}_3 \\ -- \\ 1 \end{pmatrix} \tag{106}$$

We see that this corresponds to the choices:

$$P_1^* = \begin{pmatrix} \bar{Q}_1 - \bar{Q}_3 \\ -- \\ 0 \end{pmatrix} \quad P_2^* = \begin{pmatrix} \bar{Q}_2 - \bar{Q}_3 \\ -- \\ 0 \end{pmatrix} \quad P_3^* = \begin{pmatrix} \bar{Q}_3 \\ -- \\ 1 \end{pmatrix} \tag{107}$$

I.e. the basis points P_1 and P_2 are chosen as points at infinity.

3.8 Parallel projection

Instead of considering points in space "far away" we can move the projection point of the camera away from the image plane. In the limit the projection point will be a point at infinity, with homogeneous coordinates:

$$\begin{pmatrix} \hat{P}_0 \\ -- \\ 0 \end{pmatrix} \tag{108}$$

where P_0 denotes the direction to the projection point.

If we have an affine image coordinate system with basis points:

$$P_1 = \begin{pmatrix} \hat{P}_1 \\ -- \\ 0 \end{pmatrix} \quad P_2 = \begin{pmatrix} \hat{P}_2 \\ -- \\ 0 \end{pmatrix} \quad P_3 = \begin{pmatrix} \bar{P}_3 \\ -- \\ 1 \end{pmatrix} \tag{109}$$

i.e. P_1 and P_2 are points at infinity, the matrix $\mathbf{P}$ in the projection equation:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} = \begin{pmatrix} [P_0 \; P_2^* \; P_3^* \; U] \\ \\ [P_0 \; P_1^* \; P_3^* \; U] \\ \\ [P_0 \; P_2^* \; P_1^* \; U] \end{pmatrix} = \mathbf{P} \; \mathbf{U} \tag{110}$$

will have a special structure.

The last row of $\mathbf{P}$ is given by the minors of: the 4×3 matrix:

$$(P_0 \; P_2^* \; P_1^*) = \begin{pmatrix} \hat{P}_0 & \hat{P}_1 & \bar{P}_2 \\ -- & -- & -- \\ 0 & 0 & 0 \end{pmatrix} \tag{111}$$

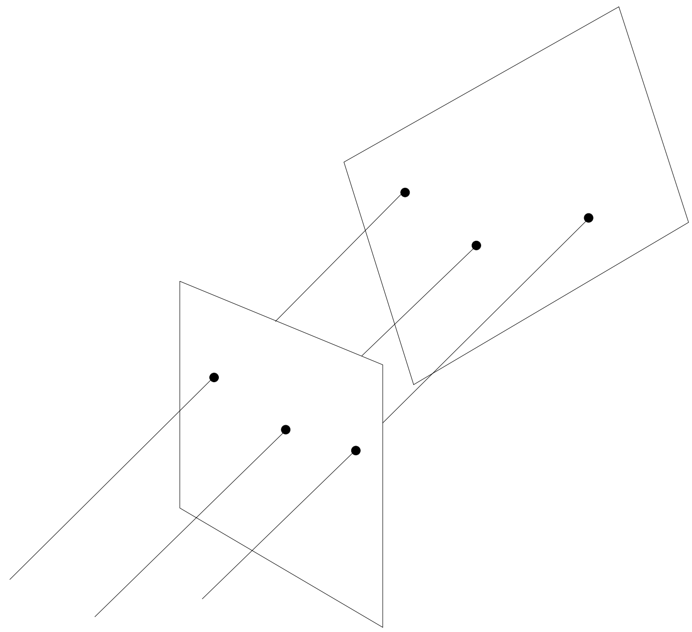

Figure 12: *Parallel projection induces an affine mapping between two planes*

It can be computed as:

$$(0, \quad 0, \quad 0, \quad [\hat{P}_0, \hat{P}_2, \hat{P}_3]) \tag{112}$$

P therefore has the structure:

$$\mathbf{P} \;=\; \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \end{pmatrix} \tag{113}$$

If the point U is contained in a plane spanned by U_1^*, U_2^*, U_3^* we have:

$$U \;=\; x_u U_1^* \;+\; y_u U_2^* +\; z_u U_3^* \tag{114}$$

We get an affine image coordinate system by choosing

$$U_1 \;=\; \begin{pmatrix} \hat{U}_1 \\ -- \\ 0 \end{pmatrix} \quad U_2 \;=\; \begin{pmatrix} \hat{U}_2 \\ -- \\ 0 \end{pmatrix} \quad U_3 \;=\; \begin{pmatrix} \bar{U}_3 \\ -- \\ 1 \end{pmatrix} \tag{115}$$

The mapping from the plane U_1^*, U_2^*, U_3^* to the image plane can then be written:

$$\begin{pmatrix} x \\ y \\ w \end{pmatrix} \;=\; \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & p_{34} \end{pmatrix} \begin{pmatrix} \hat{U}_1 & \hat{U}_2 & \bar{U}_3 \\ -- & -- & -- \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_u \\ y_u \\ w_u \end{pmatrix} \;=\; \begin{pmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ 0 & 0 & t_{33} \end{pmatrix} \begin{pmatrix} x_u \\ y_u \\ w_u \end{pmatrix}$$

which is an affine transformation between the two planes. The affine transformations therefore play very much the same role for parallel projections as the general linear transformations do for general perspective projections. In general parallel projections will lead to simplified relations and algorithms for various problems.

4 Epipolar Geometry

4.1 Planar surfaces - Projective transformations

If we have a second image plane given by the points (P_1^b, P_2^b, P_3^b) and a projection point P_0^b we can express a point P^b in that plane using image coordinates:

$$P^b \;=\; x_b P_1^b \;+\; y_b P_2^b + \; w_b P_3^b \tag{116}$$

These image coordinates can be related to the image coordinates of plane U via a perspective transformation, in the same way as previously:

$$\begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} \;=\; \mathbf{T}_{bu} \begin{pmatrix} x_u \\ y_u \\ w_u \end{pmatrix} \tag{117}$$

If we combine eq. 98 and 117 we can relate the image coordinates of image a and b as:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} \;=\; \mathbf{T}_{au}\,\mathbf{T}_{bu}^{-1} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} \tag{118}$$

We get the important result that the two images of a planar surface are related via a linear or *projective* transformation.

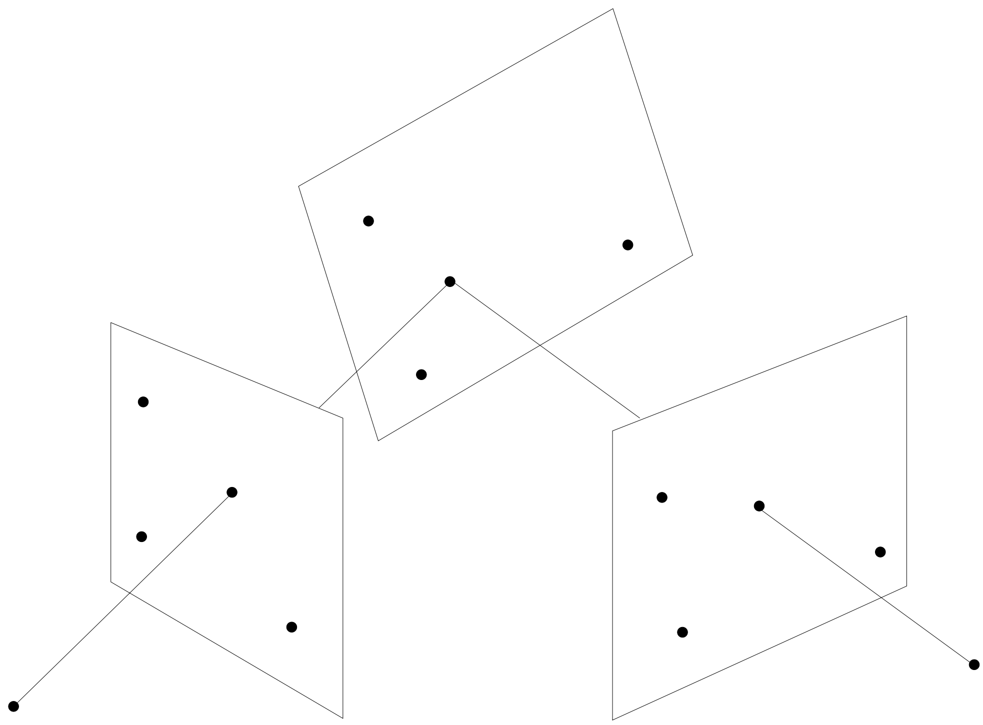

Figure 13: *A planar surface induces a projective transformation between image planes a and b*

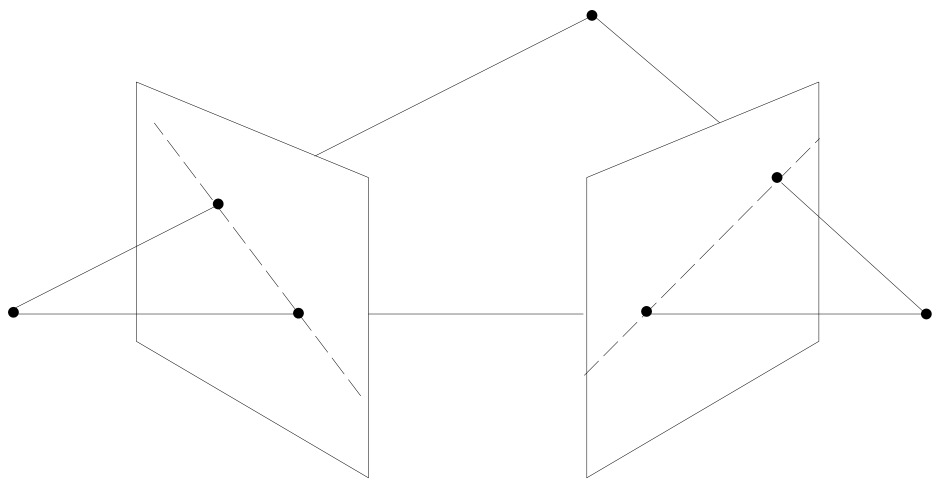

Figure 14: *Epipolar constraint: Points A_0, B_0, A, B are coplanar*

4.2 The Epipolar Constraint - The Fundamental Matrix

In the general case when points in space are not confined to a plane, the relation between image coordinates in different views is more complicated. The fact that they are related can be seen easily if we look at fig. 14. The points A_0, B_0, A and B all lie in the same plane, We therefore have

$$[A_0\ B_0\ A\ B] = 0 \tag{119}$$

Expressing A and B in image coordinates and image plane defining points:

$$\begin{aligned} A &= x_a A_1 + y_a A_2 + w_a A_3 \\ B &= x_b B_1 + y_b B_2 + w_b B_3 \end{aligned} \tag{120}$$

If we use this in eq. 119 we get:

$$(x_a\ y_a\ w_a) \begin{pmatrix} [A_0\ B_0\ A_1\ B_1] & [A_0\ B_0\ A_1\ B_2] & [A_0\ B_0\ A_1\ B_3] \\ [A_0\ B_0\ A_2\ B_1] & [A_0\ B_0\ A_2\ B_2] & [A_0\ B_0\ A_2\ B_3] \\ [A_0\ B_0\ A_3\ B_1] & [A_0\ B_0\ A_3\ B_2] & [A_0\ B_0\ A_3\ B_3] \end{pmatrix} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} =$$

$$= (x_a,\ y_a,\ w_a)\ \mathbf{F} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = 0$$

We see that image coordinates in image a and b are constrained by a bi-linear relation defined by a 3×3 matrix $\mathbf{F}$ which is known as the *fundamental matrix*. The constraint relation is known as the *epipolar constraint* and obviously plays a fundamental role in relating image points across different views.

The fundamental matrix is built up from the image plane defining vectors $A_1 \ldots B_3$ which is also the case for the projection matrices, $\mathbf{P^a}$ and $\mathbf{P^b}$:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} = \mathbf{P^a}\, U \qquad \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = \mathbf{P^b}\, U$$

The fundamental matrix can be related to the projection matrices. In order to do this we write:

$$\mathbf{P^a} = \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \end{pmatrix} \qquad \mathbf{P^b} = \begin{pmatrix} b_1^T \\ b_2^T \\ b_3^T \end{pmatrix} \tag{121}$$

where a_i and b_i are the 4 dimensional row vectors in the 3×4 matrices $\mathbf{P^a}$ and $\mathbf{P^b}$ respectively.

$$\begin{aligned} x_a &= a_1^T\, U & x_b &= b_1^T\, U \\ y_a &= a_2^T\, U & y_b &= b_2^T\, U \\ w_a &= a_3^T\, U & w_b &= b_3^T\, U \end{aligned} \tag{122}$$

These equations can be combined to give:

$$\begin{aligned} (x_a\, a_3^T - w_a\, a_1^T)\, U &= 0 \\ (y_a\, a_3^T - w_a\, a_2^T)\, U &= 0 \\ (x_b\, b_3^T - w_b\, b_1^T)\, U &= 0 \\ (y_b\, b_3^T - w_b\, b_2^T)\, U &= 0 \end{aligned} \tag{123}$$

This is a system of four equations for the four-vector U. Since it is homogeneous, the system determinant must vanish, and we get:

$$\begin{bmatrix} (x_a\, a_3^T - w_a\, a_1^T) \\ (y_a\, a_3^T - w_a\, a_2^T) \\ (x_b\, b_3^T - w_b\, b_1^T) \\ (y_b\, b_3^T - w_b\, b_2^T) \end{bmatrix} = [x_a\, a_3 - w_a\, a_1,\; y_a\, a_3 - w_a\, a_2,\; x_b\, b_3 - w_b\, b_1,\; y_b\, b_3 - w_b\, b_2] =$$

$$= (x_a, y_a, w_a) \begin{pmatrix} [a_2 a_3 b_2 b_3] & -[a_2 a_3 b_1 b_3] & [a_2 a_3 b_1 b_2] \\ -[a_1 a_3 b_2 b_3] & [a_1 a_3 b_1 b_3] & -[a_1 a_3 b_1 b_2] \\ [a_1 a_2 b_2 b_3] & -[a_1 a_2 b_1 b_3] & [a_1 a_2 b_1 b_2] \end{pmatrix} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} =$$

$$= (x_a,\ y_a,\ w_a)\ \mathbf{F} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = 0 \tag{124}$$

This quadratic form in the image coordinates is identical to that in eq. 119. The matrix is just the fundamental matrix expressed in terms of the row vectors $a_1 \ldots b_3$ of the projection matrices.

4.3 Epipolar points, lines and planes

The fundamental matrix is a representation of the relative positions of two cameras. The bilinear form of the epipolar constraint eq. 119 means that it is linear in both image coordinates. If we define the vector,

$$l_b^T = (x_a, y_a, w_a)\ \mathbf{F} \tag{125}$$

we have:

$$l_b^T \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = 0 \tag{126}$$

This is the equation of a line in the image plane b and it is called the *epipolar line.*The fundamental matrix and a point in image a, defines a line in image b, such that the corresponding point in image b is constrained to lie on that line. Correspondingly we can define an epipolar line in image a using the image point in image b and the fundamental matrix. In this way, knowledge of the fundamental matrix helps to solve the problem of matching points in image a and image b. The five points A_0 B_0 A B and U all lie in the same plane, called the *epipolar plane.* The epipolar lines have the geometric interpretation as the intersection of the epipolar plane with the image planes a and b. (see. fig. 14

It is clear from fig. 14 that all epipolar lines of an image intersect in a common point called the *epipolar point.* This point is the intersection of the line connecting the projection points A_0 and B_0 with the respective image plane. If we denote the epipolar point in image a as A_e, we have for an arbitrary point P

$$[A_0\ B_0\ A_e\ P] = 0 \tag{127}$$

since the points A_0 B_0 A_e all lie on the same line.

If we express the epipolar point in image coordinates,

$$A_e = x_a^e A_1 + y_a^e A_2 + w_a^e A_3 \tag{128}$$

and use eq. 127 taking the arbitrary point P as B_1, B_2 and B_3 respectively, we get:

$$x_b^e\ [A_0\ B_0\ A_1\ B_i] + y_b^e\ [A_0\ B_0\ A_2\ B_i] + w_b^e\ [A_0\ B_0\ A_3\ B_i] = 0 \quad i = 1, 2, 3 \tag{129}$$

which can be expressed as:

$$\mathbf{F}^T \begin{pmatrix} x_a^e \\ y_a^e \\ w_a^e \end{pmatrix} = 0 \tag{130}$$

In very much the same way, we can derive

$$\mathbf{F}\begin{pmatrix} x_b^e \\ y_b^e \\ w_b^e \end{pmatrix} = 0 \tag{131}$$

The epipolar points in images a and b are the nullspaces of the matrices $\mathbf{F}^T$ and $\mathbf{F}$ respectively. This means that the fundamental matrix must be singular, i.e.

$$[\mathbf{F}\ ] = 0 \tag{132}$$

4.4 Computing the Fundamental matrix

The fundamental matrix $\mathbf{F}$ is 3×3 and therefore has 9 elements. Since it is homogeneous, the elements are only defined up to an arbitrary scale factor. The fundamental matrix is therefore determined by 8 parameters.

For each pair of matched points, the epipolar constraint 124 gives a linear equation in the unknown F-matrix elements. For n matched points we get the linear system:

$$\begin{aligned} x_a^1 x_b^1 f_{11} + x_a^1 y_b^1 f_{12} + \ldots + w_a^1 w_b^1 f_{33} &= 0 \\ x_a^2 x_b^2 f_{11} + x_a^2 y_b^2 f_{12} + \ldots + w_a^2 w_b^2 f_{33} &= 0 \\ &\vdots \\ x_a^n x_b^n f_{11} + x_a^n y_b^n f_{12} + \ldots + w_a^n w_b^n f_{33} &= 0 \end{aligned} \tag{133}$$

Given 8 points we can therefore in general compute the F-matrix using just linear equations. However, if we use the non-linear singularity constraint, $[\mathbf{F}\ ] = 0$ we need only 7 points at the price of a more complicated algorithm.

4.5 Paralell projection - linear constraints

In the case of paralell projection the fundamental matrix and the epipolar constraint can be simplified considerably.In the paralell projection case the projection points of the two cameras are at infinity and we have using an affine image coordinate system:

$$\begin{aligned} A_0 &= \begin{pmatrix} \hat{A}_0 \\ -- \\ 0 \end{pmatrix} & A_1 &= \begin{pmatrix} \hat{A}_1 \\ -- \\ 0 \end{pmatrix} & A_2 &= \begin{pmatrix} \hat{A}_2 \\ -- \\ 0 \end{pmatrix} \\ B_0 &= \begin{pmatrix} \hat{B}_0 \\ -- \\ 0 \end{pmatrix} & B_1 &= \begin{pmatrix} \hat{B}_1 \\ -- \\ 0 \end{pmatrix} & B_2 &= \begin{pmatrix} \hat{B}_2 \\ -- \\ 0 \end{pmatrix} \end{aligned} \tag{134}$$

we get that:

$$[A_0\ B_0\ A_1\ B_1] = [A_0\ B_0\ A_1\ B_2] = [A_0\ B_0\ A_2\ B_2] = [A_0\ B_0\ A_2\ B_1] = 0$$

And the fundamental matrix takes on the simple form:

$$\mathbf{F} = \begin{pmatrix} 0 & 0 & [A_0\ B_0\ A_1\ B_3] \\ 0 & 0 & [A_0\ B_0\ A_2\ B_3] \\ [A_0\ B_0\ A_3\ B_1] & [A_0\ B_0\ A_3\ B_2] & [A_0\ B_0\ A_3\ B_3] \end{pmatrix}$$

The epipolar constraint will therefore be linear in the cartesian image coordinates:

$$(x_a,\ y_a,\ w_a)\ \mathbf{F} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} =$$

$$= x_a w_b f_{13} + y_a w_b f_{23} + w_a x_b f_{31} + w_a y_b f_{32} + w_a w_b f_{33} =$$

$$= \bar{x}_a f_{13} + \bar{y}_a f_{23} + \bar{x}_b f_{31} + \bar{y}_b f_{32} + f_{33} = 0$$

If we have three cameras, a, b and c we can use this to write the coordinates of a point in one image as a linear combination of the coordinates of the point in the other two. Suppose we have determined the fundamental matrices between cameras a and b, $\mathbf{F}^{ab}$ and cameras a and c, $\mathbf{F}^{ac}$. We can then write the linear system of equations:

$$\bar{x}_a f_{13}^{ab} + \bar{y}_a f_{23}^{ab} + \bar{x}_b f_{31}^{ab} + \bar{y}_b f_{32}^{ab} + f_{33}^{ab} = 0$$

$$\bar{x}_a f_{13}^{ac} + \bar{y}_a f_{23}^{ac} + \bar{x}_c f_{31}^{ac} + \bar{y}_c f_{32}^{ac} + f_{33}^{ac} = 0$$

If we have measured the coordinates in images b and c, we can solve for x_a and y_a and the solution will be a linear combination of the coordinates, x_b, y_b, x_c, y_c. This means that for point sets and paralell projection, any view can be written as a linear combination of two arbitrary views.

4.6 Calibrated Cameras - The Essential Matrix

In the preceding chapter we saw that using an arbitrary coordinate system, the projection equation can be written:

$$\begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} = \mathbf{A}\ \mathbf{R_a}\ (\mathbf{I} \ \mid\ -\bar{A}_0)\ U \qquad (135)$$

Likewise we have for camera b:

$$\begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = \mathbf{B}\ \mathbf{R_b}\ (\mathbf{I} \ \mid\ -\bar{B}_0)\ U \qquad (136)$$

Multiplying both sides of these equations we can write them as :

$$(\mathbf{A}\ \mathbf{R_a})^{-1} \begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} = (\mathbf{I} \ \mid\ -\bar{A}_0)\ U \qquad (\mathbf{B}\ \mathbf{R_b})^{-1} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = (\mathbf{I} \ \mid\ -\bar{B}_0)\ U (137)$$

Let $\mathbf{F_0}$ be the fundamental matrix associated with projection matrices $(\mathbf{I} \mid -\bar{A}_0)$ and $(\mathbf{I} \mid -\bar{B}_0)$ Note that it depends only on the viewpoints A_0 B_0

Denote:

$$p_a{}^T = (x_a,\ y_a,\ w_a) \qquad p_b{}^T = (x_b,\ y_b,\ w_b) \tag{138}$$

We have the epipolar constraint:

$$((\mathbf{AR_a})^{-1}p_a)^T\ \mathbf{F_0}\ (\mathbf{BR_b})^{-1}p_b = 0 \tag{139}$$

which can be written as:

$$p_a{}^T\mathbf{A}^{-1^T}(\mathbf{R_a}^{-1})^T\ \mathbf{F_0}\ \mathbf{R_b}^{-1}\mathbf{B}^{-1}p_b = 0 \tag{140}$$

or

$$p_a{}^T\mathbf{A}^{-1^T}\mathbf{R_a}\ \mathbf{F_0}\ \mathbf{R_b}^{-1}\mathbf{B}^{-1}p_b = 0 \tag{141}$$

Where we have made use of the fact that $(\mathbf{R_a}^{-1})^T = \mathbf{R_a}$ since $\mathbf{R_a}$ is orthogonal.

We obviously have :

$$\mathbf{F} = \mathbf{A}^{-1^T}\mathbf{R_a}\ \mathbf{F_0}\ \mathbf{R_b}^{-1}\mathbf{B}^{-1} \tag{142}$$

where $\mathbf{F}$ is the F-matrix of the original uncalibrated system. If we use the first camera as the reference frame, we have $\mathbf{R_a} = \mathbf{I}$, $\bar{A}_0 = (0,0,0)^T$ $\mathbf{R_b} = \mathbf{R}$ and $\bar{B}_0 = (X_0, Y_0, Z_0)^T$. The F-matrix $\mathbf{F_0}$ is now built from the projection matrices:

$$(\mathbf{I} \mid -\bar{A}_0) = \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad (\mathbf{I} \mid -\bar{B}_0) = \begin{pmatrix} b_1^T \\ b_2^T \\ b_3^T \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & \bar{X}_0 \\ 0 & 1 & 0 & \bar{Y}_0 \\ 0 & 0 & 1 & \bar{Z}_0 \end{pmatrix}$$

and can be computed acc. to eq. 124 as:

$$\mathbf{F_0} = \begin{pmatrix} [a_2a_3b_2b_3] & -[a_2a_3b_1b_3] & [a_2a_3b_1b_2] \\ -[a_1a_3b_2b_3] & [a_1a_3b_1b_3] & -[a_1a_3b_1b_2] \\ [a_1a_2b_2b_3] & -[a_1a_2b_1b_3] & [a_1a_2b_1b_2] \end{pmatrix} = \begin{pmatrix} 0 & -\bar{Z}_0 & \bar{Y}_0 \\ \bar{Z}_0 & 0 & -\bar{X}_0 \\ -\bar{Y}_0 & \bar{X}_0 & 0 \end{pmatrix}$$

and we get:

$$\mathbf{F} = \mathbf{A}^{-1^T}\ \mathbf{F_0}\ \mathbf{R}^{-1}\ \mathbf{B}^{-1} = \mathbf{A}^{-1^T}\ \mathbf{E}\ \mathbf{B}^{-1}$$

where the matrix:

$$\mathbf{E} = \mathbf{F_0}\ \mathbf{R}^{-1}$$

is called the *essential matrix*

For uncalibrated cameras with arbitrary image coordinate systems, the image coordinates are related by the epipolar constraint

$$(x_a,\ y_a,\ w_a)\ \mathbf{F} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} = 0 \tag{143}$$

If the cameras are calibrated, we can use normalized image coordinates

$$\begin{pmatrix} x'_a \\ y'_a \\ w'_a \end{pmatrix} = \mathbf{A}^{-1} \begin{pmatrix} x_a \\ y_a \\ w_a \end{pmatrix} \qquad \begin{pmatrix} x'_b \\ y'_b \\ w'_b \end{pmatrix} = \mathbf{B}^{-1} \begin{pmatrix} x_b \\ y_b \\ w_b \end{pmatrix} \tag{144}$$

and get the relation:

$$(x'_a,\ y'_a,\ w'_a)\ \mathbf{E} \begin{pmatrix} x'_b \\ y'_b \\ w'_b \end{pmatrix} = 0 \tag{145}$$

In the same way as for the fundamental matrix, the essential matrix can be computed solving a linear system, given at least 8 points. If we look at the structure of the **E**-matrix however, we see that it depends only on the relative rotation and translation of the two cameras. The rotation is described by three rotation angles. The translation is also described by three parameters, but since the epipolar constraint relation 145 is homogeneous in these parameters, they can only be determined up to an arbitrary scale factor. This means that only the two parameters describing the direction of translation can be determined resulting in a total of five parameters for the **E**-matrix. Using non-linear methods for determination of rotation and translation we therefore only need five points. In this case we will get multiple solutions however.

Given that we have determined relative rotation **R** and direction of translation, i.e. $\bar{P}_0$ up to arbitrary scale, we can use the projection equations and compute the position $\bar{U}$, of the point in 3-D relative to the cameras. Note that since $\bar{P}_0$ can only be determined up to arbitrary scale, the same applies to $\bar{U}$. This can be seen as a consequence of the fact that the image coordinates do not change if both $\bar{P}_0$ and $\bar{U}$ are multiplied by an arbitrary number.

6 Invariants

In the previous section we saw that the 3-D reconstruction from observed image data can only be computed up to a certain linear transformation depending on the calibration and nature of the cameras. There are therefore two questions remaining before we can talk about a reconstruction:

- How do we represent the 3-D reconstruction ?
- How do we compute it ?

These questions are of course intimately related and we start with the first one. Since the reconstruction is only up to a linear transformation the representation must be such that if $U_1, U_2 \ldots U_n$ and $\mathbf{T}U_1, \mathbf{T}U_2 \ldots \mathbf{T}U_n$ are two reconstructions they should have the same representation. It should of course also be possible to compute all possible solutions from the representation. The first requirement says that the representation should be *invariant* w.r.t the transformation $\mathbf{T}$ and the second requirement says that it should be *complete*. We are therefore looking for functions $I_1 \ldots I_k$ such that

$$I_k(U_1 \ldots U_n) \;=\; I_k(\mathbf{T}U_1 \ldots \mathbf{T}U_n) \tag{165}$$

with the property that $(U_1 \ldots U_n)$ can be computed from the values of $I_1 \ldots I_k$.

6.1 Distances, Areas and Ratios

Depending on the subgroup we are considering, we will have different invariants. The invariants of a certain group are of course also invariants of all its subgroups. If we consider the euclidean group in 1-D we only have translations along a line, i.e:

$$\bar{x}' \;=\; \bar{x} \;+\; \bar{t} \tag{166}$$

The distance between two points is easily seen to be an invariant for this transformation, i.e.

$$I_e \;=\; |\bar{x}_1' - \bar{x}_2'| \;=\; |\bar{x}_1 - \bar{x}_2'| \tag{167}$$

If we consider affine transformations in 1-D we have

$$\bar{x}' \;=\; a\bar{x} \;+\; b \tag{168}$$

The distance between two points is transformed as:

$$|\bar{x}_1' - \bar{x}_2'| \;=\; a|\bar{x}_1 - \bar{x}_2'| \tag{169}$$

i.e. it is not invariant since it depends on a. However if we take the ratio of two distances we get:

$$I_a \;=\; \frac{|\bar{x}_1' - \bar{x}_2'|}{|\bar{x}_1' - \bar{x}_3'|} \;=\; \frac{|\bar{x}_1 - \bar{x}_2|}{|\bar{x}_1 - \bar{x}_3|} \tag{170}$$

i.e. an invariant.

The distance between two points can be expressed as a determinant: of homogeneous coordinates:

$$I_e \;=\; |\bar{x}_1 - \bar{x}_2| \;=\; \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \\ 1 & 1 \end{bmatrix} \tag{171}$$

For the affine invariant ratio we get:

$$I_a \;=\; \frac{|\bar{x}_1 - \bar{x}_2|}{|\bar{x}_1 - \bar{x}_3|} \;=\; \frac{\begin{bmatrix} \bar{x}_1 & \bar{x}_2 \\ 1 & 1 \end{bmatrix}}{\begin{bmatrix} \bar{x}_1 & \bar{x}_3 \\ 1 & 1 \end{bmatrix}} \tag{172}$$

Using homogeneous coordinates, we can write the affine transformation as:

$$\begin{pmatrix} \bar{x}' \\ 1 \end{pmatrix} \;=\; \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{x} \\ \bar{1} \end{pmatrix} \tag{173}$$

This is a subgroup of the general linear group.

$$\begin{pmatrix} x' \\ w' \end{pmatrix} \;=\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix} \tag{174}$$

The distance between two points can be expressed as:

$$|\bar{x}_1 - \bar{x}_2| \;=\; \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \\ 1 & 1 \end{bmatrix} \;=\; \frac{1}{w_1 w_2} \begin{bmatrix} x_1 & x_2 \\ w_1 & w_2 \end{bmatrix} \tag{175}$$

Under the general linear transformation, distance is transformed as:

$$|\bar{x}'_1 - \bar{x}'_2| \;=\; \begin{bmatrix} \bar{x}'_1 & \bar{x}'_2 \\ 1 & 1 \end{bmatrix} \;=\; \frac{1}{w'_1 w'_2} \begin{bmatrix} x'_1 & x'_2 \\ w'_1 & w'_2 \end{bmatrix} \;=\; \frac{1}{w'_1 w'_2} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ w_1 & w_2 \end{bmatrix} \;=$$

$$=\; \frac{1}{w'_1 w'_2} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{w_1 w_2} \begin{bmatrix} \bar{x}_1 & \bar{x}_2 \\ 1 & 1 \end{bmatrix} \;=\; \frac{1}{w'_1 w'_2} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \frac{1}{w_1 w_2} |\bar{x}_1 - \bar{x}_2|$$

If we form the double ot *cross ratio* we get:

$$\begin{aligned} I_p \;&=\; \frac{|\bar{x}'_1 - \bar{x}'_2|\,|\bar{x}'_3 - \bar{x}'_4|}{|\bar{x}'_1 - \bar{x}'_3|\,|\bar{x}'_2 - \bar{x}'_4|} \;=\; \frac{\begin{bmatrix} x'_1 & x'_2 \\ w'_1 & w'_2 \end{bmatrix} \begin{bmatrix} x'_3 & x'_4 \\ w'_3 & w'_4 \end{bmatrix}}{\begin{bmatrix} x'_1 & x'_3 \\ w'_1 & w'_3 \end{bmatrix} \begin{bmatrix} x'_2 & x'_4 \\ w'_2 & w'_4 \end{bmatrix}} \;= \\ &=\; \frac{|\bar{x}_1 - \bar{x}_2|\,|\bar{x}_3 - \bar{x}_4|}{|\bar{x}_1 - \bar{x}_3|\,|\bar{x}_2 - \bar{x}_4|} \;=\; \frac{\begin{bmatrix} x_1 & x_2 \\ w_1 & w_2 \end{bmatrix} \begin{bmatrix} x_3 & x_4 \\ w_3 & w_4 \end{bmatrix}}{\begin{bmatrix} x_1 & x_3 \\ w_1 & w_3 \end{bmatrix} \begin{bmatrix} x_2 & x_4 \\ w_2 & w_4 \end{bmatrix}} \end{aligned} \tag{176}$$

which is invariant over general linear or *projective* transformations. Note that the construction of the cross ratio is such that all scale factors w_i and the determinant of the transformation matrix cancels.

Going from the euclidean to affine and general linear group we get that distance, ratios of distances and cross ratios of distances are invariants. Note that when using homogeneous coordinates to express the invariants we must fix the scale factor in the euclidean and affine cases. In the projective case however, the scale factor is arbitrary.

In the 2-D case with points in the plane we consider the area enclosed by three points with cartesian coordinates $\bar{p}_1, \bar{p}_2$ and $\bar{p}_3$ where $p_i^T = (\bar{x}_i, \bar{y}_i)$. This area can be expressed as:

$$S = |\bar{p}_2 - \bar{p}_1| \; |\bar{p}_3 - \bar{p}_1| \; sin(\alpha) \tag{177}$$

where α is the angle between the lines $p_1 p_2$ and $p_1 p_3$. This is just the vector product of the two vectors $\bar{p}_2 - \bar{p}_1$ and $\bar{p}_3 - \bar{p}_1$, which can be expressed as the determinant:

$$S = [\bar{p}_2 - \bar{p}_1 \; \bar{p}_3 - \bar{p}_1] \tag{178}$$

The euclidean transformation in 2-D is a rotation and translation.

$$\bar{p}' = \mathbf{R}\bar{p} + \bar{t} \tag{179}$$

The area S is then transformed as:

$$\begin{aligned} S' &= [\bar{p}_2' - \bar{p}_1' \; \bar{p}_3' - \bar{p}_1'] = [\mathbf{R}(\bar{p}_2 - \bar{p}_1) \; \mathbf{R}(\bar{p}_3 - \bar{p}_1)] = \\ &= [\mathbf{R}] \, [\bar{p}_2 - \bar{p}_1 \; \bar{p}_3 - \bar{p}_1] = [\bar{p}_2 - \bar{p}_1 \; \bar{p}_3 - \bar{p}_1] = S \end{aligned} \tag{180}$$

where the last equality comes from the fact that $[\mathbf{R}] = 1$ since rotation matrices are orthonormal. Surface area is therefore an invariant under euclidean transformations, just as distances.

Using the linearity properties of determinants we can develop the expression for the surface area as:

$$S = [\bar{p}_2 - \bar{p}_1 \; \bar{p}_3 - \bar{p}_1] = [\bar{p}_2 \; \bar{p}_3] - [\bar{p}_2 \; \bar{p}_1] - [\bar{p}_1 \; \bar{p}_3] = \begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix} \tag{181}$$

The euclidean invariant surface area can therefore be expressed as a determinant of homogeneous coordinates in a similar way as distance:

$$I_e = S = \begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix} \tag{182}$$

The euclidean transform can be written using homogeneous coordinates as:

$$\begin{pmatrix} \bar{p}' \\ -- \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{R} & | & \bar{t} \\ -- & -- & -- \\ \bar{0}^T & | & 1 \end{pmatrix} \begin{pmatrix} \bar{p} \\ -- \\ 1 \end{pmatrix} \tag{183}$$

This is a special case of the affine transformation where:

$$\begin{pmatrix} \bar{p}' \\ -- \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{A} & | & \bar{b} \\ -- & -- & -- \\ \bar{0}^T & | & 1 \end{pmatrix} \begin{pmatrix} \bar{p} \\ -- \\ 1 \end{pmatrix} = \mathbf{T} \begin{pmatrix} \bar{p} \\ -- \\ 1 \end{pmatrix} \tag{184}$$

with $\mathbf{A}$ a general 2×2 matrix.

If the affine transformation is applied to four points $\bar{p}_1 \ldots \bar{p}_4$ we can compute the ratio of surface areas:

$$I_e \;=\; \frac{\begin{bmatrix} \bar{p}'_1 & \bar{p}'_2 & \bar{p}'_3 \\ 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} \bar{p}'_1 & \bar{p}'_2 & \bar{p}'_4 \\ 1 & 1 & 1 \end{bmatrix}} \;=\; \frac{[\mathbf{T}] \begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix}}{[\mathbf{T}] \begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}'_4 \\ 1 & 1 & 1 \end{bmatrix}} \;=\; \frac{\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_4 \\ 1 & 1 & 1 \end{bmatrix}} \tag{185}$$

which is an affine invariant in the same way as ratios of distances in 1-D.

Finally we can consider general linear transformations of point coordinates:

$$p' \;=\; \mathbf{T}p \tag{186}$$

where p is the homogeneous coordinate vector $(x, y, w)^T$ of a point in the plane, and $|bfT$ is a general 3×3 transformation matrix. . Using five points we can form the cross ratio of areas:

$$I_p \;=\; \frac{[p'_1\ p'_2\ p'_5]\ [p'_3\ p'_4\ p'_5]}{[p'_1\ p'_3\ p'_5]\ [p'_2\ p'_4\ p'_5]} \;=\; \frac{[\mathbf{T}]\ [p_1\ p_2\ p_5]\ [\mathbf{T}]\ [p_3\ p_4\ p_5]}{[\mathbf{T}]\ [p_1\ p_3\ p_5]\ [\mathbf{T}]\ [p_2\ p_4\ p_5]} \;=\; \frac{[p_1\ p_2\ p_5]\ [p_3\ p_4\ p_5]}{[p_1\ p_3\ p_5]\ [p_2\ p_4\ p_5]}$$

which is invariant over general linear or projective transformations. Note that the combination of points in the determinants is chosen so that the scale factors of the homogeneous coordinates cancel, i.e. the expression for the cross ratio is independent of the scale factors of the homogeneous coordinates.

The method of expressing euclidean, affine and projective invariants as determinants, ratios and cross ratios of determinants respectively can be generalized to any dimension of space.

6.2 Affine Coordinates

The purpose of introducing invariants was to be able to represent the equivalence classes of points sets that are generated by various linear transformations. In order to do this we shall see how invariants of a certain transformation can be used to generate coordinates for the points that are independent of that transformation.

Given three points in the plane with cartesian coordinates $\bar{p}_1, \bar{p}_2$ and $\bar{p}_3$ we have seen previously that we can use these as a basis for all other points $\bar{p}_n$:

$$\bar{p}_n \;=\; \bar{p}_1 \;+\; \alpha_n(\bar{p}_2 - \bar{p}_1) \;+\; \beta_n(\bar{p}_3 - \bar{p}_1) \tag{187}$$

The coordinates α_n, β_n can be computed by forming the determinants:

$$[\bar{p}_n - \bar{p}_1\ \bar{p}_3 - \bar{p}_1] \;=\; \alpha_n[\bar{p}_2 - \bar{p}_1\ \bar{p}_3 - \bar{p}_1] \;+\; \beta_n[\bar{p}_3 - \bar{p}_1\ \bar{p}_3 - \bar{p}_1] \;=$$

$$=\; \alpha_n[\bar{p}_2 - \bar{p}_1\ \bar{p}_3 - \bar{p}_1]$$

$$[\bar{p}_n - \bar{p}_1\ \bar{p}_2 - \bar{p}_1] \;=\; \alpha_n[\bar{p}_2 - \bar{p}_1\ \bar{p}_2 - \bar{p}_1] \;+\; \beta_n[\bar{p}_3 - \bar{p}_1\ \bar{p}_2 - \bar{p}_1] \;=$$

$$=\; \beta_n[\bar{p}_3 - \bar{p}_1\ \bar{p}_2 - \bar{p}_1]$$

We get:

$$\alpha_n = \frac{[\bar{p}_n - \bar{p}_1,\ \bar{p}_3 - \bar{p}_1]}{[\bar{p}_2 - \bar{p}_1,\ \bar{p}_3 - \bar{p}_1]} = \frac{\begin{bmatrix} \bar{p}_1 & \bar{p}_3 & \bar{p}_n \\ 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix}}$$

$$\beta_n = \frac{[\bar{p}_2 - \bar{p}_1,\ \bar{p}_n - \bar{p}_1]}{[\bar{p}_2 - \bar{p}_1,\ \bar{p}_3 - \bar{p}_1]} = \frac{\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_n \\ 1 & 1 & 1 \end{bmatrix}}{\begin{bmatrix} \bar{p}_1 & \bar{p}_2 & \bar{p}_3 \\ 1 & 1 & 1 \end{bmatrix}} \tag{188}$$

which are actually affine invariants. This means that the coordinates α_n, β_n will not be affected by any affine transformation of the point set. They can therefore be used to characterize the point set in a way that is independent of any affine transformation. The points $\bar{p}_1, \bar{p}_2$ and $\bar{p}_3$ will have affine coordinates $(0,0)^T, (1,0)^T$ and $(0,1)^T$ respectively. By applying an arbitrary affine transformation to the coordinates (α_n, β_n) we obtain a specific member of the affine equivalence class of point sets. The representation (α_n, β_n) is therefore complete in the sense that all information about the point set can be generated from it. These coordinates are called *affine coordinates*.

6.3 Projective Coordinates

Using the homogeneous coordinates p_1, p_2 and p_3 of three points in the plane as a basis, we can express the homogeneous coordinates of any other point in the plane as:

$$p_n = a p_1 + b p_2 + c p_3 \tag{189}$$

In order for this expression to be unique, we know that we have to fix the scale factor of the three basis points. However, using four points , we shall see that it is possible to define a basis in a way that is independent of the specific scale factors. Moreover, we will see that the coordinates in this basis will actually be projective invariants.

In very much the same way as in the affine case we can form determinants by adding points to both sides of this equation and using the fact that a determinant vanishes if the same point occurs twice: Adding points 23, 13 and 12 respectively we get:

$$\begin{aligned} [p_n\ p_2\ p_3] &= a\,[p_1\ p_2\ p_3] \\ [p_n\ p_1\ p_3] &= -b\,[p_1\ p_2\ p_3] \\ [p_n\ p_1\ p_2] &= c\,[p_1\ p_2\ p_3] \end{aligned} \tag{190}$$

Writing $[i\ j\ k]$ for the determinant $[p_i\ p_j\ p_k]$ we have using eq. 189:

$$[1\ 2\ 3]\,p_n = [2\ 3\ n]\,p_1 - [1\ 3\ n]\,p_2 + [1\ 2\ n]\,p_3 \tag{191}$$

Note that this expression is actually valid independent of the scale factors of the homogeneous coordinates.

We now introduce a 4:th point and define the vectors:

$$\begin{aligned} p_1^* &= & - & [2\ 3\ 4\ ]\, p_1 \\ p_2^* &= & & [1\ 3\ 4\ ]\, p_2 \\ p_3^* &= & - & [1\ 2\ 4\ ]\, p_3 \end{aligned} \tag{192}$$

This set of vectors all have a common scale factor in the sense that if any p_i is scaled, all p_i^*'s will be scaled with the same factor. They can therefore be used a basis without the need to consider scale factors. They are called the projective basis for the set of points $p_1 \ldots p_4$. Replacing the p_i's in Eq. **??** with the projective basis p_i^* we get:

$$\begin{aligned} [1\ 2\ 3\ ]\, p_n &= \frac{[2\ 3\ n\ ]}{[2\ 3\ 4\ ]}\, p_1^* \;+\; \frac{[1\ 3\ n\ ]}{[1\ 3\ 4\ ]}\, p_2^* \;+\; \frac{[1\ 2\ n\ ]}{[1\ 2\ 4\ ]}\, p_3^* \;= \\ &= \alpha_n p_1^* \;+\; \beta_n p_2^* \;+\; \gamma_n p_3^* \end{aligned} \tag{193}$$

where $\alpha_n \ldots \gamma_n$ are the homogeneous projective coordinates for the point p_n in the projective basis. The absolute projective coordinates are the three projective invariants:

$$\begin{aligned} I_1 &= \frac{\alpha_n}{\gamma_n} = \frac{[2\ 3\ n\ ]\,[1\ 2\ 4]}{[2\ 3\ 4\ ]\,[1\ 2\ n\ ]} \\ I_2 &= \frac{\beta_n}{\gamma_n} = \frac{[1\ 3\ n\ ]\,[1\ 2\ 4]}{[1\ 3\ 4\ ]\,[1\ 2\ n\ ]} \end{aligned} \tag{194}$$

for $n > 4$. The projective coordinates $\alpha_i \ldots \gamma_i$ will be unit 3- vectors for $i = 1 \ldots 3$ and point 4 can be seen to have projective coordinates $(1, 1, 1, 1)$

p_1	p_2	p_3	p_4	p_n
1	0	0	1	α_n
0	1	0	1	β_n
0	0	1	1	γ_n

A The Determinant

A.1 Definition

The determinant is a real valued function defined on the coordinates of n vectors in n-space, or equivalently an $n \times n$ matrix.

$$det: \ x_1, \ x_2 \ldots x_n \longrightarrow R \qquad x_i \in R^n \tag{202}$$

We will denote the determinant as $[x_1, \ x_2 \ldots x_n]$. It has the following properties.

- If $\mathbf{I}$ is the identity matrix:

$$[\mathbf{I}] \ = \ 1$$

- The determinant is anti-symmetric in the vectors x_i, i.e it changes sign whenever two vectors are permutated:

$$[\ x_1, \ x_2 \ldots x_i \ldots x_j \ldots x_n] \ = \ - \ [\ x_1, \ x_2 \ldots x_j \ldots x_i \ldots x_n]$$

- It is multilinear, i.e. for all vectors x_i:

$$[\ x_1, \ x_2 \ldots \alpha a \ + \ \beta b \ldots x_n] \ = \ \alpha[\ x_1, \ x_2 \ldots a \ \ldots x_n] \ + \ \beta[\ x_1, \ x_2 \ldots b \ \ldots x_n]$$

From these properties we can deduce a general algebraic expression for the determinant in terms of the components $x_{i,j}$ of the vectors x_i. Denoting by e_j the j:th unit vector in R^n we have:

$$x_i \ = \ \sum_{j=1}^{n} x_{i,j} e_j \tag{203}$$

Using the multilinearity property for x_1 we get:

$$[\ x_1, \ x_2 \ldots x_n] \ = \ \sum_j x_{1,j} \ [\ e_j, \ x_2 \ldots x_n] \tag{204}$$

Repeating this for all x_i:

$$[\ x_1, \ x_2 \ldots x_n] \ = \ \sum_{j_1} x_{1,j_1} \ \sum_{j_2} x_{2,j_2} \ \ldots \ \sum_{j_n} x_{n,j_n} \ [\ e_{j_1}, \ e_{j_2} \ldots e_{j_n}] \tag{205}$$

Note that the anti-symmetry property implies that the determinant vanishes whenever two vectors are equal:

$$x_i \ = \ x_j \quad ==> \quad [\ x_1, \ x_2 \ldots x_i \ldots x_j \ldots x_n] \ + \ [\ x_1, \ x_2 \ldots x_j \ldots x_i \ldots x_n] \ =$$

$$= \ 2[\ x_1, \ x_2 \ldots x_i \ldots x_j \ldots x_n] \ = \ 0 \tag{206}$$

This means that of all the determinants $[\ e_{j_1}, \ e_{j_2} \ldots e_{j_n}]$ we only have to consider those where all indices $j_1, j_2 \ldots j_n$ are different, i.e. permutations of the determinant of the unit matrix $\mathbf{I} = [\ e_1, \ e_2 \ldots e_n]$ which will be + or -1. The sign depends on whether the permutation is odd or even.

$$[\ x_1, \ x_2 \ldots x_n] \ = \ \sum_{j_1, j_2 \cdots j_n} x_{1,j_1} \ x_{2,j_2} \ \ldots \ x_{n,j_n} \ sign(j_1, j_2 \ldots j_n) \tag{207}$$

where the sum is over all distinct combinations $j_1, j_2 \ldots j_n$.

A.2 Development Along a Row-vector

There is a convenient way of expressing the determinant of a matrix recursively in terms of linear combinations of sub-determinants,. The determinant $[\ e_j,\ x_2 \ldots x_n]$ in eq, 204 can be written explicitly:

$$[\ e_j,\ x_2 \ldots x_n] \ = \ \begin{bmatrix} 0 & x_{2,1} & \ldots & x_{n,1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{2,j} & \ldots & x_{n,j} \\ \vdots & \vdots & & \vdots \\ 0 & x_{2,n} & \ldots & x_{n_n} \end{bmatrix} \tag{208}$$

Using eq. 207 this determinant can itself be developed as:

$$[\ e_j,\ x_2 \ldots x_n] \ = \ \sum_{j_2 \cdots j_n} x_{2,j_2} \ \ldots \ x_{n,j_n} \ [\ e_j,\ e_{j_2} \ldots e_{j_n}] \ =$$

$$= \ \sum_{j_2 \cdots j_n \neq j} x_{2,j_2} \ \ldots \ x_{n,j_n} \ sign(j, j_2 \ldots j_n) \ =$$

$$= \ (-1)^{j-1} \begin{bmatrix} x_{2,1} & \ldots & x_{n,1} \\ \vdots & & \vdots \\ x_{2,j-1} & \ldots & x_{n,j-1} \\ & & \\ x_{2,j+1} & \ldots & x_{n,j+1} \\ \vdots & & \vdots \\ x_{2,n} & \ldots & x_{n_n} \end{bmatrix} \ = \ (-1)^{j-1}[\mathbf{X}_{1,j}] \tag{209}$$

Where $\mathbf{X}_{1,j}$ is the $n-1 \times n-1$ matrix formed by deleting column 1 and row j from the matrix $\mathbf{X} = (\ x_1,\ x_2 \ldots x_n)$ Eq. 204 can therefore be written as:

$$[\ x_1,\ x_2 \ldots x_n] \ = \ \sum_j (-1)^{j-1} \ x_{1,j} \ [\mathbf{X}_{1,j}] \tag{210}$$

As an example for a 3×3 matrix we have:

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \ = \ a_1 \begin{bmatrix} b_2 & c_2 \\ b_3 & c_3 \end{bmatrix} \ - \ a_2 \begin{bmatrix} b_1 & c_1 \\ b_3 & c_3 \end{bmatrix} \ + \ a_3 \begin{bmatrix} b_1 & c_1 \\ b_2 & c_2 \end{bmatrix}$$

By permutation of the columns we see that the expansion can be made along an arbitrary row vector i.

$$[\ x_1,\ x_2 \ldots x_n] \ = \ \sum_j (-1)^{i+j} \ x_{i,j} \ [\mathbf{X}_{i,j}] \tag{211}$$

Where $\mathbf{X}_{i,j}$ is the $n-1 \times n-1$ matrix formed by deleting column i and row j from the matrix $\mathbf{X}$

A.3 Laplace Expansions

If we use the multilinearity property for x_1 and x_2 we can express the determinant similarly to 204:

$$[\, x_1,\ x_2 \dots x_n] \; = \; \sum_i \; x_{1,i} \sum_j \; x_{2,j} \; [\, e_i \; e_j, \; x_3 \dots x_n] \tag{212}$$

Using the fact that:

$$[\, e_i \; e_j, \; x_3 \dots x_n] \; = \; -[\, e_j \; e_i, \; x_3 \dots x_n] \tag{213}$$

we get:

$$[\, x_1, \; x_2 \dots x_n] \; = \; \sum_{j>i} \begin{bmatrix} x_{1,i} & x_{2,i} \\ x_{1,j} & x_{2,j} \end{bmatrix} [\, e_i \; e_j, \; x_3 \dots x_n] \tag{214}$$

Analogously to eq. 209 we can expand:

$$[\, e_i \; e_j, \; x_3 \dots x_n] \; = \; \sum_{j_3 \cdots j_n} x_{3,j_2} \; \dots \; x_{n,j_n} \; [\, e_i \; e_j, \; e_{j_3} \dots e_{j_n}] \; =$$

$$= \; \sum_{j_3 \cdots j_n \neq i,j} x_{3,j_3} \; \dots \; x_{n,j_n} \; sign(i,j,j_3 \dots j_n) \tag{215}$$

This last quantity can be identified as the determinant of the $n-2 \times n-2$ matrix formed by deleting rows i and j and columns 1 and 2. The relation can easily be generalized to arbitrary columns k and l. If we call this determinant $[\mathbf{X}_{k,i,l,j}]$ and use $(x_k, x_l)_{i,j}$ to denote the subdterminant formed by rows i and j of the matrix (x_k, x_l), we can write:

$$[\, x_1, \; x_2 \dots x_n] \; = \; \sum_{j>i} (-1)^{i+j} (x_k, x_l)_{i,j} \; [\mathbf{X}_{k,i,l,j}] \tag{216}$$

As an example for a 4×4 matrix we have:

$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{bmatrix} =$$

$$= \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} c_3 & d_3 \\ c_4 & d_4 \end{bmatrix} - \begin{bmatrix} a_1 & b_1 \\ a_3 & b_3 \end{bmatrix} \begin{bmatrix} c_2 & d_2 \\ c_4 & d_4 \end{bmatrix} +$$

$$+ \begin{bmatrix} a_1 & b_1 \\ a_4 & b_4 \end{bmatrix} \begin{bmatrix} c_2 & d_2 \\ c_3 & d_3 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \begin{bmatrix} c_1 & d_1 \\ c_4 & d_4 \end{bmatrix} -$$

$$- \begin{bmatrix} a_2 & b_2 \\ a_4 & b_4 \end{bmatrix} \begin{bmatrix} c_1 & d_1 \\ c_3 & d_3 \end{bmatrix} + \begin{bmatrix} a_3 & b_3 \\ a_4 & b_4 \end{bmatrix} \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \end{bmatrix}$$

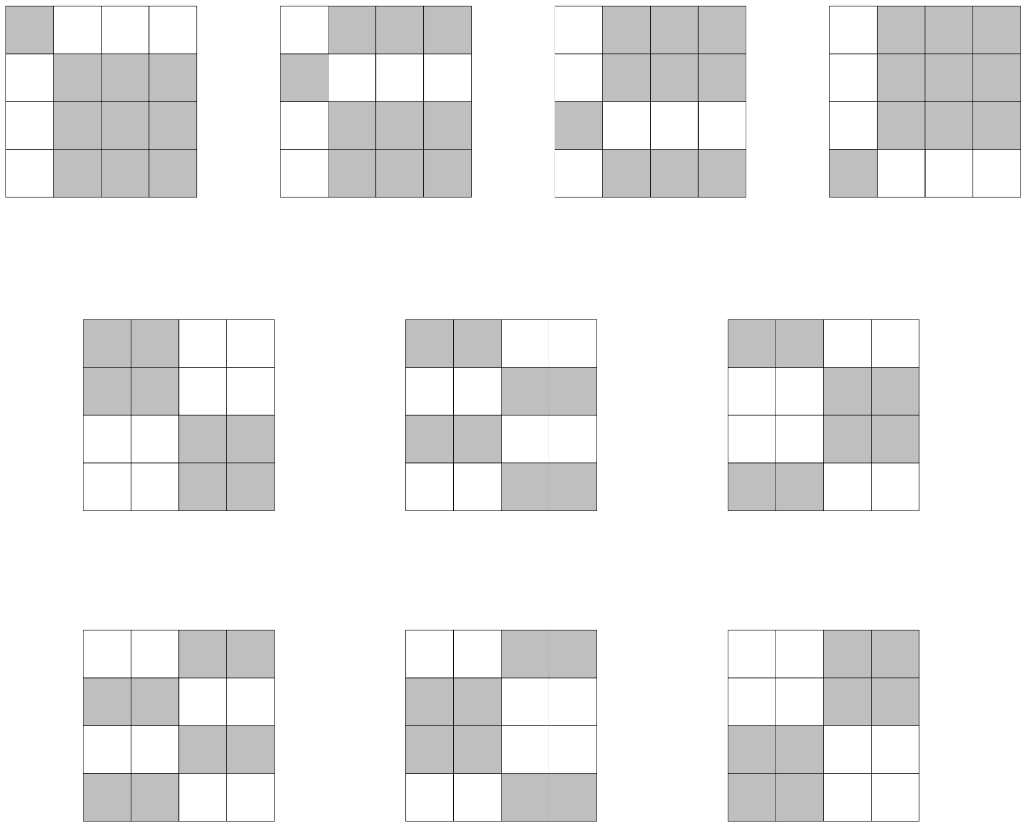

Figure 13: *Minors used in Laplace expansions*

This relation and the previous single column expansion are known as *Laplace expansions* of the determinant. It an of course be generalized to arbitrary numbers of columns $\leq n$. Fig. A.3 illustrates the minors used in the two examples of Laplace expansion

Since the computation of a determinant means picking an element from each column and different rows, it is easy to see that the determinant does not change if the matrix is transposed, i.e:

$$[\mathbf{X}^T] \;=\; [\mathbf{X}] \tag{217}$$

All Laplace expansions can therefore be made in the rows instead of the columns

A.4 The Product Rule

By taking the product of two matrices we get a new matrix. The determinant of the product matrix can then be related to the determinant of the original matrices in a very simple way.

Suppose we have two $n \times n$ matrices:

$$\mathbf{A} = (a_1,\ a_2,\ \ldots\ a_n) \qquad \mathbf{B} = (b_1,\ b_2,\ \ldots\ b_n) \quad a_i, b_j \ \in\ R^n \tag{218}$$

The determinant of their product can then be computed as:

$$\begin{aligned} [\mathbf{AB}] \ &= \ [\sum_{i_1} b_{1,i_1} a_{i_1}\ \ldots \sum_{i_n} b_{n,i_n} a_{i_n}] \ = \ \sum_{i_1 \ldots i_n} b_{1,i_1}\ \ldots\ b_{n,i_n}\ [a_{i_1}\ \ldots\ a_{i_n}] \ = \\ &= \sum_{i_1 \ldots i_n} b_{1,i_1}\ \ldots\ b_{n,i_n}\ sign(i_1, i_2 \ldots i_n)\ [a_1\ \ldots\ a_n] \ = [a_1\ \ldots\ a_n]\ [b_1\ \ldots\ b_n] \end{aligned} \tag{219}$$

The determinant involving the a_i:s vanishes whenever two vectors are equal. It is therefore always equal to $[\mathbf{A}]$ up to sign. The sum involving the b_i:s is then taken over distinct i_k:s only and is therefore just $[\mathbf{B}]$ and we get the important rule:

$$[\mathbf{AB}] \ = \ [\mathbf{A}]\,[\mathbf{B}] \tag{220}$$

A.5 Rectangular Matrices

If

$$\mathbf{A} = (a_1,\ a_2,\ \ldots\ a_p) \qquad \mathbf{B} = (b_1,\ b_2,\ \ldots\ b_p) \quad a_i, b_j \ \in\ R^n \tag{221}$$

are rectangular $n \times p$ matrices, with $p > n$ their product

$$\mathbf{AB}^T \ = \ (a_1,\ a_2,\ \ldots\ a_p) \begin{pmatrix} b_1^T \\ \\ b_2^T \\ \vdots \\ b_p^T \end{pmatrix} \tag{222}$$

will be a $n \times n$ square matrix. If we evaluate the determinant of this product we get:

$$\begin{aligned} [\mathbf{AB}^T] \ &= \ [\sum_{i_1}^{p} b_{i_1,1}\ a_{i_1}\ \sum_{i_2}^{p} b_{i_2,2}\ a_{i_2}\ \ldots \sum_{i_n}^{p} b_{i_n,n}\ a_{i_n}] \ = \\ & \sum_{i_1 \ldots i_n}^{p} b_{i_1,1}\ \ldots\ b_{i_n,n}\ [a_{i_1}\ \ldots\ a_{i_n}] \end{aligned} \tag{223}$$

Note that the summation indices ranges from 1 to $p > n$. The determinant $[a_{i_1}\ \ldots\ a_{i_n}]$ is just a subdeterminant of the matrix $\mathbf{A}$, where n columns out of p have been selected. There will totally be $\binom{p}{n}$ subdeterminants. If we denote such a subdeterminant with the columns ordered lexicographically, i.e. $l_1 < l_2 < \ldots l_n$ as $[\mathbf{A}]_{l_1, l_2 \ldots l_n}$ and denote $\Pi_{l_1, l_2 \ldots l_n}$ as the set of permutations of the indices $l_1, l_2 \ldots l_n$

$$[\mathbf{AB}^T] \ = \sum_{l_1 < \ldots < l_n}\ \sum_{i_1 \ldots i_n \in \Pi_{l_1, l_2 \ldots l_n}} b_{i_1,1}\ \ldots\ b_{i_n,n}\ \ sign(i_1, i_2 \ldots i_n)\ [\mathbf{A}]_{l_1, l_2 \ldots l_n} \tag{224}$$

but

$$\sum_{i_1 \ldots i_n \in \Pi_{l_1, l_2 \ldots l_n}} b_{i_1,1} \ \ldots \ b_{i_n,n} \ \ sign(i_1, i_2 \ldots i_n) \ = \ [\mathbf{B}]_{l_1, l_2 \ldots l_n} \tag{225}$$

which implies:

$$[\mathbf{AB}^T] \ = \ \sum_{l_1 < \ldots < l_n} [\mathbf{A}]_{l_1, l_2 \ldots l_n} \ [\mathbf{B}]_{l_1, l_2 \ldots l_n} \tag{226}$$

A.6 Solving Linear systems

The solution to linear systems can be written explicitly using determinants.

Suppose we have the linear equation:

$$b \ = \ \mathbf{A} q \tag{227}$$

where b and x are n-vectors and $\mathbf{A}$ an $n \times n$ matrix.

As before we have:

$$\mathbf{A} \ = \ (a_1, \ a_2, \ \ldots \ a_n) \tag{228}$$

Let $q \ = \ (q_1 \ \ldots \ q_n)^T$ We can the write b as:

$$b \ = \ \sum_j q_j \ a_j \tag{229}$$

Replacing the vector a_i with b in the determinant for $\mathbf{A}$ we get:

$$\begin{aligned} &[\ a_1, \ a_2, \ldots a_{i-1}, \ b, \ a_{i+1}, \ldots a_n] \ = \ [\ a_1, \ a_2, \ldots a_{i-1}, \ \sum_j q_j \ a_j, \ a_{i+1}, \ldots a_n] \ = \\ &= \ \sum_j q_j \ [\ a_1, \ a_2, \ldots a_{i-1}, a_j, \ a_{i+1}, \ldots a_n] \ = \ q_i \ [\ a_1, \ a_2, \ldots a_n] \end{aligned} \tag{230}$$

Where the last equality is due to the fact that the determinant involving the a_i:s vanishes unless $j = i$. We therefore get:

$$q_i \ = \ \frac{[\ a_1, \ a_2, \ldots a_{i-1}, \ b, \ a_{i+1}, \ldots a_n]}{[\ a_1, \ a_2, \ldots a_n]} \tag{231}$$