TECHNICAL UNIVERSITY OF CIVIL ENGINEERING
BUCHAREST

PROCEEDINGS OF THE 21ST WORKSHOP ON MATHEMATICS, COMPUTER SCIENCE AND TECHNICAL EDUCATION

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
VOLUME 7 / 2024

Bucharest, May 24-26, 2024

ISSN 2601-9299
ISSN-L 2601-9299

CONTENTS

CERTAIN TRANSCENDENTAL ELEMENTS IN CERTAIN EXTENSIONS OF LOCAL FIELDS

SEVER ACHIMESCU AND VICTOR ALEXANDRU

ABSTRACT. Let $(K, |\cdot|)$ be a complete field with respect to a discrete absolute value $|\cdot|$ with a finite residue field. Let T be a transcendental element over K belonging to a completion of a fixed algebraic closure of K to which $|\cdot|$ extends uniquelly. In [1] we defined $\Delta_K(T)$ an analogous of the notion of the discriminant of an algebraic element. For many elements T we notice that $\Delta_K(T) = 0$. The goal of this paper is to find more examples of T with $\Delta_K(T) \neq 0$.

Mathematics Subject Classification (2010): 11S99, 13A18, 12J05, 12J10, 12J25.
Key words:

1. Introduction

In [1] on the one hand we defined an invariant analogous to the discriminant over a local field K for transcendental elements of $\widetilde{\overline{K}} = \Omega$ over K and on the other hand we gave an example of a transcendental element with non-zero invariant. In this paper we continue and we give another such an example. The below *Background material* is quoted from [1].

2. Background material

By a local field $(K\,, |\cdot|)$ we understand a complete field with respect to a discrete absolute value $|\cdot|$ with a finite residue field. For example, if $charK = 0$ then K is isomorphic to a finite extension of the p-adic field $\mathbf{Q_p}$ and $|\cdot|$ is the unique extension to K of the p-adic absolute value $\overline{\mathbf{Q}}_\mathbf{p}$, normalized such that $|p| = \frac{1}{p}$. It is well known (see [1]) that $|\cdot|$ uniquelly extends to a fixed algebraic closure $\overline{K}$ of $K$and further by continuity to $\widetilde{\overline{K}} = \Omega$, the completion of $\overline{K}$ with respect to $|\cdot|$.

Since any $\sigma \in Gal(\overline{K}/K)$ is an isometry with respect to $|\cdot|$, it follows that σ uniquelly extends to a continuous K-automorphism of Ω. Let us denote $G_K = Gal_{cont}(\Omega/K) \simeq Gal(\overline{K}/K)$. Let $T \in \Omega$. Let $C_K := \{\sigma(T)\,,\, \sigma \in G_K\}$ be the orbit of T with respect to the action of G_K on Ω. Recall from [3], Theorem 3.5 that if T is transcendental over K then C_K is an infinite compact subset of Ω.

In [4] at page 29 it is associated a chain to each $T \in \Omega$ defined in terms of the distances from T to its conjugates over K, that is $\{\sigma(T)\,,\, \sigma \in G_K = Gal_{cont}(\Omega/K) \cong Gal(\overline{K}/K)\}$. The orbit of T with respect to the action of the group G_K is $C_K(T) := \{\sigma(T)\,,\, \sigma \in G_K\}$. $C_K(T)$ is always a compact set and it is a finite set if and only if $T \in \overline{K}$. Let us denote $B[T, \epsilon] = \{\beta \in \Omega\,,\, |\beta - T| \leq \epsilon\}$

Let $N(K,T,\epsilon)$ the number of disjoint such closed balls of radius ϵ covering $C_K(T)$. The function

$$(0,\infty) \longrightarrow \mathbf{N} - \{0\}$$
$$\epsilon \longrightarrow N(K,T,\epsilon)$$

is a decreasing step function. It is bounded if and only if $T \in \overline{K}$. Its image is an increasing sequence $1 = N_1 < N_2 < ...$ whch is infinite if and only if $T \in \Omega - \overline{K}$.

Let $\epsilon_j = inf\{\epsilon > 0\,,\, N(K,T,\epsilon) = N_j\}$. Since each $\sigma \in G_K$ is an isometry, each of the N_j balls of radius ϵ_j covering $C_K(T)$ is covered by the same number of balls of radius ϵ_{j+1} which intersect $C_K(T)$. It follows that $N_j | N_{j+1}$, for all j. If $T \in \Omega - \overline{K}$ we obtain the infinite chain

$$N_K(T) = \begin{pmatrix} \epsilon_1 & \epsilon_2 & ... \\ N_1 & N_2 & ... \end{pmatrix}$$

Note that $\epsilon_1 = sup\{|T - \sigma(T)|, \,, \sigma \in G_K\}$ is the diameter of C_K and $\epsilon_1 > \epsilon_2 > ...\epsilon_n > ...$ and $lim_{n\to\infty}\epsilon_n = 0$. The sequence $(\epsilon_n)_{n\geq 1}$ is said to be *the fundamental sequence* associated to T. For $T = \alpha \in \overline{K}$ one obtains a finite chain

$$N_K(\alpha) = \begin{pmatrix} \epsilon_1 & \epsilon_2 & ... & \epsilon_{l'_K}(\alpha) \\ N_1 & N_2 & ... & N_{l'_K}(\alpha) \end{pmatrix}$$

where $\epsilon_{l'_K}(\alpha) = 0$ and $N_{l'_K}(\alpha) = deg_K(\alpha)$.

Let $B_1^m, B_2^m, ...B_{N_m}^m$ be a partition of $C_K(T)$ with balls of radius ϵ_m, that is $B_1^m = \{z \in C_K(T)\,,\, |z - T| \leq \epsilon_m\}$. Any $\sigma \in G_K$ permutes $B_1^m, B_2^m, ...B_{N_m}^m$, and if we denote $H_m = \{\sigma \in G_K\,,\, \sigma(B_1^m) = B_1^m\} \subseteq G_K$ we have $N_m = [G_K : H_m]$. Let $T_i \in B_i^m$ arbitrarily fixed with $T_1 = T$. Then $|T - T_i| > \epsilon_m$ for all $i \geq 2$. For example, if R_m is a set of representatives for $(G_K/H_m)_{left}$ then we can pick $T_i = \sigma_i T\,,\, \sigma_i \in R_m$. Let us notice that the number $x_m := \prod_{i=2}^{N_m} |T - T_i|$ does not depend on the choice of $T_i \in B_i^m$, it depends on the choice of T (in fact, of $C_K(T)$) only. Indeed, for $T_i' \in B_i^m$ we have $|T - T_i'| = |T - T_i + T_i - T_i'| = |T - T_i|$ since $|T_i - T_i'| < |T - T_i|$.

The sequence $(x_m^{\frac{1}{N_m}})_m$ converges and its limit is an invariant of T (and of $C_K(T)$).

Let us denote

$$\Delta_K(T) := lim_{m\to\infty} x_m^{\frac{1}{N_m}}$$

Then $\Delta_K(T)$ can be considered an analogous of the discriminant of $T = \alpha \in \overline{K}$ according to the following observation: $x_m^{N_m} = \prod_{i\neq j} |T_j - T_i|$ since we can take $T_i = \sigma_i T$, for a suitable $\sigma_i \in R_m$ and the automorphisms of G_K are also isometries with respect to $|\cdot|$.

Now we give a method for computing the numerical invariants $\Delta_K(T)$ in certain cases.

3. Main Result

Let ξ_{p^n} denote a primitive nth root of unity and let $L = \cup_n \mathbf{Q}_p(\xi_{p^n})$. It is known that $Gal(\mathbf{Q}_p(\xi_{p^n}/\mathbf{Q}_p) \cong U(\mathbf{Z}_{p^n})$ and $[\mathbf{Q}_p(\xi_{p^n} : \mathbf{Q}_p] = \phi(p^n) = p^{n-1}(p-1)$. For each $\sigma \in Gal(\mathbf{Q}_p(\xi_{p^n}/\mathbf{Q}_p)$ and for each n there exists i satisfying $\sigma(\xi_{p^n}) = \xi_{p^n}^i$, $1 \leq i \leq n$ and $(i, p^n) = 1$.

Proposition 1

Let $T=\Sigma_{i=0}^{\infty}a_i\xi_{p^i}$ with $a_0=1$ and $a_i\in\mathbf{Z}$ satisfying $|a_{i+1}|_p<|a_i|_p(\frac{1}{p})^{\frac{1}{p-1}}$ for all $i\geq 0$. Then T is a generic element for $\tilde{L}$, that is $\tilde{L}=\widetilde{\mathbf{Q}_p(T)}=\widetilde{\mathbf{Q}_p[T]}$.

Lemma

Let $\sigma\in Gal(L/\mathbf{Q}_p)=Gal_{cont}(\tilde{L}/\mathbf{Q}_p)$ and let i satisfying $\sigma(\xi_{p^{i-1}})=\xi_{p^{i-1}}$ (it is known that such an i exists). Then $|\sigma(\xi_{p^i})-\xi_{p^i}|_p=(\frac{1}{p})^{\frac{1}{p-1}}$.

proof of Lemma

Let $a,b\in\mathbf{Z}$ such that $\sigma(\xi_{p^i})=\xi_{p^i}^a$ and $a=1+bp^{i-1}$. We have

$$|\sigma(\xi_{p^i})-\xi_{p^i}|_p=|\xi_{p^i}(1-\xi_{p^i}^{bp^{i-1}})|_p=|(1-\xi_{p^i}^{bp^{i-1}})|_p=(\frac{1}{p})^{\frac{1}{p-1}}$$

the last equality is to be proved similarly to the equality $|1-\xi_p^b|_p=(\frac{1}{p})^{\frac{1}{p-1}}$ which follows from the fact that the factors of the product $\Pi_{i=1}^{p-1}(1-\xi_p^i)=p$ have the same p-adic absolute value, being conjugates over $\mathbf{Q}_p$, and $|p|_p=\frac{1}{p}$.

proof of Proposition 1

Let $\sigma\in Gal(L/\mathbf{Q}_p)=Gal_{cont}(\tilde{L}/\mathbf{Q}_p)$ be arbitrarly fixed.

Since $|a_{i+s}|_p\leq|a_i|_p(\frac{1}{p})^{\frac{1}{p-1}}$ for each $s\geq 1$ we obtain that $|\sigma(T)-T|_p=|\sigma(\Sigma_{k=0}^{i}a_k\xi_{p^k})-\Sigma_{k=0}^{i}a_k\xi_{p^k}|_p=|a_i|_p(\frac{1}{p})^{\frac{1}{p-1}}$ thus $\sigma(T)\neq T$. Since $\sigma\in Gal(L/\mathbf{Q}_p)=Gal_{cont}(\tilde{L}/\mathbf{Q}_p)$ was arbitrarly fixed we conclude that

$$\sigma(T)\neq T\,,\,\forall\sigma\in Gal(L/\mathbf{Q}_p)=Gal_{cont}(\tilde{L}/\mathbf{Q}_p)$$

Now the conclusion that T is a generic element follows from the article [2].

Proposition 2

Let ξ_{p^n} denote a primitive nth root of unity and let $L=\cup_n\mathbf{Q}_p(\xi_{p^n})$. Let

$$T=1+p^p\xi_p+p^{p^2}\xi_{p^2}+p^{p^3}\xi_{p^3}+...$$

Then

$$\Delta_K(T)\neq 0$$

.

p̂roof

Note that T satisfies the hypothesis of *Proposition 1.* Indeed, $a_i=p^{p^i}\Rightarrow|a_i|_p=(\frac{1}{p})^{p^i}$ and it follows that $|a_{i+1}|_p<|a_i|_p(\frac{1}{p})^{\frac{1}{p-1}}$.

Now we use the notation ϵ_i (the fundamental sequence associated to T), N_i (the number of disjoint closed balls of radius ϵ_i covering the orbit of T), $x_i=\Pi_{k=2}^{N_i}|T-T_k|_p$ introduced in the previous (the second) paragraph. We have that $\epsilon_{i+1}=|a_{i+1}|_p(\frac{1}{p})^{\frac{1}{p-1}}$, $N_{i+1}=p^{i-1}(p-1)$, $x_{i+1}=|\epsilon_1^{p-1}\epsilon_2^p...\epsilon_{i+1}^p|_p=|a_1^{p-1}a_2^p...a_{i+1}^p|_p(\frac{1}{p})^{\frac{i+1}{p-1}}$.

Let us denote v the exponential valuation corresponding to the absolute value $|\cdot|_p$. We have that $v(a_1^{p-1}a_2^p...a_{i+1}^p) = (p-1)+p(p^2+p^3+...+p^{i+1}) = p-1+p^3\frac{p^3-1}{p-1}$ thus

$$x_{i+1}^{\frac{1}{N_{i+1}}} = (\frac{1}{p})^{\frac{p-1+p^3\frac{p^i-1}{p-1}+\frac{i+1}{p-1}}{p^{i-1}(p-1)}}$$

Finally,

$$\Delta_K(T) := lim_{i\to\infty}x_i^{\frac{1}{N_i}} = (\frac{1}{p})^{\frac{p^3}{(p-1)^2}} \neq 0$$

References

[1] S. Achimescu, V.Alexandru and C.S.Andronescu, *A notion analogous to the discriminant for transcendental elements in certain extensions of local fields*, RSMUP 143(2020), 105-112.

[2] V.Alexandru, N.Popescu, A.Zaharescu, *On the closed subfields of Cp*, Journal of Number Theory, 68.2(1998), 131-150.

[3] A.Iovita, A.Zaharescu, *Completions of r.a.t.-valued fields*, Journal of Number Theory, 50(1995), 202-205.

[4] A.Popescu, N.Popescu, M.Vajaitu, A.Zaharescu, *Chains of metric invariants over a local field*, Acta Arithmetica 103.1 (2002).

[5] A.Robert, *A Course in p-adic Analysis* Springer GTM 198.

[6] L.Washington *Introduction to Cyclotomic Fields* Springer GTM 83.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania

Email address: sachimescu@yahoo.com

Department of Mathematics, University of Bucharest, Bucharest, Romania

Email address: vralexandru@yahoo.com

NON-NEWTONIAN APPROACHES IN DIFFERENTIAL GEOMETRY

MUHITTIN EVREN AYDIN, AYKUT HAS AND BEYHAN YILMAZ

Abstract. In this conference paper, we study the multiplicative rectifying submanifolds in the multiplicative Euclidean space. We first consider the multiplicative rectifying curves and charaterize in terms of the multiplicative curvatures. A classification for such curves is also given. Then, as higher dimensional objects, we deal with the multiplicative rectifying submanifolds, characterizing via multiplicative concurrent vector fields. The multiplicative rectifying hypersurfaces are completely classified.

Mathematics Subject Classification (2010): 53B25, 11U10.
Key words: Rectifying curve, rectifying submanifold, multiplicative calculus, multiplicative Euclidean space.

1. Introduction

During the recent years there has been an increasing interest to perform the tools of non-Newtonian calculi in the study of objects in geometry. In this sense, there are many geometrical studies using fractional calculus, see [3, 6, 7, 21, 35]. Moreover, after the comprehensive books of Georgiev and his collaborators [17, 18, 19], authors have started to use multiplicative calculus ([22, 23]) in the study of geometrical objects. For example, see [12, 25, 28].

Besides geometry, there are many applications of multiplicative calculus in dynamical systems [1, 2, 31], in economics [13, 15, 30], and in image analysis [16, 27]. The techniques from multiplicative calculus are also used in complex analysis [8, 10, 33], in differential equations [9, 34, 36], in numerical analysis [26, 29, 37], in algebra [11, 14], in variational analysis [32], and in spectral and Dirac system theories [20, 24, 38].

As far as the authors know at this moment, apart from multiplicative curves and surfaces, the study of multiplicative submanifolds is not considered. In this paper we initiate the study of such submanifolds in the multiplicative Euclidean space and give the multiplicative counterparts of the well-known Gauss-Weingarten formulas. We introduce the notions of multiplicative induced connection, second fundamental form, shape operator and normal connection. In addition, we provide the features of these maps. The notion of multiplicative rectifying submanifold is defined. Then, we recall the results (without proofs) on multiplicative rectifying curves and submanifolds of higher dimension.

The proofs are given in the studies [4, 5].

2. Multiplicative rectifying curves

Let $\mathbb{R}_*^n = \{(x_1, ..., x_n) : x_1, ..., x_n > 0\}$ and $\mathbf{x}, \mathbf{y} \in \mathbb{R}_*^n$. A positive-definite scalar product on $\mathbb{R}_*^n$ is defined by

$$\langle \mathbf{x}, \mathbf{y} \rangle_* = x_1 \cdot_* y_1 +_* ... +_* x_n \cdot_* y_n = e^{\log x_1 \log y_1 + ... + \log x_n \log y_n}.$$

We call $(\mathbb{R}_*^n, \langle , \rangle_*)$ *multiplicative Euclidean space* denoted by $\mathbb{E}_*^n$.

Given a multiplicative biregular curve $\mathbf{x}(s)$ in $\mathbb{E}_*^3$, $s \in I \subset \mathbb{R}_*$, where $\{\mathbf{t}(s), \mathbf{n}(s), \mathbf{b}(s)\}$ is the multiplicative Frenet frame.

Definition 2.1. *We call $\mathbf{x}(s)$ a multiplicative rectifying curve if the multiplicative principal normal component is 1, namely $\langle \mathbf{x}(s), \mathbf{n}(s) \rangle_* = 1$, for every $s \in I$.*

With this definition, we write

$$\mathbf{x}(s) = \lambda(s) \cdot_* \mathbf{t}(s) +_* \mu(s) \cdot_* \mathbf{b}(s),$$

where $\lambda(s) = \langle \mathbf{x}(s), \mathbf{t}(s) \rangle_*$ and $\mu(s) = \langle \mathbf{x}(s), \mathbf{b}(s) \rangle_*$.

In what follows, we characterize the multiplicative rectifying curves in terms of the multiplicative curvatures and position vector.

Proposition 2.2. *If $\mathbf{x}(s) \subset \mathbb{E}_*^3$ with $\kappa \neq 1$ is a multiplicative rectifying curve, then nowhere the multiplicative torsion is 1 and*

$$\langle \mathbf{x}(s), \mathbf{t}(s) \rangle_* = s +_* a = as, \quad \langle \mathbf{x}(s), \mathbf{b}(s) \rangle_* = b, \quad a, b \in \mathbb{R}_*, b \neq 1.$$

The converse statement is true as well.

Proposition 2.3. *If $\mathbf{x}(s) \subset \mathbb{E}_*^3$ is a multiplicative rectifying curve, then the multiplicative ratio of the multiplicative curvatures is*

$$\tau(s) /_* \kappa(s) = c \cdot_* s +_* d = e^{c \log s + \log d}, \quad c, d \in \mathbb{R}_*, c \neq 1.$$

The converse statement is true as well.

Proposition 2.4. *If $\mathbf{x}(s) \subset \mathbb{E}_*^3$ is a multiplicative rectifying curve with $\kappa > 1$, then*

$$\rho(s)^{2_*} = s^{2_*} +_* e^c \cdot_* s +_* e^d, \quad c, d \in \mathbb{R}, d > 0,$$

or equivalently,

$$e^{(\log \rho(s))^2} = e^{(\log s)^2 + c \log s + d}.$$

The converse statement is true as well.

Proposition 2.5. *If $\mathbf{x}(s) \subset \mathbb{E}_*^3$ with $\kappa \neq 1$ is a multiplicative rectifying curve, then $\rho(s)$ is nonconstant and $\|\mathbf{x}^{\perp_*}(s)\|_*$ is constant. The converse statement is true as well.*

Theorem 2.6. *Let $\mathbf{x}(s) \subset \mathbb{E}_*^3$ with $\kappa \neq 1$ be a multiplicative rectifying curve and $\mathbb{S}_*^2$ the multiplicative sphere of radius e. Then, there is a multiplicative reparametrization of $\mathbf{x}(s)$ such that*

$$\mathbf{x}(\tilde{s}) = (a \cdot_* \sec_* \tilde{s}) \cdot_* \mathbf{y}(\tilde{s}), \quad a \in \mathbb{R}_*, a > 1,$$

where $\mathbf{y}(\tilde{s})$ is a parameterized curve lying in $\mathbb{S}_*^2$ by multiplicative arc length. The converse statement is true as well.

3. Multiplicative rectifying submanifolds

Let M be a multiplicative Riemannian manifold and $F : M \to \mathbb{E}_*^m$ a multiplicative isometric immersion. Denote by X and Y some multiplicative tangent vector fields to M. If $\widetilde{\nabla}$ is a multiplicative Riemann connection on $\mathbb{E}_*^m$, then a multiplicative orthogonal decomposition called *multiplicative Gauss formula* is given by

$$\widetilde{\nabla}_{*X}Y = \nabla_{*X}Y +_* h(X,Y),$$

where $\nabla_{*X}Y$ and $h(X,Y)$ are multiplicative tangent and multiplicative normal components of $\widetilde{\nabla}_{*X}Y$ to M. We call h *multiplicative second fundamental form*. Moreover, when $h = 1$, M is called *multiplicative totally geodesic*.

Proposition 3.1. *(1) ∇_* is an multiplicative Riemannian connection on M.*
(2) h is symmetric and multiplicative bilinear.

Proposition 3.2. *An n-dimensional multiplicative totally geodesic submanifold in $\mathbb{E}_*^m$ is an open part of an multiplicative linear subspace $\mathbb{E}_*^n$ of $\mathbb{E}_*^m$. The converse is also true.*

Let ξ be a multiplicative normal vector field to M in $\mathbb{E}_*^m$. A multiplicative orthogonal decomposition of $\widetilde{\nabla}_{*X}\xi$ called *multiplicative Weingarten formula* is uniquely given by

$$\widetilde{\nabla}_{*\widetilde{X}}\xi = -_* A_\xi(X) +_* D_{*X}\xi, \tag{3.1}$$

where $A_\xi(X)$ and $D_{*X}\xi$ are multiplicative tangent and multiplicative normal components of $\widetilde{\nabla}_{*X}\xi$ to M. We call the operator A *multiplicative shape operator*.

Proposition 3.3. *For every multiplicative tangent X, Y and multiplicative normal vector field ξ to M,*
(1) $A_\xi(X)$ is multiplicative linear and

$$\langle h(X,Y), \xi\rangle_* = \langle A_\xi(X), Y\rangle_*,$$

(2) D_* is a multiplicative nature connection satisfying

$$D_{*X}(\langle \xi, \eta\rangle_*) = \langle D_{*X}\xi, \eta\rangle_* +_* \langle \xi, D_{*X}\eta\rangle_*.$$

Definition 3.4. *Let $F : M \to \mathbb{E}_*^m$ be a multiplicative immersion of a multiplicative Riemannian manifold to $\mathbb{E}_*^m$. M is called an multiplicative rectifying submanifold if*

$$\langle F(p), Im\ h_p\rangle_* = 1, \quad \forall p \in M.$$

If also $F^{T_*}(p) \neq 1$ and $F^{N_*}(p) \neq 1$, for every $p \in M$ then M is said to be proper.

Theorem 3.5. *Let $F : M \to \mathbb{E}_*^m$ be a multiplicative isometric immersion of a multiplicative Riemannian manifold to $\mathbb{E}_*^m$ with $F^{N_*} \neq 1$. If M is a proper multiplicative rectifying submanifold then F^{T_*} is multiplicative concurrent on M. The converse is also true.*

Theorem 3.6. *Let $F : M \to \mathbb{E}_*^m$ be a multiplicative immersion to $\mathbb{E}_*^m$ of a multiplicative Riemannian manifold of dimension $(m-1)$. Then, M is a multiplicative rectifying manifold if and only if it is an open part of a multiplicative hyperplane.*

References

[1] D. Aniszewska and M. Rybaczuk, *Analysis of the multiplicative Lorenz system*, Chaos, Solitons and Fractals **25** (2005), 79–90.

[2] D. Aniszewska and M. Rybaczuk, *Lyapunov type stability and Lyapunov exponent for exemplary multiplicative dynamical systems*, Nonlinear Dyn. **54** (2008), 345–354.

[3] M. E. Aydin, A. Has, and B. Yilmaz, *Applications of fractional calculus in equiaffine geometry: plane curves with fractional order*, Math. Meth. Appl. Sci. **44(17)** (2021), 13659–13669.

[4] M. E. Aydin, A. Has, and B. Yilmaz, *A non-Newtonian approach in differential geometry of curves: multiplicative rectifying curves*, Bull. Korean Math. Soc. **61(2)** (2024), 849–866.

[5] M. E. Aydin, A. Has, and B. Yilmaz, *Multiplicative rectifying submanifolds of multiplicative Euclidean space*, Math. Meth. Appl. Sci. (2024), 1–11. DOI: 10.1002/mma.10329.

[6] D. Baleanu and S. I. Vacaru, *Constant curvature coefficients and exact solutions in fractional gravity and geometric mechanics*, Cent. Eur. J. Phys. **9(5)** (2011), 1267-1279.

[7] D. Baleanu and S. I. Vacaru, *Fractional almost Kähler-Lagrange geometry*, Nonlinear Dyn. **64** (2011), 365-373.

[8] A. E. Bashirov, E. M. Kurpınar, and A. Özyapıcı, *Multiplicative calculus and its applications*, J. Math. Anal. Appl. **337(1)** (2008), 36–48.

[9] A. E. Bashirov, E. Mısırlı, Y. Tandoğdu, and A. Özyapıcı, *On modeling with multiplicative differential equations*, Appl. Math. J. Chinese Univ., **26(4)** (2011), 425-438.

[10] A. E. Bashirov and S. Norozpour, *On an alternative view to complex calculus*, Math. Meth. Appl. Sci. **41** (2018), 7313– 7324.

[11] A. F. Cakmak and F. Basar, *Some new results on sequence spaces with respect to non-Newtonian calculus*, J. Inequal. Appl. **228** (2012), 17 pages.

[12] H. Ceyhan, Z. Ozdemir, and I. Gok, *Multiplicative generalized tube surfaces with multiplicative quaternions algebra*, Math. Meth. Appl. Sci. **47** (2024), 9157–9168.

[13] F. Cordova-Lepe, *The multiplicative derivative as a measure of elasticity in economics*, TEMAT-Theaeteto Atheniensi Mathematica **2(3)** (2006), 7 pages.

[14] F. Cordova-Lepe, R. Del Valle, and K. Vilches, *A new approach to the concept of linearity. Some elements for a multiplicative linear algebra*, Int. J. Comp. Math. **97(1-2)** (2020), 109-119.

[15] D. A. Filip and C. Piatecki, *A non-newtonian examination of the theory of exogenous economic growth*, Math. Aeterna **4** (2014), 101–117.

[16] L. Florack and H. van Assen, *Multiplicative calculus in biomedical image analysis*, J. Math. Im. Vis. **42** (2012), 64–75.

[17] S. G. Georgiev, *Multiplicative Differential Geometry*, Chapman and Hall/CRC., New York, 2022.

[18] S. G. Georgiev and K. Zennir, *Multiplicative Differential Calculus*, Chapman and Hall/CRC., New York, 2022.

[19] S. G. Georgiev, K. Zennir, and A. Boukarou, *Multiplicative Analytic Geometry*, Chapman and Hall/CRC., New York, 2022.

[20] S. Goktas, H. Kemaloglu, and E. Yilmaz, *Multiplicative conformable fractional Dirac system*, Turk. J. Math. **46(3)** (2022), 973 – 990.

[21] U. Gozutok, H. A. Coban, and Y. Sagiroglu, *Frenet frame with respect to conformable derivative*, Filomat **33(6)** (2019), 1541-1550.

[22] M. Grossman and R. Katz, *Non-Newtonian Calculus*, Lee Press, Pigeon Cove Massachussets, 1972.

[23] M. Grossman, *Bigeometric Calculus: A System with a Scale-Free Derivative*, Archimedes Foundation, Massachusetts, 1983.

[24] T. Gulsen, E. Yilmaz, and S. Goktas, *Multiplicative Dirac system*, Kuwait J. Sci. **49(3)** (2022), 1-11.

[25] A. Has and B. Yilmaz, *A non-Newtonian magnetic curves in multiplicative Riemann manifolds*, Phys. Scr. **99(4)** (2024), 045239.

[26] E. Misirli and Y. Gurefe, *Multiplicative Adams Bashforth-Noulton methods*, Numer. Algor. **57** (2011), 425–439.
[27] M. Mora, F. Cordova-Lepe, and R. Del-Valle, *A non-Newtonian gradient for contour detection in images with multiplicative noise*, Pattern Recognition Letters **33(10)** (2012), 1245–1256.
[28] S. K. Nurkan, I. Gurgul, and M. K. Karacan, *Vector properties of geometric calculus*, Math. Meth. Appl. Sci. **46** (2023), 17672–17691.
[29] A. Ozyapici, M. Riza, B. Bilgehan, and A. E. Bashirov, *On multiplicative and Volterra minimization methods*, Numer. Algor. **67** (2014), 623–636.
[30] H. Ozyapici, I. Dalcı, and A. Ozyapıcı, *Integrating accounting and multiplicative calculus: an effective estimation of learning curve*, Comp. Math. Org. Th. **23** (2017), 258–270.
[31] M. Rybaczuk, A. Kedzia, and W. Zielinski, *The concept of physical and fractal dimension II. The differential calculus in dimensional spaces*, Chaos Solitons and Fractals **12(13)** (2001), 2537-2552.
[32] D. F. M. Torres, *On a non-Newtonian calculus of variations*, Axioms **10(3)** (2021), 171.
[33] A. Uzer, *Multiplicative type complex calculus as an alternative to the classical calculus*, Comp. Math. Appl. **60(10)** (2010), 2725–2737.
[34] M. Waseem, M. A. Noor, F. A. Shah, and K. I. Noor, *An efficient technique to solve nonlinear equations using multiplicative calculus*, Turk. J. Math. **428(2)** (2018), 679-691.
[35] T. Yajima, S. Oiwa, and K. Yamasaki, *Geometry of curves with fractional-order tangent vector and Frenet-Serret formulas*, FCAA **21** (2018), 1493–1505.
[36] N. Yalcin, *The solutions of multiplicative Hermite differential equation and multiplicative Hermite polynomials*, Rend. Circ. Mat. Palermo, II. Ser **70** (2021), 9–21.
[37] M. Yazici and H. Selvitopi, *Numerical methods for the multiplicative partial differential equations*, Open Math. **15** (2017), 1344–1350.
[38] E. Yilmaz, *Multiplicative Bessel equation and its spectral properties*, Ricerche mat (2021). https://doi.org/10.1007/s11587-021-00674-1.

Department of Mathematics, Faculty of Science, Firat University, Elazig, Turkey
Email address: meaydin@firat.edu.tr

Department of Mathematics, Faculty of Science, Kahramanmaras Sutcu Imam University, Kahramanmaras, Turkey
Email address: ahas@ksu.edu.tr

Department of Mathematics, Faculty of Science, Kahramanmaras Sutcu Imam University, Kahramanmaras, Turkey
Email address: beyhanyilmaz@ksu.edu.tr

EUCLIDEAN SUBMANIFOLDS WITH INFINITESIMAL VARIATIONS AND BENDINGS

ŞTEFAN-CEZAR BROSCĂŢEANU

Abstract. The fundamental equations and the fundamental theorem of infinitesimal variations are recalled and we present alternative proofs of some formulae. As applications, in low dimensional case, we consider infinitesimal bending of rectifying curves.

Mathematics Subject Classification (2010): 53A07, 53B25, 53A04.
Key words: Infinitesimal variation, infinitesimal bending.

1. Preliminaries

1.1. **Infinitesimal Variations.** In [3] the notions of infinitesimal variation and infinitesimal bending of a Euclidean submanifold are introduced.

Let $f : M^n \to \mathbb{R}^m$ be an isometric immersion of a Riemannian manifold into an Euclidian space. A *smooth variation* of f is a smooth map $\mathcal{F} : I \times M^n \to \mathbb{R}^m$, where $0 \in I \subset \mathbb{R}$ is an open interval, such that $f_t = \mathcal{F}(t, \cdot) : M^n \to \mathbb{R}^m$ is an immersion for any $t \in I$ and $f_0 = f$. The *variational vector field* of a variation $\mathcal{F}$ of f is the section $\mathcal{T} \in \Gamma(f^*T\mathbb{R}^m)$ defined by

$$\mathcal{T} = \mathcal{F}_* \frac{\partial}{\partial t}\Big|_{t=0} = \tilde{\nabla}_{\frac{\partial}{\partial t}} f_t\Big|_{t=0}.$$

An infinitesimal variation is the infinitesimal analogue of an isometric variation, being variations that preserve lengths but just up to the first order.

Definition 1.1. [3] A smooth variation $\mathcal{F} : I \times M^n \to \mathbb{R}^m$ of an isometric immersion $f : M^n \to \mathbb{R}^m$ is called an *infinitesimal variation* if it satisfies the condition

$$\frac{\partial}{\partial t}\Big|_{t=0} \langle f_{t*}X, f_{t*}Y\rangle = 0, \tag{1.1}$$

for any $X, Y \in \mathfrak{X}(M)$.

Let $C : I \to O(m)$ be a smooth family of orthogonal transformations of $\mathbb{R}^m$ and let $v : I \to \mathbb{R}^m$ be a smooth map such that $(C(0), v(0)) = (I, 0)$. Then, one defines an isometric variation $\mathcal{F}$ of f by

$$\mathcal{F}(t, x) = C(t)f(x) + v(t),$$

for all $(t, x) \in I \times M^n$.

$\mathcal{F}$ is called a *trivial isometric variation*.

Definition 1.2. [3] A section $\mathcal{T}$ of $f^*T\mathbb{R}^m$ is called *infinitesimal bending* of a isometric immersion $f : M^n \to \mathbb{R}^m$ if the condition

$$\langle \tilde{\nabla}_X \mathcal{T}, f_* Y\rangle + \langle f_* X, \tilde{\nabla}_Y \mathcal{T}\rangle = 0 \tag{1.2}$$

holds for any tangent vector fields $X, Y \in \mathfrak{X}(M)$.

1.2. **The Associated Pair.** An infinitesimal bending $\mathcal{T} \in \Gamma(f^*T\mathbb{R}^*)$ of an isometric immersion $f : M^n \to \mathbb{R}^m$ together with its second fundamental form $h : TM \times TM \to N_f M$ determines an *associate pair* of tensors $(\beta, \mathcal{E})$ [3] to $\mathcal{T}$, where $\beta : TM \times TM \to N_f M$ is symmetric and $\mathcal{E} : TM \times N_f M \to N_f M$ satisfies the compatibility condition

$$\langle \mathcal{E}(X, \eta), \xi\rangle + \langle \mathcal{E}(X, \xi), \eta\rangle = 0, \tag{2.1}$$

for any $X \in \mathfrak{X}(M)$ and $\eta, \xi \in \Gamma(N_f M)$.

Let $L \in \Gamma(Hom(TM, f^*T\mathbb{R}^m))$ be the tensor defined by

$$LX = \tilde{\nabla}_X \mathcal{T} = \mathcal{T}_* X,$$

for any $X \in \mathfrak{X}(M)$. Also, consider $B : TM \times TM \to f^*T\mathbb{R}^m$ to be the tensor given by

$$B(X, Y) = (\tilde{\nabla}_X L)Y = \tilde{\nabla}_X LY - L\nabla_X Y, \tag{2.2}$$

for any $X, Y \in \mathfrak{X}(M)$. The flatness of the ambient space and

$$B(X, Y) = \tilde{\nabla}_X \tilde{\nabla}_Y \mathcal{T} - \tilde{\nabla}_{\nabla_X Y} \mathcal{T}$$

yield that B is symmetric.

The tensor $\beta : TM \times TM \to N_f M$ defined by

$$\beta(X, Y) = (B(X, Y))_{N_f M}$$

is also symmetric.

Let $\mathrm{y} \in \Gamma(Hom(N_f M, TM))$ be given by

$$\langle \mathrm{y}_\eta, X\rangle + \langle \eta, LX\rangle = 0.$$

One defines the tensor $\mathcal{E} : TM \times N_f M \to N_f$ by

$$\mathcal{E}(X, \eta) = h(X, \mathrm{y}_\eta) + (LA_\eta X)_{N_f M}.$$

where $N_f M$ is called the *normal bundle*.

1.3. **The Fundamental Equations.** The pair of tensors associated to an infinitesimal bending satisfy a set of three equations that form the *fundamental system of equations* of an infinitesimal variation. The term fundamental means that they are the integrability condition of the system of differential equations whose solutions yield an infinitesimal bendings.

Proposition 1.3. [3] *The pair $(\beta, \mathcal{E})$ associated to an infinitesimal bending $\mathcal{T}$ satisfies the following system of three equations:*

$$A_{\beta(Y,Z)} X + B_{h(Y,Z)} X = A_{\beta(X,Z)} Y + B_{h(X,Z)} Y, \tag{3.1}$$

$$(\nabla^\perp_X \beta)(Y, Z) - (\nabla^\perp_Y \beta)(X, Z) = \mathcal{E}(Y, h(X, Z)) - \mathcal{E}(X, h(Y, Z)), \tag{3.2}$$

(3.3) $(\nabla^{\perp}_X\beta)(Y,\eta)-(\nabla^{\perp}_Y\mathcal{E})(X,\eta)=\beta(X,A_\eta Y)-\beta(A_\eta X,Y)+h(X,B_\eta Y)-h(B_\eta X,Y)$,

for all $X,Y,Z\in\mathfrak{X}(M)$ *and* $\eta\in\Gamma(N_fM)$.

Moreover, equation (3.2) *is equivalent to*

(3.4) $$(\nabla_X B_\eta)Y-(\nabla_Y B_\eta)X-B_{\nabla^{\perp}_X\eta}Y+B_{\nabla^{\perp}_Y\eta}X=A_{\mathcal{E}(X,\eta)}Y-A_{\mathcal{E}(Y,\eta)}X,$$

for all $X,Y,Z\in\mathfrak{X}(M)$ *and* $\eta\in\Gamma(N_fM)$.

1.4. **The Fundamental Theorem.** The Fundamental Theorem for infinitesimal variation is the following:

Theorem 1.4. [3] *Let* $f:M^n\to\mathbb{R}^m$ *be an isometric immersion of a simply connected Riemannian manifold. Let* $\beta:TM\times TM\to N_fM$ *be a symmetric tensor and let the tensor* $\mathcal{E}:TM\times N_fM\to N_fM$ *satisfy the compatibility condition* (2.1). *If the pair* $(\beta,\mathcal{E})\neq 0$ *satisfies* (3.1), (3.2) *and* (3.3), *then there is a unique infinitesimal bending* $\mathcal{T}$ *of* f *having* $(\beta,\mathcal{E})$ *as associated pair.*

2. Some Alternative Proofs

In the book [3] is mentioned that to prove the fundamental system of equation of an infinitesimal variation there is an alternative classical way, more precisely starting from the fact that, from the relation

(5.1) $$\frac{\partial}{\partial t}\big|_{t=0}g_t=0,$$

where g_t are the metrics induced by the infinitesimal variation $f_t=f+t\mathcal{T}$, one obtains for any $X,Y,Z,W\in\mathfrak{X}(M)$, the relations

(5.2) $$\frac{\partial}{\partial t}\big|_{t=0}\nabla^t_X Y=0$$

and

(5.3) $$\frac{\partial}{\partial t}\big|_{t=0}g_t\big(R^t(X,Y)Z,W\big)=0,$$

for any $X,Y,Z,W\in\mathfrak{X}(M)$

These relations will be used to compute the derivatives with respect to t at $t=0$ of the Gauss, Codazzi and Ricci equation for f_t. We point out that the relation (5.1) is satisfied by the metrics g_t induced by the infinitesimal variation.

We did calculations in this direction, more precisely we obtained (5.2) from (5.1); after that, we have obtained (5.3) from (5.2).

Summarizing, we will prove the following

Problem. *Given an infinitesimal bending* $\mathcal{T}$ *of a submanifold* $f:M^n\to\mathbb{R}^m$, *let* $\mathcal{F}:I\times M^n\to\mathbb{R}^m$ *be a variation of* $f=f_0$ *by immersions* $f_t:M^n\to\mathbb{R}^m$ *with variational vector field* $\mathcal{T}$, *show that the Levi-Civita connection of the metrics* g_t *induced by* f_t *for* $t\in I$ *satisfies*

(5.2) $$\frac{\partial}{\partial t}\big|_{t=0}\nabla^t_X Y=0$$

and

$$\frac{\partial}{\partial t}\big|_{t=0} g_t\big(R^t(X,Y)Z,W\big)=0. \tag{5.3}$$

Proof. Let $(x^1,...,x^n)$ be local coordinates. Then the relation to prove (5.2) can be written equivalently as

$$\frac{\partial}{\partial t}\big|_{t=0}\nabla^t_{\dfrac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j}=0. \tag{5.4}$$

We have

$$\nabla^t_{\dfrac{\partial}{\partial x^i}}\frac{\partial}{\partial x^j}=(\Gamma_t)^k_{ij}\frac{\partial}{\partial x^k},$$

with $(\Gamma_t)^k_{ij}$ the Christoffel symbols.

They are expressed by

$$(\Gamma_t)^k_{ij}=\frac{1}{2}(g_t)^{ks}\big(\frac{\partial(g_t)_{js}}{\partial x^i}+\frac{\partial(g_t)_{is}}{\partial x^j}-\frac{\partial(g_t)_{ij}}{\partial x^s}\big).$$

Then

$$\frac{\partial(\Gamma_t)^k_{ij}}{\partial t}=\frac{1}{2}\frac{\partial(g_t)^{ks}}{\partial t}\big(\frac{\partial(g_t)_{js}}{\partial x^i}+\frac{\partial(g_t)_{is}}{\partial x^j}-\frac{\partial(g_t)_{ij}}{\partial x^s}\big)+\frac{1}{2}(g_t)^{ks}\frac{\partial}{\partial t}\big(\frac{\partial(g_t)_{js}}{\partial x^i}+\frac{\partial(g_t)_{is}}{\partial x^j}-\frac{\partial(g_t)_{ij}}{\partial x^s}\big).$$

But, using (5.1), we have:

$$\frac{\partial}{\partial t}\frac{\partial(g_t)_{ij}}{\partial x^s}=\frac{\partial}{\partial x^s}\frac{\partial(g_t)_{ij}}{\partial t}\to 0\quad(t\to 0)\Rightarrow$$
$$\Rightarrow\lim_{t\to 0}\frac{\partial}{\partial t}\big(\frac{\partial(g_t)_{js}}{\partial x^i}+\frac{\partial(g_t)_{is}}{\partial x^j}-\frac{\partial(g_t)_{ij}}{\partial x^s}\big)=0.$$

We have $\delta^k_j=(g_t)^{ks}(g_t)_{js}$, which implies

$$\frac{\partial(g_t)^{ks}}{\partial t}(g_t)_{js}=-(g_t)^{ks}\frac{\partial(g_t)_{js}}{\partial t}.$$

But $\dfrac{\partial(g_t)_{js}}{\partial t}\to 0(t\to 0)$ and $(g_t)_{js}$ inversable.

So, we obtained

$$\lim_{t\to 0}\frac{\partial(\Gamma_t)^k_{ij}}{\partial t}=0,$$

which implies the relation (5.4), and then we proved (5.2).

Next we will prove, by using (5.2) that

$$\frac{\partial}{\partial t}\big|_{t=0} g_t(R^t(X,Y)Z,W)=0. \tag{5.3}$$

We have

$$\frac{\partial}{\partial t}\big[g_t(R^t(X,Y)Z,W)\big]=\frac{\partial g_t}{\partial t}\big(R^t(X,Y)Z,W\big)+g_t\big(\frac{\partial}{\partial t}R^t(X,Y)Z,W\big).$$

The first term of the right side goes to 0, for $t \to 0$. We will prove that the second term of the right side also goes to 0, for $t \to 0$.

$$g_t(R^t(X,Y)Z,W) = g_t(\nabla^t_X\nabla^t_Y Z - \nabla^t_Y\nabla^t_X Z - \nabla^t_{[X,Y]}Z, W) =$$
$$= -g_t(\nabla^t_Y Z, \nabla^t_X W) + g_t(\nabla^t_X Z, \nabla^t_Y W) - g_t(\nabla^t_{[X,Y]}Z, W).$$

Then

$$g_t(\frac{\partial}{\partial t}R^t(X,Y)Z,W) = -g_t(\frac{\partial}{\partial t}\nabla^t_Y Z, \nabla^t_X W) - g_t(\nabla^t_Y Z, \frac{\partial}{\partial t}\nabla^t_X W) +$$
$$+g_t(\frac{\partial}{\partial t}\nabla^t_X Z, \nabla^t_Y W) + g_t(\nabla^t_X Z, \frac{\partial}{\partial t}\nabla^t_Y W) - g_t(\frac{\partial}{\partial t}\nabla^t_{[X,Y]}Z, W) - g_t(\nabla^t_{[X,Y]}Z, \frac{\partial}{\partial t}W).$$

But, according to (5.2), $\lim_{t\to 0}\frac{\partial}{\partial t}\nabla^t_Y Z = 0$ and then we have similar relation, i.e,

$$\lim_{t\to 0}\frac{\partial}{\partial t}\nabla^t_X W = \lim_{t\to 0}\frac{\partial}{\partial t}\nabla^t_X Z = \lim_{t\to 0}\frac{\partial}{\partial t}\nabla^t_Y W = \lim_{t\to 0}\frac{\partial}{\partial t}\nabla^t_{[X,Y]}Z = 0.$$

It follows that $\lim_{t\to 0}\frac{\partial}{\partial t}g_t\big(R^t(X,Y)Z,W\big) = 0$ and then (5.3) is proven.

Remark 2.1. The relation (5.3) was obtained globally, not by using local coordinates, as in the proof of (5.2).

3. Infinitesimal Bending of a Rectifying Curve

We will focus in this section on low-dimensional case, i.e. the manifold is a curve. We consider its infinitesimal bending and the associated infinitesimal bending field.

Let $c(s)$ be a unit speed curve in $\mathbb{E}^3$, parameterized by arc-length s. A family of curves c_ξ is an *infinitesimal bending of the curve* c if is of the form

$$c_\xi(s) = c(s) + \xi z(s),$$

with ξ =parameter≥ 0, z the *infinitesimal bending field of* c, having certain properties defined next.

One can choose z such that $||z'(s)|| = 1$ and defined in the points of curve c, i.e.

$$z(s) = \alpha_1(s)t(s) + \alpha_2(s)n(s) + \alpha_3(s)b(s),$$

with $\{t, n, b\}$ the Frenet-Serret frame of c.

According to [7] z is an *infinitesimal bending field* of c if and only if

$$\alpha_1'(s) = k(s)\alpha_2(s), \tag{6.1}$$

where $k(s)$ is the curvature of c.

Also, according to [7], if $k_\xi(s), \tau_\xi(s)$ are the curvature and the torsion of $c_\xi(s)$, respectively, the following relations hold:

$$k_\xi = k + \xi[k'\alpha_1 + \alpha_2'' + (k^2 - \tau^2)\alpha_2 - 2\tau\alpha_3' - \alpha_3\tau'], \tag{6.2}$$

$$\tau_\xi = \tau + \xi\{\tau'\alpha_1 + k(\alpha_3' + 2\tau\alpha_2) + [\frac{1}{k}(2\tau\alpha_2' + \tau'\alpha_2 + \alpha_3'' - \tau^2\alpha_3)]'\}. \tag{6.3}$$

On the other hand, rectifying curves form a special class of curves, with nice geometrical properties. More precisely, a unit speed space curve $c(s)$ with $k(s) \neq 0$, for any $s \in I$, is a *rectifying curve* (i.e. its position vector always lies in its rectifying plane) ([2], [5]) if and only if there exist the constants $a, b \in \mathbb{R}, a \neq 0$, such that

$$\frac{\tau(s)}{k(s)} = as + b, \quad s \in I. \tag{6.4}$$

We will consider the following

Problem. *Under which conditions the infinitesimal bending c_ξ of a rectifying curve c is also rectifying?*

We consider some particular cases. The general problem is not solved yet.

Remark 3.1. If s is the arc-length of the curve c and s^* is the arc-length of c_ξ, then

$$s^* = \frac{1}{\sqrt{1+\xi^2}}s. \tag{6.5}$$

We have $c_\xi(s) = c(s) + \xi z(s)$, with $||c'(s)|| = 1$. One can choose $z(s)$, with $||z'(s)|| = 1$. By definition of the bending vector field,

$$\langle c'(s), z'(s)\rangle = 1;$$

then

$$||c_\xi'(s)||^2 = \langle c'(s) + \xi z'(s), c'(s) + \xi z'(s)\rangle = 1 + \xi^2,$$

which implies (6.5).

According to (6.4), it follows that c_ξ is rectifying if and only if is

$$\frac{\tau_\xi(s^*)}{k_\xi(s^*)} = \frac{\tau_\xi(s)}{k_\xi(s)} = As^* + B = A \cdot \frac{1}{\sqrt{1+\xi^2}}s + B, \tag{6.6}$$

$A, B \in \mathbb{R}, A \neq 0$.

We apply (6.2)-(6.6) to give answers to our problem, in the following particular cases. For each case we just state the result, the proofs can be found in [1].

Case 1. $z(s) = \alpha_1(s)t(s) + \alpha_2(s)n(s) + \alpha_3(s)b(s)$, with $\alpha_1(s) = \alpha_1 = \text{constant} \neq 0$, $\alpha_2(s) = 0$, $\alpha_3(s) = \alpha_3 = \text{constant}$, $\alpha_1^2 + \alpha_3^2 \neq 0$, $k(s) = k = \text{constant}$.

Theorem 3.2. [1] *Let $c(s)$ be an unit speed rectifying curve of constant curvature and c_ξ its infinitesimal bending curve, with z the infinitesimal bending vector field along the tangent, i.e. $z(s) = \alpha_1 t(s)$, $\alpha_1 = constant \neq 0$. Then c_ξ is a rectifying curve.*

Case 2. $z(s) = \alpha_1(s)t(s) + \alpha_2(s)n(s) + \alpha_3(s)b(s)$, with $\alpha_1(s) = \text{constant}$, $\alpha_2(s) = 0$, $\alpha_3(s) = \text{constant}$, $\alpha_1^2 + \alpha_3^2 \neq 0$, $\tau(s) = \tau = \text{constant}$.

Theorem 3.3. [1] *Let $c(s)$ be an unit speed rectifying curve of constant torsion and c_ξ its infinitesimal bending curve, with z the infinitesimal bending vector field along the binormal, i.e. $z(s) = \alpha_3 b(s)$, $\alpha_3 = constant \neq 0$. Then c_ξ is a rectifying curve.*

Remark 3.4. i) $k(s)$ and $\tau(s)$ of a rectifying curve $c(s)$ cannot be constant at the same time, because $\dfrac{\tau(s)}{k(s)}$ is a linear function of s for a rectifying curve.

ii) In both previous cases 1 and 2 we considered $\alpha_2 = 0$ in order to simplify relations (6.2) and (6.3).

Some situations when $\alpha_2(s) = \alpha_2 = \text{constant} \neq 0$ will be discussed next.

Case 3. $k(s) = k = \text{constant}$, $\alpha_1(s), \alpha_3(s)$ - linear functions, $\alpha_2(s) = \alpha_2 = \text{constant} \neq 0$. Then c rectifying does not imply c_ξ rectifying.

Case 4. $k(s) = k = \text{constant}$, $\alpha_1(s) -$ linear function, $\alpha_2(s) = \alpha_2 = \text{constant} \neq 0$, $\alpha_3(s) = \alpha_3 = \text{constant}$. Then c rectifying does not imply c_ξ rectifying.

Other possible cases:

- similar with case 3 and case 4, but $\tau(s) = \tau = \text{constant}$; these cases imply $k(s) = k = \text{constant}$, which is not possible for rectifying curves, according to Remark 3.4.
- a more general case $\alpha_2(s) \neq 0$ (not necessarily constant) together with other geometrical restrictions.

Besides the above considerations, some rectifying curves on special surfaces (for example spherical surfaces) can be considered and their infinitesimal bending will be studied.

References

[1] S.-C. Broscateanu, *Infinitesimal bendings of rectifying curves*, submitted.

[2] B.-Y. Chen, *When does the position vector of a space curve always lie in its rectifying plane?*, American Math. Monthly **110** (2003), 147-152.

[3] M. Dajczer and M.I. Jimenez, *Infinitesimal Variations of Submanifolds*, Ensaios Matematicos **35**, Sociedade Brasileira de Matematica, 2021.

[4] M. Dajczer, M.I. Jimenez and Th. Vlachos, *Conformal infinitesimal variations of Euclidean hypersurfaces*, Annali di Matematica Pura ed Applicata **2** (2021), 743-768.

[5] M. Jianu, S. Achimescu, L. Daus, A. Mihai, O.-A. Roman and D. Tudor, *Characterization of rectifying curves by their involutes and evolutes*, Mathematics **9** (2021), 3077.

[6] M.I. Jimenez, *Infinitesimal bendings of complete Euclidean hypersurfaces*, Manuscripta Mathematica **157** (2018), 513-527.

[7] M. S. Najdanovici, *Infinitesimal bending influence on the Willmore energy of curves*, Filomat **29(10)** (2015), 2411-2419.

Interdisciplinary Doctoral School, Faculty of Mathematics and Computer Science, Transilvania University of Braşov, Romania

Email address: stefan.broscateanu@unitbv.ro

Karl Theodor Wilhelm Weierstrass
1815-1897

Father of modern analysis

WEIERSTRASS APPROXIMATION THEOREM REVISITED

Gheorghe Bucur
Faculty of Mathematics and Computer Science
University of Bucharest, Romania
E-mail: gheorghebucur42@gmail.com
Gavriil Păltineanu
Department of Mathematics and Computer Science
Technical University of Civil Engineering of Bucharest, Romania
E-mail: gavriil.paltineanu@gmail.com

Abstract: From the multitude of proofs of Weierstrass approximation theorem of continuous functions with algebraic polynomials, this paper presents three significant types of proofs, namely: Weierstrass proof, that uses singular integrals, Lebesgue's proof that uses the uniform approximation with algebraic polynomials of the function $|x|$ and Bernstein's proof based on probability theory.

By Weierstrass approximation theorems there are commonly understood the two results published in 1885, in his seventies, where he shows that any continuous function on a real compact interval can be uniformly approximated by an algebraic polynomial, and, accordingly, any continuous periodic function of period $T = 2\pi$ can be uniformly approximated by a trigonometric polynomial. In fact, the two theorems are equivalent, in the sense that one implies the other and reciprocally. See for example Remark 1.1.7 from [3].

These results somewhat counterbalance Weierstrass famous example from 1861, of a continuous function on a real interval, which is not differentiable at any point of this interval; this example shows that some continuous functions can be "very non-smooth", while from Weierstrass approximation theorems it follows that a continuous function can be uniformly approximated by "very smooth" functions (i.e., polynomial functions).

It is well known that an entire function (a function that allows a power series expansion) can be uniformly approximated on any interval from the domain of convergence of its power series by an algebraic polynomial (the corresponding Taylor polynomial). However, it is also known that only a restricted class of functions can be expanded in a power series; such a function must be, among other requirements, infinitely differentiable.

Weierstrass theorem widely expands the class of functions that can be uniformly approximated by polynomials from entire to continuous functions.

Weierstrass theorems are so important – both from theoretical and practical point of view – that even since they were published numerous mathematicians proposed various proofs for them. These proofs are extremely instructive, as their basic ideas can be successfully applied in other problems of mathematical analysis.

In his excellent survey[1], Allan Pinkus classifies these proofs in three groups.

The first one is composed of the proofs based on singular integrals; the original proof of Weierstrass, as well as the proofs given by Picard, Fejér, Landau and de la Vallée Poussin belong to this class.

The second group contains the proofs based on the uniform approximation of some particular functions, such as the proofs given by Runge/Phragmén, Lebesgue, Mittag Leffler and Lerch.

The proofs that do not belong to the first two groups form the third one; among these, we mention the proofs given by Bernstein, Volterra and Lerch.

All these proofs appeared up to 1913. Half of a century after, in 1964, H.Kuhn publishes in Archiv der Mathematik an "elementary" proof of the theorem concerning the approximation by algebraic polynomials based on Bernoulli's inequality. This proof is considered by Allan Pinkus as being the simplest and the most elegant ever since.

In what follows, we will present some of the most representative proofs of the Weierstrass approximation theorem of continuous functions by algebraic polynomials.

1. From the first group we present the Weierstrass's proof in the modern version [1].

For any bounded, uniformly continuous function $f:\mathbb{R}\to\mathbb{R}$ and any $h>0$, we define the function $S_h(f):\mathbb{R}\to\mathbb{R}$ by:

$$S_h(f)(x)=\frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}f(u)\cdot e^{-\left(\frac{u-x}{h}\right)^2}du\,. \tag{1}$$

Theorem 1. *Let $f:\mathbb{R}\to\mathbb{R}$ be a bounded, uniformly continuous function. Then $S_h(f)$ converges uniformly to f as $h\downarrow 0$, i.e. $\forall\varepsilon>0,\ \exists h_0>0$ such that for any $0<h<h_0$ we have:*

$$\left|S_h(f)(x)-f(x)\right|<\varepsilon,\ \forall x\in\mathbb{R}\,.$$

Proof. First we will show that:

$$\frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}e^{-\left(\frac{u-x}{h}\right)^2}du=1\,.$$

Indeed, if we change the variable $\frac{u-x}{h}=v$, we get:

$$\frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}e^{-\left(\frac{u-x}{h}\right)^2}du=\frac{h}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}e^{-v^2}=\frac{1}{\sqrt{\pi}}\cdot\sqrt{\pi}=1\,.$$

Therefore we have:

$$f(x)=\frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}f(x)\cdot e^{-\left(\frac{u-x}{h}\right)^2}du\,. \tag{2}$$

[1] Allan Pinkus, *The Weierstass Approximation Theorems*, Survey in Approximation Theory, Vol 1, 2005, pp 1-37

From (1) and (2) it follows:

$$\left|S_h(f)(x)-f(x)\right| \le \frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}\left|f(u)-f(x)\right|\cdot e^{-\left(\frac{u-x}{h}\right)^2}du.$$

Since the function f is bounded and uniformly continuous on $\mathbb{R}$, it turns out, on the one hand, that there is $M>0,$ such that $|f(t)|<M,\ \forall t\in\mathbb{R}$, and on the other hand that: $\forall\varepsilon>0,\ \exists\delta_\varepsilon>0$ with the property $\left|f(x)-f(y)\right|<\frac{\varepsilon}{2}$, for any $x,y\in\mathbb{R}$ such that $|x-y|<\delta_\varepsilon$.

Further we have:

$$\left|S_h(f)(x)-f(x)\right| \le \frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{-\infty}^{\infty}\left|f(u)-f(x)\right|\cdot e^{-\left(\frac{u-x}{h}\right)^2}du \le$$

$$\le \frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{|u-x|\le\delta_\varepsilon} + \frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{|u-x|\ge\delta_\varepsilon} \le$$

$$\le \frac{\varepsilon}{2}+\frac{1}{h\cdot\sqrt{\pi}}\cdot\int_{|u-x|\ge\delta_\varepsilon}\left|f(u)-f(x)\right|\cdot e^{-\left(\frac{u-x}{h}\right)^2}du \le$$

$$\le \frac{\varepsilon}{2}+\frac{2\cdot M}{h\cdot\sqrt{\pi}}\cdot\int_{|u-x|\ge\delta_\varepsilon} e^{-\left(\frac{u-x}{h}\right)^2}du.$$

If we do it again the change of variable $\frac{u-x}{h}=v$, we get:

$$\left|S_h(f)(x)-f(x)\right| \le \frac{\varepsilon}{2}+\frac{2\cdot M}{\sqrt{\pi}}\cdot\int_{|v|\ge\frac{\delta_\varepsilon}{h}} e^{-v^2}dv\ =$$

$$= \frac{\varepsilon}{2}+\frac{2\cdot h\cdot M}{\delta_\varepsilon\sqrt{\pi}}\cdot\int_{|v|\ge\frac{\delta_\varepsilon}{h}}\frac{\delta_\varepsilon}{h}\cdot e^{-v^2}dv \le \frac{\varepsilon}{2}+\frac{2\cdot h\cdot M}{\delta_\varepsilon\sqrt{\pi}}\cdot\int_{|v|\ge\frac{\delta_\varepsilon}{h}}|v|\cdot e^{-v^2}dv \le$$

$$\le \frac{\varepsilon}{2}+\frac{4\cdot h\cdot M}{\delta_\varepsilon\sqrt{\pi}}\cdot\int_0^{\infty} v\cdot e^{-v^2}dv = \frac{\varepsilon}{2}+\frac{2\cdot h\cdot M}{\delta_\varepsilon\sqrt{\pi}}.$$

If we suppose that $0<h<h_0=\frac{\varepsilon\cdot\delta_\varepsilon\cdot\sqrt{\pi}}{4\cdot M}$ then it results:

$$\left|S_h(f)(x)-f(x)\right| \le \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,\ \forall 0<h<h_0. \tag{3}$$

With these preparations we can now prove the Weierstrass approximation theorem of continuous functions with algebraic polynomials, namely:

Theorem 2. *Let* $f:[a,b]\to\mathbb{R}$ *be a continuous function. Then for any* $\varepsilon>0$ *there exists an algebraic polynomial* P *such that:*

$$\left|f(x)-P(x)\right|<\varepsilon,\quad \forall x\in[a,b]\,.$$

We extend the function f to a uniformly continuous and bounded function on $\mathbb{R}$ as follows:

$$\overline{f}(x)=\begin{cases}0 & \text{if } x\in(-\infty,a-1)\\ f(a)\cdot(x-a+1) & \text{if } x\in[a-1,a)\\ f(x) & \text{if } x\in[a,b]\\ f(b)\cdot(-x+b+1) & \text{if } x\in(b,b+1]\\ 0 & \text{if } x\in(b+1,\infty)\end{cases}.$$

Obviously, there exist $R>0$ such that $\overline{f}(x)=0,$ if $|x|>R.$ Let $M>0$ with the property $\left|\overline{f}(x)\right|<M,\ \forall x\in\mathbb{R}$ and let $\varepsilon>0$ arbitrary. From (3) it follows that there exists $h_0>0$ such that:

$$\left|S_{h_0}\overline{f}(x)-\overline{f}(x)\right|<\frac{\varepsilon}{2}$$

where:

$$S_{h_0}\overline{f}(x)=\frac{1}{h_0\cdot\sqrt{\pi}}\cdot\int_{-R}^{R}\overline{f}(u)\cdot e^{-\left(\frac{u-x}{h_0}\right)^2}du\,.$$

For $|x|<R$ and $|u|<R$ we have: $\frac{-2\cdot R}{h_0}<\frac{u-x}{h_0}<\frac{2\cdot R}{h_0}$. Taking into account of the expansion into the Taylor' series of the function e^{-v^2} we deduce that there exists $n_\varepsilon\in\mathbb{N}^*$ such that:

$$\left|\frac{1}{h_0\cdot\sqrt{\pi}}\cdot e^{-\left(\frac{u-x}{h_0}\right)^2}-\frac{1}{h_0\cdot\sqrt{\pi}}\cdot\sum_{k=0}^{n_\varepsilon}\frac{(-1)^k}{k!}\cdot\left(\frac{u-x}{h_0}\right)^{2\cdot k}\right|<\frac{\varepsilon}{4\cdot R\cdot M}\ .$$

Further we have:

$$\left|\frac{1}{h_0\cdot\sqrt{\pi}}\cdot\int_{-R}^{R}f(u)\cdot e^{-\left(\frac{u-x}{h_0}\right)^2}du-\frac{1}{h_0\cdot\sqrt{\pi}}\cdot\int_{-R}^{R}f(u)\cdot\sum_{k=0}^{n_\varepsilon}\frac{(-1)^k}{k!}\cdot\left(\frac{u-x}{h_0}\right)^{2\cdot k}du\right|<$$

$$<\frac{\varepsilon}{4\cdot R\cdot M}\cdot M\cdot 2\cdot R=\frac{\varepsilon}{2}\,.$$

If we denote by:

$$P(x)=\frac{1}{h_0\cdot\sqrt{\pi}}\cdot\int_{-R}^{R}f(u)\cdot\sum_{k=0}^{n_\varepsilon}\frac{(-1)^k}{k!}\cdot\left(\frac{u-x}{h_0}\right)^{2\cdot k}du\,,$$

then P is a polynomial and:

$$\left|S_{h_0}\overline{f}(x)-P(x)\right|<\frac{\varepsilon}{2},\ \forall|x|<R\,.$$

If $x\in[a,b]$ then we have:

$$|f(x)-P(x)|\le\left|\overline{f}(x)-S_{h_0}\overline{f}(x)\right|+\left|S_{h_0}\overline{f}(x)-P(x)\right|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon\,,$$

and thus the theorem is proved.

2. From the second group we present the Lebesgue's proofs.

Let $f:\ [0,1]\to\mathbb{R}$ be a continuous function. First of all, we replace the function f by a spline function g of first order such that:

$$|f(x)-g(x)|<\frac{\varepsilon}{2},\ \forall x\in[0,1]\,.$$

Indeed, by the uniform continuity property of f we may consider a system $(x_i)_{i\le n}$ of real numbers:

$$0=x_0<x_1<\ldots<x_{i-1}<x_i<\ldots<x_n=1$$

such that the oscilation ω_i of the function f on the interval $[x_{i-1},x_i]$ is smaller then $\frac{\varepsilon}{2}$, i.e. $M_i-m_i<\frac{\varepsilon}{2}$, where $M_i=\sup\{f(t);\ t\in[x_{i-1},x_i]\}$ and $m_i=\inf\{f(t);\ t\in[x_{i-1},x_i]\}$.

Then we consider the affine functions $g_i:[0,1]\to\mathbb{R}$, given by:

$$g_i(x)=f(x_{i-1})+\frac{f(x_i)-f(x_{i-1})}{x_i-x_{i-1}}\cdot(x-x_{i-1})$$

and we remark that $g_i(x_{i-1})=f(x_{i-1}),\ g_i(x_i)=f(x_i)$, hence $g_i(x_{i-1}),g_i(x_i)$ belong to the interval $[m_i,M_i]$ and therefore for any $x\in[x_{i-1},x_i]$ we have:

$$g_i(x),f(x)\in[m_i,M_i],\ |g_i(x)-f(x)|<\frac{\varepsilon}{2}\,.$$

Further we denote g the function defined on the interval $[0,1]$ by:

$$g(x)=g_1(x)+\sum_{i=1}^{n}\left[g(x_{i+1})-g(x_i)\right]\cdot h(x-x_i)\,,\tag{4}$$

where h is the Heaviside function on $\mathbb{R}$:

$$h(t)=\begin{cases}1,\ \text{if}\ t\ge 0\\ 0,\ \text{if}\ t\le 0\end{cases}.$$

It is easy to verify that for any $i\in\{1,2,...,n\}$ the function g_i is the restriction of the function g to $[x_{i-1},x_i]$ and therefore

$$|f(x)-g(x)|<\frac{\varepsilon}{2},\ \forall x\in[0,1]\,.$$

Let's note that the spline function given by (4) can also be written in the form:

$$g(x) = a \cdot x + b + \sum_{i=1}^{n-1} c_i (x - x_i)_+^1 , \tag{5}$$

where:

$$x_+^1 = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

and $a \cdot x + b = g_1(x)$. Indeed, since $g_{i+1}(x) - g_i(x)$ is a polynomial of degree 1 that vanishes at x_i, it is necessarily of de form $c_i \cdot (x - x_i)$. On the other hand, since $2 \cdot x_+^1 = |x| + x$, the form (5) may be rewritten as:

$$g(x) = A \cdot x + B + \sum_{i=1}^{n-1} C_i |x - x_i| . \tag{6}$$

If we will show that the function $|x|$ can be uniformly approximated on the interval $[-1,1]$ by a polynomial, then, it will turn out that g can be approximated by a polynomial, that is Weierstrass theorem. Indeed, let's assume that the function $|x|$ can be uniformly approximated by polynomials on the interval $[-1,1]$, i.e. for any $\eta > 0$ there exists a polynomial p satisfying $\big|\, |x| - p(x) \,\big| < \eta,\ \forall x \in [-1,1]$. Then

$$\big|\, |x - x_i| - p(x - x_i) \,\big| < \eta,\ \forall x \in [0,1] \text{ (since } 0 \leq x_i \leq 1).$$

Further we have:

$$\left| g(x) - \left[Ax + B + \sum_{i=1}^{n-1} C_i p(x - x_i) \right] \right| = \left| \sum_{i=1}^{n-1} C_i \left[|x - x_i| - p(x - x_i) \right] \right| \leq$$

$$\leq \sum_{i=1}^{n-1} C_i \big|\, |x - x_i| - p(x - x_i) \big| .$$

For $\eta = \dfrac{\varepsilon}{2 \cdot (n-1) \cdot \sum_{i=1}^{n-1} C_i}$ it results:

$$\left| g(x) - \left[Ax + B + \sum_{i=1}^{n-1} C_i p(x - x_i) \right] \right| \leq \frac{\varepsilon}{2} .$$

Finally we have:

$$\left| f(x) - \left(A \cdot x + B + \sum_{i=1}^{n-1} C_i p(x - x_i) \right) \right| \leq$$

$$\leq |f(x) - g(x)| + \left| g(x) - \left(A \cdot x + B + \sum_{i=1}^{n-1} C_i p(x - x_i) \right) \right| <$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,\ \forall x \in [0,1] .$$

It remains to show that the function $|x|$ can be uniformly approximated by polynomials on the interval $[-1,1]$. For this, H. Lebesgue proceeded as follows:

$$|x|=\sqrt{1-(1-x^2)}=1+\sum_{n=1}^{\infty}(-1)^n\frac{(2\cdot n-3)!!}{2^n\cdot n!}\cdot(1-x^2)^n\ .$$

If we denote by $z=1-x^2$ then we have:

$$|x|=\sqrt{1-z}=1+\sum_{n=1}^{\infty}(-1)^n\frac{(2\cdot n-3)!!}{2^n\cdot n!}\cdot z^n,\ \ |z|<1.$$

On the other hand, we note that the series with positive terms $\sum_{n=1}^{\infty}\frac{(2\cdot n-3)!!}{2^n\cdot n!}$ is convergent as it results from Raabe's test. Indeed, we have:

$$\lim_{n\to\infty}n\cdot\left(\frac{u_n}{u_{n+1}}-1\right)=n\cdot\left(\frac{2\cdot(n+1)}{2\cdot n-1}-1\right)=\frac{3}{2}>1\ .$$

As:

$$\left|(-1)^n\frac{(2\cdot n-3)!!}{2^n\cdot n!}\cdot z^n\right|\le\frac{(2\cdot n-3)!!}{2^n\cdot n!},\ \forall z\in[-1,1],$$

we deduce that series $1+\sum_{n=1}^{\infty}(-1)^n\frac{(2\cdot n-3)!!}{2^n\cdot n!}\cdot z^n$ is uniformly convergent on the interval $[-1,1]$.

On the other hand, $-1\le 1-x^2\le 1\Rightarrow -\sqrt{2}\le x\le\sqrt{2}$.

Obvious, the sequence of polynomial:

$$P_n(x)=1+\sum_{k=1}^{n}(-1)^k\frac{(2\cdot n-3)!!}{2^n\cdot n!}\cdot(1-x^2)^k$$

is uniformly convergent to $|x|$ on $[-1,1]$.

An ingenious argument to obtain a sequence of polynomials which uniformly approximate the function $|x|$ was presents by N. Bourbaki [1949, pp. 55].

More precisely we have:

Theorem 1. *For any* $\varepsilon>0$ *there is a polynomial* $p:\mathbb{R}\to\mathbb{R}$ *such that* $p(0)=0$ *and*

$$\big|\ |x|-p(x)\ \big|<\varepsilon,\ \textit{for}\ |x|\le 1.$$

Proof. We define a recursive sequence of polynomials as follows:

$$p_{n+1}(x)=p_n(x)+\frac{1}{2}\Big[x^2-p_n^2(x)\Big],\ \ x\in\mathbb{R},\ \ n\in\mathbb{N}^*$$

$$p_0\equiv 0\ .$$

Obviously we have $p_n(0)=0,\ \forall n\in\mathbb{N}$.

Further on, we will show, by mathematical induction, that:

$$0\le p_n(x)\le|x|,\ \forall x\in[-1,1]. \tag{7}$$

Indeed, for $n=0$ the statement is obvious. We assume the true statement for n and prove it for $n+1$. So we assume that $p_n(x)\ge 0$ and $x^2\ge p_n^2(x),\ \forall x\in[-1,1]$, from where it results on one side that $p_{n+1}(x)\ge 0$ and on the other side that $p_{n+1}(x)\le|x|\ \ \forall x\in[-1,1]$.

This last inequality results from the relationship:

$$|x|-p_{n+1}(x)=\Big[|x|-p_n(x)\Big]\cdot\left\{1-\frac{1}{2}\Big[|x|+p_n(x)\Big]\right\} \tag{8}$$

and from the inequalities $|x|+p_n(x)\le 2\cdot|x|\le 2$ and $1-\frac{1}{2}\left[|x|+p_n(x)\right]\ge 1-1=0$.

Next, we will show by mathematical induction that:

$$|x|-p_n(x)\le|x|\cdot\left(1-\frac{|x|}{2}\right)^n,\ \forall x\in[-1,1],\ \forall n\in\mathbb{N} \tag{9}$$

Indeed, for $n=0$ the inequality is true. Assuming that the inequality is true for n, from (8) it follows:

$$|x|-p_{n+1}(x)\le\left[|x|-p_n(x)\right]\cdot\left(1-\frac{|x|}{2}\right)\le|x|\cdot\left(1-\frac{|x|}{2}\right)^{n+1}.$$

For any $0<\varepsilon\le 1$, from (8) we deduce that there exists $n_\varepsilon\in\mathbb{N}^*$ such that:

$$0\le|x|-p_n(x)\le|x|\le\varepsilon,\ \forall x\in[-\varepsilon,\varepsilon],$$

$$0\le|x|-p_n(x)\le\left(1-\frac{\varepsilon}{2}\right)^n\le\varepsilon,\ \text{if}\ \varepsilon\le|x|\le 1,\ \ \forall n\ge n_\varepsilon$$

Therefore we have:

$$0\le|x|-p_n(x)\le\varepsilon,\ \forall x\in[-1,1],\ \forall n\ge n_\varepsilon\ .$$

Remark 1. For any $M>0$ and any $\varepsilon>0$ there is a polynomial $p:\mathbb{R}\to\mathbb{R}$ such that $p(0)=0$ and $\left||x|-p(x)\right|<\varepsilon,\ \text{for}\,|x|\le M$.

The statement follows from the inequality:

$$\left|M\cdot\left|\frac{x}{M}\right|-p(x)\right|<M\cdot\varepsilon,\ \text{for}\ |x|\le M$$

Corollary 1. (Lebesgue). *Let X be a compact Hausdorff space and let $\mathcal{A}\subset C(X,\mathbb{R})$ be a closed subalgebra .Then $\mathcal{A}$ is a lattice.*

Proof. It is enough to show that $f\in\mathcal{A}$ implies $|f|\in\mathcal{A}$. If we denote by $M=\|f\|$, then from Remark 1.1 it follows that there exists a polynomial p such that:

$$\left|\,|f(x)|-p[f(x)]\,\right|<\varepsilon,\ \forall x\in X.$$

Since $p(f)\in\mathcal{A}$, it results that $|f|\in\bar{\mathcal{A}}=\mathcal{A}$.

3. From the third group we present the Bernstein's proof based on probability theory (See [2]).

A random variable X is a Bernoulli random variable with parameter $p\in[0,1]$ if it has the following distribution:

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

More precisely, this means the following: let $0\le p\le 1$ be an arbitrary number, the probability space $\{1,0\}$ with the $\sigma-$algebra $\{\phi,\{1\},\{0\},\{1,0\}\}$ and the probability q_p defined as such:

$$q_p\left(\{1\}\right)=p,\ \ q_p\left(\{0\}\right)=1-p,\ \ q_p\left(\{0,1\}\right)=1,\ \ q_p\left(\{\varphi\}\right)=0.$$

On the space $\Omega=\{1,0\}^{\mathbb{N}}$ we consider the $\sigma-$algebra product $\mathcal{B}$ and the product measure denoted by P_p on this $\sigma-$algebra. An element $\omega\in\Omega$ may be regard as a sequence $\omega_1,\omega_2,...,\omega_n,...$where $\omega_i\in\{0,1\}$ for any $i\in\mathbb{N}$. If we denote by X_n the $n-$th projection of Ω on $\{1,0\}$ then X_n is a random variable.

We have:

$$P_p\left(X_n=1\right)=p,\ \ P_p\left(X_n=0\right)=1-p,\ \ \forall n\in\mathbb{N}$$

and by definition of a product measure the sequence $\left(X_n\right)_n$ is independent.

Therefore $\left(X_n\right)_n$ is an i.i.d. sequence.

We remark that the average of any X_n is p and the variance (dispersion) is $\sigma^2(X_n)=p\cdot\left(1-p\right)$. To be self-contained we prove the inequality:

$$P_p\left(\left|\frac{S_n}{n}-p\right|\geq\varepsilon\right)\leq\frac{p\cdot\left(1-p\right)}{n\cdot\varepsilon^2}\leq\frac{1}{n\cdot\varepsilon^2},\ \ \forall n\in\mathbb{N},\tag{1}$$

where $S_n=X_1+X_2+....+X_n$.

Indeed, we have, using Chebyshev's inequality:

$$P_p\left[\left(\frac{S_n}{n}-p\right)^2\geq\varepsilon^2\right]\leq\frac{1}{\varepsilon^2}\cdot\int\left(\left(\frac{S_n}{n}-p\right)^2dP_p\right)=\frac{1}{\varepsilon^2}\cdot\int\left(\frac{\sum_{i=1}^{n}\left(X_i-p\right)}{n}\right)^2dP_p=$$

$$=\frac{1}{n\cdot\varepsilon^2}\cdot\int\frac{\left(\sum_{i=1}^{n}\left(X_i-p\right)\right)^2}{n}dP_p=\frac{\sigma^2(X_1)}{n\cdot\varepsilon^2}=\frac{p\cdot(1-p)}{n\cdot\varepsilon^2}\leq\frac{1}{n\cdot\varepsilon^2}.$$

We fix $n\in\mathbb{N}^*$ and for any $k=\overline{0,n}$ we denote by $\mathcal{F}_k$ the set of all subsets F of the set $\{1,2,...,n\}$ each of them containing k elements. Certainly the cardinal of $\mathcal{F}_k$ is C_n^k. If for any $\mathcal{F}\in\mathcal{F}_k$ we denote by:

$$\Omega_F=\left(\bigcap_{i\in F}\left(X_i=1\right)\cap\bigcap_{j\notin F}\left(X_j=0\right)\right)$$

then we have:

$$\left(S_n=k\right)=\bigcup_{F\in\mathcal{F}_k}\Omega_F,\ \ P_p\left(\Omega_p\right)=p^k\cdot(1-p)^{n-k}$$

and since the events $\left(\Omega\right)_{F\in\mathcal{F}_k}$ are pairewise disjoint we get:

$$P_p\left(S_n=k\right)=C_n^k\cdot p^k\cdot\left(1-p\right)^{n-k}=P_p\left(\frac{S_n}{n}=\frac{k}{n}\right),\ \ k\in\{0,1,...,n\}.\tag{2}$$

Let $f:[0,1]\to\mathbb{R}$ be a continuous function. The random variable $f\circ\left(\frac{S_n}{n}\right)$ has the distribution:

$$\begin{pmatrix} f\left(\frac{k}{n}\right) \\ C_n^k\cdot p^k\cdot(1-p)^{n-k}\end{pmatrix}_{k=0,1,\ldots,n}$$

and the average:

$$M\left[f\left(\frac{S_n}{n}\right)\right]=\int f\left(\frac{S_n}{n}\right)dP_p=\sum_{k=0}^{n}C_n^k\cdot p^k\cdot(1-p)^{n-k}\cdot f\left(\frac{k}{n}\right)$$

that is exactly the Bernstein polynomial attached to the function f, which we denote by $B_n^p(f)$. Now we can prove Weierstrass theorem i.e.:

For any continuous function $f:[0,1]\to\mathbb{R}$ and any $\varepsilon>0$ there exists a polynomial P_ε such that:

$$|P_\varepsilon(x)-f(x)|<\varepsilon,\quad \forall x\in[0,1].$$

Proof. Since the function f is uniformly continuous it results that, for any $\varepsilon>0$ there exists $\delta_\varepsilon>0$ such that $|f(x)-f(y)|<\varepsilon$ if $|x-y|<\delta_\varepsilon$. Further we denote by $\|f\|$ the uniform norm of f and if we use inequality (1) we get:

$$B_n^p(f)-f(p)=\int f\left(\frac{S_n}{n}\right)dP_p-f(p)=\int\left(f\left(\frac{S_n}{n}\right)-f(p)\right)dP_p.$$

$$\left|B_n^p(f)-f(p)\right|\le\int\left|f\left(\frac{S_n}{n}\right)-f(p)\right|dP_p=$$

$$=\int_{\left|\frac{S_n}{n}-p\right|<\delta_\varepsilon}\left|f\left(\frac{S_n}{n}\right)-f(p)\right|dP_p+\int_{\left|\frac{S_n}{n}-p\right|\ge\delta_\varepsilon}\left|f\left(\frac{S_n}{n}\right)-f(p)\right|dP_p\le$$

$$\le\frac{\varepsilon}{2}+2\cdot\|f\|\cdot P_p\left(\left|\frac{S_n}{n}-p\right|\ge\delta_\varepsilon\right)\le\frac{\varepsilon}{2}+2\cdot\|f\|\cdot\frac{1}{n\cdot\delta_\varepsilon^2}.$$

Taking now $n\in\mathbb{N}$ such that $2\cdot\|f\|\cdot\frac{1}{n\cdot\delta_\in^2}<\frac{\varepsilon}{2}$ then the polynomial

$$P_\varepsilon(p)=B_n^p(f),\ \forall p\in[0,1]$$

satisfies the required condition. Indeed, the natural number n depends on ε and δ_ε in the above inequality and it does not depend on p in the inequality (1).

References

[1] Anton, R.: Shep-internet

[2] Bass, R.F.: *Real analysis for graduate students*, Version 4.3, July, 2022.

[3] Bucur, I., Paltineanu, G.: *Topics in uniform Approximation of continuous functions,* Birkhäuser, Frontiers in Mathematics, 2020.

[4] Pinkus, A.: *The Weierstass Approximation Theorems*, Survey in Approximation Theory, Vol.1, pp. 1-37, 2005.

[5] Výborný, R.: *The Weierstrass theorem on polynomial approximation,* Mathematica Bohemica, 130 (2005), No 2, pp. 161-165, 2004.

[6] Weierstrass, K.: *Uber die analytische Darstellbarkeit sogennanter willkurlicher Functiononen reller Argumente*, Sitzungsherichre der Acad. Berlin, pp. 633-639, pp. 789-805, 1885.

MULTIPLICATIVE CHARACTERISTIC FUNCTIONS CALCULUS AND APPLICATIONS

Ileana Bucur
Department of Mathematics and Computer Science
Technical University of Civil Engineering of Bucharest, Romania
E-mail: bucurileana@yahoo.com

Abstract. Very often, in Mathematics, the object we are dealing with is decomposed into other objects much more simpler to be understandable. In this paper we introduce the concept of "multiplicative characteristic function" on arbitrary set X and we give a general representation theorem of any function $f: X \to [0,1]$ as a uniform limit of finite products of such type of functions.
This result may be fruitfully used in measure theory, distribution theory and so on. Some consequences in uniform approximation of continuous functions are present here.
Mathematical Subject Classification: Primary 41A10; Sec 46J10.
Key words: multiplicative characteristic functions, Uryson family, $V-$family.

Introduction

If X is an arbitrary set and $A \subset X$, then for any $\alpha \in (0,1)$ we define the function α_A given by:

$$\alpha_A(x) = \begin{cases} \alpha & \text{if } x \in A \\ 1 & \text{if } x \in X \setminus A \end{cases}$$

and we call it an $\alpha-$characteristic function of A.

In section 1, Theorems 1.1 and 1.2, we show that any $f: X \to [0,1]$ is uniformly approximated by functions of the form:

$$\prod_{1 \le i \le n} \alpha_{[f \le \alpha^i]}.$$

In section 2, where X is a topological normal space and $C_1(X)$ is the set of all continuous functions on X with values in $[0,1]$ we consider a family $\mathcal{U} \subset C_1(X)$ such that for any $\alpha, \delta \in (0,1)$ and any closed, disjoint subsets F', F'' of X there exists $\varphi \in \mathcal{U}$ such that:

$$\alpha - \delta \le \varphi \text{ on } X,\ \varphi \le \alpha + \delta \text{ on } F',\ 1 - \delta \le \varphi \text{ on } F''$$

and we call it *a **multiplicative Uryson family.***

In Theorem 2.1 it is shown that any function $f \in C_1(X)$ is uniformly approximated by finite products of elements of $\mathcal{U}$. Applications to Von Neumann density theorem are presented in section 3.

1. Multiplicative characteristic functions Calculus

In this section we introduce a type of simply function on arbitrary set X with values in the interval $I=[0,1]$ of $\mathbb{R}$ such that any function $f: X \to [0,1]$ may be uniformly approximated by finite products whose factors are such simply functions.

For beginning we fix a real number $\alpha \in (0,1)$. If A is an arbitrary subset of X we denote by α_A the function on X defined by:

$$\alpha_A(x) = \begin{cases} \alpha & \text{if } x \in A \\ 1 & \text{if } x \in X \setminus A \end{cases}$$

and we call it the $\alpha-$characteristic function of A. We have:

$$\alpha \le \alpha_A \le 1 \text{ and } \alpha_A \le \alpha_B \text{ if } A \supset B .$$

The following trivial assertions allow the uniform approximation of any f as above by means of products of $\alpha-$characteristic functions.

Lemma 1.1. *If $A_1, A_2, ..., A_n$ is a finite system of subsets in X and $\alpha \in (0,1)$ then we have:*

$$\left(\prod_{i \le n} \alpha_{A_i}\right)(x) = \alpha^{card\{j, j \le n;\, x \in A_j\}}$$

and, particularly if $A_1 \supset A_2 \supset ... \supset A_n$ then:

$$\left(\prod_{i \le n} \alpha_{A_i}\right)(x) = \alpha^k \text{ if } x \in A_k \setminus A_{k+1},\ k = 0,1,...,n,$$

with the convention $A_0 = X,\ A_{n+1} = \phi$.

Lemma 1.2. *If $\alpha \in (0,1)$ and $(A_i)_{1 \le i \le n}$ is a decreasing finite system of subsets of X we will denote by:*

$$\varphi = \prod_{1 \le i \le n} \alpha_{A_i}, \quad \varphi' = \prod_{1 \le i \le n-1} \alpha_{A_i}, \quad \varphi'' = \prod_{2 \le i \le n} \alpha_{A_i} .$$

With this notations we have:

a) $\varphi = \alpha^n$ on A_n, $\varphi = 1$ on $X \setminus A_1$, $\alpha^n \le \varphi \le 1$ on X

b) $\alpha^n \le \varphi \le \varphi' \le \varphi'' \le 1,\ \varphi'' - \varphi \le 1 - \alpha$

c) If for any $i = 1,2,...,n-1$, we consider an arbitrary function ω_i on X such that $\alpha_{A_i} \le \omega_i \le \alpha_{A_{i+1}}$ and we denote by: $\tilde{\varphi} = \prod_{1 \le i \le n-1} \omega_i$ then we have:

$$\varphi' \le \tilde{\varphi} \le \varphi" .$$

Proof. The assertions *a*) and *c*) follow directly from the definition of $\alpha-$characteristic functions as well as the first part of assertion *b*) taking into account that we have:

$$\alpha \le \alpha_{A_i} \le \alpha_{A_{i+1}} \le 1, \quad \forall i = 1,2,...,n-1.$$

For the second part of *b*) we remark that we have:

$$\varphi'' - \varphi = \varphi'' \cdot (1 - \alpha_{A_i}) \le 1 - \alpha_{A_i} \le 1 - \alpha.$$

Notations. If $f:X\to\mathbb{R}$ is an arbitrary function and α,β are real numbers such that $\alpha<\beta,$ as usually we denote:

$$[f\geq\alpha]=f^{-1}\big([\alpha,\infty)\big)=\{x\in X;\,f(x)\geq\alpha\}$$
$$[f<\alpha]=f^{-1}\big((-\infty,\alpha)\big)=\{x\in X;\,f(x)<\alpha\}$$
$$[\alpha\leq f<\beta]=\{x\in X;\ \alpha\leq f(x)<\beta\}\ \text{and so on}.$$

For any function $f:X\to\mathbb{R}$, any $\alpha\in(0,1)$ and any $n\in\mathbb{N}^*$, we shall denote by $f_\alpha,\ f'_\alpha,\ f''_\alpha$ the functions associated to f defined by:

$$f_\alpha=\prod_{1\leq i\leq n}\alpha_{[f\leq\alpha^i]},\ f'_\alpha=\prod_{1\leq i\leq n-1}\alpha_{[f\leq\alpha^i]}\ ,\ f''_\alpha=\prod_{2\leq i\leq n}\alpha_{[f\leq\alpha^i]}\ .$$

Theorem 1.1. *Let $\alpha\in\mathbb{R},\ 0<\alpha<1,$ let $n=n_\alpha$ be a natural number such that $\alpha^{n-1}<1-\alpha$ and for any $f:X\to[0,1]$ let $f_\alpha,f'_\alpha,f''_\alpha$ the functions associated to f as above. Then we have:*

$$f\leq f_\alpha\leq f'_\alpha\leq f''_\alpha;\ f''_\alpha-f\leq 2\cdot(1-\alpha).$$

Proof. First of all, we remark that the system $(A_i)_{i\leq n}$ of subsets of X given by $A_i=[f\leq\alpha^i]$ is decreasing and using Lemma 1.2 we get:

$$f_\alpha\leq f'_\alpha\leq f''_\alpha;\ f''_\alpha-f_\alpha\leq 1-\alpha\ .$$

Now we show that $f\leq f_\alpha$ and $f_\alpha-f\leq 1-\alpha$. Indeed, if we look to the following partition of X:

$$X=(X\setminus A_1)\cup(A_1\setminus A_2)\cup\ldots\cup(A_{n-1}\setminus A_n)\cup A_n,$$

we have:

1) $x\in X\setminus A_1\Rightarrow x\in X\setminus A_i,\ i=1,2,\ldots,n\ \Rightarrow f_\alpha(x)=1$ and $\alpha<f(x)\leq 1$ and therefore
$$f(x)\leq f_\alpha(x),$$
$$f_\alpha(x)-f(x)=1-f(x)\leq 1-\alpha$$
2) $x\in A_k\setminus A_{k+1}\Rightarrow f_\alpha(x)=\alpha^k$ (see Lemma 1.1), $\alpha^{k+1}<f(x)\leq\alpha^k$ and therefore:
$f(x)\leq f_\alpha(x),$
$f_\alpha(x)-f(x)=\alpha^k-f(x)\leq 1-f(x)\leq\alpha^k-\alpha^{k+1}=\alpha^k\cdot(1-\alpha)\leq 1-\alpha,\ \forall k\leq n-1$
3) $x\in A_n\Rightarrow x\in A_i,\ i=1,2,\ldots,n\ \Rightarrow f_\alpha(x)=\alpha^n,\ f(x)\leq\alpha^n$ and therefore:
$$f(x)\leq f_\alpha(x),$$
$$f_\alpha(x)-f(x)=\alpha^n-f(x)\leq\alpha^n<1-\alpha\ .$$
From the above considerations we deduce:

$$f\leq f_\alpha,\ \ f_\alpha-f\leq 1-\alpha,\ \ f''_\alpha-f=(f''_\alpha-f_\alpha)+(f_\alpha-f)\leq 2\cdot(1-\alpha).$$

Theorem 1.2. *Let* $f: X \to [0,1]$ *be an arbitrary function, let* $(\alpha_n)_n$ *be a sequence in* $(0,1)$ *such that* $\lim_{n\to\infty} \alpha_n = 1$. *If for any* $k \in \mathbb{N}$ *we denote by* $f_{\alpha_k}, f'_{\alpha_k}, f''_{\alpha_k}$ *the functions associated to* f *and to the number* $\alpha = \alpha_k$, *as in Theorem 1.1 then we have:*

$$f \le f_{\alpha_k} \le f'_{\alpha_k} \le f''_{\alpha_k} \quad \forall k \in \mathbb{N}$$

and the sequences $\left(f_{\alpha_k}\right)_k, \left(f'_{\alpha_k}\right)_k, \left(f''_{\alpha_k}\right)_k$ *are uniformly convergent to* f on X.

The proof directly comes from Theorem 1.1 using the inequalities:

$$f''_\alpha - f \le 2 \cdot (1 - \alpha_k),\ k \in \mathbb{N}.$$

2. Uniform approximation in topological spaces

In this section, X will be a topological normal space, $C_1(X) = C_1$ will be the set of all continuous functions on X with value in $I = [0,1] \subset \mathbb{R}$.

Definition 2.1. *In [1] we introduced so called* ***Uryson family*** *in* $C_1(X)$ *as being a subset* $\mathcal{U}$ *of* C_1 *such that for any* $\varepsilon \in (0,1)$ *and any two disjoint closed subsets* F', F'' *of* X *there exists* $u \in \mathcal{U}$ *such that:*

$$u(x) \le \varepsilon,\ \forall x \in F' \text{ and } u(x) \ge 1 - \varepsilon,\ \forall x \in F''$$

and we have shown that:

$$\overline{co}(\mathcal{U}) = C_1$$

where $\overline{co}(\mathcal{U})$ means the closed, convex covering of the set $\mathcal{U}$ (closure with respect to the uniform convergence). The famous Stone-Weierstrass Theorem on algebras of continuous functions as well other density results come directly from the above assertions [1], [2].

Now we introduce a new class of continuous functions and we prove a density result under "multiplicative form".

Definition 2.2. *A subset* $\mathcal{U} \subset C_1$ *is called* $m-$***Uryson family*** *(multiplicative Uryson family) if for any* $\alpha, \delta \in (0,1)$ *and any two disjoint closed subsets* F', F'' *of* X *there exists* $u \in \mathcal{U}$ *such that:*

$$\alpha - \delta \le u(x), \forall x \in X;\ \ |u(x) - \alpha| \le \delta, \forall x \in F';\ \ |1 - u(x)| \le \delta, \forall x \in F''$$

or equivalently:

$$\alpha - \delta \le u(x), \forall x \in X;\ \ u(x) \le \alpha + \delta, \forall x \in F';\ \ 1 - \delta \le u(x) \forall x \in F''\ .$$

Remark 2.1. *Any multiplicative Uryson family is an Uryson family.*

Theorem 2.1. *If* $\mathcal{U}$ *is a multiplicative Uryson family on* X *and* $\mathcal{U}$ *is stable with respect to the product operation i.e.* $u, v \in \mathcal{U} \Rightarrow u \cdot v \in \mathcal{U}$, *then* $\mathcal{U}$ *is uniformly dense in* $C_1(X)$.

Proof. Let $\varepsilon \in (0,1)$ be arbitrarily small, let $\alpha = 1 - \varepsilon$ and $n \in \mathbb{N}$ be such that $\alpha^{n-1} < \varepsilon$. If $f \in C_1(X)$ we denote as in Theorem 1.1 $A_i = [f \le \alpha^i]$ and we consider the functions $f_{\alpha_k}, f'_{\alpha_k}, f''_{\alpha_k}$ associated to f:

$$f_\alpha = \prod_{1 \le i \le n} \alpha_{[f \le \alpha^i]},\ f'_\alpha = \prod_{1 \le i \le n-1} \alpha_{[f \le \alpha^i]}\ ,\ f''_\alpha = \prod_{2 \le i \le n} \alpha_{[f \le \alpha^i]}\ .$$

By Theorem 1.1 we have:

$$f \le f_\alpha \le f'_\alpha \le f''_\alpha;\ f''_\alpha - f \le 2\cdot\varepsilon \text{ or equivalently } f''_\alpha \le f + 2\cdot\varepsilon .$$

For any $0<\delta<\alpha$ we denote $f'_{\alpha,\delta}, f''_{\alpha,\delta}$ the positive functions:

$$f'_{\alpha,\delta} = \prod_{1\le i\le n-1}\left(\alpha_{A_i} - \delta\right);\quad f''_{\alpha,\delta} = \prod_{2\le i\le n}\left(\alpha_{A_i} + \delta\right).$$

Obviously the functions $f'_{\alpha,\delta}, f''_{\alpha,\delta}$ tend uniformly on X to the functions f'_α, f''_α when δ tends to 0. Hence for a sufficiently small δ we have:

$$f-\varepsilon \le f'_\alpha - \varepsilon \le f'_{\alpha,\delta} \le f'_\alpha \le f''_\alpha \le f''_{\alpha,\delta} \le f''_\alpha + \varepsilon \le f + 3\cdot\varepsilon . \tag{1}$$

Now, for any $i=1,2,...,n-1$ we consider the closed, disjoint subsets F', F'' given by:

$$F' = A_{i+1} = [f \le \alpha^{i+1}],\ F'' = \overline{[f > \alpha^i]} \subset [f \ge \alpha^i].$$

By hypotheses $\mathcal{U}$ is a multiplicative Uryson family, so we choose $u_i \in \mathcal{U}$ such that:

$$\alpha - \delta \le u_i \text{ on } X,\ u_i \le \alpha + \delta \text{ on } A_{i+1},\ u_i \ge 1-\delta \text{ on } \overline{X \setminus A_i} = \overline{[f > \alpha^i]}.$$

It is almost obvious that we have:

$$\alpha_{A_i} - \delta \le u_i \le \alpha_{A_{i+1}} + \delta,\ i = 1,2,...,n-1$$

and therefore:

$$f'_{\alpha,\delta} = \prod_{1\le i\le n-1}(\alpha_{A_i} - \delta) \le \prod_{1\le i\le n-1} u_i \le \prod_{1\le i\le n-1}(\alpha_{A_{i+1}} + \delta) = f''_{\alpha,\delta}\ . \tag{2}$$

From the relations (1) and (2) we get:

$$f - 3\cdot\varepsilon < f - \varepsilon \le f'_{\alpha,\delta} \le \prod_{1\le i\le n-1} u_i \le f''_{\alpha,\delta} \le f + 3\cdot\varepsilon;\quad \left| f - \prod_{1\le i\le n-1} u_i \right| \le 3\cdot\varepsilon\ .$$

Definition 2.3. *A subset $\mathcal{U}$ of $C_1(X)$ is called a $V-$family if it has the properties:*

1) $f\cdot g \in \mathcal{U},\ \forall f,g \in \mathcal{U}$
2) $1-f \in \mathcal{U},\ \forall f \in \mathcal{U}$

Remark 2.1. It is not difficult to show that if $c_0 \in (0,1)$ then the subset:

$$M = \left\{\left(1-c_0^n\right)^m;\ n,m \in \mathbb{N}^*\right\}$$

is dense in $[0,1]$.

Similarly, if we consider an arbitrary function $f: X \to (\alpha,\beta)$, where $0<\alpha\le f\le\beta<1$, then any function constant c, $c\in[0,1]$ is uniformly approximated by functions of de form:

$$\left(1-f^n\right)^m;\ n,m \in \mathbb{N}^*.$$

Lemma 2.1. *Let $\mathcal{V} \subset C_1(X)$ be a $V-$family which contains the constant functions $c\in[0,1]$ and which is also an Uryson family. Then $\mathcal{V}$ is a multiplicative Uryson family.*

Proof. Let F', F'' be two closed, disjoint subsets of X and for any $\varepsilon\in(0,1)$ let $t\in\mathcal{V}$ be such that:

$$1-\varepsilon \le t(x),\ \forall x \in F'';$$
$$t(x) \le \varepsilon,\ \forall x \in F';$$
$$(1-\varepsilon)^2 \le (1-\varepsilon)\cdot t(x),\ \forall x \in F'';$$
$$(1-\varepsilon)\cdot t(x) \le 1-\varepsilon,\ \forall x \in X\ .$$

For an arbitrary number $\beta \in (0,1)$ the function $s = \beta \cdot (1-\varepsilon) \cdot t \in \mathcal{V}$ satisfies the relations:

$\beta \cdot (1-\varepsilon)^2 \le s(x),\ \forall x \in F'';$

$s(x) \le \beta \cdot (1-\varepsilon),\ \forall x \in X;$

$s(x) \le \beta \cdot (1-\varepsilon) \cdot \varepsilon,\ \forall x \in F';$

$s(x) \le \varepsilon,\ \forall x \in F';$

$1-\beta \cdot (1-\varepsilon) \le 1-s(x),\ \forall x \in X;$

$1-s(x) \le 1-\beta \cdot (1-\varepsilon)^2,\ \forall x \in F'';$

$1-s(x) \ge 1-\varepsilon,\ \forall x \in F';$

$(1-\beta)+\beta \cdot \varepsilon \le 1-s(x),\ \forall x \in X;$

$1-s(x) \le 1-\beta+\beta \cdot \varepsilon \cdot (2-\varepsilon),\ \forall x \in F'';$

$1-s(x) \ge 1-\varepsilon,\ \forall x \in F'.$

If α, δ are two arbitrary numbers in $(0,1)$ and taking $\beta = 1-\alpha$ in the previous considerations we have:

$\alpha - \delta \le \alpha + \beta \cdot \varepsilon \le 1-s(x),\ \forall x \in X;$

$1-s(x) \le \alpha + \beta \cdot \varepsilon \cdot (2-\varepsilon) \le \alpha + 2 \cdot \varepsilon,\ \forall x \in F'';$

$1-s(x) \ge 1-\varepsilon,\ \forall x \in F'.$

The proof is finish if in the above considerations we put $\varepsilon = \frac{\delta}{2}$.

The following assertion is coming directly from Theorem 2.1 applying the above Lemma.

Theorem 2.2. *If $\mathcal{V} \subset C_1(X)$ is a $V-$family which contains the constant functions $c \in [0,1]$ and which is also an Uryson family, then $\mathcal{V}$ is uniformly dense in $C_1(X)$.*

3. On the Von-Neumann density theorem on compact spaces

We remark that Von Neumann have introduced the class of continuous functions called in the present paper $V-$family in his article [5] on Probabilistic logics. He stated, without proof a density theorem for such set of functions. The first proof of Von Neumann theorem was given by R.I. Jewett in [4]. Here we recover rapidly their result as a corollaries of our general results which may be successfully used also in the Distribution Theory, Measure Theory and so on.

Throughout this section X will be a compact Hausdorff space and $C_1(X)$ the set of all continuous functions on X with values in $I = [0,1]$, and essentially we prove the following Von-Neumann assertion.

Theorem 3.1. *Let $\mathcal{V} \subset C_1(X)$ be a $V-$family which contains the constant functions $c \in [0,1]$ and which separates the points of X. Then $\mathcal{V}$ is uniformly dense in $C_1(X)$.*

Proof. We admit for the moment that the uniform closure of $\mathcal{V}$ is a lattice with respect to the pointwise order relation (see Appendix Corollary 3.1). In this case the proof reduces, using Theorem 2.2, to show that $\mathcal{V}$ is an Uryson class.

Let F', F'' be two arbitrary disjoint, closed subsets of X and let $(x', x'') \in F' \times F''$ be an arbitrary point. Since $\mathcal{V}$ separates the points of X there exists $v \in \mathcal{V}$ such that $v(x'') < v(x')$.

Using the Appendix, Lemma 3.1., there exist two natural numbers n,m such that the function $u_{x',x''}$ from $\mathcal{V}$ given by:

$$u_{x',x''}=\left(1-v^n(x)\right)^m$$

satisfies the relations:

$$u_{x',x''}(x')<\varepsilon;\quad u_{x',x''}(x'')>1-\varepsilon\ .$$

Certainly the sets $[u_{x',x''}<\varepsilon]$ and $[u_{x',x''}>1-\varepsilon]$ are open neighborhoods of the point x', respectively x''. If we fix x' in F' and we consider for each point $x''\in F''$ the function $u_{x',x''}$ as before, then, by compactness reasons, there exists a finite system $x_1'',x_2'',...,x_n''$ in F'' such that:

$$\bigcup_{i=1}^{n}[u_{x',x_i''}>1-\varepsilon]\supset F''\ .$$

Obviously the function:

$$u_{x'}=\underset{i=1}{\overset{n}{\vee}}u_{x',x_i''}$$

belongs to $\mathcal{V}$ and we have:

$$u_{x'}(x')<\varepsilon;\quad u_{x'}>1-\varepsilon \text{ on } F''.$$

If for any point $x'\in F''$ we consider, as before, a function $u_{x'}\in\mathcal{V}$ such that:

$$u_{x'}(x')<\varepsilon;\quad u_{x'}>1-\varepsilon \text{ on } F''.$$

then, since F' is compact and $F'\subset\bigcup_{x'\in F'}[u_{x'}<\varepsilon]$ we may choose a finite system $x_1',...,x_k'$ in F' such that $F'\subset\bigcup_{i=1}^{k}[u_{x_i'}<\varepsilon]$. Obviously, the function $u\in\mathcal{V}$ given by:

$$u=\underset{i=1}{\overset{k}{\wedge}}u_{x_i'}$$

satisfies the relations:

$$u\le\varepsilon \text{ on } F';\ \ u\ge 1-\varepsilon \text{ on } F''\ .$$

Appendix

We remark that the most assertions here comes from [4], but the proofs were considerably improved by G. Paltineanu and I. Bucur in [6], [2].

Lemma 3.1. *If a,b are real numbers such that $0\le a<b\le 1$, then, for any $\varepsilon\in(0,1)$ there exist two natural numbers n,m such that the function $\varphi:[0,1]\to[0,1]$ given by:*

$$\varphi(x)=\left(1-x^n\right)^m$$

have the properties:

$$\varphi\ge 1-\varepsilon \text{ on } [0,a] \text{ and } \varphi\le\varepsilon \text{ on } [b,1]\ .$$

For the proof see [4], [6], [2].

Lemma 3.2. *Let $\mathcal{P}_2$ be the smallest $V-$family of $C_1\left(I^2\right)$ containing the projections p_1,p_2 of I^2 in I i.e.:*

$$p_1(x,y)=x,\ \ p_2(x,y)=y,\ \ \forall x,y\in I,$$

where $I=[0,1]\subset\mathbb{R}$. *Then* $\mathcal{P}_2$ *is an Uryson family of* $C_1\left(I^2\right)$.

For the proof see [4], [6], [2].

Theorem 3.2. *The above family* $\mathcal{P}_2$ *is uniformly dense in* $C_1\left(I^2\right)$.

The assertion follows directly from Theorem 2.2., applying Lemma 3.2.

Corollary 3.1. *If* $\mathcal{V}$ *is a* $V-$*family of* $C_1(X)$ *which containing the constant functions* $c\in[0,1]$ *and separates the points of* X, *then the uniform closure of* $\mathcal{V}$ *is a lattice with respect to the pointwise order relation in* $C_1(X)$.

Proof. Let $f_1,f_2\in\mathcal{V}$ and let $\mathcal{V}(f_1,f_2)$ the smallest $V-$family of $C_1(X)$ containing the functions f_1,f_2. A reason of Halmos's type, in the measure theory, show that we have:

$$\mathcal{V}(f_1,f_2)=\left\{h(f_1,f_2);\ h\in\mathcal{P}_2(I^2)\right\};\quad \overline{\mathcal{V}(f_1,f_2)}\supset\left\{h(f_1,f_2);\ h\in\overline{\mathcal{P}_2(I^2)}\right\}.$$

Particularly, taking $h\in\overline{\mathcal{P}_2(I^2)}$ as follows $h(x,y)=\inf\{x,y\}$ we get $h(f_1,f_2)\in\overline{\mathcal{V}(f_1,f_2)}$ and therefore the function on X:

$$x\to h[f_1(x),f_2(x)]=\inf\{f_1(x),f_2(x)\}$$

belongs to $\overline{\mathcal{V}(f_1,f_2)}$ i.e. it belongs to $\overline{\mathcal{V}}$.

References

[1] Bucur, I.: *A version of Stone-Weierstrass theorem on a G.Paltineanu's result,* Proceedings of the Mathematical and Educational Symposium of Department of Mathematics and Computer Science, UTCB, 2014.

[2] Bucur, I., Paltineanu, G.: *Topics in Uniform Approximation of Continuous Functions*, Frontiers in Mathematics, Springer International Publishing, 2020.

[3] Gillman L., Jerison, M.: *Rings of continuous functions*, Van Nostrand, Princeton, N.J., 1960.

[4] Jewett, R.I.: *A variation on the Stone –Weierstrass theorem*, Proc. Amer. Math. Soc., 14, pp. 690-693, 1963.

[5] von Neumann, J.: *Probabilistic logics and the synthesis of reliable organisms from unreliable components,* Automata Studies, Princeton Univ. Press, Princeton, N.J., pp. 93-94, 1956.

[6] Paltineanu, G., Bucur, I.: *Some Density Theorems in the Set of Continuous Functions with Values in the Unit Interval,* Mediterr. J. Math., 14: 44, DOI 10.1007, 2017.

[7] Prolla, J.B.: *On Von Neumann's variation of the Weierstrass-Stone theorem*, Numer. Funct. Anal. Optim., pp. 133-4, 349-353, 1992.

SOLVING LINEAR SYSTEMS USING MONTE CARLO METHODS

DANIEL CIUIU

Abstract. In this paper we will solve linear systems with positive defined matrices using Monte Carlo methods.

First we estimate the minimum eigenvalue by Monte Carlo methods, or we estimate a lower bound for this eigenvalue. The lower bound can be determined using the eventual diagonal domination, or using the characteristic polynomial.

Next we determine the region of the solution and we determine using Monte Carlo methods the minimum of the classical function involved in relaxation methods.

Mathematics Subject Classification (2010): 65C05, 91G60, 65F10, 65F15.
Key words: Positive defined matrices, relaxation methods, Monte Carlo.

1. Introduction

In the relaxation methods a key theorem is the following [4].

Theorem 1.1. *Solving the linear system $A * x = b$, where the matrix A is symmetric positive defined is equivalent to finding the minimum of $F(x) = \frac{1}{2} * \langle x, A * x\rangle - \langle x, b\rangle$.*

The minimum/ maximum of a function Ψ on the domain D is obtained as follows.

(1) Simulate $nrsim$ uniform random variables on D or on another domain D_1 such that there exists a one-to-one continuous function $\phi : D_1 \to D$.
(2) If we generate Y_i uniform in D_1 we compute $X_i = \phi(Y_i)$.
(3) Compute $\Psi(X_i)$ and determinate minimum/ maximum.

In [1] there is presented how we can compute the minimum eigen value of a symmetric positive defined matrix using the Monte Carlo method. We have simulated normal vectors with zero expectation and the matrix as variance-covariance matrix, and points on the unit sphere as orthogonal regression coefficients. The minimum eigenvalue is the minimum variance of errors.

A special case of matrix A is when the values on the same diagonal of A are the same. This is the case of the transition matrix Q for the constructed absorbing Markov chain for the reliability of hammock networks [2].

If the linear system arrise from linear regression the multicolinearity can be detected [3]. It means that the we have a strong debendence between explanatory variables and the determinant of matrix is closed to zero.One of the method to detect multicolinearity is the BKW=Belsley—Kuh—Weisch test, which detect the ill conditioned linear system. We compute the square root of the condition number, the condition index

$$CI = \sqrt{\frac{\lambda_n}{\lambda_1}}. \tag{1.1}$$

If CI is between 10 and 30 we have moderate multicolinearity, while for $CI > 30$ we have severe multicolinearity.

2. Methodology

If the matrix of the linear system, A, is symmetric positive defined, the matrix is diagonalisable and the eigenvalues are positive. Given U the matrix of the corresponnding eigenvectors on collumns, we have

$$D = U' * A * U, \ hence \tag{2.1}$$

$$A = U * D * U', \tag{2.1'}$$

where D is the diagonal matrix of eigenvalues.

The linear system $A * x = b$ can be written as

$$U * D * U' * x = b, \ or \tag{2.2}$$

$$D * y = c, \tag{2.2'}$$

where $y = U' * x$ and $c = U' * b$.

Because U is a rotation, the Euclidean norm is the same:

$$\begin{cases} \|y\| = \|x\| \\ \|c\| = \|b\| \end{cases} . \tag{2.3}$$

Solving the linear system $(2.2')$ we obtain

$$y_i = \frac{c_i}{\lambda_i}, \tag{2.4}$$

where λ_i are the eigenvalues of A. It results that

$$\|y\| \leq \frac{\|c\|}{\lambda_1}, \tag{2.5}$$

where λ_1 is the minimum eigenvalue.

Therefore we generate feasible solutions in the cube that contains the sphere with center zero and radius $\frac{\|c\|}{\lambda_1}$, and we compute the minimum of the function from Theorem 1.1.

Remark 2.1. *We can substitute the above radius* $\frac{\|c\|}{\lambda_1}$ *by any value* $\frac{\|c\|}{\widetilde{\lambda}}$ *such that* $\widetilde{\lambda} \leq \lambda_1$.

In the same manner we conclude that

$$\|y\| \geq \frac{\|c\|}{\lambda_n}, \tag{2.5'}$$

where λ_n is the maximum eigenvalue. In this way we can simulate the feasible solution on the spheric crown

$$\frac{\|b\|}{\lambda_n} \leq \|y\| \leq \frac{\|b\|}{\lambda_1}, \tag{2.5"}$$

because $\|c\| = \|b\|$ due to the rotation U.

We can simulate feasible solutions inside the sphere of radius $\frac{\|b\|}{\lambda_1}$ according (2.5), or on spheric crown of radius $\frac{\|b\|}{\lambda_n}$ and $\frac{\|b\|}{\lambda_1}$ according (2.5"). To simplify the simulation of the feasible solutions we can replace the sphere of radius $\frac{\|b\|}{\lambda_1}$ by its circumscribed cube, and the of radius $\frac{\|b\|}{\lambda_n}$ by its inscribed cube.

If A is diagonal dominant, we can consider $\widetilde{\lambda}$ as the minimum difference between the absolute value of the diagonal element on the lines and the sum of absolute values of other elements:

$$\widetilde{\lambda}_1 = \min_{i=\overline{1,n}} \left\{ |A_{ii}| - \sum_{j \neq i} |A_{ij}| \right\}. \tag{2.6}$$

The maximum eigenvalue is

$$\widetilde{\lambda}_1 = trace\,(A) - (n-1)\,\widetilde{\lambda}_1, \tag{2.6'}$$

where $trace\,(A)$ is the first Viète sum, i.e. the sum of elements on the diagonal.

If we compute the characteristic polynomial

$$P\,(\lambda) = \lambda^n - S_{n-1} * \lambda^{n-1} + S_{n-2} * \lambda^{n-2} + ... + (-1)^{n-1} * S_1 * \lambda + (-1)^n * S_0, \tag{2.7}$$

where S_i are the sums of Viète. These sums are computed as in classical manner:

(1) S_{n-1} is the trace of matrix.
(2) S_0 is the determinant of the matrix.
(3) The other sums, S_{n-j} for $j = \overline{2, n-1}$, are computed as sums of j successive indices of lines and same for collumns starting with $1, ..., j$ and ending with $n-j+1, ..., n$.

After we have computed the characteristic polynomial P we compute the minimiser of λ_1 as

$$\widetilde{\lambda} = \min_{i=\overline{1,n}} \sqrt[i]{\left| \frac{S_0}{n * S_i} \right|}, \tag{2.8}$$

where $S_n = 1$, and we use the inequality of thriangle.

The presented methodology can be used in linear regression problems, where the minimum squares method yields to a linear system that can be solved by Cholesky method or by relaxation methods [3].

If the matrix A has same values on diagonals we can compute the sums of Viète S_j by computing only the first determinant Δ_j from the classical Jacobi method to reduce a quadratic form to the canonic form, and

$$S_j = (j+1)\,\Delta_j. \tag{2.9}$$

To estimate maximum eigenvalue of A we take into account the wellknown formula for distance between a point M and a line d having the direction $\overrightarrow{v}$

$$dist\,(M, d) = \frac{\|\overrightarrow{MN} \times \overrightarrow{v}\|}{\|\overrightarrow{v}\|}, \tag{2.10}$$

where N is a point on d.

Tacking into account that $\|\overrightarrow{MN} \times \overrightarrow{v}\| = \|\overrightarrow{MN}\| * \|\overrightarrow{v}\| * \sin\theta$, where θ is the angle between $\overrightarrow{MN}$ and $\overrightarrow{v}$, we can write

$$(2.10') \quad \|\overrightarrow{MN} \times \overrightarrow{v}\| = \sqrt{\|\overrightarrow{MN}\|^2 * \|\overrightarrow{v}\|^2 - \|\overrightarrow{MN}\|^2 * \|\overrightarrow{v}\|^2 * \cos^2\theta} = \sqrt{\|\overrightarrow{MN}\|^2 * \|\overrightarrow{v}\|^2 - \langle\overrightarrow{MN} \cdot \overrightarrow{v}\rangle^2},$$

we obtain

$$(2.10") \qquad dist\,(M,d) = \frac{\sqrt{\|\overrightarrow{MN}\|^2 * \|\overrightarrow{v}\|^2 - \langle\overrightarrow{MN} \cdot \overrightarrow{v}\rangle^2}}{\|\overrightarrow{v}\|}.$$

Because the denominator in finding minimum and baximum eigenvalue of the above formula, $\|\overrightarrow{v}\|$, is equal to one, the minimum sum of squares is equal to λ_1, and the maximum sum of squares is S_{max} such that the sum of squares of distances between the $nrsim$ points and the line containing origin and with direction given by λ_n (maximum eigenvalue) is $trace\,(A) - S_{max} = trace\,(A) - \lambda_n$: trace of A is the average of square norms of the points, and average square distance is minimum if and only if the sum of squares of $\langle x, coef\rangle$ is maximum. Therefore the expected dot product between the vector of coefficients and normal vectors has the minimum λ_1 and the maximum λ_n.

The maximum eigenvalue is

$$(2.11) \qquad \widetilde{\lambda}_n = \max_{i=\overline{1,n}} \sqrt[n-i]{|n * S_i|}.$$

If we simulate the feasible solutions in the spheric crown or in the region between the inscribed cube in the small sphere, with radius $\frac{\|b\|}{\lambda_n}$ and the circuscribed cube of big sphere, with radius $\frac{\|b\|}{\lambda_1}$, the ratio between the mentained zone and the big sphere/ cube is

$$(2.12) \qquad ratio = 1 - \left(\frac{\lambda_1}{\lambda_n}\right)^n.$$

Therefore if the multicolinearity is found by the BKW test, we have no improvement from cube/ sphere case to "between cubes"/ spheric crown: we subtract a very small number from one. Due to the exponent n it results that for the same ratio of eigenvalues (hence the same condition number) the above improvement decrease with the dimension of the system n.

3. Applications

Example 3.1. *Solve the linear system*

$$\begin{cases} 9x + 3y - z = 2 \\ 3x + 7y + 2z = 1 \\ -x + 2y + 8z = 3 \end{cases}.$$

The solution of above system obtained with Excel is $x = 0.31034$, $y = -0.11671$ and $z = 0.44297$. With simple relaxation method for given maximal error $\max|r_i| = 0.001$ we obtain after 18 steps $x = 0.31025$, $y = -0.11664$ and $z = 0.44287$. The real error is 0.000587.

The matrix is diagonal dominant, hence it is positive defined.

With suprarelaxation with $\omega = 1.07977$ we obtain in the same conditions as above after 15 steps $x = 0.31028$, $y = -0.11665$ and $z = 0.44296$, and the real error is 0.000442.

The eigenvalues of A are $\lambda_1 = 3.70017$, $\lambda_2 = 9.0852$ and $\lambda_3 = 11.21463$. The condition index is $CI = 1.74093$.

With our $C++$ programe we obtain results presented in Table 1 that follows

TABLE 1. Results for the above system by the Monte Carlo methods

Values method	Monte Carlo	Diagonal dominance	Viète
λ_1	3.66656 3.7346 3.65425 3.63408	2	2.12994
λ_3	11.09817 11.26555 10.88328 11.29358	20	72
Max radius	1.02048 1.00188 1.02392 1.0296	1.87083	1.75669
Min radius	0.33714 0.33213 0.19849 0.33131	0.18708 0.10801	0.05197 0.03
Solution on cube	$\begin{pmatrix} 0.304 \\ -0.10175 \\ 0.43433 \end{pmatrix}$	$\begin{pmatrix} 0.31302 \\ -0.10637 \\ 0.46313 \end{pmatrix}$	$\begin{pmatrix} 0.29454 \\ -0.10546 \\ 0.43351 \end{pmatrix}$
Solution on sphere	$\begin{pmatrix} 0.31338 \\ -0.11431 \\ 0.45005 \end{pmatrix}$	$\begin{pmatrix} 0.32619 \\ -0.11285 \\ 0.42153 \end{pmatrix}$	$\begin{pmatrix} 0.3182 \\ -0.09335 \\ 0.44561 \end{pmatrix}$
Solution between cubes	$\begin{pmatrix} 0.32674 \\ -0.16568 \\ 0.45521 \end{pmatrix}$	$\begin{pmatrix} 0.30922 \\ -0.12265 \\ 0.45948 \end{pmatrix}$	$\begin{pmatrix} 0.31255 \\ -0.10924 \\ 0.43153 \end{pmatrix}$
Solution on spheric crown	$\begin{pmatrix} 0.30255 \\ -0.1154 \\ 0.44722 \end{pmatrix}$	$\begin{pmatrix} 0.30056 \\ -0.13112 \\ 0.43377 \end{pmatrix}$	$\begin{pmatrix} 0.303966 \\ -0.11252 \\ 0.4537 \end{pmatrix}$
Minimum F on cube	-0.91578	-0.91397	-0.91542
Minimum F on sphere	-0.91615	-0.91307	-0.91358
Minimum F between cubes	-0.91005	-0.91538	-0.9158
Minimum F on spheric crown	-0.91608	-0.91435	-0.91566

We have simulated 10000 normal vectors with variance-covariance matrix A, 10000 points on unit sphere and 1 million feasible solutions. The possible solutions are simulated inside the sphere with radius $\frac{1}{\lambda_1}$ or inside the circumscribed cube of this sphere, or betweem the spheres of radius $\frac{1}{\lambda_3}$ and $\frac{1}{\lambda_1}$ (spheric crown), or between the inscribed cube in the sphere of radius $\frac{1}{\lambda_3}$ and the circumscribed cube of the sphere of radius $\frac{1}{\lambda_1}$.

In the above table we have computed four values for λ_1 and four values for λ_3 in the Monte Carlo case, because we simulate normal vectors for each case of possible solution: cube, sphere, between cubes and spheric crown. The same number of values are computed for the maximum/ minimum radius of the spheric crown.

The minimum radius is computed in the case "between cubes" in different manner than in the other case: the solution is between the **inscribed cube** in the small sphere and the circumscribed cube of the big sphere .

We mention that in the case of solution x minimum of $F(x)$ is equal to $-0.5*\langle b, x\rangle = -0.916445623342175$. We notice that in the last four lines the minimum is closed to the real one.

Example 3.2. *Determine the linear regression of the dynamics of bank deposits Y in terms of the dynamics of income X_1 and the dynamics of passive interest rate X_2* [3].

The matrix of linear system is

$$A = \begin{pmatrix} 1 & 1.068 & 3.984 \\ 1.068 & 2.1636 & 4.2696 \\ 3.984 & 4.2696 & 15.8984 \end{pmatrix},$$

and the right sides is

$$b = \begin{pmatrix} 0.452 \\ 1.27 \\ 1.8344 \end{pmatrix}.$$

The linear regression obtained in [3] is

$Y = -3.7869 + 0.7572 * X_1 + 0.861 * X_2$.

The matrix is not diagonal dominant, hence we can not apply the suprarelaxation method. We can neither apply (2.6) for estimating λ_1

The eigenvalues of A are $\lambda_1 = 0.00154$, $\lambda_2 = 0.94901$ and $\lambda_3 = 18.11144$. $CI = 108.32511$, i.r. severe multicolinearity.

If we use in the same conditions as in Example 3.1 the simple relaxation method (maximum error 0.001) we obtain after **2243 steps** $A_0 = -3.17723$, $A_1 = 0.75911$ and $A_2 = 0.70771$. The real error is 0.000999.

If we apply the Monte Carlo method to obtain the minimum/ maximum eigenvalue simulating 10000 normal vectors and 10000 sets of coefficients, we obtain $\lambda_1 = 0.00189$ and $\lambda_3 = 19.0873$. They are close to the above real λ_1 and λ_3 obtained with Scilab. But if we simulate 1 million feasible solutions we obtain strange results: $A_0 = 53.58246$, $A_1 = 1.44435$ and $A_2 = -13.63316$. The minimum of F is 2.490955 instead of -0.4147. We have simulated 1 million feasible solutions, as above, in the cube.

Analogously, if we estimate the values λ_1 and λ_3 using Viète sums we obtain $\lambda_1 = 0.00154$ and $\lambda_3 = 57.186$. Even λ_1 is identical with five digits, we obtain also strange results: $A_0 = -48.92698$, $A_1 = 2.52771$ and $A_2 = 13.80325$. Minimum of F is 38.266.

4. Conclusions

In our $C++$ program to estimate the minimum eigenvalue by Monte Carlo methods we have estimated 10000 random vectors with variance-covariance matrix A and 10000 possible coefficients of orthogonal regression on unit sphere.

When we simulate feasible solutions inside the cube we simulate components uniform in $[-R, R]$. When we simulate points inside the sphere with radius R we simulate the n components uniform on $\left[0, \sqrt{R^2 - R_{crt}^2}\right]$, where initially $R_{crt} = R$, and it is updated after each simulated component.

For the sphere of radius R (for instance $R = 1$ as for minimum/ maximum eigenvalue) we simulate first $n-1$ components as above, and the last one has the absolute value $\sqrt{R^2 - R_{crt}^2}$, and the sign is chosen randomly.

Same random sign is chosen for "between cubes", after we simulate the absolute value between the cube lines lenghts uniformly.

For spheric crown we generate first $n-1$ components as for the outside sphere, and for the last one we check first if we are in the spheric crown with these components: $R_{crt} > R_{min}$. In this case we continue as in the "inside sphere" case. Otherwise the absolute value of the last component is uniform between the reminder radius $\sqrt{R^2 - R_{crt}^2}$, but with $R = R_{min}$ and $R = R_{max}$.

We notice that when we have severe multicolinearity the relaxation methods (if aviable - for instance if we have not diagonal dominance as in Example 3.2 we can not apply suprarelaxation method) need

many steps for a given error. As the number of 2243 steps in this case. With Monte Carlo methods we obtain strange results due to the large radius of the sphere, 1221.82922 using real λ_1 (with Scilab) and 1107.59869 if we estimate $\lambda_1 = 0.00205$ by Monte Carlo methods, respectively 1477.42737 if we estimate $\lambda_1 = 0.00154$ by Viète sums.

When we simulate one billion feasible solutions we obtain reasonable but not closed solution. In the Mobte Carlo estimation of $\lambda_1 = 0.00162$ we obtain $F_{min} = 0.11585$, but the run time is over 3 hours.

In the case of diagonal dominance we have used for λ_n the formula $(2.6')$. An open problem is to estimate λ_n more effective (smaller than using above formula).

References

[1] C. Costinescu and D. Ciuiu, *The Monte Carlo Methods to find eigenvalues and eigenvectors*, in Proceedings of International Conference Trends and Challenges in Applied Mathematics, June 20-23 2007, TUCB, 157-160.

[2] M. Jianu, D. Ciuiu, L. Dăuş and M. Jianu, *Markov Chain Method for Computing the Reliability of Hammock Networks*, Probability in the Engineering and Informational Sciences, **36** (2022), 276-293.

[3] N. Jula and D. Jula, *Modelare economică: Modele econometrice şi de optimizare*, Mustang, Bucharest, 2010.

[4] G. Păltineanu, P. Matei and G.D. Mateescu, *Analiză numerică*, Conspress, Bucharest 2010.

[5] I. Văduva, *Modele de simulare*, Bucharest University Printing House, 2004.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania

Center for Mathematical Modeling, National Institute for Economic Research "Costin C. Kiriţescu", Romanian academy

Email address: dciuiu@yahoo.com

BASIC 'MORPHISMS' IN INTERVAL ANALYSIS

Rodica-Mihaela Dăneţ
Technical University of Civil Engineering Bucharest, Romania
E-mail: rodica.danet@gmail.com
Marian-Valentin Popescu
Technical University of Civil Engineering Bucharest, Romania
E-mail: popescu.marianvalentin@gmail.com
Nicoleta Popescu
E-mail: popescu.snicoleta@gmail.com

Abstract: In 2013, motivated by the idea of rewriting Functional Analysis in a different framework than the historical one, the first author introduced *interval-spaces* (abbreviated as *i-spaces*), each i-space *IE* being associated with a real ordered vector space *E,* as the set of all closed intervals of *E.* Algebraic operations in *IE* were introduced naturally. Also, an order relation was introduced in *IE*, namely the *weak order.* At first we wanted to prove Hahn-Banach type Theorems (both 'Existence' and 'Extension' Theorems). Initially, the value space of the 'morphisms' was $\mathbb{R}$, then in our successive works it was a complete Dedekind vector space *F* and later even an i-space, in particular $I\mathbb{R}$. These 'morphisms' are the *i-*(*sub*)*linear functionals*, the *i-*(*sub*)*linear operators*, the (*sub*)*linear i-functionals* and the (*sub*)*linear i-operators.* The difficulty of our proofs stems from the fact that an i-space is not a vector space, first of all, because not every interval admits a symmetric with respect to the addition introduced in the i-spaces. In this paper, after a brief study of the lattice structure of $I\mathbb{R}$, we will consider sets of certain 'sublinear morphisms', endowed with the pointwise ordering, and we will show that Zorn's Lemma can be applied. The goal is to apply these results in the demonstration of some Hahn-Banach Existence type Theorems.

1. Introduction

Initially*, Interval Analysis* (also known as *Interval Arithmetic*, *Interval Mathematics* or *Interval Computation*) was a method developped by mathematicians since 1950s and 1960s as an approach to putting bounds on rounding errors and measurement errors in mathematical computation and thus developing numerical methods that yield reliable results. Practically, in Interval Analysis, instead of real numbers, one operates with closed intervals of real numbers. There were several mathematicians forerunners of these studies, probably the most important work being written in 1958, by T. Sunaga. The first book on modern Interval Analysis, titled *Interval Analysis,* was written by R.E. Moore in 1966, considered the year of birth of this branch of mathematics. After the appearance of this book, groups from different countries (for example, German School, Bulgarian School, Romanian School) started to investigate systematically the theory and applications of Interval Analysis. Among the contributors, we mention: U. Kulish, S. Markov, R. Anguelov. In 2013, at the seventh Positivity Conference (a Zannen Centennial Conference), Leiden University, the Netherlands, july 22-26, the first author of this paper introduced (see [4] and [5]) the *interval-spaces* (abbreviated as *i-spaces*), as spaces of closed intervals in arbitrary ordered vector spaces. Also she studied ([5], [3] and [6]) the problem of the the extension of some *i-linear functionals*. In two previous papers, [7] and [8], we studied the extension of *i-linear operators* with values in a Dedekind complete lattice space. Three important results were proved: a Hahn-Banach Existence type Theorem, a Mazur-Orlicz Existence type Theorem and a Hahn-Banach Extension type Theorem. We also introduced and studied *normed interval-spaces* – see [10]. Then in the category of these spaces, we considered *bounded interval-operators* and *bounded*

interval-functionals and we formulated results of the Hahn-Banach Extension type Theorem for *bounded interval-functionals*. As is known, the Hahn-Banach Extension Theorem is one of the fundamental results of Functional Analysis. Our main goal was to extend this study to the case where the i-linear operators are in fact, linear i-operators, that is, they act between two i-spaces. In this sense, we mention that the *bounded interval-operators* previously mentioned are some operators that act between two interval-spaces, more precisely between two normed interval-spaces. We mention that the main difficulty to work in the i-space setting comes from the fact that these spaces, endowed with the natural algebraic operations, are not vector spaces.

Note that hereafter we will use the terms of "linear space" and "vector space" as synonyms.

Next, we will specify what the **basic** "**objects**" are introduced in Interval Analysis.

First were considered **intervals on the *real line***: $[\underline{x},\overline{x}]=\{x \mid \underline{x}\le x\le \overline{x}\}$, where $\underline{x}$ and $\overline{x}$ are real numbers. Let's denote the interval above with $[x]$ and call its extremities $\underline{x}$ and $\overline{x}$, the *endpoints* of $[x]$.

More generally, we considered in [4] intervals in an arbitrary *ordered vector space* E. We call the set $IE=\{[x]=[\underline{x},\overline{x}] \mid \underline{x}\le\overline{x} \text{ in } E\}$ the ***interval-set*** (in short, ***i-set***) associated to E. Of course $E\subset IE$ because any $a\in E$ can be written as $[a]=[a,a]$. We will call this interval a ***degenerate interval***. We will call any interval $[a]\in IE\setminus E$ a ***nondegenerate interval.***

The *i*-set IE becomes an ***interval-space*** (in short, ***i-space***) if it is endowed with the following algebraic operations:

1. **the (Minkovski) addition**, defined by $[a]\oplus[b]=\{x+y \mid x\in[a] \text{ and } y\in[b]\}$, that is,

$$[a]\oplus[b]=\left[\underline{a}+\underline{b},\ \overline{a}+\overline{b}\right] \text{ if } [a]=[\underline{a},\overline{a}]\in IE \text{ and } [b]=\left[\underline{b},\overline{b}\right]\in IE;$$

2. **the scalar multiplication with reals**, defined by

$\alpha\cdot[a]=\{\alpha x \mid x\in[\underline{a},\overline{a}]\}$, that is, $\alpha\cdot[a]=\begin{cases}[\alpha\underline{a},\alpha\overline{a}], & \text{if } \alpha\in\mathbb{R},\ \alpha\ge 0\\ [\alpha\overline{a},\alpha\underline{a}], & \text{if } \alpha\in\mathbb{R},\ \alpha<0\end{cases}$, where $[a]=[\underline{a},\overline{a}]\in IE$.

Sometimes, we will denote $\alpha\cdot[a]$ by $\alpha[a]$. It is known that, endowed with these algebraic operations, IE is ***not*** a (real) *vector space*. More precisely, $(IE,\oplus)$ is a commutative monoid, with the *neutral* (or *identity*) element $\mathbf{0}$ ($\mathbf{0}=[0,0]$, sometimes also denoted by $[0]$), but it is ***not*** a *group*, because a nondegenerate closed interval has *no* inverse with respect to the addition, that is, *has* ***no*** *opposite*. Indeed, by way of contradiction, suppose that for the order interval $[a]=[\underline{a},\overline{a}]$ with $\underline{a}<\overline{a}$ in E, there exists an inverse $[b]=\left[\underline{b},\overline{b}\right]$. Hence $[\underline{a},\overline{a}]\oplus\left[\underline{b},\overline{b}\right]=\mathbf{0}$, that is, $\underline{a}+\underline{b}=0$ and $\overline{a}+\overline{b}=0$. Therefore, $\underline{b}=-\underline{a}$ and $\overline{b}=-\overline{a}$. But $\underline{b}\le\overline{b}$ implies that $-\underline{a}\le-\overline{a}$ or, equivalently, $\overline{a}\le\underline{a}$, which contradicts that $\underline{a}<\overline{a}$. We also notice that the scalar multiplication in IE has the following properties:

$$\alpha(\beta[a])=(\alpha\beta)[a], \text{ if } [a]\in IE \text{ and } \alpha,\ \beta\in\mathbb{R}; \tag{1}$$

$$1\cdot[a]=[a] \text{ for each } [a]\in IE; \tag{2}$$

$$\alpha([a]\oplus[b])=\alpha[a]\oplus\alpha[b], \text{ if } [a],[b]\in IE \text{ and } \alpha\in\mathbb{R}; \tag{3}$$

$$(\alpha+\beta)[a]=\alpha[a]\oplus\beta[a],\ \text{if } [a]\in IE \text{ and } \alpha,\ \beta\in\mathbb{R},\ \text{with } \alpha\beta>0. \tag{4}$$

Hence, an *i*-space is **not** a linear space but it is a *quasilinear space in the sense of Markov* (and *in the sense of Aseev*, too) - see, for example, [5, section 2] for definitions, [1], [11], [12] and [14].

We also consider the ***subtraction*** in IE: $[a]\ominus[b]=[a]\oplus(-[b])$, where $-[b]=(-1)\cdot[b]$. If $[a]=[\underline{a},\overline{a}]$, then $[a]\ominus[a]=[-(\overline{a}-\underline{a}),\overline{a}-\underline{a}]$, that is, it is a ***symmetric interval.*** If $\underline{a}<\overline{a}$ (that is, $[a]$ is a nondegenerate interval), then $[a]\ominus[a]\neq\mathbf{0}$ and again we conclude that IE is not a vector space.

Among the most known *order relations* between $[a]=[\underline{a},\overline{a}]$ and $[b]=[\underline{b},\overline{b}]\in IE$ we cite:

1) the ***strong order***: $[a]\prec[b]\Leftrightarrow\overline{a}<\underline{b}$ or $[a]=[b]$ (note that $[a]=[b]\Leftrightarrow\underline{a}=\underline{b}$ and $\overline{a}=\overline{b}$);

2) the ***containment order***: $[a]\leq[b]\Leftrightarrow\underline{b}\leq\underline{a}$ and $\overline{a}\leq\overline{b}$ (that is, $[a]\subseteq[b]$);

3) the ***weak order***: $[a]\leq[b]\Leftrightarrow\underline{a}\leq\underline{b}$ and $\overline{a}\leq\overline{b}$ (used, for example, in [4]).

We say that the *i*-space *IE* is an ***ordered interval-space*** (in short, an ***ordered i-space*****)** if it is endowed with an order relation "$\leq$" such that for any $[x],[y],[z]\in IE$:

1. $[x]\leq[y]$ implies $[x]\oplus[z]\leq[y]\oplus[z]$;

2. For any scalar $\alpha\geq0$, $[x]\leq[y]$ implies $\alpha\cdot[x]\leq\alpha\cdot[y]$.

In the following we will use the ***weak order*** in all *i*-spaces. Endowed with this order relation, any *i*-space becomes an *ordered i-space*.

Also, in what follows we will consider an arbitrary real ordered vector space E and its associated interval-space *IE*. An ***interval-subspace*** of IE or, in short, an ***i-subspace*** of IE, is a nonempty set IG of IE, closed under the algebraic operations (this meaning that for any $[u]$, $[v]\in IG$ and $\alpha\in\mathbb{R}$, we have $[u]\oplus[v]\in IG$ and $\alpha[u]\in IG$.) Obviously $\mathbf{0}=[0,0]\in IG$ (because for any $[u]\in IG$, taking $\alpha=0$, it follows that $\mathbf{0}=0\cdot[u]\in IG$).

Further **we will introduce *interval-lattices*** and ***interval-Dedekind complete lattices***, notions that are necessary for the rest of the work. However, we will not delve the study of these interval-spaces, as this is not the purpose of the work.

An i-subset *IA* of an i-space *IF* is said to be ***i-order bounded*** (*in IF*) if there exist two intervals $[m]$ and $[M]$ in *IF* such that $[m]\leq[a]\leq[M]$ for all $[a]\in IA$.

An i-subset *IA* of an i-space *IF* is called ***i-Dedekind complete*** if for every non-empty i-subset $IB\subseteq IA$ such that *IB* is i-order bounded in *IA*, both sup *IB* and inf *IB* exist and are elements of *IA*. We say that an i-space *IF* is an ***i-Dedekind complete space*** if *IF* is an i-Dedekind complete subset of *IF*.

An ***interval-lattice*** (in short, ***i-lattice***) is an i-space *IF* such that for any pair of intervals $[x],[y]\in IF$, there exist (with respect to the weak order "$\leq$"):

1) their *supremum* (denoted $[x]\vee[y]$, or equivalently, $\sup\{[x],[y]\}$) in *IF*, and

2) their *infimum* (denoted $[x]\wedge[y]$, or equivalently, $\inf\{[x],[y]\}$) in *IF*.

We note that, unlike the case of vector lattices, in the definition of i-lattices, we must demand both the existence of the supremum of any two elements, as well as the existence of their infimum. Indeed for example, the relation that would correspond to equality $x\wedge y=-\big((-x)\vee(-y)\big)$ from vector lattices, would be written formally $[x]\wedge[y]=-\big([-x]\vee[-y]\big)$ and would not always be true. The first argument would be the non-existence, for some intervals, of the symmetric element with respect to the addition introduced in the i-spaces. (We remember that for any interval $[b]$, the notation $-[b]$ means $-[b]=(-1)\cdot[b]$, that is, if $[b]=\left[\underline{b},\overline{b}\right]$, then $-[b]=\left[-\overline{b},-\underline{b}\right]$.)

An i-lattice *F* is called ***interval-Dedekind complete lattice*** (in short, ***i-Dedekind complete lattice***) if it is an i-Dedekind complete space.

Any bounded nonempty subset *S* of the set of real numbers (and more generally, any bounded nonempty subset *S* of a Dedekind complete vector lattice *F*) has an infimum and a supremum.
If *S* is *not lower bounded* (respectively, *not upper bounded*), one often formally writes $\inf S=-\infty$ (respectively, $\sup S=\infty$).

Proposition 1. a) *If F is a vector lattice*, *then IF endowed with the weak order is an i-lattice and if* $[a]=\left[\underline{a},\overline{a}\right]$ *and* $[b]=\left[\underline{b},\overline{b}\right]\in IF$ *, then*:

1) $[a]\wedge[b]=\left[\underline{a}\wedge\underline{b},\overline{a}\wedge\overline{b}\right]$,
2) $[a]\vee[b]=\left[\underline{a}\vee\underline{b},\overline{a}\vee\overline{b}\right]$.

b) More generally, if $\left(\underline{a}_j\right)_{j\in J}$ and $\left(\overline{a}_j\right)_{j\in J}$ are two families of elements of *F* such that $\underline{a}_j\le\overline{a}_j$ for all $j\in J$, and if there exist $\inf\limits_{j\in J}\left(\underline{a}_j\right)$, $\inf\limits_{j\in J}\left(\overline{a}_j\right)$, $\sup\limits_{j\in J}\left(\underline{a}_j\right)$, $\sup\limits_{j\in J}\left(\overline{a}_j\right)$ then also there exist $\inf\limits_{j\in J}\left[\underline{a}_j,\overline{a}_j\right]$ and $\sup\limits_{j\in J}\left[\underline{a}_j,\overline{a}_j\right]$ and, moreover, the following equalities are valid:

1) $\inf\limits_{j\in J}\left[\underline{a}_j,\overline{a}_j\right]=\left[\inf\limits_{j\in J}\left(\underline{a}_j\right),\inf\limits_{j\in J}\left(\overline{a}_j\right)\right]$;
2) $\sup\limits_{j\in J}\left[\underline{a}_j,\overline{a}_j\right]=\left[\sup\limits_{j\in J}\left(\underline{a}_j\right),\sup\limits_{j\in J}\left(\overline{a}_j\right)\right]$.

Lemma 2. *Let F be an ordered vector space and A a nonempty subset of F. Then the set A is order bounded in F if and only if the i-set IA is i-order bounded in IF* (*where IF is endowed with the weak order*).

Proposition 3. *If F is a Dedekind complete vector lattice, then the i-space IF is an i-Dedekind complete lattice.*

2. Preliminaries

Let IE be an i-space, and let IG be an i-subspace of IE.

Definition 1. a) The ***null set*** of IE is the set $\mathcal{O}=\{[-b,b]\mid b\geq 0,\ b\in E\}$ of all symmetric intervals in IE.

b) The ***null part*** $\mathcal{O}_{IG}$ of IG is the set $\mathcal{O}\cap IG$.

The null set $\mathcal{O}$ of IE is an i-subspace of IE. Obviously $[a]\ominus[a]\in\mathcal{O}$ for all $[a]\in IE$. If $[a]$ is a degenerate interval, then $[a]\ominus[a]=\mathbf{0}\in\mathcal{O}$. We will denote by $[o]$ the generic element of $\mathcal{O}$ (that is, for each $[o]\in\mathcal{O}$, there exists $[a]\in IE$, with $[o]=[a]\ominus[a]$). Obviously $\mathcal{O}$ is closed under the algebraic operations on IE.

Also $\mathcal{O}_{IG}=\{[u]\ominus[u]\,|\,[u]\in IG\}$. (Indeed, we notice that for all $[v]\in\mathcal{O}_{IG}$, $[v]=\left[\frac{1}{2}v\right]\ominus\left[\frac{1}{2}v\right]$ and obviously, $\left[\frac{1}{2}v\right]\in IG$, because IG is an i-subspace of IE.)

Next, we will specify what the **basic** "**morphisms**" are that we introduced in Interval Analysis.

Definition 2. If IE is an *i*-space, and IG is an *i*-subspace of IE, we say (see [5]) that a functional $f:IG\to\mathbb{R}$ is an ***interval-linear functional*** (in short, ***i-linear functional***) if:

1. $f([x]\oplus[y])=f([x])+f([y])$ for all $[x],[y]\in IG$ (that is, f is an *interval-additive functional*, or in short, an *i-additive functional*);

2. $f(\alpha[x])=\alpha f([x])$ for all $[x]\in IG$ and $\alpha\in\mathbb{R}$ (that is, f is an *interval-homogeneous functional*, or in short, an *i-homogeneous functional*).

Remark 1. (Properties of an *i*-linear functional)

a) If $f:IG\to\mathbb{R}$ is an *i*-linear functional and $[o]\in\mathcal{O}_{IG}$, then $f([o])=0$. (Indeed, if $[o]=[-a,a]$, then $[o]=[-o]$ and thus $f([o])=f((-1)\cdot[o])=-f([o])$, that is, $f([o])=0$.)

b) If $f:IG\to\mathbb{R}$ is an *i*-linear functional, and $[a]\in IG,[o]\in\mathcal{O}_{IG}$, then $f([a]\oplus[o])=f([a])$. (Obviously, this results from "a)".)

Definition 3. If IE is an *i*-space and IG is an i-subspace of IE, we say that a real-valued function $s:IG\to\mathbb{R}$ is an ***interval-sublinear functional*** (in short, an ***i-sublinear functional***) if:

1) $s([x]\oplus[y])\leq s([x])+s([y])$ for all $[x],[y]\in IG$ (that is, s is an *i-subadditive functional*);

2) $s(\alpha[x])=\alpha s([x])$ for all $[x]\in IG$ and $\alpha>0$ (that is, s is an *i-positively homogeneous functional*);

3) $s([x]\oplus[o])=s([x])$ for all $[x]\in IG$ and $[o]\in\mathcal{O}_{IG}$.

Remark 2. Note that we assume "3)" for an i-sublinear functional since any i-linear functional (that checks "b)" from Remark 1) must be an i-sublinear functional.

Remark 3. (Properties of an i-sublinear functional) Let IG be an i-subspace of an i-space IE $\left(IG \subseteq IE\right)$. Let also $s : IG \to \mathbb{R}$ be an i-sublinear functional. Then:

a) $s\left(\left[o\right]\right) = 0$ for all $\left[o\right] \in \mathcal{O}_{IG}$;

b) $s\left(0 \cdot \left[a\right]\right) = 0$ for all $\left[a\right] \in IG$.

Remark 4. Note that from now on, F will be an arbitrary Dedekind complete vector lattice. Also, IE will be an interval-space and $IG \subseteq IE$ an arbitrary interval-subspace. In the Definition 2 and Definition 3, we defined what it means an *i-linear functional* $f : IG \to \mathbb{R}$ and an *i-sublinear functional* $s : IG \to \mathbb{R}$.

Remark 5. It is obvious that (see [7] and [8]) **we can extend these definitions,** if *instead of real-valued functions*, we consider *operators with values in a Dedekind complete vector lattice.* Thus we will define the notions of *interval-linear operator* (in short, *i-linear operator*) and *interval-sublinear operator* (in short *i-sublinear operator*), respectively. We note not only that the conditions that defined those concepts remain the same, but also, that the same properties from Remark 1 and Remark 3 hold, respectively.

Definition 4. If IE is an i-space, IG is an i-subspace of IE and F is a *Dedekind complete vector lattice,* an operator $L : IG \to F$ is called an *interval-linear operator* or, in short, an *i-linear operator*, if

a) $L\left(\left[a\right] \oplus \left[b\right]\right) = L\left(\left[a\right]\right) + L\left(\left[b\right]\right)$ for all $\left[a\right], \left[b\right] \in IG$ (that is, L is an *i-additive operator*);

b) $L\left(\alpha\left[a\right]\right) = \alpha L\left(\left[a\right]\right)$ for all $\left[a\right] \in IG$ and $\alpha \in \mathbb{R}$ (that is, L is an *i-homogeneous operator*).

Remark 6. (Properties of an i-linear operator) If $L : IG \to F$ is an i-linear operator as in Definition 4, then:

1) $L\left(\left[o\right]\right) = 0$ for all $\left[o\right] \in \mathcal{O}_{IG}$;

2) $L\left(\left[x\right] \oplus \left[o\right]\right) = L\left(\left[x\right]\right)$ for all $\left[x\right] \in IG$ and $\left[o\right] \in \mathcal{O}_{IG}$.

Definition 5. If IE is an i-space, IG is an i-subspace of IE and F is a *Dedekind complete vector lattice,* an operator $S : IG \to F$ is called an *interval-sublinear operator* or, in short, an *i-sublinear operator*, if

a) $S\left(\left[x\right] \oplus \left[y\right]\right) \leq S\left(\left[x\right]\right) + S\left(\left[y\right]\right)$ for all $\left[x\right], \left[y\right] \in IG$ (that is, S is an *i-subadditive operator*);

b) $S\left(\alpha\left[x\right]\right) = \alpha S\left(\left[x\right]\right)$ for all $\left[x\right] \in IG$ and $\alpha > 0$ (that is, S is an *i-positively homogeneous operator*);

c) $S\left(\left[x\right] \oplus \left[o\right]\right) = S\left(\left[x\right]\right)$ for all $\left[x\right] \in IG$ and $\left[o\right] \in \mathcal{O}_{IG}$.

Remark 7. Notice that we assume "c)" for an i-sublinear operator since any i-linear operator, which obviously is an i-sublinear operator, satisfies "2)", from Remark 6.

Remark 8. (Properties of an i-sublinear operator) Let $S:IG\to F$ be an i-sublinear operator as in Definition 5. Then:

d) $S\left(\left[o\right]\right)=0$ for all $\left[o\right]\in\mathcal{O}_{IG}$;

e) $S\left(0\cdot\left[a\right]\right)=0$ for all $\left[a\right]\in IG$;

In [5] the first author generalized in Interval Analysis the well-known Hahn-Banach Existence Theorem, the Mazur-Orlicz Existence Theorem and its consequence, the Hahn-Banach Extension Theorem, all this for functionals defined on i-spaces. For example, an obtained result is the following.

Theorem 4. (a Hahn-Banach Existence type Theorem) *Let IE be an arbitrary i-space and $IG\subseteq IE$ an i-subspace. Let also $s:IG\to\mathbb{R}$ be an i-sublinear functional. Then there exists an i-linear functional $l:IG\to\mathbb{R}$ such that $l\left(\left[v\right]\right)\le s\left(\left[v\right]\right)$ for all $\left[v\right]\in IG$, that is, $l\le s$ on IG.*

We mention that in the following we slightly changed the name of the notions from [10], to better mark that not only the domains of the multivalued functions that appear are *i-spaces*, but, very importantly, the spaces in which they take values are *i-spaces,* too. Thus, for example, instead of *i-*(*sub*)*linear functional*, we will say (*sub*)*linear i-functional* and instead of *i-*(*sub*)*linear operator* we will say (s*ub*)*linear i-operator.*

Definition 6. (see [10] and the new terminology) Let IE and IF be two i-spaces, IG an *i-subspace of* IE, and $T:IE\to IF$ a multivalued function. We say that T is *:*

1) an ***additive interval-operator*** (in short, an ***additive i-operator*****)** if

$$\left[T\left(\left[x\right]\oplus\left[y\right]\right)\right]=\left[T\left(\left[x\right]\right)\right]\oplus\left[T\left(\left[y\right]\right)\right] \text{ for all } \left[x\right],\left[y\right]\in IG;$$

2) a ***homogeneous interval-operator*** (in short, a ***homogeneous i-operator***) if

$$\left[T\left(\alpha\left[x\right]\right)\right]=\alpha\left[T\left(\left[x\right]\right)\right] \text{ for all } \left[x\right]\in IG \text{ and } \alpha\in\mathbb{R}.$$

(Notice that $\left[T\left(\left[x\right]\oplus\left[y\right]\right)\right],\left[T\left(\left[x\right]\right)\right]$ and $\left[T\left(\left[y\right]\right)\right]$ are from *IF*, that is, all these sets are closed intervals in *F*.) *If $F=\mathbb{R}$, and consequently, $IF=I\mathbb{R}$, instead of additive i-operator* (*homogeneous i-operator, respectively*) *we will say* ***additive i-functional*** (***homogeneous i-functional***, *respectively*).

Lemma 5. *An* ***interval*** $\left[a\right]\in IE$ *is* ***symmetric*** *if and only if* $\left[a\right]=-\left[a\right]$ (that is, $\left[a\right]=\left(-1\right)\cdot\left[a\right]$).

Proposition 6. (see [10] and the new terminology) *Let IE and IF be two i-spaces, IG an i-subspace of IE, and $T:IG\to IF$ an operator. If $\left[o\right]\in\mathcal{O}_{IG}$ and $T:IG\to IF$ is an additive i-operator and, simultaneous, a homogeneous i-operator, then $\left[T\left(\left[o\right]\right)\right]\in\mathcal{O}_{IF}$.*

Definition 7. (see [10] and the new terminology) Let IE and IF be two i-spaces and IG an *i-subspace of IE*. We say that a multimap $T:IG\to IF$ is a ***linear interval-operator*** (in short, a ***linear i-operator***) if:

1) $\left[T\left(\left[x\right]\oplus\left[y\right]\right)\right]=\left[T\left(\left[x\right]\right)\right]\oplus\left[T\left(\left[y\right]\right)\right]$ for all $\left[x\right],\left[y\right]\in IG$ (T is an additive i-operator);

2) $\left[T\left(\alpha\left[x\right]\right)\right]=\alpha\left[T\left(\left[x\right]\right)\right]$ for all $\left[x\right]\in IG$ and $\alpha\in\mathbb{R}$ (T is a homogeneous i-operator).

3) $\left[T\left(\left[o\right]\right)\right]=\mathbf{0}\left(=\left[0_F,0_F\right]\right)$ for all $\left[o\right]\in IG$.

If $F=\mathbb{R}$, and consequently, $IF=I\mathbb{R}$, instead of *linear i-operator,* we will say ***linear i-functional.***

We remark that according to Proposition 6, the hypothesis "3)" in Definition 7, makes sense, because $\mathbf{0}\in\mathcal{O}_{IF}$.

Example 1. (a *linear i-operator between two i-spaces* - see [10] and the new terminology) For any fixed λ in $\mathbb{R}$, define $T_\lambda : I\mathbb{R}\to I\mathbb{R}$ by

$$\left[T_\lambda\left(\left[x\right]\right)\right]=\begin{cases}\lambda\left[x\right], \text{ if } \left[x\right]\in I\mathbb{R}\setminus\mathcal{O}_{I\mathbb{R}}\\ \mathbf{0} \quad , \text{ if } \left[x\right]\in\mathcal{O}_{I\mathbb{R}}\end{cases}.$$

Then T_λ is a *linear i-operator.*

Proposition 7. (see [10] and the new terminology) *Let IE and IF be two i-spaces and IG an i-subspace of IE. If $T:IG\to IF$ is a linear i-operator, then $\left[T\left(\left[x\right]\oplus\left[o\right]\right)\right]=\left[T\left(\left[x\right]\right)\right]$ for all $\left[x\right]\in IG$ and $\left[o\right]\in\mathcal{O}_{IG}$.*

In the following we consider that the i-space IF is ordered with so-called *weak order* "$\leq$" (see, for example, [9]). More precisely, if $\left[a\right]=\left[\underline{a},\overline{a}\right]$ and $\left[b\right]=\left[\underline{b},\overline{b}\right]$ are in IF, then

$$\left[a\right]\leq\left[b\right] \text{ if } \underline{a}\leq\underline{b} \text{ and } \overline{a}\leq\overline{b}.$$

Definition 8. (see [10] and the new terminology) Let IE and IF be two i-spaces, and IG an *i-subspace of IE*. We say that an operator $S:IG\to IF$ is:

1) a ***subadditive interval-operator*** (in short, a ***subadditive i-operator***) if

$$\left[S\left(\left[x\right]\oplus\left[y\right]\right)\right]\leq\left[S\left(\left[x\right]\right)\right]\oplus\left[S\left(\left[y\right]\right)\right] \text{ for all } \left[x\right],\left[y\right]\in IG;$$

2) a ***positively homogeneous interval-operator*** (in short, a ***positively homogeneous i-operator***) if

$$\left[S\left(\alpha\left[x\right]\right)\right]=\alpha\left[S\left(\left[x\right]\right)\right] \text{ for all } \left[x\right]\in IG \text{ and } \alpha>0.$$

(Notice that $\left[S\left(\left[x\right]\oplus\left[y\right]\right)\right]$, $\left[S\left(\left[x\right]\right)\right]$ and $\left[S\left(\left[y\right]\right)\right]$ are from IF, that is, all these sets are closed intervals in F.) If $F=\mathbb{R}$, and consequently, $IF=I\mathbb{R}$, instead of *subadditive i-operator* (*positively homogeneous i-operator*, respectively) we will say ***subadditive i-functional*** (***positively homogeneous i-functional***, respectively).

Definition 9. (see [10] and the new terminology) Let IE and IF be two i-spaces, and IG an *i-subspace of IE*. We say that a multivalued function $S:IG\to IF$ is a ***sublinear interval-operator*** (in short, a ***sublinear i-operator***) if:

1) S is a *subadditive i-operator*;

2) S is a *positively homogeneous i-operator*;

3) $\left[S\left(\left[x\right]\oplus\left[o\right]\right)\right]=\left[S\left(\left[x\right]\right)\right]$ for all $\left[x\right]\in IG$ and $\left[o\right]\in\mathcal{O}_{IG}$.

If $F=\mathbb{R}$, and consequently, $IF=I\mathbb{R}$, instead of *sublinear i-operator*, we will say ***sublinear i-functional***.

We notice that the hypothesis "3)" in the Definition 9 is related to the remark that if IE and IF are two i-spaces, and IG is an i-subspace of IE, then any linear i-operator $T: IG \to IF$ (that obviously must be a sublinear i-operator) satisfies the equality

$$\left[T\left(\left[x\right]\oplus\left[o\right]\right)\right]=\left[T\left(\left[x\right]\right)\right]$$

for all $\left[x\right]\in IG$ and $\left[o\right]\in\mathcal{O}_{IG}$, according to Proposition 7**.**

Example 2. (a *sublinear i-functional between two i-spaces*) Define $S: I\mathbb{R}\to I\mathbb{R}$ by

$$\left[S\left(\left[x\right]\right)\right]=\left[\underline{x}+\overline{x},\left|\underline{x}+\overline{x}\right|\right] \text{ if } \left[x\right]=\left[\underline{x},\overline{x}\right]\in I\mathbb{R}.$$

Then *S* is a *sublinear i-functional.*

Proposition 8. (*Properties of sublinear i-operators between two i-spaces*) If IE and IF are two i-spaces, IG is an i-subspace of IE, and $S: IG\to IF$ *is* a *sublinear i-operator, then:*

4) $\left[S\left(\left[o\right]\right)\right]=\mathbf{0}\ \left(\in\mathcal{O}_{IF}\right)$ *for all* $\left[o\right]\in\mathcal{O}_{IG}$;

5) $\left[S\left(0\left[x\right]\right)\right]=\mathbf{0}$ *for all* $\left[x\right]\in IG$.

3. Main part
Pointwise ordering in certain sets of sublinear 'morphisms' in Interval Analysis. Hahn-Banach Existence type Theorems in Interval Analysis

In this paragraph, IE will be an i-space, IG an *i-subspace of* IE and *F* a Dedekind complete vector lattice.

We will ***first*** refer to *i-sublinear functionals* (such as $p: IG\to\mathbb{R}$), respectively to *i-linear functionals* $(l: IE\to\mathbb{R})$. ***Next*** we will consider *i-sublinear operators* $(S: IE\to F)$, and *i-linear operators* $(L: IE\to F)$. ***Finally*** we will refer to *sublinear i-functionals* $(P: IG\to I\mathbb{R})$.

1) *The case of i-sublinear functionals*

We mention that in [5] we introduced the *pointwise ordering on a set* $\mathcal{S}$ of i-sublinear functionals.

In what follows we will fix an i-sublinear functional *p* on an i-subspace *IG* of the i-space *IE* (with *E* an arbitrary ordered vector space). We consider the collection $\mathcal{S}\ \left(=\mathcal{S}_p\right)$ of all i-sublinear functionals *q* on *IG* such that $q\left(\left[v\right]\right)\le p\left(\left[v\right]\right)$ for all $\left[v\right]\in IG$. We remark that $\mathcal{S}$ is a nonempty set because, obviously, $p\in\mathcal{S}$. Consider in $\mathcal{S}$ the *pointwise ordering* "$\le$", that is, for $q_1,q_2\in\mathcal{S}$, $q_1\le q_2\Leftrightarrow q_1\left(\left[v\right]\right)\le q_2\left(\left[v\right]\right)$ for all $\left[v\right]\in IG$.

We consider $\mathcal{T}$ *a nonempty subset of* $\mathcal{S}$**.**

Now recall some order concepts.

An *upper bound* (respectively, a *lower bound*) of $\mathcal{T}$ is an element $q^*\in\mathcal{S}$ such that $q\le q^*$ (respectively, $q^*\le q$) for all $q\in\mathcal{T}$.

A *maximal element* (respectively, a *minimal element*) of $\mathcal{T}$ is an element $q^* \in \mathcal{T}$ such that $q^* \le q$ (respectively, $q \le q^*$) implies $q = q^*$.

The set $\mathcal{S}$ is *inductively ordered from above* (respectively, *from below*) if each totally ordered subset $\mathcal{T}$ of $\mathcal{S}$ has an upper (respectively, a lower) bound. We are interested in this concept because we will apply it to prove a Hahn-Banach Existence type Theorem for i-linear functionals defined on a i-subspace *IG* of an i-space *IE* (see Theorem 4 above), via the following lemma.

Lema 9. (Zorn's lemma) *If* $(\mathcal{S}, \le)$ *is a non-empty ordered set such that* $\mathcal{S}$ *is inductively ordered from above* (*respectively, from below*), *then* $\mathcal{S}$ *has a least one maximal* (*respectively, minimal*) element.

Proposition 10. (see [5, Proposition 2]) *The set* $\mathcal{S}(=\mathcal{S}_P)$ *above is inductively ordered from below* (*that is, each totally ordered subset of* $\mathcal{S}$ *has a lower bound*).
For *proof,* suppose that $\mathcal{T} = \{q_j\}_{j \in J}$ is a totally ordered subset of $\mathcal{S}$ and define $q : IG \to \mathbb{R}$ by $q([v]) = \inf_{j \in J} q_j([v])$ for all $[v] \in IG$. It is shown by the way of contradiction that *q* is *well-defined.* (We begin the proof or this, assuming that there exists $[u] \in IG$ such that $q[u] = -\infty$.) Then we show that $q \in \mathcal{S}$. So it follows that *q* is a lower bound of $\mathcal{T}$. It follows that $\mathcal{S}$ is inductively ordered from below.

2) ***The case of i-sublinear operators***

In what follows, *E* will be an arbitrary ordered vector space and *F* will be an arbitrary Dedekind complete vector lattice.

The following result appeared, without proof, in [7]. For proof, we will add a *smallest element* $(-\infty)$ to the Dedekind complete vector lattice *F*, obtaining $F \cup \{-\infty\}$. This is among those mathematical constructs that are certainly very useful. We came across this technique, for example, in Walter Rudin's popular book, that is, the *Analysis Course* which he has taught to first year graduate students at the University of Wisconsin since 1962 – see for example [13]**.** The first author of this paper also used this technique of widening a Dedekind complete vector lattice with a "*largest*" *element* and respectively, a "*smallest*" *element* in her book *Some techniques for the existence and the extension of linear operators* - see [2], pp. 17-18, and chapter 2 in its part dedicated to the problems concerning the extension of a (positive) linear operator and the existence of a linear operator.

Of course, if we are working with these simbols, we need to be familiar with the rules. For example, $(-\infty) + y = -\infty$ for any *y* in *F.* The calculation conventions in $F \cup \{-\infty\}$ are the same as those in $\mathbb{R} \cup \{-\infty\}$.

Proposition 11. *Let IE be an i-space, IG an i-subspace of IE and F a Dedekind complete vector lattice.* We will fix an i-sublinear operator $P : IG \to F$ and *w*e will consider the collection $\mathcal{S}(=\mathcal{S}_P)$ of all i-sublinear operators $Q : IG \to F$, such that $Q([v]) \le P([v])$ for all $[v] \in IG$. Consider in $\mathcal{S}$ the

pointwise ordering "≤", that is, for $Q_1, Q_2 \in \mathcal{S}$, $Q_1 \le Q_2 \Leftrightarrow Q_1[v] \le Q_2[v]$ for all $[v] \in IG$. *Then the set* $\mathcal{S}$ *is inductively ordered from below* (*that is, each totally ordered subset* $\mathcal{T}$ *of* $\mathcal{S}$ *has a lower bound*).

Proof. Firstly we remark that $\mathcal{S}$ is a nonempty set because, obviously, $P \in \mathcal{S}$. Suppose that $\mathcal{T} = \{Q_j\}_{j \in J}$ is a totally ordered subset of $\mathcal{S}$ and define $Q : IG \to F$ by $Q([v]) = \inf_{j \in J} Q_j([v])$ for all $[v] \in IG$. We consider $F \cup \{-\infty\}$ instead of *F.*

1. Firstly we remark that *Q is well-defined.* Indeed by the *way of contradiction* suppose that there exists $[u] \in IG$ such that $Q([u]) = -\infty$. Then it follows that

$$Q([u] \oplus [v]) = -\infty \text{ for all } [v] \in IG \tag{5}$$

(Indeed, let ε be arbitrary in *F.* Then there exists $j_\varepsilon \in J$ such that

$$Q_{j_\varepsilon}([u]) < \varepsilon . \tag{6}$$

Since Q_{j_ε} is an i-subadditive operator it follows that $Q([u] \oplus [v]) \le Q_{j_\varepsilon}([u] \oplus [v]) \overset{(6)}{<} \varepsilon + Q_{j_\varepsilon}([v])$ for all $[v] \in IG$. Hence

$$Q([u] \oplus [v]) < \varepsilon + Q_{j_\varepsilon}([v]) \text{ for all } [v] \in IG. \tag{7}$$

By choosing ε sufficiently small, from (7) it follows (5)).

But (5) implies that for

$$[o] = [u] \ominus [u] \in \mathcal{O}_{IG},\ Q[o] = -\infty . \tag{8}$$

(Indeed, because for any $[v] \in IG$, this i-subspace also contains $[v] \ominus [u]$, and for all $j \in J$, Q_j is an i-sublinear operator, it follows that

$$Q([v]) = \inf_{j \in J} Q_j([v]) = \inf_{j \in J} Q_j([v] \oplus [o]) = \inf_{j \in J} Q_j([v] \oplus ([u] \ominus [u])) = \inf_{j \in J} Q_j([u] \oplus ([v] \ominus [u])) \overset{(5)}{=} -\infty.$$

Now put $[o]$ instead of $[v]$ and it follows (8)).

But (8) contradicts with $Q([o]) = \inf_{j \in J} Q_j([o]) = 0$.

(Here we used again that for all $j \in J$, Q_j is an i-sublinear operator, and apply axiom 'c)' from Definition 5.) The above mentioned contradiction shows us that $Q([v]) > -\infty$ for all $[v] \in IG$, that is, Q is well-defined. Obviously, Q will be a lower bound of $\mathcal{T}$ in $\mathcal{S}$ if $Q \in \mathcal{S}$.

2. Now we will prove that $Q \in \mathcal{S}$. Because $Q([v])$ is defined as an infimum of $Q_j([v])$, $j \in J$, and for all $j \in J$, $Q_j \in \mathcal{S}$, it follows that:

a) $Q([v]) \leq P([v])$ for all $[v] \in IG$;

b) $Q([v]) \oplus [o] = Q([v])$ for all $[v] \in IG$ *and* $[o] \in \mathcal{O}_{IG}$;

c) Q is an *i-positively homogeneous operator*;

It remains to prove only that Q is an *i-subadditive operator*, that is,

$$Q([v_1] \oplus [v_2]) \leq Q([v_1]) + Q([v_2]),$$

for all $[v_1], [v_2] \in IG$.

(To prove this we mention that

$$Q([v_1] \oplus [v_2]) = \inf_{j \in J} Q_j([v_1] \oplus [v_2]) \leq \inf_{j \in J} (Q_j[v_1] + Q_j[v_2]). \tag{9}$$

Now we will prove that

$$\inf_{j \in J} (Q_j[v_1] + Q_j[v_2]) \leq Q([v_1]) + Q([v_2]). \tag{10}$$

We know that $Q([v_1]) > -\infty$ and $Q([v_2]) > -\infty$.

Given any $\varepsilon > 0$ in F, there exist $j_1, j_2 \in J$ such that

$$\inf_{j \in J} (Q_j[v_1]) + \varepsilon > Q_{j_1}[v_1] \text{ and } \inf_{j \in J} (Q_j[v_2]) + \varepsilon > Q_{j_2}[v_2].$$

Therefore, since $\mathcal{T}$ is a totally ordered subset of $\mathcal{S}$, it follows that

$$Q([v_1]) + Q([v_2]) + 2\varepsilon = \inf_{j \in J} (Q_j[v_1]) + \inf_{j \in J} (Q_j[v_2]) + 2\varepsilon > Q_{j_1}[v_1] + Q_{j_2}[v_2]$$

$$\geq \begin{cases} Q_{j_1}[v_1] + Q_{j_1}[v_2], \text{ if } Q_{j_1} \leq Q_{j_2} \text{ in } \mathcal{T} \\ Q_{j_2}[v_1] + Q_{j_2}[v_2], \text{ if } Q_{j_2} \leq Q_{j_1} \text{ in } \mathcal{T} \end{cases} \geq \inf_{j \in J} (Q_j[v_1] + Q_j[v_2]).$$

Since $\varepsilon > 0$ in F is arbitrary and F is an *Archimedean vector lattice* because it is a *Dedekind complete vector lattice*, it follows that the inequality (10) is true.

From (9) and (10) it follows that Q is an i-subadditive operator.)

Therefore, according to properties a), b), c) of Q mentioned above, we deduce that $Q \in \mathcal{S}$. This shows us that Q is a lower bound of $\mathcal{T} = (Q_j)_{j \in J}$ in $\mathcal{S}$. So, we infer that $\mathcal{S}$ (endowed with the *pointwise ordering*) is inductively odered from below. □

Using Proposition 11, we can prove the following Hahn-Banach Existence type Theorem – see [7, Theorem 5, without proof].

Theorem 12. (the operatorial form of a Hahn-Banach Existence Theorem in Interval Analysis, for the case that the functions that appear take values in a Dedekind complete vector lattice). *Let IE be an i-space, IG an i-subspace of IE, and F a Dedekind complete vector lattice. For every i-sublinear operator $S: IG \to F$ there exists an i-linear operator $T: IE \to F$ such that $T \le S$ on IG (that is, $T([v]) \le S([v])$ for all $[v] \in IG$).*

Proof. Denote by $\mathcal{S}$ the (nonempty) set of all i-sublinear operators $Q: IG \to F$ such that $Q \le S$. We consider $F \cup \{-\infty\}$. From Proposition 11, it follows that $\mathcal{S}$ endowed with the *pointwise ordering* is inductively ordered from below. Therefore, by Zorn's lemma, there exists at least one minimal element $L \in \mathcal{S}$, that is, L is an i-sublinear operator on IG such that

$$L([v]) \le S([v]) \quad \text{for all } [v] \in IG.$$

It remains to be demonstrated that L is an i-linear operator on IG. For this we note that:

1) $L = H$, where $H: IG \to F$ is defined by

$$H([v]) = \inf_{\substack{\alpha>0 \\ [u]\in IG}} \left(L([v] \oplus \alpha[u])\right) - \alpha L([u]) \text{ for all } [v] \in IG.$$

Indeed, H is *well-defined* (that is, $H([v]) > -\infty$ for all $[v] \in IG$), H is an *i-sublinear operator, $H \in \mathcal{S}$, and $H \le L$*; but L is a *minimal element* in $\mathcal{S}$, and then $L \le H$ on IG. So, $H = L$ on IG.

2) $L([v_1]) = H([v_1]) \le L([v_1] \oplus \alpha([v_2])) - \alpha L([v_2])$ for all $\alpha > 0$ and for all $[v_1], [v_2] \in IG$. Taking $\alpha = 1$, it follows that

$$L([v_1]) \le L([v_1] \oplus [v_2]) - L([v_2]).$$

3) $L(\lambda[v]) = \lambda L([v])$ for all $[v] \in IG$ and $\lambda < 0$.

Consequently, L is an i-linear operator on IG *and* $L([v]) \le S([v])$ for all $[v] \in IG$.

So we will choose $T = L$. □

3) *The case of sublinear i-operators*

Proposition 13. *Let IE be an i-space, IG an i-subspace of IE, and $P: IG \to I\mathbb{R}$ a sublinear i-functional. We endow with the pointwise ordering "$\le$", any family of sublinear i-functionals between IG and $I\mathbb{R}$, that is, for any two such sublinear i-functionals Q_1, Q_2, we assume that $Q_1 \le Q_2$ if $[Q_1[v]] \le [Q_2[v]]$ (in $I\mathbb{R}$) for all $[v] \in IG$. We consider the collection $\mathcal{S} = \mathcal{S}_P$ of all sublinear i-functionals $Q: IG \to I\mathbb{R}$ such that $Q \le P$. Then, $\mathcal{S}$ is inductively ordered from below* (*that is, each totally ordered subset $\mathcal{T}$ of $\mathcal{S}$ has a lower bound*).

For *proof*, firstly we remark that $\mathcal{S}$ is a nonempty set because, obviously, $P \in \mathcal{S}$.

Suppose that $\mathcal{T}=\left(\left[Q_j\right]\right)_{j\in J}$ is a totally ordered subset of $\mathcal{S}$.

Define $Q: IG\to I\mathbb{R}$ by $\left[Q\left(\left[v\right]\right)\right]=\inf\limits_{j\in J}\left[Q_j\left(\left[v\right]\right)\right]$ for all $\left[v\right]\in IG$. We will prove that:

1) Q is *well-defined,* that is, $-\infty\notin\left[Q\left(\left[v\right]\right)\right]$ for all $\left[v\right]\in IG$;

2) $Q\in\mathcal{S}$;

3) Q is a lower bound of $\mathcal{T}$.

Obviously, "2)" implies "3)". So it remains to demonstrate "1)" and "2)". We also mention that it is useful to write the intervals that give the values of Q and of each Q_j, by using their endpoints. For example, to prove "1)", by the *way of contradiction,* suppose that there exists $\left[u\right]\in IG$ such that $-\infty\in\left[Q\left(\left[u\right]\right)\right]$.
This means that

$$\underline{Q\left(\left[u\right]\right)}=-\infty,$$

where $\left[Q\left(\left[u\right]\right)\right]=\left[\underline{Q\left(\left[u\right]\right)},\overline{Q\left(\left[u\right]\right)}\right]$.

References

[1] Aseev, S.M.: *Quasilinear operators and their application in the theory of multivalued mappings*, Proceedings of the Steklov Institute of Mathematics **2**, 23-52 (1986).

[2] Dăneț, R-M.: *Some techniques for the existence and the extension of linear operators* (in Romanian), Publishing House of the Romanian Academy, Bucharest, pp 17-18 (1993).

[3] Dăneț, R-M.: *From vector spaces to interval-spaces via the technique of the auxiliary sublinear functionals*, Proceedings of the 13th Workshop of Scientific Communications, Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, May 23, 2015, Matrix Rom Publishing House, Bucharest, 38-43 (2015).

[4] Dăneț, R.-M.: *A Mazur-Orlicz Type Theorem in Interval Analysis and its Consequences*, Proceedings of the seventh Positivity conference, 22-26 July, 2013, Leiden.

[5] Dăneț, R-M: *A Mazur-Orlicz Type Theorem in Interval Analysis and its Consequences*, Ordered Structures and Applications, Positivity VII, Trends in Mathematics, 127-159, 2016, Springer International Publishing.

[6] Dăneț, R-M.: *A new example that shows how we can correct the "defect" of the addition in any interval-space,* Proceedings of Mathematics and Educational Symposium of Department of Mathematics and Computer Science, the 2nd edition, Technical University of Civil Engineering Bucharest, May 28, 2016, Matrix Rom Publishing House, Bucharest, 45-50 (2016).

[7] Dăneț, R-M., Popescu, M.-V., Popescu, N.: *The technique of the auxiliary i-sublinear operator in interval analysis. 1.*, Proceedings of the 17th Workshop on Mathematics, Computer Science and Tehnical Education, Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, June 20, 2020, volume 3/2020, Matrix Rom Publishing House, Bucharest, 32-37 (2020)

[8] Dăneţ, R-M., Popescu, M-V., Popescu, N.: *The technique of the auxiliary i-sublinear operator in interval analysis. 2.*, Proceedings of the 18th Workshop on Mathematics, Computer Science and Technical Education, Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, May 29, 2021, volume 4/2021, Matrix Rom Publishing House, Bucharest, 16-23 (2021).

[9] Dăneţ, R-M., Popescu, M.-V., Popescu, N.: *Order relations in interval-spaces*, Proceedings of the 19th Workshop on Mathematic Computer Science and Technical Education, Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, June 11, 2022, volume 5/2022, Matrix Rom Publishing House, Bucharest, 15-29 (2022).

[10] Dăneţ, R-M., Popescu, M-V., Popescu, N.: *The setting of normed interval spaces*, Proceedings of the 20th Workshop on Mathematic Computer Science and Technical Education, Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, May 27, 2023, Matrix Rom Publishing House, Bucharest, volume **6**(2023), 26-38 (2023).

[11] Markov, S.: *On the algebra of intervals and convex bodies*, Journal of Universal Computer Science **4**, 34-47 (1998).

[12] Markov, S.: *Quasilinear spaces and their relation to vector spaces*, Electronic Journal on Mathematics of Computation, 2-1 (2005).

[13] Rudin, W: *Real and Complex Analysis*, International Student edition, *McGRAW-Hill,* London, NewYork, Sidney, Toronto, Mexico, Johannesburg, Panama, Singapore and *MLADINSKA KNJIGA,* Ljubljana, 1970, available at https://sowndarmath.files.wordpress.com/2017/10/real-and-complex-analysis-w-rudin.pdf

[14] Yilmaz, Y., Çakan, S., Aytekin, S.: *Topological quasilinear spaces*, Abstract and Applied Analysis **2012**, Article ID 951374, 10 pages (2012).

COUNTING LATTICE PATHS TO FIND THE COEFFICIENTS OF RELIABILITY POLYNOMIALS

MARILENA JIANU AND LEONARD DĂUŞ

Abstract. The first two coefficients of the reliability polynomial of hammock networks can be calculated by counting the lattice paths with steps of the form $(1,1)$ and $(1,-1)$, restricted in a strip along the x-axis. In this paper, we use the instruments of linear algebra (eigenvalues and eigenvectors) to count the lattice paths in a hammock network. As a consequence, some identities involving trigonometric power sums are obtained.

Mathematics Subject Classification (2010): 05C50, 05C31, 05C38, 94C15.
Key words: lattice paths, reliability polynomial, eigenvalues, eigenvectors.

1. Introduction

Lattice paths are a fundamental concept in mathematics, especially within the field of combinatorics and its applications in various areas such as network reliability and optimization. They are often used to model a variety of real-world scenarios where movement or decision-making follows a discrete set of possibilities.

In its essence, a lattice path involves a sequence of moves from a lattice point to another, subject to certain constraints. A lattice path in $\mathbb{Z}^d$ with steps in the set S (a subset of $\mathbb{Z}^d$) is a sequence $v_0, v_1, \ldots, v_k \in \mathbb{Z}^d$ such that each consecutive difference $v_i - v_{i-1}$ lies in S [13]. Lattice paths play an important role in combinatorics, being an useful instrument for studying various combinatorial objects. For instance, if we denote by e_i the unit coordinate vector in $\mathbb{Z}^d$, the number of lattice paths with steps in $S = \{e_1, \ldots, e_d\}$ from the origin to a point $a = (a_1, \ldots, a_d) \in \mathbb{N}^d$ is given by the multinomial coefficient $\binom{a_1+\ldots+a_d}{a_1,\ldots,a_d}$. Another classical result (see [14]) is related to the *Dyck paths* - the lattice paths in the 2-dimensional space $\mathbb{Z}^2$ with steps in $S = \{(1,1),(1,-1)\}$ joining the point $(0,0)$ to the point $(2n,0)$ and never passing below the x-axis. The number of Dyck paths is equal to the Catalan number: $C_n = \frac{1}{n+1}\binom{2n}{n}$.

The study of lattice paths not only enriches our understanding of combinatorial structures but also offers insights into practical problems in network theory. Lattice paths can be used to represent different connections within a network, and analyzing these paths can help in calculating the reliability of the network. For instance, the *hammock networks* introduced in [12] are a regular type of networks highly appreciated for their robustness. In spite of their regularity, the computation of the reliability polynomial of hammock networks of large dimensions is a difficult task. Consequently, different methods of approximation, involving beta distribution [2], cubic splines [4], Hermite interpolation polynomial [9] or Markov chains [11] were investigated. The full Hermite polynomial provides an accurate approximation if the first two coefficients of the reliability polynomial

are known and these coefficients can be calculated by counting some special type of lattice paths. In the present paper, following the ideas from [8], we find a recurrence relation to compute the number of lattice paths of length $k+1$ based on the lattice paths of length k. The recurrent computing process plays an important role in applied mathematics, having numerous applications in various domains (see, for instance [6, 7]). An explicit expression is found from the recurrence relation, such that we can use the instruments of linear algebra (eigenvalues and eigenvectors) to count the lattice paths in a hammock network.

2. The main result

We recall that a hammock network of length l and width w of the first (or second) kind is a probabilistic graph whose vertices are all the even (or odd, respectively) lattice points (x, y) in the rectangle $0 \leq x \leq l$, $0 \leq y \leq w$ and whose (undirected) edges are of the form $\{(x, y), (x+1, y+1)\}$, or $\{(x, y), (x+1, y-1)\}$. The vertices on the y-axis are called *source nodes*, while those on the vertical line $x = l$ are called *terminus nodes*. Each edge is supposed to be operational with the probability p, while the nodes are supposed to be always operational. The reliability of the network is the probability to have at least one lattice path made up of operational edges, connecting a source point to a terminus point. This probability is expressed by the following polynomial in p:

$$h(p) = \sum_{k=1}^{wl} N_k p^k (1-p)^{wl-k}, \tag{2.1}$$

where N_k is the number of subgraphs with exactly k edges, containing at least one lattice path that joins a source point to a terminus point. Obviously, such a subgraph has at least l edges, so $N_k = 0$ for every $k = 1, \ldots, l-1$, and N_l is equal to the number of lattice paths with steps of the form $(1, 1)$ and $(1, -1)$, connecting even / odd points on the y-axis with even / odd points on the straight line $x = l$ and never passing below the x-axis, or above the horizontal line $y = w$. We call them lattice paths of length l confined in the strip $0 \leq y \leq w$. There are two kinds of such lattice paths: even (corresponding to networks of the first kind) and odd (for networks of the second kind). We use the following notations:

$N(w, l)$ = the number of all *Dyck-type* lattice paths of length l confined in the strip $0 \leq y \leq w$.

$N_e(w, l)$ = the number of even *Dyck-type* lattice paths of length l confined in the strip $0 \leq y \leq w$.

$N_o(w, l)$ = the number of odd *Dyck-type* lattice paths of length l confined in the strip $0 \leq y \leq w$.

Due to the symmetry of the hammock network, the number of odd and even lattice paths is the same when at least one of the network's dimensions l and w is odd, hence $N_e(w, l) = N_o(w, l) = \frac{1}{2} N(w, l)$ in this case.

We begin by counting the even lattice paths for some fixed width of the network, w. For every $i = 1, 2, \ldots, l$, and $j = 0, 1, \ldots, w$, we denote by $x_{i,j}$ the number of even Dyck

paths of length i starting at a point on the y-axis and ending up at the point (i, j). Let

$$x_{0,j} = \begin{cases} 0 & \text{if } j \text{ is odd} \\ 1 & \text{if } j \text{ is even.} \end{cases}$$

be the even lattice points on the y-axis.

Then, for every $i = 1, 2, \ldots$, one can write:

$$x_{i,j} = \begin{cases} x_{i-1,1} & \text{if } j = 0 \\ x_{i-1,j-1} + x_{i-1,j+1} & \text{if } 1 \leq j \leq w-1 \\ x_{i-1,w-1} & \text{if } j = w. \end{cases} \tag{2.2}$$

Note that $x_{i,j} = 0$ if $i + j$ is odd.

Let $A = (a_{i,j})_{i,j=0,\ldots,w}$ be the square matrix of dimension $w + 1$ defined by:

$$a_{ij} = \begin{cases} 1, & \text{if } |i-j| = 1 \\ 0, & \text{if } |i-j| \neq 1 \end{cases} \qquad i, j = 0, 1, \ldots, w. \tag{2.3}$$

If we denote by $X^{(i)}$ the (column) vector of components $x_{i,j}$, $j = 0, \ldots, w$, then the next recurrence relation follows by (2.2) :

$$X^{(i)} = AX^{(i-1)}, \text{ for all } i = 1, 2, \ldots. \tag{2.4}$$

We denote by $U = (11\ldots1)^T$, $U_e = (0101\ldots)^T$, and $U_o = (1010\ldots)^T$. By (2.4) we obtain that

$$X^{(i)} = A^i U_e, \text{ for all } i = 1, 2, \ldots, \tag{2.5}$$

so the number of even Dyck-type lattice paths of length l confined in the strip $0 \leq y \leq w$ is $N_e(w, l) = \sum_{j=0}^{w} x_{l,j} = U^T A^l U_e$.

In the same way, if we take as initial vector the vector of odd points on y-axis, $X^{(0)} = U_o$, then the number of odd paths is obtained: $N_o(w, l) = U^T A^l U_o$.

As noted above, the number of even / odd Dyck paths is equal to the first nonzero coefficient N_l of the reliability polynomial (2.1) for a hammock network of the first kind, or of the second kind, respectively (see [10]). The authors found in [9] the expression of the first nonzero coefficient for hammock networks with $l \geq w+1$. We have proved above a general result, for any dimensions l, w:

Theorem 2.1. *The number of lattice paths with steps in $S = \{(1, 1), (1, -1)\}$ joining points on y-axis to points on the line $x = l$ and never passing below x-axis or above the straight line $y = w$ is*

$$N(w, l) = U^T A^l U, \tag{2.6}$$

the number of even paths is

$$N_e(w, l) = \sum_{j=0}^{w} x_{l,j} = U^T A^l U_e, \tag{2.7}$$

and the number of odd paths is

$$N_o(w,l) = \sum_{j=0}^{w} x_{l,j} = U^T A^l U_o, \tag{2.8}$$

where $A = (a_{i,j})_{i,j=0,\ldots,w}$ *is the square matrix defined by equation* (2.3), $U = (111\ldots1)^T$, $U_e = (1010\ldots)^T$, *and* $U_o = (0101\ldots)^T$.

Let us have a better look at the matrix A:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots & 0 & 0 \\ 1 & 0 & 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & 1 & \ldots & 0 & 0 \\ 0 & 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 1 & 0 \end{pmatrix}. \tag{2.9}$$

One can remark that A is the adjacency matrix of the path graph $P_n = (V, E)$ (where $n = w + 1$) having the set of vertices $V = \{0, 1, \ldots, w\}$ and the set of edges $E = \{(0,1), (1,2), \ldots, (w-1, w)\}$.

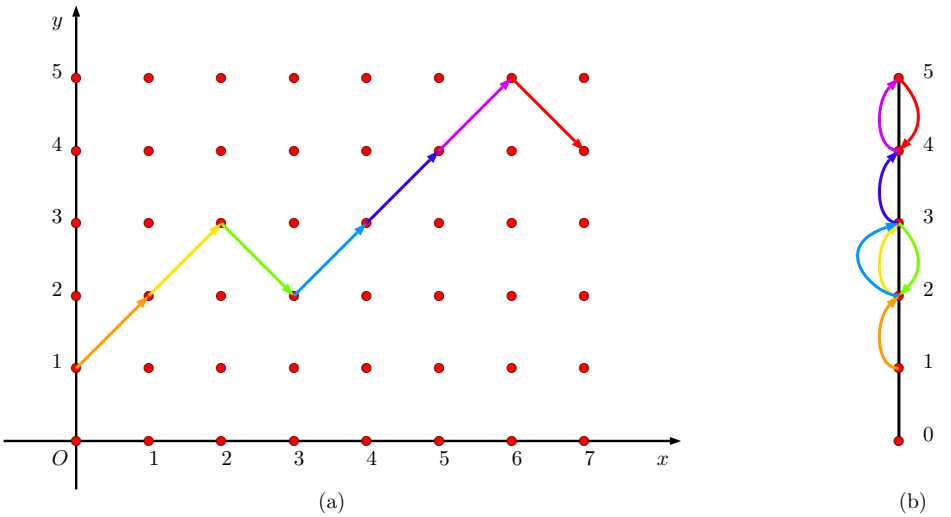


FIGURE 1. (a) A Dyck path of length $l = 7$ with $0 \le y \le w = 5$, from the point $(0,1)$ to the point $(7,4)$; (b) The corresponding walk of length 7 on the path graph $P_{w+1} = P_6$: 1 - 2 - 3 - 2 - 3 - 4 - 5 - 4

For every $i, j = 0, \ldots, w$, the element $a_{i,j}^{(k)}$ of the matrix A^k is equal to the number of walks of length k on the path graph P_n joining the the vertex i to the vertex j [3, Lemma 2.5]. Obviously (see Figure 1), this number also counts the Dyck-type lattice paths confined in the strip $0 \le y \le w$, connecting the points $(0, i)$ and (k, j) (it is 0

whenever $i+j+k$ is odd). We also remark that if $i=j$ then the element $a_{i,i}^{(k)}$ of the matrix A^k is equal to the number of closed walks of length k, evaluated in [5].

The eigenvalues of the matrix A are (see [1, Theorem 3.7]):

$$\lambda_j = 2\cos\frac{j\pi}{n+1}, \quad j=1,2,\ldots,n, \tag{2.10}$$

and the corresponding eigenvectors

$$\mathrm{u}^{(j)} = \sqrt{\frac{2}{n+1}}\left(\sin\frac{j\pi}{n+1}, \sin\frac{2j\pi}{n+1}, \ldots, \sin\frac{nj\pi}{n+1}\right)^T \quad j=1,2,\ldots,n \tag{2.11}$$

form an orthonormal basis of $\mathbb{R}^n$. We denote by C the orthogonal (and symmetric) matrix whose columns are the eigenvectors $\mathrm{u}^{(1)}, \mathrm{u}^{(2)}, \ldots, \mathrm{u}^{(n)}$ and by D the diagonal matrix $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Since $A = CDC$, the powers of the matrix A can be calculated as $A^k = CD^kC$ and the next theorem follows.

Theorem 2.2. *The number of lattice paths with steps in $S = \{(1,1),(1,-1)\}$ joining points on y-axis to points on the line $x=l$ and never passing below x-axis or above the straight line $y=w$ is*

$$N(w,l) = \frac{2^{l+1}}{w+2}\sum_{k=0}^{\lfloor \frac{w}{2} \rfloor} \frac{\cos^l \frac{(2k+1)\pi}{w+2}}{\tan^2 \frac{(2k+1)\pi}{2(w+2)}}. \tag{2.12}$$

If at least one of the numbers w, l is odd, then the number of even paths is equal to the number of odd paths:

$$N_e(w,l) = N_o(w,l) = \frac{1}{2}N(w,l) = \frac{2^l}{w+2}\sum_{k=0}^{\lfloor \frac{w}{2} \rfloor} \frac{\cos^l \frac{(2k+1)\pi}{w+2}}{\tan^2 \frac{(2k+1)\pi}{2(w+2)}}. \tag{2.13}$$

If w and l are both even, then the number of even paths is

$$N_e(w,l) = \frac{2^{l+1}}{w+2}\sum_{k=0}^{\frac{w}{2}} \frac{\cos^l \frac{(2k+1)\pi}{w+2}}{\sin^2 \frac{(2k+1)\pi}{w+2}}, \tag{2.14}$$

while the number of odd paths in this case is

$$N_o(w,l) = \frac{2^{l+1}}{w+2}\sum_{k=0}^{\frac{w}{2}} \frac{\left(2+\cos\frac{(2k+1)\pi}{w+2}\right)\cos^{l+1}\frac{(2k+1)\pi}{w+2}}{\sin^2 \frac{(2k+1)\pi}{w+2}}, \tag{2.15}$$

Proof. We denote by $n = w+1$ and let A be the matrix (2.9). By Theorem 2.1 and using (2.10) and (2.11) we can write:

$$\begin{aligned} N(w,l) &= U^T A^l U = U^T C D^l C U \\ &= \frac{2^{l+1}}{n+1} \sum_{r=1}^{n} \cos^l \frac{r\pi}{n+1} \sum_{i,j=1}^{n} \sin \frac{ri\pi}{n+1} \sin \frac{rj\pi}{n+1} \\ &= \frac{2^{l+1}}{n+1} \sum_{r=1}^{n} \cos^l \frac{r\pi}{n+1} \left(\sum_{j=1}^{n} \sin \frac{rj\pi}{n+1} \right)^2. \end{aligned}$$

Since

$$\sum_{j=1}^{n} \sin \frac{rj\pi}{n+1} = \begin{cases} 0 & \text{if } r \text{ is even} \\ \frac{\cos \frac{r\pi}{2(n+1)}}{\sin \frac{r\pi}{2(n+1)}} & \text{if } r \text{ is odd,} \end{cases}$$

the formula (2.12) follows.

If w is odd, then the symmetric of an even path w.r.t. the straight line $y = \frac{w}{2}$ is an odd path, and conversely, the symmetric of an odd path is an even path. Similarly, if l is odd, we consider the symmetry w.r.t. the straight line $x = \frac{l}{2}$, so we have as many even as odd lattice paths in this case.

Suppose that w and l are both even. By Theorem 2.1 and using (2.10) and (2.11) we can write:

$$\begin{aligned} N_e(w,l) &= U^T A^l U_e = U_e^T C D^l C U_e \\ &= \frac{2^{l+1}}{n+1} \sum_{r=1}^{n} \cos^l \frac{r\pi}{n+1} \sum_{i,j=0}^{\frac{n-1}{2}} \sin \frac{r(2i+1)\pi}{n+1} \sin \frac{r(2j+1)\pi}{n+1} \\ &= \frac{2^{l+1}}{n+1} \sum_{r=1}^{n} \cos^l \frac{r\pi}{n+1} \left(\sum_{j=0}^{\frac{n-1}{2}} \sin \frac{r(2j+1)\pi}{n+1} \right)^2. \end{aligned}$$

Since

$$\sum_{j=0}^{\frac{n-1}{2}} \sin \frac{r(2j+1)\pi}{n+1} = \begin{cases} 0 & \text{if } r \text{ is even} \\ \frac{1}{\sin \frac{r\pi}{n+1}} & \text{if } r \text{ is odd,} \end{cases}$$

the formula (2.14) follows.

Finally, since $N_o(w,l) = N(w,l) - N_e(w,l)$, the formula (2.15) follows. □

3. Applications. Trigonometric sums

Knowing that the number of even paths, $N_e(w,l)$ is equal to the first coefficient of the reliability polynomial of hammock networks of the first kind, we compare the results in Theorem 2.2 to the formulas obtained in [9, Theorem 5] for the case $l \leq w+1$:

i) If l is odd, then

$$N_e(w,l) = 2^{l-1}(w+2) - (l+1)\binom{l}{\frac{l-1}{2}} \tag{3.1}$$

ii) If l is even and w is odd, then

$$N_e(w,l) = 2^{l-1}(w+2) - \frac{2l+1}{2}\binom{l}{\frac{l}{2}} \tag{3.2}$$

iii) If l and w are both even, then

$$N_e(w,l) = 2^{l-1}(w+2) - l\binom{l}{\frac{l}{2}} \tag{3.3}$$

Using the formulas (2.13) and (3.1), and the notations $w = m-2$, $l = 2n-1$, the next corollary follows:

Corollary 3.1. *For any positive integers m, n such that $m \geq 2n$ we have:*

$$\sum_{k=1}^{\lfloor \frac{m}{2} \rfloor} \frac{\cos^{2n-1}\frac{(2k-1)\pi}{m}}{\tan^2\frac{(2k-1)\pi}{2m}} = \frac{m^2}{2} - \frac{mn}{2^{2n-2}}\binom{2n-1}{n}.$$

Using the formulas (2.13) and (3.2), and the notations $w = 2p-1$, $l = 2n$, the next corollary follows:

Corollary 3.2. *For any positive integers n, p such that $p > n$ we have:*

$$\sum_{k=1}^{p} \frac{\cos^{2n}\frac{(2k-1)\pi}{2p+1}}{\tan^2\frac{(2k-1)\pi}{2(2p+1)}} = \frac{(2p+1)^2}{2} - \frac{(2p+1)(4n+1)}{2^{2n+1}}\binom{2n}{n}.$$

Finally, using the relations (2.14) and (3.3), and the notations $w = 2p$, $l = 2n$, the next corollary follows:

Corollary 3.3. *For any positive integers n, p such that $p > n$ we have:*

$$\sum_{k=1}^{p} \frac{\cos^{2n}\frac{(2k-1)\pi}{2p}}{\sin^2\frac{(2k-1)\pi}{2p}} = p^2 - \frac{np}{2^{2n-1}}\binom{2n}{n}.$$

References

[1] Bapat, R.B.: *Graphs and Matrices*, Springer, 2010

[2] Cowel, S.R., Hoară, S., Beiu, V.: Experimenting with beta distributions for approximating hammocks' reliability. In: Intelligent Methods in Computing, Communications and Control (ICCCC2020), AISC, vol. 1243, pp. 70–81. Springer, Cham, Swizerland (2021).

[3] Biggs, N.: *Algebraic Graph Theory*, Cambridge University Press, London, 1974.

[4] Cristescu, G., Drăgoi, V.F.: Efficient approximation of two-terminal networks reliability polynomials using cubic splines. IEEE Trans. Reliab., **70**(3), 1193-1203 (2021).

[5] da Fonseca, C.M., Glasser, M.L., Kowalenko, V.: Basic trigonometric power sums with applications. Ramanujan Journal, **42**, 401-428 (2017).

[6] Dăneţ, R.-M.; Dilimoţ- Niţă, S.-V.; Popescu, M.-V.: Mathematics in Physics of Buildings (1) Harmonic Series, Euler's Constant, Recurrence Sequenc., A XV-a Conferinţă cu participare internaţională, "Confort, Eficienţă, Conservarea energiei şi Protecţia mediului" (CECEPM - XV), Technical University of Civil Engineering of Bucharest, Faculty of Building Services, 26-27 November 2008.
[7] Dăneţ, R.-M.; Dilimoţ- Niţă, S.-V.; Popescu, M.-V.: Sequences and series applied in the physics for buildings. Mathematical Modelling in Civil Engineering Scientific Bulletin of the Technical University of Civil Engineering Bucharest **5**(3), 36-50 (2009).
[8] Dăuş, L., Beiu, V., Cowell, S., Poulin, P.: Brick-wall lattice paths and applications. Tech. Rep. arXiv, 17 Apr. 2018, http://arxiv.org/abs/1804.05277
[9] Dăuş, L., Jianu, M.: Full Hermite interpolation of the reliability of a hammock network. Appl. Anal. Discrete Math., **14**, 198–220 (2020).
[10] Dăuş, L., Jianu, M.: The shape of the reliability polynomial of a hammock network. In: Intelligent Methods in Computing, Communications and Control (ICCCC2020), AISC, vol. 1243, pp. 93–105. Springer, Cham, Switzerland (2021).
[11] Jianu, M., Ciuiu, D., Dăuş, L., Jianu, M.: Markov chain method for computing the reliability of hammock networks. Probab. Eng. Inf. Sci., **36**(2), 276–293 (2022).
[12] Moore, E.F. and Shannon, C.E.: Reliable circuits using less reliable relays Part I, *J. Frankl. Inst.* **262**(3): 191-208, (1956).
[13] Stanley, R.: Enumerative Combinatorics, Volume 1 (2 ed.). Cambridge University Press, New York, (2012).
[14] Stanley, R.: Catalan Numbers. Cambridge University Press, New York, (2015).

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: marilena.jianu@utcb.ro

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: leonard.daus@utcb.ro

A NOTE ON PUISEUX SERIES

GHIOCEL GROZA

Abstract. By theorem on Puiseux series it follows that an algebraic function is represented as a Puiseux series which converges on a suitable domain. The goal of this note is to introduce, in special cases, an algorithm to find the coefficients of the Puiseux series associated to an algebraic function. The method is described by means of an example. Finally, some connections with fractional calculus are given.

Mathematics Subject Classification (2010): 14H05, 39B32, 26A33, 30B10.
Key words: Algebraic functions, Puiseaux series, Caputo derivatives.

1. Puiseaux series and algebraic function

Let m be a positive integer. A formal power series of the form $\sum\limits_{n=n_0}^{+\infty} a_n x^{n/m}$, where $x^{1/m}$ is a formal symbol and n_0 is an integer, is called a *Puiseux series* (see, for example, [3], [4] or [7]). If $F(X,Y) \in \mathbb{C}[X,Y]$ is a polynomial, then a function $y = y(x)$, $y : I \subset \mathbb{C} \to \mathbb{C}$, such that $F(x, y(x)) = 0$, for every $x \in I$, is called an *algebraic function* (see, for instance, [3] or [7]).

The theorem on Puiseux series shows that an algebraic function is represented as a Puiseux series which converges on a suitable domain. A special case of this theorem is given below.

Theorem 1.1. *(see* [4]*, p. 126) Let* $F(X,Y) = Y^n + A_1(X)Y^{n-1} + ... + A_{n-1}(X)Y + A_n(X) \in \mathbb{C}[X,Y]$ *be a polynomial such that* $F(0,0) = 0$*. Then there exist a natural number* $m \geq 1$ *and a series* $\phi \in \mathbb{C}[[T]]$ *such that* $\phi(0) = 0$*, the series converges in a neighborhood of the origin and*

$$F(T^m, \phi(T)) = 0 \text{ in } \mathbb{C}[[T]].$$

We introduce, in special cases, an algorithm to find the coefficients of the Puiseux series associated to an algebraic function. The details of the method is described by using a particular case.

Consider, as in [7] p. 20,

$$F(X,Y) = Y^3 + XY^2 - (X + X^2)Y + X^2 + 2X^3.$$

By Theorem 1.1 there exist a positive integer m, a positive real number B and Puiseaux series of the form

$$y(X) = X^{j_0/m} \sum_{j=0}^{+\infty} a_j X^{j/m},\ j_0 \geq 1,\ a_0 \neq 0, \tag{1.1}$$

such that, for every $x \in [0, B)$,

$$F(x, y(x)) = 0. \tag{1.2}$$

By (1.1), for every $x \in [0, B)$ it follows that

$$y^2(x) = x^{2j_0/m} \sum_{j=0}^{+\infty} b_j x^{j/m},\ y^3(x) = x^{3j_0/m} \sum_{j=0}^{+\infty} c_j x^{j/m}, \tag{1.3}$$

where

$$b_j = \sum_{\substack{k_1+k_2=j \\ k_1,k_2 \geq 0}} a_{k_1} a_{k_2},\ c_j = \sum_{\substack{k_1+k_2+k_3=j \\ k_1,k_2,k_3 \geq 0}} a_{k_1} a_{k_2} a_{k_3} = \sum_{k=0}^{j} a_k b_{j-k} \tag{1.4}$$

By substituting (1.1) and (1.3) into (1.2), for every j, we get

$$E_j(j_0, m) = c_{j-3j_0} + b_{j-2j_0-m} - a_{j-m-j_0} - a_{j-2m-j_0} + \delta_{2m,j} + 2\delta_{3m,j} = 0, \tag{1.5}$$

where $\delta_{i,j}$ is the Kronecker symbol and the coefficients a_k, b_k and c_k are equal to zero, if $k < 0$. Hence, for $j \leq 1$, $E_j(j_0, m) = 0$, for every j_0 and m. Then $E_2(j_0, m) = -a_{2-m-j_0} + \delta_{2m,2} = 0$, which implies $m = 1 = j_0$ and $a_0 = 1$. Similarly, for $m > 1$, it follows that $E_3(j_0, m) = c_{3-3j_0} - a_{3-m-j_0} = 0$. Hence we find $j_0 = 1$, $m = 2$ and $a_0 = c_0 = a_0^3$. Thus $a_0 = \pm 1$ and there are three cases to consider. (A) $m = 1, j_0 = 1, a_0 = 1$. (B) $m = 2, j_0 = 1, a_0 = 1$. (C) $m = 2, j_0 = 1, a_0 = -1$.

In case A, by (1.5), we get the recurrence formula

$$a_n = c_{n-1} + b_{n-1} - a_{n-1} + \delta_{2,n+2} + 2\delta_{3m,n+2},\ n \geq 1.$$

Hence, by (1.4), it follows that $a_1 = a_0^3 + a_0^2 - a_0 + 2 = 3$, $a_2 = 3a_0^2 a_1 + 2a_0 a_1 - 3 = 12$, $a_3 = 84$, and so on.Thus there exists a positive constant B_1 such that, for every $x \in [0, B_1)$, $y_1(x) := x(1 + 3x + 12x^2 + 84x^3 + 651x^4 + 5520x^5 + 49332x^6 + 457977x^7 + 4372716x^8 + 42666672x^9 + ...)$ is a solution of Equation 1.2.

In case B, we obtain

$$c_{j-3} + b_{j-4} - a_{j-3} - a_{j-5} + \delta_{4,j} + 2\delta_{6,j} = 0.$$

Since

$$c_k = 3a_0^2 a_k + 3a_0 b_k^* + c_k^*,$$

where

$$b_k^* = \sum_{\substack{k_1+k_2=k \\ k_1,k_2 \geq 1}} a_{k_1} a_{k_2};\ c_k^* = \sum_{r=1}^{k-1} a_r b_{k-r}^*$$

it follows the recurrence formula

$$a_n = \frac{1}{3a_0^2 - 1}(-3a_0 b_n^* - c_n^* - b_{n-1} + a_{n-2} - \delta_{4,n+3} - 2\delta_{6,n+3}),\ n \geq 1. \tag{1.6}$$

FIGURE 1. The solutions of Equation 1.2 in a neighborhood of the origin

By (1.6) we find $a_1 = \frac{1}{2}(-b_0 - 1) = -1$, $a_2 = \frac{1}{2}(-3b_2^* - b_1 + a_0) = \frac{1}{2}(-3 + 2 + 1) = 0$, $a_3 = \frac{1}{2}(-3b_3^* - c_3^* - b_2 + a_1 - 2) = \frac{0+1-1-1-2}{2} = -1.5$ and so on. Thus there exists a positive constant B_2 such that, for every $x \in [0, B_2)$, $y_2(x) := x^{0.5}(1 - x^{0.5} - 1.5x^{1.5} - 3x^2 - 6x^{2.5} - 15.375x^3 - 42x^{3.5} - 115.5x^4 - 325.5x^{4.5} + ...)$ is a solution of the equation (1.2).

Similarly, in case C, we obtain the solution $y_3(x) := -x^{0.5}(1 + x^{0.5} + 1.5x^{1.5} - 3x^2 + 6x^{2.5} - 15.375x^3 + 42x^{3.5} - 115.5x^4 + 325.5x^{4.5} + ...)$. The solutions are represented in Figure 1.

2. Elements of fractional calculus

A list of books which cover a wide variety of topics from fractional calculus contains [2], [8], [9] and [11]. We start with some definitions.

Definition 2.1. *The Riemann - Liouville fractional integral operator of order $\alpha \geq 0$ of a real function $f \in C_{-1}(a, +\infty)$ (i.e. there exist $\rho > -1$ and a continuous function $g : [a, +\infty) \to \mathbb{R}$ such that $f(x) = (x-a)^\rho g(x)$, for every $x \in (a, +\infty)$) is defined as*

$$J_a^\alpha f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)} \displaystyle\int_a^x (x-s)^{\alpha-1} f(s) ds\,, & \alpha > 0, \\ f(x)\,, & \alpha = 0, \end{cases}$$

for $x \geq a$.

The Riemann - Liouville integral operator has the following properties:

$$J^\alpha \left(\lambda_1 f_1(x) + \lambda_2 f_2(x)\right) = \lambda_1 J^\alpha f_1(x) + \lambda_2 J^\alpha f_2(x)\,,\ \forall \lambda_1, \lambda_2 \text{ constants},$$

$$J^\alpha J^\beta f(x) = J^{\alpha+\beta} f(x), \quad \forall \alpha, \beta \geq 0.$$

Definition 2.2. *Let $n = \lceil \alpha \rceil$ (where $\lceil x \rceil = \min\{z \in \mathbb{Z} : z \geq x\}$ denotes the ceiling function). The Caputo fractional derivative of order $\alpha \geq 0$ of a function $f \in C_{-1}^{(n)}([a, +\infty))$ (that is $f \in C^{(n-1)}([a, +\infty)$ and $f^{(n)} \in C_{-1}([a, +\infty))$) is defined as*

$$D_a^\alpha f(x) = J^{n-\alpha} f^{(n)}(x) = \begin{cases} \dfrac{1}{\Gamma(n-\alpha)} \displaystyle\int_a^x \frac{f^{(n)}(s)}{(x-s)^{1+\alpha-n}} ds\,, & \alpha \notin \mathbb{N} \\ f^{(\alpha)}(x)\,, & \alpha \in \mathbb{N} \end{cases}.$$

Example 2.3. Let $\beta \geq 0$ and $f(x) = (x-a)^\beta$. Then

$$D_a^\alpha f(x) = \begin{cases} 0, & \text{if } \beta \in \{0, ..., n-1\} \\ \frac{\Gamma(\beta+1)}{\Gamma(\beta+1-\alpha)}(x-a)^{\beta-\alpha} & \text{if } \beta \in \mathbb{N} \text{ and } \beta \geq n \\ & \text{or } \beta \notin \mathbb{N} \text{ and } \beta > n-1, \end{cases} \tag{2.1}$$

where $n = \lceil \alpha \rceil$.

A series of the form

$$\sum_{n=0}^{+\infty} a_n(x-x_0)^{n\alpha}, x \geq x_0,$$

where a_n are real numbers, $\alpha > 0$, $x, x_0 \in \mathbb{R}$ is called an *α-fractional* power series at x_0. A real function f is said to be *representable as an α - fractional Taylor series on $I = [x_0, b)$* at x_0, if it can be written as the sum of a fractional power series:

$$f(x) = \sum_{n\geq 0} a_n(x-x_0)^{n\alpha}, \forall x \in I. \tag{2.2}$$

If $I = [x_0, b]$, then f is said to be representable as an α - fractional Taylor series on I if there exists $b' > b$ and f is representable into an α - fractional Taylor series on $[x_0, b')$

For $I = [a, b)$ or $I = [a, b]$, denote by $\mathcal{T}_{a,\alpha}(I)$ the set of all functions $f : I \to \mathbb{R}$ which are representable into an α- fractional Taylor series on I if, for every $x_0 \in I$, f is representable as an α - fractional Taylor series on $I = [x_0, b)$ at x_0.

For any $\alpha > 0$ and $n \in \mathbb{N}$, we denote by $D_a^{n\circ\alpha} f = \underbrace{D_a^\alpha \circ D_a^\alpha \circ \ldots \circ D_a^\alpha}_{n-times} f$. If $f \in \mathcal{T}_{a,\alpha}(I)$, by (2.1) , it follows that $D_a^{n\circ\alpha} f = D_a^{\alpha_1} f$, where $\alpha_1 = n\alpha$. It is known the following result (see [1], [5], [6] and [10]):

Theorem 2.4. *Let $\sum_{n\geq 0} a_n t^{n\alpha}$, $\alpha \in (0, 1]$ be a fractional power series, let $R \geq 0$ be the radius of convergence of the power series $\sum_{n\geq 0} a_n x^n$ and $r = \begin{cases} R^{\frac{1}{\alpha}}, \text{ if } R < \infty \\ \infty, \text{ if } R = \infty . \end{cases}$ Then*
i) If $R > 0$, then the series $\sum_{n\geq 0} a_n t^{n\alpha}$ converges absolutely and uniformly on $[0, b]$, for all $b \in (0, r)$ and there exists a positive integer $N(b)$ such that

$$|a_n| \leq b^{-n\alpha}, \text{ for all } n \geq N(b);$$

ii) If $f : [0, r) \to \mathbb{R}$ is the sum of the fractional power series, $f(t) = \sum_{n\geq 0} a_n t^{n\alpha}$, $\forall t \in [0, r)$, then f is continuous and there exists the Caputo derivative, $D^\alpha f : [0, r) \to \mathbb{R}$. Moreover, the series of the Caputo derivatives, $\sum_{n\geq 0} a_n D^\alpha (t^{n\alpha}) = \sum_{n\geq 1} a_n \frac{\Gamma(n\alpha+1)}{\Gamma((n-1)\alpha+1)} t^{(n-1)\alpha}$ converges absolutely and uniformly on $[0, b]$, $\forall b \in (0, r)$ and

$$D^\alpha f(t) = \sum_{n\geq 0} a_n D^\alpha (t^{n\alpha}) = \sum_{n\geq 1} a_n \frac{\Gamma(n\alpha+1)}{\Gamma((n-1)\alpha+1)} t^{(n-1)\alpha}, \forall t \in [0, r).$$

iii) The coefficients of the α - fractional Taylor series (2.2) are given by

$$a_n = \frac{D^{n\alpha} f(x_0)}{\Gamma(n\alpha+1)}.$$

Leibnitz' formula in the fractional setting is more complicated than in the classical case (see, for example, [2], p. 33). However, in the case when the functions are representable as an α-fractional Taylor series it follows a simpler form.

If $n, k_1, ..., k_s$ are nonnegative integers and $k_1+k_2+...+k_s = n$, denote $\left(\begin{array}{c} n \\ k_1 \ ... \ k_s \end{array} \right)_\alpha := \frac{\Gamma(n\alpha+1)}{\Gamma(k_1\alpha+1)...\Gamma(k_s\alpha+1)}$.

Theorem 2.5. *(Leibniz' formula for Caputo derivaives). Assume that $f, g \in \mathcal{T}_{a,\alpha}(I)$. Then,*

$$(D_a^{k\alpha} fg)(a) = \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \ \ k-j \end{array} \right)_\alpha (D_a^{j\alpha} f)(a)(D_a^{(k-j)\alpha} g)(a).$$

Proof. It is enough to consider $a = 0$. Then, by Theorem 2.4, we may write

$$f(t) = \sum_{k=0}^{\infty} a_k t^{k\alpha}, \ g(t) = \sum_{k=0}^{\infty} b_k t^{k\alpha}.$$

Hence

$$f(t)g(t) = \sum_{k=0}^{\infty} c_k t^{k\alpha}, \text{ where } c_k = \sum_{j=0}^{k} a_j b_{k-j}.$$

Since, by Theorem 2.4, $a_k = \frac{D_0^{k\alpha} f(0)}{\Gamma(k\alpha+1)}$, we find

$$c_k = \frac{(D_0^{k\alpha} fg)(0)}{\Gamma(k\alpha+1)} = \sum_{j=0}^{k} \frac{(D_0^{j\alpha} f)(0)}{\Gamma(j\alpha+1)} \frac{(D_0^{(k-j)\alpha} g)(0)}{\Gamma((k-j)\alpha+1)}.$$

Hence we get

$$(D^{k\alpha} fg)(0) = \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \ \ k-j \end{array} \right)_\alpha (D_0^{j\alpha} f)(0)(D_0^{(k-j)\alpha} g)(0),$$

which implies the theorem. □

Remark 2.6. By Theorem 2.5 we find

$$(D_a^{k\alpha} f^2)(a) = \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \ \ k-j \end{array} \right)_\alpha (D_a^{j\alpha} f)(a)(D_a^{(k-j)\alpha} f)(a).$$

and

$$(D_a^{k\alpha} f^3)(a) = \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \ \ k-j \end{array} \right)_\alpha (D_a^{j\alpha} f^2)(a)(D_a^{(k-j)\alpha} f)(a)$$

$$= \sum_{j=0}^{k} \left(\begin{array}{c} k \\ j \ \ k-j \end{array} \right)_\alpha \sum_{i=0}^{j} \left(\begin{array}{c} j \\ i \ \ j-i \end{array} \right)_\alpha (D_a^{i\alpha} f)(a)(D_a^{(j-i)\alpha} f)(a)(D_a^{(k-j)\alpha} f)(a)$$

$$= \sum_{j=0}^{k} \sum_{i=0}^{j} \left(\begin{array}{c} k \\ i \ \ j-i \ \ k-j \end{array} \right)_\alpha (D_a^{i\alpha} f)(a)(D_a^{(j-i)\alpha} f)(a) D_a^{(k-j)\alpha} f)(a).$$

Thus by computing the Caputo derivatives in Equation 1.2 and by using Theorem 2.4 we find a system with respect to the variables a_i which implies the recurrence formula obtained in Section 1.

References

[1] El-Ajou, A., Arqub, O. A., Zhour, Z. A. and Momani, S.: New results on fractional power series: theories and applications, *Entropy* **15** 298 (2013), 5305–5323.

[2] Diethelm, K.: *The analysis of fractional differential equations. An application-oriented exposition using differential operators of Caputo type*, Springer, Berlin, 2010.

[3] Eichler, M.: *Introduction to the theory of algebraic numbers and functions*, Academic Press New York, 1966.

[4] Fischer, G.: *Plane algebraic curves*, American Mathematical Society, 2001.

[5] Groza, G. and Jianu, M.: Functions represented into fractional Taylor series, in *ITM Web of Conference* **29**, 010117 (2019), https://doi.org/10.1051/itmconf/ 20192901017.

[6] Groza, G., Jianu, M. and Mierluş-Mazilu, I.: Some properties of the functions representable as fractional power series, *Mathematics* **12**, 961(2024) https://doi.org/ 10.3390/math12070961.

[7] Jung, H. W. E.: *Einführung in die theorie der algebraischen funktionen einer veränderlichen*, Walter de Gruyter & Co, Berlin, 1925.

[8] Kilbas, A. A., Srivastava, H. M. and Trujillo, J. J.: *Theory and applications of fractional differential equations*, Elsevier, Amsterdam, 2006.

[9] Miller, K. S. and Ross, B.: *An introduction to the fractional calculus and fractional differential equations*, John Willey&Sons, New York, 1993.

[10] Odibat, Z. and Shawagfeh, N.: Generalized Taylor's formula, *Appl. Math. Comput.* **186**(1) (2007), 286-293.

[11] Podlubny, I.: *Fractional differential equations*, Academic Press, New York, 1999.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania

Email address: grozagc@yahoo.com

ADVANCING STEM EDUCATION: DIGITAL TOOLS AND PEDAGOGICAL TRAINING FOR HEI EDUCATORS (DIGISTEM)

Ion Mierluș Mazilu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei 124, sector 2, 38RO-020396 Bucharest, Romania
E-mail: ion.mierlusmazilu@utcb.ro
Ștefania Constantinescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: stefania.constantinescu@utcb.ro
Alice Anghelescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: alice.anghelescu@utcb.ro

Abstract: The integration of digital tools and effective pedagogical training in STEM education is crucial for enhancing the teaching capabilities of higher education institution (HEI) educators. This research paper, titled "Advancing STEM Education: Digital Tools and Pedagogical Training for HEI Educators," addresses the necessity for developing comprehensive methodologies and platforms to support educators in adopting digital resources and improving pedagogical practices. The study begins with a thorough educational research work and needs analysis, identifying current gaps and challenges faced by STEM educators in digital competence and pedagogical training. Based on these findings, the DigiSTEM Methodology (PR1) is developed, incorporating best practices for integrating digital tools into STEM education and refining pedagogical approaches. To facilitate the practical application of this methodology, the DigiSTEM Platform (PR2) is established, offering a user-friendly interface for educators to access, implement, and share digital resources. Guidelines for using digital tools and resources in STEM education (PR3) are provided, ensuring educators can effectively navigate and utilize these resources to enhance teaching and learning outcomes.

Furthermore, the development of a Massive Open Online Course (MOOC) offers scalable training opportunities, equipping educators with the necessary skills and knowledge to leverage digital tools in their STEM curricula. This MOOC is designed to provide comprehensive training, from basic digital literacy to advanced pedagogical strategies. The research employs a mixed-methods approach, including surveys, interviews, and observations, to gather data from a diverse group of STEM educators and stakeholders. The results highlight significant improvements in digital and pedagogical competence among participants, demonstrating the effectiveness of the DigiSTEM initiative.

Overall, this paper contributes to the advancement of STEM education by providing a structured framework and practical tools for enhancing digital and pedagogical skills among HEI educators. It offers actionable insights and recommendations for HEIs to support ongoing professional development and foster a culture of innovation in STEM teaching.

Mathematics Subject Classification (2010): 97B40, 97C70, 97D40, 68N99.

Key words: STEM Education, Digital Tools, Educational Technology, Online Learning Platforms, MOOC (Massive Open Online Course), Digital Resources, Instructional Design, Innovative Teaching Practices.

1. Introduction

STEM education is undoubtedly crucial in equipping students with the necessary skills to succeed in an increasingly technology-driven society. The amalgamation of Science, Technology, Engineering, and Mathematics not only cultivates critical thinking and problem-solving abilities but also fosters innovation and creativity. However, the effectiveness of STEM education hinges on the proficiency of Higher Education Institution (HEI) educators in integrating digital tools and employing pedagogical strategies adeptly.

In recent years, the landscape of education has undergone a significant transformation with the advent of digital technologies. These technologies offer unprecedented opportunities to enhance teaching and learning experiences. From interactive simulations to immersive virtual reality environments, digital tools have the potential to make complex concepts more accessible and engaging for students. However, realizing this potential requires educators who are not only adept at utilizing these tools but also proficient in pedagogical techniques that promote effective learning.

Unfortunately, many HEI educators face challenges in both realms. While some may lack the technical expertise to leverage digital tools effectively, others may struggle with implementing pedagogical approaches that foster deeper understanding and retention of STEM concepts. As a result, there is a pressing need for comprehensive methodologies and platforms that can support educators in bridging these gaps and enhancing their teaching capabilities.

Enter DigiSTEM—a holistic initiative designed to empower HEI educators with the necessary skills and tools to excel in STEM education. At its core, DigiSTEM recognizes the symbiotic relationship between digital literacy and pedagogical proficiency. By integrating best practices from both domains, DigiSTEM aims to equip educators with a multifaceted approach to teaching STEM subjects.

This paper embarks on a journey to explore the necessity for such methodologies and platforms in advancing STEM education. Through a thorough examination of existing challenges and gaps faced by HEI educators, we aim to underscore the critical role that DigiSTEM can play in addressing these issues. By providing educators with practical resources, guidelines, and training opportunities, DigiSTEM seeks to foster a culture of innovation and excellence in STEM teaching.

In the pages that follow, we will delve deeper into the components of the DigiSTEM initiative, examining how the DigiSTEM Methodology, DigiSTEM Platform, and guidelines for using digital tools can revolutionize STEM education. Furthermore, we will explore the development and implementation of a Massive Open Online Course (MOOC) tailored to the needs of HEI educators, offering scalable training opportunities to enhance their digital and pedagogical skills.

Ultimately, this paper aims to shed light on the transformative potential of DigiSTEM in advancing STEM education. By providing educators with the necessary support and resources, we can ensure that future generations are equipped with the skills and knowledge needed to thrive in a rapidly evolving world. Through collaboration and innovation, we can pave the way for a brighter future—one where STEM education transcends boundaries and empowers individuals to reach their full potential.

2. Background and Literature Review

In the dynamic landscape of STEM education, the integration of digital tools and pedagogical training stands as a pivotal challenge for educators worldwide. A comprehensive review of existing literature sheds light on the multifaceted nature of this challenge, highlighting the gaps and deficiencies that hinder the effective implementation of digital technologies and pedagogical strategies in STEM classrooms.

Digital tools hold immense promise for revolutionizing the learning experience, offering interactive simulations, multimedia resources, and collaborative platforms that can engage students in ways previously unimaginable. However, despite the abundance of these resources, many educators find themselves ill-equipped to harness their full potential. Research indicates that a significant proportion of STEM educators lack the requisite digital competence to navigate and integrate these tools seamlessly into their teaching practices [1]. Without adequate training and support, educators may struggle to leverage digital technologies effectively, limiting their ability to create engaging and immersive learning experiences for their students.

Moreover, the issue extends beyond mere technical proficiency to encompass pedagogical training—a domain where STEM education often falls short. While educators may possess a deep understanding of their respective STEM disciplines, they may not necessarily possess the pedagogical skills needed to translate this knowledge into effective teaching strategies. Traditional pedagogical approaches, which prioritize rote memorization and passive learning, are ill-suited to the dynamic and interdisciplinary nature of STEM subjects [2]. As a result, students may disengage from the learning process, leading to suboptimal learning outcomes and a failure to cultivate the critical thinking and problem-solving skills that are essential for success in STEM fields.

Furthermore, the intersection of digital competence and pedagogical training presents unique challenges that demand innovative solutions. Educators must not only be proficient in using digital tools but also adept at integrating them seamlessly into their instructional practices in a manner that enhances rather than detracts from the learning experience. This requires a nuanced understanding of pedagogical principles and instructional design, coupled with hands-on experience in leveraging digital technologies to achieve educational objectives [3].

In light of these challenges, there is a pressing need for comprehensive professional development initiatives that address the dual requirements of digital competence and pedagogical training in STEM education. Such initiatives must provide educators with the knowledge, skills, and resources needed to navigate the digital landscape effectively while fostering innovative pedagogical approaches that promote active learning and student engagement.

The literature review underscores the critical importance of addressing the challenges faced by STEM educators in digital competence and pedagogical training. By understanding the multifaceted nature of these challenges and their implications for teaching and learning, stakeholders can work towards developing holistic solutions that empower educators to thrive in the digital age and foster a culture of innovation in STEM education.

3. Methodology

The methodology section of this paper outlines a robust approach to understanding the challenges faced by STEM educators regarding digital competence and pedagogical training. By employing a mixed-methods approach, which includes surveys, interviews, and observations, this study seeks to conduct a comprehensive needs analysis, thereby laying the groundwork for the development of the DigiSTEM Methodology.

Surveys serve as a valuable tool for gathering quantitative data from a large sample of STEM educators. By administering surveys, researchers can assess the current state of digital competence and pedagogical training among educators, as well as identify specific areas of strength and weakness. Surveys offer a standardized means of collecting data, allowing for comparisons across different demographic groups and educational contexts.

In addition to surveys, interviews provide an opportunity to delve deeper into the experiences and perspectives of individual educators. Through semi-structured interviews, researchers can explore the nuances of educators' challenges and aspirations, gaining insights that may not be

captured through quantitative measures alone. Interviews allow for rich, detailed narratives that illuminate the complexities of educators' professional practices and beliefs.

Complementing surveys and interviews, observations offer a window into educators' classroom practices and interactions with digital technologies. By observing educators in action, researchers can gain firsthand insights into how digital tools are integrated into teaching and learning activities. Observations provide valuable context for interpreting survey and interview data, helping researchers to understand the real-world implications of educators' digital competence and pedagogical approaches.

Drawing on the data collected through surveys, interviews, and observations, researchers can identify common themes and patterns, as well as areas of divergence and discrepancy. These findings form the basis for the development of the DigiSTEM Methodology, which seeks to integrate best practices for incorporating digital tools into STEM education and refining pedagogical approaches. By synthesizing insights from across the mixed-methods analysis, the DigiSTEM Methodology aims to provide educators with a structured framework for enhancing their teaching practices in the digital age.

Tthe methodology employed in this study offers a rigorous and comprehensive approach to understanding the challenges faced by STEM educators and developing solutions to address them. By combining surveys, interviews, and observations, researchers can gain a multifaceted understanding of educators' needs and experiences, informing the development of the DigiSTEM Methodology and ultimately contributing to the advancement of STEM education.

4. The DigiSTEM Initiative

The DigiSTEM Initiative represents a transformative approach to addressing the challenges faced by STEM educators in integrating digital tools and refining pedagogical practices. Comprising three core components— the DigiSTEM Methodology (PR1), the DigiSTEM Platform (PR2), and guidelines for using digital tools and resources in STEM education (PR3)— this initiative offers a comprehensive framework and practical tools to enhance teaching and learning outcomes in STEM classrooms.

At the heart of the DigiSTEM Initiative lies the DigiSTEM Methodology (PR1), a structured framework that integrates best practices for incorporating digital tools into STEM education and refining pedagogical approaches. Drawing on insights from the needs analysis conducted among STEM educators, the DigiSTEM Methodology provides a roadmap for educators to enhance their digital competence and pedagogical skills. By offering a systematic approach to leveraging digital technologies in teaching and learning, the DigiSTEM Methodology empowers educators to create engaging and interactive learning experiences that foster deeper understanding and retention of STEM concepts.

In tandem with the DigiSTEM Methodology, the DigiSTEM Platform (PR2) serves as a user-friendly interface for accessing and implementing digital resources. Designed with educators' needs in mind, the DigiSTEM Platform offers a curated collection of digital tools and resources tailored to STEM education. From interactive simulations to multimedia presentations, educators can easily explore and integrate a wide range of digital resources into their instructional practices. Moreover, the DigiSTEM Platform facilitates collaboration and knowledge sharing among educators, enabling them to exchange ideas and best practices for incorporating digital technologies into their teaching.

To complement the DigiSTEM Methodology and Platform, guidelines for using digital tools and resources in STEM education (PR3) provide educators with practical recommendations and strategies for effectively navigating and utilizing these tools. These guidelines offer insights into how educators can leverage digital technologies to enhance teaching and learning outcomes, from designing engaging learning activities to assessing student progress. By

offering clear and actionable guidance, the guidelines ensure that educators can make the most of the digital resources available to them, maximizing the impact of their instructional practices.

Collectively, the components of the DigiSTEM Initiative— the DigiSTEM Methodology, Platform, and guidelines— represent a holistic approach to advancing STEM education. By providing educators with a structured framework and practical tools, the DigiSTEM Initiative empowers them to enhance their teaching practices and cultivate the critical thinking and problem-solving skills needed for success in the 21st century. As educators embrace digital technologies and innovative pedagogical approaches, the DigiSTEM Initiative holds the promise of transforming STEM education and preparing students to thrive in an increasingly digital world.

5. Implementation and Evaluation

The implementation and evaluation of the DigiSTEM initiative mark a critical phase in its journey towards enhancing STEM education. Through the development of a Massive Open Online Course (MOOC) tailored to educators' varying skill levels, the initiative aims to equip educators with the necessary skills to effectively integrate digital tools into their STEM curricula. Furthermore, the effectiveness of the initiative is rigorously evaluated through surveys, interviews, and observations, providing valuable insights into the impact of DigiSTEM on educators' digital and pedagogical competence.

The development of the MOOC represents a pivotal step in the implementation of the DigiSTEM initiative. By offering a flexible and accessible learning platform, the MOOC caters to educators' diverse needs and preferences. From foundational concepts of digital literacy to advanced pedagogical strategies, the MOOC covers a wide range of topics designed to empower educators with the knowledge and skills needed to leverage digital tools effectively. Through interactive modules, engaging activities, and real-world examples, educators are provided with practical guidance and resources to enhance their teaching practices in STEM subjects.

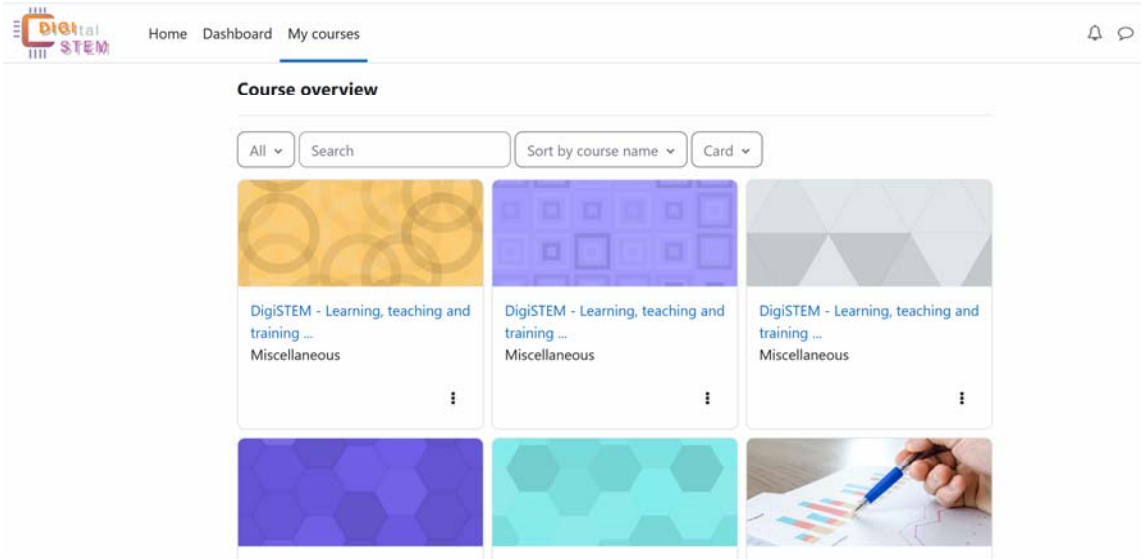
Moreover, the MOOC serves as a catalyst for fostering a community of practice among STEM educators. By facilitating collaboration and knowledge sharing, educators can learn from each other's experiences and exchange best practices for incorporating digital technologies into their teaching. This collaborative approach not only enhances the learning experience but also promotes ongoing professional development and innovation in STEM education.

In addition to the development of the MOOC, the effectiveness of the DigiSTEM initiative is rigorously evaluated through a variety of methods, including surveys, interviews, and observations. These evaluation methods provide valuable insights into the impact of DigiSTEM on educators' digital and pedagogical competence, as well as their perceptions of the initiative's effectiveness. By collecting data from a diverse group of participants, researchers can assess the extent to which DigiSTEM has succeeded in achieving its objectives and identify areas for improvement.

The results of the evaluation demonstrate significant improvements in digital and pedagogical competence among participants, highlighting the effectiveness of the DigiSTEM initiative in addressing the challenges faced by STEM educators. Educators report feeling more confident in their ability to integrate digital tools into their teaching practices and employ innovative pedagogical strategies to enhance student learning outcomes. Furthermore, observations of educators' classroom practices reveal tangible evidence of the impact of DigiSTEM on teaching and learning, underscoring the transformative potential of the initiative.

The implementation and evaluation of the DigiSTEM initiative represent a significant milestone in the advancement of STEM education. By providing educators with the necessary

skills and support to leverage digital tools effectively, DigiSTEM is poised to revolutionize teaching and learning in STEM subjects. Through ongoing evaluation and refinement, the initiative will continue to evolve and adapt to meet the evolving needs of educators and students alike, ensuring that STEM education remains at the forefront of innovation and excellence.



Platform Moodle Digistem EU

6. Conclusions

In the ever-evolving landscape of STEM education, the DigiSTEM initiative emerges as a beacon of innovation and excellence, offering educators a transformative approach to enhancing their digital and pedagogical skills. Through a structured framework and practical tools, DigiSTEM empowers HEI educators to navigate the complexities of the digital age and cultivate engaging and immersive learning experiences for their students. As we draw this journey to a close, it is essential to reflect on the profound impact of DigiSTEM and consider recommendations for supporting ongoing professional development and fostering a culture of innovation in STEM teaching.

The DigiSTEM initiative represents a significant step forward in addressing the challenges faced by HEI educators in integrating digital tools and refining pedagogical practices. By providing educators with a comprehensive methodology, user-friendly platform, and practical guidelines, DigiSTEM equips them with the knowledge and skills needed to excel in STEM education. From interactive simulations to collaborative projects, educators are empowered to create dynamic learning environments that inspire curiosity, creativity, and critical thinking among their students.

Furthermore, the effectiveness of the DigiSTEM initiative is underscored by the significant improvements in digital and pedagogical competence observed among participants. Through rigorous evaluation methods, including surveys, interviews, and observations, researchers have documented the transformative impact of DigiSTEM on educators' teaching practices and student learning outcomes. These findings serve as a testament to the value of investing in professional development initiatives that prioritize digital literacy and pedagogical excellence in STEM education.

Moving forward, it is imperative for HEIs to prioritize ongoing professional development and foster a culture of innovation in STEM teaching. Recommendations for supporting educators in this endeavor include:

1. Providing access to continued training and resources: HEIs should offer educators access to workshops, seminars, and online courses focused on digital literacy and innovative pedagogical strategies. By investing in professional development opportunities, HEIs can empower educators to stay abreast of emerging trends and technologies in STEM education.
2. Fostering collaboration and knowledge sharing: HEIs should create platforms and networks that facilitate collaboration and knowledge sharing among educators. By fostering a culture of collaboration, HEIs can encourage educators to share best practices, collaborate on research projects, and support each other in their professional growth.
3. Integrating digital literacy into teacher preparation programs: HEIs should integrate digital literacy into teacher preparation programs to ensure that future educators are equipped with the skills and knowledge needed to leverage digital tools effectively. By incorporating digital literacy training into pre-service teacher education, HEIs can prepare educators to meet the challenges of the digital age head-on.

In conclusion, the DigiSTEM initiative represents a paradigm shift in STEM education, offering educators a structured framework and practical tools for enhancing their digital and pedagogical skills. By empowering educators with the necessary skills and support, DigiSTEM contributes to the advancement of STEM education and prepares students for success in the 21st century. As we look to the future, it is essential for HEIs to prioritize ongoing professional development and foster a culture of innovation in STEM teaching, ensuring that educators have the resources and support they need to thrive in an increasingly digital world..

References

[1] Borba, M. C., Askar, P., Engelbrecht, J., Gadanidis, G., Llinares, S. and Aguilar, M. S. *'Blended Learning, E-learning and Mobile Learning in Mathematics Education'*, ZDM Mathematics Education, 48 (2016), 589-610.

[2] Clark, R. E. and Mayer, R. E. *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*. 4th edn. (Wiley, 2016).

[3] http://digistem.eu/index.php/en/

[4] https://learndigisteam.com/

[5] https://stemeducationjournal.springeropen.com/

[6] https://www.ed.gov/stem

[7] https://www.britannica.com/topic/STEM-education

[8] https://www.khanacademy.org

ENERGIZE TOMORROW: SERVICE-LEARNING FOR SUSTAINABLE ENERGY ACCESS (GIRLS – GENERATION FOR INNOVATION, RESILIENCE, LEADERSHIP AND SUSTAINABILITY)

Ion Mierluș Mazilu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei 124, sector 2, 38RO-020396 Bucharest, Romania
E-mail: ion.mierlusmazilu@utcb.ro
Ștefania Constantinescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: stefania.constantinescu@utcb.ro
Alice Anghelescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: alice.anghelescu@utcb.ro

Abstract: This paper explores a service-learning project aimed at empowering high school students with the knowledge and skills necessary to ensure access to affordable, reliable, sustainable, and modern energy for all, in alignment with the United Nations Sustainable Development Goal 7. This project involves university students from the Technical University of Civil Engineering who present a prototype of an energy-independent solar house, integrating surface geothermal energy with classical installations, to inspire and educate high school students on sustainable energy practices.

Through this service-learning initiative, the project bridges academic knowledge and real-world applications, fostering an understanding of the significance and implementation of renewable energy technologies. The high school students are engaged in interactive educational modules that cover the principles of solar and geothermal energy, mathematical modeling of energy systems, and practical activities to calculate and optimize energy efficiency. This hands-on approach not only enhances their comprehension of sustainable energy but also develops their problem-solving and analytical skills.

The paper details the project's design, methodology, and implementation, highlighting the integration of mathematical concepts in the analysis and optimization of energy systems. It also assesses the impact of the project through pre- and post-project surveys, evaluating knowledge gain and skill development among the high school participants. Challenges encountered during the project and the solutions devised are discussed, providing insights for future service-learning initiatives.

The findings underscore the project's success in promoting sustainable energy education and its potential to influence policy and practice in the field of renewable energy. The conclusion offers recommendations for expanding such projects and emphasizes the crucial role of education in advancing global energy sustainability.

Mathematics Subject Classification (2010): 00A71, 97M50.

Key words: Sustainable Energy, Service-Learning, Renewable Energy Education, Energy-Independent House, Mathematical Modeling, Energy Efficiency.

1. Introduction

The imperative to transition to renewable energy sources has never been more urgent. The escalating impacts of climate change and the finite nature of fossil fuels underscore the necessity for sustainable energy solutions. The global community faces a critical juncture where immediate and decisive actions are required to mitigate environmental degradation and ensure energy security for future generations.

Renewable energy sources, such as solar, wind, and geothermal, offer promising alternatives to traditional fossil fuels. These sources are not only abundant and sustainable but also have minimal environmental impact compared to coal, oil, and natural gas. Transitioning to renewable energy can significantly reduce greenhouse gas emissions, decrease air pollution, and promote economic stability by reducing dependence on imported fuels.

However, the shift to a sustainable energy paradigm necessitates widespread public understanding and acceptance. Education plays a pivotal role in this transition, as it equips individuals with the knowledge and skills needed to adopt and advocate for renewable energy technologies. By instilling a deep understanding of sustainable practices in the younger generation, we can cultivate a workforce capable of driving innovation and implementation in the renewable energy sector.

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project is conceived within this context. This initiative seeks to bridge the gap between academic learning and real-world application by engaging high school students in a comprehensive renewable energy education program. Through hands-on activities, interactive modules, and the development of a prototype energy-independent solar house, the project aims to inspire and equip students with the tools necessary to contribute to a sustainable future.

The project is anchored in the belief that education can catalyze change. To this end, it has set forth the following primary objectives:

- Empower high school students with knowledge and skills related to sustainable energy. The foundation of the project is to provide students with a robust understanding of renewable energy sources and technologies. This includes theoretical knowledge about solar and geothermal energy, as well as practical skills in designing, implementing, and optimizing these systems. By demystifying the science behind renewable energy, the project aims to empower students to become informed advocates and practitioners.
- Demonstrate the practical application of renewable energy technologies. Knowledge without application is of limited value. The project emphasizes the practical aspects of renewable energy through the construction and analysis of a prototype energy-independent solar house. This hands-on experience allows students to see firsthand how renewable energy systems operate, how they can be integrated into existing infrastructures, and what challenges must be overcome to optimize their efficiency.
- Foster an understanding of the principles of solar and geothermal energy. Solar and geothermal energy are two of the most promising renewable energy sources. By delving into the principles of these technologies, students gain insights into their mechanisms, advantages, and limitations. Understanding these principles is crucial for developing innovative solutions and for appreciating the broader context of energy sustainability.
- Enhance problem-solving and analytical skills through interactive learning. The project is designed to be interactive and engaging, encouraging students to actively participate in their learning process. Through activities such as energy efficiency calculations, system optimization exercises, and the use of mathematical modeling, students develop critical thinking and analytical skills. These skills are not only essential for understanding and implementing renewable energy technologies but are also broadly applicable across various scientific and engineering disciplines.

Service-learning combines academic learning with community service, creating a mutually beneficial environment where students apply classroom knowledge to address real-world challenges. In the context of the "Energize Tomorrow" project, service-learning serves several vital purposes:

- Bridging Theory and Practice: Students can connect theoretical knowledge with practical application, enhancing their understanding and retention of complex concepts.
- Community Engagement: By working on projects that benefit their communities, students develop a sense of responsibility and civic engagement. They see the direct impact of their work, which can be incredibly motivating and fulfilling.
- Skill Development: Service-learning projects require students to develop a range of skills, from technical abilities related to renewable energy systems to soft skills such as teamwork, communication, and problem-solving.

The project aligns with the United Nations Sustainable Development Goals (SDGs), particularly SDG 7: "Ensure access to affordable, reliable, sustainable, and modern energy for all." By educating high school students about sustainable energy, the project contributes to global efforts to increase the adoption of renewable energy technologies. It also supports SDG 4: "Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all," by providing students with high-quality, relevant education that prepares them for future challenges.

The methodological approach of the "Energize Tomorrow" project involves a blend of educational strategies and practical activities. The project is structured into two main educational modules:

- Introduction to Renewable Energy: This module provides foundational knowledge about solar and geothermal energy, including their scientific principles, technological applications, and environmental benefits. Mathematical modeling is introduced as a tool for analyzing and optimizing energy systems.
- Practical Activities: This module focuses on hands-on learning through the construction of a prototype energy-independent solar house. Students engage in activities such as designing energy systems, calculating energy efficiency, and using optimization techniques to improve system performance.

To assess the impact of the project, pre- and post-project surveys are conducted to evaluate knowledge gain and skill development among the students. This assessment helps identify the strengths and areas for improvement in the project, ensuring its continuous enhancement and effectiveness.

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project represents a forward-thinking approach to education, aiming to equip the next generation with the knowledge and skills necessary to lead the transition to sustainable energy. By integrating academic learning with practical application and community service, the project not only enhances students' understanding of renewable energy but also fosters a sense of responsibility and engagement. As we face the pressing challenges of climate change and energy security, such educational initiatives are crucial for building a sustainable future.

2. Integrating Knowledge and Practice

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project is built on a comprehensive educational framework designed to bridge the gap between theoretical knowledge and practical application. This integrated approach ensures that high school students not only learn about sustainable energy concepts but also apply them in real-world scenarios. The educational modules and practical applications form the core of this initiative, providing a detailed roadmap for fostering a deep understanding of renewable energy technologies.

Introduction to Renewable Energy

The journey begins with a foundational course on renewable energy. This module aims to instill a solid grasp of the basic principles of renewable energy, focusing on two primary sources: solar and geothermal energy.

Solar Energy

Solar energy is one of the most accessible and abundant renewable energy sources. Students learn about the science behind solar power, including how solar panels convert sunlight into electricity through photovoltaic cells. Key concepts such as solar irradiance, efficiency, and energy conversion are discussed.

To make the learning process engaging, students participate in interactive demonstrations that illustrate how solar panels work. They also analyze case studies of solar energy implementations in different parts of the world, understanding both the technological and socio-economic impacts of these projects.

Geothermal Energy

Geothermal energy, derived from the Earth's internal heat, is another significant renewable energy source. This section covers the basic principles of geothermal energy, including the types of geothermal resources (such as hot water springs and geothermal reservoirs) and the technologies used to harness this energy.

Students explore how geothermal heat pumps and geothermal power plants operate. They also examine the environmental benefits and challenges associated with geothermal energy, comparing its efficiency and sustainability with other energy sources.

Mathematical Modeling in Renewable Energy

A critical component of this module is the introduction of mathematical modeling. Students learn how to use mathematical tools to analyze and optimize renewable energy systems. They are taught to create models that predict energy output, efficiency, and economic feasibility of solar and geothermal projects. This analytical skill is crucial for understanding the complexities of energy systems and for making data-driven decisions.

Practical Activities

The second module focuses on hands-on learning experiences, allowing students to apply the theoretical knowledge gained in Module 1 to practical scenarios.

Prototype Energy-Independent Solar House

At the heart of this module is the construction and analysis of a prototype energy-independent solar house. This practical project serves as a tangible representation of the principles discussed in the classroom.

Design and Construction

Students collaborate with university mentors to design a solar house that integrates both solar and geothermal energy systems. They start with the architectural design, considering factors such as building orientation, insulation, and energy efficiency. The design phase includes:

- Site Analysis: Assessing the optimal location for solar panel installation to maximize sunlight exposure.
- System Integration: Planning the integration of solar panels, batteries for energy storage, and geothermal heat pumps.
- Material Selection: Choosing sustainable materials for construction to enhance the energy efficiency and environmental sustainability of the house.

Energy System Installation

Once the design is finalized, students move on to the installation phase. They work in teams to install solar panels, set up the electrical system, and integrate geothermal heat pumps. This hands-on experience provides invaluable insights into the practical aspects of renewable energy systems.

Students also learn about the challenges of system integration, such as ensuring the compatibility of different technologies and optimizing the energy flow between solar and geothermal systems. Troubleshooting sessions are conducted to address any issues that arise during the installation process.

Data Collection and Analysis

After the solar house is operational, students engage in data collection and analysis. They monitor the energy production and consumption, analyzing parameters such as efficiency, output fluctuations, and system performance. This data is used to refine the mathematical models developed in Module 1, enhancing the accuracy of predictions and optimizations.

Students also explore ways to improve the system's performance through optimization techniques. They experiment with different configurations and settings, using tools like linear programming and simulation models to find the most efficient solutions.

Interactive Learning and Problem-Solving

Throughout both modules, the project emphasizes interactive learning and problem-solving. Students participate in workshops, group discussions, and hands-on activities that encourage active engagement and collaboration.

Workshops and Seminars

Regular workshops and seminars are conducted by university faculty and industry experts. These sessions cover advanced topics in renewable energy, such as the latest technological advancements, policy implications, and economic considerations. Students have the opportunity to ask questions, engage in discussions, and gain deeper insights into the field.

Group Projects and Case Studies

Group projects and case studies are integral to the learning process. Students work in teams to analyze real-world renewable energy projects, identify challenges, and propose solutions. This collaborative approach fosters teamwork, critical thinking, and analytical skills.

Problem-Solving Exercises

Interactive problem-solving exercises are designed to challenge students and enhance their analytical abilities. These exercises involve scenarios such as optimizing the energy output of a solar farm, designing an off-grid energy system for a rural community, and analyzing the economic feasibility of different renewable energy projects.

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project exemplifies an innovative educational approach that seamlessly integrates knowledge and practice. By engaging high school students in both theoretical and practical aspects of renewable energy, the project not only imparts essential skills and knowledge but also inspires a new generation of environmentally conscious individuals.

The detailed educational modules and practical activities ensure that students gain a comprehensive understanding of sustainable energy technologies. The hands-on experience of building and analyzing a prototype energy-independent solar house provides a real-world context that reinforces classroom learning. Through interactive learning and problem-solving exercises, students develop critical skills that are essential for tackling the global challenges of energy sustainability.

By empowering students with the tools and knowledge needed to promote sustainable energy practices, the "Energize Tomorrow" project contributes to the broader goal of achieving affordable, reliable, sustainable, and modern energy for all. This integrated approach to education serves as a model for future initiatives aimed at fostering a sustainable future through informed and engaged youth.

3. Impact, Challenges, and Future Directions

Impact Assessment

The success of the "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project hinges on its ability to impart knowledge, skills, and inspiration to high school students. Evaluating the impact of the project is crucial for understanding its effectiveness and for identifying areas for improvement.

To gauge the educational impact, pre- and post-project surveys are administered to participating students. These surveys assess students' baseline knowledge and attitudes towards renewable energy before the project begins and measure any changes in knowledge, skills, and perceptions after its completion.

Pre-Project Surveys These surveys collect data on students' initial understanding of renewable energy concepts, their interest in the subject, and their awareness of the global energy crisis. Questions are designed to assess both theoretical knowledge and practical awareness.

Post-Project Surveys After the project's completion, students are surveyed again to measure any changes in their knowledge and attitudes. The post-project surveys include questions that mirror those in the pre-project surveys, allowing for direct comparison. Additionally, they include open-ended questions that invite students to reflect on their experiences and provide qualitative feedback.

One of the key metrics for evaluating the project's impact is the knowledge gain among students. This is measured by comparing pre- and post-project survey results. Significant improvements in test scores indicate that the educational modules have successfully enhanced students' understanding of renewable energy concepts.

Additionally, students' ability to apply theoretical knowledge to practical problems is assessed through their performance in hands-on activities and group projects. Successful completion of these tasks demonstrates that students can translate classroom learning into real-world applications.

Skill Development The project aims to equip students with practical skills related to renewable energy technologies. This includes technical skills in designing and installing solar and geothermal systems, as well as analytical skills in mathematical modeling and optimization.

Students' skill development is evaluated through their participation in the construction and analysis of the prototype energy-independent solar house. Observations of their ability to work collaboratively, solve problems, and adapt to new challenges provide insights into their practical competencies.

Another important aspect of the project's impact is student engagement and motivation. The hands-on, interactive nature of the project is designed to foster a deep interest in renewable energy and sustainability.

Student engagement is measured through their active participation in workshops, discussions, and practical activities. High levels of enthusiasm and involvement suggest that the project has successfully captured their interest and inspired a commitment to sustainable practices.

Challenges and Solutions

Implementing a service-learning project of this scale is not without its challenges. Identifying and addressing these challenges is essential for ensuring the project's success and for improving future iterations.

One of the primary challenges encountered during the project is the technical complexity of renewable energy systems. Designing and installing a prototype energy-independent solar house requires a high level of technical expertise and problem-solving skills.

To address these challenges, the project incorporates comprehensive training sessions led by university faculty and industry experts. These sessions cover the technical aspects of solar and

geothermal energy systems, providing students with the necessary knowledge and skills to tackle technical issues.

Additionally, the project encourages collaborative problem-solving, where students work in teams to brainstorm solutions and share expertise. This approach not only enhances their technical abilities but also fosters teamwork and communication skills.

Coordinating a project that involves multiple stakeholders, including university students, high school students, faculty, and community members, presents logistical challenges. Ensuring smooth communication, scheduling, and resource allocation is crucial for the project's success.

To address these challenges, the project employs a detailed project management plan that outlines roles, responsibilities, and timelines. Regular meetings and progress reports help keep all stakeholders informed and aligned. Additionally, contingency plans are developed to handle any unforeseen issues that may arise during the project.

Maintaining high levels of student engagement throughout the project can be challenging, especially given the complex and technical nature of the subject matter. To keep students motivated and interested, the project incorporates a variety of interactive and hands-on activities.

These activities are designed to be both educational and enjoyable, ensuring that students remain engaged and enthusiastic. By providing opportunities for creative problem-solving and real-world applications, the project maintains students' interest and fosters a positive learning environment.

Future Directions

Building on the success of the "Energize Tomorrow" project, there are several future directions that can be explored to expand its impact and reach.

The educational modules can be expanded to include additional renewable energy technologies, such as wind and hydroelectric power. By broadening the curriculum, students can gain a more comprehensive understanding of the diverse range of renewable energy sources available.

Furthermore, integrating topics such as energy storage, smart grids, and energy policy can provide students with a holistic view of the energy landscape and the various factors that influence the adoption of renewable energy technologies.

The project can be scaled to involve more schools and communities, increasing its reach and impact. Partnerships with other educational institutions, non-profit organizations, and industry partners can provide additional resources and support for expanding the project.

Additionally, developing online resources and virtual modules can make the project accessible to a wider audience, allowing students from different regions to participate and benefit from the program.

To better understand the long-term impact of the project, follow-up studies can be conducted to track students' progress and career paths. These studies can provide valuable insights into how the knowledge and skills gained through the project influence students' future academic and professional choices.

By maintaining connections with project alumni and supporting their continued involvement in renewable energy initiatives, the project can create a lasting network of individuals committed to advancing sustainable energy practices.

The success of the "Energize Tomorrow" project highlights the importance of education in promoting sustainable energy practices. By sharing the project's outcomes and best practices with policymakers and educators, the project can advocate for the inclusion of renewable energy education in school curricula nationwide.

Furthermore, engaging with local and national governments to support policies that promote renewable energy adoption and education can help create an enabling environment for sustainable energy initiatives.

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project represents a pioneering effort to integrate academic knowledge with practical application in the field of renewable energy. By empowering high school students with the knowledge and skills necessary to promote sustainable energy practices, the project addresses critical global challenges and fosters a new generation of environmentally conscious individuals.

Through comprehensive educational modules, hands-on activities, and interactive learning experiences, the project successfully bridges the gap between theory and practice. The evaluation of the project's impact, along with the identification and resolution of challenges, provides valuable insights for future iterations and expansions.

As the project looks towards the future, the potential for growth and increased impact is significant. By expanding the curriculum, scaling the project, and advocating for policy changes, "Energize Tomorrow" can continue to inspire and educate students, contributing to a sustainable energy future for all.

4. Conclusions

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project stands as a significant and innovative initiative designed to bridge the gap between academic learning and practical application in the field of renewable energy. By empowering high school students with the knowledge and skills necessary to ensure access to affordable, reliable, sustainable, and modern energy, the project aligns with the United Nations Sustainable Development Goal 7 and fosters a new generation of environmentally conscious and technically proficient individuals.

The primary objective of the project was to educate and inspire high school students about sustainable energy practices. Through the presentation of a prototype of an energy-independent solar house and the integration of surface geothermal energy with classical installations, university students from the Technical University of Civil Engineering successfully demonstrated the practical applications of renewable energy technologies. This hands-on approach not only captivated the interest of the high school students but also provided them with a tangible example of how theoretical knowledge can be translated into real-world solutions.

The educational modules, designed to cover the principles of solar and geothermal energy, were instrumental in enhancing the students' understanding of these concepts. By incorporating mathematical modeling of energy systems and practical activities to calculate and optimize energy efficiency, the project ensured that students were not only passive recipients of information but active participants in the learning process. This interactive approach significantly improved their problem-solving and analytical skills, preparing them for future challenges in the field of sustainable energy.

Evaluating the impact of the project was a critical component of its success. Pre- and post-project surveys revealed a significant increase in the students' knowledge and skills related to renewable energy. The data collected showed that students gained a deeper understanding of the technical aspects of solar and geothermal energy systems, as well as a heightened awareness of the importance of sustainable energy practices.

Moreover, the project successfully engaged students on an emotional and motivational level. The interactive and hands-on nature of the activities fostered a genuine interest in renewable energy and inspired many students to consider careers in this field. The enthusiasm and dedication observed during the project activities were clear indicators of its positive impact on student engagement and motivation.

The implementation of the project was not without its challenges. Technical complexities, logistical issues, and maintaining student engagement were some of the hurdles encountered. However, these challenges were effectively addressed through comprehensive training sessions, detailed project management plans, and the incorporation of engaging and interactive activities. The collaborative problem-solving approach adopted by the project team proved to be particularly effective in overcoming technical challenges, fostering teamwork, and enhancing the overall learning experience.

Looking forward, the potential for expanding and refining the "Energize Tomorrow" project is immense. There are several avenues through which the project can grow to have an even greater impact.

Expanding the educational modules to include additional renewable energy technologies, such as wind and hydroelectric power, would provide students with a broader understanding of the renewable energy landscape. Incorporating topics like energy storage, smart grids, and energy policy would also give students a more comprehensive view of the factors influencing the adoption and implementation of renewable energy technologies.

To increase its reach, the project can be scaled to involve more schools and communities. Partnerships with other educational institutions, non-profit organizations, and industry partners can provide the necessary resources and support for this expansion. Additionally, developing online resources and virtual modules can make the project accessible to a wider audience, allowing students from different regions to participate and benefit from the program.

Conducting follow-up studies to track the progress and career paths of students who participated in the project would provide valuable insights into its long-term impact. These studies could help identify how the knowledge and skills gained through the project influence students' future academic and professional choices, and inform improvements to the project design.

Sharing the outcomes and best practices of the "Energize Tomorrow" project with policymakers and educators can advocate for the inclusion of renewable energy education in school curricula nationwide. Engaging with local and national governments to support policies that promote renewable energy adoption and education can create an enabling environment for sustainable energy initiatives.

The "Energize Tomorrow: Service-Learning for Sustainable Energy Access" project represents a pioneering effort to integrate academic knowledge with practical application in the field of renewable energy. Its success in empowering high school students with the knowledge and skills necessary to promote sustainable energy practices underscores the importance of education in addressing global energy challenges.

As the project moves forward, the lessons learned and the achievements made provide a strong foundation for future growth and impact. By continuing to inspire and educate students, "Energize Tomorrow" can contribute to the development of a sustainable energy future for all, demonstrating the profound impact that service-learning initiatives can have on both individuals and communities.

References

[1] Venkata Rao, Ravipudi & Keesari, H. S. (2022). "Introduction to Renewable Energy Systems". Optimization in Renewable Energy Systems". Publisher Elsevier: 1-26.

[2] Smith, A. (2018). "Mathematics for Sustainable Energy Analysis." Journal of Renewable Energy Education 10(2): 123-135.

[3] Brown, C. (2019). "Optimization Techniques for Renewable Energy Systems." Sustainable Energy Journal 5(3): 245-260.

[4] Johnson, B. (2017). "Data Analysis Methods for Renewable Energy Applications." Renewable Energy Data Science 8(4): 321-335.
[5] White, D. (2016). "Control Theory in Renewable Energy Systems." Journal of Sustainable Energy Control 12(1): 45-56.
[6] Green, E. (2021). "Financial Modeling for Renewable Energy Investments." Renewable Energy Finance Review 15(2): 187-200.
[7] Technical University of Civil Engineering. Available online: https://dmi.utcb.ro/girls/ (accessed on 1 June 2024).
[8] GIRLS Project. Available online: https://girlsproject.eu/ (accessed on 1 June 2024).

A SURVEY ABOUT BARYCENTRIC LAGRANGE INTERPOLATION

Ion Mierluș Mazilu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei 124, sector 2, 38RO-020396 Bucharest, Romania
E-mail: ion.mierlusmazilu@utcb.ro
Ștefania Constantinescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: stefania.constantinescu@utcb.ro
Alice Anghelescu
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest
Bd. Lacul Tei, sector 2, 38RO-020396 Bucharest, Romania
E-mail: alice.anghelescu@utcb.ro

Abstract: This paper presents the barycentric Lagrange polynomial interpolation and sequential barycentric Lagrage interpolation. Some considerations are made on the advantages of the method versus the Lagrange and Newton polynomial interpolations regarding the size order of the calculations. In addition, two algorithms are described: one for barycentric Lagrange interpolation and another for sequential barycentric interpolation. Finally, some examples of applications of this method are presented. The bibliography produces an overview of the field, of the interest of researchers in this direction, and of the relevance of this domain.
Mathematics Subject Classification (2010): 65D05, 65D25.
Key words: polynomial Lagrange and Newton interpolation, barycentric formula, sequential barycentric interpolation.

1. Introduction

Let $f:[a,b]->R$ be a function and (n+1) distinct points called interpolation nodes:

$$X=(x_0,x_1,...,x_n)\ ,\ x_i\neq x_j\ ,\ x_i\in[a,b]\ (\forall)\left(i\neq j\ ,\ i,\ j\in\overline{0,n}\right). \tag{1}$$

The values of the function f in the nodes are known,

$$(\forall)x_i\ \ f(x_i)=f_i\ ,i=\overline{0,n} \tag{2}$$

but the form of the function f is not known, or if it is known, it is very complicated. The interpolation problem consists in determining a function $g:[a,b]->R$ of the known form with the property that

$$(\forall)x_i\ \ f(x_i)=g(x_i),i=\overline{0,n} \tag{3}$$

in nodes.

Formulated this way, the problem can have several solutions or none [14]. The most used class of interpolation functions is the class of polynomials of degree n.

The paper presents barycentric Lagrange interpolation, sequential barycentric interpolation highlighting the advantages of the method from the point of view of calculation efficiency. Algorithms corresponding to the presented methods are given, and, finally domains where

these methods could be applied. The bibliography produces an overview of the field, of the interest of researchers in this direction, and of the relevance of this domain.

2. Polynomial interpolation

2.1. Lagrange interpolation

We consider X given by equation (1) and F given by equation (4):

$$F=\left(f_0,f_1,...,f_n\right),\left(\forall\right)x_i\ f(x_i)=f_i,i=\overline{0,n}\ . \tag{4}$$

Then there exists and is unique P_n a polynomial of degree n that interpolates the function f in nodes x_i, called Lagrange's interpolation polynomial [14].

Now we note

$$P_n(x)=\ell_0(x)f_0(x)+\ell_1(x)f_1(x)+...+\ell_n(x)f_n(x) \tag{5}$$

where

$$\ell_j(x)=\frac{\prod\limits_{\substack{k=0\\k\neq j}}^{n}\left(x-x_k\right)}{\prod\limits_{\substack{k=0\\k\neq j}}^{n}\left(x_j-x_k\right)},\ \ell_j(x_k)=\begin{cases}1 & j=k\\ 0 & \text{otherwise}\end{cases},\ j,k=\overline{0,n}\ . \tag{6}$$

If f is a function of class $(n+1)$, $f\in C^{n+1}\left[a,b\right]$, then the approximation error can be evaluated

$$\left|E\left(f,x\right)\right|=\left|f(x)-P_n(x)\right|=\frac{\left|f^{(n+1)}\left(\xi_x\right)\right|}{\left(n+1\right)!}\left|\prod_{i=0}^{n}\left(x-x_i\right)\right|\le\frac{M}{\left(n+1\right)!}\left|\prod_{i=0}^{n}\left(x-x_i\right)\right| \tag{7}$$

where $M>0$ such that $\left|f^{(n+1)}\left(x\right)\right|\le M\ \left(\forall\right)x\in\left[a,b\right]$ and $\xi_x\in\left[a,b\right]$.

Remarks:

1. If f is a polynomial of degree n then P_n coincides with f on $[a,b]$.
2. Polynomial interpolation, except when the degree of the polynomial is small, it is not indicated if the interpolation nodes are equidistant recommending to use Chebyshev nodes.
3. For practical calculations, many authors highlight the following deficiencies of the method [1]:
4. Each evaluation of $P_n(x)$ requires $O(n^2)$[1] elementary operations: additions and multiplications.
5. Adding a new given pair (x_{n+1},f_{n+1}) a new calculation from scratch.
6. The calculation is numerically unstable. The interpolating Lagrange polynomial tends to oscillate between the interpolating nodes. Those oscillations, called the Runge phenomenon, can in some cases get worse as the number of interpolation nodes increases. It is a well-known classical result that the oscillation caused by the *Runge phenomenon*[2] can be minimized by interpolating at the Cebyshev nodes [13].

For these reasons it is considered that the Lagrange form of P_n is mainly a theoretical one used as tool for proving theorems. For calculations, it is recommended to use Newton's

[1]The complexity of the algorithms: We will consider the complexity of algorithms only from the aspect of the number of elementary operations - time complexity (not from the point of view of occupied memory, space complexity!), considering: addition, subtraction, multiplication, division, modulo operations, Boolean operations, divisions and assignments, as being elementary. For this, the order of magnitude of the number of elementary operations is determined. The order of magnitude of h is defined as the set of functions $O(h)=\left\{t:N\to R_+\left|\left(\exists\right)c>0,\right|\left(\exists\right)k_0\in N\ a.e.\ t(k)\le c\cdot h(k),\ \left(\forall\right)k>k_0\right\}$

[2] Runge's phenomenon is a problem of oscillation at the edges of an interval that occurs when using polynomial interpolation with polynomials of high degree over a set of equispaced interpolation points.

formula, which requires only *O(n)* elementary operations for each point for which P_n is evaluated together with some numbers, which are independent of the point *x*.

2.2. Newton interpolation

Considering P_n of the form [1]

$$P_n(x) = a_0 + a_1(x - x_0) + ... + a_n(x - x_0)(x - x_1)(x - x_2)...(x - x_{n-1}) \quad (5')$$

where

$$\begin{array}{lllll}
a_0 = f[x_0] & & & & \\
& a_1 = f[x_0, x_1] & & & \\
f[x_1] & & a_2 = f[x_0, x_1, x_2] & & \\
& f[x_1, x_2] & & a_3 = f[x_0, x_1, x_2, x_3] & \ddots \\
f[x_2] & & f[x_1, x_2, x_3] & \vdots & a_n = f[x_0, x_1, x_2, ..., x_n] \\
& f[x_2, x_3] & & \vdots & \\
f[x_3] & \vdots & \vdots & f[x_{n-3}, ..., x_n] & \iddots \\
\vdots & \vdots & f[x_{n-2}, x_{n-1}, x_n] & & \\
\vdots & f[x_{n-1}, x_n] & & & \\
f[x_n] & & & &
\end{array}$$

and

$$f\left[x_j, x_{j+1}, ..., x_{k-1}, x_k\right] = \frac{f\left[x_{j+1}, ..., x_{k-1}, x_k\right] - f\left[x_j, x_{j+1}, ..., x_{k-1}\right]}{x_k - x_j} \quad (8)$$

and $f\left[x_j\right] = f_j$

being necessary $O(n^2)$ elementary operations.

3. Barycentric Lagrange formula

3.1. First barycentric interpolation formula

If we consider that

$$\ell_j(x) = \frac{\ell(x)}{\left(x - x_j\right)}, \text{ where } \ell(x) = \left(x - x_0\right)\left(x - x_1\right)...\left(x - x_n\right), \; j = \overline{0,n} \; . \quad (9)$$

We can define the *barycentric weigths* by

$$w_j = \frac{1}{\prod\limits_{\substack{k=0 \\ k \neq j}}^{n}\left(x_j - x_k\right)} \quad j = \overline{0,n} \quad (10)$$

or

$$w_j = \frac{1}{\ell'\left(x_j\right)} \text{ , and } \ell_j\left(x\right) = \ell\left(x\right)\frac{w_j}{x - x_j} \text{ , } \quad j = \overline{0,n} \, . \quad (11)$$

Thus

$$P_n(x) = \ell(x)\sum_{j=0}^{n}\frac{w_j}{x - x_j}f_j \; . \quad (12)$$

Now the Lagrange interpolation requires $O(n^2)$ elementary operations for calculating some quantities independent of *x*, the numbers w_j, followed by $O(n)$ elementary operations for evaluating $P_n(x)$ once these numbers are known. Rutishauser [18] called the form in equation (12) the *first form of the barycentric interpolation formula*.

3.2. Second barycentric interpolation formula

Considering the function to be interpolated f=1, then f_j=1 , $j=\overline{0,n}$, equation (12) becomes

$$1=\sum_{j=0}^{n}\ell_j(x)=\ell(x)\sum_{j=0}^{n}\frac{w_j}{x-x_j},\ \ell(x)=\frac{1}{\sum_{j=0}^{n}\frac{w_j}{x-x_j}}\ . \tag{13}$$

Substituting $\ell(x)$ from equation (13) in equation (12) *the second barycentric interpolation formula* is obtained

$$P_n(x)=\frac{\sum_{j=0}^{n}\frac{w_j}{x-x_j}f_j}{\sum_{j=0}^{n}\frac{w_j}{x-x_j}}\ . \tag{12'}$$

Remark: The barycentric formula (12') has a symmetry that allows updating the weights w_j in *O(n)* elementary operations to incorporate another pair of data (x_{n+1},f_{n+1}), thus:

a) Calculate $w_j'=\frac{w_j}{x_j-x_{n+1}},\ (\forall)\ j=\overline{0,n}$

b) Calculate $w_{n+1}'=\frac{1}{\prod_{k=0}^{n}(x_{n+1}-x_k)}$ for another $(n+1)$ elementary operations.

4. Sequential Interpolation

In [13] the following theorem is proved which allows that using the barycentric interpolation formula, new nodes

$$\overline{X}=(x_{n+1},x_{n+2},...,x_{n+k}),\ x_i\in[0,1],\ (\forall)i=\overline{n+1,n+k},\ x_i\neq x_j,\ (\forall)\left(i\neq j,\ i,\ j\in\overline{n+1,n+k}\right)$$

can be added to the existing ones by updating the interpolator corresponding to the data *(X,F)* through sequential interpolation including the barycentric weights, w_j, already calculated.

Theorem [13] *For* $n\geq 0$, *let* $P_n(x)$ *and* $Q_{n+1}(x)$ *be two barycentric interpolators, where* $P_n(x)$ *is defined as in equation (12) and* $Q_{n+1}(x)$ *is constructed with an additional dummy node and* $x_{n+1}^*\in[0,1]$, $x_{n+1}^*\neq x_j$, $j=\overline{0,n}$ *and* $f_{n+1}^*=P_n\left(x_{n+1}^*\right)$. *Then*

$$Q_{n+1}(x)-P_n(x)=\ell(x)\sum_{j=0}^{n+1}w_jf_j\ . \tag{13}$$

*Remark*s: 1. The right hand of previous equation depesns on (but not x_{n+1}^*) and $\sum_{j=0}^{n+1}w_jf_j$ that depens on X and $X\cup\left\{x_{n+1}^*\right\}$ *and* $F\cup\left\{f_{n+1}^*\right\}$.

2. The interval [0,1] is not a restriction and it can be adapted to any finite interval $[a,b]\subset R$.

3. New node can be chosen according to the relationship

$$x_{n+1}=\arg\max_{x\in[0,1]}\left|Q_{n+1}(x)-P_n(x)\right| \tag{14}$$

and X becames $X\cup x_{n+1}$. The search for x_{n+1} is correctly posed because the problem is to maximize a bounded function on a compact set.

5. Algorithms

5.1. Algorithm to calculate the barycentric weigths

Input X, F

$V_0=0$

for j=1 to n do

for k=0 to j-1 do $V_k=(x_k-x_j)V_k$ *endfor*

$V_j=(x_j-x_0)\ (x_j-x_1)\ \ldots\ (x_j-x_{j-1})$

endfor

for j=0 to n do $w_j=1/V_j$ *endfor*

5.2. Algorithm for computing the new barycentric weights [1]

Input $X\cup\{x_{n+1}\}$

$w_0^{(0)}=1$

for j=1 to n do

for k=0 to j-1 do

$$w_k^{(j)}=\left(x_k-x_j\right)w_k^{(j-1)}$$

endfor

$$w_k^{(j)}=\prod_{k=0}^{j-1}\left(x_j-x_k\right)$$

endfor

for j=0 to n do

$$w_j^{(j)}=\frac{1}{w_j^{(j)}}$$ *endfor*

return

5.3. Algorithms for sequential interpolation

Input: initial values for X, F, k=number of new nodes

$X=\left(x_0,x_1,\ldots,x_n\right)$, $x_i\in[0,1]$, $(\forall)i=\overline{0,n}$, $x_i\neq x_j$, $(\forall)\left(i\neq j,\ i,\ j\in\overline{0,n}\right)$

k new nodes

$\overline{X}=\left(x_{n+1},x_{n+2},\ldots,x_{n+k}\right)$, $x_i\in[0,1]$, $(\forall)i=\overline{n+1,n+k}$, $x_i\neq x_j$, $(\forall)\left(i\neq j,\ i,\ j\in\overline{n+1,n+k}\right)$

Compute $\ell_n(x)=\prod_{i=0}^{n}\left(x-x_i\right)$

$j=0$

Do

$$x_{n+j+1}=\arg\max_{x\in[0,1]}\left|\ell_{n+j}(x)\right|$$

$$\ell_{n+j+1}(x)=\ell_{n+j}(x)\left(x-x_{n+j+1}\right)$$

$j:=j+1$

while $j\geq k$

Return $\overline{X}=\left(x_{n+1},x_{n+2},\ldots,x_{n+k}\right)$, $x_i\in[0,1]$, $(\forall)i=\overline{n+1,n+k}$

Stop!

6. Some applications filds of barycentric Lagrange interpolation

6.1. Estimarea constantei Lebesgue[3]

A good measure of the behavior of the interpolation problem is given by the Legesgue constant defined [1] using the norm of P_n

$$\Lambda_n = \|P_n\| = \sup_{f \in C[a,b]} \frac{\|P_n\|}{\|f\|} \ , \ where \ \|f\| = \max_{x \in [a,b]} |f(x)| \tag{15}$$

C[a,b] being the space of all continuous functions on [*a, b*]. In [15] it is shown that this number can be determined by

$$\Lambda_n = \max_{x \in [a,b]} \sum_{j=0}^{n} |l_j(x)| \ . \tag{16}$$

A lower bound on the Lebesgue constant can be calculated using the barycentric weights as follows [10]

$$\Lambda_n \geq \frac{1}{2n^2} \frac{\max_{0 \leq j \leq n} |w_j|}{\min_{0 \leq j \leq n} |w_j|} \ . \tag{17}$$

6.2. Differentiation of Polynomial Interpolants

In [1] it is proved that

$$\ell_j'(x_j) = \frac{\frac{w_j}{w_i}}{x_i - x_j} \ , \ \ell_j''(x_i) = -2 \frac{\frac{w_j}{w_i}}{x_i - x_j} \left[\sum_{k \neq i} \frac{\frac{w_k}{w_i}}{x_i - x_k} - \frac{1}{x_i - x_j} \right] \tag{18}$$

and

$$\ell_j'(x_j) = -\sum_{i \neq j} \ell_j'(x_j) \ , \ \ell_j''(x_j) = -\sum_{i \neq j} \ell_j''(x_j) \ , \ D_{ij}^{(1)} = \ell_j'(x_i) \text{ and } D_{ij}^{(2)} = \ell_j''(x_i) \tag{19}$$

$D^{(1)}$ and $D^{(2)}$ first- and second-order differentiation matrices obtained by interpolating *(X,F)* data, then differentiating the interpolant at the grid points.

6.3. Spectral Methods for Differential Equations

The barycentric Lagrange forms play an important role in spectral collocation methods in the numerical solution of differential equations [7].

6.4. Fast Multipole Methods

Rokhlin [1,17] and Greengard [8] are invented fast algorithms for evaluating certain sums [3]. In [6], the multipole methods are applied to the evaluation of interpolating polynomials of high degrees by means of the first form of the barycentric formula (12). Tests based on the second formula (12') are described in [2].

6.5. Barycentric Lagrange Interpolation Method for Solving Love's Integral Equations

In [19] is presented the barycentric Lagrange interpolation method for solving Love's integral equations

$$u(x) = 1 + \int_{-1}^{1} \frac{d}{\pi \left(d^2 + (x-t)^2 \right)} u(t) dt \ , \ x \leq 1 \tag{20}$$

that appears in electrostatic systems.

[3] The Lebesgue constants (depending on a set of nodes and of its size) give an idea of how good the interpolant of a function (at the given nodes) is in comparison with the best polynomial approximation of the function (the degree of the polynomials are fixed). The Lebesgue constant for polynomials of degree at most n and for the set of n+ 1 nodes X is generally denoted by $\Lambda_n(X)$

References

[1] Berrut, J.-P., Trefethen, L.N.: Barycentric Lagrange Interpolation, *SIAM REVIEW* (2004) *Society for Industrial and Applied Mathematics,* Vol. 46, No. 3, pp. 501–517.

[2] Blochliger, M.: *Ein Spezialfall der Multipolmethode f¨ur baryzentrische Formeln*, Diplomathesis, University of Fribourg, Switzerland, 1998.

[3] Boyd, J.P.: Multipole expansions and pseudospectral cardinal functions: A new generalization of the Fast Fourier Transform, *J. Comp. Phys*., 102 (1992), pp. 184–186.

[4] Chapra, S.C., Canale, R.P.: *Numericl methods for Engineering*, Six Edition, Mc Graw Hill, 2010.

[5] Danet, N.: *Analiza numerica cu aplicatii*, Editura Matrix, Bucuresti, 2005.

[6] Dutt, A.M., Gu, M., Rokhlin, V.: Fast algorithms for polynomial interpolation, integration, and differentiation, *SIAM J. Numer. Anal., 33* (1996), pp. 1689–1711.

[7] Eriksson, K., Estep, D., Hansbo, P., Johnson, C.: *Introduction to Computational Methods for Differential Equations*, Clarendon Press, Oxford, 1995.

[8] Greengard, L.: *The Rapid Evaluation of Potential Fields in Particle Systems*, MIT Press, Cambridge, MA, 1988.

[9] Groza, G.: *Analiza numerica*, Editura Matrix, Bucuresti, 2005.

[10] Heinrichs, W.: *Strong convergence estimates for pseudospectral methods*, Appl. Math., 37 (1992), pp. 401–417.

[11] Henrici, P.: Barycentric formulas for interpolating trigonometric polynomials and their conjugates, *Numer. Math*., 33 (1979), pp. 225–234.

[12] Higham, N.J.: The numerical stability of barycentric Lagrange interpolation, *IMA J. Numer* 24(4), (2004).

[13] Hugo Maruri-Aguilar, H., Trandafir, P.-C.: Sequential Barycentric Interpolation, in *mODA 9 –Advances in Model-Oriented Design and Analysis, Physiac-Verlag, A Springer Company - Proceedings of the 9th International, Workshop in Model-Oriented Design and Analysis held in Bertinoro, Italy,* June 14–18, 2010.

[14] Paltineanu, G., Matei, P., Trandafir, R.: *Bazele Analizei Numerice*, Editura Printech, Bucuresti, 2001.

[15] Powell, M.J.D.: *Approximation Theory and Methods*, Cambridge University Press, Cambridge, UK, 1981.

[16] Schwarz, H.R.: *Numerische Mathematik*, 4th ed., Teubner, Stuttgart, 1997; English translation of the 2nd edition, *Numerical Analysis: A Comprehensive Introduction*, Wiley, New York, 1989.

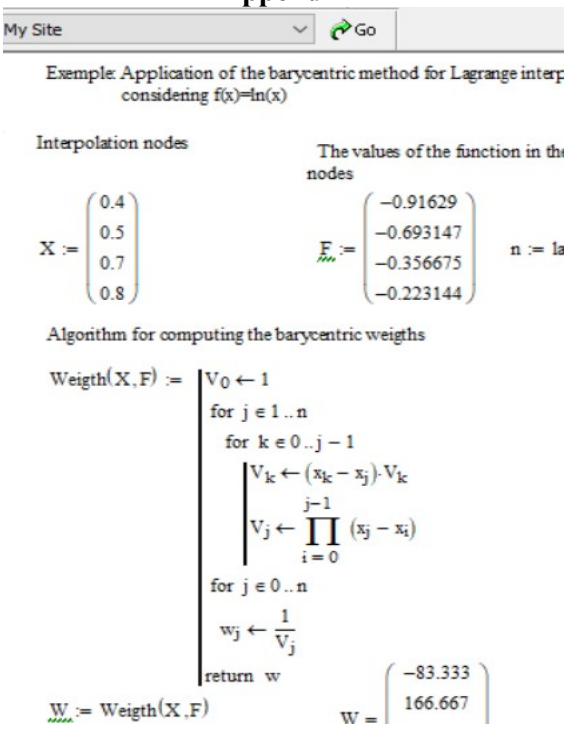
[17] Rokhlin, V.: Rapid solution of integral equations of classical potential theory, *J. Comput. Phys.*, 60 (1985), pp. 187–207.

[18] Rutishauser, H.: *Vorlesungen ¨uber numerische Mathematik*, Vol. 1, Birkh¨auser, Basel, Stuttgart, 1976; English translation, *Lectures on Numerical Mathematics*, Walter Gautschi, ed., Birkh¨auser, Boston, 1990.

[19]. Shoukralla, E.S., Ahemed, B.M.: Barycentric Lagrange Interpolation Method for Solving Love's Integral Equations, Research Square, https://doi.org/10.21203/rs.3.rs-2882853/v1, Springer, Boundary Value Problems, 9 Jul 2023.

[20] Szabados, J., Vertesi, P.: *Interpolation of Functions*, World Scientific, Singapore, 1990.

Appendix


My Site
Go
Exemple: Application of the barycentric method for Lagrange interp
considering f(x)=ln(x)
Interpolation nodes
The values of the function in th
nodes
X :=
0.4
0.5
0.7
0.8
F :=
−0.91629
−0.693147
−0.356675
−0.223144
n := la
Algorithm for computing the barycentric weigths
Weigth(X, F) :=
V0 ← 1
for j ∈ 1 .. n
for k ∈ 0 .. j − 1
Vk ← (xk − xj)·Vk
Vj ← ∏ (xj − xi), i = 0 to j−1
for j ∈ 0 .. n
wj ← 1/Vj
return w
W := Weigth(X, F)
W =
−83.333
166.667

ZEROS OF POLYNOMIALS OVER LOCALLY COMPACT FIELDS

SEVER ANGEL POPESCU

In memory of my Teacher and Master, Dr. doc. Nicolae Popescu

Abstract. Let K be a non-discrete locally compact (commutative) field and let $P \in K[X]$ be a non-constant polynomial. In this note we give two criteria for the existence of a zero for P in the initial field K. Then, we apply these criteria to simply prove that the complex number field is algebraically closed and, under the truth of a conjecture, that any polynomial equation $P(X) = 0$, with $P \in K[X]$, is solvable by radicals.

Mathematics Subject Classification (2020): Primary 12J10, 12E12; Secondary 12J25, 12J99.
Key words: locally compact field, zeros of polynomials, Fundamental Theorem of Algebra, radical extension.

1. Introduction

We started from a brilliant idea of C. Fefferman [1] and we tried to use it in the much more general context of a non-discrete locally compact (commutative) field K. From [6], Chapter 1, §2 we know that K is a rank one valued field (K, φ), where $\varphi : K \to \mathbb{R}_{\geq 0}$ is its corresponding multiplicative absolute value. The structure of such fields are well known (see [6], Chapter 1, §2). Namely, if φ is Archimedean, then $(K, \varphi) \simeq (\mathbb{R}, |\cdot|^s)$ or $(K, \varphi) \simeq (\mathbb{C}, |\cdot|^s)$, where $|\cdot|$ is the usual absolute value and $0 < s \leq 1$. If φ is not Archimedean, then $(K, \varphi) \simeq (L, \varphi_L)$ or $(K, \varphi) \simeq (S, \psi_X)$, where L is a finite extension of $\mathbb{Q}_p$, the p-adic number field for a prime number p, φ_L is the unique extension to L of the p-adic absolute value φ_p of $\mathbb{Q}_p$, and S is a finite extension of the power series field $\mathbb{F}_p((X))$, with ψ_X the unique extension to S of the X-adic valuation of $\mathbb{F}_p((X))$. For a prime number p, as usual, $\mathbb{F}_p$ is the finite field with p elements. A more elementary treatment of this subject one can find in [2], Chapter 2, or in [4], Chapter 3.

In Section 2 we give the basic results (Theorem 2.2 and Theorem 2.4) and two applications of them, a simple proof for the Fundamental Theorem of Algebra (Corollary 2.5) and, under the truth of a conjecture, an elementary proof for the solvability by radicals of a polynomial equation $P(X) = 0$, $P \in K[X]$ (Corollary 2.6). In particular, there exists a unique non-discrete locally compact field that contains the radicals of all its elements, the complex number field. We also prove (under the truth of the same conjecture) that $Gal(L_P/K)$ is a solvable group, where $P \in K[X]$ and L_P is a decomposition field of P (Corollary 2.7).

2. The main results

In the following we fix a non-discrete locally compact field (K, φ), where φ is a canonical multiplicative absolute value (see Section 1), we also fix an algebraic closure $\overline{K}$ of K, and a non-constant polynomial,

$$P(x) = c_n x^n + c_{n-1}x^{n-1} + ... + c_0 \in K[x], c_n \neq 0.$$

Lemma 2.1. *With the above notation and assumptions, the function $x \to \varphi(P(x))$, defined on K with values in $\mathbb{R}_+$, has an absolute minimum point x_0 in K, that is there exists x_0 in K such that*

$$\inf_{x\in K} \varphi(P(x)) = \varphi(P(x_0)).$$

Proof. Since

$$\varphi(P(x)) = \varphi(x)^n \varphi\left(c_n + \frac{c_{n-1}}{x} + ... + \frac{c_0}{x^n}\right),$$

we can find $M > 0$ such that $\varphi(P(x)) \geq \varphi(c_0)$ for any $x \in K$ with $\varphi(x) > M$. Since $x \to \varphi(P(x))$ is a continuous function, there exists at least an element $x_0 \in B[0, M] = \{x \in K : \varphi(x) \leq M\}$, so that

$$\inf_{x\in B[0,M]} \varphi(P(x)) = \varphi(P(x_0)),$$

because $B[0, M]$ is a compact subspace in K. Let x be in $K \setminus B[0, M]$. Then,

$$\varphi(P(x)) \geq \varphi(c_0) = \varphi(P(0)) \geq \varphi(P(x_0)).$$

Thus $\varphi(P(x)) \geq \varphi(P(x_0))$ for any $x \in K$. □

Theorem 2.2. *Let (K, φ) be $(\mathbb{R},|\cdot|^s)$ or $(\mathbb{C},|\cdot|^s)$, $0 < s \leq 1$, where $|\cdot|$ is the usual absolute value on $\mathbb{R}$ and $\mathbb{C}$ respectively, and let P be a non-constant polynomial of degree n in $K[x]$. Let x_0 be an absolute minimum in K for the continuous function $x \to \varphi(P(x))$, $x \in K$ (see Lemma 2.1), and let*

$$P(x) = a_0 + a_j(x - x_0)^j + \cdots + a_n(x - x_0)^n, \ a_j \neq 0, \tag{2.1}$$

be the Taylor expansion of P at x_0. Then, x_0 is a zero for the equation $P(x) = 0$ if and only if the equation

$$x^j + \frac{a_0}{a_j} = 0 \tag{2.2}$$

has a root in K.

Proof. We see that if $P(x_0) = a_0 = 0$, then the equation (2.2) has the root $x = 0$ in K. Conversely, let us assume now that $a_0 \neq 0$, that is we assume that x_0 is not a root for P, and let $y_0 \in K$ be a root of the equation (2.2). We take a small $\varepsilon > 0$ and we compute

$$P(x_0 + \varepsilon y_0) = a_0 + a_j\varepsilon^j y_0^j + a_{j+1}\varepsilon^{j+1}y_0^{j+1} + \cdots + a_n\varepsilon^n y_0^n.$$

Since $y_0^j = -\frac{a_0}{a_j}$, we can also write:

$$P(x_0 + \varepsilon y_0) = a_0 - a_0\varepsilon^j - a_0\frac{a_{j+1}}{a_j}\varepsilon^{j+1}y_0 - \cdots - a_0\frac{a_n}{a_j}\varepsilon^n y_0^{n-j}.$$

Hence,

$$\varphi(P(x_0+\varepsilon y_0)) = \varphi(a_0)\varphi\left(1-\varepsilon^j-\varepsilon^{j+1}R(\varepsilon)\right),$$

where $R(x)$ is a polynomial in $K[x]$. Since $\varphi(x) = |x|^s$, $0 < s \leq 1$, we see that $\varphi\left(1-\varepsilon^j-\varepsilon^{j+1}R(\varepsilon)\right) < 1$ for ε small enough. Thus $\varphi(P(x_0+\varepsilon y_0)) < \varphi(a_0)$ for such an ε. But this is a contradiction, because $\varphi(a_0) = \varphi(P(x_0))$ and x_0 is an absolute minimum in K for $x \to \varphi(P(x))$, $x \in K$. □

Definition 2.3. *Let (K,φ) be a non-Archimedean locally compact non-discrete field. With the above notation, let x_0 be an absolute minimum point in K for the function $x \to \varphi(P(x))$, $x \in K$. We say that x_0 satisfies property (P) if the coefficients $a_0, a_j, a_{j+1}, \cdots, a_n$ of the polynomial $P(x)$ in its Taylor expansion (2.1) verify the following condition:*

$$\varphi(a_{j+k}) < \frac{\varphi(a_j)^{\frac{j+k}{j}}}{\varphi(a_0)^{\frac{k}{j}}} \tag{2.3}$$

for any $k = 1, 2, \cdots, n-j$. If $a_0 = 0$, we write $\frac{1}{+0} = +\infty$ and, in this case, property (P) is obviously satisfied.

Conjecture C. *For a non-Archimedean locally compact non-discrete field (K,φ), always does exist an absolute minimum point x_0 (for the continuous function $x \to \varphi(P(x))$, $x \in K$) that satisfies property (P).*

Theorem 2.4. *Let (K,φ) be a non-Archimedean locally compact non-discrete field, and let P be a non-constant polynomial of degree n in $K[x]$. Let x_0 be an absolute minimum point for the continuous function $x \to \varphi(P(x))$, $x \in K$, that satisfies property (P). Then, x_0 is a root of the equation $P(x) = 0$ if and only if the equation*

$$x^j + \frac{a_0}{a_j} = 0 \tag{2.4}$$

has at least one root in K. Here a_0, a_j are coefficients of P in its Taylor expansion from (2.1)

Proof. If x_0 is a root of P in K, then $a_0 = 0$ and $x = 0$ is a root in K for the equation (2.4). Conversely, we assume that $a_0 = P(x_0) \neq 0$ and $q \in K$ is a root of the equation (2.4), that is

$$q^j = -\frac{a_0}{a_j}. \tag{2.5}$$

Let us come back to (2.1) and compute

$$P(x_0+q) = -a_0\left(\frac{a_{j+1}}{a_j}q + \cdots + \frac{a_n}{a_j}q^{n-j}\right).$$

Thus,

$$\varphi(P(x_0+q)) = \varphi(a_0)\varphi\left(\sum_{k=1}^{n-j}\frac{a_{j+k}}{a_j}q^k\right). \tag{2.6}$$

Since

$$\varphi\left(\sum_{k=1}^{n-j}\frac{a_{j+k}}{a_j}q^k\right)\leq \max_{1\leq k\leq n-j}\left\{\frac{\varphi(a_{j+k})}{\varphi(a_j)}\varphi(q)^k\right\},$$

and since

$$\varphi(q)=\left[\frac{\varphi(a_0)}{\varphi(a_j)}\right]^{\frac{1}{j}}$$

(see (2.5)), we conclude that

$$\frac{\varphi(a_{j+k})}{\varphi(a_j)}\varphi(q)^k=\frac{\varphi(a_{j+k})\varphi(a_0)^{\frac{k}{j}}}{\varphi(a_j)^{\frac{j+k}{j}}}<1$$

for any $k=1,2,\cdots,n-j$ (see (2.3)). Hence,

$$\varphi(P(x_0+q))<\varphi(a_0)=\varphi(P(x_0)),$$

a contradiction, because x_0 is an absolute minimum point in K for $x\to\varphi(P(x))$, $x\in K$. Therefore, if the equation (2.5) has a root q in K, then $a_0=P(x_0)=0$. □

Corollary 2.5. (*The Fundamental Theorem of Algebra*) *The complex number field* $\mathbb{C}$ *is algebraically closed.*

Proof. (See also [1] for a similar idea.) We directly apply Theorem 2.2 and the fact that the equation $X^j+a_0/a_j=0$ always has all its solutions in $\mathbb{C}$, because $\exp(z)\in\mathbb{C}$ if $z\in\mathbb{C}$ and radicals of positive real numbers are also real numbers. □

Let K be an arbitrary field. We say that R/K is a *finite radical extension* of K if $R=L_s$, so that

$$K=L_0\subset L_1\subset\ldots\subset L_s$$

is a tower of fields, such that $L_i=L_{i-1}(\theta_i)$, with θ_i a root of an equation of the following type,

$$X^{n_i}+b_{i-1}=0,\ b_{i-1}\in L_{i-1},\ i=1,2,\ldots,s.$$

We say that a polynomial $P\in K[X]$ is *solvable* (*by radicals*) if its decomposition field L_P is contained in a finite radical extension R of K.

Corollary 2.6. *Now we assume that Conjecture C is true. Let* (K,φ) *be a non-Archimedean, non-discrete locally compact field, and let* $P(X)=0$ *be a polynomial equation with* $P\in K[X]$. *Then* $P(X)=0$ *has all its solutions in a finite radical extension* $R\subset\overline{K}$, *where* $\overline{K}$ *is a fixed algebraic closure of* K. *Thus,* P *is solvable by radicals.*

Proof. We can assume that $n=\deg_K P\geq 2$. Since (K,φ) is a complete field, φ can be uniquely extended to $\overline{K}$ by an absolute value $\overline{\varphi}$. In what follows we will denote $\overline{\varphi}$ also by φ.

Let us write again formula (2.1) in which we highlight an arbitrary finite extension T of K in which $P(X)$ can be written as

$$P(X)=a_0^{(T)}+a_{j_T}^{(T)}(X-x_0^{(T)})^{j_T}+a_{j_T+1}^{(T)}(X-x_0)^{j_T+1}+\ldots+a_n^{(T)}(X-x_0^{()})^n, \tag{2.7}$$

where $x_0^{(T)}$ is an arbitrary element in T and $a_{j_T}^{(T)} \neq 0$. We choose now $x_0^{(T)} \in T$ with property (P), such that

$$\inf_{x \in T} \varphi(P(x)) = \varphi(P(x_0^{(T)})) \tag{2.8}$$

(see Lemma 2.1 and Conjecture C).

If $P(x_0^{(K)}) = 0$ (here $T = K$), we are done. Let us assume that $P(x_0^{(K)}) \neq 0$ and so, from Theorem 2.4 (we use here the fact that Conjecture C is true), we see that any root α_{j_K} of the equation $X^{j_K} + a_0^{(K)}/a_{j_K}^{(K)} = 0$ is not in $K \overset{not}{=} K_0$. Let us take such a root $\alpha_{j_{K_0}} \in \overline{K}$, and let us consider $K_1 = K_0[\alpha_{j_{K_0}}]$, and

$$\inf_{x \in K_1} \varphi(P(x)) = \varphi(P(x_0^{(K_1)}))$$

for an element $x_0^{(K_1)} \in K_1$, which satisfies property (P). Thus,

$$\varphi(P(x_0^{(K_1)})) \leq \varphi(P(x_0^{(K_0)})).$$

If $P(x_0^{(K_1)}) = 0$, we are done. If not, we construct $K_2 = K_1[\alpha_{j_{K_1}}]$, where $\alpha_{j_{K_1}}$ is a fixed root of the equation $X^{j_{K_1}} + a_0^{(K_1)}/a_{j_{K_1}}^{(K_1)} = 0$, and so on. We see that

$$K = K_0 \subset K_1 \subset \ldots \subset K_m \subset \ldots \subset \overline{K} \tag{2.9}$$

and

$$\varphi(P(x_0^{(K_0)})) \geq \varphi(P(x_0^{(K_1)})) \geq \ldots \geq \varphi(P(x_0^{(K_m)})) \geq \ldots, \tag{2.10}$$

where $x_0^{(K_j)} \in K_j$ is an absolute minimum point of $x \to \varphi(P(x))$, $x \in K_j$, which satisfies property (P) for $j = 0, 1, ...$ Let $\alpha_{j_{K_s}} \in \overline{K}$ be the above fixed solution of the equation

$$X^{j_{K_s}} + a_0^{(K_s)}/\alpha_{j_{K_s}}^{(K_s)} = 0, \ s \in \{0, 1, \ldots\}. \tag{2.11}$$

Since $P'(x_0^{(K_h)}) = 0$ for any $h \in \{0, 1, \ldots\}$ where "we are not done" ($j_{K_h} > 1$, otherwise the equation (2.11) has a solution in K_h, and so, "we are done"), we see that there exists a smallest k such that

$$\varphi(P(x_0^{(K_k)})) = \varphi(P(x_0^{(K_{k+1})})) = \ldots = \varphi(P(x_0^{(K_m)})) = \ldots$$

Since $\varphi(P(x_0^{(K_{k+1})})) = \varphi(P(x_0^{(K_k)}))$, the absolute minimum point $x_0^{(K_{k+1})} \in K_{k+1}$, which satisfies property (P), can be taken to be $x_0^{(K_k)} \in K_k$, so we can write, instead of the equation,

$$P(X) = a_0^{(K_{k+1})} + a_{j_{K_{k+1}}}^{(K_{k+1})}(X - x_0^{(K_{k+1})})^{j_{K_{k+1}}} + ...,$$

the previous equation:

$$P(X) = a_0^{(K_k)} + a_{j_{K_k}}^{(K_k)}(X - x_0^{(K_k)})^{j_{K_k}} + ...,$$

so $\alpha_{j_{K_{k+1}}}^{(K_{k+1})}$ can be taken to be $\alpha_{j_{K_k}}^{(K_k)} \in K_k\left[\alpha_{j_{K_k}}^{(K_k)}\right] = K_{k+1}$ and again, applying Theorem 2.4 (we use here the fact that Conjecture C is true), we get that $P(x_0^{(K_k)}) = P(x_0^{(K_{k+1})}) = 0$.

Thus, the equation $P(X) = 0$ has a solution in the radical extension K_k/K. Hence, we can write $P(X) = (X - x_0^{(K_k)})Q(X)$ with $Q \in K_k[X]$, and we continue in the same way as above with Q instead of P and K_k instead of K, etc. Therefore, the equation $P(X) = 0$ has all its solutions in a finite radical extension R/K. □

Corollary 2.7. (*see also* [5]*, IV, Corollary 5 with an alternative proof*) *We also assume here that Conjecture C is true. Let (K, φ) be a non-discrete locally compact field and let P be a non-constant polynomial in $K[X]$. Let $L_P \subset \overline{K}$ be the decomposition field of P. Then $Gal(L_P/K)$ is a solvable group.*

Proof. We reconsider the tower of subfields (2.9),

$$K = K_0 \subset K_1 \subset \ldots \subset K_k \subset \overline{K}$$

such that $K_{j+1} = K_j[\alpha_{j_{K_j}}]$, where $\alpha_{j_{K_j}}$ is a root of the equation

$$X^{j_{K_j}} + a_0^{(K_j)}/a_{j_{K_j}} = 0$$

and $a_0^{(K_j)}$, $a_{j_{K_j}} \neq 0$, with $j_{K_j} > 1$, are the first two coefficients in the expansion (2.7) of $P(x)$ for $T = K_j$. We also assume that K_k contains all the roots of P (see the proof of Corollary 2.6). Let us denote by U_m the set of all the m-th roots of unity in $\overline{K}$ and denote by $L_{-1} = K_0 = K$, $L_0 = K_0[U_{j_{K_0}}]$, $L_1 = L_0[\alpha_{j_{K_0}}]$, $L_2 = L_1[U_{j_{K_1}}]$, $L_3 = L_2[\alpha_{j_{K_1}}], \ldots$ We see that

$$L_{-1} \subset L_0 \subset L_1 \subset \ldots \subset L_{2k+1},$$

where L_j/K is a normal extension and, from [3], Theorem 6.2, we see that

$Gal(L_j/L_{j-1})$ is a cyclic group for any $j = 0, 1, ..., 2k+1$. Thus $Gal(L_{2k+1}/K)$ is a solvable group. Moreover, the decomposition field L_P of P is contained in L_{2k+1}. Since

$$Gal(L_P/K) \simeq Gal(L_{2k+1}/K)/Gal(L_{2k+1}/L_P)$$

and since $Gal(L_{2k+1}/K)$ is a solvable group, we see that $Gal(L_P/K)$ is also a solvable group. □

References

[1] C. Fefferman, *An Easy Proof of the Fundamental Theorem of Algebra*, The American Mathematical Monthly, Vol. **74**, No. 7 (Aug.-Sept. 1967).

[2] G. Groza, A. Popescu, *Extensions of valued fields* (Romanian), Editura Academiei Române, Bucharest, 2011.

[3] S. Lang, *Algebra*, Springer-Verlag, New-York, 2002.

[4] N. Popescu, *Elements of Algebraic Number Theory* (Romanian), Bucharest University Press, 2021.

[5] J. P. Serre, Local Fields, Springer, 1979.

[6] A. Weil, *Basic Number Theory*, Springer, 1967.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, B-ul Lacul Tei 122, sector 2, Bucharest 020396, Bucharest, Romania.

Current address: Soseaua Stefan cel Mare 26, Bloc 24A, Sc. A, Ap. 20, Et. 5, Sector 2, Bucharest 020143, Romania.

Email address: angel.popescu@gmail.com

SEISMIC RESPONSE ANALYSIS OF BUILDINGS BY APPLICATION OF LAPLACE TRANSFORM

Narcisa Teodorescu
Technical University of Civil Engineering Bucharest
E-mail: narcisa.teodorescu@utcb.ro
Simona Cristina Nartea
Technical University of Civil Engineering Bucharest
E-mail: cristina.nartea@utcb.ro

Abstract: This article explores how the Laplace transform can be applied to the analysis of building motion and behaviour under earthquake action. The main steps of this analysis are presented, including the application of the Laplace transform to the differential equations, obtaining the transfer function and interpreting the building response in the Laplace domain.
Mathematics Subject Classification (2010): 44A10.
Key words: differential equation, Laplace transform.

1. Introduction

Structural safety is a major concern in the design and construction of buildings and infrastructure, and the assessment of structural behaviour under extreme conditions or emergency situations is crucial. The Laplace transform can be used in this context to analyse the structural response of buildings under different types of massive loads or in emergency situations such as earthquakes or other catastrophic events. A concrete example is the analysis of the behaviour of a building in case of a strong earthquake.

2. Theoretical Fundamentals

Analyzing the pole points and zeros of a transfer function in the Laplace domain is a crucial step in understanding the behavior of a system. The poles and zeros provide insights into the stability, frequency response, and transient characteristics of the system.
1.*Understand the Transfer Function*

The transfer function is represented in the Laplace domain as a ratio of polynomials in s (Laplace variable).

General form $H(s) = \frac{N(s)}{D(s)} = \frac{(s-z_1)(s-z_2)\dots(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_n)}$.

2.*Identify Zeros and Poles*

Zeros: The values of s that make the numerator $N(s)$ equal to zero. We denote these values by $z_1, z_2, \dots, z_m$

Poles: The values of s that make the denominator $D(s)$ equal to zero. We denote these values by $p_1, p_2, \dots, p_n$

3.*Plot Pole-Zero Diagram*

- Use a complex plane to plot the locations of zeros and poles.

- Represent zeros with "o" and poles with "x" on the plane.

4.*Analyze Pole-Zero Locations*

- *Stability:* For a system to be stable, all poles must have negative real parts.
- *Frequency Response:* Zeros and poles affect the system's response to different frequencies. Zeros influence amplification, while poles influence attenuation.

5.*Stability Analysis:*

- If all poles have negative real parts, the system is stable.
- If some poles have positive real parts, the system is unstable.
- If poles lie on the imaginary axis, the system may exhibit oscillatory behavior.

6.*Natural Frequency and Damping Factor:*

- Poles with real negative and imaginary non-zero sides can represent the natural modes of vibration of the system.
- The distance between poles and zeros, as well as the angles between them, can provide information about the damping factor and natural frequency of the system.

7.*Impact of Zeros:*

- Zeros indicate frequencies where the system responds strongly. They can affect the system's behavior in terms of amplification or attenuation.

8.*Impact of Poles:*

- Identify the dominant poles, which have the most significant influence on the system's behavior.
- Poles determine the system's time response characteristics, such as decay rate and oscillatory behavior.

9.*Optimization:*

- Engineers may optimize system performance by adjusting the locations of poles and zeros.

3. Mathematical Model and Examples

Suppose we have a building with a certain stiffness and structural characteristics, and an earthquake is modeled as a force or input motion acting on the building. The motion and behaviour of buildings under earthquake action can be described by a system of differential equations that model the dynamics of structures during an earthquake. These equations can be quite complex and depend on the specific characteristics of the building as well as the characteristics of the earthquake. A simplification of this model can be represented by a system of differential equations for motion in the vertical plane. We will consider the vertical movement of one floor of the building.

The equation of motion of the floor under the action of seismic forces can be expressed as a second order differential equation:

$$m\frac{d^2x}{dt^2} + c\frac{dx}{dt} + kx = F(t).$$

where

- The first term is the inertial force ($m\frac{d^2x}{dt^2}$)
- The second term is the damping force ($c\frac{dx}{dt}$)
- The third term is the restoring force due to structural stiffness (kx)
- $F(t)$ is the seismic force function of time.

This differential equation describes the motion and behaviour of the floor during an earthquake. It can be used to analyse the response of the building to earthquake action by solving the differential equation as a function of time.
To apply the Laplace transform to the differential equation describing the motion and behaviour of the building under earthquake action, we can follow the usual steps in the Laplace transform of each term of the equation.
Laplace transform steps

1. Apply Laplace transform on each term
- $\mathcal{L}\left(m\frac{d^2x}{dt^2}\right)$
- $\mathcal{L}\left(c\frac{dx}{dt}\right)$
- $\mathcal{L}(kx)$
- $\mathcal{L}\big(F(t)\big)$

2. Use the Laplace transform and noting the terms depending on the variables from Laplace domain
- $m \cdot s^2X(s) - msx(0) - mx'(0)$
- $c \cdot sX(s) - cx(0)$
- $k \cdot X(s)$
- $F(s)$

where $X(s)$ is the Laplace transform of the movement depending on time

$F(t)$ is the Laplace transform of the seismic force
s is the complex variable from the Laplace domain.

3. Replace those expressions in the original differential equation

$$m \cdot s^2X(s) - msx(0) - mx'(0) + c \cdot sX(s) - cx(0) + k \cdot X(s) = F(s)$$

4. Simplify and solve for $X(s)$

$$(ms^2 + cs + k)X(s) = F(s) + msx(0) + mx'(0) + cx(0)$$
$$X(s) = \frac{F(s) + msx(0) + mx'(0) + cx(0)}{ms^2 + cs + k}.$$

This transfer function can be used to analyse the response of the building in the frequency domain and to obtain useful information about vibration, resonance and stability modes under seismic conditions.

Example of analyzing the pole points and zeros of the transfer function in the Laplace domain

Transfer Function: $H(s) = \frac{s+2}{(s+1)(s+3)}$.

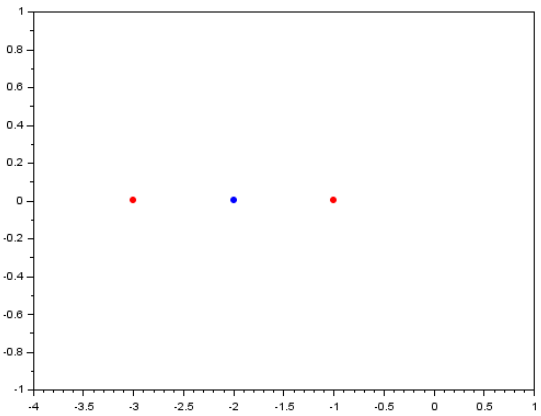
1.Identify Zeros and Poles

- Zeros (s_z): The zero occurs where the numerator $s + 2$ is equal to zero. $s_z = -2$.
- Poles (s_p): The poles occur where the denominator $(s + 1)(s + 3)$ is equal to zero. $s_p = -1, -3$.

2.Write Transfer Function in Factored Form

- $H(s) = \frac{s+2}{(s+1)(s+3)}$

3.Plot Pole-Zero Diagram



- Plot zeros with “blue”
- Plot poles with “red”

4.Analyze Pole-Zero Locations

- Stability: All poles ($s_p = -1, -3$) have negative real parts, so the system is stable.
- Frequency Response: Zeros at $s_z = -2$ may influence the system’s response.

5.Pole-Zero Ratio

- The system has two poles and one zero, making it a first-order system.

6.Region of convergence (ROC)

- The ROC is to the right of the rightmost, which is $s = -3$.

7. Stability Analysis

- The system is stable as all poles have negative real parts.

8. Dominant poles

- In this case, the dominant poles are the ones with the most negative real parts which is $s = -3$.

9.Transfer function properties

- Evaluate properties like damping ratio, natural frequency and setting time based on pole locations.

10.Impact of zeros

- The zero at $s = -2$ indicates a frequency where the system may respond strongly.

11.Impact of poles

- The poles at s= $-1, -3$ determine the system's time response characteristics.

12.Optimization

- Engineers may optimize system performance by adjusting the locations of poles and zeros.

13.Simulation and Validation

- Simulate the system's response using software tools and validate the analysis against experimental data.

4. Conclusions

The application of the Laplace Transform in seismic response analysis of buildings provides a robust framework for understanding and predicting the dynamic behavior of structures under earthquake excitations. By transforming the time-domain differential equations governing structural dynamics into the Laplace domain, engineers can leverage the mathematical simplicity of algebraic equations to analyze complex systems more effectively. Here are the key takeaways from using the above method:

- The Laplace Transform simplifies the solution of the system's differential equations, converting them into algebraic equations that are easier to handle and solve. This is particularly useful for linear time-invariant systems where the transform aids in deriving the system's transfer function.
- The analysis of poles and zeros in the Laplace domain allows for a thorough assessment of the system's stability. All poles with negative real parts indicate a stable system, which is essential for ensuring the structural integrity under seismic loads. The frequency response analysis reveals how the building will behave across different seismic frequencies, identifying potential resonant frequencies that could amplify the response.
- The insights gained from the Laplace Transform analysis contribute to improved structural safety. By understanding how buildings respond to seismic excitations, engineers can design structures that are better equipped to withstand earthquakes, minimizing the risk of catastrophic failures.

In summary, while the Laplace Transform is a powerful tool for structural analysis, its application should be done with careful consideration of system linearity, model assumptions, and the specific challenges posed by nonlinear and dynamic structural behavior. Integrating

Laplace Transform analysis with experimental data and other analytical methods provides a more comprehensive approach to ensuring structural safety.

References

[1] Dai-lin Li, Jia-jian Lin, Guo-dong Shi, Jun-liang Zhang, Hong Li, Xin Zhang: Research on Laplace transform of stress wave propagation relaxation function in multi-layer media, *Case Studies in Construction Materials* **18** (2023), pp. 1-13.
[2] Ibrahim, R.A.: *Structural Dynamics: Applications in Earthquake and Wind Engineering*. Boca Raton, FL: CRC Press, 2009.
[3] Ileana Toma, Viorel Petrehus, Narcisa Teodorescu, Iuliana Popescu, *Matematici Speciale, curs si aplicatii*, Editura Conspress, ISBN 978-973-100-243-9, Bucuresti, (350pp), 2012.

DRIFT ANALYSIS OF THE GAMBLER'S RUIN PROBLEM

DANIEL TUDOR, DAN ANDREI TUDOR AND DAN CARAGHEORGHEOPOL

Abstract. We develop drift criteria for the state classification of an irreducible Markov Chain, and we apply these to the well-known Gambler's Ruin Problem.

Mathematics Subject Classification (2010): 60J10.
Key words: markov chains, drift conditions, Foster-Lyapunov functions.

1. The Gambler's Ruin problem and Markov Chains

The Gambler's ruin problem poses the following game: A player flips a possibly biased coin and every time the coin lands on heads, the player is awarded with one dollar. If the coin lands on tails, the player loses one dollar. If at any point the player is left without any dollars, the overseer of the game immediately gives the player one dollar, so that the player can keep playing the game. This game can be described as a Markov Chain with state space $S = \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and transition probabilities $p_{i,i+1} = p$, $p_{i,i-1} = q = 1 - p$ for $i \geq 1$ and $p_{01} = 1$. We shall assume that $p, q \in (0, 1)$. This chain is *irreducible* as it can go with positive probability from any state to any other state in some finite number of steps.

According to [4], a state i is recurrent if $\mathbb{P}(T_i < \infty | X_0 = i) = 1$ and transient otherwise, where $T_i := \inf\{n \geq 1 : X_n = i\}$. It is called *positive* recurrent if it is recurrent and $\mathbb{E}(T_i | X_0 = i) < \infty$ and it is called *null* recurrent if it is recurrent and $\mathbb{E}(T_i | X_0 = i) = \infty$. In the case of an *irreducible* Markov Chain, the states are either all *(positive/null) recurrent* or all *transient*. In our current paper, we aim to describe the chain's structure with the aid of local conditions at each state, referring to the idea of *drift*. The following section is devoted to introducing this concept and the criteria for the chain's state structure. Throughout the paper, we assume that the sequence of random variables $(X_n)_{n \geq 0}$ is an irreducible Markov Chain.

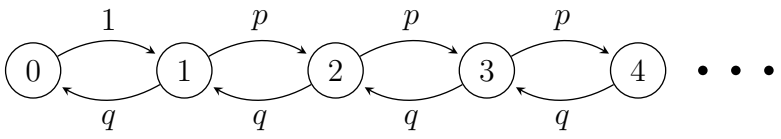


Figure 1. State transition diagram of the Gambler's Ruin problem

2. Theoretical Background

We start this section by stating without proof three important general theorems that regard the stability of the Markov Chain. These two theorems will then be reshaped to be able to discuss the problem at hand easily. We assume the Markov Chain is irreducible and is defined on an arbitrary state space S for the scope of these two theorems. We are however interested only on the theory when $S = \mathbb{N}_0$.

Theorem 2.1 (Foster [2])**.** *The Markov Chain is positive recurrent if there exists some non-negative function $f : S \to \mathbb{R}_{\geq 0}$ such that*

$$\sum_{j \in S} p_{ij} f(j) \leq f(i) - \epsilon \qquad \text{for } i \notin A, \tag{2.1}$$

$$\sum_{j \in S} p_{ij} f(j) < \infty \qquad \text{for } i \in A, \tag{2.2}$$

for some $\epsilon > 0$ and finite set A.

Theorem 2.2 (Fayolle-Malyshev-Menshikov [1])**.** *The Markov Chain is transient if there exists some non-negative function f such that*

$$\sum_{j \in S} p_{ij} f(j) \geq f(i) + \epsilon \qquad \text{for } i \notin A \tag{2.3}$$

for some $\epsilon > 0$ and set A, and for some $d > 0$, $|f(i) - f(j)| > d$ implies that $p_{ij} = 0$.

Theorem 2.3 (Mertens et al. [5])**.** *The Markov Chain is recurrent if there exists some function $f : S \to \mathbb{R}$ such that*

$$\sum_{j \in S} p_{ij} f(j) \leq f(i) \qquad \text{for } i \notin A, \tag{2.4}$$

for some finite set A, and $\lim_{i \to \infty} f(i) = \infty$.

We now turn our attention specifically to the case when $S = \mathbb{N}_0$. We can now introduce the concept of a drift, which arises in the upcoming criteria.

Definition 2.4. *The drift at state i is defined as $\gamma_i := \sum_{j \in \mathbb{N}_0} p_{ij}(j - i)$.*

A sufficient condition for positive recurrence, using Foster's theorem, is given by Pakes.

Lemma 2.5 (Pakes' Lemma [6])**.** *The Markov Chain is positive recurrent if $\gamma_i < \infty$ for all $i \in \mathbb{N}_0$ and there exists $\epsilon > 0$ such that $\gamma_i \leq -\epsilon$ for all $i \geq N$, for some $N \in \mathbb{N}_0$.*

Proof. We have, $\sum_{j \in \mathbb{N}_0} p_{ij}\, j \leq i - \epsilon$, for all $i \geq N$. The first condition of the lemma ensures that $\sum_{j \in \mathbb{N}_0} p_{ij}\, j < \infty$ for all $i \in \mathbb{N}_0$. Therefore, the conditions of Foster's theorem are satisfied with $f(i) = i$ and $A = \{0, 1, \ldots, N - 1\}$, so the chain is positive recurrent. □

The following theorem is presented here without proof. We will use it to showcase a null recurrence criterion involving drifts.

Theorem 2.6 (Kaplan's criterion [3])**.** *The Markov Chain is not positive recurrent if $\gamma_i < \infty$ for all $i \geq 0$, for some $N \in \mathbb{N}$, $\gamma_i \geq 0$ for all $i \geq N$, and for some $d > 0$, $i - j > d$ implies that $p_{ij} = 0$.*

Let us now, using the above theorems, construct a local criterion for null-recurrence using drifts.

Proposition 2.7. *The Markov Chain is null recurrent if $\gamma_i < \infty$ for all $i \geq 0$, there exists $N \in \mathbb{N}$ such that $\gamma_i = 0$ for all $i \geq N$, and for some $d > 0$, $i - j > d$ implies that $p_{ij} = 0$.*

Proof. The statement follows from Kaplan's criterion and Theorem 2.3 applied with $f(i) = i$ for all i and $A = \{0, 1, \ldots, N-1\}$. □

Finally, by combining some of the above theorems, let us now showcase a criterion for transience involving drifts.

Proposition 2.8. *The Markov Chain is transient if $\gamma_i < \infty$ for all $i \geq 0$, there exists $N \in \mathbb{N}$ such that $\gamma_i = 0$ for all $i \geq N$, and for some $d > 0$, $|i - j| > d$ implies that $p_{ij} = 0$.*

Proof. This is immediate from Kaplan's criterion and Theorem 2.2 taken with $f(i) = i$ for all i and $A = \{0, 1, \ldots, N-1\}$. □

3. Structure of the Gambler's Ruin Problem

We can now turn our attention back to the Gambler's Ruin Problem. We now have all the necessary tools to completely describe its behaviour. A quick computation of the drift at i yields $\gamma_i = p - q$ for all $i \geq 1$ and $\gamma_0 = 1$.

If $p < q$, then in view of Lemma 2.5, the chain is positive recurrent, by taking $N = 1$ and $\epsilon = q - p > 0$.

If $p = q$, then $\gamma_i = 0$ for all $i \geq 1$. In addition, if $i - j > 1$, we have $p_{ij} = 0$. Therefore, in view of Proposition 2.7, the chain is null recurrent.

If $p > q$, then, since $|i - j| > 1$ implies $p_{ij} = 0$, we apply Proposition 2.8 with $\epsilon = p - q$ and $N = 1$ to conclude that the chain is transient.

References

[1] G. Fayolle, V. Malyshev, M. Menshikov, Topics in the Constructive Theory of Countable Markov Chains, 1995. 10.1017/CBO9780511984020.

[2] F.G. Foster, *On the Stochastic Matrices Associated with Certain Queuing Processes*, The Annals of Mathematical Statistics, Ann. Math. Statist. 24(3), pp. 355-360, 1953.

[3] M. Kaplan, *A sufficient condition of nonergodicity of a Markov chain (Corresp.)*, in IEEE Transactions on Information Theory, vol. 25, no. 4, pp. 470-471, 1979.

[4] J.R. Norris, *Markov Chains*, ch. 1, 1997.

[5] J.-F. Mertens, E. Samuel-Cahn, and S. Zamir, *Necessary and sufficient conditions for recurrence and transience of Markov Chains, in terms of inequalities*, Journal of Applied Probability, vol. 15, no. 4, pp. 848–851, 1978.

[6] A.G. Pakes, *Some conditions for ergodicity and recurrence of Markov Chains*, Operations Res 17, pp. 1058-1061, 1969.

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: daniel.tudor@utcb.ro

Graduate in Mathematics, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
Email address: d.tudor@student.vu.nl

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: dancaraghe@gmail.com

EXPLICIT CONSTRUCTIONS OF MARKOV CHAINS FROM UNIFORM RANDOM VARIABLES

DAN ANDREI TUDOR, DAN CARAGHEORGHEOPOL AND DANIEL TUDOR

Abstract. We give an explicit measure-theoretic construction of discrete-time Markov Chains with arbitrary initial distributions and transition matrices, starting from independent uniform random variables.

Mathematics Subject Classification (2010): 60J10.
Key words: markov chains, uniform random variables, measure theory.

1. Introduction and background

In the present paper, we aim to construct a discrete-time Markov Chain on an explicit probability space. The main idea is that we sample from a sequence of independent uniform random variables, and each realization determines the following jump of the Markov Chain. The current section gives two background results: we give the definition of a Markov Chain, in accordance with [2] and show that there exists a probability space allowing for an infinite sequence of independent uniform random variables [1, Theorem 4.11.1].

First, let us define what a discrete-time Markov Chain is. Consider a *countable* set S, which we call the *state space*. Any $i \in S$ is called a *state*. The vector $\boldsymbol{\lambda} = (\lambda_i : i \in S)$ is a *distribution* on S if $\lambda_i \geq 0$ for all $i \in S$ and $\sum_{i \in S} \lambda_i = 1$. A matrix $P = (p_{ij} : i, j \in S)$ is called a *transition matrix* if every row $(p_{ij} : j \in S)$ is a distribution. With these, we can now define a discrete-time Markov Chain.

Definition 1.1. *Let S be a state space, $\boldsymbol{\lambda}$ a distribution, and P a transition matrix on S. Then a sequence of random variables $(X_n)_{n \geq 0}$ is a Markov Chain with initial distribution $\boldsymbol{\lambda}$ and transition matrix P, denoted as Markov$(\boldsymbol{\lambda}, P)$, if for all $n \in \mathbb{N}_0$ and $i_0, \ldots, i_n \in S$,*

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) = \lambda_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n}.$$

The following theorem ensures the existence of a probability space allowing for an infinite sequence of independent uniform random variables [1, Theorem 4.11.1].

Theorem 1.2. *Let $([0,1), \mathscr{B}([0,1)), \lambda)$ be the underlying probability space of one uniform random variable on $[0,1)$, where $\mathscr{B}([0,1))$ is the Borel σ-algebra on $[0,1)$ and λ is the Lebesgue measure. Then on the measurable space $([0,1)^\infty, \mathscr{B}([0,1))^\infty)$, there exists a unique probability measure $\mathbb{P}$ and a sequence of independent uniform random variables $(U_n)_{n \geq 0}$, each taking values from $([0,1)^\infty, \mathscr{B}([0,1))^\infty)$ to $([0,1), \mathscr{B}([0,1)))$, such that*

$$\mathbb{P}(U_0 \in B_0, \ldots, U_n \in B_n) = \prod_{i=0}^{n} \lambda(B_i)$$

for all $B_i \in \mathscr{B}([0,1))$ *and all* $n \in \mathbb{N}_0$.

The above theorem is presented in more generality in [1], and actually ensures the existence of an infinite sequence of independent random variables that follow any desired distribution. For our present work, we shall stick to an infinite sequence of independent uniform random variables.

2. Explicit construction

Now that we have established the existence of a probability space that supports a sequence of independent uniform random variables on $[0,1)$, we can now define, on the probability space $([0,1)^\infty, \mathscr{B}([0,1))^\infty, \mathbb{P})$, where $\mathbb{P}$ is given by Theorem 1.2, a Markov Chain. We assume without loss of generality that $S = \mathbb{N}_0$.

Suppose we are given an initial distribution $\boldsymbol{\lambda}$ and a transition matrix P. Define intervals

$$\begin{aligned} \Lambda_0 &= [0, \lambda_0) \\ \Lambda_i &= \left[\sum_{k=0}^{i-1} \lambda_k, \sum_{k=0}^{i} \lambda_k\right) \qquad \text{for } i = 1, 2, 3, \ldots \end{aligned}$$

Similarly, for each $i \in \mathbb{N}_0$, define intervals

$$\begin{aligned} A_{i0} &= [0, p_{i0}) \\ A_{ij} &= \left[\sum_{k=0}^{j-1} p_{ik}, \sum_{k=0}^{j} p_{ik}\right) \qquad \text{for } j = 1, 2, 3, \ldots \end{aligned}$$

Note that $\lambda(\Lambda_i) = \lambda_i$ for all $i \in \mathbb{N}_0$ and $\lambda(A_{ij}) = p_{ij}$ for all $i, j \in \mathbb{N}_0$. Define

$$\begin{aligned} X_0 &= \sum_{i_0 \in S} i_0 \mathbb{1}_{\{U_0 \in \Lambda_{i_0}\}}, \\ X_1 &= \sum_{i_1 \in S} \sum_{i_0 \in S} i_1 \mathbb{1}_{\{U_1 \in A_{i_0 i_1}\}} \mathbb{1}_{\{U_0 \in \Lambda_{i_0}\}}, \\ &\cdots\cdots\cdots \\ X_n &= \sum_{i_n \in S} \sum_{i_{n-1} \in S} \cdots \sum_{i_0 \in S} i_n \mathbb{1}_{\{U_n \in A_{i_{n-1} i_n}\}} \mathbb{1}_{\{U_{n-1} \in A_{i_{n-2} i_{n-1}}\}} \cdots \mathbb{1}_{\{U_0 \in \Lambda_{i_0}\}}, \\ &\cdots\cdots\cdots \end{aligned} \tag{2.1}$$

for all $n \in \mathbb{N}_0$. One can see that each interval Λ_i and A_{ij} is in $\mathscr{B}([0,1))$, so each indicator function is measurable. The product and sum of measurable functions are measurable, and so is the limit of a sequence of measurable functions. Thus each X_n is measurable, making the random variables well-defined.

Each of the random variables now lives on the probability space $([0,1)^\infty, \mathscr{B}([0,1))^\infty, \mathbb{P})$ and takes values into the state space $S = \mathbb{N}_0$. We now show that this sequence of

random variables satisfies Definition 1.1 of a Markov Chain. For every $n \in \mathbb{N}_0$, and every $i_0, \ldots, i_n \in S$, we have

$$\begin{aligned}
&\mathbb{P}(X_0 = i_0, X_1 = i_1, \ldots, X_n = i_n) \\
&\qquad \overset{(2.1)}{=} \mathbb{P}(U_0 \in \Lambda_{i_0}, U_1 \in A_{i_0 i_1}, \ldots, U_n \in A_{i_{n-1} i_n}) \\
&\qquad = \mathbb{P}(U_0 \in \Lambda_{i_0}) \cdot \mathbb{P}(U_1 \in A_{i_0 i_1}) \cdot \cdots \cdot \mathbb{P}(U_n \in A_{i_{n-1} i_n}) \\
&\qquad = \lambda(\Lambda_{i_0}) \lambda(A_{i_0 i_1}) \ldots \lambda(A_{i_{n-1} i_n}) \\
&\qquad = \lambda_{i_0} p_{i_0 i_1} \ldots p_{i_{n-1} i_n},
\end{aligned}$$

where we use the independence of the uniform random variables. This shows that $(X_n)_{n \geq 0}$ is Markov$(\boldsymbol{\lambda}, P)$. As no assumptions have been made on the initial distribution $\boldsymbol{\lambda}$ and the transition matrix P, the above provides an explicit construction for all Markov Chains defined on a countably infinite state space. If the state space is finite, the construction can be easily adjusted by running the intervals only on the finite state space.

If one is instead given only a state space S (in our case $\mathbb{N}_0$) and the interval $[0, 1)$ is partitioned into measurable sets Λ_i for $i \in S$, and also into measurable sets A_{ij} for each $i, j \in S$, one can still construct a Markov Chain in the same way. The initial distribution is derived as $\boldsymbol{\lambda} = (\lambda(\Lambda_i) : i \in S)$ and the transition matrix as $P = (\lambda(A_{ij}) : i, j \in S)$. Then, defining the random variables as in (2.1) ensures that the sequence is Markov$(\boldsymbol{\lambda}, P)$.

3. Conclusive remarks

With the present paper, we have shown explicitly how to construct a Markov Chain on a probability space that allows for a sequence of independent uniform random variable, as seen in Theorem 1.2. This explicit construction is the basis of computer simulation of Markov Chains, as most statistical packages have a uniform random variable sampler.

References

[1] R.B. Ash and C. Doleans-Dade, *Probability and Measure Theory*, 2nd edition, 2000.
[2] J.R. Norris, *Markov Chains*, ch. 1, 1997.

Graduate in Mathematics, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
Email address: d.tudor@student.vu.nl

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: dancaraghe@gmail.com

Department of Mathematics and Computer Science, Technical University of Civil Engineering Bucharest, Bucharest, Romania
Email address: daniel.tudor@utcb.ro

AN INTRODUCTION TO SOLVING EULER SPHERICAL TRIANGLES WITH MATHCAD

Mariana Zamfir
Department of Mathematics and Computer Science
Technical University of Civil Engineering Bucharest, Romania
E-mail: zamfirvmariana@yahoo.com

Abstract: Engineers use the solution of spherical triangles every day in their work of constructing buildings, bridges and roads; astronomers use it to determine altitude and azimuth of a star, or to find amplitude and local apparent time of sunrise and sunset for a point on the earth; surveyors use it constantly to find all sorts of heights, distances and directions; and navigators use it to compute latitude, longitude and course at sea.
Mathematics Subject Classification (2010): 51M04; 14J29.
Key words: great circle; small circle; spherical triangle; solution of spherical triangle; MathCAD.

1. Great and small circles on a sphere. Spherical triangles on a sphere

The present work is constructed to show how MathCAD can be used as an alternative teaching and computing tool to explore a chapter of Geometry, "Spherical Trigonometry".

On beginning this paper, it is advisable to remind several elementary definitions and notations of spherical geometry theory. We hope that this short overview will be further useful even for the less experienced reader.

A *sphere* is a solid bounded by a surface of which all the points are equally distant from a fixed interior point called the *center* of the sphere. A straight line which joins any point of the sphere to the center is called *radius*. A straight line passed through the center and terminated at both ends by the sphere is called *diameter*.

A *plane section* of a sphere is a figure whose boundary is the intersection of a plane with a sphere. The boundary of every plane section of a sphere is a *circle*. The circle in which a sphere is cut by a plane is known as a *great circle* when the plane passes through the center of the sphere, and as a *small circle* when the cutting plane does not pass through the center of the sphere. Therefore, the centre and the radius of every great circle on a sphere are the same as the center and the radius of its sphere.

Since the shape of the Earth is approximately that of a sphere having radius $R \approx 6371$ km, the Equator and the meridians on the terrestrial globe are great circles and the parallels of latitude are small circles.

In general, only one great circle can be passed through two given points on a sphere, since these points and the center of the sphere determine a plane. However, when the two points are the extremities of a diameter of the sphere an infinite number of great circles can be drawn through them (e.g. the meridians passing through the North and the South Poles).

When only one great circle can be drawn through two given points on a sphere, the two points divide the great circle into two arcs of differing lengths. The length of the shorter arc of the great circle is the shortest distance that can be drawn on the sphere from the one point to the other.

On an arbitrary surface, *geodesics* are simply the curves corresponding to the shortest length between two given points on the surface. On a sphere, great circles define geodesics, because the shortest path between any two points on the sphere is the distance measured along the arc of the great circle between them.

A small circle always has smaller radius than its sphere, and an arc of small circle does not represent the shortest path between two points on a sphere.

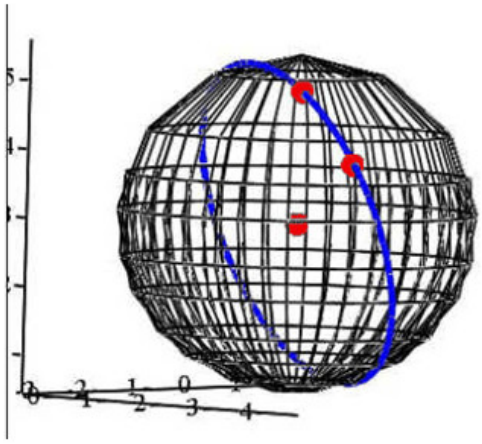


Fig. 1.1. Great circle

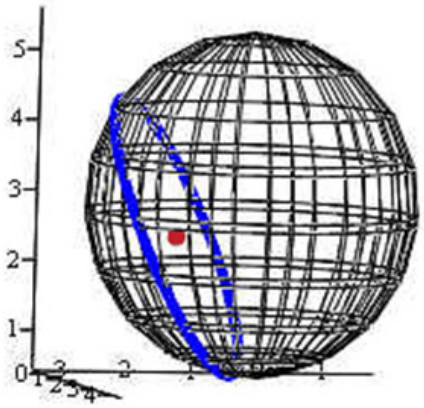


Fig. 1.2. Small circle

A *spherical triangle* is a portion formed on a sphere which is bounded by three arcs of great circles. The bounding arcs of great circles which form a spherical triangle are the *sides* of the triangle; the points of intersection of the sides are the *vertices* of the triangle; the angles formed by the sides at the points where they meet are the *angles* of the spherical triangle.

The magnitude of an angle of a spherical triangle is that of the plane angle formed by the two tangents to the sides of the angle at its vertex and it is expressed in angular measure (degrees-minutes-seconds or radians). The same as in plane geometry, an arc of a great circle is measured by the angle that it subtends at the center of its sphere. Hence, the magnitude of a side of a spherical triangle will be expressed in angular measure (degrees-minutes-seconds or radians) and not in linear measure (meters or kilometers).

As in the case of plane triangles, the upper-case letters A, B, C are generally used to denote the vertices (and angles) of a spherical triangle, and the lower-case letters a, b, c are used to denote the corresponding sides (arcs of great circles) of a spherical triangle.

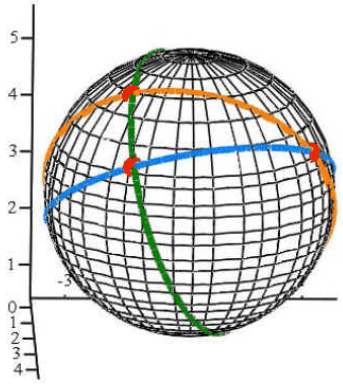


Fig. 1.3. Spherical triangle

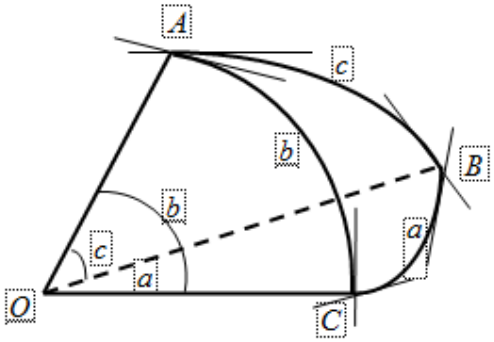


Fig. 1.4. Six elements of spherical triangle

A spherical triangle can be:

- *right-angled triangle* (one of its angles is equal to 90°).
- *quadrantal triangle* (one of its sides is equal to 90°).
- *oblique triangle* (not a right-angled or a quadrantal one).

- *isosceles triangle* (two of the sides and two of the angles opposite are equal).
- *equilateral triangle* (all three sides and all three angles are equal).

Observation 1.1. A spherical triangle may have two right angles, three right angles, two obtuse angles, three obtuse angles, two sides each a quadrant, or three sides equal to a quadrant. For example, on the terrestrial globe, the spherical triangle bounded by an arc on the Equator which is equal to a quadrant and by two arcs of the meridians joining the extremities of the former arc to the North Pole, has all three right angles and each of its three sides equal to a quadrant.

In general, we shall consider only *Euler spherical triangles* formed on a sphere, for which each of the sides and each of the angles are less than 180°.

Proposition 1.1. ([3], [4]) For an Euler spherical triangle ABC formed on a sphere, the following important properties are established:

1) $a,\ b,\ c < 180^\circ;\ A,\ B,\ C < 180^\circ$.

2) $0^\circ < a + b + c < 360^\circ$.

3) $180^\circ < A + B + C < 540^\circ$.

4) If $a < b < c$, then $A < B < C$.

5) The sum of two sides is greater than the third side ($a + b > c,\ a + c > b,\ b + c > a$).

6) If two angles are equal, the opposite sides are equal; and conversely.

7) If two angles are unequal, the opposite sides are unequal, and the greater side is opposite to the greater angle; and conversely.

8) If the triangle is isosceles, the angles at its base are equal.

2. Solution of oblique spherical triangles

Two spherical triangles can be *symmetrically equal* if each of the six elements (three sides and three angles) of one of them is equal to each of the six elements of the other. Since spherical triangles are three dimensional figures on a sphere, symmetrically equal does not necessarily mean congruent. Two spherical triangles are *congruent* only if they can be applied one to the other so as to coincide.

Proposition 2.1. ([4]) Two spherical triangles lied on the same sphere, or on equal spheres, are symmetrically equal if they have three of their elements equal:

- Three sides of the one are respectively equal to the three sides of the other.
- Three angles of the one are respectively equal to the three angles of the other.
- Two sides and their included angle of the one are respectively equal to two sides and the included angle of the other.
- Two angles and their included side of the one are respectively equal to two angles and the included side of the other.
- Two sides and the angle opposite to one of the sides of a triangle are respectively equal to two sides and the angle opposite to one of the sides of the other triangle.
- Two angles and the opposite side to one of the angles of a triangle are respectively equal to two angles and the opposite side to one of the angles of the other triangle.

Spherical trigonometry treats of the various relations connecting the sides and angles of a spherical triangle. In every spherical triangle on a sphere the three sides and the three angles constitute its six parts or elements. The *solution of spherical triangles* is an important process of spherical trigonometry by which, when the values of any three of these six elements are known, the values of the other three remaining elements can generally be found by formulas of spherical trigonometry, which are given next.

We now proceed to the solution of oblique spherical triangles in general. Accordingly, six cases are considered as follows:

1) Case 1: Given the three sides (**Case SSS**).
2) Case 2: Given the three angles (**Case AAA**).
3) Case 3: Given two sides and their included angle (**Case SAS**).
4) Case 4: Given two angles and their included side (**Case ASA**).
5) Case 5: Given two sides and an angle opposite one of them (**Case SSA**).
6) Case 6: Given two angles and a side opposite one of them (**Case AAS**).

Observation 2.1. To solve spherical triangles for which the given parts are three elements, the only cases in which there is any uncertainty are the last two (Case SSA and Case AAS). When an element of a spherical triangle is determined using the Law of Sines (Case SSA and Case AAS), there is often some difficulty in determining whether the element found is of the first quadrant or of the second quadrant, hence if we have double solutions, or if we have only one solution, or if we have no solution. We proceed to show in this paper what conditions are necessary that there may be two triangles, or only one triangle, or even when the spherical triangle is impossible.

We will expose these ambiguous cases in spherical geometry in the following theorems:

Theorem 2.1. (**Case SSA**) ([4])

To construct a spherical triangle ABC when two sides $BC = a$, $AC = b$, and the angle A opposite to a are the known parts, the following situations may be possible:

1) If $\sin a < \sin b \sin A$, then $\sin B = \dfrac{\sin b \sin A}{\sin a} > 1$ (i.e. B has an impossible value) and there is no spherical triangle that satisfies the given conditions.

2) If $\sin a = \sin b \sin A$, then $\sin B = \dfrac{\sin b \sin A}{\sin a} = 1$ (i.e. $B = 90^\circ$) and there is only one triangle ABC which has the given elements.

3) If $\sin a > \sin b \sin A$, then $\sin B = \dfrac{\sin b \sin A}{\sin a} < 1$ (i.e. there are two admissible values of B obtained by the above equality: one a positive acute angle B and the other its supplement, $180^\circ - B$, or there is no value of B) and it follows:

a) when $A < 90^\circ$, then:

- there are two spherical triangles if $a < b$ and $a < 180^\circ - b$
- there is only one spherical triangle if a lies between b and $180^\circ - b$
- there is no spherical triangle if $a > b$ and $a > 180^\circ - b$

b) when $A > 90^\circ$, then:

- there are two spherical triangles if $a > b$ and $a > 180^\circ - b$
- there is only one spherical triangle if a lies between b and $180^\circ - b$
- there is no spherical triangle if $a < b$ and $a < 180^\circ - b$.

Remark 2.1. ([4]) When $\sin a > \sin b \sin A$, if both values B and $180^\circ - B$ be such that $A - B$ and $a - b$ have like signs, then there are two spherical triangles which satisfy the given conditions; if only one of the values of B verifies the above condition, there is only one triangle; if neither of the values of B satisfies this condition, then the triangle is impossible.

The following case is quite analogous to the first, and is treated in the same manner, and gives rise to the same ambiguities:

Theorem 2.2. (**Case AAS**) ([4])

To solve a spherical triangle ABC when two angles A, B, and the side $BC = a$ opposite to A are given, the following situations may be possible:

1) If $\sin A < \sin a \sin B$, then $\sin b = \dfrac{\sin a \sin B}{\sin A} > 1$ (i.e. the side b does not exist) and there is no spherical triangle which satisfies the given conditions.

2) If $\sin A = \sin a \sin B$, then $\sin b = \dfrac{\sin a \sin B}{\sin A} = 1$ (i.e. $b = 90^\circ$) and there is only one triangle ABC which has the given elements.

3) If $\sin A > \sin a \sin B$, then $\sin b = \dfrac{\sin a \sin B}{\sin A} < 1$ (i.e. there are two admissible values of b furnished by the above equality: one a positive acute side b and the other its supplement, $180^\circ - b$, or there is no value of b) and it follows:

a) when $a < 90^\circ$, then:

- there are two spherical triangles if $A < B$ and $A < 180^\circ - B$
- there is only one spherical triangle if A lies between B and $180^\circ - B$
- the triangle required is impossible if $A > B$ and $A > 180^\circ - B$

b) when $a > 90^\circ$, then:

- there are two spherical triangles if $A > B$ and $A > 180^\circ - B$
- there is only one spherical triangle if A lies between B and $180^\circ - B$
- the triangle required is impossible if $A < B$ and $A < 180^\circ - B$.

3. Relations between the trigonometrical functions of the sides and the angles of a spherical triangle

In this section some relations between the sides and the angles of any spherical triangle will be exposed. These relations will be used in the solution of practical numerical problems.

The Spherical Law of Cosines for sides (are relations between three sides and an angle; [3], [4]):

$$\begin{aligned} \cos(a) &= \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A) \\ \cos(b) &= \cos(a)\cos(c) + \sin(a)\sin(c)\cos(B) \\ \cos(c) &= \cos(a)\cos(b) + \sin(a)\sin(b)\cos(C) \end{aligned}.$$

The Spherical Law of Cosines for angles (are relations between three angles and a side; [3]. [4]):

$$\begin{aligned} \cos(A) &= -\cos(B)\cos(C) + \sin(B)\sin(C)\cos(a) \\ \cos(B) &= -\cos(A)\cos(C) + \sin(A)\sin(C)\cos(b) \\ \cos(C) &= -\cos(A)\cos(B) + \sin(A)\sin(B)\cos(c) \end{aligned}.$$

The Spherical Law of Sines (are relations between two angles and tow sides; [3], [4]):

$$\frac{\sin(a)}{\sin(A)} = \frac{\sin(b)}{\sin(B)} = \frac{\sin(c)}{\sin(C)}.$$

The Cotangent four-part formulas (are relations between two angles and tow sides; [3]):

$$\begin{aligned}
\sin(A)\cot(B) &= \sin(c)\cot(b) - \cos(c)\cos(A)\\
\sin(A)\cot(C) &= \sin(b)\cot(c) - \cos(b)\cos(A)\\
\sin(B)\cot(A) &= \sin(c)\cot(a) - \cos(c)\cos(B)\\
\sin(B)\cot(C) &= \sin(a)\cot(c) - \cos(a)\cos(B)\\
\sin(C)\cot(A) &= \sin(b)\cot(a) - \cos(b)\cos(C)\\
\sin(C)\cot(B) &= \sin(a)\cot(b) - \cos(a)\cos(C)
\end{aligned} .$$

Napier's Analogies (are relations between five elements; [3], [4]):

$$\tan\left(\frac{a+b}{2}\right) = \tan\left(\frac{c}{2}\right)\cdot\frac{\cos\left(\frac{A-B}{2}\right)}{\cos\left(\frac{A+B}{2}\right)} \qquad \tan\left(\frac{a-b}{2}\right) = \tan\left(\frac{c}{2}\right)\cdot\frac{\sin\left(\frac{A-B}{2}\right)}{\sin\left(\frac{A+B}{2}\right)}$$

$$\tan\left(\frac{A+B}{2}\right) = \cot\left(\frac{C}{2}\right)\cdot\frac{\cos\left(\frac{a-b}{2}\right)}{\cos\left(\frac{a+b}{2}\right)} \qquad \tan\left(\frac{A-B}{2}\right) = \cot\left(\frac{C}{2}\right)\cdot\frac{\sin\left(\frac{a-b}{2}\right)}{\sin\left(\frac{a+b}{2}\right)}$$

$$\tan\left(\frac{b+c}{2}\right) = \tan\left(\frac{a}{2}\right)\cdot\frac{\cos\left(\frac{B-C}{2}\right)}{\cos\left(\frac{B+C}{2}\right)} \qquad \tan\left(\frac{b-c}{2}\right) = \tan\left(\frac{a}{2}\right)\cdot\frac{\sin\left(\frac{B-C}{2}\right)}{\sin\left(\frac{B+C}{2}\right)}$$

$$\tan\left(\frac{B+C}{2}\right) = \cot\left(\frac{A}{2}\right)\cdot\frac{\cos\left(\frac{b-c}{2}\right)}{\cos\left(\frac{b+c}{2}\right)} \qquad \tan\left(\frac{B-C}{2}\right) = \cot\left(\frac{A}{2}\right)\cdot\frac{\sin\left(\frac{b-c}{2}\right)}{\sin\left(\frac{b+c}{2}\right)}$$

$$\tan\left(\frac{c+a}{2}\right) = \tan\left(\frac{b}{2}\right)\cdot\frac{\cos\left(\frac{C-A}{2}\right)}{\cos\left(\frac{C+A}{2}\right)} \qquad \tan\left(\frac{c-a}{2}\right) = \tan\left(\frac{b}{2}\right)\cdot\frac{\sin\left(\frac{C-A}{2}\right)}{\sin\left(\frac{C+A}{2}\right)}$$

$$\tan\left(\frac{C+A}{2}\right) = \cot\left(\frac{B}{2}\right)\cdot\frac{\cos\left(\frac{c-a}{2}\right)}{\cos\left(\frac{c+a}{2}\right)} \qquad \tan\left(\frac{C-A}{2}\right) = \cot\left(\frac{B}{2}\right)\cdot\frac{\sin\left(\frac{c-a}{2}\right)}{\sin\left(\frac{c+a}{2}\right)} .$$

Delambre-Gauss six-part rules (are relations between all the six elements of the triangle and may be used as check formulas; [3], [4]). The results obtained should always be checked:

$$\sin\left(\frac{A}{2}\right)\sin\left(\frac{b+c}{2}\right) = \sin\left(\frac{a}{2}\right)\cos\left(\frac{B-C}{2}\right) \qquad \sin\left(\frac{A}{2}\right)\cos\left(\frac{b+c}{2}\right) = \cos\left(\frac{a}{2}\right)\cos\left(\frac{B+C}{2}\right)$$

$$\cos\left(\frac{A}{2}\right)\sin\left(\frac{b-c}{2}\right) = \sin\left(\frac{a}{2}\right)\sin\left(\frac{B-C}{2}\right) \qquad \cos\left(\frac{A}{2}\right)\cos\left(\frac{b-c}{2}\right) = \cos\left(\frac{a}{2}\right)\sin\left(\frac{B+C}{2}\right)$$

$$\sin\left(\frac{B}{2}\right)\sin\left(\frac{c+a}{2}\right) = \sin\left(\frac{b}{2}\right)\cos\left(\frac{C-A}{2}\right) \qquad \sin\left(\frac{B}{2}\right)\cos\left(\frac{c+a}{2}\right) = \cos\left(\frac{b}{2}\right)\cos\left(\frac{C+A}{2}\right)$$

$$\cos\left(\frac{B}{2}\right)\sin\left(\frac{c-a}{2}\right) = \sin\left(\frac{b}{2}\right)\sin\left(\frac{C-A}{2}\right) \qquad \cos\left(\frac{B}{2}\right)\cos\left(\frac{c-a}{2}\right) = \cos\left(\frac{b}{2}\right)\sin\left(\frac{C+A}{2}\right)$$

$$\sin\left(\frac{C}{2}\right)\sin\left(\frac{a+b}{2}\right)=\sin\left(\frac{c}{2}\right)\cos\left(\frac{A-B}{2}\right) \qquad \sin\left(\frac{C}{2}\right)\cos\left(\frac{a+b}{2}\right)=\cos\left(\frac{c}{2}\right)\cos\left(\frac{A+B}{2}\right)$$

$$\cos\left(\frac{C}{2}\right)\sin\left(\frac{a-b}{2}\right)=\sin\left(\frac{c}{2}\right)\sin\left(\frac{A-B}{2}\right) \qquad \cos\left(\frac{C}{2}\right)\cos\left(\frac{a-b}{2}\right)=\cos\left(\frac{c}{2}\right)\sin\left(\frac{A+B}{2}\right).$$

4. An MathCAD algorithm for solving a spherical triangle in Case SSA

Example 4.1. A MathCAD worksheet for solving a spherical triangle $A\,B\,C$ whose parts are $BC = a = 78^\circ 53'20''$, $AC = b = 28^\circ 14'30''$, and $A = 90^\circ$.

Solution. Case SSA: This is an ambiguous case and there may be two solutions

ORIGIN ≡ 1

The convert function for sides and angles in radians:

$$F(d, m, s) := \frac{\pi}{180} \cdot \left(d + \frac{m}{60} + \frac{s}{60^2}\right)$$

The convert function for sides and angles in DMS (degrees-minutes-seconds):

$$M(x) := \begin{pmatrix} d\left(\text{floor}\left(\frac{180}{\pi} \cdot x\right)\right) \\ \text{floor}\left(60 \cdot m\left(\frac{180}{\pi} \cdot x\right)\right) \\ \text{round}\left(60 \cdot m\left(60 \cdot \left(\frac{180}{\pi} \cdot x\right)\right)\right) \end{pmatrix}$$

where $d(x) \equiv \begin{cases} x & \text{if } x > 0 \\ 180 + x & \text{otherwise} \end{cases}$ $\qquad m(x) \equiv x - \text{floor}(x)$

$$\text{round}(x) \equiv \text{if}(x - \text{floor}(x) < 0.5, \text{floor}(x), \text{ceil}(x))$$

The magnitude of the side a in radians: $a := F(78, 53, 20)$ $\qquad a = 1.377$ rad

The magnitude of the side b in radians: $b := F(28, 14, 30)$ $\qquad b = 0.493$ rad

The measure of the angle A in radians: $A := F(90, 0, 0)$ $\qquad A = 1.570$ rad

For the ambiguous case SSA, we compute:

$$\sin(a) - \sin(b) \cdot \sin(A) = 0.508$$

$$a - b = 0.884 \qquad a - (\pi - b) = -1.272$$

Since $\sin(a) - \sin(b) \cdot \sin(A) > 0$, $A = 90^\circ$, $b < a < \pi - b$, there is only one value of the angle B and thus only one spherical triangle $A\,B\,C$.

Using Law of Sines, solve $\sin(B)$:

$$\sin B := \frac{\sin(b) \cdot \sin(A)}{\sin(a)} \qquad \sin B = 0.482$$

Measure of the angle B in radians: $B := \text{asin}(\sin B)$ $\qquad B = 0.503$ rad

Measure of the angle B in DMS: $M(B) = \begin{pmatrix} 28 \\ 49 \\ 52 \end{pmatrix}$

Using Napier's Analogies, solve the angle C:

$$C := 2 \cdot \operatorname{atan}\left(\frac{1}{\tan\left(\frac{A+B}{2}\right)} \cdot \frac{\cos\left(\frac{a-b}{2}\right)}{\cos\left(\frac{a+b}{2}\right)} \right)$$

Measure of the angle C in radians: $C = 1.465$ rad

Measure of the angle C in DMS: $M(C) = \begin{pmatrix} 83 \\ 56 \\ 41 \end{pmatrix}$

Using Napier's Analogies, solve the side c:

$$c := 2 \cdot \operatorname{atan}\left(\tan\left(\frac{a+b}{2}\right) \cdot \frac{\cos\left(\frac{A+B}{2}\right)}{\cos\left(\frac{A-B}{2}\right)} \right)$$

Measure of the side c in radians: $c = 1.350$ rad

Measure of the side c in DMS: $M(c) = \begin{pmatrix} 77 \\ 21 \\ 51 \end{pmatrix}$

Check: Law of Sines or Delambre-Gauss six-part rules may be used as check formulas:

$$\frac{\sin(a)}{\sin(A)} - \frac{\sin(b)}{\sin(B)} = 0 \qquad \frac{\sin(a)}{\sin(A)} - \frac{\sin(c)}{\sin(C)} = 0 \qquad \text{or}$$

$$\cos\left(\frac{A}{2}\right) \cdot \cos\left(\frac{b-c}{2}\right) - \cos\left(\frac{a}{2}\right) \cdot \sin\left(\frac{B+C}{2}\right) = 0 .$$

References

[1] Adams, A., Rogers, D.: *Elements for Computer Graphics*, McGrow-Hill, 1990.

[2] Adams, A., Billow, L.: *Descriptive Geometry and Geometric Modelling*, Holt, Rinehart and Winston, 1988.

[3] Kells, L.M., Kern, W.F., Bland, J.R.: *Plane And Spherical Trigonometry*, McGraw-Hill Book Co., New-York and London, 1940.

[4] Murray, D.A.: *Spherical trigonometry for colleges and secondary schools,* Longmans, Green and Co., New-York, 1908.

[5] Zamfir, M.: *The geometry of the sphere with MathCad*, Proceedings of 12-th Workshop of Scientific Communications, Department of Mathematics and Computer Science, Technical University of Civil Engineering, Bucharest, Romania, Matrix Rom, pp. 119-122, 2013.

[6] Zamfir, M.: *Trigonometric formulas for solving Euler spherical triangles with MathCad*, Proceedings of 14-th Workshop of Scientific Communications, Department of Mathematics and Computer Science, Technical University of Civil Engineering, Bucharest, Romania, Matrix Rom, pp. 131-136, 2017.