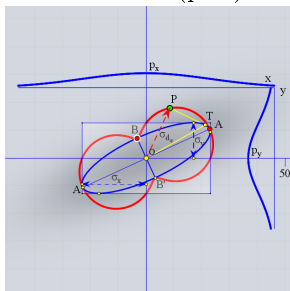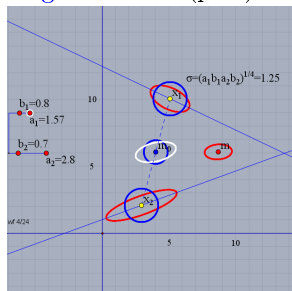# Cinderella Animations
Wolfgang Förstner, Bonn, 2024
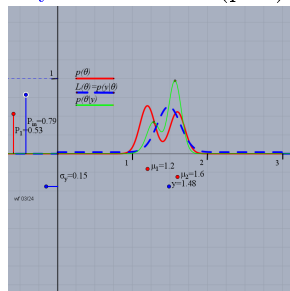
Pedal Curve (p. 3)
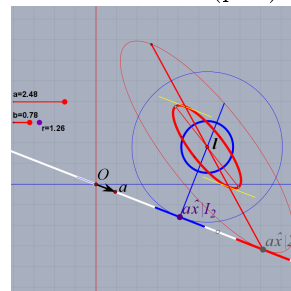

Weighted Mean (p. 5)

Bayes Estimation (p. 5)

ML Estimation (p. 6)

Point on Line (p. 7)

Uncertain Plane (p. 8)
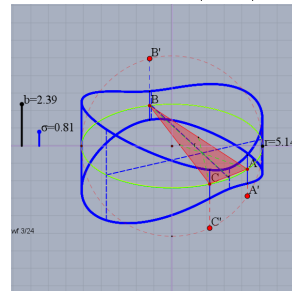
Uncertain Circle (p. 10)
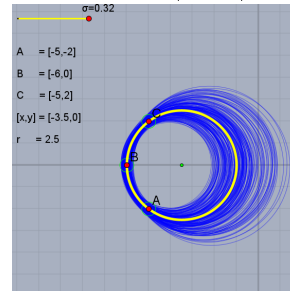
Gauge and Loop-Closing (p. 11)
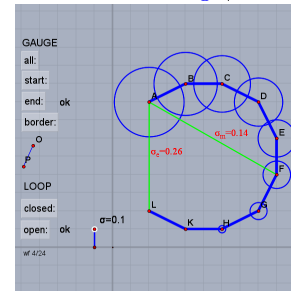
Gauge and Point Cloud (p. 11)
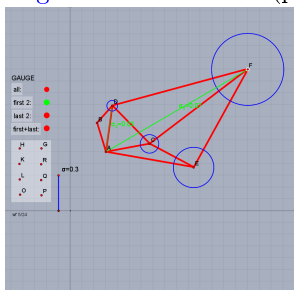
CovM Comparison (p. 12)
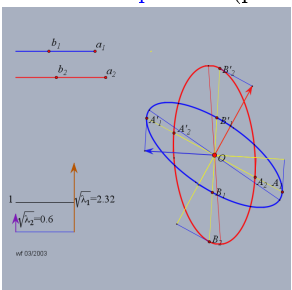
Classification Boundary (p. 14)
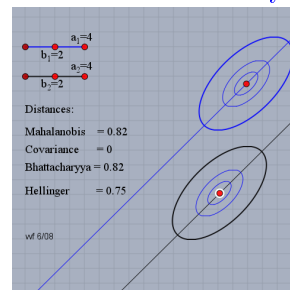
P4P Precision (p. 18)
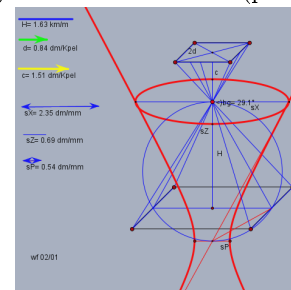
P4P Correlation (p. 18)
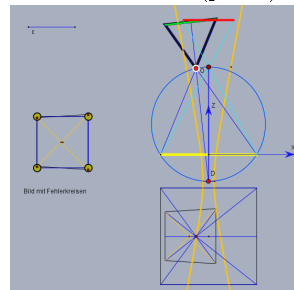
Epipolar Geometry (p. 20)
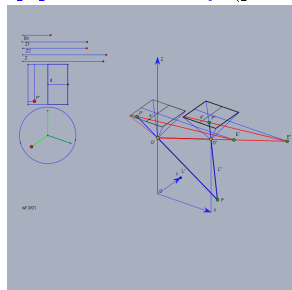
Depth of Field (p. 20)
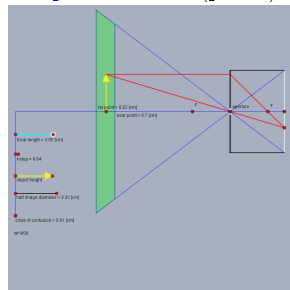
Trifocal Contraint (p. 22)
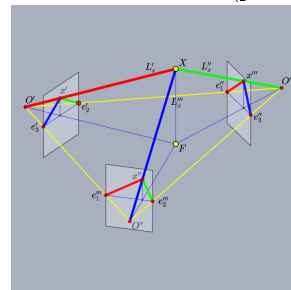
Crossratio (p. 24)
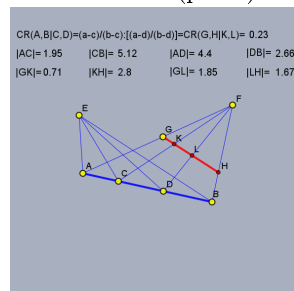
Rectification (p. 24)

Motion as Rotation (p. 27)
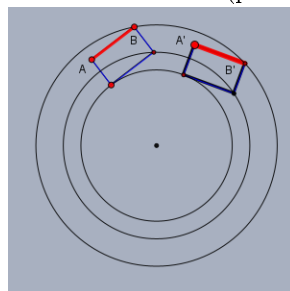
Double Reflection (p. 28)

Complex Multiplication (p. 29)

Four Bearings (p. 30)

Moving a Piano (p. 30)

Contents

1 Preface

This collection of animations initially served as illustrations in my lectures. They emerged since more than two decades. The topics are mainly statistics, estimation theory, and geometry. They all are constructions enabled by the free software Cinderella which provides interactive graphics, and html-export. In the original version Cinderella.1 (1998) the animations needed to be constructed, mainly with compass and ruler. The new version of Cinderella.2 includes a script language, which allows complex programming, contains additional plotting routines, and - as a side effect - an easy control of font placement especially of moving elements. These features allowed to enhance the appearance of the old animations, and an easy generation of new ones.

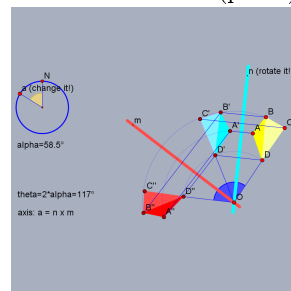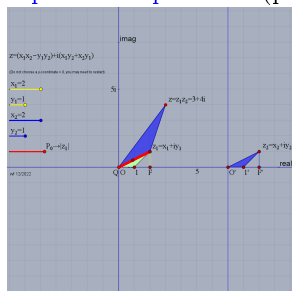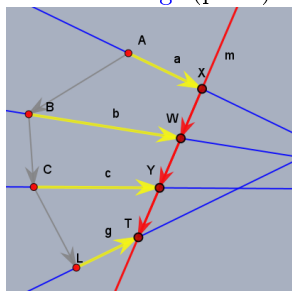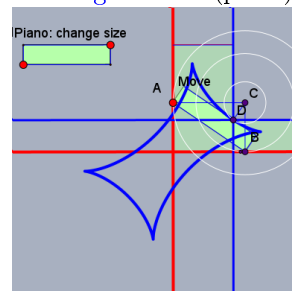The animations are of different complexity, some refer to simple geometric relations, some to more advanced concept of estimation theory. They partly were illustrations of notes in the accompanying Collected Notes, to which then links are provided.

The animations are not fool proof: In case you get stuck, just restart the animation.

The first page provides links to the animations and to some additional explanations below.

One may download the html-files of the Cinderella animations and use them locally, if you copy the Javascript-file KetCindyPlugin.jar into the same directory.

You also may refer to the repository github to download all Cinderella Animations together with a version of this file, which allows you to use the animations locally, i.e., without having access to the internet. You then also obtain all source code of the animations and the explanatory html-text, which you might use freely.

2 Pedal Curve of Standard Ellipse

This animation illustrates the uncertainty of two random variables, say, the coordinates of a 2D point. The basics are also given in (Förstner and Wrobel, 2016, Sect. 2.4.4.2), from which the following text is taken:

"If two independent random variables are normally distributed according to

$$\underline{x} \sim \mathcal{N}(\mu_x, \sigma_x^2) \qquad \underline{y} \sim \mathcal{N}(\mu_y, \sigma_y^2), \tag{1}$$

their joint density function is

$$p_{xy}(x, y) = g_x(x \mid \mu_x, \sigma_x^2) \, g_y(y \mid \mu_y, \sigma_y^2) \tag{2}$$

$$= \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\left(\frac{x - \mu_x}{\sigma_x}\right)^2 + \left(\frac{y - \mu_y}{\sigma_y}\right)^2\right)}. \tag{3}$$

With the vectors

$$\boldsymbol{x} = \left[\begin{array}{c} x \\ y \end{array}\right] \qquad \mu = \left[\begin{array}{c} \mu_x \\ \mu_y \end{array}\right] \tag{4}$$

and the 2×2 matrix,

$$\Sigma = \left[\begin{array}{cc} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{array}\right], \tag{5}$$

this can be written as

$$g_{xy}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi\sqrt{|\boldsymbol{\Sigma}|}} e^{-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})}. \tag{6}$$

If the 2×2 matrix Σ is a general symmetric positive definite matrix

$$\Sigma = \left[\begin{array}{cc} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{array}\right], \tag{7}$$

the two random variables are dependent. The correlation coefficient,

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \in [-1, 1], \tag{8}$$

measures the degree of linear dependency. If $\rho_{xy} = 0$, the two random variables are *uncorrelated*, and if they are normally distributed, they are *independent*, since then $\mathrm{p}_{xy}(x, y) = \mathrm{p}_x(x)\,\mathrm{p}_y(y)$. The 2D normal distribution is an elliptic bell-shaped function and can be visualized by one of its contour lines, cf. Fig. 1. The *standard ellipse*, sometimes also called standard error ellipse, is defined by

uncorrelated, independent random variables

standard ellipse

$$(\boldsymbol{x} - \boldsymbol{\mu})^\mathsf{T} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = 1 \,. \tag{9}$$
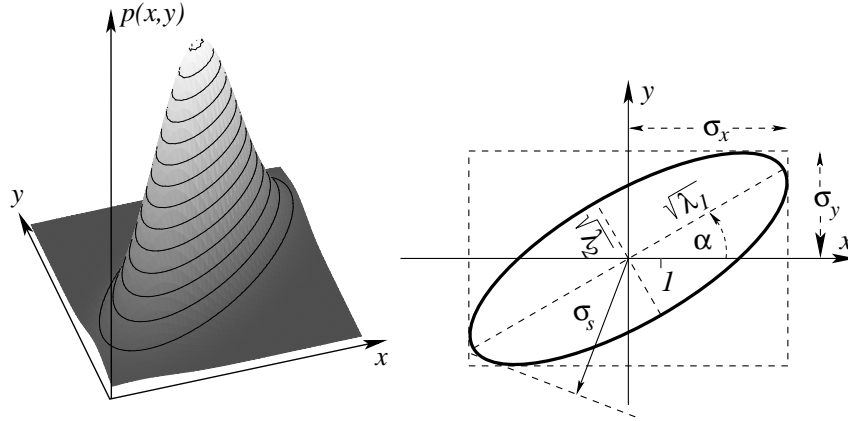
Figure 1: General 2D normal or Gaussian distribution, centred at the origin. **Left:** density function. **Right:** standard ellipse. Actual values: $\mu_x = \mu_y = 0$, $\sigma_x = 4.9$, $\sigma_y = 3.2$, $\rho = 0.7$

The standard ellipse allows the visualization of important properties of the uncertain point:

- The standard ellipse is centred at $\boldsymbol{\mu}_x$.

- The bounding box has size $2\sigma_x \times 2\sigma_y$.

- The semi-axes are the square roots of the eigenvalues λ_i of the covariance matrix, namely $\sigma_{\max} = \sqrt{\lambda_1}$ and $\sigma_{\min} = \sqrt{\lambda_2}$, which are the square roots of the eigenvalues of $\boldsymbol{\Sigma}$,

$$\sigma_{\max,\min}^2 = \frac{1}{2}(\sigma_x^2 + \sigma_y^2) \pm \frac{1}{2}\sqrt{(\sigma_x^2 - \sigma_y^2)^2 + 4\sigma_{xy}^2} \,. \tag{10}$$

- If the two coordinates are correlated, the major axis is not parallel to the coordinate system. The angle α is given by

$$\alpha = \frac{1}{2}\mathrm{atan2}\left(2\sigma_{xy}, \sigma_x^2 - \sigma_y^2\right) \in (-\pi/2, +\pi/2] \tag{11}$$

using a two-argument version of the arctan function.
The sign of the angle follows the sign of the correlation coefficient ρ_{xy} or the covariance σ_{xy}.

- The standard deviation σ_s of a distance s between the point $\boldsymbol{\mu}_x$ and a fixed point in an arbitrary direction, indicated here by an arrow, is given by the distance of $\boldsymbol{\mu}_x$ from the tangent to the standard ellipse perpendicular to that direction. This shows that the minor and the major axes of the standard ellipse give the minimum and the maximum of the directional uncertainty of the point.

4

In higher dimensions, (9) represents an ellipsoid or a hyper-ellipsoid $\mathcal{E}$. The probability $S = P(\underline{x} \in \mathcal{E})$ that a random point lies within the standard ellipsoid depends on the dimension as shown in the first line of Table 1, and rapidly diminishes with the dimension.

Instead of showing the standard ellipse or standard ellipsoid, we therefore can show the *confidence ellipse* or *confidence ellipsoid*. The confidence ellipsoid is the k-fold standard ellipsoid, such that the probability $P(\underline{x} \in \mathcal{E}(k))$ that a sample lies within the ellipsoid is a certain prespecified value S

$$\mathcal{E}(k): \quad (\boldsymbol{x} - \boldsymbol{\mu})^{\mathsf{T}} \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = k^2, \quad P(\underline{x} \in \mathcal{E}(k)) = S. \tag{12}$$

The standard ellipse is identical to the confidence ellipse for $k = 1$: $\mathcal{E} = \mathcal{E}(1)$. For the dimension $d = 1$ and a probability $P(\underline{x} \in \mathcal{E}(k)) = S = 0.9973$, we would obtain $k = 3$. Here the ellipse reduces to the interval $[-k\sigma_x, +k\sigma_x]$.

For $S = 95\%$, $S = 99\%$ and $S = 99.9\%$, the values $k(S)$ determined from the right equation in (12) are given in Table 1 for different dimensions.

Table 1: Confidence regions. First row: Probabilities $P(\underline{x} \in \mathcal{E})$ for different dimensions d of a random vector $\underline{x}$. **Other rows:** Factor $k(S)$ for the confidence ellipsoids $\mathcal{E}(k(S))$ for $S = 0.95, 0.99, 0.999$ and for different dimensions d.

d	1	2	3	4	5	10	20	50	100
$P(\underline{x} \in \mathcal{E})$	0.68	0.40	0.20	0.09	$3.7 \cdot 10^{-2}$	$1.7 \cdot 10^{-4}$	$1.7 \cdot 10^{-10}$	$1.2 \cdot 10^{-33}$	$1.8 \cdot 10^{-80}$
$k(0.95)$	1.96	2.45	2.80	3.08	3.33	4.28	5.60	8.22	11.2
$k(0.99)$	2.58	3.03	3.37	3.64	3.88	4.82	6.13	8.73	11.6
$k(0.999)$	3.29	3.72	4.03	4.30	4.53	5.44	6.73	9.31	12.2

...”

Go to the beginning.

3 Weighted Mean of two 2D Points with arbitrary Covariance Matrix

The animation allows to explore the effect of the covariance matrices of two 2D points, with general covariance matrices, onto the maximum likelihood (ML) estimate of the mean of the two points and the comparison with the arithmetic mean of the two points including the loss in accuracy when assuming a wrong covariance matrix. The concept is described in the note Accuracy of the Mean when using a Wrong Covariance Matrix. Details referring to the Cinderella animation there are given in Subsection 11.6.

Go to the beginning.

4 Bayes Estimation with multi-model prior and Outliers

The animation allows to explore the effect of prior information onto the statistically optimal estimation of a parameter, and refers to the note Bayes- and Maximum-Likelihood-Estimation. Details are given in Subsection 5.6.

The prior is a mixture of two Gaussians, having the same standard deviation $\sigma = 0.1$. You interactively may change the two mean values μ_1 and μ_1 and the probability $P_1 = P(\underline{\mu} = \mu_1)$ for the mean of the first Gaussian. You may simulate a single Gaussian by either choosing $\mu_1 = \mu_2$, or by setting $P_1 = 0$ or $P_1 = 1$. Observe, in case the red dot for changing P_1 is negative or above 1, the value is set to 0 or 1, respectively.

The likelihood $L(\theta) = p(y \mid \theta)$ is a mixture of a Gaussian $\mathcal{N}(y \mid x, \sigma_y^2)$ and a uniform distribution $\mathcal{U}(y \mid 0, 3)$. The value y has the meaning of the observed value. You may change the observation y and its standard deviation σ_y moving the two blue dots. You also may change the probability P_{in} that the observation is an inlier. Again, if the blue

dot controlling $P_{\text{in}} = \text{P}(y\ \texttt{is\ inlier})$ is below 0 or above 1, we either do not know where the observation is between 0 and 3, or we know for sure it is an inlier, respectively.

The posterior density generally shows two local maxima. The global maximum usually is taken as the best – maximum a posteriori (MAP) – estimate. However, in case the smaller maximum is still large, it might be adequate, to report both maxima including the ratio of the posterior density.

In case one allows outliers, it is interesting to observe when the observation has an negligible effect onto the estimate, and how the prior density influences the maximum/maxima of the posterior density.

Go to the beginning.

5 Effect of the Covariance Matrix of Observations onto ML-Estimates

The animation allows to explore the effect of the covariance matrix of a 2D observation onto the maximum likelihood (ML) estimate of a single parameter. In the context of note Pre-calibration and in-situ Self-calibration with Correlated Observations it allows to visualize Rao's Lemma, cited in Subsection 9.3.2.

The Gauss–Markov model for the animation is given by[1]

$$\mathbb{E}(\boldsymbol{y}) = \boldsymbol{x}\theta\,, \quad \text{with} \quad |\boldsymbol{x}| = 1 \quad \text{and} \quad \mathbb{D}(\boldsymbol{y}) = \Sigma_{yy} = W_{yy}^{-1}\,. \tag{13}$$

The model is linear in the parameter. The variance of the estimated parameter is

$$\sigma_{\hat{\theta}}^2 = (\boldsymbol{x}^\mathsf{T} W_{yy} \boldsymbol{x})^{-1}\,. \tag{14}$$

In the animation we provide two cases for the covaraince matrix

1. Isotropic uncertainty with

$$\textsf{case } I: \quad \Sigma_{yy} = r^2 I_2\,. \tag{15}$$

   The radius r of the (circular) standard ellipse (blue) can be controlled interactively.

2. General uncertainty with

$$\textsf{case } \Sigma: \quad \Sigma_{yy} = R(\phi)\text{Diag}([s_1, s_2])R^\mathsf{T}(\phi)\,. \tag{16}$$

   The semiaxes s_1 and s_2 and the direction ϕ of the standard ellipse (red) can be controlled interactively.

Choosing the principal axes of the standard ellipse parallel or perpendicular to the vector $\boldsymbol{x}$, allows to demonstrate Rao's lemma (Rao, 1967).

Choosing the direction of $\boldsymbol{x}$ along the first diagonal $\boldsymbol{x} = [1,1]^\mathsf{T}/\sqrt{2}$, or equivalently $y_2 = y_1$, allows to mimic the estimation of the mean of the two values y_1 and y_2, since the constraint for the fitted observations is $\hat{y}_2 - \hat{y}_1 = 0$. The optimal estimate for the fitted observations $\hat{\boldsymbol{y}} \mid I = \boldsymbol{x}\hat{\theta} \mid I_2$ is at the point $\hat{\boldsymbol{y}}/2 = [y_1/2, y_2/2]^\mathsf{T}$ confirming the classical arithmetic mean. Changing the red ellipse allows to reach any point on the white line, i.e., for any given mean m there always is a covariance matrix, such that the weighted mean of y_1 and y_2 is identical to $\hat{\mu} = m$.

The next animation demonstrates this for the 2D case.

Go to the beginning.

---

[1] In the context of this collection of Cinderella Animations we allow the notation in the animation be adapted, either $(\boldsymbol{\theta}, X, \boldsymbol{y})$ or $(\boldsymbol{x}, A, \boldsymbol{l})$, the design matrices being $X = \boldsymbol{x}$ and $A = \boldsymbol{a}$ respectively.

6   Checking Point-Line Incidence

The animation shows the effect of the geometric relation of an uncertain point and an uncertain line onto the test statistic for checking the incidence of the point and the line.

The line ℓ is generated by the join of two points $\mathcal{A}$ and $\mathcal{B}$:

$$\ell = \mathcal{A} \wedge \mathcal{B} : \quad \mathbf{l} = \mathbf{x}_A \times \mathbf{x}_B \,. \tag{17}$$

We assume the following uncertainties:

$$\Sigma_{AA} = \sigma_A^2 I_2 \quad \Sigma_{BB} = \sigma_B^2 I_2 \,, \tag{18}$$

and

$$\Sigma_{CC} = R(\psi) \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix} R^{\mathsf{T}}(\psi) \,, \tag{19}$$

where the parameters $\boldsymbol{x}_A, \boldsymbol{x}_B, \sigma_A, \sigma_B, a, b$, and ψ can be given interactively. The construction of the standard hyperbola of the joining line is based on the centroid representation of the line

$$\underline{\ell} : \quad (\boldsymbol{x}_O, \alpha, \sigma_q, \sigma_\alpha) \,, \tag{20}$$

see Fig. 2, taken from Förstner and Wrobel (2016, p. 374). All parameters can easily be derived from the given uncertain points $\underline{\mathcal{A}}$ and $\underline{\mathcal{B}}$, though we do not need all.
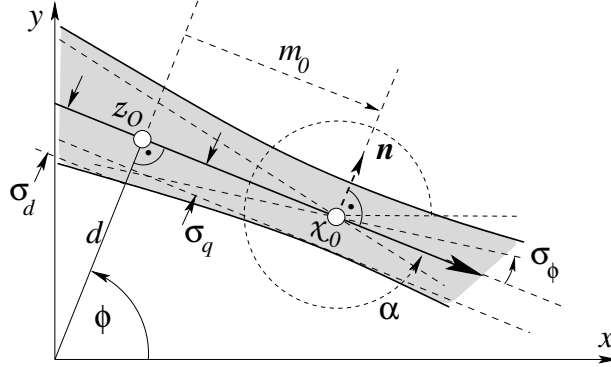
Figure 2: Uncertain 2D line and its representation in the real plane. Centroid χ_0, direction α of the line, direction ϕ of the normal $\boldsymbol{n}$, $\phi = \alpha + \pi/2$, distance d to the origin, foot point $\boldsymbol{z}_O$ of the origin, distance m_0 of the centroid from the foot point, standard deviations $\sigma_\alpha = \sigma_\phi$, σ_d, and σ_q: of the direction α, the distance d and the position of the centroid across the line, respectively. The standard deviation σ_ϕ is visualized as the angle between the (mean) line and one of the asymptotic lines of the hyperbola, from Förstner and Wrobel (2016, p. 374)

The center $\boldsymbol{x}_O$ if the line is given by

$$\boldsymbol{x}_O = \frac{\sigma_B^2 \boldsymbol{x}_A + \sigma_A^2 \boldsymbol{x}_B}{\sigma_A^2 + \sigma_B^2} \,, \tag{21}$$

which is the point

$$\boldsymbol{x}(\lambda) = (1 - \lambda)\boldsymbol{x}_A + \lambda \boldsymbol{X}_B \tag{22}$$

with minimal uncertainty

$$\lambda = \operatorname{argmin}_\lambda (1 - \lambda)^2 \sigma_A^2 + \lambda^2 \sigma_B^2 \,. \tag{23}$$

thus

$$\Sigma_{xx}(\lambda) = \frac{\sigma_A^2 \sigma_B^2}{\sigma_A^2 + \sigma_B^2} I_2 \tag{24}$$

The direction α is given by construction. Its standard deviation results from

$$\sigma_\alpha = \frac{\sqrt{\sigma_A^2 + \sigma_B^2}}{|\boldsymbol{x}_A - \boldsymbol{x}_B|}, \tag{25}$$

being an approximation, sufficient for small uncertainties. We do not use it in the animation. The standard deviation σ_q is given by

$$\sigma_q = \frac{\sigma_A \sigma_B}{\sqrt{\sigma_A^2 + \sigma_B^2}}. \tag{26}$$

The hyperbola is constructed by the four end points of the two standard intervals at $\mathcal{A}$ and $\mathcal{B}$ and one endpoint of the standard interval at the midpoint $\boldsymbol{x}_M = (\boldsymbol{x}_A + \boldsymbol{x}_B)/2$, which is given by $\sqrt{\sigma_A^2 + \sigma_B^2}/2$ across the line (not shown in the animation).

The statistical test $z(\mathcal{A}, \mathcal{B}, \mathcal{C})$ of the incidence of a point, here $\mathcal{C}$ and a line l, which is constructed by two points, here by $\mathcal{A}$ and $\mathcal{B}$ thus

$$\iota(\mathcal{A}, \mathcal{B}, \mathcal{C}) := \iota(\mathcal{C}, \mathcal{A} \wedge \mathcal{B}) \tag{27}$$

depends on the sequence of the three points, when applying variance propagation to determine the uncertainty of $\mathit{l} = \mathcal{A} \wedge \mathcal{B}$, due to linearization effects for the line, which are different for $z(\mathcal{B}, \mathcal{C}, \mathcal{A})$ and $z(\mathcal{C}, \mathcal{A}, \mathcal{B})$. This is why we check the incidence by checking the collinearity of the three points based on the determinant

$$D = 2A = \det([\mathbf{x}_A, \mathbf{x}_B, \mathbf{x}_C]) \tag{28}$$

which is invariant to permutations of the three points (except for the sign). We therefore obtain the test statistic

$$z = \frac{D}{\sigma_D} \tag{29}$$

with the variance of the determinant

$$\sigma_D^2 = |\boldsymbol{x}_C - \boldsymbol{x}_B|^2 \sigma_A^2 + |\boldsymbol{x}_C - \boldsymbol{x}_A|^2 \sigma_B^2 + \{(\boldsymbol{x}_B - \boldsymbol{x}_A)^\perp\}^\mathsf{T} \Sigma_{CC} (\boldsymbol{x}_B - \boldsymbol{x}_A)^\perp, \tag{30}$$

where the operator $\perp$ rotates a vector by 90°. The sign of z is positive, in case $\mathcal{C}$ lies on the left-hand side of the line $\mathcal{AB}$.

Go to the beginning.

7 Visualization of the Uncertainty of a Plane through Three Points

The uncertain plane is described in the note 19 in Centroid Form of an Uncertain Plane. Here we provide a visualization for a simple case.

7.1 General Setup

A plane has three degrees of freedom. The three parameters together with the six parameters of the corresponding covariance matrix can be used to represent an uncertain plane $\underline{\mathcal{A}}$ in a centroid form:

$$\underline{\mathcal{A}}: \quad \{\boldsymbol{X}_O, \mathcal{Q}, \sigma_q, \sigma_\alpha, \sigma_\beta\}, \tag{31}$$

where

1. the centre $\boldsymbol{X}_O$ of the uncertain plane is the point where the uncertainty across the plane is minimum,

2. the rotation matrix $\mathcal{Q} = [\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{N}]$ specifies a local frame, with the first two axes span the plane, and the third axis provides the normal $\boldsymbol{N}$ of the plane, and

3. the standard deviation σ_q is the uncertainty of the position of $\boldsymbol{X}$ along the normal and the two standard deviations σ_α and σ_β are the uncertainty of the normal, namely uncertainty of rotations around $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$.
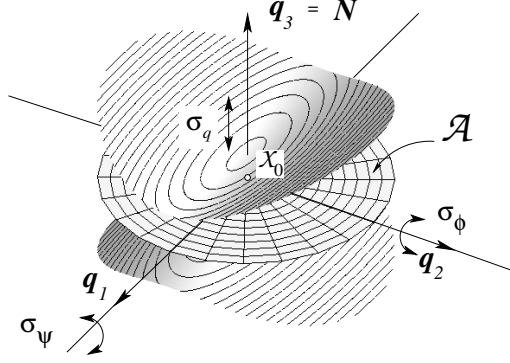
Figure 3: Uncertain plane. The mean plane is shown as a circular disc in the middle, containing the centre point $\mathcal{X}_0$ of the standard hyperboloid of two sheets indicating the uncertainty perpendicular to the plane. At the centre point $\mathcal{X}_0$, the uncertainty across the plane is smallest. The mutually perpendicular 3D lines through $\mathcal{X}_0$ with directions $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are the axes of maximal and minimal rotational uncertainty of the plane. Isolines of uncertainty perpendicular to the plane are ellipses with their large semi-axis in the direction of $\boldsymbol{q}_1$. Fig. adapted from Förstner and Wrobel (2016, p. 378)

7.2 Uncertain Plane through three Co-Circular Points

For the animation, we assume three uncertain points $\underline{\mathcal{X}}_1$, $\underline{\mathcal{X}}_2$, and $\underline{\mathcal{X}}_3$ lie on a horizontal circle. They have the same accuracy with a standard deviation of σ in all directions. Depending on the form of the triangle, the plane through the three points has a different uncertainty. This is shown by the upper and lower bound of the vertical standard interval of predicted points on the circle

The three points have the coordinates

$$\underline{\boldsymbol{X}}_i = R \left[\begin{array}{c} \cos \alpha_i \\ \sin \alpha_i \\ \underline{Z}_i \end{array} \right] \quad \text{with} \quad i = 1, 2, 3, \quad \text{and} \quad \underline{Z}_i \sim \mathcal{N}(0, \sigma^2). \tag{32}$$

The plane is assumed to be given by

$$Z_i = \theta_1 + \theta_2 X_i + \theta_3 Y_i \quad \text{with} \quad i = 1, 2, 3. \tag{33}$$

The parameters can be derived from the equation system

$$A\boldsymbol{\theta} = \boldsymbol{Z} \quad \text{with} \quad A = \left[\begin{array}{ccc} 1 & X_1 & Y_1 \\ 1 & X_2 & Y_2 \\ 1 & X_3 & Y_3 \end{array} \right], \ \boldsymbol{Z} = \left[\begin{array}{c} Z_1 \\ Z_2 \\ Z_3 \end{array} \right] \quad \text{and} \quad \boldsymbol{\theta} = \left[\begin{array}{c} \theta_1 \\ \theta_2 \\ \theta_3 \end{array} \right] \tag{34}$$

Thus, since

$$\mathbb{D}(\underline{\boldsymbol{Z}}) = \sigma^2 I_3 \tag{35}$$

The covariance matrix of the parameters is

$$\Sigma_{\theta\theta} = \sigma^2 A^{-1} A^{-\mathsf{T}}. \tag{36}$$

For an arbitrary point

$$\left[\begin{array}{c} X \\ Y \end{array} \right](t) = R \left[\begin{array}{c} \cos t \\ \sin t \end{array} \right] \tag{37}$$

on the circle, its vertical variance therefore is

$$\sigma_Z^2(t) = [1, R\cos t, R\sin t]\Sigma_{\theta\theta} \left[\begin{array}{c} 1 \\ R\cos t \\ R\sin t \end{array} \right], \tag{38}$$

which yields the two curves

$$\boldsymbol{X}(t) = r \begin{bmatrix} \cos t \\ \sin t \\ \pm \sigma_Z(t) \end{bmatrix} \tag{39}$$

collecting the upper and lower border of the vertical 1D standard intervals of the points on the circle.

The point where the uncertainty across the plane is minimum is the centroid of the three points

$$\boldsymbol{x}_O = \sum_{i=1}^{3} \boldsymbol{x}_i/3, \quad \text{with} \quad \boldsymbol{x}_i = \begin{bmatrix} X_i \\ Y_i \end{bmatrix}. \tag{40}$$

The directions in which the plane has its best and worst rotational uncertainty is given by the principal axes of the central moment matrix.

$$M = \sum_{i=1}^{3} \boldsymbol{x}_i \boldsymbol{x}_i^{\mathsf{T}} - 3\boldsymbol{x}_O \boldsymbol{x}_O^{\mathsf{T}}. \tag{41}$$

They are indicated by the two mutually perpendicular dashed blue lines.

Go to the beginning.

8 Uncertain Circle through three Points

Given three 2D points $\boldsymbol{x}_i$ with isotropic and homogeneous uncertainty of their coordinates, we visualize uncertainty of the circle through these points.

The circle-constraint

$$g_i(\boldsymbol{x}_i; \boldsymbol{x}_0, r) = 0 \quad \text{with} \quad g_i = (x_i - x_0)^2 + (y_i - y_0)^2 - r^2 \tag{42}$$

can be written as linear problem in the parameters

$$\boldsymbol{\theta} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ x_0^2 + y_0^2 - r^2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_0 \\ y_0 \\ r \end{bmatrix} = \begin{bmatrix} u \\ v \\ \sqrt{u^2 + v^2 - w} \end{bmatrix} \tag{43}$$

The parameters (u, v, w) can be derived from the linear equation system

$$M\boldsymbol{\theta} = \boldsymbol{m} \quad \text{with} \quad M = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{m} = \begin{bmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ x_3^2 + y_3^2 \end{bmatrix}. \tag{44}$$

leading to

$$\boldsymbol{\theta} = M^{-1}\boldsymbol{m}. \tag{45}$$

For visualizing the uncertainty, we draw K samples on the standard ellipsoid of the input data

$$\begin{bmatrix} \underline{\boldsymbol{x}}_1 \\ \underline{\boldsymbol{x}}_2 \\ \underline{\boldsymbol{x}}_3 \end{bmatrix}_k = \underline{\boldsymbol{z}}_k \sim \mathcal{N}(\mu_z, \sigma^2 I_6) \quad \text{with} \quad \mu_z = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \boldsymbol{x}_3 \end{bmatrix}, \tag{46}$$

and derive the K parameters $(\boldsymbol{x}_0, r)_k$. They cover an uncertainty region, as can be seen in the animation.

Go to the beginning.

9 Effect of Gauge Choice onto a Polygon with and without Loop-Closing and a Triangulated Network

Thwo Cinderella animations allows to investigate the effect of choosing a certain gauge for representing the uncertainty of a point cloud, namely a polygon and a triangulated network.

In addition, the effect of loop closing onto the internal precision of the polygon can be explored.

The description of the underlying estimation model also is used for the animation of choosing a gauge in a triangulated geodetic network.

9.1 The gauge of a point cloud

Classical coordinate transformations change the coordinates, either due to a possibly scaled motion of the points or due to a change of the coordinate system.

Often, we estimate the coordinates of a point cloud which is defined only up to a similarity transformation (or another transformation), e.g., when deriving the coordinates from a set of calibrated cameras using a, what is called free bundle adjustment. Then one needs to fix the coordinate system in some way, e.g., by fixing a minimal set of parameters (4 in 2D, 7 in 3D) in order to avoid the normal equation system to become singular. The fixing parameters define what is called the *gauge* (or, mostly in German, the *datum*). The internal precision of the point cloud only depends on the configuration of the images and the observed points, not on the choice of the parameters fixing the coordinate system. Angles and distance ratios are invariant to this choice. Therefore, one is *free* in choosing the gauge, or the datum, why one – in the context of bundle adjustment – also refers to *free bundle adjustment*. Changing the fixing parameters changes the uncertainty of the parameters, but not the uncertainty of angles and distance ratios.

We distinguish these two cases: changing the coordinates is called a K-Transformation, changing the fixing parameters is called a gauge or S-transformation (S standing for similarity).

The basic concepts for handling the gauge or the datum are described in (Förstner and Wrobel, 2016, Sect. 4.5.1), from which the following figure is taken. The Cinderella
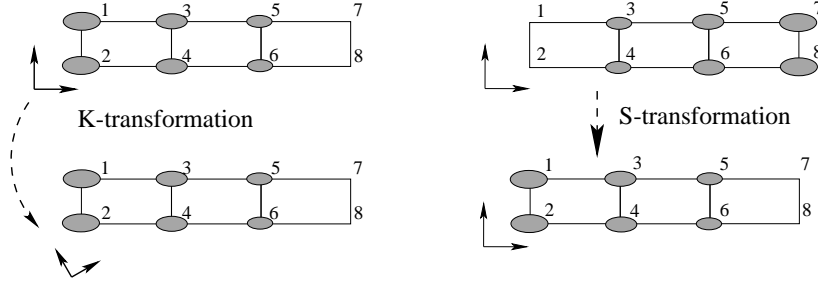
Figure 4: Coordinate or K-transformation and gauge or S-transformation. **Left:** K-transformation changes the reference coordinate system of the coordinates. The gauge is not changed, which here is defined by point pair (7,8). **Right:** S-transformations change the gauge of a covariance matrix, here from point pair (1,2) to point pair (7,8). The coordinates are not changed

animation allows to explore the effect of choosing a specific gauge onto the uncertainty of the points of a polygon and onto the angles and distance ratios derived from these coordinates.

The description of the underlying estimation model is given in Gauge Choice and Loop Closing of Uncertain Polygons and is consistently using complex numbers for coordinates, scaled rotations and angles and distance ratios.

9.2 Loop-Closing in SLAM

The increase of the variances with the length of a path, being of third order since it approximately is a doubly integrated white noise sequence, requires to exploit possibilities of loop closing, i.e., identifying correspondences with objects seen – possibly long – before. This may dramatically reduce the uncertainty. In our case of a polygon with observed angle and length ratios, we assume, that such observations are possible at the starting and at the end point of a polygon. The effect can be observed in the animation.

The animation Gauge in triangulation applies the same idea to a triangulated planar point cloud, representing a geodetic network, where in each triangle all three scaled rotations between neighbouring sides are assumed to be observed.

Go to the beginning.

10 Comparing two Covariance Matrices

This Cinderella Animation visualizes the comparison of two covariance matrices.

The comparison of two covariance matrices occurs in the following situations, see Fig. 5:

- A new (version of an) algorithms provide estimates as an old (reference) algorithm for the same parameter vector together with their covariance matrices, say Σ and C. We want to know, whether one algorithm yields better results than the other, i.e., whether $\Sigma \leq C$.

- A user has specified the minimum uncertainty of parameters, say C which needs to be achieved by some method. Either the method itself provides a covariance matrix Σ for the estimated parameters, or we are able to derive the covariance matrix by some simulation, e.g., by randomly perturbating the observations K times and from the K resulting parameters derive the empirical covariance matrix. Again we want to know whether $\Sigma \leq C$.

- We want to know how similar two covariance matrices are, e.g., measured by a distance $d(\Sigma, C)$ between the two covariance matrices.
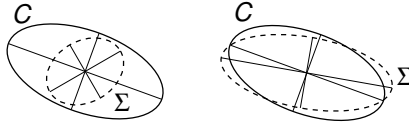
Figure 5: Comparing a covariance matrix Σ with a criterion matrix C: the standard ellipsoid of the covariance matrix Σ is required to lie completely in the standard ellipsoid of the criterion matrix C, as in the left figure, or is required to be close to C, as in the right figure. From (Förstner and Wrobel, 2016, Fig. 4.10)

In all cases we want to compare two covariance matrices, say Σ and C. The comparison should be invariant w.r.t. to an affine transformation, especially w.r.t. to a similarity or an individual scaling of the parameters. The comparison can be based on the eigenvalues of ΣC^{-1}. They indicate the ratios of variances $\sigma^2_{f(\theta)}$ of arbitrary functions $f(\boldsymbol{\theta})$, especially provide the maximum and minimum ratios of the variances.

The Cinderella animation allows to explore the maximum and minimum of the ratio r

$$r_{\max,\min} = \sigma^{(C)}_{f(\theta)} / \sigma^{(\Sigma)}_{f(\theta)} \Big|_{\max,\min} = \sqrt{\lambda_{\max,\min}(\Sigma C^{-1})} \tag{47}$$

of the standard deviations as a function of the two covariance matrices C and Σ.

The basic concept for such a comparison can be found in (Förstner and Wrobel, 2016, Sect. 4.6.2.3). We provide this text here and adapted it to our context.

Acceptability of the Precision

Users may require a certain precision to be achieved by the observation process. We assume they specify this precision by some reference or *criterion matrix* $C := \Sigma_{\hat{\theta}\hat{\theta}}^{\text{ref}}$. Then we need to compare the achieved theoretical covariance $\Sigma := \Sigma_{\hat{\theta}\hat{\theta}}$ with the a priori specified covariance matrix (Baarda, 1973). We first discuss the comparison and then methods for specification.

The comparison of two covariance matrices is only meaningful if they refer to the same gauge. Otherwise, they need to be transformed into the same gauge, cf. the Cinderella . We may require that the achieved precision is consistently better than the reference or that it is similar to it on average.

Acceptability. The acceptability of the achieved precision may be based on the individual empirical standard deviations, requiring

$$\sigma_{\hat{\theta}_u} \leq \sigma_{\hat{\theta}_u}^{\text{ref}} \qquad \text{for a prespecified reference set of } u \,, \tag{48}$$

where the reference standard deviations are taken from the diagonal elements of C. Obviously, the mutual dependencies of the parameters are not taken into account in this comparison.

Following Baarda (1973), we therefore could require any function (value) $s(\hat{\boldsymbol{\theta}})$ of the parameters to be more precise when determined with the covariance matrix Σ than when determined with the criterion matrix C. This can be formally written as

$$\sigma_s^{(\Sigma)} \leq \sigma_s^{(C)} \,. \tag{49}$$

This is visualized in Fig. 5, left.

With the Jacobian $\boldsymbol{e} = \partial f / \partial \hat{\boldsymbol{\theta}}$, this leads to the requirement $\boldsymbol{e}^{\mathsf{T}} \Sigma \boldsymbol{e} \leq \boldsymbol{e}^{\mathsf{T}} C \boldsymbol{e}$ or to

$$r(\boldsymbol{e}) = \frac{\boldsymbol{e}^{\mathsf{T}} \Sigma \boldsymbol{e}}{\boldsymbol{e}^{\mathsf{T}} C \boldsymbol{e}} \leq 1 \,. \tag{50}$$

Therefore the maximal eigenvalue $\lambda_{\max}$ of the generalized eigenvalue problem

$$\Sigma \boldsymbol{e} = \lambda C \boldsymbol{e} \tag{51}$$

needs to be less than 1,

$$\lambda_{\max}(C^{-1}\Sigma) \leq 1 \,. \tag{52}$$

The analysis can also be performed on a sub-vector of $\hat{\boldsymbol{\theta}}$. If the parameters are constrained, a regular S-transformation needs to be performed to arrive at two regular and comparable covariance matrices, see Remarks on the Equivalence of Gauge- or S-transformations and Reducing Homogeneous Coordinates, Subsection 15.3.2.

Distance of two covariance matrices. We can also determine the average distance of two $U \times U$ covariance matrices using the eigenvalues λ_u of $C^{-1}\Sigma$, which can be interpreted as the ratios of variances determined from Σ and C, respectively, and averaging their deviations from 1. By taking logarithms, we arrive at the average deviation of the ratios of the variances

$$d^2(\Sigma, C) = \frac{1}{U} \sum_{u=1}^{U} \log^2 \lambda_u(\Sigma C^{-1}) \geq 0 \tag{53}$$

from 1.

This can be shown to be a metric between two covariance matrices (Förstner and Moonen, 1999). From this we can determine the average deviation $d/2$ of the ratio of the standard deviations from 1, a value of e.g., 0.1 indicating the standard deviations differ by 10% on average.

Whereas $\lambda_{\max}$ in (52) tells the worst case, the squared distance d^2 in (53) tells the average logarithm of the ratio of two variances determined with Σ instead of with C. However, if the role of Σ and C are exchanged, the maximum eigenvalue will be replaced by the minimum eigenvalue, but the distance d^2 remains invariant.

Go to the beginning.

11 Classification Boundary of two Gaussians

11.1 The optimal decision boundary

The optimal decision boundary between two Gaussians is a conic in general.

Assume, we have the densities for the vector valued n-dimensional variable $\underline{\boldsymbol{x}}$ for the two classes ω_i

$$\underline{\boldsymbol{x}} \mid \omega_i = \mathrm{g}(\boldsymbol{x}; \boldsymbol{\mu}_i, \Sigma_i) \tag{54}$$

Then the best decision is given by comparing the posterior densities

$$\mathrm{P}(\omega_1 \mid \boldsymbol{x}) > \mathrm{P}(\omega_1 \mid \boldsymbol{x}) \mapsto \widehat{\omega} = \omega_1, \text{ else } \widehat{\omega} = \omega_2 \,. \tag{55}$$

This leads to criterium

$$\mathrm{p}(\boldsymbol{x} \mid \omega_1)\mathrm{P}(\omega_1) \geq \mathrm{p}(\boldsymbol{x} \mid \omega_2)\mathrm{P}(\omega_2) \mapsto \widehat{\omega} = \omega_1, \text{ else } \widehat{\omega} = \omega_2 \,. \tag{56}$$

or using the logarithm of the likelihood functions $L(\boldsymbol{x} \mid \omega_i) = \ln \mathrm{p}(\boldsymbol{x} \mid \omega_i)$ to the decision function

$$d(\boldsymbol{x}) = L(\boldsymbol{x} \mid \omega_1) + \mathrm{P}(\omega_1) - L(\boldsymbol{x} \mid \omega_2) - \ln \mathrm{P}(\omega_2) \underset{\omega_2}{\overset{\omega_1}{\lessgtr}} 0 \,. \tag{57}$$

For Gaussians we have the log-likelihood function

$$L(\boldsymbol{x} \mid \omega_1) = -\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu}_i)^{\mathsf{T}}\Sigma_i^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_i) - \frac{n}{2}\ln 2\pi - \frac{1}{2}\ln |\Sigma_i| \tag{58}$$

Hence, the decision function reads as

$$d(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{\mu}_2)^{\mathsf{T}}\Sigma_2^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_2) - (\boldsymbol{x} - \boldsymbol{\mu}_1)^{\mathsf{T}}\Sigma_1^{-1}(\boldsymbol{x} - \boldsymbol{\mu}_1) + c \underset{\omega_2}{\overset{\omega_1}{\lessgtr}} 0 \tag{59}$$

with the constant

$$c = \ln \frac{\mathrm{P}(\omega_1)}{\mathrm{P}(\omega_2)} + \ln \frac{|\Sigma_2|}{|\Sigma_1|} \,, \tag{60}$$

which is a quadratic function in $\boldsymbol{x}$, hence a conic. Its form depends on the quadratic part

$$\boldsymbol{x}^{\mathsf{T}}(\Sigma_2^{-1} - \Sigma_1^{-1})\boldsymbol{x} \tag{61}$$

It is a linear decision boundary, in case $\Sigma_1 = \Sigma_2$.

In the animation we assume

$$c = 0 \quad \text{or} \quad \frac{\mathrm{P}(\omega_2)}{\mathrm{P}(\omega_1)} = \frac{|\Sigma_2|}{|\Sigma_1|} \,, \tag{62}$$

which leads to Gaussians with the same height, such that the standard ellipses represent contours of the same density. We draw as background-color the posterior probability

$$\mathrm{P}(\omega_1|\boldsymbol{x}) = \sigma(d(\boldsymbol{x})) \approx \frac{1}{2}(1 + \mathrm{erf}(d(\boldsymbol{x})), \tag{63}$$

see Jordan (1995), large values shown in blue, small values in yellow. Especially, we use a polynomial approximation

$$\mathrm{erf}(x) \approx \frac{x}{|x|}\left(1 - \frac{1}{(1 + 0.278393|x| + 0.230389x^2 + 0.000972|x|^3 + 0.078108x^4)^4}\right) \tag{64}$$

with a maximum error of $5 \ 10^{-4}$, which avoids numerical problems for large x.[2] The decision function are all $\boldsymbol{x}$ where $d(\boldsymbol{x}) < 0.1\sqrt{d + 0.3}$ with the Mahalanobis distance d of the two densities from (80). The variable threshold leads to an approximately constant width of the graph of the decision boundary and avoids a singularity for $d = 0$.

Remark: The first version of the animation geometrically constructed the decision boundary by taking that conic, which passes through five intersection points of corresponding contour lines. This often lead to a singularity for $\boldsymbol{\Sigma}_1 \approx \boldsymbol{\Sigma}_2$ since Cinderella does not provide a conic for strictly collinear points. ◇

11.2   The Bayes Error

The Bayes error of a classifier is the sum of the misclassification errors. Let the function

$$r(\boldsymbol{x}) = \min\{\mathrm{P}(\omega_1 \mid \boldsymbol{x}), \mathrm{P}(\omega_2 \mid \boldsymbol{x})\} \tag{65}$$

Then the Bayes error is defined

$$\begin{aligned} \epsilon &= \mathbb{E}(r(\underline{\boldsymbol{x}})) = \int r(\boldsymbol{x})\mathrm{p}(dx)\mathrm{d}\boldsymbol{x} & (66) \\ &= \mathrm{P}_1 \int_{R_1} \mathrm{p}_1(\boldsymbol{x})\mathrm{d}\boldsymbol{x} + \mathrm{P}_2 \int_{R_2} \mathrm{p}_2(\boldsymbol{x})\mathrm{d}\boldsymbol{x} & (67) \\ &= \mathrm{P}_1\epsilon_1 + \mathrm{P}_2\epsilon_2 & (68) \end{aligned}$$

where the $\mathrm{P}_i\epsilon_i$ are the probabilities for making an error when deciding for class ω_i. Since the regions R_i are separated by a conic there is not analytical way to determine the Bayes error. However, there exist upper bounds.

11.3   The Bhattacharyya distance

An upper bound for the Bayes error is the Bhattacharyya distance ϵ_u:

$$\epsilon \leq \epsilon_u \tag{69}$$

with

$$\epsilon_u = \sqrt{\mathrm{P}_1\mathrm{P}_2} \int \sqrt{\mathrm{p}_1(\boldsymbol{x})\mathrm{p}_2(\boldsymbol{x})}\mathrm{d}\boldsymbol{x} \tag{70}$$

Observe, the Bhattacharyya distance is no metric, though the name *distance* suggests this.

In case $\mathrm{P}_1 = \mathrm{P}_2$ and Gaussian densities we obtain

$$\epsilon_u = \epsilon_{u,\mu} + \epsilon_{u,\Sigma} \tag{71}$$

with the contribution due to the difference of the mean values

$$\epsilon_{u,\mu} = \frac{1}{8}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^\mathsf{T} \left(\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right)^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1), \tag{72}$$

which is a quarter of the squared Mahalanobis distance of the mean values.

The contribution due to the difference of the covariance matrices is given by

$$\epsilon_{u,\Sigma} = \frac{1}{2}\ln\frac{\left|\frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2}\right|}{\sqrt{|\boldsymbol{\Sigma}_1| \ |\boldsymbol{\Sigma}_2|}} \tag{73}$$

Assume the n-dimensional case

$$\boldsymbol{\Sigma}_1 = \boldsymbol{I}_n \quad \text{and} \quad \boldsymbol{\Sigma}_2 = \boldsymbol{\Lambda} = \mathrm{Diag}([\lambda_j]) \quad \text{with} \quad j = 1, ..., n \tag{74}$$

This can always be achieved by simultaneous diagonalization of two general covariance matrices, see Fig. 6. One can show, that if $\boldsymbol{\Lambda}$ and A are the eigenvalue and eigenvector

[2]see https://en.wikipedia.org/wiki/Error_function
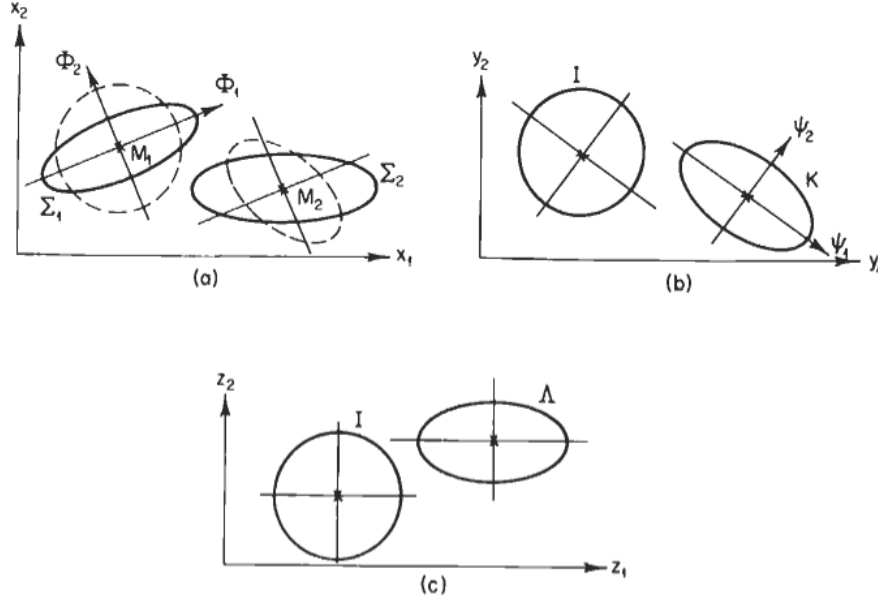
Figure 6: Simultaneous Diagonalization, from (Fukunaga, 1990, p. 32). (a) Original situation with major axes of Σ_1 in the coordinate system S_x. (b) After whitening Σ_1 and applying the linear transformation to Σ_2, this leads $\overline{\Sigma}_1 = I$ and to $\overline{\Sigma}_2$ described in the coordinate system S_y. (c) Rotating the coordinate system does not change the form of $\overline{\overline{\Sigma}}_1$, but allows to have the main diagonals of $\overline{\overline{\Sigma}}_2$ parallel to the coordinate system Z_z

matrices of $\Sigma_1^{-1}\Sigma_2$ from

$$\Sigma_1^{-1}\Sigma_2 A = A\mathbf{\Lambda} \tag{75}$$

then

$$A^{\mathsf{T}}\Sigma_1 A = I \quad \text{and} \quad A^{\mathsf{T}}\Sigma_2 A = \mathbf{\Lambda}. \tag{76}$$

We will use this relation later.

Hence

$$\det \Sigma_1 = 1\,, \quad \det \Sigma_2 = \prod_j \lambda_j \quad \text{and} \quad \det((\Sigma_1 + \Sigma_2)/2) = \prod_j \frac{1}{2}(1 + \lambda_j) \tag{77}$$

Therefore

$$\epsilon_{u,\Sigma} = \frac{1}{2} \ln \prod_j \frac{1}{2}\left(\sqrt{\lambda_j} + \frac{1}{\sqrt{\lambda_j}}\right) \geq 0 \tag{78}$$

The lower limit only is reached in case all $\lambda_j = 1$. We also can write the covariance related error as

$$\epsilon_{u,\Sigma} = \frac{1}{2} \sum_{j=1}^{n} \ln \frac{1}{2}\left(\sqrt{\lambda_j} + \frac{1}{\sqrt{\lambda_j}}\right) \geq 0 \tag{79}$$

11.4 Metric properties of the Bhattacharyya distance

The Bhattacharyya distance is no metric. We check the two components.

11.4.1 The Mahalanobis distance is not metric

The Mahalanobis distance

$$d(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \Sigma_1\Sigma_2) = \sqrt{(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^{\mathsf{T}}(\Sigma_1 + \Sigma_2)^{-1}(\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)} \tag{80}$$

is no metric.

16

A counter example is sufficient. Take the three 1D distributions

$$p_1 = g(x; 0, 1), \quad p_2 = g(x; 1, 1), \quad \text{and} \quad p_3 = g(x; 2, a^2). \tag{81}$$

The Mahalanobis distance specializes to

$$d(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2) = \frac{|\mu_2 - \mu_1|}{\sqrt{\sigma_1^2 + \sigma_2^2}}. \tag{82}$$

Then, with $\sigma_3^2 = a^2$ we have the distances

$$d_{12} = \frac{1}{\sqrt{2}}, d_{23}(a) = \frac{1}{\sqrt{1 + a^2}}, \quad \text{and} \quad d_{13}(a) = \frac{2}{\sqrt{1 + a^2}}. \tag{83}$$

The function

$$f(a) = d_{12} + d_{23}(a) - d_{13}(a) \tag{84}$$

should be non-negative for all a in order that the triangle inequality is fulfilled. However, we have

$$f(a) = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{1 + a^2}} - \frac{2}{\sqrt{1 + a^2}} = \frac{\sqrt{1 + a^2} + \sqrt{2} - 2\sqrt{2}}{\sqrt{2}\sqrt{1 + a^2}} = \frac{\sqrt{1 + a^2} - \sqrt{2}}{\sqrt{2}\sqrt{1 + a^2}} \tag{85}$$

This only is non-negative for $a^2 \geq 1$. Hence, for all variances $\sigma_3^2 = a^2 < 1$ the triangle equation is not fulfilled.

11.4.2   The covariance component is a metric

The distance related to the covariance matrices is given by

$$d_B(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) = \sqrt{2\epsilon_{u,\mathbf{\Sigma}}} = \sqrt{\ln \frac{\left| \frac{\mathbf{\Sigma}_1 + \mathbf{\Sigma}_2}{2} \right|}{\sqrt{|\mathbf{\Sigma}_1| \ |\mathbf{\Sigma}_2|}}}. \tag{86}$$

I did not find a reference stating whether this function is a metric. However, simulations with random $n \times n$ covariance matrices ($n = 2, 4, 10$), suggest, that this is a metric between two covariance matrices.

Alternatively, a Riemann metric measuring the distance of two covariance matrices is given by

$$d_C^2(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2) = \sum_{j=1}^{n} \ln^2 \lambda_j(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2), \tag{87}$$

see Calvo and Oller (1990, eq. (5)) and Förstner and Moonen (1999).

In case of simultaneously diagonalized covariance matrices the two metrics simplify to

$$d_B^2(I, \mathbf{\Lambda}) = \sum_{i=1}^{n} \ln \left(\frac{1}{2} \left(\sqrt{\lambda_j} + \frac{1}{\sqrt{\lambda_j}} \right) \right) \quad \text{and} \quad d_C^2(I, \mathbf{\Lambda}) = \sum_{j=1}^{n} \ln^2 \lambda_j, \tag{88}$$

We now compare the two functions

$$f(x) = \ln \left(\frac{1}{2} \left(\sqrt{x} + \frac{1}{\sqrt{x}} \right) \right) \quad \text{and} \quad g(x) = \frac{1}{8} \ln^2 x^2 \tag{89}$$

graphically, see Fig. 7 The Taylor expansions at $x = 1$ up to fifth order are

$$T_f = \frac{(x-1)^2}{8} - \frac{(x-1)^3}{8} + \frac{7(x-1)^4}{64} - \frac{3(x-1)^5}{32} \tag{90}$$

and

$$T_g = \frac{(x-1)^2}{8} - \frac{(x-1)^3}{8} + \frac{11(x-1)^4}{96} - \frac{5(x-1)^5}{48} \tag{91}$$

with the difference

$$T_f - T_g = -\frac{(x-1)^4}{192} + \frac{(x-1)^5}{96}, \tag{92}$$

indicating the two functions have the same derivatives at $x = 1$ up to third order.

Obviously, both metrics, though different, are very similar.
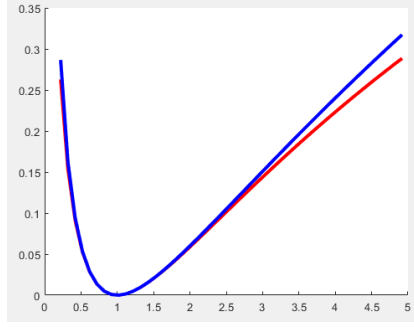
Go to the beginning.

Figure 7: Distances between two covariance matrices. Red: covariance contribution to Bhattacharyya distance. Blue: metric distance of covariance matrices

12 Uncertainty of Pose from Four Points

This text refers to two Cinderella Animations exploring the uncertainty of the spatial resection based on four control points of a horizontal square in normal position, i.e., where the ground plane and the image plane are parallel, drawn in taking position for an aerial camera, the Z-axis pointing upwards.

Two aspects of the uncertainty of the pose of the camera are analysed:

- The uncertainty of the reconstructed projection centre is visualized as a function of the controllable configuration, especially by the standard ellipse of the horizontal (X) and vertical (Z) coordinates.

- The correlation between translation and rotation parameters usually is quite high. This is demonstrated by showing the difference of a pure horizontal motion and a combined rotation and translation, which mutually compensate the influence onto the image coordinates, except for a second order effect.

As to be expected, the horizontal uncertainty of the camera pose may be quite large, however, when reprojecting image points to the surface around the control points, their horizontal accuracy is not influenced by the pose of the camera, since it can be interpreted as an interpolation between the control points.

12.1 The configuration

The basic geometry is specified by

1. height H of the camera above ground

2. principal distance c (focal length, and

3. effective viewing angle, specified by the image coordinates $\pm d$.

see Fig. 8.

12.2 The theoretical precision of the camera pose

We assume all image coordinates are measured with the same standard deviation σ.

With half the effective viewing angle

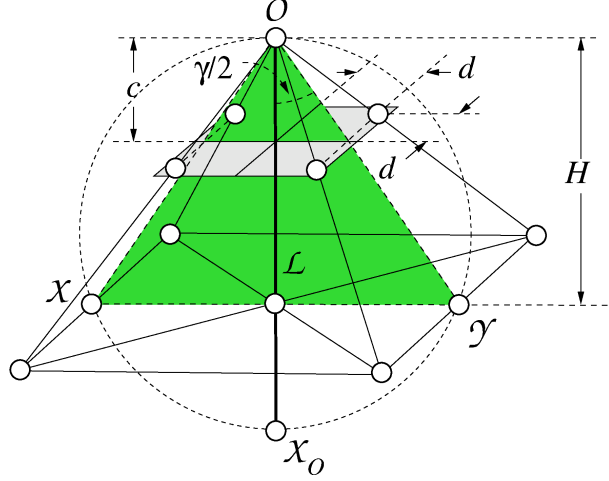$$\frac{\gamma}{2} = \arctan \frac{d}{c}, \tag{93}$$

18

Figure 8: Standard configuration with four coplanar points. The optical axis $\mathcal{L}$ is an uncertain 3D line. The Z-axis is assumed to point upwards. The line has its centre at $\mathcal{X}_O$, where the uncertainty of points on $\mathcal{L}$ across the line is minimum. When arguing in the XZ-plane, this point is the intersection of the circle through $\mathcal{O}$ and the two midpoints $\mathcal{X}$ and $\mathcal{Y}$ of the ground square on the X-axis. Here, the camera is shown in viewing position with the image plane being $Z = c$ with $c < 0$, whereas in the animation it is shown in taking position $(c > 0)$

we have the accuracy of the parameters of the pose $\mathcal{O}(X, Y, Z)$ of the camera

$$\sigma_X = \sigma_Y \quad = \quad \frac{\sqrt{2}}{2} \frac{Z_O}{c} \sqrt{1 + \frac{1}{\sin^4 \frac{\gamma}{2}}} \quad \sigma_0 \tag{94}$$

$$\sigma_Z \quad = \quad \frac{Z_O}{2d} \ \sigma_0 \tag{95}$$

$$\sigma_\omega = \sigma_\phi \quad = \quad \frac{\sqrt{2}}{2} \frac{c}{d^2} \ \sigma_0 \tag{96}$$

$$\sigma_\kappa \quad = \quad \frac{1}{2d} \ \sigma_0 \,. \tag{97}$$

and the correlation between the horizontal position and the angles

$$\rho_{X\phi} = -\rho_{Y\omega} = \frac{1}{\sqrt{1 + \sin^4 \frac{\gamma}{2}}} \,, \tag{98}$$

see Förstner and Wrobel (2016, p. 532).

The viewing axis $\mathcal{L}$ is an uncertain line, whose standard hyperboloid, when projected into the XZ-plane, is shown in the animation.

Due to circular symmetry around the Z-axis, the uncertain optical axis has a single point $\mathcal{X}_O$ where its uncertainty across the line is minimum. The distance $\overline{\mathcal{OX}_O}$ is identical with the diameter of the circle through the projection centre $\mathcal{O}$ and, say argumenting in the XZ-plane, the two midpoints $\mathcal{X}$ and $\mathcal{Y}$. This is plausible, since rotation the camera around $\mathcal{X}_O$ does not change the angle under which these two points are seen by the camera.

This property can be explored with the Cinderella animation on the correlation. There such a combined motion, an in a first order approximation horizontal motion on the circle while rotating the camera to fix the view to the point $\mathcal{X}_O$, can be realized. It shows, that specified tolerance circles for the image coordinates cohere with a hyperbolic tolerance region around the viewing line.

The low uncertainty of the projection line in the vicinity of the given control points indicates, that - given the surface around these control points and an image point, its re-

projection to the ground is quite precise. This is plausible, as its XY-coordinates actually are interpolated between the control points, and the – possibly large – uncertainty of the position of the camera has no effect.

Go to the beginning.

13   Epipolar Geometry of the Normal Case

The animation allows to explore the epipolar geometry for the case that the relative rotation between the two cameras is zero, i.e., the two image planes are parallel.

The configuration can be interactively controlled, by changing:

1. The principal distance c (focal length), including its sign, hence, allowing to choose the taking position or the viewing position of the two cameras.

2. The length of the basis B_X, related to a world camera coordinate parallel to the coordinate system of the first camera.

3. The heights Z' and Z'' of the two cameras, together with the control of B_X allowing to realize sideward and forward motion.

4. The 3D position of a scene point P is given by its image points P' in the first image and its depth Z.

5. For making the imaging process visually plausible, one may change the size d of the image.

6. For better visualization the complete 3D configuration may be viewed by changing the azimuth and elevation angle of the viewing direction.

The animation shows all elements of the epipolar geometry, especially all geometric elements sitting on the epipolar plane:

1. the two projection centres O' and O'',

2. the scene point P,

3. its image points P' and P'',

4. the two projection rays L' and L'',

5. the two epipolar lines e' and e'', and

6. the two epipoles E' and E''.

The variability allows to explore the geometry in detail, but also to choose a perspective which may be used as a template for a generic but fixed graphic.

Go to the beginning.

14   Depth of Field

The animation allows to explore the depth of field of a real camera with a finite aperture, which provides insight into the range of distances an object needs to have in order that its image shows a limited blur.

Though Wikipedia provides a nice description of the geometry, including deriving the necessary equations, and there exist apps which allow a direct calculation of the depth of field, it might be useful to be able to explore the geometry interactively, though the animation only working in a limited range of real configurations, especially w.r.t. the real lengths of focal length and distance to object.

If the camera would be a ideal pin hole camera with an infinitely small pinhole, all points of the scene could be sharply mapped to some image plane - at the cost of an

infinite exposure time and the assumption the scene and the camera would not perform any motions. Real pinhole cameras have a pinhole with a non-zero diameter which needs to be very small to achieve sharp images. Real cameras use lens systems to allow for a larger hole while the same time producing sharp images – under certain conditions, to be discussed.

The geometry of the optical projection with an ideal lens (without distortions) enforces the relation of the scene point P and its image point P' to follow two constraints:

1. The collinearity of the scene point P, the image point P' and the centre O of the (ideal) lens.

2. Rays parallel to the optical axis converge and intersect in the camera side focal point F, whose distance to the optical centre O is the focal length, defined by the optical characteristics of the lens.

As a consequence, points P which are not at infinity map to image points P' which have a larger distance to the optical centre O than the focal point F. Therefore focussing, manual or automatic, is required when taking real images. Since the mapping is a rigorous perspective one, planes in the scene are mapped to planes in the camera space. For reasons of construction, the image plane usually is perpendicular to the optical axis. This is, why only scene points on a plane, say A, perpendicular to the optical axis can be simultaneously be mapped to an adequately chosen (by focussing) image plane, say A'.

For a given focussing, i.e., a fixed image plane A', all scene points outside the corresponding scene plane are not mapped sharply, i.e., show a certain blur. If one allows a certain, small, blur to be acceptable, e.g., below 0.05 mm, or below 1 pixel, then points in the vicinity of the scene plane A can be mapped to the image plane within this tolerance. Hence, there is a near limit and a far limit for scene points which can be mapped with acceptable blur. The range of distances between near and far limit is called the depth of field.

The depth of field, i.e., the range between near and far limit depends on the following three parameters:

1. the focal length f, usually given in mm or in pixel,

2. the allowable blur σ, again given in mm or in pixel,

3. the diameter of the aperture a, usually given via the f-stop s, being the number by which the focal length needs to be divided to get a, i.e., we have $a = f/s$, and

4. the intended distance Z of the scene point to the camera, which practically is controlled by the focussing.

The animation allows to control

1. the focal length f

2. the aperture, via the f-stop,

3. the maximum blur,

4. the horizontal position of the scene point, and in addition, for increasing plausibility of the drawing,

5. the height Z of the scene point,

6. the size d of the image area, and

7. the position of the camera via shifting F (for constructive reasons changing the distance to the scene point).

Go to the beginning.

15 Trifocal Constraint for Image Triplet

The animation addresses the geometry of the image triplet. The following text is taken from (Förstner and Wrobel, 2016, Sect. 14.1).

"There are several reasons to analyse image triplets:

- Given three images, say 1, 2, and 3, the relative orientation of two pairs of them, say $(1, 2)$ and $(2, 3)$, does not tell us anything about the mutual scale of the resulting photogrammetric models.

- The relative orientation of three images gives constraints on all image coordinates involved. This is in contrast to the relative orientation of an image pair, which only gives constraints in one direction, in the normal case for the y-coordinates, while the x-coordinates cannot be checked at all.

- The relative orientation of image triplets can be based on both *corresponding points* and *corresponding lines*, in contrast to image pairs, where corresponding lines give no constraint for the relative orientation, see Fig. 9. Analogously to the image pair,
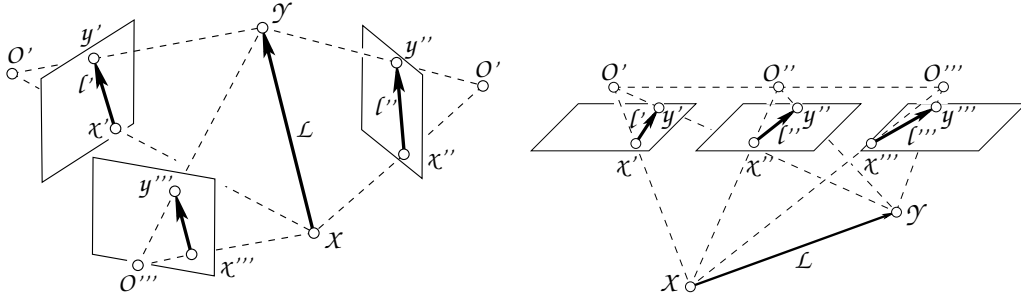

Figure 9: The relative orientation of the image triplet is fully captured in the trifocal tensor. **Left:** general configuration. **Right:** Collinear projection centres. The trifocal tensor allows us to establish constraints between corresponding points (χ', χ'', χ''') and lines (l', l'', l''') in three images. It also allows us to predict points and lines from two given points and lines in all feasible configurations without determining the 3D point $\mathcal{X}$ or 3D line $\mathcal{L}$. This is possible also if the epipolar geometry would not be sufficient, namely if the 3D points and the three projection centres $(\mathcal{O}', \mathcal{O}'', \mathcal{O}''')$ are coplanar as in the important case of collinear projection centres as shown in the right figure

the constraints between corresponding points and lines, respectively are linear in homogeneous coordinates of the entities. In addition, the constraints are linearly dependent on 27 parameters, collected in what is called the *trifocal tensor* T. The trifocal tensor is a $3 \times 3 \times 3$ array representing the complete geometry of the image triplet, in full analogy to the nine parameters of the fundamental matrix F which captures the complete geometry of the image pair.

- The prediction of points and lines from two images in the third can also be based on the trifocal tensor. This prediction is linear in its parameters and in the homogeneous coordinates of the points and lines. Similarly to the determination of the epipolar line in the case of the image pair, the prediction of a point or line in a third image can be performed without first determining the 3D point or 3D line.

  In the important situation of collinear projection centres (Fig. 9, right), the prediction of the point χ''' in image 3 based on the points χ' and χ'' in the first two images, obtained from the epipolar geometry of image pairs $(1, 2)$ and $(1, 3)$, is not possible using corresponding epipolar lines, as they are identical in the third image and do not give a unique prediction. Prediction using the trifocal tensor does not have this deficiency, as it implicitly works with the 3D point.

The Coplanarity Constraints of the Image Triplet

We start with the generic situation where the three projection centres are not collinear. The three projection centres O', O'', and O''' then uniquely span the *trifocal plane*, see Fig. 10

Let the three projections be given by

$$\chi' = \mathcal{P}'(\mathcal{X}): \quad \mathbf{x}' = \mathsf{P}_1 \mathbf{X} \tag{99}$$

$$\chi'' = \mathcal{P}''(\mathcal{X}): \quad \mathbf{x}'' = \mathsf{P}_2 \mathbf{X} \tag{100}$$

$$\chi''' = \mathcal{P}'''(\mathcal{X}): \quad \mathbf{x}''' = \mathsf{P}_3 \mathbf{X}, \tag{101}$$

with projection matrices $\mathsf{P}_t = \mathsf{K}_t \mathsf{R}_t [I_3 | -\boldsymbol{Z}_t] = [A_t | \boldsymbol{a}_t], t \in \{1, 2, 3\}$.
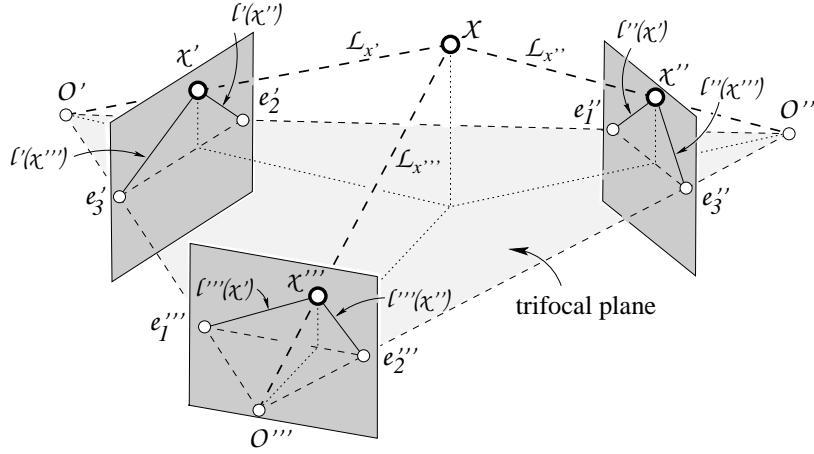

Figure 10: Geometry of the image triplet with points. Three projection centres O', O'', and O'''. Three image points χ', χ'', and χ'''. Six epipoles, e.g., $e_3' = \mathcal{P}'(O''')$. Three projection rays $\mathcal{L}_{x'}$, $\mathcal{L}_{x''}$, and $\mathcal{L}_{x'''}$. Six epipolar lines, e.g., $l'(\chi''') = \mathcal{P}'(\mathcal{L}_{x'''})$. If the 3D point $\mathcal{X}$ is outside the trifocal plane, the two epipolar lines in each image have a unique intersection point, namely the image point of $\mathcal{X}$. Otherwise, if the 3D point is on the trifocal plane, the two epipolar lines in each image are identical and do not yield a unique intersection point

Remark: Elements of the three images are denoted either by primes or by numbers. If these are variables where the order does not matter, we use $t = 1, 2, 3$. If we refer to pairs or triplets of indices, where the order is important, we use $i, j, k \in \{1, 2, 3\}$ in order to avoid double indices. ⋄

In the general case, the prediction of points can be based on the epipolar geometry of two image pairs. We have three fundamental matrices,

$$\mathsf{F}_{ij} = A_i^{-\mathsf{T}} \mathsf{S}_{\mathsf{b}_{ij}} A_j^{-1} \qquad (i, j) \in \{(1, 2), (2, 3), (3, 1)\}, \tag{102}$$

with $\mathbf{b}_{ij} = \boldsymbol{Z}_j - \boldsymbol{Z}_i$.

Let the two points χ'' and χ''' in the second and the third images be given. Then the intersection of the epipolar lines in the first image,

$$\mathbf{l}'(\mathbf{x}'') = \mathsf{F}_{12}\mathbf{x}'' \qquad \mathbf{l}'(\mathbf{x}''') = \mathsf{F}_{13}\mathbf{x}''', \tag{103}$$

yields the predicted point $\chi' = l'(\chi'') \times l'(\chi''')$, thus

$$\mathbf{x}' = \mathsf{F}_{12}\mathbf{x}'' \times \mathsf{F}_{13}\mathbf{x}'''. \tag{104}$$

Similar expressions can be found for predicting image points in the other two images.

This method of prediction only works if the 3D point and the three projection centres are not coplanar, or − equivalently − if the 3D point is not on the trifocal plane. Then the two projection lines $\mathcal{L}_{x''}$ and $\mathcal{L}_{x'''}$ lie in the trifocal plane, the two epipolar lines $l'(\chi'')$

and $l'(\chi''')$ are identical, and therefore the intersection point $\chi' = l'(\chi'') \cap l'(\chi''')$ is not unique. Practically, even 3D points close to the trifocal plane cause numerical uncertainties in this type of prediction.

Unfortunately, this unfavourable situation occurs often, especially in image sequences, where consecutive projection centres are collinear or nearly collinear. If the three projection centres are collinear, they and the 3D points are always coplanar, so the prediction of an image point using the epipolar geometry of two pairs of images fails for all 3D points. This can easily be visualized: Any three projection rays which are coplanar lead to three 3D intersection points, $\mathcal{X}_{12} = \mathcal{L}_{x'} \cap \mathcal{L}_{x''}$, $\mathcal{X}_{23}$ and $\mathcal{X}_{31}$. These three points need not be identical. Thus the epipolar constraints are fulfilled, in spite of the fact that the three image points are not corresponding. The epipolar constraints thus are only necessary but not sufficient conditions for the correspondence of three image points.

But a prediction of a point and thus a constraint can be achieved. This can be seen by first determining the 3D point, e.g., $\mathcal{X}_{23}$, by triangulation and then projecting it into the other image, which results in the constraint $\chi' \equiv \mathcal{P}'(\mathcal{X}_{23})$. ... "

The animation is the basis for Fig. 10 and allows to move the point $\mathcal{X}$ in 3D, especially into the trifocal plane and visually verify the above mentioned geometric relations.

Go to the beginning.

16 Crossratio

Two animations deal with the cross ratio, an invariant of projective mappings, or homographies.

1. The first, on Crossratio, allows to explore the principles of invariants of four colliner points and four concurrent lines.

2. The second, on rectification, is an application using the invariance of the crossratio for images of planar objects using what is called the paper-strip-method.

The following text on the basics of the cross-ratio is taken from Förstner and Wrobel (2016), sect. 6.4.3.

"Angles between lines and the ratio of distances between collinear point pairs are not preserved under collineations. This can be seen in the example of Fig. 11.

However, four collinear points have an invariant under projective transformation: the cross ratio, which can be transferred to four concurrent lines and also to a pencil of four concurrent planes. The cross ratio can be used to describe more general configurations by their invariants.

16.1 Cross Ratio of Four Collinear Points

The basic configuration for the cross ratio is four collinear points.

Definition 16.1: *Cross ratio of four collinear points.* The cross ratio $\mathrm{CR}(\chi_1, \chi_2, \chi_3, \chi_4)$ of four collinear points with line coordinates (x_1, x_2, x_3, x_4) is defined as

$$\mathrm{CR}(\chi_1, \chi_2, \chi_3, \chi_4) = \frac{x_1 - x_3}{x_2 - x_3} : \frac{x_1 - x_4}{x_2 - x_4} \qquad (105)$$

If a point is at infinity, the rule $\infty/\infty = 1$ is used. $\diamond$

We now have the following theorem:

Theorem 16.1: *Invariance of cross ratio.* The cross ratio $\mathrm{CR}(\chi_1, \chi_2, \chi_3, \chi_4)$ of four collinear points is invariant under collineations $\mathcal{H}$. Thus if $\chi_i' = \mathcal{H}(\chi_i), i = 1, 2, 3, 4$, then

$$\mathrm{CR}(\chi_1', \chi_2', \chi_3', \chi_4') = \mathrm{CR}(\chi_1, \chi_2, \chi_3, \chi_4). \qquad (106)$$
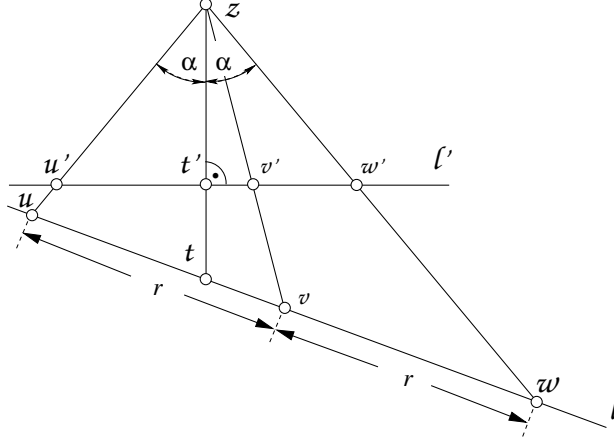
Figure 11: Non-invariance of the distance ratio under perspectivity. The three collinear points u, v, and w on line ℓ having the same distance $r = \overline{uv} = \overline{vw}$ are mapped to the line ℓ' via the perspective projection with projection centre z, leading to the points u', v', and w'. Obviously, the distances between the image points are different: $\overline{u'v'} \neq \overline{v'w'}$. The midpoint v of $\overline{uw}$ is not mapped to the midpoint v' of $\overline{u'w'}$. Now, imagine the point v is the centre of a circle with radius $\overline{uv}$ lying in the plane through ℓ perpendicular to the drawing plane. Its image in the plane through ℓ', again orthogonal to the drawing plane, will be an ellipse: Obviously, the centre of an ellipse, which is the image of a circle, is not the image of the centre of that circle

The proof exploits the fact that each of the four indices appears twice in the cross ratio, once in the numerator, once in the denominator.

Given four points, there are 24=4! permutations for their sequence. Thus it is possible to define 24 different cross ratios. However, six of them are distinct generally, but mutually functionally dependent. If one cross ratio is λ, we have the six different values for cross ratios of four points:

$$\lambda, \quad \frac{1}{\lambda}, \quad 1 - \lambda, \quad \frac{1}{1 - \lambda}, \quad \frac{\lambda}{1 - \lambda}, \quad \frac{1 - \lambda}{\lambda}. \tag{107}$$

Example 16.1: *Inferring distances between collinear points.* Assume, in Fig. 12,
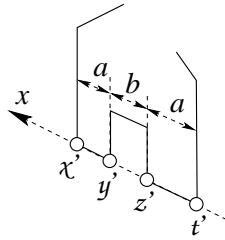

Figure 12: Example for the use of the invariance of the cross ratio. Due to the assumed symmetry – the distance of the door from its two neighbouring facade borders is identical to some unknown value a and the assumed knowledge about the width of the building we can infer the true width of the door in the scene from the image points

we have observed the collinear image points x', y', z' and t' and know that the door is in the centre of the facade, which has a width of $w = 10$ m. Then, using the cross ratio we can determine the width b of the door and its distance a from the right and left wall from the two equations:

$$\mathrm{CR}(x, y, z, t) = \frac{a + b}{b} : \frac{2a + b}{a + b} = \mathrm{CR}(x', y', z', t') \qquad 2a + b = 10 \,[\mathrm{m}] \tag{108}$$

where the cross ratio $\mathrm{CR}(x', y', z', t')$ can be determined from image measurements. ◇

Mirror symmetric configurations in a plane are characterized by (1) the existence of a symmetry axis and (2) the property that lines through symmetric point pairs are parallel.

Now, we regard points on such a line: a point χ, its mirror point χ', the mid-point y of both, which is the point of symmetry, and the point χ_∞. Then their cross ratio is $\mathrm{CR}(\chi, \chi', y, \chi_\infty) = -1$ and we say the configuration is *harmonic*. We therefore use the definition:

Definition 16.2: *Harmonic points.* Four points on a line are harmonic if their cross ratio is -1. $\diamond$

16.2 Cross Ratio of Four Concurrent Lines

The cross-ratio transfers to a pencil of four rays.

Definition 16.3: *Cross ratio of four concurrent lines.* The cross ratio $\mathrm{CR}(l_1, l_2, l_3, l_4)$ of four concurrent lines with directions $(\phi_1, \phi_2, \phi_3, \phi_4)$ is defined as

$$\mathrm{CR}(l_1, l_2, l_3, l_4) = \frac{\sin(\phi_1 - \phi_3)}{\sin(\phi_2 - \phi_3)} : \frac{\sin(\phi_1 - \phi_4)}{\sin(\phi_2 - \phi_4)} \,. \tag{109}$$

$\diamond$

We can see this from Fig. 13: The coordinate differences $x_i - x_j$ on line m and the sine of the direction differences $\sin(\phi_i - \phi_j)$ are related by the area F of the triangle $(z\ \chi_i\ \chi_j)$ with sides s_i, s_j via $2F = h(x_i - x_j) = s_i s_j \sin(\phi_i - \phi_j)$, which allows us to develop (109) from (105).
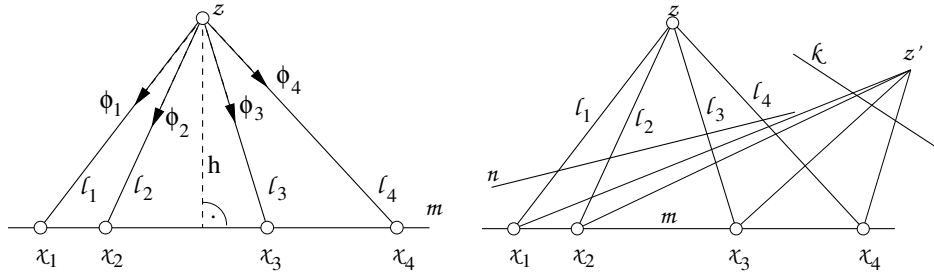

Figure 13: Cross ratio of points and lines. **Left:** The collinear points $\chi_i, i = 1, 2, 3, 4$ and the concurrent lines $l_i, i = 1, 2, 3, 4$, are related by the central point z, having the distance h from the line m. The set of points and the set of lines have the same cross ratio. **Right:** Concurrent lines allow us to transfer the cross ratio from the original points $\chi_i, i = 1, 2, 3, 4$ to the intersection points of the lines $l_i, i = 1, 2, 3, 4$ with the line n or via the lines through z' to the intersection points with the line k

Given a fifth line m not identical to l_i and not passing through their intersection point, the cross ratio of the four concurrent lines l_i can be computed by

$$\mathrm{CR}(l_1, l_2, l_3, l_4) = \frac{|\mathbf{m}, \mathbf{l}_1, \mathbf{l}_2|}{|\mathbf{m}, \mathbf{l}_3, \mathbf{l}_2|} : \frac{|\mathbf{m}, \mathbf{l}_1, \mathbf{l}_4|}{|\mathbf{m}, \mathbf{l}_3, \mathbf{l}_4|} \,. \tag{110}$$

This cross ratio is also the cross ratio $\mathrm{CR}(\chi_1, \chi_2, \chi_3, \chi_4)$ of the four intersection points $\chi_i = l_i \cap m$ of the lines l_i with m.

Proof: Without loss of generality, we can choose the line $\mathbf{m} = [0, 1, 0]$ to be the x-axis, and the intersection point of the lines not to lie on the y-axis, e.g. at $\mathbf{x} = [0, 1, 1]^\mathsf{T}$. The intersection points $\chi_i = l_i \cap m$ of the lines with the x-axis are assumed to have coordinates x_i. Then the determinants are

$$|\mathbf{m}, \mathbf{l}_i, \mathbf{l}_j| = \left| \mathbf{m}, \mathsf{S}(\mathbf{x}) \begin{bmatrix} x_i \\ 0 \\ 1 \end{bmatrix}, \mathsf{S}(\mathbf{x}) \begin{bmatrix} x_j \\ 0 \\ 1 \end{bmatrix} \right| = \begin{vmatrix} 0 & 1 & 1 \\ 1 & x_i & x_j \\ 0 & -x_i & -x_j \end{vmatrix} = x_j - x_i \,, \tag{111}$$

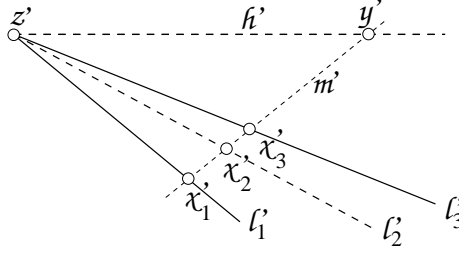which completes the proof. $\diamond$

Figure 14: Horizon of a plane with equidistant coplanar lines

Example 16.2: *Image of the horizon from the image of three equidistant parallel lines.* Given is an image with three lines l_i', $i = 1, 2, 3$, which in 3D are coplanar, parallel, and equidistant, see Fig. 14. The task is to determine the image h' of the line at infinity of the plane. We give two solutions, a constructive one and an algebraic one:

1. We first determine the intersection point $z' = l_1' \cap l_2'$ of two of the image lines, l_i'. z' is the image of the point at infinity of the set of 3D lines. Then we take an arbitrary line m' passing through the three lines l_i', leading to three intersection points $x_i' = m' \cap l_i'$. We now construct a point $y' \in m'$ such that $\text{CR}(x_1', x_3', x_2', y') = -1$. Then the sought line is $h' = z' \wedge y'$.

2. The construction can be used to derive a closed form solution (Schaffalitzky and Zisserman, 2000),
$$\mathbf{h}' = [\mathbf{m}', \mathbf{l}_1', \mathbf{l}_2']\mathbf{l}_3' - [\mathbf{m}', \mathbf{l}_2', \mathbf{l}_3']\mathbf{l}_1' \,, \tag{112}$$
where m' is an arbitrary line not identical to l_i and not passing through the intersection point of the three lines.

$\diamond$

..."

16.3 Image-Rectification with the Paper-Strip-Method

The paper-strip-method aims at transferring a point P, say from an image to a map, based on four given corresponding points $(ABCD)$ and $(A'B'C'D')$, exploiting the invariance of the cross-ratio of concurrent lines relating the point to the four given points. In the animation we have the two identities

$$\text{CR}(BA, BD|BC, BP) = \text{CR}(B'A', B'D'|B'C', B'P') \quad \text{and} \tag{113}$$
$$\text{CR}(CA, CD|CB, CP) = \text{CR}(C'A', C'D'|C'B', C'P') \tag{114}$$

for the directions of the concurrent lines from B and B', and C and C' repectively. They are realized by the markers on the paper strip: they are used to transfer the cross-ratios $\text{CR}(Ba, Bd|Bc, Bp)$ and $\text{CR}(CA, CD|CB, CP)$ in the image to the correspoinding corss-ratios anforcing identities $\{a, c, d, p\} = \{a', c', d', p'\}$ and $\{a, b, d, p\} = \{a', b', d', p'\}$ of the markers by placing the paperstrip in the map adequately.

Go to the beginning.

17 2D/3D Motion as Rotation/Screw

Any 2D motion $\mathbf{x}' = R\mathbf{x} + \mathbf{t}$ can be represented as a rotation around some point $\mathbf{x}_0$. In both cases the number of degrees of freedom is three, wither the translation (2) and the rotation (1) or the point of rotation (2) and the rotation (1). The rotation around $\mathbf{x}_0$ is given by

$$\mathbf{x}' = R(\mathbf{x} - \mathbf{x}_0) + \mathbf{x}_0 = R\mathbf{x} + (\mathbf{x}_0 - R^\mathsf{T}\mathbf{x}_0) = R\mathbf{x} + (I - R^\mathsf{T})\mathbf{x}_0 \tag{115}$$

which leads to the relations

$$\mathbf{t} = (I - R^\mathsf{T})\mathbf{x}_0 \quad \text{and} \quad \text{if } R \neq I \text{ then } \mathbf{x}_0 = (I - R^\mathsf{T})^{-1}\mathbf{t} \,. \tag{116}$$

Hence for pure translation the centre of rotation is not defined.

Similarly, any 3D motion $\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$ can be represented as a screw motion, which combines a rotation around a 3D line and a translation by s along this line. In both cases the number of degrees of freedom is six, either the translation (3) and the rotation (3), or the 3D line (4) and the rotation angle (1) and the translation along the line (1).

With the normalized rotation axis $\boldsymbol{r}$, with $|\boldsymbol{r}| = 1$, following $\boldsymbol{r} = \boldsymbol{R}\boldsymbol{r}$, we have the rotation around the 3d line, specified by $(\boldsymbol{x}_0, \boldsymbol{r})$

$$\boldsymbol{x}' = R_{r,\theta}(\boldsymbol{x} - \boldsymbol{x}_0) + \boldsymbol{x}_0 + s\,\boldsymbol{r} = R_{r,\theta}\boldsymbol{x} + (\boldsymbol{x}_0 + s\,\boldsymbol{r} - R^{\mathsf{T}}\boldsymbol{x}_0) \tag{117}$$

Hence, we have a constraint for the translation

$$\boldsymbol{t} = s\,\boldsymbol{r} + (I - R^{\mathsf{T}})\boldsymbol{x}_0\,, \tag{118}$$

which needs to be partitioned into a component $\boldsymbol{t}_r$ parallel to $\boldsymbol{r}$ and a component $\boldsymbol{t}_p$ perpendicular to $\boldsymbol{r}$, which can be achieved using:

$$\boldsymbol{t} = \boldsymbol{t}_{\parallel} + \boldsymbol{t}_{\perp}\,, \quad \text{with} \quad \boldsymbol{t}_{\parallel} = s\,\boldsymbol{r} \quad s = \boldsymbol{t} \cdot \boldsymbol{r}\,, \quad \text{and} \quad \boldsymbol{t}_{\perp} = \boldsymbol{t} - \boldsymbol{t}_r\,. \tag{119}$$

Then the constraint for $\boldsymbol{x}_0$ is

$$\boldsymbol{t}_{\perp} = (I - R^{\mathsf{T}})\boldsymbol{x}_0 \tag{120}$$

thus

$$\boldsymbol{x}_0 = (I - R^{\mathsf{T}})^{-1}\boldsymbol{t}_{\perp} = (I - R^{\mathsf{T}})^{-1}(\boldsymbol{t} - (\boldsymbol{t} \cdot \boldsymbol{r})\boldsymbol{r}) \tag{121}$$

Again, for pure translation, the centre of the rotation is not defined.

Observe, that $\boldsymbol{x}_0$ is the point on the 3D line closest to the origin, thus only has two degrees of freedom, and that the rotation axis also only has two degrees of freedom. Together with the rotation angle θ and the translation s along the 3D line, the degrees of freedom of the screw motion is six.

The animation can be interpreted in 2D and 3D. In 2D it shows the planar motion of a rectangle. In 3D, the figure can be interpreted as a projection along the axis of rotation, such that the rotation axis passes through point $\boldsymbol{x}_0$ perpendicular to the viewing plane, the quadrangle can have arbitrary depths along the axis of rotation, and one can imagine the motion by s along this line.

Go to the beginning.

18 Double Reflection

The animation illustrates the definition of a rotation as a double reflection at two planes through the origin. The animation is found in a lecture using quaternions for representing the rotation, (slides: https://www.ipb.uni-bonn.de/html/teaching/3dcs-wf-2020/05a-3D-CS-Bsc-Rotations-as-two-Reflections-using-Quaternions.pdf, video https://www.youtube.com/watch?v=oSK3VCtr7_w).

The topic is intensely discussed: An easy approach can be found in https://www.euclideanspace.com/maths/geometry/affine/reflection/quaternion/index.htm. An animation can be found in https://demonstrations.wolfram.com/RotationAsProductOfTwoReflections/, however, not discussing the redundancy of the representation. Also the animation, similar to this one https://www.geogebra.org/m/SpzcdDS6, does not discuss the freedom in choosing the two planes.

The animation here shows how a 2D or 3D rotation around the origin can be represented by reflections at two planes passing the origin and visualizes the ambiguity of choosing a pair of mirrors.

If the angle between the normals of the two lines (in 2D) or planes (in 3D) is α, then the rotation angle is $\theta = 2\alpha$ and the rotation point/axis is the intersection point/line of

the two lines/planes. This directly can be seen from the animation: The 2D figure can also be interpreted as a projection of a 3D configuration along the rotation axis, or – equivalently – onto a plane perpendicular to the rotation axis.

A 2D rotation is defined by one parameter, namely the angle, while a 3D rotation is defined by three parameters, namely the direction of the rotation axis (2 parameters) and the rotation angle (1 parameter).

Specifying two normals for the rotation, hence, appears redundant, since in 2D one needs two angles for specifying the two normals, while in 3D one needs 4 parameters to specify the two normals.

However, as can be seen from the animation, a joint common rotation of the two planes around their intersection line does not influence the rotation. This explains the redundancy of the specification of a rotation with two normals.

Mirroring at planes with normalized normals $\boldsymbol{n}$ and $\boldsymbol{m}$ are

$$M = I - 2\boldsymbol{m}\boldsymbol{m}^\mathsf{T} \quad \text{and} \quad N = I - 2\boldsymbol{n}\boldsymbol{n}^\mathsf{T} \tag{122}$$

Concatenation yields

$$\begin{aligned} Q = NM &= (I - 2\boldsymbol{n}\boldsymbol{n}^\mathsf{T})(I - 2\boldsymbol{m}\boldsymbol{m}^\mathsf{T}) \tag{123} \\ &= I - 2\boldsymbol{n}\boldsymbol{n}^\mathsf{T} - 2\boldsymbol{m}\boldsymbol{m}^\mathsf{T} + 4\boldsymbol{n}\boldsymbol{n}^\mathsf{T}\boldsymbol{m}\boldsymbol{m}^\mathsf{T} \tag{124} \end{aligned}$$

If Q is a rotation, then it must fulfil $Q Q^\mathsf{T} = I$. We actually have

$$Q Q^\mathsf{T} = (NM)(NM)^\mathsf{T} = N(MM)N = N(I)N = I. \tag{125}$$

since the reflection matrices are symmetric and involutory, thus fulfil $M = M^\mathsf{T}$ and $M^2 = I$.[3] The vector $\boldsymbol{c} = \boldsymbol{n} \times \boldsymbol{m}$ is invariant:

$$Q\boldsymbol{c} = (I - 2\boldsymbol{n}\boldsymbol{n}^\mathsf{T} - 2\boldsymbol{m}\boldsymbol{m}^\mathsf{T} + 4\boldsymbol{n}\boldsymbol{n}^\mathsf{T}\boldsymbol{m}\boldsymbol{m}^\mathsf{T})(\boldsymbol{n} \times \boldsymbol{m}) = \boldsymbol{n} \times \boldsymbol{m} \tag{126}$$

since both vectors $\boldsymbol{n}$ and $\boldsymbol{m}$ are perpendicular to $\boldsymbol{c}$. Hence $\boldsymbol{c}$ is a point on the rotation axis, whose normalized direction vector is

$$\boldsymbol{r} = \frac{\boldsymbol{n} \times \boldsymbol{m}}{|\boldsymbol{n} \times \boldsymbol{m}|}. \tag{127}$$

Go to the beginning.

19 Scaled Rotation as Complex Multiplication

The animation allows to inspect the relation between the multiplication of two complex numbers and the scaled rotation as a special motion. The explanation can be found in the slides : There the 2D similarity transformation $\boldsymbol{x}' = \lambda R(\phi)\boldsymbol{x} + \boldsymbol{t}$ is shown to be a linear regression $z = ax + b$ using complex numbers for all entries, e.g., $\boldsymbol{x} = [x, y]^\mathsf{T} \leftrightarrow z = x + \mathrm{i}y$, identifying the role of the scaled rotation $Q = \lambda R(\phi)$ and the polar representation $a = \lambda \exp(\mathrm{i}\phi)$.

A more general application can be found in the animation on choosing the gauge of a planar polygon or in a planar geodetic network, Sect. 9, with details in Gauge Choice and Loop Closing of Uncertain Polygons, Sect. 16.2 ff.

Go to the beginning.

_____

[3] They are Householder matrices.

20 Reconstructing the Straight Path of a Vehicle from Four Bearings

Discussing variations of determining the pose of moving objects I came across the problem to reconstruct the path of a vehicle moving on a straight line with constant velocity when at constant time intervals observing the azimuth, via bearings from four varying but known positions. The solution can be found geometrically and allows an intuitive visualization.

Much later I learned that there exists the 4-bearings method to solve this problem, see the blog https://www.mathscinotes.com/2015/04/computing-a-ships-course-from-four-bearings/. I also found a much easier geometric solution, described in https://ricojansen.nl/downloads/the_four_bearings_method_v2,Kuikueg.pdf. The document also describes geometric method for the case of bearings from a static point where only three bearings are necessary.

Go to the beginning.

21 Moving a Piano around a Rectangular Corner

Moving a piano around a corner of two corridors requires the corridors to have a minimum width, depending on the size of the piano. Assuming the piano's ground plan is a rectangle with known size, and given the width of one corridor then the width of the other corridor needs to have a minimum width.

The mathematical solution is given in Necessary Width of a Corridor to Transport a Piano around a Corner. The construction used for the animation is based on the fact that each planar motion is a rotation around some point. Here, we know that two of the piano's corners A and B move along the two axes, hence the centre of rotation C is the intersection of the perpendicular lines through the points, where the piano is touching the axes. The point D of the other side of the piano closest to the centre C of the circle provides the minimal width of one corridor given the width of the other corridor.

Go to the beginning.

References

Baarda, W. (1973). *S-Transformations and Criterion Matrices*, Volume 5/1 of *Publication on Geodesy, New Series*. Netherlands Geodetic Commission. 13

Calvo, M. and J. M. Oller (1990). A distance between multivariate normal distributions based in an embedding into the siegel group. *Journal of Multivariate Analysis 35* (2), 223–242. 17

Förstner, W. and B. Moonen (1999). A Metric for Covariance Matrices. In F. Krumm and V. S. Schwarze (Eds.), *Festschrift for Erik W. Grafarend on the occasion of his 60th birthday. Also appeared in: Geodesy - The Challenge of the 3rd Millennium (2003, with editors Professor Dr. Erik W. Grafarend, Dr. Friedrich W. Krumm,Dr. Volker S. Schwarze, ISBN: 978-3-642-07733-3 (Print) 978-3-662-05296-9 (Online))*, pp. 113–128. 13, 17

Förstner, W. and B. P. Wrobel (2016). *Photogrammetric Computer Vision – Statistics, Geometry, Orientation and Reconstruction*. Springer. 3, 7, 9, 11, 12, 19, 22, 24

Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition* (2nd ed. ed.). Academic Press. 16

Jordan, M. I. (1995). Why the logistic function? A tutorial discussion on probabilities and neural networks. https://cs.nyu.edu/~roweis/csc2515-2006/readings/whylogistic.pdf, last visited 12.4.2024. 14

Rao, R. C. (1967). Least squares theory using an estimated dispersion matrix and its application to measurement of signals. In *Proc. Fifth Berkeley Symp. on Math. Statist. and Prob., Vol. 1*, pp. 355–372. Univ. of Calif. Press. Lemma 5a. 6

Schaffalitzky, F. and A. Zisserman (2000). Planar Grouping for Automatic Detection of Vanishing Lines and Points. *Image and Vision Computing 18*, 647–658. 27