# KoG

No. 15. (2011)
ISSN 1331-1611

SCIENTIFIC-PROFESSIONAL JOURNAL OF CROATIAN SOCIETY FOR GEOMETRY AND GRAPHICS

Official publication of the Croatian Society for Geometry and Graphics publishes scientific and professional papers from the fields of geometry, applied geometry and computer graphics.

**Founder and Publisher**

Croatian Society for Geometry and Graphics



**Design**

Miroslav Ambruš-Kiš

**Layout**

Sonja Gorjanc, Ema Jurkin

**Cover Illustration**

F. V. Holi, *photography*

**Print**

"O-TISAK", d.o.o., Zagreb

**URL address**

http://www.hdgg.hr/kog
http://hrcak.srce.hr

**Edition**

150

**Published annually**

**Guide for authors**

Please, see the last page

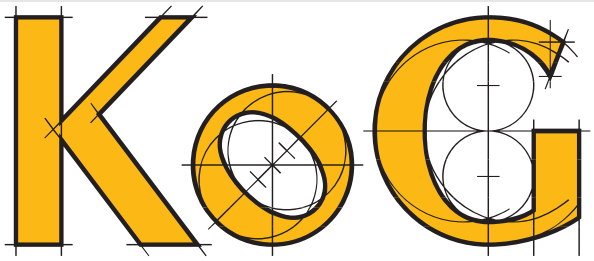
**KoG** is cited in: Mathematical Reviews, MathSciNet, Zentralblatt für Mathematik

---

This issue has been financially supported by The Ministry of Science, Education and Sport of the Republic of Croatia.

ISSN 1331–1611

No. 15
Zagreb, 2011



SCIENTIFIC AND PROFESSIONAL JOURNAL OF
CROATIAN SOCIETY FOR GEOMETRY AND GRAPHICS

# CONTENTS

## ORIGINAL SCIENTIFIC PAPERS



## PROFESSIONAL PAPERS

ISSN 1331–1611

KoG

BROJ 15
Zagreb, 2011

ZNANSTVENO-STRUČNI ČASOPIS
HRVATSKOG DRUŠTVA ZA GEOMETRIJU I GRAFIKU

# SADRŽAJ

## IZVORNI ZNANSTVENI RADOVI



## STRUČNI RADOVI

Original scientific paper
Accepted 18. 12. 2011.

JÁNOS PALLAGI, BENEDEK SCHULTZ, JENÖ SZIRMAI

# Equidistant Surfaces in $\mathbf{H}^2\times\mathbf{R}$ Space

**Equidistant Surfaces in $\mathbf{H}^2\times\mathbf{R}$ Space**

**ABSTRACT**

After having investigated the equidistant surfaces ("perpendicular bisectors" of two points) in $\mathbf{S}^2\times\mathbf{R}$ space (see [6]) we consider the analogous problem in $\mathbf{H}^2\times\mathbf{R}$ space from among the eight Thurston geometries. In [10] the third author has determined the geodesic curves, geodesic balls of $\mathbf{H}^2\times\mathbf{R}$ space and has computed their volume, has defined the notion of the geodesic ball packing and its density. Moreover, he has developed a procedure to determine the density of the geodesic ball packing for generalized Coxeter space groups of $\mathbf{H}^2\times\mathbf{R}$ and he has applied this algorithm to them.

In this paper we introduce the notion of the equidistant surface to two points in $\mathbf{H}^2\times\mathbf{R}$ geometry, determine its equation and we shall visualize it in some cases. The pictures have been made by the Wolfram Mathematica software.

**Key words:** non-Euclidean geometries, geodesic curve, geodesic sphere, equidistant surface in $\mathbf{H}^2\times\mathbf{R}$ geometry

**MSC 2010:** 53A35, 51M10, 51M20, 52C17, 52C22

**Ekvidistantne plohe u prostoru $\mathbf{H}^2\times\mathbf{R}$**

**SAŽETAK**

Nakon istraživanja ekvidistanthih ploha ("okomitih simetrala" dviju točaka) u prostoru $\mathbf{S}^2\times\mathbf{R}$ (vidi [6]), razmatramo analogni problem u prostoru $\mathbf{H}^2\times\mathbf{R}$ iz osam Thurstonovih geometrija. U radu [10] treći je autor odredio geodetske krivulje i kugle prostora $\mathbf{H}^2\times\mathbf{R}$ te definirao pojam popunjavanja geodetskim kuglama i njegovu gustoću. Pored toga, razvio je metodu određivanja gustoće popunjavanja geodetskim kuglama za generalizirane Coxeterove grupe prostora $\mathbf{H}^2\times\mathbf{R}$ i primijenio taj algoritam na njih. U ovom radu uvodimo pojam ekvidistantne plohe dviju točaka u geometriji $\mathbf{H}^2\times\mathbf{R}$, određujemo njihovu jednadžbu i vizualiziramo neke slučajeve. Slike su napravljene u Wolframovom programu Mathematica.

**Ključne riječi:** neeuklidske geometrije, geodetska krivulja, geodetska sfera, ekvidistantna ploha u $\mathbf{H}^2\times\mathbf{R}$ geometriji

## 1 Basic notions of $\mathbf{H}^2\times\mathbf{R}$ geometry

The $\mathbf{H}^2\times\mathbf{R}$ geometry is one one of the eight simply connected 3-dimensional maximal homogeneous Riemannian geometries. This Seifert fibre space is derived by the direct product of the hyperbolic plane $\mathbf{H}^2$ and the real line $\mathbf{R}$. The points are described by $(P,p)$ where $P\in\mathbf{H}^2$ and $p\in\mathbf{R}$.

In [2] E. Molnár has shown, that the homogeneous 3-spaces have a unified interpretation in the projective 3-sphere $\mathcal{PS}^3(\mathbf{V}^4,V_4,\mathbf{R})$. In our work we shall use this projective model of $\mathbf{H}^2\times\mathbf{R}$ and the Cartesian homogeneous coordinate simplex $E_0(\mathbf{e}_0)$,$E_1^\infty(\mathbf{e}_1)$,$E_2^\infty(\mathbf{e}_2)$, $E_3^\infty(\mathbf{e}_3)$, $(\{\mathbf{e}_i\}\subset\mathbf{V}^4$ with the unit point $E(\mathbf{e}=\mathbf{e}_0+\mathbf{e}_1+\mathbf{e}_2+\mathbf{e}_3))$ which is distinguished by an origin $E_0$ and by the ideal points of coordinate axes, respectively. Moreover, $\mathbf{y}=c\mathbf{x}$ with $0<c\in\mathbf{R}$ (or $c\in\mathbb{R}\setminus\{0\}$) defines a point $(\mathbf{x})=(\mathbf{y})$ of the projective 3-sphere $\mathcal{PS}^3$ (or that of the projective space $\mathcal{P}^3$ where opposite rays $(\mathbf{x})$ and $(-\mathbf{x})$ are identified). The dual system $\{(e^i)\}\subset V_4$ describes the simplex planes, especially the plane at infinity $(e^0)=E_1^\infty E_2^\infty E_3^\infty$, and generally, $v=u\frac{1}{c}$ defines a plane $(u)=(v)$ of $\mathcal{PS}^3$ (or that of $\mathcal{P}^3$, respectively). Thus $0=\mathbf{x}u=\mathbf{y}v$ defines the incidence of point $(\mathbf{x})=(\mathbf{y})$ and plane $(u)=(v)$, as $(\mathbf{x})\mathrm{I}(u)$ also denotes it. Thus $\mathbf{H}^2\times\mathbf{R}$ can be visualized in the affine 3-space $\mathbf{A}^3$ (so in $\mathbf{E}^3$) as well.

The point set of $\mathbf{H}^2\times\mathbf{R}$ in the projective space $\mathcal{P}^3$, are the following open cone solid (see Fig. 1-2):

$$\mathbf{H}^2\times\mathbf{R}:=$$
$$\left\{X(\mathbf{x}=x^i\mathbf{e}_i)\in\mathcal{P}^3:-(x^1)^2+(x^2)^2+(x^3)^2<0<x^0,\,x^1\right\}.$$

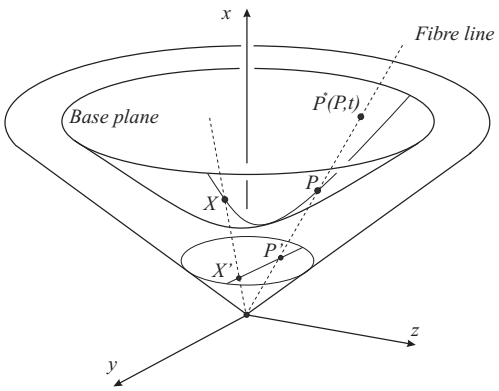



Figure 1: *Projective model of* $\mathbf{H}^2\times\mathbf{R}$ *geometry*

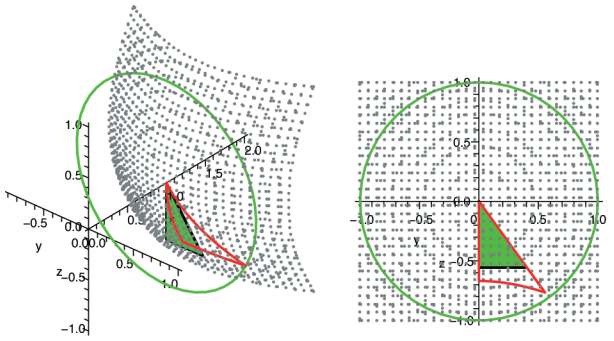


Figure 2: *The connection between Cayley-Klein model of the hyperbolic plane and the "base plane" of the model of* $\mathbf{H}^2\times\mathbf{R}$ *geometry.*

In this context E. Molnár [2] has derived the infinitesimal arc-length square at any point of $\mathbf{H}^2\times\mathbf{R}$ as follows

$$\begin{aligned}(ds)^2 =& \frac{1}{-x^2+y^2+z^2}\cdot\Big[(x^2+y^2+z^2)(dx)^2+\\ &2dxdy(-2xy)+2dxdz(-2xz)+\\ &(x^2+y^2-z^2)(dy)^2+\\ &2dydz(2yz)(x^2-y^2+z^2)(dz)^2\Big].\end{aligned}\tag{1}$$

By introducing the new $(t,r,\alpha)$ coordinates in (2), our formula becomes simplier in (3): $-\pi<\alpha\leq\pi$ and $r\geq 0$ with $t\in\mathbf{R}$ the fibre coordinate. The proper points can be described by the following equations:

$$\begin{aligned}&x^0=1,\ \ x^1=e^t\cosh r,\\ &x^2=e^t\sinh r\cos\alpha,\ \ x^3=e^t\sinh r\sin\alpha.\end{aligned}\tag{2}$$

We apply the usual Cartesian coordinates for the visualization and further computations, i.e. $x=x^1/x^0, y=x^2/x^0, z=x^3/x^0$. So the infinitesimal arc length square with coordinates $(t,r,\alpha)$ at any proper point of $\mathbf{H}^2\times\mathbf{R}$ - and the symmetric metric tensor $g_{ij}$ obtained from it - are the following:

$$(ds)^2=(dt)^2+(dr)^2+\sinh^2 r(d\alpha)^2,\tag{3}$$

$$g_{ij}:=\begin{pmatrix}1&0&0\\0&1&0\\0&0&\sinh^2 r\end{pmatrix}.\tag{4}$$

By the usual method of the differential geometry we have obtained the equation system of the geodesic curves [5]:

$$\begin{aligned}&x(\tau)=e^{\tau\sin v}\cosh(\tau\cos v),\\ &y(\tau)=e^{\tau\sin v}\sinh(\tau\cos v)\cos u,\\ &z(\tau)=e^{\tau\sin v}\sinh(\tau\cos v)\sin u,\\ &-\pi<u\leq\pi,\ \ -\frac{\pi}{2}\leq v\leq\frac{\pi}{2}.\end{aligned}\tag{5}$$

**Remark 1.1** *The starting point of our geodesics can be chosen at* $(1,1,0,0)$ *by the homogeneity of* $\mathbf{H}^2\times\mathbf{R}$.

**Definition 1.2** *The distance* $d(P_1,P_2)$ *between the points* $P_1$ *and* $P_2$ *is defined by the arc length* $s=\tau$ *in (5) of the geodesic curve from* $P_1$ *to* $P_2$.

**Definition 1.3** *The geodesic sphere of radius* $\rho$ *(denoted by* $S_{P_1}(\rho)$*) with center at the point* $P_1$ *is defined as the set of all points* $P_2$ *in the space with the condition* $d(P_1,P_2)=\rho$. *We also require that the geodesic sphere is a simply connected surface without selfintersection in* $\mathbf{H}^2\times\mathbf{R}$ *space (see Fig. 3).*

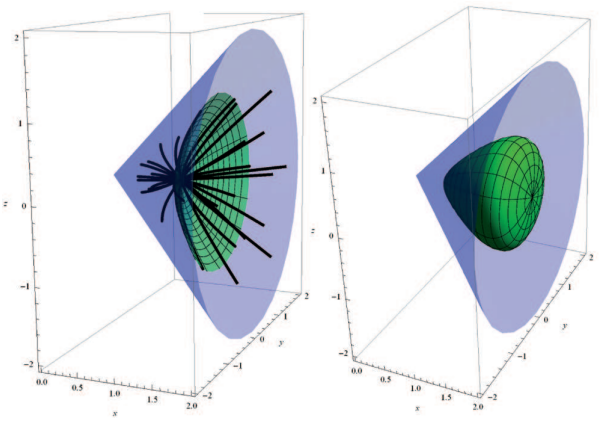


Figure 3: *Geodesics with varying parameters and the "base-hyperboloid" in the cone and a geodesic sphere with radius* $\frac{2}{3}$ *centered at* $(1,1,0,0)$.

### 1.1 Equidistant surfaces in $\mathbf{H}^2\times\mathbf{R}$ geometry

One of our further goals is to visualize and examine the Dirichlet-Voronoi cells of $\mathbf{H}^2\times\mathbf{R}$ where the faces of the DV-cells are equidistant surfaces. The definition below comes naturally.

**Definition 1.4** *The equidistant surface* $\mathcal{S}_{P_1P_2}$ *of two arbitrary points* $P_1,P_2\in\mathbf{H}^2\times\mathbf{R}$ *consists of all points* $P'\in\mathbf{H}^2\times\mathbf{R}$, *for which* $d(P_1,P')=d(P',P_2)$. *Moreover, we require that this surface is a simply connected piece without selfintersection in* $\mathbf{H}^2\times\mathbf{R}$ *space.*

It can be assumed by the homogeneity of $\mathbf{H}^2\times\mathbf{R}$ that the starting point of a given geodesic curve segment is $P_1(1,1,0,0)$. The other endpoint will be given by its homogeneous coordinates $P_2(1,a,b,c)$. We consider the geodesic curve segment $\mathcal{G}_{P_1P_2}$ and determine its parameters $(\tau,u,v)$ expressed by $a,b,c$. We obtain by equation system (5) the following identity :

$$\sqrt{a^2-b^2-c^2}=e^{\tau\sin v}\tag{6}$$

If we substitute this into (5), the equation system can be solved for $(\tau,u,v)$.

$$\tau=\frac{\log\sqrt{a^2-b^2-c^2}}{\sin v},\quad \text{if } v\neq 0. \tag{7}$$

$$v=\arctan\left(\frac{\log\sqrt{a^2-b^2-c^2}}{\operatorname{arccosh}\left(\frac{a}{\sqrt{a^2-b^2-c^2}}\right)}\right), \tag{8}$$

if $P_2(a,b,c)$ does not lie on the axis $[x]$ i.e. $(b,c)\neq(0,0)$.

$$\tan u=\frac{z(\tau)}{y(\tau)}=\frac{c}{b}\Rightarrow u=\arctan\left(\frac{c}{b}\right). \tag{9}$$

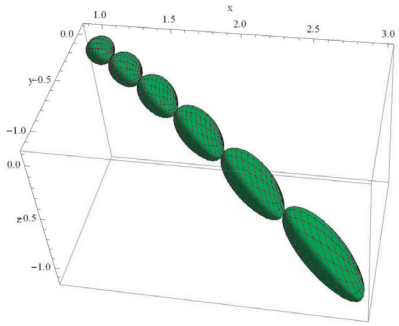


Figure 4: *Touching geodesic spheres of radius* $\frac{1}{10}$ *centered on the geodesic curve with starting point* $(1,1,0,0)$ *and parameters* $u=\frac{\pi}{4}$, $v=\frac{\pi}{3}\neq 0$.

**Remark 1.5** *If* $P_2\in[x]$, *then* $v=\frac{\Pi}{2}$ *and* $u=0$, *and the geodesic curve is an Euclidean line segment between* $P_1$ *and* $P_2$. *If* $v=0$, *then* $\tau=\operatorname{arccosha}$ *and the two points are on the same hyperboloid surface. These special cases will be discussed in section 3 in terms of the equidistant surfaces belonging to them.*

It is clear that $X\in\mathcal{S}_{P_1P_2}$ iff $d(P_1,X)=d(X,P_2)\Rightarrow d(P_1,X)=d(X^{\mathcal{F}},P_2^{\mathcal{F}})$, where $\mathcal{F}$ is a composition of isometries which maps $X$ onto $(1,1,0,0)$, and then by (7) the length of the geodesic (e.g. the distance between the two points) is comparable to $d(P_1,X)$. This method leads to the implicit equation of the equidistant surface of two proper points $P_1(1,a,b,c)$ and $P_2(1,d,e,f)$ in $\mathbf{H}^2\times\mathbf{R}$:

$$\begin{aligned}&\mathcal{S}_{P_1P_2}(x,y,z)\Rightarrow\\&4\operatorname{arccosh}^2\left(\frac{ax-by-cz}{\sqrt{a^2-b^2-c^2}\sqrt{x^2-y^2-z^2}}\right)+\\&\log^2\left(\frac{a^2-b^2-c^2}{x^2-y^2-z^2}\right)=\\&=4\operatorname{arccosh}^2\left(\frac{dx-ey-fz}{\sqrt{d^2-e^2-f^2}\sqrt{x^2-y^2-z^2}}\right)+\\&\log^2\left(\frac{d^2-e^2-f^2}{x^2-y^2-z^2}\right).\end{aligned} \tag{10}$$

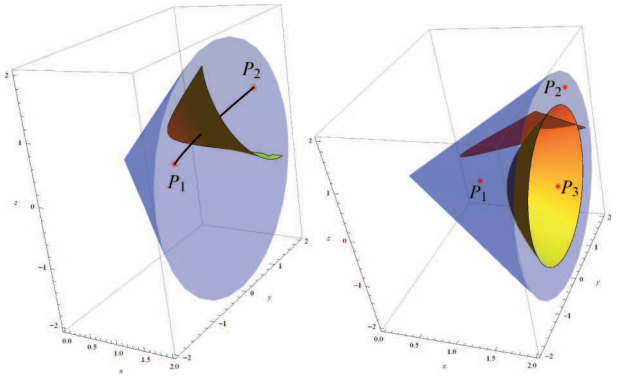



Figure 5: *Equidistant surfaces with* $P_1(1,1,0,0)$ *and* $P_2(1,2,1,1)$, *and the two special cases.*

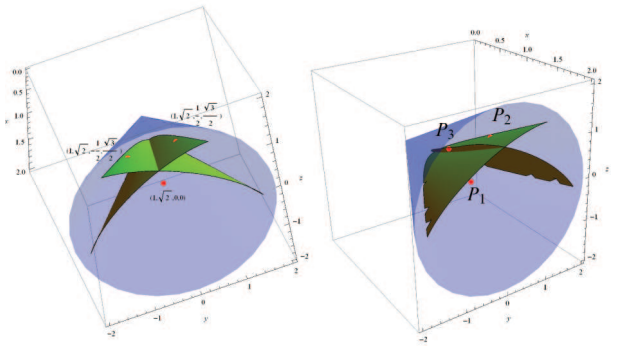



Figure 6: *Equidistant surfaces to points* $P_1(1,\sqrt{2},0,0)$, $P_2(1,\sqrt{2},\frac{1}{2},\frac{\sqrt{3}}{2})$ *and* $P_1(1,\sqrt{2},0,0)$, $P_3(1,\sqrt{2},-\frac{1}{2},\frac{\sqrt{3}}{2})$.

### 1.2 Some observations

We introduce the next denotations to simplify the equation (10): $\mathbf{a}=\overrightarrow{OP_1}$, $\mathbf{b}=\overrightarrow{OP_2}$ and $\mathbf{x}=\overrightarrow{OX}$. We define the scalar product for all vectors $\mathbf{u}(u_1,u_2,u_3)$ and $\mathbf{v}(v_1,v_2,v_3)$ by the following equation:

$$\langle\mathbf{u},\mathbf{v}\rangle=-u_1v_1+u_2v_2+u_3v_3,$$

moreover, we introduce the denotation $|\mathbf{v}|=\sqrt{-\langle\mathbf{v},\mathbf{v}\rangle}$ similarly to the $\mathbf{S}^2\times\mathbf{R}$ space (see [6]).

With these denotations, the equation of the surface becomes shorter and gives important informations about equidistant surfaces:

$$\begin{aligned}&\operatorname{arccosh}^2\left(\frac{-\langle\mathbf{a},\mathbf{x}\rangle}{|\mathbf{a}||\mathbf{x}|}\right)+\log^2\left(\frac{|\mathbf{a}|}{|\mathbf{x}|}\right)=\\&\operatorname{arccosh}^2\left(\frac{-\langle\mathbf{x},\mathbf{b}\rangle}{|\mathbf{x}||\mathbf{b}|}\right)+\log^2\left(\frac{|\mathbf{b}|}{|\mathbf{x}|}\right).\end{aligned}$$

The last step is to notice that $\operatorname{arccosh}\left(\frac{-\langle \mathbf{a},\mathbf{x}\rangle}{|\mathbf{a}||\mathbf{x}|}\right)$ is the hyperbolic disance between points $\mathbf{a}$ and $\mathbf{x}$ in the projective model of the hyperbolic plane. So let $\varepsilon = d_h(\mathbf{a},\mathbf{x})$ and $\delta = d_h(\mathbf{x},\mathbf{b})$. The final form of the equation is the following:

$$\varepsilon^2 + \log^2(|\mathbf{a}||\mathbf{x}|^{-1}) = \delta^2 + \log^2(|\mathbf{b}||\mathbf{x}|^{-1}) \qquad (11)$$

**Remark 1.6** *This formula also describes the equidistant surface of* $\mathbf{S}^2\times\mathbf{R}$ *with the usual Euclidean scalar product, vector length and angle formula (see [6]).*

It is now easy to examine some special cases: when $|\mathbf{a}| = |\mathbf{b}|$, the equidistant surface consists of those points of an Euclidean plane in our model, which are inner points of the cone (e.g. proper point of $\mathbf{H}^2\times\mathbf{R}$). Another special case appears when $\mathbf{a}$ and $\mathbf{b}$ are on the same *fibre*. In this case ($\delta = \varepsilon$) the equidistant surface is the "positive side" of a hyperboloid of two sheets.

Our projective method gives us a way of investigation the $\mathbf{H}^2\times\mathbf{R}$ space, which suits to study and solve similar problems (see [10]). In this paper we have examined only some problems, but analogous questions in $\mathbf{H}^2\times\mathbf{R}$ geometry or, in general, in other homogeneous Thurston geometries are timely (see [11], [8], [9]).

## References


[1] A. M. MACBEATH, The classification of non-Euclidean plane crystallographic groups. *Can. J. Math.,* **19** (1967), 1192–1295.

[2] E. MOLNÁR, The projective interpretation of the eight 3-dimensional homogeneous geometries. *Beiträge zur Algebra und Geometrie,* **38** (1997) No. 2, 261–288.

[3] E .MOLNÁR, I. PROK, J. SZIRMAI, Classification of tile-transitive 3-simplex tilings and their realizations in homogeneous spaces. *Non-Euclidean Geometries, János Bolyai Memorial Volume*, Ed. A. PREKOPA and E. MOLNÁR, Mathematics and Its Applications **581**, Springer (2006), 321–363.

[4] E. MOLNÁR, J. SZIRMAI, Symmetries in the 8 homogeneous 3-geometries. *Symmetry: Culture and Science*, Vol. **21 No. 1-3** (2010), 87-117.

[5] E. MOLNÁR, B. SZILÁGYI, Translation curves and their spheres in homogeneous geometries. *Publicationes Math. Debrecen*, Vol. **78/2** (2011), 327–346.

[6] J. PALLAGI, B. SCHULTZ, J. SZIRMAI, Visualization of geodesic curves, spheres and equidistant surfaces in $\mathbf{S}^2\times\mathbf{R}$ space. *KoG* **14**, (2010), 35-40.

[7] P. SCOTT, The geometries of 3-manifolds. *Bull. London Math. Soc.*, **15** (1983) 401–487. (Russian translation: Moscow "Mir" 1986.)

[8] J. SZIRMAI, The optimal ball and horoball packings to the Coxeter honeycombs in the hyperbolic d-space. *Beiträge zur Algebra und Geometrie*, **48** No. 1 (2007), 35–47.

[9] J. SZIRMAI, The densest geodesic ball packing by a type of **Nil** lattices. *Beiträge zur Algebra und Geometrie*, **48** No. 2 (2007), 383–398.

[10] SZIRMAI, J. Geodesic ball packings in $\mathbf{H}^2\times\mathbf{R}$ space for generalized Coxeter space groups. *Mathematical Communications,* to appear (2011).

[11] J. SZIRMAI, Geodesic ball packings in $\mathbf{S}^2\times\mathbf{R}$ space for generalized Coxeter space groups. *Beiträge zur Algebra und Geometrie*, **52**(2011), 413–430.

[12] W. P. THURSTON (and S. LEVY, editor), *Three-Dimensional Geometry and Topology*. Princeton University Press, Princeton, New Jersey, Vol **1** (1997).



**János Pallagi**
e-mail: jpallagi@math.bme.hu

**Benedek Schultz**
e-mail: schultz.benedek@gmail.com

**Jenö Szirmai**
e-mail: szirmai@math.bme.hu

Budapest University of Technology and Economics,
Institute of Mathematics, Department of Geometry
H-1521 Budapest, Hungary


**Acknowledgement:** We thank Prof. Emil Molnár for helpful comments to this paper.

Original scientific paper
Accepted 21. 12. 2011.

**MILJENKO LAPAINE**

# Mollweide Map Projection

**Mollweide Map Projection**

**ABSTRACT**

Karl Brandan Mollweide (1774-1825) was German mathematician and astronomer. The formulas known after him as Mollweide's formulas are shown in the paper, as well as the proof "without words". Then, the Mollweide map projection is defined and formulas derived in different ways to show several possibilities that lead to the same result. A generalization of Mollweide projection is derived enabling to obtain a pseudocylindrical equal-area projection having the overall shape of an ellipse with any prescribed ratio of its semiaxes. The inverse equations of Mollweide projection has been derived, as well.

The most important part in research of any map projection is distortion distribution. That means that the paper continues with the formulas and images enabling us to get some filling about the linear and angular distortion of the Mollweide projection.

Finally, several applications of Mollweide projections are represented, with the International Cartographic Association logo as an example of one of its successful applications.

**Key words:** Mollweide, Mollweide's formula, Mollweide map projection

**MSC 2010:** 51N20, 01A55, 51-03, 51P05, 86A30

**Mollweideova kartografska projekcija**

**SAŽETAK**

Karl Brandan Mollweide (1774-1825) bio je njemački matematičar i astronom. U ovom radu prikazane su formule nazvane po njemu kao Mollweideove formule, a uz njih "dokaz bez riječi". Zatim je definirana Mollweideova kartografska projekcija uz izvod formula na nekoliko različitih načina kako bi se pokazalo da postoji više mogućnosti koje vode do istoga rezultata. Izvedena je generalizacija Mollweideove projekcije koja omogućava dobivanje pseudocilindričnih ekvivalentnih (istopovršinskih) projekcija smještenih u elipsu s bilo kojim unaprijed zadanim odnosnom njezinih poluosi. Izvedene su i inverzne jednadžbe Mollweideove projekcije.

Najvažniji dio istraživanja svake kartografske projekcije je ustanovljavanje razdiobe deformacija. Stoga su u radu dane formule i grafički prikazi koji daju uvid u razdiobu linearnih i kutnih deformacija Mollweideove projekcije.

Na kraju je prikazano nekoliko primjena Mollweideove projekcije. Među njima je i logotip Međunarodnoga kartografskog društva, kao jedan od primjera njezine uspješne primjene.

**Ključne riječi:** Mollweide, Mollweideova formula, Mollweideova kartografska projekcija

## 1 Mollweide's Formulas

In trigonometry, Mollweide's formula, sometimes referred to in older texts as Mollweide's equations, named after Karl Mollweide, is a set of two relationships between sides and angles in a triangle. It can be used to check solutions of triangles.

Let $a$, $b$, and $c$ be the lengths of the three sides of a triangle. Let $\alpha$, $\beta$, and $\gamma$ be the measures of the angles opposite those three sides respectively. Mollweide's formulas state that

$$\frac{a+b}{c} = \frac{\cos\frac{\alpha-\beta}{2}}{\sin\frac{\gamma}{2}} \quad \text{and} \quad \frac{a-b}{c} = \frac{\sin\frac{\alpha-\beta}{2}}{\cos\frac{\gamma}{2}}.$$

Each of these identities uses all six parts of the triangle - the three angles and the lengths of the three sides.

These trigonometric identities appear in Mollweide's paper *Zusätze zur ebenen und sphärischen Trigonometrie* (1808). A proof without words of these identities (see Fig. 1) is given in DeKleine (1988) and Nelsen (1993).

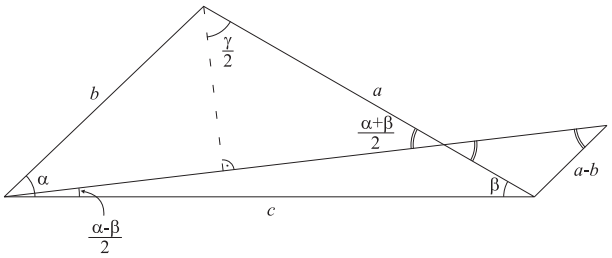



Figure 1: *Mollweide equation - Proof without Words. According DeKleine (1988).*

One of the more puzzling aspects is why these equations should have become known as the Mollweide equations since in the 1808 paper in which they appear Mollweide refers the book by Antonio Cagnoli (1743-1816) *Traité de Trigonométrie Rectiligne et Sphérique, Contenant des Méthodes et des Formules Nouvelles, avec des Applications à la Plupart des Problêmes de l'astronomie* (1786) which contains the formulas. However, the formulas go back to Isaac Newton, or even earlier, but there is no doubt that Mollweide's discovery was made independently of this earlier work (URL1).

## 2 Mollweide Map Projection Equations

Pseudocylindrical map projections have in common straight parallel lines of latitude and curved meridians. Until the 19th century the only pseudocylindrical projection with important properties was the sinusoidal or Sanson-Flamsteed. The sinusoidal has equally spaced parallels of latitude, true scale along parallels, and equivalency or equal-area. As a world map, it has disadvantage of high distortion at latitudes near the poles, especially those farthest from the central meridian (Fig. 2).

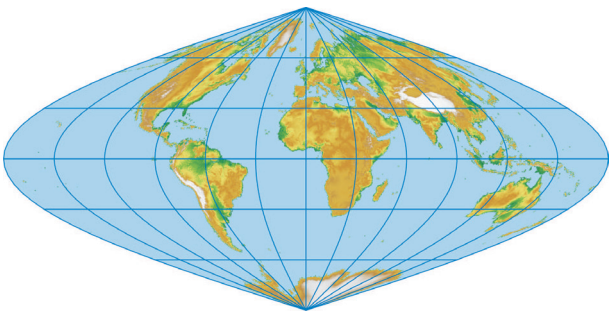


Figure 2: *Sanson or Sanson-Flamsteed or Sinusoidal projection*

In 1805, Mollweide announced an equal-area world map projection that is aesthetically more pleasing than the sinusoidal because the world is placed in an ellipse with axes in a 2:1 ratio and all the meridians are equally spaced semiellipses. The Mollweide projection was the only new pseudocylindrical projection of the nineteenth century to receive much more than academic interest (Fig. 3).

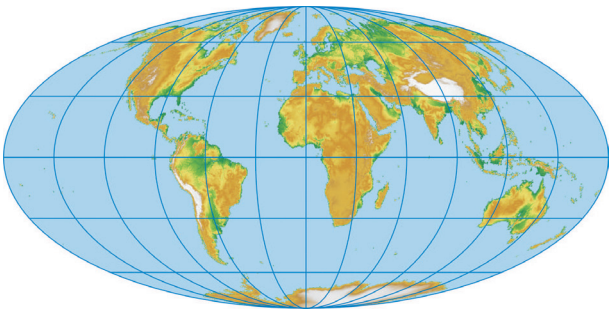


Figure 3: *Mollweide projection*

Mollweide presented his projection in response to a new globular projection of a hemisphere, described by Georg Gottlieb Schmidt (1768-1837) in 1803 and having the same arrangement of equidistant semiellipses for meridians. But Schmidt's curved parallels do not provide the equal-area property that Mollweide obtained (Snyder, 1993).

O'Connor and Robertson (URL1) stated that Mollweide produced the map projection to correct the distortions in the Mercator projection, first used by Gerardus Mercator in 1569. While the Mercator projection is well adapted for sea charts, its very great exaggeration of land areas in high latitudes makes it unsuitable for most other purposes. In the Mercator projection the angles of intersection between the parallels and meridians, and the general configuration of the land, are preserved but as a consequence areas and distances are increasingly exaggerated as one moves away from the equator. To correct these defects, Mollweide drew his elliptical projection; but in preserving the correct relation between the areas he was compelled to sacrifice configuration and angular measurement. The Mollweide projection lay relatively dormant until J. Babinet reintroduced it in 1857 under the name homalographic. The projection has been also called the Babinet, homalographic, homolographic and elliptical projection. It is discussed in many articles, see for example Boggs (1929), Close (1929), Feeman (2000), Philbrick (1953), Reeves (1904) and Snyder (1977) and books or textbooks by Fiala (1957), Graur (1956), Kavrajskij (1960), Kuntz (1990), Maling (1980), Snyder (1987, 1993), Solov'ev (1946) and Wagner (1949). The well known equations of the Mollweide projections read as follows:

$$x = \sqrt{2}R\sin\beta \tag{1}$$

$$y = \frac{2\sqrt{2}}{\pi}R\lambda\cos\beta \tag{2}$$

$$2\beta + \sin 2\beta = \pi\sin\varphi. \tag{3}$$

In these formulas $x$ and $y$ are rectangular coordinates in the plane of projection, $\varphi$ and $\lambda$ are geographic coordinates of the points on the sphere and $R$ is the radius of the sphere to be mapped. The angle $\beta$ is an auxiliary angle that is connected with the latitude $\varphi$ by the relation (3). For given latitude $\varphi$, the equation (3) is a transcendental equation in $\beta$. In the past, it was solving by using tables and interpolation method. In our days, it is usually solved by using some iterative numerical method, like bisection or Newton-Raphson method.

### 2.1 First approach

A half of the sphere with the radius $R$ should be mapped onto the disk with the radius $\rho$ (adopted from Borčić, 1955). If we request that the area of the hemisphere is equal to the area of the disk, than there is the following relation:

$$2R^2\pi = \rho^2\pi \tag{4}$$

from where we have

$$\rho = \sqrt{2}R. \tag{5}$$

Let the circle having the radius $\rho$ be the image of the meridians with the longitudes $\lambda = \pm\frac{\pi}{2}$. From Fig. 4 we see that the rectangular coordinates $x_0$ and $y_0$ of any point $T_0$ belonging to this circle can be written like this:

$$x_0 = \rho\sin\beta \tag{6}$$

$$y_0 = \rho\cos\beta. \tag{7}$$

Due to the request that the projection should be pseudocylindrical, the abscise $x = x_0$ for any point with the same latitude regardless of the longitude the relation (1) holds.

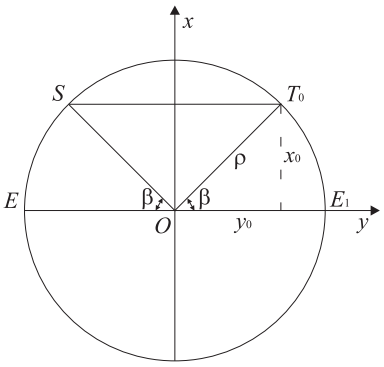



Figure 4: *Derivation of Mollweide projection equations*

On the other hand, the ordinate $y$ will depend on the latitude and longitude. According to the equal-area condition, the following relation exists:

$$y_0 : y = \frac{\pi}{2} : \lambda. \tag{8}$$

By using (8) and (5), the relation (7) goes into (2). In order to finish the derivation, we need to find the relation between the auxiliary angle $\beta$, and the latitude $\varphi$. According to the equal-area condition, the area $SEE_1T_0$ should be equal to the area of the spherical segment between the equator and the parallel of latitude $\varphi$, which is mapped as the straight-line segment $ST_0$:

$$\Delta OST_0 + 2OT_0E_1 = R^2\pi\sin\varphi,$$

that is

$$\frac{\rho^2}{2}\sin(\pi - 2\beta) + \frac{2\rho}{2}\beta\rho = R^2\pi\sin\varphi \tag{9}$$

from where we have (3).

### 2.2 Second approach

Given the earth's radius $R$, suppose the equatorial aspect of an equal-area projection with the following properties:

- A world map is bounded by an ellipse twice broader than tall
- Parallels map into parallel straight lines with uniform scale
- The central meridian is a part of straight line; all other ones are semielliptical arcs.

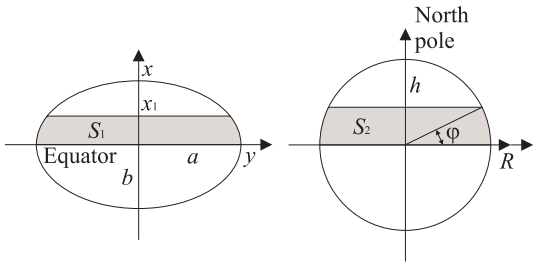



Figure 5: *Second approach to derivation of Mollweide projection equations*

Suppose an earth-sized map; let us define two regions, $S_1$ on the map and $S_2$ on the earth, both bounded by the equator and a parallel (URL2). The equal-area property can be used to calculate $x$ for given $\varphi$. Given $x$ and $\lambda$, $y$ can be calculated immediately from the ellipse equation, since horizontal scale is constant.
Equation of ellipse centred in origin, with major axis on $y$-axis is:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \text{ or}$$

$$y^2 = b^2\left(1 - \frac{x^2}{a^2}\right).$$

For $0 \le x \le a$

$$y = \frac{b}{a}\sqrt{a^2 - x^2}.$$

Area between $y$-axis and parallel mapped into $x = x_1$ is

$$S_1 = 2\int_0^{x_1} y dx = 2\frac{b}{a}\int_0^{x_1} \sqrt{a^2 - x^2} dx$$

Let $x = a\sin\beta, 0 \le x \le a, 0 \le \beta \le \frac{\pi}{2}, dx = a\cos\beta d\beta$, then

$$\int \sqrt{a^2 - x^2} dx = \int \sqrt{a^2(1 - \sin^2\beta)}\, a\cos\beta d\beta =$$

$$a^2 \int \cos^2\beta d\beta.$$

Since $\cos^2\alpha = \frac{1+\cos 2\alpha}{2}$

$$a^2 \int \cos^2\beta d\beta = a^2 \int \frac{1+\cos 2\beta}{2} d\beta =$$

$$= \frac{a^2}{2}\left(\int d\beta + \int \cos 2\beta d\beta\right) = \frac{a^2}{2}\left(\beta + \frac{\sin 2\beta}{2}\right) + C$$

$$S_1 = 2\frac{b}{a}\left(\frac{a^2}{2}\left(\beta + \frac{\sin 2\beta}{2}\right) + C\right)_0^{\beta} =$$

$$= \frac{ab}{2}(2\beta + \sin 2\beta) = 2R^2(2\beta + \sin 2\beta)$$

for some $0 \le \beta \le \frac{\pi}{2}$, corresponding to $x_1 = a\sin\beta$ and because of $ab\pi = 4R^2\pi$.

On a sphere, the area between the equator and parallel $\varphi$ is

$$S_2 = 2\pi Rh = 2\pi R^2 \sin\varphi$$

$S_1 = S_2 \Rightarrow 2R^2(2\beta + \sin 2\beta) = 2\pi R^2 \sin\varphi$, i.e. (3).

The auxiliary angle $\beta$ must be found by interpolation or successive approximation. Finally, since horizontal scale is uniform, and $ab\pi = 4R^2\pi, b = 2a$ and $a = \sqrt{2}R$ we have (1). Due to the relation

$$y : \lambda = \frac{b}{a}\sqrt{a^2 - x^2} : \pi$$

$$y = \frac{2\lambda}{\pi}\sqrt{2R^2 - x^2} = 2\sqrt{2R^2 - 2R^2\sin^2\beta}\frac{\lambda}{\pi},$$ i.e. (2) holds.

### 2.3 Third approach

From the theory of map projections we know that general equations of pseudocylindrical projections have the form:

$$x = x(\varphi) \tag{10}$$

$$y = y(\varphi, \lambda) \tag{11}$$

Furthermore, for equal-area pseudocylindrical projection holds (Borčić, 1955)

$$y = \frac{R^2\cos\varphi}{\frac{dx}{d\varphi}}\lambda \tag{12}$$

Let us suppose that a half of the sphere has to be mapped onto a disc with the boundary

$$x^2 + y^2 = \rho^2.$$

In order to have an equal-area mapping of the half of the sphere with the radius $R$ onto a disc with the radius $\rho$ we should have

$$2R^2\pi = \rho^2\pi$$

from where

$$\rho^2 = 2R^2.$$

That implies

$$x^2 + y^2 = 2R^2.$$

Taking into account (12) for $\lambda = \pm\frac{\pi}{2}$

$$y = \pm\frac{R^2\pi}{2}\frac{\cos\varphi}{\frac{dx}{d\varphi}}$$

$$x^2 + \frac{R^4\pi^2}{4}\frac{\cos^2\varphi}{\left(\frac{dx}{d\varphi}\right)^2} = 2R^2.$$

That is a differential equation that could be solved by the method of separation of variables:

$$2\sqrt{2R^2 - x^2}dx = R^2\pi\cos\varphi d\varphi$$

where the sign + has been chosen. After integration we can get

$$2\int \sqrt{2R^2 - x^2}dx = R^2\pi\sin\varphi + C$$

By the appropriate substitution in the integral on the left side, or just looking to any mathematical manual we can get the following:

$$2\cdot\frac{1}{2}\left(x\sqrt{2R^2 - x^2} + 2R^2\arcsin\frac{x}{R\sqrt{2}}\right) = R^2\pi\sin\varphi + C$$

For $\varphi = 0, x = 0$ and $C = 0$.
Therefore we have

$$x\sqrt{2R^2 - x^2} + 2R^2\arcsin\frac{x}{R\sqrt{2}} = R^2\pi\sin\varphi. \tag{13}$$

By substitution (1), (13) goes to (3), while (12) can be written as

$$y = \frac{2\lambda}{\pi}\sqrt{2R^2 - x^2}, \text{ which is equivalent to (2).}$$

**Remark 1**
Although the applied condition was that a half of the sphere has to be mapped onto a disc, the final projection equations hold for the whole sphere and give its image situated into an ellipse.

**Remark 2**
In references, the Mollweide projection is always defined by equations (1)-(3), which means by using an auxiliary angle or parameter. My equation (13) shows that there is no need to use any auxiliary parameter. There exists the direct relation between the x-coordinate and the latitude φ.

**Remark 3**
The method applied in this chapter can be applied in derivation of other pseudocylindrical equal-area projections, as are e.g. Sanson projection, Collignon projection or even cylindrical equal-area projection.

## 3 Generalization of Mollweide Projection

Let us consider the shape of the Mollweide projection of the whole sphere. From the equations (1) and (2), by elimination of β it is easy to obtain the equation of a meridian in the projection

$$\left(\frac{x}{\sqrt{2}R}\right)^2 + \left(\frac{\pi y}{2\sqrt{2}R\lambda}\right)^2 = 1. \tag{14}$$

It is obvious that for a given λ (14) is the equation of an ellipse. It follows that the semiaxis $a$ is constant, while $b$ depends on the longitude λ. If we take $\lambda = \pi$, than $b = 2\sqrt{2}R$, and

$$a : b = 1 : 2 \tag{15}$$

and that is the ratio of semiaxes in the Mollweide projection. The question arises: is it possible to find out a pseudocylindrical equal-area projection that will give the whole word in an arbitrary ellipse satisfying any given ratio $a : b$ or $b : a$? The answer is yes, and we are going to proof it. Let us denote $\mu = b : a$. First of all, the area of an ellipse with the semiaxes $a$ and $b = \mu a$ should be equal to the area of the whole sphere:

$$ab\pi = \mu a^2\pi = 4R^2\pi.$$

This is equivalent with

$$b = \frac{4R^2}{a}, \mu = \frac{4R^2}{a^2} \text{ or } a = \frac{2R}{\sqrt{\mu}}. \tag{16}$$

Now, the equation of the ellipse with the centre in the origin and with the semiaxes $a$ and $b$ reads

$$\frac{x^2}{a^2} + \frac{y^2}{\mu^2 a^2} = 1, \text{ or}$$

$$y^2 = \mu^2(a^2 - x^2). \tag{17}$$

Furthermore, the projection should be cylindrical and equal-area, which is generally expressed by (12). If we substitute (12) into (17), taking into account that $\lambda = \pi$, after some minor transformation we can get the following differential equation with separated variables

$$R^2\pi\cos\varphi d\varphi = \mu\sqrt{a^2 - x^2}dx. \tag{18}$$

Integral of the left side of the equation is elementary, while for that on the right side we need a substitution

$$x = a\sin\beta. \tag{19}$$

This leads to the equation

$$\pi\cos\varphi d\varphi = 4\cos^2\beta d\beta.$$

The application of the trigonometric identity

$$\cos^2\beta = \frac{1 + \cos 2\beta}{2}$$

gives us the following differential equation that is ready for integration:

$$\pi\cos\varphi d\varphi = 2(1 + \cos 2\beta)d\beta.$$

After integration, we obtain

$$\pi\sin\varphi = 2\beta + \sin 2\beta + C, \tag{20}$$

where $C$ is an integration constant. By using the natural conditions $\varphi = 0, x = 0$ and $\beta = 0$ we obtain $C = 0$. In that way, the final form of (5.8) is again the known relation (3). From (18) and (19) we have

$$\frac{dx}{d\varphi} = \frac{a^2\pi\cos\varphi}{4\sqrt{a^2 - x^2}} = \frac{a\pi\cos\varphi}{4\cos\beta} = \frac{R}{\sqrt{\mu}}\frac{\pi\cos\varphi}{2\cos\beta}$$

and taking into account (12)

$$y = \frac{4R^2}{a\pi}\lambda\cos\beta = \mu a\frac{\lambda}{\pi}\cos\beta = 2R\sqrt{\mu}\frac{\lambda}{\pi}\cos\beta,$$

while

$$x = a\sin\beta = \frac{2R}{\sqrt{\mu}}\sin\beta.$$

Let us summarize:

$$x = \frac{2R}{\sqrt{\mu}} \sin\beta.$$

$$y = 2R\sqrt{\mu}\frac{\lambda}{\pi}\cos\beta.$$

$$2\beta + \sin 2\beta = \pi \sin\varphi.$$

These are equations defining the generalized Mollweide projection onto an ellipse of any given ratio $\mu = b : a$ of its semiaxes.

***Example 1.***
Let us take $\mu = 1$, that is $a = b$, which means that we have a bounding circle. According to (16) $a = b = 2R$.

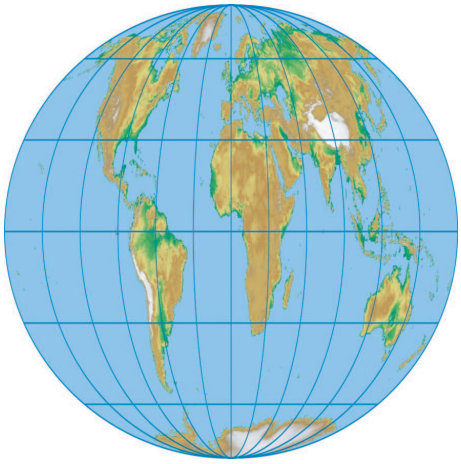


Figure 6: *Generalized Mollweide projection onto a disc*

***Example 2.***
Let us take $\mu = 2$, that is $b = 2a$. According to (16) $a = \sqrt{2}R, b = 2\sqrt{2}R$, and we are able to recognize the classic Mollweide projection (Fig. 3).

***Example 3.***
Let us define the ratio $\mu$, by the condition that the linear scale along the equator equals 1. From the theory of map projections it is known that the linear scale along parallels is given by

$$n = \frac{\sqrt{G}}{R\cos\varphi},$$

where

$$G = \left(\frac{\partial x}{\partial\lambda}\right)^2 + \left(\frac{\partial y}{\partial\lambda}\right)^2.$$

In our case $x = x(\varphi)$, which means that

$$\frac{\partial x}{\partial\lambda} = 0.$$

The condition

$n = 1$ for $\varphi = 0$

goes to

$$\frac{\partial y}{\partial\lambda} = R\cos\varphi = R.$$

Now,

$$\frac{\partial y}{\partial\lambda} = \frac{2R\sqrt{\mu}}{\pi}\cos\beta = R.$$

and from there and $\beta = 0$ due to $\varphi = 0$ we have

$$\sqrt{\mu} = \frac{\pi}{2}, \text{ or } \mu = \tfrac{\pi^2}{4}.$$

Finally, $a = \frac{4R}{\pi}$, $b = R\pi$ and

$$x = \frac{4}{\pi}R\sin\beta$$

$$y = R\lambda\cos\beta$$

$$2\beta + \sin 2\beta = \pi\sin\varphi.$$

It is easy to see that the linear scale in the direction of meridian is also 1 throughout the equator in this version of Mollweide projection (Fig. 7). See also Bromley (1965).

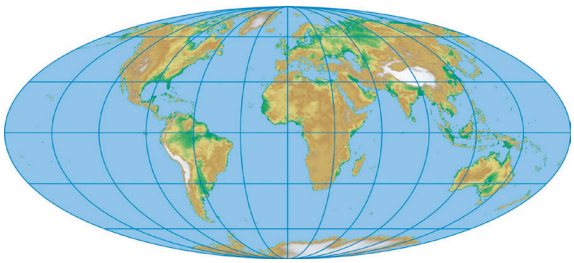


Figure 7: *Generalized Mollweide projection without linear distortions along the equator*

**Remark 4**
The same approach can be applied to find a generalized Mollweide projection satisfaying the condition $n = 1$ for $\varphi = \varphi_0$, where $0 \le \varphi_0 \le \frac{\pi}{2}$.

## 4 Inverse Equations of Mollweide Projection

The inverse equations of any map projections read as follows:

$$\varphi = \varphi(x, y)$$

$$\lambda = \lambda(x, y).$$

The computation of $\varphi$ and $\lambda$ from given $x$ and $y$ in Mollweide projection is straightforward. In fact, for the given $x$ from (1) we can get the auxiliary angle $\beta$

$$\sin\beta = \frac{x}{\sqrt{2}R}$$

Then, from (3) we have

$$\sin\varphi = \frac{1}{\pi}(2\beta + \sin 2\beta)$$

and from (2)

$$\lambda = \frac{\pi y}{2\sqrt{2}R\cos\beta}.$$

## 5 Distribution of Distortions in Mollweide Projection

For the Mollweide projection given by equations (1)–(3) it can be derived in the straightforward manner:

$$\tan\varepsilon = \frac{2\tan\beta}{\pi}\lambda$$

$$m = \frac{\pi\cos\varphi}{2\sqrt{2}\cos\beta\cos\varepsilon}$$

$$n = \frac{2\sqrt{2}\cos\beta}{\pi\cos\varphi}$$

$$2\tan\frac{\omega}{2} = \sqrt{m^2 + n^2 - 2},$$

where

$\varepsilon$ is defined by $\varepsilon = \theta - \frac{\pi}{2}$, and $\theta$ is the angle between a meridian and a parallel in the plane of projection
$m$ is a linear scale along meridian
$n$ is a linear scale along parallel
$\omega$ is a maximal angular distortion at a point.

The scale of the area $p = 1$ by definition.

The distribution of distortion of Mollweide projection has been investigated and represented in tabular and/or graphical form by several authors (Behrmann, 1909, Solov'ev, 1946, Graur, 1956, Fiala, 1957, Maling 1980).

The linear scale along parallels depends on latitude only. The linear scale along meridians depends both on latitude and longitude. The only standard parallels are 40°44'12"N and S. The only two points with no distortion are the intersections of the central meridian and standard parallels.

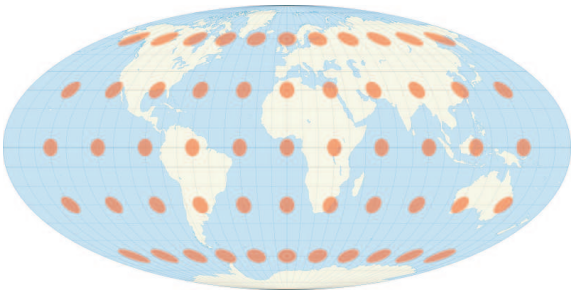


Figure 8: *The Mollweide projection with Tissot's indicatrices of deformation (URL5)*

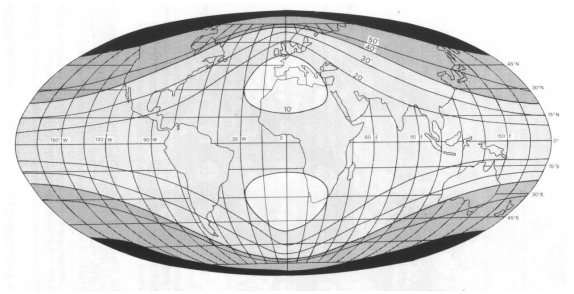


Figure 9: *Mollweide projection for the whole word, showing isograms for maximum angular deformation at* 10°, 20°, 30°, 40° *and* 50°. *Parts of the world map where* $\omega > 80°$ *are shown in black (Maling, 1980; Canters and Crols, 2011).*

## 6 Some Applications of Mollweide Projection

For those who would like to research the Mollweide projection in more detail, I would recommend the following web-sites: URL2, URL3 and URL6. Although, due it carefully, due to some incorrect statements occurring on the Internet.

Mollweide's projection has been extremely influential. Besides the developments by Goode (URL7), derived works include the interrupted Sinu-Mollweide projection by A. K. Philbrick (1953), other aspect maps like Bartholomew's *Atlantis*, and simple rescaling by reciprocal factors which preserve its features - e.g., making the equator a standard parallel free of distortion (Bromley, 1965), or making the whole map circular instead of elliptical as indicating in the Chapter 3.

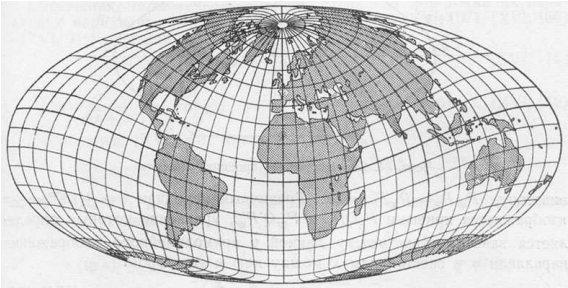


Figure 10: *Oblique aspect of the Mollweide projection (Solov'ev, 1946, Kavrajskij, 1960*

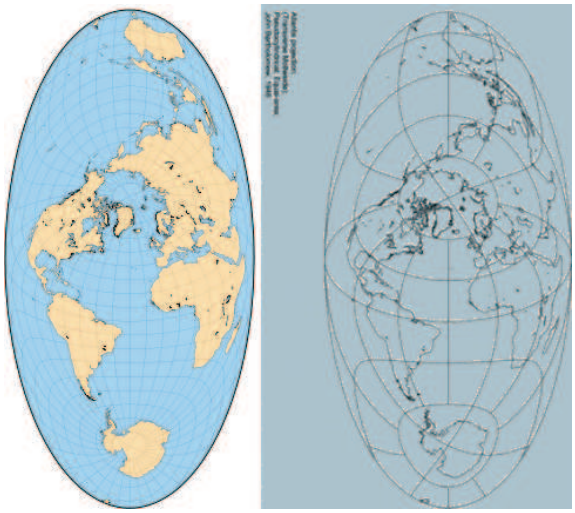


Figure 11: *The Atlantis Map (Bartholomew, 1948), Transversal aspect of the Mollweide projection (URL3)*

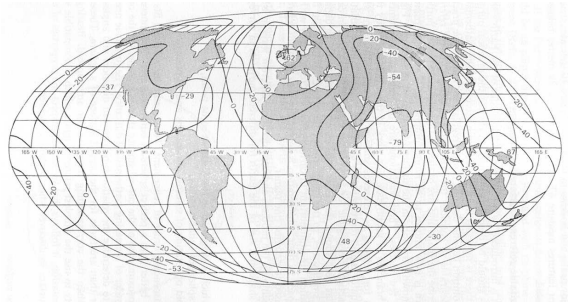


Figure 12: *Inferred contours of the geoid (in metres) for the whole word, based upon Kuala's analysis of variations in gravity potential with both latitude and longitude (Maling 1980)*

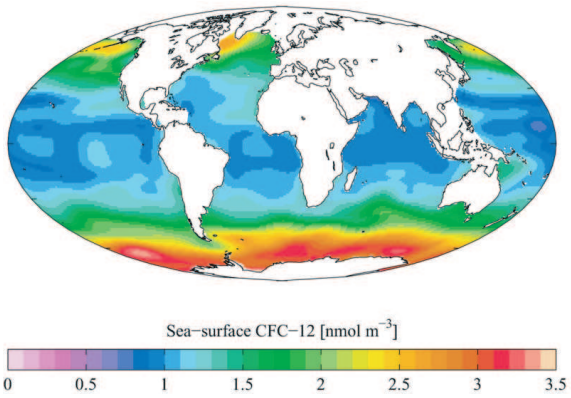



Figure 13: *Sea-surface freon levels measured by the Global Ocean Data Analysis Project. Projected using the Mollweide projection (URL5).*

Figure 14: *The Map Room - A weblog about maps (URL8)*

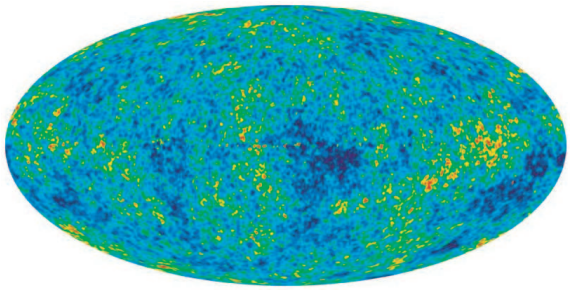


Figure 15: *Full-sky image of Cosmic Microwave Background as seen by the Wilkinson Microwave Anisotropy Probe (URL5).*

**Remark 5**

The Mollweide and Hammer projections are occasionally confused, since they are both equal-area and share the elliptical boundary; however, the latter design has curved parallels and is not pseudocylindrical (Fig. 16).

The logo of the International Cartographic Associtaion (ICA) has the world in Mollweide projection in its central part (Fig. 17). The mission of the ICA is to promote the discipline and profession of cartography in an international context.

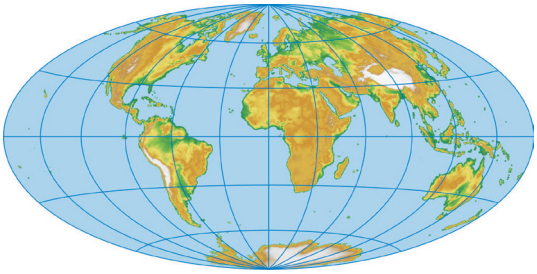


Figure 16: *Hammer projection (URL 9)*


Figure 17: *ICA logo (URL4)*

## 7 Conclusions

German mathematician and astronomer Karl Brandan Mollweide (1774-1825) is known for trigonometric formulae and map projections named after him. It is possible to derive his projection equations in different ways. One can choose the classic approach without using calculus, another using integrals or the third one, which consists of establishing and solving a differential equation.

Furthermore, it is possible to generalize the Mollweide projection in order to provide pseudocylindrical equal-area projections which represent the entire Earth in an ellipse with any prescribed ratio of its semiaxes. The original Mollweide projection has the ratio of 2:1. Inverse equations of Mollweide projection also exist.

The paper also provides the formulae and illustrations of the distortion distribution in the Mollweide projection. Considering several applications of Mollweide projections represented in the paper, it is obvious that even though the map projection is more than 200 years old, it still has numerous applications. For example, the International Cartographic Association has used it in its logo since it was founded 50 years ago.

## References

[1] BEHRMANN, W. (1909): Zur Kritik der flächentreuen Projektionen der ganzen Erde und einer Halbkugel. Sitzungsberichte der Mathematisch-Physikalischen Klasse der Königlich-Bayerischen Akademie der Wissenschaften zu München, Bayerische Akademie der Wissenschaften München.

[2] BOGGS, S. W. (1929): A New Equal-Area Projection for World Maps, The Geographical Journal 73 (3), 241-245.

[3] BORČIĆ, B. (1955): Matematička kartografija, Tehnička knjiga, Zagreb.

[4] BROMLEY, R. H. (1965): Mollweide Modified, The Professional Geographer, 17 (3), 24.

[5] CANTERS, F., CROLST, T. (2011): Low-error Equal-Area Map Projections for Global Mapping if the Terrestrial Envirnoment, CO-302, ICC, Paris.

[6] CLOSE, C. (1929): An Oblique Mollweide Projection of the Sphere, The Geographical Journal 73 (3), 251-253.

[7] DEKLEINE, H. A. (1988): Proof without Words: Mollweide's Equation, Mathematics Magazine 61 (5) (1988), 281.

[8] FEEMAN, F. G. (2000): Equal Area World Maps: A Case Study, SIAM Review 42 (1), 109-114.

[9] FIALA, F. (1957): Mathematische Kartographie, Veb Verlag Technik, Berlin.

[10] GRAUR, A. V. (1956): Matematicheskaja kartografija, Izdatel'stvo Leningradskogo Universiteta.

[11] KAVRAJSKIJ, V. V. (1960): Izbrannye trudy, Tom II, Vypusk 3, Izdanie Upravlenija nachal'nika Gidrograficheskoj sluzhby VMF.

[12] KUNTZ, E. (1990): Kartennetzentwurfslehre, 2. Auflage, Wichmann, Karlsruhe.

[13] MALING, D. H. (1980): Coordinate Systems and Map Projections, Georg Philip and Son Limited, London.

[14] MOLLWEIDE, K. B. (1805): Ueber die vom Prof. Schmidt in Giessen in der zweyten Abtheilung seines Handbuchs der Naturlehre S. 595 angegeben Projection der Halbkugelfläche, Zach's Monatliche Correspondenz zur Beförderung der Erd- und Himmels-Kunde, vol. 12, Aug., 152-163.

[15] MOLLWEIDE'S PAPER (1808): Zusätze zur ebenen und sphärischen Trigonometrie.

[16] NELSEN, R. B. (1993): Proofs without Words, Exercises in Visual Thinking, Classroom Resource Materials, No. 1, The Mathematical Association of America, p. 36.

[17] PHILBRICK, A. K. (1953): An Oblique Equal Area Map for World Distributions, Annals of the Association of American Geographers 43 (3), 201-215.

[18] REEVES, E. A. (1904): Van der Grinten's Projection, The Geographical Journal 24 (6), 670-672.

[19] SCHMIDT, G. G. (1801-3): Projection der Halbkugelfläche, in his Handbuch der Naturlehre, Giessen.

[20] SNYDER, J. P. (1977): A Comparison of Pseudocylindrical Map Projections, The American Cartographer, Vol. 4, No. 1, 59-81.

[21] SNYDER, J. P. (1987): Map Projections - A Working Manual, U.S. Geological Survey Professional Paper 1395, Washington.

[22] SNYDER, J. P. (1993): Flattening the Earth, The University of Chicago Press, Chicago and London.

[23] SOLOV'EV, M. D. (1946): Kartograficheskie proekcii, Geodezizdat, Moscow.

[24] WAGNER, K. (1949): Kartographischenetzentwürfe, Bibliographisches Institut Leipzig.

[25] URL1: MacTutor History of Mathematics, O'Connor, J. J., Robertson, E. F.: Karl Brandan Mollweide
http://www-history.mcs.st-andrews.ac.uk/Biographies/Mollweide.html

[26] URL2: Carlos A. Furuti
http://www.progonos.com/furuti

[27] URL3: Map Projections, Instituto de matematica, Universidade Federal Fluminense, Rio de Janeiro, Rogrio Vaz de Almeida Jr, Jonas Hurrelmann, Konrad Polthier and Humberto Jos Bortolossi
http://www.uff.br/mapprojections/mp_en.html

[28] URL4: International Cartographic Associtaion – ICA Logo download and design guidelines
http://icaci.org/logo

[29] URL5: Mollweide projection on wikipedia
http://en.wikipedia.org/wiki/Mollweide_projection

[30] URL6: Equal-area maps
http://www.equal-area-maps.com/

[31] URL7: Goode_homolosine_projection
http://en.wikipedia.org/wiki/Goode_homolosine_projection

[32] URL8: The Map Room – A weblog about maps
http://www.maproomblog.com/

[33] URL9: Hammer projection
http://en.wikipedia.org/wiki/Hammer_projection

**Miljenko Lapaine**
Faculty of Geodesy University of Zagreb
Kačićeva 26, 10000 Zagreb, Croatia
e-mail: mlapaine@geof.hr

Original scientific paper
Accepted 21. 12. 2011.

MÁRTA SZILVÁSI-NAGY*
SZILVIA BÉLA**

# B-spline Patches Fitting on Surfaces and Triangular Meshes

**B-spline Patches Fitting on Surfaces and Triangular Meshes**

**ABSTRACT**

In this paper a technique for the construction of quartic polynomial B-spline patches fitting on analytical surfaces and triangle meshes is presented. The input data are curvature values and principal directions at a given surface point which can be computed directly, if the surface is represented by a vector function.

In the case of discrete surface representation, i.e. on a triangle mesh the required input data are computed from a circular neighborhood of a specified triangle facet. Such a surface patch may replace a well defined region of the mesh, and can be used e.g. in re-triangulation, mesh-simplification and rendering algorithms.

**Key words:** B-spline surface, local surface approximation, principal curvatures, triangle mesh

**MSC 2010:** 65D17, 65D05, 65D07, 68U05, 68U07

**B-splajn dijelovi koji pristaju na plohe i triangularne mreže**

**SAŽETAK**

U ovom se radu prikazuje metoda za konstrukciju kvartnog polinoma B-splajn dijela podesnog za analitičke plohe i mreže trokuta. Ulazni podaci su vrijednosti zakrivljenosti i glavni smjerovi u danoj točki plohe, koji se mogu izravno računati za plohu zadanu vektorskom funkcijom.

Za slučj diskretne reprezentacije plohe, tj. za triangularnu mrežu, odgovarajući ulazni podaci računaju se iz kružne okoline određenog trokuta mreže. Takvi dijelovi mogu zamijeniti dobro definirano područje mreže, i mogu se upotrijebiti npr. u retriangulaciji, simplifikaciji mreže i renderiranju.

**Ključne riječi:** B-splajn ploha, localna aproksimacija plohe, glavne zakrivljenosti, mreža trokuta

## 1 Introduction

Surface patches matching free form surfaces, triangular meshes or point-based surfaces are widely used in computer graphics and in many applications. Different types of patches have been developed to reconstruct the surface geometry.

A technique in [16] which generates a hole-free, piecewise linear approximation to point-based surfaces uses circular and elliptical planar surface segments, so-called splats, for surface reconstruction and high-quality rendering. A circular splat is given by its center, its normal vector and its radius. Elliptical splats need two additional vectors to define the major and minor axes replacing the radius, they can adapt to the local curvature of the surface. Large number of linear splats are needed to represent the shape of most smooth models.

Quadrics are defined in [4] for mesh simplification algorithms which produce an approximation composed of fewer triangles that preserves surface shape. The quadrics characterize the local shape of the surface, they are elongated in directions of low curvature and thin in directions of high curvature. Minimization of a quadric error metric generates a triangulation with optimal triangle shape. In [2] quadratic and cubic splats are computed using a moving least squares procedure. They provide good quality and high rendering speed using fewer primitives than linear splats. In [5] a rendering primitive, called differential point is introduced with embedded curvature information in the vicinity of the actual point. The method leads to a more sparsely surface representation, to accelerated shading, to a point-based simplification technique, and to a better quality of rendering than a pure splat-based approach. The inputs are NURBS surface or polygonal mesh. A differential point is constructed from a sample point and principal curvatures and principal directions. Practically, a local surface is defined implicitly in the neighborhood of the point by a

* Supported by a joint project between the TU Berlin and the BUTE.
** Supported by the grant TÁMOP - 4.2.2.B-10/1–2010-0009.

set of osculating circle arcs in normal planes passing through given tangent vectors such that the distance of the circle arc from the surface or mesh is less than a given tolerance. The necessary curvature values for a mesh are estimated by the method of Taubin ([15]).

A method for fitting NURBS surfaces for cloud-of-points data representing rotational surfaces is shown in [1]. First, a scalar valued B-spline function is fitted to the data, then it is converted to a parametric NURBS. Three dimensional object matching is the tool for fitting NURBS surfaces to point based surfaces or to an other NURBS in [7]. Two intrinsic surface properties, the Gaussian and the mean curvatures are used for matching, and an optimal rigid body transformation is developed. A $C^2$-continuous spline surface is constructed to triangular meshes in [6]. The construction is made in two phases. First, a so called guide surface is constructed from vertices and boundary data, then it is modified such that the final surface has a good shape also in the case of triangulation with isolated extraordinary vertices.

Trigonometric surface patches are constructed from curvature data matching a neighborhood of a face of a triangular mesh in [14] and a neighborhood on an analytical surface [13]. The principal directions and curvature values of the meshed surface are estimated by the method developed in [11] and [12]. This is a face based method, different from the vertex based algorithms used in the most papers in the large literature dealing with discrete differential geometry ([8], [9], [10]).

In this paper the construction of a uniform polynomial B-spline surface patch of $4\times 4$ degree from given curvature values is presented. These curvatures are the principal curvatures of a base surface at a given point. They determine its osculating circles in the two normal planes through the principal directions. The construction is made in two phases. First, the surface interpolates the two circle arcs by its middle parameter curves, and approximates four additional surface points at its corner points. Then the approximation is improved by correcting some boundary data while minimizing an error between the B-spline patch and the given base surface.

In Section 2 the approximation of a circular arc by a fourth degree B-spline curve is analyzed. In Section 3 the computation of a B-spline surface patch of $4\times 4$ degree is presented from input data, which allow to fit the patch on a base surface. In Section 4 the generation of the input data from a base surface is shown in both, analytical and discrete representations. Examples are shown in Section 5.

## 2 Approximation of a circular arc by a fourth degree B-spline curve

The uniform polynomial B-spline curve is represented by the vector function

$$\mathbf{g}(t)=\left[t^4\ t^3\ t^2\ t\ 1\right]\mathbf{M}^4\left[\mathbf{p}_0\ \mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3\ \mathbf{p}_4\right]^T,\ 0\le t\le 1, \quad (1)$$

where the coefficient matrix is

$$\mathbf{M}^4=\frac{1}{24}\begin{bmatrix}1&-4&6&-4&1\\-4&12&-12&4&0\\6&-6&-6&6&0\\-4&-12&12&4&0\\1&11&11&1&0\end{bmatrix}. \quad (2)$$

Let the circular arc of radius $\rho$ and central angle $2\alpha$ be given in the $xz$ coordinate plane parametrized as follows

$$\mathbf{c}(t)=\mathbf{i}\,\rho\sin(\alpha(2t-1))+\mathbf{k}\,\rho\cos(\alpha(2t-1)),\ t\in[0,1]. \quad (3)$$

Three points $\mathbf{c}(0)$, $\mathbf{c}(0.5)$, $\mathbf{c}(1)$ and two tangent vectors at the end points $\dot{\mathbf{c}}(0)$ and $\dot{\mathbf{c}}(1)$ will be interpolated by solving the system of linear equations

$$\mathbf{c}(0)=\mathbf{g}(0),\ \mathbf{c}(0.5)=\mathbf{g}(0.5),\ \mathbf{c}(1)=\mathbf{g}(1),$$
$$\dot{\mathbf{c}}(0)=\dot{\mathbf{g}}(0),\ \dot{\mathbf{c}}(1)=\dot{\mathbf{g}}(1)$$

for the unknown control points $\mathbf{p}_i$, $i=0\ldots4$, where $\dot{\mathbf{c}}$ and $\dot{\mathbf{g}}$ denote the derivatives according to the parameter $t$. A unique (symbolical) solution exists, and the B-spline curve with the computed control points approximates the given circular arc with an error

$$\int_0^1(\mathbf{c}(t)-\mathbf{g}(t))^2dt<10^{-18},\ 2\alpha\le\frac{\pi}{3}. \quad (4)$$

In the examples the relative error with respect to the arc length is even smaller. The limit for the central angle is a usual limit also in classical approximations.

## 3 Computation of a B-spline surface patch of $4\times 4$ degree from given geometric data (symbolical solution)

The polynomial B-spline surface patch of $4\times4$ degree with uniform knot vector is described by the vector function

$$\mathbf{r}(u,v)=\left[u^4\ u^3\ u^2\ u\ 1\right]\cdot\mathbf{M}\cdot\mathbf{B}\cdot\mathbf{M}^T\cdot\left[v^4\ v^3\ v^2\ v\ 1\right]^T, \quad (5)$$
$$(u,v)\in[0,1]\times[0,1].$$

The geometric data are the points of the control net:
$\mathbf{B}=[\mathbf{b}[i,j]]\ ,\quad i=0,\ldots4,\quad j=0,\ldots4.$
The prescribed input data in our surface construction are the data of two circular arcs lying in two orthogonal planes, which will be interpolated by the middle parameter curves.

Four additional boundary data are four corner points of the required patch. For the parameter curve $\mathbf{r}(u,0.5)$ three points of a circular arc

$\mathbf{r}(0,0.5) = M11$, $\mathbf{r}(0.5,0.5) = M$, $\mathbf{r}(1,0.5) = M12$

and two tangent vectors at the starting and end points $M11$ and $M12$,

$\mathbf{r}_u(0,0.5) = T11$, $\mathbf{r}_u(1,0.5) = T12$

respectively, are given. Similarly, the other parameter curve $\mathbf{r}(0.5,v)$ is determined by three points of a circular arc

$\mathbf{r}(0.5,0) = M21$, $(\mathbf{r}(0.5,0.5) = M)$, $\mathbf{r}(0.5,1) = M22$

and by two prescribed tangent vectors at the points $M21$ and $M22$

$\mathbf{r}_v(0.5,0) = T21$, $\mathbf{r}_v(0.5,1) = T22$.

The parametrization of these arcs is the same as that of the curve $\mathbf{c}(t)$ in (3). $\mathbf{r}_u$ and $\mathbf{r}_v$ denote the partial derivatives of the function $\mathbf{r}(u,v)$ according to $u$ and $v$, respectively. The position vectors of the points $Mij$ are denoted by the same letter, $Tij$ denote the corresponding tangent vectors.
The four corner points of the patch denoted by $P00$, $P10$, $P01$ and $P11$ according to their parameter values are also prescribed (Fig. 1). These are in all 13 input data.

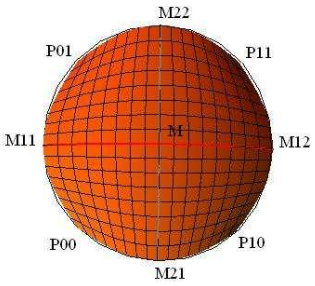



Figure 1: *The middle parameter curves of the surface patch pass through $M11$, $M$, $M12$ and $M21$, $M$, $M22$, respectively, the corner points are $P00$, $P10$, $P01$ and $P11$.*

The B-spline patch has 25 unknown control points therefore, 12 additional data are necessary for the computation. These are generated from the previous data in the following way.

$$\begin{aligned}
\mathbf{r}_u(0,0) &= -\mathbf{r}_v(0,0) = \frac{1}{4}(M21 - M11)\\
\mathbf{r}_u(1,0) &= \mathbf{r}_v(1,0) = \frac{1}{4}(M12 - M21)\\
\mathbf{r}_u(1,1) &= -\mathbf{r}_v(1,1) = \frac{1}{4}(M12 - M22)\\
\mathbf{r}_u(0,1) &= \mathbf{r}_v(0,1) = \frac{1}{4}(M22 - M11)
\end{aligned} \tag{6}$$

Four twist vectors at the corner points are determined by the change of the first partial derivatives while moving from a corner point into the corresponding midpoint along a boundary curve.

$$\begin{aligned}
\mathbf{r}_{uv}(0,0) &= (\mathbf{r}_v(0.5,0) - \mathbf{r}_v(0,0) + \mathbf{r}_u(0,0.5) - \mathbf{r}_u(0,0)) \cdot \lambda\\
\mathbf{r}_{uv}(1,0) &= (-\mathbf{r}_v(0.5,0) + \mathbf{r}_v(1,0) + \mathbf{r}_u(1,0.5) - \mathbf{r}_u(1,0)) \cdot \lambda\\
\mathbf{r}_{uv}(1,1) &= (-\mathbf{r}_v(0.5,1) + \mathbf{r}_v(1,1) - \mathbf{r}_u(1,0.5) + \mathbf{r}_u(1,1)) \cdot \lambda\\
\mathbf{r}_{uv}(0,1) &= (\mathbf{r}_v(0.5,1) - \mathbf{r}_v(0,1) - \mathbf{r}_u(0,0.5) + \mathbf{r}_u(0,1)) \cdot \lambda
\end{aligned} \tag{7}$$

Here the first partial derivatives on the right hand sides are expressed by the prescribed points $Mij$ according to the equations (6). The free parameter $\lambda$ changes the lengths of the twist vectors which strongly influence the shape of the surface patch. It will be determined by minimizing an error function. Putting all these conditions into a system of equations does not result in a solution for the unknown control points of the required patch however, all the equations are linear ones. Instead of this, the computation is organized according to the following strategy.
The matrix of the control points $\mathbf{B} = [\mathbf{b}[i,j]]$, $i = 0,\ldots 4$, $j = 0,\ldots 4$ will be partitioned, and the control points will be computed in three phases. Figure 2 shows which control points are computed in one step by marking them with the same symbol.

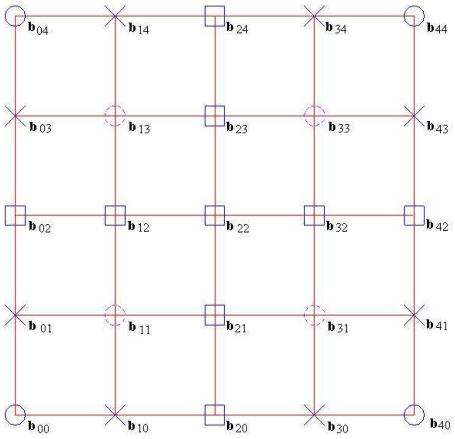



Figure 2: *Partition of the matrix of the control points.*

In the **first step** control points on the boundaries are computed from the data of four boundary curves as the solution of four systems of linear equations by the interpolation method of circular arcs in Section 2. Each boundary curve is determined by three points and two tangent vectors at the end points. Each system of linear equations result in five control points. The data and the corresponding solutions are as follows, while the boundary curves of the patch follow in counter clockwise direction.
For the boundary curve $v = 0$ the data are $P00$, $M21$, $P10$, $\mathbf{r}_u(0,0)$, $\mathbf{r}_u(1,0)$. The solutions are $\mathbf{p}_i^{u0}$, $i = 0,\ldots 4$.

For the boundary curve $u=1$ the data are $P10, M12, P11$, $\mathbf{r}_v(1,0), \mathbf{r}_v(1,1)$. The solutions are $\mathbf{p}_i^{1v}, i=0,\dots 4$.
For the boundary curve $v=1$ the data are $P01, M22, P11$, $\mathbf{r}_u(0,1), \mathbf{r}_u(1,1)$. The solutions are $\mathbf{p}_i^{u1}, i=0,\dots 4$.
For the boundary curve $u=0$ the data are $P00, M11, P01$, $\mathbf{r}_v(0,0), \mathbf{r}_v(0,1)$. The solutions are $\mathbf{p}_i^{0v}, i=0,\dots 4$.
Not all these solutions will be placed into the matrix of the control points $\mathbf{B}$.
In the **second step** the system of linear equations from the interpolation conditions for the middle parameter curves is solved.
The data are
$\mathbf{r}(0.5,0)=M21$, $\mathbf{r}(0.5,0.5)=M$, $\mathbf{r}(0.5,1)=M22$,
$\mathbf{r}_v(0.5,0)=T21$, $\mathbf{r}_v(0.5,1)=T22$,
$\mathbf{r}(0,0.5)=M11$, $\mathbf{r}(1,0.5)=M12$,
$\mathbf{r}_u(0,0.5)=T11$, $\mathbf{r}_u(1,0.5)=T12$
The unique symbolical solution of this system are nine control points $\mathbf{b}[i,2], i=0\dots 4$ and $\mathbf{b}[2,j], j=0,1,3,4$ expressed with these data and the remaining sixteen control points.
Now, twelve from the sixteen control points will be replaced by the points of the solution in the first step as follows.

$$\mathbf{b}[1,0]=\mathbf{p}_1^{u0},\ \mathbf{b}[3,0]=\mathbf{p}_3^{u0},\ \mathbf{b}[4,1]=\mathbf{p}_1^{1v},\ \mathbf{b}[4,3]=\mathbf{p}_3^{1v}$$
$$\mathbf{b}[1,4]=\mathbf{p}_1^{u1},\ \mathbf{b}[3,4]=\mathbf{p}_3^{u1},\ \mathbf{b}[0,1]=\mathbf{p}_1^{0v},\ \mathbf{b}[0,3]=\mathbf{p}_3^{0v}$$
$$\mathbf{b}[0,0]=\frac{1}{2}(\mathbf{p}_0^{u0}+\mathbf{p}_0^{0v}),\ \mathbf{b}[4,0]=\frac{1}{2}(\mathbf{p}_4^{u0}+\mathbf{p}_0^{1v}),$$
$$\mathbf{b}[4,4]=\frac{1}{2}(\mathbf{p}_4^{u1}+\mathbf{p}_4^{1v}),\ \mathbf{b}[0,4]=\frac{1}{2}(\mathbf{p}_4^{0v}+\mathbf{p}_0^{u1}).$$

We note that the interpolation conditions for the corner points are not satisfied, but "relaxed" by the last four equations. They will be corrected by minimizing an error function in the last step.
In the **third step** the last four control points are computed from the system of four linear equations expressed by the twist vectors at the corner points which are computed earlier from the prescribed data in (7). The solution of the system of the linear equations are the control points

$\mathbf{b}[1,1], \mathbf{b}[3,1], \mathbf{b}[3,3], \mathbf{b}[1,3]$.

Finally, the matrix $\mathbf{B}=[\mathbf{b}[i,j]]\ ,\quad i=0,\dots 4,\quad j=0,\dots 4$ is expressed by the prescribed data and the free scalar parameter $\lambda$.
In the **fourth step** the free parameter $\lambda$ is determined from an error function expressing the squared sum of distances between the prescribed and computed corner points.

$$\begin{aligned} d(\lambda) =&(\mathbf{r}(0,0)-P00)^2+(\mathbf{r}(1,0)-P10)^2 \\ &+(\mathbf{r}(1,1)-P11)^2+(\mathbf{r}(0,1)-P01)^2 \end{aligned} \tag{8}$$

This function is quadratic in $\lambda$, the minimization results in a unique value of it.

In **Example 1** two circular arcs of radius 2 and central angle $2\alpha=\pi/3$ are given in the $xz$ and $yz$ coordinate plane, respectively. The four corner points are rotated end points of these circular arcs around the axis $z$. The computed B-spline patch interpolates the given arcs within the integrated error (see (4)) of $10^{-8}$, the interpolation error at the corner points (see (8)) is within $10^{-16}$. We note that the interpolation error of the patch to the data of the given circular arcs is practically zero (less than $10^{-28}$), and it is independent from the value of $\lambda$. The length of the twist vectors determined by $\lambda$ effects on the error at the corner points very strongly. The computed value of $\lambda$ by minimizing the distance function in (8) is 0.89, while the value $\lambda=0$ results in a very large error of 1.2.

In **Example 2** the data are similar but they determine a hyperbolic surface patch. The analysis shows similar results to that in Example 1.

The Figures 3 and 4 show the middle parameter curves $\mathbf{r}(u,0.5)$ and $\mathbf{r}(0.5,v)$ of the resulting patches on the left hand side, and the patches with the corresponding control points on the right. The whole control net cannot be shown clearly, therefore they are omitted in the figures.

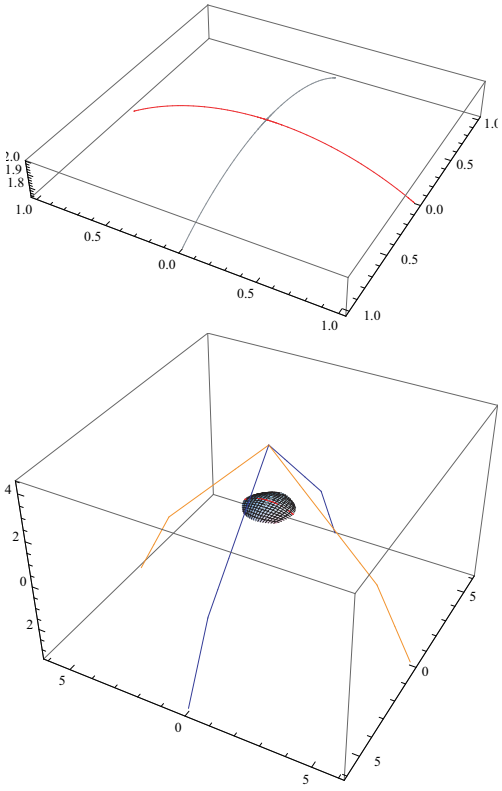



Figure 3: *Example 1.*

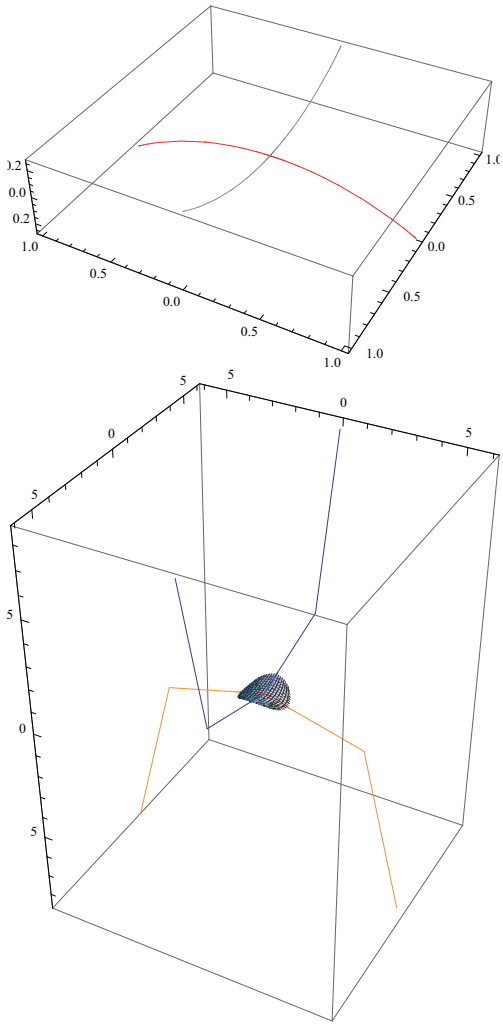


Figure 4: *Example 2*.

## 4 Solution of the fitting problem; computing the input data of the B-spline patch

A B-spline patch of $4 \times 4$ degree fitting on a base surface will be determined by two osculating circle arcs of the base surface lying in two normal planes through the principal directions at the actual surface point, and four additional surface points in a neighborhood of this point. The neighborhood is determined by the arc length of the osculating circle arcs given by the user. Then by measuring the given arc length on the surface from the point in normal planes rotating around the surface normal a circular neighborhood is constructed. In the case of a triangular mesh the neighborhood is constructed around the barycentric center of a specified face in the mesh.

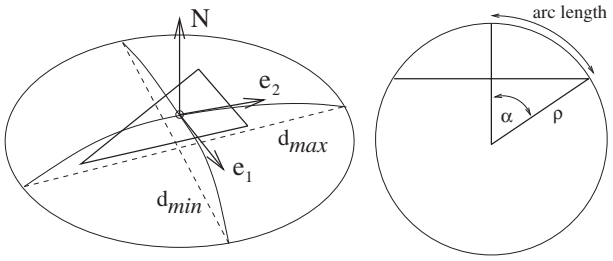



Figure 5: *Circular neighborhood and an osculating circle in a normal plane*.

In Figure 5 such a circular neighborhood is shown with the two osculating circle arcs of the shortest and largest chord lengths denoted by $d_{min}$ and $d_{max}$, respectively. They determine at a regular surface point the orthogonal principal directions $\mathbf{e}_1$ and $\mathbf{e}_2$, if $d_{min} \neq d_{max}$. The central angle $2\alpha$ and the radius $\rho$ of an osculating circle arc in a normal plane is computed from the given arc length $s$ and the corresponding chord length $d$ in the following way ([11], [12]).
From $\rho\alpha = s$ and $\rho \sin\alpha = \frac{d}{2}$ and the approximation

$$\sin\alpha \approx \alpha - \frac{\alpha^3}{6}, 0 < \alpha << 1$$

follows that

$$\alpha \approx \sqrt{(1 - \frac{d}{2s})6}, \rho \approx \frac{s}{\alpha} \text{ if } \alpha \neq 0, \ \kappa_n \approx \frac{\alpha}{s}.$$

The tangent vectors at the end points of the arc are computed with the parametrization in the local coordinate system determined by the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{N}$ according to the vector equation (3) in Section 2.
In the case of an analytical surface the principal curvatures and principal directions are computed directly from the known equation of the surface ([3]), though the construction of the circular neighborhood (more precisely the points of its boundary) is computed by a discrete method, while measuring the given arc length along a polygonal line which approximates the surface curve lying in an intersecting normal plane.
From the constructed circular neighborhood five interpolation points and four tangent vectors to the circular arcs (see the second step in the computation of the control points in Section 3) and four surface points on the boundary (determining the corner points of the patch) are used as direct input data in the matrix $\mathbf{B}$ of the control points. The control net is expressed symbolically by these input data and a scalar parameter $\lambda$. After replacing the actual numerical values into the matrix $\mathbf{B}$, only the computation of $\lambda$ is necessary in the concrete examples (see the fourth step in Section 3).
Figure 6 shows the data of a B-spline patch matching a circular neighborhood on the base surface. The interpolation

points on the patch boundary $Mij$ and $Pij$ $(i = 0,1,\ j = 0,1)$ are end points of the surface curves in the corresponding normal planes computed as polygonal lines on the surface by measuring the given arc length along them. The tangent vectors $Tij_u$ and $Tij_v$ $(i = 0,1,\ j = 0,1)$ at the corner points are determined as described in the equations (6) in Section 3.

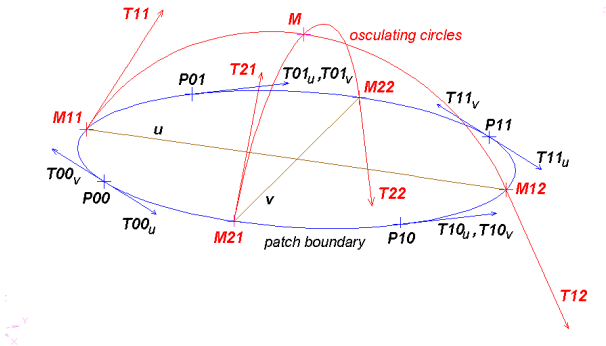



Figure 6: *Input data of a patch computed in a circular neighborhood of a base surface.*

## 5 Examples

The next two examples are computed with analytical base surfaces. In this case the principal directions are computed exactly from the vector functions representing the surfaces. The number of the computed points on the boundary of the circular patch is 72 in both examples.

In **Example 3** (Fig. 7) the base surface is a cylinder of radius 10, the circular neighborhood around a surface point is constructed with the arc length 5. The generated B-spline surface interpolates the data points of the two middle parameter curves $\mathbf{r}(u,0.5)$ and $\mathbf{r}(0.5,v)$ practically with zero error (less than $10^{-27}$). The minimized error of the sum of squared distances between the prescribed and computed corner points of the patch (see in (4)) is within the relative error of 4% with respect to the given arc length measured on the surface around the given point.

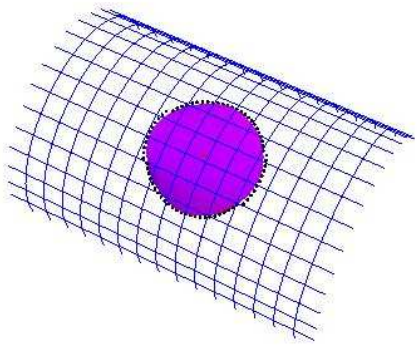


Figure 7: *Example 3.*

The results are similar in **Example 4** (Fig. 8) computed with a torus. The radius of its meridian circle is 10, the "radius" of the circular patch is 5, and the interpolation error at the corner points is here 4% too. Of course, the error is larger with growing neighborhoods to be approximated.

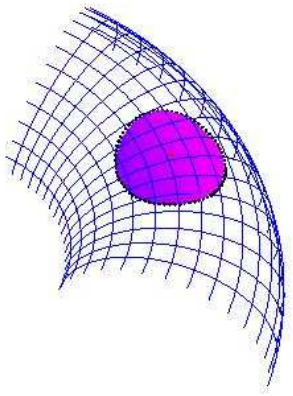


Figure 8: *Example 4.*

The B-spline patches in the next two examples are computed on triangulated surfaces, on so-called "synthetic meshes" generated by a triangulation of the parameter domain of the given surfaces. The principal curvatures and directions are estimated by the method described in Section 4. The circular neighborhood is computed around a chosen triangle face, more exactly around its barycentric center point in 48 normal planes.

In **Example 5** the mesh is a triangulated cylinder of radius 1 (Fig. 9). The circular neighborhood is constructed with the given arc length of 0.6. The estimation error in the computation of the principal curvatures is less than $10^{-3}$. The constructed B-spline patch interpolates the data given in the principal normal sections for the middle parameter curves practically with zero error. The error at the corner points is approximately $2 \cdot 10^{-2}$.

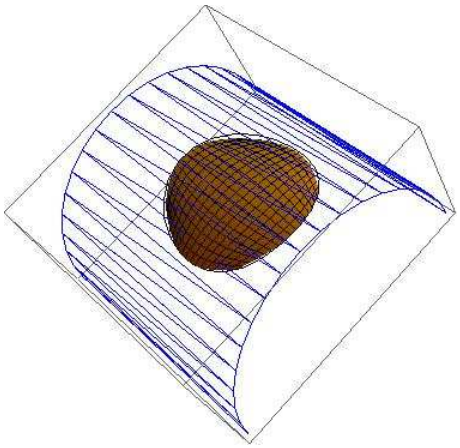


Figure 9: *Example 5.*

In the case of a synthetic mesh of a torus shown in **Example 6**, the results are better due to the dense triangulation. The error at the corner points of the patch is approximately $10^{-2}$.

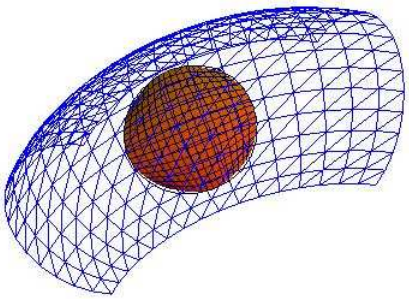


Figure 10: *Example 6*.

We note that the presented construction of a circular neighborhood is working also on a "bad triangulation" shown on the cylinder, where the long, thin triangles have no vertices in the actual neighborhood. This is due to the face based estimation of normal curvatures and to an appropriate polyhedral data structure representing the mesh, which provides effective computation of intersections with planes.

**Example 7** shows a "real" mesh of a sphere, i.e. a triangulation generated from measured points on the surface of a sphere offered for test purposes. The computed B-spline patch matches very well a relatively large neighborhood shown in Figure 11.

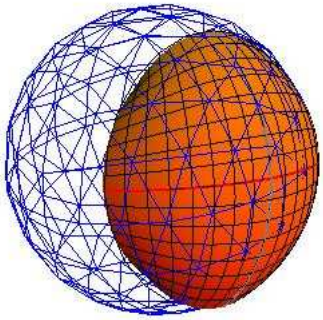


Figure 11: *Example 7*.

## 6 Conclusions

Construction of a uniform polynomial B-spline surface patch of $4 \times 4$ degree from geometric input data has been presented, which are suitable for the solution of a fitting problem, how a given neighborhood of a point on a given surface can be approximated by such a B-spline patch. The well chosen geometric data define the control points of the B-spline patch, and they can be computed on analytic and meshed surfaces by constructing a circular neighborhood around a specified point. The measurement of this neighborhood is a user specified arc length, which is measured on the surface around the surface point in normal sections resulting from a discrete description of the boundary of the neighborhood. In the examples the approximation of specified neighborhoods by the generated B-spline patches has been shown with error estimation. The figures, the symbolic computation of the control points of the B-spline patch and patch fitting with error estimation have been made with the algebraic program package Wolfram Mathematica. The curvature estimation and construction of the circular neighborhoods on triangular meshes have been computed with a program developed by the first author in the program language Java.
Further research is necessary for improving the approximation around the boundaries of the computed surface patches and for estimating the measurement of matched neighborhoods with prescribed error tolerances.

## References

[1] S -H. Bae, B. K. Choi, NURBS surface fitting using orthogonal coordinate transform for rapid product development. *Computer-Aided Design* 34 (2002) 683-690.

[2] N. K. Bolla, P. J. Narayanan, Algebraic splats representation for point based models. *IEEE Sixth Indian Conference on Computer Vision, Graphics & Image Processing* 2008, 71-78.

[3] M. P. Do Carmo, 1976, *Differential Geometry of Curves and Surfaces*. Prentice-Hall, Englewood Cliffs, NJ.

[4] P. S. Heckbert, M. Garland, Optimal triangulation on quadric-based surface simplification. *Computational Geometry* 14 (1999) 49-65.

[5] A. Kalaiah, A. Varshney, Modeling and rendering of points with local geometry. *IEEE Transactions on Visualization and Computer Graphics* 9 (2003) 100-129.

[6] K. Karciauskas, J. Peters, Assembling curvature continuous surfaces from triangular patches. *Computers & Graphics* 33 (2009) 204=210.

[7] K. H. Ko, T. Maekawa, N. M. Patrikalakis, An algorithm for optimal free-form object matching. *Computer-Aided Design* 35 (2003) 913-923.

[8] M. MEYER, M. DESBRUN, P. SCHRÖDER, A. H. BARR, Discrete differential-geometry operators for triangulated 2-manifolds, in *Visualization and Mathematics III*.(H.C. Hege and K. Polthier, Ed.) pp. 35-57, Springer Verlag 2003.

[9] K. POLTHIER, M. SCHMIES, Straightest geodesics on polyhedral surfaces, in *Mathematical visualization* (K. Polthier and H.C. Hege Ed.) pp.391-409, Springer Verlag 1998.

[10] S. RUSINKIEWICZ, Estimating curvatures and their derivatives on triangle meshes, *Proc. 3DPVT'04* 2004, 486-493.

[11] M. SZILVASI-NAGY, About curvatures on triangle meshes. *KoG* 10 (2006) 13-18.

[12] M. SZILVASI-NAGY, Face-based estimations of curvatures on triangle meshes, *Journal for Geometry and Graphics* 12(2008), No. 1, 63-73.

[13] M. SZILVÁSI-NAGY, Construction of circular splats on analytic surfaces. In *Fifth Hungarian Conference on Computer Graphics and Geometry*, Budapest 2010, 55-57.

[14] M. SZILVÁSI-NAGY, Surface patches constructed from curvature data. *KoG* 14 (2010) 29-34.

[15] G. TAUBIN, Estimating the tensor of curvature of a surface from a polyhedral approximation, in *ICCV'95 Proceedings of the Fifth International Conference on Computer Vision, IEEE Computer Society*, Washington DC, USA, 902.

[16] J. WU, L. KOBBELT, Optimized sub-sampling of point sets for surface splatting. *In: Eurographics 2004 (Eds. M.-P. Cani, M. Slater)* 23 (2004), 643-652.

**Márta Szilvási-Nagy**
e-mail: szilvasi@math.bme.hu
**Szilvia Béla**
email: belus@math.bme.hu

Dept. of Geometry
Budapest University of Technology and Economics
H-1521 Budapest, Hungary

Original scientific paper
Accepted 19. 12. 2011.

NORMAN JOHN WILDBERGER

# Universal Hyperbolic Geometry III: First Steps in Projective Triangle Geometry

**Universal Hyperbolic Geometry III: First Steps in Projective Triangle Geometry**

**ABSTRACT**

We initiate a triangle geometry in the projective metrical setting, based on the purely algebraic approach of universal geometry, and yielding in particular a new form of hyperbolic triangle geometry. There are three main strands: the Orthocenter, Incenter and Circumcenter hierarchies, with the last two dual. Formulas using ortholinear coordinates are a main objective. Prominent are five particular points, the $b$, $z$, $x$, $h$ and $s$ points, all lying on the Orthoaxis $A$. A rich kaleidoscopic aspect colours the subject.

**Key words:** universal hyperbolic geometry, triangle geometry, projective geometry, bilinear form, ortholinear coordinates, incenter, circumcenter, orthoaxis

**MSC 2010:** 51M10, 14N99, 51E99

**Univerzalna hiperbolička geometrija III: Prvi koraci u projektivnoj geometriji trokuta**

**SAŽETAK**

Na temelju algebarskog pristupa univerzalne geometrije, uvodimo geometriju trokuta u projektivno-metrički okvir. To rezultira jednim novim oblikom hiperboličke geometrije trokuta. Tri su glavne okosnice: hijerarhije ortocentara, središta upisanih i središta opisanih kružnica, od kojih su posljednje dvije dualne. Primjena ortolinearnih koordinata u formulama ima bitnu ulogu. Istaknuto je pet posebnih točaka ($b$, $z$, $x$, $h$ i $s$) koje leže na ortogonalnoj osi $A$. Bogato, kaleidoskopsko gledište karakterizira obradu teme.

**Ključne riječi:** univerzalna hiperbolička geometrije, geometrija trokuta, projektivna geometrija, bilinearna forma, ortolinearne koordinate, središte upisane kružnice, središte opisane kružnice, ortogonalna os

## 1 Introduction

Recently there has been a revival of interest in classical geometry and in particular the study of triangles ([6], [7], [9], [10], [12], [13], [14]). This paper introduces triangle geometry into the framework of *Universal Hyperbolic Geometry (UHG)* ([18], [19]) and beyond; in the context of a general metrical structure on the projective plane. The basic measurements of *quadrance* and *spread* replace the usual notions of *distance* and *angle*, and these depend on a general bilinear form. Hyperbolic geometry provides the motivation and is used for the illustrations. The approach is purely algebraic and works over any field not of characteristic two; the reader may easily keep the fundamental example of the *rational number field* foremost in mind. Ultimately this theory is a natural consequence of Rational Trigonometry ([15], [16], [17]).

Triangle geometry in this setting has features that resemble and also contrast with classical hyperbolic geometry, studied and described in [1], [2], [3], [4], [5], [11] and [21]. The *Orthocenter hierarchy*, involving *Altitudes, Orthic triangles,* the *Orthic axis*, the *Double triangle,* and the *Orthoaxis*, on which the important $s, h, x, b$ and $z$ *points* are to be found, is primary. The *Incenter* and *Circumcenter hierarchies* are precisely dual, and their existences depend on number theoretic conditions, unlike the usual Euclidean situation. The former contains the *Incenters*, *Bilines* (analogs of *vertex* or *angle bisectors*), *Bipoints, Apollonius points*, *Centrian lines*, *Sight lines, Contact points, Gergonne points* and *Nagel points* etc. The latter contains *Circumlines, Midpoints, Midlines (*analogs of *perpendicular bisectors*), *Medians, Centroids*, *Sound points*, *Tangent lines*, *Jay lines* and *Wren lines* etc. *Duality* pervades the subject; interchanging points and lines, sides and vertices, and quadrance and spread.

This paper is largely self-contained; we start with a general introduction to universal metrical projective geometry. When we study a triangle $\overline{a_1a_2a_3}$, it will prove convenient to use a linear transformation to change coordinates, so that we may assume that $a_1 = [1:0:0]$, $a_2 = [0:1:0]$ and $a_3 = [0:0:1]$, with the *orthocenter* represented by $h = [1:1:1]$. With these *ortholinear coordinates* the bilinear form is given by a pair of inverse symmetric projective matrices:

$$\mathbf{B} = \begin{bmatrix} a & 1 & 1 \\ 1 & b & 1 \\ 1 & 1 & c \end{bmatrix}, \quad \mathbf{A} = \mathbf{B}^{-1} = \begin{bmatrix} 1-bc & c-1 & b-1 \\ c-1 & 1-ac & a-1 \\ b-1 & a-1 & 1-ab \end{bmatrix}. \qquad (1)$$

*This shifts projective triangle geometry from the study of a general triangle under a particular bilinear form to the study of a particular triangle under a general bilinear form,* giving a simpler and more general theory.

Formulas will be our main aims; most of these depend on the three parameters $a, b, c$ occurring in (1), and hopefully will provide a solid platform for further investigations. They also suggest a possible alternative to trilinear coordinates in affine/Euclidean triangle geometry. This paper introduces a rich theory which has many additional relationships and remarkable aspects which will be further studied in the coming years.

### 1.1 Projective linear algebra and Universal geometry

In this section we introduce the main objects: *(projective) points* and *lines*, via projective linear algebra. This is linear algebra with vectors and matrices defined only up to non-zero scalar multiples. We write the usual vectors and matrices with round brackets, while projective vectors and projective matrices, in square brackets, are by definition unchanged if we multiply all coordinates simultaneously by a non-zero number. So while $\overrightarrow{v} \equiv (3,1,2) \equiv \begin{pmatrix} 3 & 1 & 2 \end{pmatrix}$ represents a usual row vector (or $1 \times 3$ matrix), the corresponding projective row vector is $a = \begin{bmatrix} 3 & 1 & 2 \end{bmatrix}$. By definition $a$ is also equal to $\begin{bmatrix} -3 & -1 & -2 \end{bmatrix}$ or to $\begin{bmatrix} 6 & 2 & 4 \end{bmatrix}$.

We will generally use bold labels to represent projective matrices: while

$$A = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & -1 & -11 \\ 0 & 2 & -2 \\ 0 & 0 & 6 \end{pmatrix}$$

denote ordinary matrices, the corresponding *projective matrices are*

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 12 \\ 0 & 9 & 3 \\ 0 & 0 & 3 \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 3 & -1 & -11 \\ 0 & 2 & -2 \\ 0 & 0 & 6 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 3 & -1 & -11 \\ 0 & 2 & -2 \\ 0 & 0 & 6 \end{bmatrix}.$$

Inverses are easier to compute in the projective setting, since determinants in the denominator can be dispensed with: for example $\mathbf{A}^{-1} = \mathbf{B}$, so that integer arithmetic only is required. While in general projective matrices cannot be added, they can be multiplied!

We now introduce additional notation and terminology that allows us to work consistently with both row and column vectors horizontally. A *non-zero* projective row vector $a$ will be written in either of the following forms:

$$a \equiv \begin{bmatrix} x & y & z \end{bmatrix} \equiv [x : y : z]$$

and will be called a **(projective) point**. A *non-zero* projective column vector $L$ will be written as

$$L \equiv \begin{bmatrix} l \\ m \\ n \end{bmatrix} \equiv \langle l : m : n \rangle$$

and will be called a **(projective) line**. The point $a \equiv [x : y : z]$ and the line $L \equiv \langle l : m : n \rangle$ are **incident** precisely when $lx + my + nz = 0$; equivalently $a$ **lies on** $L$, or $L$ **passes through** $a$. The corresponding matrix equation is

$$aL \equiv \begin{bmatrix} x & y & z \end{bmatrix} \begin{bmatrix} l \\ m \\ n \end{bmatrix} = [x : y : z] \langle l : m : n \rangle = 0. \tag{2}$$

Three or more points are **collinear** precisely when they all lie on a line $L$, and three or more lines are **concurrent** precisely when they all pass through a point $a$.

The **join** $a_1 a_2$ of distinct points $a_1 \equiv [x_1 : y_1 : z_1]$ and $a_2 \equiv [x_2 : y_2 : z_2]$ is the line

$$\begin{aligned} a_1 a_2 &\equiv [x_1 : y_1 : z_1] \times [x_2 : y_2 : z_2] \\ &\equiv \langle y_1 z_2 - y_2 z_1 : z_1 x_2 - z_2 x_1 : x_1 y_2 - x_2 y_1 \rangle. \end{aligned}$$

The **meet** $L_1 L_2$ of distinct lines $L_1 \equiv \langle l_1 : m_1 : n_1 \rangle$ and $L_2 \equiv \langle l_2 : m_2 : n_2 \rangle$ is the point

$$\begin{aligned} L_1 L_2 &\equiv \langle l_1 : m_1 : n_1 \rangle \times \langle l_2 : m_2 : n_2 \rangle \\ &\equiv [m_1 n_2 - m_2 n_1 : n_1 l_2 - n_2 l_1 : l_1 m_2 - l_2 m_1]. \end{aligned}$$

These operations, using the usual Euclidean cross product, are well-defined, and will be used repeatedly in this paper. The symbol $\times$ in the linear algebra context avoids confusion with matrix multiplication.

Then $a_1 a_2$ *is the unique line incident with both* $a_1$ *and* $a_2$, *and* $L_1 L_2$ *is the unique point incident with both* $L_1$ *and* $L_2$. A complete symmetry or *duality* between points and lines is a key feature of this subject.

We also recall a few more definitions from [18] and [19]. A **side** $\overline{a_1 a_2} \equiv \{a_1, a_2\}$ is a set of two points. A **vertex** $\overline{L_1 L_2} \equiv \{L_1, L_2\}$ is a set of two lines. A **triangle** $\overline{a_1 a_2 a_3} \equiv \{a_1, a_2, a_3\}$ is a set of three non-collinear points, and a **trilateral** $\overline{L_1 L_2 L_3} \equiv \{L_1, L_2, L_3\}$ is a set of three non-concurrent lines.

A triangle $\overline{a_1 a_2 a_3}$ determines an **associated trilateral** $\overline{L_1 L_2 L_3}$, where $L_1 \equiv a_2 a_3$, $L_2 \equiv a_1 a_3$ and $L_3 \equiv a_1 a_2$. Symmetrically a trilateral $\overline{L_1 L_2 L_3}$ determines an associated triangle $\overline{a_1 a_2 a_3}$, where $a_1 \equiv L_2 L_3$, $a_2 \equiv L_1 L_3$ and $a_3 \equiv L_1 L_2$. The triangle $\overline{a_1 a_2 a_3}$ has three sides, namely $\overline{a_1 a_2}$, $\overline{a_2 a_3}$ and $\overline{a_1 a_3}$, as well as three vertices, namely $\overline{L_1 L_2}$, $\overline{L_2 L_3}$ and $\overline{L_1 L_3}$. In this paper we concentrate on triangles.

### 1.2 Projective bilinear forms

We now introduce a metrical structure on our three-dimensional vector space; this will be done via a symmetric bilinear form $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 \equiv \overrightarrow{v}_1 A \overrightarrow{v}_2^T$ given by an invertible symmetric $3 \times 3$ matrix $A$, where $\overrightarrow{v}_1$ and $\overrightarrow{v}_2$ are ordinary row vectors, and $T$ denotes transpose. We wish to transfer this bilinear form to projective points and lines: let's start with perpendicularity. Recall that vectors $\overrightarrow{v}_1, \overrightarrow{v}_2$ are perpendicular precisely when $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = 0$.

Denote by $\mathbf{A}$ and $\mathbf{B}$ the projective matrices associated to $A$ and its inverse matrix $B$ respectively. Points $a_1$ and $a_2$ are **perpendicular** precisely when $a_1 \mathbf{A} a_2^T = 0$, and in this case we write $a_1 \perp a_2$. This is a symmetric relation. Dually, lines $L_1$ and $L_2$ are **perpendicular** precisely when $L_1^T \mathbf{B} L_2 = 0$; we write $L_1 \perp L_2$. It is useful to restate these relations by introducing a formal notion of duality: the projective point $a$ and the projective line $L$ are **dual** precisely when

$$L = a^\perp \equiv \mathbf{A} a^T \quad \text{or equivalently} \quad a = L^\perp \equiv L^T \mathbf{B}.$$

So two points, or two lines, are perpendicular precisely when *one is incident with the dual of the other*. It now follows that $a_1 \perp a_2$ precisely when $a_1^\perp \perp a_2^\perp$, since the latter condition is

$$\begin{aligned} 0 &= \left(\mathbf{A} a_1^T\right)^T \mathbf{B} \left(\mathbf{A} a_2^T\right) = \left(a_1 \mathbf{A}^T\right) \mathbf{B} \left(\mathbf{A} a_2^T\right) \\ &= a_1 \left(\mathbf{A}\mathbf{B}\right) \left(\mathbf{A} a_2^T\right) = a_1 \mathbf{A} a_2^T. \end{aligned} \tag{3}$$

A point $a$ is **null** precisely when it is perpendicular to itself, that is, when $a \mathbf{A} a^T = 0$. Dually a line $L$ is **null** precisely when it is perpendicular to itself, that is, when $L^T \mathbf{B} L = 0$.

Our main interests are *hyperbolic* and *elliptic geometries*, which arise respectively from the special cases

$$\mathbf{A} = \mathbf{J} \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = \mathbf{B}, \quad \mathbf{A} = \mathbf{I} \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{B}. \tag{4}$$

But other possibilities are also of interest, and *for triangle geometry also important*, as we shall soon see.

### 1.3 Visualization

The Figures in this paper all come from hyperbolic geometry: we represent the point $a \equiv [x : y : z]$ by the affine point $[X, Y] \equiv [x/z, y/z]$, and the line $L \equiv \langle l : m : n \rangle$ by the linear equation $lX + mY + n = 0$, which would be the hyperbolic line $(l : m : -n)$ in [18]. Null points are those $a$ for which $x^2 + y^2 - z^2 = 0$; the corresponding affine points lie on the **null circle** $X^2 + Y^2 = 1$, always in blue. Null lines are tangent to this null circle. The duality becomes exactly the *projective polarity* between points and lines associated with the null circle.

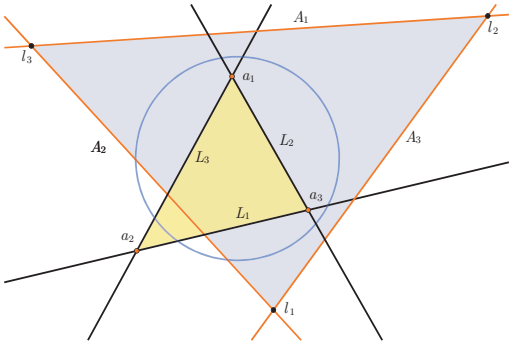



Figure 1: *A Triangle $\overline{a_1 a_2 a_3}$ and its Dual triangle $\overline{l_1 l_2 l_3}$*

We will adopt the general convention that triangle geometry constructs associated to a particular triangle are Capitalized (a familiar idea for German readers). So Figure 1 shows a Triangle $\overline{a_1 a_2 a_3}$, in yellow, with the notation we will consistently use: the **Points** of the triangle are $a_1, a_2, a_3$, the **Lines** are $L_1 \equiv a_2 a_3$, $L_2 \equiv a_1 a_3$, $L_3 \equiv a_1 a_2$, the **Dual points** are $l_1 \equiv L_1^\perp$, $l_2 \equiv L_2^\perp$, $l_3 \equiv L_3^\perp$, the **Dual lines** are $A_1 \equiv a_1^\perp$, $A_2 \equiv a_2^\perp$, $A_3 \equiv a_3^\perp$, and the **Dual triangle** is $\overline{l_1 l_2 l_3}$, in light blue. Points and their dual lines are generally pictured with the same colour.

### 1.4 Quadrance and spread

An inverse pair of symmetric projective matrices $\mathbf{A}$ and $\mathbf{B}$ give us more than perpendicularity: they allow the *introduction of metrical quantities into algebraic geometry*. This has been a blind spot in the history of the subject!

The **quadrance** $q(a_1, a_2)$ between points $a_1$ and $a_2$, and the **spread** $S(L_1, L_2)$ between lines $L_1$ and $L_2$, are the respective numbers

$$\begin{aligned} q(a_1, a_2) &\equiv 1 - \frac{\left(a_1 \mathbf{A} a_2^T\right)^2}{\left(a_1 \mathbf{A} a_1^T\right)\left(a_2 \mathbf{A} a_2^T\right)} \quad \text{and} \\ S(L_1, L_2) &\equiv 1 - \frac{\left(L_1^T \mathbf{B} L_2\right)^2}{\left(L_1^T \mathbf{B} L_1\right)\left(L_2^T \mathbf{B} L_2\right)}. \end{aligned} \tag{5}$$

While the numerators and denominators of these expressions depend on choices of representative vectors and matrices for $a_1, a_2, \mathbf{A}, L_1, L_2$ and $\mathbf{B}$, the *quotients are independent of scaling*, so the overall expressions are indeed well-defined projectively.

Clearly $q(a, a) = 0$ and $S(L, L) = 0$, while $q(a_1, a_2) = 1$ precisely when $a_1 \perp a_2$, and dually $S(L_1, L_2) = 1$ precisely when $L_1 \perp L_2$. An argument similar to (3) shows that for points $a_1$ and $a_2$,

$$S\left(a_1^\perp, a_2^\perp\right) = q(a_1, a_2). \tag{6}$$

Quadrance and spread are undefined if one or both of the points or lines involved is null. We will adopt the zero denominator convention: *statements involving a fraction with zero in the denominator are empty*, and a variant: *statements involving a proportion with all entries zero are empty*.

**Example 1** *In the hyperbolic case, the quadrance between* $a_1 \equiv [x_1 : y_1 : z_1]$ *and* $a_2 \equiv [x_2 : y_2 : z_2]$ *is*

$$q(a_1,a_2) \equiv 1 - \frac{(x_1x_2 + y_1y_2 - z_1z_2)^2}{(x_1^2 + y_1^2 - z_1^2)(x_2^2 + y_2^2 - z_2^2)}$$
$$= -\frac{(y_1z_2 - y_2z_1)^2 + (z_1x_2 - z_2x_1)^2 - (x_1y_2 - y_1x_2)^2}{(x_1^2 + y_1^2 - z_1^2)(x_2^2 + y_2^2 - z_2^2)} \quad (7)$$

*and the spread between* $L_1 \equiv \langle l_1 : m_1 : n_1 \rangle$ *and* $L_2 \equiv \langle l_2 : m_2 : n_2 \rangle$ *is*

$$S(L_1,L_2) \equiv 1 - \frac{(l_1l_2 + m_1m_2 - n_1n_2)^2}{(l_1^2 + m_1^2 - n_1^2)(l_2^2 + m_2^2 - n_2^2)}$$
$$= -\frac{(m_1n_2 - m_2n_1)^2 + (n_1l_2 - n_2l_1)^2 - (l_1m_2 - l_2m_1)^2}{(l_1^2 + m_1^2 - n_1^2)(l_2^2 + m_2^2 - n_2^2)}. \quad (8)$$

**Example 2** *In the elliptic case, the quadrance between* $a_1 \equiv [x_1 : y_1 : z_1]$ *and* $a_2 \equiv [x_2 : y_2 : z_2]$ *is*

$$q(a_1,a_2) \equiv 1 - \frac{(x_1x_2 + y_1y_2 + z_1z_2)^2}{(x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2)}$$
$$= \frac{(y_1z_2 - y_2z_1)^2 + (z_1x_2 - z_2x_1)^2 + (x_1y_2 - y_1x_2)^2}{(x_1^2 + y_1^2 + z_1^2)(x_2^2 + y_2^2 + z_2^2)} \quad (9)$$

*and the spread between* $L_1 \equiv \langle l_1 : m_1 : n_1 \rangle$ *and* $L_2 \equiv \langle l_2 : m_2 : n_2 \rangle$ *is*

$$S(L_1,L_2) \equiv 1 - \frac{(l_1l_2 + m_1m_2 + n_1n_2)^2}{(l_1^2 + m_1^2 + n_1^2)(l_2^2 + m_2^2 + n_2^2)}$$
$$= \frac{(m_1n_2 - m_2n_1)^2 + (n_1l_2 - n_2l_1)^2 + (l_1m_2 - l_2m_1)^2}{(l_1^2 + m_1^2 + n_1^2)(l_2^2 + m_2^2 + n_2^2)}. \quad (10)$$

**Theorem 1 (Null quadrance/spread)** *If* $a_1$ *and* $a_2$ *are distinct points, then* $q(a_1,a_2) = 0$ *precisely when* $a_1a_2$ *is a null line. If* $L_1$ *and* $L_2$ *are distinct lines, then* $S(L_1,L_2) = 0$ *precisely when* $L_1L_2$ *is a null point.*

**Proof.** We prove the first statement, the second follows by duality. Suppose that $A$ is a $3 \times 3$ invertible symmetric matrix with $B$ the adjugate matrix (the inverse of $A$ up to a scalar), so we may write

$$A \equiv \begin{pmatrix} a & b & c \\ b & d & f \\ c & f & g \end{pmatrix}, \quad B \equiv \begin{pmatrix} dg - f^2 & cf - bg & bf - cd \\ cf - bg & ag - c^2 & bc - af \\ bf - cd & bc - af & ad - b^2 \end{pmatrix}.$$

Since $L \equiv a_1a_2$ is a null line precisely when $L^TBL = 0$, the theorem is a consequence of the following remarkable identity in the various variables, involving only vectors and the usual linear algebra:

$$\Big((x_1,y_1,z_1)A(x_1,y_1,z_1)^T\Big)\Big((x_2,y_2,z_2)A(x_2,y_2,z_2)^T\Big) -$$
$$-\Big((x_1,y_1,z_1)A(x_2,y_2,z_2)^T\Big)^2$$
$$= (y_1z_2 - y_2z_1, z_1x_2 - z_2x_1, x_1y_2 - x_2y_1) \cdot B \cdot$$
$$\cdot (y_1z_2 - y_2z_1, z_1x_2 - z_2x_1, x_1y_2 - x_2y_1)^T. \qquad \square$$

In the paper [17] we show that this general projective metrical geometry *obeys exactly the same main trigonometric laws as those of Universal Hyperbolic Geometry* as set out in the paper [18], independent of the quadratic form. In particular *the laws of trigonometry for hyperbolic and elliptic geometries, which are both projective theories, are exactly identical*. This is indeed Universal Geometry.

### 1.5 Linear transformations and the Fundamental theorem of projective geometry

A bilinear form $\overrightarrow{v}_1 \cdot \overrightarrow{v}_2 = \overrightarrow{v}_1 A \overrightarrow{v}_2^T$ is transformed when we change coordinates. Suppose we have an invertible linear transformation $T(\overrightarrow{v}) \equiv \overrightarrow{v}M = \overrightarrow{w}$ on three-dimensional space, acting on row vectors via right multiplication by an invertible $3 \times 3$ matrix $M$, with inverse matrix $N$, so that $\overrightarrow{w}N = \overrightarrow{v}$. Define a new bilinear form $\odot$ by

$$\overrightarrow{w}_1 \odot \overrightarrow{w}_2 \equiv (\overrightarrow{w}_1N) \cdot (\overrightarrow{w}_2N) = (\overrightarrow{w}_1N)A(\overrightarrow{w}_2N)^T$$
$$= \overrightarrow{w}_1 \left(NAN^T\right) \overrightarrow{w}_2^T.$$

So the matrix $A$ for the original bilinear form $\cdot$ becomes the matrix $NAN^T$ for the new bilinear form $\odot$.

The linear transformation $T$ acting on row vectors induces a projective transformation $\mathbf{T}$ on one-dimensional subspaces, which are essentially (projective) points, as well as two-dimensional subspaces, which are essentially (projective) lines. Let $\mathbf{M}$ and $\mathbf{N}$ be the projective matrices associated to $M$ and $N$. On points, we define $\mathbf{T}(a) = a\mathbf{M}$. To see how $\mathbf{T}$ acts on lines, we use duality; the point $a$ is incident with the line $L$ precisely when $aL = 0$, which is precisely when $(a\mathbf{M})(\mathbf{N}L) = 0$, so we require that $\mathbf{T}(L) \equiv \mathbf{N}L$. In this way incidence is preserved when we apply a linear transformation to both points and lines.

The notion of perpendicularity is also modified: the points $a_1$ and $a_2$ are $\odot$**-perpendicular** precisely when $a_1\mathbf{N}$ and $a_2\mathbf{N}$ are perpendicular, in other words precisely when

$a_1\left(\mathbf{NAN}^T\right)a_2^T = 0$, while the lines $L_1$ and $L_2$ are **$\odot$-perpendicular** precisely when $L_1^T\left(\mathbf{M}^T\mathbf{BM}\right)L_2 = 0$. The inverse pair of symmetric projective matrices

$$\widetilde{\mathbf{A}} = \mathbf{NAN}^T \quad \text{and} \quad \widetilde{\mathbf{B}} = \mathbf{M}^T\mathbf{BM}$$

determine new notions of duality: $a^\perp = \widetilde{\mathbf{A}}a^T$ and $L^\perp = L^T\widetilde{\mathbf{B}}$ as well as new quadrances and spreads:

$$\widetilde{q}(a_1,a_2) \equiv 1 - \frac{\left(a_1\widetilde{\mathbf{A}}a_2^T\right)^2}{\left(a_1\widetilde{\mathbf{A}}a_1^T\right)\left(a_2\widetilde{\mathbf{A}}a_2^T\right)} \quad \text{and}$$

$$\widetilde{S}(L_1,L_2) \equiv 1 - \frac{\left(L_1^T\widetilde{\mathbf{B}}L_2\right)^2}{\left(L_1^T\widetilde{\mathbf{B}}L_1\right)\left(L_2^T\widetilde{\mathbf{B}}L_2\right)}. \tag{11}$$

Recall that the *Fundamental theorem of projective geometry* in this setting is really basic linear algebra: a general linear transformation of three-dimensional space maps any three linearly independent vectors $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3$ to any other three vectors. If in addition we are given a fourth vector $\overrightarrow{v}_4 = \lambda_1\overrightarrow{v}_1 + \lambda_2\overrightarrow{v}_2 + \lambda_3\overrightarrow{v}_3$ with none of $\lambda_1,\lambda_2,\lambda_3$ zero, then we can send $\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3$ respectively to $(1/\lambda_1,0,0)$, $(0,1/\lambda_2,0)$, $(0,0,1/\lambda_3)$, so that $\overrightarrow{v}_4$ is sent to $(1,1,1)$. When we view this projectively, we have essentially a proof of the Fundamental theorem: we can construct a projective linear transformation that sends four generic projective points $a_1,a_2,a_3$ and $a_4$ (no three collinear) respectively to $[1:0:0]$, $[0:1:0]$, $[0:0:1]$ and $[1:1:1]$.

### 1.6 An example with the basic Triangle

We illustrate these abstractions in a concrete example. Our basic Triangle shown in Figure 2 comes from the hyperbolic plane where the points originally have the *approximate values*:

$$a_1 \approx [-0.4 : 0.4 : 1], a_2 \approx [-0.7 : -0.4 : 1], a_3 \approx [0.1 : 0.1 : 1]$$

corresponding to the affine points $A_1 \approx [-0.4, 0.4]$, $A_2 \approx [-0.7, -0.4]$, $A_3 \approx [0.1, 0.1]$. The following calculations are subject to round-off and approximation.

The Orthocenter, using formulas for hyperbolic geometry altitudes, is $h \approx [-0.286886 : 0.217349 : 1]$. Now

$$(x,y,z)\begin{pmatrix} -0.4 & 0.4 & 1 \\ -0.7 & -0.4 & 1 \\ 0.1 & 0.1 & 1 \end{pmatrix} = (-0.2869, 0.2173, 1)$$

has the solution $(x,y,z) \approx (0.586371, 0.117125, 0.296503)$. We conclude that the transformation $T(v) = vN$ where

$$N \equiv \begin{pmatrix} 0.586371 & 0 & 0 \\ 0 & 0.117125 & 0 \\ 0 & 0 & 0.296503 \end{pmatrix}\begin{pmatrix} -0.4 & 0.4 & 1 \\ -0.7 & -0.4 & 1 \\ 0.1 & 0.1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} -0.2345484 & 0.2345484 & 0.586371 \\ -0.0819875 & -0.04685 & 0.117125 \\ 0.0296503 & 0.0296503 & 0.296503 \end{pmatrix}$$

sends $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ to multiples of $(-0.4,0.4,1)$, $(-0.7,-0.4,1)$, $(0.1,0.1,1)$ respectively, and also $(1,1,1)$ to $(-0.2869,0.2173,1)$.

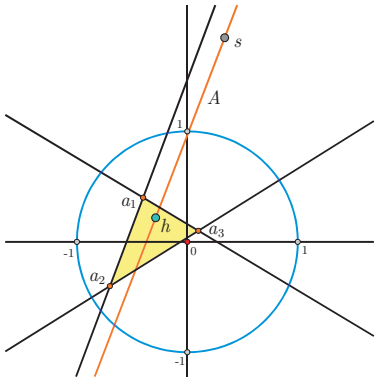



Figure 2: *Basic triangle $\overline{a_1a_2a_3}$ with Orthocenter $h$ Orthostar $s$, and Orthoaxis $A$*

The inverse projective matrix $\mathbf{N}^{-1} = \mathbf{M}$ projectively sends the points $a_1,a_2,a_3$ to $[1:0:0], [0:1:0], [0:0:1]$ and $h$ to $[1:1:1]$. Recalling the definition of $\mathbf{J}$ in (4), the bilinear form in the new standard coordinates is given (approximately) by the pair of projective inverse matrices $\mathbf{A} = \mathbf{NJN}^T$ and $\mathbf{B} = \mathbf{M}^T\mathbf{JM}$, which are

$$\mathbf{A} \approx \begin{bmatrix} -0.2338 & -0.0604 & -0.1739 \\ -0.0604 & -0.00480 & -0.0385 \\ -0.1739 & -0.0385 & -0.0862 \end{bmatrix} \quad \text{and}$$

$$\mathbf{B} \approx \begin{bmatrix} -0.7173 & 1 & 1 \\ 1 & -6.745 & 1 \\ 1 & 1 & -1.692 \end{bmatrix}$$

so that

$$a \approx -0.7173, \quad b \approx -6.745, \quad c \approx -1.692.$$

As an application, let's look at an important point associated to the Triangle $\overline{a_1a_2a_3}$ called the *Orthostar* $s = [a+2 : b+2 : c+2]$. In our example this would be the point $[1.2827 : -4.745 : 0.308]$, and to convert that back into the original projective or hyperbolic coordinates, we would multiply by $\mathbf{N}$ to get

$$[1.2827 : -4.745 : 0.308]\mathbf{N} \approx [0.0973 : 0.5322 : 0.2877] \approx$$
$$\approx [0.34 : 1.85 : 1]$$

which agrees approximately with the affine value for $s$ of $[0.34, 1.85]$ in Figure 2. In the same spirit, the *Orthoaxis* $A \equiv hs$ would have standard coordinates

$$[1:1:1]\times[a+2:b+2:c+2] = \langle c-b:a-c:b-a\rangle \approx$$
$$\approx \langle 5.053 : 0.9747 : -6.0277\rangle .$$

Since this is a line, to convert back to the original coordinates we would multiply by $M$ on the left:

$$M\begin{bmatrix} 5.053 \\ 0.9747 \\ -6.0277 \end{bmatrix} \approx \begin{bmatrix} -47.08 \\ 18.02 \\ -17.42 \end{bmatrix} \approx \begin{bmatrix} 2.702 \\ -1.03 \\ 1.0 \end{bmatrix}$$

giving the line $2.702X - 1.03Y + 1 = 0$ with projective coordinates $\langle 2.702 : -1.03 : 1\rangle$ or hyperbolic coordinates $(-2.702 : 1.03 : 1)$. The Orthoaxis $A$ appears in Figure 2 as the orange line.

### 1.7 Midpoints, midlines, bilines and bipoints

There are four more important metrical concepts that play a big role in projective triangle geometry. A side $\overline{ab}$ has a **midpoint** $m$ precisely when $m$ is a point lying on $ab$ which satisfies $q(a,m) = q(m,b)$, and it has a **midline** $M$ precisely when $M$ is a line passing through a midpoint, perpendicular to the corresponding line $ab$ of the side. Midlines are called *perpendicular bisectors* in Euclidean geometry; we prefer the more compact terminology, which emphasizes the duality between midpoints and midlines. Figure 3 shows our standard Triangle $\overline{a_1a_2a_3}$ that we will be using throughout this paper, together with its six Midpoints $m$ and six Midlines $M$.

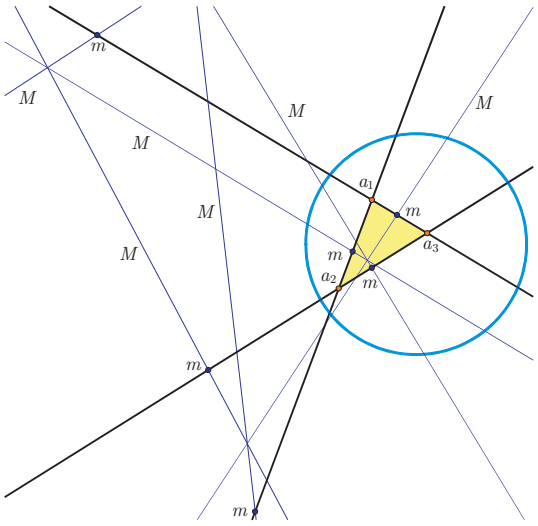



Figure 3: *Midpoints $m$ and Midlines $M$ of the Triangle $\overline{a_1a_2a_3}$*

Dually a vertex $\overline{KL}$ has a **biline** $B$ precisely when $B$ is a line passing through $KL$ which satisfies $S(K,B) = S(B,L)$, and it has a **bipoint** $b$ precisely when $b$ is a point lying on a biline, perpendicular to the corresponding point $KL$ of the vertex. Bilines are called *angle* or *vertex bisectors* in Euclidean geometry. Bipoints have no Euclidean analogs.

Figure 4 shows the six Bilines $B$ and four of the six Bipoints $b$ of our standard Triangle. Both Figures 3 and 4 have interesting collinearities and concurrences that the reader might like to observe; we will explore these later.

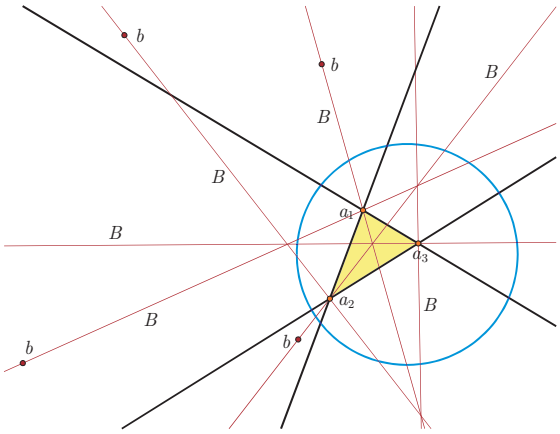



Figure 4: *Bilines B and Bipoints b*

We will see that the existence of midpoints and bilines depends on certain quadratic equations having solutions, with the consequence that sides and vertices generally have zero or two midpoints, or bilines. In a general triangle there are then several possibilities about which sides and vertices have midpoints or bilines. In future work we will explore interesting variants to these concepts which partially replace them when they do not exist.

## 2 Ortholinear coordinates

### 2.1 The Orthocenter theorem

Here is a main theorem which will be pivotal in our approach to triangle geometry in this general projective setting. There has recently been renewed interest in the Orthocenter in hyperbolic geometry ([8]); deservedly so.

**Theorem 2 (Orthocenter theorem)** *Suppose that $\overline{a_1a_2a_3}$ is a triangle which is not a right triangle, so that no two of the three lines $L_1 \equiv a_2a_3$, $L_2 \equiv a_1a_3$ and $L_3 \equiv a_1a_2$ are perpendicular. Then the **altitude lines** (or just **altitudes**) $N_1 \equiv a_1L_1^\perp$, $N_2 \equiv a_2L_2^\perp$ and $N_3 \equiv a_3L_3^\perp$ are defined and concurrent. Their common meet, the **Orthocenter** $h$, does not lie on $L_1, L_2$ or $L_3$.*

**Proof.** If $\overline{a_1a_2a_3}$ is not a right triangle, then none of the points $a_1, a_2, a_3$ are dual to the opposite lines $L_1, L_2, L_3$, so the three altitudes $N_1 \equiv a_1L_1^\perp$, $N_2 \equiv a_2L_2^\perp$ and $N_3 \equiv a_3L_3^\perp$ are well-defined. Set $h \equiv N_1N_2$, with the idea of proving

that $N_3$ is also incident with $h$. Now $h$ does not lie on any of the lines $L_1, L_2$ or $L_3$, since otherwise $\overline{a_1a_2a_3}$ would be a right triangle, contrary to our assumption. From the Fundamental theorem of projective geometry, we can apply a linear transformation to change coordinates so that

$$a_1 = [1:0:0], a_2 = [0:1:0], a_3 = [0:0:1], h = [1:1:1].$$

It follows that

$$\begin{aligned} L_1 &= a_2a_3 = [0:1:0] \times [0:0:1] = \langle 1:0:0 \rangle, \\ L_2 &= a_1a_3 = [1:0:0] \times [0:0:1] = \langle 0:1:0 \rangle, \\ L_3 &= a_1a_2 = [1:0:0] \times [0:1:0] = \langle 0:0:1 \rangle, \end{aligned} \tag{12}$$

and

$$\begin{aligned} N_1 &= a_1h = [1:0:0] \times [1:1:1] = \langle 0:1:-1 \rangle, \\ N_2 &= a_2h = [0:1:0] \times [1:1:1] = \langle 1:0:-1 \rangle. \end{aligned}$$

Suppose that the inverse projective matrix $\mathbf{B}$ for the quadratic form in these new coordinates is

$$\mathbf{B} \equiv \begin{bmatrix} a & d & e \\ d & b & f \\ e & f & c \end{bmatrix}.$$

Then since $L_1 \perp N_1$

$$\langle 1:0:0 \rangle^T \mathbf{B} \langle 0:1:-1 \rangle = [d-e] = 0$$

and since $L_2 \perp N_2$

$$\langle 0:1:0 \rangle^T \mathbf{B} \langle 1:0:-1 \rangle = [d-f] = 0.$$

From these two equations we deduce that $e = f$, so that also

$$\langle 0:0:1 \rangle^T \mathbf{B} \langle 1:-1:0 \rangle = [e-f] = 0,$$

which implies that $a_3h = \langle 1:-1:0 \rangle$ is indeed perpendicular to $L_3$. So $N_3 = a_3L_3^{\perp} = a_3h$ passes through $h$, which does not lie on $L_1, L_2$ or $L_3$. □

**Theorem 3 (Ortholinear forms)** *If $a_1 = [1:0:0]$, $a_2 = [0:1:0]$, $a_3 = [0:0:1]$ and $h = [1:1:1]$ is the orthocenter of $\overline{a_1a_2a_3}$, then either*

$$\mathbf{B} = \begin{bmatrix} a & 1 & 1 \\ 1 & b & 1 \\ 1 & 1 & c \end{bmatrix} \quad \text{or} \quad \mathbf{B} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}.$$

*The second possibility occurs precisely when $\overline{a_1a_2a_3}$ is a **fully right** triangle: any two of its lines are perpendicular.*

**Proof.** This follows from the proof of the previous theorem: the orthocenter being $h$ implies that $d = e = f$. So up to a re-scaling, the possibilities are either $d = e = f = 1$ or $d = e = f = 0$.

Let us now consider the second alternative: where

$$\mathbf{B} = \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$$

and each of $a, b, c$ is non-zero by assumption. This then yields the dual points of the triangle to be $\langle 1:0:0 \rangle^T \mathbf{B} = [1:0:0]$, and also $[0:1:0]$ and $[0:0:1]$. The dual points are then exactly the same as the original points, so this is a fully right triangle: all three points and lines are mutually perpendicular. □

To summarize, we state the following result.

**Theorem 4 (Ortholinear coordinates)** *If the triangle $\overline{a_1a_2a_3}$ is not a right triangle, then we may change coordinates so that the bilinear form is given by the pair of projective matrices*

$$\mathbf{B} = \begin{bmatrix} a & 1 & 1 \\ 1 & b & 1 \\ 1 & 1 & c \end{bmatrix}, \quad \mathbf{A} = \mathbf{B}^{-1} = \begin{bmatrix} 1-bc & c-1 & b-1 \\ c-1 & 1-ac & a-1 \\ b-1 & a-1 & 1-ab \end{bmatrix} \tag{13}$$

*which depend only on the three numbers $a, b, c$, and so that $a_1, a_2, a_3$ and the orthocenter $h$ have the forms*

$$a_1 \equiv [1:0:0], \; a_2 \equiv [0:1:0], \; a_3 \equiv [0:0:1], \; h \equiv [1:1:1].$$

We say refer to this as the **standard bilinear form**, and that $\overline{a_1a_2a_3}$ is the **standard triangle,** or just the **Triangle**. The coordinates of this framework are called **ortholinear coordinates**. *We will henceforth assume that we have made this choice of coordinates.*

The duals of the Altitudes $N_1 = \langle 0:1:-1 \rangle$, $N_2 = \langle 1:0:-1 \rangle, N_3 = \langle 1:-1:0 \rangle$ are the **Altitude points**

$$\begin{aligned} n_1 &= N_1^T \mathbf{B} = [0:b-1:1-c], \\ n_2 &= N_2^T \mathbf{B} = [a-1:0:1-c], \\ n_3 &= N_3^T \mathbf{B} = [a-1:1-b:0]. \end{aligned} \tag{14}$$

The dual of the Orthocenter $h = [1:1:1]$ is the **Ortholine**

$$H = \mathbf{A}[1:1:1]^T = \langle b+c-bc-1 : a+c-ac-1 : a+b-ab-1 \rangle.$$

**Theorem 5 (Null points/lines)** *The point $p \equiv [x:y:z]$ in Ortholinear coordinates is a null point precisely when*

$$\begin{aligned} &(1-bc)x^2 + (1-ac)y^2 + (1-ab)z^2 + \\ &+ 2(c-1)xy + 2(b-1)xz + 2(a-1)yz = 0. \end{aligned}$$

*The line $L \equiv \langle l:m:n \rangle$ is a null line precisely when*

$$al^2 + bm^2 + cn^2 + 2lm + 2ln + 2mn = 0.$$

**Proof.** These follow by using (13) to expand the respective conditions

$$[x:y:z]\,\mathbf{A}\,[x:y:z]^T = 0 \text{ and}$$

$$\langle l:m:n\rangle^T\,\mathbf{B}\,\langle l:m:n\rangle = 0. \qquad \square$$

**Corollary 1** *Using ortholinear coordinates, the Points $a_1 \equiv [1:0:0]$, $a_2 \equiv [0:1:0]$ and $a_3 \equiv [0:0:1]$ are null points precisely when $bc = 1, ac = 1$ and $ab = 1$ respectively, and the Lines $L_1 \equiv \langle 1:0:0\rangle$, $L_2 \equiv \langle 0:1:0\rangle$ and $L_3 \equiv \langle 0:0:1\rangle$ are null lines precisely when $a = 0, b = 0$ and $c = 0$ respectively.*

Define

$$D \equiv abc - a - b - c + 2. \tag{15}$$

Then it is straightforward to check that

$$\det B = \det \begin{pmatrix} a & 1 & 1 \\ 1 & b & 1 \\ 1 & 1 & c \end{pmatrix} = D \quad \text{and}$$

$$\det A = \det \begin{pmatrix} 1-bc & c-1 & b-1 \\ c-1 & 1-ac & a-1 \\ b-1 & a-1 & 1-ab \end{pmatrix} = -D^2.$$

**Theorem 6 (Triangle quadrances and spreads)** *Using Ortholinear coordinates, the quadrances $q_1 \equiv q(a_2,a_3)$, $q_2 \equiv q(a_1,a_3)$, $q_3 \equiv q(a_1,a_2)$ and spreads $S_1 \equiv S(L_2,L_3)$, $S_2 \equiv S(L_1,L_3)$, $S_3 \equiv S(L_1,L_2)$ of the standard Triangle $\overline{a_1a_2a_3}$ are*

$$q_1 = \frac{-Da}{(ab-1)(ac-1)},\ q_2 = \frac{-Db}{(ab-1)(bc-1)},$$
$$q_3 = \frac{-Dc}{(ac-1)(bc-1)}$$

*and*

$$S_1 = \frac{bc-1}{bc},\ S_2 = \frac{ac-1}{ac},\ S_3 = \frac{ab-1}{ab}.$$

*These numbers also satisfy*

$$1 - q_1 = \frac{(a-1)^2}{(ab-1)(ac-1)},$$
$$1 - q_2 = \frac{(b-1)^2}{(ab-1)(bc-1)},$$
$$1 - q_3 = \frac{(c-1)^2}{(ac-1)(bc-1)} \tag{16}$$

*and*

$$1 - S_1 = \frac{1}{bc},\ 1 - S_2 = \frac{1}{ac},\ 1 - S_3 = \frac{1}{ab}. \tag{17}$$

**Proof.** These are straightforward computations. $\square$

Although it will play only a small role in this paper, we also introduce the most important number associated to the Triangle, and a formula for it in terms of $a, b, c$.

**Theorem 7 (Triangle quadrea)** *The **quadrea** $\mathcal{A}$ of the triangle $\overline{a_1a_2a_3}$ is*

$$\mathcal{A} \equiv q_2q_3S_1 = q_1q_3S_2 = q_1q_2S_3 = \frac{D^2}{(ab-1)(ac-1)(bc-1)}.$$

**Proof.** This follows directly from the formulas of the previous theorem. $\square$

We cannot help but point out an important trigonometric formula that follows from this: the *Extended Spread law* asserts that

$$\frac{S_1}{q_1} = \frac{S_2}{q_2} = \frac{S_3}{q_3} = \frac{\mathcal{A}}{q_1q_2q_3}.$$

### 2.2 Cevians, traces, Desargues theorem and Canonical lines

Consider a variable point $p \equiv [x:y:z]$ distinct from the Points $a_1, a_2, a_3$ of the Triangle $\overline{a_1a_2a_3}$. The lines $a_1p, a_2p, a_3p$ are the **Cevian lines**, or just **Cevians,** of $p$. These are

$$a_1p = [x:y:z] \times [1:0:0] = \langle 0:z:-y\rangle,$$
$$a_2p = [x:y:z] \times [0:1:0] = \langle z:0:-x\rangle,$$
$$a_3p = [x:y:z] \times [0:0:1] = \langle y:-x:0\rangle.$$

The points $t_1 \equiv (a_1p)L_1, t_2 \equiv (a_2p)L_2, t_3 \equiv (a_3p)L_1$ are the **trace points**, or just **traces**, of $p$. These are

$$t_1 = \langle 0:z:-y\rangle \times \langle 1:0:0\rangle = [0:y:z],$$
$$t_2 = \langle z:0:-x\rangle \times \langle 0:1:0\rangle = [x:0:z],$$
$$t_3 = \langle y:-x:0\rangle \times \langle 0:0:1\rangle = [x:y:0].$$

**Theorem 8 (Desargues theorem)** *Suppose that $p \equiv [x:y:z]$ is a point that does not lie on any of the Lines of the triangle, with traces $t_1, t_2, t_3$. Then the points $g_1 \equiv (t_2t_3)L_1, g_2 \equiv (t_1t_3)L_2, g_3 \equiv (t_1t_2)L_3$ are collinear, and their join is the line $S(p) \equiv \langle yz : xz : xy\rangle$.*

**Proof.** Using the formulas above for the traces, we compute

$$\begin{aligned} g_1 \equiv (t_2t_3)L_1 &= \langle -yz:xz:xy\rangle \times \langle 1:0:0\rangle \\ &= [0:xy:-xz] = [0:y:-z], \\ g_2 \equiv (t_1t_3)L_2 &= \langle yz:-xz:xy\rangle \times \langle 0:1:0\rangle \\ &= [xy:0:-yz] = [x:0:-z], \\ g_3 \equiv (t_1t_2)L_3 &= \langle yz:xz:-xy\rangle \times \langle 0:0:1\rangle \\ &= [xz:-yz:0] = [x:-y:0]. \end{aligned}$$

We have used the fact that $x,y,z$ are all non-zero, by assumption, to cancel these common factors as they occur. The points $g_1,g_2,g_3$ are collinear since

$$\det\begin{pmatrix} 0 & y & -z \\ x & 0 & -z \\ x & -y & 0 \end{pmatrix} = 0$$

and their join is

$$[0:y:-z]\times[x:0:-z] = \langle yz:xz:xy\rangle \equiv S(p). \qquad \square$$

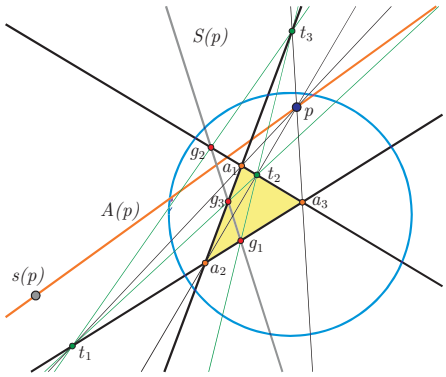



Figure 5: *Cevians, traces and the lines $S(p)$ and $A(p)$*

Associated to $p$ is the dual of the line $S(p)$:

$$\begin{aligned} s(p) \equiv S(p)^{\perp} &= \langle yz:xz:xy\rangle^T \mathbf{B} \\ &= [xy+xz+ayz : xy+yz+bxz : xz+yz+cxy]. \end{aligned}$$

Furthermore, the join of $p$ and $s(p)$

$$\begin{aligned} A(p) \equiv ps(p) = \langle & y^2z - yz^2 + cxy^2 - bxz^2 : \\ & xz^2 - x^2z + ayz^2 - cyx^2 : \\ & x^2y - xy^2 + bzx^2 - azy^2 \rangle \end{aligned}$$

is the **canonical line** of the generic point $p$. This is an interesting and important construction that is not available in Euclidean geometry, and it has many applications. In the special case when $p = h$, the Orthocenter of $\overline{a_1a_2a_3}$, the canonical line $A \equiv A(h)$ will be called the *Orthoaxis* of the triangle, and will be seen to be the most important line in triangle geometry.

There is also a dual formulation: consider a line $M$ distinct from the Lines $L_1,L_2,L_3$. The points $L_1M, L_2M, L_3M$ are the **Menelaus points** of $M$. If $M \equiv \langle l:m:n\rangle$ then the Menelaus points are

$$L_1M = [0:n:-m],\ L_2M = [n:0:-l],\ L_3M = [m:-l:0].$$

The lines $T_1 \equiv (L_1M)\,a_1, T_2 \equiv (L_2M)\,a_2, T_3 \equiv (L_3M)\,a_3$ are the **trace lines** of $M$, these are

$$T_1 = \langle 0:m:n\rangle,\ T_2 = \langle l:0:n\rangle,\ T_3 = \langle l:m:0\rangle.$$

**Theorem 9 (Desargues dual theorem)** *Suppose that $M = \langle l:m:n\rangle$ is a line that does not pass through any of the Points of the triangle, with trace lines $T_1,T_2,T_3$. Then the lines $(T_2T_3)\,a_1, (T_1T_3)\,a_2, (T_1T_2)\,a_3$ are concurrent, and they pass through the point $[mn:ln:lm]$.*

**Proof.** This is dual to the previous theorem. $\square$

Note that the transforms implicit in both these theorems are of the form $x:y:z \to x^{-1}:y^{-1}:z^{-1}$ which makes it clear that they are inverses of each other.

### 2.3 Existence of midpoints and bilines

**Theorem 10 (Side midpoints)** *Suppose that $p_1$ and $p_2$ are non-null, non-perpendicular points, forming a non-null side $\overline{p_1p_2}$. Then $\overline{p_1p_2}$ has a non-null midpoint $m$ precisely when $1-q(p_1,p_2)$ is a square, and in this case there are exactly two perpendicular midpoints $m$.*

**Proof.** We suppose without loss of generality that $p_1 = a_1 \equiv [1:0:0]$ and $p_2 = a_2 \equiv [0:1:0]$ so that by the Triangle quadrances and spreads theorem

$$1 - q(p_1,p_2) = \frac{(c-1)^2}{(bc-1)(ac-1)}.$$

By assumption each of $c-1, bc-1$ and $ac-1$ are nonzero. An arbitrary point $m$ on $ab = \langle 0:0:1\rangle$ has the form $m = [x:y:0]$, which is null precisely when $(bc-1)x^2 + (ac-1)y^2 + 2(1-c)xy = 0$, by the Null point theorem.

Assuming that $m$ is non-null, we compute that

$$\begin{aligned} q(p_1,m) &= \frac{Dcy^2}{(bc-1)((bc-1)x^2+(ac-1)y^2+2(1-c)xy)} \\ q(p_2,m) &= \frac{Dcx^2}{(ac-1)((bc-1)x^2+(ac-1)y^2+2(1-c)xy)}. \end{aligned}$$

By assumption $\overline{p_1p_2}$ is non-null, so by the Corollary to the Null points/lines theorem, $c \neq 0$, and so the above expressions are equal precisely when $x^2(bc-1) = y^2(ac-1)$ has a solution, which occurs precisely when $1-q(p_1,p_2)$ is a square. In fact if

$$\frac{1}{(bc-1)(ac-1)} = r^2 \qquad (18)$$

then the two midpoints are $m = [(ac-1)r : \pm 1 : 0]$, and they are perpendicular, since

$$\begin{aligned} &[(ac-1)r:1:0]\,\mathbf{A}\,[(ac-1)r:-1:0]^T \\ &= (ac-1)\left(1-(bc-1)(ac-1)r^2\right) = 0. \end{aligned} \qquad \square$$

We refer to the pair of midpoints $m$ of a side as **opposites**. It follows that the dual midline $M$ of a midpoint $m$ passes through the opposite midpoint. While the next theorem is dual to the previous one, we give a direct proof.

**Theorem 11 (Vertex bilines)** *Suppose that $L_1$ and $L_2$ are non-null non-perpendicular lines forming a non-null vertex $\overline{L_1L_2}$. Then $\overline{L_1L_2}$ has a non-null biline $B$ precisely when $1-S(L_1,L_2)$ is a square, and in this case there are exactly two perpendicular bilines $B$.*

**Proof.** We suppose without loss of generality that $L_1 = \langle 1:0:0\rangle$ and $L_2 = \langle 0:1:0\rangle$, so that from the Triangle quadrances/spreads theorem

$$1-S(L_1,L_2) = \frac{1}{ab}.$$

An arbitrary line through $L_1L_2 = [0:0:1]$ has the form $B = \langle l:m:0\rangle$, which by the Null spread theorem is null precisely when $al^2+bm^2+2lm = 0$, and then

$$S(L_1,B) = \frac{(ab-1)m^2}{a(al^2+bm^2+2lm)} \quad \text{and}$$
$$S(L_2,B) = \frac{(ab-1)l^2}{b(al^2+bm^2+2lm)}.$$

By assumption $\overline{L_1L_2}$ is non-null, so by the Corollary to the Null points/lines theorem, $ab-1 \neq 0$, and so the above expressions are equal precisely when $l^2a = m^2b$ has a solution, which occurs precisely when $1-S(L_1,L_2)$ is a square. In fact if

$$\frac{1}{ab} = w^2 \tag{19}$$

then the two bilines are $B = \langle l:m:0\rangle = \langle bw:\pm 1:0\rangle$ and they are perpendicular since

$$\langle bw:1:0\rangle^T \mathbf{B} \langle bw:-1:0\rangle = b\left(abw^2-1\right) = 0. \qquad \square$$

We refer to the pair of bilines $B$ of a vertex as **opposites**. It follows that the dual bipoint $b$ of a biline $B$ lies on the opposite biline.

## 3 Orthocenter Hierarchy

We now initiate our study of triangle geometry constructions involving *perpendicularity*. The focus is on the *Orthocenter* $h$ and various other key points that are related to the most important line in the subject: the *Orthoaxis* $A$. The computations are based on ortholinear coordinates; finding meets and joins, which essentially amount to taking cross products; and finding duals, either by multiplying transposes of points by $\mathbf{A}$ (on the left) or transposes of lines by $\mathbf{B}$ (on the right). Our goal is to establish formulas for important points and lines to facilitate the understanding of relationships between them: the reader is encouraged to follow along and check our computations, which are mostly elementary.

### 3.1 Triangle lines, dual points, dual lines

We start with a review of the basic Triangle $\overline{a_1a_2a_3}$, whose Points $a$ and Lines $L$ are

$$a_1 = [1:0:0],\ a_2 = [0:1:0],\ a_3 = [0:0:1] \quad \text{and}$$
$$L_1 = \langle 1:0:0\rangle,\ L_2 = \langle 0:1:0\rangle,\ L_3 = \langle 0:0:1\rangle.$$

The **Dual points** $l_1 \equiv L_1^\perp$, $l_2 \equiv L_2^\perp$, $l_3 \equiv L_3^\perp$ are the duals of the Lines $L$, and the **Dual lines** $A_1 \equiv a_1^\perp$, $A_2 \equiv a_2^\perp$, $A_3 \equiv a_3^\perp$ are the duals of the Points $a$. These are

$$l_1 = [a:1:1],\ l_2 = [1:b:1],\ l_3 = [1:1:c] \quad \text{and}$$
$$A_1 = \langle 1-bc : c-1 : b-1\rangle,$$
$$A_2 = \langle c-1 : 1-ac : a-1\rangle,$$
$$A_3 = \langle b-1 : a-1 : 1-ab\rangle.$$

The **Altitudes** are $N_1 \equiv a_1l_1$, $N_2 \equiv a_2l_2$, $N_3 \equiv a_3l_3$, and the **Altitude dual points** are $n_1 \equiv A_1L_1$, $n_2 \equiv A_2L_2$, $n_3 \equiv A_3L_3$. These are, as established previously,

$$N_1 = \langle 0:1:-1\rangle,\ N_2 = \langle 1:0:-1\rangle,\ N_3 = \langle 1:-1:0\rangle \text{ and}$$
$$n_1 = [0:b-1:1-c],\ n_2 = [a-1:0:1-c],$$
$$n_3 = [a-1:1-b:0].$$

The dual of the Orthocenter $h = [1:1:1]$ is the Ortholine

$$H = \langle b+c-bc-1 : a+c-ac-1 : a+b-ab-1\rangle.$$

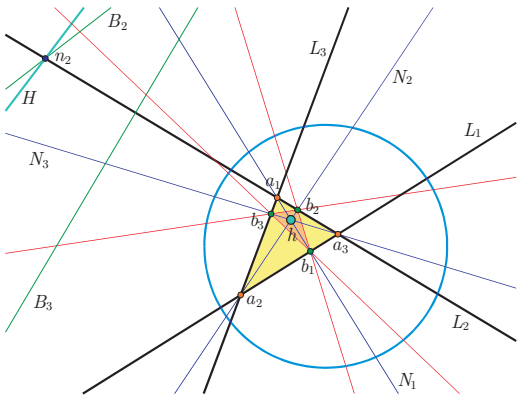



Figure 6: *Altitudes, Base points, Orthocenter h, Orthline H and Orthic triangle*

The **Base points** $b_1 \equiv N_1L_1$, $b_2 \equiv N_2L_2$, $b_3 \equiv N_3L_3$ are the meets of corresponding Altitudes $N$ and Lines $L$, and the **Base lines** $B_1 \equiv n_1l_1$, $B_2 \equiv n_2l_2$, $B_3 \equiv n_3l_3$ are their duals. These are

$$b_1 = [0:1:1],\ b_2 = [1:0:1],\ b_3 = [1:1:0] \quad \text{and}$$
$$B_1 = \langle b+c-2 : a(1-c) : a(1-b)\rangle,$$
$$B_2 = \langle b(1-c) : a+c-2 : b(1-a)\rangle,$$
$$B_3 = \langle c(1-b) : c(1-a) : a+b-2\rangle.$$

The **Orthic lines** $C_1 \equiv b_2b_3, C_2 \equiv b_1b_3, C_3 \equiv b_1b_2$ are the joins of Base points $b$, and the **Orthic points** $c_1 \equiv B_2B_3$, $c_2 \equiv B_1B_3, c_3 \equiv B_1B_2$ are the meets of Base lines $B$. These are

$$C_1 = \langle -1:1:1\rangle,\ C_2 = \langle 1:-1:1\rangle,\ C_3 = \langle 1:1:-1\rangle,$$
$$c_1 = [2-a:b:c],\ c_2 = [a:2-b:c],\ c_3 = [a:b:2-c].$$

The **Orthic triangle** $\overline{b_1b_2b_3}$ is perspective with the Triangle $\overline{a_1a_2a_3}$, with center of perspectivity the Orthocenter $h$, since the Altitudes are the lines of perspectivity.

**Theorem 12 (Triangle Base center)** *The **Orthic dual triangle** $\overline{c_1c_2c_3}$ is perspective with the Triangle $\overline{a_1a_2a_3}$, and the center of perspectivity is the **Base center** $b = [a:b:c]$.*

**Proof.** We compute the lines

$$a_1c_1 = [1:0:0]\times[2-a:b:c] = \langle 0:-c:b\rangle,$$
$$a_2c_2 = [0:1:0]\times[a:2-b:c] = \langle c:0:-a\rangle,$$
$$a_3c_3 = [0:0:1]\times[a:b:2-c] = \langle -b:a:0\rangle$$

and check that these are all incident with $b \equiv [a:b:c]$. □

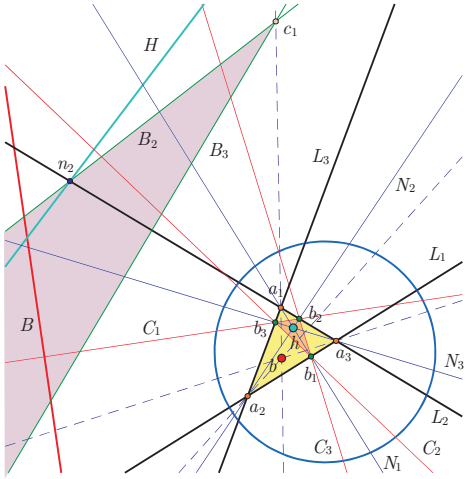



Figure 7: *Orthic dual triangle $\overline{c_1c_2c_3}$ and Base center $b$*

Note that there is a bit of duplication of symbols here, the letter $b$ being used in the same formula with two different meanings, hopefully without undue confusion. The Base center is an important triangle point, as we shall see; its dual is the **Base axis**

$$B \equiv \langle a-b-c+2bc-abc:$$
$$-a+b-c+2ac-abc:$$
$$-a-b+c+2ab-abc\rangle.$$

Figure 7 shows the three Orthic lines $C_1, C_2, C_3$ but only one of the Orthic points, namely $c_1$, since the other points are off the screen. The mathematical symmetry between points and lines is not respected by our biology; lines tend to be more visible, while points are simpler.

### 3.2 Orthic axis and Orthoaxis (they are different!)

The **Desargues points** $g_1 \equiv C_1L_1$, $g_2 \equiv C_2L_2$, $g_3 \equiv C_3L_3$ are the meets of corresponding Orthic lines $C$ and Lines $L$, and the **Desargues lines** $G_1 \equiv c_1l_1$, $G_2 \equiv c_2l_2$, $G_3 \equiv c_3l_3$ are the joins of corresponding Orthic points and Dual points. These are

$$g_1 = [0:1:-1],\ g_2 = [1:0:-1],\ g_3 = [1:-1:0] \quad \text{and}$$
$$G_1 = \langle b-c:a+ac-2:2-a-ab\rangle,$$
$$G_2 = \langle 2-ba-b:c-a:b+ab-2\rangle,$$
$$G_3 = \langle c+bc-2:2-ac-c:a-b\rangle.$$

**Theorem 13 (Triangle orthic axis)** *The Desargues points $g_1, g_2, g_3$ are collinear, and lie on the **Orthic axis** $S \equiv \langle 1:1:1\rangle$. The Desargues lines $G_1, G_2, G_3$ are concurrent, and pass through the **Orthostar** $s \equiv [a+2:b+2:c+2]$. The Orthic axis and Orthostar are dual.*

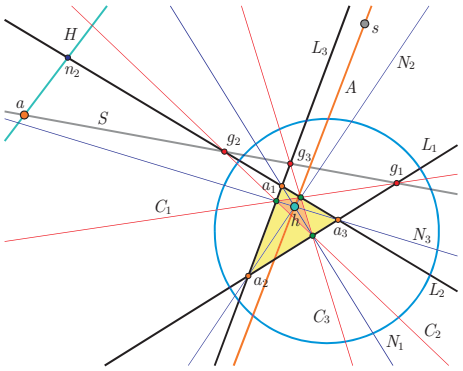



Figure 8: *Desargues points g, Orthic axis S, Orthoaxis A*

**Proof.** The Desargues points $g_1, g_2, g_3$ are collinear either by Desargues theorem applied to the Cevian triangle $\overline{b_1b_2b_3}$ of the Orthocenter $h$, or directly since

$$\det\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} = 0.$$

Their join, the Orthic axis, is

$$S \equiv g_1g_2 = [0:1:-1]\times[1:0:-1] = \langle -1:-1:-1\rangle$$
$$= \langle 1:1:1\rangle.$$

Dually the Desargues lines $G_1, G_2, G_3$ are concurrent, which we can check by evaluating the corresponding determinant. The common point through which they pass is the Orthostar

$$s \equiv S^{\perp} = \langle 1:1:1\rangle^T B = [a+2:b+2:c+2].$$

This is clearly dual to the Orthic axis. □

We now come to the feature attraction of this paper: the **Orthoaxis** $A$ is the join of the Orthocenter $h$ and the Orthostar $s$, or equivalently the canonical line of $h$. It is

$$A \equiv hs = [1:1:1] \times [a+2:b+2:c+2] \\ = \langle c-b : a-c : b-a \rangle .$$

Note that the Orthoaxis $A$ is perpendicular to the Orthic axis $S$, since the Orthostar $s$ lies on the Orthoaxis. *The Orthoaxis is the most important line in projective triangle geometry.*

The **Orthoaxis point** $a \equiv HS$ is the dual of the Orthoaxis; it is

$$a = A^{\perp} = \langle c-b : a-c : b-a \rangle^T \mathbf{B} \\ = [(a-1)(b-c) : (b-1)(a-c) : (c-1)(a-b)] .$$

**Theorem 14 (Base center on Orthoaxis)** *The Base center $b$ lies on the Orthoaxis $A$.*

**Proof.** We check incidence between the Base center $b = [a:b:c]$ and the Orthoaxis $A = \langle c-b : a-c : b-a \rangle$:

$$bA = [a:b:c] \langle c-b : a-c : b-a \rangle \\ = [a(c-b) + b(a-c) + c(b-a)] = 0. \qquad \square$$

The **AntiOrthic lines** $T_1 \equiv a_1 g_1, T_2 \equiv a_2 g_2, T_3 \equiv a_3 g_3$ are the joins of corresponding Points $a$ and Desargues points $g$, and the **AntiOrthic points** $t_1 \equiv A_1 G_1, t_2 \equiv A_2 G_2, t_3 \equiv A_3 G_3$ are the meets of corresponding Dual lines $A$ and Desargues lines $G$. They have the form

$$T_1 = \langle 0:1:1 \rangle , \; T_2 = \langle 1:0:1 \rangle , \; T_3 = \langle 1:1:0 \rangle \quad \text{and}$$
$$t_1 = [2 : b+1 : c+1] ,$$
$$t_2 = [a+1 : 2 : c+1] ,$$
$$t_3 = [a+1 : b+1 : 2] .$$

The **AntiBase points** $e_1 \equiv T_2 T_3, e_2 \equiv T_1 T_3, e_3 \equiv T_1 T_2$ are the meets of AntiOrthic lines $T$, and the **AntiBase lines** $E_1 \equiv t_2 t_3, E_2 \equiv t_1 t_3, E_3 \equiv t_1 t_2$ are the joins of AntiOrthic points $t$. They have the form

$$e_1 = [-1:1:1] , \; e_2 = [1:-1:1] , \; e_3 = [1:1:-1]$$

and

$$E_1 = \langle b+c+bc-3 : (1-c)(a+1) : (1-b)(a+1) \rangle ,$$
$$E_2 = \langle (1-c)(b+1) : (a+c+ac-3) : (1-a)(b+1) \rangle ,$$
$$E_3 = \langle (1-b)(c+1) : (1-a)(c+1) : (a+b+ab-3) \rangle .$$

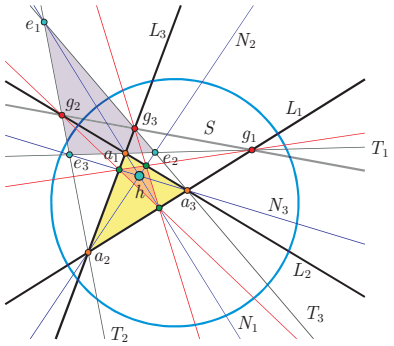



Figure 9: *AntiBase points and the AntiOrthic triangle* $\overline{e_1 e_2 e_3}$

**Theorem 15 (AntiOrthic perspectivity)** *The **AntiOrthic triangle** $\overline{e_1 e_2 e_3}$ and the Triangle $\overline{a_1 a_2 a_3}$ are perspective from the Orthocenter $h$.*

**Proof.** This is equivalent to the statement that the Altitudes $N$ pass through the corresponding AntiBase points $e$. For example $N_1$ is incident with $e_1$, since $e_1 N_1 = [-1:1:1] \langle 0:1:-1 \rangle = 0$, and similarly $N_2$ is incident with $e_2$, and $N_3$ is incident with $e_3$. $\square$

It is perhaps worth mentioning that one can work out formulas for these various constructs directly in hyperbolic geometry in terms of the coordinates of a general triangle. However this proves rather taxing; even the Orthocenter involves for each coefficient a homogeneous polynomial of degree six with 24 terms. The system presented here punches far above its weight, as the relative simplicity of the formulas so far confirms.

### 3.3 Parallels and the Double triangle

In universal hyperbolic geometry, the notion of parallel is more specialized than in classical hyperbolic geometry. We do not refer to two lines (or two points) as being parallel. Rather we refer to **a line $P$ through a point $a$ being parallel to a line $L$** : it means that $P$ is perpendicular to the altitude from $a$ to $L$. In Euclidean geometry this is like defining parallel lines to be "perpendicular to a perpendicular": a local definition rather than a global one. This motivates the important construction of the Double triangle of a Triangle.

The **Parallel lines** $P_1 \equiv a_1 n_1$, $P_2 \equiv a_2 n_2$, $P_3 \equiv a_3 n_3$ are the joins of corresponding Points $a$ and Altitude points $n$, and the **Parallel points** $p_1 \equiv A_1 N_1, p_2 \equiv A_2 N_2, p_3 \equiv A_3 N_3$ are their duals: meets of corresponding Dual lines $A$ and

Altitudes $N$. These are

$$P_1 = \langle 0 : c-1 : b-1 \rangle,\ P_2 = \langle c-1 : 0 : a-1 \rangle,$$
$$P_3 = \langle b-1 : a-1 : 0 \rangle \quad \text{and}$$
$$p_1 = [2-b-c : 1-bc : 1-bc],$$
$$p_2 = [1-ac : 2-a-c : 1-ac],$$
$$p_3 = [1-ab : 1-ab : 2-a-b].$$

The **Double points** $d_1 \equiv P_2P_3$, $d_2 \equiv P_1P_3$, $d_3 \equiv P_1P_2$ are the meets of Parallel lines $P$, and the **Double lines** $D_1 \equiv p_2p_3$, $D_2 \equiv p_1p_3$, $D_3 \equiv p_1p_2$ are the joins of Parallel points $p$. These are

$$d_1 = [1-a : b-1 : c-1],\ d_2 = [a-1 : 1-b : c-1],$$
$$d_3 = [a-1 : b-1 : 1-c] \quad \text{and}$$
$$D_1 = \langle a+2b+2c-bc-abc-3 : (ac-1)(b-1) : (ab-1)(c-1) \rangle,$$
$$D_2 = \langle (bc-1)(a-1) : 2a+b+2c-ac-abc-3 : (ab-1)(c-1) \rangle,$$
$$D_3 = \langle (bc-1)(a-1) : (ac-1)(b-1) : 2a+2b+c-ab-abc-3 \rangle.$$

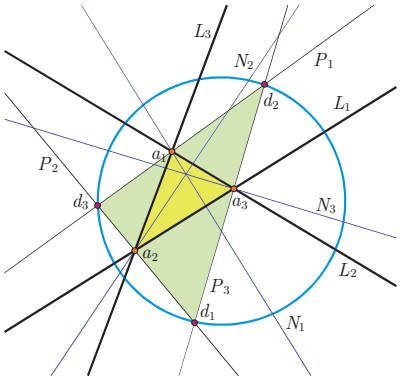



Figure 10: *Parallel lines and the Double triangle* $\overline{d_1d_2d_3}$

The triangle $\overline{d_1d_2d_3}$ is the **Double triangle** of the Triangle $\overline{a_1a_2a_3}$. The following theorems seem remarkable.

**Theorem 16 (Double triangle midpoint)** *The Points $a_1, a_2, a_3$ are midpoints of the Double triangle $\overline{d_1d_2d_3}$.*

**Proof.** Using the expressions above, we compute that

$$q(a_3, d_1) = \frac{D(a+b-2)}{4a+4b+4c-ab-ac-bc-a^2-b^2-c^2+abc^2+ab^2c+a^2bc-4abc-5}$$
$$= q(a_3, d_2)$$

where $D$ is the determinant defined in (15), and where a common factor of $(ab-1)$ in the numerator and denominator has been cancelled provided that $ab \neq 1$. So $a_3$ is a midpoint of $\overline{d_1d_2}$. Similarly $a_1$ is a midpoint of $\overline{d_2d_3}$, and $a_2$ is a midpoint of $\overline{d_1d_3}$. □

**Theorem 17 (Double triangle null points)** *The Double triangle $\overline{d_1d_2d_3}$ has a null point precisely when all of its points are null points, and this occurs precisely when the quadrea $\mathcal{A}$ of the Triangle $\overline{a_1a_2a_3}$ is equal to 1.*

**Proof.** Using the Null point theorem, $d_1 = [1-a : b-1 : c-1]$ is null precisely when

$$0 = (1-bc)(1-a)^2 + (1-ac)(b-1)^2 + (1-ab)(c-1)^2$$
$$+ (2c-2)(1-a)(b-1) + (2b-2)(1-a)(c-1) +$$
$$+ (2a-2)(b-1)(c-1).$$

After expanding and simplifying, the right hand side is the symmetric expression

$$5-4a-4b-4c+ab+ac+bc+a^2+b^2+c^2+$$
$$+4abc-abc^2-ab^2c-a^2bc.$$

This same expression arises for the nullity of $d_2 = [a-1 : 1-b : c-1]$ and $d_3 = [a-1 : b-1 : 1-c]$, so if one point of the Double triangle is null, so are the other two.

Using the Triangle quadrea theorem, the difference between the quadrea $\mathcal{A}$ of the Triangle and 1 is

$$\frac{(abc-b-c-a+2)^2}{(ab-1)(ac-1)(bc-1)} - 1 = \frac{5-4a-4b-4c+ab+ac+bc+a^2+b^2+c^2+4abc-abc^2-ab^2c-a^2bc}{(bc-1)(ac-1)(ab-1)}.$$

So the Double triangle has null points precisely when the quadrea $\mathcal{A}$ is equal to 1. □

In our standard example, the Triangle has approximate quadrea $\mathcal{A} \approx 1.04$, which explains why the points of the Double triangle in Figure 10 appear close to being null points.

**Theorem 18 (Double triangle perspectivity)** *The Double triangle $\overline{d_1d_2d_3}$ and the Triangle $\overline{a_1a_2a_3}$ are perspective from a point, the **Double point,** or $x$ **point,** which is $x \equiv [a-1 : b-1 : c-1]$. The $x$ point lies on the Orthoaxis $A$.*

**Proof.** We compute the lines

$$a_1d_1 = [1:0:0] \times [1-a : b-1 : c-1] = \langle 0 : 1-c : b-1 \rangle,$$
$$a_2d_2 = [0:1:0] \times [a-1 : 1-b : c-1] = \langle c-1 : 0 : 1-a \rangle,$$
$$a_3d_3 = [0:0:1] \times [a-1 : b-1 : 1-c] = \langle 1-b : a-1 : 0 \rangle.$$

These lines are concurrent (compute a determinant), and the common meet is

$$x = \langle 0 : 1-c : b-1 \rangle \times \langle c-1 : 0 : 1-a \rangle$$
$$= \left[ (a-1)(c-1) : (b-1)(c-1) : (c-1)^2 \right]$$
$$= [a-1 : b-1 : c-1].$$

We have cancelled a common factor $c-1$; should this be zero, then the three lines are concurrent since two of them are equal. The $x$ point lies on the Orthoaxis $A \equiv hs$ since $[a-1:b-1:c-1]$ is a (projective) linear combination of the Orthocenter $h \equiv [1:1:1]$ and the Orthostar $s \equiv [a+2:b+2:c+2]$. □

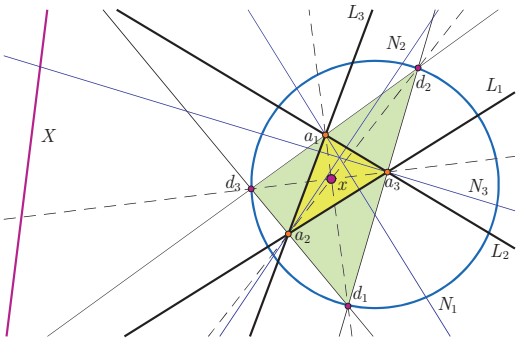



Figure 11: *The Double point, or x point*

The dual of the $x$ point is the $X$ **line**

$$X = \langle a-2b-2c+3bc-abc+1 : \\ -2a+b-2c+3ac-abc+1 : \\ -2a-2b+c+3ab-abc+1 \rangle.$$

**Theorem 19 (Double dual triangle perspectivity)** *The Double triangle $\overline{d_1d_2d_3}$ and the Dual triangle $\overline{l_1l_2l_3}$ are perspective from a point, the **Double dual point,** or $z$ **point**, which is $z \equiv [a+1:b+1:c+1]$. The $z$ point lies on the Orthoaxis.*

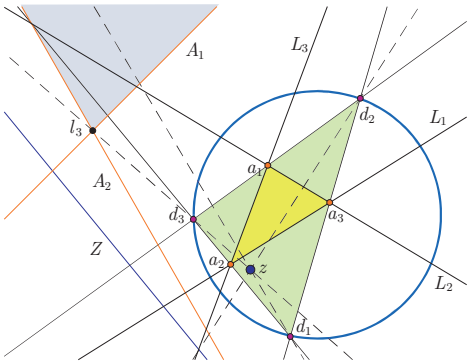



Figure 12: *The Double dual point, or z point*

**Proof.** We compute the lines

$$\begin{aligned} l_1d_1 &= [a:1:1] \times [1-a:b-1:c-1] \\ &= \langle c-b:1-ac:ab-1 \rangle, \\ l_2d_2 &= [1:b:1] \times [a-1:1-b:c-1] \\ &= \langle bc-1:a-c:1-ab \rangle, \\ l_3d_3 &= [1:1:c] \times [a-1:b-1:1-c] \\ &= \langle 1-bc:ac-1:b-a \rangle. \end{aligned}$$

These lines are concurrent and their common meet is $z \equiv [a+1:b+1:c+1]$. This point lies on the Orthoaxis since it is a (projective) linear combination of the Orthocenter $h \equiv [1:1:1]$ and the Orthostar $s \equiv [a+2:b+2:c+2]$. □

The dual of the $z$ point is the $Z$ **line**

$$Z \equiv \langle a+bc-abc-1 : b+ac-abc-1 : c+ab-abc-1 \rangle.$$

The **AltDual lines** $K_1 \equiv a_1n_1$, $K_2 \equiv a_2N_2$, $K_3 \equiv a_3n_3$ are the joins of corresponding Points $a$ and Altitude points $n$, and the **AltDual points** $k_1 \equiv A_1N_1$, $k_2 \equiv A_2N_2$, $k_3 \equiv A_3N_3$ are meets of corresponding Dual lines $A$ and Altitudes $N$. These are

$$\begin{aligned} &K_1 = \langle 0 : c-1 : b-1 \rangle, \; K_2 = \langle c-1 : 0 : a-1 \rangle, \\ &K_3 = \langle b-1 : a-1 : 0 \rangle \quad \text{and} \\ &k_1 = [2-b-c : 1-bc : 1-bc], \\ &k_2 = [1-ac : 2-a-c : 1-ac], \\ &k_3 = [1-ab : 1-ab : 2-a-b]. \end{aligned}$$

Clearly the **AltDual triangle** $\overline{k_1k_2k_3}$ is perspective to the Triangle, since the Altitudes pass through both Points and AltDual points, but in addition, as Figure 13 suggests, it is also in perspective with the Double triangle. The center of perspectivity, the **AltDual point** $k$, does not generally lie on the Orthoaxis. We leave this result to the reader, as well as the following exercise investigating other perspective centers of various secondary triangles associated to our basic Triangle.

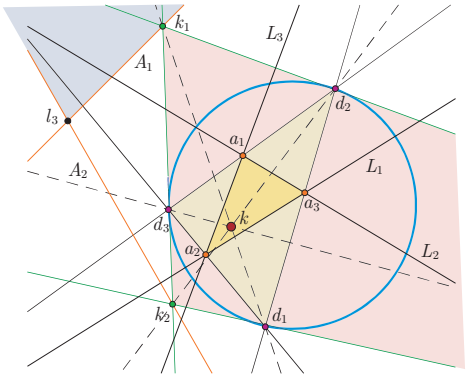



Figure 13: *AltDual triangle $\overline{k_1k_2k_3}$ and the AltDual point k*

**Exercise 1** *Show that the following pairs of triangles are perspective, and find the centers of perspectivity: i) $\overline{c_1c_2c_3}$ and $\overline{p_1p_2p_3}$, ii) $\overline{b_1b_2b_3}$ and $\overline{d_1d_2d_3}$, iii) $\overline{b_1b_2b_3}$ and $\overline{l_1l_2l_3}$, iv) $\overline{c_1c_2c_3}$ and $\overline{d_1d_2d_3}$, v) $\overline{c_1c_2c_3}$ and $\overline{l_1l_2l_3}$, vi) $\overline{k_1k_2k_3}$ and $\overline{d_1d_2d_3}$. Can you find more?*

### 3.4 Special points on the Orthoaxis

There are five interesting, and related, points on the Orthoaxis $A = \langle c-b : a-c : b-a \rangle$, namely the points $z = [a+1 : b+1 : c+1]$, $b = [a : b : c]$, $x = [a-1 : b-1 : c-1]$, $h = [1 : 1 : 1]$ and $s = [a+2 : b+2 : c+2]$. Of course there very well may be more! In Figure 14 we see them in this particular order.

**Theorem 20 (Orthoaxis harmonic ranges)** *The points $z, b, x, h$ form a harmonic range. The points $z, b, h, s$ also form a harmonic range.*

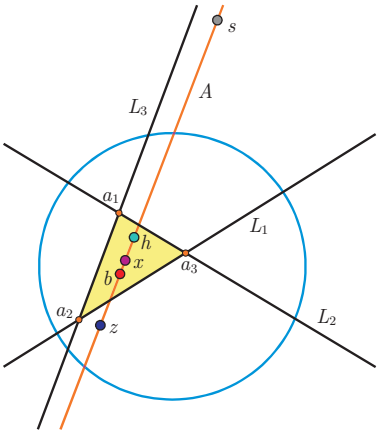



Figure 14: *Orthoaxis A and points $z, b, x, h$ and $s$*

**Proof.** Recall that the points $z, b, x, h$ form a harmonic range precisely when the cross ratio $R(z, x : b, h) = -1$. Define the following vectors which represent each of the five points:

$$\overrightarrow{z} = (a+1, b+1, c+1),$$
$$\overrightarrow{b} = (a, b, c),$$
$$\overrightarrow{x} = (a-1, b-1, c-1),$$
$$\overrightarrow{h} = (1, 1, 1),$$
$$\overrightarrow{s} = (a+2, b+2, c+2).$$

Then $\overrightarrow{z} + \overrightarrow{x} = 2\overrightarrow{b}$ and $\overrightarrow{z} - \overrightarrow{x} = 2\overrightarrow{h}$ so that $b$ and $h$ are harmonic conjugates with respect to $z$ and $x$, so that $R(z, x : b, h) = -1$. Similarly $\overrightarrow{z} + \overrightarrow{h} = \overrightarrow{s}$ and $\overrightarrow{z} - \overrightarrow{h} = \overrightarrow{b}$ so that $b$ and $s$ are harmonic conjugates with respect to $z$ and $h$, so that $R(z, h : b, s) = -1$. □

There are three more cross ratios naturally determined by the five points. We will leave it to the reader to check that in addition

$$R(z, x : b, s) = -3, R(b, h : x, s) = -1/2, R(z, h : x, s) = -2.$$

**Theorem 21 (Second double triangle perspectivity)** *The Triangle $\overline{a_1a_2a_3}$ and the double triangle of the Double triangle $\overline{d_1d_2d_3}$ are perspective from the **Second double point**, or $y$ **point**.*

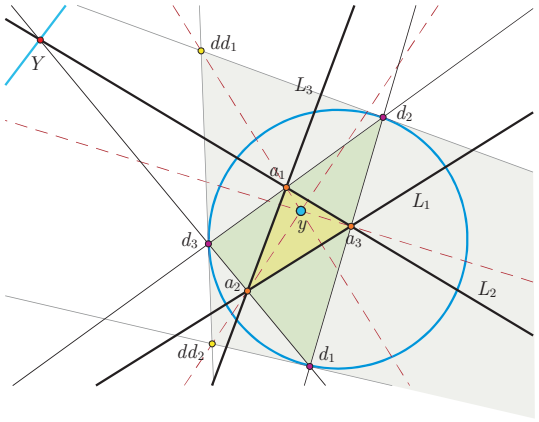



Figure 15: *The double of the Double triangle and the $y$ point*

**Proof.** We leave the proof to the reader. The formula for $y$ is somewhat lengthy to write out. □

It is worth pointing out that in general the $y$ point does *not* lie on the Orthoaxis $A$, although it often, as in our example, gets very close! It is also worth noting that the obvious pattern does not appear to continue; the double of the double of the Double triangle is *not* in general perspective with the original Triangle.

## 4 Incenter Hierarchy

Although the Incenter and Circumcenter hierarchies are exactly dual, we treat first the former, which is closer to the Euclidean situation and so more familiar. It is also simpler, on account of the difference between (18) and (19).

### 4.1 Bilines, Incenters and Apollonians

From the Vertex bilines theorem, Bilines of the Triangle exist precisely when the spreads $S_1, S_2, S_3$ have the property that $1-S_1, 1-S_2, 1-S_3$ are all squares. In Figure 16, where we are working approximately, this amounts to each of these quantities being positive, which they are, and so there are six Bilines $B$, two opposite ones for each vertex, and six dual Bipoints $b$, two opposite ones lying on each Dual line.

From the Triangle quadrances and spreads theorem, Bilines exist precisely when we can find $u, v, w$ satisfying the **quadratic relations**

$$\frac{1}{bc} = u^2, \quad \frac{1}{ac} = v^2, \quad \frac{1}{ab} = w^2. \tag{20}$$

But there is also a **cubic relation**

$$\frac{1}{abc} = uvw \tag{21}$$

which we are able to impose, by taking the product of the three quadratic relations (20), and possibly changing the sign of any or all of $u,v,w$ (they must all be non-zero).

These relations (20) and (21) will play an essential role in what follows; any triple $\{u,v,w\}$ satisfying them gives rise to three more such triples: namely $\{u,-v,-w\},\{-u,v,-w\}$ and $\{-u,-v,w\}$. So there is a *fourfold Klein-type symmetry* occurring here. Another implication is that we also have the relations

$$u = avw,\ v = buw,\ w = cuv. \tag{22}$$

We saw at the end of the proof of the Vertex bilines theorem, that Bilines for the vertex $\overline{L_1L_2}$ are $B = \langle bw : \pm 1 : 0\rangle$. We see now from (22) that these can be rewritten as $\langle v : u : 0\rangle$ and $\langle v : -u : 0\rangle$.

So the Bilines $B$ are

$$\begin{aligned}
&\langle 0 : w : v\rangle, \langle 0 : w : -v\rangle \quad \text{through} \quad a_1,\\
&\langle w : 0 : u\rangle, \langle w : 0 : -u\rangle \quad \text{through} \quad a_2,\\
&\langle v : u : 0\rangle, \langle v : -u : 0\rangle \quad \text{through} \quad a_3.
\end{aligned}$$

The Bipoints $b$ are dual, and are

$$\begin{aligned}
&[v+w : v+bw : w+cv],\ [v-w : v-bw : -w+cv]\\
&\qquad \text{on} \quad A_1,\\
&[u+aw : u+w : w+cu],\ [u-aw : u-w : -w+cu]\\
&\qquad \text{on} \quad A_2,\\
&[u+av : v+bu : u+v],\ [u-av : v-bu : -u+v]\\
&\qquad \text{on} \quad A_3.
\end{aligned}$$

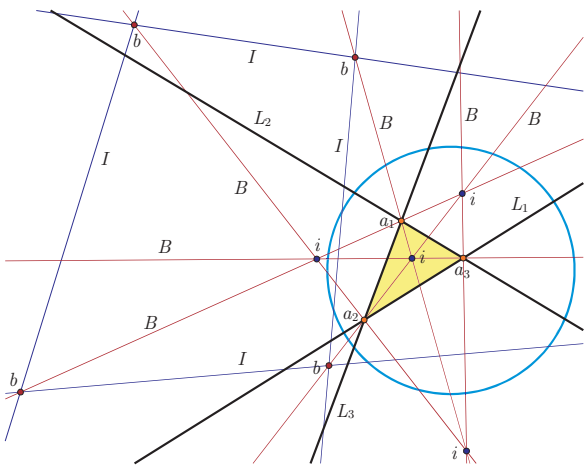



Figure 16: *Bilines, Bipoints, Incenters and Inlines*

**Theorem 22 (Incenters)** *Bilines $B$ are concurrent in threes, meeting at four* ***Incenters*** *$i_0,i_1,i_2$ and $i_3$. Bipoints $b$ are collinear in threes, joining on four* ***Inlines*** *$I_0,I_1,I_2$ and $I_3$.*

**Proof.** The following triples of Bilines $B$ are concurrent:

$$\begin{aligned}
&\langle 0 : w : -v\rangle,\ \langle w : 0 : -u\rangle,\ \langle v : -u : 0\rangle\\
&\qquad \text{through} \quad i_0 \equiv [u : v : w],\\
&\langle 0 : w : -v\rangle,\ \langle w : 0 : u\rangle,\ \langle v : u : 0\rangle\\
&\qquad \text{through} \quad i_1 \equiv [-u : v : w],\\
&\langle 0 : w : v\rangle,\ \langle w : 0 : -u\rangle,\ \langle v : u : 0\rangle\\
&\qquad \text{through} \quad i_2 \equiv [u : -v : w],\\
&\langle 0 : w : v\rangle,\ \langle w : 0 : u\rangle,\ \langle v : -u : 0\rangle\\
&\qquad \text{through} \quad i_3 \equiv [u : v : -w].
\end{aligned}$$

We check this by computing

$$\begin{aligned}
&\det\begin{pmatrix} 0 & w & -v\\ w & 0 & -u\\ v & -u & 0\end{pmatrix}\\
=&\det\begin{pmatrix} 0 & w & -v\\ w & 0 & u\\ v & u & 0\end{pmatrix}\\
=&\det\begin{pmatrix} 0 & w & v\\ w & 0 & -u\\ v & u & 0\end{pmatrix}\\
=&\det\begin{pmatrix} 0 & w & v\\ w & 0 & u\\ v & -u & 0\end{pmatrix} = 0.
\end{aligned}$$

The corresponding meets are $\langle 0 : w : -v\rangle \times \langle w : 0 : -u\rangle = \left[-uw : -vw : -w^2\right] = [u : v : w] \equiv i_0$ and similarly for the other Incenters. The situation with Bipoints $b$ is dual. □

At this point, it appears that there is no intrinsic reason to prefer one Incenter over the others; our notation seems somewhat arbitrary. However it is possible that this symmetry may eventually be broken.

**Apollonian points** $a$ are meets of Bilines $B$ and corresponding Lines $L$. There are six; they have the form

$$\begin{aligned}
&[0 : v : w],\ [0 : v : -w],\\
&[u : 0 : w],\ [u : 0 : -w],\\
&[u : v : 0],\ [u : -v : 0]
\end{aligned} \tag{23}$$

on $L_1,L_2,L_3$ respectively. The **Apollonian lines** $A$ are joins of Bipoints $b$ and corresponding Dual points $l$.

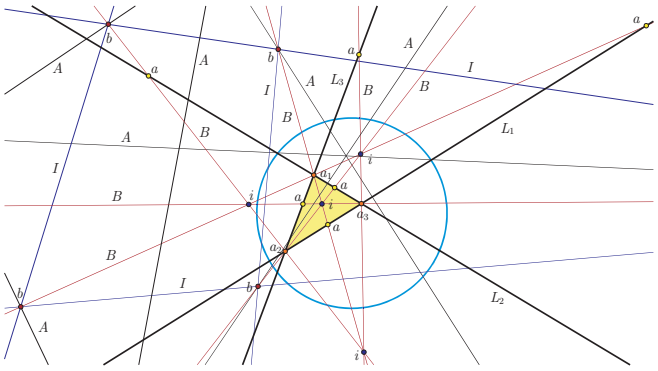


Figure 17: *Apollonian points a and lines A*

**Theorem 23 (Apollonian harmonic conjugates)** *The two Apollonian points on a side of the Triangle are harmonic conjugates with respect to the two Points of that side.*

**Proof.** Consider the side $\overline{a_2a_3}$ with Apollonian points $[0:w:v]$ and $[0:w:-v]$. Then it is a standard fact that the projective points determined by the vectors $(0,v,w)$ and $(0,v,-w)$ are harmonic conjugates with respect to those determined by the vectors $(0,v,0)$ and $(0,0,w)$. □

We leave the dual result concerning Apollonian lines to the reader.

A famous property of the Apollonian points in the Euclidean case is that the three circles built from pairs of these as diameters meet at the two *Isodynamic points*. This property is modified in the projective setting, by introducing an *important variant of a circle*. Given two points $a$ and $b$, the **Thaloid** of the side $\overline{ab}$ is the locus of a point $p$ satisfying the property that $pa \perp pb$, or equivalently

$$S(pa,pb) = 1. \tag{24}$$

It is straightforward that this is a conic. It is not generally a (metrical) circle, but shares some of its properties.

An **Apollonian Thaloid** is a Thaloid of a side consisting of two Apollonian points, both on a Line of the triangle. There are three Apollonian Thaloids, one for each side of the Triangle.

**Theorem 24 (Isodynamic points)** *If two Apollonian Thaloids meet at a point s, then the third does too.*

**Proof.** The equations of the three Thaloids are obtained directly from the forms of the Apollonians in (23) and the defining relation (24):

$$bcx^2 - acy^2 - 2bzx + 2azy = 0,$$
$$acy^2 - abz^2 - 2cxy + 2bxz = 0,$$
$$abz^2 - bcx^2 - 2ayz + 2cxy = 0.$$

If we add these three equations we get zero on both sides, so they are dependent. So if two Thaloids have a common point $s$, then this is shared by the third. □

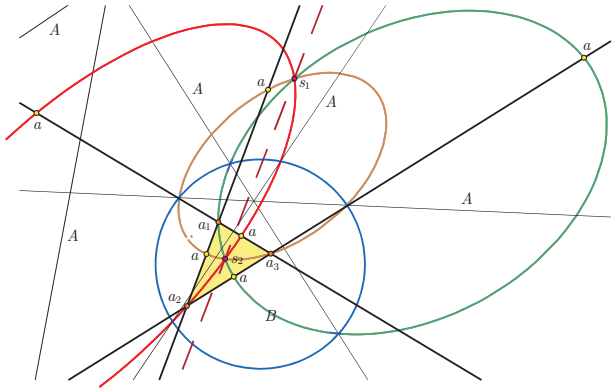


Figure 18: *Apollonian Thaloids and Isodynamic points* $s_1, s_2$

Such a common point of the three Apollonian Thaloids is an **Isodynamic point**; Figure 18 shows two such: $s_1$ and $s_2$, together with the line through them. In the Euclidean case this is the **Brocard line**, which also passes through the orthocenter—here in the projective situation that is not generally the case, even though it appears to in this example. In the Euclidean case the centers of the three Apollonian circles are collinear, falling on the **Lemoine line**: something analogous happens here but it requires additional ideas which we leave for another occasion.

### 4.2 Centrians, InCentrians, Contact points and Incircles

**Theorem 25 (Centrian lines)** *The Apollonian points a are collinear in threes, joining on four* ***Centrian lines*** *J. The Apollonian lines A are concurrent in threes, meeting at four* ***Centrian points*** *j.*

**Proof.** The following triples of Apollonian points are collinear:

$[0:v:-w]$, $[u:0:-w]$, $[u:-v:0]$ on $J_0 \equiv \langle vw:uw:uv\rangle$,
$[0:v:-w]$, $[u:0:w]$, $[u:v:0]$ on $J_1 \equiv \langle -vw:uw:uv\rangle$,
$[0:v:w]$, $[u:0:-w]$, $[u:v:0]$ on $J_2 \equiv \langle vw:-uw:uv\rangle$,
$[0:v:w]$, $[u:0:w]$, $[u:-v:0]$ on $J_3 \equiv \langle vw:uw:-uv\rangle$.

The collinearities may easily be checked by computing determinants. The corresponding meets are $[0:v:-w] \times [u:0:-w] = \langle vw:uw:uv\rangle \equiv J_0$ and similarly for the other Centrian lines. The situation with the Apollonian lines $A$ is dual; here are the formulas for the Centrian points:

$$j_0 \equiv [uv+uw+avw : uv+vw+buw : uw+vw+cuv],$$
$$j_1 \equiv [uv+uw-avw : uv-vw+buw : uw-vw+cuv],$$
$$j_2 \equiv [uv-uw+avw : uv+vw-buw : vw-uw+cuv],$$
$$j_3 \equiv [uw-uv+avw : vw-uv+buw : uw+vw-cuv]. \quad \square$$

The four Incenters and the four Centrian points are corresponding, since the three Bilines that meet in an Incenter also give rise to the three Apollonian points lying on a particular Centrian line, which is dual to a particular Centrian point. The **InCentrian lines** are joins of corresponding Incenters $i$ and Centrian points $j$. These have the form

$$i_0 j_0 = \langle cv-bw : aw-cu : bu-av\rangle,$$
$$i_1 j_1 = \langle cv-bw : aw+cu : -bu-av\rangle,$$
$$i_2 j_2 = \langle -cv-bw : aw-cu : bu+av\rangle,$$
$$i_3 j_3 = \langle cv+bw : -aw-cu : bu-av\rangle.$$

While it is easy to check that the Incenters lie on these InCentrian lines, showing that the Centrian points do so requires the quadratic relations. (The reader is encouraged to check this).

The **InCentrian points** are meets of corresponding Inlines $I$ and Centrian lines $J$, and are dual to the InCentrian lines.

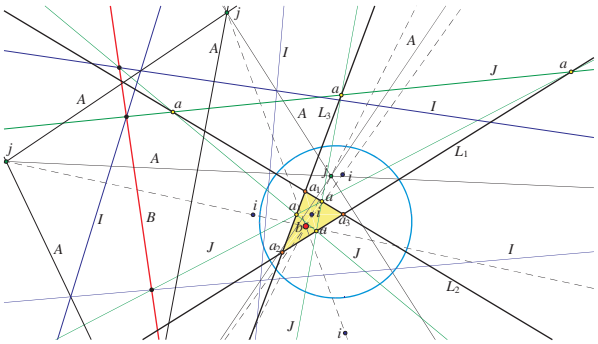


Figure 19: *Centrians, Incentrians and Base center b*

**Theorem 26 (InCentrian center)** *The four InCentrian lines are concurrent, and meet at the Base center b. The four InCentrian points are collinear, and join on the Base axis B.*

**Proof.** The InCentrian line $i_0 j_0 = \langle cv-bw : aw-cu : bu-av\rangle$ passes through $b = [a:b:c]$ since

$$[a:b:c]\,\langle cv-bw : aw-cu : bu-av\rangle$$
$$= [(cv-bw)a + (aw-cu)b + (bu-av)c] = 0.$$

Similarly $j_1 i_1$, $j_2 i_2$, $j_3 i_3$ also pass through $b$. The second statement follows by duality. $\square$

The **InDual** lines are joins of Incenters $i$ and Dual points $l$. They may also be described as altitudes from Incenters to the Lines, and there are 12. The InDual lines associated to the Incenters are:

$$\langle v-w : aw-u : u-av\rangle,\ \langle v-bw : w-u : bu-v\rangle,$$
$$\langle cv-w : w-cu : u-v\rangle \quad \text{to} \quad i_0 = [u:v:w],$$
$$\langle v-w : aw+u : -u-av\rangle,\ \langle v-bw : w+u : -bu-v\rangle,$$
$$\langle cv-w : w+cu : -u-v\rangle \quad \text{to} \quad i_1 = [-u:v:w],$$
$$\langle -v-w : aw-u : u+av\rangle,\ \langle -v-bw : w-u : bu+v\rangle,$$
$$\langle -cv-w : w-cu : u+v\rangle \quad \text{to} \quad i_2 = [u:-v:w],$$
$$\langle v+w : -aw-u : u-av\rangle,\ \langle v+bw : -w-u : bu-v\rangle,$$
$$\langle cv=w : -w-cu : u-v\rangle \quad \text{to} \quad i_3 = [u:v:-w].$$

The **InDual points** are meets of Inlines $I$ and Lines $L$, and are dual to InDual lines.

The **Contact points** are meets of corresponding InDual lines and Lines $L$; there are a total of 12. They may also be described as the bases of the altitudes from the Incenters to the Lines. The Contact points associated to the Incenters are:

$$[0:u-av:u-aw],\ [v-bu:0:v-bw],$$
$$[w-cu:w-cv:0] \quad \text{to} \quad i_0 = [u:v:w],$$
$$[0:u+av:u+aw],\ [v+bu:0:v-bw],$$
$$[w+cu:w-cv:0] \quad \text{to} \quad i_1 = [-u:v:w],$$
$$[0:u+av:u-aw],\ [v+bu:0:v+bw],$$
$$[w-cu:w+cv:0] \quad \text{to} \quad i_2 = [u:-v:w],$$
$$[0:u-av:u+aw],\ [v-bu:0:v+bw],$$
$$[w+cu:w+cv:0] \quad \text{to} \quad i_3 = [u:v:-w].$$

The **Contact lines** are joins of corresponding InDual points and Dual points $l$. Figure 20 shows the InDual lines and Contact points. The latter are intimately connected with important conics associated to the Triangle—the Incircles.

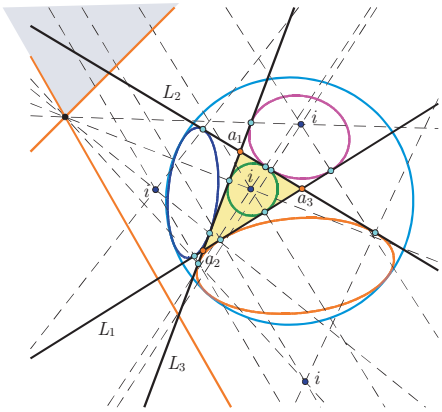



Figure 20: *InDual lines, Contact points and Incircles*

A **circle** is the locus of a point $p$ satisfying $q(p,c) = K$ for some fixed point $c$ and some fixed number $K$: this is a conic. Circles centered at the Incenters and passing through the associated Contact points are tangent to the Lines at these points, and are called **Incircles**; there are in this case four, and these are also shown in Figure 20.

Note that in this particular case the two Incircles whose centers are outside the null circle have a quite different character from the two with interior centers. The former are often called 'curves of constant width' in the classical literature, and are tangent to the null circle at the points where the dual of the center (in this case an Inline) meets it. Circles can take on different forms, appearing in our affine view also as hyperbolas outside the null circle, tangent to it at these same points. See [18] for some pictures; also the video *UnivHypGeom25: Geometer's Sketchpad and Visualizing circles in Universal Hyperbolic Geometry* in the YouTube playlist [20].

### 4.3 Sight lines, Gergonne and Nagel points

A **Sight line** is the join of a Contact point with the Point $a$ opposite to the Line that it lies on. A **Sight point** is the dual of a Sight line. There are 12 Sight lines; three associated to each Incenter, and four incident with each Point. The Sight lines associated to the Incenters are:

$$\begin{aligned}
&\langle 0 : u-aw : -u+av\rangle,\ \langle v-bw : 0 : -v+bu\rangle,\\
&\qquad \langle w-cv : -w+cu, 0\rangle \quad \text{to} \quad i_0 = [u:v:w],\\
&\langle 0 : -u-aw : u+av\rangle,\ \langle v-bw : 0 : -v-bu\rangle,\\
&\qquad \langle w-cv : -w-cu, 0\rangle \quad \text{to} \quad i_1 = [-u:v:w],\\
&\langle 0 : u-aw : -u-av\rangle,\ \langle -v-bw : 0 : v+bu\rangle,\\
&\qquad \langle w+cv : -w+cu, 0\rangle \quad \text{to} \quad i_2 = [u:-v:w],\\
&\langle 0 : u+aw : -u+av\rangle,\ \langle v+bw : 0 : -v+bu\rangle,\\
&\qquad \langle -w-cv : w+cu, 0\rangle \quad \text{to} \quad i_3 = [u:v:-w].
\end{aligned}$$

**Theorem 27 (Gergonne points)** *The three Sight lines associated to an Incenter meet at a* ***Gergonne point*** *$g$. These are:*

$$\begin{aligned}
g_0 &= [-abcu+acv+abw-1 : bcu-abcv+abw-1 :\\
&\qquad bcu+acv-abcw-1],\\
g_1 &= [abcu+acv+abw+1 : -bcu-abcv+abw+1 :\\
&\qquad -bcu+acv-abcw+1],\\
g_2 &= [-abcu-acv+abw+1 : bcu+abcv+abw+1 :\\
&\qquad bcu-acv-abcw+1],\\
g_3 &= [-abcu+acv-abw+1 : bcu-abcv-abw+1 :\\
&\qquad bcu+acv+abcw+1].
\end{aligned}$$

**Proof.** We check that $g_0$ as defined is incident with the Sight line $\langle 0 : u-aw : -u+av\rangle$ by computing

$$\begin{aligned}
&[-abcu+acv+abw-1 : bcu-abcv+abw-1 :\\
&\qquad bcu+acv-abcw-1]\,\langle 0 : u-aw : -u+av\rangle\\
&= \left[-a\left(v-w+abw^2-acv^2-buw+cuv\right)\right] = 0
\end{aligned}$$

using the quadratic relations and (22). The computations for the other Sight lines and $g_1, g_2, g_3$ are similar. □

An **InGergonne line** is the meet of a corresponding Incenter $i$ and Gergonne point $g$. An **InGergonne point** is the join of a corresponding Inline and Gergonne line $G$. The four InGergonne lines are

$$\begin{aligned}
g_0 i_0 &= \langle v-w-bcuv+bcuw : w-u+acuv-acvw :\\
&\qquad u-v-abuw+abvw\rangle,\\
g_1 i_1 &= \langle w-v+bcuv-bcuw : -w-u-acuv-acvw :\\
&\qquad u+v+abuw+abvw\rangle,\\
g_2 i_2 &= \langle v+w+bcuv+bcuw : u-w-acuv+acvw :\\
&\qquad -v-u-abuw-abvw\rangle,\\
g_3 i_3 &= \langle v+w+bcuv+bcuw : -u-w-acuv-acvw :\\
&\qquad u-v-abuw+abvw\rangle.
\end{aligned}$$

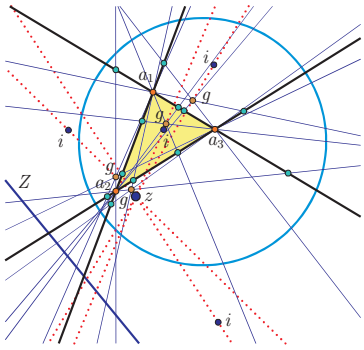



Figure 21: *Sight lines, Gergonne points $g$, InGergonne lines and the $z$ point*

**Theorem 28 (InGergonne center)** *The four InGergonne lines are concurrent, and meet at the z point.*

**Proof.** We check that $g_0i_0$ passes through $z = [a+1 : b+1 : c+1]$ by computing

$$\begin{aligned}
&[a+1 : b+1 : c+1] \cdot \\
&\langle v-w-bcuv+bcuw : w-u+acuv-acvw : u-v-abuw+abvw\rangle \\
&= [av-bu-aw+cu+bw-cv-abuw+acuv+abvw \\
&\quad -bcuv-acvw+bcuw] \\
&= (a-b)(cuv-w)+(c-a)(buw-v)+(b-c)(avw-u) = 0
\end{aligned}$$

where we have used the relations (22). Similarly $g_1i_1$, $g_2i_2$, $g_3i_3$ also pass through the $z$ point. □

**Theorem 29 (Nagel points)** *The following triples of Sight lines are concurrent. Each triple involves one Sight line associated to each of the Incenters, and so is associated to the Incenter with which it does not share a Sight line:*

$$\begin{aligned}
&\langle 0 : -u-aw : u+av\rangle, \quad \langle -v-bw : 0 : v+bu\rangle, \\
&\quad \langle -w-cv : w+cu : 0\rangle \quad \text{to} \quad i_0 = [u : v : w], \\
&\langle 0 : u-aw : -u+av\rangle, \quad \langle v+bw : 0 : -v+bu\rangle, \\
&\quad \langle w+cv : -w+cu : 0\rangle \quad \text{to} \quad i_1 = [-u : v : w], \\
&\langle 0 : u+aw : -u+av\rangle, \quad \langle v-bw : 0 : -v+bu\rangle, \\
&\quad \langle w-cv : -w-cu : 0\rangle \quad \text{to} \quad i_2 = [u : -v : w], \\
&\langle 0 : u-aw : -u-av\rangle, \quad \langle v-bw : 0 : -v-bu\rangle, \\
&\quad \langle w-cv : -w+cu : 0\rangle \quad \text{to} \quad i_3 = [u : v : -w].
\end{aligned}$$

*The points where these triples meet are the* ***Nagel points***

$$\begin{aligned}
n_0 &= [abcu+acv+abw+1 : bcu+abcv+abw+1 : \\
&\quad bcu+acv+abcw+1], \\
n_1 &= [abcu-acv-abw+1 : bcu-abcv-abw+1 : \\
&\quad bcu-acv-abcw+1], \\
n_2 &= [-abcu+acv-abw+1 : -bcu+abcv-abw+1 : \\
&\quad -bcu+acv-abcw+1], \\
n_3 &= [-abcu-acv+abw+1 : -bcu-abcv+abw+1 : \\
&\quad -bcu-acv+abcw+1].
\end{aligned}$$

**Proof.** We check that $n_0$ as defined is incident with $\langle 0 : -u-aw : u+av\rangle$ by computing

$$\begin{aligned}
&[abcu+acv+abw+1 : bcu+abcv+abw+1 : \\
&\quad bcu+acv+abcw+1]\,\langle 0 : -u-aw : u+av\rangle \\
&= \left[a\left(v-w-abw^2+acv^2-buw+cuv\right)\right] = 0
\end{aligned}$$

using the quadratic relations and (22). The computations for the other Sight lines and $n_1, n_2, n_3$ are similar. □

The joins of Incenters $i$ and corresponding Nagel points $n$ are the **InNagel lines**. They are

$$\begin{aligned}
n_0i_0 &= \langle w-v-bcuv+bcuw : u-w+acuv-acvw : \\
&\quad v-u-abuw+abvw\rangle, \\
n_1i_1 &= \langle v-w+bcuv-bcuw : w+u-acuv-acvw : \\
&\quad -u-v+abuw+abvw\rangle, \\
n_2i_2 &= \langle -v-w+bcuv+bcuw : -u+w-acuv+acvw : \\
&\quad v+u-abuw-abvw\rangle, \\
n_3i_3 &= \langle -v-w+bcuv+bcuw : u+w-acuv-acvw : \\
&\quad -u+v-abuw+abvw\rangle.
\end{aligned}$$

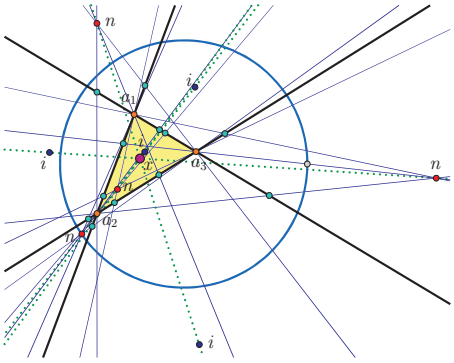



Figure 22: *Nagel points n, InNagel lines and the x point*

**Theorem 30 (InNagel center)** *The four InNagel lines are concurrent, and meet at the x point.*

**Proof.** We check that $n_0i_0$ passes through $x = [a-1 : b-1 : c-1]$ by computing

$$\begin{aligned}
&[a-1 : b-1 : c-1]\,\langle w-v-bcuv+bcuw : \\
&\quad u-w+acuv-acvw : v-u-abuw+abvw\rangle \\
&= [bu-av+aw-cu-bw+cv+abuw-acuv-abvw+ \\
&\quad +bcuv+acvw-bcuw] = 0
\end{aligned}$$

as in the proof of the InGergonne center theorem. The other incidences are similar. □

The joins of corresponding Gergonne points and Nagel points are the **Gergonne-Nagel lines**. They are

$$\begin{aligned}
g_0n_0 &= \langle a(bc-1)(bw-cv) : b(ac-1)(cu-aw) : \\
&\quad c(ab-1)(av-bu)\rangle, \\
g_1n_1 &= \langle -a(bc-1)(bw-cv) : b(ac-1)(cu+aw) : \\
&\quad -c(ab-1)(av+bu)\rangle, \\
g_2n_2 &= \langle -a(bc-1)(bw+cv) : -b(ac-1)(cu-aw) : \\
&\quad c(ab-1)(av+bu)\rangle, \\
g_3n_3 &= \langle a(bc-1)(bw+cv) : -b(ac-1)(cu+aw) : \\
&\quad -c(ab-1)(av-bu)\rangle.
\end{aligned}$$

**Theorem 31 (Gergonne-Nagel center)** *The four Gergonne-Nagel lines are concurrent, and meet at the* ***Gergonne-Nagel center****, or u* ***point,*** *which is*

$$u = [(ac-1)(ab-1) : (bc-1)(ab-1) : (bc-1)(ac-1)].$$

**Proof.** We may check directly that the $u$ point defined as above does indeed lie on each Gergonne-Nagel line: this does not require use of the quadratic or cubic relations. □

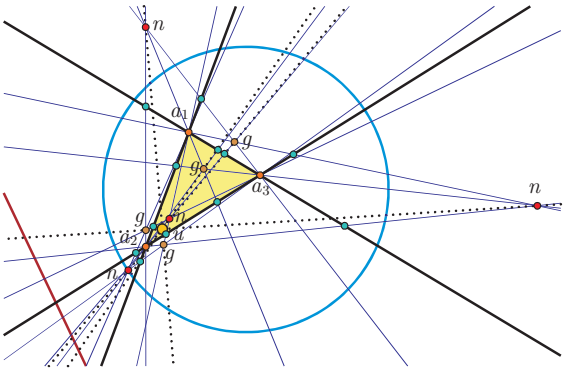



Figure 23: *Gergonne-Nagel lines and the Gergonne-Nagel center: the u-point*

We leave it as an exercise for the reader to establish that the $u$ point lies on the Orthoaxis precisely when the numbers $a, b, c$ are not distinct.

## 5 Circumcenter Hierarchy

There is a fundamental duality between the Incenter and Circumcenter hierarchies, since from (6) a point $m$ is a midpoint of a side $\overline{ab}$ precisely when its dual line $M = m^{\perp}$ is a biline of the dual vertex $\overline{a^{\perp}b^{\perp}}$. So by dualizing we can transform all known facts about the Incenter hierarchy of a triangle to the Circumcenter hierarchy of the dual triangle, and vice versa. This is a striking difference between projective Triangle geometry and the more familiar Euclidean version, and sheds also some light on the latter.

*All the results of this section are consequences of the dual results established in the previous section, after some book-keeping between the two hierarchies.*

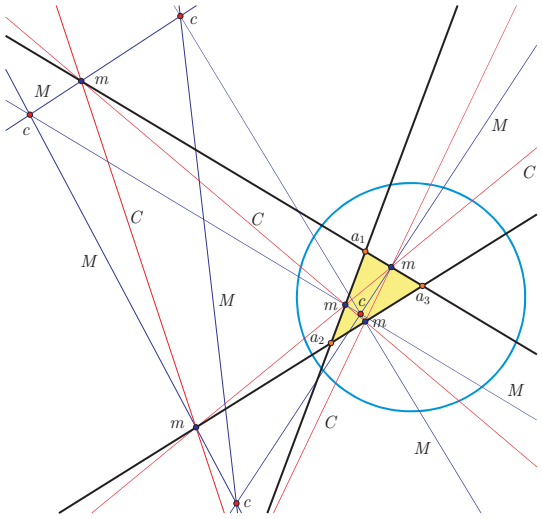



Figure 24: *Midpoints m, Midlines M, Circumcenters c, Circumlines C*

We will now assume that the triangle $\overline{a_1a_2a_3}$ has **Midpoints** $m$, and so also **Midlines** $M$. This occurs precisely when $1-q_1, 1-q_2, 1-q_3$ are all squares, and in this case there are six Midpoints, two on each side. This is equivalent to the Dual triangle $\overline{l_1l_2l_3}$ having bilines. Each Midline $M$ passes through a Midpoint $m$, since any two Midpoints of a side are perpendicular.

**Theorem 32 (Circumlines)** *Midpoints m are collinear in threes, joining on four* ***Circumlines*** *C. Midlines M are concurrent in threes, meeting at four* ***Circumcenters*** *c.*

**Median lines** (or just **medians**) $D$ are joins of corresponding Midpoints $m$ and Points $a$. There are six Medians, two passing through each Point. The duals are the **Median points** $d$, the meets of corresponding Midlines $M$ and Dual lines $A$.

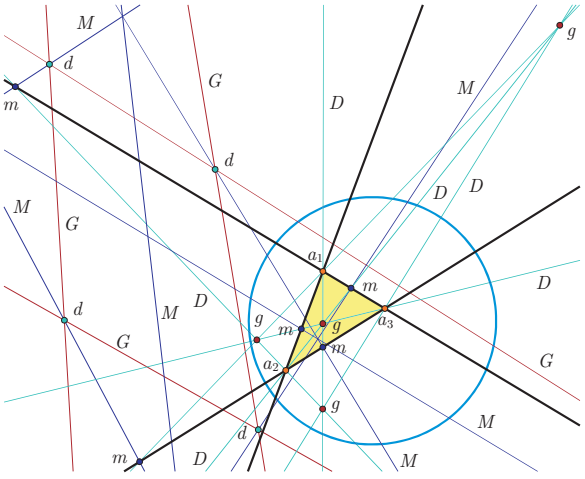



Figure 25: *Medians D, Median points d, Centroids g, Centroid lines G*

**Theorem 33 (Median harmonic conjugates)** *The two Median lines through a vertex of the Triangle are harmonic conjugates with respect to the two Lines of that vertex.*

**Theorem 34 (Centroids)** *The Median lines D are concurrent in threes, meeting at four* ***Centroid points*** *g. The Median points d are collinear in threes, joining on four* ***Centroid lines*** *G.*

A **Median Thaloid** is a Thaloid of a side consisting of two Median points, both on a Dual line of the Triangle. There are three Median Thaloids.

**Theorem 35 (Isostatic points)** *If two Median Thaloids meet at a point r, then the third does too.*

Such a common point is an **Isostatic point**; Figure 26 shows the three Median Thaloids as well as two Isostatic points: $r_1$ and $r_2$, and their join.

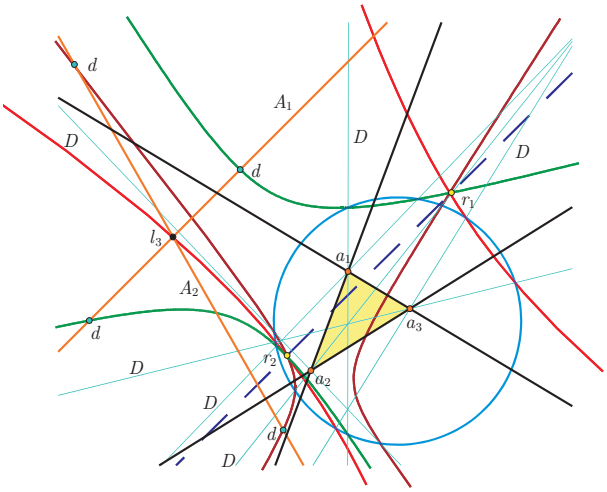



Figure 26: *Median Thaloids and Isostatic points $r_1$ and $r_2$*

The four Circumlines and the four Centroid lines are corresponding, since the three Midpoints that join in a Circumline also give rise to the three Median lines passing through a particular Centroid line, which is dual to a particular Centroid line. **CircumCentroid points** are meets of corresponding Circumlines and Centroid lines. **CircumCentroid lines** are partially analogous to *Euler lines*, being joins of Circumcenters and Centroids. The next result shows that the $z$ point might also be called the **Euler center**!

**Theorem 36 (CircumCentroid axis)** *The four CircumCentroid points are collinear, and join on the Z line. The four CircumCentroid lines are concurrent, and meet at the z point.*

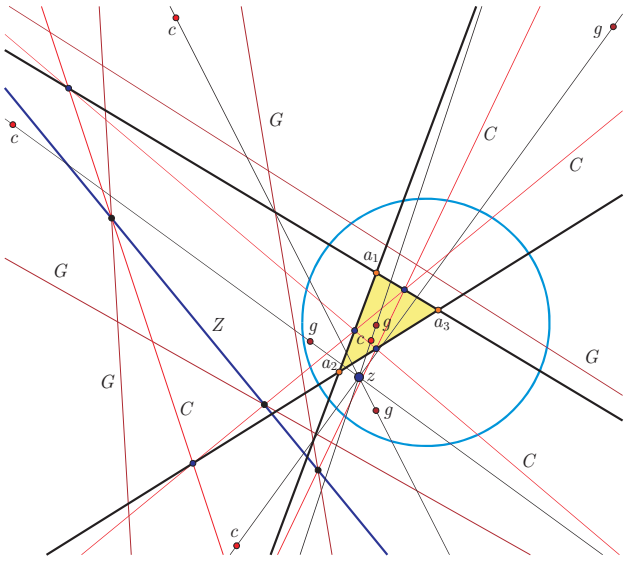



Figure 27: *CircumCentroids, the Z line and the z point*

**CircumDual points** are meets of Circumlines and Dual lines. There are twelve CircumDual points, four on each Dual line, three on each Circumline. **CircumDual lines** are duals of CircumDual points. **Tangent lines** are joins of corresponding CircumDual points and Points; there are twelve. Tangent points are duals of Tangent lines.

A **Circumcircle** is a circle centered at a Circumcenter passing through one, hence all of the Points. These are shown in Figure 28, with CircumDual points and Tangent lines, which really are tangent to the Circumcircles.

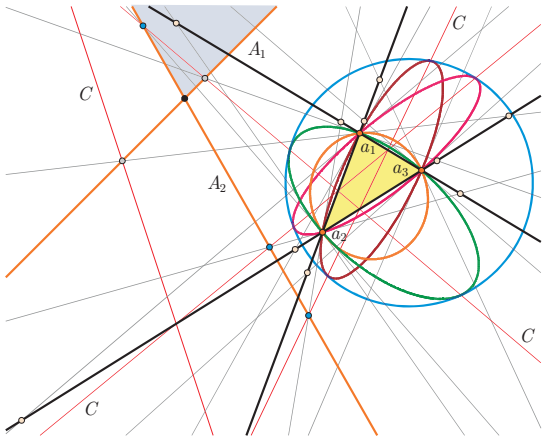



Figure 28: *CircumDuals points, Tangent lines, Circumcircles and Sound points*

**Sound points** are meets of Tangent lines and Lines; there are twelve, three associated to each Circumline. They are also shown in Figure 28. **Sound lines** are the duals of Sound points.

**Theorem 37 (Jay lines)** *The three Sound points associated to a particular Circumline C join on a* ***Jay line*** *J.*

There are four Jay lines, and **Jay points** *j* are their duals. **CircumJay points** are meets of Circumlines and associated Jay lines; there are four, and **CircumJay lines** are their duals; joins of Circumcenters and Jay points.

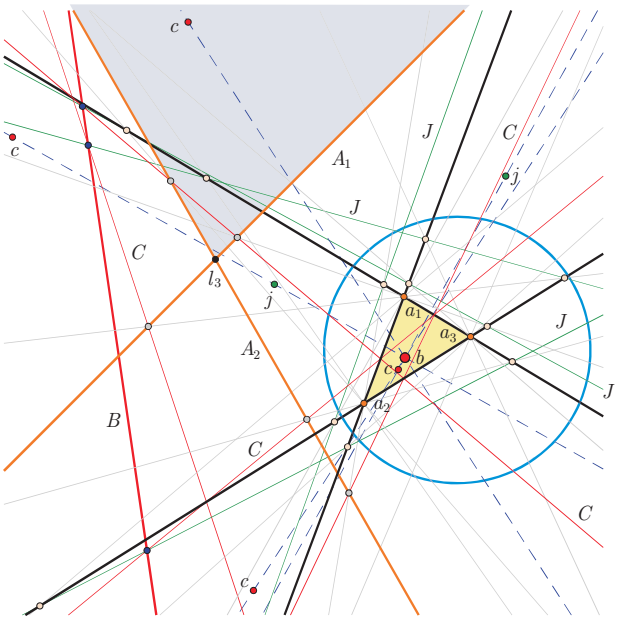



Figure 29: *Jay points and lines, CircumJay points and lines, and the Base center b*

**Theorem 38 (CircumJay center)** *The four CircumJay points join on the Base axis B. The four CircumJay lines meet at the Base center b.*

**Theorem 39 (Wren lines)** *Sound points associated to different Circumlines are collinear in threes, and join on four* ***Wren lines*** *W.*

The duals of the Wren lines are **Wren points** *w*. To each Wren line we associate the Circumline not associated to the Sound points on it. **CircumWren points** are the meets of Circumlines and associated Wren lines, and **CircumWren lines** are their duals.

**Theorem 40 (CircumWren center)** *The four CircumWren points join on the Orthic axis S. The four CircumWren lines meet at the Orthostar s.*

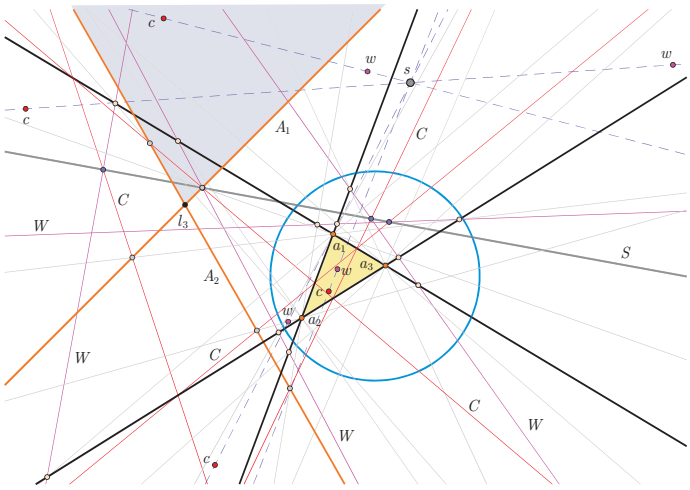



Figure 30: *Wren points and lines, CircumWren points and lines, and the Orthostar s*

A **JayWren point** is the meet of associated Jay lines and Wren lines; there are four. The duals are the **JayWren lines**.

**Theorem 41 (JayWren center)** *The four JayWren points join on the* ***JayWren axis****, or the V* ***line****. The four JayWren lines meet at the* ***JayWren center****, or v* ***point****.*

This is a good point to remark that in the projective situation there are remarkable additional constructions, that are available at times when midpoints and bilines may not exist, which allow a wide extension of many of the theorems in this paper. This topic will be developed elsewhere.

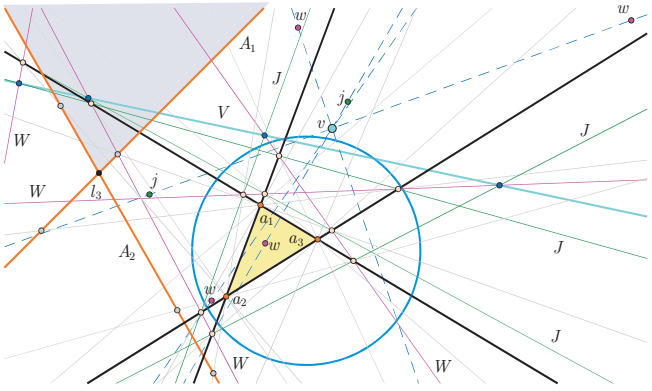



Figure 31: *JayWren points and lines, JayWren center v and axis V*

## 6 Bridging between the Incenter and Circumcenter hierarchies

Although we have so far emphasized the complete duality between the Incenter and Circumcenter hierarchies, it is also the case that there are numerous remarkable connections between the two. We give a brief indication of this with three examples, leaving proofs to another occasion.

We assume we have a (generic) triangle $\overline{a_1a_2a_3}$ with *both Bilines and Midpoints* (so both hierarchies exist). This is, at least approximately, the situation with our example Triangle.

**Theorem 42** *The JayWren center v, the Gergonne-Nagel center u and the Orthocenter h are collinear.*

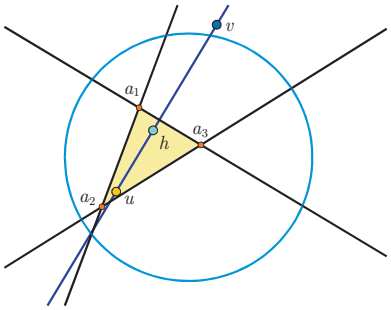



Figure 32: *Jay-Wren-Gergonne-Nagel axis*

**Theorem 43 (InCirc joins/meets)** *The* 16 ***InCirc joins*** *joining the four Incenters i and the four Circumcenters c meet also four at a time at four* ***InCirc centers*** *r.*

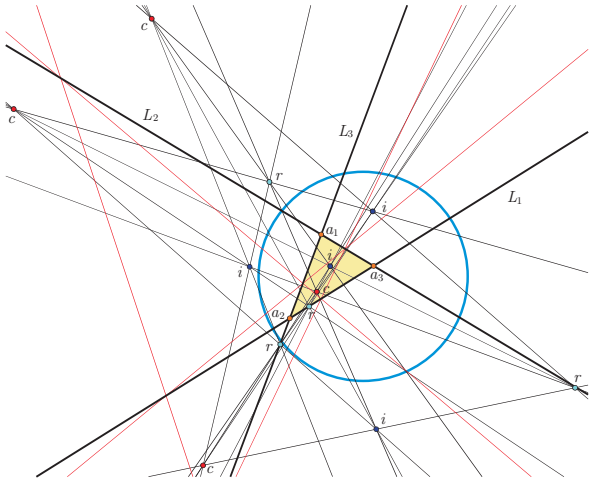



Figure 33: *InCirc joins and InCirc centers*

The four Incenters *i*, four Circumcenters *c* and four InCirc centers *r* form a pleasant symmetrical configuration of 12 points.

**Theorem 44 (InCentroid joins/meets)** *The* 16 ***InCentroid joins*** *joining the four Incenters i and the four Centroid points g meet two at a time at* 24 ***InCentroid meets*** *which lie,* 8 *each, on the Lines.*

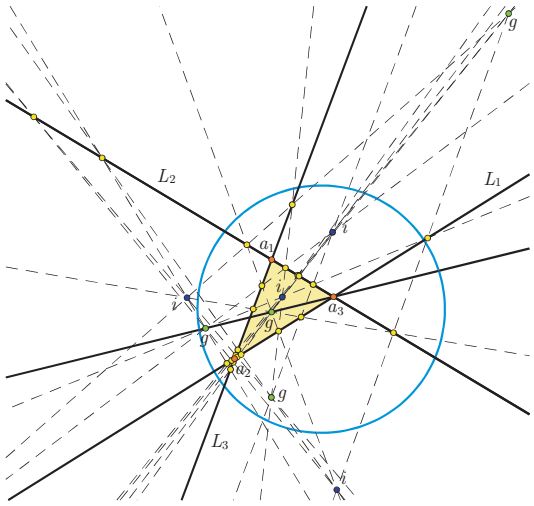



Figure 34: *InCentroid lines and meets*

An interesting direction is to ponder the implications of this work for classical Euclidean triangle geometry. There are also many further phenomenon in the projective setting with no obvious affine/Euclidean parallel, which will be studied in future papers. The author will shortly post Triangle Geometry GSP worksheets on his UNSW website. Hopefully these, together with the formulas and pictures in this paper, will empower and encourage the reader to make his/her own discoveries!

## References

[1] W. BENZ, *Classical Geometries in Modern Contexts: Geometry of Real Inner Product Spaces*, Birkhäuser, Basel, 2007.

[2] O. BOTTEMA, On the medians of a triangle in hyperbolic geometry, *Can. J. Math.* **10** (1958), 502–506.

[3] H. BRAUNER, *Geometrie Projektiver Räume I, II*, Bibliographisches Institut, Mannheim, 1976.

[4] H. S. M. COXETER, *Non-Euclidean Geometry*, 6th ed., Mathematical Association of America, Washington D. C., 1998.

[5] O. DEMIREL and E. SOYTURK, The hyperbolic Carnot theorem in the Poincaré disk model of hyperbolic geometry, *Novi Sad J. Math.* **38**(2) (2008), 33–39.

[6] M. J. GREENBERG, *Euclidean and Non-Euclidean Geometries: Development and History*, 4th ed., W. H. Freeman and Co., San Francisco, 2007.

[7] R. HARTSHORNE, *Geometry: Euclid and Beyond*, Springer, New York, 2000.

[8] N. IVANOV, Arnol'd, the Jacobi Identity and Orthocenters, *Amer. Math. Monthly*, **118**(1) (2011), 41–65.

[9] C. KIMBERLING, *Triangle Centers and Central Triangles*, vol. 129 Congressus Numerantium, Utilitas Mathematica Publishing, Winnepeg, MA, 1998.

[10] C. KIMBERLING, *Encyclopedia of Triangle Centers*, `http://faculty.evansville.edu/ck6/encyclopedia/ETC.html`.

[11] A. PRÉKOPA and E. MOLNAR, eds., *Non-Euclidean Geometries: János Bolyai Memorial Volume*, Springer, New York, 2005.

[12] N. SONMEZ, A trigonometric proof of the Euler theorem in hyperbolic geometry, *Algebras Groups Geom.* **26**(1) (2009), 75–79.

[13] A. A. UNGAR, Hyperbolic barycentric coordinates, *Aust. J. Math. Anal. Appl.* **6**(1) (2009), 1–35.

[14] A. A. UNGAR, *Hyperbolic Triangle Centers: The Special Relativistic Approach*, FTP 166, Springer Dordrecht, 2010.

[15] N. J. WILDBERGER, *Divine Proportions: Rational Trigonometry to Universal Geometry*, Wild Egg Books, Sydney, 2005. `http://wildegg.com`.

[16] N. J. WILDBERGER, Affine and projective metrical geometry, *arXiv: math/0612499v1*, 2006, to appear, *J. of Geometry*.

[17] N. J. WILDBERGER, Chromogeometry and Relativistic Conics, *KoG* **13** (2009), 43–50.

[18] N. J. WILDBERGER, Universal Hyperbolic Geometry I: Trigonometry, *arXiv: math/0909.1377v1,* 2009.

[19] N. J. WILDBERGER, Universal Hyperbolic Geometry II: A pictorial overview, *KoG* **14** (2010), 3–24.

[20] N. J. WILDBERGER, *UnivHypGeom*: `http://www.youtube.com/playlist?list=PL6ACFCC19EA82CA71&feature=plcp`

[21] H. E. WOLFE, *Introduction to Non-Euclidean Geometry*, Holt, Rinehart and Winston, New York, 1945.

**Norman John Wildberger**
School of Mathematics and Statistics UNSW
Sydney 2052 Australia
e-mail: n.wildberger@unsw.edu.au

Original scientific paper
Accepted 23. 12. 2011.

GÜNTER WALLNER
FRANZ GRUBER

# Interactive Modeling and Subdivision of Flexible Equilateral Triangular Mechanisms

**Interactive Modeling and Subdivision of Flexible Equilateral Triangular Mechanisms**

**ABSTRACT**

Based on the requests from architects, we developed a system which allows to interactively design and subdivide flexible triangular surfaces. Due to economical reasons the number of different types of building elements should be small. For that reason we only use equilateral base triangles of unique size with the possibility of subdivision. To allow to interactively move vertices and to ensure constant edge length we use force-directed methods instead of inverse kinematics. This paper describes the data structure, the algorithm and the influence of subdivision on the kinematic flexibility of the mesh.

**Key words:** subdivision, uniform 1-to-4 split, flexible mechanism, force directed algorithm

**MSC 2000:** 00A67, 65D17

**Interaktivno modeliranje i razdioba fleksibilnog mehanizma jednakostraničnih trokuta**

**SAŽETAK**

Prema zahtjevima arhitekata razvijamo sustav koji dozvoljava interaktivnu tvorbu i razdiobu fleksibilne triangulirane plohe. Broj različitih sastavnih dijelova iz ekonomskih razloga treba biti mali. Zbog toga koristimo samo sukladne jednakostranične bazne trokute s mogućnošću razdiobe. Kako bismo dozvolili interaktivno kretanje vrhova i osigurali konstantnost duljine bridova umjesto inverzne kinematike koristimo metodu upravljanja silom. Rad opisuje strukturu podataka, algoritam i utjecaj razdiobe na kinematičku fleksibilnost mreže.

**Ključne riječi:** razdioba, uniformna 1-4 podjela, fleksibilni mahanizam, algoritam upravljanja silom

## 1 Introduction

In order to design the booth of the University of Applied Arts Vienna at the Vienna Fair 2011[1] (see Figure 1), the Department of Geometry was asked how to interactively model a surface consisting of equilateral triangles of different sizes. This surface, covered with sound-absorbing material (see Figure 2), was floating above the booth like a cloud and was therefore called *Acoustic Cloud* by responsible architect Juliana Herrero. For the arrangement of the triangles she required a tool which allows her to individually subdivide triangles of the mesh and to interactively move vertices while keeping the side lengths of the triangles constant. This yields a triangular mechanism whose kinematic behavior – to our knowledge – can not be simulated with existing software packages. Just as well, the theoretical solution of the task is almost impossible. Therefore, we developed a tool which attempts a numerical solution by using force-directed methods. In this paper we present the algorithm and discuss the kinematic restrictions which come along with the subdivision of the surface. Depending on the subdivision levels of adjacent triangles subdivision can enhance flexibility or not, as can be seen from the example depicted in Figure 6 and Figure 9.

Before we will describe our algorithm in Section 2 we will review related work. However, since this paper is mainly concerned with the practical application of a force-directed algorithm to a kinematic problem, we will restrict us to a short overview. Usually, force directed algorithms are associated with the drawing of graphs, where they are used to position the nodes of the graph in an aesthetically pleasing way by assigning forces to edges and nodes. Historically, the work of Tutte [9] can be considered as one of the first force-directed algorithms. Tutte only used springs of ideal length zero and no repulsive forces. Eades [3] and Fruchterman and Reingold [4] used spring forces, similar to those in Hooke's law. Both methods apply repulsive forces between all nodes and attractive forces to nodes connected by edges. However, force-directed algorithms have been applied to other domains as well.

[1] The Vienna Fair is an international contemporary art fair focusing on Central and Eastern Europe. `http://www.viennafair.at`

Figure 1: *The booth of the University of Applied Arts Vienna at the Vienna Fair 2011. (Photo by Virgil Widrich)*

Figure 2: *A close-up of the* Acoustic Cloud. *(Photo by Virgil Widrich)*

Examples include the work of Quinn and Breuer [8] who describe a force-directed method to place components on printed circuit boards. Provot [7] used a force-directed algorithm to simulate the behavior of cloth by approximating it by a deformable surface composed of a network of masses and springs. Djidjev [1] published a force-directed method to smooth unstructured triangular and tetrahedral meshes. Recently, Gruber et al. [5] described a method which optimizes given grids with rectangular topology on an arbitrary parametric double-curved surface in regard to orthogonality and, optionally, locally almost constant grid size.

## 2 Algorithm

### 2.1 Basic Data Structure

In the following we consider an initial equilateral triangular mesh $M = (V,T)$ where $V$ is a set of vertices and $T$ is a set of equilateral non-subdivided (Level 0) triangles with side-length $a$. At this point we should stress that equilaterality is not a necessity and the algorithm can be easily adapted to arbitrary triangular meshes. Additionally a set of edges $E$ is stored. Basically, a vertex $v$ is defined by its position in $\mathbb{R}^3$ and an edge $e = (v_0, v_1)$ connects vertex $v_0$ with $v_1$. An edge also stores references to its adjacent triangles $n_0$ and $n_1$. Note, that a triangle is only considered adjacent if both vertices $v_0$ and $v_1$ are corners of the triangle.

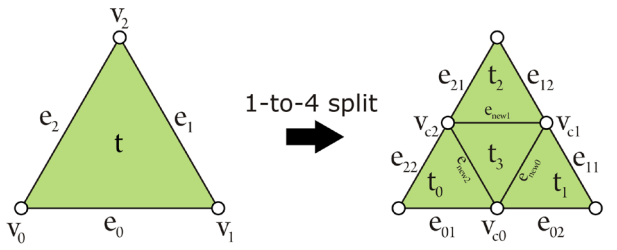



Figure 3: *Subdivision of a triangle with a 1-to-4 split.*

A triangle $t = (v_0, v_1, v_2)$ is defined by the three corners $v_0$, $v_1$ and $v_2$. Furthermore, references to the three edges $e_0 = (v_0, v_1)$, $e_1 = (v_1, v_2)$ and $e_2 = (v_2, v_0)$ are stored. These data structures are shown in Appendix A in Listing 1, Listing 2 and Listing 3 along with further variables which will be described later.

### 2.2 Subdivision

For subdivision we use a uniform 1-to-4 split, which divides a triangle $t^i$ at subdivision level $i$ into four triangles $t_0^{i+1}$, $t_1^{i+1}$, $t_2^{i+1}$ and $t_3^{i+1}$ as depicted in Figure 3. The 1-to-4 split is a commonly used method for subdivision and remeshing of triangles, e.g., Loop subdivision [6] and Butterfly subdivision [2] both use 1-to-4 refinement. To recursively traverse the hierarchy it is necessary that $t^i$ stores pointers to these four child triangles which in turn have to store a pointer to their parent $t^i$. The level itself is also stored in the data structure of the triangle.

A uniform 1-to-4 split introduces new vertices $v_{c0}$, $v_{c1}$ and $v_{c2}$ at the center of the edge $e_0^i$, $e_1^i$ and $e_2^i$ of $t^i$. These vertices are only added to $V$ if a vertex with the same coordinates does not already exist. We will refer to $e_j^i$, $j \in \{0,1,2\}$ as the *support edge* of $v_{cj}$ and to $v_{cj}$ as the *center vertex* of $e_j^i$.

Further, each edge $e_j^i$, $j \in \{0,1,2\}$ will be divided into two parts $e_{j1}^{i+1}$ and $e_{j2}^{i+1}$. $e_j^i$ stores references to these two child edges which in turn store a pointer to their parent $e_j^i$. Also, $e_j^i$ stores a pointer to the corresponding center vertex $v_{cj}$. These six edges are added to $E$ along with the three new edges $e_{new0}^{i+1}$, $e_{new1}^{i+1}$ and $e_{new2}^{i+1}$ of $t_3^{i+1}$. In the following we

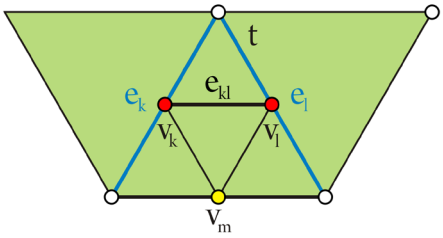



Figure 4: *Two inflexible center vertices $v_k, v_l$ (red) of a subdivided triangle $t$ induce a further inflexible vertex $v_m$ (yellow).*

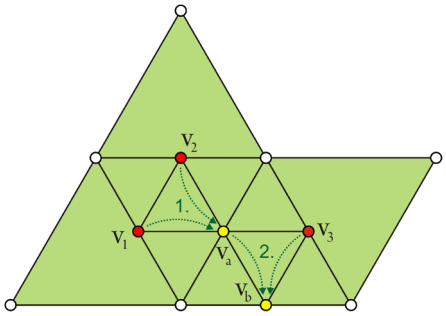



Figure 5: *An example for propagation of inflexibility: 1. the two inflexible vertices $v_1$ and $v_2$ will cause $v_a$ to become inflexible. 2. $v_a$ together with the already existing inflexible vertex $v_3$ induce a further inflexible point $v_b$.*

will call an edge $e^0$ with only one neighboring triangle (either $n_0$ or $n_1$) a *border edge*.

### 2.3 Kinematic Restrictions

Different subdivision levels between adjacent triangles restrict the movability of the mechanism in certain ways, i.e, just because an edge has been subdivided it does not mean that it is allowed to fold at its center vertex.

#### 2.3.1 Inflexible Vertices

We will refer to a vertex as *flexible* if the corresponding support edge is allowed to fold, otherwise *inflexible*. All vertices which are corners of a triangle at level 0 are flexible. The flexibility of vertices $v_{cj}$, $j \in \{0,1,2\}$ introduced during subdivision is algorithmically determined by performing the following test. If there already exists a vertex $v^*$ with the same coordinates then $v^*$ is marked as flexible. Otherwise, we check if the corresponding support edge $e_j$ is a border edge. If it is, $v_{cj}$ is flexible, if it is not $v_{cj}$ is inflexible. In other words, a center vertex $v$ with support edge $e$ is inflexible if and only if one of the adjacent triangles of $e$ is subdivided, otherwise $v$ is flexible.

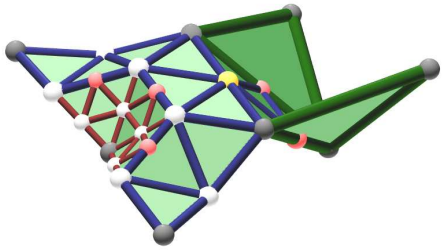


Figure 6: *An example of a 3D mechanism. Different subdivision levels result in inflexible vertices (shown in red and yellow). Such vertices prevent folding of their support edge and therefore restrict the movability of the mechanism in regard to interactive displacement.*

#### 2.3.2 Induced Inflexible Vertices

However, the test in Section does not capture all kinematic restrictions because it does not take into account the configuration of the neighborhood. Certain configurations can cause a propagation of inflexibility through the mesh, in other words, inflexible vertices can induce further inflexible vertices what we therefore call *induced inflexible vertices* in the following. To detect these induced inflexible vertices we run a second test after the above mentioned test.

If a triangle $t$ is subdivided and two of the three center vertices $v_k$ and $v_l$ of its edges are inflexible (as shown in Figure 4) then the third center vertex $v_m$ also becomes inflexible because of the following reason. If $v_k$ is inflexible then the corresponding support edge $e_k$ is not allowed to fold, the same holds true for $v_l$ and its support edge $e_l$. Therefore the edge $e_{kl}$ between $v_k$ and $v_l$ is also rigid, even if the sub-triangles of $t$ are further subdivided. $e_k$ and $e_l$ together with $e_{kl}$ as a distance piece between them make the triangle $t$ rigid.

The introduction of an induced inflexible vertex can lead to further induced inflexible vertices. A simple example is shown in Figure 5. To handle this possible series of reactions, the test has to be repeated until no further induced inflexible vertices are found. Figure 6 shows a 3D example of a surface with different subdivision levels and the different types of inflexible vertices.

### 2.4 Force-directed Placement

An important feature of the tool is that the surface can be interactively modeled by moving vertices. Since a theoretical calculation with inverse kinematic of such a mechanism is almost impossible, we attempt a numerical solution with force-directed methods. The iterative force-directed algorithm is responsible for arranging the vertices in a way that

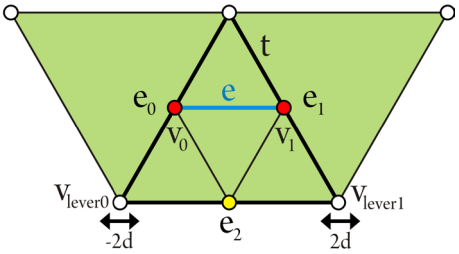



Figure 7: *In order to keep the distance between two inflexible vertices $v_0$ and $v_1$ constant, their displacements have to be applied twice to the corresponding vertices $v_{lever0}$ and $v_{lever1}$ instead to $v_0$ and $v_1$.*

guarantees that edges keep their predefined ideal length. In case of equilateral triangles the $idealLength = 0.5^i \cdot a$, where $i$ is the subdivision level of the edge and $a$ is the side length of a level 0 triangle. However, the set of edges $E$ contains more information than necessary for the simulation. Therefore, we derive a minimal set of edges $F \subseteq E$ which completely describes the flexible mechanism.

To determine the edges of $F$, we initially loop only through edges $e$ which do not have a parent[2]. In order to decide which parts of the edge hierarchy of $e$ belong to $F$, we start the following recursive process. If $e$ does not have any children, the edge itself belongs to $F$. Otherwise, we check if the center vertex of $e$ is inflexible in which case $e$ is also part of $F$. However, if the center vertex is flexible then the process is repeated recursively for $e = e_{c0}$ and $e = e_{c1}$ where $e_{c0}$ and $e_{c1}$ are the two child edges of $e$.

For each vertex $v$ a displacement vector $v^{disp}$ is stored which is set to zero at the beginning of each iteration. Afterward, the algorithm loops through each edge $e = (v_0, v_1) \in F$ and calculates the deviation $\Delta$ from the actual edge length $\Delta = e_{ideal} - \|v_{01}\|$ where $v_{01} = v_1^{pos} - v_0^{pos}$. If $v_0$ is flexible then $d := \varepsilon \cdot \Delta \cdot \frac{v_{01}}{\|v_{01}\|}$ is subtracted from $v_0^{disp}$, where $\varepsilon > 0$ is a small constant. Similarly $d$ is added to $v_1^{disp}$ if $v_1$ is flexible.

No displacement is added to inflexible points because their position has to remain at the center of their rigid support edge. This calculation will be performed later. However, if both vertices of an edge are inflexible, as depicted in Figure 7, a problem arises because the edge would not have the effect of a distance piece anymore. We solved the problem by transferring the edges' distance keeping force onto other suitable vertices. Let $e_0$ be the support edge of $v_0$, $e_1$ the support edge of $v_1$ and $t$ the triangle which contains $e_0$ and $e_1$. Further, let $e_2$ be the third edge of $t$. Then the displacements caused by $e$ will be applied twice to the corresponding vertices $v_{lever0}$ and $v_{lever1}$ of $e_2$.

[2]Note, that this are not only edges on subdivision level 0.

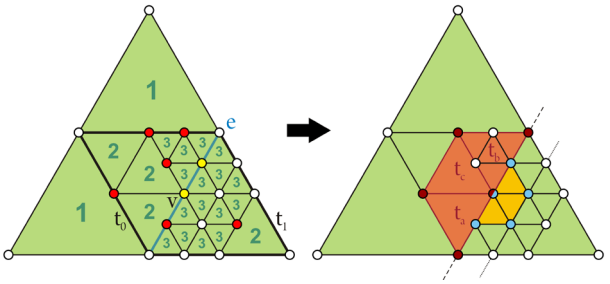



Figure 8: *An example for the planarity condition of an (induced) inflexible vertex v with support edge e and adjacent triangles $t_0$ and $t_1$: The left image shows the subdivision level of triangles and the flexibility of vertices. The right image shows the trapezoids implied by v for $t_0$ (red) and $t_1$ (orange). The corners of each separate trapezoid together with v have to be coplanar. The different sizes of the trapezoids are a result from the different subdivisions at both sides of e. The smaller orange trapezoid allows folding the mesh at the dotted line, which would not be the case if the three adjacent level 2 triangles would have been chosen for the trapezoid.*

Once all edges have been processed the aggregated displacement of each vertex is added to its position. Finally, the calculation of the position of inflexible vertices is performed. This calculation has to be carried out in a hierarchical top to down manner to make sure that the center vertex of the parent of an edge $f$ is already at the right position before calculating the position of a possible center vertex of $f$.

These steps are summarized in Listing 4 in Appendix A.

### 2.4.1 Planarity Condition and Convergence Speed

Because the algorithm described above only takes edge lengths and not planarity into account it will converge slowly to the solution. However, an (induced) inflexible vertex will not only make its support edge rigid but will also make parts of the surface planar. We take advantage of this theoretical fact to improve the converge speed of the algorithm. Let us assume that $v$ is an (induced) inflexible vertex with $e$ being its support edge and let $t_0^i$ and $t_1^i$ be the adjacent triangles of $e$ at level $i$. In the following let us consider $t_0^i$ to be subdivided (the same process is performed for $t_1^i$ if it is subdivided). Figure 8 shows a possible configuration. Inside the hierarchy of $t_0^i$ we search for three triangles incident to $v$ with the following properties:

1. All three triangles are on the same subdivision level $m$
2. There is no higher level $k > m$ which fulfills condition (1), i.e., $m$ is maximal

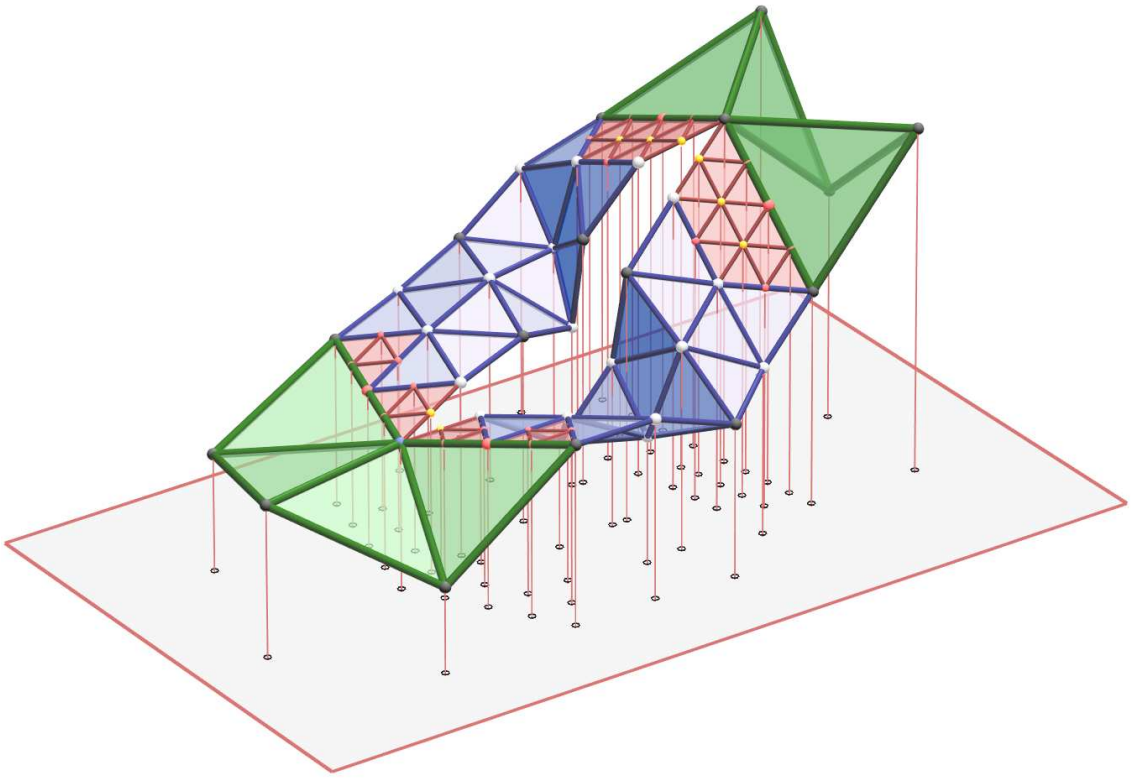


Figure 9: *An example of a mesh with subdivisions up to level 2 which was designed with the described tool (level 0 = green, level 1 = blue and level 2 = red). Whereas triangles on subdivision level 0 and 1 have high flexibility, the level 2 triangles – as chosen in this example – have no further influence on the kinematic behavior of the mechanism. Inflexible vertices are shown in red and induced inflexible vertices are colored yellow.*

All vertices of these three triangles have to be coplanar, because the support edge $e$ of $v$ can be thought of as a frame joint for the two triangles $t_a$ and $t_b$ as seen in Figure 8. Together with the remaining triangle $t_c$ they form an isosceles trapezoid. $v$ stores these vertices in an array called trapezoidVertices0 in the data structure. After displacement of the vertices in the force-directed algorithm, the vertices of this array are orthogonally projected onto their regression plane. This procedure increases the speed of the algorithm considerably.

#### 2.4.2 Termination Condition

The steps outlined above are repeated iteratively until all displacements are below a certain threshold, mathematically $v_{disp} < \varepsilon \quad \forall v \in V$.

## 3 Operations

An important aspect in the development of the system was interactivity. Users should be able to easily work with the mesh in order to facilitate the design process. Therefore all operations can be performed with a single mouse click together with keyboard shortcuts.

The system allows to append edges of side length $a$ to corners of level 0 triangles. If adding a new edge yields a new base triangle (i.e., a triangle at subdivision level 0) then this triangle is automatically added to $T$. Only edges with no adjacent triangles can be removed. Existing triangles – independent of their subdivision level – can be subdivided with a 1-to-4 split, as described in Section 2.2. When deleting a triangle two cases can occur: If a triangle $t^i$ with $i = 0$ (base triangle) is deleted then edges not adjacent to any remaining triangles are deleted automatically as well. In case $i > 0$ all four triangles of the parent triangle from $t^i$, including edges and vertices which are no longer part of a remaining triangle, are deleted in order to guarantee a well-defined subdivision hierarchy. Dragging a vertex with the left (right) mouse button moves it horizontally (vertically). Once the position of a vertex changes the arrangement of the mesh has to be recalculated which is handled by the force-directed algorithm described in Section 2.4.

## 4 Conclusions

In this paper we described a tool to interactively design triangular mechanisms by extending the mesh, subdividing triangles, and moving vertices (Figure 9 shows an example). Although we restricted ourselves – due to the initial artistic concept – to equilateral triangles it should be noted that the algorithm can be easily modified to general triangles.

To simulate the kinematic behavior of the mesh we used an iterative force directed algorithm. As it turned out, subdivisions must not necessarily increase the flexibility, because in many cases inflexibility propagates quite easily. This means that subdivision has to be applied well-considered in order to raise flexibility. From a designer's point of view, this may be very restrictive. In general the flexibility can be improved if one allows small deviations from the ideal edge length. Experiments showed that ignoring the planarity condition (see Section 2.4.1) and allowing for the edge length to deviate only by 1% from its predefined ideal length, increases the flexibility already considerably.

## A Data Structures and Pseudo Code

```
struct sVertex
{
  float[3] position;
  float[3] displacement;

  bool flexible;
  bool inducedInflexible;

  sEdge *supportEdge;

  Array<sVertex*> trapezoidVertices0;
  Array<sVertex*> trapezoidVertices1;
};
```

Listing 1: Vertex data structure

```
struct sEdge
{
  sVertex *v0;
  sVertex *v1;
  sVertex *center;

  sEdge *c0;  // child 0
  sEdge *c1;  // child 1
  sEdge *parent;

  sTriangle* n0;
  sTriangle* n1;

  float idealLength;
  int level;
};
```

Listing 2: Edge data structure

```
struct sTriangle
{
  sVertex* v0;
  sVertex* v1;
  sVertex* v2;

  sEdge* e0;
  sEdge* e1;
  sEdge* e2;

  sTriangle* c0;  // child 0
  sTriangle* c1;  // child 1
  sTriangle* c2;  // child 2
  sTriangle* c3;  // child 3
  sTriangle* parent;

  int level;
};
```

Listing 3: Triangle data structure

```
foreach (sVertex v)
  v.displacement = (0,0,0);

foreach (sEdge e) {
  float[3] l = e.v0.position - e.v1.position;
  float Δ = e.idealLength - l.length();

  l.normalize();

  if (e.v0.flexible == true)
    e.v0.displacement += -ε*Δ*l;
  if (e.v1.flexible == true)
    e.v1.displacement +=  ε*Δ*l;

  if (e.v0.flexible==false &&
      e.v1.flexible==false) {
    // find lever0 and lever1, see Fig. 6
    lever0.displacement += -2*ε*Δ*l;
    lever1.displacement +=  2*ε*Δ*l;
  }
}

foreach (sVertex v)
  v.position += v.displacement;

calcPositionOfInflexiblePoints();

foreach (trapezoid t) {
  Plane regPlane = calcRegressionPlane(t);
  foreach (Vertex v in t.vertices)
    regPlane.orthoProject(v);
}
```

Listing 4: Pseudocode for the force-directed algorithm from Section 2.4

## References

[1] H. DJIDJEV, Force-Directed Methods For Smoothing Unstructured Triangular And Tetrahedral Meshes. *Proceedings of the 9th International Meshing Roundtable* (2000), 395–406

[2] N. DYN, J. GREGORY, D. LEVIN. A BUTTER, A Butterfly Subdivision Scheme for Surface Interpolation with Tension Control. *ACM Trans. Graph.* **9** (1990), 160–169

[3] P. A. EADES, A heuristic for graph drawing. *Congressus Numerantium* **42** (1984), 149–160

[4] T. M. J. FRUCHTERMAN, E. M. REINGOLD, Graph Drawing by Force-directed Placement. *Software - Practice and Experience* **21** (1991), 1129–1164

[5] F. GRUBER, G. WALLNER, G. GLAESER, Force Directed Near-Orthogonal Grid Generation on Surfaces. *Journal for Geometry and Graphics* **14** (2010), 135–145

[6] C. LOOP., Smooth subdivision surfaces based on triangles. *Masters Thesis*, Utah University (1987)

[7] X. PROVOT, Deformation constraints in a mass-spring model to describe rigid cloth behavior. *Graphics Interface* (1995), 147–154

[8] N. QUINN, M. BREUER, A force directed component placement procedure for printed circuit boards. *IEEE Transactions on Circuits and Systems* (1979), 377–388

[9] W. T. TUTTE, How to draw a graph. *Proceedings of The London Mathematical Society* **13** (1963), 743–767

**Günter Wallner**
e-mail: guenter.wallner@uni-ak.ac.at

**Franz Gruber**
e-mail: franz.gruber@uni-ak.ac.at

University of Applied Arts Vienna
Department of Geometry
1010 Vienna, Austria

Stručni rad
Prihvaćeno 13. 04. 2011.

KRISTIAN SABO
SANJA SCITOVSKI

# Lokacija objekata u ravnini*

**Location of Objects in a Plane**

**ABSTRACT**

In the paper we consider a direct and the inverse location problem in the plane. Thereby we use different distance-like functions with appropriate illustrations. Several examples from various areas of applications are given.

**Key words:** data clustering, location problem, k-means, k-median, optimization

**MSC 2010:** 62H30, 68T10, 90C26, 90C27, 91C20, 47N10

**Lokacija objekata u ravnini**

**SAŽETAK**

U radu razmatramo izravni i obratni problem lokacije objekata u ravnini uz korištenje različitih kvazimetričkih funkcija s odgovarajućim ilustracijama. Dano je nekoliko primjera iz različitih područja primjena.

**Ključne riječi:** grupiranje podataka, klasteri, problem lokacije, $k$-sredina, $k$-median, optimizacija

## 1 Uvod

U radu razmatramo sljedeći problem lokacije: Uz pretpostavku da su poznate lokacije objekata $c_1,\dots,c_k$ u ravnini, dani skup točaka $S=\{a_i=(x_i,y_i)\colon i=1,\dots,m\}$, $(m\gg k)$ treba grupirati (razdijeliti) na $k$ nepraznih disjunktnih skupova (klastera) $\pi_1,\dots,\pi_k$, tako da $j$-tom klasteru $\pi_j$ pripadnu one točke koje su u nekom smislu najbliže $j$-tom centru $c_j$. Alternativno, ovaj problem mogao bi se postaviti i na drugi način. Uz pretpostavku da je poznata particija $\Pi=\{\pi_1,\dots,\pi_k\}$ sastavljena od nepraznih disjunktnih podskupova (klastera) zadanog skupa $S\subset\mathbb{R}^2$, treba odrediti centre $c_1,\dots,c_k\in\mathbb{R}^2$ klastera $\pi_1,\dots,\pi_k$.

Odgovarajući obratni problem bio bi traženje optimalne lokacije centara $c_1^\star,\dots,c_k^\star\in\mathbb{R}^2$, na osnovi kojih bi mogli definirati i odgovarajuće optimalne klastere $\pi_1^\star,\dots,\pi_k^\star$.

[1] engl. $k$-means problem, $k$-median problem

Broj svih particija $m$-članog skupa $S$ sastavljenih od nepraznih disjunktnih skupova $\pi_1,\dots,\pi_k$ jednak je Stirlingovom broju druge vrste $\left\{{m\atop k}\right\}$ (vidi [10], str. 257 ili [22]), gdje je

$$\left\{{m\atop k}\right\}=\frac{1}{k!}\sum_{j=1}^{k}(-1)^{k-j}\binom{k}{j}j^m,$$

koji u praktičnim primjenama može biti izuzetno velik (vidi [17]). Problem traženja optimalne particije spada u NP-teške probleme (vidi [7]) nekonveksne optimizacije općenito nediferencijabilne funkcije više varijabli, koja najčešće posjeduje značajan broj stacionarnih točaka.

Postoji opsežna recentna literatura iz ovog područja s različitim i brojnim primjenama, a može se pronaći pod nazivom *problem k-sredina* ili *problem k-medijana*[1] [11, 12, 16–18]. Tako se mogu pronaći brojne primjene u poljoprivredi (primjerice, razvrstavanje oranica prema plodnosti zemljišta), biologiji (primjerice, klasifikacija kukaca u grupe), genetici, medicini, prometu, kod biranja lokacije građevinskih objekata, kod razumijevanja klimatskih kretanja, u upravljanju (primjerice, rangiranje gradova i općina za potrebe financijske podrške), u poslovanju, u društvenim i humanističkim znanostima itd. Navedimo nekoliko konkretnih primjera:

**Primjer 1.** (Primjena u potresnom inženjerstvu)

*Metode grupiranja podataka često se koriste u različitim inženjerskim primjenama. Navedimo jednu primjenu u za život važnom potresnom inženjerstvu. Na osnovi seizmoloških podataka iz prethodnog razdoblja klasterskom analizom moguće je procijeniti mjesta nepouzdana za lokaciju građevinskih objekata. U radu [21] prikazana je primjena klasterske analize u procjenjivanju oštećenja cjevovoda (voda, nafta, plin) i prepoznavanju područja nastajanja visokih oštećenja na njima. Ta mjesta uglavnom su problematična područja (seizmički kritična) i visoko rizična u nastajanju nedostataka i kvarova na cjevovodima. Razumijevanje razloga nastajanja oštećenja na tim mjestima može poboljšati prevenciju i ublažavanje oštećenja cjevovoda.*

* Rad je napisan uz potporu Ministarstva znanosti, tehnologije i športa Republike Hrvatske u okviru projekta 235-2352818-1034.

**Primjer 2.** (Pretraživanje teksta)

*Postoji opsežna literatura o primjeni klasterske analize kod pretraživanja teksta (vidi primjerice [7, 14]), pri čemu uzorak od jedne ili više riječi treba pronaći u nekom dokumentu. U ovom slučaju općenito se radi o izuzetno velikim skupovima podataka visoke dimenzije, a rezultat pretrage očekuje se u realnom vremenu.*

**Primjer 3.**

(Detekcija opasnih mjesta u prometu)

*U doktorskoj dizertaciji [11] navodi se primjer primjene klasterske analize kod detekcije opasnih mjesta na autocesti "New Jersey Turnpike". Na osnovi podataka o mjestima i vrsti prometnih nezgoda detektiraju se kritične lokacije.*

**Primjer 4.** (Definiranje izbornih jedinica)

*Pretpostavimo da su poznati podaci* $(a_i, w_i)$, $i = 1, \ldots, n$ *o lokacijama naselja* $a_i$ *s brojem glasača* $w_i$. *Treba odrediti izborne jedinice* $\pi_1, \ldots, \pi_k$, *tako da za svaku izbornu jedinicu* $\pi_j$ *vrijedi*

$$(i) \quad (1-p)\tfrac{m}{k} \le |\pi_j| \le (1+p)\tfrac{m}{k},$$
$$(ii) \quad d(c_j, a_i) \le r, \quad \forall a_i \in \pi_j,$$

*gdje su* $c_1, \ldots, c_k$ *centri izbornih jedinica,* $m = \sum_{i=1}^{n} w_i$, $a\ p > 0$ *i* $r > 0$ *su zadani brojevi. Uvjet* $(i)$ *osigurava ravnomjernost broja glasača po izbornim jedinicama do na* $p\%$, *a uvjet* $(ii)$ *definira maksimalnu udaljenost centra* $c_j$ *izborne jedinice* $\pi_j$ *do naselja u toj izbornoj jedinici.*

**Primjer 5.**

*Promatramo skup korisnika koje na neki način treba povezati s još neizgrađenim objektima kao što su primjerice željezničke stanice, sportski kompleksi, knjižnice, antene ili supermarketi. Ovdje se prirodno pojavljuje problem određivanja optimalne lokacije objekata, tako da objekti u nekom smislu budu što bliže korisnicima. Osim toga dodatno možemo zahtijevati da troškovi povezivanja budu primjereno raspodijeljeni na sve korisnike te isto tako da prihodi koje ostvaruju objekti budu primjereno rasopodijeljeni na sve objekte i na taj način svi korisnici kao i objekti budu zadovoljni. U tom slučaju navedeni optimizacijski problem može se promatrati kao specijalni oblik problema lokacije, koji je u engleskom govornom području poznat pod nazivom* "facility location games" *(vidi primjerice [3, 9]).*

## 2 Izravni problem lokacije

Funkciju $d\colon \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_+$, koja ima barem svojstvo pozitivne definitnosti

$$d(x,y) \ge 0 \quad \& \quad d(x,y) = 0 \Leftrightarrow x = y,$$

i pomoću koje se lako računa centar $c$ svakog diskretnog skupa $\pi \subset \mathbb{R}^2$

$$c = \operatorname*{argmin}_{z \in \mathrm{conv}(\pi)} \sum_{a \in \pi} d(z,a), \tag{1}$$

zovemo *kvazimetrička funkcija*[2] (vidi primjerice [13, 20]). Najpoznatija kvazimetrička funkcija je tzv. *kvazimetrička funkcija najmanjih kvadrata* $d_{LS}(a,b) = \|a-b\|_2^2$, $a, b \in \mathbb{R}^2$, koja osim svojstva pozitivne definitnosti ima i svojstvo simetričnosti, ali ne zadovoljava nejednakosti trokuta. Centar $c$ diskretnog skupa $\pi \subset \mathbb{R}^2$ u ovom slučaju je centroid $c = \frac{1}{|\pi|} \sum_{a \in \pi} a$ (u fizici i mehanici težište ili Steinerova točka) skupa $\pi$.

Za dani skup točaka $S = \{a_i = (x_i, y_i) \in \mathbb{R}^2 \colon i = 1, \ldots, m\} \subset \mathbb{R}^2$ izravni problem lokacije možemo definirati na dva načina:

*A*: Ako su poznati centri $c_1, \ldots, c_k \in \mathbb{R}^2$ $(k \ll m)$, treba odrediti particiju skupa $S$ sastavljenu od nepraznih disjunktnih skupova $\pi_1, \ldots, \pi_k$, tako da bude

$$a_i \in \pi_j \Leftrightarrow d(c_j, a_i) \le d(c_s, a_i), \quad \forall s = 1, \ldots, k; \tag{2}$$

*B*: Ako je poznata particija $\Pi = \{\pi_1, \ldots, \pi_k\}$ skupa $S$ sastavljena od nepraznih disjunktnih skupova, treba odrediti centre $c_1, \ldots, c_k \in \mathbb{R}^2$, $(k \ll m)$ klastera prema (1).

Primjerice ako su u Primjeru 4 poznati centri $c_1, \ldots, c_k$ izbornih jedinica, a treba definirati izborne jedinice $\pi_1, \ldots, \pi_k$, radi se o rješavanju izravnog problema *A*. Korištenjem različitih kvazimetričkh funkcija dobivaju se razni principi – kriteriji grupiranja. Osim spomenute kvazimetričke funkcije najmanjih kvadrata $d_{LS}$, u literaturi [2, 4, 7, 11, 13, 15, 17] često se koristi

[2] Oznaka $v = \operatorname*{argmin}_{u \in D} f(u)$ znači da funkcija $f$ u točki $v \in D$ postiže svoju najmanju vrijednost na skupu $D$

(*i*) $d_1(a,b) = \|a-b\|_1 = |x_1-x_2| + |y_1-y_2|$,
$a=(x_1,y_1), b=(x_2,y_2)$
(Manhattan ili taxicab udaljenost)

(*ii*) $d_\infty(a,b) = \|a-b\|_\infty = \max\{|x_1-x_2|, |y_1-y_2|\}$,
$a=(x_1,y_1), b=(x_2,y_2)$
(Čebiševljeva udaljenost)

(*iii*) $d_B(a,b) = x_1 \ln\frac{x_1}{x_2} + y_1 \ln\frac{y_1}{y_2} - x_1 - y_1 + x_2 + y_2$,
$a=(x_1,y_1), b=(x_2,y_2)$
(Bregmanova generalizirana I-udaljenost
ili Kullbach-Leiblerova udaljenost).

Spomenimo da kvazimetričke funkcije $d_1$ i $d_\infty$ zadovoljavaju još dodatno svojstva simetričnosti i nejednakosti trokuta te su zbog toga prave metričke funkcije, za razliku od Bregmanove generalizirane I-udaljenosti $d_B$ za koju se lako vidi da ne zadovoljava niti jedno od spomenutih svojstava metrike. Navedena Bregmanova generalizirana I-udaljenosti $d_B$ samo je jedna iz klase kvazimetričkih funkcija koje su u literaturi poznate kao Bregmanove udaljenosti. Važno svojstvo Bregmanovih udaljenosti je lako računanje centra diskretnog skupa u smislu formule (1), a za mnoge ovakve kvazimetričke funkcije za centar diskretnog skupa mogu se dobiti i eksplicitne formule (vidi [13]). U literaturi postoji veliki broj različitih primjena klasterske analize koje se zasnivaju na korištenju Bregmanovih udaljenosti kao što su primjerice teorija informacija, klasifikacija teksta, obrada signala, analiza govora itd. Više o tome se može naći u [1].

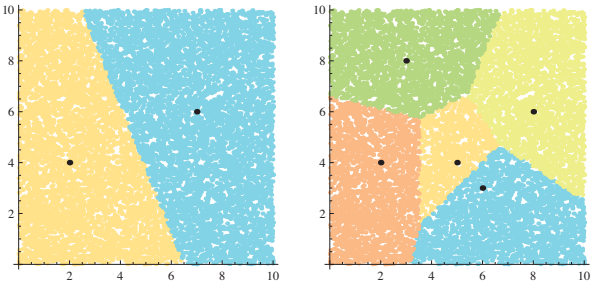
### 2.1 Izravni problem lokacije uz $d_{LS}$ kvazimetričku funkciju

Za dani skup točaka $S \subset \mathbb{R}^2$ potražit ćemo rješenje izravnog problema lokacije $A$ i izravnog problema $B$ uz korištenje $d_{LS}$ kvazimetričke funkcije. Ako su poznati centri $c_1,\dots,c_k \in \mathbb{R}^2$, particiju $\Pi = \{\pi_1,\dots,\pi_k\}$ određujemo *principom minimalnih udaljenosti* tako da bude (vidi [7, 13, 17, 19])

$$a_i \in \pi_j \Leftrightarrow \|c_j - a_i\|_2 \le \|c_s - a_i\|_2, \forall s = 1,\dots,k. \qquad (3)$$

Za $k=2$ ovaj problem (vidi [19]) svodi se na određivanje simetrale dužine $\overline{c_1c_2}$, a može se eksplicitno riješiti kao što je prikazano na Slici 1.

a) $m=10\,000$, $k=2$ b) $m=10\,000$, $k=5$



Slika 1: *Grupiranje skupa $S$ u klastere*

Pokazuje se da u općem slučaju ovaj problem vodi na konstrukciju Voronoijevih dijagrama (vidi [7, 13, 15, 17]). O konstrukciji Voronoijevih dijagrama vidi također [8]. Na Slici 1 ovaj problem ilustriran je primjenom vlastitog `Mathematica-modula` za $m=10\,000$ podataka u ravnini i $k=2$, odnosno $k=5$ centara.

Ako je poznata particija $\Pi = \{\pi_1,\dots,\pi_k\}$ skupa $S = \{a_i=(x_i,y_i)\in\mathbb{R}^2 : i=1,\dots,m\} \subset \mathbb{R}^2$, centre klastera $c_1,\dots,c_k$ dobivamo na sljedeći način (vidi [17])

$$c_j = \operatorname*{argmin}_{z\in \mathrm{conv}(\pi_j)} \sum_{a\in\pi_j} \|z-a\|_2^2 =$$
$$\left( \tfrac{1}{|\pi_j|} \sum_{a_i\in\pi_j} x_i, \tfrac{1}{|\pi_j|} \sum_{a_i\in\pi_j} y_i \right). \qquad (4)$$

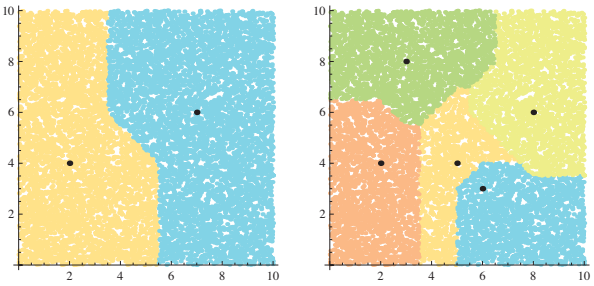
### 2.2 Izravni problem lokacije uz $d_1$ metričku funkciju

Za dani skup točaka $S \subset \mathbb{R}^2$ potražit ćemo rješenje izravnog problema lokacije $A$ i izravnog problema $B$ uz korištenje $d_1$ metričke funkcije. Ako su poznati centri $c_1,\dots,c_k \in \mathbb{R}^2$, particiju $\Pi = \{\pi_1,\dots,\pi_k\}$ određujemo *principom minimalnih udaljenosti* tako da bude

$$a_i \in \pi_j \Leftrightarrow \|c_j - a_i\|_1 \le \|c_s - a_i\|_1, \forall s = 1,\dots,k. \qquad (5)$$

Na Slici 2 ovaj problem ilustriran je primjenom vlastitog `Mathematica-modula` za $m=10\,000$ podataka u ravnini i $k=2$, odnosno $k=5$ centara.

a) $m=10\,000$, $k=2$ b) $m=10\,000$, $k=5$



Slika 2: *Grupiranje skupa $S$ u klastere*

Ako je poznata particija $\Pi = \{\pi_1, \ldots, \pi_k\}$ skupa $S = \{a_i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \ldots, m\} \subset \mathbb{R}^2$, centre klastera $c_1, \ldots, c_k$ dobivamo na sljedeći način (vidi [17])

$$c_j = \operatorname*{argmin}_{z \in \operatorname{conv}(\pi_j)} \sum_{a_i \in \pi_j} \|z - a\|_1 = \left( \operatorname*{med}_{a_i \in \pi_j} x_i, \operatorname*{med}_{a_i \in \pi_j} y_i \right), \tag{6}$$

gdje je $\operatorname*{med}_{a_i \in \pi_j} x_i$ oznaka za medijan niza kojeg čine apscise svih točaka iz klastera $\pi_j$.

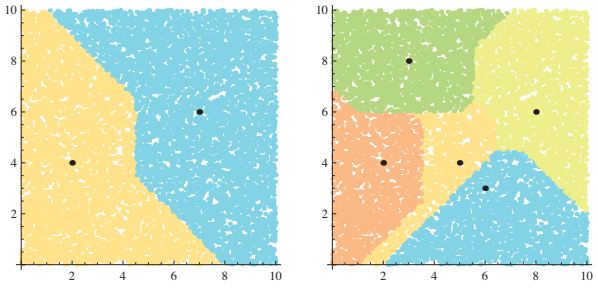
### 2.3 Izravni problem lokacije uz $d_\infty$ metričku funkciju

Za dani skup točaka $S \subset \mathbb{R}^2$ potražit ćemo rješenje izravnog problema lokacije $A$ i izravnog problema $B$ uz korištenje $d_\infty$ metričke funkcije. Ako su poznati centri $c_1, \ldots, c_k \in \mathbb{R}^2$, particiju $\Pi = \{\pi_1, \ldots, \pi_k\}$ određujemo *principom minimalnih udaljenosti* tako da bude

$$a_i \in \pi_j \Leftrightarrow \|c_j - a_i\|_\infty \leq \|c_s - a_i\|_\infty, \quad \forall s = 1, \ldots, k. \tag{7}$$

Na Slici 3 ovaj problem ilustriran je primjenom vlastitog `Mathematica-modula` za $m = 10\,000$ podataka u ravnini i $k = 2$, odnosno $k = 5$ centara.

a) $m = 10\,000$, $k = 2$ b) $m = 10\,000$, $k = 5$



Slika 3: *Grupiranje skupa $S$ u klastere*

Ako je poznata particija $\Pi = \{\pi_1, \ldots, \pi_k\}$ skupa $S = \{a_i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \ldots, m\} \subset \mathbb{R}^2$, centre klastera $c_1, \ldots, c_k$ dobivamo na sljedeći način

$$c_j = \operatorname*{argmin}_{z \in \operatorname{conv}(\pi_j)} \sum_{a \in \pi_j} \|z - a\|_\infty. \tag{8}$$

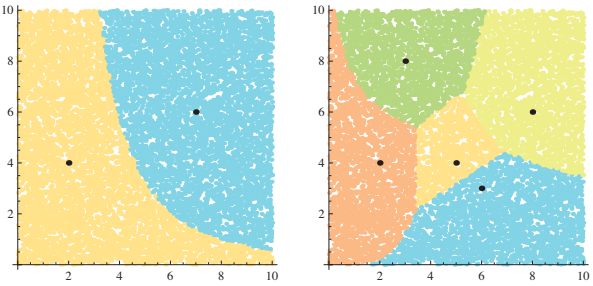
### 2.4 Izravni problem lokacije uz Bregmanovu generaliziranu I-udaljenost

Za dani skup točaka $S \subset \mathbb{R}^2$ potražit ćemo rješenje izravnog problema lokacije $A$ i izravnog problema $B$ uz korištenje kvazimetričke funkcije $d_B$ koja je u literaturi poznata kao Bregmanova generalizirana I-udaljenost. Ako su poznati centri $c_1, \ldots, c_k \in \mathbb{R}^2$, particiju $\Pi = \{\pi_1, \ldots, \pi_k\}$ određujemo *principom minimalnih udaljenosti* tako da bude

$$a_i \in \pi_j \Leftrightarrow d_B(c_j, a_i) \leq d_B(c_s, a_i), \forall s = 1, \ldots, k. \tag{9}$$

Na Slici 4 ovaj problem ilustriran je primjenom vlastitog `Mathematica-modula` za $m = 10\,000$ podataka u ravnini i $k = 2$, odnosno $k = 5$ centara. Primijetimo da su rubovi dobivenih disjunktnih skupova krivulje.

a) $m = 10\,000$, $k = 2$ b) $m = 10\,000$, $k = 5$



Slika 4: *Grupiranje skupa $S$ u klastere*

Ako je poznata particija $\Pi = \{\pi_1, \ldots, \pi_k\}$ skupa $S = \{a_i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \ldots, m\} \subset \mathbb{R}^2$, centre klastera $c_1, \ldots, c_k$ dobivamo primjenom geometrijske sredine na podatke iz klastera

$$c_j = \operatorname*{argmin}_{z \in \operatorname{conv}(\pi_j)} \sum_{a \in \pi_j} d_B(z, a) = \left( \left( \prod_{a_i \in \pi_j} x_i \right)^{1/|\pi_j|}, \left( \prod_{a_i \in \pi_j} y_i \right)^{1/|\pi_j|} \right). \tag{10}$$

### 2.5 `Mathematica-modul`

Sve prikazane ilustracije izrađene su vlastitim *Mathematica–modulom* `Particija[a, c, d]`. Modulu se predaje lista podataka `a`, lista centara `c` i ranije definirana kvazimetrička funkcija `d`. Za svaki element `a[[i]]` liste `a` modul pronalazi najbliži centar `c[[j]]`, a nakon toga element `a[[i]]` pridružuje klasteru `pi[[j]]`. Elementi svakog klastera prikazuju se točkicama jedne boje. Centri klastera označeni su crnim točkama.

```
In[1]:= Particija[a_, c_, d_] :=
Module[{m=Length[a], k=Length[c], pi,tab,min,imin},
 slc=ListPlot[c,
     PlotStyle -> {Black, AbsolutePointSize[5]}];
                  (* separacija *)
 pi = Table[{}, {j, k}];
 Do[
    tab = Table[d[c[[j]], a[[i]]], {j, k}];
    min = Min[tab];
    imin = Position[tab, min][[1]];
```

```
 Do[
    If[imin=={j}, pi[[j]]=Append[pi[[j]], a[[i]]]],
       {j, k}],
       {i, m}];
                  (* crtanje *)
 tab = Table[
        ListPlot[pi[[j]],
          PlotStyle->{Opacity[.5], Hue[.13 j^2]}],
           {j, k}];
Show[tab,slc, AspectRatio->Automatic, ImageSize->200]
     ]
```

Lista podataka `a` treba biti sastavljena od objekata oblika $\{x, y\}$. To znači da podaci mogu dolaziti iz proizvoljno odabranog područja $\Omega \subset \mathbb{R}^2$. Primjerice, naredbom

```
RandomReal[10, {10000, 2}]
```

definira se 10000 uniformno distribuiranih slučajnih točaka iz kvadrata $[0,10]\times[0,10]$. Lista centara `c` odabire se po volji. Kvazimetrička funkcija `d` definira se kao funkcija dviju varijabli. Primjerice, kvazimetrička funkcija najmanjih kvadrata $d_{LS}$ definira se naredbom

```
d[x_, y_] := Norm[x - y, 2]^2
```

a metrička funkcija $d_1$ naredbom

```
d[x_, y_] := Norm[x - y, 1]
```

Podatke, kvazimetričku funkciju i poziv modula, primjerice za slučaj prikazan na Slici 2b, možemo implementirati na sljedeći način.

```
In[2]:= SeedRandom[3]
        a = RandomReal[10, {10000, 2}];
        c = {{5,4}, {6,3}, {8,6}, {2,4}, {3,8}};
        d[x_, y_] := Norm[x - y, 1]
        Particija[a, c, d]
```

Pri tome naredba `SeedRandom[3]` osigurava da ćemo svakim pokretanjem programa dobiti iste slučajne brojeve. Izvođenje programa za navedene primjere traje $1-2$ sekunde.

## 3 Obratni problem lokacije

Zadan je skup točaka $S = \{a_i = (x_i, y_i) \in \mathbb{R}^2 : i = 1,\dots,m\} \subset \mathbb{R}^2$ u ravnini i neka kvazimetrička funkcija $d\colon \mathbb{R}^2\times\mathbb{R}^2 \to \mathbb{R}_+$. Treba pronaći optimalne centre $c_1^\star,\dots,c_k^\star \in \mathbb{R}^2$ tako da bude

$$F(c_1^\star,\dots,c_k^\star) = \min_{c_1,\dots,c_k\in \mathrm{conv}(S)} F(c_1,\dots,c_k), \qquad (11)$$

gdje je $F\colon \mathbb{R}^{2k}\to\mathbb{R}$

$$F(c_1,\dots,c_k) = \sum_{i=1}^{m} \min_{1\le j\le k} d(c_j, a_i). \qquad (12)$$

Poznavajući centre $c_1^\star,\dots,c_k^\star$ principom minimalnih udaljenosti moguće je definirati optimalnu particiju $\Pi^\star = \{\pi_1^\star,\dots,\pi_k^\star\}$ (problem *A* u *t.2*). Primjerice problem određivanja optimalnih izbornih jedinica iz Primjera 4 jedan je obrnuti problem lokacije, a može se definirati na sljedeći način.

$$F(c_1,\dots,c_k) = \sum_{i=1}^{n} w_i \min_{1\le j\le k} d(c_j, a_i) \quad \longrightarrow \min,$$

uz uvjete:

$(i)$ $\quad (1-p)\frac{m}{k} \le |\pi_j| \le (1+p)\frac{m}{k}, \quad j = 1,\dots,k$

$(ii)$ $\quad d(c_j, a_i) \le r, \quad \forall a_i \in \pi_j, \forall j = 1,\dots,k.$

Kao što smo već ranije spomenuli, ovaj problem u literaturi se može pronaći pod nazivom problem $k$-sredina ili problem $k$-medijana, a u općem slučaju radi se o problemu traženja globalnog minimuma višedimenzionalne nediferencijabilne funkcije (vidi [5, 6]) koja može imati veći broj lokalnih minimuma (vidi primjerice [11, 15]). O metodama za rješavanje ovog problema može se vidjeti primjerice u [16, 18].

Najpoznatiji postupak za rješavanje ovog problema je *algoritam $k$-sredina* (vidi [7, 11, 13, 15, 17, 19, 20]), kojim nažalost možemo pronaći lokalni minimum kriterijske funkcije cilja. Ipak, višestrukim pokretanjem ovog algoritma s različitim početnim aproksimacijama, možemo pronaći optimalno rješenje (vidi [15]). Niže navodimo skicu algoritma $k$-sredina za opći slučaj kada je $S \subset \mathbb{R}^n$.

**Algoritam 1.**
**(Standardni algoritam $k$-sredina)**[3]

Korak 1: Izabrati $\Pi = \{\pi_1,\dots,\pi_k\}$;

Korak 2: Izračunati: $\theta = (c_1,\dots,c_k)$, pri čemu je $c_j = \operatorname{argmin}_{z\in\mathbb{R}^n} \sum_{a_i\in\pi_j} d(z, a_i)$;
Izračunati $F_1 = F(\theta)$;

Korak 3: Pomoću centara iz $\theta$ prema principu minimalnih udaljenosti formirati novu particiju $\mathcal{N} = \{\nu_1,\dots,\nu_k\}$;
Izračunati nove centre: $\zeta = (\zeta_1,\dots,\zeta_k)$, pri čemu je $\zeta_j = \operatorname{argmin}_{z\in\mathbb{R}^n} \sum_{a_i\in\nu_j} d(z, a_i)$;
Izračunati $F_2 = F(\zeta)$;

Korak 4: Ako je $F_2 < F_1$, staviti $\theta = \zeta$; $F_1 = F_2$ i prijeći na Korak 3; u suprotnom STOP.

[3]Odgovarajuća programska podrška dostupna je na adresi: `http://www.mathos.hr/oml/software.htm`

## Literatura


[1] M. R. Ackermann, J. Blömer, *Coresets and Approximate Clustering for Bregman Divergences*, Proceedings of the 20th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'09), 1088-1097, Society for Industrial and Applied Mathematics (SIAM), 2009

[2] A. Banerjee, S. Merugu, I. S. Dhillon, J. Ghosh, *Clustering with Bregman divergences*, Journal of Machine Learning Research, **6**(2005), 1705–1749

[3] P. Chardaire, *Facility location optimization and cooperative games*, PhD thesis, University of East Anglia, Norwich, UK, 1998.

[4] B. Divjak, *Bilješke o taxicab geometriji*, KoG **5**(2000), 5-9

[5] D. E. Finkel, C. T. Kelley, *Additive scaling and the `DIRECT` algorithm*, J. Glob. Optim. **36**(2006), 597-–608

[6] C. A. Floudas, C. E. Gounaris, *A review of recent advances in global optimization*, J. Glob. Optim. **45**(2009), 3-–38

[7] G. Gan, C Ma, J. Wu, *Data Clustering: Theory, Algorithms, and Applications*, SIAM, Philadelphia, 2007.

[8] Ž. Gjuranić, *Modeliranje terena pomoću Delaunayjeve triangulacije*, KoG **11**(2007), 49-52

[9] M. X. Goemans, M. Skutella, *Cooperative facility location games*, Journal of Algorithms **50**(2004) 194–214

[10] K. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics*, Addison - Wesley, Boston, 2003.

[11] C. Iyigun, *Probabilistic Distance Clustering*, Dissertation, Graduate School – New Jersey, Rutgers, 2007.

[12] C. Iyigun, A. Ben-Israel, *A generalized Weiszfeld method for the multi-facility location problem*, Operations Research Letters **38**(2010), 207-214

[13] J. Kogan, *Introduction to Clustering Large and High-Dimensional Data*, Cambridge University Press, 2007.

[14] J. Kogan, C. Nicholas, M. Wiacek, *Hybrid clustering with divergences*. In: M. W. Berry and M. Castellanos (Eds.), Survey of Text Mining: Clustering, Classification, and Retrieval, Second Edition, Springer, 2007.

[15] F. Leisch, *A toolbox for K-centroids cluster analysis*, Computational Statistics & Data Analysis **51**(2006), 526-544

[16] A. M. Rodríigues-Chia, I. Espejo, Z. Drezner, *On solving the planar k-centrum problem with Euclidean distances*, European Journal of Operational Research **207**(2010), 1169-1186

[17] K. Sabo, R. Scitovski, I. Vazler, *Grupiranje podataka: klasteri*, Osječki matematički list **10**(2010), 149-176

[18] A. Schöbel, D. Scholz, *The big cube small cube solution method for multidimensional facility location problems*, Computers & Operations Research **37**(2010), 115–122

[19] H. Späth, *Cluster–Formation und- Analyse*, R. Oldenburg Verlag, München, 1983.

[20] M. Teboulle, *A unified continuous optimization framework for center-based clustering methods*, Journal of Machine Learning Research **8**(2007), 65-102

[21] S. Toprak, E. Nacaroglu, O. A. Cetin, A. C. Koc, *Pipeline Damage Assessment Using Cluster Analysis*, TCLEE 2009: Lifeline Earthquake Engineering in a Multihazard Environment, pp. 1-8, (doi 10.1061/41050(357)78)

[22] D. Veljan, *Kombinatorna i diskretna matematika*, Algoritam, Zagreb, 2001.



**Kristian Sabo**
Odjel za matematiku
Sveučilište J. J. Strossmayera u Osijeku
e-mail: ksabo@mathos.hr

**Sanja Scitovski**
Risnjačka 7, Osijek
e-mail: sanja@mathos.hr


**Zahvala:** Zahvaljujemo anonimnom recenzentu, koji je svojim primjedbama i sugestijama u znatnoj mjeri pomogao da ovaj tekst bude bolji.

Professional paper
Accepted 29. 11. 2011.

ANA SLIEPČEVIĆ
IVANA BOŽIĆ

# Perspective Collineation and Osculating Circle of Conic in PE-plane and I-plane

**Perspective Collineation and Osculating Circle of Conic in PE-plane and I-plane**

**ABSTRACT**

All perspective collineations in a real affine plane are classified according to a constant cross-ratio and the position of the center and axis. A special attention will be given to the conditions which basic elements of perspective collineation have to fulfill in order to obtain the touch or osculation or hyperosculation of two conics. On the affine models of an isotropic and pseudo - Euclidean plane the osculating circle of a conic is constructed by using perspective collineations.

**Key words:** perspective collineation, homology, elation, constant cross - ratio, conic, osculating circle

**MSC 2010:** 51M15

**Perspektivna kolineacija i oskulacijska kružnica konike u PE-ravnini i I-ravnini**

**SAŽETAK**

Sve perspektivne kolineacije u realnoj afinoj ravnini klasificiraju se s obzirom na karakterističnu konstantu te položaj središta i osi. Pokazuje se, kako odabrati temeljne elemente perspektivne kolineacije kako bi se neka konika i njezina slika dodirivale u jednoj ili dvije točke, oskulirale se ili hiperoskulirale. Na afinim se modelima izotropne i pseudoeuklidske ravnine pomoću perspektivne kolineacije konstruiraju oskulacijske kružnice konika.

**Ključne riječi:** perspektivna kolineacija, homologija, elacija, karakteristična konstanta, konika, oskulacijska kružnica

## 1 Introduction

The transformation of a real projective plane known as a **collineation** maps points to points, lines to lines and preserves the incidence relation. Any collineation that has one range of invariant points ($o$) and a pencil of invariant lines ($S$) is called a **perspective collineation**. The fixed line $o$ is called the **axis** and the fixed point $S$ is called the **center** of the perspective collineation. All lines joining the pairs of corresponding points are called **rays** and pass through the center $S$. The intersection point of two corresponding lines lies on the axis $o$. Every perspective collineation is a projective transformation since the cross rations of four distinct points of a range of points and four distinct lines of a pencil of lines are invariant. Here are some basic properties of a perspective collineation ([1], [2]).

**Teorem 1.** *A perspective collineation is uniquely determined by its center $S$, axis $o$ and one pair of corresponding points $A, A_1$. (Instead of corresponding points a pair of corresponding lines can be given.)*

**Teorem 2.** *If $A, A_1$ denote a pair of corresponding points, $S$ the center of perspective collineation and $K$ the intersection point of the ray $SA$ and the axis $o$ ($K = SA \cap o$), then the cross ratio $(AA_1, KS)$ is constant. This constant cross ratio is marked by $k$ and generally, is a real nonzero number.*

**Teorem 3.** *All perspective collineations form a group under the operation of composition.*

## 2 Classification of plane perspective collineations

According to the mutual position of the center and axis all perspective collineations are divided into two subsets: **homologies and elations**. A perspective collineation is called an elation if its axis $o$ and center $S$ are incident, otherwise it is called a homology. In each of these subsets the affine elations and affine homologies known as a point reflection, line reflection and translation are extracted. (Table 1)

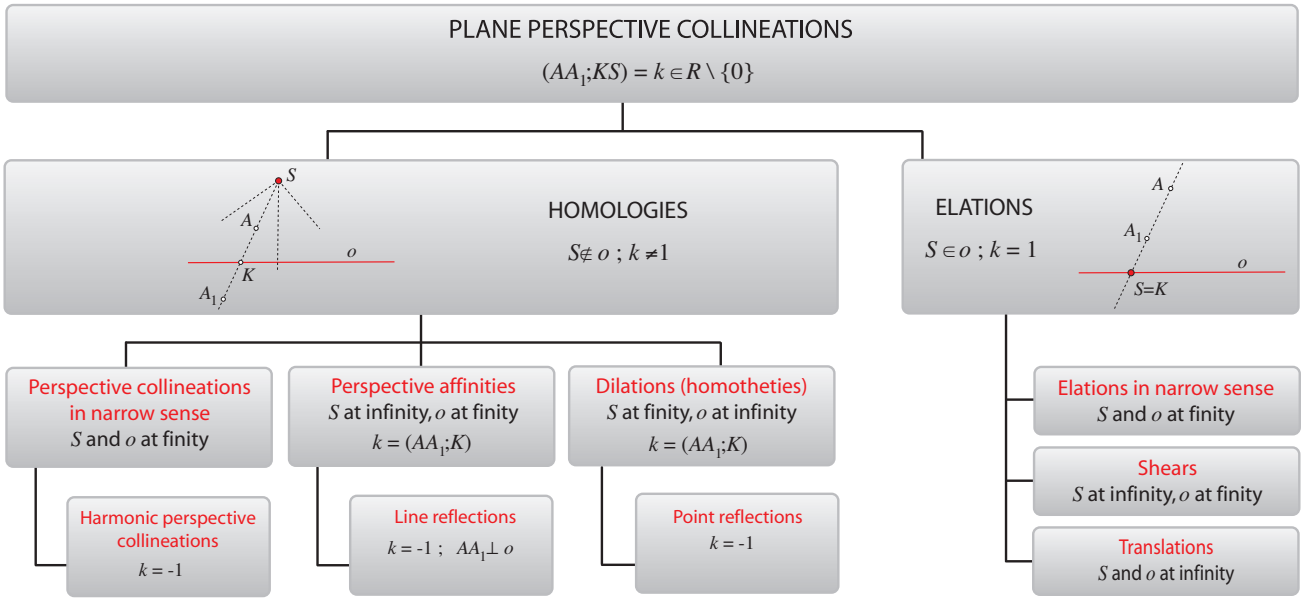



Table 1

All the cases may be summarized as follows:

*Homologies* may be classified into:

2.1 **Perspective collineations in narrow sense** are homologies with the finite center $S$ and finite axis $o$, the constant cross ratio of a perspective collineation is a real nonzero number. An involutive perspective collineation in narrow sense is called a *harmonic perspective collineation*, its constant cross ratio equals $-1$.

2.2 **Perspective affinities** are homologies with the center $S$ at infinity and the finite axis $o$. The constant cross ratio of a perspective affinity is division ratio $k = (AA_1;K)$ where $A$, $A_1$ is a pair of corresponding points and $K$ is the intersection point of the ray $AA_1$ and the axis $o$ ($K = AA_1 \cap o$). A division ratio of three collinear points is an invariant of a perspective affinity. An involutive perspective affinity is called *line reflection*. Consequently, its constant cross ratio equals $-1$.

2.3 **Perspective similarities or dilations or homotheties** are homologies with the finite center $S$ and axis $o$ at infinity. The constant cross ratio of a perspective similarity is the division ratio $k = (AA_1;S)$ where $A, A_1$ is a pair of corresponding points. An involutive perspective similarity is called a *point reflection*. Consequently, its constant cross ratio equals $-1$.

Elations are perspective collineations for which the center and axis are incident, that is $S = K$. The constant cross ratio of all elations is equal to 1.

The *elations* may be classified as follows:

2.4 **Elations in narrow sense**- with the finite center and finite axis

2.5 **Shears**- with the center at infinity

2.6 **Translations** - with the center and axis at infinity

Shears and translations map the line at infinity to itself. Thus, they are affine transformations. An affine transformation preserves division ratio of three collinear points.

## 3 Construction of the osculating circle of a conic at an arbitrary point

The order of a conic is an invariant of perspective collineation, i.e., a perspective collineation maps conics into conics. Affine transformations preserve the line at infnity, hence they map a conic into a conic of the same type (i.e. ellipse is mapped into ellipse, hyperbola is mapped into hyperbola and parabola is mapped into parabola). Two conics intersect in four points, some of which may coincide or be real or imaginary. If two real intersection points coincide, the conics $c$ and $c_1$ *touch* at this so-called touching point. If

three real intersection points coincide, $c$ and $c_1$ are *osculating conics* at this point. If four real intersection points coincide, $c$ and $c_1$ are *hyperosculating conics* at this point. A special attention will be given to the conditions which basic elements of perspective collineation have to fulfill in order to obtain the touch or osculation or hyperosculation of two conics.
If a conic $c$ touches the axis $o$ at a point $A$ or passes through the center $S$ ($S \notin o$) of a perspective collineation, then the conic $c$ will be mapped into a conic $c_1$ which touches the conic $c$ at the point $A$ or at the point $S$, respectively.
If a conic $c$ touches the axis $o$ at a point $A$ and passes through the center $S$ ($S \notin o$) of a perspective collineation, the points $S$ and $A$ are the touching points of the conics $c$ and $c_1$. A conics with two common touching points can also be obtained if the point $S$ and the line $o$ are a pole and a polar with respect to a conic $c$. In this case the intersections of the axis and the conic $c$ are common points of tangency for $c$ and $c_1$. If the center of a perspective collineation is within the conic $c$ than the intersection points will be a pair of conjugate imaginary points.
If the conics $c$ and $c_1$ are osculating conics, they determine an elation with the common tangent at the point of tangency as its axis (the point of tangency is different from the center of the elation). Also if a conic $c$ passes through the center of an elation and doesn't touch the axis, then the conics $c$ and $c_1$ are osculating conics.
If a conic $c$ touches axis $o$ at the center $S$ of elation then conic $c$ will be mapped into hyperosculating conic $c_1$.

All these aforementioned facts provide that by using the appropriate perspective collineation for given conic $c$ it is possible to construct an osculating or hyperosculating conic or a conic $c_1$ which touches the conic $c$ at one or two points.
By applying a perspective collineation an osculating circle of a conic in the Euclidean plane and on the projective models of some projective - metric planes is constructed in [4] and [5]. By using an elation the same constructions can be made on the affine models of the pseudo - Euclidean and isotropic plane.

### 3.1 Pseudo - Euclidean plane

The ordered triple $\{f, I, J\}$ is called the absolute figure of the pseudo - Euclidean plane where $I$ and $J$ are two real absolute points on the absolute line $f$. According to their position with respect to the absolute figure, the proper conics of the pseudo - Euclidean plane may be divided into nine types [4]. A circle is a conic incidental with both absolute points.

Let the absolute figure of PE - plane be $\{f, I, J\}$ where the line $f$ is a line at infinity and the points $I$ and $J$ are the points at infinity on perpendicular lines $i$ and $j$. Let the pseudo - Euclidean ellipse $c$ be presented by a circle in the Euclidean sense. It needs to construct the osculating circle of $c$ at its arbitrarily chosen point $A$. By using the appropriate elation the given ellipse $c$ can be mapped into its osculating circle $c_1$ at the given point $A$. The construction is carried out in steps:
The point $A$ is selected for the center of an elation. The intersection points of the rays $AI$ and $AJ$ with the ellipse $c$ are marked by $I'$ and $J'$. $I', I$ and $J', J$ are the pairs of corresponding points of that elation. The line $f' = I'J'$ corresponds to the absolute line $f$. The lines $f$ and $f'$ intersect at the point at infinity, thus the axis $o$ is parallel to the line $f'$ and passes through the center $S$. The elation $(A, o, I, I')$ maps the given conic $c$ into an osculating circle $c_1$ (Figure 1).

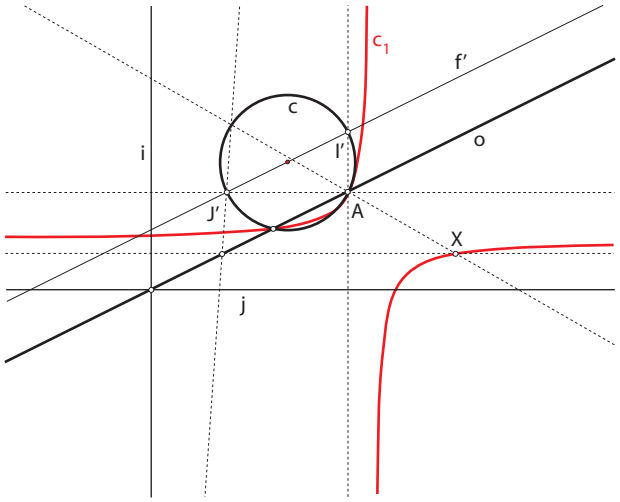



Figure 1

### 3.2 Isotropic plane

The ordered pair $\{f, F\}$ is called the absolute figure of the isotropic plane where the point $F$ is called the absolute point on the absolute line $f$ [3]. Let the affine model of an isotropic plane with the absolute figure at infinity be given. Let the absolute point $F$ be on the line $f_F$. A circle of the isotropic plane is a conic touching $f$ at $F$. Let a conic $c$ and the tangent at a point $A$ on the conic are given.There are two ways to find an elation that will map the given conic $c$ into the osculating circle $c_1$ at the point $A$:

The first way is to take the tangent $a$ of the conic $c$ in the point $A$ for the axis of an elation, and then find

the center $S$ of the elation on the axis. If the tangent $a$ is taken for the axis of an elation then the tangent $f'$ to the conic $c$ corresponds to the absolute tangent $f$ of the osculating conic $c_1$ passes through the intersection point of the lines $a$ and $f$. The point of tangency $F'$ of $f'$ and $c$ and the absolute point $F$ are a pair of corresponding points, and the ray $FF'$ intersects the axis $a$ at the center $S$ of an elation (Figure 2).

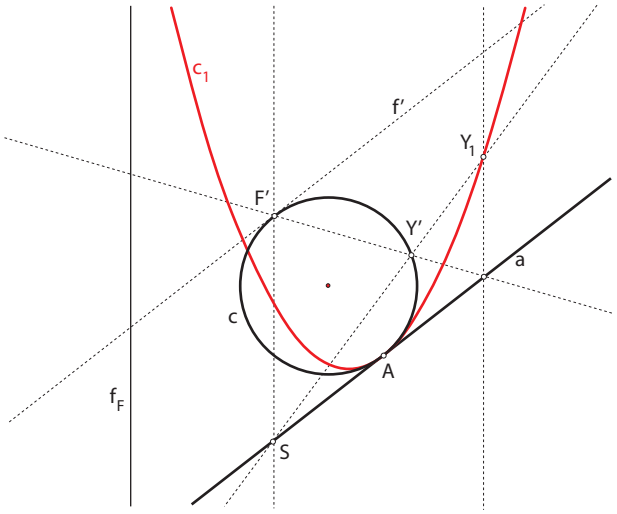



Figure 2

The second way is to take the point $A$ for the center of elation and then find the axis of the elation. If the point $A$ is taken for the center of an elation, then $AF$ is the ray the elation. The point $F'$ is the intersection point of the ray $AF$ and the conic $c$. The line pair $f, f'$ is a pair of corresponding lines. The axis $o$ passes through the point $A$ and intersection point of the lines $f'$ and $f$, thus it is parallel to $f'$. (Figure 3).

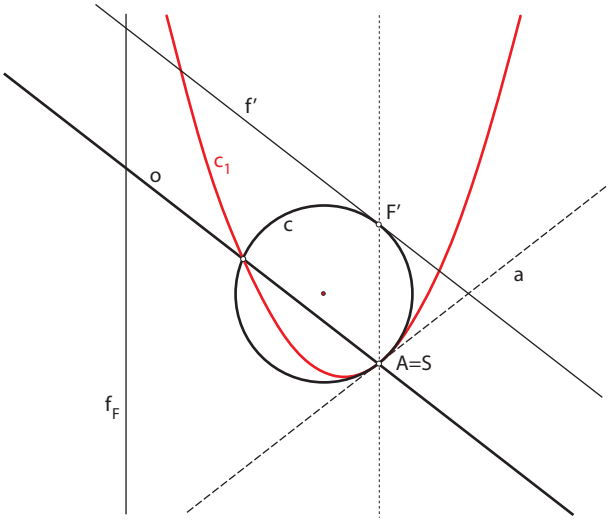



Figure 3

## References

[1] Cederberg J.N., *A course in modern geometries*, Springer-Verlag, New York, 2005.

[2] Palman D., *Projektivna geometrija*, Školska knjiga, Zagreb, 1984.

[3] Sliepčević A., Katić - Žlepalo M., Osculating circles of conics in the isotropic plane, Slovenský časopis pre geometriu a grafiku, 5 (2008), No. 10, 21-26

[4] Sliepčević A., Kovačević N., Hyperosculating circles of conics in the pseudo - euclidean plane, manuscript

[5] Weiss G., Sliepčević A., Osculating circles of conics in Cayley - Kleine planes, KoG, No.13 (2009), 7-12

**Ana Sliepčević**
Faculty of Civil Engineering, University of Zagreb
Kačićeva 26, 10000 Zagreb, Croatia
e-mail: anas@grad.hr

**Ivana Božić**
Department of Civil Engineering
Polytechnic of Zagreb
V. Holjevca 15, 10010 Zagreb, Croatia
e-mail: ivana.bozic@tvz.hr

Professional paper
Accepted 30. 11. 2011.

MARTINA TRIPLAT HORVAT
MILJENKO LAPAINE
DRAŽEN TUTIĆ

# Application of Bošković's Geometric Adjustment Method on Five Meridian Degrees

*On the Occasion of 300th Anniversary of the Birth of Ruđer Josip Bošković*

**Application of Bošković's Geometric Adjustment Method on Five Meridian Degrees**

**ABSTRACT**

In this paper, the first method of adjustment, proposed by Josip Ruđer Bošković, is described in detail, on the example on five meridian degrees. Bošković sets three conditions on the data of the lengths of the meridian degrees to calculate corrections that would fix all degrees in order to get a better estimate of true values. The conditions that have to be satisfied are explained by geometric method which Bošković described in all his studies. For the purpose of this paper, in the process of computing these five meridian degrees, data from Bošković original book have been used.

Geometric solution, described by Bošković himself, is not easy to understand at first, as this is noted by other authors who have studied Bošković's method as well. Hence, geometric description of the Bošković's method is shown in analytical form as well.

**Key words:** Josip Ruđer Bošković, geometric adjustment method

**MSC 2010:** 01A50, 86A30, 62-03, 51-03, 62A01, 62P99, 62J05, 41A10

**Primjena Boškovićeve geometrijske metode izjednačenja na pet meridijanskih stupnjeva**

**SAŽETAK**

U ovom radu detaljno je prikazana prva metoda izjednačenja, koju je osmislio Josip Ruđer Bošković, na primjeru pet stupnjeva meridijana. Bošković je izračunao popravke kojima bi popravio duljine meridijanskih stupnjeva i na taj način dobio što bolje procjene njihovih pravih vrijednosti. Postavljajući tri uvjeta tom prilikom formirao je svoju metodu izjednačenja koju je primijenio na podatke o duljinama meridijanskih stupnjeva. Uvjeti koji moraju biti zadovoljeni objašnjeni su geometrijskom metodom kakvu Bošković opisuje u svim svojim djelima. U postupku računanja, koja su provedena u ovom radu na pet meridijanskih stupnjeva, korišteni su podaci iz Boškovićevih originalnih djela.

Geometrijsko rješenje kako ga je Bošković opisao nije odmah lako razumljivo, što su uočili i drugi autori koji su proučavali Boškovićevu metodu. Stoga će geometrijski opisana Boškovićeva metoda biti također prikazana i u analitičkom obliku.

**Ključne riječi:** Josip Ruđer Bošković, geometrijska metoda izjednačenja

## 1 Introduction

Ruđer Josip Bošković (Dubrovnik, 18th May 1711 - Milan, 13th February 1787) began to publish theses on Earth's shape and size as a young scientist. These issues were a major scientific problem of the 18th century. In 1739, when he was only 28, he published two dissertations: *De veterum argumentis pro telluris sphaericitate* (On the arguments of the ancients for the sphericity of the Earth) and *Dissertatio de telluris figura* (A dissertation on the shape of the Earth).

During the 18th century scientists were having a great discussion about the question whether the Earth was oblate or oblong (prolate) at the poles. In the late 17th century, Newton proved that the Earth should be flattened at the poles because of its rotation. Domenico Cassini assumed the opposite, that Earth had the shape of an egg so, at the end of

17th and the beginning of 18th century he conducted comprehensive geodetic observations to prove his assumption.

During this period there were two basic methods for determining the Earth's figure: pendulum experiments and the determination of the meridian arc length. The idea of the second method was to determine the length of the meridian arc that corresponded to one degree of latitude. French Academy carried out the measurements during 1730s to test theoretical interpretations of the Earth's figure.

## 2 Bošković's Thoughts on the Shape of the Earth

Bošković thought that the irregularity of the shape of the Earth could be examined in the way to "exactly determine two meridian degrees in different longitudes, but in the same latitude" [1]. Bošković wanted to measure a meridian degree at some latitude, on which another meridian degree is accurately determined, but at a different longitude. Pope Benedict XIV gave, Cardinal Valenti, the State Secretary of the Holy See, permission for Bošković to perform "astronomical and geographical journey" along the meridian from Rome to Rimini in the Papal State because Rome and Rimini are approximately at the same meridian. The area comprised of two meridian degrees (Fig. 1). The length of the middle meridian degree, between Rome and Rimini, could be compared with the length of the measured meridian degree in the south France. This length of the meridian degree was measured ten years earlier in Perpignan, by Jacques Cassini and Nicolas Louis de Lacaille (De la Caille). In Fig. 1, the labels $A$, $B$, $C$, $D$, $E$, $F$, $G$, $H$, $I$ and $L$ show points of the chain of triangles between Rome and Rimini with the names of the hills on which points are located.

Bošković chose Christopher Maire for companion. In 1755 Bošković and Maire published the first results of those measurements and the analysis of measured data in the book *De Litteraria Expeditione per Pontificiam ditionem ad dimentiendas duas Meridiani gradus et corrigendam mappam geographicam* (A scientific journey through the Papal State with the purpose of measuring two degrees of meridian and correcting a geographical map) in more than 500 pages.

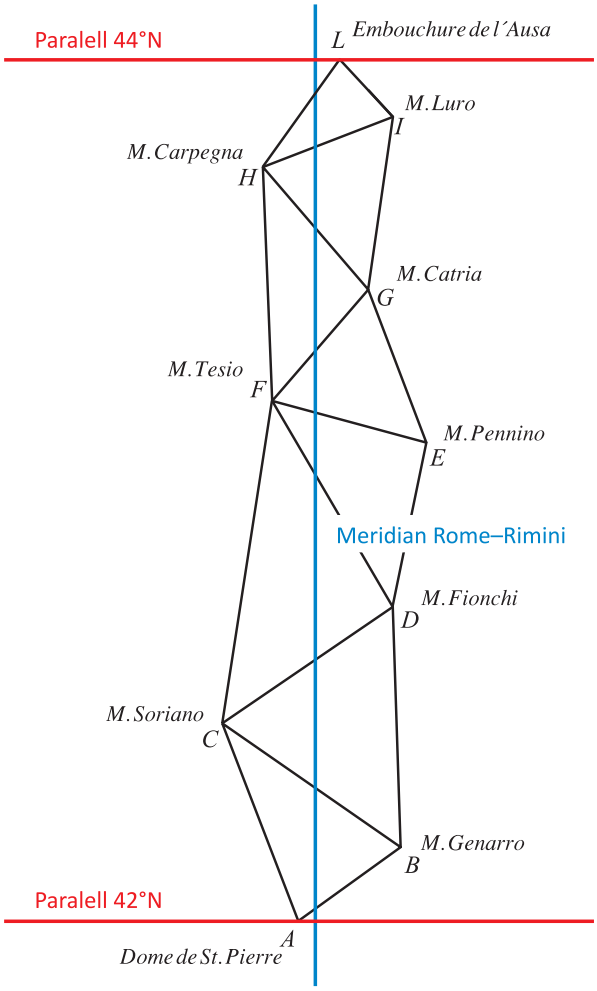



Figure 1: *Arc of the meridian Rome-Rimini and the system of the triangle chain between them (as in Fig. 2 in the first annex from the book* Voyage astronomique ...*).*

Bošković published his results of geodetic measurements three more times:

- in 1757, in the abstract form for the journal of the Academy of Bologne
- in 1760, in the supplement of the Benedict Stay's poem and
- in 1770, in French translation of his major geodetic work *Voyage astronomique...* in which he gave an example of adjustment.

In 1760 Bošković started processing the results of measurements of meridian degrees that were conducted after 1736.

In order to accurately determine the figure of the Earth, Bošković, in his first attempt to determine ellipticity, compared five arc lengths of one meridian degree, which he considered to be sufficiently accurate and in his second attempt he compared nine meridian degrees. The measurements of the five meridian degrees were carried out in Quito, in the Cape of Good Hope, in Paris, in Lapland and his own, carried out in Rome, Italy [2].

Whereas astronomical and geodetic measurements are liable with errors caused by various sources, Bošković was aware that the causes of errors could not be fully eliminated during the construction of instruments and measurements. When comparing mentioned degrees of meridian, Bošković could not determine such an ellipsoid consistent with all the measurements. He decided to determine corrections that would fix all degrees and get a better estimate of true values. He formed his own adjustment method of the results of measurements proposing three conditions for determining the corrections.

## 3 Bošković's Adjustment Method

In his works Bošković presented only geometric approach to obtain corrections. In the summary of his main book [1] Bošković stated that he used algebraic approach only to derive short formulas based on the geometric solution, which immediately gave the solution.

Bošković has set the problem in the following way. It is necessary to calculate the mean ellipticity (fr. *ellipticit*, lat. *ellipticitate*) of all meridian degrees, which are mutually compared. Taking into account the relation that has to have differences of (compared) meridian degrees as well as the laws of probabilities regarding to corrections, it is necessary to adjust the degrees to be reduced to this relation. To obtain such mean, which is not simply arithmetic mean, but tied by certain law of fortuitous combinations and the calculus of probabilities, and considering a certain number of meridian degrees, the corrections have to be found and applied to each measured meridian degree, taking into account the following three conditions [1], [3], [5]:

1. The differences of the meridian degrees are proportional to the differences of the versed sines[1] of double latitudes
2. The sum of the positive corrections is equal to the sum of the negative ones (by their absolute values) and
3. The absolute sum of all the corrections, positive as well as negative, is the least possible one, in case in which first two conditions are fulfilled.

Bošković formed the first condition out of the requirements of the law of balance, or the assumption that the Earth had the shape of an ellipsoid. Starting from the laws of gravity, Isaac Newton, in the second half of the 17th century, claimed that as the result of the Earth's rotation around its axis and the mutual attraction between the planets's mass, Earth should be flattened at the poles. Newton (1726) [4] in his *Principia* in Volume 3, Proposition 20 says: "Whence arises this Theorem, that the increase of weight in passing from the equator to the poles is nearly as the versed sine of double the latitude; or, which comes to the same thing, as the square of the right sin of the latitude; and the arcs of the degrees of latitude in the meridian increase nearly in the same proportion.".

The second condition emerged from the fact that the deflections of the pendulum or observer's errors that increase or decrease the meridian degrees have the same degree of probability, or the errors with positive and negative signs are equally probable. To fulfil the second condition, the sum of all the values of corrections should be equalized to zero and Bošković said that this was the only condition in which the sum of the positive can be equalized with the sum of the negative ones.

The third condition is necessary to approximate the measured values as much as possible because the measurement errors are probably very small. Also, this third condition, Bošković proposed because the solution was not completely defined with the first two conditions [3], [5].

In the process of computation, performed for the purpose of this paper, measured values shown in Table 1 are used. The arc lengths of meridian and their corresponding latitudes (2nd and 4th column of Table 1) have been taken out of Bošković's original books [1], [3], [5]. Other values from the Table 1 are calculated from these values. Instead of versed sines of double latitudes, a half of their values are taken (according to [2]).

[1] versed sines - reversed sin, a trigonometric function of an angle or arc which are not in use today, and is defined as $\mathrm{versin}\varphi = 1 - \cos\varphi$.

| **Place of the measured meridian degree** | **Latitude of the measured degree** | $^1/_2$versin **multiplied by 10 000** | **Arc length [toise]** | **Differences from the first degree measured in Quito [toise]** |
|---|---|---|---|---|
| Quito | 0° 0' | 0 (*AA*) | 56 751 (*Aa*) | 0 (*aa*) |
| Cape of Good Hope | 33° 18' | 3014 (*AB*) | 57 037 (*Bb*) | 286 (*Ob*) |
| Rome | 42° 59' | 4648 (*AC*) | 56 979 (*Cc*) | 228 (*Pc*) |
| Paris | 49° 23' | 5762 (*AD*) | 57 074 (*Dd*) | 323 (*Qd*) |
| Lapland | 66° 19' | 8387 (*AE*) | 57 422 (*Ee*) | 671 (*Re*) |

Table 1: *Five measured meridian degrees and their latitudes (names of places and the values in the second and fourth columns are taken from* [3]*).*

Bošković's procedure for determining the corrections can be described as following. Versed sines of double latitudes of five meridian degrees are put in a rectangular coordinate system on the abscissa (Fig. 2). The points $B$, $C$, $D$ and $E$ are drawn as if they were measured at the equator from the point $A$. On the vertical axis $aA$, $bB$, $cC$, $dD$ and $eE$ segments are put. They represent the five lengths of the corresponding measured arcs (in toises[2] per degree) that need to be fixed. These segments are perpendicular to $AF$. The size of the $AF$ can be considered as one unit in length, while the $A$, $B$, $C$, $D$ and $E$ are represent the five values of the $sin^2\varphi$ marked on the unit interval. Equator passes through point $A$, and the North Pole through point $F$.

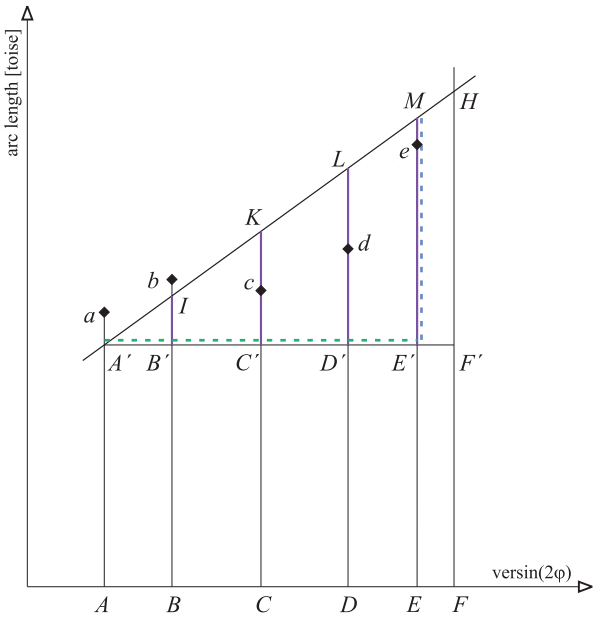



Figure 2: *Proportionality for the degrees*.

Bošković said that any straight line which intersects those segments could determine a degree which would satisfy the first condition. If we draw a straight line $A'F'$ through the point $A'$, which is parallel to $AF$, the determined meridian degrees will intersect that straight line at points $B'$, $C'$, $D'$ and $E'$. Values $E'M$, $D'L$, $C'K$, $B'I$ and $A'A'$ (zero), given in such way, represent the differences of the degrees according to a degree at the equator and they are proportional to the values $A'E'$, $A'D'$, $A'C'$, $A'B'$ and $A'A'$ (zero) or a versed sines $AE$, $AD$, $AC$, $AB$ and $AA$ (zero) [3]. The proportionality for the measured degree at the point $e$ with dashed (blue and green) lines is shown in Figure 2.
Bošković's first condition can be expressed analytically in such way that the corrections can be expressed with the equation of the first degree which includes two (unknown) sizes, values $k$ and $l$, as well as the degrees of meridian and corresponding versed sines of double latitudes:

$$L_i + v_i = k\text{versin}2\varphi_i + l, \quad (i = 1,2,\cdots,n), \tag{1}$$

where is
$l$ - the length of a meridian degree at the equator
$k$ - the excess of a degree at the pole over one at the equator,
$L_i$ - the length of an arc at location $i$
$\varphi_i$ - the latitude of the midpoint of the arc at location $i$
$v_i$ - the corrections of a meridian degree.

In order to determine the first point of the required straight line the second condition, that the sum of the positive corrections is equal to the sum of the negative ones (by their absolute values), will be applied.
The ordinate segments $eM$, $dL$, $cK$, $bI$ and $aA'$ are corrections with positive (blue segments in Fig. 3) or negative sign (red segments in Fig. 3). Their sign depends on whether the points $e$, $d$, $c$, $b$ and $a$ are located on one or the other side of straight line $A'H$ in relation to $AF$. The second condition will be satisfied when the line passes through the common centre of gravity $G$ of those points. According to the centre of the gravity, the sum of the distances of all points on one side, in any direction, is equal to the

[2]Toise - the old measure unit which are not in use today, 1 toise = 1.949 m.

sum of the distances of all points on the opposite side of the straight line [3].

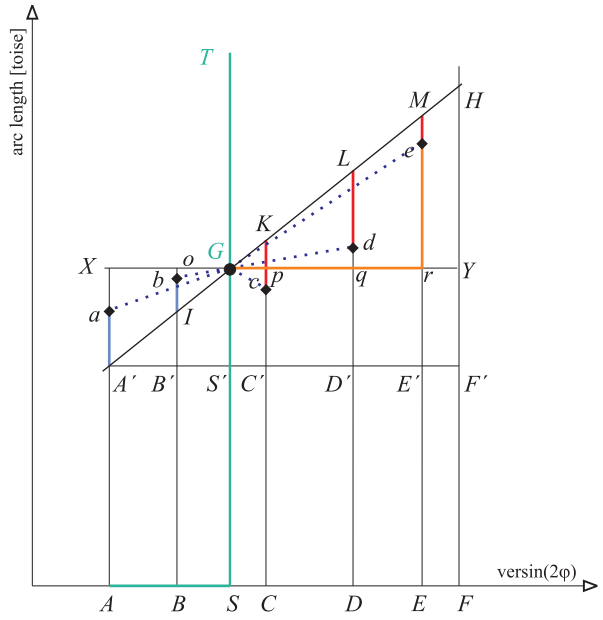



Figure 3: *Simplified Fig. 7 from Bošković's first annex with auxiliary sizes needed for calculation.*

The coordinates of the barycentre $G$ are determined by the values $AS$ and $SG$ (values shown in Fig. 3 with green lines). Latitude which corresponds to the position of the point $G$ is defined by the distance $AS$. The value $AS$ was calculated by dividing the sum of values in the third column of Table 1 with the number of the measurements.

$$AS = \frac{AA + AB + AC + AD + AE}{5} = 4362.2 \tag{2}$$

The value $SG$ defines position of point $G$ completely. This value is applied perpendicularly to $AF$, and is calculated as the arithmetic mean of the values that are in the fourth column of Table 1.

$$SG = \frac{Aa + Bb + Cc + Dd + Ee}{5} = 57052.6 \tag{3}$$

$AS$ and $SG$ define the first meridian degree which defines the straight line. Through point $G$ the infinite number of straight lines can be drawn that will satisfy the first two conditions.

Analytically, second condition can be written as follows

$$\sum_{j=1}^{m} v_j^+ = -\sum_{k=1}^{u} v_k^-, \tag{4}$$

where
$\sum_{j=1}^{m} v_j^+$ – the sum of $m$ positive corrections,
$\sum_{k=1}^{u} v_k^-$ – the sum of $u$ negative corrections.

The second condition derived from the expression (4) can also be written in the following form

$$\sum_{i=1}^{n} v_i = 0, \tag{5}$$

where
$n = m + u$ – is the number of measurements.

In order to fully determine the straight line it is necessary to determine one more point. In this case we can use the third condition.
To satisfy the third condition it is necessary to determine in which order the moving straight line, which rotates around the point $G$, passes each point. Visualize a line with a starting position $SGT$ (as it is shown with light green line in the Fig. 3) that rotates clockwise around the point $G$. In the beginning it will close a small angle, and corrections will be big. By rotating the straight line, the absolute values of corrections will decrease until the line reaches any point $a$, $b$, $c$, $d$ or $e$. When the line falls into any of these points, the correction in that point will be cancelled. As soon as the line passes that point, correction which corresponds to that particular point will change sign and it will begin to grow. At the same time, corrections of the other points to which line has not yet come will continue to decrease. Accordingly, the absolute sum of all corrections will decrease until the sum of the corrections that increase is not greater than the sum of the corrections that decrease. This conclusion comes from the fact that the sums of the corrections with the positive and negative signs are the same if we consider their absolute value and each of the sums contains half of the total sum.
The order in which the moving straight line passes through any point can be determined numerically. Bošković, however specifies that calculating will not be necessary. Structure itself, provided it is true, will be sufficient to determine the order in which the line passes through each point.
The line $XY$ is drawn parallel to $AF$ thought the point $G$ (Fig. 3). Perpendiculars to the $AF$ which pass through points $A$, $B$, $C$, $D$ and $E$ will intersect the line $XY$ in $o$, $p$, $q$ and $r$, respectively. If we visualize that the barycenter G is the center of a rectangular coordinate system, the angles $SGY$, $YGT$, $TGX$ and $XGS$ form the quadrants of this coordinate system. Firstly, it should be determined in which quadrant every point is, since each of them are left or right in relation to $SGT$, depending on whether their versed sines are smaller or larger than the $AS$. Also, each point must be below or above $XGY$ depending on whether its degree is smaller or larger than the $SG$ [3].

| | Quito ($a$) | Cape of Good Hope ($b$) | Rome ($c$) | Paris ($d$) | Lapland ($e$) |
|---|---|---|---|---|---|
| A | 4362.2 ($AS$) | 1348.2 ($BS$) | -286.8 ($CS$) | -1400.8 ($DS$) | -4025.8 ($ES$) |
| B | 301.6 ($Xa$) | 15.6 ($ob$) | 73.6 ($pc$) | -21.4 ($qd$) | -369.4 ($re$) |
| C=A/B | 14.5 | 86.4 | -3.9 | 65.4 | 10.9 |

Table 2: *Auxiliary computation for solving the third condition.*

Also the tangents of the angles which make $GS$ or $GT$ to the line that passes through each point (shown with dashed blue line in Fig. 3) should be found. If the line passes through the point $e$ then the $re$ is in relation to $Gr$, as the radius according to the tangent of the angle $reG$, or $eGT$. This relation can be written as follows (shown with orange lines in Fig. 3):

$$\tan reG = \frac{Gr}{re}. \tag{6}$$

Since $Gr$ (equals to $ES$) is the residual of versed sines of $e$ and versed sines of $AS$ (values in Table 2, row A) and $re$ is the residual of degrees $Ee$ and $SG$ (values in Table 2, row B), these values should be divided and used to sort a series of numbers (values in Table 2, row C). The series begins with a growing range of positive numbers (points contained in the first and the third quadrant of the coordinate system), and continues with a descending range of negative numbers (points that are contained in the second and the fourth quadrant of the coordinate system).

To determine all required values from Table 2 it is necessary to calculate the distances $aX$, $bo$, $cp$, $dq$, and $er$ (Table 2, row B) to the straight line $XY$. These distances are calculated as residuals of $NG$ and values $aa$, $Ob$, $Pc$, $Qd$ and $Re$ (values in the fifth column, Table 1). The value $NG$ equals to the arithmetic mean of the numbers of the fifth column in Table 1 and amounts to 301.6.

Distances $AS$, $BS$, $CS$, $DS$ and $ES$ of points $a$, $b$, $c$, $d$ and $e$ in relation to the straight line $SGT$ are equal to the difference of the size $AS$ from the sizes $AA$, $AB$, $AC$, $AD$ and $AE$ (the third column in Table 1). The calculated distances are shown in Table 2, row A.

Tangents of the angles (Table 2, row C) are calculated as the ratio of the size of the line A and B in Table 2. Arranging a series with the resulting values from the row C, the rotating line intersects the points in the order $e, a, d, b, c$.

To find the position of the straight line that will satisfy the required minimum (the sum of all corrections are at least possible) the absolute values of the sizes from row A in Table 2 are summed in the mentioned order until the sum of the values do not pass half of their total sum.

The first absolute value for the point $e$ equals 4025.8 and it is less than the half of the total absolute sum which is 5710. Then the $e$ value is added to the following value from the series ($a$) 4362.2. Their sum equals 8388 which exceeds half of the total absolute sum. When the rotating straight line passes the point $a$, the size exceeds half of their total absolute sum. In this way we get the minimum that is required.

Analytically this condition can be written as follows:

$$\sum_{i=1}^{n} |v_i| = minimum. \tag{7}$$

By finding the straight line that satisfies the conditions defined by two meridian degrees all other degrees and their corrections can be found.

To determine the corrections $bi$, $ck$, $dl$ and $em$, the value $fV$ (the residual between the length of the degree at the pole and at the equator) needs to be calculated, using the ratio:

$$aN : NG = af : fV \;\Rightarrow\; fV = \frac{NG \cdot af}{aN} = 691.4. \tag{8}$$

Respecting the requirement of proportionality with the residuals of the sizes $Oi$, $Pk$, $Ql$ and $Rm$ of the $Ob$, $Pc$, $Qd$ and $Re$, the corrections of certain meridian degrees can be calculated. The quantities $Oi$, $Pk$, $Ql$, and $Rm$ (Table 3, column 2) are the distances from the straight line $aV$ to the line $af$. The procedure for their calculation is shown on the example of the values $Oi$, which is obtained from the ratio:

$$af : aO = fV : Oi \;\Rightarrow\; Oi = \frac{aO \cdot fV}{af} = 208.4. \tag{9}$$

And the values $Ob$, $Pc$, $Qd$ and $Re$ (Table 3, column 3 and Table 1, row 5) are the differences of certain meridian degrees from the first meridian ($Aa$) because it stays without correction.

| | Quito (*a*) | Cape of Good Hope (*b*) | Rome (*c*) | Paris (*d*) | Lapland (*e*) |
|---|---|---|---|---|---|
| Distances of straight lines *aV* i *af* [toise] | 0 | 208.4 | 321.3 | 398.4 | 579.9 |
| Residual from the first degree [toise] | 0 | 286 | 228 | 323 | 671 |
| Corrections [toise] | 0 | -77.6 | +93.3 | +75.4 | -91.1 |

Table 3: *Calculated values of corrections*.

Sum of the corrections (Table 3, column 4) with a positive and with a negative sign are equal in absolute value and equals 168.7 toises which satisfies the second condition. Five meridian arc lengths and straight line that satisfies conditions are shown in Figure 4. Equation of a straight line that satisfies all conditions can be written as follows:

$$L = 0.07\text{versin}2\varphi + 56751. \qquad (10)$$

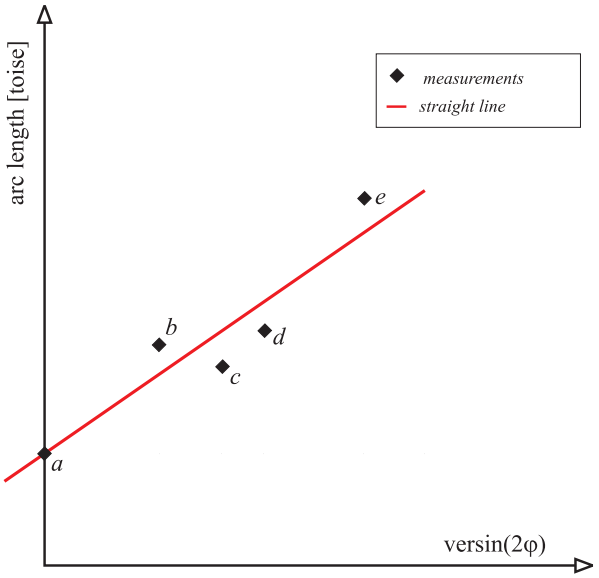



Figure 4: *Five meridian arc lengths and Bošković's straight line that satisfies conditions*.

## 4 Conclusions

The straight line which passes through points *a* and *G* gives the smallest sum of all corrections (satisfying the third condition), and the sum of positive and negative corrections is equal in its absolute value (satisfying the second condition). Point *a* remains without correction since the straight line, that gives the minimum, passes exactly through that point.

Bošković defined and applied the principle in which the measured values can be approximated with linear function. With defined conditions, unknowns and adjusted values can be determined. With equation of a straight line, Bošković linked the determined length of meridian degrees and versed sines of their corresponding latitudes, and as unknown sizes he determined values of parameters of the straight line (*k* and *l*).

In this paper we presented a method of adjustment proposed by Bošković and developed by Laplace. Today this method is mostly known as $L_1$-norm method. It uses only one condition (the third one) out of the three proposed by Bošković.
Difficulties in implementation of $L_1$-norm occurred when it was needed to adjust a greater number of unknowns. With the development of computer technology and algorithms based on linear programming this is no longer insuperable problem.

Today the most used method of adjustment is the method of least squares, also known as $L_2$-norm method. Bošković's adjustment method certainly cannot replace the now well established method of least squares, but this does not exclude the possibility that these two methods can be used parallel and complement each other. The most significant contribution of Bošković's method, and also its most important application is the detection of gross errors in measurements which gives excellent results.

## References

[1] BOSCOVICH, R. J. (1757): *De litteraria expeditione per pontificiam ditionem*. De Bononiensi Scientiarum et Artium Instituto atque Academia Commentarii, Tomus Quartus, 353-394. Typis Lalii a Vulpe Instituti Scientiarum Typographi, Bononiae.

[2] MAIRE, CH., BOSCOVICH, R. J. (1755): *De litteraria expeditione per pontificiam ditionem ad dimetiendos duos meridiani gradus et corrigendam mappam geographicam*. Typographio Palladis, Roma.

[3] MAIRE, CH., BOSCOVICH, R. J. (1770): *Voyage astronomique et ggraphique, dans l'tat de l'Eglise, entrepris par l'ordre et sous les auspicer du pape Benoit XIV, pour mesurer deux dgrs du meridian, & corriger la Carte de l'Etat ecclsiastique*. N. M. Tilliar, Paris.

[4] NEWTON, I. (1726): *Philosophiae Naturalis Principia Mathematica*. Translated, 1846; as Newtons Principia the Mathematical Principles of Natural Philosophy, by Sir Isaac Newton; Translated Into English by Andrew Motte, ed. Daniel Adee, New York.

[5] STAY, B. (1760): *Philosophiae recentioris a Benedicto Stay versibus tradita libri X Cum adnotationibus, et supplementis P. Rogerii Josephi Boscovich S. J.*, Tomus II. Typis, et sumptibus Nicolai et Marci Palearini, Roma.

**Martina Triplat Horvat**
e-mail: mthorvat@geof.hr
**Miljenko Lapaine**
e-mail: mlapaine@geof.hr
**Dražen Tutić**
e-mail: dtutic@geof.hr
Faculty of Geodesy University of Zagreb
Kačićeva 26, 10000 Zagreb, Croatia

Stručni rad
Prihvaćeno 27. 12. 2011.

MIRELA KATIĆ ŽLEPALO
BORIS UREMOVIĆ

# Primjena kotirane projekcije u određivanju obima iskopa građevinske jame

**Application of Elevational Projection in Defining Scope of Construction Pit Excavation.**

**ABSTRACT**

One of the frequent problems that the civil engineering experts have to deal with during planning the process of building is to define the scope of construction pit excavation. The most common method to solve this problem is to use the so-called elevational projection. In this article we show the phases of defining the scope of construction pit excavation together with one example from the civil engineering practice and how to use some tasks from basic theory of elevational projection to solve such problems.

**Key words:** elevational projection, scope of construction pit excavation

**MSC 2010:** 51N05

**Primjena kotirane projekcije u određivanju obima iskopa građevinske jame**

**SAŽETAK**

Jedan od čestih problema s kojim se građevinski inženjeri susreću tijekom planiranja procesa izgradnje je definiranje obima iskopa građevinske jame. Za rješavanje navedenog problema najčešće se koristi kotirana projekcija. U ovom članku navode se faze rješavanja obima iskopa građevinske jame i pokazuje se jedan primjer iz prakse te kako se nizom elementarnih položajnih zadataka riješenih u kotiranoj projekciji dolazi do rješenja.

**Ključne riječi:** kotirana projekcija, obim iskopa građevinske jame

## 1 Uvod

Kotirana projekcija je metoda ortogonalnog projiciranja na horizontalnu ravninu pri čemu je točka određena svojom projekcijom (tlocrtom) i kotom. Kota točke je mjerni broj koji predstavlja udaljenost te točke od horizontalne ravnine na koju projiciramo i izražava se u metrima. Ravninu projekcije uobičajeno uzimamo tako da sve njezine točke imaju kotu 0, smatramo da je ona na nultoj nadmorskoj visini i zovemo nultom horizontalnom ravninom, [1], [2].

Ravnina se u kotiranoj projekciji prikazuje projekcijom svojih slojnica, a to su presječnice te ravnine s horizontalnim ravninama. Sve slojnice jedne ravnine su međusobno paralelne. Analogno se teren prikazuje projekcijama svojih slojnica koje su također dobivene zamišljenim presjekom terena s horizontalnim ravninama, tj. to su linije na kojima sve točke terena imaju istu kotu. Nagib ravnine definira se kao tangens priklonog kuta te ravnine, [1], [2].

Ovim člankom želimo pokazati jedan konkretan primjer gdje se kotirana projekcija primjenjuje u praksi, a to je određivanje obima iskopa građevinske jame.

Izgradnja svake građevine sastoji se od više različitih procesa povezanih u jednu logičnu cjelinu. Prije početka izgradnje građevine svakako se mora izvršiti iskop materijala čija složenost i obim, pa u konačnici i organizacija izvođenja ovisi o različitim faktorima, od kojih su najčešći oblik i veličina građevine, konfiguracija terena ili vrsta tla.

Razlikujemo dva osnovna slučaja. Prvi slučaj se odnosi na građevine koje nemaju ukopanih (podrumskih) dijelova, već je jedino potrebno iskopati manje količine materijala na mjestima gdje je potrebno izgraditi temelje građevine. Drugi slučaj je kad građevina ima polu-ukopane ili ukopane dijelove, koji iziskuju iskop veće količine materijala odnosno zahtijevaju iskop "građevinske jame". Predmet ovog rada je pokazati upotrebu kotirane projekcije u određivanju obima takvih iskopa, a pokazat će se i jedan primjer iz građevinske prakse.

## 2 Određivanje obima iskopa građevinske jame

### 2.1 Osnovni koraci

Plan iskopa građevinske jame je sastavni dio Projekta organizacije građenja, odnosno sheme organizacije gradilišta. Obim iskopa je potrebno odrediti u fazi planiranja iz-

gradnje građevine kako bi se u sklopu sheme organizacije gradilišta mogli odrediti slobodni prostori za postavljanje ostalih privremenih objekata kao što su uredi, skladišta, otvorene deponije materijala, prometne površine, nepokretni strojevi (toranjske dizalice), itd.

Svaki problem određivanja obima iskopa građevinske jame je jedinstven, jer ovisi o različitim faktorima. Bez obzira na tu činjenicu, postupak rješavanja tog problema je uvijek isti, te se može postaviti u tri osnovna koraka; (1) analiziranje projektne dokumentacije – tlocrta i presjeka građevine; (2) analiziranje posebnih uvjeta i podataka o okolini; (3) određivanje obima iskopa građevinske jame na osnovu navedenih podataka.

### 2.2 Analiza projektne dokumentacije

Analizom projektne dokumentacije, koja među ostalim sadrži i tlocrte i nacrte građevine, određuju se osnovni podaci o iskopu kao što su tlocrtni oblik dna iskopa, dubina iskopa ili potreba za iskopom u fazama. Tlocrtni oblik dna iskopa kojeg prikazujemo u kotiranoj projekciji ovisi o obliku i vrsti temelja građevine. Rub dna iskopa paralelno slijedi rub dna temelja građevine na određenoj udaljenosti (najčešće 0,5m do 1m) i na taj način dobiva se manipulativni prostor koji je potreban zbog lakše izgradnje.

Dubina iskopa je najčešće jednaka po cijeloj tlocrtnoj površini građevine, uz određene izuzetke ovisne o vrsti temelja ili određenim drugim posebnostima projekta, npr. konfiguraciji terena. Vrsta temelja utječe na dubinu ovisno o tome je li ispod građevine predviđena jedinstvena temeljna ploča (tada je dubina iskopa konstantna) ili su za temeljenje građevine predviđeni trakasti temelji s temeljnim stopama (tada se predviđaju posebna produbljenja iskopa na njihovim mjestima).

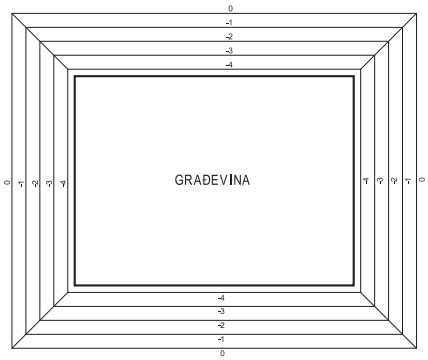
Ostale posebnosti projekta koje mogu imati utjecaja na dubinu iskopa su razna projektna rješenja (npr. otvori za dizala ili razne crpne stanice) koja iziskuju dodatna (najčešće manja) produbljenja iskopa.

### 2.3 Analiza posebnih uvjeta i podataka o okolini

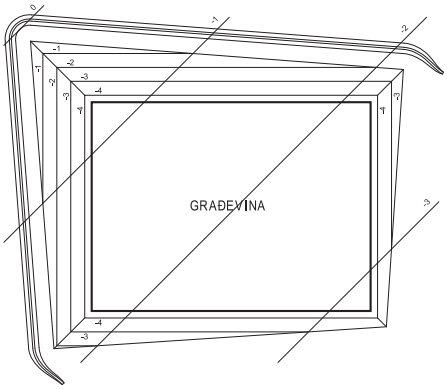
Postoji mnogo različitih uvjeta, koji mogu utjecati na obim iskopa materijala, a najvažniji su: (1) konfiguracija terena; (2) vrsta materijala koji se mora iskopati; (3) veličina i oblik građevinske čestice – parcele; (4) položaj i veličina postojećih i budućih objekata poput građevina, prometnica ili drveća, [4].

#### 2.3.1 Konfiguracija terena

Konfiguracija terena utječe prvenstveno na dubinu i oblik ruba iskopa. Ukoliko je površina terena horizontalna, utjecaj je najčešće malen, dubina iskopa je jednaka po cijeloj površini tlocrta građevine, a rub iskopa je uvijek na jednakoj udaljenosti od građevine. Na Slici 1 prikazan je u kotiranoj projekciji iskop na ravnom terenu nadmorske visine 0m. Dno iskopa treba biti na -4m nadmorske visine. Oko čitavog ruba zamišljenog tlocrta građevine dodano je 0,5 m manipulativnog prostora. Kosine iskopa su u ovom slučaju četiri ravnine nagiba $n = 1$ koje se međusobno sijeku u četiri presječnice, te se s terenom sijeku na koti 0. Presječnica dviju ravnina je pravac čije točke dobivamo presijecanjem istoimenih slojnica tih dviju ravnina, a presjek kosine iskopa s terenom dobiva se presijecanjem slojnica ravnine s istoimenim slojnicama terena. Na Slici 2 prikazana je kotirana projekcija građevinske jame sa sličnim tlocrtom građevine, dno iskopa također treba biti na -4m nadmorske visine kao i u prethodnom slučaju, ali sada teren nije horizontalan. U ovom slučaju dubina iskopa se mijenja, a rub iskopa je na različitim udaljenostima od građevine. Rub iskopa dobivamo presijecanjem slojnica ravnina iskopa s istoimenim slojnicama terena i to je topografska linija. Zbog nagiba terena potrebno je voditi računa da se predvidi mogućnost slijevanja oborinskih voda u građevinsku jamu, te da se izgradnjom zaštitnih kanala ili dodatnih konstrukcija (produbljenja za crpljenje unutar građevinske jame) omogući njihovo slobodno otjecanje ili njihovo sigurno zahvaćanje i odvodnja. Ovdje je oko ruba iskopa na udaljenosti od 1m predviđen zaštitni kanal koji zbog pada terena ima svoj prirodan pad i ispuste u skladu s konfiguracijom (padom) terena.




Slika 1: *Oblik građevinske jame ukoliko je teren horizontalan*




Slika 2: *Oblik građevinske jame ukoliko je teren pod nagibom*

#### 2.3.2 Vrsta materijala

Vrsta materijala koja se mora iskopati prvenstveno određuje nagib kosine iskopa. S obzirom na geomehanička svojstva materijala – tla, u praksi se za nagib kosine uzima vrijednost tangensa priklonog kuta 45°, tj. nagib ravnine $n = \tan\alpha = 1$. Taj nagib je prikladan kod iskopa u šljunčanom tlu ili u zemlji. Kod iskopa u čvršćem materijalu (npr. kameni materijal ili šljunak s udjelom gline) prikloni kut kosine može biti i veći od 45°, a iznimno za iskope u čvrstim stijenama može biti i okomit. S obzirom da su u prirodi moguće razne kombinacije materijala, točan nagib kosine mora se odrediti geomehaničkim ispitivanjima, [3], [4].

#### 2.3.3 Veličina i oblik građevinske čestice – parcele

Veličina i oblik građevinske čestice utječu najviše na veličinu iskopa, npr. ukoliko se građevina nalazi preblizu ruba čestice, pa nije moguće izvesti iskop u potpunosti s predviđenim nagibom kosine, potrebno je ili povećati nagib kosine iskopa (ukoliko to geomehaničke karakteristike dozvoljavaju) ili u potpunosti zamijeniti kosinu iskopa s potpornom konstrukcijom za zaštitu građevinske jame od urušavanja (armirano-betonska (AB) dijafragma, čelične talpe i sl.).

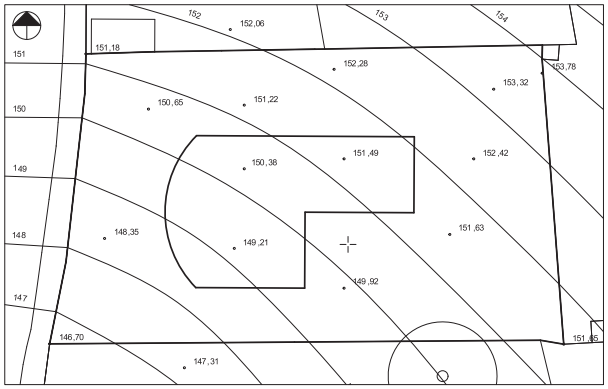
#### 2.3.4 Položaj i veličina postojećih i budućih objekata

Ukoliko na građevinskoj parceli ili izvan nje postoje objekti poput građevina ili prometnica, potrebno je izvesti iskop na način da se ne ugrozi sigurnost korištenja navedenih objekata ili same građevinske jame. Potrebna sigurnost će se postići na način da se osiguraju dovoljne udaljenosti ruba iskopa od svih postojećih i budućih objekata ili prometnica, ili da se smanje nagibi kosina iskopa, a sve kako navedeni objekti svojom težinom ne bi ugrozili postojanost kosine iskopa, [3].

### 2.4 Primjer određivanja veličine i oblika iskopa građevinske jame – primjer iz građevinske prakse

#### 2.4.1 Uvod – geodetska podloga

Geodetske podloge postojećeg stanja sadrže razne podatke, poput visinskih kota terena i objekata, rubova katastarskih čestica, rubova postojećih objekata, položaja prometnica, decimetarske mreže, oznake smjera sjevera, itd. Uobičajeno je da slojnice terena nisu ucrtane u podlogu, osim u slučajevima kada je za tim posebno izražena potreba. Položaj slojnica terena se može odrediti ručno interpolacijom ili pomoću raznih specijaliziranih računalnih programa kao na predmetnoj podlozi ispod (Slika 3). Ručna interpolacija značila bi npr. za slojnicu terena 151 približnim računom naći nekoliko točaka koje su na koti 151 te ih spojiti topografskom linijom. U primjeru na Slici 3 to bi bilo između točaka s kotama 150,65 i 151,22, zatim između 150,38 i 151,49 itd.




Slika 3: *Geodetska podloga građevinske čestice s ucrtanim položajem građevine i osnovnim podacima*

#### 2.4.2 Korak 1. – Analiza projektne dokumentacije

Analizom projektne dokumentacije za odabranu građevinu, utvrđeni su sljedeći podaci potrebni za definiranje obima iskopa građevinske jame:

- način temeljenja građevine je temeljna ploča jednake debljine ispod cjelokupne površine građevine
- donja visinska kota temeljne ploče je jednaka po cijeloj površini i iznosi +147,50 mnv
- ispod temeljne ploče po cijeloj površini se nalazi tamponski sloj šljunka debljine 50cm
- na predmetnoj građevini nisu projektirani nikakvi dodatni elementi, koji bi zahtijevali dodatna lokalna produbljenja građevinske jame, jer zgrada nema dizalo i nema crpne stanice za otpadne vode
- kada se u obzir uzmu dubina objekta, debljina tamponskog sloja šljunka ispod temeljne ploče, i sve posebnosti projektnog rješenja građevine, donja kota iskopa građevinske jame se nalazi na visini +147,00 mnv
- s obzirom na tehnologiju izgradnje, potreban je manipulativni prostor širine 1m uz temeljnu ploču, te je za toliko potrebno povećati dno iskopa građevinske jame

#### 2.4.3 Korak 2. – Analiza posebnih uvjeta i podataka o okolini

Analizom posebnih uvjeta i podataka o okolini određeni su sljedeći podaci bitni za iskop građevinske jame:

- teren je jednoliko nagnut na cjelokupnoj površini građevinske parcele, a prosječni nagib terena je približno 20% ($n = \tan\alpha = \frac{1}{5}$)
- najviša točka građevinske parcele ima visinsku kotu približno +154,10 mnv, a najniža točka ima visinsku kotu približno +146,70 mnv

- s obzirom na položaj objekta, konfiguraciju terena i hidrološke uvjete, nije predviđena izrada dodatnih konstrukcija za prihvat oborinskih voda

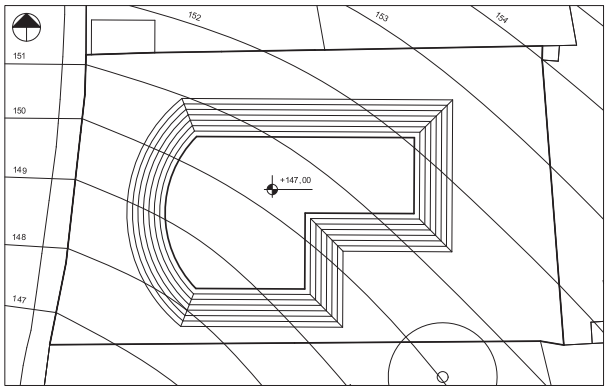
- geomehaničkim ispitivanjima utvrđeno je da se ispod cjelokupne površine građevinske parcele nalazi sloj zaglinjenog šljunka debljine 6m, te da se cjelokupna građevinska jama nalazi unutar navedenog sloja

- nagib ravnine, odnosno plohe kosine iskopa građevinske jame će s obzirom na vrstu tla (zaglinjeni šljunak) biti postavljen na $n = 2$

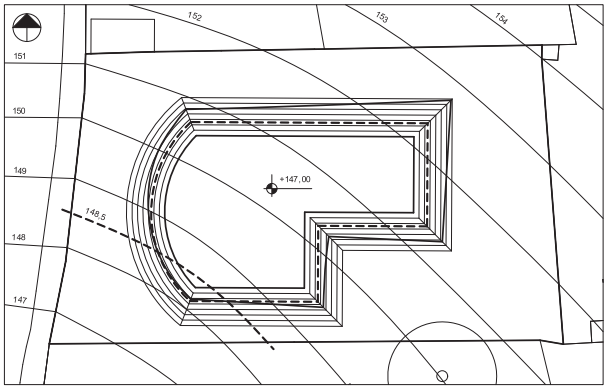
- građevina se nalazi na dovoljno velikoj udaljenosti od rubova parcele i svih susjednih postojećih i budućih objekata, pa je iskop građevinske jame moguće napraviti u cijelosti s istim nagibom kosine ($n = 2$), bez izgradnje dodatnih potpornih konstrukcija

2.4.4 Korak 3. – izrada plana iskopa građevinske jame

Na osnovu geodetske podloge, te analize projektne dokumentacije i posebnih uvjeta i podataka o okolini, što je detaljno objašnjeno u 2.4.2 i 2.4.3 napravljen je osnovni plan iskopa građevinske jame (Slika 4), a potom je određen rub iskopa kao presječna krivulja dobivena presijecanjem istoimenih slojnica ploha iskopa građevinske jame i terena (Slika 5). Za potrebe izrade presječne krivulje bilo je potrebno interpolirati dodatnu slojnicu terena na koti +148,50 mnv, kako bi se detaljnije mogla odrediti presječna krivulja na zakrivljenoj plohi iskopa na zapadnoj strani objekta.




Slika 4: *Osnovni plan iskopa građevinske jame*




Slika 5: *Presječna krivulja između ploha iskopa građevinske jame i terena*

2.4.5 Određivanje ostatka sheme organizacije gradilišta

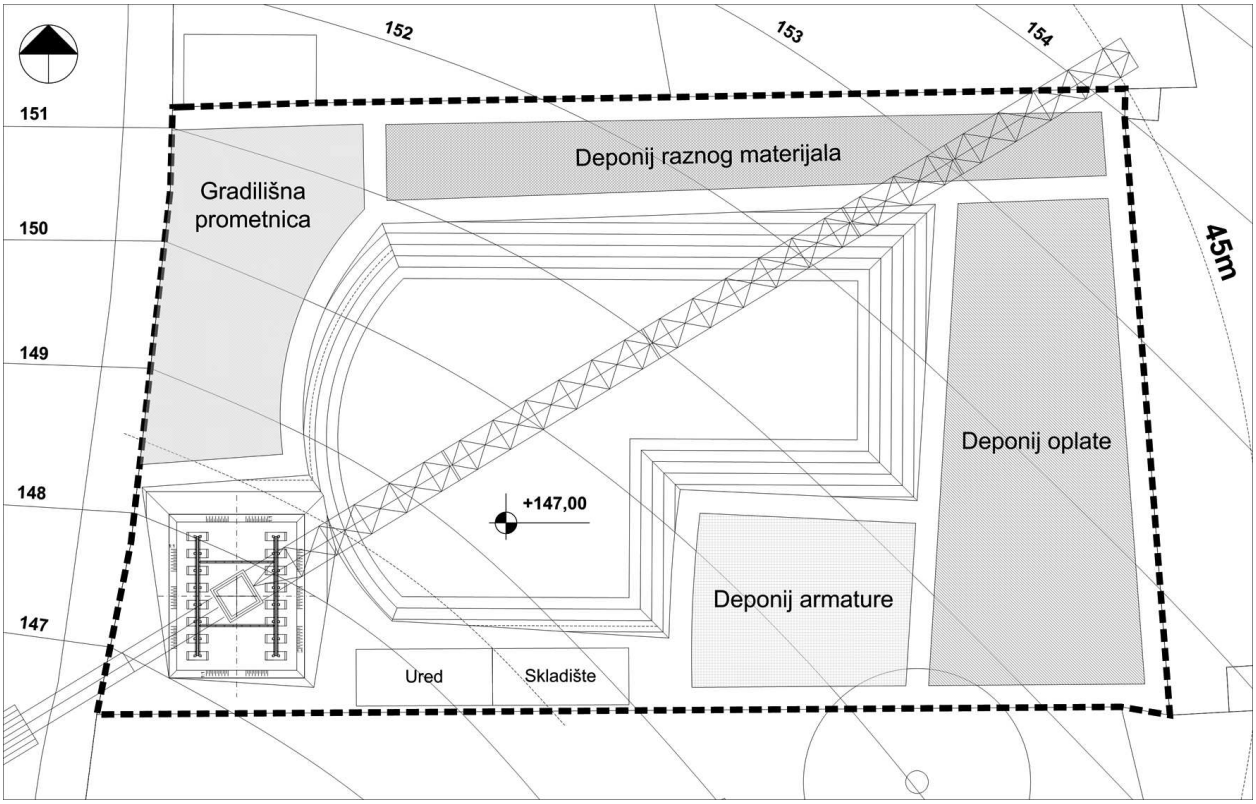
Nakon određivanja presječne krivulje između ploha iskopa građevinske jame i terena, moguće je pristupiti izradi ostatka sheme organizacije gradilišta, pa se u sljedećim koracima uz iskop građevinske jame u shemu moraju ucrtati položaji i dimenzije sljedećih trajnih i privremenih objekata:

- toranjska dizalica
- sve prometne površine
- privremeni objekti za radnike na gradilištu (uredi, sanitarni objekti, itd.)
- privremeni skladišni objekti (materijal i alat)
- privremene deponije materijala
- instalacije (električna energija, vodovod, kanalizacija)

Sve navedeno možemo vidjeti u dovršenoj shemi organizacije gradilišta na Slici 6.

## 3 Zaključak

Određivanje obima iskopa građevinske jame u praksi se radi pomoću kotirane projekcije. Poznato je da se kotirana projekcija primjenjuje kod rješavanja mnogih zadataka u niskogradnji (projektiranje cesta, željeznica, trasiranje...), no ovim člankom htjeli smo pokazati da se kotirana projekcija primjenjuje i u organizaciji gradilišta. Radi se zapravo o nizu elementarnih položajnih zadataka riješenih u kotiranoj projekciji kao što su postavljanje ravnine određenog nagiba, određivanje presječnica dviju ravnina, konstruiranje presječnice ravnine i terena, interpolacija slojnica itd. Možemo dakle zaključiti da je geometrija u srži mnogih zadataka s kojima se projektanti svakodnevno susreću u praksi.




Slika 6: *Shema organizacije gradilišta*

## Literatura

[1] I. BABIĆ, S. GORJANC, A. SLIEPČEVIĆ, V. SZIROVICZA, *Nacrtna geometrija – zadaci*, HDGG, Zagreb, 2007.

[2] J. KOS-MODOR, E. JURKIN, N. KOVAČEVIĆ, *Kotirana projekcija – skripta iz nacrtne geometrije za RGN fakultet*, HDGG, Zagreb, 2010.

[3] Z. LINARIĆ, *Leksikon strojeva i opreme za proizvodnju građevinskih materijala: Učinci za strojeve i vozila pri zemljanim radovima*, Biblioteka Mineral, Busines Media Croatia, Zagreb, 2007

[4] Z. LINARIĆ, *Tehnologija građenja I – zemljani radovi*, elektronski udžbenik za Tehničko veleučilište u Zagrebu

**Mirela Katić Žlepalo**
e-mail: mkatic@tvz.hr
**Boris Uremović**
e-mail: boris.uremovic@tvz.hr
Tehničko veleučilište u Zagrebu
Avenija Većeslava Holjevca 15, 10000 Zagreb, Croatia

# How to get KoG?

The easiest way to get your copy of KoG is by contacting the editor's office:

Nikoleta Sudeta
nsudeta@arhitekt.hr
Faculty of Architecture
Kačićeva 26, 10 000 Zagreb, Croatia
Tel: (+385 1) 4639 219
Fax: (+385 1) 4639 465

The price of the issue is €15 + mailing expenses €5 for European countries and €10 for other parts of the world.

The amount is payable to:

| | |
|---|---|
| ACCOUNT NAME: | Hrvatsko društvo za geometriju i grafiku<br>Kačićeva 26, 10000 Zagreb, Croatia |
| IBAN: | HR862360000-1101517436 |

---

# Kako nabaviti KoG?

KoG je najjednostavnije nabaviti u uredništvu časopisa:

Nikoleta Sudeta
nsudeta@arhitekt.hr
Arhitektonski fakultet
Kačićeva 26, 10 000 Zagreb
Tel: (01) 4639 219
Fax: (01) 4639 465

Za Hrvatsku je cijena primjerka 100 KN + 10 KN za poštarinu.

Nakon uplate za:

**HDGG (za KoG), Kačićeva 26, 10000 Zagreb**
žiro račun broj **2360000-1101517436**

poslat ćemo časopis na Vašu adresu.

Ako Vas zanima tematika časopisa i rad našega društva, preporučamo Vam da postanete članom HDGG (godišnja članarina iznosi 150 KN). Za članove društva časopis je besplatan.

# INSTRUCTIONS FOR AUTHORS

**SCOPE.** "KoG" publishes scientific and professional papers from the fields of geometry, applied geometry and computer graphics.

**SUBMISSION.** Scientific papers submitted to this journal should be written in English or German, professional papers should be written in Croatian, English or German. The papers have not been published or submitted for publication elsewhere.
The manuscript with wide margins and double spaced should be sent in PDF format via e-mail to the one of the editors:

Sonja Gorjanc
sgorjanc@grad.hr

Ema Jurkin
ema.jurkin@rgn.hr

The first page should contain the article title, authors and coauthors names, affiliation, a short abstract in English, a list of keywords and the Mathematical subject classification.

**UPON ACCEPTANCE.** After the manuscript has been accepted for publication authors are requested to send its LaTeX file via e-mail to the address:

sgorjanc@grad.hr

Figures should be included in EPS or PS formats.

**OFFPRINTS.** The corresponding author and coauthors will receive hard copies of the issue free of charge and a PDF file of the article via e-mail.

---

# UPUTE ZA AUTORE

**PODRUČJE.** "KoG" objavljuje znanstvene i stručne radove iz područja geometrije, primijenjene geometrije i računalne grafike.

**UPUTSTVA ZA PREDAJU RADA.** Znanstveni radovi trebaju biti napisani na engleskom ili njemačkom jeziku, a stručni na hrvatskom, engleskom ili njemačkom. Rad ne smije biti objavljen niti predan na recenziju drugim časopisima.
Rukopis sa širokim marginama i dvostrukim proredom šalje se u PDF formatu elektronskom poštom na adresu jedne od urednica:

Sonja Gorjanc
sgorjanc@grad.hr

Ema Jurkin
ema.jurkin@rgn.hr

Prva stranica treba sadržavati naslov rada, imena autora i koautora, podatke o autoru i koautorima, sažetak na hrvatskom i engleskom, ključne riječi i MSC broj.

**PO PRIHVAĆANJU RADA.** Tekst prihvaćenog rada autor dostavlja elektronskom poštom kao LaTeX datoteku, a slike u EPS ili PS formatu, na adresu:

sgorjanc@grad.hr

**POSEBNI OTISCI.** Svaki autor i koautor dobiva po jedan tiskani primjerak časipisa i PDF datoteku svog članka.

ISSN 1331-1611