SLOVENSKÝ ČASOPIS PRE GEOMETRIU A GRAFIKU
ročník 17, číslo 33, 2020
ISSN 1336-524X

G
Slovenský časopis pre geometriu a grafiku

Ročník 17 (2020), číslo 33
ISSN 1336-524X

Vydáva:
Slovenská spoločnosť pre Geometriu a Grafiku SSGG

G je vedecký časopis pre geometriu a grafiku publikujúci originálne vedecké práce, prehľady a informatívne články, krátke referáty, odborné príspevky, analýzy, aktuality a rešerše z rôznych odvetví geometrie (elementárna, deskriptívna, konštrukčná, projektívna, analytická, diferenciálna, algebrická, počítačová, výpočtová, konečná, neeuklidovská) a topológie, geometrického modelovania a počítačovej grafiky, v oblasti základného teoretického výskumu, v oblasti výučby geometrie na všetkých typoch škôl, z histórie a metodológie vývoja geometrie, a z aplikácií geometrie a geometrických metód v rôznych vedeckých, spoločenských a technických disciplínach.

Redakcia: Slovenská spoločnosť pre Geometriu a Grafiku
IČO: 31 816 304
Ústav matematiky a fyziky, Strojnícka fakulta
Slovenská technická univerzita v Bratislave
Námestie slobody 17
812 31 Bratislava

Objednávky, reklamácie a predplatné vybavuje:
Redakcia G - SSGG
ÚMF SjF STU, Námestie slobody 17, 812 31 Bratislava
ssgg@ssgg.sk

Periodicita: Časopis vychádza dvakrát do roka v náklade 200 kusov.
Ročné predplatné bez poštovného a balného je 20,- Eur.
Evidenčné číslo EV 3228/09

Informácie a pokyny pre autorov na adrese: www.ssgg.sk

Tlačí: ForPress Nitrianske tlačiarne, s.r.o.

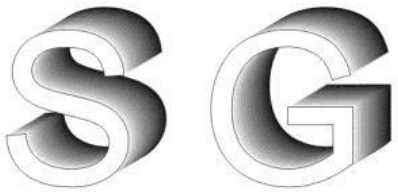
Časopis G je citovaný v: Zentralblatt für Mathematik

SLOVENSKÝ ČASOPIS PRE GEOMETRIU A GRAFIKU
SLOVAK JOURNAL FOR GEOMETRY AND GRAPHICS

ročník 17, číslo 33, 2020
volume 17, number 33, 2020

Obsah – Contents

SLOVENSKÁ SPOLOČNOSŤ



PRE GEOMETRIU A GRAFIKU

Nezisková vedecká spoločnosť pre rozvoj geometrie a počítačovej grafiky

zaregistrovaná dňa 13.5.2002 na Ministerstve vnútra SR ponúka
všetkým záujemcom individuálne alebo kolektívne členstvo.
Elektronickú prihlášku nájdete na domovskej stránke spoločnosti.

Cieľom spoločnosti je stimulovať vedecký výskum, aplikácie i pedagogickú prácu a metodiku vyučovania v oblasti geometrie a počítačovej grafiky.

Spoločnosť pôsobí na celom území Slovenskej republiky a jej poslaním je:

a) podporovať rozvoj geometrie a počítačovej grafiky a ich vzájomnej interakcie

b) presadzovať kvalitu geometrického a grafického vzdelania na všetkých typoch škôl v SR

c) spolupracovať s medzinárodnými spoločnosťami a organizáciami rovnakého zamerania

d) podieľať sa na organizácii vedeckých podujatí, konferencií, seminárov a sympózií o geometrii a počítačovej grafike

e) publikovať vedecký časopis s názvom G venovaný geometrii a grafike

f) rozvíjať vlastnú edičnú a publikačnú činnosť

g) získať priazeň a členstvo organizácií aj jednotlivcov.

Vítané sú všetky ďalšie aktivity – diskusné fórum na Internete, softvérový bazár, workshopy, e-learningové kurzy ai., ktoré možno vykonávať pod hlavičkou spoločnosti.

Spoločnosť SSGG
Ústav matematiky a fyziky
Strojnícka fakulta STU v Bratislave
Námestie slobody 17, 812 31 Bratislava, SR
e-mail: ssgg@ssgg.sk, URL: www.ssgg.sk

Note on determining approximate symmetries of planar algebraic curves with inexact coefficients

Michal Bizzarri, Miroslav Lávička, Jan Vršek

Abstrakt
Tento článok* sa venuje istým modifikáciám nedávno publikovanej metódy aproximačnej rekonštrukcie nepresných rovinných kriviek, ktoré sú považované za perturbácie istých neznámych rovinných súmerných kriviek. Vstupná krivka je určená nepresným polynómom a kroky rekonštrukcie nadväzujú na výsledky nedávno publikovaných článkov [6, 7]. Funkčnosť navrhnutého prístupu je dokumentovaná na niekoľkých konkrétnych príkladoch.

Kľúčové slová: rovinné algebraické krivky, vyhľadávanie súmerností, harmonické polynómy, Laplaceov operátor, aproximácia

Abstract
This paper* is devoted to a certain modification of the recently published method for an approximate reconstruction of inexact planar curves which are assumed to be perturbations of some unknown planar symmetric curves. The input curve is given by a perturbed polynomial and the reconstruction steps follow the results from the recently published papers [6, 7]. The functionality of the designed approach is presented on particular examples.

Keywords: planar algebraic curves, symmetry detection, harmonic polynomials, Laplace operator, approximation

1 Introduction and motivation

This paper is devoted to the symmetries of planar curves with inexact coefficients. Being symmetric is a very useful feature which many real shapes possess and symmetries in the natural world have significantly inspired people when producing tools, buildings, artwork etc. An object has symmetry if there is a transformation (such as translation, rotation, reflection etc.) that maps the object onto itself (i.e., the object has an invariance under the geometric transformation). It is very important to be able to detect symmetry in geometrical models, both from theoretical and practical point of view.

Problems dealing with symmetry detection and computation are often addressed in papers coming from applied fields such as Computer Aided Geometric Design, Pattern Recognition or Computer Vision, see [1, 5, 6] for the exhaustive list of references. In fields such as Patter Recognition or Computer Vision especially the problem of detecting similarity is essential because objects must be recognized regardless of their position and scale. In geometric modelling, symmetry is important on its own right, since it is a distinguished feature of the

*Expanded version of the contribution to the Proceedings of the Slovak–Czech Conference on Geometry and Graphics 2019 (Trenčianske Teplice, September 2019).

shape of an object. Nonetheless it is also important in terms of storing or managing images, because knowing the symmetries of an image allows the machine to reconstruct the object at a lower computational or memory cost.

The symmetry problem has been addressed simultaneously by computer science and mathematics researchers. Whereas computer science typically processes set of points or meshes, mathematics is mainly interested in objects described by equations. Recent research has focused for instance on efficient algorithms for finding congruences and symmetries of large point sets generated by 3D scans. The computation of symmetries and equivalences of rational algebraic varieties also experienced a significant increase of interest as these objects are very important in geometric modelling and related applications. One can find many papers devoted to the detection and computation of symmetries and some equivalences of curves, see e.g. [9, 8, 11, 10], or recent series of papers [1, 2, 3, 4, 5]. The problem of deterministically computing the symmetries of a given planar algebraic curve was recently studied in [6].

As mentioned before, many real world shapes exhibit a symmetry. However, in most cases this symmetry is not perfect but only approximate – which may happen, for instance, when some input error (or some error caused by numerical computations) occurs. And, of course, in this situations all subsequent exact algorithms and scenarios formulated for algebraic curves with symmetries fail. Recently, see [7], we designed an algorithm for an approximate reconstruction of an inexact planar curve which is assumed to be a perturbation of some unknown planar curve. The initial step of the reconstruction algorithm is to find a suitable approximate centre of symmetry and a particular regular m-gon to whose group of symmetries the group of symmetries of the curve is isomorphic. In this paper, we modify the part devoted to finding the approximate centre of symmetry and present an alternative approach that more closely matches the original exact algorithm based on computing with Laplace operator, cf. [6].

The rest of the paper is organized as follows. Section 2 recalls some basic facts concerning algebraic curves and their symmetries. We also recall the approach that uses the Laplace operator for determining symmetries of algebraic curves. Section 3 is devoted to the modification of the method formulated originally for exact algebraic curves. The designed method is presented on several examples in Section 4. Finally, we conclude the paper in Section 5.

2 Preliminaries

First we recall selected elementary notions, basic properties and suitable methods whose knowledge is further assumed.

2.1 Symmetric algebraic curves in plane

A *planar algebraic curve* $\mathcal{C}$ is a subset of $\mathbb{E}^2_{\mathbb{R}}$ defined as the zeroset of a polynomial $f(x, y)$. We will assume that f has real coefficients, is irreducible over $\mathbb{C}$ and $\dim_{\mathbb{R}} \mathcal{C} = 1$. Any *isometry* $\phi \in \mathbf{Iso}_2$ of $\mathbb{E}^2_{\mathbb{R}}$ possesses the form $\mathbf{x} \mapsto \mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A} \in \mathbf{O}(\mathbb{R}, 2)$ and $\mathbf{b} \in \mathbb{R}^2$. For $\det(\mathbf{A}) = 1$, or $= -1$ we speak about *direct*, or *indirect* isometries, respectively.

We write $\mathbf{Sym}(\mathcal{C})$ for the group of symmetries of the curve $\mathcal{C}$, i.e.,

$$\mathbf{Sym}(\mathcal{C}) := \{\phi \in \mathbf{Iso}_2;\ \phi(\mathcal{C}) = \mathcal{C}\}. \tag{1}$$

It is well known that $\mathbf{Sym}(\mathcal{C})$ is finite unless $\mathcal{C}$ is a union of parallel lines or a union of concentric circles. Moreover, if $\mathbf{Sym}(\mathcal{C})$ is finite then it is isomorphic to a subgroup of the group of symmetries of some regular m-gon, $m \leq \deg(\mathcal{C})$. In what follows we are interested solely in curves with a finite group of symmetries. The elements of a finite symmetry group are rotations (all of them with the same center) and reflections (axes of all of them passing through the same point).

Analogously we introduce $\mathbf{Sym}(f)$ by the relation

$$\mathbf{Sym}(f) := \{\phi \in \mathbf{Iso}_2;\ f \circ \phi = \lambda f\}, \tag{2}$$

where $\lambda \neq 0$ is a constant.

We recall the following statement, which can be efficiently used to verify whether $\phi \in \mathbf{Sym}(\mathcal{C})$, see [6] for more details:

Proposition 2.1. An isometry $\phi \in \mathbf{Sym}(\mathcal{C})$ if and only if $f(\mathbf{Ax} + \mathbf{b}) = \lambda f(\mathbf{x})$, where $\lambda = 1$ or $\lambda = -1$.

2.2 Symmetries of planar curves via harmonic polynomials

We start with recalling the exact approach which has been formulated recently. For the sake of brevity we will mention only basic steps and a generic scenario; the reader who is more interested in this topic is kindly referred to [6], where all proofs and further explanations can be found.

In general, it is not easy to find symmetries ϕ belonging to $\mathbf{Sym}(\mathcal{C})$ directly and one has to apply a suitable computational approach – for instance to find some new polynomial $h(x, y)$ such that $\mathbf{Sym}(h)$ is finite, easy to determine (i.e., easier then $\mathbf{Sym}(f)$) and $\mathbf{Sym}(\mathcal{C}) = \mathbf{Sym}(f) \subset \mathbf{Sym}(h)$. In [6], a successive application of the Laplace operator yielding the sequence

$$f \longmapsto \triangle f \longmapsto \triangle^2 f \longmapsto \cdots \longmapsto \triangle^\ell f = h, \tag{3}$$

and followed by the associated chain of groups of symmetries

$$\mathbf{Sym}(f) \subset \mathbf{Sym}(\triangle f) \subset \mathbf{Sym}(\triangle^2 f) \subset \cdots \subset \mathbf{Sym}(\triangle^\ell f) = \mathbf{Sym}(h), \tag{4}$$

was efficiently used for finding such a polynomial h. Application of this technique is justified by the fact that the *Laplace operator* as a linear mapping $\triangle : \mathbb{R}[x, y] \to \mathbb{R}[x, y]$ defined by

$$\triangle f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{5}$$

commutes with isometries, i.e., it holds

$$(\triangle f) \circ \phi = \triangle(f \circ \phi). \tag{6}$$

A polynomial h satisfying $\triangle h = 0$ is called *harmonic*. By repeatedly computing the Laplacian, cf. (3), in general we come down to either harmonic polynomials, or conic sections, or lines. All situations are discussed in the original paper, here we recall only the most interesting part, i.e.,

when one arrives at a harmonic polynomial h. We recall that if h is harmonic and $\deg(h) > 1$ then $\mathbf{Sym}(h)$ is finite.

Next, we identify $\mathbb{C}$ with $\mathbb{R}^2$ via $z = x + \mathrm{i}y \leftrightarrow (x, y)$. For a polynomial $h(x, y)$ we consider a complex function

$$g(x, y) = \partial_x h - \mathrm{i}\partial_y h, \tag{7}$$

where $\partial_x h, \partial_y h$ represent the partial derivatives of h with respect to x, y. The standard substitution

$$x = \frac{1}{2}(z + \overline{z}) \qquad \text{and} \qquad y = -\frac{\mathrm{i}}{2}(z - \overline{z}) \tag{8}$$

allows to write $g(x, y)$ as a complex function $g(z, \overline{z})$ in the complex variable z. Moreover, as h is harmonic then $g(x, y)$ satisfies the Cauchy-Riemann conditions and thus $g(x, y)$ is holomorphic and $g(z, \overline{z})$ does not depend on $\overline{z}$, i.e.,

$$g(z, \overline{z}) = g(z) = \sum_{j=0}^{\delta} b_j z^j. \tag{9}$$

The roots of $g(z)$ yield the singular points of the vector field $(\partial_x h, -\partial_y h)$. As any $\phi \in \mathbf{Sym}(h)$ maps real singular points of the considered vector field onto real singular points of this field, we finally obtain

$$\mathbf{Sym}(h) \subset \mathbf{Sym}(\Sigma), \tag{10}$$

where $\Sigma = \{\zeta_1, \ldots, \zeta_\delta\} \subset \mathbb{C}$ is the set of all roots of $g(z)$ (counted with multiplicity). Symmetries of $h(x, y)$ are then derived from Σ, resp. $g(z)$. For instance, a possible center of any rotational symmetry of $h(x, y)$ is encoded in the barycenter of Σ, i.e.,

$$\mathbf{p} = \frac{1}{\delta} \sum_{i=1}^{\delta} \zeta_i. \tag{11}$$

In addition, using Vieta's formulas on $g(z)$, one can see that the computation of the roots is not necessary and we obtain

$$\mathbf{p} = -\frac{b_{\delta-1}}{\delta b_\delta}. \tag{12}$$

Potential candidates for the rotation angle are of the type $\frac{2\pi}{m}$, where $m \leq \delta + 1 = \deg(h)$.

Similarly, a method how to determine the potential axes of symmetry of $h(x, y)$ from the coefficients of $g(z)$ is also presented in [6].

3 Formulation of the problem and modified algorithm

In paper [6] exact symmetries of algebraic curves in plane were studied. Recently, this problem has been extended in [7] also to approximate symmetries. In latter case, the input to the

algorithm is a planar curve $\mathcal{C}$ which is a perturbation of some unknown symmetric planar curve $\mathcal{C}_0$. This perturbed curve is described by a polynomial $f(x,y)$ of degree d, i.e.,

$$\mathcal{C}:\ f(x,y)=\sum_{\substack{i,j\geq 0\\ i+j\leq d}}^{d} a_{i,j}x^i y^j=0,\quad a_{ij}\in\mathbb{R}. \tag{13}$$

For various purposes, it is often useful to consider curves in the projective plane. Every affine algebraic curve of equation $f(x,y)=0$ may be completed into the projective curve of equation $F(X,Y,Z)=0$, where

$$F(X,Y,Z)=Z^d f\left(X/Z,Y/Z\right)=\sum_{\substack{i,j,k\geq 0\\ i+j+k=d}}^{d} a_{i,j}X^iY^jZ^k \tag{14}$$

is the result of the homogenization of f and $X:Y:Z$ are the homogeneous coordinates in the projective plane. Let us write $\mathbf{c}=(a_{d,0}:a_{d-1,1}:\cdots:a_{00})$ and we say that $\mathbf{c}$ represents $\mathcal{C}$. The space of all planar projective curves of degree d can be identified with the projective space $\mathbb{P}_{\mathbb{R}}^{N-1}$, where $N=\binom{d+2}{2}$.

The perturbed curve $\mathcal{C}$ possesses no symmetries. Nonetheless, the original curve $\mathcal{C}_0$ was by assumption symmetric and thus using the exact approach, recalled in the previous section, one could arrive at a distinguished point $\mathbf{p}$ (a center of any possible rotation, or a point through which the axes of reflection are passing). The following strategy for approximate reconstruction of $\mathcal{C}_0$ was suggested in [7] (for more details see the original reference):

- (a) Determine a point $\widetilde{\mathbf{p}}$ (the approximate center) and an integer m (the number of vertices of a regular polygon) from the known perturbed curve $\mathcal{C}$;
- (b) Construct a new curve $\widetilde{\mathcal{C}}$ having the symmetry of an m-gon with the center at $\widetilde{\mathbf{p}}$ and being as close as possible to the given perturbed curve $\mathcal{C}$.
- (c) Determine all the symmetries of the computed exact symmetric curve $\widetilde{\mathcal{C}}$ to obtain the approximate symmetries of the perturbed curve $\mathcal{C}$.

In this paper we focus on the crucial part of the algorithm and formulate an alternative approach for determining a suitable approximate center of symmetry $\widetilde{\mathbf{p}}$ of the resulting curve $\widetilde{\mathcal{C}}$, i.e., we will deal with step (a) solely. Computing m is not part of this modified approach – one has to consider all m from 2 to d and consequently choose the best approximation. The remaining parts of the original algorithm remain the same.

Unlike in [7], we formulate the approach based on applying a sequence of Laplacians, see (3) – which was the method used originally in paper on exact symmetries, cf. [6]. From this reason we assume that the original symmetric curve $\mathcal{C}_0$ was transformable by the chain of Laplacians to a harmonic curve satisfying (3). As another new contribution, we solve the problem using complex variables.

First we substitute (8) into $f(x,y)$ which allows to write it as a complex function $f(z,\overline{z})$ in the complex variable z in the form

$$f(z,\overline{z})=\sum_{i=0}^{d}\sum_{j=0}^{d-i} m_{i,j}z^i\overline{z}^j, \tag{15}$$

or in the following matrix form

$$f(z,\bar{z}) = (1, z, z^2, \ldots, z^d)\,\mathbf{M}\begin{pmatrix} 1 \\ \bar{z} \\ \bar{z}^2 \\ \vdots \\ \bar{z}^d \end{pmatrix}, \tag{16}$$

where $m_{i,j} = \overline{m}_{j,i} \in \mathbb{C}$, and $m_{i,i} = \overline{m}_{i,i} \in \mathbb{R}$, cf. [12] for further details.

Let us emphasize that, in this case, $\mathbf{M}$ is a Hermitian matrix with a zero submatrix $\mathbf{0}_{(d-\ell)\times(d-\ell)}$, i.e., it possesses the following structure

$$\mathbf{M} = \begin{pmatrix} m_{0,0} & \overline{m}_{1,0} & \ldots & \overline{m}_{\ell,0} & \overline{m}_{\ell+1,0} & \ldots & \overline{m}_{k,0} & \overline{m}_{k+1,0} & \ldots & \overline{m}_{d-1,0} & \overline{m}_{d,0} \\ m_{1,0} & m_{1,1} & \ldots & \overline{m}_{\ell,1} & \overline{m}_{\ell+1,1} & \ldots & \overline{m}_{k,1} & \overline{m}_{k+1,1} & \ldots & \overline{m}_{d-1,1} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots \\ m_{\ell,0} & m_{\ell,1} & \ldots & m_{\ell,\ell} & \overline{m}_{\ell+1,\ell} & \ldots & \overline{m}_{k,\ell} & 0 & \ldots & 0 & 0 \\ m_{\ell+1,0} & m_{\ell+1,1} & \ldots & m_{\ell+1,\ell} & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ m_{\ell+2,0} & m_{\ell+2,1} & \ldots & m_{\ell+2,\ell} & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{k,0} & m_{k,1} & \ldots & m_{k,\ell} & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ m_{k+1,0} & m_{k+1,1} & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ m_{d-1,0} & m_{d-1,1} & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \\ m_{d,0} & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 & \ldots & 0 & 0 \end{pmatrix},$$

The reason why $\mathbf{M}$ contains the block of zeros follows from the assumption that the chain (3) ends with a harmonic polynomial and from fact how the Laplacian operator works in complex variables. i.e.,

$$\triangle f(z,\bar{z}) = 4\frac{\partial}{\partial z}\frac{\partial f}{\partial \bar{z}}. \tag{17}$$

Therefore we obtain

$$\triangle f(z,\bar{z}) = (1, z, z^2, \ldots, z^{d-2})\,\mathbf{M}_1\begin{pmatrix} 1 \\ \bar{z} \\ \bar{z}^2 \\ \vdots \\ \bar{z}^{d-2} \end{pmatrix}, \tag{18}$$

where $\mathbf{M}_1$ is of the form

$$\mathbf{M}_1 = 4\begin{pmatrix} m_{1,1} & \dots & \ell\,\overline{m}_{\ell,1} & (\ell+1)\,\overline{m}_{\ell+1,1} & \dots & k\,\overline{m}_{k,1} & (k+1)\,\overline{m}_{k+1,1} & \dots & (d-1)\,\overline{m}_{d-1,1} \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \ell\,m_{\ell,1} & \dots & \ell^2\,m_{\ell,\ell} & \ell(\ell+1)\,\overline{m}_{\ell+1,\ell} & \dots & \ell k\,\overline{m}_{k,\ell} & 0 & \dots & 0 \\ (\ell+1)\,m_{\ell+1,0} & \dots & (\ell+1)\ell\,m_{\ell+1,\ell} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ k\,m_{k,0} & \dots & k\ell\,m_{k,\ell} & 0 & \dots & 0 & 0 & \dots & 0 \\ (k+1)\,m_{k+1,0} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (d-1)\,m_{d-1,0} & \dots & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Hence, the chain of Laplacians (3) can be replaced by the chain of matrices

$$\mathbf{M} \longmapsto \mathbf{M}_1 \longmapsto \mathbf{M}_2 \longmapsto \cdots \longmapsto \mathbf{M}_\ell, \tag{19}$$

where the matrix $\mathbf{M}_\ell$ of a harmonic polynomial of degree $k-\ell$ has the form

$$\mathbf{M}_\ell = 4^\ell \ell^2 \begin{pmatrix} \binom{\ell}{\ell} m_{\ell,\ell} & \binom{\ell+1}{\ell} \overline{m}_{\ell+1,\ell} & \dots & \binom{k}{\ell} \overline{m}_{k,\ell} & 0 & \dots & 0 \\ \binom{\ell+1}{\ell} m_{\ell+1,\ell} & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \binom{k}{\ell} m_{k,\ell} & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{20}$$

Let us recall that it was assumed that the sequence is ending with a harmonic polynomial. Then it is evident that the harmonic polynomial $h = \triangle^\ell f$ in the chain (including the values of k, ℓ) can be easily identified from the position of the block of zeros in the original matrix $\mathbf{M}$. Moreover, we will see that the center can be decoded from the matrix $\mathbf{M}$, as well.

Following the previous approach we write polynomial (9) associated to the harmonic polynomial h given by the matrix $\mathbf{M}_\ell$. It holds

$$\frac{\partial h}{\partial z} = \frac{1}{2}\left(\partial_x h - \mathrm{i}\partial_y h\right), \tag{21}$$

and thus we obtain

$$g(z) = 2\frac{\partial h}{\partial z} = 2 \cdot 4^\ell \ell^2 \sum_{i=0}^{k-\ell-1} (i+1) \binom{i+\ell+1}{\ell} m_{i+\ell+1,\ell} z^i. \tag{22}$$

Finally using expression (12) we arrive at the center of symmetry of the curve $\mathcal{C}$

$$\mathbf{p} = \frac{-1}{k-\ell-1} \cdot \frac{(k-\ell-1)\binom{k-1}{\ell} m_{k-1,\ell}}{(k-\ell)\binom{k}{\ell} m_{k,\ell}} = -\frac{m_{k-1,\ell}}{k\,m_{k,\ell}}. \tag{23}$$

Next we consider a perturbation of the original symmetric curve. This influences also the matrix $\mathbf{M}$ which contains a block of "almost zeros", now. Our goal is to identify this almost-zero-

submatrix and set it as zero matrix. This yields a new curve $\widetilde{\mathcal{C}}$ described by the equation

$$\widetilde{\mathcal{C}} : (1, z, z^2, \dots, z^d) \begin{pmatrix} m_{0,0} & \dots & \overline{m}_{\ell,0} & \dots \\ \vdots & \ddots & \vdots & \ddots \\ m_{\ell,0} & \dots & m_{\ell,\ell} & \dots \\ \vdots & \ddots & \vdots & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \bar{z} \\ \bar{z}^2 \\ \vdots \\ \bar{z}^d \end{pmatrix} = 0. \tag{24}$$

Then we continue as in the exact case, determine the point (23) and set it as the approximate center $\widetilde{\mathbf{p}}$.

Moreover, the previous result implies that the perturbation of the center is not worsen by applying the sequence of Laplacians and it respects the order of perturbation of the coefficients of the original curve. For this purpose, we recall some details dealing with the error propagation during computing with inexact quantities. Consider $A = a + \alpha$, $B = b + \beta$, where $\alpha \ll a$, $\beta \ll b$ and $|\alpha| \leq \epsilon$, $|\beta| \leq \epsilon$. Then it holds

$$\left| \frac{A}{B} - \frac{a}{b} \right| \lesssim \frac{(a+b)\epsilon}{b^2} \tag{25}$$

and we can formulate the following assertion.

Lemma 3.1. For the error ϵ_1 of the centre of the symmetric curve whose coefficients are given with maximal error ϵ it holds

$$\epsilon_1 \lesssim \frac{(m_{k_1,\ell} + m_{k,l})\epsilon}{k^2 m_{k,\ell}^2}. \tag{26}$$

The final step of the reconstruction algorithm is to find a suitable symmetric curve $\widetilde{\mathcal{C}}$ sufficiently "close" to the given perturbed curve $\mathcal{C}$ when the center $\widetilde{\mathbf{p}}$ of $\widetilde{\mathcal{C}}$ is prescribed. From this part, we may follow the approach designed in [7]. In particular, we construct a basis of all curves of degree d with the rotational symmetry of m-gon and with the center of rotation $\widetilde{\mathbf{p}}$, and compute the orthogonal projection of the perturbed curve to the space spanned by the spanned basis, see [7] for all necessary details.

To measure suitably a quality of the approximation (i.e., the deviation δ between the perturbed and the constructed curve) we will apply the standard metric used for computing the distance between the points $\widetilde{\mathbf{c}}$, $\mathbf{c}$ in the projective space of algebraic curves of degree d, in particular

$$\delta(\widetilde{\mathbf{c}}, \mathbf{c}) = \arccos\left(\frac{|\widetilde{\mathbf{c}} \cdot \mathbf{c}|}{\|\widetilde{\mathbf{c}}\| \|\mathbf{c}\|} \right), \tag{27}$$

where '$\cdot$' and $\| \ \|$ denote the standard inner product and the standard norm in the corresponding vector space. The angle δ is real-valued, and runs from 0 to $\frac{\pi}{2}$.

4 Computed examples

In this section we present the designed modification of the original approach from [7] on some commented examples.

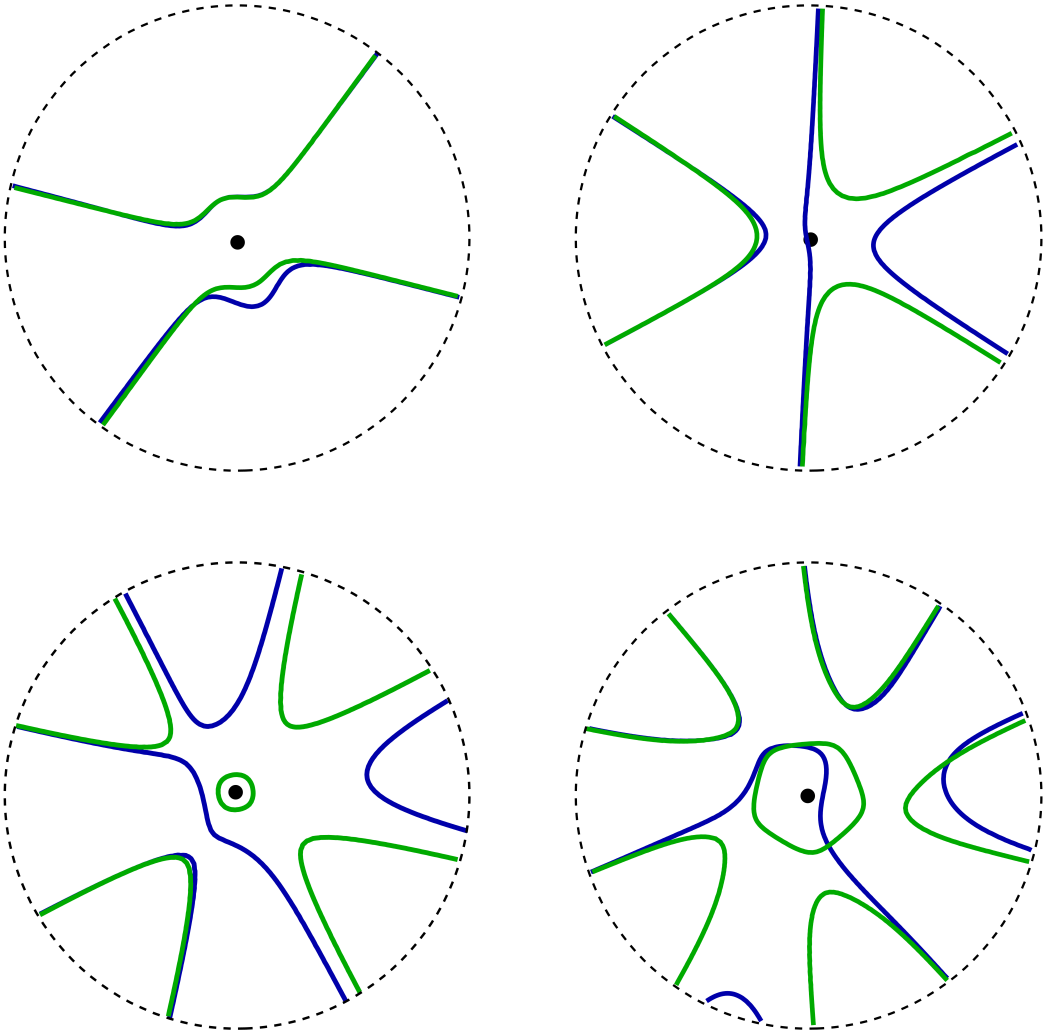


Fig. 1. Four non-symmetric curves (blue) given by perturbing the unknown symmetric curves and the corresponding closest symmetric curves (green) with the guessed centers (black) of the symmetries.

Example: Consider a quartic curve $\mathcal{C}$ given by a polynomial with floating coefficients (which is a perturbation of a certain unknown symmetric curve), see Fig. 1 (top, left)

$$f = 13x^2y^2 - 64x^3y - 167.4x^2y - 17x^4 - 136.4x^3 - 289.4x^2 + 15.4xy^3 \\ + 73xy^2 - 96.5xy - 227.1x + 13.2y^4 + 72.5y^3 + 157.7y^2 + 91.6y - 39.4.$$

First, we transform f into the complex representation, cf. (8), and use the matrix form (16) – for the sake of compactness we display the coefficients of the matrices with one decimal place only.

$$\mathbf{M} \approx \begin{pmatrix} -39.4 & -113.6+45.8\mathrm{i} & -111.8-24.1\mathrm{i} & -26.2-30\mathrm{i} & -1.1-5\mathrm{i} \\ -113.6-45.8\mathrm{i} & -65.8 & -42+6.3\mathrm{i} & -7.6-6.1\mathrm{i} & 0 \\ -111.8+24.1\mathrm{i} & -42-6.3\mathrm{i} & 0.2 & 0 & 0 \\ -26.2+30\mathrm{i} & -7.6+6.1\mathrm{i} & 0 & 0 & 0 \\ -1.1+5\mathrm{i} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Next, we find the maximal almost-zero-submatrix in $\mathbf{M}$ and create a new one with this

submatrix being exactly-zero.

$$\begin{pmatrix} -39.4 & -113.6+45.8\mathrm{i} & -111.8-24.1\mathrm{i} & -26.2-30\mathrm{i} & -1.1-5\mathrm{i} \\ -113.6-45.8\mathrm{i} & -65.8 & -42+6.3\mathrm{i} & -7.6-6.1\mathrm{i} & 0 \\ -111.8+24.1\mathrm{i} & -42-6.3\mathrm{i} & 0 & 0 & 0 \\ -26.2+30\mathrm{i} & -7.6+6.1\mathrm{i} & 0 & 0 & 0 \\ -1.1+5\mathrm{i} & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence we have $\ell = 1$ and $k = 3$ and using (23) we obtain an approximate center of symmetry

$$\mathbf{p} \doteq (-0.991, -1.074). \tag{28}$$

Subsequently projecting $\mathcal{C}$ to all curves with the symmetry of m-gon with the center $\mathbf{p}$, where $m \in \{2, 3, 4\}$, we obtain the best solution for $m = 2$, see Fig. 1. The deviation angle (27) is approximately equal to 0.35.

Example: We have a non-symmetric cubic curve $\mathcal{C}$, see Fig. 1 (top, right) given by the polynomial

$$f = -6.8x^2y + 17.1x^3 - 102.1x^2 - 53.9xy^2 + 131.7xy \\ + 138x + 2.4y^3 + 101.3y^2 - 233.3y - 3.7.$$

The matrix form (16) of the complex representation of f looks as follows (we again display the coefficients of the matrices with one decimal place only)

$$\mathbf{M} \approx \begin{pmatrix} -3.7 & 69-116.6\mathrm{i} & -50.8+32.9\mathrm{i} & 8.9-1.1\mathrm{i} \\ 69+116.6\mathrm{i} & -0.4 & -0.3 & 0 \\ -50.8-32.9\mathrm{i} & -0.3 & 0 & 0 \\ 8.9+1.1\mathrm{i} & 0 & 0 & 0 \end{pmatrix}.$$

A matrix with maximal almost-zero-submatrix in $\mathbf{M}$ being exactly-zero is

$$\begin{pmatrix} -3.7 & 69-116.6\mathrm{i} & -50.8+32.9\mathrm{i} & 8.9-1.1\mathrm{i} \\ 69+116.6\mathrm{i} & 0 & 0 & 0 \\ -50.8-32.9\mathrm{i} & 0 & 0 & 0 \\ 8.9+1.1\mathrm{i} & 0 & 0 & 0 \end{pmatrix}.$$

Therefore we have $\ell = 0$, $k = 3$ and we arrive at the approximate center of symmetry

$$\mathbf{p} \doteq (2.035, 0.972). \tag{29}$$

Finally projecting $\mathcal{C}$ to all curves with the symmetry of m-gon with the center $\mathbf{p}$, where $m \in \{2, 3\}$, we obtain the best solution for $m = 3$, see Fig. 1. The deviation angle (27) is approximately equal to 0.39.

Example: We demonstrate the presented approach on further two curves of degrees four and seven, respectively. See Fig. 1, bottom. In those cases we arrive at curves with symmetries of square and pentagon, respectively. The deviation between the given non-symmetric curves and the computed symmetric ones are 0.751 and 0.273, respectively.

5 Conclusion

In this paper, we studied and designed a certain modification of the recently presented method for an approximate reconstruction of a planar algebraic curves with inexact coefficients, being

a perturbation of some unknown (originally) symmetric planar algebraic curve. We focused solely on the initial step of the algorithm from [7] which is devoted to computing a suitable approximate centre of symmetry and a particular regular m-gon to whose group of symmetries the group of symmetries of the curve is isomorphic. This modified method suitably uses, as the algorithm for the exact case, cf. [6], the sequence of Laplacians, which is an operator reducing the degree of the input polynomial and preserving symmetries. The functionality of the designed approach was illustrated on several examples. The readers interested in this topic are kindly referred to [6, 7] where they can find more details.

Acknowledgments

The authors were supported by the project LO1506 of the Czech Ministry of Education, Youth and Sports.

References

[1] Alcázar, J. G. Efficient detection of symmetries of polynomially parametrized curves. *Journal of Computational and Applied Mathematics*, volume 255, 2014: p. 715 – 724.

[2] Alcázar, J. G., Hermoso, C., Muntingh, G. Detecting similarity of rational plane curves. *Journal of Computational and Applied Mathematics*, volume 269, 2014: p. 1 – 13.

[3] Alcázar, J. G., Hermoso, C., Muntingh, G. Detecting symmetries of rational plane and space curves. *Computer Aided Geometric Design*, volume 31 (3), 2014: p. 199 – 209.

[4] Alcázar, J. G., Hermoso, C., Muntingh, G. Symmetry detection of rational space curves from their curvature and torsion. *Computer Aided Geometric Design*, volume 33, 2015: p. 51 – 65.

[5] Alcázar, J. G., Hermoso, C., Muntingh, G. Similarity detection of rational space curves. *Journal of Symbolic Computation*, volume 85, 2018: p. 4 – 24, 41th International Symposium on Symbolic and Algebraic Computation (ISSAC16).

[6] Alcázar, J. G., Lávička, M., Vršek, J. Symmetries and similarities of planar algebraic curves using harmonic polynomials. *Journal of Computational and Applied Mathematics*, volume 357, September 2019: p. 302–318.

[7] Bizzarri, M., Lávička, M., Vršek, J. Approximate symmetries of planar algebraic curves with inexact input. *Computer Aided Geometric Design*, volume 76, January 2020, 1017942019.

[8] Brass, P., Knauer, C. Testing congruence and symmetry for general 3-dimensional objects. *Computational Geometry*, volume 27 (1), 2004: p. 3 – 11, computational Geometry - EWCG'02.

[9] Huang, Z., Cohen, F. S. Affine-invariant B-spline moments for curve matching. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 1994, p. 490 – 495.

[10] Lebmeir, P. *Feature detection for real plane algebraic curves*. PhD Thesis, Technische Universität München, 2009.

[11] Lebmeir, P., Richter-Gebert, J. Rotations, translations and symmetry detection for complexified curves. *Computer Aided Geometric Design*, volume 25(9), 2008: p. 707 – 719, classical Techniques for Applied Geometry.

[12] Tarel, J.-P., Cooper, D. B. The complex representation of algebraic curves and its simple exploitation for pose estimation and invariant recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 22(7), 2000: p. 663 – 674.

RNDr. Michal Bizzarri, PhD.

NTIS – New Technologies for the Information Society
Faculty of Applied Sciences,
University of West Bohemia
Univerzitní 8, 301 00 Plzeň, Czech Republic
e-mail: bizzarri@ntis.zcu.cz

Doc. RNDr. Miroslav Lávička, PhD.

Department of Mathematics & NTIS
Faculty of Applied Sciences,
University of West Bohemia
Univerzitní 8, 301 00 Plzeň, Czech Republic
e-mail: lavicka@kma.zcu.cz

RNDr. Jan Vršek, PhD.

Department of Mathematics & NTIS
Faculty of Applied Sciences,
University of West Bohemia
Univerzitní 8, 301 00 Plzeň, Czech Republic
e-mail: vrsekjan@kma.zcu.cz

Sierpinski's curve: a (beautiful) paradigm of recursion

Paola Magrone

Abstrakt

Článok nadväzuje na pôvodné práce Waclawa Sierpinskeho z roku 1915, kedy predstavil rekurzívnu štruktúru, ktorá nesie jeho meno, Sierpinského trojuholník**.** Jeho pôvodným zámerom bolo nájsť príklad novej množiny, krivky vytvorenej na základe známej geometrie trojuholníkov. Tento trojuholník, ktorý obsahuje gemetrickú rekurziu, bol presne definovaný v roku 1915, ale objavil sa už aj pred Sierpinským, a je doteraz referenčným bodom pre vedcov.

Kľúčové slová: Sierpinského trojuholník, rekurzia, Cantorova krivka, Jordanova krivka

Abstract

This paper focuses on the original articles written by Waclaw Sierpinski in 1915, when he introduced the recursive structure that bears his name, the Sierpinski's triangle. His first aim was to exhibit the example of a new set, a curve traced starting from the geometry of the well-known triangle. The triangle, which embodies geometric recursion, was rigorously defined in 1915, but appeared also before Sierpinski, and is still a reference point for scientists.

Keywords: Sierpinski's triangle, recursion Cantorian curve, Jordanian curve

1 New definitions for new mathematical objects

At the end of Nineteenth century, the community of mathematicians gave birth to many contributions in the theory of sets and structure of numbers: new objects were defined, such the concepts of accumulation point and limit of a sequence of numbers. The Real numbers, their definition and structure as we know and use them today, is mainly due to the work developed during those years by Richard Dedekind (1831-1916) and Georg Cantor (1845–1918). Waclaw Sierpinski (1882-1969) begun his mathematical studies in the theory of numbers, and turned to set theory after becoming acquainted with Cantor theories. Sierpinski in 1909 started teaching the first course on set theory, which gained year after year a greater importance; he gave many important contributions to the growth of this discipline and to its sistematization. He was one of the founders of the Polish mathematical school which put roots in those new theories and carried them on [3, 7, 20 tome 1].

This paper focuses on the original articles written by Sierpinski in 1915, when he introduced the recursive structure that bears his name, the Sierpinski's triangle. His aim was to exhibit the example of a curve with such surprising and counterintuitive properties, to give rise to the need to rethink the definition of curve.

This article is organized as follows: this first section is devoted to describe which were the available definition of curve at the beginning of Twentieth century. Section 2 focuses on the Sierpinski's curve itself, and its mathematical properties. Section 3 describes some examples of the use of the Sierpinski's triangle in art and science.

At the beginning of Twentieth century, there were two available definitions of a plane curve: the Cantorian and the Jordanian one. A *Cantorian curve* is a planar continuum (a set contained in the plane, closed, connected, containing more than one point) which has empty interior (each point belonging to the set is a frontier point). A *Jordanian curve* is the image of a segment of straight line via a continuous function, not necessarily biunivocal [12, pg. 90]. Cantorian and Jordanian curves are topologically invariant [18], i.e., they are transformed in Cantorian curves and Jordanian curves respectively by a continuous function, with continuous inverse.

> *As we know, there is no definition of a curve which conforms to intuition without being too vast or too narrow. The difficulty already arises for plane curves and even more so in the three-dimensional space where we are not at, even until now, drawing a limit between the notion of line and that of surface [19].*

In the articles dating 1915-16 [18,19] the author points out the urgency of giving a definition of curve which could match with intuition and would also remain sound. For this purpose he deliberately attacked these two definitions, to foster the scientific community to rethink and improve them.

> *The existence of curves filling the square shows that the definition by Jordan is too wide, because it embraces geometric figures which our intuition refuses to call line. But Cantor's definition is itself too wide. […] There exist indeed Cantorian curves in which no couple of points can be connected by a simple arc [19].*

Sierpinski refers firstly to the Peano's curve which is Jordanian, and not Cantorian for the reason it fills the unit square. The second example he refers to, is the set ("the topological sine") $A = A_1 \cup A_2$ where

$$A_1 = \{(0, y): y \in [-1, 1]\}, \quad A_2 = \{(x, y): y = \sin 1/x : x \in (0, 1/\pi]\}.$$

This set is closed, and connected, but non path connected: the points of A_1 cannot be connected with points of A_2 by a simple arc, which is the continuous image of a segment of line [13]. It is Cantorian, but not Jordanian. The optimal definition of curve could be to include both characterizations, "But even then, we come up against some very striking surprises" [19].

2 Sierpinski's curve

The aim of the paper "*Sur une courbe dont tout point est un point de ramification*" was to show a (paradoxical) example of a curve, Cantorian and Jordanian at the same time, having all points as ramification points. Sierpinski defines a *ramification point* of a continuum C to be a point $p \in C$ such that there exist three subsets (all continua) of C having in common, two by two, only p. We point out that this definition led to that of *order of ramification* of a point (given by P. Urysohn and K. Menger a few years after Sierpinski's papers [14]). Roughly speaking, the order of a point p in X is the number of lines meeting at p. To count this number one can take an arbitrary small circular neighborhood of p and count the intersections of the set X with the neighborhood.

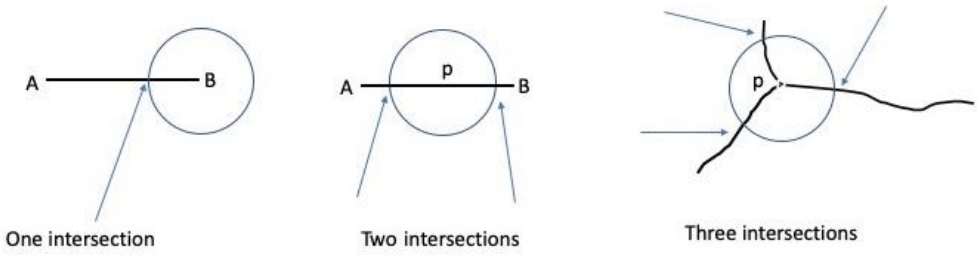


Fig. 1. Order of ramification of a point. Left: the end point B has order one. Center: point p has order two. Right, the point *p* has order three. Pictures by the author.

Sierpinski, then, constructed this paradoxical curve, starting from the triangle which bears his name. Take an equilateral triangle (which side has, for example, length 1), join the middle points of the three edges, obtaining 4 equilateral triangles (for other and more details: [4, 6, and 19]). The interior of the central triangle, the one not containing either of the three vertices A, B or C, is erased. In Fig. 2 (original drawings of Sierpinski) left, the black central triangle is the void; in the same figure, right, the second iteration; at each step the black triangles are the one representing void parts. Observe that in the vast majority of visualizations produced after 1915, by hand or by computer, colours are used in a complementary way: black for full areas, white for voids.

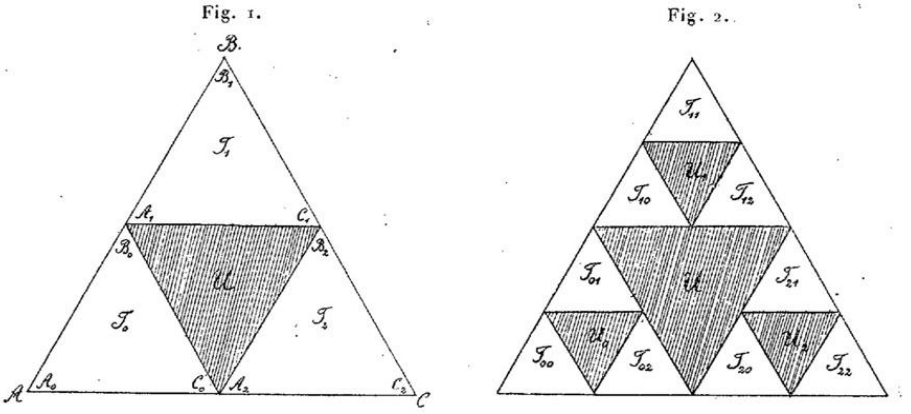


Fig. 2. From the original article of Sierpinski, 1915. Left, level 1 iteration, right, level 2 iteration.

There are now three triangles, around a central "void". Iterate the procedure on the remaining three triangles: mark the middle points of the edges, join them, obtain 4 triangles; discard the central one, keep the other three. At level one, since we obtain 4 triangles and discard one, we are left with 3^1 triangles; level two shows 3^2 triangles. At level n there will be a set of 3^n triangles, equilateral and identical. Carrying on this procedure to the limit, as n goes to infinity, the intersection of all the sets obtained at each iteration, that is $S = \cap_n S_n$ yields to the definition of Sierpinski's triangle (all the points not belonging to the discarded voids, at each step). The set S is closed, connected, contains at least one point, so it is a continuum, and has empty interior, so it matches with the definition of Cantorian curve [19]. A property clearly showed by this geometric construction is the self-similarity of S: it contains copies of itself, at many (infinite, when going to the limit) different sizes.

Sierpinski's triangle has many interesting dimension and measure properties. It has a non-integer dimension, which can be calculated as follows. The "box-counting" dimension [15,16] of a set T is the following limit (if it exists):

$$\mathrm{Dim}(T) = -\lim (\ln (N_\varepsilon)/\ln (\varepsilon)), \text{ as } \varepsilon \to 0.$$

Where N_ε is the minimum number of circles of radius ε which are necessary to cover the set T. In order to compute the dimension of the Sierpinski's triangle, proceeding as in [6], one can choose for any level n, $\varepsilon = (\sqrt{3})/3 \cdot 2^{-n}$. At level n, the number of triangles generated by the recursive process are 3^n, so $N_\varepsilon = 3^n$. The calculation of the limit yields

$$\dim(S) = -\lim_n (\ln (3^n)/\ln ((\sqrt{3})/3 \cdot 2^{-n}) = \ln 3/\ln 2.$$

Furthermore, by a straightforward calculation, it follows that the area of S is zero.

Sierpinski built, then, an actual curve C, piecewise linear, starting by the structure given by the triangle S (Fig. 3). In Fig. 3, center, the line $L^1 = S'_0 S'_1 S'_2 S'_3$ contains one edge for each of the three triangles T_1, T_2, T_3 appearing in Fig. 3, left. Iterating this procedure, as shown in Fig. 4, implies that at step n the line L^n will contain one edge of each of the 3^n triangles of the level n iteration.

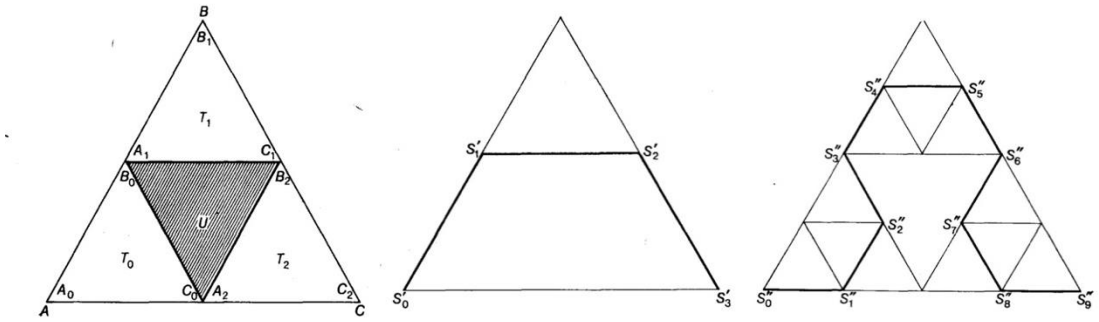


Fig. 3. The construction of the curve, first two steps. Sierpinski (1916).

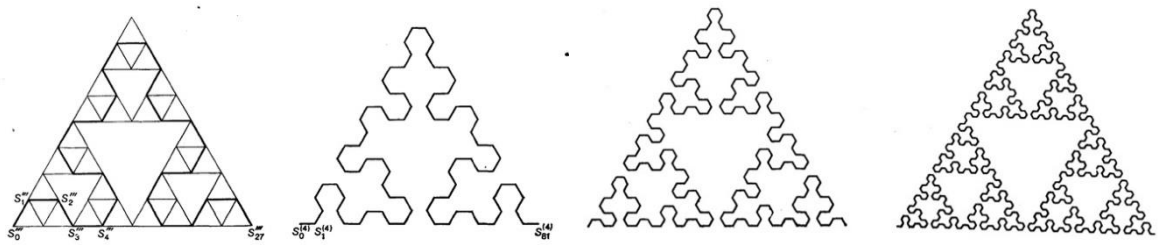


Fig. 4. The construction of the curve, further steps. Lines L^2, L^3, L^4, L^5. Sierpinski (1916).

The equation of the curve (a polygonal chain) can be written down explicitly by means of parametric equations

$$X = \phi_n(t),\ Y = \psi_n(t), \quad \text{with } t \in [0,1]$$

such that for $t = 0, 1/3^n, 2/3^n, \ldots, 1$ the curve passes by the vertexes of the n-level triangle (as in Fig. 3, Fig. 4) and that the functions $\phi_n(t)$ and $\psi_n(t)$ are linear for $t \in (i\text{-}1/3^n, i/3^n)$, $i = 1,2,\ldots, 3^n$. So a polygonal line remains defined for each level n. A straightforward calculation shows that the sequence $(\phi_n(t), \psi_n(t))$ converges uniformly to $(\phi(t), \psi(t))$ in the interval [0,1]. Since $(\phi_n(t), \psi_n(t))$ were continuous functions, the limit curve is continuous, which yields the Jordan property.

2.1 The Sierpinski's triangle and the curve are the same set of points

The segments composing the polygonal line L^n at level n, for any n, belong to the side of one of the 3^n triangles, so every point $q \in L^n$ belongs to S and, since S is closed, passing to the limit, each point $q \in R$ belongs to S, which yields $R \subseteq S$.

To prove the opposite inclusion, we recall again that the line L^n passes for each of the 3^n triangles, and that the set S is contained in these 3^n triangles. So, for every point $r \in S$ there exists a sequence of points $r_n \in R$ such that $dist\,(r, r_n) < 1/2^n$. By the compactness of [0, 1] there exists a subsequence $t_{nk} \rightarrow t^* \in [0, 1]$, such that (exploiting the uniform convergence of $(\phi_n(t), \psi_n(t))$ to $(\phi(t), \psi(t))$)

$$r_{nk} = (\phi_{nk}(t_{nk}), \psi_{nk}(t_{nk})) \rightarrow (\phi(t^*), \psi(t^*)) \in R, \quad \text{as } k \text{ goes to infinity}$$

So $r \in S$ implies $r \in R$.

This implies that the sets R and S coincide. This fact yields that the set S is a Jordanian curve and that the dimension of R and S are the same, so $d(R) = \ln 3/\ln 2$.

2.2 The order of ramification

This weird set fits perfectly with both definitions of curve. Sierpinski proved that each point, except for the three vertices A, B, C are ramification points. All other points have order 3. The idea of the proof is the following: let p be an arbitrary point, not a vertex of any of triangle in the construction of S, we need to construct three continua arriving at p and having only p as common point.

The point p is contained in an infinite sequence of triangles $T_{\alpha 1} \ldots . T_{\alpha n}$, each one contained in the previous ones; let $A_{\alpha 1} \ldots . A_{\alpha n}$ be the infinite sequence of the left vertices of those triangles (in the following lines, the letters A, B, C will denote, consistently with Fig. 2, respectively a left, up and right vertex). Joining all those $A_{\alpha 1} \ldots . A_{\alpha n}$ we obtain a polygonal curve, which we call L_A. Let us define the set D to be composed by the point p and all the points of the segments belonging to the curve L_A. It can be proved that the set D is closed, connected and infinite, so it is a continuum, and it is a subset of S. With the same procedure, let us define the set Q, consisting of the point p and a polygonal line L_B, connecting all the upper vertices $B_{\alpha 1} \ldots . B_{\alpha n}$ of the sequence of triangles containing p, and the set G consisting of p and the polygonal line L_C, connecting all the right vertices $C_{\alpha 1} \ldots . C_{\alpha n}$. The three sets, D, Q, G are continua, and one can prove that they intersect, two by two, only at p. So any point p which is neither one of the vertices of the initial triangle, nor a vertex of any n-level triangle, has ramification order 3 [18].

Since Sierpinski wanted to give an example of a curve having all ramification points, in order to remove the exception of the three vertices of the initial triangle, he then suggests to put together six triangles, to obtain an hexagon: in this way each one of the three initial vertexes meet another vertex with the same order, and become a point having ramification order 3.

3 The recursive structure and its manifold apparitions in science and art

In [21] the author underlines that the reason we continue to encounter the Sierpinksi's triangle is that it embodies recursion, and is one of the simplest recursive structures we know: as shown in previous sections, it can be easily drawn with ruler and pencil, up to a reasonable level. One of the most ancient apparitions of the recursive structure of Sierpinski is what is called "Pascal's (1623-1662) triangle", the triangular arrangement of the binomial coefficients; Sierpinski's structure reveals if even and odd coefficients are shaded in two different colours. Drawings of this pattern can be found already in 13th century Chinese mathematics books ([11, 21]), well before Pascal.

Another of the encounters, happened in 1883, and described in [21], regards Edouard Lucas (1842-1891) who created the puzzle game "The Tower of Hanoi". The graph of the possible moves of the game takes the shape of a Sierpinski's triangle.

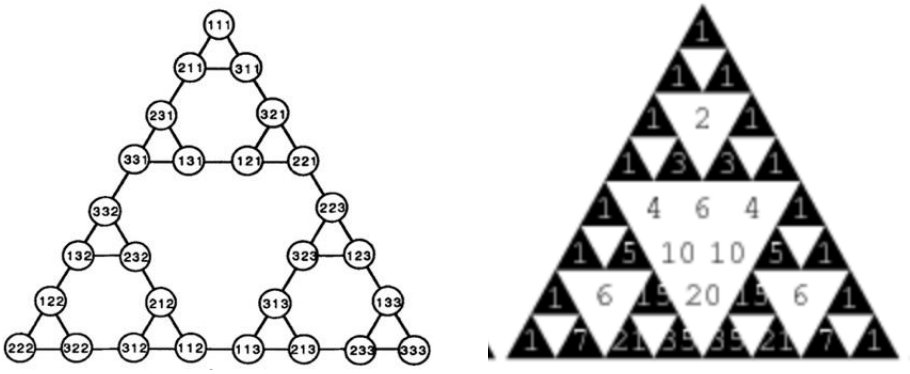


Fig. 5. Left: the graph of 3-disc Hanoi, [21]. Right: Pascal's triangle, from http://mathworld.wolfram.com/SierpinskiSieve.html.

A surprising meeting with Sierpinski's triangle could be found on a medieval floor, in central Italy, made with stone mosaic by Marmorari Romani (as they are called in scholarly literature). Clearly, since in this case the triangle is a physical object, the recursion cannot be repeated infinitely many times; on the other hand, the geometric structure is clear if the scale iteration is at least of level 3, as shown in the isolated triangle of Fig. 6, right [6].

A beautiful example of isolated triangles in golden leaf, showing level 3 and 4 of iteration, belonging to the frieze of the cloister of Saint John in Lateran in Rome been studied in [4]; the same isolated triangle has been the object of a collaboration with an inmate of a North American Jail, who planned and realized with a team of other inmates a piece of art reproducing the triangle; mathematics and the beauty of the ancient object were both an important stimulus for the realization of the project [10] .

Fig. 6. Left: Sierpinski's carpets, Santa Maria in Cosmedin, Rome.
Right: San Clemente, Rome (late 11th century) photo Alessandra Carlini.

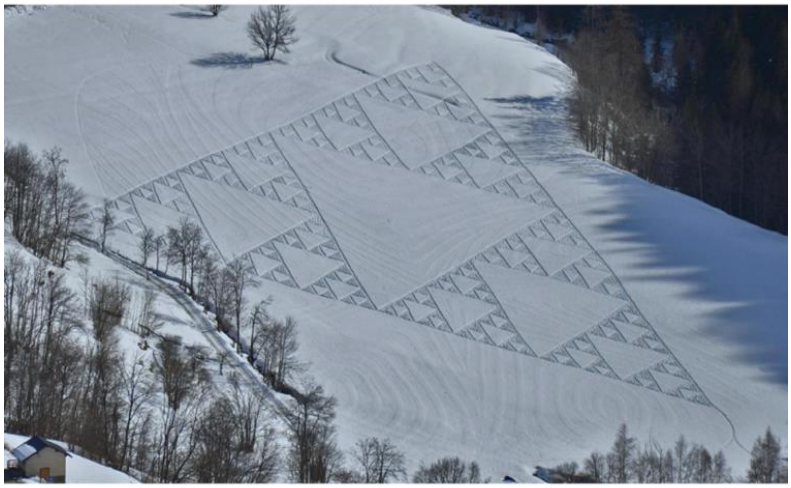


Fig. 7. A Sierpinski triangle drawn in the snow, by Simon Beck.

In Fig. 7, a Sierpinski's triangle traced in the snow, by the snow artist Simon Beck. Beck works using a compass and draws with the footprints he leaves with the snowshoes. Actually he walks along a unique, continuous line, since as can be seen in the pictures, there are areas where he must not leave footprints, and he does not "jump" from one area to the other; in other words, he "draws a line without removing the pen from the sheet".

Fig. 8. Variations on Sierpinski's triangles by Simon Beck.

During a Ted Talk he explained that he started with patterns having central simmetry, since they are the simplest to do with this technique. The Sierpinski's triangle is one of his favorites, he has reproduced it on snow and sand, with some variations, as in Fig. 8.

In Fig. 9, left, a Sierpinski's triangle drawn with a plotter led by a cellular automata (brought to Aplimat conference in 2018) designed in order to make possible the interaction with public in a science and technology museum. In picture 9, right, a composition made with five Sierpinski's triangles, slightly overlapping and arranged to form a penthagon. This image was the final project for the Math and Art course (Honors College, Ball State University). The authors produced a short animation showing the penthagon growing as the triangles become larger, with a starbust effect [1].

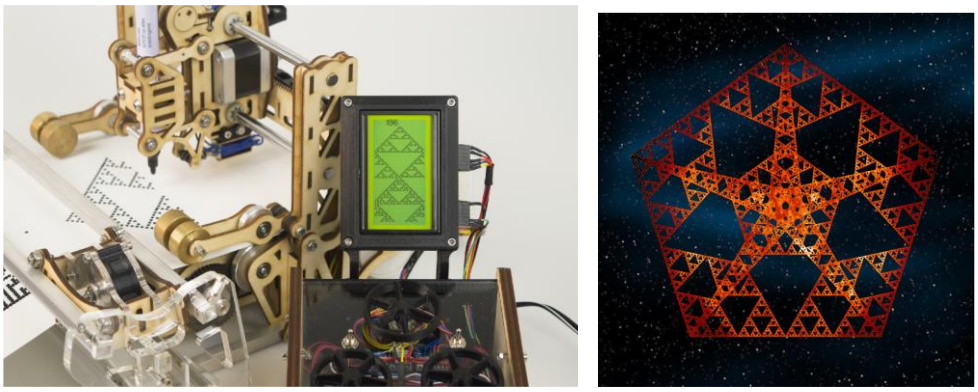


Fig. 9. Left: Plotter linked to a cellular automata drawing a Sierpinski triangle [9]. Right: composition made with five Sierpinski's triangles [1].

Picture in Fig. 10 shows an arrangement of rings [8], which leads to the shape of Sierpinski's Triangle; the iterations in Fig. 10 are from 1 to 4 and 9-10. The number of rings at step n can be computed by the formula $(3^{n+1}-3)/2$. In the interations 9 and 10 the thikness of the circles has been reduced to one fifth in order to make the picture clearer.

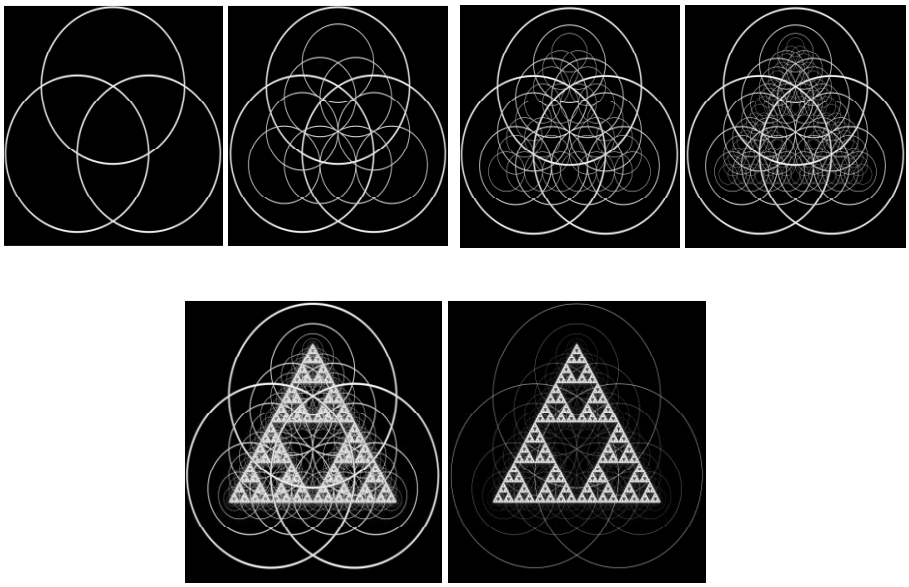


Fig. 10. An arrangement of rings bringing to the Sierpinski's triangle [8].

In Fig. 11 we observe the first three steps of an iterated function system [5]: at step 1 the unit circle is mapped into the four smaller circles (Fig.11, left); step 2 repeats the procedure for each of the four circles, and so on. The limit set is a Sierpinski's triangle (except for the small central smaller circle).

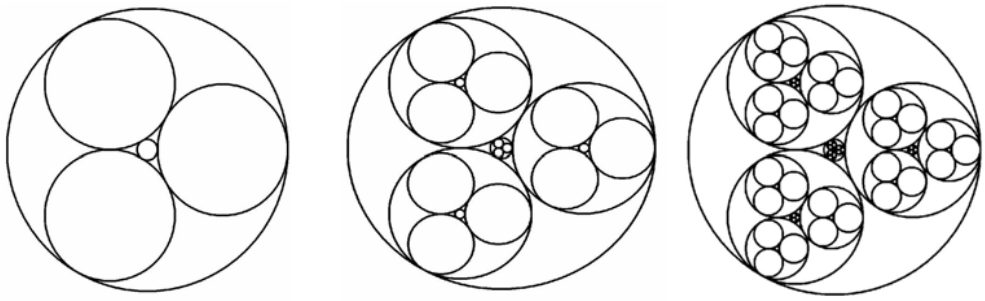


Fig. 11. The first three steps of the iterated function system in [5].

Then the author composes these transformations with a transformation of the kind $U(z) = (uz + v)/(-vz - u)$ where $|u| - |v| = 1$ (i.e., a subgroup of the group of Möbius transformations, mapping the unit circle onto itself, and its interior, onto itself). By changing the parameters u and v, the points of the circle can be rotated and the circle deformed, before applying the previous iterated function system. This allows to obtain different limit images (Fig. 12). The image shown in Fig. 13 is entitled "Sierpinski Triangle Eroding" and was exhibited during the Art Exhibit at the Joint Mathematics Meetings (of the AMS) in January 2008 in San Diego, USA; it was made overlapping a sequence of images.

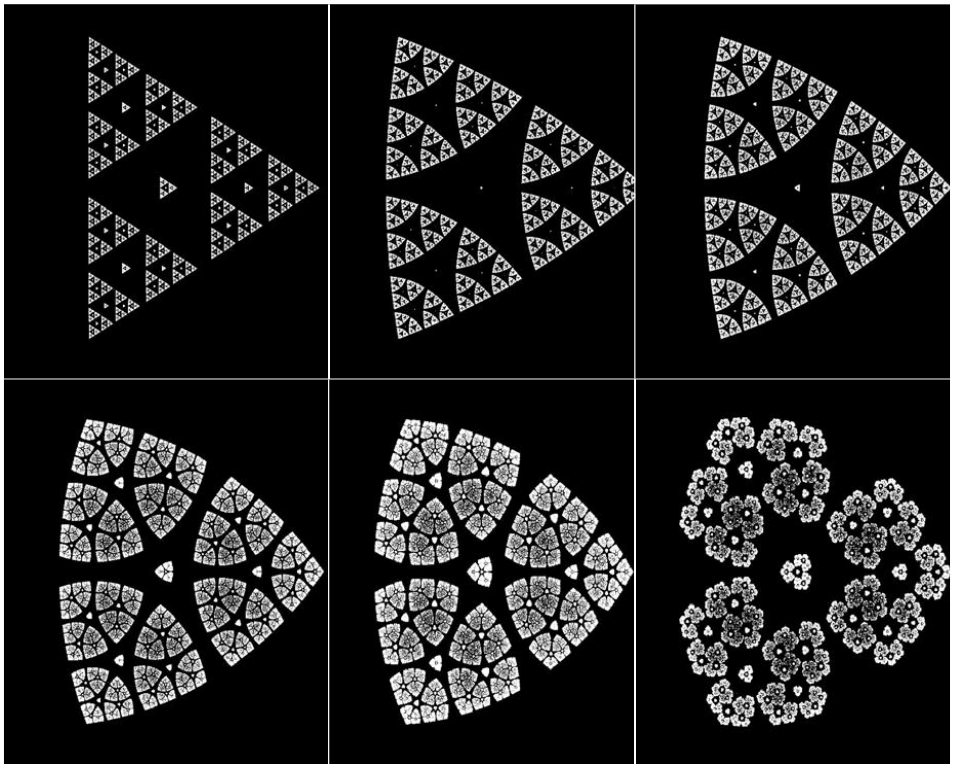


Fig. 12. Limit figures corresponding to different values of parameters.

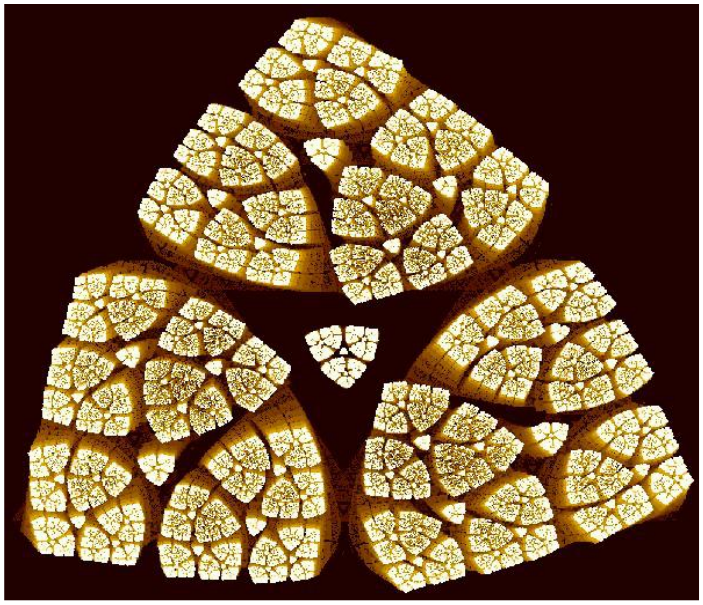


Fig. 13. "Sierpinski Triangle Eroding", part of the Art Exhibit at the Joint Mathematics Meetings (2008) obtained overlapping many images [5].

In [17] the author shows the presence of a first iteration Sierpinski's triangle in Klimt's painting "Beethoven Frieze" and a second iteration triangle on a Neolithic vessel.

In Section 2 we described the top-down construction of Sierpinski's triangle, which consists in inserting voids in a "full" object. The same set can be obtained with a down-top procedure, following the "chaos game" [2, 21]: let *A*, *B*, *C* be three points in the plane; the starting point is any point *P* inside the triangle *ABC*. Choose randomly one of the three vertices, move *P* towards the chosen vertex, for half of the distance and mark the point. Repeat this procedure, each time choosing randomly the vertex, which gives the direction towards which the point will move, and mark the arriving point, at each step. All these points accumulate on a Sierpinski triangle, as shown in Fig. 14. This shows that the Sierpinski's triangle is an "attractor" for the dynamics of chaos game. In other words, the points found step by step with the chaos game, self-assemble on the pattern of the Sierpinski's triangle.

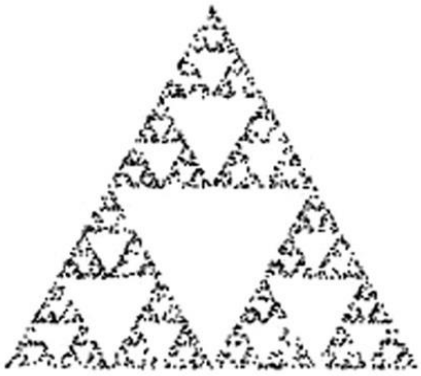


Fig. 14. Accumulating points of the down-top procedure, obtained by the Chaos Game [21].

5 Conclusions

We reported about the original articles dating 1915-16 where Waclaw Sierpinski introduced the triangle that today bears his name. The motivations that led him to think of this new mathematical object lay in the need to rethink the definition of curve. After many years the Sierpinski's triangle inspires artistic creations and still gathers the interest of the scientific community.

Acknowledgements

The paper was partially supported by *GNAMPA*-INdAM. The author thanks Alfonso Sorrentino for interesting discussions about the order of ramification.

References

[1] ACESKA R., O'BRIEN B., Artistic Excursions with the Sierpinski Triangle. In Bridges 2019 Conference Proceedings, pp. 537-540.

[2] ALLIGOOD K.T., TIM, D., SAUER, T.D., YORKE, J. A.: *Chaos: An Introduction to Dynamical Systems*. Springer, New York, 1996.

[3] AULL, C. E., LOWEN, R., *Handbook of the History of General Topology*, vol 2. Kluwer Academic Publishers, 1998.

[4] BRUNORI, P., MAGRONE P., TEDESCHINI LALLI, L.: Imperial Porphiry and Golden Leaf: Sierpinski Triangle in a Medieval Roman Cloister. In ICGG 2018 Proceedings of the 18th International Conference on Geometry and Graphics, Advances in Intelligent Systems and Computing 809, pp. 595–609, 2019.

[5] BURNS, A. M., From Sierpinski Triangle to Fractal Flowers. In Bridges 2008 Mathematics, Music, Art, Architecture, Culture, pp. 117-122.

[6] CONVERSANO, E., TEDESCHINI LALLI, L.: Sierpinsky Triangles in Stone, on Medieval Floors in Rome. In Aplimat Journal of Applied Mathematics, Vol. 4, 2011, pp. 113-122.

[7] FRYDE, M. M., "WACŁAW SIERPIŃSKI — MATHEMATICIAN." The Polish Review 8, no. 1, 1963: pp. 3-10. www.jstor.org/stable/25776446.

[8] GAROUSI M., DEZFOOLIAN, H. R., Constructing Sierpinski Triangle with Rings. In Bridges 2010: Mathematics, Music, Art, Architecture, Culture, pp .411-414.

[9] HAUDE, L., Creating an interaction with cellular automata for science and technology museums. In Aplimat 2018 - 17th Conference on Applied Mathematics, Bratislava, pp. 420-433.

[10] HAVENS, C., TEDESCHINI LALLI, L., Math. and art in prison, a collaborative effort across the ocean. In Aplimat 2020 - 19th Conference on Applied Mathematics, Bratislava, pp. 587-594.

[11] HOSCH, W. L., Pascal's triangle, https://www.britannica.com/science/Pascals-triangle

[12] JORDAN, C., *Cours d'analyse de l'École polytechnique*, tome I Calcul différentiel, Paris, Gauthier-Villars et fils Paris, 1893.

[13] KOSNIOWSKI, C., *A first course in algebraic topology*. Cambridge University Press, 1980.

[14] MANDELBROT, B. B., *The fractal geometry of nature*. Freeman, 1983.

[15] OTT, E.: *Chaos in Dynamical Systems*. Second Edition, Cambridge University Press, 2002.

[16] OTT, E.: *Attractor dimensions.* http://www.scholarpedia.org/article/Attractor_dimensions, last accessed 2020/01/20

[17] RICHTÁRIKOVÁ, D., Patterns that last. In Aplimat Journal of Applied Mathematics, Vol. 5, pp. 195-208, 2012.

[18] SIERPINSKI, W. *Sur une courbe cantorienne qui contient une image biunivoque et continue de toute courbe donnee.* C. R. Acad. Paris 162, 1916, pp. 629-632, in French; there is a more developed 1916 Russian version, printed in French translation in Oeuvres Choisies, tome II, pp. 107-119.

[19] SIERPINSKI, W., *Sur une courbe dont tout point est un point de ramification*. Compt. Rendus Acad. Sci. Paris 160, 1915, pp. 302–305, in French, there is a more developed 1916 Russian version, printed in French translation in Oeuvres Choisies, tome II, pp. 99-106.

[20] SIERPINSKI, W., *Oeuvres Choisies*. PWN, Polish Scientific Publishers, Warsaw-Poland 1974-1976.

[21] STEWART, I.: Four Encounters with Sierpinski's Gasket. In The Mathematical Intelligencer Vol. 17, n.1, 1995, Springer-Verlag New York p. 52-64.

Paola Magrone, PhD.

Department of Architecture
University Roma Tre
Via Madonna dei Monti 40, I–00184 Rome, Italy
e-mail: magrone@mat.uniroma3.it

Moving ellipses on quadrics

Hellmuth Stachel

Abstrakt

Ku každej regulárnej kvadrike 3-rozmerného euklidovského priestoru existuje troj-parametrická množina rezových rovín, avšak rozmery každej rezovej elipsy a hyperboly závisia iba od jej dvoch polosí. Preto na každej kvadrike existuje jednoparametrická množina kongruentných elíps aj hyperbol, ktoré sa môžu medzi sebou navzájom premiestňovať. Pre prípad elíps uvádzame parametrizáciu ich pohybu na elipsoidoch, hyperboloidoch a paraboloidoch. Pohyby úzko súvisia s teóriou konfokálnych kvadrík.

Kľúčové slová: konfokálne kvadriky, kužeľosečky na kvadrikách

Abstract

For each regular quadric in the Euclidean 3-space, there is a three-parameter set of cutting planes, but the size of an ellipse or hyperbola depends only on its two semiaxes. Therefore, on each quadric $\mathcal{Q}$ there exist ellipses or hyperbolas with a one-parameter set of congruent copies, which can even be moved into each other. For the case of ellipses, we present parametrizations of motions on ellipsoids, hyperboloids, and paraboloids. These motions are closely related to the theory of confocal quadrics.

Keywords: confocal quadrics, conics on quadrics

1 Introdution

There are well-known examples of conics which can be moved on quadrics. Apart from the trivial case of circles on a sphere, paraboloids are surfaces of translation, even with a continuum of translational nets of parabolas. On quadrics of revolution, each planar section can be rotated while it remains on the quadric.

What's about general quadrics $\mathcal{Q}$? There is a three-parameter family of planes which cut $\mathcal{Q}$ along a conic. However, the size of an ellipse or hyperbola depends only on its two semiaxes. This parameter count reveals that on each quadric $\mathcal{Q}$ there exist conics with a one-parameter family of congruent copies on $\mathcal{Q}$. Below, we focus on ellipses and provide parametrizations for the motion of appropriate ellipses on ellipsoids, hyperboloids, and paraboloids. The motions prove to be in close relation to the family of quadrics being confocal with $\mathcal{Q}$.

2 Moving ellipses on a triaxial ellipsoid

On each regular quadric $\mathcal{Q}$, two conics e_1 and e_2 in parallel planes are homothetic (Fig. 1). This means in the case ellipses, that they have parallel axes and the same ratio of semiaxes $a_e : b_e$. Moreover, their centers lie on the same diameter. This follows from the polarity with respect to (henceforth abbreviated as w.r.t.) $\mathcal{Q}$.

On an ellipsoid $\mathcal{E}$, we obtain the biggest ellipse within a homothetic family as the intersection with a plane through the ellipsoid's center O. On the other hand, there is a point $P \in \mathcal{E}$ with

a tangent plane τ_P parallel to the cutting planes, and the axes of the homothetic conics are parallel to the principal curvature directions at P (Fig. 1). The conics are even homothetic to the Dupin indicatrix at P. This can be confirmed, e.g., by straight forward computation using a Taylor expansion at P.

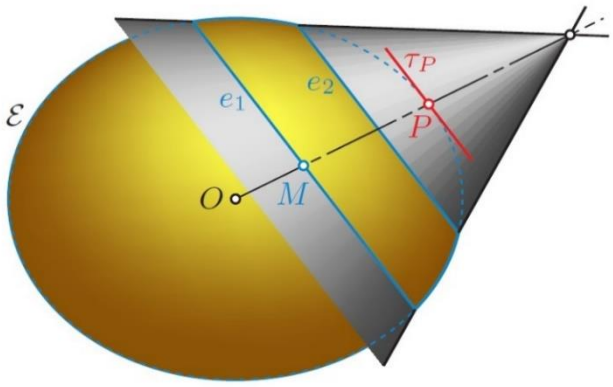


Fig. 1. Homothetic sections e_1, e_2 of the ellipsoid $\mathcal{E}$ in parallel planes.

According to the definition of the Dupin indicatrix, the ratio of the principal curvatures κ_1, κ_2 at P is reciprocal to the ratio of the squared semiaxes of the ellipses on $\mathcal{E}$ in planes parallel to τ_P, i.e.,

$$a_e : b_e = \sqrt{\kappa_1} : \sqrt{\kappa_2}, \quad \text{if} \quad \kappa_1 > \kappa_2 . \tag{1}$$

The lines of curvature on quadrics are related to confocal quadrics. Therefore, we recall some relevant properties of confocal quadrics.

2.1 Confocal central quadrics

Let $\mathcal{E}$ be a triaxial ellipsoid with semiaxes a, b, and c. The one-parameter family of quadrics being confocal with $\mathcal{E}$ is given as

$$F(x,y,z;k) := \frac{x^2}{a^2+k} + \frac{y^2}{b^2+k} + \frac{z^2}{c^2+k} - 1 = 0 , \tag{2}$$

Where $k \in \mathbb{R} \setminus \{-a^2, -b^2, -c^2\}$ serves as a parameter. In the case $a > b > c > 0$, this family includes

$$\text{for} \begin{cases} -c^2 < k < \infty & \text{triaxial ellipsoids,} \\ -b^2 < k < -c^2 & \text{one-sheeted hyperboloids,} \\ -a^2 < k < -b^2 & \text{two-sheeted hyperboloids.} \end{cases} \tag{3}$$

Confocal quadrics intersect their common planes of symmetry along confocal conics. As limits for $k \to -c^2$ and $k \to -b^2$ we obtain 'flat' quadrics, i.e., the focal ellipse and the focal hyperbola.

The confocal family sends through each point $P = (\xi, \eta, \zeta)$ outside the coordinate planes, i.e., with $\xi\eta\zeta \neq 0$, exactly one ellipsoid, one one-sheeted hyperboloid, and one two-sheeted hyperboloid. The corresponding parameters k define the three *elliptic coordinates* of P. We focus on points P of the ellipsoid $\mathcal{E}$ with $k = 0$,

$$\mathcal{E}: \frac{\xi^2}{a^2}+\frac{\eta^2}{b^2}+\frac{\zeta^2}{c^2}=1. \tag{4}$$

The two hyperboloids $\mathcal{H}_1$ and $\mathcal{H}_2$ through P with respective parameters k_1 and k_2, where

$$-a^2 < k_2 < -b^2 < k_1 < -c^2 < 0, \tag{5}$$

satisfy

$$\mathcal{H}_i: \frac{\xi^2}{a^2+k_i}+\frac{\eta^2}{b^2+k_i}+\frac{\zeta^2}{c^2+k_i}=1, \qquad i=1,2. \tag{6}$$

For given Cartesian coordinates (ξ, η, ζ) of a point P, we obtain the elliptic coordinates, i.e., the parameters of the quadrics through P, by solving $F(\xi, \eta, \zeta; k) = 0$ in (3) for k. This results in a cubic equation with three real roots. Conversely, if the tripel $(0, k_1, k_2)$ of elliptic coordinates is given, then the Cartesian coordinates (ξ, η, ζ) of the corresponding points $P \in \mathcal{E}$ satisfy

$$\begin{aligned}
\xi^2 &= \frac{a^2(a^2+k_1)(a^2+k_2)}{(a^2-b^2)(a^2-c^2)}, \\
\eta^2 &= \frac{b^2(b^2+k_1)(b^2+k_2)}{(b^2-c^2)(b^2-a^2)}, \\
\zeta^2 &= \frac{c^2(c^2+k_1)(c^2+k_2)}{(c^2-a^2)(c^2-b^2)}.
\end{aligned} \tag{7}$$

There exist 8 such points, symmetric w.r.t. the coordinate planes.

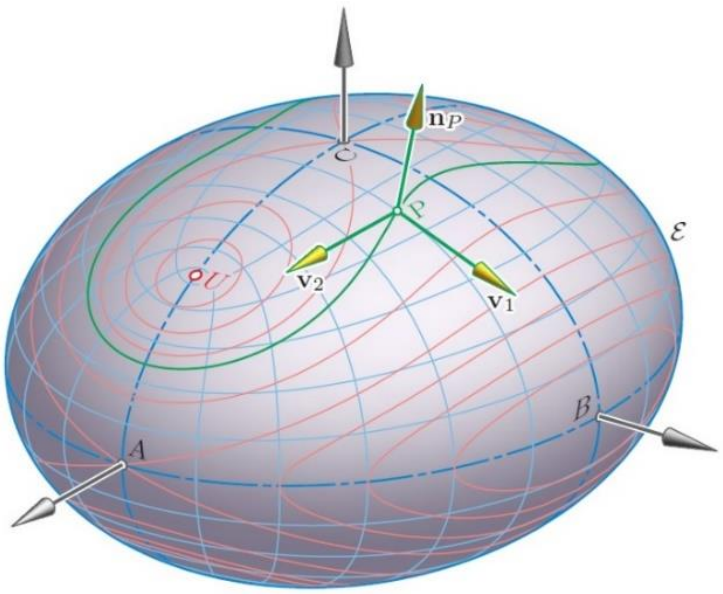


Fig. 2. Ellipsoid $\mathcal{E}$ with lines of curvature (blue), curves of constant ratio of principal curvatures $\kappa_1 : \kappa_2$ (red), principal curvature directions $\mathbf{v}_1$, $\mathbf{v}_2$ at the point P, and one *umbilic point* U with $\kappa_1 = \kappa_2$.

At each point P of the ellipsoid $\mathcal{E}$, the surface normal n_P to $\mathcal{E}$ has the direction vector

$$\mathbf{n}_P = \left(\frac{\xi}{a^2}, \frac{\eta}{b^2}, \frac{\zeta}{c^2} \right). \tag{8}$$

The surface normals of the two hyperboloids $\mathcal{H}_1$ und $\mathcal{H}_2$ through P are in direction of the vectors

$$\mathbf{v}_i := \left(\frac{\xi}{a^2 + k_i}, \frac{\eta}{b^2 + k_i}, \frac{\zeta}{c^2 + k_i} \right), \quad i = 1, 2\,. \tag{9}$$

The differences of any two of the equations in (4) and (6) yield

$$\begin{aligned} &\frac{\xi^2}{a^2(a^2 + k_i)} + \frac{\eta^2}{b^2(b^2 + k_i)} + \frac{\zeta^2}{c^2(c^2 + k_i)} = 0\,, \qquad i = 1, 2\,, \quad \text{and} \\ &\frac{\xi^2}{(a^2 + k_1)(a^2 + k_2)} + \frac{\eta^2}{(b^2 + k_1)(b^2 + k_2)} + \frac{\zeta^2}{(c^2 + k_1)(c^2 + k_2)} = 0 \end{aligned} \tag{10}$$

This is equivalent to vanishing dot products

$$\mathbf{n}_P \cdot \mathbf{v}_1 = \mathbf{n}_P \cdot \mathbf{v}_2 = \mathbf{v}_1 \cdot \mathbf{v}_2 = 0.$$

Therefore, confocal quadrics form a triply orthogonal system of surfaces. Due to a theorem of Dupin, they intersect each other along lines of curvature. The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ from (9) define the principal curvature directions at P.

2.2 Ellipses on ellipsoids

Now, we look for the biggest ellipse on $\mathcal{E}$ within a homothetic family.

Lemma 1. The semiaxes of the ellipse in the diameter plane parallel to the tangent plane τ_P at the point $P \in \mathcal{E}$ with elliptic coordinates $(0, k_1, k_2)$ are

$$a_p = \sqrt{-k_2}\,, \quad b_p = \sqrt{-k_1}\,. \tag{11}$$

Proof. The diameter plane is spanned by the direction vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ from (9). We look for $\lambda \in \mathbb{R}$ with $\lambda \mathbf{v}_i \in \mathcal{E}$, hence by (4)

$$\lambda^2 \left[\frac{\xi^2}{(a^2 + k_i)^2 a^2} + \frac{\eta^2}{(b^2 + k_i)^2 b^2} + \frac{\zeta^2}{(c^2 + k_i)^2 c^2} \right] = 1\,.$$

This condition does not change if we subtract from the term in square brackets the left-hand side of the first equation in (10), divided by k_i. Thus, we obtain

$$\lambda^2 \left[\frac{\xi^2}{(a^2 + k_i)^2 a^2} - \frac{\xi^2}{k_i (a^2 + k_i) a^2} + \ldots \right] = 1\,,$$

and, finally,

$$-\frac{\lambda}{k_i}^2\left[\frac{\xi^2}{(a^2+k_i)^2}+\frac{\eta^2}{(b^2+k_i)^2}+\frac{\zeta^2}{(c^2+k_i)^2}\right]=-\frac{\lambda}{k_i}^2\|\mathbf{v}_i\|^2=1,$$

hence, $a_P=|\lambda|\|\mathbf{v}_2\|=\sqrt{-k_2}$ and $b_P=|\lambda|\|\mathbf{v}_1\|=\sqrt{-k_1}$. These equations can already be found in [1, p. 517].

□

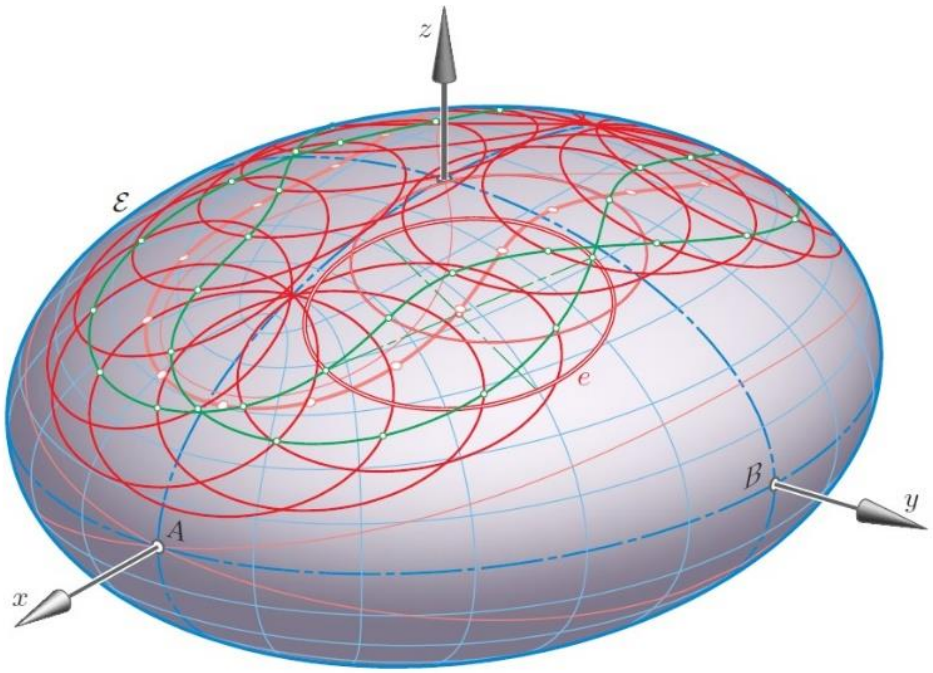


Fig. 3. Moving the ellipse e on the ellipsoid $\mathcal{E}$. The trajectories of the principal vertices of e are displayed in green.

For the motion of a given ellipse e with semiaxes (a_e, b_e), Lemma 1 implies the necessary condition

$$a_e \le a_P=\sqrt{-k_2}, \quad \text{where} \quad b<\sqrt{-k_2}<a \tag{12}$$

by virtue of (5). We infer, under inclusion of (1):

Theorem 1. If an ellipse e with semiaxes (a_e, b_e) is moving on a triaxial ellipsoid $\mathcal{E}$, then both points $P \in \mathcal{E}$ with tangent planes τ_P parallel to the plane of e move on curves with proportional elliptic coordinates $k_1 : k_2 = -a_e^2 : -b_e^2$. Along these curves also the ratio of the principal curvatures remains constant (see Fig. 2).

The ellipses of $\mathcal{E}$ in planes parallel to τ_P have their principal vertices in the plane spanned by the center O, point P, and by the principal curvature direction $\mathbf{v}_2$ from (9). Therefore, the principal vertices are located on an ellipse, for which OP and the major axis with length a_P in the plane through O determine conjugate diameters. Let $\mathbf{p}$ denote the position vector of P and $\mathbf{m} = \mu\,\mathbf{p}$ with $0 \le \mu = \sin x < 1$ that of the center M of any ellipse in e homothetic family.

Then, its major semiaxis a_e equals $a_P \cos x = a_P \sqrt{1-\mu^2}$ which results in

$$\mu^2 = 1 - \frac{a_e^2}{a_P^2} = 1 + \frac{a_e^2}{k_2}. \tag{13}$$

When during the motion of the ellipse e, the scalar μ vanishes, then its center M coincides with the center O of $\mathcal{E}$. The corresponding point P has the elliptic coordinate $k_2 = -a_e^2$. In order to continue the motion, point P has to jump to its antipode (note the example in Fig. 4).

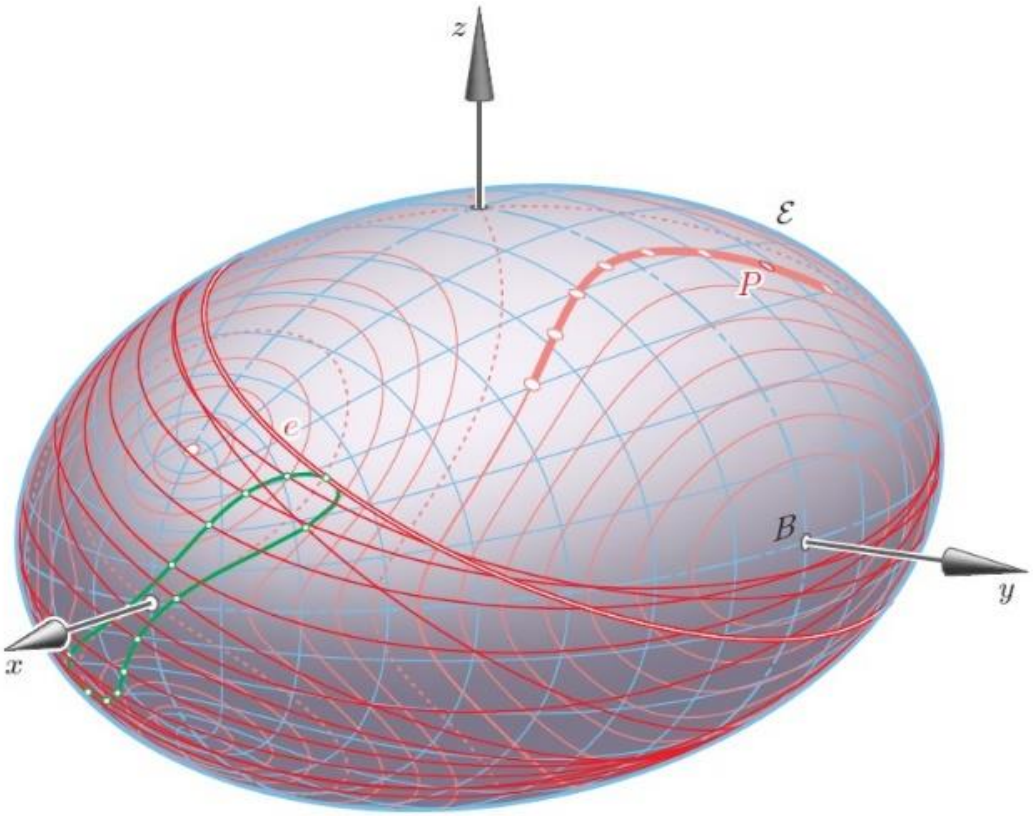


Fig. 4. Motion of the ellipse e on the ellipsoid $\mathcal{E}$ – displayed together with the trajectory of a principal vertex of e (green) and that of the corresponding point $P \in \mathcal{E}$ (red) with the tangent plane τ_P parallel to e.

In order to parametrize the motion of the ellipse e on the ellipsoid $\mathcal{E}$ (see Fig. 3), we set

$$v := \frac{k_2}{k_1} = \frac{a_e^2}{b_e^2} = \text{const., where } 1 < v < \frac{a^2}{c^2}, \tag{14}$$

and use the parameter $t = -k_2$ for representing the motion. Then, by virtue of (5), t is restricted by the interval

$$\max\{b^2, vc^2, a_e^2\} \le t \le \min\{a^2, vb^2\} \tag{15}$$

and $k_1 = t/v$. From (7) follows the parametrization $\mathbf{p}(t)$ by replacing (k_1, k_2) with $(t/v, t)$.

This implies for the trajectory of the center M of e

$$\mathbf{m}(t) = \mu(t)\,\mathbf{p}(t) \quad \text{with} \quad \mu(t) = \sqrt{1 - \frac{a_e^2}{t}}\,. \tag{16}$$

Now, we can express the motion of e in matrix form, in terms of position vectors $\mathbf{x}_m$ w.r.t. the moving space (attached to e) and $\mathbf{x}_f$ w.r.t. the fixed space (attached to $\mathcal{E}$), as

$$\mathbf{x}_f = \mathbf{m}(t) + \mathbf{M}(t)\,\mathbf{x}_m\,, \quad \text{where} \quad \mathbf{M}(t) = \left[\frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}, \frac{\mathbf{n}_P}{\|\mathbf{n}_P\|}\right]. \tag{17}$$

The three column vectors of the orthogonal matrix $\mathbf{M}(t)$ are given in (9) and (8).

Note that this parametrization is valid only for points P in the octant $x, y, z > 0$. We get a closed motion after appropriate reflections in the planes of symmetry (see Figs. 3 and 4). By the same token, algebraic properties of this motion are reported in [2].

3 Moving ellipses on a one-sheeted hyperboloid

Also on hyperboloids and paraboloids, the conics in parallel planes are homothetic. However, not in all cases the method, as used before for ellipsoids, can be applied since a point P either does not exist or lies at infinity. Moreover, paraboloids have no center O. Below, we analyse the motions of ellipses on a one-sheeted hyperboloid $\mathcal{H}_1$ and on an elliptic paraboloid $\mathcal{P}$ (see Section 4). The motion of an ellipse on a two-sheeted hyperboloid works similar to that of triaxial ellipsoids.[1]

For ellipses $e \subset \mathcal{H}_1$, there is no point $P \in \mathcal{H}_1$ with a tangent plane τ_P parallel to e. However, we find an appropriate point $\tilde{P}$ on the 'conjugate' two-sheeted hyperboloid $\mathcal{H}_2$ (Fig. 5).

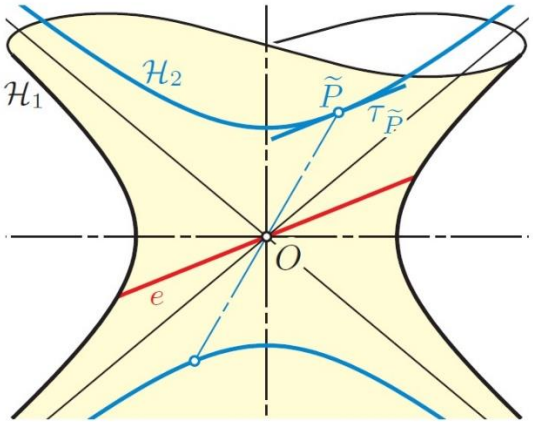


Fig. 5. For ellipses e on a one-sheeted hyperboloid $\mathcal{H}_1$, there does not exist a point $P \in \mathcal{H}_1$ with a tangent plane τ_P parallel to the plane of e.

[1] The motion of a parabola on a hyperboloid is discussed in [4, p. 355–357].

The hyperboloid $\mathcal{H}_2$ shares the asymptotic cone with $\mathcal{H}_1$, and therefore, the axes of the ellipse e are parallel to the principal curvature directions of $\mathcal{H}_2$ at $\tilde{P}$. The two hyperboloids satisfy the respective equations

$$\mathcal{H}_1 : \frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \quad \text{and} \quad \mathcal{H}_2 : -\frac{x^2}{a^2}-\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

with $a > b$. The quadrics confocal with $\mathcal{H}_2$ are given by

$$-\frac{x^2}{a^2-k}-\frac{y^2}{b^2-k}+\frac{z^2}{c^2+k}=1.$$

Again, this family sends through each point $\tilde{P}$ outside of the planes of symmetry three mutually orthogonal quadrics, one of each type. On the two-sheeted hyperboloid $\mathcal{H}_2$ with $k = 0$, we use the parameters k_0 of the ellipsoid and k_1 of the one-sheeted hyperboloid as the elliptic coordinates of $\tilde{P}$ with

$$b^2 < k_1 < a^2 < k_0 .$$

Then, similar to Lemma 1, the ellipse $e \in \mathcal{H}_1$ in the diameter plane parallel to $\tau_{\tilde{P}}$ has the semiaxes

$$a_{\tilde{P}} = \sqrt{k_0} \quad \text{and} \quad b_{\tilde{P}} = \sqrt{k_1} .$$

This is the smallest ellipse on $\mathcal{H}_1$ within the homothetic family.

If any ellipse with given semiaxes a_e and b_e is to be moved on $\mathcal{H}_1$, then the corresponding point $\tilde{P} \in \mathcal{H}_2$ has to trace a curve with proportional elliptic coordinates

$$k_0 : k_1 = a_{\tilde{P}}^2 : b_{\tilde{P}}^2 = a_e^2 : b_e^2 .$$

Similar to (7), we can parametrize the trajectory $\tilde{\mathbf{p}}(t) = (\xi, \eta, \zeta)$ of $\tilde{P}$ by $t := k_0 > a^2$, where

$$\nu := \frac{k_0}{k_1} = \frac{a_e^2}{b_e^2} = \text{const.},$$

hence $k_1 = t/\nu$ with $b^2 < k_1 < a^2$.

Now we have to find the center M of the moving ellipse on the diameter line $\left[\tilde{P}, O\right]$: For each $\tilde{P}$, the principal vertices of the ellipses in planes parallel to $\tau_{\tilde{P}}$ are placed on a hyperbola, for which the point $\tilde{P}$ and one principal vertex in the plane through O are the endpoints of conjugate diameters. If $a_e = a_{\tilde{P}} \cosh x$, then the position vector $\mathbf{m}$ of the center M of e and $\tilde{\mathbf{p}}$ of the point $\tilde{P}$ are related by $\mathbf{m} = \sinh x\, \tilde{\mathbf{p}}$. Thus, we obtain

$$\mathbf{m} = \mu\, \tilde{\mathbf{p}} \quad \text{with} \quad \mu^2 = \frac{a_e^2}{a_{\tilde{P}}^2} - 1 . \tag{18}$$

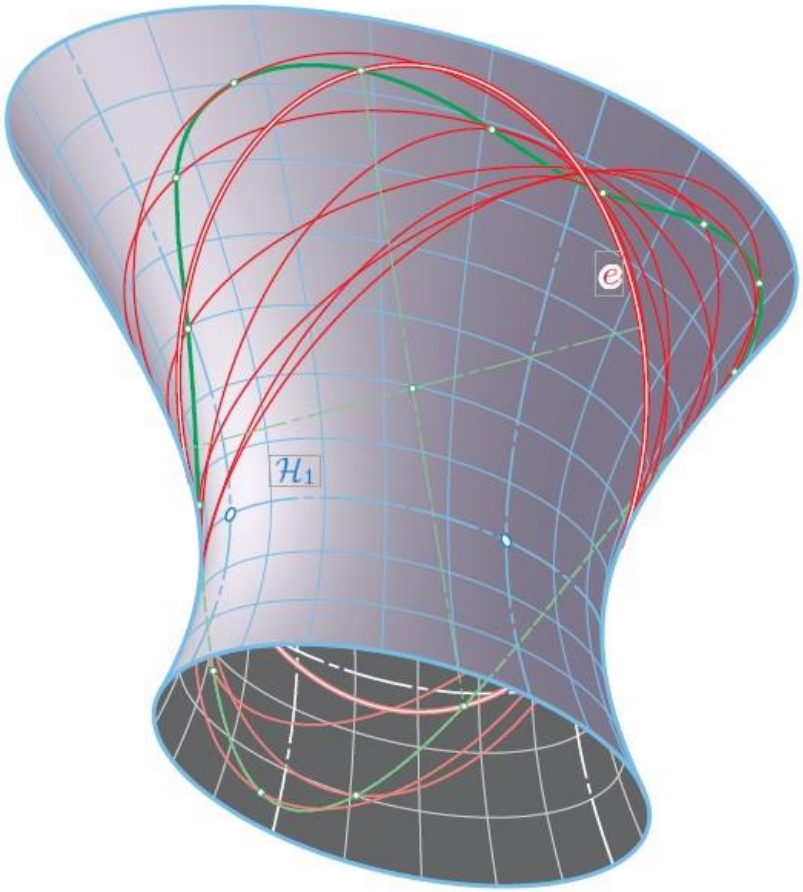


Fig. 6. Movement of the ellipse e on the one-sheeted hyperboloid $\mathcal{H}_1$. The two principal vertices of e trace the same curve (green).

This yields, similar to (17), a parametrization for the motion of the ellipse e on $\mathcal{H}_1$ (Fig. 6). As a consequence of (18), on the trajectory of $\tilde{P}$ only points with $a_{\tilde{P}}^2 = k_0 \le a_e^2$ are admitted. Therefore, the parameter $t = k_0$ runs the interval

$$\max\{a^2, vb^2\} \le t \le \min\{a_e^2, va^2\}\,.$$

In the case $a_e^2 < va^2$, the same phenomenon appears as mentioned above. When the parameter t reaches a_e^2, then, for continuing the motion of the ellipse, the point $\tilde{P}$ either has to jump to its antipode, or the scalar μ in (18) must get a negative sign.

4 Moving ellipses on an elliptic paraboloid

The quadrics being confocal with an elliptic paraboloid can be represented as

$$\frac{x^2}{a^2+k} + \frac{y^2}{b^2+k} - 2z - k = 0 \quad \text{for } k \in \mathbb{R} \setminus \{-a^2, -b^2\}\,. \tag{19}$$

In the case a > b > 0, this one-parameter family contains

$$\text{for } \begin{cases} -b^2 < k < \infty & \text{elliptic paraboloids,} \\ -a^2 < k < -b^2 & \text{hyperboloc paraboloids,} \\ k < -a^2 & \text{elliptic paraboloids.} \end{cases} \tag{20}$$

For each k, the vertex of the corresponding paraboloid has the coordinates $(0, 0, -k/2)$. The point $(0, 0, b^2/2)$ is the common focal point of the principal sections in the plane $x = 0$, and $(0, 0, a^2/2)$ is the analogue for the sections with $y = 0$. The limits for $k \to -b^2$ or $k \to -a^2$ define the two *focal parabolas* (note [4, Fig. 7.5,]).

The family of confocal parabolas sends through each point P outside the planes of symmetry $x = 0$ and $y = 0$ three surfaces, one of each type. Like before in the case of confocal central surfaces, we call the parameters of the three parabolas through P the *elliptic coordinates* of P. We focus on the elliptic paraboloid $\mathcal{P}_0$ with $k = 0$. Its points have the elliptic coordinates $(0, k_1, k_2)$, where

$$k_2 \le -a^2 \le k_1 \le -b^2 .$$

Conversely, if any point $P \in \mathcal{P}_0$ is defined by the elliptic coordinates (k_1, k_2), then its Cartesian coordinates ξ, η, ζ satisfy

$$\begin{aligned} \xi^2 &= \frac{a^2(a^2+k_1)(a^2+k_2)}{(a^2-b^2)}, \\ \eta^2 &= \frac{b^2(b^2+k_1)(b^2+k_2)}{(a^2-b^2)}, \\ \zeta &= \frac{a^2+b^2+k_1+k_2}{2}. \end{aligned} \tag{21}$$

The normal vectors $\mathbf{n}_P$ of $\mathcal{P}_0$ and $\mathbf{v}_i$ of the paraboloid $\mathcal{P}_i$ with parameter k_i, $i = 1, 2$, at the point P are (note Fig. 7)

$$\mathbf{n}_P = \begin{pmatrix} \frac{\xi}{a^2} \\ \frac{\eta}{b^2} \\ 1 \end{pmatrix}, \quad \mathbf{v}_i = \begin{pmatrix} \frac{\xi}{a^2+k_i} \\ \frac{\eta}{b^2+k_i} \\ 1 \end{pmatrix}. \tag{22}$$

Also confocal paraboloids form a triply orthogonal system of surfaces, and consequently, they intersect each other along lines of curvature. The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in (22) define the principal curvature directions at P.

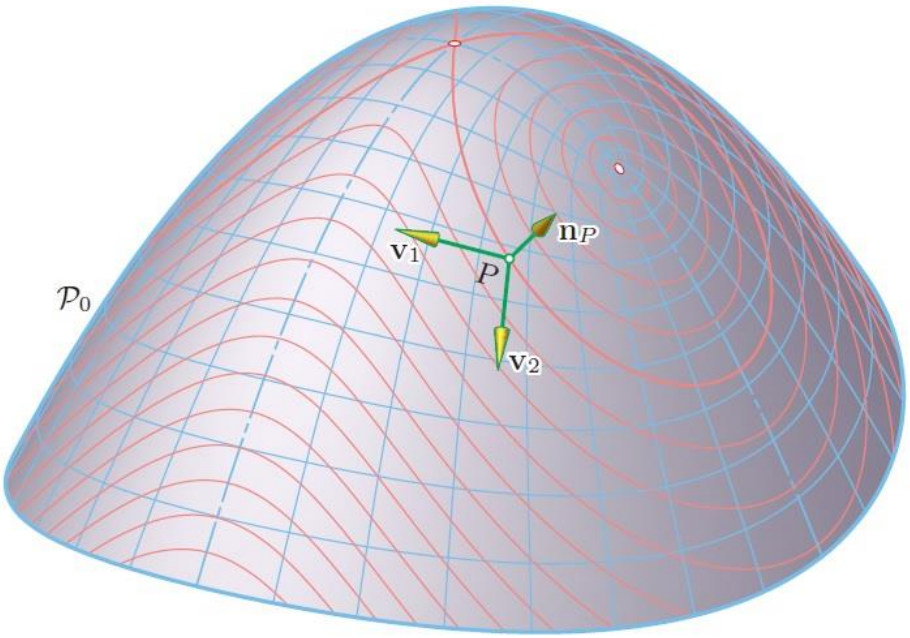


Fig. 7. Elliptic paraboloid $\mathcal{P}_0$ with lines of curvature (blue), curves of constant ratio of principal curvatures $\kappa_1 : \kappa_2$ (red), and direction vectors $\mathbf{v}_1$, $\mathbf{v}_2$ of the principal curvature tangents at the point $P \in \mathcal{P}_0$.

Also confocal paraboloids form a triply orthogonal system of surfaces, and consequently, they intersect each other along lines of curvature. The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in (22) define the principal curvature directions at P.

Lemma 2. Given a regular quadric $\mathcal{Q}_0$, let $P \in \mathcal{Q}_0$ be a point in general position with the tangent plane τ_P to $\mathcal{Q}_0$. If $\mathcal{Q}_1$ and $\mathcal{Q}_2$ are the remaining two confocal quadrics through P, the pole of τ_P w.r.t. $\mathcal{Q}_2$ is the center of curvature of the orthogonal section of $\mathcal{Q}_0$ at P through the principal curvature tangent t_P orthogonal to $\mathcal{Q}_2$.

Proof. We can verify this by straight forward computation: Based on the parametrizations of $\mathcal{Q}_0$ by elliptic coordinates (k_1, k_2), as given in (7) for central quadrics and in (21) for paraboloids, we compute the first and second fundamental form and the center of curvature (= Meusnier point) for the orthogonal section of $\mathcal{Q}_0$ through t_P (see, e.g., [3]).

A synthetic proof runs as follows: Let c be the line of intersection between the confocal quadrics $\mathcal{Q}_0$ and $\mathcal{Q}_1$. Then, c is a line of curvature for both. The developable $\mathcal{T}$ which contacts $\mathcal{Q}_0$ along c has generators orthogonal to c. Also the surface normals to $\mathcal{Q}_0$ along c form a developable $\mathcal{N}$. Its cuspidal points are the centers of curvature of the orthogonal sections of $\mathcal{Q}_0$ through the tangents to c (note [4, p. 418ff]).

At the point $P \in c$, the tangent t_P to c, the surface normal n_P to $\mathcal{Q}_0$, and the generator g_P of $\mathcal{T}$ are mutually orthogonal. Any two of them define the principal curvature directions at P for one of the three confocal quadrics. For example, the lines g_P and n_P are conjugate tangents of $\mathcal{Q}_2$, and therefore, even polar w.r.t. $\mathcal{Q}_2$.

The polarity w.r.t. $\mathcal{Q}_2$ transforms the developable $\mathcal{T}$ through g_P into a developable $\mathcal{T}'$ through n_P, while tangent planes τ_X of $\mathcal{T}$ and $\mathcal{Q}_0$ at points $X \in c$ are sent to points X' of the cuspidal edge $c_{\mathcal{T}'}$ of $\mathcal{T}'$. The poles of each plane w.r.t. the quadrics of a confocal family lie on a line orthogonal to the given plane (see, e.g., [4, p. 292]). Therefore, the $\mathcal{Q}_2$-pole X' of τ_X lies on the normal n_X of $\mathcal{Q}_0$ at X. Consequently, the cuspidal edge $c_{\mathcal{T}'}$ of $\mathcal{T}'$ is a curve on the developable $\mathcal{N}$. The polarity w.r.t. $\mathcal{Q}_2$ takes the generator $g_X \subset \mathcal{T}$ to the tangent g'_X to $c_{\mathcal{T}'}$ at X', which is also a tangent of $\mathcal{N}$.

Now we prove, that the cuspidal edge $c_{\mathcal{T}'}$ of $\mathcal{T}'$ passes through the cuspidal point $C_{\mathcal{N}}$ of $n_P \subset \mathcal{N}$:

The tangent plane τ_P to $\mathcal{T}$ at P is the limit $X \to P$ of a plane connecting the generator g_P with any point of g_X. By virtue of the polarity w.r.t. $\mathcal{Q}_2$ with $\mathcal{T} \to \mathcal{T}'$, the cuspidal point $P' \in c_{\mathcal{T}'}$ on n_P is the limit $X \to P$ of the point of intersection between n_P and any plane through g'_X. As noted before, the tangent plane $[n_X, t_X]$ along n_X to $\mathcal{N}$ is such a plane, since it passes through g'_X. However, the limit $X \to P$ of the point of intersection $n_P \cap [n_X, t_X]$ yields also the cuspidal point $C_{\mathcal{N}}$ of n_P w.r.t. the developable $\mathcal{N}$. This means, that $C_{\mathcal{N}}$ equals the pole P' of τ_P w.r.t. $\mathcal{Q}_2$. □

We apply Lemma 2 to the elliptic paraboloid $\mathcal{P}_0$. The tangent plane τ_P to $\mathcal{P}_0$ at $P = (\xi, \eta, \zeta)$ has the equation

$$\tau_P : \frac{\xi}{a^2} x + \frac{\eta}{b^2} y + z = \zeta \,.$$

Its pole w.r.t. the paraboloid $\mathcal{P}_i$ with parameter k_i is

$$C_i = \begin{pmatrix} \frac{a^2 + k_i}{a^2} \xi \\ \frac{b^2 + k_i}{b^2} \eta \\ \zeta + k_i \\ \end{pmatrix} = \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} + k_i \begin{pmatrix} \frac{\xi}{a^2} \\ \frac{\eta}{b^2} \\ 1 \end{pmatrix}. \tag{23}$$

This confirms that the principal curvatures of $\mathcal{P}_0$ at P are

$$\kappa_i = 1/\overline{PC_i} = \frac{1}{-k_i \left\| \mathbf{n}_P \right\|}, \text{ where } \kappa_1 > \kappa_2 \,. \tag{24}$$

Now we have to place a given ellipse e with semiaxes a_e and b_e, where $a_e^2 : b_e^2 = k_2 : k_1$, in a plane parallel to τ_P in the correct way on $\mathcal{P}_0$. This means, the center M of e lies on the diameter d_P of the paraboloid $\mathcal{P}_0$ and the major axis is parallel to the principal curvature tangent t_P in direction $\mathbf{v}_2$, i.e., orthogonal to the paraboloid $\mathcal{P}_2$ through P (Fig. 7).

The major axis lies in the plane ε spanned by t_P and d_P. This plane intersects $\mathcal{P}_0$ along a parabola p. Due to Meusnier's theorem, we obtain the center of curvature P^* of p at P as the pedal point of C_2 from (23) in ε. Let $\rho = \overline{PP^*}$ denote the radius of curvature at P (Fig. 8). Then the chord S_1S_2 of p parallel to t_P through the midpoint of PP^* has its midpoint S_0 on the diameter d_P and the length 2ρ.

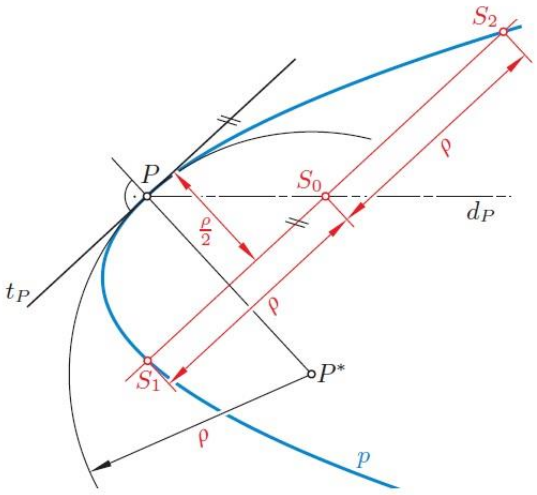


Fig. 8. For a given parabola p with point $P \in p$ and corresponding center of curvature P^*, this is a construction of the endpoints S_1, S_2 on a particular chord of p.

This follows with the help of a *shear*, i.e., a perspective affine transformation in ε with t_P as axis and the ideal point of t_P as its center. This shear transforms p into a parabola p' which osculates p at P. We can define a shear such that P becomes the vertex of p'. Then, the midpoint of PP^* if the focal point of p', and for p' the chord parallel to t_P through the focal point has the length 2ρ. Under the inverse shear, the chord is just translated parallel to t_P.

For the parabola p, the squared length of chords parallel to t_P is proportional to the distance between P and the midpoint of the chord. According to Fig. 8, in our case the factor of proportionality is known as $\overline{S_1S_2}^2 / \overline{PP_0}$. Consequently, the respective position vectors $\mathbf{p}$, $\mathbf{s}_0$, and $\mathbf{m}$ of P, S_0, and the center M of the wanted ellipse e are related by

$$\mathbf{m} = \mathbf{p} + \frac{a_e^2}{\rho^2}(\mathbf{s}_0 - \mathbf{p}) \,. \tag{25}$$

Now, we can parametrize the motion of a given ellipse e on $\mathcal{P}_0$ in the following way. By (24), the given semiaxes define the locus of points $P \in \mathcal{P}_0$ with proportional elliptic coordinates

$$v := \frac{k_2}{k_1} = \frac{a_e^2}{b^2}, \quad \text{where} \quad v > 1 \,.$$

In the same way as before, we use $t := -k_2$ as the motion parameter. Then the pair of elliptic coordinates $k_1 = t/v$ and $k_2 = t$ yields the trajectory $\mathbf{p}(t)$ of the point $P \in \mathcal{P}_0$ by (21). For each admissible t, we compute the Meusnier point C_2 by (23) and then its pedal point C^* in the plane ε, as described above. Finally, due to (25), we can find the correct position of the ellipse $e \in \mathcal{P}_0$ in a plane parallel to τ_P.

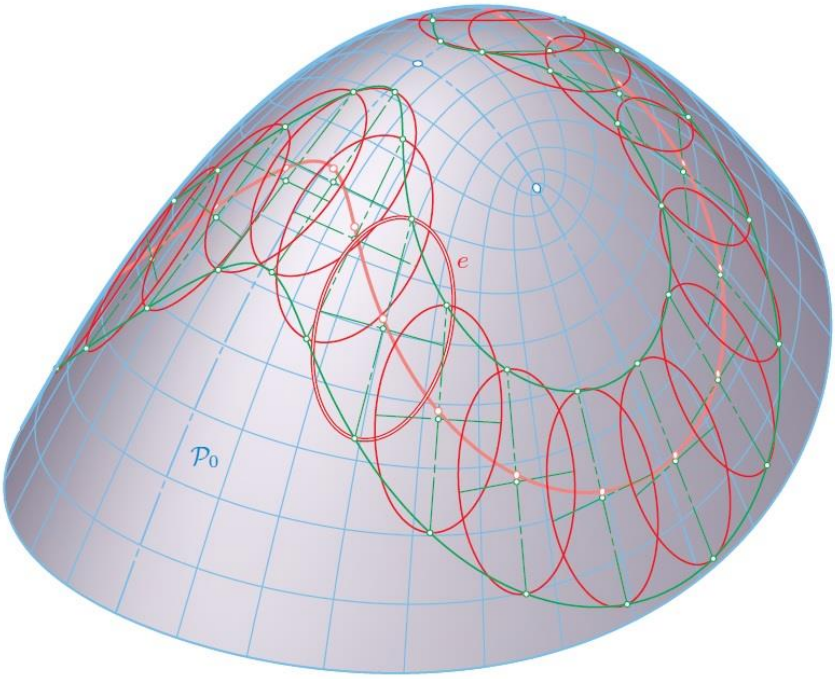


Fig. 9. Ellipse e moving on the elliptic paraboloid $\mathcal{P}_0$ – displayed together with the trajectories of the principal vertices of e (green) and the related curve of constant ratio of principal curvatures (red).

We summarize:

Theorem 2. On regular quadrics $\mathcal{Q}$, all ellipses e other than circles can be moved, except on a one-sheeted hyperboloid the gorge ellipse and on a triaxial ellipsoid the ellipse with the longest and the shortest diameter as axes. During these motions, the points $P \in \mathcal{Q}$ with a tangent plane parallel to the plane of e trace curves with a constant ratio of elliptic coordinates on $\mathcal{Q}$.

References

[1] BIANCHI, L., *Vorlesungen über Differentialgeometrie*. B. G. Teubner, Leipzig, 1899.

[2] BRAUNER, H., *Quadriken als Bewegflächen*. In Monatsh. Math. 59, 45–63, (1955).

[3] KÜHNEL, W., *Differential Geometry, Curves Surfaces Manifolds*. 3rd ed., American Mathematical Society, 2015.

[4] ODEHNAL, B., STACHEL, H., GLAESER, G., *The Universe of Quadrics*. Springer Verlag Berlin Heidelberg 2020, 606 p., 365 Figs., ISBN 978-3-662-61052-7.

Prof. Hellmuth Stachel, PhD, Dr. h. c.

Vienna University of Technology
Wiedner Hauptstr. 8-10/104, A 1040 Wien, Austria
e-mail: stachel@dmg.tuwien.ac.at

Orthogonal axonometry: How can it be determined?

Mária Vojteková, Oľga Blažeková

Abstrakt

Axonometrické zobrazenie je vhodná forma dvojrozmernej reprezentácie trojrozmerných objektov využívaná v technickom kreslení a architektúre. Cieľom je uchovať priestorový dojem scény bez skreslenia spôsobeného vzdialenosťou pozorovateľa. V tomto článku uvádzame nové možnosti určenia kolmej axonometrie, transformačné vzťahy platné medzi nimi a rovnice pre výpočet súradníc priemetu obrazu bodu v kolmej axonometrii na základe podmienok jej určenia.

Kľúčové slová: premietanie, kolmá axonometria, rovnice zobrazenia

Abstract

In technical drawing and in architecture axonometric projection is a form of a two-dimensional representation of three-dimensional objects. The goal is to preserve a spatial impression without distortion due to the distance from an observer. In this paper we give new possibilities to determine an orthogonal axonometry, transformation relations between them, and image equations of a point in orthogonal axonometry based on these options of determination.

Keywords: projection, orthogonal axometry, image equation

1 Introduction

There are well-known examples of conics which can be moved on quadrics. Apart from the trivial case of circles on a sphere, paraboloids are surfaces of translation, even with a continuum of translational nets of parabolas. On quadrics of revolution, each planar section can be rotated while it remains on the quadric. Technical drawings need to be precise, accurate and unambiguous, so engineers and technicians use orthogonal projections (Monge projection, Method of contouring, etc.). On the other hand, for most people it is hard to imagine an object from e.g. Monge mapping. Drawings in linear perspective give a feeling of reality, but there is a problem in change of size of objects depending on the distance from the observer [1]. The compromise is an axonometry with its fixed relation between sizes of objects in space and those on projected space and its good visualization.

Axonometry originated in China. Some concepts of axonometry (especially isometry) had existed in a rough empirical form for centuries well before William Farish (1759–1837), professor at Cambridge University, who was the first to provide detailed rules for isometric drawing. Farish published his ideas in 1822 in paper "On Isometrical Perspective", in which he recognized "need for accurate technical working drawings free of optical distortion" [2]. Since then axonometry became an important graphic technique for artists, architects, and engineers. It usually comes as a standard feature of CAD systems and other visual computing tools.

2 Principle of axonometry

Axonometry is a graphical procedure belonging to descriptive geometry that generates a planar image of a three-dimensional object. The term "axonometry" means "to measure along axes",

and indicates that the dimensions and scaling of the coordinate axes play a crucial role. Axonometry is a parallel projection of a space onto one plane, into which we also project base elements of the coordinate system attached to the object.

Let the Cartesian coordinate system with the origin O, axes x, y, z and planes $\pi = (x, y)$, $\nu = (x, z), \mu = (y, z)$ be given in the Euclidean space. Axonometric image plane ρ will be determined as a plane that is not parallel to any axis and let direction $\vec{s}$ be not parallel to the plane ρ. Let the positive semi axes intersect image plane ρ at the points X, Y and Z. Image of the origin O in parallel projection onto the plane ρ in direction $\vec{s}$ will be denoted O_a, images of the coordinate axes x, y, z by x_a, y_a, z_a (Fig. 1).

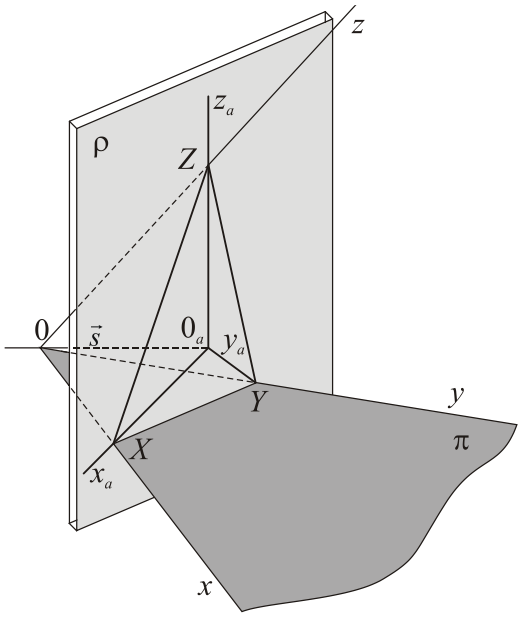


Fig. 1. Axonometric projection

Three points X, Y, Z form the axonometric triangle and lines x_a, y_a, z_a form the axonometric axial cross. The drawing plane can be identical to the plane ρ, or a plane parallel to the plane ρ. By shifting of the drawing plane in the projection direction $\vec{s}$, one changes only the size of the axonometric triangle XYZ, but the projection of the axonometric axial cross does not change. The coordinate axis z_a is usually drawn vertically. All axonometric triangles are homothetic with the centre at point O_a and they determine the same axonometry. Let the images of a measurement unit on axes x, y, z are denoted by p, q, r. The question is, whether one can situate them in the drawing plane arbitrarily. The answer is given by Pohlke´s theorem: The three line segments with a common beginning point and not contained in a line, can be considered as a parallel projection of the three adjacent edges of a cube, see for instance [3]. Following formula is valid for the values p, q, r denoted as coefficients of change

$$p^2 + q^2 + r^2 = 2 + \operatorname{cotg}^2 \theta, \tag{1}$$

where θ is the angle between the projection direction and the image plane. Proof of this famous theorem may be found for example in [3]. If the direction of projection is perpendicular to the image plane, axonometry is said to be normal or orthogonal (English literature usually refers to it as "axonometric projection") [4]; otherwise it is said to be skew.

3 Orthogonal axonometry

We worked on these theorems about orthogonal axonometry:

- Axonometric triangle XYZ is an acute-angled triangle.
- The axes x_a, y_a, z_a are altitudes in the triangle XYZ and the point O_a is its orthocenter.
- $p, q, r \in (0,1)$.
- $p^2 + q^2 + r^2 = 2.$ (2)

Proofs can be found e.g. in [5].

An orthogonal axonometry is graphically given in a drawing plane by any of the following:

a) axonometric triangle,

b) axonometric axial cross.

Our objective was to analyse various options of determining an orthogonal axonometry, to establish transformation relations between them, and to compute planar coordinates of axonometric image of an arbitrary point.

Let the Cartesian coordinate base (O', x', y') with the origin $O' = O_a$ and $y' = z_a$ be given in the image plane. Orthogonal axonometry may be determined by:

a) **values of angles $\boldsymbol{\alpha}, \boldsymbol{\beta}$** (Fig. 2); $\alpha = \pi - \sphericalangle(x_a, x')$, $\beta = \sphericalangle(y_a, x')$, $\alpha \in (0, \pi/2)$, $\beta \in (0, \pi/2)$, and $\alpha + \beta < \pi/2$. Intervals for values of angles follow from Pohlke´s theorem and the property that the axonometric triangle XYZ is acute-angled, the proof can found in [6]. This option of definition is the same as an orthogonal axonometry given graphically by an axonometric axial cross, e.g. $\sphericalangle(x_a, y_a)$, $\sphericalangle(z_a, y_a)$.

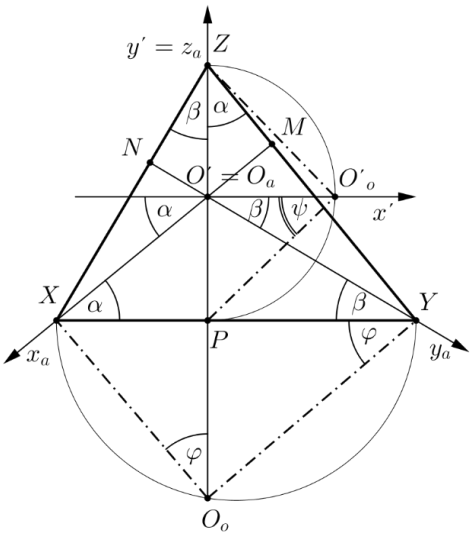


Fig. 2. Axonometric triangle and axes

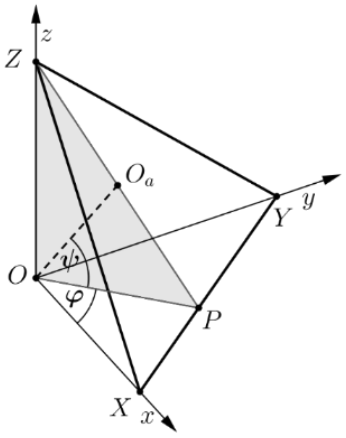


Fig. 3. Angles φ and ψ

b) **values of $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$**; $p, q, r \in (0,1)$ while only two of them are required, as the third one can be calculated from formula $p^2 + q^2 + r^2 = 2$.

c) **lengths of sides of axonometric triangle,** whereby only the ratio of them $|XY|:|YZ|:|ZX|$ is necessary, because parallel image planes intersect axes x, y, z in homothetic axonometric triangles.

d) **values of angles φ, ψ** (Fig. 3); the angles represent the second and the third spherical coordinate of the point O_a in space (the first spherical coordinate - the length of the line segment OO_a does not play any role, because parallel image planes intersect axes x, y, z in homothetic axonometric triangles). In general, the angle $\varphi \in \langle 0, 2\pi)$, without loss of generality we can assume the front view, $\varphi \in (0, \pi/2)$, and the view from top, $\psi \in (0, \pi/2)$, since the goal of projection is a good visualization. Other options have similar calculations.

It is understood that length of p, q, r may be graphically determined by the rotation of the coordinate plane $\pi = (x, y)$ to the image plane ρ around the line XY (Fig. 2), O_o is the rotated position of the origin O, O_oX and O_oY are perpendicular, and the values of p, q can be calculated as ratios of the projected and rotated images

$$p = \frac{|O_aX|}{|O_0X|}, \quad q = \frac{|O_aY|}{|O_0Y|}. \tag{3}$$

Analogously, the length of r is determined by the rotation of the plane defined by points O, P, Z to the image plane ρ around the axis z_a (Fig. 2). The line PO_o' is perpendicular to the line $O_o'Z$ and

$$r = \frac{|O_aZ|}{|O_0'Z|}. \tag{4}$$

The constructions mentioned above are described in detail in [7].
The angles $\alpha, \beta, \varphi, \psi$ are denoted in Fig. 2, being

$$\varphi = \sphericalangle XO_oP, \ \psi = \sphericalangle PO_o'O_a.$$

To deduce relations between defining options one must consider the following statements: $\sin\varphi = \frac{|XP|}{|XO_0|}$ in the triangle XO_oP, $\cos\alpha = \frac{|XP|}{|XO_a|}$ in the triangle XO_aP, $\cos\varphi = \frac{|YP|}{|YO_0|}$ in the triangle YO_oP and $\cos\beta = \frac{|YP|}{|YO_a|}$ in the triangle YO_aP give the equations

$$\sin\varphi = p\cos\alpha, \tag{5}$$

$$\cos\varphi = q\cos\beta. \tag{6}$$

Since $\cos\psi = \frac{|O_aZ|}{|O_0'Z|}$ in the triangle $O_aO_o'Z$, one may express

$$\cos\psi = r. \tag{7}$$

The length of the line segment OP can be expressed by the lengths of line segments

$$|O_0´P| = |O_0P|. \tag{8}$$

Values of the inner angles in the axonometric triangle XYZ are $\pi/2-\beta, \pi/2-\alpha, \alpha+\beta$.

The following theorems and formulas are considered in order to deduce transformation relations:

- Goniometric functions in a right-angled triangle;
- Euclid´s theorems about the altitude and the leg in a right-angled triangle;
- Sine formula and cosine formula for a triangle;
- Formulas for calculation of an area of a triangle.

We deduced each of the twelve transformation relations with the aid of previous statements. To determine angles $\alpha, \beta, \varphi, \psi$, only one goniometric function is needed because all of them are in the interval $(0, \pi/2)$.

4 Transformation relations

a) Knowing the angles α, β the values p, q, r are sought.
Considering $|XO_o|^2 = |XY| \cdot |XP|$, $|YO_o|^2 = |XY| \cdot |YP|$, $|PO_o|^2 = |XP| \cdot |YP|$ in the triangle XYO_o, $\cos\alpha = \dfrac{|XP|}{|XO_a|}$ in the triangle XPO_a, $\sin\beta = \dfrac{|XN|}{|XY|}$ in the triangle XYN and $\sin(\alpha+\beta) = \dfrac{|XN|}{|XO_a|}$ in the triangle XO_aN, we calculate

$$p^2 = \frac{|XO_a|^2}{|XO_0|^2} = \frac{|XO_a|^2}{|XY||XP|} = \frac{|XO_a|}{|XY|} \cdot \frac{1}{\cos\alpha} = \frac{|XO_a|}{\cos\alpha} \cdot \frac{\sin\beta}{|XN|} = \frac{\sin\beta}{\cos\alpha \cdot \sin(\alpha+\beta)}. \tag{9}$$

Similarly considering the triangles YPO, XYM, YO_aM we deduce

$$q^2 = \frac{\sin\alpha}{\cos\beta \cdot \sin(\alpha+\beta)}. \tag{10}$$

Using Eq. (6), (8) and relations in the triangles $PO_0´O_a, XPO_a, YPO_a$ we calculate

$$\begin{aligned} r^2 &= \cos^2\psi = 1-\sin^2\psi = 1-\frac{|PO_a|^2}{|PO_0´|^2} = \\ &= 1-\frac{|PO_a|^2}{|PO_0|^2} = 1-\frac{|PO_a||PO_a|}{|XP||YP|} = 1-\operatorname{tg}\alpha \cdot \operatorname{tg}\beta. \end{aligned} \tag{11}$$

The value r can be also obtained from (2).

b) Knowing the angles α, β the ratios of sides of axonometric triangle $|XY|:|YZ|:|ZX|$ are sought by using sine formula for the inner angles of the axonometric triangle

$$|XY|:|YZ|:|ZX| = \sin(\alpha+\beta):\sin\left(\frac{\pi}{2}-\beta\right):\sin\left(\frac{\pi}{2}-\alpha\right) = \\ = \sin(\alpha+\beta):\cos\beta:\cos\alpha\,. \tag{12}$$

c) Knowing the angles α, β the angles φ, ψ are sought.

d)

$$\operatorname{tg}^2\varphi = \frac{|PO_0|^2}{|PY|^2} = \frac{|XP|\,|YP|}{|PY|^2} = \frac{\dfrac{|XP|}{|O_aP|}}{\dfrac{|YP|}{|O_aP|}} = \frac{\operatorname{cotg}\alpha}{\operatorname{cotg}\beta}\,. \tag{13}$$

Considering equation (11) $\sin\psi$ is found

$$\sin^2\psi = 1-\cos^2\psi = \operatorname{tg}\alpha\cdot\operatorname{tg}\beta\,. \tag{14}$$

e) Knowing the values p, q, r the angles α, β are sought.
Firstly by aid of relations in the triangles $O_aPX, XO_oP, O_aPY, YO_oP$ we calculate

$$p\sin\alpha = \sin\psi\cdot\cos\varphi\,, \tag{15}$$

$$q\sin\beta = \sin\psi\cdot\sin\varphi\,. \tag{16}$$

Applying (15), (7) and (5) we obtain

$$\sin^2\alpha = \frac{\left(1-p^2\right)\left(1-r^2\right)}{p^2r^2}\,. \tag{17}$$

Likewise applying (16), (4) and (6) we find

$$\sin^2\beta = \frac{\left(1-q^2\right)\left(1-r^2\right)}{q^2r^2}\,. \tag{18}$$

f) Knowing the values p, q, r the ratios of sides of axonometric triangle $|XY|:|YZ|:|ZX|$ are sought. Substituting (17) and (18) into (12) ratios can be expressed as

$$|XY|:|YZ|:|ZX| = \frac{\sqrt{1-r^2}}{pq}:\frac{\sqrt{1-p^2}}{rq}:\frac{\sqrt{1-q^2}}{rp}\,, \tag{19}$$

$$|XY|:|YZ|:|ZX| = r\sqrt{1-r^2}:p\sqrt{1-p^2}:q\sqrt{1-q^2}\,. \tag{20}$$

g) Knowing the values p, q, r the angles φ, ψ are sought.
Applying (6) and (18) we have

$$\cos\varphi = \frac{\sqrt{1-p^2}}{r}, \text{ respectively } \sin\varphi = \frac{\sqrt{1-q^2}}{r}. \tag{21}$$

Referring to (7), it holds $\cos\psi = r$.

h) Knowing the sides of the axonometric triangle $|XY|, |YZ|, |ZX|$ or their ratio, angles α, β are sought. There are more possibilities to find them. The first one uses cosine formula for the inner angles in the axonometric triangle

$$\sin\beta = \cos\left(\frac{\pi}{2}-\beta\right) = \frac{|ZX|^2+|XY|^2-|ZY|^2}{2|XY||ZX|}, \tag{22}$$

$$\sin\alpha = \cos\left(\frac{\pi}{2}-\alpha\right) = \frac{|ZY|^2+|XY|^2-|ZX|^2}{2|XY||ZY|}. \tag{23}$$

Other possibility is to apply Heron´s formula for the area S of the triangle XYZ

$$\cos\beta = \sin\left(\frac{\pi}{2}-\beta\right) = \frac{2S}{|XY||ZX|}, \tag{24}$$

$$\cos\alpha = \sin\left(\frac{\pi}{2}-\alpha\right) = \frac{2S}{|XY||ZY|}. \tag{25}$$

i) Knowing the sides of axonometric triangle $|XY|, |YZ|, |ZX|$ or their ratio, p, q, r are sought. Applying equations (9), (24), (25) we get

$$p^2 = \frac{|YZ|^2\sqrt{|XY|^2|ZX|^2-4S^2}}{4S^2}, \tag{26}$$

$$q^2 = \frac{|ZX|^2\sqrt{|XY|^2|YZ|^2-4S^2}}{4S^2}. \tag{27}$$

The value r is obtained from (2).

j) Knowing the sides of axonometric triangle $|XY|, |YZ|, |ZX|$ or their ratio, the angles φ, ψ are sought. Applying (5), (26), (7) and (14) we express

$$\sin^2\varphi = \frac{\sqrt{|XY|^2|ZX|^2-4S^2}}{|XY|^2}, \tag{28}$$

$$\sin^2\psi = \frac{\sqrt{\left(|XY|^2|YZ|^2 - 4S^2\right)\left(|XY|^2|ZX|^2 - 4S^2\right)}}{4S^2}. \tag{29}$$

k) Knowing the angles φ, ψ the angles α, β are sought. By using (5), (6), (15) and (16) we have

$$\operatorname{tg}\alpha = \sin\psi \operatorname{cotg}\varphi, \tag{30}$$

$$\operatorname{tg}\beta = \sin\psi \operatorname{tg}\varphi. \tag{31}$$

l) Knowing the angles φ, ψ the values p, q, r are sought. By using (5), (6), (15) and (16) we calculate

$$p^2 = 1 - \cos^2\varphi \cos^2\psi, \tag{32}$$

$$q^2 = 1 - \sin^2\varphi \cos^2\psi. \tag{33}$$

Referring to (7), $\cos\psi = r$.

m) Knowing the angles φ, ψ the ratios of sides of axonometric triangle $|XY|:|YZ|:|ZX|$ are sought. Applying of (5), (6), (15), (16), (32) and (33) we express

$$\begin{aligned} &|XY|:|YZ|:|ZX| = \sin\psi : p\cos\varphi : q\sin\varphi = \\ &= \sin\psi : \sqrt{1-\cos^2\varphi\cos^2\psi}\cdot\cos\varphi : \sqrt{1-\sin^2\varphi\cos^2\psi}\cdot\sin\varphi . \end{aligned} \tag{34}$$

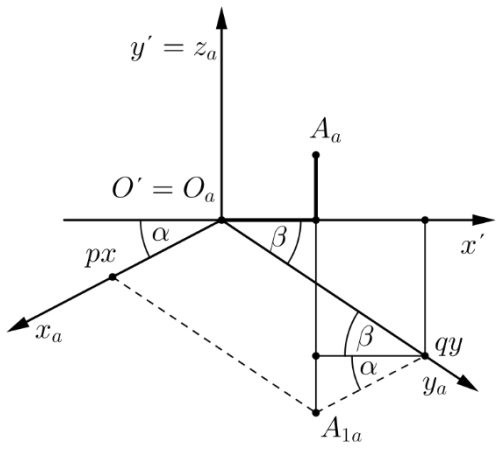


Fig. 4. Projection of the point A

5 Image equations of a point in an orthogonal axonometry

Image equations of a point with coordinates $A[x, y, z]$ into a plane with Cartesian base O', x', y' in an orthogonal axonometry (Fig. 4) are [6]:

$$\begin{aligned} x' &= -p\cos\alpha\cdot x + q\cos\beta\cdot y \ , \\ y' &= -p\sin\alpha\cdot x - q\sin\beta\cdot y + r\cdot z \ . \end{aligned} \tag{35}$$

Determining only the angles α, β and applying (9) and (10) we express

$$\begin{aligned} x' &= -\sqrt{\frac{\sin\beta\cos\alpha}{\sin(\alpha+\beta)}}\cdot x + \sqrt{\frac{\sin\alpha\cos\beta}{\sin(\alpha+\beta)}}\cdot y \ , \\ y' &= -\sqrt{\frac{\sin\beta\sin^2\alpha}{\cos\alpha\sin(\alpha+\beta)}}\cdot x - \sqrt{\frac{\sin\alpha\sin^2\beta}{\cos\beta\sin(\alpha+\beta)}}\cdot y + \sqrt{1-\mathrm{tg}\alpha\ \mathrm{tg}\beta}\cdot z \ . \end{aligned} \tag{36}$$

Determining the angles φ, ψ and applying (15), (16) the image equations are

$$\begin{aligned} x' &= -\sin\varphi\cdot x + \cos\varphi\cdot y \ , \\ y' &= -\sin\psi\cos\varphi\cdot x - \sin\psi\sin\varphi\cdot y + \cos\psi\cdot z \ . \end{aligned} \tag{37}$$

Determining the values p, q, r and applying (21) the image equations are

$$\begin{aligned} x' &= -\frac{\sqrt{1-q^2}}{r}\cdot x + \frac{\sqrt{1-p^2}}{r}\cdot y \ , \\ y' &= -\sqrt{1-r^2}\,\frac{\sqrt{1-p^2}}{r}\cdot x - \sqrt{1-r^2}\,\frac{\sqrt{1-p^2}}{r}\cdot y + r\cdot z \ . \end{aligned} \tag{38}$$

Calculation of image equations for the option of specifying orthogonal axonometry by lengths of an axonometric triangle is possible. It is, however uselessly intricate and not very often used in computer graphics.

6 Conclusions

We pointed out advantage of axonometry in comparison to Monge projection and linear perspective. Various options were shown how orthogonal axonometry could be determined regarding to image equations of a point. Our main goal was to deduce transformation relations between four considered options: the angles α, β formed by axonometric images of axes and Cartesian planar axes in the image plane; the values of coefficients of change - p, q, r; the lengths of sides of axonometric triangle or their ratio; and the angles φ, ψ determining position of the image plane in the space.

Finally, we derived formulas for calculation of coordinates of axonometric image of and arbitrary space point using various possibilities of orthogonal axonometry specification.

Acknowledgement

This paper is an output of the science project 12/PEDAS/2019, FPEDAS, University of Žilina.

References

[1] *Axonometric projections – a technical overview*, [online] Available at: https://www.compuphase.com/axometr.htm [Accessed: 20-05-2019].

[2] KRIKKE, J., *Axonometry: A Matter of Perspective*, IEEE Computer Graphics and Applications, 20(4), pp. 7-11, 2000. http://dx.doi.org/10.1109/38.851742.

[3] MEDEK, V., ZÁMOŽÍK, J., *Konštruktívna geometria pre technikov*, Alfa, Bratislava 1978.

[4] *Axonometric projection*, [online] Available at: https://www.revolvy.com/page/Axonometric-projection [Accessed: 20-05-2019].

[5] URBAN, A., *Deskriptivní geometrie I*, SNTL/ALFA, Praha 1977.

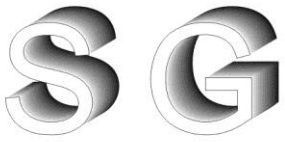
[6] ČERNÝ, J., KOČANDRLOVÁ, M., *Konstruktivní geometrie*, Vydavatelství ČVUT, Praha 1998.

[7] ČENĚK, G., MEDEK, V., *Kurz deskriptívnej geometrie pre technikov I*, Štátne nakladateľstvo technickej literatúry, Bratislava 1953.

RNDr. Mária Vojteková, PhD.
Department of Quantitative Methods
and Economic Informatics
Faculty of Operation and Economics of Transport
and Communications
University of Žilina
Univerzitná 8215/1, 010 26 Žilina, Slovak Republic
e-mail: maria.vojtekova@fpedas.uniza.sk

RNDr. Oľga Blažeková, PhD.
Department of Quantitative Methods
and Economic Informatics
Faculty of Operation and Economics of Transport
and Communications
University of Žilina
Univerzitná 8215/1, 010 26 Žilina, Slovak Republic
e-mail: olga.blazekova@fpedas.uniza.sk

SLOVENSKÁ SPOLOČNOSŤ



PRE GEOMETRIU A GRAFIKU

SLOVAK SOCIETY FOR GEOMETRY AND GRAPHICS

is a non-profit scientific organisation with the objective
to stimulate scientific research and teaching methodology
in the fields of geometry and computer graphics
and to foster international collaboration.

SSGG informs on organisation of different scientific events related to geometry and computer graphics organised in Slovakia.

SSGG provides a platform for donations and sponsorship of scientific workers in the related fields (especially young ones) in order to stimulate scientific development in these disciplines and to enhance the quality of geometry and graphics education of engineers and designers particularly.

Society is publisher of G, the first Slovak scientific journal for geometry and graphics.

All other activities dealing with dissemination of knowledge in the fields of geometry and graphics are welcome, discussion forum on Internet, software market, workshops, Internet courses and chats, etc., and can be provided within the scope of society activities.

Slovak Society for Geometry and Graphics is a collective member of ISGG - International Society for Geometry and Graphics.

SSGG
Institute of Mathematics and Physics
Faculty of Mechanical Engineering
Slovak University of Technology in Bratislava
Námestie slobody 17, 812 31 Bratislava, SR
e-mail: ssgg@ssgg.sk, URL: www.ssgg.sk

Abstracts

M. Bizzarri, M. Lávička, J. Vršek: Note on determining approximate symmetries of planar algebraic curves with inexact coefficients

This paper is devoted to a certain modification of the recently published method for an approximate reconstruction of inexact planar curves which are assumed to be perturbations of some unknown planar symmetric curves. The input curve is given by a perturbed polynomial and the reconstruction steps follow the results from the recently published papers. The functionality of the designed approach is presented on particular examples.

P. Magrone: Sierpinski's curve: a (beautiful) paradigm of recursion

This paper focuses on the original articles written by Waclaw Sierpinski in 1915, when he introduced the recursive structure that bears his name, the Sierpinski's triangle. His first aim was to exhibit the example of a new set, a curve traced starting from the geometry of the well-known triangle. The triangle, which embodies geometric recursion, was rigorously defined in 1915, but appeared also before Sierpinski, and is still a reference point for scientists.

H. Stachel: Moving ellipses on quadrics

For each regular quadric in the Euclidean 3-space, there is a three-parameter set of cutting planes, but the size of an ellipse or hyperbola depends only on its two semiaxes. Therefore, on each quadric Q there exist ellipses or hyperbolas with a one-parameter set of congruent copies, which can even be moved into each other. For the case of ellipses, we present parametrizations of motions on ellipsoids, hyperboloids, and paraboloids. These motions are closely related to the theory of confocal quadrics.

M. Vojteková, O. Blaženková: Orthogonal axonometry: How can it be determined?

In technical drawing and in architecture axonometric projection is a form of two-dimensional representation of three-dimensional objects. The goal is to preserve a spatial impression without distortion due to the distance from an observer. In this paper we give new possibilities to determine an orthogonal axonometry, transformation relations between them, and image equations of a point in orthogonal axonometry based on these options of determination.

G
Slovak Journal for Geometry and Graphics

Volume 17 (2020), No. 33
ISSN 1336-524X

Edited by:
Slovak Society for Geometry and Graphics

SSGG

G is a scientific journal covering the fields of geometry and graphics for publication of original scientific papers, review and information articles, reports, state-of-the-art reviews, communications and other contributions from all geometric disciplines (elementary, descriptive, constructive, projective, coordinate, differential, algebraic, computer, computational, finite, non-Euclidean) and topology, geometric modelling and computer graphics, in the area of the basic theoretical research, education of geometry in all types of schools, from the history and methodology of the development of geometry and on applications of geometry and geometric methods in different scientific, social or technical disciplines.

Editorial office: Slovak Society for Geometry and Graphics
IČO: 31 816 304
Faculty of Mechanical Engineering
Slovak University of Technology in Bratislava
Námestie slobody 17
812 31 Bratislava, Slovakia

Correspondence concerning subscriptions, claims and distribution:
Redakcia G - SSGG
SjF STU, Námestie slobody 17, 812 31 Bratislava, Slovakia
ssgg@ssgg.sk

Frequency: One volume per year consisting of two issues at a price of EUR 20,- per volume, not including surface mail shipment abroad.
Evidentiary number EV 3228/09

Information and instructions for authors are available at the address: www.ssgg.sk

Printed by: ForPress Nitrianske tlačiarne, s.r.o.

G is cited in: Zentralblatt für Mathematik

www.ssgg.sk