From Rytz to the Covariance Ellipsoid

Hellmuth Stachel

Technische Universität Wien
Vienna University of Technology

stachel@dmg.tuwien.ac.at — http://www.geometrie.tuwien.ac.at/stachel

Conference on Geometry: Theory and Applications, Vorau, June 3-8, 2007

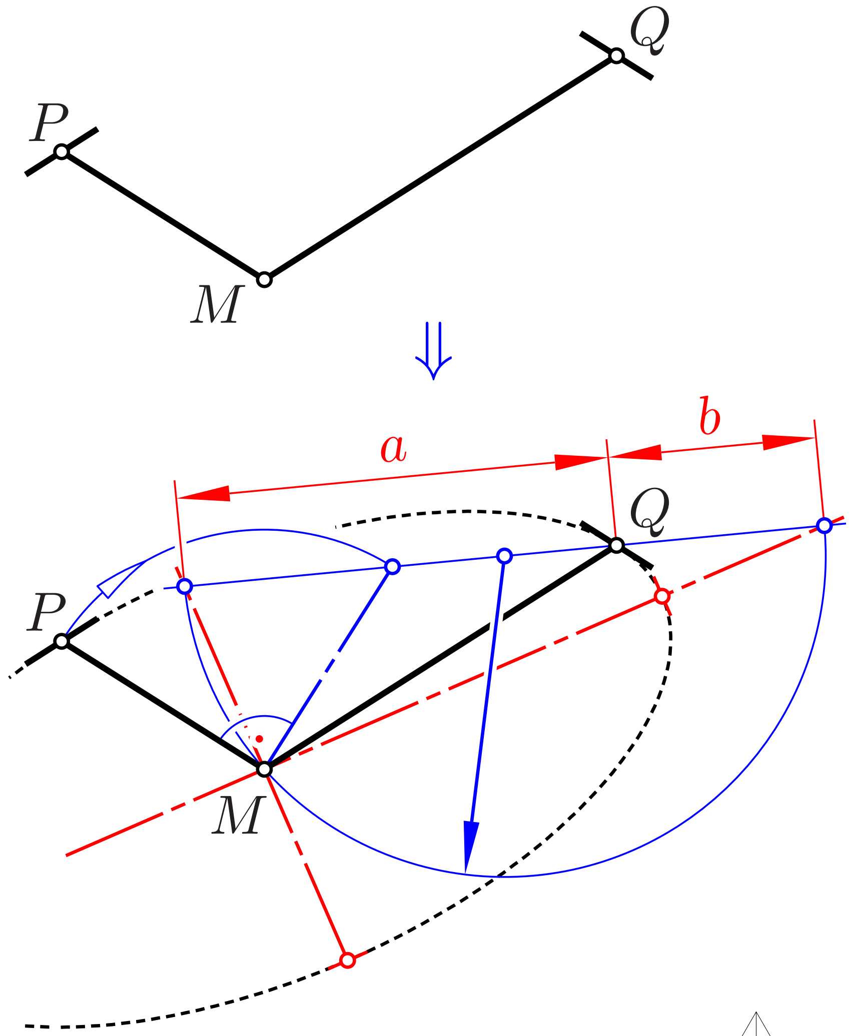
Table of contents

1. Comments on Rytz' construction

Rytz' construction[1] is a standard construction and sometimes seen as a nightmare of Descriptive Geometry.

GIVEN: Conjugate diameters MP, MQ of an ellipse.

WANTED: Axes and vertices.



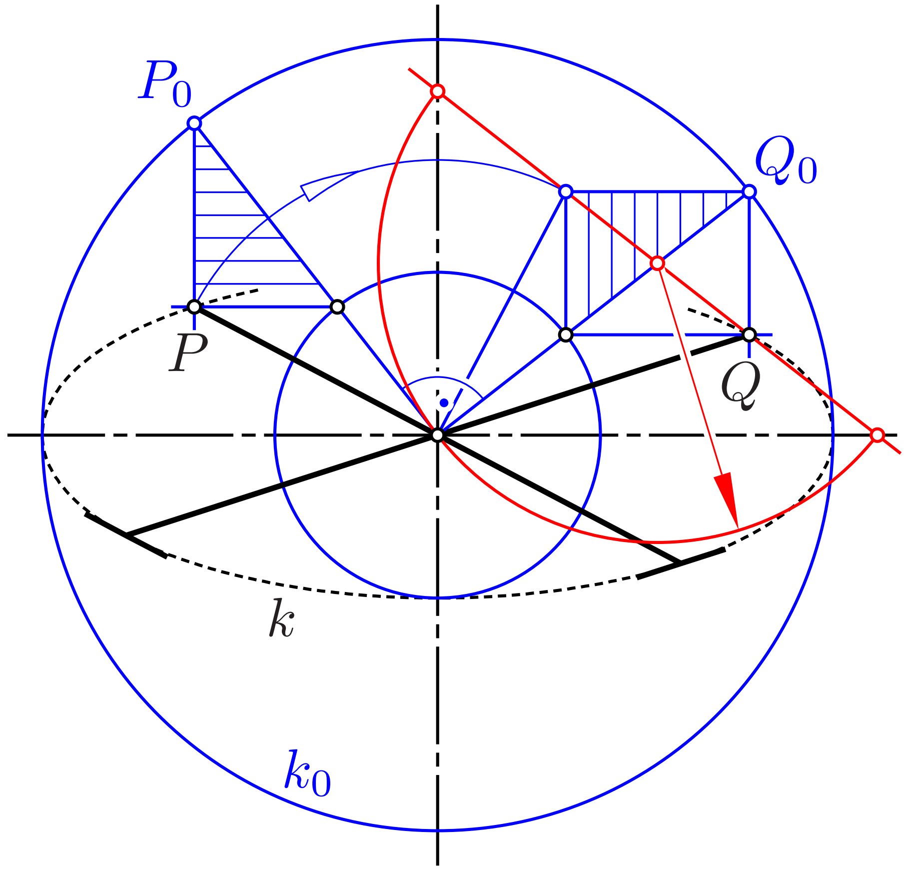
[1]) D. RYTZ, 1801–1868, teacher in Switzerland

Proving Rytz' construction

The synthetic proof uses a particular affine transformation which maps the ellipse k onto a circle k_0.

Zbeněk Nádeník (Prague, 2001) studied the history of Rytz' construction and he posed the question:

Is there a simple analytic approach to this problem ?



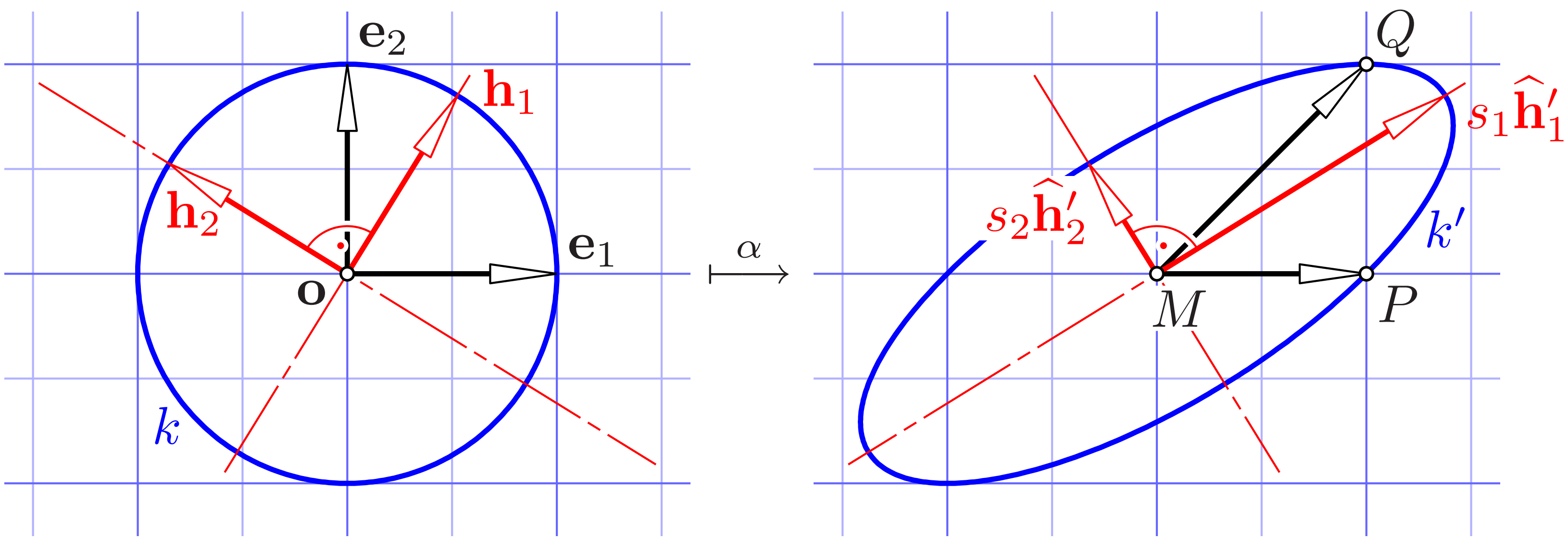
Rytz' construction = Frézier's construction?

Die Ansicht von O. Mycak, dass nicht G. Monge, sondern A.-F. Frézier der Begründer DG ist, muss man als übertrieben beurteilen. Nur in Hauptrissen: Die Vorgänger von Frézier haben zu den Konstruktionen nur die Anweisungen, aber nicht die Begründungen angegeben. Erst Frézier hat die Konstruktionen mit den Beweisen versehen (wenn auch schlecht im Falle der Achsenkonstruktion einer Ellipse aus den konjugierten Durchmessern). Das Werk von Frézier ist sehr vielsprachig, an vielen Stellen sind die Wiederholungen und die Wasserfälle von Wörtern. Dagegen erst Monge hat in die Menge der Konstruktionen ein System eingetragen. Er hat erkannt, dass aus einer kleineren Anzahl der einfachen Konstruktionen die anderen hergeleitet werden können. Das ist das grosse Verdienst von Monge, dadurch hat er die DG gebildet.

Zum 200. Jahrestag der Ausgabe von Monges "Géométrie descriptive" habe ich einen tschechischen Artikel veröffentlicht, in dem ich auch über Frézier und über Monges Assistenten Sylvestre-François Lacroix geschrieben habe.

Mit herzlichen Grüssen

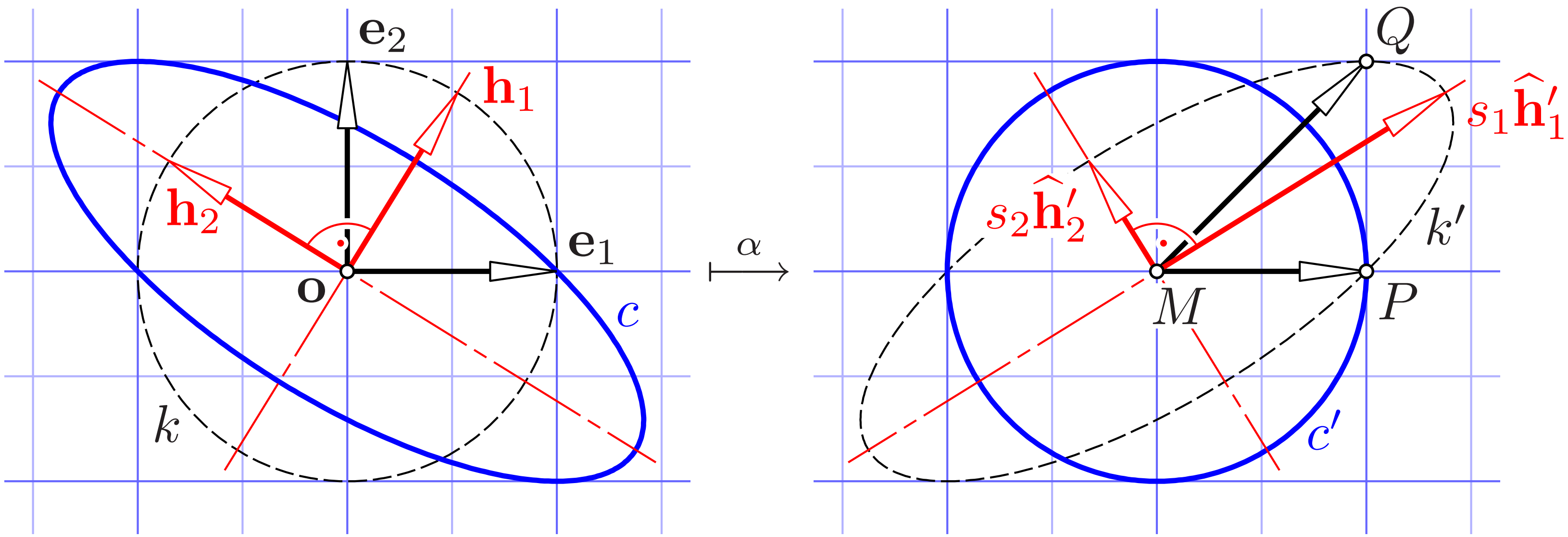
Rytz' construction gives singular values



There is an affine transformation α mapping orthonormal vectors $\mathbf{e}_1$, $\mathbf{e}_2$ onto the given conjugate diameters $\overrightarrow{MP}$, $\overrightarrow{MQ}$ and the unit circle k onto the ellipse k'.

The orthonormal unit vectors $\mathbf{h}_1$, $\mathbf{h}_2$ are mapped onto the orthogonal vectors $\mathbf{h}'_1$ and $\mathbf{h}'_1$ of lengths s_1 and s_2 (= principal distorsions or singular values) of α.

Rytz' construction gives singular values

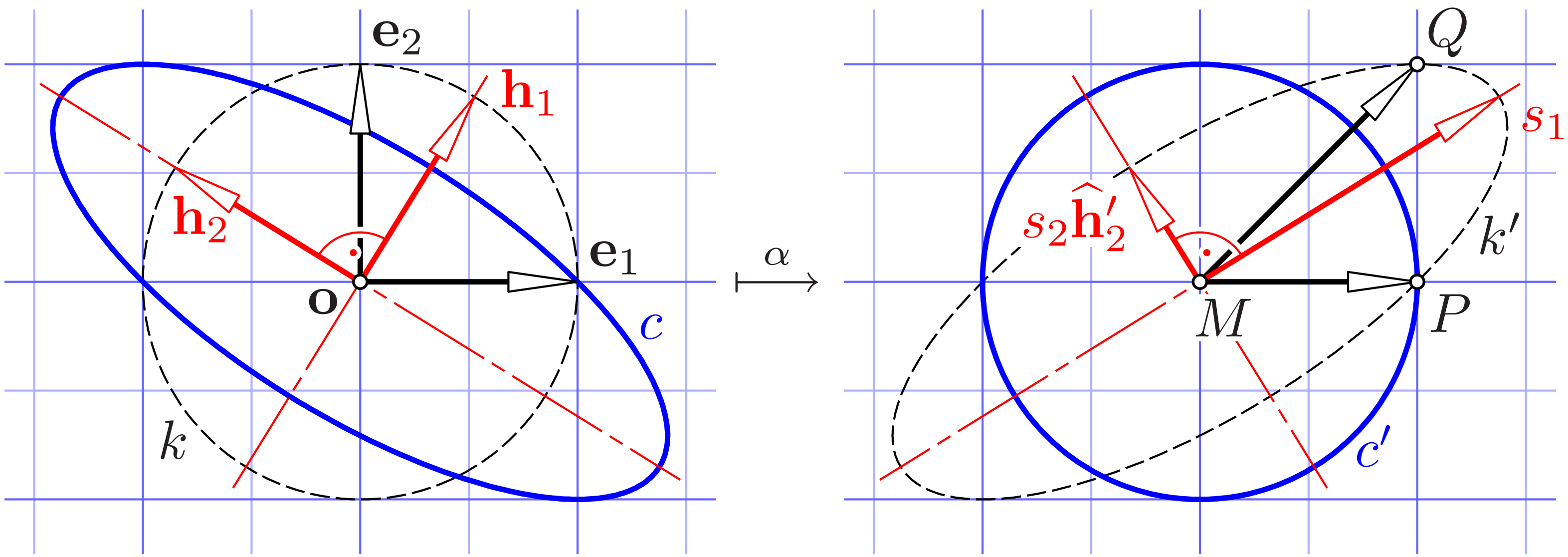


$$\alpha\colon\ \mathbf{x}' = A\,\mathbf{x}, \quad A = (\mathbf{p}\,\mathbf{q}) = \begin{pmatrix} p_x & q_x \\ p_y & q_y \end{pmatrix}$$

$$\|\mathbf{x}'\| = 1 \iff \mathbf{x}^T A^T A\,\mathbf{x} = 1 \quad \ldots\ c \qquad (A^TA = \text{Gramian})$$

$\mathbf{h}_1, \mathbf{h}_2$ are normed eigenvectors of A^TA to the eigenvalues s_1^2, s_2^2.

Rytz' construction gives singular values



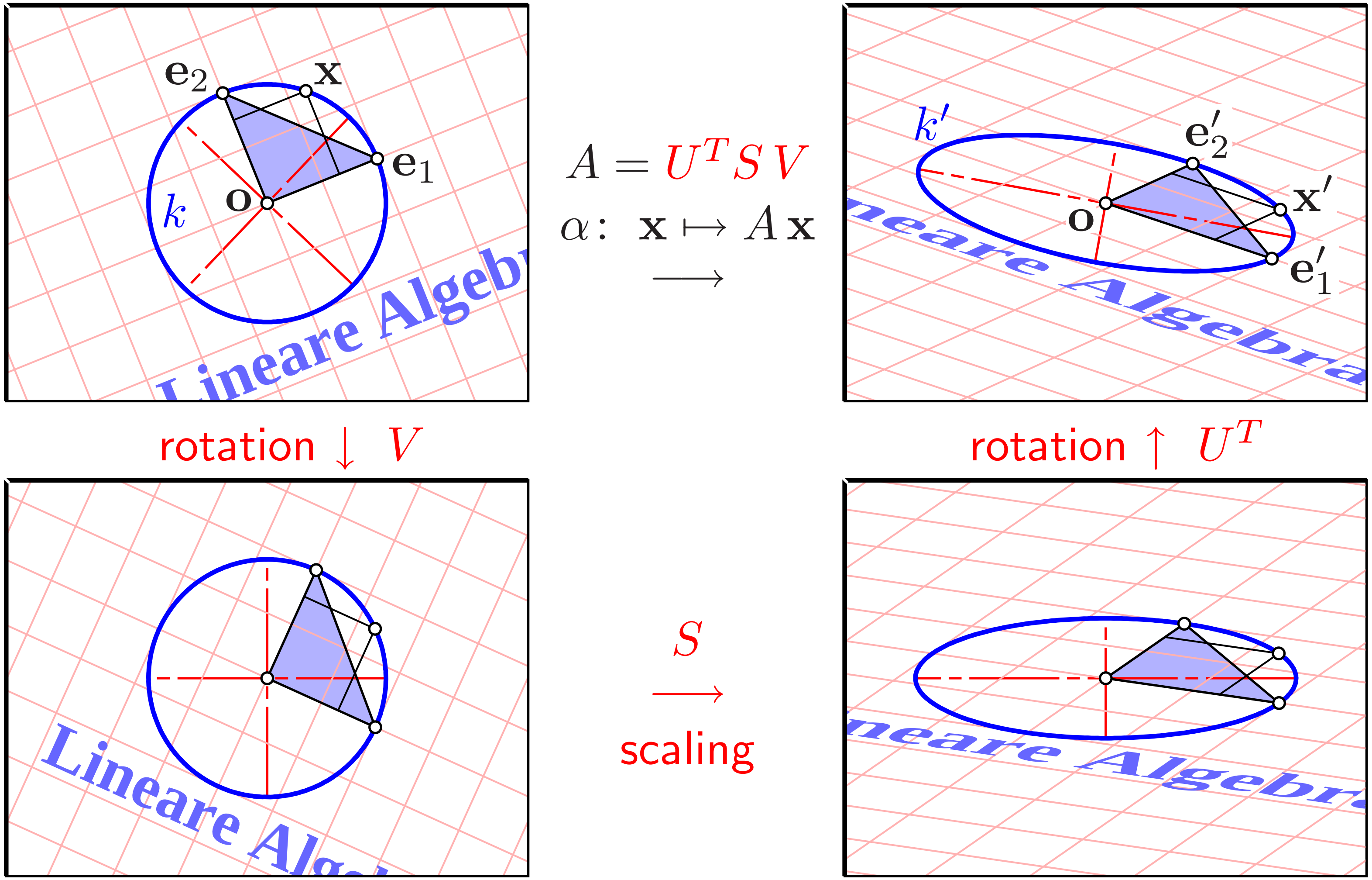
$c \ldots \mathbf{x}^T(A^TA)\,\mathbf{x} = 1$ $\qquad$ $k' \ldots \mathbf{x}'^T(AA^T)^{-1}\,\mathbf{x}' = 1$

$$s_j^2\mathbf{h}_j = A^TA\,\mathbf{h}_j \mid \mathbf{h}_i\cdot \implies s_j^2\delta_{ij} = s_j^2(\mathbf{h}_i^T\mathbf{h}_j) = (A\,\mathbf{h}_i)^T(A\,\mathbf{h}_j) = \mathbf{h}'_i\cdot\mathbf{h}'_j.$$

s_1, s_2 are principal axes of the image k' of the unit circle

$U := {}_{\widehat{H}'}T_{E'},\ V := {}_{H}T_E \implies A = U^T\begin{pmatrix} s_1 & 0 \\ 0 & s_2 \end{pmatrix} V$ with orthogonal U, V.

Singular value decomposition



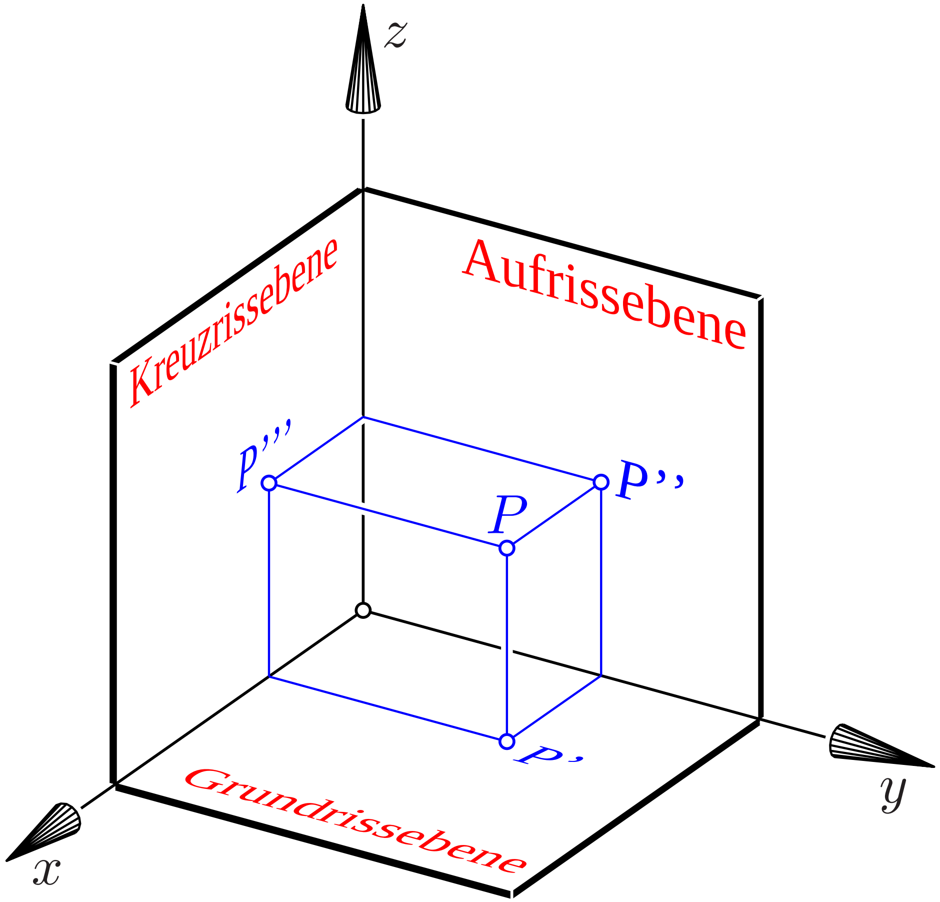
Rytz' construction gives singular values

One simple application:

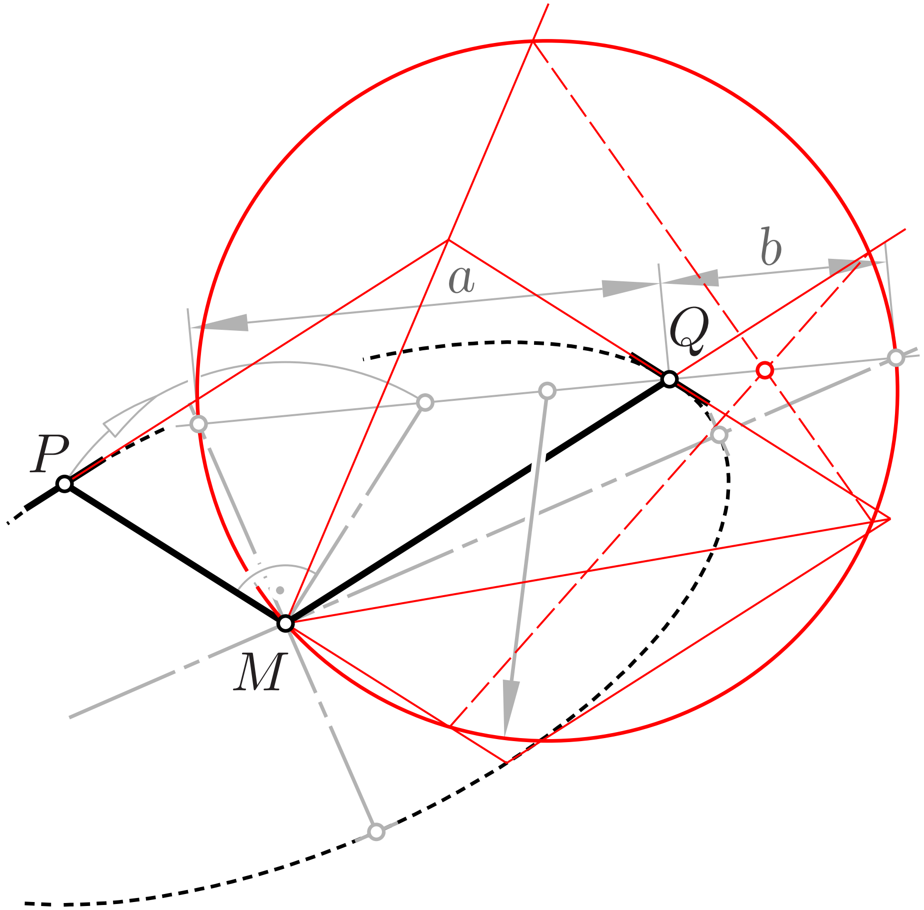
affine distorsion of text (or figures) by rotation and axial scaling

Other hints:

B. Casselman: *Mathematical illustrations: a manual of geometry and postscript.*
Cambridge University Press 2005



Rytz solves an eigenvalue problem

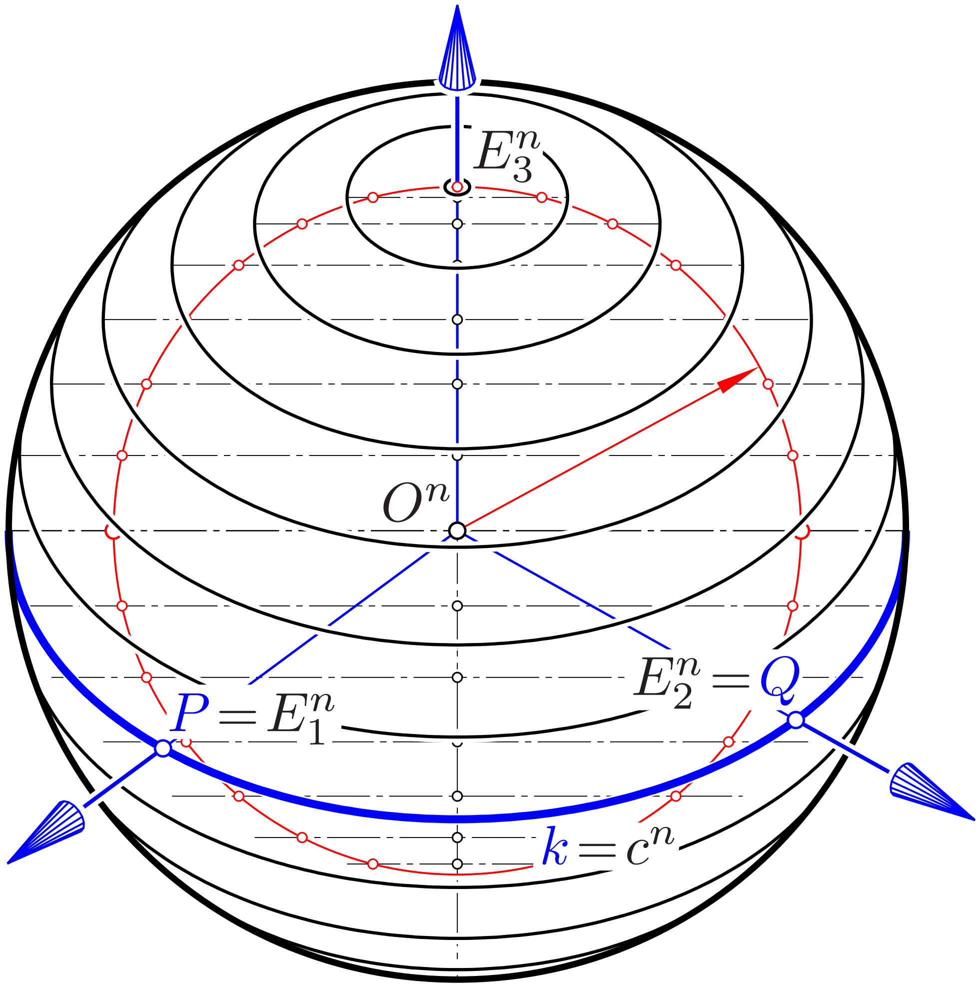


Compared with the STEINERcircle construction for the orthogonal pair in the involution of conjugate diameters:

RYTZ gives at once

- the eigenvalues and
- the eigenvectors.

A consequence of Gauß' equation



The axonometry given by the reference system $(O^n; E_1^n, E_2^n, E_3^n)$ is similar to an orthogonal view iff (expressed in complex coordinates with origin O^n)

$$e_1^2 + e_2^2 + e_3^2 = 0\,.$$

The complex coordinates of the focal points f_j of k are $f_j = \pm i e_3 \implies f_j^2 = -e_3^2\,.$

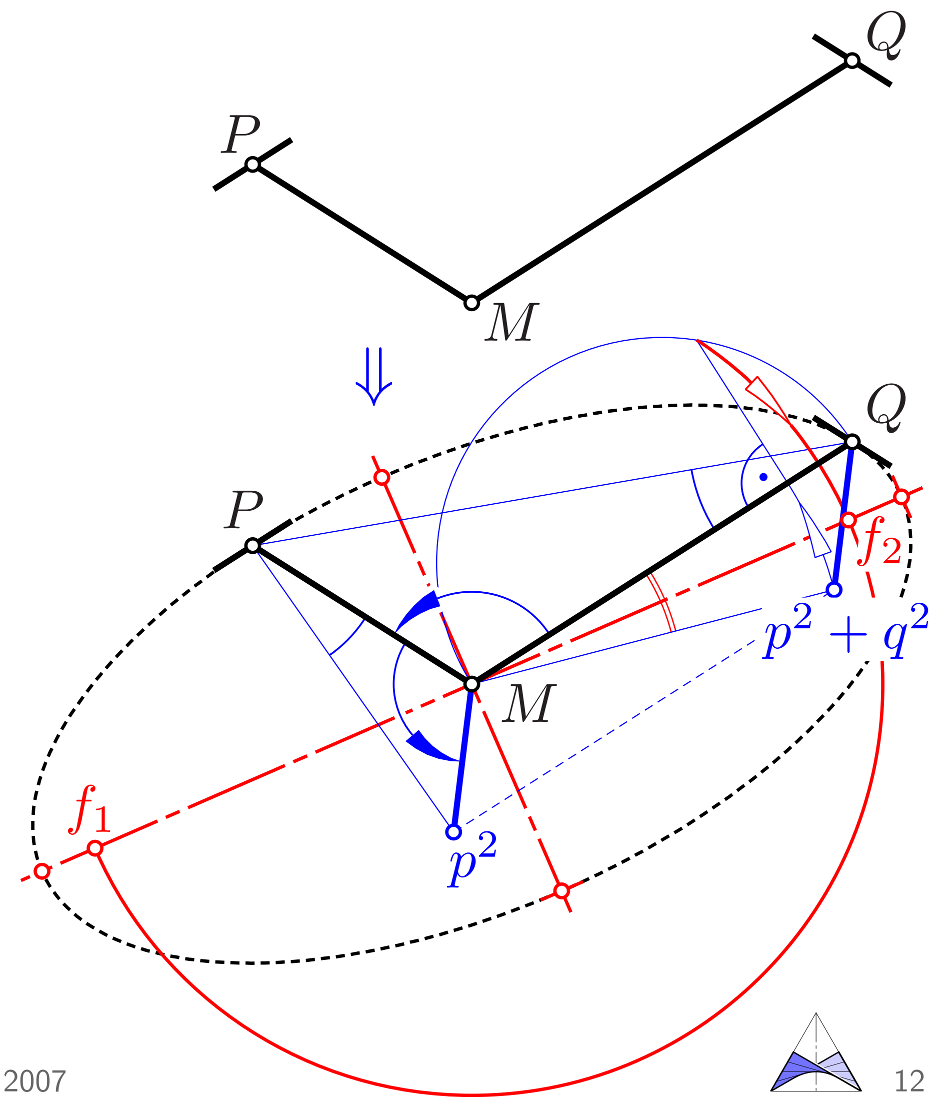
Focal points of an ellipse

GIVEN: Conjugate diameters MP, MQ of an ellipse k.

WANTED: Axes and focal points F_1, F_2 of k.

Expressed in complex coordinates with origin M:

$$f_j^2 = p^2 + q^2.$$



Focal points of an hyperbola

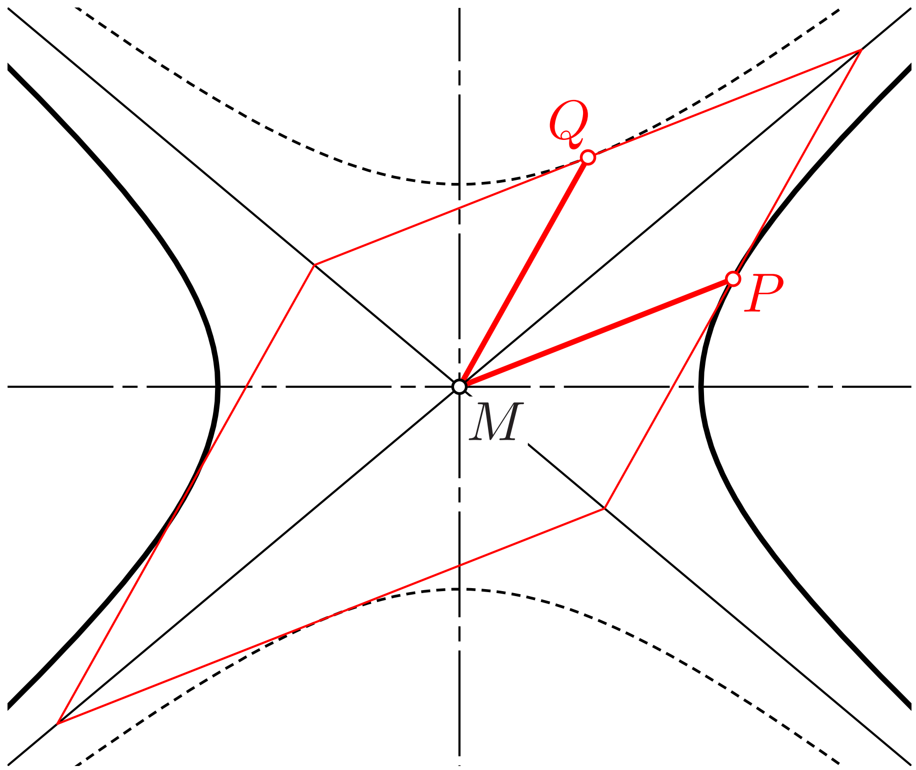
GIVEN: Conjugate diameters MP, MQ of an hyperbola.

WANTED: Focal points F_1, F_2 .

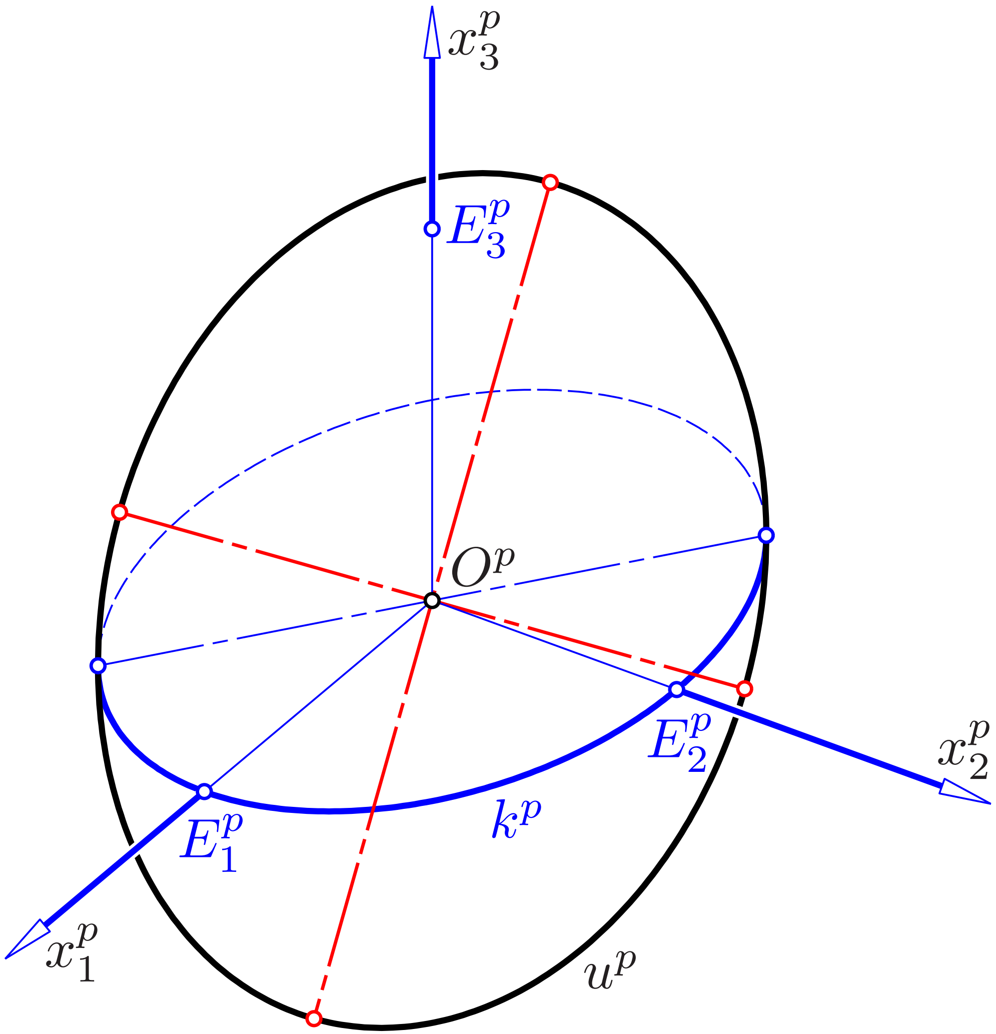
Expressed in complex coordinates with origin M:

$$f_i^2 = p^2 - q^2.$$

Proof: $P = (a\cosh t,\, b\sinh t)$, $Q = (a\sinh t,\, b\cosh t)$.



Oblique axonometry



GIVEN: Axonometric reference system $(O^p; E_1^p, E_2^p, E_3^p)$.

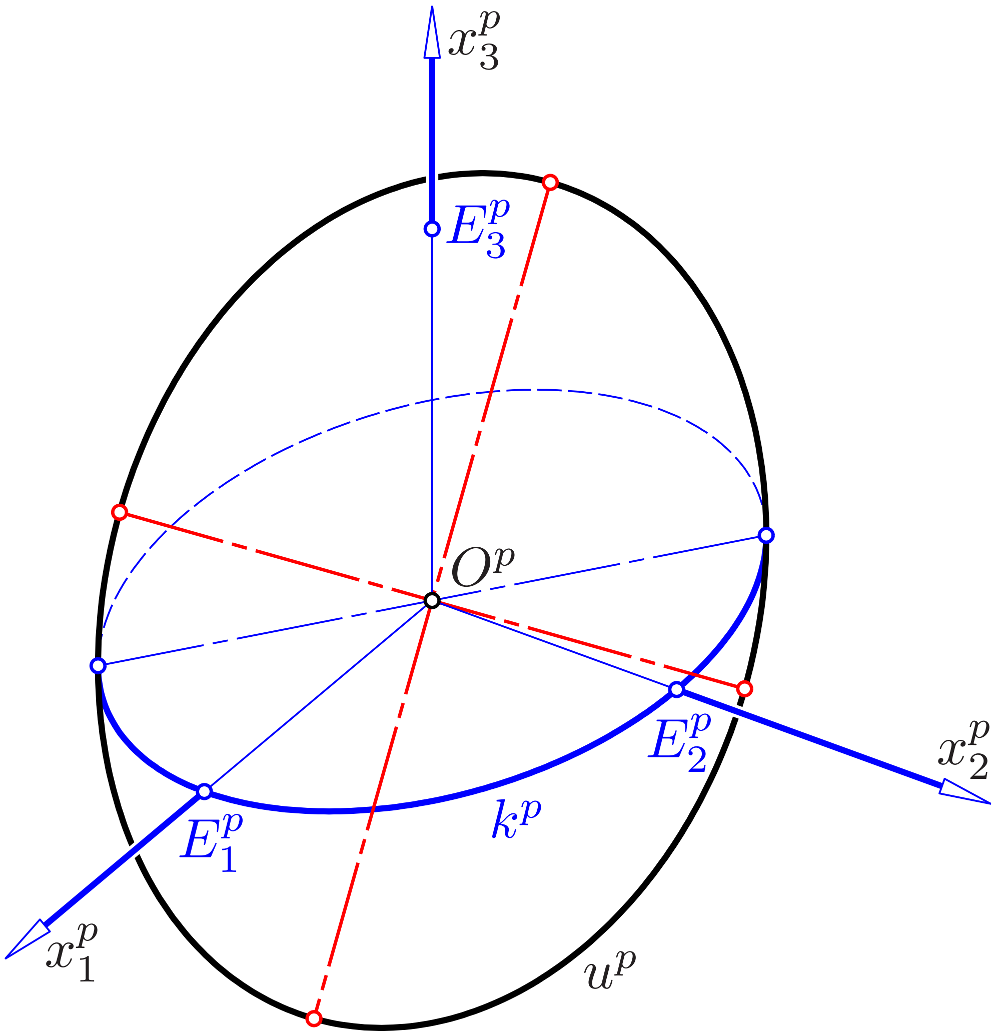
WANTED: Contour u^p of the unit sphere.

Due to E. WAELSCH, J.ber. DMV **21**, p. 24, (1912) the complex coordinates of the focal points $f_{1,2}$ of u^p obey

$$f^2 = e_1^2 + e_2^2 + e_3^2$$

provided $e_j \in \mathbb{C}$ are the complex coordinates of E_j^p.

Oblique axonometry



$$f^2 = \Big[\underbrace{\underbrace{(e_1^2 + e_2^2)}_{\text{Rytz for } k^p} + e_3^2}_{\text{second Rytz for } u^p}\Big]$$

Proof: The axonometric mapping $\mathbf{x} \mapsto \mathbf{x}'$ obeys

$$\mathbf{x}' = A\,\mathbf{x}, \quad A = (\mathbf{e}_1^p\,\mathbf{e}_2^p\,\mathbf{e}_3^p)$$

with the coordinates of E_j^p in the columns.

Oblique axonometry

For $\mathbf{x}' = A\,\mathbf{x}$ there is the singular value decomposition

$$A = U^T \begin{pmatrix} s_1 & 0 & 0 \\ 0 & s_2 & 0 \end{pmatrix} V = U^T \begin{pmatrix} s_1\,\mathbf{r}_1 \\ s_2\,\mathbf{r}_2 \end{pmatrix}, \quad s_1 \geq s_2 > 0\,.$$

$\mathbf{r}_j$ are the orthonormal row vectors of V. Then $\mathbf{r}_3$ spans the kernel of A (rays of sight). The unit circle u in the plane spanned by $\mathbf{r}_1$ and $\mathbf{r}_2$ is the true contour.

We modify the coordinate frame in the image plane by neglecting U^T. Then the aparent contour has the equation

$$u^p\colon\ \frac{x_1'^{\,2}}{s_1^2} + \frac{x_2'^{\,2}}{s_2^2} = (x_1'\ x_2') \begin{pmatrix} 1/s_1^2 & 0 \\ 0 & 1/s_2^2 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = 1\,.$$

Oblique axonometry

For $A = \begin{pmatrix} s_1\,\mathbf{r}_1 \\ s_2\,\mathbf{r}_2 \end{pmatrix} = (\mathbf{e}_1^p\,\mathbf{e}_2^p\,\mathbf{e}_3^p)$ the complex coordinates of E_j^p give

$$f^2 = e_1^2 + e_2^p + e_3^2 = s_1^2\underbrace{\mathbf{r}_1^2}_{1} - s_2^2\underbrace{\mathbf{r}_2^2}_{1} + 2i\,(\underbrace{\mathbf{r}_1\cdot\mathbf{r}_2}_{0}) = (s_1^2 - s_2^2) + 0\,i \in \mathbb{C}\,.$$

Generalization: Axonometry $\mathbb{E}^n \to \mathbb{E}^2$: Given: Unit points $E_1^p, \ldots, E_n^p$, Wanted: Contour u^p of unit sphere (= ellipse enclosing $E_1^p, \ldots, E_n^p$).

Result: The singular values s_1, s_2 of $A = (e_1^p \ldots e_n^p)$ are semi-axes of u^p, and the complex coordinates f of the focal points obey $f^2 = e_1^2 + \ldots + e_n^2$.

Oblique axonometry

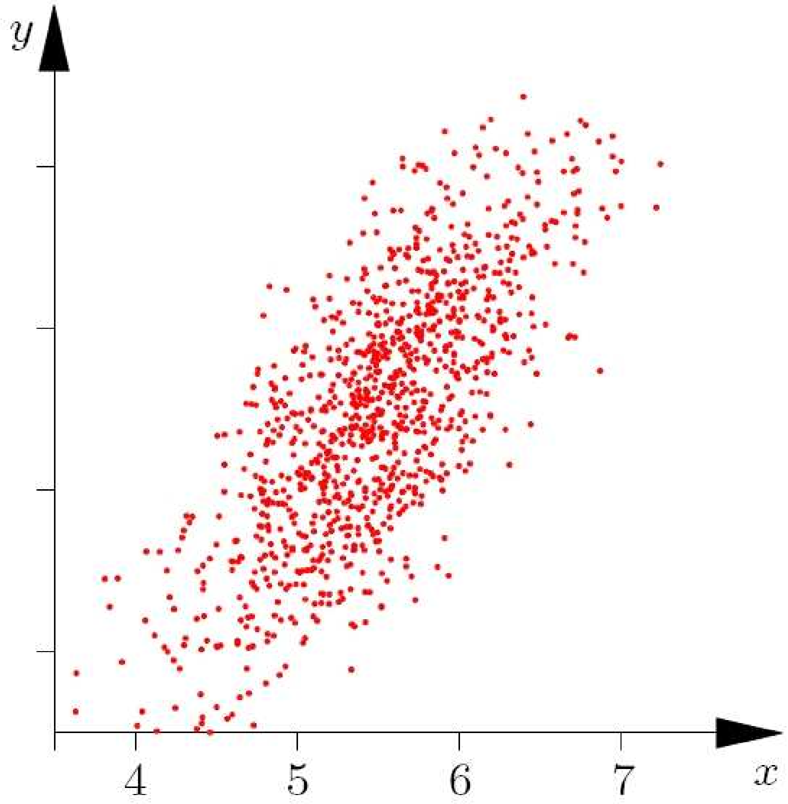
For $A = \begin{pmatrix} s_1 \, \mathbf{r}_1 \\ s_2 \, \mathbf{r}_2 \end{pmatrix} = (\mathbf{e}_1^p \, \mathbf{e}_2^p \, \mathbf{e}_3^p)$ the complex coordinates of E_j^p give

$$f^2 = e_1^2 + e_2^p + e_3^2 = s_1^2 \underbrace{\mathbf{r}_1^2}_{1} - s_2^2 \underbrace{\mathbf{r}_2^2}_{1} + 2i \, (\underbrace{\mathbf{r}_1 \cdot \mathbf{r}_2}_{0}) = (s_1^2 - s_2^2) + 0 \, i \in \mathbb{C} \, .$$

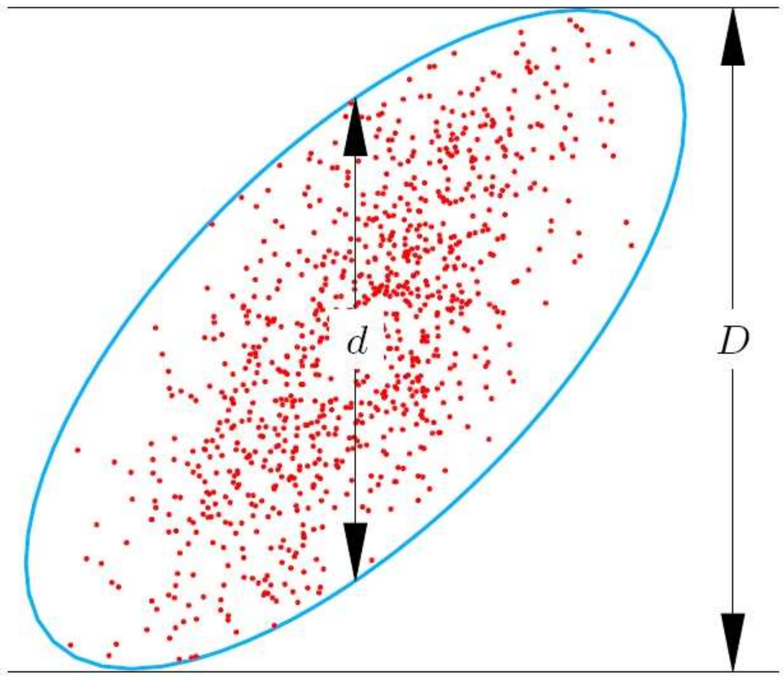
Generalization: Axonometry $\mathbb{E}^n \to \mathbb{E}^2$: GIVEN: Unit points $E_1^p, \ldots, E_n^p$,
WANTED: Contour u^p of unit sphere (= ellipse enclosing $E_1^p, \ldots, E_n^p$).

RESULT: The singular values s_1, s_2 of $A = (e_1^p \, \ldots \, e_n^p)$ are semi-axes of u^p, and the complex coordinates f of the focal points obey $f^2 = e_1^2 + \ldots + e_n^2$.

The concentration ellipse in statistics



are x_j, y_j or given data covariant?



concentration ellipse, correlation coefficient $r = \sqrt{1 - (d/D)^2}$

Eutactic points in the plane

$E_1^p, \ldots, E_n^p \in \mathbb{E}^2$ are called eutactic (SCHLAEFLI, 1901)

$\iff$ the singular values of matrix A with column vectors $(\mathbf{e}_1 \ldots \mathbf{e}_n)$ are equal, i.e., $s_1 = s_2 = s$,

$\iff$ the two row vectors $\mathbf{r}_1, \mathbf{r}_2$ of A match $\|\mathbf{r}_1\| = \|\mathbf{r}_2\| = s > 0$ and $\mathbf{r}_1 \cdot \mathbf{r}_2 = 0$, i.e., $AA^T = s^2 I_2$

$\iff$ u^p is a circle,

$\iff$ $f^2 = e_1^2 + \ldots + e_n^2 = 0$,

$\iff$ the axonometry is similar to an orthogonal view,

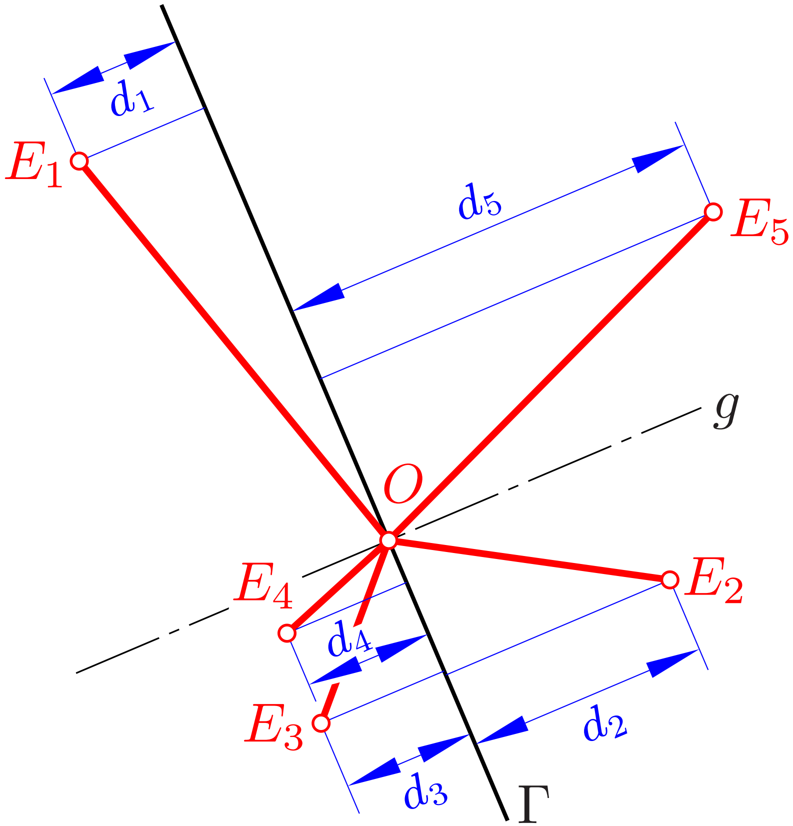
$\iff$ the axonometry is similar to an isocline view, provided $n > 3$,

$\iff$ there is no correlation between x_j and y_j, i.e., the correlation coefficient is 0 .

Another characterization of eutactic points ?

$E_1, \ldots, E_n$ are eutactic $\iff$ for any hyperplane Γ through O the sum

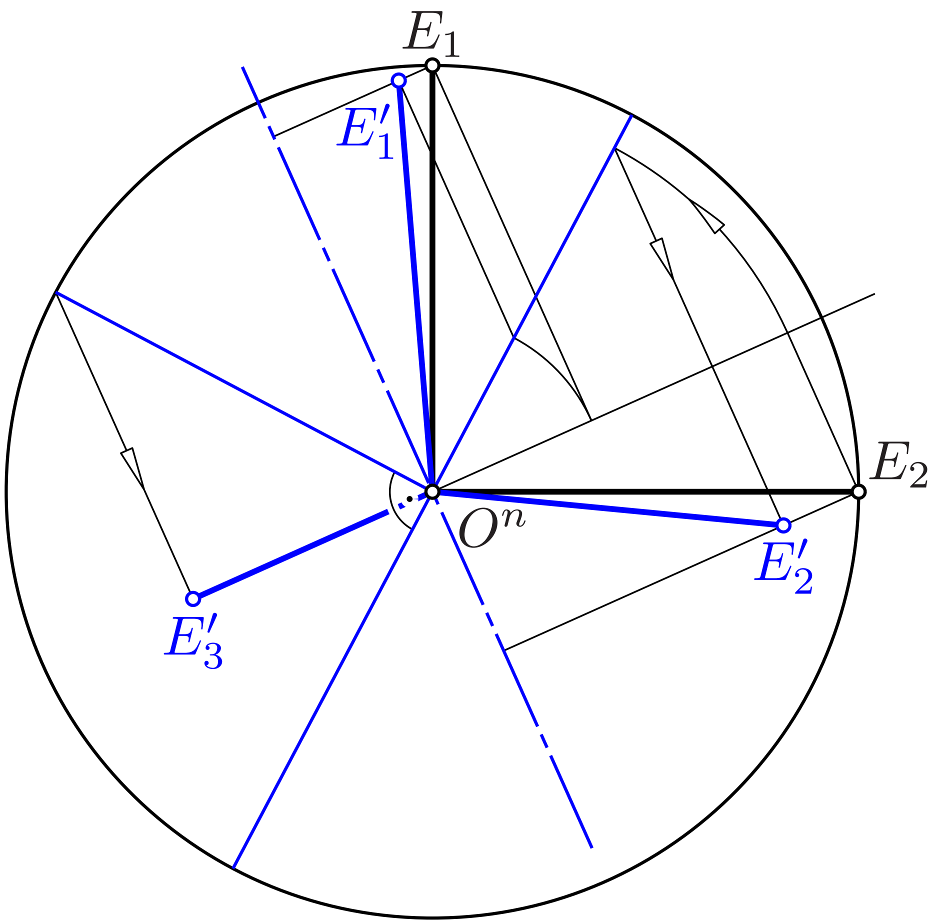
$$\sum_{j=1}^{n} \overline{E_j\,\Gamma}^{\,2} = \sum_{j=1}^{n} d_j^2 = s^2 = \text{const.}$$

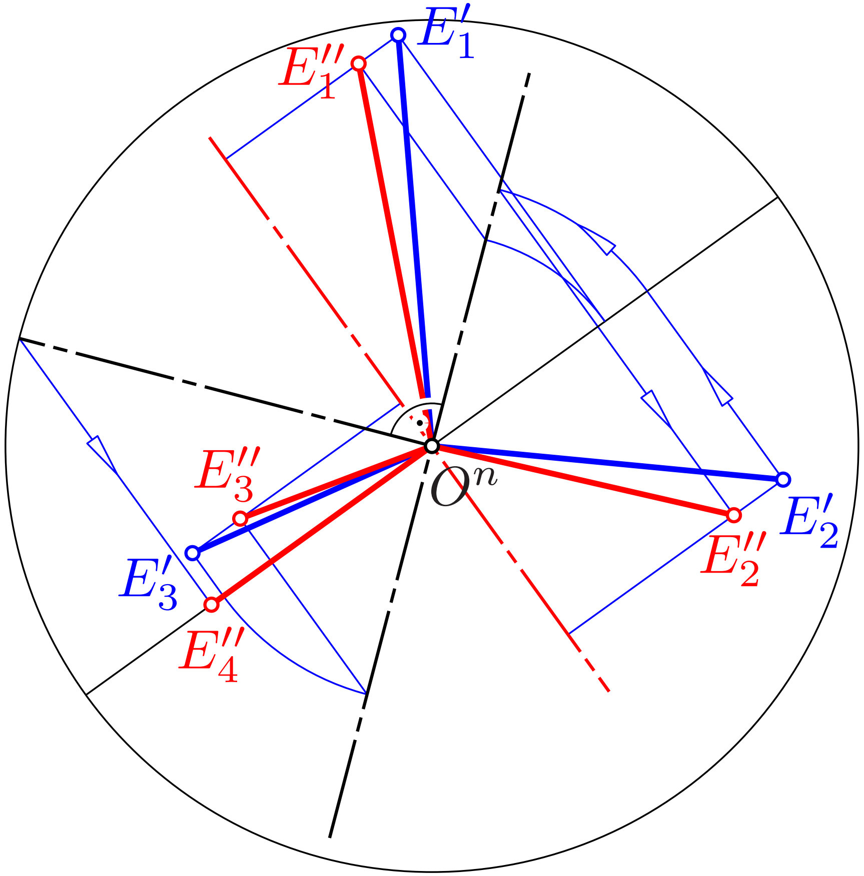


Proof: Let $\mathbf{n}$ be a unit vector orthogonal Γ and $\mathbf{r}_1, \mathbf{r}_2$ be the row vectors in $A = (\mathbf{e}_1 \ \ldots \ \mathbf{e}_n)$:

$$\sum_{j=1}^{n} (\mathbf{e}_j \cdot \mathbf{n})^2 = \underbrace{\mathbf{r}_1^2}_{s^2} n_1^2 + \underbrace{\mathbf{r}_2^2}_{s^2} n_2^2 = s^2.$$

How to obtain eutactic points iteratively ?





The ellipsoid of inertia

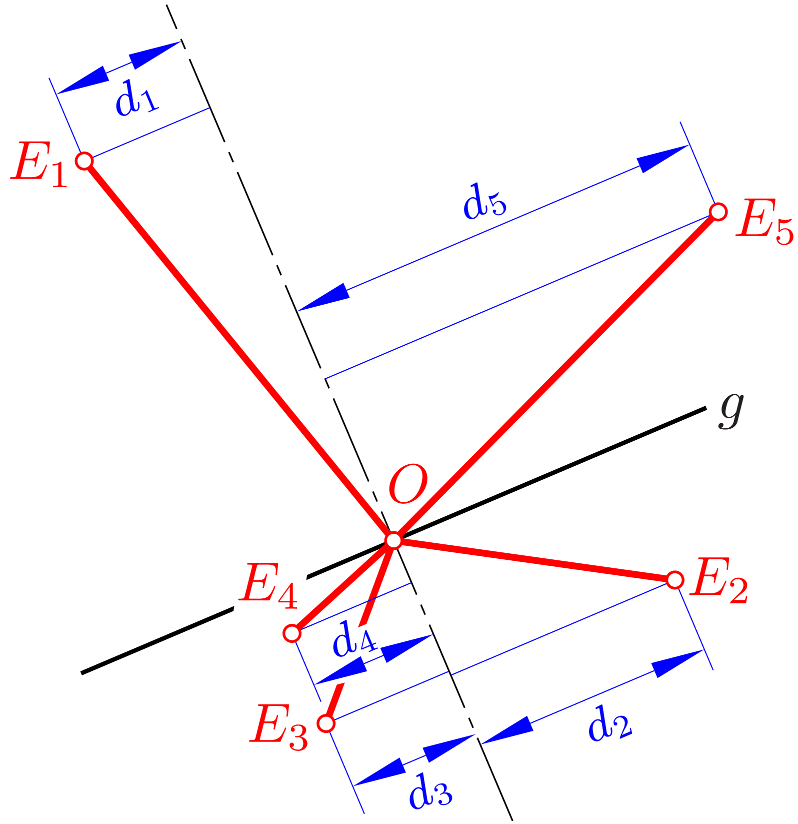
GIVEN: $E_1, \dots, E_n$.

For any line g through O project E_j orthogonally onto $E'_j \in g$ and determine the sum of squared distances

$$Q_g^2 := \sum_{j=1}^{n} \overline{OE'_j}^{\,2} = \sum_{j=1}^{n} d_j^2$$

and protract from O along g the length $\pm 1/Q_g$.

Then these points constitute the ellipsoid of inertia (or covariance ellipsoid).



The ellipsoid of inertia

Proof: $$Q_g^2 = \frac{1}{\|\mathbf{g}\|^2}\sum_{j=1}^{n}(\mathbf{e}_j\cdot\mathbf{g})^2 = \frac{1}{\|\mathbf{g}\|^2}\sum_{j=1}^{n}(\mathbf{g}^T\mathbf{e}_j)(\mathbf{e}_j^T\mathbf{g}) = \frac{1}{\|\mathbf{g}\|^2}\,\mathbf{g}^T\Big(\sum_{j=1}^{n}\mathbf{e}_j\,\mathbf{e}_j^T\Big)\mathbf{g}\,.$$

$$\mathbf{x} = \pm\frac{1}{\sqrt{Q_g}}\frac{\mathbf{g}}{\|\mathbf{g}\|} = \pm\frac{\|\mathbf{g}\|}{\sqrt{\sum_j(\mathbf{e}_j\cdot\mathbf{g})^2}}\frac{\mathbf{g}}{\|\mathbf{g}\|} = \pm\frac{1}{\sqrt{\mathbf{g}^T\big(\sum_j\mathbf{e}_j\,\mathbf{e}_j^T\big)\mathbf{g}}}\,\mathbf{g}\,.$$

These points obey

$$\mathbf{x}^T\Big(\sum\mathbf{e}_j\,\mathbf{e}_j^T\Big)\mathbf{x} = \mathbf{x}^T\underbrace{(\mathbf{e}_1\ \ldots\ \mathbf{e}_n)}_{A}\begin{pmatrix}\mathbf{e}_1^T\\ \vdots\\ \mathbf{e}_n^T\end{pmatrix}\mathbf{x} = \mathbf{x}^T(A\,A^T)\,\mathbf{x} = 1$$

The ellipsoid of inertia

Proof: $Q_g^2 = \frac{1}{\|\mathbf{g}\|^2}\sum_{j=1}^{n}(\mathbf{e}_j\cdot\mathbf{g})^2 = \frac{1}{\|\mathbf{g}\|^2}\sum_{j=1}^{n}(\mathbf{g}^T\mathbf{e}_j)(\mathbf{e}_j^T\mathbf{g}) = \frac{1}{\|\mathbf{g}\|^2}\,\mathbf{g}^T\Big(\sum_{j=1}^{n}\mathbf{e}_j\,\mathbf{e}_j^T\Big)\mathbf{g}\,.$

$$\mathbf{x} = \pm\frac{1}{\sqrt{Q_g}}\,\frac{\mathbf{g}}{\|\mathbf{g}\|} = \pm\frac{\|\mathbf{g}\|}{\sqrt{\sum_j(\mathbf{e}_j\cdot\mathbf{g})^2}}\,\frac{\mathbf{g}}{\|\mathbf{g}\|} = \pm\frac{1}{\sqrt{\mathbf{g}^T(\sum_j\mathbf{e}_j\,\mathbf{e}_j^T)\mathbf{g}}}\,\mathbf{g}\,.$$

These points obey

$$\mathbf{x}^T\big(\sum\mathbf{e}_j\,\mathbf{e}_j^T\big)\,\mathbf{x} = \mathbf{x}^T\,\underbrace{(\mathbf{e}_1\;\ldots\;\mathbf{e}_n)}_{A}\begin{pmatrix}\mathbf{e}_1^T\\ \vdots\\ \mathbf{e}_n^T\end{pmatrix}\mathbf{x} = \mathbf{x}^T(A\,A^T)\,\mathbf{x} = 1$$

The ellipsoid of inertia

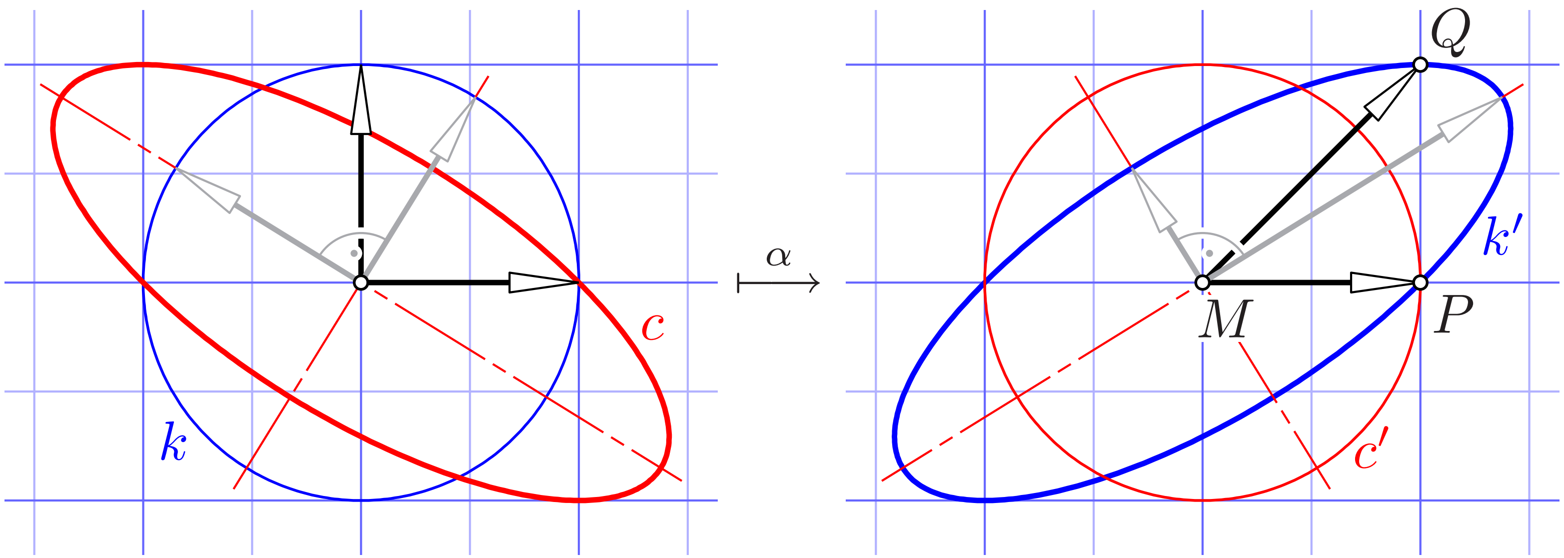
as

$$\mathbf{x}^T(A\,A^T)\,\mathbf{x} = \frac{\mathbf{g}^T(\sum_j \mathbf{e}_j\,\mathbf{e}_i^T)\mathbf{g}}{\mathbf{g}^T(\sum_j \mathbf{e}_j\,\mathbf{e}_i^T)\mathbf{g}} = 1\,.$$

We use the singular value decomposition of A:

$$\begin{aligned} AA^T &= (\mathbf{e}_1\,\ldots\,\mathbf{e}_n)\begin{pmatrix}\mathbf{e}_1\\ \vdots\\ \mathbf{e}_n\end{pmatrix} = (U^T S\,V)(V^T S^T U) = U^T(SS^T)U = \\ &= U^T\begin{pmatrix} s_1 & & 0 & \\ & \ddots & & 0\\ 0 & & s_m & \end{pmatrix}\begin{pmatrix} s_1 & & 0\\ & \ddots & \\ 0 & & s_m\\ & 0 & \end{pmatrix} = \begin{pmatrix} s_1^2 & & 0\\ & \ddots & \\ 0 & & s_m^2\end{pmatrix}. \end{aligned}$$

The ellipsoid of inertia



The semiaxes of the ellipsoid of inertia are $1/s_1, \ldots, 1/s_m$ (like c), the contour u^p has semiaxes $s_1, \ldots, s_m$ (like k).

The ellipsoid of inertia is polar to u^p with respect to the unit sphere.

Eutactic points in $\mathbb{E}^m$

GIVEN: Axonometry $\mathbb{E}^n \to \mathbb{E}^m, \ n > m \geq 2$.

$E_1, \ldots, E_n \in \mathbb{E}^m$ are called eutactic with respect to O

$\iff$ the singular values of matrix A with column vectors $(\mathbf{e}_1 \ \ldots \ \mathbf{e}_n)$ are equal, i.e., $s_1 = \ldots = s_m = s$,

$\iff$ the row vectors $\mathbf{r}_1, \ldots, \mathbf{r}_m$ are of length s and pairweise orthogonal, i.e., $AA^T = s^2 I_m$

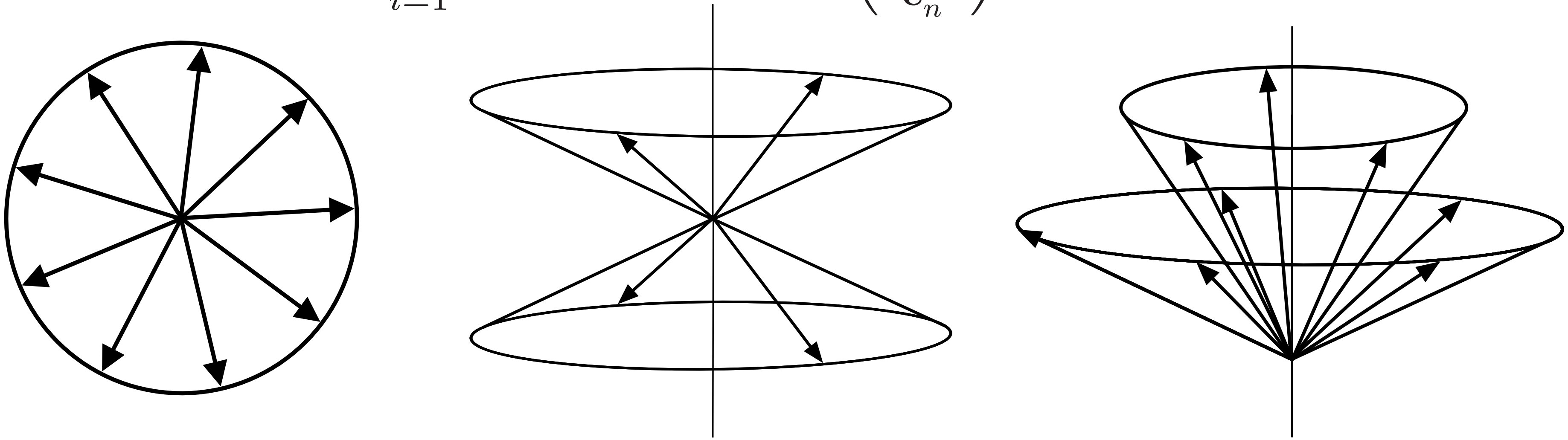
$\iff$ u^p and the ellipsoid of inertia are spheres,

$\iff$ the axonometric image is similar to an orthogonal view or $(n > 3)$ isocline view.

Almost-orthonormal vector systems

The vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n \in \mathbb{R}^m$, $n > m$, are called almost-orthonormal[1]

$$\iff \sum_{i=1}^{n} \mathbf{e}_i \, \mathbf{e}_i^T = (\mathbf{e}_1 \ \ldots \ \mathbf{e}_n) \begin{pmatrix} \mathbf{e}_1^T \\ \vdots \\ \mathbf{e}_n^T \end{pmatrix} = AA^T = I_m \, .$$



[1] D. Haase, H.S.: *Almost-orthonormal Vector Systems.* Beitr. Algebra Geom. **37**, 367–381 (1996)

Almost-orthonormal vector systems

$$\sum_i (\mathbf{e}_i\, \mathbf{e}_i^T) = I_n \mid \cdot\, \mathbf{x} \implies \mathbf{x} = \sum_i \mathbf{e}_i\, \mathbf{e}_i^T \mathbf{x} = \sum_i \mathbf{e}_i (\mathbf{e}_i^T \mathbf{x}) = \sum_i (\underbrace{\mathbf{x} \cdot \mathbf{e}_i}_{\widetilde{x}_i})\, \mathbf{e}_i$$

$(\mathbf{e}_1, \ldots, \mathbf{e}_n)$ is a generating system of $\mathbb{R}^m$. For $\mathbf{x} = \sum x_i\, \mathbf{e}_i$ we use $(x_1, \ldots, x_n)$ as redundant (and of course not unique) coordinates of $\mathbf{x} \in \mathbb{R}^m$. We call the particular coordinates $\widetilde{x}_i = (\mathbf{x} \cdot \mathbf{e}_i)$ distinguished.

$$\mathbf{x} = \sum_i x_i\, \mathbf{e}_i \mid \cdot\, \mathbf{y} \implies \mathbf{x} \cdot \mathbf{y} = \sum_i x_i\, (\underbrace{\mathbf{y} \cdot \mathbf{e}_i}_{\widetilde{y}_i}) = \sum_i x_i\, \widetilde{y}_i = \sum_i \widetilde{x}_i\, y_i\,.$$

The dot product has standard form when computed from distinguished coordinates.

Almost-orthonormal vector systems

$$\widetilde{x}_i = \mathbf{x} \cdot \mathbf{e}_i = \sum_j x_j \, \mathbf{e}_j \cdot \mathbf{e}_i = \sum (\mathbf{e}_i \cdot \mathbf{e}_j) \, x_j \implies \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix} = \underbrace{(\mathbf{e}_i \cdot \mathbf{e}_j)}_{G} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The Gramian maps redundant coordinates onto distinguished coordinates.

Any endomorphism f of $\mathbb{R}^m$ can be represented by a distinguished matrix $\widetilde{C}$.

- f is selfadjoint $\iff$ $\widetilde{C}$ is symmetric.
- f is isometric $\iff$ $\widetilde{C}$ is orthogonal.

Almost-orthonormal vector systems

$$\widetilde{x}_i = \mathbf{x} \cdot \mathbf{e}_i = \sum_j x_j \, \mathbf{e}_j \cdot \mathbf{e}_i = \sum (\mathbf{e}_i \cdot \mathbf{e}_j) \, x_j \implies \begin{pmatrix} \widetilde{x}_1 \\ \vdots \\ \widetilde{x}_n \end{pmatrix} = \underbrace{(\mathbf{e}_i \cdot \mathbf{e}_j)}_{G} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

The Gramian maps redundant coordinates onto distinguished coordinates.

Any endomorphism f of $\mathbb{R}^m$ can be represented by a distinguished matrix $\widetilde{C}$.

- f is selfadjoint $\iff$ $\widetilde{C}$ is symmetric.
- f is isometric $\iff$ $\widetilde{C}$ is orthogonal.

Almost-orthonormal vector systems

After embedding $\mathbb{R}^m$ in $\mathbb{R}^n$, $\mathbf{e}_1, \dots, \mathbf{e}_n$ are orthogonal projections of an orthonormal basis of $\mathbb{R}^n$ into $\mathbb{R}^m$.

The distinguished coordinates are the $\mathbb{R}^n$-coordinates of the vector $\mathbf{x} \in \mathbb{R}^m \subset \mathbb{R}^n$.