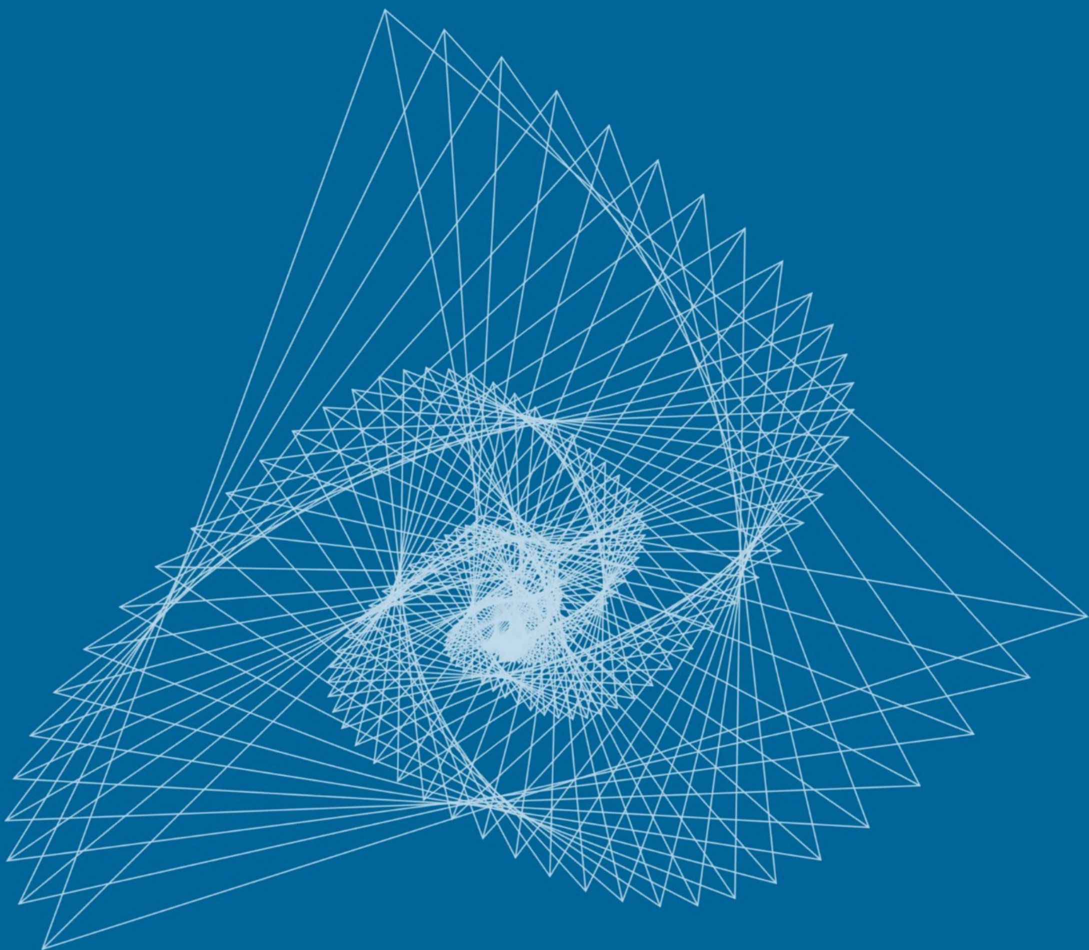
VOL. 10, NO. 2

URAL MATHEMATICAL JOURNAL

N.N. Krasovskii Institute of Mathematics and Mechanics of
the Ural Branch of Russian Academy of Sciences and
Ural Federal University named after the first President of Russia B.N.Yeltsin

ISSN: 2414-3952



2024

ISSN 2414-3952
(Online)

Ural Mathematical Journal

Vol. 10, No. 2, 2024

https://umjuran.ru

Electronic Periodical Scientific Journal
Founded in 2015

The Journal is registered by the Federal Service for Supervision in the Sphere of Communication, Information Technologies and Mass Communications
Certificate of Registration of the Mass Media Эл № ФС77-61719 of 07.05.2015

Founders

N.N. Krasovskii Institute of Mathematics and Mechanics of the Ural Branch of Russian Academy of Sciences

Ural Federal University named after the first President of Russia B.N. Yeltsin

Contact Information

16 S. Kovalevskaya str., Ekaterinburg, Russia, 620990
Phone: +7 (343) 375-34-73 Fax: +7 (343) 374-25-81
Email: secretary@umjuran.ru
Web-site: https://umjuran.ru

TABLE OF CONTENTS

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 4–14

DOI: 10.15826/umj.2024.2.001

SPECTRAL EXPANSION FOR SINGULAR BETA STURM–LIOUVILLE PROBLEMS

Bilender P. Allahverdiev

Department of Mathematics, Khazar University,
41 Mahsati Str., AZ1096 Baku, Azerbaijan;

UNEC-Azerbaijan State University of Economics,
6 Istiglaliyyat Str., Baku, Azerbaijan
bilenderpasaoglu@gmail.com

Hüseyin Tuna

Department of Mathematics, Mehmet Akif Ersoy University,
15030 Burdur, Turkey;

UNEC-Azerbaijan State University of Economics,
6 Istiglaliyyat Str., Baku, Azerbaijan
hustuna@gmail.com

Yüksel Yalçınkaya

Department of Mathematics, Süleyman Demirel University,
32260 Isparta, Turkey
matyuksel@hotmail.com

Abstract: In this study, beta Sturm–Liouville problems are discussed. For such equations, the spectral function is established in the singular case. A spectral expansion is given with the help of this function.

Keywords: Sturm–Liouville theory, Fractional derivatives and integrals, Spectral expansion.

1. Introduction

Fractional derivatives are mathematical operations that describe derivatives with non-integer degrees, extending the traditional concept of derivatives with integer degrees. These derivatives are part of a branch of mathematics often referred to as "fractional analysis" or "fractional calculus". The application areas of fractional derivatives are quite wide. For example, mathematical models expressed with fractional derivatives are used in fields such as electromagnetism, diffusion processes, and semiconductor physics. In addition, the concepts of fractional derivatives can be applied during the analysis of some fractal structures or complex systems.

In 2014, Khalil et al. defined conformable fractional derivatives and integrals by using classical derivative methods [9]. Later, Atangana et al. defined the beta fractional derivative and created a model of the famous river blindness disease based on Caputo and beta derivatives [3]. Martinez et al. have created analytical solutions of the space-time generalized nonlinear Schrödinger equation,

[0]2010 Mathematical Subject Classificition: 34B24, 26A33, 34L05.

including the beta derivative, using the sub-equation method [13]. Beta derivative has a particular applicability, particularly within the fields of biology and medicine [4, 5].

The spectral expansion of differential equations is a method of converting a given mathematical expression into a simpler and more easily solvable form. This method is particularly useful for solving difficult or complex differential equations. Expansion makes it possible to obtain analytical solutions or to obtain more effective solutions by numerical methods. The expansion of differential equations is important for numerical solutions as well as for obtaining analytical solutions. To solve a differential equation with numerical methods, it may often be possible to take an expanded equation in a simpler way and then solve this simplified equation numerically. The expansion of differential equations has many applications in mathematics, engineering, physics, and other branches of science. It is an indispensable tool, especially for obtaining analytical or numerical solutions to complex and real-world problems. In [1], the authors proved the existence of the spectral function for the singular conformable Sturm–Liouville problem.

The congruent fractional Sturm–Liouville problem is an extended version of the Sturm–Liouville theory and deals with differential equations involving fractional derivatives. While traditional Sturm–Liouville theory determines eigenvalues and eigenfunctions by examining quadratic linear differential equations, the fractional Sturm–Liouville problem includes fractional derivatives in equations. Fractional Sturm–Liouville problems often require eigenvalues and eigenfunctions to be obtained by analytical or semi-analytical expressions. Solving such equations may require spectral analysis methods that are often used for eigenvalue problems. Some researchers examine the solution of differential equations with fractional derivatives by dealing with eigenvalue problems such as fractional Sturm–Liouville problems and standard Sturm–Liouville problems [2, 6–8, 11, 14].

The computation and properties of fractional derivatives are generally more complex compared to integer-order derivatives. They expand the properties of traditional derivatives, and certain rules such as the fractional chain rule apply. Fractional derivatives can yield meaningful and valuable results for specific classes of functions. When calculating fractional derivatives, a method closely related to the integral operation is used. Special fractional derivative operators are used to calculate the fractional derivative of the function. These operators have some special properties and rules.

In this paper, singular beta Sturm–Liouville equations defined as

$$-T_\beta^2 y + v(\zeta)y = \mu y, \quad \zeta \in (0,\infty), \tag{1.1}$$

where μ is a complex eigenvalue parameter, $v(.)$ is a real-valued function defined on $[0,\infty)$, and $v \in L^1_{\beta,loc}(0,\infty)$, were considered. Using Levitan's method [11], the spectral function was established for such equations. A spectral expansion theorem was proved with the help of this function.

2. Preliminaries

Definition 1 [3, 13]. *Let $0 < \beta \le 1$ and $\sigma : [0,\infty) \to \mathbb{R} := (-\infty,\infty)$ be a function. The beta derivative of σ is defined by*

$$T_\beta\sigma(\zeta) = \frac{d^\beta\sigma(\zeta)}{dt^\beta} := \lim_{\varepsilon\to 0}\frac{\sigma(\zeta+\varepsilon(\zeta+1/\Gamma(\beta))^{1-\beta})-\sigma(\zeta)}{\varepsilon}.$$

As is known, fractional derivatives do not have the basic properties of the classical derivative (such as the derivative of the product, the derivative of the division). However, the beta derivative has the basic properties of the ordinary derivative and is therefore an extension of the conformable derivative.

Theorem 1 [13]. *Let σ, ω be beta differentiable functions for $\zeta > 0$ and $(0 < \beta \le 1)$. The following relations hold*:

(i)

$$T_\beta(\lambda\sigma+\delta\omega)=\lambda T_\beta\sigma+\delta T_\beta\omega, \quad \text{for all } \mu,\delta\in\mathbb{R},$$

(ii)

$$T_\beta(\sigma\omega)=\sigma T_\beta(\omega)+\omega T_\beta(\sigma),$$

(iii)

$$T_\beta(\frac{\sigma}{\omega})=\frac{\omega T_\beta(\sigma)-\sigma T_\beta(\omega)}{\omega^2},$$

(iv)

$$T_\beta(\sigma(\zeta))=\left(\zeta+\frac{1}{\Gamma(\beta)}\right)^{1-\beta}\frac{df(\zeta)}{dt},$$

(v)

$$T_\beta(\zeta^n)=(\zeta+\frac{1}{\Gamma(\beta)})^{1-\beta}n\zeta^{n-1}, \quad \mathbb{N}:=\{1,2,3,...\}.$$

P r o o f. The proof is clear, so we omit it. □

Definition 2. *Let $\sigma:[a,\infty)\to\mathbb{R}$, be a given function, then the beta-integral of σ is:*

$${}_aI^\beta(\sigma(\zeta))=\int_a^\zeta\left(t+\frac{1}{\Gamma(\beta)}\right)^{\beta-1}\sigma(t)dt,$$

where $0<\beta\leq 1$ and

$$({}^bT_\beta\sigma)(\zeta)=\lim_{\zeta\to b^-}({}^bT_\beta\sigma)(\zeta).$$

Theorem 2. *Let σ,ω be beta-differentiable functions. Then, the following relation holds*

$$\int_a^b\sigma(\zeta)T_\beta(\omega)(\zeta)d_\beta\zeta=\sigma(\zeta)\omega(\zeta)\big|_a^b-\int_a^b\omega(\zeta)T_\beta(\sigma)(\zeta)d_\beta\zeta.$$

P r o o f. By Theorem 1 the proof is clear. □

Let

$$L^2_\beta(0,b):=\left\{\sigma:\Big(\int_0^b|\sigma(\zeta)|^2d_\beta(\zeta)\Big)^{1/2}<\infty\right\}.$$

Then $L^2_\beta(0,b)$ is a Hilbert space endowed with the inner product

$$\langle\sigma,\omega\rangle:=\int_0^b\sigma\left(\zeta\right)\overline{\omega\left(\zeta\right)}d_\beta\zeta, \quad \sigma,\omega\in L^2_\beta(0,b).$$

The β-*Wronskian* of σ and ω is defined by

$$W_\beta(\sigma,\omega)(\zeta)=p(\zeta)\big[\sigma(\zeta)T_\beta\omega(\zeta)-\sigma(\zeta)T_\beta\omega(\zeta)\big], \quad \zeta\in[0,b].$$

Theorem 3. *Let A be an operator defined as $A\{t_i\}=\{y_i\}$, where*

$$y_i=\sum_{k=1}^\infty a_{ik}t_k \quad and \quad i\in\mathbb{N}.$$

If

$$\sum_{i,k=1}^\infty|a_{ik}|^2<+\infty \tag{2.1}$$

then A is a compact operator in l^2 [12].

3. Regular beta Sturm–Liouville problem

Consider the following regular problem

$$-T_{\beta}^{2}y(\zeta)+v(\zeta)y(\zeta)=\mu y(\zeta),\quad 0<\zeta<b<\infty, \tag{3.1}$$
$$y(0,\mu)\cos\theta+T_{\beta}y(0,\mu)\sin\theta=0, \tag{3.2}$$
$$y(b,\mu)\cos\gamma+T_{\beta}y(b,\mu)\sin\gamma=0,\quad \gamma,\theta\in\mathbb{R}, \tag{3.3}$$

where $v(\cdot)$ is a real-valued function defined on $[0,\infty)$, μ is a complex eigenvalue parameter, and $v\in L^{1}_{\beta,loc}(0,\infty)$, where

$$L^{1}_{\beta,loc}(0,\infty):=\left\{\sigma:[0,\infty)\to\mathbb{C}:\int_{0}^{b}|\sigma(\zeta)e|\,d_{\beta}(\zeta)<\infty,\ \forall b\in(0,\infty)\right\}.$$

We denote by $\phi(\zeta,\mu)$ and $\psi(\zeta,\mu)$ two solutions of (3.1) satisfying

$$\phi(0,\mu)=\sin\theta,\quad T_{\beta}\phi(0,\mu)=-\cos\theta, \tag{3.4}$$
$$\psi(b,\mu)=\sin\gamma,\quad T_{\beta}\psi(b,\mu)=-\cos\gamma. \tag{3.5}$$

Then *Green's function of* (3.1)–(3.3) is defined as

$$G(\zeta,t,\mu)=\frac{1}{W(\phi,\psi)}\begin{cases}\psi(\zeta,\mu)\phi(t,\mu), & 0\le t<\zeta,\\ \phi(\zeta,\mu)\psi(t,\mu), & \zeta<t<b.\end{cases} \tag{3.6}$$

Without loss of generality we can assume that $\mu=0$ is not an eigenvalue of (3.1)–(3.3). From (3.6), we find

$$G(\zeta,t)=G(\zeta,t,0)=\frac{1}{W(\phi,\psi)}\begin{cases}\psi(\zeta)\phi(t), & 0\le t<\zeta,\\ \phi(\zeta)\psi(t), & \zeta<t<b.\end{cases}$$

Theorem 4. *$G(\zeta,t)$ is a beta Hilbert–Schmidt kernel, i.e.,*

$$\int_{0}^{b}\int_{0}^{b}|G(\zeta,t)|^{2}d_{\beta}(\zeta)d_{\beta}(t)<+\infty.$$

P r o o f. From (3.6), we infer that

$$\int_{0}^{b}d_{\beta}(\zeta)\int_{0}^{\zeta}|G(\zeta,t)|^{2}d_{\beta}(t)<+\infty,$$

and

$$\int_{0}^{b}d_{\beta}(\zeta)\int_{\zeta}^{b}|G(\zeta,t)|^{2}d_{\beta}(t)<+\infty$$

since $\phi,\psi\in L^{2}_{\beta}(0,b)$. Hence we obtain

$$\int_{0}^{b}\int_{0}^{b}|G(\zeta,t)|^{2}d_{\beta}(\zeta)d_{\beta}(t)<+\infty. \tag{3.7}$$

□

Theorem 5. *The operator $\digamma$ defined as*

$$(\digamma\sigma)(\zeta)=\int_0^b G(\zeta,t)\sigma(t)d_\beta(t)$$

is compact and self-adjoint on $L^2_\beta(0,b)$.

P r o o f. Let $\varphi_i=\varphi_i(t)$ $(i\in\mathbb{N})$ be an orthonormal basis of $L^2_\beta(0,b)$. Define

$$t_i=(\sigma,\varphi_i)=\int_0^b \sigma(t)\overline{\varphi_i(t)}d_\beta(t),$$
$$y_i=(\omega,\varphi_i)=\int_0^b \omega(t)\overline{\varphi_i(t)}d_\beta(t),$$
$$a_{ik}=\int_0^b\int_0^b G(\zeta,t)\varphi_i(\zeta)\overline{\varphi_k(t)}d_\beta(\zeta)d_\beta(t)\quad (i,k\in\mathbb{N}).$$

Then, $L^2_\beta(0,b)$ is mapped isometrically onto l^2. By this mapping, $\digamma$ transforms into A on l^2, (3.7) is translated into (2.1). By Theorems 3 and 4, the operator A is compact. Therefore the operator $\digamma$ is compact.

Let $\sigma,\omega\in L^2_\beta(0,b)$. Then we see that

$$(\digamma\sigma,\omega)=\int_0^b(\digamma\sigma)(\zeta)\overline{\omega(\zeta)}d_\beta(\zeta)=\int_0^b\int_0^b G(\zeta,t)\sigma(t)\overline{\omega(\zeta)}d_\beta(\zeta)d_\beta(t)$$
$$=\int_0^b\sigma(\zeta)\overline{\left(\int_0^b G(t,\zeta)\omega(t)d_\beta(t)\right)}d_\beta(\zeta)=\int_0^b\sigma(\zeta)\overline{\left(\int_0^b G(\zeta,t)\omega(t)d_\beta(t)\right)}d_\beta(\zeta)=(\sigma,\digamma\omega),$$

due to $G(\zeta,t)=G(t,\zeta)$. □

4. Eigenfunction expansion

Let $\mu_{m,b}$ $(m\in\mathbb{N})$ denote the eigenvalues of (3.1)–(3.3) and $\phi_{m,b}(\zeta)=\phi(\zeta,\mu_{m,b})$ are the corresponding eigenfunctions. By virtue of Theorem 5 and the Hilbert–Schmidt theorem [10], we infer that

$$\int_0^b|\sigma(\zeta)|^2\,d_\beta(\zeta)=\sum_{m=1}^{\infty}\frac{1}{\gamma^2_{m,b}}\int_0^b|\sigma(\zeta)\phi_{m,b}(\zeta)|^2d_\beta(\zeta),$$

where $\sigma(.)\in L^2_\beta(0,b)$ and

$$\gamma^2_{m,b}=\int_0^b\phi^2_{m,b}(\zeta)d_\beta(\zeta).$$

Set

$$\rho_b(\mu)=\begin{cases}-\sum\limits_{\mu<\mu_{m,b}<0}\dfrac{1}{\gamma^2_{m,b}}, & \text{for } \mu\le 0,\\ \sum\limits_{\mu<\mu_{m,b}<0}\dfrac{1}{\gamma^2_{m,b}}, & \text{for } \mu\ge 0.\end{cases}$$

Then we obtain

$$\int_0^b|\sigma(\zeta)|^2\,d_\beta(\zeta)=\int_{-\infty}^{\infty}|\Upsilon(\mu)|^2\,d_{\rho_b}(\mu),\tag{4.1}$$

which is called the *Parseval equality*, where

$$\Upsilon(\mu)=\int_0^b\sigma(\zeta)\phi(\zeta,\mu)d_\beta(\zeta).$$

Lemma 1. *For any $\tau > 0$, there exists a positive constant $P = P(s)$ not depending on b such that*

$$\bigvee_{-R}^{R}\{\rho_b(\mu)\} = \sum_{-R\le\mu_{m,b}<R}\frac{1}{\gamma_{m,b}^2} = \rho_b(R) - \rho_b(-R) < P, \tag{4.2}$$

where $\bigvee$ denotes the total variation.

P r o o f. Let $\sin\theta \neq 0$. By (3.4), there exists a positive number k nearby 0 such that

$$\Big(\frac{1}{k}\int_0^k \phi(\zeta,\mu)d_\beta\zeta\Big)^2 > \frac{1}{2}\sin^2\theta \tag{4.3}$$

due to $\phi(\zeta,\mu)$ is continuous at 0. Let us define $\sigma_k(t)$ by

$$\sigma_k(\zeta) = \begin{cases} \dfrac{1}{k}, & 0\le\zeta\le k,\\ 0, & \zeta > k. \end{cases}$$

Combining (4.1), (4.2) and (4.3), we conclude that

$$\begin{aligned}\int_0^k \sigma_k^2(\zeta)d_\beta\zeta &= \frac{1}{k^2\beta}(k+\frac{1}{\Gamma(\beta)})^\beta = \int_{-\infty}^{\infty}\Big(\frac{1}{k}\int_0^k\phi(\zeta,\mu)d_\alpha\zeta\Big)^2 d_{\rho_b}(\mu)\\ &\ge \int_{-R}^{R}\Big(\frac{1}{k}\int_0^k\phi(\zeta,\mu)\,d_\beta\zeta\Big)^2 d_{\rho_b}(\mu) > \frac{1}{2}\sin^2\theta\left\{\rho_b(R)-\rho_b(-R)\right\}.\end{aligned}$$

If $\sin\theta = 0$, then $\sigma_k(\zeta)$ is defined by

$$\sigma_k(t) = \begin{cases} \Big(\dfrac{1}{k}\Big)^2, & 0\le\zeta\le k,\\ 0, & \zeta > k. \end{cases}$$

The proof of the lemma follows from Parseval's equality. ☐

Let ρ be any nondecreasing function on $-\infty < \mu < \infty$. We will denote by $L^2_\rho(\mathbb{R})$ the Hilbert space of all functions $\sigma : \mathbb{R}\to\mathbb{R}$ measurable with respect to the Lebesgue–Stieltjes measure defined by ρ, with the condition

$$\int_{-\infty}^{\infty}\sigma^2(\mu)\,d_\rho(\mu) < \infty$$

and with the inner product

$$(\sigma,\omega)_\rho := \int_{-\infty}^{\infty}\sigma(\mu)\,\omega(\mu)\,d_\rho(\mu).$$

Theorem 6. *For Problem (3.1)–(3.2), there exists a nondecreasing function $\rho(\mu)$ $(-\infty < \mu < \infty)$ with the following properties.*

(i) *If σ is a real-valued function and $\sigma \in L^2_\beta(0,\infty)$, then there exists a function $\Upsilon \in L^2_\rho(\mathbb{R})$ satisfying*

$$\lim_{b\to\infty}\int_{-\infty}^{\infty}\Big\{\Upsilon(\mu) - \int_0^b\sigma(\zeta)\phi(\zeta,\mu)d_\beta(\zeta)\Big\}d_\rho(\mu) = 0, \tag{4.4}$$

and the Parseval equality

$$\int_0^\infty \sigma^2(\zeta)d_\beta(\zeta) = \int_{-\infty}^{\infty}\Upsilon^2(\mu)d_\rho(\mu) \tag{4.5}$$

holds.

(ii) *The integral*

$$\int_{-\infty}^{\infty} \Upsilon(\mu)\phi(\zeta,\mu)d_{\rho}(\mu),$$

converges to σ *in* $L^2_\beta(0,\infty)$. *That is,*

$$\lim_{n\to\infty}\int_0^\infty \left\{\sigma(\zeta)-\int_{-n}^{n}\Upsilon(\mu)\phi(\zeta,\mu)d_{\rho}(\mu)\right\}^2 d_\beta(\zeta)=0.$$

P r o o f. (i) Suppose that:

1) The real-valued function $\sigma_\xi(.)$ vanishes outside the interval $[0,\xi]$, where $\xi<b$.

2) $\sigma_\xi(\zeta)$ and $T_\beta\sigma_\xi(\zeta)$ are continuous.

3) $\sigma_\xi(\zeta)$ satisfies (3.2).

By (4.1), we deduce that

$$\int_0^\xi \sigma^2_\xi(\zeta)d_\rho(\zeta)=\int_{-\infty}^{\infty}\Upsilon^2_\xi(\mu)d_\rho(\mu), \tag{4.6}$$

where

$$\Upsilon_\xi(\mu)=\int_0^\xi \sigma_\xi(\zeta)\phi(\zeta,\mu)d_\beta(\zeta). \tag{4.7}$$

Since $\phi(t,\mu)$ satisfies the equation (3.1), we see that

$$\phi(\zeta,\mu)=\frac{1}{\mu}\left[-T^2_\beta\phi(\zeta,\mu)+v(\zeta)\phi(\zeta,\mu)\right].$$

By (4.7), we get

$$\Upsilon_\xi(\mu)=\frac{1}{\mu}\int_0^\xi \sigma_\xi(\zeta)\left[-T^2_\beta\phi(\zeta,\mu)+v(\zeta)\phi(\zeta,\mu)\right]d_\beta(\zeta).$$

Since $\sigma_\xi(\zeta)$ and $\phi(\zeta,\mu)$ satisfy the boundary condition (3.4) and $\sigma_\xi(\zeta)$ vanishes in a neighborhood of the point ξ, we get

$$\Upsilon_\xi(\mu)=\frac{1}{\mu}\int_0^b \phi(\zeta,\mu)\left[-T^2_\beta\sigma_\xi(\zeta)+v(\zeta)\sigma_\xi(\zeta)\right]d_\beta(\zeta),$$

via the integration by parts.

For any finite $R>0$, by using (4.1), we get

$$\begin{aligned}\int_{|\mu|>R}\Upsilon^2_\xi(\mu)d_{\rho_b}(\mu) &\le \frac{1}{R^2}\int_{|\mu|>R}\left\{\int_0^b\left[\phi(\zeta,\mu)\left[-T^2_\beta\sigma_\xi(\zeta)+v(\zeta)\sigma_\xi(\zeta)\right]\right]d_\beta(\zeta)\right\}^2 d_{\rho_b}(\mu)\\ &\le \frac{1}{R^2}\int_{-\infty}^{\infty}\left\{\int_0^b\left[\phi(\zeta,\mu)\left[-T^2_\beta\sigma_\xi(\zeta)+v(\zeta)\sigma_\xi(\zeta)\right]\right]d_\beta(\zeta)\right\}^2 d_{\rho_b}(\mu)\\ &= \frac{1}{R^2}\int_0^\xi\left[-T^2_\beta\sigma_\xi(\zeta)+v(\zeta)\sigma_\xi(\zeta)\right]^2 d_\beta(\zeta).\end{aligned}$$

From (4.6), we see that

$$\left|\int_0^\xi \sigma^2_\xi(\zeta)d_\beta(\zeta)-\int_{-R}^{R}\Upsilon^2_\xi(\mu)d_{\rho_b}(\mu)\right|\le\frac{1}{R^2}\int_0^\xi\left[-T^2_\beta\sigma_\xi(\zeta)+v(\zeta)\sigma_\xi(\zeta)\right]^2 d_\beta(\zeta). \tag{4.8}$$

By Lemma 1, we see that $\{\rho_b(\mu)\}$ is bounded. By Helly's theorems [10], we can find a sequence $\{b_{n_k}\}$ such that the sequence $\rho_{b_{n_k}}(\mu)$ converges $(b_{n_k} \to \infty)$ to a monotone function $\rho(\mu)$. Passing to the limit as $b_{n_k} \to \infty$ in (4.8), we get

$$\left| \int_0^\xi \sigma_\xi^2(\zeta) d_\beta(\zeta) - \int_{-R}^{R} \Upsilon_\xi^2(\mu) d_\beta(\mu) \right| \leq \frac{1}{R^2} \int_0^\xi \left[-T_\beta^2 \sigma_\xi(\zeta) + \phi(\zeta)\sigma_\xi(\zeta) \right]^2 d_\beta(\zeta).$$

Hence, letting $R \to \infty$, we obtain

$$\int_0^\xi \sigma_\xi^2(\zeta) d_\beta(\zeta) = \int_{-\infty}^{\infty} \Upsilon_\xi^2(\mu) d_\rho(\mu).$$

Assume that σ is an arbitrary real-valued function on $L_\beta^2(a,\infty)$. Then there exists a sequence $\{\sigma_s(\zeta)\}$ satisfying the conditions 1)–3) and such that

$$\lim_{s\to\infty} \int_0^\infty (\sigma(\zeta) - \sigma_s(\zeta))^2 \, d_\beta(\zeta) = 0. \tag{4.9}$$

Let

$$\Upsilon_\tau(\mu) = \int_0^\infty \sigma_\tau(\zeta)\phi(\zeta,\mu) d_\beta(\zeta).$$

Then, we have

$$\int_0^\infty \sigma_\tau^2(\zeta) d_\beta(\zeta) = \int_{-\infty}^{\infty} \Upsilon_\tau^2(\mu) d_\rho(\mu).$$

By (4.9), we see that $\sigma_s(\zeta)$ is a Cauchy sequence, i.e.,

$$\int_0^\infty (\sigma_{\tau_1}(\zeta) - \sigma_{\tau_2}(\zeta))^2 \, d_\beta(\zeta) \to 0$$

as $\tau_1, \tau_2 \to \infty$. Thus we have

$$\int_{-\infty}^{\infty} (\Upsilon_{\tau_1}(\mu) - \Upsilon_{\tau_2}(\mu))^2 \, d_\rho(\mu) = \int_0^\infty (\sigma_{\tau_1}(t) - \sigma_{\tau_2}(t))^2 \, d_\beta(\zeta) \to 0$$

as $\tau_1, \tau_2 \to \infty$. Therefore, there exists a limit function Υ satisfying

$$\int_0^\infty \sigma^2(\zeta) d_\beta(\zeta) = \int_{-\infty}^{\infty} \Upsilon^2(\mu) d_\rho(\mu),$$

by the completeness of the space $L_\rho^2(\mathbb{R})$.

Now, we show that K_τ defined as

$$K_\tau(\mu) = \int_0^\tau \sigma(\zeta)\phi(\zeta,\mu) d_\beta(\zeta)$$

converges to Υ as $\tau \to \infty$. Assume that ω is another function in $L_\beta^2(0,\infty)$. Similarly, $\Omega(\mu)$ can be defined by ω. Then we have

$$\int_0^\infty (\sigma(\zeta) - \omega(\zeta))^2 \, d_\beta(\zeta) = \int_{-\infty}^{\infty} \{\Upsilon(\mu) - \Omega(\mu)\}^2 \, d_\rho(\mu).$$

Now set

$$\omega(\zeta) = \begin{cases} \sigma(\zeta), & \zeta \in [0,\tau], \\ 0, & \zeta \in (\tau,\infty). \end{cases}$$

Then we have
$$\int_{-\infty}^{\infty}\{\Upsilon(\mu)-K_{\tau}(\mu)\}^{2}d_{\rho}(\mu)=\int_{\tau}^{\infty}\sigma^{2}(\zeta)d_{\beta}(\zeta)\to 0\ (\tau\to\infty).$$

(ii) Suppose that $\sigma,\omega\in L_{\beta}^{2}(0,\infty)$ and $\Upsilon(\mu),\Omega(\mu)$ are their Fourier transforms, respectively. Then $\Upsilon\mp\Omega$ are the transforms of $\sigma\mp\omega$. From (4.5), we obtain

$$\int_{0}^{\infty}[\sigma(\zeta)+\omega(\zeta)]^{2}d_{\beta}(\zeta)=\int_{-\infty}^{\infty}(\Upsilon(\mu)+\Omega(\mu))^{2}d_{\rho}(\mu), \tag{4.10}$$
$$\int_{0}^{\infty}[\sigma(\zeta)-\omega(\zeta)]^{2}d_{\beta}(\zeta)=\int_{-\infty}^{\infty}(\Upsilon(\mu)-\Omega(\mu))^{2}d_{\rho}(\mu). \tag{4.11}$$

Combining (4.10) and (4.11), we conclude that

$$\int_{0}^{\infty}\sigma(\zeta)\omega(\zeta)d_{\beta}(\zeta)=\int_{-\infty}^{\infty}\Upsilon(\mu)\Omega(\mu)d_{\varrho}(\mu). \tag{4.12}$$

Define
$$\sigma_{\varsigma}(\zeta)=\int_{-\varsigma}^{\varsigma}\Upsilon(\mu)\phi(\zeta,\mu)d_{\rho}(\mu),$$

where Υ is defined in (4.4) and ς is a positive number. Let $\omega(.)$ be a function which is equal to zero outside the finite interval $[0,\tau]$. Hence

$$\begin{aligned}\int_{0}^{\tau}\sigma_{\varsigma}(\zeta)\omega(\zeta)d_{\beta}(\zeta)&=\int_{0}^{\tau}\Big\{\int_{-\varsigma}^{\varsigma}\Upsilon(\mu)\phi(\zeta,\mu)d_{\rho}(\mu)\Big\}\omega(\zeta)d_{\beta}(\zeta)\\&=\int_{-\varsigma}^{\varsigma}\Upsilon(\mu)\Big\{\int_{0}^{\tau}\phi(\zeta,\mu)\omega(\zeta)d_{\beta}(\zeta)\Big\}d_{\rho}(\mu)=\int_{-\varsigma}^{\varsigma}\Upsilon(\mu)\Omega(\mu)d_{\rho}(\mu).\end{aligned} \tag{4.13}$$

From (4.12), we get
$$\int_{0}^{\infty}\sigma_{\varsigma}(\zeta)\omega(\zeta)d_{\beta}(\zeta)=\int_{-\infty}^{\infty}\Upsilon(\mu)\Omega(\mu)d_{\rho}(\mu). \tag{4.14}$$

By (4.13) and (4.14), we have

$$\int_{0}^{\infty}(\sigma(\zeta)-\sigma_{\varsigma}(\zeta))\,\omega(\zeta)d_{\beta}(\zeta)=\int_{|\mu|>\varsigma}\Upsilon(\mu)\Omega(\mu)d_{\rho}(\mu).$$

From the Cauchy–Schwarz inequality, we see that

$$\begin{aligned}\left|\int_{0}^{\infty}(\sigma(\zeta)-\sigma_{\varsigma}(\zeta))\,\omega(\zeta)d_{\beta}(\zeta)\right|^{2}&\leq\int_{|\mu|>\varsigma}\Upsilon^{2}(\mu)d_{\rho}(\mu)\int_{|\mu|>\varsigma}\Omega^{2}(\mu)d_{\rho}(\mu)\\&\leq\int_{|\mu|>\varsigma}\Upsilon^{2}(\mu)d_{\rho}(\mu)\int_{-\infty}^{\infty}\Omega^{2}(\mu)d_{\rho}(\mu).\end{aligned} \tag{4.15}$$

Let
$$\omega(\zeta)=\begin{cases}\sigma(\zeta)-\sigma_{\varsigma}(\zeta), & \zeta\in[0,\tau]\,,\\ 0, & \zeta\in(\tau,\infty).\end{cases}$$

From (4.15), we obtain

$$\int_{0}^{\infty}(\sigma(\zeta)-\sigma_{\varsigma}(\zeta))^{2}d_{\beta}(\zeta)\leq\int_{|\mu|>\varsigma}\Upsilon^{2}(\mu)d_{\rho}(\mu).$$

Letting $\varsigma\to\infty$ gives the desired result due to the right-hand side does not depend on τ. □

Example 1. If we take $\beta = 1$ in (1.1), then we obtain the ordinary Sturm–Liouville problem defined by

$$-y'' + v(\zeta)y = \mu y, \quad \zeta \in (0, \infty),$$

where μ is a complex eigenvalue parameter, $v(.)$ is a real-valued function defined on $[0, \infty)$, and $v \in L^1_{loc}(0, \infty)$. Then Theorem 6 gives the spectral expansion for this problem (see [11]).

Example 2. Consider the following problem

$$\begin{gathered} -T^2_\beta y(\zeta) - ky(\zeta) = \mu y(\zeta), \quad 0 < \zeta < \infty, \\ y(0) = 0, \end{gathered} \tag{4.16}$$

where k is a constant. It is clear that

$$\phi(\zeta, \mu) = \frac{\sin\left(\int_0^\zeta \sqrt{\mu + k} d_\beta \zeta\right)}{\sqrt{\mu + k}}$$

is the solution of (4.16). By Theorem 6, we obtain

$$\Upsilon(\mu) = \int_0^\infty \sigma(\zeta) \frac{\sin\left(\int_0^\zeta \sqrt{\mu + k} d_\beta \zeta\right)}{\sqrt{\mu + k}} d_\beta(\zeta),$$

and

$$\sigma(\zeta) = \int_{-\infty}^\infty \Upsilon(\mu) \frac{\sin\left(\int_0^\zeta \sqrt{\mu + k} d_\beta \zeta\right)}{\sqrt{\mu + k}} d_\rho(\mu).$$

5. Conclusion

The present study is devoted to the discussion of beta Sturm–Liouville problems. In the context of such equations, the spectral function was established in the singular case. A spectral expansion was derived with the aid of this function. The Titchmarh–Weyl theory for this type of equations may be the subject of future research.

Acknowledgements

The authors are highly grateful to the referee for his/her useful suggestions.

REFERENCES

1. Allahverdiev B. P., Tuna H., Yalçınkaya Y. Spectral expansion for singular conformable Sturm–Liouville problem. *Math. Commun.*, 2020. Vol. 25. P. 237–252. URL: https://hrcak.srce.hr/244270
2. Allahverdiev B. P., Tuna H., Yalçınkaya Y. Conformable fractional Sturm–Liouville equation. *Math. Meth. Appl. Sci.,* 2019. Vol. 42, No. 10. P. 3508–3526. DOI: 10.1002/mma.5595
3. Atangana A., Alqahtani R. T. Modelling the spread of river blindness disease via the Caputo fractional derivative and the beta-derivative. *Entropy,* 2016. Vol. 18, No. 2. Art. no. 40. DOI: 10.3390/e18020040
4. Atangana A., Goufo E. F. D. On the mathematical analysis of Ebola hemorrhagic fever: deathly infection disease in West African countries. *BioMed Res. Int.*, 2014. Art. no. 261383. DOI: 10.1155/2014/261383

5. Atangana A., Oukouomi Noutchie S. C. Model of break-bone fever via beta-derivatives. *BioMed Res. Int.*, 2014. Art. no. 523159. DOI: 10.1155/2014/523159
6. Aydemir K., Mukhtarov O. Sh. A new type Sturm–Liouville problem with an abstract linear operator contained in the equation. *Quaest. Math.,* 2022. Vol. 45, No. 12. P. 1931–1948. DOI: 10.2989/16073606.2021.1979681
7. Cetinkaya F. A. On a conformable fractional wave equation. *J. Natural App. Sci.,* 2018. Vol. 22, No. 3. P. 1110–1113. DOI: 10.19113/sdufenbed.430296
8. Gülşen T., Yılmaz E., Kemaloğlu H. Conformable fractional Sturm–Liouville equation and some existence results on time scales. *Turkish J. Math.,* 2018. Vol. 42, No. 3. Art. no. 48. P. 1348–1360. DOI: 10.3906/mat-1704-120
9. Khalil R., Horani M. Al., Yousef A., Sababheh M. A new definition of fractional derivative. *J. Computat. Appl. Math.,* 2014. Vol. 264. P. 65–70. DOI: 10.1016/j.cam.2014.01.002
10. Kolmogorov A. N., Fomin S. V. *Introductory Real Analysis.* New York: Dover Publ. Inc., 1970. 393 p.
11. Levitan B. M., Sargsjan I. S. *Sturm–Liouville and Dirac Operators*. Ser. Math. Appl. (Soviet Series). Dordrecht: Kluwer Academic Publishers Group, 1991. 350 p. DOI: 10.1007/978-94-011-3748-5
12. Naimark M. A. *Linear Differential Operators*, 2nd edn. New York: Ungar, 1968. 156 p.
13. Yépez-Martínez H., Gómez-Aguilar J. F., Baleanu D. Beta-derivative and sub-equation method applied to the optical solitons in medium with parabolic law nonlinearity and higher order dispersion. *Optik,* 2018. Vol. 155. P. 357–365. DOI: 10.1016/j.ijleo.2017.10.104
14. Wang W.-C. Some notes on conformable fractional Sturm–Liouville problems. *Boundary Value Probl.*, 2021. Vol. 2021. Art. no. 103. DOI: 10.1186/s13661-021-01581-y

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 15–24

DOI: 10.15826/umj.2024.2.002

REACHABLE SET OF SOME DISCRETE SYSTEM WITH UNCERTAIN LIU DISTURBANCES[1]

Boris I. Ananyev

Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences,
16 S. Kovalevskaya Str., Ekaterinburg, 620108, Russian Federation

abi@imm.uran.ru

Abstract: The paper considers the problem of finding the reachable set for a linear system with determinate and stochastic Liu's uncertainties. As Liu's uncertainties, we use uniformly distributed ordinary uncertain values defined in some uncertain space and independent of one another. This fact means that the state vector of the system becomes infinite-dimensional. As determinate uncertainties, we consider feedback controls and unknown initial states. Besides, there is a constraint in the form of a sum of uncertain expectations. The initial estimation problem reduces to a determinate multi-step problem for matrices with a fixed constraint at the right end of the trajectory. This reduction requires some information on Liu's theory. We give necessary and sufficient conditions for the finiteness of a target functional in the obtained determinate problem. We provide a numerical example of a two-dimensional two-step system.

Keywords: Uncertainty theory, Uncertain values, Feedback controls, Attainable set, Lagrange multipliers.

1. Introduction

Baoding Liu's uncertainty theory has been widely developed in the last decade [9, 12, 13]. Elements of the theory are used in control theory, mathematical programming, financial mathematics, robotics, and other areas of applied mathematics. Liu notes in his book that "uncertainty theory has become a branch of axiomatic mathematics for modeling belief degrees."

It should be noted that Liu's theory is only one of the possible approaches to describing and accounting for uncertainty. Such approaches include various versions of probability theory, Zadeh's fuzzy set theory, interval analysis, and chaos theory [1, 6, 8, 11]. Possibility theory is actively developed as an alternative to probability in [10]. The theory of guaranteed estimation [7], based on a set-theoretic description of uncertainty, has also gained wide popularity. Of course, Liu's theory overlaps with the theories mentioned above.

This paper presents an extended version of the lecture given at the XIV All-Russian Conference on Control Problems [2]. We consider the estimation problem for discrete time Liu's processes described by the linear equations

$$x_k = (A_k x_{k-1} + B_k v_k)(1 + \lambda_k \xi_k), \quad x_k \in \mathbb{R}^n, \quad k \in 1:m, \tag{1.1}$$

where $|\lambda_k| \le 1$ are real numbers; $v_k = K_k x_{k-1}$ are uncertain feedback controls; ξ_k are ordinary uncertain values uniformly distributed on $[-1, 1]$, independent one of another, and defined on the N-space $(\Omega, \mathcal{F}, N)$, where $\mathcal{F}$ is a σ-algebra, and N is the uncertainty measure (function) of the set. The following constraints are also given:

$$J(x_0, \mathbb{K}) = \sum_{k\in 1:m} \mathrm{E}\left(v_k' R_k v_k + x_{k-1}' Q_k x_{k-1}\right) \le 1, \quad x_0 \in \mathbf{X}_0, \tag{1.2}$$

[1]The work was performed as part of research conducted in the Ural Mathematical Center with the financial support of the Ministry of Science and Higher Education of the Russian Federation (Agreement number 075-02-2024-1377).

where E is the uncertain expectation, $v_k = K_k x_{k-1}$, $\mathbb{K} = K_{1:m}$, R_k and Q_k are symmetric matrices of appropriate dimension, and $\mathbf{X}_0$ is a convex compact set in $\mathbb{R}^n$. One can see that the state vector x_k depends on elementary event $\omega \in \Omega$ because of uncertain values $\xi_k(\omega)$. So, estimating the reachable set at the terminal time m becomes the problem of infinite dimension for system (1.1) under constraints (1.2). Note that an estimation problem for a determinate system with an uncertain matrix was studied in [5]. First of all, let us recall some facts on Liu's theory.

2. Necessary facts on Liu's theory

Given a measurable space $(\Omega, \mathcal{F})$, where Ω is an arbitrary set and $\mathcal{F}$ is a σ-algebra of subsets of Ω, the *uncertain measure* N is defined on $\mathcal{F}$ to satisfy the following axioms:

1. Normality: $N(\Omega) = 1$.
2. Duality: $N(A) + N(A^c) = 1$ for any event $A \in \mathcal{F}$, where $A^c = \Omega \setminus A$.
3. Subadditivity: for any sequence $A_i \in \mathcal{F}$,

$$N\Big(\bigcup_{i\in\mathbb{N}} A_i\Big) \leq \sum_{i\in\mathbb{N}} N(A_i).$$

Any uncertain measure satisfies the relations $0 \leq N(A) \leq 1$ and $N(A) \leq N(B)$ for all $A, B \in \mathcal{F}$ such that $A \subset B$. But it is not a measure in ordinary sense [4]. Of course, any probability measure P satisfies axioms 1–3 and is, therefore, an uncertainty measure. Every measurable function $\xi : \Omega \to \mathbb{R}$ is called an *uncertain variable*. The independence of the family ξ_t, $t \in T$, of uncertain variables is defined as follows:

$$N\Big(\bigcap_{t\in T} \xi_t^{-1}(B_t)\Big) = \bigwedge_{t\in T} N\left(\xi_t^{-1}(B_t)\right)$$

for any $B_t \in \mathcal{B}$, where $\mathcal{B}$ is the Borelian σ-algebra on $\mathbb{R}$ and $\bigwedge_{t\in T} a_t = \inf_{t\in T} a_t$. The independence of uncertain variables was generalized to an arbitrary set T of indexes t in [3].

If ξ is an uncertain variable, then its distribution function is defined by the formula $F_\xi(x) = N(\xi \leq x)$ for all $x \in \mathbb{R}$, which corresponds to the similar notion in probability theory. It is proved (see [3, 9]) that a nondecreasing function $F : \mathbb{R} \to [0, 1]$ is a distribution function for some uncertain variable if and only if the following properties hold:

(1) $F \not\equiv 1$;

(2) $F \not\equiv 0$;

(3) the condition $F(x) = 1$ for all $x > x^*$ must imply that $F(x^*) = 1$.

For any distribution function satisfying 1–3, one can build a so-called *ordinary* uncertain variable $\xi(x) = x$ on the measure space $(\mathbb{R}, \mathcal{B})$ in the following way. First, consider a family of sets $\mathcal{L}$ consisting of semi-infinite intervals $(-\infty, x]$, their complements (x, ∞), the empty set, and the entire space $\mathbb{R}$. One can define the uncertainty measure N on $\mathcal{L}$ as follows: $N((-\infty, x]) = F(x)$, $N((x, \infty)) = 1 - F(x)$, and $F(\varnothing) = 0$. After that, for $B \in \mathcal{B}$, the uncertainty measure N_ξ is defined on $\mathcal{B}$ by the formula

$$N_\xi(B) = \begin{cases} \inf\limits_{B \subset \bigcup\limits_{i\in\mathbb{N}} A_i} \sum\limits_{i\in\mathbb{N}} N(A_i) & \text{if} \quad \inf\limits_{B \subset \bigcup\limits_{i\in\mathbb{N}} A_i} \sum\limits_{i\in\mathbb{N}} N(A_i) < 0.5, \\ 1 - \inf\limits_{B^c \subset \bigcup\limits_{i\in\mathbb{N}} A_i} \sum\limits_{i\in\mathbb{N}} N(A_i) & \text{if} \quad \inf\limits_{B^c \subset \bigcup\limits_{i\in\mathbb{N}} A_i} \sum\limits_{i\in\mathbb{N}} N(A_i) < 0.5, \\ 0.5 & \text{in other cases.} \end{cases} \tag{2.3}$$

Here, the infimums are taken over all sequences $A_i \in \mathcal{L}$ that cover the corresponding sets. We see that $N_\xi = N$ on $\mathcal{L}$.

An ordinary uncertain variable ξ uniformly distributed on $[-1, 1]$ has its distribution function

$$F_\xi(x) = \begin{cases} 0, & \text{if} \quad x \le -1; \\ (x+1)/2, & \text{if} \quad x \in (-1,1); \\ 1, & \text{if} \quad x \ge 1. \end{cases}$$

Such uncertain variables belong to the class of regular uncertain variables, for which there is an interval (a, b) where $F_\xi(x)$ is continuous and strictly increasing. Besides, $\lim_{x\to a} F_\xi(x) = 0$ and $\lim_{x\to b} F_\xi(x) = 1$. Here, it is possible that $a = -\infty$ and $b = \infty$.

The mathematical expectation of an uncertain variable ξ is defined by the formula

$$\mathrm{E}\xi = \int_0^\infty N(\xi \ge x)dx - \int_{-\infty}^0 N(\xi \le x)dx$$

if at least one of the integrals is finite. Since the function $1 - F_\xi(x)$ differs from the function $N(\xi \ge x)$ only at countably many points, we have

$$\mathrm{E}\xi = \int_0^\infty (1 - F_\xi(x))dx - \int_{-\infty}^0 F_\xi(x)dx = \int_0^\infty x dF_\xi(x) + \int_{-\infty}^0 x dF_\xi(x)$$

using integration by parts. For variables with regular distribution functions, we have

$$\int_0^\infty x dF_\xi(x) = \int_{F_\xi(0)}^1 F_\xi^{-1}(\alpha)d\alpha, \quad \int_{-\infty}^0 x dF_\xi(x) = \int_0^{F_\xi(0)} F_\xi^{-1}(\alpha)d\alpha. \tag{2.4}$$

From (2.3), we obtain $\mathrm{E}\xi = 0$ for the ordinary uniformly distributed uncertain variable ξ. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called strictly increasing if $f(u_1, \ldots, u_n) \ge f(v_1, \ldots, v_n)$ for $u_i \ge v_i$ and $f(u_1, \ldots, u_n) > f(v_1, \ldots, v_n)$ for $u_i > v_i$.

The following theorem is often used in applications.

Theorem 1 [12, Theorem 2.6]. *Let $[u; v] = [u_1; \ldots; u_m; v_1; \ldots; v_n]$ be independent uncertain variables with regular distribution functions $[F_u; F_v] = [F_{u_1}; \ldots; F_{u_m}; F_{v_1}; \ldots; F_{v_n}]$, respectively. If the function $f(u, v)$ is strictly increasing in u and strictly decreasing in v, then the uncertain variable $\xi = f(u, v)$ has the inverse distribution function*

$$\begin{gathered} F_\xi^{-1}(a) = f\left(F_u^{-1}(a), F_v^{-1}(1-a)\right), \quad a \in (0,1), \\ F_u^{-1} = [F_{u_1}^{-1}; \ldots; F_{u_m}^{-1}], \quad F_v^{-1} = [F_{v_1}^{-1}; \ldots; F_{v_n}^{-1}]. \end{gathered} \tag{2.5}$$

Corollary 1. *For any uncertain variable ξ, $\mathrm{E}(a\xi) = a\mathrm{E}\xi$ for all $a \in \mathbb{R}$. For any regular and independent uncertain variables ξ and η with finite mathematical expectations, $\mathrm{E}(a\xi + b\eta) = a\mathrm{E}\xi + b\mathrm{E}\eta$ for all $a, b \in \mathbb{R}$.*

Indeed,

$$N(a\xi \le x) = N(\xi \le x/a) = F_\xi(x/a), \quad a > 0,$$

and

$$\mathrm{E}(a\xi) = \int_{-\infty}^\infty x dF_\xi(x/a) = a\mathrm{E}\xi.$$

If $a < 0$, then $N(a\xi \leq x) = N(\xi \geq x/a) = 1 - F_\xi(x/a)$ N-almost everywhere and

$$\mathrm{E}(a\xi) = \int_{-\infty}^{\infty} x d(1 - F_\xi(x/a)) = a\mathrm{E}\xi.$$

Moreover,

$$F^{-1}_{\xi+\eta} = F^{-1}_\xi + F^{-1}_\eta, \quad F^{-1}_{-\xi}(a) = F^{-1}_\xi(1-a)$$

if ξ and η are regular. Unfortunately, the linear property of mathematical expectation is not valid for arbitrary uncertain variables.

It must be kept in mind that uncertain variables ξ and η with identical distribution functions $F_\xi \equiv F_\eta \equiv F$ may have different distributions N_ξ and N_η on $\mathbb{R}$. For example, the ordinary uncertain variable ξ uniformly distributed on $[-1, 1]$ has the uncertainty measure $N_\xi(\xi = x) = F(x) \bigwedge (1 - F(x)) \neq 0$ for all $x \in (-1, 1)$ by (2.3). On the other hand, the uniformly distributed uncertain variable η with probability $N_\eta((a, b]) = (b - a)/2$, $a, b \in [-1, 1]$, has $N_\eta(\eta = x) = 0$ for all $x \in (-1, 1)$.

In contrast to probability theory, the distribution N_ξ of an ordinary uncertain variable cannot be analytically expressed in terms of the distribution function $F_\xi(x) = N(\xi \leq x)$. Additionally, a function identically equal to a constant on $\mathbb{R}$ cannot be a distribution function in probability theory but in Liu's theory (see [3]). The results of the following lemma were proved in [13, Example 1.6] but for completeness, we present a proof, which is, moreover, simpler.

Lemma 1. *For an ordinary uncertain variable ξ uniformly distributed on $[-1, 1]$, $\mathrm{E}(\xi^2 + b\xi) = 1/3$ for all $|b| \geq 2$ and $\mathrm{E}\xi^2 = 7/24$.*

P r o o f. Let $b \geq 2$. The function $f(x) = x^2 + bx$ strictly increases on $[-1, 1]$ from $1 - b$ to $1 + b$. If $\eta = f(\xi)$, then $F^{-1}_\eta = f(F^{-1}_\xi)$ by (2.4). Therefore,

$$\mathrm{E}f(\xi) = \int_0^1 F^{-1}_\eta(a)da = \int_0^1 \left((2a-1)^2 + b(2a-1)\right) da = (2a-1)^3/6 + b(2a-1)^2/4\Big|_0^1 = 1/3$$

by (2.4). If $b \leq -2$, then the function $f(x) = x^2 + bx$ strictly decreases on $[-1, 1]$ from $1 - b$ to $1 + b$. We have $\eta = f(\xi)$ and $F^{-1}_\eta(a) = f(F^{-1}_\xi(1 - a))$ by Theorem 1. So, $\mathrm{E}\eta = 1/3$ as well.

Now let $\eta = \xi^2$. Compute F_η via (2.3). Define $x_1 = -\sqrt{x}$ and $x_2 = \sqrt{x}$. We have

$$F_\eta(x) = N_\xi(x_1 \leq \xi \leq x_2).$$

Since $[x_1, x_2] = (-\infty, x_2] \bigcap [x_1, \infty)$, we have

$$F_\xi(x_2) \bigwedge (1 - F_\xi(x_1)) = (x_2 + 1)/2 \geq 1/2.$$

For the complement of $[x_1, x_2]$, we obtain $[x_1, x_2]^c = (-\infty, x_1) \bigcup (x_2, \infty)$ and

$$F_\xi(x_1) + 1 - F_\xi(x_2) = 1 - \sqrt{x}.$$

Therefore,

$$F_\eta(x) = \begin{cases} 0 & \text{if} \quad x < 0, \\ 0.5 & \text{if} \quad \sqrt{x} \in [0, 1/2], \\ \sqrt{x} & \text{if} \quad \sqrt{x} \in (1/2, 1], \\ 1 & \text{if} \quad \sqrt{x} > 1. \end{cases}$$

Finally,

$$\mathrm{E}\eta = \int_0^1 x dF_\eta(x) = \int_{1/4}^1 \sqrt{x} dx/2 = 7/24.$$

□

In what follows, we need the distribution function of $\eta = (1 + \lambda\xi)^2$, $|\lambda| \le 1$. It can be found by Theorem 1.10 from [13]. Define

$$x_1 = -(\sqrt{x} + 1)/\lambda, \quad x_2 = (\sqrt{x} - 1)/\lambda.$$

Using the theorem, we obtain

$$F_\eta(x) = \begin{cases} 0 & \text{if} \quad x < (1-\lambda)^2, \\ F_\xi(x_2) \bigwedge (1 - F_\xi(x_1)) & \text{if} \quad F_\xi(x_2) \bigwedge (1 - F_\xi(x_1)) < 0.5, \\ F_\xi(x_2) - F_\xi(x_1) & \text{if} \quad F_\xi(x_2) - F_\xi(x_1) > 0.5, \\ 0.5 & \text{otherwise,} \end{cases}$$

for $0 < \lambda \le 1$. By the formula for $F_\xi(x)$, we have

$$F_\eta(x) = \begin{cases} 0 & \text{if} \quad x < (1-|\lambda|)^2, \\ (\sqrt{x} - 1)/(2|\lambda|) + 1/2 & \text{if} \quad x \in [(1-|\lambda|)^2, (1+|\lambda|)^2], \\ 1, & \text{if} \quad x > (1+|\lambda|)^2. \end{cases} \tag{2.6}$$

If $-1 \le \lambda < 0$, then $x_2 < x_1$ and $\eta = (|\lambda|\xi - 1)^2$. We come to formula (2.6) as well. The inverse function for the regular uncertain variable η has the form

$$F_\eta^{-1}(x) = (|\lambda|(2x - 1) + 1)^2, \quad x \in [0, 1].$$

3. Statement of the problem

Definition 1. *A set $\mathcal{X}_m$ is called the reachable set for system (1.1) under constraints (1.2) if it consists of uncertain variables $\eta \in \mathbb{R}^n$ for which there exists a family $(x_0, \mathbb{K})$ satisfying (1.2) and such that $x_m = \eta$ with equations (1.1) satisfied. The equalities are considered N-almost everywhere.*

The *problem* is to find the reachable set $\mathcal{X}_m$. Let

$$V_m(\eta) = \min \big\{ J(x_0, \mathbb{K}) : x_m = \eta,\, x_0 \in \mathbf{X}_0,\, K_k \in \mathbb{R}^{q \times n} \big\}. \tag{3.1}$$

Then

$$\mathcal{X}_m = \{\eta : V_m(\eta) \le 1\}.$$

Note that if $0 \in \mathbf{X}_0$, then always $0 \in \mathcal{X}_m$. Further, we exclude the uncertain variables η for which $V_m(\eta) = -\infty$. This can be so because the matrices R_k and Q_k in (1.2) can be nonpositive definite.

4. Main results

First, we transform the computation of $V_m(\eta)$ into an equivalent deterministic optimal control problem.

Let $X_k = \mathrm{E}(x_k x_k')$. Since $x_k \in \mathbb{R}^n$, the matrix $x_k x_k'$ belongs to $\mathbb{R}^{n \times n}$ and its elements are uncertain variables. Therefore, X_k is a symmetric matrix for all $k \in 1 : m$.

Theorem 2. *Let $\Xi = E\eta\eta'$. If the minimization problem* (3.1) *has a finite value $V_m(\eta)$, then it is equivalent to the following deterministic optimal control problem*:

$$\begin{gathered}\mathcal{V}_m(\Xi) = \min\left\{\mathbf{J}(x_0,\mathbb{K}) : X_m = \Xi,\ x_0 \in \mathbf{X}_0,\ K_k \in \mathbb{R}^{q\times n}\right\},\\ \mathbf{J}(x_0,\mathbb{K}) = \sum_{k\in 1:m} \operatorname{tr}\left(\left(K_k' R_k K_k + Q_k\right) X_{k-1}\right),\end{gathered} \tag{4.1}$$

where tr means the trace of a matrix, under the recurrent relations

$$\begin{gathered}X_k = \nu_k U_k(K_k, X_{k-1}), \quad X_0 = x_0 x_0', \quad \nu_k = \mu_k/\mu_{k-1},\\ U_k(K_k, X) = (A_k + B_k K_k) X (A_k + B_k K_k)',\\ \mu_k = \int_0^1 (|\lambda_k|(2x-1)+1)^2 \dots (|\lambda_1|(2x-1)+1)^2\, dx.\end{gathered} \tag{4.2}$$

The functionals $J(x_0,\mathbb{K})$ and $\mathbf{J}(x_0,\mathbb{K})$ coincide and therefore $\mathcal{V}_m(\Xi) = V_m(\eta)$.

P r o o f. Using equality (1.1), we have

$$\begin{gathered}x_k x_k' = U_k\left(K_k, x_{k-1}x_{k-1}'\right)(1+\lambda_k\xi_k)^2\\ = (A_k + B_k K_k)\dots(A_1 + B_1 K_1) x_0 x_0' (A_1 + B_1 K_1)' \dots (A_k + B_k K_k)'\\ \times (1+\lambda_k\xi_k)^2 \dots (1+\lambda_1\xi_1)^2, \text{ and, therefore,}\\ X_k = (A_k + B_k K_k)\dots(A_1 + B_1 K_1) x_0 x_0' (A_1 + B_1 K_1)' \dots (A_k + B_k K_k)'\\ \times \mathrm{E}((1+\lambda_k\xi_k)^2 \dots (1+\lambda_1\xi_1)^2).\end{gathered} \tag{4.3}$$

By Theorem 1 and formulas (2.4), (2.5), and (2.6), we see that

$$\mathrm{E}((1+\lambda_k\xi_k)^2 \dots (1+\lambda_1\xi_1)^2) = \mu_k$$

in (4.2). Thus, $X_0 = x_0 x_0'$ and $X_k = \nu_k U_k(K_k, X_{k-1})$ for $k \in 1:m$. The coincidence of $J(x_0,\mathbb{K})$ and $\mathbf{J}(x_0,\mathbb{K})$ follows from (4.3) and the equalities

$$J(x_0,\mathbb{K}) = \sum_{k\in 1:m} \mathrm{E} x_{k-1}' \left(K_k' R_k K_k + Q_k\right) x_{k-1}$$

$$= \sum_{k\in 1:m} \mathrm{E}\operatorname{tr}\left(\left(K_k' R_k K_k + Q_k\right) x_{k-1}x_{k-1}'\right) = \sum_{k\in 1:m} \operatorname{tr}\left(\left(K_k' R_k K_k + Q_k\right) X_{k-1}\right) = \mathbf{J}(x_0,\mathbb{K}).$$

□

Remark 1. We can compute $\mathrm{E}(1+\lambda\xi)^2 = 1+\lambda^2\mathrm{E}(2/\lambda\xi+\xi^2) = 1+\lambda^2/3$ by Lemma 1. Suppose that $|\lambda_k| \equiv 1$. Then $\mu_k = 4^k \int_0^1 x^{2k} dx = 4^k/(2k+1)$. Therefore, $\lim_{k\to\infty} \nu_k = 4$ in this case.

Remark 2. Note that the properties $R_k \geq 0$ and $Q_k \geq 0$ were not used in the proof of Theorem 2. If these properties hold, then we have $V_m(\eta) = \mathcal{V}_m(X_m) \geq 0$.

Corollary 2. *The reachable set is $\mathcal{X}_m = \{\eta : \mathcal{V}_m(E\eta\eta') \leq 1\}$.*

Problem (4.1) is determinate. So, we seek the minimum of the smooth functional $\mathbf{J}(x_0,\mathbb{K})$ in (4.1) under the equality conditions (4.2) with given $X_m = \Xi$. According to the Kuhn–Tucker theorem, we form the Lagrange function

$$\mathcal{L} = \mathbf{J}(x_0,\mathbb{K}) + \sum_{k\in 1:m} \operatorname{tr}\left(H_k(\nu_k U_k(K_k, X_{k-1}) - X_k)\right) + \operatorname{tr}(\Gamma(X_m - \Xi)), \quad X_0 = x_0 x_0',$$

where the symmetric matrices $H_{1:m}$ and Γ are the Lagrange multipliers. Let us write necessary optimality conditions (x_0 is fixed):

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial K_k} &= 2R_kK_kX_{k-1} + 2\nu_kB_k'H_kA_kX_{k-1} + 2\nu_kB_k'H_kB_kK_kX_{k-1} = 0, \quad k \in 1:m,\\ \frac{\partial\mathcal{L}}{\partial X_{k-1}} &= K_k'R_kK_k + \nu_k(A_k+B_kK_k)'H_k(A_k+B_kK_k) - H_{k-1} + Q_k = 0, \quad k \in 2:m,\\ \frac{\partial\mathcal{L}}{\partial X_m} &= -H_m + \Gamma = 0.\end{aligned} \tag{4.4}$$

Here, we use the formula of differentiation

$$\frac{\partial\,\mathrm{tr}\,(A'B)}{\partial A} = B$$

for matrices of appropriate dimensions. Define

$$L_k = R_k + \nu_kB_k'H_kB_k, \quad M_k = \nu_kB_k'H_kA_k. \tag{4.5}$$

To resolve equalities equalities (4.4), we set

$$K_k = -L_k^+M_k + Y_k - L_k^+L_kY_k.$$

This expression satisfies the equation $L_kK_k = -M_k$ if and only if $L_kL_k^+M_k = M_k$. Here A^+ is the pseudoinverse matrix and $Y_k \in \mathbb{R}^{q\times n}$ is an arbitrary matrix. Substituting the expression for K_k into the second row of (4.4), we obtain the equation

$$H_{k-1} = \nu_kA_k'H_kA_k - M_k'L_k^+M_k + Q_k, \quad H_m = \Gamma, \quad k \in 1:m. \tag{4.6}$$

This equation determines H_0 for $k=1$. We come to the conclusion.

Theorem 3. *Let the value $V_m(\eta) = \mathcal{V}_m(\Xi)$ be finite and be reached at the pair $(x_0, \mathbb{K}^0)$. Then there exist symmetric matrices $H_{1:m}$ and Γ satisfying equations (4.6) with the matrix coefficients given in (4.5) and*

$$L_kL_k^+M_k = M_k \quad \textit{and} \quad L_k \geq 0. \tag{4.7}$$

Moreover, optimal matrices have the form

$$K_k^0 = -L_k^+M_k + Y_k - L_k^+L_kY_k,$$

where the matrices $Y_k \in \mathbb{R}^{q\times n}$ are arbitrary. The optimal value is

$$V_m(\eta) = \mathcal{V}_m(\Xi) = \min\left\{-\mathrm{tr}(\Gamma\Xi) + x_0'H_0x_0 : x_0 \in \mathbf{X}_0,\, X_m = \Xi\right\}.$$

P r o o f. From (4.1), we know that

$$\begin{aligned}\mathbf{J}(x_0,\mathbb{K}) &= \sum_{k\in 1:m} \mathrm{tr}\left(\left(K_k'R_kK_k + Q_k\right)X_{k-1}\right)\\ &= \sum_{k\in 1:m}\left\{\mathrm{tr}\left(\left(K_k'R_kK_k + Q_k\right)X_{k-1}\right) + \mathrm{tr}\left(H_kX_k\right) - \mathrm{tr}\left(H_{k-1}X_{k-1}\right)\right\} - \mathrm{tr}(\Gamma\Xi) + x_0'H_0x_0.\end{aligned} \tag{4.8}$$

Substituting X_k and H_{k-1} from (4.2) and (4.6) into (4.8), we can write the cost functional as follows:

$$\begin{aligned}\mathbf{J}(x_0,\mathbb{K}) &= \sum_{k\in 1:m} \operatorname{tr}\left(\left(K_k' L_k K_k + M_k' K_k + K_k' M_k + M_k' L_k^+ M_k\right) X_{k-1}\right) - \operatorname{tr}(\Gamma\Xi) + x_0' H_0 x_0 \\ &= \sum_{k\in 1:m} \operatorname{tr}\Big(\left(K_k + L_k^+ M_k\right)' L_k \left(K_k + L_k^+ M_k\right) X_{k-1}\Big) - \operatorname{tr}(\Gamma\Xi) + x_0' H_0 x_0. \end{aligned} \tag{4.9}$$

Here $L_k \geq 0$. If $L_p \ngeq 0$ for some $p \in 1:m$, then there exist a vector h and a number $\alpha < 0$ such that $L_p h = \alpha h$. Let $N = [\underbrace{h,\dots,h}_{\text{n vectors}}]$. Then $L_p N = \alpha N$. If

$$K_k = -L_k^+ M_k, \quad k \neq p, \quad K_p = -L_p^+ M_p + \delta N/\sqrt{|\alpha|},$$

then

$$\lim_{\delta\to\infty} \mathbf{J}(x_0,\mathbb{K}) = -\infty.$$

□

These relations are sufficient for optimality.

Theorem 4. *Equations (4.6) along with relations (4.2), (4.5), (4.7) are sufficient for finiteness of values $\mathcal{V}_m(\Xi) = V_m(\eta) > -\infty$, and optimal values K_k^0 with corresponding minimum are specified in Theorem 3. The system contains $2mn(n+1)/2$ equations with the same quantity of variables, namely, $mn(n+1)/2$ variables $H_{0:m-1}$, $n(n+1)/2$ variables Γ, and $(m-1)n(n+1)/2$ variables $X_{1:m-1}$.*

Indeed, if relations (4.2), (4.5), (4.7) are valid, then

$$\mathbf{J}(x_0,\mathbb{K}) \geq -\operatorname{tr}(\Gamma\Xi) + x_0' H_0 x_0$$

according to (4.9).

Corollary 3. *Consider the matrix*

$$\mathbf{H}_0 = \prod_{k\in 0:m-1} (A_{m-k} + B_{m-k} K_{m-k}^0).$$

It follows from (4.3) that $X_m = \mu_m \mathbf{H}_0 X_0 \mathbf{H}_0'$. Therefore, $\operatorname{tr}(\Gamma X_m) = \mu_m x_0' \mathbf{H}_0' \Gamma \mathbf{H}_0 x_0$. The reachable set is

$$\mathcal{X}_m = \{\eta : \min\{x_0'(H_0 - \mu_m \mathbf{H}_0' \Gamma \mathbf{H}_0) x_0 : x_0 \in \mathbf{X}_0,\ \mu_m \mathbf{H}_0 X_0 \mathbf{H}_0' = \Xi = E\eta\eta'\} \leq 1\}.$$

The minimization is provided here under inequalities (4.7).

5. Example

Consider the 2-dimensional system (1.1) in which

$$n = m = 2, \quad q = 1, \quad \lambda_1 = 0.2, \quad \lambda_2 = -0.1,$$
$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}, \quad B_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad B_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Under constraints (1.2), we have

$$R_1 = p, \quad R_2 = 4, \quad Q_1 = p\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \quad Q_2 = p\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},$$

where $p \geq -1$ is a numeric parameter. There are 12 variables and the same number of equations. Let

$$X_1 = (x_{1ij}), \quad H_k = (h_{kij}), \quad \Gamma = (\gamma_{ij}), \quad x_0 = (x_{0i}), \quad K_k = (k_{ki}), \quad M_k = (m_{ki}).$$

For $k = 1$, we obtain

$$h_{011} = p\nu_1 h_{111}/(p + \nu_1 h_{111}) + p, \quad h_{012} = 0, \quad h_{022} = p, \quad k_{11} = -\nu_1 h_{111}/(p + \nu_1 h_{111}), \quad k_{12} = 0,$$

where

$$\nu_1 = \mu_1 = 1 + \lambda_1^2/3 = 76/75 = 1.0133.$$

From equation (4.2), we have $x_{111} = \nu_1(1 + k_{11})^2 x_{01}^2$and $x_{1i2} = 0$. For $k = 2$, we have $\xi_{ij} = \nu_2(1 + k_{21})^2 x_{111}$ for any i and j. Let $\xi_{ij} = \xi$ and $k_{22} = 0$. On the other hand, since $H_2 = \Gamma$, we have $k_{21} = -\nu_2\gamma/(4 + \nu_2\gamma)$, where $\gamma = \gamma_{11} + \gamma_{22} + 2\gamma_{12}$. Finally, from equation (4.6) for $k = 2$, we obtain

$$h_{111} = \nu_2\gamma + p - \nu_2^2\gamma^2/(4 + \nu_2\gamma), \quad h_{112} = h_{122} = 0.$$

Here

$$\mu_2 = \int_0^1 (|\lambda_1|(2x - 1) + 1)^2(|\lambda_2|(2x - 1) + 1)^2 dx = 1.0434, \quad \nu_2 = \mu_2/\mu_1 = 1.0297.$$

These equations imply that an uncertain variable $\eta = (\eta_1, \eta_2)$ belongs to $\mathcal{X}_2$ if and only if $E\eta_1^2 = E\eta_2^2 = E\eta_1\eta_2$. A family of uncertain variables of the form $\eta_1 = \eta_2 = \alpha$, where α is some uncertain variable, satisfies these conditions.

Consider the value $\xi = Fx_{01}^2$, where $F = \nu_1\nu_2(1 + k_{11})^2(1 + k_{21})^2$. Let x_0 be fixed and $f(\gamma) = h_{011}/F - \gamma$. Then

$$V_2(\eta) = \mathcal{V}_2(\Xi) = \xi \min\{f(\gamma) : L_i \geq 0\} + px_{02}^2$$

for $p > 0$, where $L_2 = 4 + \nu_2\gamma$ is the increasing linear function of γ. We see that the expression $f(\gamma) = h_{011}/F - \gamma$ in the braces depends only on γ, but the entire problem is to minimize the function $\mathcal{V}_2(\Xi)$ of two variables γ and x_{02} under nonlinear constraints and the equality condition $\xi = Fx_{01}^2$. Let $x_{02} = 1$ and $|x_{01}| \leq 1$ for simplicity. We set $p = 0.5$, for example. Computing the minimum of the smooth convex function $f(\gamma)$, we have $\min f(\gamma) = 0.4316$, and it is achieved at $\gamma^0 = -0.4316$. All the constraints are satisfied. So, $\mathcal{V}_2(\Xi) = 0.4316\xi + 0.5 \leq 1$ or $\xi \in [0, 1.1584]$. As matrix Γ, it is possible to take any symmetric matrix with $\gamma^0 = -0.4316$. If $p \downarrow 0$, then $\xi \in [0, b(p)]$, where $b(p) \to \infty$. If $-1 \leq p < 0$, then the value $V_m(\eta) = \mathcal{V}_m(\Xi) \leq 0$ is also finite for all $\xi \geq 0$. This means that $\xi \in [0, \infty)$.

6. Conclusion

- The discrete-time estimation problem has been considered for one class of uncertain Liu processes whose equations include unknown deterministic parameters subject to a priori constraints.
- The initial value problem is reduced to a deterministic multi-step problem for matrices with a fixed constraint at the right end of the trajectory.
- Necessary and sufficient conditions for the finiteness of the objective functional in the deterministic problem are obtained.

• A numerical solution of the initial value problem is considered with an example.

• In the general case, since the expectation has no, generally speaking, property of additivity, the reduction of problems with uncertain Liu disturbances to determinate ones is difficult. The received determinate problem is also unusual because it deals with implicit matrix equations.

• The ordinary uniformly distributed uncertain variables in this paper can be easily replaced by any regular and independent Liu variables.

• A similar problem for continuous systems will be considered in the future.

REFERENCES

1. Ananyev B. I. A state estimation of Liu equations. *AIP Conf. Proc.*, 2015. Vol. 1690. Art. no. 040005. DOI: 10.1063/1.4936712
2. Ananyev B. I. State estimation of some class of discrete Liu's processes. In: *Proc. XIV All-Russian Conference on Control Problems (VSPU-2024), Moscow, June 17–20, 2024*. Moscow: IPU RAS, 2024. P. 604–608. URL: https://vspu2024.ipu.ru/prcdngs (in Russian)
3. Ananyev B. I. On some complements to Liu's theory. *Proc. Steklov Inst. Math.*, 2024. Vol. 325, No. Suppl. 1. P. S1–S16. DOI: 10.1134/S0081543824030015
4. Bulinskii A. V., Shiryaev A. N. *Teoriya sluchainikh processov* [The Theory of Random Processes]. Moscow: Fizmatlit, 2003. 400 p. (in Russian)
5. Chernousko F. L. An estimate for the reachable sets of linear systems with an indefinite matrix. *Dokl. Akad. Nauk*, 1996. Vol. 349, No. 1. P. 32–34. (in Russian)
6. Hou Y., Yang Sh. Uncertain programming with recourse. *J. Chem. Pharmaceutical Res.*, 2015. Vol. 7, No. 3. P. 2366–2372.
7. Kurzhanski A., Varaiya P. *Dynamics and Control of Trajectory Tubes: Theory and Computation.* SCFA. Boston: Birkhäuser, 2014. 445 p. DOI: 10.1007/978-3-319-10277-1
8. Liu B. Toward uncertain finance theory. *J. Uncertain. Anal. Appl.*, 2013. Vol. 1. Art. no. 1. 15 p. DOI: 10.1186/2195-5468-1-1
9. Liu B. *Uncertainty Theory.* 4-th ed. Berlin, Heidelberg: Springer, 2015. 487 p. DOI: 10.1007/978-3-662-44354-5
10. Pytyev Yu. P. *Vozmozhnost kak alternativa veroyatnosti. Matematicheskiye i empiricheskiye osnovy, prilozheniya. 2-e izd.* [Possibility as the Alternative of Probability. Mathematical and Emperical Grounds, Applications. 2-nd ed.]. Moscow: Fizmatlit, 2016. 600 p. (in Russian)
11. Sheng L., Zhu Y., Yan H., Wang K. Uncertain optimal control approach for CO_2 mitigation problem. *Asian J. Control*, 2017. Vol. 19, No. 6. P. 1931–1942. DOI: 10.1002/asjc.1524
12. Yao K. *Uncertain Differential Equations.* Berlin, Heidelberg: Springer, 2016. 158 p. DOI: 10.1007/978-3-662-52729-0
13. Zhu Y. *Uncertain Optimal Control.* Berlin, Heidelberg: Springer, 2019. 208 p. DOI: 10.1007/978-981-13-2134-4

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 25–36

DOI: 10.15826/umj.2024.2.003

REDUCING GRAPHS BY LIFTING ROTATIONS OF EDGES TO SPLITTABLE GRAPHS

Vitaly A. Baransky[†], **Valentin V. Zuev**[††], **Tatiana A. Senchonok**[†††]

Ural Federal University,
51 Lenina av., Ekaterinburg, 620000, Russian Federation

[†]Vitaly.Baransky@urfu.ru [††]Valentin.Zuev@urfu.ru [†††]Tatiana.Senchonok@urfu.ru

Abstract: A graph G is splittable if its set of vertices can be represented as the union of a clique and a coclique. We will call a graph H a splittable ancestor of a graph G if the graph G is reducible to the graph H using some sequential lifting rotations of edges and H is a splittable graph. A splittable r-ancestor of G we will call its splittable ancestor whose Durfey rank is r. Let us set $s = (1/2)(\operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda))$, where $\operatorname{hd}(\lambda)$ and $\operatorname{tl}(\lambda)$ are the head and the tail of a partition λ. The main goal of this work is to prove that any graph G of Durfey rank r is reducible by s successive lifting rotations of edges to a splittable r-ancestor H and s is the smallest non-negative integer with this property. Note that the degree partition $\operatorname{dpt}(G)$ of the graph G can be obtained from the degree partition $\operatorname{dpt}(H)$ of the splittable r-ancestor H using a sequence of s elementary transformations of the first type. The obtained results provide new opportunities for investigating the set of all realizations of a given graphical partition using splittable graphs.

Keywords: Integer partition, Graphical partition, Degree partition, Splittable graph, Rotation of an edge.

1. Introduction

Everywhere we mean by a graph a simple graph, i. e., a graph without any loops and multiple edges. We will adhere to the terminology and notation from [1, 2, 6].

An *integer partition*, or simply, a *partition* is a non-increasing sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of non-negative integers that contains only a finite number of non-zero components (see [1]).

Let $\operatorname{sum}\lambda$ denote the sum of all components of a partition λ and called it the *weight* of the partition λ. It is often said that a partition λ is a partition of the non-negative integer $n = \operatorname{sum}\lambda$. The *length* $\ell(\lambda)$ of a partition λ is the number of its non-zero components. For convenience, a partition λ will often be written as $\lambda = (\lambda_1, \dots, \lambda_t)$, where $t \geq \ell(\lambda)$, i. e., we will omit zeros, starting from some zero component without forgetting that the sequence is infinite.

We will say that the partition $(\lambda_1, \dots, \lambda_i - 1, \dots, \lambda_j + 1, \dots)$ is obtained from the partition $(\lambda_1, \dots, \lambda_i, \dots, \lambda_j, \dots)$ by an elementary transformation of the first type. An elementary transformation of the second type is a reduction of some partition component by 1.

A partition can conveniently be depicted as a Ferrers diagram, which can be thought of as a set of square boxes of the same size (see the example below in Fig. 1). We will use Cartesian notation for Ferrers diagrams.

For each partition λ, we will consider a *conjugate partition* λ^* whose components are equal to the number of boxes in the corresponding rows of the Ferrers diagram of the partition λ.

We determine the *rank* $r(\lambda)$ of the partition λ by setting $r(\lambda) = \max\{i | \lambda_i \geq i\}$. Obviously, the rank $r = r(\lambda)$ of a partition λ is equal to the number of boxes on the main diagonal of the Ferrers diagram of this partition.

As the *head* $\operatorname{hd}(\lambda)$ we take the partition that is obtained from the partition λ by reducing the first r components by the same number $r - 1$ and zeroing all components with numbers $r + 1$, $r + 2, \dots$ (for an example, see the diagram in Fig. 2).

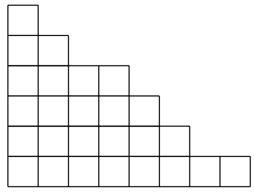


Figure 1. The Ferrers diagram of the partition $(6, 5, 4, 4, 3, 2, 1, 1)$.

As the *tail* $\mathrm{tl}(\lambda)$ we take a partition for which the Ferrers diagram of the conjugate partition is obtained from the Ferrers diagram of the partition λ by deleting the first r columns, i. e. the Ferrers diagram of the partition $\mathrm{tl}(\lambda)^*$ is located to the right of the Durfey square (see Fig. 2).

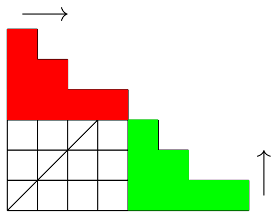


Figure 2. The head $\mathrm{hd}(\lambda) = (3, 2, 1, 1)$ and the tail $\mathrm{tl}(\lambda) = (4, 2, 1)$ of the partition $(6, 5, 4, 4, 3, 2, 1, 1)$.

The theory of partitions is one of the classical areas of combinatorics. Its foundations were laid by L. Euler. Information about the achievements of the theory of partitions can be found in [1].

This work continues the cycle of researches by V.A. Baransky, T.A. Koroleva, T.A. Senchonok and V.V. Zuev, which uses the method of elementary transformations for partitions and the associated method of rotating edges in graphs. Using these methods, new results were obtained on some details of the structure of the lattice of partitions and the properties of graphical partitions, including maximal graphical partitions. Results were also obtained on the connection of graphs with threshold graphs, and an important class of bipartite-threshold graphs was considered (a brief overview of the results obtained is contained in [2]).

Let (x, v, y) be a triple of different vertices of a graph $G = (V, E)$ such that $xv \in E$ and $vy \notin E$. We call such a triple

1) *lifting* if $\deg(x) \leq \deg(y)$,
2) *lowering* if $\deg(x) \geq 2 + \deg(y)$,
3) *preserving* if $\deg(x) = 1 + \deg(y)$.

A transformation φ of a graph G such that $\varphi(G) = G - xv + vy$, i. e., the edge xv is first removed from G and then the edge vy is added, is called *a rotation of the edge* (in the graph G *around vertex* v), corresponding to the triple (x, v, y).

The rotation of an edge in a graph G corresponding to the triple (x, v, y) is called

1) *lifting* if the triple (x, v, y) is lifting,
2) *lowering* if the triple (x, v, y) is lowering,
3) *preserving* if the triple (x, v, y) is preserving.

We will consider the cases where $\deg(x) = 1$ or $\deg(y) = 0$ are admissible, i. e., after the edge is rotated, an isolated vertex may appear, or the edge will rotate in the graph G with the addition of a new isolated vertex. Note that a rotation of an edge in a graph G is lifting if and only if its inverse rotation is lowering.

A graph G is called *splittable* (see, for example, [6]) if its set of vertices can be represented as the union of a clique and a coclique.

These graphs were introduced in [3], where it was shown, that G is splittable if and only if it does not have an induced subgraph isomorphic to one of the three forbidden graphs C_4, C_5, or $2K_2$.

R.I. Tyshkevich used splittable graphs to study unigraphic partitions, i.e., graphic partitions that have a unique realization up to isomorphism and isolated vertices [9].

Many other characterizations and properties of splittable graphs have been discovered (see [7, Ch. 8–9] and [8]). Among them is the fact that whether a graph G is splittable can be determined from its degree sequence $\mathrm{dpt}(G)$ [5]. In our terminology, such a condition is equivalent to the following equality $\mathrm{sum}(\mathrm{hd}(\lambda)) = \mathrm{sum}(\mathrm{tl}(\lambda))$, where $\lambda = \mathrm{dpt}(G)$ [2]. Note that the graph G is threshold if and only if $\mathrm{hd}(\lambda)) = \mathrm{tl}(\lambda)$.

We will call a graph H a splittable ancestor of a graph G if the graph G is reducible to the graph H using some sequential lifting rotations of edges and H is a splittable graph. Note that the graph G can be obtained from H by sequentially performing lowering rotations of edges. It is important to note that therefore $\mathrm{dpt}(H)$ can be obtained from $\mathrm{dpt}(G)$ using elementary transformations of the first type. This means that the partition $\mathrm{dpt}(H)$ lies above the partition $\mathrm{dpt}(G)$ in the lattice of all partitions of the weight $\mathrm{sum}(\mathrm{dpt}(G))$.

A *Durfey rank* of a graph G is the rank (i. e., Durfey rank) of its degree partition, i. e., the number of boxes on the main diagonal of the Ferrers diagram of $\mathrm{dpt}(G)$.

Let G be an arbitrary graph with vertex set V, r is the Durfey rank and n is the cardinality of vertices of G. Let q be a natural number such that $1 \le q < n$.

An ordered pair (V_1, V_2) of subsets of a set V will be called a *2-decomposition of rank* q of the set V if $|V_1| = q$, $|V_2| = n - q$ and $V = V_1 \bigsqcup V_2$, i. e., V is the disjoint union of the sets V_1 and V_2 (here the sets V_1 and V_2 do not intersect). We will sometimes omit the words "rank q" if we know what rank we are talking about. The sets V_1 and V_2 will be called the first and second components of the 2-decomposition, respectively.

In this work, we will consider 2-decompositions of rank r of the set V, where r is the Durfey rank of G, i. e., at $q = r$.

Among 2-decompositions (V_1, V_2) of rank r of the set V we select special 2-decompositions, which we will call *principal 2-decompositions* of the graph G, for which all vertices of the set V_1 have degrees greater than or equal to r, and all vertices of the set V_2 have degrees less than or equal to r.

Let the degree partition of the graph G have the form

$$\lambda = \mathrm{dpt}(G) = (\lambda_1, \dots, \lambda_r, \lambda_{r+1}, \dots, \lambda_n),$$

r is the rank of the partition λ and n is the number of vertices. Let us order the set of vertices $V = \{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ of the graph G in such a way that

$$\lambda_1 = \deg\, v_1 \ge \dots \ge \lambda_r = \deg\, v_r \ge r \ge \lambda_{r+1} = \deg\, v_{r+1} \ge \dots \ge \lambda_n = \deg\, v_n.$$

We can obtain the principal 2-decomposition (V_1, V_2) of the graph G by setting

$$V_1 = \{v_1, \dots, v_r\} \quad \text{and} \quad V_2 = \{v_{r+1}, \dots, v_n\}.$$

Let $u \in V_1$, $v \in V_2$ and vertices u and v have the same degrees equal to r. Let us move on to a new principal 2-decomposition (V_1', V_2') of the graph G by setting

$$V_1' = V_1 - u + v \quad \text{and} \quad V_2' = V_2 - v + u.$$

This procedure we will called a *procedure of exchanging vertices of degree* r from the sets V_1 and V_2. It transforms the principal 2-decomposition (V_1, V_2) to the principal 2-decomposition (V_1', V_2') of the graph G. It is clear that using sequences of exchanges of vertices of degree r from any principal 2-decomposition (V_1, V_2) one can obtain all principal 2-decompositions of the graph G. It is easy to see that the principal 2-decompositions of a graph G can differ from each other only by vertices of degree r in the first and the second components.

Let (V_1, V_2) be an arbitrary 2-decomposition of the set V of vertices of a graph G. Then the set E of all edges of the graph G connecting vertices from V_1 with vertices from V_2 will be called a *section* of the graph G corresponding to the 2-decomposition (V_1, V_2). We will call a bipartite graph (V_1, E, V_2) the *sandwich subgraph* of the graph G corresponding to the 2-decomposition (V_1, V_2).

We have $\mathrm{hd}(\lambda) \leq \mathrm{tl}(\lambda)$ by virtue of the ht-criterion [2], where $\mathrm{hd}(\lambda)$ and $\mathrm{tl}(\lambda)$ are the head and the tail of the partition λ, and the integer $\mathrm{sum}\,\mathrm{tl}(\lambda) - \mathrm{sum}\,\mathrm{hd}(\lambda)$ is even. Let us set

$$s = \frac{1}{2}(\mathrm{sum}\,\mathrm{tl}(\lambda) - \mathrm{sum}\,\mathrm{hd}(\lambda)).$$

Let r is the Durfey rank of a graph G. A *splittable r-ancestor* of G we will called its splittable ancestor whose Durfey rank is r.

The main goal of this study is to prove the following theorem.

Theorem 1. *Let G be an arbitrary graph whose Durfey rank is equal to r and $\lambda = \mathrm{dpt}\, G$.*

1. *Let (V_1, V_2) be a principal 2-decomposition of the graph G. Then the graph G is reduced to a splittable r-ancestor $H' = (K(V_1), E', V_2)$ by means of some sequential execution of s lifting edge rotations, and s is the smallest non-negative integer with this property.*
2. *Let (V_1', V_2') be a non-principal 2-decomposition of the set of vertices V of the graph G and the graph G is reducible to some splittable r-ancestor of the form $H' = (K(V_1'), E', V_2')$ by sequentially performing of t lifting rotations of edges. Then $t > s$.*

We see that any graph of Durfey rank r is reducible by s successive lifting rotations of edges to a splittable graph of Durfey rank r, and s is the smallest non-negative integer with this property.

Let a splittable graph $H' = (K(V_1), E', V_2)$ be obtained from a graph G using some sequential execution of s lifting edge rotations, where (V_1, V_2) is some 2-decomposition of the set of vertices of the graph G. Then the graph G can be obtained from the splittable graph H' using an inverse sequence consisting of s lowering edge rotations. Therefore, the degree partition $\mathrm{dpt}(G)$ of the graph G can be obtained from the degree partition $\mathrm{dpt}(H')$ of the graph H' using a sequence of s elementary transformations of the first type [2].

Let r is the Durfey rank of a graph G. Its *closest splittable r-ancestor* is a splittable graph H', which has Durfey rank r and which can be obtained from the graph G by some sequential execution of s lifting rotations of edges.

Next, we present an algorithm (see Algorithm 1 and Lemma 6) for finding all closest splittable r-ancestors of a graph G.

Corollary 1. *Let G be a graph of Durfey rank r. Then the graph G is obtainable from some of its closest splittable r-ancestor using a sequence consisting of*

$$s = \frac{1}{2}(\mathrm{sum}\,\mathrm{tl}(\mathrm{dpt}(G)) - \mathrm{sum}\,\mathrm{hd}(\mathrm{dpt}(G)))$$

lowering rotations of edges, and the degree partition $\mathrm{dpt}(G)$ of the graph G is obtainable from the degree partition $\mathrm{dpt}(H')$ of the graph H' using a sequence of s elementary transformations of the first type.

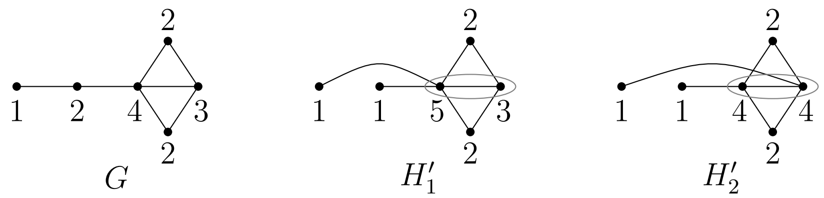


Figure 3. The graph G with Durfey rank 2 and its two non-isomorphic closest splittable 2-ancestors.

Example 1. Figure 3 shows an example of a graph G of Durfey rank 2 and its two non-isomorphic closest splittable 2-ancestors, each obtained from G by a single lifting rotation of an edge. Note that here $s = 1$ and the clique V_1 in the graphs H_1' and H_2' is two-element. It is easy to check that for each $i = 1, 2$ the degree partition $\operatorname{dpt}(G)$ of the graph G is obtained from the degree partition $\operatorname{dpt}(H_i')$ of the graph H_i' using one elementary transformation of the first type.

For a graph G of Durfey rank r, consider the *family of all closest splittable r-ancestors*, consisting of pairwise non-isomorphic graphs without isolated vertices. Let us denote this family by $CSrA(G)$.

Let $\lambda = \operatorname{dpt}(G)$. Let $CSrA(\lambda)$ denote a family of graphs that is equal to the union of families $CSrA(G)$, when G runs through all realizations of the partition λ on the set V, i. e., this is the set of all closest splittable r-ancestors of all realizations (up to isomorphism and isolated vertices) of the partition λ.

It is useful to note that the operation swap of switching edges in alternating 4-cycles does not change the set V of vertices of the graph [4], therefore all realizations of the partition λ can be considered up to isomorphism and isolated vertices on some single set V.

Note that it would be interesting to find a fairly simple description of the family $CSrA(\lambda)$, since from the graphs of this family one can obtain, by Corollary 1, all realizations of the partition λ using sequences consisting of s lowering rotations of edges. This fact makes it possible to study the family of all realizations of the partition λ without using switching edges operation [4] in graphs.

2. Proof of the main results

We first present four auxiliary lemmas and one algorithm.

Lemma 1. *Let $H = (K(V_1), E_1, V_2)$ be a splittable graph, $\mu = \operatorname{dpt}(H)$, V_1 be a clique of cardinality $r = r(\mu)$, consisting of elements of degrees $\mu_1, \ldots, \mu_r$ greater than or equal to r, and V_2 be a coclique consisting of elements of degrees $\mu_{r+1}, \ldots, \mu_n$ less than or equal to r, where n is the number of elements of the graph H. Then* $\operatorname{sum}\operatorname{hd}(\mu) = \operatorname{sum}\operatorname{tl}(\mu)$.

P r o o f. Let us remove all edges of the form $e = uv$ from the graph H, where $u, v \in V_1$. We obtain a bipartite graph $H_1 = (V_1, E_1, V_2)$, which is a sandwich subgraph of the graph H and for which $\operatorname{dpt}_{H_1}(V_1) = \operatorname{hd}(\mu)$ and $\operatorname{dpt}_{H_1}(V_2) = \operatorname{tl}^*(\mu)$ (see [2]). Therefore, we have $\operatorname{sum}\operatorname{hd}(\mu) = |E_1| = \operatorname{sum}\operatorname{tl}^*(\mu) = \operatorname{sum}\operatorname{tl}(\mu)$. □

Lemma 2. *Let (V_1, V_2) be an arbitrary principal 2-decomposition of a graph G whose Durfey rank is r. Let $e = vx$ be an edge of the graph G such that $v, x \in V_2$. Then there is a vertex $y \in V_1$ for which the triple (x, v, y) is lifting (see Fig. 4). Let us denote by H the graph that obtainable from the graph G using the lifting rotation of edge corresponding to this triple. Then for the graph H we have*

- $\deg u \geq r$ *for any vertex $u \in V_1$;*

- $\deg u \le r$ *for any vertex* $u \in V_2$;
- *the Durfey rank of the graph* H *is equal to* r *and the pair* (V_1, V_2) *remains a principal 2-decomposition for graph* H;
- $\operatorname{sum}\operatorname{hd}(\lambda) + 1 = \operatorname{sum}\operatorname{hd}(\eta)$, $\operatorname{sum}\ \operatorname{tl}(\lambda) - 1 = \operatorname{sum} tl(\eta)$, *where* $\eta = \operatorname{dpt}(H)$.

P r o o f. Since $v \in V_2$, we have $\deg v \le r$. If v is adjacent to all vertices from V_1, then, taking into account the edge $e = vx$, we obtain $\deg v \ge r + 1$ which is contradictory. Therefore, there is a vertex $y \in V_1$ that is not adjacent to v and for which obviously holds $\deg y \ge r \ge \deg x$ (see Fig. 4). Therefore, the triple (x, v, y) is lifting. It is clear that for a graph H obtained from a graph G using

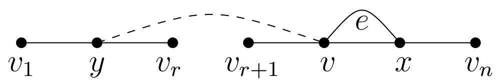


Figure 4. The lifting rotation of the edge $e = vx$ in the graph G.

the lifting rotation corresponding to this triple, the conclusions of the lemma are satisfied in the obvious way. □

Lemma 3. *Let* (V_1, V_2) *be an arbitrary principal 2-decomposition of a graph* G *whose Durfey rank is equal to* r. *Let the vertices* $y, v \in V_1$ *of the graph* G *be distinct and not adjacent. Then there is a vertex* $x \in V_2$ *for which the triple* (x, v, y) *is lifting (see Fig. 5). Let us denote by* H *the graph that obtainable from the graph* G *using the lifting rotation of edge corresponding to this triple. Then for the graph* H *holds*

- $\deg u \ge r$ *for any vertex* $u \in V_1$;
- $\deg u \le r$ *for any vertex* $u \in V_2$;
- *Durfey rank of the graph* H *is equal to* r *and the pair* (V_1, V_2) *remains a principal 2-decomposition for graph* H;
- $\operatorname{hd}(\lambda) + 1 = \operatorname{sum}\operatorname{hd}(\eta)$, $\operatorname{sum}\operatorname{tl}(\lambda) - 1 = \operatorname{sum}\operatorname{tl}(\eta)$, *where* $\eta = \operatorname{dpt}(H)$.

P r o o f. If v is not adjacent to all vertices from V_2, then by virtue of the equality $|V_1| = r$ we have $\deg v < r$, which is contradictory. Therefore, there is a vertex $x \in V_2$ that is adjacent to v and for which it obviously holds $\deg y \ge r \ge \deg x$ (see Fig. 5).

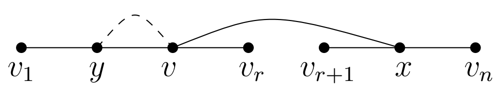


Figure 5. The lifting rotation of edge $e = vx$ in the graph G.

Therefore, the triple (x, v, y) is lifting. It is clear that for a graph H obtained from a graph G using the lifting rotation corresponding to this triple, the conclusions of the lemma are satisfied in the obvious way. □

Let (V_1, V_2) be an arbitrary 2-decomposition of rank r of the set V of vertices of a graph G.

Let by $W_1(G, V_1, V_2)$ we denote the set of all pairs of non-adjacent distinct vertices from V_1, and by $W_2(G, V_1, V_2)$ we denote the set of all pairs of adjacent vertices from V_2, i. e., the number of pairs of vertices from V_2 that are connected by edges of the graph G.

Through

$$w(G, V_1, V_2) = w_1(G, V_1, V_2) + w_2(G, V_1, V_2)$$

we will denote the *weight* of the 2-decomposition (V_1, V_2) of the graph G, where

$$w_1 = |W_1(G, V_1, V_2)| \text{ and } w_2 = |W_2(G, V_1, V_2)|.$$

Let (V_1, V_2) be any principal 2-decomposition of a graph G. Based on Lemmas 2 and 3, the following algorithm obviously leads to the construction of a splittable r-ancestors of the graph G of the form $H' = (K(V_1), E', V_2)$ using $t = w(G, V_1, V_2)$ lifting rotations of edges.

Algorithm 1. Let G be an arbitrary graph of Durfey rank r, $\lambda = \deg G$ and (V_1, V_2) be a principal 2-decomposition.

1. Let $H_0 = G$.
2. Let the graph H_i be constructed from the graph H_0 using i lifting rotations of edges, where

$$0 \leq i < w(G, V_1, V_2) \text{ and } w(H_i, V_1, V_2) = w(G, V_1, V_2) - i.$$

   Perform any of the following two actions (a) or (b).

   (a) If there is an edge $e = vx$ of the graph H_i such that $v, x \in V_2$, then by Lemma 2 there is a vertex $y \in V_1$ for which the triple (x, v, y) is lifting. Let us denote by H_{i+1} the graph that is obtained from the graph H_i using the lifting rotation of edge corresponding to this triple.
   (b) If there are two distinct non-adjacent vertices $y, v \in V_1$ of the graph H_i, then by Lemma 3 there is a vertex $x \in V_2$ for which the triple (x, v, y) is lifting. Let us denote by H_{i+1} the graph that is obtained from the graph H_i using the lifting rotation of edge corresponding to this triple.

3. Step 2 perform t times, where $t = w(G, V_1, V_2)$. As a result, a splittable r-ancestor $H_t = (K(V_1), E_t, V_2)$ of the graph G will be constructed.

Proving Theorem 1, we will establish along the way that using Algorithm 1 we can find all the closest splittable r-ancestors of the graph G.

Let (V_1, V_2) be an arbitrary 2-decomposition of rank r of the set V of vertices of a graph G. Then $|V_1| = r$ and $|V_2| = n - r$, where n is the number of vertices of the graph G.

Let $u \in V_1$ and $v \in V_2$. Let by $w(G, u \in V_1, v \in V_2)$ we denote the sum of the number of vertices from V_1 that are not adjacent to u and distinct from u, as well as the number of vertices from V_2 adjacent to v. We will call this integer by a *contribution* of the pair of vertices u and v to the weight $w(G, V_1, V_2)$ of the 2-decomposition (V_1, V_2) of the graph G.

For an arbitrary vertex z of the graph G, let $D_1(z)$ and $D_2(z)$ denote, respectively, the number of vertices from V_1 and V_2 adjacent to vertex z. Let us also put $d_1(z) = |D_1(z)|$ and $d_2(z) = |D_2(z)|$. Then obviously $\deg z = d_1(z) + d_2(z)$.

Lemma 4. *Let (V_1', V_2') be an arbitrary 2-decomposition of rank r of the set V of vertices of a graph G, where r is the Durfey rank of this graph. Let $u \in V_1'$ and $v \in V_2'$. Let us put $V_1'' = V_1' - u + v$ and $V_2'' = V_2' - v + u$. (This procedure we will call, as before, the exchanging vertices in 2-decomposition.) Then the 2-decomposition (V_1'', V_2'') has rank r and it holds*

1) *if $\deg u < \deg v$ in graph G, then $w(G, V_1'', V_2'') < w(G, V_1', V_2')$;*
2) *if $\deg u = \deg v$ in graph G, then $w(G, V_1'', V_2'') = w(G, V_1', V_2')$.*

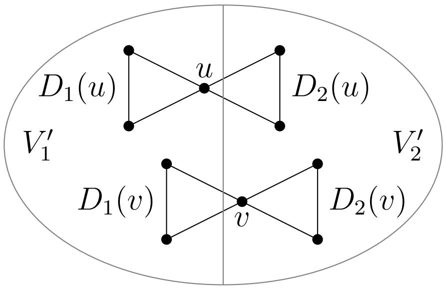


Figure 6. Sets $D_1(u)$ and $D_1(v)$, as well as sets $D_2(u)$ and $D_2(v)$ may intersect.

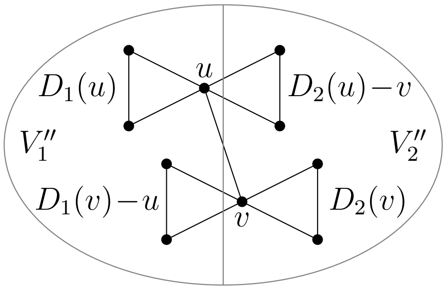


Figure 7. Sets $D_1(u)$ and $D_1(v) - u$, as well as sets $D_2(u) - v$ and $D_2(v)$ may intersect.

P r o o f. Let us first consider two cases.
1 case. Let vertices u and v be not adjacent in the graph G. Then (see Fig. 6)

$$w(G, u \in V_1', v \in V_2') = r - 1 - d_1(u) + d_2(v) = r - 1 - d_1(u) + \deg v - d_1(v),$$
$$w(G, v \in V_1'', u \in V_2'') = r - 1 - d_1(v) + d_2(u) = r - 1 - d_1(v) + \deg u - d_1(u).$$

2 case. Let vertices u and v be adjacent in the graph G. Then (see Fig. 7)

$$w(G, u \in V_1', v \in V_2') = r - 1 - d_1(u) + d_2(v) = r - 1 - d_1(u) + \deg v - d_1(v),$$
$$w(G, v \in V_1'', u \in V_2'') = r - 1 - (d_1(v) - 1) + (d_2(u) - 1)$$
$$= r - 1 - d_1(v) + 1 + \deg u - d_1(u) - 1 = r - 1 - d_1(v) + \deg u - d_1(u).$$

Thus, in each of the two cases considered, following equalities are satisfied

$$w(G, u \in V_1', v \in V_2') = r - 1 - d_1(u) + \deg v - d_1(v),$$
$$w(G, v \in V_1'', u \in V_2'') = r - 1 - d_1(v) + \deg u - d_1(u).$$

Finally, let's look at two cases.

1. Let $\deg u < \deg v$ in the graph G. Then, by virtue of the two equalities obtained, we have

$$w(G, u \in V_1', v \in V_2') > w(G, v \in V_1'', u \in V_2''),$$

   i. e., the contribution of vertices u and v decreased when moving from the 2-decomposition (V_1', V_2') to the 2-decomposition (V_1'', V_2''). This implies $w(G, V_1'', V_2'') < w(G, V_1', V_2')$.
2. Let $\deg u = \deg v$ in graph G. Then we have

$$w(G, u \in V_1', v \in V_2') = w(G, v \in V_1'', u \in V_2'').$$

   This implies $w(G, V_1'', V_2'') = w(G, V_1', V_2')$. □

Lemma 4 has a simple meaning: if you exchange a vertex of a lower degree from V_1' in a 2-decomposition (V_1', V_2') with a vertex of a higher degree from another component of this 2-decomposition, then the weight will decrease when moving from a 2-decomposition (V_1', V_2') to a new 2-decomposition (V_1'', V_2'').

Lemma 4 implies

Corollary 2. 1. *Non-negative integers $w(G, V_1, V_2)$ are the same for all principal 2-decompositions (V_1, V_2) of the graph G.*

2. *Non-negative integer $w(G, V_1, V_2)$ for principal 2-decomposition (V_1, V_2) of the graph G is less than the same form integer for any non-principal 2-decomposition of rank r.*

P r o o f. It is enough to note that principal 2-decompositions can differ only in the location of vertices of degree r in their components. In addition, any non-principal 2-decomposition of rank r comes to a principal 2-decomposition of rank r using a certain sequence of operations of exchanging vertices. □

Lemma 5. *Let (V_1, V_2) be an arbitrary 2-decomposition of rank q of the set V of vertices of a graph G, where $1 \le q < n$ and n is the cardinality of V. Then any rotation of an edge in the graph G can change the weight $w(G, V_1, V_2)$ of the 2-decomposition (V_1, V_2) by no more than 1 when moving to a new graph.*

P r o o f. Let the rotation of the edge $e = xv$ correspond to a triple (x, v, y). Vertices v and y are different and not adjacent. The old edge $e = xv$ and the new edge $f = vy$ cannot lie in different sets V_1 and V_2, since they have a common vertex v incident to them. This obviously implies the statement of the lemma. □

P r o o f of Theorem 1. Let (V_1, V_2) be a principal 2-decomposition of a graph G. Let $w_1 = w_1(G, V_1, V_2)$ be the number of all pairs of distinct non-adjacent vertices from V_1, $w_2 = w_2(G, V_1, V_2)$ be the number of all pairs of adjacent vertices from V_2. Then

$$w = w(G, V_1, V_2) = w_1(G, V_1, V_2) + w_2(G, V_1, V_2) = w_1 + w_2.$$

Algorithm 1 reduce the graph G to a splittable r-ancestor of the form $H' = (K(V_1), E', V_2)$ by using w lifting rotations of edges.

By Lemmas 2 and 3 we have

$$\operatorname{sum}\operatorname{hd}(\lambda) + (w_1 + w_2) = \operatorname{sum}\operatorname{hd}(\mu),$$
$$\operatorname{sum}\operatorname{tl}(\lambda) - (w_1 + w_2) = \operatorname{sum}\operatorname{tl}(\mu),$$

where $\mu = \operatorname{dpt}(H')$.

Since the splittable graph H' satisfies the conditions of Lemma 1, we obtain

$$\operatorname{sum}\operatorname{hd}(\mu) = \operatorname{sum}\operatorname{tl}(\mu),$$

which implies

$$\operatorname{sum}\operatorname{hd}(\lambda) + (w_1 + w_2) = \operatorname{sum}\operatorname{tl}(\lambda) - (w_1 + w_2).$$

Therefore,

$$2(w_1 + w_2) = \operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda) = 2s,$$

i. e., $s = w_1 + w_2 = w$.

Let (V_1', V_2') be an arbitrary 2-decomposition of the graph G. Suppose that the graph G is reduced to a splittable graph $H' = (K(V_1'), E_1, V_2')$ by t lifting rotations of edges, where V_1' is a clique of cardinality r and V_2' is a coclique. Then obviously $w(H', V_1', V_2') = 0$.

As t lifting rotations of the edges change the weight of 2-decomposition (V_1', V_2') from $w(G, V_1', V_2')$ to 0, by Lemma 5 the following holds:

$$t \geq w(G, V_1', V_2').$$

Let's look at two cases.

1 case. If (V_1', V_2') is the principal 2-decomposition of the graph G, then the resulting inequality, due to the fact that $w = s$, gives $t \geq s$ and the proof of statement 1) of the theorem is completed.

2 case. Let (V_1', V_2') be a non-principal 2-decomposition of the graph G. Then, taking into account Corollary 2, we obtain

$$t \geq w(G, V_1', V_2') > w(G, V_1, V_2) = s,$$

where (V_1, V_2) is an arbitrary of the principal 2-decompositions of the set V of vertices of the graph G. The proof of statement 2) is also completed. □

Lemma 6. *Any closest splittable r-ancestor of a graph G can be obtained by some application of Algorithm 1.*

P r o o f. Let $H' = (K(V_1'), E', V_2')$ be some closest splittable r-ancestor of the graph G, i. e., it can be obtained from the graph G using a sequence of s lifting rotations of edges.

Then, by virtue of what was established in the proof of the theorem, the 2-decomposition (V_1', V_2') is the principal 2-decomposition of the graph G (here $t = s$). It is clear that in a sequence of s lifting rotations of edges transforming G to H', each lifting rotation must decrease the weight of the 2-decomposition (V_1', V_2') by exactly 1, i. e., it must be performed in accordance with step 2 of Algorithm 1. □

Algorithm 1 we will call *the algorithm for reducing a graph G to a closest splittable r-ancestor.* Of course, different implementations of this algorithm may produce different closest splittable r-ancestors of the original graph G (see, for example, Fig. 3).

3. Conclusion

In conclusion, we note that in connection with Corollary 2 the following two problems are of interest.

Firstly, we give a necessary definition. Let μ and λ be graphical partitions of the same weight $2m$ such that μ dominates λ. Let $\operatorname{height}(\mu, \lambda)$ denote the height of the partition μ over the partition λ in the lattice of all partitions of weight $2m$, which is equal to the length of the shortest sequence of elementary transformations of the first type transforming μ into λ (see [2]).

Problem 1. *Let λ be a graphical partition of rank r. Find all graphical partitions μ of rank r that dominate partition λ such that* $\operatorname{sum}\mu = \operatorname{sum}\lambda$*,*

$$\operatorname{sum}\operatorname{hd}(\mu) = \operatorname{sum}\operatorname{tl}(\mu) \quad \textit{and} \quad \operatorname{height}(\mu, \lambda) = \frac{1}{2}(\operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda)).$$

Note that the condition $\operatorname{sum}\operatorname{hd}(\mu) = \operatorname{sum}\operatorname{tl}(\mu)$ means that any realization of the partition μ is a splittable graph (see, for example, [2]). The condition

$$\operatorname{height}(\mu, \lambda) = \frac{1}{2}(\operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda))$$

means that using some s lowering rotations of edges for any realization of the partition μ leads to a realization of the partition λ, where

$$s = \frac{1}{2}(\operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda)).$$

Problem 2. *Let λ be a graphical partition of rank r.*

1. *For a given graph G of Durfey rank r, find the family $CSrA(G)$ of all its closest splittable r-ancestors.*
2. *Find the family $CSrA(\lambda)$ of all splittable graphs, each of which is the closest splittable r-ancestor for some realization of the partition λ.*
3. *Find a family of closest splittable r-ancestors of some realizations of the partition λ such that*
    - *every realization of the partition λ can be obtain (up to isomorphism and isolated vertices) from element of this family by sequentially applying s lowering rotations of edges where $s = (1/2)(\operatorname{sum}\operatorname{tl}(\lambda) - \operatorname{sum}\operatorname{hd}(\lambda))$,*
    - *this family has the smallest possible number of elements.*

The work [2] gives an example of a partition $\lambda = (4, 3, 2, 2, 2, 1)$, for which $r = 2$ and $s = 1$ such that each of its realizations can be obtained from a common splittable 2-ancestor using a single lowering rotation of an edge.

In conclusion, let us give another example that shows that one splittable r-ancestor may not be sufficient to obtain all realizations of a given partition λ of rank r by sequentially applying s lowering rotations of edges.

Example 2. Let $\lambda = (3, 3, 2, 2, 1, 1)$. Then $r = 2$, $\operatorname{hd}(\lambda) = (2, 2)$, $\operatorname{tl}(\lambda) = (4, 2)$, $s = 1$.

It is easy to check that the partition λ has 5 pairwise non-isomorphic realizations G_1, G_2, G_3, G_4, G_5 without isolated vertices and these realizations have exactly 2 non-isomorphic closest splittable 2-ancestors H_1' and H_2' (see Fig. 8 and Fig. 9). Here V_1 consists of two vertices of the highest degree, and V_2 consists of four remaining vertices (note that for the graph H_1' in V_2 there is one vertex of zero degree).

In G_2 we have $t_1 = 1$ and $t_2 = 0$, and in G_1, G_3, G_4, G_5 we have $t_1 = 0$ and $t_2 = 1$.

It is easy to check that with respect to the principal 2-decomposition (V_1, V_2)

- graph G_1 has H_1' as exactly one closest splittable 2-ancestor;
- graphs G_2 and G_3 have exactly 2 closest splittable 2-ancestors H_1' and H_2';
- graphs G_4 and G_5 have H_2' as exactly one closest splittable 2-ancestor.

Note also that graph H_1' can be obtained from graph H_2' using a single lifting rotation of an edge, and $\deg H_2' = (4, 3, 2, 1, 1, 1)$ can be obtained from $\deg H_1' = (4, 3, 2, 2, 1)$ using one elementary transformation of the first type.

Note that the graph H_1' is a threshold graph [6], since the partition $(4, 3, 2, 2, 1)$ has the same tail and head [2], and the graph H_2' is not a threshold graph, since its degree partition $(4, 3, 2, 1, 1, 1)$ is not a maximum graphical partition.

It is clear that to obtain all realizations of the partition $\lambda = (3, 3, 2, 2, 1, 1)$ by applying a single lowering rotation of an edge, we need to use both graphs H_1' and H_2'. Here H_1' and H_2' are not common closest splittable 2-ancestors of graphs G_1 and G_5.

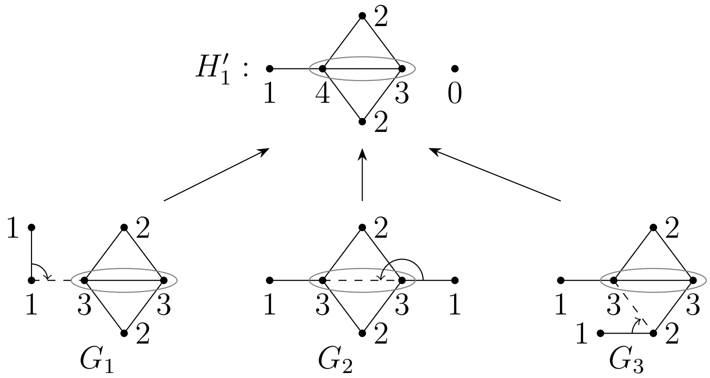


Figure 8. The common closest splittable 2-ancestor of graphs G_1, G_2 and G_3.

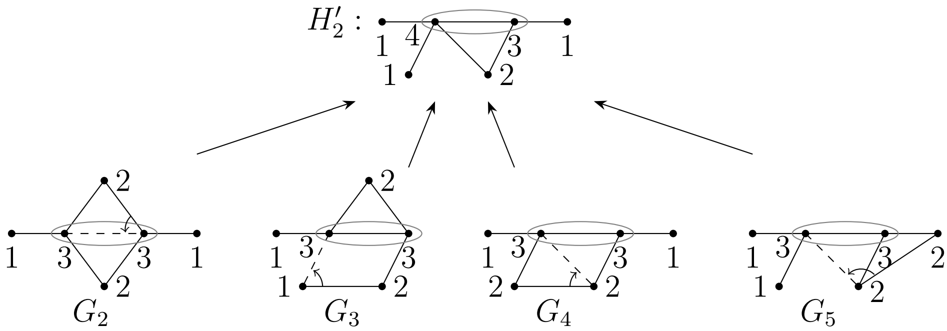


Figure 9. The common closest splittable 2-ancestor of graphs G_2, G_3, G_4 and G_5.

Acknowledgements

The authors thank the referees for giving several valuable suggestions.

REFERENCES

1. Andrews G. E. *The Theory of Partitions.* Cambridge: Cambridge Univ. Press, 1976. 255 p. DOI: 10.1017/CBO9780511608650
2. Baransky V. A., Senchonok T. A. Around the Erdös–Gallai criterion. *Ural Math. J.,* 2023. Vol. 9, No. 1. P. 29–48. DOI: 10.15826/umj.2023.1.003
3. Foldes S., Hammer P. L. Split graphs. In: *Proc. of the 8th South-Eastern Conf. on Comb.: Graph Theory and Computing*, 1977. P. 311–315.
4. Fulkerson D. R., Hoffman A. J., McAndrew M. H. Some properties of graphs with multiple edges. *Canad. J. Math.*, 1965. Vol. 17. P. 166–177. DOI: 10.4153/CJM-1965-016-2
5. Hammer P. L., Simeone B. The splittance of a graph. *Combinatorica,* 1981. Vol. 1. P. 275–284. DOI: 10.1007/BF02579333
6. Mahadev N. V. R., Peled U. N. *Threshold Graphs and Related Topics.* Ser. Annals of Discr. Math., vol. 56. Amsterdam: North-Holland Publ. Co., 1995. 542 p.
7. Merris R. *Graph Theory.* New York: J. Wiley & Sons Inc., 2001. 256 p. DOI: 10.1002/9781118033043
8. Merris R. Split graphs. *European J. Combin.*, 2003. Vol. 24. P. 413–430. DOI: 10.1016/S0195-6698(03)00030-1
9. Tyshkevich R. I. Decomposition of graphical sequences and unigraphs. *Discrete Math.,* 2000. Vol. 220. P. 201–238. DOI: 10.1016/S0012-365X(99)00381-7

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 37–48

DOI: 10.15826/umj.2024.2.004

COMPLETELY REACHABLE ALMOST GROUP AUTOMATA[1]

David Fernando Casas Torres

Ural Federal University,
51 Lenina Str., Ekaterimburg, 620000, Russian Federation

dkasastorres@urfu.ru

Abstract: We consider finite deterministic automata such that their alphabets consist of exactly one letter of defect 1 and a set of permutations of the state set. We study under which conditions such an automaton is completely reachable. We focus our attention on the case when the set of permutations generates a transitive imprimitive group.

Keywords: Deterministic finite automata, Transition monoid, Complete reachability, Permutation group.

1. Introduction

A deterministic finite automaton is said to be *completely reachable* if every non-empty subset of states is the image of the whole state set by the action of some word. Such automata appeared in the study of descriptional complexity of formal languages [11] and in relation to the Černý conjecture [8]. A systematic study of completely reachable automata was initiated in [5] and [6], and continued in [4]; in these papers Bondar and Volkov developed a characterization of completely reachable automata that relied on the construction of a series of digraphs.

One of the main results by Ferens and Szykuła [9] was an algorithm of polynomial time complexity, with respect to the number of states and letters, to decide whether a given automaton was completely reachable. This seemed to solve the complexity problem for this kind of automata. A different approach was proposed by Volkov and the author [7] for the special case of automata with two letters. There the characterization relied on whether one of the letters acted as a complete cyclic permutation of the states and how the other letter acted on certain subsets of states.

In this paper we give an approach to the generalization of the result in [7] by allowing that all the letters except one act as permutations of the set state and studying how the additional non-permutation letter acts on non-trivial blocks of imprimitivity if there are any.

The study of automata where all letters but one are permutations is by no means new. This kind of automata is presented with different names. In [1], they are called *almost-permutation automata* and are used to present an example of a series of slowly synchronzing automata with a sink state. In [2], automata are under the disguise of transformation semigroups and are called *near permutation*. In [3], they are called *almost-group automata*; there it is proved that these automata synchronize with high probability. Finally, in [12], the non-permutation letters are the identity except in a subset of states where they have the same image. There, it is proved that if no equivalence relation is preserved under the action of the letters, then the automaton is synchronizing. Among these papers, we would like to highlight the work done in [10] where the primitivity of a group of permutations of a state set has been tightly related to the complete reachability of the automata

[1]This work was supported by the Ministry of Science and Higher Education of the Russian Federation (project FEUZ-2023-0022).

generated by adding a non permutation letter. Thus, the result presented here approaches this theory from the other side where the group is transitive but not primitive and we suggest a condition to ensure that the automaton generated is completely reachable.

In Section 2 we present the definitions and notation used in this paper. Then in Section 3 we present and prove a necessary condition for almost group automata to be completely reachable. In Section 4 we describe a set of directed graphs useful in the discussion of complete reachability and prove a partial sufficient condition related with the one discussed in the Section 3.

2. Preliminaries

A *deterministic finite automaton*, or simply an automaton, is usually defined as a triple

$$\mathcal{A} = \langle Q, \Sigma, \delta \rangle,$$

where Q, the states, and Σ, the alphabet, are finite sets and $\delta : Q \times \Sigma \to Q$ is the transition function. For each letter in alphabet of the automaton $a \in \Sigma$ we can define the function $\delta_a : Q \to Q$ where $\delta_a(q) = \delta(q, a)$, hence each letter can be considered individually as a function of Q on itself or a *transformation* over Q. This observation makes reasonable for us to use the following notation: for every $q \in Q$ and $a \in \Sigma$ we will denote $\delta(q, a) := q \cdot a$. Derived from this we can say that an automaton can be specified using just its set of states and the action of each letter in this set; that is why from now on we will define automata as pairs of the state set and the alphabet.

A *word* of an automaton is a finite sequence of letters over its alphabet; this includes the empty word. The set of all words over the alphabet Σ is denoted by Σ^*. We can extend the action of letters to words recursively in the following way: if $w \in \Sigma^*$, $a \in \Sigma$ and $q \in Q$, then $q \cdot wa := (q \cdot w) \cdot a$, and the action of the empty word is the identity function. Furthermore, the action of words can be applied to subsets of states: if $P \subseteq Q$ and $w \in \Sigma^*$, then

$$P \cdot w := \{p \cdot w \mid \text{ for every } p \in P\}.$$

A subset of states $P \subseteq Q$ is called *reachable* if there is a word $w \in \Sigma^*$ such that its image is exactly P, that is, $Q \cdot w = P$. An automaton is said to be *synchronizing* if at least one singleton is reachable, i.e., there is a state $q \in Q$ and a word $w \in \Sigma^*$ such that $Q \cdot w = \{q\}$. An automaton is *completely reachable* if every non-empty subset of states is reachable.

Let $\mathcal{A} = \langle Q, \Sigma \rangle$ be an automaton and $w \in \Sigma^*$ an arbitrary word. The excluded set of w denoted by $\mathrm{excl}(w)$ is the set of states that have no preimages by w. The *defect* of w is the size of its excluded set. In the case the defect is 0, the word w represents a permutation of the set of states. Since a word is a total function, if the defect of w is bigger than 0, then there must be states with the same image. These images are the *duplicated* states of the word; the set of these states will be denoted by $\mathrm{dupl}(w)$. When any of these sets, $\mathrm{excl}()$ or $\mathrm{dupl}()$, is a singleton we will make no distinction between the set and the state inside it. Additionally, for the case of words of defect 1 we know that exactly two states must have the same image; we will call this pair of states the *collapsed set* of the word, denoted by $\mathrm{coll}()$.

Some transformations over a set of states Q can be bijective, thus permutations. Thanks to this we can use some terminology of the theory of permutation groups. Recall that the set of all the bijective transformations of a finite set Q on itself is denoted by S_Q, also called the *symmetric group*. Let $G \subseteq S_Q$ be a group of permutations of Q. This group is said to be *transitive* if for every pair of states $q, p \in Q$ there is a permutation $\sigma \in G$ such that $p \cdot \sigma = q$. Our main subject of study are automata and their words, thus, except when specified, when we talk about arbitrary permutations of the set of states we assume that there is allways a word that produces it. This is, if we mention the permutation $g \in S_Q$, we assume that there is a word w such that $q \cdot w = q \cdot g$ for every $q \in Q$.

A non-empty subset $B \subseteq Q$ is said to be a *block* of the group if and only if for every $\sigma \in G$ either $B \cdot \sigma = B$ or $B \cdot \sigma \cap B = \emptyset$. The singletons and Q itself are, always, blocks, these are called *trivial*. A permutation group $G \subseteq S_Q$ is said to be *primitive* if it is transitive and the only blocks are the trivial ones; otherwise the group is said to be *imprimitive*. In this article when we talk about a block of imprimitivity, unless stated the contrary, it will always be non-trivial. If a transitive group $G \subset S_Q$ has a block of imprimitivity $B \subseteq Q$, all the images of B by elements of G are also blocks of imprimitivity and form a partition of Q. This partition of subsets is called a *system of imprimitivity* of the group G over Q.

A *directed graph* Γ is a pair $(V(\Gamma), E(\Gamma))$, where $V(\Gamma)$ is the vertex set and $E(\Gamma) \subseteq V(\Gamma) \times V(\Gamma)$ is the set of directed edges. To each edge we can assign one or more labels from some set L. We will denote an edge $(s,t) \in E(\Gamma)$ labelled with w as $s \xrightarrow{w} t$. Since in this paper we consider only directed graphs, from now we omit the word "directed". For reference, the first and second components of an edge will be called the *source* and *tail* respectively. A *path* of a graph is a set of edges $e_1, e_2, \ldots, e_m$, $m \geq 1$, such that for every $1 \leq i < m$, the tail of e_i is the same as the source of e_{i+1}. Vertices $p, q \in V(\Gamma)$ are *strongly connected* if there is a path from p to q and from q to p. Also we consider each vertex as strongly connected with itself. A strongly connected component of a graph is a maximal subset of vertices such that all its vertices are strongly connected to each other.

An automaton can be represented as a labelled graph, where the vertex set is the states set and for each state p and letter a, there is a labelled edge $p \xrightarrow{a} p \cdot a$. This is the *underlying* graph of the automaton.

As we mentioned in the introduction, in [10] the following characterization of primitive permutation groups is given. Here $[n] := \{1, 2 \ldots, n\}$ and if S is a set of transformations, then $\langle S \rangle$ is the transformation semigroup generated by S. Recall that a transformation f is *idempotent* if $f^2 = f$.

Theorem 1 [10, Theorem 3.1]. *Let G be a permutation group on $[n]$ with $n \geq 3$. Then G is primitive if and only if for each [idempotent] transformation $f : [n] \to [n]$ of* defect 1 *every non-empty subset $A \subseteq [n]$ is reachable in $\langle G \cup \{f\} \rangle$.*

This theorem presents a characterization of primitive groups. In the language of automata it states that in the presence of a set of permutation letters that generates a primitive group, the addition of any transformation of defect 1 suffices to obtain a completely reachable automaton. Here we study a related case. We would like to know what happens when the group generated by the permutation letters is transitive but not primitive. We will see how this situation is not that forgiving and it requires a more complex relation between the group generated by the permutations and the transformation of defect 1. The results proved in this article are closely related to the ones presented in [7] for automata with just one permutation letter and one with defect 1. The work presented in this article and in [7] is based on the work made in [6]; there the main actor is a graph constructed in several steps. We will briefly explain the construction in Section 4.

For the rest of this paper we will consider automata $\mathcal{A} = \langle Q, \Sigma \rangle$, where $\Sigma = \Sigma_0 \cup \{a\}$ and:

- the set of letters $\Sigma_0 \subset S_Q$, are all permutations of Q,
- the generated subgroup $G = \langle \Sigma_0 \rangle$ is transitive,
- the letter a has defect 1.

The excluded state of the only letter of defect 1 will be denoted by e, i.e., $\mathrm{excl}(a) = e$. Unless specified otherwise, the group generated by all permutation letters is denoted as G. We will call automata with these characteristics *almost group* automata.

Let $r \in \mathrm{coll}(a)$ be one of the two states collapsed by a. There is a permutation that sends e to r; call it $\sigma \in G$. The transformation σa has defect 1, $e = \mathrm{excl}(\sigma a)$, and $e \in \mathrm{coll}(\sigma a)$. Consider the automaton $\overline{\mathcal{A}} = \langle Q, \Sigma_0 \cup \{\sigma a\}\rangle$. Note that $\mathcal{A}$ is completely reachable if and only if $\overline{\mathcal{A}}$ is completely reachable. Therefore, there is no loss of generality when we add the condition that $e \in \mathrm{coll}(a)$ from the beginning. When this happens we call the automaton *standardized*. This change will simplify the arguments we use for the rest of the article.

For any automaton a subset of states, $P \subseteq Q$ is *invariant* by a transformation $w \in \Sigma^*$, or *w-invariant*, if $P \cdot w \subseteq P$. The condition to get complete reachability in the binary case is that no subset of states that *represents* a subgroup of the cyclic group is invariant by letter of defect 1. The cosets of any subgroup generate a partition of the group that contains this subgroup. There is a parallel situation in the case of blocks of imprimitivity, they form a partition of the set of states. There are subgroups for each of these blocks of imprimitivity that let them invariant[2]. This is the main reason to consider systems of imprimitivity. These partitions of the set of states are the closest to represent subgroups of the group acting on the states.

3. A necessary condition

First, we begin by proving that complete reachability implies that some blocks can not be a-invariant.

Proposition 1. *Let $\mathcal{A}$ be a standardized almost group automaton. If $\mathcal{A}$ is completely reachable, then G is transitive and if there is at least one block of imprimitivity then no block of imprimitivity that contains $e = \mathrm{excl}(a)$ is invariant by a.*

P r o o f. First, let us prove that the condition for the group generated by the set of permutations to be transitive is necessary. For every word $w \in \Sigma^*$ it is true that $e \in \mathrm{excl}(wa)$; furthermore,

$$|\mathrm{excl}(w)| \leq |\mathrm{excl}(wa)| \leq |\mathrm{excl}(w)| + 1.$$

This is, the action of adding a to a word either increases by one or keeps the defect of the resulting word. Note that adding a permutation does not modify the defect of any word. Hence, in order to reach the subsets $Q \setminus \{q\}$ for every $q \in Q$, it is necessary that there exists a permutation $\sigma_q \in G$ such that $e \cdot \sigma_q = q$. Let $p, q \in Q$ be an arbitrary pair of states. By the previously said, if $\mathcal{A}$ is completely reachable, then there are two permutations $\sigma_p, \sigma_q \in G$ such that $e \cdot \sigma_p = p$ and $e \cdot \sigma_q = q$. Finally, note that $p \cdot \sigma_p^{-1}\sigma_q = q$. Thus, G is transitive.

For a subset of states $S \subset Q$, we denote by $\overline{S}$ its complement, i.e., $Q \setminus S$.

The proof that no block of imprimitivity is a-invariant will be by contradiction. Suppose that B is a block of imprimitivity that contains e and is a-invariant. This block belongs to a system of imprimitivity. Let $w \in \Sigma^*$ be the shortest word that reaches the complement of a block of this system, say $\overline{C}$. If $w = w'\,b$ with $b \in \Sigma_0$, then $Q \cdot w' = \overline{C} \cdot b^{-1}$, the complement of a block of imprimitivity. This contradicts the condition of w being the shortest word. Hence, w can not end in a permutation.

As a consequence, the word w ends with the letter a, i.e., $w = w'\,a$. Recall that $Q \cdot w'a \subset Q \cdot a$ and $e \notin Q \cdot a$, thus $e \notin Q \cdot w$ and we conclude that $Q \cdot w = \overline{B}$.

Since B is a-invariant, we can conclude that its complement is also a-invariant. And since every $q \in \overline{B}$ has a preimage by a then this letter acts as a permutation of $\overline{B}$. Therefore $Q \cdot w' = Q \cdot w'a = \overline{B}$ what, again, contradicts the supposition of w being the shortest. There is no other type of letter in which the word w could finish, then we end with an absurd. This situation came from supposing that B is a-invariant, thus we have our proposition. $\square$

By the preceding proof we have:

[2]Considerations of this are treated ahead in the paper.

Corollary 1. *If there is a block of imprimitivity that contains e and is invariant by a, then its complement is not reachable.*

4. Rystsov graphs of almost group automata

4.1. The structure of Rystsov graphs

In [6] and [4] Bondar and Volkov presented a characterization of completely reachable automata. The characterization relies on the construction of a graph that can be constructed from an arbitrary automaton. This graph is a generalization of the ideas presented by Igor Rystsov in [13]. That is why we will call these graphs as *Rystsov graphs* of the automata.

The Rystsov graph of an automaton $\mathcal{A}$, denoted by $\Gamma(\mathcal{A})$, is constructed in an inductive way. This means that in order to construct $\Gamma(\mathcal{A})$ we first assemble a graph called $\Gamma_1(\mathcal{A})$, verify if we can continue and if that is the case from $\Gamma_1(\mathcal{A})$ we continue with the construction of $\Gamma_2(\mathcal{A})$ and so on. This series of graphs is guaranteed to always finish, the final graph is the Rystsov graph of $\mathcal{A}$. For the construction of the partial graphs we make use of the sets of words $W_k(\mathcal{A}) \subset \Sigma^*$, defined as the subset of all words of defect k for $k \geq 1$.

The first step is to construct the graph $\Gamma_1(\mathcal{A})$ where its vertex set is $Q_1 := Q$ and its edge set is defined as:

$$E_1(\mathcal{A}) := \{\mathrm{excl}(w) \xrightarrow{w} \mathrm{dupl}(w) \in Q_1 \times Q_1 \mid w \in W_1(\mathcal{A})\}.$$

Example 1. Consider the automaton $\mathcal{E}_{18} := \langle\{1, 2, \ldots, 18\}, \{a, b, c\}\rangle$, where b and c are permutations with the following cyclic representation:

$$\begin{aligned} b :=&(1, 11, 13, 5, 7, 17)(2, 10, 14, 4, 8, 16)(3, 12, 15, 6, 9, 18),\\ c :=&(1, 3, 2)(4, 5, 6)(7, 13)(8, 16)(9, 15)(10, 14)(11, 17)(12, 18),\end{aligned}$$

and the transformation a has defect 1. The following representation of a puts the respective image under each state and omits the states that do not change:

$$\begin{pmatrix} 1\,2\,5\,6\,8 \\ 6\,8\,6\,5\,2 \end{pmatrix}.$$

Note that the $\mathrm{excl}(a) = 1$, $\mathrm{dupl}(a) = 6$ and $\mathrm{coll}(a) = \{1, 5\}$. The group generated by $\{b, c\}$ is transitive and the blocks of imprimitivity that contain the state 1 are the sets

$$\{1, 5\}, \quad \{1, 2, 3, 4, 5, 6\}.$$

We can see $1 \xrightarrow{a} 6$, $2 \xrightarrow{ac^2} 5$ and $1 \xrightarrow{ab^3a} 3$ are edges in $E_1(\mathcal{E}_{18})$.

We continue the definitions. For any automaton $\mathcal{A}$ consider the following subset of states:

$$D_1(\mathcal{A}) := \{p \in Q_1 \mid p = \mathrm{dupl}(w) \quad \& \quad e = \mathrm{excl}(w) \text{ for } w \in W_1(\mathcal{A})\}.$$

These are the states directly connected to e in $\Gamma_1(\mathcal{A})$, that is, tails of the edges with e as source. The following lemma states that all the edges in $\Gamma_1(\mathcal{A})$ are images by G of these initial edges.

Lemma 1. *If $q \to p \in E(\Gamma_1(\mathcal{A}))$, then there are $\sigma_q \in G$ and $d \in D_1(\mathcal{A})$ such that $e \cdot \sigma_q = q$ and $d \cdot \sigma_q = p$. Or, what is equivalent, there is a permutation $\sigma_q \in G$ such that $p \cdot \sigma_q^{-1} \in D_1(\mathcal{A})$.*

P r o o f. If $q \to p$ is an edge of $\Gamma_1(\mathcal{A})$, then there is a word of defect 1, call it $w \in W_1(\mathcal{A})$ such that $\mathrm{excl}(w) = q$ and $\mathrm{dupl}(w) = p$; this happens due to the definition of $\Gamma_1(\mathcal{A})$. Remember that G is transitive, thus there is a permutation $\sigma_q \in G$ such that $e \cdot \sigma_q = q$.

The word $w\sigma_q^{-1}$ has defect 1 and $\mathrm{excl}(w\sigma_q^{-1}) = \mathrm{excl}(w) \cdot \sigma_q^{-1}$, at the same time $\mathrm{dupl}(w\sigma_q^{-1}) = \mathrm{dupl}(w) \cdot \sigma_q^{-1}$. Thus $\mathrm{excl}(w\sigma_q^{-1}) = q \cdot \sigma_q^{-1} = e$ and $p \cdot \sigma_q^{-1} \in D_1(\mathcal{A})$. ☐

This lemma also tell us that in order to compute $\Gamma_1(\mathcal{A})$ it is sufficient to calculate $D_1(\mathcal{A})$, and then apply to the generated edges permutations that send e to each of the different states of the automaton. In our running example the initial edges of $\Gamma_1(\mathcal{E}_{18})$ are shown in Figure 1. The

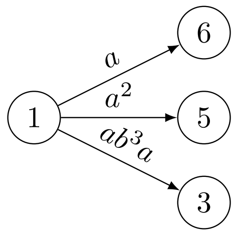


Figure 1. The initial edges of $\Gamma_1(\mathcal{E}_{18})$.

strongly connected component that contains 1 is shown in Figure 2 (we omitted the labels to avoid confusion).

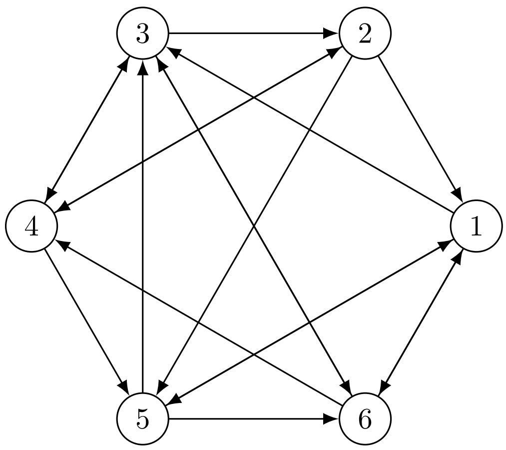


Figure 2. A strongly connected component of $\Gamma_1(\mathcal{E}_{18})$.

Now, let ${C_e}^{[1]} \subseteq Q_1$ be the vertex set of the strongly connected component of $\Gamma_1(\mathcal{A})$ that contains e.

Lemma 2. *The set $C_e^{[1]}$ is a block of imprimitivity.*

P r o o f. If $\Gamma_1(\mathcal{A})$ is strongly connected then $C_e^{[1]} = Q$ and the proposition is true. Then let us assume $\Gamma_1(\mathcal{A})$ is not strongly connected and $C_e^{[1]}$ is a proper subset of Q.

Let $\sigma \in G$ be a permutation such that $e \cdot \sigma = \mathrm{dupl}(a) = d$, then the edge

$$d \xrightarrow{a\sigma} d \cdot \sigma \in E_1.$$

If we repeat the application of σ, the produced words have all defect 1, and there is an $i \geq 1$ such that $d \cdot \sigma^i = e$. Then, $C_e^{[1]}$ is not a singleton since at least $e, d \in C_e^{[1]}$.

Considering this we will prove first that for any $\rho \in G$, the subset ${C_e}^{[1]} \cdot \rho$ is also a strongly connected component.

Let $p, q \in {C_e}^{[1]}$ be two arbitrary states. In $\Gamma_1(\mathcal{A})$ there is a path:

$$p \xrightarrow{w_1} t_1 \xrightarrow{w_2} t_2 \ldots t_{k-1} \xrightarrow{w_k} q$$

where every w_i is a word of defect 1. Since permutations act well on excluded and duplicated states, then:

$$p \cdot \rho \xrightarrow{w_1\rho} t_1 \cdot \rho \xrightarrow{w_2\rho} t_2 \cdot \rho \ldots t_{k-1} \cdot \rho \xrightarrow{w_k\rho} q \cdot \rho$$

is a path in ${C_e}^{[1]} \cdot \rho$; in the same way we can prove the existence of a path connecting $q \cdot \rho$ with $p \cdot \rho$, making ${C_e}^{[1]} \cdot \rho$ a strongly connected component.

What is left is to prove that ${C_e}^{[1]}$ and its images by permutations of G are blocks of imprimitivity. Let $\rho \in G$ be an arbitrary permutation. Suppose $x \in {C_e}^{[1]} \cap {C_e}^{[1]} \cdot \rho$ and let $y \in {C_e}^{[1]}$ and $z \in {C_e}^{[1]} \cdot \rho$ be two different states. By the definition of strongly connected component there are paths from y to x, from x to z and, going back, from z to x, and from x to y. Then

$${C_e}^{[1]} = {C_e}^{[1]} \cdot \rho.$$

This makes ${C_e}^{[1]}$ a block of imprimitivity. □

We continue the inductive construction of the graph $\Gamma(\mathcal{A})$. Once we get $\Gamma_k(\mathcal{A})$, $k \geq 1$, if one of the following alternatives happens then the construction will be stopped and $\Gamma(\mathcal{A}) := \Gamma_k(\mathcal{A})$: either the graph is strongly connected, or all the strongly connected components are not *big enough* (we will address the meaning of this in a moment). If none of these two possibilities happen, then we proceed to construct $\Gamma_{k+1}(\mathcal{A})$. The new vertex set Q_{k+1} will consists of the strongly connected components of $\Gamma_k(\mathcal{A})$; thus, each vertex is a collection of vertices of the set Q_k.

In order to define the edges of this new graph, we need to properly define when a strongly connected component is *big enough*. For this note that each vertex of $\Gamma_2(\mathcal{A})$ is a subset of states (even considering singletons) and the vertices of $\Gamma_3(\mathcal{A})$ would be collections of subsets of states and so on. With this in mind for $k \geq 2$ let $V \in Q_k$ be a vertex of $\Gamma_k(\mathcal{A})$, define the *foliage* of V, and denote it by $\mathrm{leaf}(V)$, as follows: for $V \in Q_2$, its foliage is the set itself, i.e., $\mathrm{leaf}(V) := V$, and for $k > 2$,

$$\mathrm{leaf}(V) := \bigcup_{x \in V} \mathrm{leaf}(x).$$

At the end $\mathrm{leaf}(V)$ is a subset of states. A vertex V of $\Gamma_{k+1}(\mathcal{A})$, or, what is the same, a strongly connected component of $\Gamma_k(\mathcal{A})$, is *big enough* if $|\mathrm{leaf}(V)| \geq k+1$. Thus, we stop the construction if none of the would be vertices of $\Gamma_{k+1}(\mathcal{A})$ have more of $k+1$ states in their foliages. (The term "foliage" is borrowed from [4], where the definition of the vertex sets of the graphs Γ takes form of a rooted tree.) Suppose that this is not the case and we can continue the process, then we can define a new set of edges:

$$E_{k+1} := \{C \xrightarrow{w} D \in Q_{k+1} \times Q_{k+1} \mid C \neq D, \text{ there is a } w \in W_{k+1}(\mathcal{A}),$$
$$\mathrm{excl}(w) \subseteq \mathrm{leaf}(C),\ \mathrm{dupl}(w) \cap \mathrm{leaf}(D) \neq \emptyset\}.$$

The edge set of $\Gamma_{k+1}(\mathcal{A})$ will be the edges of $\Gamma_k(\mathcal{A})$ that connect different vertices of Q_{k+1} together with the set E_{k+1}. For a more detailed discussion of the construction of $\Gamma(\mathcal{A})$ we recommend the reader [4, Section 3].

We have the following theorem that characterizes completely reachable automata:

Theorem 2. *[6] If an automaton $\mathcal{A} = \langle Q, \Sigma \rangle$ is such that the graph $\Gamma(\mathcal{A})$ is strongly connected and $\Gamma(\mathcal{A}) = \Gamma_k(\mathcal{A})$, then $\mathcal{A}$ is completely reachable; more precisely, for every non-empty subset $P \subseteq Q$, there is a product w of words of defect at most k such that $P = Q \cdot w$.*

In the case that the group G is primitive over Q, from Lemma 2 we can see that $\Gamma_1(\mathcal{A})$ will be strongly connected and by Theorem 2 it immediately follows that $\mathcal{A}$ is completely reachable. That is why, from now on the group G will, besides being transitive, have at least a block of imprimitivity.

Example 2. Recall the automaton $\mathcal{E}_{18}$ presented in Example 1. We have seen that $C_e^{[1]} = \{1, 2, 3, 4, 5, 6\}$, and the other strongly connected components are the sets $B_2 = \{7, 8, 9, 10, 11, 12\}$ and $B_3 = \{13, 14, 15, 16, 17, 18\}$. Since these sets have more than two elements, we can continue the construction of $\Gamma(\mathcal{E}_{18})$. Accordingly to the previously said, the vertex set of $\Gamma_2(\mathcal{E}_{18})$ is $Q_2 = \{C_e^{[1]}, B_2, B_3\}$. Consider the word $w := ab^3aca$, note that $\mathrm{excl}(w) = \{1, 3\}$ and $\mathrm{dupl}(w) = \{8, 6\}$, hence the edge $C_e^{[1]} \xrightarrow{w} B_2 \in E_2$. If we add b twice more we have:

$$\begin{aligned} \mathrm{excl}(wb) &= \{11, 12\}, \quad \mathrm{dupl}(wb) = \{9, 16\} \\ \mathrm{excl}(wbb) &= \{13, 15\}, \quad \mathrm{dupl}(wbb) = \{18, 2\}. \end{aligned}$$

Thus adding the edges $B_2 \xrightarrow{wb} B_3$ and $B_3 \xrightarrow{wbb} C_e^{[1]}$ to E_2. These are enough to conclude, thanks to Theorem 2, that $\mathcal{E}_{18}$ is completely reachable.

We will extend the results given by Lemma 1 and Lemma 2. Following the previous notation denote the strongly connected component that contains e in the graph $\Gamma_k(\mathcal{A})$ as $C_e^{[k]}$.

Lemma 3. *If the foliages of the vertices in $\Gamma_k(\mathcal{A})$ form a system of imprimitivity of G over Q, then $Y \to Z \in E_{k+1}$ if and only if there is a permutation $\sigma \in G$ and a set $X \in Q_k$ such that $\mathrm{leaf}(Y) = \mathrm{leaf}({C_e}^{[k]}) \cdot \sigma$; $\mathrm{leaf}(X) \cdot \sigma = \mathrm{leaf}(Z)$ and ${C_e}^{[k]} \to X \in E_{k+1}$.*

P r o o f. Since permutations respect the defect of any word and act well on excluded and duplicated sets, the converse is easy to see.

Now, if $Y \xrightarrow{w} Z \in E_{k+1}$, with $w \in W_{k+1}(\mathcal{A})$, then $\mathrm{excl}(w) \subset \mathrm{leaf}(Y)$ and $\mathrm{dupl}(w) \cap \mathrm{leaf}(Z) \neq \emptyset$. Let $w = u\,a\,\sigma$ with $\sigma \in \Sigma^*$ as a permutation, this is, σ is the longest word generating a permutation after the last appearance of the letter a in w. Since permutations do not increase the defect of a word, then $ua \in W_{k+1}(\mathcal{A})$ and $\mathrm{excl}(w\sigma^{-1}) = \mathrm{excl}(ua)$. From the last affirmation we can conclude that $\mathrm{excl}(ua) \subseteq \mathrm{leaf}(Y) \cdot \sigma^{-1}$.

Since, by hypothesis, $\mathrm{leaf}(Y)$ is a block of imprimitivity then also it is $\mathrm{leaf}(Y) \cdot \sigma^{-1}$. Recall that $e \in \mathrm{excl}(ua)$ thus $\mathrm{excl}(ua) \subseteq {C_e}^{[k]} = \mathrm{leaf}(Y) \cdot \sigma^{-1}$. Using the same argument we can conclude that $\mathrm{leaf}(X) = \mathrm{leaf}(Z) \cdot \sigma^{-1}$. $\square$

Lemma 4. *If the foliages of the vertices in $\Gamma_k(\mathcal{A})$ form a system of imprimitivity, then the foliage of ${C_e}^{[k+1]}$ is a block of imprimitivity of G over Q.*

P r o o f. If each of the foliages of the vertices of $\Gamma_k(\mathcal{A})$ forms a system of imprimitivity, then the foliage of ${C_e}^{[k+1]}$ is just the union of blocks of imprimitivity.
We can use an argument similar to the one used in the proof of Lemma 2 to prove that the image by any $\sigma \in G$ of the foliage of ${C_e}^{[k+1]}$ is also a strongly connected component and a block of imprimitivity. $\square$

Lemma 3 and Lemma 4 form the proof by induction the following result.

Proposition 2. *For any $k \geq 1$, the foliages of the vertices of each $\Gamma_k(\mathcal{A})$ form a system of imprimitivity.*

Note that for any $k \geq 1$ if it happens that $\mathrm{leaf}({C_e}^{[k]}) = Q$ then $\Gamma_k(\mathcal{A})$ is strongly connected and $\mathcal{A}$ is completely reachable. Now we will prove that if this is not the case for any k, then some block of imprimitivity that contains e is invariant by a. We will use the following set:

$$D_k(\mathcal{A}) := \{p \in Q \mid p \in \mathrm{dupl}(w) \text{ for some } w \in \Sigma^* \\ \text{such that } |\mathrm{excl}(w)| \leq k \text{ and } e \in \mathrm{excl}(w) \subseteq \mathrm{leaf}({C_e}^{[k-1]})\}.$$

The set of states duplicated by words of defect less than k such that their excluded set is contained in $C_e^{[k]}$.

4.2. Intermezzo

Before we continue, it is necessary to present some definitions and results related to the theory of permutation groups that are used in the rest of this section. Let Q be a finite set and $G \subseteq S_Q$ be a subgroup of permutations of Q. For any non-empty subset $P \subset Q$ consider the set of permutations:

$$\mathrm{St}_G(P) := \{\sigma \in G \mid P \cdot \sigma = P\},$$

that is, the set of permutations of G that preserve P set-wise. It can be easily proved that $\mathrm{St}_G(P)$ is a subgroup of G. Let us call it the *stabilizer* of the subset P.

Now consider an arbitrary but fixed system of imprimitivity of G over Q, call it $\mathfrak{B}$. The following fact is well known and we will omit the proof.

Proposition 3. *Let G be a group of permutations of a finite set Q. Suppose that G is transitive and $\mathfrak{B}$ is a system of imprimitivity. If $B, C \in \mathfrak{B}$ are two different blocks of imprimitivity then $\mathrm{St}_G(B)$ and $\mathrm{St}_G(C)$ are conjugate subgroups of G.*

For a subgroup H of a group G, the *core* of H, denoted by $\mathrm{Cr}(H)$, is the intersection of all the conjugates of H in G, i.e.,

$$\mathrm{Cr}(H) := \bigcap_{\sigma \in G} \sigma^{-1} H \sigma.$$

Note that this subgroup is normal for G.

Resuming with the transitive group G of permutations of Q, Proposition 3 tells us that for every system of imprimitivity $\mathfrak{B}$ of Q all the stabilizers of the blocks are conjugate. Hence, the following definition makes sense.

Definition 1. *Let G be a subgroup of permutations of Q and $\mathfrak{B}$ be a system of imprimitivity of Q. The core of $\mathfrak{B}$, denoted by $\mathrm{Cr}(\mathfrak{B})$, is the intersection of all the stabilizers of the blocks in $\mathfrak{B}$.*

In some occasions it is more convenient to work with blocks of imprimitivity, hence to talk about the core of a block of imprimitivity. If $\mathfrak{B}$ is a system of imprimitivity and $B \in \mathfrak{B}$ is a block, we denote $\mathrm{Cr}(B) := \mathrm{Cr}(\mathfrak{B})$. For our purposes we look for the core of certain blocks of imprimitivity to act in a transitive way on said blocks. We can ensure this if said core acts transitively on at least one of the blocks.

Proposition 4. *Let G be a group of permutations of the finite set Q. Suppose that G is transitive and $\mathfrak{B}$ is a system of imprimitivity. If $B \in \mathfrak{B}$ is a block and $\mathrm{Cr}(\mathfrak{B})$ acts transitively on B, then this core is also transitive on all the blocks of $\mathfrak{B}$.*

P r o o f. Let $C \in \mathfrak{B}$ be a different block of $\mathfrak{B}$, besides let $p, q \in C$ be two different states. We aim to prove that there is a permutation $\sigma \in \mathrm{Cr}(\mathfrak{B})$ such that $p \cdot \sigma = q$. Being G transitive, there is a permutation $\tau \in G$ such that $C \cdot \tau = B$. Let $r, s \in B$ be such that $p \cdot \tau = r$ and $q \cdot \tau = s$. By hypothesis, there is a permutation $\rho \in \mathrm{Cr}(\mathfrak{B})$ such that $r \cdot \rho = s$. Thus

$$p \cdot \tau\rho\tau^{-1} = q.$$

Since the core is normal in G we can conclude that $\tau\rho\tau^{-1} \in \mathrm{Cr}(\mathfrak{B})$. □

4.3. Non-reachability and invariance

In this part we see that for some almost-group automata not being completely reachable implies there is at least one block of imprimitivity invariant by the letter of defect 1.

Before the main proposition we present a technical lemma. Since in the following lemma k is arbitrary but fixed, $C_e^{[k]}$ will be referred just as C_e.

Lemma 5. *Let $\mathcal{A} = \langle Q, \Sigma_0 \cup \{a\}\rangle$ be an almost-group automaton. If in $\Gamma_k(\mathcal{A})$ there is an edge $C_e \to X$ and $\mathrm{Cr}(\mathrm{leaf}(C_e))$ is transitive for C_e, then for every state $q \in \mathrm{leaf}(X)$, there exists a word v of defect k such that $\mathrm{excl}(v) \subset \mathrm{leaf}(C_e)$ and $q \in \mathrm{dupl}(v)$.*

P r o o f. The edge $C_e \to X$ is produced by a word w such that $\mathrm{excl}(w) \subset \mathrm{leaf}(C_e)$ and $\mathrm{dupl}(w) \cap \mathrm{leaf}(X) \neq \emptyset$. Let $p \in \mathrm{dupl}(w) \cap \mathrm{leaf}(X)$ be arbitrary. Since $\mathrm{Cr}(C_e)$ is transitive, by Proposition 4 there is a permutation $\sigma \in \mathrm{Cr}(C_e)$ such that $p \cdot \sigma = q$. At the same time it is true that $C_e \cdot \sigma = C_e$, since the core is a subset of $\mathrm{St}_G(C_e)$. Therefore we have that $\mathrm{excl}(w\sigma) \subset \mathrm{leaf}(C_e)$ and $q \in \mathrm{dupl}(w\sigma)$. □

Using the Lemma 3 we also can conclude:

Corollary 2. *If in $\Gamma_k(\mathcal{A})$ there is an edge $X \xrightarrow{w} Y$ and $\mathrm{Cr}(\mathrm{leaf}(C_e))$ is transitive for C_e. Then for every state $q \in \mathrm{leaf}(Y)$, there exists a word v of defect k such that $\mathrm{excl}(v) \subseteq \mathrm{leaf}(X)$ and $q \in \mathrm{dupl}(v)$.*

With these two lemmas, we are ready for the main result of this part:

Theorem 3. *Let $\mathcal{A} = \langle Q, \Sigma_0 \cup \{a\}\rangle$ be an almost-group automaton. Suppose $\Gamma(\mathcal{A})$ is not strongly connected. This means for some $k \geq 1$ it happens that $\Gamma(\mathcal{A}) = \Gamma_k(\mathcal{A})$; and ${C_e}^{[k]} = {C_e}^{[j]}$ for every $j \geq k$. Besides this, suppose that for every $\ell \leq k$ the cores $\mathrm{Cr}({C_e}^{[\ell]})$ are transitive on ${C_e}^{[\ell]}$. Then $\mathrm{leaf}({C_e}^{[k]})$ is invariant by a.*

P r o o f. We will use a, structurally, similar proof of the same fact for binary automata presented in [7]. Suppose that ${C_e}^{[k]} = {C_e}^{[k+1]}$. By induction on $0 \leq \ell \leq k$ we will prove that

$$\mathrm{leaf}({C_e}^{[\ell]}) \cdot a \subseteq \mathrm{leaf}({C_e}^{[k]}).$$

For $\ell = 0$ take ${C_e}^{[0]} = \{e\}$ hence the proposition is true in this case.

Our *first induction hypothesis* is that $\mathrm{leaf}({C_e}^{[\ell]}) \cdot a \subseteq \mathrm{leaf}({C_e}^{[k]})$. By the construction of $\Gamma_{\ell+1}(\mathcal{A})$, for any $p \in \mathrm{leaf}({C_e}^{[\ell+1]})$ there is a $X_m \in Q_\ell$ such that $p \in \mathrm{leaf}(X_m)$ and there is a path:

$${C_e}^{[\ell]} \to X_1 \to X_2 \to \cdots \to X_m$$

in $\Gamma_\ell(\mathcal{A})$.

Now, by induction on the length of the path (the number $m > 1$) the idea is to prove that $\mathrm{leaf}(X_m) \cdot a \subseteq \mathrm{leaf}({C_e}^{[k]})$.

If $m = 1$, since there is an edge ${C_e}^{[\ell]} \to X_1$ we use Lemma 5 to ensure that for $p \in \mathrm{leaf}(X_1)$ there is a word $w \in W_\ell(\mathcal{A})$ such that $\mathrm{excl}(w) \subseteq \mathrm{leaf}({C_e}^{[\ell]})$ and $p \in \mathrm{dupl}(w) \cap \mathrm{leaf}(X_1)$. The defect of wa is at most $\ell + 1 \leq k + 1$ and by the *first induction hypothesis* $\mathrm{excl}(wa) \subseteq \mathrm{leaf}({C_e}^{[k]})$ and

$$p \cdot a \in \mathrm{dupl}(wa) \subseteq D_{k+1}(\mathcal{A}) \subseteq {C_e}^{[k+1]} = {C_e}^{[k]},$$

proving what we wanted.

Now suppose that $m > 1$ and $\mathrm{leaf}(X_{m-1}) \cdot a \subseteq \mathrm{leaf}({C_e}^{[k]})$, i.e., the *second induction hypothesis*. By the Corollary 2 for $p \in \mathrm{leaf}(X_m)$ there is a word $w \in W_\ell(\mathcal{A})$ such that $\mathrm{excl}(w) \subseteq \mathrm{leaf}(X_{m-1})$ and $p \in \mathrm{dupl}(w)$. If we apply the same argument as before, but this time using the *second induction hypothesis* we can conclude that

$$p \cdot a \in \mathrm{dupl}(wa) \subseteq D_{k+1}(\mathcal{A}) \subseteq {C_e}^{[k+1]} = {C_e}^{[k]},$$

again, as intended.

Since ${C_e}^{[\ell+1]}$ is a strongly connected component of $\Gamma_\ell(\mathcal{A})$, thus its foliage is the union of the respective foliages of its vertices. We have proved that

$$\mathrm{leaf}({C_e}^{[\ell+1]}) \cdot a \subseteq \mathrm{leaf}({C_e}^{[k+1]}) = \mathrm{leaf}({C_e}^{[k]}).$$

$\square$

The previous theorem proves that for certain almost group automata not being completely reachable is equivalent to having a non-trivial imprimitivity block that is invariant under the letter of defect 1.

5. Conclusion

We considered automata with an alphabet such that there is exactly one letter of defect 1 and the other letters are permutations over the state set. We found a necessary and sufficient condition to decide whether these automata are completely reachable. We saw that if the group generated by the permutations is primitive, then the automaton is completely reachable. On the other case, if the group is transitive and it has non trivial blocks of imprimitivity the condition depends on the behaviour of the letter of defect one over certain blocks of imprimitivity. The author believes that the additional condition stated in Theorem 3, the one stating the transitivity of the cores on the blocks of imprimitivity, can be omitted but more work on this direction must be done. In any case these results generalize what was presented in [7] where the alphabet was binary since the automata presented in that article are almost group and the group generated by the permutation letter is the cyclic one, which is abelian and thus the additional condition is given. Once decided whether or not an automaton is completely reachable, the next interesting question is to find a bound to the shortest word required to reach subsets of size $1 \leq k < n$. In [9] it is stated that this bound is at most $2n(n-k)$; but we believe that due to the strict structure of the considered automata the bound can be improved. Nevertheless this problem is open by the moment.

REFERENCES

1. Ananichev D., Vorel V. A new lower bound for reset threshold of binary synchronizing automata with sink. *J. Autom. Lang. Comb.*, 2019. Vol. 24, No. 2–4. P. 153–164. DOI: 10.25596/jalc-2019-153

2. André J. M. Near permutation semigroups. *Semigroups and Languages*, 2004. World Scientific. P. 43–53. DOI: 10.1142/9789812702616_0003
3. Berlinkov M. V., Nicaud C. Synchronizing almost-group automata. *Internat. J. Found. Comput. Sci.*, 2020. Vol. 31, No. 08. P. 1091–1112. DOI: 10.1142/S0129054120420058
4. Bondar E., Casas D., Volkov M. Completely reachable automata: An interplay between automata, graphs, and trees. *Internat. J. Found. Comput. Sci.*, 2023. Vol. 34, No. 06. P. 655–690. DOI: 10.1142/S0129054123450053
5. Bondar E. A., Volkov M. V. Completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 9777: Descriptional Complexity of Formal Systems (DCFS 2016), C. Câmpeanu, F. Manea, J. Shallit (eds)*. Cham: Springer, 2016. International Publishing. P. 1–17. DOI: 10.1007/978-3-319-41114-9_1
6. Bondar E. A., Volkov M. V. A characterization of completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 11088: Developments in Language Theory (DLT2018), M. Hoshi, S. Seki (eds)*. Cham: Springer, 2018. P. 145–155. DOI: 10.1007/978-3-319-98654-8_12
7. Casas D., Volkov M.V. Binary completely reachable automata. In: *Lecture Notes in Comput. Sci., vol. 13568: LATIN 2022: Theoretical Informatics, Castañeda A., Rodríguez-Henríquez F. (eds)*. Cham, 2022. Cham: Springer. P. 345–358. DOI: 10.1007/978-3-031-20624-5_21
8. Don H. The Černý conjecture and 1-contracting automata. *Electron. J. Combin.*, 2016. Vol. 23, No. 3. Art. no. 3.12. DOI: 10.37236/5616
9. Ferens R., Szykulła M. Completely reachable automata: a polynomial algorithm and quadratic upper bounds. In: *LIPIcs, vol. 261: 50th International Colloquium on Automata, Languages, and Programming (ICALP 2023), K. Etessami, U. Feige, G. Puppis (eds)*. Dagstuhl, Germany: Dagstuhl Publishing. P. 59:1–59:17. DOI: 10.4230/LIPIcs.ICALP.2023.59
10. Hoffmann S. New characterizations of primitive permutation groups with applications to synchronizing automata. *Inform. and Comput.*, 2023. Vol. 295(A). Art. no. 105086. DOI: 10.1016/j.ic.2023.105086
11. Maslennikova M.I. *Reset Complexity of Ideal Languages.* 2014. 12 p. arXiv:1404.2816v1 [cs.FL]
12. Rystsov I. Estimation of the length of reset words for automata with simple idempotents. *Cybernet. Systems Anal.*, 2000. Vol. 36, No. 3. P. 339–344. DOI: 10.1007/BF02732984
13. Rystsov I., Szykuła M. *Primitive Automata that are Synchronizing.* 2023. 13 p. arXiv:2307.01302v1 [cs.FL]

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 49–59

DOI: 10.15826/umj.2024.2.005

STATISTICAL CONVERGENCE IN TOPOLOGICAL SPACE CONTROLLED BY MODULUS FUNCTION

Parthiba Das[†], **Susmita Sarkar**[††], **Prasenjit Bal**[†††]

Department of Mathematics, ICFAI University Tripura,
Kamalghat, Agartala, 799022, India

[†]parthivdas1999@gmail.com [††]susmitamsc94@gmail.com [†††]balprasenjit177@gmail.com

Abstract: The notion of f-statistical convergence in topological space, which is actually a statistical convergence's generalization under the influence of unbounded modulus function is presented and explored in this paper. This provides as an intermediate between statistical and typical convergence. We also present many counterexamples to highlight the distinctions among several related topological features. Lastly, this paper is concerned with the notions of s^f-limit point and s^f-cluster point for a unbounded modulus function f.

Keywords: Asymptotic density, f-statistical convergence, f-statistical limit point, f-statistical cluster point.

1. Introduction

Statistical density was initially introduced in Zygmund's 1935 monograph [18]. Extending on the concept using statistical density, Fast [12] (along with Schohenberg [17]) in 1951 broadened the definition of convergence to include statistical convergence. Let $\mathbb{N}$ denote the set of natural numbers, and $A \subseteq \mathbb{N}$. The notation $\delta(A)$ signifies the natural density or asymptotic density of set A [12], defined as

$$\delta(A) = \lim_{n\to\infty} \frac{|\{k \le n : k \in A\}|}{n}.$$

A real sequence $\{x_n : n \in \mathbb{N}\}$ is considered statistically convergent to a point l (see [17]) if, for every $\epsilon > 0$

$$\delta(\{n \in \mathbb{N} : |x_n - l| \ge \epsilon\}) = 0.$$

Subsequent to the contributions of Fridy [13] and Connor [9] in the realm of statistical convergence, other mathematicians have displayed considerable interest in this domain. In 2008, Maio and Kočinac [15] extended the notion to statistical convergence in topological spaces. In a topological space (X,τ), a sequence $\{x_n : n \in N\}$ is deemed statistically convergent to a point l if, for every neighborhood U of l,

$$\delta(\{n \in \mathbb{N} : x_n \notin U\}) = 0.$$

This form of convergence has proven to be highly valuable across various fields, particularly in the examination of open cover classes and selection principles [1–4, 10, 14, 16].

To establish a notion of convergence that sits between the ordinary convergence and statistical convergence many authors produced several approches. In 2012, Bhunia et al. [6] (see also Çolak et al. [7, 8]) enhanced the idea of s-convergence for real sequences by imposing a limitation on asymptotic density up to order α, where $0 < \alpha \le 1$. Through the utilization of asymptotic density of order α, a more stringent convergence criterion is introduced, surpassing statistical convergence but remaining less stringent than the conventional convergence in a topological space. As a direct outcome of this exploration, a novel class of open covers, denoted as $s^\alpha - \Gamma$, emerges. With the same purpose the unbounded modulus function is used in this paper and the concept of f-statistical convergence has been extended to the topological point of view. The class of modulus functions, as described by the given conditions, is a set of functions from the positive real numbers to the positive real numbers. Let's break down the key properties:

1. *Zero at Zero.* The function $f(x)$ is equal to zero if and only if x is equal to zero. This implies that the function is zero only at the origin.

2. *Subadditivity.* For any positive real numbers x and y, the function satisfies the property $f(x+y) \leq f(x) + f(y)$. This condition is known as subadditivity, indicating that the function's values do not grow faster than the sum of its individual parts. It's a form of the triangle inequality.

3. *Monotonicity.* The function is increasing, meaning that as the input increases, the function values also increase. Mathematically, if $a < b$, then $f(a) \leq f(b)$.

4. *Right-Continuous at Zero.* The function $f(x)$ is continuous from the right at $x = 0$. This implies that as the input approaches zero from the positive side, the function values approach the limit without any sudden jumps or discontinuities.

This class of modulus functions appears to capture functions that exhibit properties similar to those of the absolute value function. The conditions ensure a certain level of behavior for the function, making it well-behaved and suitable for various mathematical applications.

Function f is unbounded if

$$\lim_{x \to \infty} f(x) = \infty.$$

For an unbounded modulus function f, f-density of a set $A \subseteq \mathbb{N}$ is denoted by $\delta^f(A)$ and is defined as [5]

$$\delta^f(A) = \lim_{n \to \infty} \frac{f(|\{k \leq n : k \in A\}|)}{f(n)}.$$

In this paper we have explored that the function f is very useful to control the rate at which statistical convergence occurs. We extend the concept of s-convergence to s^f-convergence in topological environment and explore several attributes of this convergence criteria. In last section we investigate some properties of s^f-limit points and s^f-cluster points.

2. Preliminaries

In this paper, a space X is defined as a topological space X with topology τ. No separation axioms have been granted, unless otherwise stated. For standard ideas, symbols, and terminology, we refer to [11]. For the convenience of the readers, this section includes certain required concepts.

The unbounded modulus functions defined on the set $\mathbb{N}$ of all natural numbers are the modulus functions taken into consideration in this study. Therefore, right continuous at zero and zero to zero property are disregarded. The modulus function $f : \mathbb{N} \to \mathbb{R}^+$, defined as $f(n) = \log(1+n)$, is regarded as such in the majority of the cases. It is obvious to note that this modulus function is unbounded.

Definition 1 [5]. *For an unbounded modulus function f, f-density of a set $A \subseteq \mathbb{N}$ is denoted by $\delta^f(A)$ and is defined as*

$$\delta^f(A) = \lim_{n \to \infty} \frac{f(|\{k \leq n : k \in A\}|)}{f(n)}.$$

Using the concept Bhardwaj et al. [5] extended the concept of statistical convergence of a real sequence upto s^f-convergence.

Definition 2 [5]. *A real sequence $\{x_n : n \in \mathbb{N}\}$ is considered s^f-convergent to a point l if, for every $\epsilon > 0$,*

$$\delta^f(\{n \in \mathbb{N} : |x_n - l| \geq \epsilon\}) = 0,$$

where f is an unbounded modulus function.

In [15], the concept of statistical dense sub-sequence, s-limit point of a sequence and s-cluster point of a sequence are discussed.

Definition 3 [15]. *A subsequence $\mathcal{V} = \{x_{n_k} : k \in \mathbb{N}\}$ of the sequence $\{x_n : n \in \mathbb{N}\}$ is called a statistically dense if*

$$\delta(\{n_k : x_{n_k} \in \mathcal{V}\}) = 1.$$

Definition 4 [15]. *A point x is said to be a statistical limit point of a sequence $\{x_n : n \in \mathbb{N}\}$ in a space X, if there is a set $\{n_1 < n_2 < ... < n_k < ...\} \subset \mathbb{N}$ whose asymptotic density is not zero (which means that it is greater than zero or does not exist) such that*

$$\lim_{k\to\infty} x_{n_k} = x.$$

Definition 5 [15]. *A point x is called a statistical cluster point of a sequence $\{x_n : n \in \mathbb{N}\}$ if for each neighborhood U of x the asymptotic density of the set $\{n \in \mathbb{N} : x_n \in U\}$ is positive.*

3. On f-statistical convergence

Definition 6. *Let $f : \mathbb{N} \to \mathbb{R}$ be an unbounded modulus function and (X, τ) be a topological space. A sequence $\{x_n : n \in \mathbb{N}\}$ in X will be called f-statistical convergent (in short s^f-convergent) to $x \in X$, if for every neighborhood U of x,*

$$\delta^f(\{n \in \mathbb{N} : x_n \notin U\}) = 0,$$

$$\text{i.e.,} \quad \lim_{n\to\infty} \frac{f(|\{k \leq n : x_n \notin U\}|)}{\{f(n)\}} = 0.$$

From the study of Maio and Kočinac [15], we know that every convergent sequence is statistical convergent but converse is not true. Since for a finite set F, $\delta^f(F) = 0$, therefore usual convergence implies the s^f-convergence. For any unbounded modulus function f and $A \subseteq \mathbb{N}$, if $\delta^f(A) = 0$, then $\delta(A) = 0$. Thus every s^f-convergent sequence is statistical convergent. Moreover, concept of statistical convergence coincides with the concept of s^f-convergence if the unbounded modulus function under consideration is $f(n) = n$.

Example 1. There is a sequence in topological space which is statistical convergent but not s^f-convergent and a sequence which is s^f-convergent but not convergent.

Let (X, τ) be a topological space where $X = \{1, 2\}$ and $\tau = \mathcal{P}(X)$. Consider the sequence $\{x_n : n \in \mathbb{N}\}$ where

$$x_n = \begin{cases} 1, & \text{if} \quad n = m^2 \text{ for some } m \in \mathbb{N}. \\ 2, & \text{otherwise.} \end{cases}$$

Let the function $f(n) = \log(1 + n)$ be the unbounded modulus function under consideration. Neighborhoods of 2 are $U_1 = \{2\}$ and $U_2 = X$. Here,

$$\delta(\{n \in \mathbb{N} : x_n \notin U_1\}) = \delta(\{m^2 : m \in \mathbb{N}\}) = 0$$

and

$$\delta(\{n \in \mathbb{N} : x_n \notin U_2\}) = \delta(\emptyset) = 0.$$

Therefore,

$$x_n \overset{s-\lim}{\longrightarrow} 2.$$

But for the neighborhood $U_1 = \{2\}$ of 2, we have

$$\delta^f(\{n \in \mathbb{N} : x_n \notin U_1\}) = \frac{1}{2} \neq 0.$$

Also, for the neighborhood $V = \{1\}$ of 1, we have

$$\delta^f(\{n \in \mathbb{N} : x_n \notin V\}) = 1 \neq 0.$$

So, $\{x_n : n \in \mathbb{N}\}$ is neither s^f-convergent to 1 nor s^f-convergent to 2.

In the same space consider the sequence $\{y_n : n \in \mathbb{N}\}$ where

$$y_n = \begin{cases} 1, & \text{if} \quad n = m^m \text{ for some } m \in \mathbb{N}, \\ 2, & \text{otherwise.} \end{cases}$$

Then $\{y_n : n \in \mathbb{N}\}$ is s^f-convergent to 2 but not convergent.

Thus we have the following diagram (see Fig. 1).

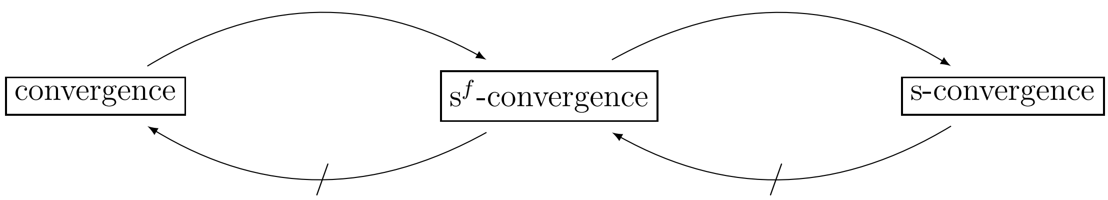


Figure 1. Types of convergence and the relationship between them.

Definition 7. *A sequence $\{x_n : n \in \mathbb{N}\}$ in a topological space X is said to s^f_* convergent to $x_0 \in X$ if there exists $A \subseteq \mathbb{N}$ with $\delta^f(A) = 1$ such that*

$$\lim_{m \to \infty,\ m \in A} x_n = x_0.$$

Example 2. There exists a sequence $\{x_n : n \in \mathbb{N}\}$ which is s^f_*-convergent but not s^f-convergent.

Let us assume a topological space (X, τ) where $X = \{a, b\}$ and $\tau = \{\emptyset, X, \{a\}, \{b\}\}$. Again we construct a sequence $\{x_n : n \in \mathbb{N}\}$ where

$$x_n = \begin{cases} a, & \text{if} \quad n \in 2\mathbb{N}, \\ b, & \text{otherwise.} \end{cases}$$

Let $f(n) = \log(1+n)$ be the modulus function under consideration then, for the neighbourhood $\{a\}$ of 'a', we have

$$\delta^f(\{n \in \mathbb{N} : x_n \notin \{a\}\}) = \delta^f(\{\mathbb{N} \setminus 2\mathbb{N}\}) = \lim_{n \to \infty} \frac{f(n+1)}{f(2n+1)} = \lim_{n \to \infty} \frac{\log(n+2)}{\log(2n+2)} = 1.$$

And for the neighbourhood $\{b\}$ of 'b', we have

$$\delta^f(\{n \in \mathbb{N} : x_n \notin \{a\}\}) = \delta^f(\{2\mathbb{N}\}) = \lim_{n\to\infty} \frac{f(n)}{f(2n)} = \lim_{n\to\infty} \frac{\log(n+1)}{\log(2n+1)} = 1.$$

Therefore $\{x_n : n \in \mathbb{N}\}$ neither s^f-convergent to a nor s^f-convergent to 'b'.
On the other hand $2\mathbb{N} \subseteq \mathbb{N}$ such that $\delta^f(2\mathbb{N}) = 1$ and $\{x_n : n \in \mathbb{N}\} = \{a, a, \dots\}$

$$\lim_{n\to\infty,\ n\in 2\mathbb{N}} x_n = a, \quad \Rightarrow x_n \overset{s_*^f-\lim}{\to} a.$$

Although s_*^f-convergence does not imply the s^f-convergence of a sequence, the s^f-convergence of a sequence implies its s_*^f-convergence in a first countable space.

Theorem 1. *In a first countable space, if a sequence $\{x_n : n \in \mathbb{N}\}$ in X s^f-converges to x, then this sequence s_*^f-converges to x.*

P r o o f. Let (X, τ) be a first countable topological space and $\{x_n : n \in \mathbb{N}\}$ be a sequence in X which s^f-converges to x. Since X is first countable, there exists countable decreasing local base $U_{1,x} \supseteq U_{2,x} \supseteq U_{3,x} \supseteq \dots$ at the point x. Now consider a set $A_i = \{n \in \mathbb{N} : x_n \in U_{i,x}\}$ for every $i \in \mathbb{N}$. Then we have $A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$.

Again we know that sequence $\{x_n : n \in \mathbb{N}\}$ is s^f-convergent then

$$\begin{aligned} &\delta^f(\{n \in \mathbb{N} : x_n \notin U_{i,x}\}) = 0, \quad \forall i \in \mathbb{N} \\ &\implies \delta^f(\{A_i^c\}) = 0, \quad \forall i \in \mathbb{N} \\ &\implies \delta^f(\{A_i\}) = 1, \quad \forall i \in \mathbb{N}. \end{aligned}$$

Let $m_1 \in A_1$ be arbitrary. Since $\delta^f(A_2) = 1$, we can find a $m_2 \in A_2$ such that $m_2 > m_1$ and such that

$$\frac{f(|\{A_2(n)\}|)}{f(n)} = \frac{f(|\{k \in A_2 : k \leq n\}|)}{f(n)} > \frac{1}{2} = 1 - \frac{1}{2}, \quad \text{for all} \quad n \geq m_2.$$

In similar way if we obtain $m_1 < m_2 < \dots < m_i \in A_i$, such that for every $n \geq m_i$ then,

$$\frac{f(|\{A_i(n)\}|)}{f(n)} = \frac{f(|\{k \in A_i : m \leq n\}|)}{f(n)} > 1 - \frac{1}{i}.$$

Now we define a set $A \subseteq \mathbb{N}$ as for each $m \leq m_1$ and $m \in A$; if $i \geq 1$ and for $m_i < m \leq m_{i+1}$, $m \in A$ if and only if $m \in A_i$. Let $A = \{n_1 < n_2 < \dots\}$. For all $n \in \mathbb{N}$ such that $m_i \leq n \leq m_{i+1}$, we have

$$\begin{aligned} \frac{f(|\{A(n)\}|)}{f(n)} &\geq \frac{f(|\{A_i(n)\}|)}{f(n)} \geq 1 - \frac{1}{i}, \\ \textit{i.e.,} \quad \lim_{n\to\infty} \frac{f(|\{A(n)\}|)}{f(n)} &\geq \lim_{n\to\infty} \frac{f(|\{A_i(n)\}|)}{f(n)} \geq \lim_{n\to\infty} 1 - \frac{1}{i}, \\ \textit{i.e.,} \quad \lim_{n\to\infty} \frac{f(|\{A(n)\}|)}{f(n)} &= 1, \ \Rightarrow \delta^f(A) = 1. \end{aligned}$$

Now, let V be a neighbourhood of x and $U_i \subseteq V$. If we take $n \in A$, $n \geq m_i$ then there exists $j \geq i$ such that we get $m_j \leq n \leq m_{j+1}$. So by the construction of A, $n \in A_j$. Therefore, for each $n \in A$, $n \geq m_i$ we get $x_n \in U_j$ and $x_n \in U_j \subseteq U_i \subseteq V$,

$$\textit{i.e.,} \quad \lim_{n\to\infty,\ n\in A} x_n = x.$$

Thus $\{x_n : n \in \mathbb{N}\}$ s_*^f-converges to x. □

Example 3. s^f-limit of an s^f-convergent sequence may not be unique.

Let us assume a topological space (X, τ) where $X = \{a, b, c\}$, $\tau = \{\emptyset, X, \{b, c\}, \{a\}\}$ and $f(n) = \log(1 + n)$ be the unbounded modulus function under consideration. We construct a sequence $\{x_n : n \in \mathbb{N}\}$ where

$$x_n = \begin{cases} a, & \text{if} \quad n = m^m \quad \text{for some} \quad m \in \mathbb{N}, \\ b, & \text{otherwise.} \end{cases}$$

Open neighbourhoods of b are $U_1 = X$ and $U_2 = \{b, c\}$.
For the neighbourhood U_1, $\{n \in \mathbb{N} : x_n \notin U_1\} = \emptyset$. So, $\delta^f(\{n \in \mathbb{N} : x_n \notin U_1\}) = 0$.
For the neighbourhood U_2 we have $\{n \in \mathbb{N} : x_n \notin U_2\} = \delta^f(\{n^n : n \in \mathbb{N}\})$.

Therefore

$$\delta^f(\{n^n : n \in \mathbb{N}\}) = \lim_{n\to\infty} \frac{\log(n+1)}{\log(n^n+1)}$$

$$= \lim_{n\to\infty} \frac{n^n + 1}{(n+1)n^n(1+\log(n))} = \lim_{n\to\infty} \frac{n^n(1+1/n^n)}{n^n n(1+1/n)(1+\log(n))} = 0.$$

Therefore, $\{x_n : n \in \mathbb{N}\}$ is f-statistical convergent sequence and $x_n \overset{s^f-\lim}{\rightarrow} b$.

Since, neighbourhood of b is the only neighbourhood of c, we can say for every neighbourhood U of c also

$$\delta^f(\{n \in \mathbb{N} : x_n \notin U\}) = 0.$$

Thus

$$x_n \overset{s^f-\lim}{\rightarrow} c.$$

Thus the limit of an s^f-convergent sequence may not be unique.

Theorem 2. *In a Hausdorff space any s^f-convergent sequence has a unique limit.*

P r o o f. Let $\{a_n : n \in \mathbb{N}\}$ be a s^f-convergent sequence in a topological space (X, τ) and

$$a_n \overset{s^f-\lim}{\rightarrow} a, \quad a_n \overset{s^f-\lim}{\rightarrow} b.$$

Since, X is Hausdorff space then there exist open sets G and H such that $a \in G, b \in H$ and $G \cap H = \emptyset$. But $\{a_n : n \in \mathbb{N}\}$ is an s^f-convergent sequence which s^f-converges to both a and b. Therefore

$$\delta^f(\{n \in \mathbb{N} : a_n \notin G\}) = 0, \quad \delta^f(\{n \in \mathbb{N} : a_n \notin H\}) = 0.$$

Since, $G \cap H = \emptyset$ and $H \subseteq X \setminus G$. Now,

$$\delta^f(\{n \in \mathbb{N} : a_n \in H\}) \le \delta^f(\{n \in \mathbb{N} : a_n \in X \setminus G\}) = \delta^f(\{n \in \mathbb{N} : a_n \notin G\}) = 0.$$

So,

$$\delta^f(\{n \in \mathbb{N} : a_n \in H\}) = 0$$

and hence

$$\delta^f(\{n \in \mathbb{N} : a_n \notin H\}) = 1.$$

This contradicts the fact that

$$\delta^f(\{n \in \mathbb{N} : a_n \notin H\}) = 0.$$

Hence in Hausdorff space any s^f-convergent sequence has the unique limit. □

Proposition 1. *In a discrete topological space (X, τ), let $p, q \in X$. $h : \mathbb{N} \to \mathbb{N}$ be an one-one function and f be the unbounded modulus function under consideration. Then the sequence*

$$x_n = \begin{cases} p, & \text{if} \quad n = h(k) \quad \text{for some} \quad k \in \mathbb{N}, \\ q, & \text{otherwise} \end{cases}$$

is s^f-convergent to q if

$$\lim_{n\to\infty} \frac{f(n)}{f \circ h(n)} \to 0$$

and does not converges otherwise.

P r o o f. In the mentioned topological space, $\{q\}$ is the smallest neighbourhood of q. To show the s^f-convergence of the sequence $\{x_n : n \in \mathbb{N}\}$, it is enough to show that

$$\delta^f(\{n \in \mathbb{N} : x_n \notin \{q\}\}) = 0.$$

Now,

$$\delta^f(\{n \in \mathbb{N} : x_n \notin \{q\}\}) = \lim_{n\to\infty} \frac{f(|\{k \le n : x_n \notin \{q\}\}|)}{f(n)}$$
$$= \lim_{n\to\infty} \frac{f(|\{h(k) \le n : k \in \mathbb{N}\}|)}{f(n)} = \lim_{n\to\infty} \frac{f(n)}{f \circ h(n)}.$$

Hence the proposition is true. □

Example 4. Subsequence of an s^f-convergent sequence may not be s^f-convergent.

Let us assume a topological space (X, τ) where $X = \{a, b, c\}$ and $\tau = \{\emptyset, X, \{b, c\}, \{a\}\}$ and a sequence $\{x_n : n \in \mathbb{N}\}$ where,

$$x_n = \begin{cases} b, & \text{if} \quad n = m^m \quad \text{for some} \quad m \in \mathbb{N}, \\ a, & \text{otherwise.} \end{cases}$$

Now for every open neighbourhood U of a , we get

$$\{n \in \mathbb{N} : x_n \notin U\} \subseteq \{m^m : m \in \mathbb{N}\}.$$

Let us consider the unbounded modulus function $f(x) = \log(1 + x)$ then we have,

$$\delta^f(\{n \in \mathbb{N} : x_n \notin \{a\}\}) = 0.$$

So, $\{x_n : n \in \mathbb{N}\}$ is an s^f-convergent sequence. Again we construct a subsequence

$$x_{n_i} = \begin{cases} x_{i^i}, & \text{if} \quad i \quad \text{is odd}, \\ x_{((i-1)^{i-1}+1)}, & \text{if} \quad i \quad \text{is even.} \end{cases}$$

Now for the open neighbourhoods of $\{a\}$ of a, we have

$$\delta^f(\{i \in \mathbb{N} : x_{n_i} \notin \{a\}\}) = \delta^f(\{2n : n \in \mathbb{N}\}) = 1 \neq 0.$$

Similarly for the open neighbourhood of $\{b, c\}$ we have

$$\delta^f(\{i \in \mathbb{N} : x_{n_i} \notin \{b, c\}\}) = \delta^f(\{1, 3, 5, \dots\}) = 1 \neq 0.$$

Therefore,

$$x_n \overset{s^f-\lim}{\nrightarrow} a, \quad x_n \overset{s^f-\lim}{\nrightarrow} b \quad \text{and} \quad x_n \overset{s^f-\lim}{\nrightarrow} c.$$

So, $\{x_{n_i} : i \in \mathbb{N}\}$ is not s^f-convergent sequence. □

Definition 8. *A subsequence $B = \{x_{n_k} : k \in \mathbb{N}\}$ of any sequence $A = \{x_n : n \in \mathbb{N}\}$ is called statistically f-dense (or s^f-dense) if $\delta^f(n_k : x_{n_k} \in B) = 1$.*

Theorem 3. *In a topological space (X, τ), a sequence $\{x_n : n \in \mathbb{N}\}$ is s^f-convergent if and only if each of its s^f-dense subsequence is s^f-convergent.*

P r o o f. Suppose (X, τ) be a topological space and $\{x_n : n \in \mathbb{N}\}$ be a sequence for which every s^f-dense subsequence is s^f-convergent. But

$$\lim_{n\to\infty} \frac{f(|\{k \leq n : x_k \in \{x_n : n \in \mathbb{N}\}\}|)}{f(n)} = \lim_{n\to\infty} \frac{f(n)}{f(n)} = 1$$

for every unbounded modulus function f. So $\{x_n : n \in \mathbb{N}\}$ is s^f-dense in itself. Therefore, $\{x_n : n \in \mathbb{N}\}$ is s^f-convergent.

Conversely, let $\{x_n : n \in \mathbb{N}\}$ be a s^f-convergent sequence of a topological space (X, τ) and a subsequence $\{x_{n_k} : k \in \mathbb{N}\}$ is s^f-dense but not s^f-convergent. Therefore, there exists a point $p \in X$ and a neighbourhood U of p such that $\delta^f(\{k \in \mathbb{N} : x_{n_k} \notin U\}) \neq 0$. Now,

$$\begin{aligned}\lim_{n\to\infty} \frac{f(|\{n \in \mathbb{N} : x_n \notin U\}|)}{f(n)} &\geq \lim_{n\to\infty, k\to\infty} \frac{f(|\{n_k \in \mathbb{N} : x_{n_k} \notin U\}|)}{f(n)} \\ &= \lim_{k\to\infty} \frac{f(|\{n_k \in \mathbb{N} : x_{n_k} \notin U\}|)}{f(|n_k|)} \times \lim_{n\to\infty} \frac{f(|n_k|)}{f(n)} \neq 0.\end{aligned}$$

Since,

$$\lim_{k\to\infty} \frac{f(|\{k \in \mathbb{N} : x_{n_k} \notin U\}|)}{f(k)} \neq 0$$

and $\{x_{n_k} : k \in \mathbb{N}\}$ is s^f-dense then

$$\text{i.e.} \quad \lim_{n\to\infty} \frac{f(|n_k|)}{f(n)} = 1.$$

So, we get

$$\delta^f(\{n \in \mathbb{N} : x_n \notin U\}) \geq \delta^f(\{k \in \mathbb{N} : x_{n_k} \notin U\}) \neq 0.$$

Therefore $\{x_n : n \in \mathbb{N}\}$ is not s^f-convergent sequence, which is a contradiction. So $\{x_{n_k} : k \in \mathbb{N}\}$ must be s^f-convergent. □

4. f-statistical limit point, f-statistical cluster point

In this section we extend the concept of statistical limit point to s^f-limit point by incorporating an unbounded modulus function f.

Definition 9. *In a topological space (X, τ), a point x_0 is called a f-statistical limit point (in short s^f-limit point) of a sequence $\{x_n : n \in \mathbb{N}\}$ if there exists a subsequence $\mathcal{V} = \{x_{n_k} : k \in \mathbb{N}\}$ such that $\delta^f\{n_k : k \in \mathbb{N}$ and $x_{n_k} \in \mathcal{V}\} > 0$ and*

$$\lim_{k\to\infty} x_{n_k} = x_0.$$

Definition 10. *In a topological space (X, τ), a point x_0 is called f-statistical cluster point (in short s^f-cluster point) of any sequence $\{x_n : n \in \mathbb{N}\}$ if for each neighbourhood U of x_0, $\delta^f\{n \in \mathbb{N} : x_n \in U\} > 0$.*

We denote the set of all f-statistical limit points and f-statistical cluster points by Λ_f and Θ_f, respectively.

Theorem 4. *For a sequence $\{x_n : n \in \mathbb{N}\}$ in a topological space (X, τ), $\Lambda_f(x_n) \subset \Theta_f(x_n)$.*

P r o o f. Let (X, τ) be a topological space, $\{x_n : n \in \mathbb{N}\}$ be a sequence and any point p be f-statistical limit point. Therefore, $p \in \Lambda_f(x_n)$. Then there exists a subsequence $\{x_{n_k} : k \in \mathbb{N}\}$, where $\{n_k : k \in \mathbb{N}\}$ have a positive δ^f-density and

$$\lim_{k\to\infty} x_{n_k} = p.$$

Now,

$$\delta^f(\{n_k : k \in \mathbb{N}\}) = \alpha \text{ (say)} > 0$$

and for every neighbourhood U of p, $\{n_k : x_{n_k} \notin U\} = F$ (say) is finite. But,

$$(\{n \in \mathbb{N} : x_n \in U\}) \supset (\{n_k : k \in \mathbb{N}\}) \setminus F.$$

Since F is a finite set,

$$\lim_{n\to\infty} \frac{f(|F|)}{f(n)} = 0$$

and f is a modulus function,

$$\begin{aligned} \delta^f(\{k_k : k \in \mathbb{N}\} \setminus F) &= \lim_{n\to\infty} \frac{f(|\{n_k : k \in \mathbb{N}\} \setminus F|)}{f(n)} = \lim_{n\to\infty} \frac{f(|\{n_k : k \in \mathbb{N}\} \setminus F|)}{f(n)} + \lim_{n\to\infty} \frac{f(|F|)}{f(n)} \\ &\geq \lim_{n\to\infty} \frac{f(|\{n_k : k \in \mathbb{N}\} \setminus F| + |F|)}{f(n)} = \lim_{n\to\infty} \frac{f(|\{n_k : k \in \mathbb{N}\}|)}{f(n)} = \delta^f(\{n_k : k \in \mathbb{N}\}). \end{aligned}$$

Therefore,

$$\delta^f(\{n \in \mathbb{N} : x_n \in U\}) \geq \delta^f(\{n_k : k \in \mathbb{N}\}) = \alpha > 0.$$

Therefore $p \in \Theta_f(x_n)$. So, $\Lambda_f(x_n) \subset \Theta_f(x_n)$. □

Theorem 5. *In a topological space (X, τ), for a sequence $\{x_n : n \in \mathbb{N}\}$, the set $\Theta_f(x_n)$ is a closed set.*

P r o o f. Let $\{x_n : n \in \mathbb{N}\}$ be a sequence in a topological space (X, τ). Let U be an arbitrary neighbourhood of point the point $x_0 \in \overline{\Theta_f(x_n)}$. So $U \cap \Theta_f(x_n) \setminus \{x_0\} \neq \emptyset$. Then we can choose another point $x_0' \in U \cap \Theta_f(x_n)$, where x_0' is a f-statistical cluster point. Then there exist a neighbourhood V of a point x_0' such that $V \subset U$ and

$$\delta^f(\{n \in \mathbb{N} : x_n \in V\}) = \alpha > 0.$$

Obviously,

$$\{n \in \mathbb{N} : x_n \in U\} \supset \{n \in \mathbb{N} : x_n \in V\}$$

and hence

$$\delta^f(\{n \in \mathbb{N} : x_n \in U\}) \supset \delta^f(\{n \in \mathbb{N} : x_n \in V\}) = \alpha > 0.$$

It means that $\delta^f(\{n \in \mathbb{N} : x_n \in U\})$ is not a set that has zero δ^f-density, i.e., $x_0 \in \Theta_f(x_n)$. So $\overline{\Theta_f(x_n)} = \Theta_f(x_n)$. Hence $\Theta_f(x_n)$ is a closed set. □

Theorem 6. *In a topological space (X,τ), if there exist two sequence $\{x_n : n \in \mathbb{N}\}, \{y_n : n \in \mathbb{N}\}$ such that $\delta^f(\{n \in \mathbb{N} : x_n \neq y_n\}) = 0$, then $\Theta_f(x_n) = \Theta_f(y_n)$ and $\Lambda_f(x_n) = \Lambda_f(y_n)$.*

P r o o f. Let $\{x_n : n \in \mathbb{N}\}$ and $\{y_n : n \in \mathbb{N}\}$ be two sequence of a topological space (X,τ). Suppose that q be any f-statistical cluster point with respect to $\{x_n : n \in \mathbb{N}\}$ sequence. So, for every neighbourhood U of q,

$$\delta^f(\{n \in \mathbb{N} : x_n \in U\}) > 0.$$

We have

$$\lim_{n\to\infty} \frac{f(|\{n \in \mathbb{N} : x_n \in U\}|)}{f(n)} > 0$$

and

$$\{n \in \mathbb{N} : x_n \in U\} \setminus \{n \in \mathbb{N} : x_n \neq y_n\} \subseteq \{n \in \mathbb{N} : y_n \in U\}.$$

Since $\delta^f\{n \in \mathbb{N} : x_n \neq y_n\} = 0$ then we get $\delta_f(\{n \in \mathbb{N} : y_n \in U\}) > 0$. This means that the set $\{n \in \mathbb{N} : y_n \in U\}$ is not a set that has zero δ_f-density so q is also f-statistical cluster point with respect to $\{y_n : n \in \mathbb{N}\}$ sequence. Therefore $\Theta_f(x_n) \subset \Theta_f(y_n)$. It is easy to see that $\Theta_f(y_n) \subset \Theta_f(x_n)$ from symmetry. Finally we have $\Theta_f(x_n) = \Theta_f(y_n)$. The equality $\Lambda_f(x_n) = \Lambda_f(y_n)$ can be shown in a similar way. □

5. Conclusion

An unbounded modulus function can help to manage the rate of statistical convergence up to a great extend. In a first countable space, s^f_*-convergence does not entail s^f-convergence, although s^f-convergence requires s^f_*-convergence. An s^f-convergent sequence posses a unique limit in a Hausdörff space. A sequence is s^f-convergent if and only if each of its s^f-dense subsequence is s^f-convergent. The set $\Lambda_f(x_n)$ of all f-statistical limit points of a sequence $\{x_n\}$ is a subset of the set $\Theta_f(x_n)$ of all f-statistical cluster points of that sequence. Moreover the collection of all f-statistical cluster points forms a closed set in related topological space.

Acknowledgements

The authors of the paper are thankful to the editor and the reviewers for their insightful comments which has immensely improved the quality and representation of the paper.

REFERENCES

1. Bal P., Rakshit D. A variation of the class of statistical γ covers. *Topol. Algebra Appl.*, 2023. Vol. 11, No. 1. Art. no. 20230101. DOI: 10.1515/taa-2023-0101
2. Bal P., Rakshit D., Sarkar S. Countable compactness modulo an ideal of natural numbers. *Ural Math. J.*, 2023. Vol. 9, No. 2. P. 28–35. DOI: 10.15826/umj.2023.2.002
3. Bal P., Rakshit D., Sarkar S. On strongly star semi-compactness of topological spaces. *Khayyam J. Math.*, 2023. Vol. 9, No. 1. P. 54–60.
4. Bal P. A countable intersection like characterization of Star-Lindelöf spaces. *Researches in Math.*, 2023. Vol. 31, No. 2. P. 3–7.
5. Bhardwaj B.K., Dhawan S. Korovkin type approximation theorems via f-statistical convergence. *J. Math. Anal.*, 2018. Vol. 9., No. 2. P. 99–117.
6. Bhunia S., Das P., Pal S.K. Restricting statistical convergence. *Acta Math. Hung.*, 2012. Vol. 134, No. 1–2. P. 153–161. DOI: 10.1007/s10474-011-0122-2

7. Çolak R. Statistical convergence of order α. In: *Modern Methods in Analysis and Its Applications, M. Mursaleen, ed.* New Delhi: Anamaya Publ., 2010. P. 121–129.
8. Çolak R., Bektaş Ç. A. λ-statistical convergence of order α. *Acta Math. Sci., Ser. B, Engl. Ed.*, 2011. Vol. 31, No. 3. P. 953–959. DOI: 10.1016/S0252-9602(11)60288-9
9. Connor J. S. The statistical and strong p-Cesaro convergence of sequences. *Analysis*, 1988. Vol. 8, No. 1–2. P. 47–63. DOI: 10.1524/anly.1988.8.12.47
10. Das P. Certain types of open covers and selection principles using ideals. *Houston J. Math.*, 2013, Vol. 39, No. 2. P. 637–650.
11. Engelking R. *General Topology.* Berlin: Heldermann Verlag, 1989. 529 p.
12. Fast H. Sur la convergence statistique. *Colloq. Math.*, 1951. Vol. 2, No. 3–4. P. 241–244. (in French)
13. Fridy J. A. On statistical convergence. *Analysis*, 1985. Vol. 5, No. 4. P. 301–313. DOI: 10.1524/anly.1985.5.4.301
14. Lahiri B. K., Das P. I and I^* convergence in topological spaces. *Math. Bohem.*, 2005. Vol. 130, No. 2. P. 153–160. DOI: 10.21136/MB.2005.134133
15. Maio G. D., Kočinac L. D. R. Statistical convergence in topology. *Topology Appl.*, 2008. Vol. 156, No. 1. P. 28–45. DOI: 10.1016/j.topol.2008.01.015
16. Šalát T. On statistically convergent sequences of real numbers. *Math. Slovaca*, 1980. Vol. 30, No. 2. P. 139–150.
17. Schoenberg I. J. The integrability of certain functions and related summability methods. *Amer. Math. Monthly*, 1959. Vol. 66, No. 5. P. 361–375. DOI: 10.2307/2308747
18. Zygmund A. *Trigonometrical Series.* Monogr. Mat., Warszawa, vol. 5. Warszawa: PWN-Panstwowe Wydawnictwo Naukowe, 1935. 332 p.

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 60–68

DOI: 10.15826/umj.2024.2.006

GRAPHICAL PROPERTIES OF CLUSTERED GRAPHS

Sambanthan Gurunathan

Department of Mathematics, Rajalakshmi Engineering College,
Thandalam, Chennai - 602 105, India

gurunathan.s@rajalakshmi.edu.in

Thangaraj Yogalakshmi

Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology(VIT),
Vellore - 632 014, Tamil Nadu, India

yogalakshmithangaraj@vit.ac.in

Abstract: Clustering is a strategy for discovering homogeneous clusters in heterogeneous data sets based on comparable structures or properties. The number of nodes or links that must fail for a network to be divided into two or more sub-networks is known as connectivity. In addition to being a metric of network dependability, connectivity also serves as an indicator of performance. The Euler graph can represent almost any issue involving a discrete arrangement of objects. It can be analyzed using the recent field of mathematics called graph theory. This paper discusses the properties of clustered networks like connectivity and chromaticity. Further, the structure of the antipodal graph in the clustered network has been explored.

Keywords: Clustered graph, Euler graph, Antipodal graph.

1. Introduction

Clustering is a widely used technique to identify similar data or objects from the enormous one. It is used in genetic technology to group the genes with similar expression patterns [16]. Ustumbas [17] used tripartite graph clustering in a social network as a cluster of three types of data objects simultaneously and produced useful information for a recommended system.

A graph is commonly used to depict the architecture of an interconnection network. When constructing the topology of interconnection networks, there are various mutually exclusive needs [1]. Network dependability is an important consideration while planning the architecture of an interconnection network. The ability of a system to deliver services that can be legitimately trusted is known as dependability [2, 6]. A network's connectivity is defined as the number of nodes or links that must fail for the network to be partitioned into two or more distinct sub-networks. Network connection assesses a network's resilience and capacity to continue operating despite the presence of certain failing components. A higher node or link connection improves the network's resiliency to failure. Connectivity is not only a metric of network dependability but also a measure of performance. The number of links that must be crossed to reach a target node is reduced as connectivity improves. Because technical limits limit the number of connections per node to a limited value, designing a network with better connectivity and a consistent number of connections per node is critical.

It is implicitly assumed when using these measurements that any subset of the network components (channels or processors) might fail at the same moment. However, in certain networks, it is acceptable to assume that various subsets of network components do not break simultaneously. Classic connectivities may not be accurate gauges of dependability for these networks. Restricted connectivities may, of course, be used as a model to assess network dependability. Harary [9] developed the notion of conditional connection by demanding some characteristic for unconnected components of $G - F$, motivated by the inadequacies of existing connectivity measures. Esfahanian et. al [7] developed another similar study in which the authors provide extensions of edge connectivity by establishing specific constraints that detached components must meet. For applications where parallel algorithms might operate on sub-networks with a certain topology, requiring certain features for disconnected components is very critical. Restricting the defective sets to a certain class, on the other hand, was inspired by the fact that interconnection networks (which are typically represented as graphs) might include diverse components with varying levels of reliability [5, 8, 10, 13, 14, 18, 19].

In a computer network, which is known as digital telecommunications network, graphs are structured into clusters. Electronic links are employed within the same cluster where as optical links are employed between the cluster communications. These data links may be wire or optic cables or wireless media. Interconnection networks are strenuous to work with in abstract terms. This motivated many researchers to propose new improved network graphs arguing the benefits and performance evaluation in different contexts.

Yogalakshmi et. al [11] introduced the clustered graphs and found the degree-based topological indices for the same. Further, the interrelation of the indices was discussed. In [12], the distance-based topological indices of a clustered graph were computed. The clustered graph was derived from a complete tripartite graph $K_{r,s,t}$, $r \geq s \geq t$ with partite sets R, S, and T by converting each vertex as a complete graph with order equal to the degree of a corresponding vertex. The adjacency condition was preserved as in the tripartite graph.

This paper discusses the properties of clustered graphs such as connectivity, Euler property, antipodal graph, chromatic number, and chromatic index.

2. Preliminaries

Let $G = (V, E)$ be a connected simple graph with $|V| = p$ and $|E| = q$, where p and q are finite. Vertices u and v are adjacent if there is an edge $e \in E$ joining u and v, and the edge e is said to be incident with u and v. The number of edges incident with a vertex u is called the degree of the vertex and is denoted by $d(u)$. The maximum and minimum graph degrees are denoted by Δ and δ, respectively. The vertex connectivity of G denoted by $\kappa(G)$ and edge connectivity of G denoted by $\lambda(G)$ are the minimum number of vertices and edges, respectively, that need to be removed from a connected graph to make it disconnected. The distance between any two vertices u and v is the length of any shortest path connecting them. The distance to the vertex farthest away from a vertex u is its eccentricity $\epsilon(u)$. The lowest and greatest eccentricities are known as the radius $\mathfrak{R}(G)$ and diameter $\mathfrak{D}(G)$ of a graph, respectively. The central vertex is one with $\epsilon(u) = \mathfrak{R}(G)$. A vertex is diametrical or peripheral if $\epsilon(u) = \mathfrak{D}(G)$. A graph is self-centered if all its vertices are central. An Euler graph is a graph that contains an Eulerian circuit. A graph is Eulerian if and only if the degree of each vertex is even.

Definition 1 [4]. *A k-partite graph is one whose vertex set can be partitioned into k subsets, or parts so that no edge has both ends in the same part. A k-partite graph is complete if any two vertices in different parts are adjacent.*

Theorem 1 [4]. *A nontrivial connected graph G is Eulerian if and only if every vertex of G has an even degree.*

Definition 2 [3]. *The antipodal graph of a graph G denoted by $A(G)$ is the graph on the same vertices as G, two vertices being adjacent if the distance between is equal to the diameter of G. A graph is said to be antipodal if it is the antipodal graph $A(H)$ of some graph H.*

Definition 3 [15]. *A graph is called super-edge connected if every minimum edge cut consists of edges incident with a vertex of minimum degree.*

Definition 4 [7]. *The restricted edge connectivity $\lambda'(G)$ is the minimum cardinality of an edge cut S in a graph G with the property that $G - S$ contains no isolated vertices.*

Definition 5 [11]. *A cluster is an n-vertex graph with maximum adjacency between the vertices. A clustered set is the collection of all clusters that all have the same degree and there is no adjacency between the clusters.*

3. Properties of clustered graphs

Let $H = (V, E)$ be a graph, and let R, S, and T be clustered sets having r, s, and t clusters of maximum adjacency, respectively. Let vertices of the clustered sets R, S, and T of the clustered graph H be

$$\begin{aligned}V(R) = \{&A_{11}, A_{12}, A_{13}, \dots, A_{1t}, A_{1(t+1)}, \dots, A_{1(t+s)}, A_{21}, A_{22}, \dots, A_{2t}, A_{2(t+1)}, \dots,\\ &A_{2(t+s)}, \dots, A_{r1}, A_{r2}, A_{r3}, \dots, A_{rt}, A_{r(t+1)}, \dots, A_{r(t+s)}\};\\ V(S) = \{&B_{11}, B_{12}, B_{13}, \dots, B_{1t}, B_{1(t+1)}, \dots, B_{1(t+r)}, B_{21}, B_{22}, \dots, B_{2t}, B_{2(t+1)}, \dots,\\ &B_{2(t+r)}, \dots, B_{s1}, B_{s2}, B_{s3}, \dots, B_{st}, B_{s(t+1)}, \dots, B_{s(t+r)}\};\\ V(T) = \{&C_{11}, C_{12}, C_{13}, \dots, C_{1s}, C_{1(s+1)}, \dots, C_{1(s+r)}, C_{21}, C_{22}, \dots, C_{2s}, C_{2(s+1)}, \dots,\\ &C_{2(s+r)}, \dots, C_{t1}, C_{t2}, C_{t3}, \dots, C_{ts}, C_{t(s+1)}, \dots, C_{t(s+r)}\}.\end{aligned}$$

Then, $H = CL(r, s, t)$ is said to be a clustered graph if its vertex set V can be partitioned into three nonempty disjoint subsets $V(R)$, $V(S)$, and $V(T)$ of vertices of the clustered sets R, S, and T preserving the adjacency relation $(A_{i(t+j)}, B_{j(t+i)})$, (B_{jk}, C_{kj}), and $(C_{k(s+i)}, A_{ik})$ for $1 \le i \le r$, $1 \le j \le s$, and $1 \le k \le t$ as the complete tripartite graph $K_{r,s,t}$.

Now, the clustered graph $CL(r, s, t)$ is constructed as in Figure 1. The number of vertices n and the number edges m of the clustered graph are $2(rs + rt + st)$ and

$$\frac{1}{2}[r^2 s + r^2 t + s^2 r + s^2 t + t^2 r + t^2 s + 6rst],$$

respectively [11].

Example 1. Figure 2 shows the clustered graph $CL(3, 2, 2)$ as an example.

Theorem 2. *The connectivity number $\kappa(H)$ of a clustered graph H is $s + t$.*

P r o o f. By Whitney's inequality

$$\begin{aligned}&\kappa(H) \le \lambda(H) \le \delta(H);\\ \textit{i.e.,}\quad &\kappa(H) \le \lambda(H) \le s + t.\end{aligned}$$

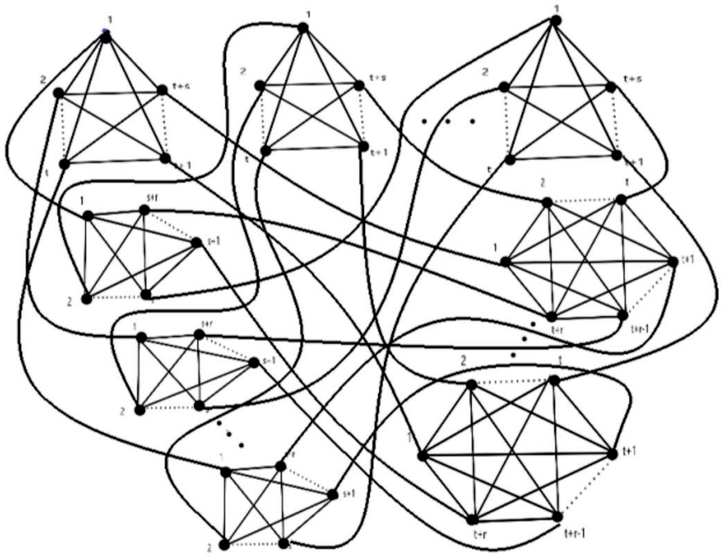


Figure 1. A clustered graph $CL(r, s, t)$.

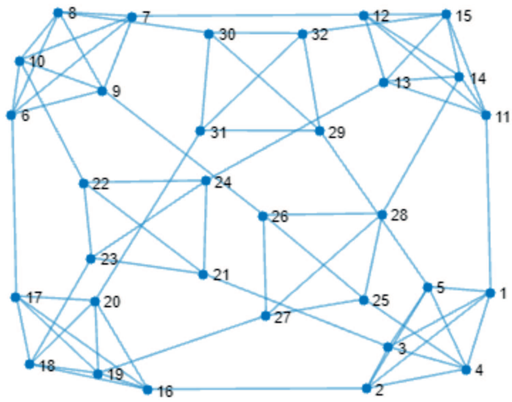


Figure 2. The clustered graph $CL(3, 2, 2)$.

Since the minimal degree is $s + t$, each cluster is adjacent to at least $s + t$ clusters, according to the adjacency criterion. As a result, detaching a cluster requires at least $s + t$ vertices. Therefore,

$$\kappa(H) = s + t.$$

□

Corollary 1. *The edge connectivity number $\lambda(H)$ of a clustered graph H is $s + t$.*

Theorem 3. *If $H(V, E)$ is a clustered graph and S is a subset of V, then*

$$2 \leq c(H - S) \leq |S|.$$

In other words, there are exactly two components in $(H - S)$ if $|S| = \delta(H)$.

P r o o f. The connectivity number of a clustered graph H is $s + t = \delta(H)$. Removal of $(s + t)$ vertices (or edges) disconnects only one cluster from the graph. Hence, by the adjacency condition of a clustered graph, the number of components is two. □

Corollary 2. *A clustered graph is a restricted edge- (λ')-connected graph.*

P r o o f. By Theorem 3, the number of components in $H - S$ is 2. Hence, it is λ'-connected. □

Corollary 3. *A clustered graph is super-edge connected.*

P r o o f. The clusters in the clustered set R contain vertices of minimum degree. By the structure of the clustered graph, the edges joining the vertices of each cluster in the clustered set R are edge cuts. Each edge in the edge cut is incident with a minimum degree vertex. Hence, the graph is super-edge connected. □

Theorem 4. *If each cluster in the clustered sets has minimum adjacency of all vertices of the clustered graph, then the minimum number of edges is* $[(rs + st + rt)(s + t)]$.

P r o o f. In a clustered graph H, we have

$$\kappa(H) = s + t \quad \text{and} \quad \kappa(H) \leq \delta(H).$$

Then the graph must have $(n/2)(s + t)$ edges, i.e., $[(rs + st + rt)(s + t)]$ edges. □

Theorem 5. *The distance between vertices in a clustered graph H is at most 4.*

P r o o f. For $1 \leq i \leq r$, $1 \leq l \leq s$, and $1 \leq k \leq t$, let A_i, B_l, and C_k be the clusters in the clustered sets R, S, and T, respectively:

$$\begin{aligned} R &= \{A_1, A_2, A_3, \ldots, A_i, \ldots, A_r\}, \\ S &= \{B_1, B_2, B_3, \ldots, B_l, \ldots, B_s\}, \\ T &= \{C_1, C_2, C_3, \ldots, C_k, \ldots, C_t\}. \end{aligned}$$

Vertices in each cluster are adjacent to each other. Any cluster in one of the clustered sets R, S, and T is adjacent to every cluster in the other clustered sets.

Consider a cluster A_i in the clustered set R. There are $(s + t)$ adjacent vertices in Ai, where s vertices are adjacent to s-clusters in S and t vertices are adjacent to t-clusters in T. Take a vertex v in A_i. Clearly, v is adjacent to a cluster B_l in S or C_k in T. Consequently, each vertex in B_l or C_k may be reached from v by a path of length at most 2. Furthermore, the path length grows by one from each vertex in A_i other than v. Similarly, the distance between a vertex in one clustered set and any vertex in another clustered set is at most three.

Now, consider vertices u in A_i and v in A_j with $i \neq j$. There is no adjacency between A_i and A_j for any $1 \leq i, j \leq r$. The path from u to v must pass through any cluster in S or T. The length of the $u - v$ path is 3 or 4.

Hence, the distance between vertices in the clustered graph is at most 4. □

Corollary 4. *In a clustered graph H, the radius is*

$$\mathfrak{R} = \begin{cases} 4, & r \geq s \geq t > 1, \\ 3, & \text{otherwise}, \end{cases}$$

and the diameter is

$$\mathfrak{D} = \begin{cases} 3, & r = s = t = 1, \\ 4, & \text{otherwise}. \end{cases}$$

Corollary 5. (i) *Clustered graph H is self-centered graph for either $r \geq s \geq t > 1$ or $r = s = t = 1$.*

(ii) *Complement of the clustered graph has diameter 2.*

Corollary 6. *In clustered graph antipodal vertices lies in the same clustered set.*

P r o o f. The diameter of clustered graph is 4. If v is a antipodal vertex of u then $d(u, v) = 4$. Let $u \in A_i$ in R then by Theorem 5, $v \in A_j$ in R for some j with $i \neq j$ such that $d(u, v) = 4$. Similarly it holds for the vertices in the clustered sets S and T.

Theorem 6. *The antipodal graph $A(H)$ of a clustered graph is disconnected, and the components are balanced partite graphs or isolated vertices.*

P r o o f. Case (i): $r = s = t = 1$. In this case, the clustered graph is isomorphic to C_6 and has a unique antipodal point for each vertex. Hence, the antipodal graph is disconnected.

Case (ii): $r \geq s \geq t > 1$. It is obvious that the clustered sets R, S, and T each have several clusters. According to Corollary 6, the antipodal vertex is in the same clustered set. Except for one vertex in each A_j, all vertices of the clusters A_j are the antipodal points of a vertex v in A_i, where $j \neq i$. Consequently, none of vertices in A_i is adjacent to the vertex v. It results in an r-partite graph, where r is the number of clusters in R. A similar argument demonstrates that the antipodal points of vertices in the clustered sets S and T form s-partite and t-partite graphs.

Case(iii): $r \geq s \geq t$ and $t = 1$. If $t = 1$, then the clustered set T contains only one cluster and its vertices have eccentricity 3. If $s = t = 1$, then both S and T contain one cluster each with vertices of eccentricity 3. Vertices in the clustered set R have eccentricity 4. The antipodal points does not exist for vertices with eccentricity 3. Since $V(A(H)) = V(H)$, the antipodal graph has isolated points. □

Theorem 7. *A clustered graph H is Eulerian if and only if r, s, and t are either all odd or all even.*

P r o o f. In a clustered graph, the possible degrees of vertices are $r + s$, $r + t$, and $s + t$. It is known that the sum of two numbers is even if both are either odd or even. According to the Eulerian criterion, the degree of all vertices is even if and only if r, s, and t are either all odd or all even. □

Corollary 7. *A clustered graph H is Eulerian if and only if its underlying complete tripartite graph is Eulerian.*

Theorem 8. *The complement $\bar{H}$ of a clustered graph H is not a Eulerian graph.*

P r o o f. The number of vertices in a clustered graph is $2(rs + rt + st)$, which is even. The degree of a vertex v in the complement graph is $\bar{d_v} = n - 1 - d_v$. It is clear that $\bar{H}$ is not Eulerian. □

Theorem 9. *If r, s, and t are neither all odd nor all even, then a spanning Eulerian subgraph exists.*

P r o o f. As the values of r, s, and t are neither all odd nor all even, two cases arise.

Case (i): there are *two odd and one even values.* Let r and s be odd and t be even. Let T be the clustered set containing clusters of degree $s + r$. Similarly, let the clustered sets S and R contain clusters of degree $t + r$ and $t + s$, respectively. Remove the edges adjacent with the clustered sets S

and R that lead to the degrees $t+r-1$ and $t+s-1$ of those adjacent vertices. There are t edges that arise from T to S and T to R, respectively. The vertices of the cluster B_1 in the clustered set S that are connected to T are $\{B_{11}, B_{12}, \dots, B_{1t}\}$, and we remove $t/2$ edges between the vertices alternatively. Continue this process for the remaining $s-1$ clusters in the clustered set S. Extend the same process to the clustered set R. Finally, the degrees of vertices are $s+r$, $t+r-1$, and $t+s-1$, which are even. Hence, the Eulerian spanning subgraph is derived.

Case (ii): there are *one odd and two even values*. Let r and s be even and t be odd. Let T, S, and R be clustered sets containing clusters of degree $s+r$, $t+r$, and $s+t$, respectively. The vertices $B_{i(t+j)}$, $i = 1, 2, \dots, s$, $j = 1, 2, \dots, r$, in the clustered set S are adjacent to the vertices $A_{j(t+i)}$ in the clustered set R. Remove the edges $A_{i(t+i)}B_{i(t+i)}$, $i = 1, 2, \dots, (s-1)$, and the edges $A_{i(t+s)}B_{s(t+i)}$, $i = s, (s+1), \dots, r$. The vertices B_{ij} and A_{kj}, $i = 1, 2, \dots, s$, $j = 1, 2, \dots, t$, $k = 1, 2, \dots, r$, are connected to the clustered set X. Remove the edges between these vertices (within the cluster) alternatively. The degrees of the vertices are $s+r$, $t+r-1$, and $t+s-1$, which are even. The proof is complete. ☐

Corollary 8. *A clustered graph H is a super-Eulerian graph for r, s, and t neither all odd nor all even.*

Theorem 10. *The chromatic number of a clustered graph is $s+r$.*

P r o o f. The maximum degree in a clustered graph is $s+r$. The bound for the chromatic number of a graph is $\Delta \le \chi \le \Delta + 1$. Let the set of colors be $\{\alpha_1, \alpha_2, \dots, \alpha_i\}$, $s+r \le i \le s+r+1$.

Define a color function f as

$$\begin{array}{lll} f(C_{ij}) = \alpha_j, & 1 \le i \le t, \quad 1 \le j \le s+r, \\ f(A_{ij}) = \alpha_j, & 1 \le i \le r, \quad 1 \le j \le t+s, \\ f(B_{ij}) = \alpha_{(i+j)}, & (i+j) \mod (r+s), \quad 1 \le i \le s, \quad 1 \le j \le t+r. \end{array}$$

Clearly, the range of f is $\{1, 2, 3, \dots, r+s\}$. Hence, the theorem is proved. ☐

Theorem 11. *The chromatic index of the clustered graph is $s+r$.*

P r o o f. By Vizing's theorem,

$$\chi'(H) \le 1 + \Delta(H).$$

Clearly, $\Delta(H) = r+s$; i.e.,

$$\chi'(H) \le 1 + r + s.$$

To prove the claim, it is enough to prove the existence of a function from the edge set of H to the color set $\mathcal{C}$, and $|\mathcal{C}| = r+s$. The edges can be categorized into two types by the structure of a clustered graph. One is within the clusters and another is between the clustered sets. Considering edges within the cluster, we obtain a complete graph. Moreover, the clusters in the clustered sets R, S, and T are of order $t+s$, $t+r$, and $s+r$, respectively. It is well known that

$$\chi'(K_n) = \begin{cases} \Delta, & n \text{ is even}, \\ \Delta + 1, & n \text{ is odd}. \end{cases}$$

Case (i): $\Delta(H)$ is even. If $\Delta(H) = s+r$ is even, then the edges within the clustered set T are assigned $r+s-1$ colors. It is clear that the degrees of the clustered sets S and R are less than or equal to the degree of the clustered set T. Therefore, all the edges within the clusters of R, S, and T are assigned at most $r+s-1$ colors. The edges between the clustered sets are now assigned $(r+s)$th color. Hence, $\chi'(H) = r+s$.

Case (ii): $\Delta(H)$ is odd. Since $r+s$ is odd, the matching is not perfect and $r+s$ colors are needed. Let $\mathcal{C} = \{\alpha_1, \alpha_2, \ldots, \alpha_{s+r}\}$ be the color set, and let E be the edge set of the clustered graph. To assign colors between the clustered sets, define a function $f\colon E \to C$ as

$$\begin{array}{lll} f(C_{ij}B_{ji}) = \alpha_{p+i+j-2}, & (p+i+j-2) \bmod (r+s), & 1 \le i \le t, \quad 1 \le j \le s, \\ f(C_{i(s+j)}A_{ji}) = \alpha_{p+s+j+i-2}, & (p+s+j+i-2) \bmod (r+s), & 1 \le i \le t, \quad 1 \le j \le r, \\ f(B_{i(t+j)}A_{j(t+i)}) = \alpha_{p+t+i+j-2}, & (p+t+i+j-2) \bmod (r+s), & 1 \le i \le s, \quad 1 \le j \le r, \end{array}$$

where $p = (r+s+1)/2$.

In the clustered set T, edges adjacent to the clustered set R are assigned r colors, and s colors assign the edges adjacent to S with repetition of t clusters. There is a maximal matching of size $(s+r-1)/2$, and each cluster of T has $(r+s)(r+s-1)/2$ edges. Therefore, each cluster in the clustered set T is assigned $r+s$ colors. Since there is no adjacency between the clusters in the same clustered set, we assign the same $r+s$ colors to the clusters. Moreover, the clusters in the clustered set R and S are assigned at most $r+s$ colors. Thus, all the edges of H can be assigned at most $r+s$ colors. Hence, $\chi'(H) = r+s$. □

4. Conclusion

The concept of conditional connectivity is introduced in response to the shortcomings of the traditional connectivity measure by requiring some property for disconnected components. Similarly, edge connectivity was created by specifying certain measures of a network's robustness. Certain properties of disconnected components are required in applications where parallel algorithms can run on subnetworks with a given topology. Euler graph has reached the pinnacle of achievement in numerous circumstances arising in computer science, physical science, communication science, economics, and many other fields. It can be used to represent almost any issue involving discrete arrangements of objects where the focus is on the relationships between the objects rather than their internal properties. This paper discussed graph properties of clustered graphs like connectivity, Eulerian property, and chromaticity. Further properties of the clustered graphs will be incorporated in the future.

REFERENCES

1. Abd-El-Barr M. Topological network design: A survey. *J. Netw. Comp. Appl.*, 2009. Vol. 32, No. 3. P. 501–509. DOI: 10.1016/j.jnca.2008.12.001
2. Al-Kuwaiti M., Kyriakopoulos N., Hussein S. A comparative analysis of network dependability, fault-tolerance, reliability, security, and survivability. *IEEE Communications Surveys & Tutorials*, 2009. Vol. 11, No. 2. P. 106–124. DOI: 10.1109/SURV.2009.090208
3. Aravamudhan R., Rajendran B. On antipodal graphs. *Discrete Math.*, 1984. Vol. 49, No. 2. P. 193–195. DOI: 10.1016/0012-365X(84)90117-1
4. Chartrand G., Lesniak L., Zhang P. *Graphs & Digraphs*, 5rd ed. New York: Chapman & Hall, 2010. 598 p. DOI: 10.1201/b14892
5. Chen Y.-C., Tan J. J. M., Hsu L.-H., Kao S.-S. Super-connectivity and super-edge-connectivity for some interconnection networks. *Appl. Math. Comput.*, 2003. Vol. 140, No. 2–3. P. 245–254. DOI: 10.1016/S0096-3003(02)00223-0
6. Elghazel W., Bahi J., Guyeux C., Hakem M., Medjaher K., Zerhouni N. Dependability of wireless sensor networks for industrial prognostics and health management. *Compu. Ind.*, 2015. Vol. 68. P. 1–15. DOI: 10.1016/j.compind.2014.10.004

7. Esfahanian A.-H., Hakimi S. L. On computing a conditional edge-connectivity of a graph. *Inform. Process. Lett.*, 1988. Vol. 27, No. 4. P. 195–199. DOI: 10.1016/0020-0190(88)90025-7
8. Fàbrega J., Fiol M. A. On the extraconnectivity of graphs. *Discrete Math.*, 1996. P. 155, No. 1–3. P. 49–57. DOI: 10.1016/0012-365X(94)00369-T
9. Harary F. Conditional connectivity. *Networks*, 1983. Vol. 13, No. 3. P. 347–357. DOI: 10.1002/net.3230130303
10. Guo L., Liu R., Guo X. Super connectivity and super edge connectivity of the Mycielskian of a graph. *Graphs Combin.*, 2012. Vol. 28, No. 2. P. 143–147. DOI: 10.1007/s00373-011-1032-3
11. Gurunathan S., Yogalakshmi T. Statistical analysis on the topological indices of clustered graphs. In: *Proc. Int. Conf. on Paradigms of Communication, Computing and Data Sciences. Algorithms for Intelligent Systems. Dua M., Jain A.K., Yadav A., Kumar N., Siarry P. (eds.).* Singapore: Springer, 2022. P. 379–388. DOI: 10.1007/978-981-16-5747-4_33
12. Gurunathan S., Yogalakshmi T., Balasubramanian K. Topological characterization of statistically clustered networks for molecular similarity analysis. *J. Math. Chem.*, 2023. Vol. 61. P. 859–876. DOI: 10.1007/s10910-022-01438-4
13. Latifi S., Hegde M., Naraghi-Pour M. Conditional connectivity measures for large multiprocessor systems. *IEEE Trans. Comput.*, 1994. Vol. 43, No. 2. P. 218–222. DOI: 10.1109/12.262126
14. Li X., Mao Y. *A Survey on the Generalized Connectivity of Graphs.* 2012. 51 p. arXiv:1207.1838v10 [math.CO]
15. Lü M., Chen G.-L., Xu J.-M. On super edge-connectivity of Cartesian product graphs. *Networks*, 2007. Vol. 49, No. 2. P. 152–157. DOI: 10.1002/net.20149
16. Shen C., Liu Y. A tripartite clustering analysis on microRNA, gene and disease model. *J. Bioinf. Comp. Biol.*, 2012. Vol. 10, No. 1. Art. no. 1240007. DOI: 10.1142/S0219720012400070
17. Ustunbas Y., Öguducu S. G. A recommendation model for social resource sharing systems based on tripartite graph clustering. In: *2011 European Intelligence and Security Informatics Conference, 12–14 September 2011, Athens, Greece.* IEEE Xplore, 2011. P. 378–381. DOI: 10.1109/EISIC.2011.56
18. Wang M., Li Q. Conditional edge connectivity properties, reliability comparisons and transitivity of graphs. *Discrete Math.*, 2002 Vol. 258, No. 1–3. P. 205–214. DOI: 10.1016/S0012-365X(02)00299-6
19. Xu J.-M., Wang J.-W., Wang W.-W. On super and restricted connectivity of some interconnection networks. *Ars Combin.*, 2010. Vol. 94, No. 6. P. 25–32.

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 69–80

DOI: 10.15826/umj.2024.2.007

$\mathcal{I}$-STATISTICAL CONVERGENCE OF COMPLEX UNCERTAIN SEQUENCES IN MEASURE

Amit Halder[†], **Shyamal Debnath**[††*]

Department of Mathematics, Tripura University (A Central University), Suryamaninagar-799022, Agartala, India

[†]mit2905halder@gmail.com, [††]shyamalnitamath@gmail.com

Abstract: The main aim of this paper is to present and explore some of properties of the concept of $\mathcal{I}$-statistical convergence in measure of complex uncertain sequences. Furthermore, we introduce the concept of $\mathcal{I}$-statistical Cauchy sequence in measure and study the relationships between different types of convergencies. We observe that, in complex uncertain space, every $\mathcal{I}$-statistically convergent sequence in measure is $\mathcal{I}$-statistically Cauchy sequence in measure, but the converse is not necessarily true.

Keywords: $\mathcal{I}$-convergence, $\mathcal{I}$-statistical convergence, Uncertainty theory, Complex uncertain variable.

1. Introduction

In the real world, there are different kinds of uncertainty. So, it makes perfect sense to investigate the behavior of uncertain phenomena. To address some aspects of this uncertain phenomena, Liu [12] introduced initially the uncertainty theory in 2007. After that, it has been studied in various fields of mathematics like calculus, set theory, graph theory, sequence and series, etc. In [12] Liu initially proposed the idea of uncertain variables as a functions from measurable space to the set of real numbers ($\mathbb{R}$). Peng [15] later expanded it to include complex uncertain variables.

In the fundamental theory of mathematics, the significance of sequence convergence is highly pivotal which is also one of the most important fields of mathematics. Furthermore, one of the most important aspects of uncertainty theory is the convergence of uncertain variable sequences. For the first time in uncertainty theory, Liu [12] established several convergence notions of uncertain variable sequences, such as convergence almost surely, convergence in measure, convergence in distribution, and convergence in mean.

Following that, by using complex uncertain variables, Chen et al. [1] introduced the concept of convergence of complex uncertain sequences and then numerous researchers have subsequently expanded this idea, including Saha et al. [17], Debnath and Das [2], and You and Yan [23]. The concept of Cauchy convergence in measure and in mean was recently presented by Wu and Xia [24].

On the other hand, in 1951, Fast [8] and Steinhaus[21] extended the concepts of convergence of a real sequence to statistical convergence independently and after that, it was studied by Fridy [9] and many other famous researchers. Later Kostyrko et al. [11] introduced a new concept of convergence namely $\mathcal{I}$-convergence, which is a generalization of statistical convergence.

Savas and Das [19] further expanded the notion of statistical convergence and $\mathcal{I}$-convergence to include $\mathcal{I}$-statistical convergence. This extension prompted further explorations in the field by researchers such as Savas and Das [20], Debnath and Debnath [5], Debnath and Rakshit [6], Mursaleen et al. [13], Savas et al. [18], Esi et al. [7], and numerous others.

*Corresponding author

Tripathy and Nath [22] introduced the concept of statistical convergence for complex uncertain sequences within the framework of uncertainty theory. Then many other researchers like Nath and Tripathy [14], Roy et al. [16], Debnath and Das [3, 4], and Kisi [10] have successfully applied the concept of generalized convergence of sequences in uncertainty theory.

Inspired by the above works, in this paper, we introduce the concepts of $\mathcal{I}$-statistical convergence in measure of complex uncertain sequences and study some of its properties. We also introduce the notion of $\mathcal{I}$-statistical Cauchy sequence in measure and identifying the relationship between $\mathcal{I}$-statistical convergence in measure and $\mathcal{I}$-statistical Cauchy sequence in measure.

2. Definitions and preliminaries

The generalized convergence notions and the theory of uncertainty, which will be utilized throughout the study, are defined and findings are presented in this section.

Definition 1 [11]. *Consider a non-empty set S. An ideal on S is defined as a family of subsets $\mathcal{I}$ that satisfies the following conditions:*

(*i*) *The empty set, ϕ, belongs to $\mathcal{I}$.*
(*ii*) *For any $U, V \in \mathcal{I}$, the union of U and V, denoted as $U \cup V$, is also in $\mathcal{I}$.*
(*iii*) *For any $U \in \mathcal{I}$ and any subset $V \subset U$, V is a member of $\mathcal{I}$.*

An ideal $\mathcal{I}$ is called non-trivial if $\mathcal{I} \neq \{\phi\}$ and $S \notin \mathcal{I}$. A non-trivial ideal $\mathcal{I}$ is called an admissible ideal in S if and only if $\big\{\{s\} : s \in S\big\} \subset \mathcal{I}$.

Example 1. (i) $\mathcal{I}_f := \{$The set of all finite subsets of $\mathbb{N}$ forms a non-trivial admissible ideal$\}$.
(ii) $\mathcal{I}_d := \{$The set of all subsets of $\mathbb{N}$ whose natural density is zero forms a non-trivial admissible ideal$\}$.

Definition 2 [11]. *Consider a non-empty set S. A family of subsets $\mathcal{F}$, which is a subset of the power set $P(S)$, is called a filter on S if and only if the following conditions are satisfied:*

(*i*) *The empty set ϕ is not a member of $\mathcal{F}$.*
(*ii*) *For any subsets U and V in $\mathcal{F}$, their intersection $U \cap V$ is also included in $\mathcal{F}$.*
(*iii*) *If U is a member of $\mathcal{F}$ and V is a superset of U, then V is also a member of $\mathcal{F}$.*

Now, let $\mathcal{I}$ be an admissible ideal. The filter $\mathcal{F}(\mathcal{I})$ associated with the ideal $\mathcal{I}$ is defined as

$$\mathcal{F}(\mathcal{I}) = \{S \setminus U : U \in \mathcal{I}\}.$$

Definition 3 [9]. *A real sequence (x_m) is said to be statistically convergent to $\ell \in \mathbb{R}$ provided that for each $\varepsilon > 0$ we have*

$$\lim_{m\to\infty} \frac{1}{m}\big|\{k \le m : |x_k - \ell| \ge \varepsilon\}\big| = 0, \quad m \in \mathbb{N}.$$

Definition 4 [11]. *A real sequence (x_m) is said to be $\mathcal{I}$-convergent to $\ell \in \mathbb{R}$, if for every $\varepsilon > 0$, we have*

$$\{m \in \mathbb{N} : |x_m - \ell| \ge \varepsilon\} \in \mathcal{I}.$$

The usual convergence of sequences is a special case of $\mathcal{I}$-convergence ($\mathcal{I}$=$\mathcal{I}_f$-the ideal of all fimite subsets of $\mathbb{N}$). The statistical convergence of sequences is also a special case of $\mathcal{I}$-convergence. In this case,

$$\mathcal{I} = \mathcal{I}_d = \Big\{A \subseteq \mathbb{N} : \lim_{m\to\infty} \frac{|\, A \cap \{1, 2, ..., m\}\,|}{m} = 0\Big\},$$

where $|A|$ is the cardinality of the set A.

Definition 5 [18]. *A real sequence (x_m) is said to be $\mathcal{I}$-statistically convergent to $\ell \in \mathbb{R}$, if for every $\varepsilon > 0$, and every $\delta > 0$,*

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : |x_k - \ell| \geq \varepsilon\}\big| \geq \delta\Big\} \in \mathcal{I}.$$

For $\mathcal{I} = \mathcal{I}_f$, $\mathcal{I}$-statistical convergence coincides with statistical convergence.

Definition 6 [12]. *Let $\mathcal{P}$ be a σ-algebra on a non-empty set Υ. If the set function $\mathcal{X}$ on Υ satisfies the following axioms, it is referred to be an uncertain measure*:

(*i*) *The first axiom, which deals with normality, is*: $\mathcal{X}\{\Upsilon\} = 1$;

(*ii*) *The second, which deals with duality, is*: $\mathcal{X}\{\Xi\} + \mathcal{X}\{\Xi^c\} = 1$ *for any* $\Xi \in \mathcal{P}$;

(*iii*) *The third, which deals with subadditivity, is*: *for every countable sequence of* $\{\Xi_m\} \in \mathcal{P}$,

$$\mathcal{X}\Big\{\bigcup_{m=1}^{\infty} \Xi_m\Big\} \leq \sum_{m=1}^{\infty} \mathcal{X}\{\Xi_m\}.$$

An uncertainty space is denoted by the triplet $(\Upsilon, \mathcal{P}, \mathcal{X})$, and an event is denoted by each member Ξ in $\mathcal{P}$. For an uncertain measure of a compound event, Liu defines a product uncertain measure as follows:

$$\mathcal{X}\Big\{\prod_{r=1}^{\infty} \Xi_r\Big\} = \bigwedge_{r=1}^{\infty} \mathcal{X}\{\Xi_r\}.$$

Definition 7 [15]. *A complex uncertain variable is represented by a variable ζ in the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ if and only if both its real part ξ and imaginary part η are uncertain variables. Here, ξ and η correspond to the real and imaginary components of the complex variable $\zeta = \xi + i\eta$, respectively.*

Definition 8 [1]. *A complex uncertain sequence (ζ_m) is said to be convergent in measure to ζ if for every $\varepsilon > 0$,*

$$\lim_{m\to\infty} \mathcal{X}\big\{\|\zeta_m(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big\} = 0.$$

Definition 9 [22]. *A complex uncertain sequence (ζ_m) is said to be statistically convergent in measure to ζ if for any given positive values of ε, δ, we have*

$$\lim_{m\to\infty} \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}\big(\|\zeta_k(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\big\}\big| = 0$$

and we write $\zeta_m \xrightarrow{S^{M_s}} \zeta$.

Definition 10. *A complex uncertain sequence (ζ_m) is said to be $\mathcal{I}$-convergent in measure to ζ if for any given positive values of ε, δ, we have*

$$\big\{m \in \mathbb{N} : \mathcal{X}\big(\|\zeta_m(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\big\} \in \mathcal{I}.$$

and we write $\zeta_m \xrightarrow{M_s(\mathcal{I})} \zeta$.

In this paper, $\mathcal{I}$ is taken to be an admissible ideal.

3. Main results

Definition 11. *A complex uncertain sequence (ζ_m) is considered to be $\mathcal{I}$-statistically convergent in measure to ζ if, for any given positive values of $\varepsilon, \delta, \upsilon$, there exists a set satisfying the condition*

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}\big(\|\zeta_k(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\big\}\big| \geq \upsilon\Big\} \in \mathcal{I}.$$

This is denoted as $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$.

Example 2. Consider the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ to be $\{\varrho_1, \varrho_2, \cdots\}$ with power set and $\mathcal{X}\{\Upsilon\} = 1$, $\mathcal{X}\{\phi\} = 0$ and

$$\mathcal{X}\{\Xi\} = \begin{cases} \sup\limits_{\varrho_m \in \Xi} \dfrac{3}{(2m+1)}, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi} \dfrac{3}{(2m+1)} < \dfrac{1}{2}, \\ 1 - \sup\limits_{\varrho_m \in \Xi^c} \dfrac{3}{(2m+1)}, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi^c} \dfrac{3}{(2m+1)} < \dfrac{1}{2}, \quad \text{for} \quad m = 1, 2, 3, \dots, \\ \dfrac{1}{2}, & \text{otherwise} \end{cases}$$

and $\zeta_m(\varrho)$ (the complex uncertain variables) are defined by

$$\zeta_m(\varrho) = \begin{cases} mi, & \text{if} \quad \varrho = \varrho_m, \quad m = 1, 2, 3, \dots, \\ 0, & \text{otherwise} \end{cases}$$

and $\zeta \equiv 0$. Take $\mathcal{I}$=$\mathcal{I}_d$.

For $m \geq 3$ and small positive values of $\varepsilon, \delta, \upsilon$ we get,

$$\begin{aligned} &\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\} \\ &= \Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\varrho : \|\zeta_k(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\} \\ &= \Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\{\varrho_k\} \geq \delta\}\big| \geq \upsilon\Big\} \\ &= \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \frac{3}{2k+1} \geq \delta\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I}. \end{aligned}$$

Therefore the sequence (ζ_m) is $\mathcal{I}$-statistically convergent in measure to ζ.

Theorem 1. *If $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$ and $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta^*$ then $\mathcal{X}\{\zeta = \zeta^*\} = 1$.*

P r o o f. Let $\varepsilon, \delta > 0$ and $0 < \upsilon < 1$, then

$$G = \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| < \frac{\upsilon}{3}\Big\} \in \mathcal{F}(\mathcal{I}),$$

and

$$H = \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta^*\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| < \frac{\upsilon}{3}\Big\} \in \mathcal{F}(\mathcal{I}).$$

Since $G \cap H \in \mathcal{F}(\mathcal{I})$ and $\phi \notin \mathcal{F}(\mathcal{I})$ this implies $G \cap H \neq \phi$. Let $m \in G \cap H$. Then

$$\frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| < \frac{\upsilon}{3}$$

and
$$\frac{1}{m}\Big|\Big\{k \le m : \mathfrak{X}\Big(\|\zeta_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big) \ge \frac{\delta}{2}\Big\}\Big| < \frac{\upsilon}{3}.$$
Therefore
$$\frac{1}{m}\Big|\Big\{k \le m : \mathfrak{X}(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}) \ge \frac{\delta}{2} \quad \text{or} \quad \mathfrak{X}(\,\|\,\zeta_k - \zeta^*\,\|\ge \frac{\varepsilon}{2}) \ge \frac{\delta}{2}\Big\}\Big| < \upsilon < 1.$$
Thus there exists some $k \le m$ such that
$$\mathfrak{X}\Big(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}\Big) < \frac{\delta}{2} \quad \text{and} \quad \mathfrak{X}\Big(\|\zeta_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big) < \frac{\delta}{2}.$$
Therefore
$$\mathfrak{X}(\|\zeta - \zeta^*\| \ge \varepsilon) \le \mathfrak{X}\Big(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}\Big) + \mathfrak{X}\Big(\|\zeta_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big) < \delta.$$
Hence we get the result. □

Theorem 2. *Elementary properties are valid*:

(*i*) $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta \iff \zeta_m - \zeta \xrightarrow{S^{M_s}(\mathcal{I})} 0$;

(*ii*) $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta \implies c\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} c\zeta$, *where* $c \in \mathbb{C}$;

(*iii*) $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$ *and* $\zeta^*_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta^* \implies \zeta_m + \zeta^*_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta + \zeta^*$;

(*iv*) $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$ *and* $\zeta^*_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta^* \implies \zeta_m - \zeta^*_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta - \zeta^*$.

P r o o f. Let $\varepsilon, \delta, \upsilon$ be any positive real numbers. For (*i*), (*ii*), the proofs are straight forward and so omitted.

(*iii*) It is obvious from the inequality
$$\mathfrak{X}\Big(\|(\zeta_k + \zeta^*_k) - (\zeta + \zeta^*)\| \ge \varepsilon\Big) \le \mathfrak{X}\Big(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}\Big) + \mathfrak{X}\Big(\|\zeta^*_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big).$$
We have
$$\begin{aligned} &\Big\{k \le m : \mathfrak{X}(\|(\zeta_k - \zeta) + (\zeta^*_k - \zeta^*)\| \ge \varepsilon) \ge \delta\Big\} \\ &\subseteq \Big\{k \le m : \mathfrak{X}\Big(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}\Big) \ge \frac{\delta}{2}\Big\} \cup \Big\{k \le m : \mathfrak{X}\Big(\|\zeta^*_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big) \ge \frac{\delta}{2}\Big\}. \end{aligned}$$
Therefore
$$\begin{aligned} &\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\big\{k \le m : \mathfrak{X}(\|(\zeta_k + \zeta^*_k) - (\zeta + \zeta^*)\| \ge \varepsilon) \ge \delta\big\}\big| \ge \upsilon\Big\} \\ &\subseteq \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \le m : \mathfrak{X}\Big(\|\zeta_k - \zeta\| \ge \frac{\varepsilon}{2}\Big) \ge \frac{\delta}{2}\Big\}\Big| \ge \frac{\upsilon}{2}\Big\} \\ &\cup \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \le m : \mathfrak{X}\Big(\|\zeta^*_k - \zeta^*\| \ge \frac{\varepsilon}{2}\Big) \ge \frac{\delta}{2}\Big\}\Big| \ge \frac{\upsilon}{2}\Big\} \in \mathcal{I}. \end{aligned}$$
This implies
$$\zeta_m + \zeta^*_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta + \zeta^*.$$

(*iv*) The reason it was left out was because it was equivalent to the proof of (*iii*) above. □

Theorem 3. *If the complex uncertain sequences* (ζ_m), (ζ_m^*) *are* $\mathcal{I}$*-statistically convergent in measure to* ζ *and* ζ^**, respectively, and there exist positive numbers* $p_1, p, q_1,$ *and* q *such that* $p_1 \leq \|\zeta_m\|, \|\zeta\| \leq p$ *and* $q_1 \leq \|\zeta_m^*\|, \|\zeta^*\| \leq q$ *for any* n*, then*

(*i*) $(\zeta_m \zeta_m^*)$ *is* $\mathcal{I}$*-statistically convergent in measure to* $\zeta\zeta^*$.
(*ii*) (ζ_m / ζ_m^*) *is* $\mathcal{I}$*-statistically convergent in measure to* ζ/ζ^*.

P r o o f. Let (ζ_m), (ζ_m^*) are $\mathcal{I}$-statistically convergent in measure to ζ and ζ^*, respectively, where (ζ_m), (ζ_m^*) both are complex uncertain sequences. For $p, q > 0$ and any given positive values of $\varepsilon, \delta, \upsilon$, we obtain

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \ \mathfrak{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2q}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I},$$
$$\Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \ \mathfrak{X}\Big(\|\zeta_k^* - \zeta^*\| \geq \frac{\varepsilon}{2p}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I}.$$

Now

$$\begin{aligned}
\mathfrak{X}\Big(\|\zeta_m\zeta_m^* - \zeta\zeta^*\| \geq \varepsilon\Big) &= \mathfrak{X}\Big(\|\zeta_m\zeta_m^* - \zeta_m\zeta^* + \zeta_m\zeta^* - \zeta\zeta^*\| \geq \varepsilon\Big)\\
&\leq \mathfrak{X}\Big(\|\zeta_m\zeta_m^* - \zeta_m\zeta^*\| \geq \frac{\varepsilon}{2}\Big) + \mathfrak{X}\Big(\|\zeta_m\zeta^* - \zeta\zeta^*\| \geq \frac{\varepsilon}{2}\Big)\\
&\leq \mathfrak{X}\Big(p\|\zeta_m^* - \zeta^*\| \geq \frac{\varepsilon}{2}\Big) + \mathfrak{X}\Big(q||\zeta_m - \zeta|| \geq \frac{\varepsilon}{2}\Big)\\
\implies \mathfrak{X}\Big(\|\zeta_m\zeta_m^* - \zeta\zeta^*\| \geq \varepsilon\Big) &\leq \mathfrak{X}\Big(\|\zeta_m^* - \zeta^*\| \geq \frac{\varepsilon}{2p}\Big) + \mathfrak{X}\Big(||\zeta_m - \zeta|| \geq \frac{\varepsilon}{2q}\Big).
\end{aligned}$$

Then for small number $\delta > 0$,

$$\begin{gathered}
\{k \leq m : \mathfrak{X}(\|\zeta_k\zeta_k^* - \zeta\zeta^*\| \geq \varepsilon) \geq \delta\}\\
\subseteq \Big\{k \leq m : \mathfrak{X}\Big(\|\zeta_k^* - \zeta^*\| \geq \frac{\varepsilon}{2p}\Big) \geq \frac{\delta}{2}\Big\} \cup \Big\{k \leq m : \mathfrak{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2q}\Big) \geq \frac{\delta}{2}\Big\}\\
\implies \frac{1}{m}\big|\{k \leq m : \mathfrak{X}(\|\zeta_k\zeta_k^* - \zeta\zeta^*\| \geq \varepsilon) \geq \delta\}\big|\\
\leq \frac{1}{m}\Big|\Big\{k \leq m : \mathfrak{X}\big(\|\zeta_k^* - \zeta^*\| \geq \frac{\varepsilon}{2p}\big) \geq \frac{\delta}{2}\Big\}\Big| + \frac{1}{m}\Big|\Big\{k \leq m : \mathfrak{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2q}\Big) \geq \frac{\delta}{2}\Big\}\Big|.
\end{gathered}$$

For small number $\upsilon > 0$,

$$\begin{gathered}
\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathfrak{X}(\|\zeta_k\zeta_k^* - \zeta\zeta^*\| \geq \varepsilon) \geq \delta\}\big| \geq \upsilon\Big\}\\
\subseteq \Big\{m \in \mathbb{N} :\leq \frac{1}{m}\Big|\Big\{k \leq m : \mathfrak{X}\Big(\|\zeta_k^* - \zeta^*\| \geq \frac{\varepsilon}{2p}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\}\\
\cup\Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathfrak{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2q}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I}.
\end{gathered}$$

Hence the sequence $(\zeta_m\zeta_m^*)$ converges $\mathcal{I}$-statistical in measure to $\zeta\zeta^*$.

(*ii*) It was left out because it is similar to the (*i*) proof above. □

Theorem 4. *If a sequence* (ζ_m) *is* $\zeta_m \xrightarrow{M_s(\mathcal{I})} \zeta$ *then it is* $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$.

P r o o f. It is evidently true. □

But in general, the converse may not be true.

Example 3. Consider the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ to be $\{\varrho_1, \varrho_2, ...\}$ with power set and $\mathcal{X}\{\Upsilon\} = 1$, $\mathcal{X}\{\phi\} = 0$ and

$$\mathcal{X}\{\Xi\} = \begin{cases} \sup\limits_{\varrho_m \in \Xi} \dfrac{m\beta_m}{2m+1}, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi} \dfrac{m\beta_m}{2m+1} < \dfrac{1}{2}, \\ 1 - \sup\limits_{\varrho_m \in \Xi^c} \dfrac{m\beta_m}{2m+1}, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi^c} \dfrac{m\beta_m}{2m+1} < \dfrac{1}{2}, \\ \dfrac{1}{2}, & \text{otherwise,} \end{cases} \quad \text{for} \quad m = 1, 2, 3, \dots,$$

where

$$\beta_m = \begin{cases} 1, & \text{if} \quad m = k^2, \quad k \in \mathbb{N}, \\ 0, & \text{otherwise,} \end{cases} \quad \text{for} \quad m = 1, 2, 3, \dots.$$

Furthermore, $\zeta_m(\varrho)$ (the complex uncertain variables) are defined by

$$\zeta_m(\varrho) = \begin{cases} (m+1)i, & \text{if} \quad \varrho = \varrho_m, \\ 0, & \text{otherwise,} \end{cases} \quad \text{for} \quad m = 1, 2, 3, \dots,$$

and $\zeta \equiv 0$. Take $\mathcal{I}$=$\mathcal{I}_f$.

For every positive value of ε, we obtain

$$\mathcal{X}\big(\{\varrho \in \Upsilon : \|\zeta_m(\varrho) - \zeta(\varrho)\| \geq \varepsilon\}\big) = \mathcal{X}(\varrho_m) = \frac{m\beta_m}{2m+1}.$$

Then

$$\{m \in \mathbb{N} : \mathcal{X}\big(\|\zeta_m - \zeta\| \geq \varepsilon\big) \geq \delta\} = \Big\{m \in \mathbb{N} : \frac{m\beta_m}{2m+1} \geq \delta\Big\} \notin \mathcal{I}_f.$$

Now

$$\frac{1}{m}\big|\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta\| \geq \varepsilon) \geq \delta\}\big| \leq \frac{\sqrt{m}}{m} = \frac{1}{\sqrt{m}}.$$

Then

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k - \zeta\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\} \subseteq \Big\{m \in \mathbb{N} : \frac{1}{\sqrt{m}} \geq \upsilon\Big\} \in \mathcal{I}_f.$$

Hence the sequence (ζ_m) is not $\mathcal{I}$-convergent in measure to $\zeta \equiv 0$ but it is $\mathcal{I}$-statistically convergent in measure to $\zeta \equiv 0$.

Theorem 5. *For any sequence*(ζ_m)*,* $\zeta_m \xrightarrow{S^{M_s}} \zeta$ *implies* $\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta$*.*

P r o o f. Let $\zeta_m \xrightarrow{S^{M_s}} \zeta$. Then for each $\varepsilon, \delta > 0$

$$\lim_{m \to \infty} \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k - \zeta\| \geq \varepsilon\big) \geq \delta\}\big| = 0.$$

So for every $\upsilon > 0$,

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k - \zeta\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\}$$

is a finite set and since $\mathcal{I}$ is an admissible ideal, it must belong to $\mathcal{I}$. Hence

$$\zeta_m \xrightarrow{S^{M_s}(\mathcal{I})} \zeta.$$

But in general, the converse may not hold. $\square$

Example 4. Let

$$\mathbb{N} = \bigcup_{j=1}^{\infty} D_j,$$

where

$$D_j = \left\{2^{j-1}k : 2 \text{ does not divide } k,\ k \in \mathbb{N}\right\}$$

be the decomposition of $\mathbb{N}$ such that each D_j is infimite and $D_j \cap D_k = \phi$, for $j \neq k$. Let $\mathcal{I}$ be the class of all subsets of $\mathbb{N}$ that can intersect only fimite number of $D_j{}'s$. Then $\mathcal{I}$ is a nontrivial admissible ideal of $\mathbb{N}$ [11].

Now we consider the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ to be $\{\varrho_1, \varrho_2, ...\}$ with power set and $\mathcal{X}\{\Upsilon\} = 1$, $\mathcal{X}\{\phi\} = 0$ and

$$\mathcal{X}\{\Xi\} = \begin{cases} \sup\limits_{\varrho_m \in \Xi} \beta_m, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi} \beta_m < \dfrac{1}{2}, \\ 1 - \sup\limits_{\varrho_m \in \Xi^c} \beta_m, & \text{if} \quad \sup\limits_{\varrho_m \in \Xi^c} \beta_m < \dfrac{1}{2}, \\ \dfrac{1}{2}, & \text{otherwise}, \end{cases}$$

where

$$\beta_m = \frac{1}{j+1}, \quad \text{if} \quad m \in D_j \quad \text{for} \quad m = 1, 2, 3, \ldots.$$

Furthermore, $\zeta_m(\varrho)$ (the complex uncertain variables) are defined by

$$\zeta_m(\varrho) = \begin{cases} (m+1)i, & \text{if} \quad \varrho = \varrho_m, \\ 0, & \text{otherwise}, \end{cases} \quad \text{for} \quad m = 1, 2, 3, \ldots,$$

and $\zeta \equiv 0$.
For $m \in \mathbb{N} \setminus D_1$ and any positive number ε, we have

$$\mathcal{X}\left(\left\{\varrho \in \Upsilon : \|\zeta_m(\varrho) - \zeta(\varrho)\| \geq \varepsilon\right\}\right) = \mathcal{X}(\varrho_m) = \beta_m.$$

Then

$$\lim_{m \to \infty} \frac{1}{m} \left|\left\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta\| \geq \varepsilon) \geq \delta\right\}\right| \neq 0.$$

Thus the sequence (ζ_m) is not statistically convergent in measure to $\zeta \equiv 0$.
Again

$$\left\{m \in \mathbb{N} : \mathcal{X}\left(\|\zeta_m - \zeta\| \geq \varepsilon\right) \geq \delta\right\} = \left\{m \in \mathbb{N} : \beta_m \geq \delta\right\} \in \mathcal{I}.$$

Therefore the sequence (ζ_m) is $\mathcal{I}$-convergent in measure to $\zeta \equiv 0$. By Theorem 4, the sequence (ζ_m) is $\mathcal{I}$-statistically convergent in measure to $\zeta \equiv 0$.

Theorem 6. *(ζ_m) is $\mathcal{I}$-statistically convergent in measure to ζ if each of its subsequences is $\mathcal{I}$-statistically convergent in measure to ζ.*

P r o o f. Assume that (ζ_m) does not $\mathcal{I}$-statistically convergent in measure to ζ. Consequently, there are positive constants $\varepsilon, \delta, \upsilon$ such that

$$A = \Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\} \notin \mathcal{I}.$$

As $\mathcal{I}$ is an admissible ideal, it implies that the set A must be infinite.

Let $A = \{m_1 < m_1 < \cdots < m_k < \cdots\}$. Let $\zeta_k^* = \zeta_{m_k},\ k \in \mathbb{N}$. Then $(\zeta_k^*)_{k\in\mathbb{N}}$ is a subsequence of (ζ_m) which is not $\mathcal{I}$-statistical convergent in measure to ζ, we have got a contradiction.

But in general, the converse may not hold.

Example 5. Consider the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ to be $\{\varrho_1, \varrho_2, \cdots\}$ with power set and $\mathcal{X}\{\Upsilon\} = 1$, $\mathcal{X}\{\phi\} = 0$ and

$$\mathcal{X}\{\Xi\} = \begin{cases} \sup\limits_{\varrho_m\in\Xi} \dfrac{1}{m}, & \text{if} \quad \sup\limits_{\varrho_m\in\Xi} \dfrac{1}{m} < \dfrac{1}{2}, \\ 1 - \sup\limits_{\varrho_m\in\Xi^c} \dfrac{1}{m}, & \text{if} \quad \sup\limits_{\varrho_m\in\Xi^c} \dfrac{1}{m} < \dfrac{1}{2}, \\ \dfrac{1}{2}, & \text{otherwise}, \end{cases} \quad \text{for} \quad m = 1, 2, 3, \ldots,$$

and $\zeta_m(\varrho)$ (the complex uncertain variables) are defined by

$$\zeta_m(\varrho) = \begin{cases} (m+1)i, & \text{if} \quad \varrho = \varrho_{m=k^2}, \quad k \in \mathbb{N}, \\ 0, & \text{otherwise}, \end{cases} \quad \text{for} \quad m = 1, 2, 3, \ldots,$$

and $\zeta \equiv 0$. Take $\mathcal{I}$=$\mathcal{I}_d$.

Clearly, the sequence (ζ_m) is $\mathcal{I}$-statistically convergent in measure to $\zeta \equiv 0$. But the subsequence $(\zeta_{m=k^2})$, $k \in \mathbb{N}$ is not $\mathcal{I}$-statistically convergent in measure to $\zeta \equiv 0$.

Definition 12. *A complex uncertain sequence, denoted as (ζ_m), is called $\mathcal{I}$-statistically Cauchy sequence in measure if, for any given ε and δ (both greater than zero), there exists a natural number m_0 such that, for any $\upsilon > 0$, we have*

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\{k \leq m : \mathcal{X}\big(\|\zeta_k(\varrho) - \zeta_{m_0}(\varrho)\| \geq \varepsilon\big) \geq \delta\}\big| \geq \upsilon\Big\} \in \mathcal{I}.$$

Theorem 7. *A complex uncertain sequence (ζ_m) is $\mathcal{I}$-statistically Cauchy sequence in measure if it is $\mathcal{I}$-statistically convergent in measure to ζ.*

P r o o f. Let the complex uncertain sequence (ζ_m) be $\mathcal{I}$-statistically convergent in measure to ζ. Then for $0 < \upsilon < 1$ and every positive number ε, δ, we have

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I}.$$

Then for $0 < \upsilon < 1$,

$$G = \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| < \upsilon\Big\} \in \mathcal{F}(\mathcal{I}).$$

Since $G \in \mathcal{F}(\mathcal{I})$ and $\phi \notin \mathcal{F}(\mathcal{I})$, so $G \neq \phi$. Let $m \in G$. Then

$$\frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| < \upsilon < 1.$$

So, there exists some $m_0 \leq m$ such that

$$\mathcal{X}\Big(\|\zeta_{m_0} - \zeta\| \geq \frac{\varepsilon}{2}\Big) < \frac{\delta}{2}.$$

We have

$$\big\{k \leq m : \mathcal{X}\big(\|\zeta_k - \zeta_{m_0}\| \geq \varepsilon\big) \geq \delta\big\} \subset \Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}$$

which implies

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\big(\|\zeta_k - \zeta_{m_0}\| \geq \varepsilon\big) \geq \delta\Big\}\Big| \geq \upsilon\Big\}$$
$$\subset \Big\{m \in \mathbb{N} : \frac{1}{m}\Big|\Big\{k \leq m : \mathcal{X}\Big(\|\zeta_k - \zeta\| \geq \frac{\varepsilon}{2}\Big) \geq \frac{\delta}{2}\Big\}\Big| \geq \upsilon\Big\} \in \mathcal{I}.$$

Hence the sequence (ζ_m) is $\mathcal{I}$-statistically Cauchy sequence in measure. □

But in general, the converse may not hold.

Example 6. Consider the uncertainty space $(\Upsilon, \mathcal{P}, \mathcal{X})$ to be equivalent to set of real number $\mathbb{R}$ with $\Xi_m = (m, \infty)$, for $m = 1, 2, 3, \ldots$ and

$$\mathcal{X}\{\Xi\} = \begin{cases} 0, & \text{if } \ \Xi = \phi \text{ or } \Xi \text{ is upper bounded,} \\ \dfrac{1}{2}, & \text{if both } \Xi \text{ and } \Xi^c \text{ are upper unbounded,} \\ 1, & \text{if } \ \Xi = \Upsilon \text{ or } \Xi^c \text{ is upper bounded.} \end{cases}$$

Furthermore, $\zeta_m(\varrho)$ (the complex uncertain variables) are defined by

$$\zeta_m(\varrho) = \begin{cases} i, & \text{if } \ \varrho \in \Xi_m, \\ 0, & \text{if } \ \varrho \notin \Xi_m, \end{cases} \quad \text{for} \quad m = 1, 2, 3, \ldots,$$

and $\zeta \equiv 0$. Take $\mathcal{I}$=$\mathcal{I}_d$. Now

$$\big\{\varrho \in \Upsilon : \|\zeta_m(\varrho) - \zeta_{m_0}(\varrho)\| \geq \varepsilon\big\} = \begin{cases} (m_0, m], & \text{if } \ 0 < \varepsilon \leq 1, \\ \phi, & \text{if } \ \varepsilon > 1. \end{cases}$$
$$\implies \mathcal{X}\big(\big\{\varrho \in \Upsilon : \|\zeta_m(\varrho) - \zeta_{m_0}(\varrho)\| \geq \varepsilon\big\}\big) = 0$$
$$\implies \lim_{m \to \infty} \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta_{m_0}\| \geq \varepsilon) \geq \delta\big\}\big| = 0.$$

Therefore

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta_{m_0}\| \geq \varepsilon) \geq \delta\big\}\big| \geq \upsilon\Big\} \in \mathcal{I}.$$

Again,

$$\mathcal{X}\big(\big\{\varrho \in \Upsilon : \|\zeta_m(\varrho) - \zeta(\varrho)\| \geq \varepsilon\big\}\big) = \mathcal{X}(\Xi_m) = 1.$$

So,

$$\lim_{m \to \infty} \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta\| \geq \varepsilon) \geq \delta\big\}\big| \neq 0.$$

Therefore

$$\Big\{m \in \mathbb{N} : \frac{1}{m}\big|\big\{k \leq m : \mathcal{X}(\|\zeta_k - \zeta\| \geq \varepsilon) \geq \delta\big\}\big| \geq \upsilon\Big\} \notin \mathcal{I}.$$

Hence the sequence (ζ_m) is not $\mathcal{I}$-statistically convergent in measure to $\zeta \equiv 0$ but it is $\mathcal{I}$-statistically Cauchy sequence in measure.

4. Conclusion

This paper mainly contributes to the study of $\mathcal{I}$-statistical convergence in measure of complex uncertain sequences, by establishing some of its properties. Also, we define $\mathcal{I}$-statistical Cauchy sequence in measure and study the relationship among them. It is possible to generalize and apply these concepts and results to future research in this area.

Acknowledgment

The first author wishes to acknowledge the *Council of Scientific and Industrial Research, India* for their fellowships funding granted under the *CSIR-JRF* scheme (File No: 09/0714(11674)/2021-EMR-I) for the preparation of this manuscript.

REFERENCES

1. Chen X., Ning Y., Wang X. Convergence of complex uncertain sequences. *J. Intell. Fuzzy Syst.*, 2016. Vol. 30, No. 6. P. 3357–3366. DOI: 10.3233/IFS-152083
2. Debnath S., Das B. Statistical convergence of order α for complex uncertain sequences. *J. Uncertain Syst.*, 2021. Vol. 14, No. 2. Art. no. 2150012. DOI: 10.1142/S1752890921500124
3. Debnath S., Das B. On λ-statistical convergence of order α for complex uncertain sequences. *Int. J. Gen. Syst.*, 2022. Vol. 52, No. 2. P. 191–202. DOI: 10.1080/03081079.2022.2132490
4. Debnath S., Das B. On rough statistical convergence of complex uncertain sequences. *New Math. Nat. Comput.*, 2023. Vol. 19, No. 1. P. 1–17. DOI: 10.1142/S1793005722500454
5. Debnath S., Debnath J. On $\mathcal{I}$-statistically convergent sequence spaces defined by sequence of Orlicz functions using matrix transformation. *Proyecciones J. Math.*, 2014. Vol. 33, No. 3. P. 277–285. DOI: 10.4067/S0716-09172014000300004
6. Debnath S., Rakshit D. On $\mathcal{I}$-statistical convergence. *Iranian J. Math. Sci. Inform.*, 2018. Vol. 13, No. 2. P. 101–109. DOI: 10.7508/ijmsi.2018.13.009
7. Esi A., Debnath S., Saha S. Asymptotically double λ_2-statistically equivalent sequences of interval numbers. *Mathematica*, 2020. Vol. 62(85), No. 1. P. 39–46. DOI: 10.24193/mathcluj.2020.1.05
8. Fast H. Sur la convergence statistique. *Colloq. Math.*, 1951. Vol. 2, No. 3–4. P. 241–244. (in French)
9. Fridy J. A. On statistically convergence. *Analysis*, 1985. Vol. 5, No. 4. P. 301–313. DOI: 10.1524/anly.1985.5.4.301
10. Kişi Ö. $S_\lambda(\mathcal{I})$-convergence of complex uncertain sequences. *Mat. Stud.*, 2019. Vol. 51, No. 2. P. 183–194.
11. Kostyrko P., Wilczyński W., Šalát T. $\mathcal{I}$-convergence. *Real Anal. Exchange*, 2000/2001. Vol. 26, No. 2. P. 669–686.
12. Liu B. *Uncertainty Theory*, 4th ed. Berlin, Heidelberg: Springer-Verlag, 2015. 487 p. DOI: 10.1007/978-3-662-44354-5
13. Mursaleen M., Debnath S., Rakshit D. $\mathcal{I}$-statistical limit superior and $\mathcal{I}$-statistical limit inferior. *Filomat*, 2017. Vol. 31, No. 7. P. 2103–2108. DOI: 10.2298/FIL1707103M
14. Nath P. K., Tripathy B. C. Convergent complex uncertain sequences defined by Orlicz function. *Ann. Univ. Craiova-Math. Comput. Sci. Ser.*, 2019. Vol. 46, No. 1. P. 139–149.
15. Peng Z. *Complex Uncertain Variable.* PhD, Tsinghua University, 2012.
16. Roy S., Tripathy B. C., Saha S. Some results on p-distance and sequence of complex uncertain variables. *Commun. Korean. Math. Soc.*, 2020. Vol. 35, No. 3. P. 907–916. DOI: 10.4134/CKMS.c200026
17. Saha S., Tripathy B. C., Roy S. On almost convergence of complex uncertain sequences. *New Math. Nat. Comput.*, 2016. Vol. 16, No. 03. P. 573–580. DOI: 10.1142/S1793005720500349
18. Savas E., Das P. A generalized statistical convergence via ideals. *App. Math. Lett.*, 2011. Vol. 24, No. 6. P. 826–830. DOI: 10.1016/j.aml.2010.12.022
19. Savas E., Das P. On $\mathcal{I}$-statistically pre-Cauchy sequences. *Taiwanese J. Math.*, 2014. Vol. 18, No. 1. P. 115–126. DOI: 10.11650/tjm.18.2014.3157
20. Savas E., Debnath S., Rakshit D. On $\mathcal{I}$-statistically rough convergence. *Publ. Inst. Math.*, 2019. Vol. 105, No. 119. P. 145–150. DOI: 10.2298/PIM1919145S

21. Steinhaus H. Sur la convergence ordinaire et la convergence asymptotique. *Colloq. Math.*, 1951. Vol. 2, No. 1. P. 73–74. (in French)

22. Tripathy B. C., Nath P. K. Statistical convergence of complex uncertain sequences. *New Math. Nat. Comput.*, 2017. Vol. 13, No. 3. P. 359–374. DOI: 10.1142/S1793005717500090

23. You C., Yan L. Relationships among convergence concepts of uncertain sequences. *Comput. Model. New Technol.*, 2016. Vol. 20, No. 3. P. 12–16.

24. Wu J., Xia Y. Relationships among convergence concepts of uncertain sequences. *Inform. Sci.*, 2012. Vol. 198. P. 177–185. DOI: 10.1016/j.ins.2012.02.048

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 81–91
DOI: 10.15826/umj.2024.2.008

TAUBERIAN THEOREM FOR GENERAL MATRIX SUMMABILITY METHOD

Bidu Bhusan Jena

Faculty of Science (Mathematics), Sri Sri University,
Cuttack 754006, Odisha, India
bidumath.05@gmail.com

Priyadarsini Parida

Department of Mathematics, Kuntala Kumari Sabat Women's College,
Balasore 756003, Odisha, India
priyadarsiniparida1@gmail.com

Susanta Kumar Paikray

Department of Mathematics, Veer Surendra Sai University of Technology,
Burla 768018, Odisha, India
skpaikray_math@vssut.ac.in

Abstract: In this paper, we prove certain Littlewood–Tauberian theorems for general matrix summability method by imposing the Tauberian conditions such as slow oscillation of usual as well as matrix generated sequence, and the De la Vallée Poussin means of real sequences. Moreover, we demonstrate $(\bar{N}, p_n)$ and $(C, 1)$ — summability methods as the generalizations of our proposed general matrix method and establish an equivalence relation connecting them. Finally, we draw several remarks in view of the generalizations of some existing well-known results based on our results.

Keywords: Matrix summability, Weighted mean, Cesàro mean, Slow oscillation, Tauberian theorem.

1. Introduction and motivation

The study of Littlewood–Tauberian theorems has long been central to mathematical analysis, particularly in the theory of summability and asymptotic analysis. Tauberian theory was initially introduced by Tauber [27] and serves to establish essential connections between summability methods and classical convergence. Numerous researchers, including Littlewood [20], Hardy and Littlewood [11], Landau [19], Schmidt [25], and Hardy [10] contemplated Tauberian hypotheses for Abel and Cesàro summability means. Mostly, they concentrated on forcing the conditions on $(n\Delta u_n)$ to recuperate convergence of (u_n) out of its Abel and Cesàro summability means. A few researchers, like Jakimovski [12] and Szász [26] are focused on imposing the conditions on the arithmetic means of $(n\Delta u_n)$, which is later signified by $(V_n^0 \Delta u_n)$.

The authors Jakimovski [12] and Szász [26] has obtained the convergence of sequences (u_n) via a Littlewood–Tauberian theorem for Cesàro summability method under the conditions on the oscillatory behavior of the generated sequence $(V_n^0 \Delta u_n)$. Also, some authors namely, Çanak and Totur (see, [5, 7, 8]), Dik [9], Totur and Dik [28], Çanak (see, [2–4]), Jena *et al.* (see, [13–16]), Parida *et al.* (see, [22–24]) have proved Tauberian theorems for Abel and Cesàro summability methods by using Tauberian conditions, the oscillatory behaviour of the generator sequences $(V_n^0 \Delta u_n)$ and the De la Vallée Poussin means.

Moreover, it is well known that if the $(\bar{N}, p_n)$-summability exists for a sequence (u_n), then the (J, p)-summability also exists. However, the (J, p)-summability does not imply $(\bar{N}, p_n)$-summability. During last decades of twentieth century, Tietz and Trautner [30] proved the converse part (that is, $(J, p) \to (\bar{N}, p_n)$) under certain Tauberian conditions. Subsequently, Kratz and Stadtmüller [17] proved $o-$ and $O-$types of Tauberian theorems for (J, p)-summability of sequence (u_n) based on various conditions. Again also, Kratz and Stadtmüller [18] obtained a new result of Tauberian theorem for (J, p)-summability method under some specific conditions for the sequence (p_n).

A few authors like, Hardy [10], Tietz [29], Tietz and Zeller [31], Ananda-Rau [1], Móricz and Rhoades [21] have obtained Tauberian theorems via $(\bar{N}, p_n)$-summability method based on different kinds of Tauberian conditions for the sequences (u_n) (or $\omega_{n,p}^{(0)}(u)$). Recently, Çanak and Totur [6] also proved Tauberian theorems via $(\bar{N}, p_n)$-summability means under slow oscillation of the generated sequences $(V_n^0 \Delta u_n)$ and De la Vallée Poussin means. Furthermore, the authors Tietz [29], Tietz and Zeller [31] proved a Tauberian theorem for $(\bar{N}, p_n)$-summability method under the conditions of controlling the oscillatory behavior of sequence (u_n). Again, Móricz and Rhoades [21] established necessary and sufficient conditions for $(\bar{N}, p_n)$-summablility of the sequence (u_n) to be convergent, and they also proved the one-sided boundedness Tauberian theorem for $(\bar{N}, p_n)$-summability method of the sequence $(\omega_{n,p}^{(0)}(u))$ under certain specific Tauberian conditions. In fact, $(\omega_{n,p}^{(0)}(u)) = O(1)$ is a Tauberian condition for the $(\bar{N}, p_n)$-summability method which was earlier given by Hardy (see, [10, Theorem 67]).

In view of the above mentioned literature, our motivation stems from the desire to generalize and unify these results by considering a broad class of matrix summability methods. Such generalizations are particularly valuable because they encapsulate various existing methods as special cases, offering a unified framework to analyze convergence and summability. A key focus of this work is on the interplay between matrix summability and Tauberian conditions, such as the slow oscillation of sequences and the behavior of the De la Vallée Poussin means. By introducing these conditions, we aim to bridge the gap between traditional Tauberian theory and modern summability methods, capturing the subtleties of sequences generated by matrices. Our investigation also demonstrates how well-known summability methods like $(\bar{N}, p_n)$ and $(C, 1)$ emerge as specific instances of the proposed general matrix method. This establishes an equivalence that enhances the applicability and depth of summability theory. Ultimately, our results not only generalize classical theorems but also provide new perspectives and tools for exploring broader classes of summability and their implications.

2. Preliminaries

Let us consider $u = (u_n)$ be a real sequence and $(A = a_{n,k})$ be an infinite lower triangular matrix with non-negative entries. The infinite matrix mean of (u_n) is

$$\sigma_{n,k}^{(1)}(u) = \sum_{k=0}^{n} a_{n,k} u_n, \quad n, k = 0, 1, 2, \ldots . \tag{2.1}$$

It is noticed that, if $a_{n,k} = p_k / P_n$ in (2.1), then it reduces to the weighted $(\bar{N}, p_n)$ mean of (u_n) and it is denoted by

$$\sigma_{p_n}^{(1)}(u) = \frac{1}{P_n} \sum_{k=0}^{n} p_k u_n, \quad n, k = 0, 1, 2, \ldots . \tag{2.2}$$

Subsequently, if $a_{n,k} = 1/(n+1)$ in (2.1), then it reduces to the Cesàro $(C,1)$ mean of (u_n) and it is denoted by

$$\sigma_C^{(1)}(u) = \frac{1}{n+1}\sum_{k=0}^{n} u_n, \quad n,k = 0,1,2,\ldots.$$

We shall also use the notation,

$$\Delta(u_n) = u_{n+1} - u_n.$$

A real sequence (u_n) is summable to s via infinite matrix mean $\sigma_{n,k}^{(1)}(u)$, if

$$\lim_{n\to\infty} \sigma_{n,k}^{(1)}(u) = s. \tag{2.3}$$

Note that, if the limit of sequence

$$\lim_{n\to\infty} (u_n) = s$$

exists, then (2.3) also exists. However, the converse is not true in general.

Next, in view of the proposed matrix summability method the converse part can be achieved under certain conditions, called Tauberian conditions, and the associated results are known as Tauberian theorems.

For each non-negative integer $m \geq 0$, we define

$$\sigma_{n,k}^{m}(u) = \begin{cases} \displaystyle\sum_{k=0}^{n} a_{n,k}\sigma_{n,k}^{(m-1)} u_n, & m \geq 1, \\ u_n, & m = 0. \end{cases}$$

The difference between u_n and the infinite matrix summability mean $(\sigma_{n,k}^{(1)}(u))$ is known as matrix Kronecker identity (or matrix generator sequences $V_{n,k}^{(0)}(\Delta u)$), and is given by

$$u_n - \sigma_{n,k}^{(1)}(u) = V_{n,k}^{(0)}(\Delta u) = \sum_{k=1}^{n} \bar{a}_{n,k}\Delta u_k, \tag{2.4}$$

where

$$\bar{A} = \bar{a}_{n,k} = \sum_{r=k}^{n} a_{n,r}, \quad n,k = 0,1,2,\ldots.$$

Similarly, we define $V_{n,k}^{(m)}(\Delta u)$ for each non-negative integer $m \geq 0$ as

$$V_{n,k}^{m}(\Delta u) = \begin{cases} \displaystyle\sum_{k=0}^{n} a_{n,k}V_{n,k}^{(m-1)}(\Delta u), & m \geq 1, \\ V_{n,k}^{(0)}(\Delta u), & m = 0. \end{cases}$$

Moreover, for the matrix De la Vallée Poussin means of (u_n), we may define

$$\tau_{n,[\lambda n],k}(u_n) = \sum_{k=n+1}^{[\lambda n]} (a_{[\lambda n],k} - a_{n,k})u_k, \quad \lambda \in (1,\infty)$$

and

$$\tau_{n,[\lambda n],k}(u_n) = \sum_{k=[\lambda n]+1}^{n} (a_{n,k} - a_{[\lambda n],k})u_k, \quad \lambda \in (0,1).$$

A sequence $u = (u_n)$ is oscillating slowly [4], if

$$\lim_{\lambda \to 1^+} \limsup_n \max_{n \leq k \leq [\lambda n]} |u_k - u_n| = 0. \tag{2.5}$$

An equivalent reformulation of (2.5) can be given as follows:

$$\lim_{\lambda \to 1^-} \limsup_n \max_{[\lambda n] \leq k \leq n} |u_n - u_k| = 0.$$

3. Some auxiliary lemmas

Before we establish the Tauberian theorems via our purposed mean, first we need the following lemmas.

Lemma 1 [9]. *The sequence $u = (u_n)$ is oscillating slowly if and only if $V^{(0)}_{n,k}(\Delta u)$ is bounded and oscillating slowly.*

Lemma 2. *Let $u = (u_n)$ be a sequence of real numbers*

(i) *for $\lambda > 1$,*

$$u_n - \sigma^{(1)}_{n,k}(u) = \left(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right)\left(\sigma^{(1)}_{[\lambda n],k}(u) - \sigma^{(1)}_{n,k}(u)\right) - \sum_{k=n+1}^{[\lambda n]} \left(a_{[\lambda n],k} - a_{n,k}\right)(u_k - u_n),$$

(ii) *for $0 < \lambda < 1$,*

$$u_n - \sigma^{(1)}_{n,k}(u) = \left(\bar{a}_{n,[\lambda n]} - \bar{a}_{[\lambda n],[\lambda n]}\right)\left(\sigma^{(1)}_{n,k}(u) - \sigma^{(1)}_{[\lambda n],k}(u)\right) - \sum_{k=[\lambda n]+1}^{n} \left(a_{n,k} - a_{[\lambda n],k}\right)(u_n - u_k).$$

P r o o f. For $\lambda > 1$, from the definition of de la Vallée Poussin means of (u_n), we have

$$\begin{aligned}\tau_{n,[\lambda n],k}(u_n) &= \sum_{k=n+1}^{[\lambda n]} (a_{[\lambda n],k} - a_{n,k})u_k = \left(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right)\sigma^{(1)}_{[\lambda n],k}u_k - \left(\bar{a}_{[\lambda n],n} - \bar{a}_{n,n}\right)\sigma^{(1)}_{n,k}u_k \\ &= \sigma^{(1)}_{n,k}u_k + \left(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right)\left(\sigma^{(1)}_{[\lambda n],k}(u) - \sigma^{(1)}_{n,k}(u)\right).\end{aligned}$$

The difference $(\tau_{n,[\lambda n],k}(u_n) - \sigma^1_{n,k}(u_n))$ can be written as

$$\tau_{n,[\lambda n],k}(u_n) - \sigma^1_{n,k}(u_n) = \left(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right)\left(\sigma^{(1)}_{[\lambda n],k}(u) - \sigma^{(1)}_{n,k}(u)\right). \tag{3.1}$$

Subtracting $(\sigma^{(1)}_{n,k}(u))$ from both sides of the identity

$$u_n = \tau_{n,[\lambda n],k}(u_n) - \sum_{k=n+1}^{[\lambda n]} \left(a_{[\lambda n],k} - a_{n,k}\right)(u_k - u_n),$$

we have

$$u_n - \sigma^{(1)}_{n,k}(u) = \left(\tau_{n,[\lambda n],k}(u_n) - \sigma^{(1)}_{n,k}(u)\right) - \sum_{k=n+1}^{[\lambda n]} \left(a_{[\lambda n],k} - a_{n,k}\right)(u_k - u_n). \tag{3.2}$$

Considering equations (3.1) and (3.2), we have

$$u_n - \sigma^{(1)}_{n,k}(u) = \left(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right)\left(\sigma^{(1)}_{[\lambda n],k}(u) - \sigma^{(1)}_{n,k}(u)\right) - \sum_{k=n+1}^{[\lambda n]} \left(a_{[\lambda n],k} - a_{n,k}\right)(u_k - u_n).$$

Next, for $0 < \lambda < 1$, the remaining part, that is, Lemma 2 (ii) can be proved in the similar lines of the proof of Lemma 2 (i). Thus, we skip the details. □

4. Main results

In this section, we establish four theorems along with their associated corollaries. The first theorem proves a Tauberian theorem under the infinite matrix mean of order 1, specifically under the $(A,1)$-summability mean, based on the slow oscillation of the sequence $u=(u_n)$. The second theorem extends this result to the infinite matrix mean of order m that is, under (A,m)- summability mean, also relying on the slow oscillation of the sequence $u=(u_n)$. The third theorem establishes and proves a Tauberian theorem under the infinite matrix mean of order 1 but focuses on the slow oscillation of a generalized sequence $V=V^{(0)}_{n,k}(\Delta u)$. Similarly, the fourth theorem generalizes this result to the infinite matrix mean of order m again utilizing the slow oscillation of the generalized sequence $V=V^{(0)}_{n,k}(\Delta u)$. Subsequently, we state and prove three corollaries that demonstrate how the results recover or extend earlier established results in the literature. This structured progression highlights the depth and generality of the Tauberian theorems under various summability means.

Theorem 1. *If $u=(u_n)$ is matrix summable to s, and so also oscillating slowly, then $u_n\to s$ as $n\to\infty$.*

P r o o f. Suppose $u=(u_n)$ is matrix summable to s, this implies that $(\sigma^{(1)}_{n,k}(u))$ is matrix summable to s. From equation (2.4), we have

$$V^{(0)}_{n,k}(\Delta u)=\sum_{k=1}^{n}\bar{a}_{n,k}\Delta u_k,$$

which is also matrix summable to zero.

As $V^{(0)}_{n,k}(\Delta u)$ is oscillating slowly, so from Lemma 1 and Lemma 2 (i), we get

$$\begin{aligned}V^{(0)}_{n,k}(\Delta u)-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))&=\left(\bar{a}_{[\lambda n],[\lambda n]}-\bar{a}_{n,[\lambda n]}\right)\left(\sigma^{(1)}_{[\lambda n],k}(V^{(0)}_{n,k}(\Delta u))-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))\right)\\&\quad-\sum_{k=n+1}^{[\lambda n]}\left(a_{[\lambda n],k}-a_{n,k}\right)\left(V^{(0)}_{n,k}(\Delta u)-V^{(0)}_{n,n}(\Delta u)\right).\end{aligned}$$

Now,

$$\begin{aligned}\left|V^{(0)}_{n,k}(\Delta u)-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))\right|&\le\left|\bar{a}_{[\lambda n],[\lambda n]}-\bar{a}_{n,[\lambda n]}\right|\left|\sigma^{(1)}_{[\lambda n],k}(V^{(0)}_{n,k}(\Delta u))-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))\right|\\&\quad+\left|\sum_{k=n+1}^{[\lambda n]}(a_{[\lambda n],k}-a_{n,k})\left(V^{(0)}_{n,k}(\Delta u)-V^{(0)}_{n,n}(\Delta u)\right)\right|.\end{aligned}\tag{4.1}$$

Next, taking lim sup to both sides of equation (4.1) as $n\to\infty$, and $\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))$ being convergent, so in view of equation (2.5), we obtain

$$\begin{aligned}&\limsup_n\left|V^{(0)}_{n,k}(\Delta u)-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))\right|\\&\le\limsup_n\max_{n+1\le k\le[\lambda n]}\left|\sum_{k=n+1}^{[\lambda n]}(a_{[\lambda n],k}-a_{n,k})\left(V^{(0)}_{n,k}(\Delta u)-V^{(0)}_{n,n}(\Delta u)\right)\right|.\end{aligned}\tag{4.2}$$

Letting $\lambda\to 1^+$ in (4.2), we have

$$\limsup_n\left|V^{(0)}_{n,k}(\Delta u)-\sigma^{(1)}_{n,k}(V^{(0)}_{n,k}(\Delta u))\right|\le 0.$$

This implies that

$$V_{n,k}^{(0)}(\Delta u) = o(1) \quad (n \to \infty).$$

Moreover, since (u_n) is matrix summable to s and $V_{n,k}^{(0)}(\Delta u) = o(1)$, consequently,

$$\lim_{n\to\infty} u_n = s.$$

This completes the proof. □

Remark 1. *For $\lambda \to 1^+$, Theorem 1 can also be proved by using (2.4) and Lemma 2.*

Theorem 2. *If $u = (u_n)$ is (A, m) summable to s and so also (u_n) is oscillating slowly, then $\lim_{n\to\infty} u_n = s$.*

P r o o f. Let $u = (u_n)$ be oscillating slowly, then $\sigma_{n,k}^m(u)$ is oscillating slowly (by Lemma 1). Since $u = (u_n)$ is (A, m) summable to s,

$$\lim_{n\to\infty} \sigma_{n,k}^m(u) = s. \tag{4.3}$$

We next write,

$$\sigma_{n,k}^m(u) = \sigma_{n,k}^1(u)(\sigma_{n,k}^{m-1}(u)). \tag{4.4}$$

Clearly from equations (4.3) and (4.4), $u = (u_n)$ is $(A, m-1)$ summable to s. Again, $(\sigma_{n,k}^{m-1}(u))$ is oscillating slowly (by Lemma 1). Thus, we have

$$\lim_{n\to\infty} \sigma_{n,k}^{m-1}(u) = s,$$

and continuing in this way, we get $\lim_{n\to\infty}(u_n) = s$. □

Corollary 1. *If $u = (u_n)$ is $(\bar{N}, p_n)$ summable to s and so also (u_n)oscillating slowly, then $u_n \to s$ as $n \to \infty$.*

P r o o f. If we substitute

$$a_{n,k} = \frac{p_k}{P_n},$$

then the matrix transform reduces to the weighted transform $(\bar{N}, p_n)$, where (p_k) is the positive real sequence and $P_n = \sum_{k=0}^{n} p_k$. Next, the proof is similar to that of the proof of Theorem 1, thus we skip the details. □

Corollary 2. *If $u = (u_n)$ is $(C, 1)$-summable to s and so also (u_n) is oscillating slowly, then $u_n \to s$ as $n \to \infty$.*

P r o o f. If we substitute

$$a_{n,k} = \frac{1}{n+1},$$

then the matrix transform reduces to the Cesàro transform $(C, 1)$. Next, the proof is similar to that of the proof of Theorem 1, thus we skip the details. □

Corollary 3. *The following embedding equivalence relations are valid for a sequence* (u_n):

(i) *Matrix summability* $(A,1) \Longrightarrow (\bar{N},p_n)$*-summability*;

(ii) $(\bar{N},p_n)$*-summability* $\Longrightarrow (C,1)$*-summability*;

(iii) *Matrix summability* $(A,1) \Longrightarrow (C,1)$*-summability.*

P r o o f. (i) Suppose $u=(u_n)$ is $(A,1)$ summable to s. This implies that

$$\sigma^{(1)}_{n,k}(u) = \sum_{k=0}^{n} a_{n,k}u_k$$

is convergent to s under matrix summability. If we substitute $a_{n,k}=p_k/P_n$, then $\sigma^{(1)}_{n,k}(u)$ mean reduces to $\sigma^{(1)}_{p_n}(u)$ mean, which is the the weighted mean (2.2). Thus, $(\sigma^{(1)}_{p_n}(u))$ is convergent to s.

(ii) If $u=(u_n)$ is $(\bar{N},p_n)$ summable to s. This implies that

$$\sigma^{(1)}_{n,p_n}(u) = \frac{1}{P_n}\sum_{k=0}^{n} p_k u_k$$

is convergent to s under $(\bar{N},p_n)$-summability mean. If we substitute $p_n=1$, then $\sigma^{(1)}_{p_n}(u)$ mean reduces to $\sigma^{(1)}_{C}(u)$ mean, which is the Cesàro mean. Thus, $(\sigma^{(1)}_{C}(u))$ is convergent to s.

(iii) If $u=(u_n)$ is matrix summable to s. This implies that

$$\sigma^{(1)}_{n,k}(u) = \sum_{k=0}^{n} a_{n,k}u_n$$

is convergent to s under matrix summability. If we substitute $a_{n,k}=1/(n+1)$, then $(\sigma^{(1)}_{n,k}(u))$ mean reduces to $(\sigma^{(1)}_{C}(u))$ mean, which is the Cesàro mean. Thus, $(\sigma^{(1)}_{C}(u))$ is convergent to s. $\square$

Theorem 3. *If* $u=(u_n)$ *is matrix summable to* s *and* $V=V^{(0)}_{n,k}(\Delta u)$ *is oscillating slowly, then* $u_n \to s$ *as* $n\to\infty$.

P r o o f. As $u=(u_n)$ is matrix summable to s, so $\sigma^{(1)}_{n,k}(u)$ is also matrix summable to s. Therefore by equation (2.4), $V^{(0)}_{n,k}(\Delta u)$ is matrix summable to zero. Also, considering the identity under (2.4) for $V^{(0)}_{n,k}(\Delta u)$, we fairly say that $V(V^{(0)}_{n,k}(\Delta u))$ is matrix summable to zero. Consequently, $V(V^{(0)}_{n,k}(\Delta u))$ is oscillating slowly by Lemma 1.

Next by Lemma 2 (i),

$$\begin{aligned} &V(V^{(0)}_{n,k}(\Delta u)) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u))) = \\ &\big(\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\big)\big(\sigma^{(1)}_{[\lambda n],k}(V(V^{(0)}_{n,k}(\Delta u))) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))\big) \\ &- \sum_{k=n+1}^{[\lambda n]} (a_{[\lambda n],k} - a_{n,k})\big(V(V^{(0)}_{n,k}(\Delta u)) - V(V^{(0)}_{n,n}(\Delta u))\big). \end{aligned} \tag{4.5}$$

Also by (4.5),

$$
\begin{aligned}
&\left|V(V^{(0)}_{n,k}(\Delta u)) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))\right| \\
\leq &\left|\bar{a}_{[\lambda n],[\lambda n]} - \bar{a}_{n,[\lambda n]}\right|\left|\sigma^{(1)}_{[\lambda n],k}(V(V^{(0)}_{n,k}(\Delta u))) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))\right| \\
&+\left|\sum_{k=n+1}^{[\lambda n]}(a_{[\lambda n],k} - a_{n,k})\big(V(V^{(0)}_{n,k}(\Delta u)) - V(V^{(0)}_{n,n}(\Delta u))\big)\right|.
\end{aligned}
\tag{4.6}
$$

Now taking $\limsup$ to both sides of equation (4.7) as $n \to \infty$, and $\sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))$ being convergent, so in view of equation (2.5), we have

$$
\begin{aligned}
&\limsup_{n}\left|V(V^{(0)}_{n,k}(\Delta u)) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))\right| \\
\leq \limsup_{n} \max_{n+1\leq k\leq[\lambda n]}&\left|\sum_{k=n+1}^{[\lambda n]}\left(a_{[\lambda n],k} - a_{n,k}\right)\left(V(V^{(0)}_{n,k}(\Delta u) - V(V^{(0)}_{n,n}(\Delta u)))\right)\right|.
\end{aligned}
\tag{4.7}
$$

Letting $\lambda \to 1^{+}$, we have

$$\limsup_{n}\left|V(V^{(0)}_{n,k}(\Delta u)) - \sigma^{(1)}_{n,k}(V(V^{(0)}_{n,k}(\Delta u)))\right| \leq 0.$$

Thus, $V(V^{(0)}_{n,k}(\Delta u)) = o(1)$ as $n \to \infty$. Since (u_n) is matrix summable to s and $V(V^{(0)}_{n,k}(\Delta u)) = o(1)$ as $n \to \infty$, $\lim\limits_{n\to\infty} u_n = s$. This completes the proof. □

Theorem 4. *If (u_n) is (A, m) summable to s and $V^{(0)}_{n,k}(\Delta u)$ is oscillating slowly, then $\lim_{n\to\infty} u_n = s$.*

P r o o f. As $V^{(0)}_{n,k}(\Delta u)$ is oscillating slowly, setting $u = (u_n)$ in place of $V^{(0)}_{n}(\Delta u)$, $\sigma^{(m)}_{n,k}(V^{(0)}_{n,k}(\Delta u))$ is oscillating slowly (by Lemma 1). Again, as $V^{(0)}_{n,k}(\Delta u)$ is (A, m) summable to s, so by Theorem 3, we have

$$\lim_{n\to\infty} \sigma^{(m)}_{n,k}(V^{(11)}_{n,k}(\Delta u)) = s. \tag{4.8}$$

By definition,

$$\sigma^{(m)}_{n,k}(V^{(0)}_{n,k}(\Delta u)) = \sigma^{(0)}_{n,k}(V^{(0)}_{n,k}(\Delta u))(\sigma^{(m-1)}_{n,k}(V^{(0)}_{n,k}(\Delta u)). \tag{4.9}$$

From (4.8) and (4.9), we have $V^{(0)}_{n,k}(\Delta u)$ is $(A, m-1)$ summable to s.

Again by Lemma 1, since $(\sigma^{(m-1)}_{n,k}(V^{(0)}_{n,k}(\Delta u)))$ is oscillating slowly, so we have by Theorem 3

$$\lim_{n\to\infty} \sigma^{(m-1)}_{n,k}(V^{(0)}_{n,k}(\Delta u)) = s.$$

Continuing in this way, we get $\lim\limits_{n\to\infty}(V^{(0)}_{n,k}(\Delta u)) = s$. □

5. Concluding remarks and observations

In this concluding part of our investigation, we draw several observations and further remarks concerning various results which we have established in this article.

Remark 2. *If $\bar{a}_{n,k} = p_k/P_n$, then matrix generator sequence $(V_{n,k}^{(0)}(\Delta u_n))$ reduces to the weighted generator sequence $(V_{n,p_n}^{(0)}(\Delta u_n))$, that is,*

$$V_{n,k}^{(0)}(\Delta u_n) = \sum_{k=1}^{n} \bar{a}_{n,k} \Delta u_k, \quad n,k = 0,1,2,...$$

reduces to

$$V_{n,p_n}^{(0)}(\Delta u_n) = \frac{1}{P_n} \sum_{r=k}^{n} p_r \Delta u_r.$$

Remark 3. *If $p_r = 1$ and $\sum_{r=k}^{n} P_r = n+1$, then the weighted generator sequence $V_{n,p_n}^{(0)}(\Delta u_n)$ reduces to the Cesàro generator sequence $V_C^{(0)}(\Delta u_n)$, that is,*

$$V_{n,p_n}^{(0)}(\Delta u_n) = \frac{1}{P_n} \sum_{r=k}^{n} P_k \Delta u_k$$

reduces to

$$V_C^{(0)}(\Delta u_n) = \frac{1}{n+1} \sum_{k=0}^{n} k \Delta u_n.$$

Remark 4. *If $\bar{a}_{n,k} = k/n+1$, then the matrix generator sequence $(V_{n,k}^{(0)}(\Delta u_n))$ reduces to the Cesàro generator sequence $V_{n,c}^{(0)}(\Delta u_n)$, that is,*

$$V_{n,k}^{(0)}(\Delta u_n) = \sum_{r=k}^{n} \bar{a}_{n,r} \Delta u_r, \quad n,r = 0,1,2,...$$

reduces to

$$V_C^{(0)}(\Delta u_n) = \frac{1}{n+1} \sum_{k=1}^{n} k \Delta u_n.$$

Remark 5. *If $u = (u_n)$ is $(\bar{N}, p_n)$ summable to s and $V = V_{n,p_n}^{(0)}(\Delta u)$ is oscillating slowly, then $u_n \to s$ as $n \to \infty$.*

Remark 6. *If $u = (u_n)$ is $(C,1)$ summable to s and $V = V_C^{(0)}(\Delta u)$ is oscillating slowly, then $u_n \to s$ as $n \to \infty$.*

Acknowledgements

Authors are thankful to the reviewers for their insightful comments and suggestions which improved the quality and presentation of the paper

REFERENCES

1. Ananda-Rau K. An example in the theory of summation of series by Riesz's typical means. *Proc. Lond. Math. Soc.*, 1930. Vol. s2-30, No. 1. P. 367–372. DOI: 10.1112/plms/s2-30.1.367
2. Çanak İ. A theorem on the Cesàro summability method. *Comput. Math. Appl.*, 2011. Vol. 61, No. 4. P. 1162–1166. DOI: 10.1016/j.camwa.2010.12.065
3. Çanak İ. A theorem for convergence of generator sequences. *Comput. Math. Appl.*, 2011. Vol. 61, No. 2. P. 408–411. DOI: 10.1016/j.camwa.2010.11.017
4. Çanak İ. An extended Tauberian theorem for the $(C, 1)$ summability method. *Appl. Math. Lett.*, 2008. Vol. 21, No. 1. P. 74–80. DOI: 10.1016/j.aml.2007.02.012
5. Çanak İ., Totur Ü. Some Tauberian conditions for Cesàro summability method. *Math. Slovaca*, 2012. Vol. 62, No. 2. P. 271–280. DOI: 10.2478/s12175-012-0008-y
6. Çanak İ., Totur Ü. Some Tauberian theorems for the weighted mean methods of summability. *Comput. Math. Appl.*, 2011. Vol. 62, No. 6. P. 2609–2615. DOI: 10.1016/j.camwa.2011.07.066
7. Çanak İ., Totur Ü. A condition under which slow oscillation of a sequence follows from Cesàro summability of its generator sequence. *Appl. Math. Comput.*, 2010. Vol. 216, No. 5. P. 1618–1623. DOI: 10.1016/j.amc.2010.03.017
8. Çanak İ., Totur Ü. A note on Tauberian theorems for regularly generated sequences. *Tamkang J. Math.*, 2008, Vol. 39, No. 2. P. 187–191. DOI: 10.5556/j.tkjm.39.2008.29
9. Dik M. Tauberian theorems for sequences with moderately oscillatory control modulo. *Math. Moravica*, 2001. Vol. 5. P. 57–94. DOI: 10.5937/MatMor0105057D
10. Hardy G. H. *Divergent Series*. Oxford: Clarendon Press, 1949. 396 p.
11. Hardy G. H., Littlewood J. E. Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive. *Proc. Lond. Math. Soc.*, 1914. Vol. s2-13, No. 1. P. 174–191. DOI: 10.1112/plms/s2-13.1.174
12. Jakimovski A. On a Tauberian theorem by O. Szász. *Proc. Amer. Math. Soc.*, 1954. Vol. 5, No. 1. P. 67–70. DOI: 10.2307/2032107
13. Jena B. B., Paikray S. K., Misra U. K. Inclusion theorems on general convergence and statistical convergence of $(L, 1, 1)$-summability using generalized Tauberian conditions. *Tamsui Oxf. J. Inf. Math. Sci.*, 2017. Vol. 31, No. 1. P. 101–115.
14. Jena B. B., Paikray S. K., Misra U. K. A Tauberian theorem for double Cesàro summability method. *Int. J. Math. Math. Sci.*, 2016. Vol. 2016. Art. no. 2431010. P. 1–4. DOI: 10.1155/2016/2431010
15. Jena B. B., Paikray S. K., Misra U. K. A proof of Tauberian theorem for Cesàro summability method. *Asian J. Math. Comput. Res.*, 2016. Vol. 8, No. 3. P. 272–276.
16. Jena B. B., Paikray S. K., Parida P., Dutta H. Results on Tauberian theorem for Cesàro summable double sequences of fuzzy numbers. *Kragujevac J. Math.*, 2020. Vol. 44, No. 4. P. 495–508. DOI: 10.46793/KgJMat2004.495J
17. Kratz W., Stadtmüller U. Tauberian theorems for J_p-summability. *J. Math. Anal. Appl.*, 1989. Vol. 139, No. 2. P. 362–371. DOI: 10.1016/0022-247X(89)90113-3
18. Kratz W., Stadtmüller U. Tauberian theorems for general J_p-methods and a characterization of dominated variation. *J. Lond. Math. Soc. (2)*, 1989. Vol. s2–39, No. 1. P. 145–159. DOI: 10.1112/jlms/s2-39.1.145
19. Landau E. Über einen Satz des herrn Littlewood. *Rend. Circ. Matem. Palermo*, 1913. Vol. 35. P. 265–276. DOI: 10.1007/BF03015606 (in German)
20. Littlewood J. E. The converse of Abel's theorem on power series. *Proc. Lond. Math. Soc.*, 1911. Vol. s2-9, No. 2. P. 434–448.
21. Móricz F., Rhoades B. E. Necessary and sufficient Tauberian conditions for certain weighted mean methods of summability. *Acta Math. Hung.*, 1995. Vol. 66. P. 105–111. DOI: 10.1007/BF01874356
22. Parida P., Paikray S. K., Jena B. B. Statistical Tauberian theorems for Cesàro integrability mean based on post-quantum calculus. *Arab. J. Math.*, 2020. Vol. 9. P. 653–663. DOI: 10.1007/s40065-020-00284-z
23. Parida P., Paikray S. K., Jena B. B. Tauberian theorems for statistical Cesàro summability of function of two variables over a locally convex space. In: *Recent Advances in Intelligent Information Systems and Applied Mathematics (ICITAM 2019), O. Castillo and al. (eds)*. Stud. Comput. Intell., vol 863. Cham: Springer, 2020. P. 779–790. DOI: 10.1007/978-3-030-34152-7_60

24. Parida P., Paikray S. K., Dutta H., Jena B. B., Dash M. Tauberian theorems for Cesàro summability of n-th sequences. *Filomat*, 2018. Vol. 32, No. 11. P. 3993–4004. DOI: 10.2298/FIL1811993P
25. Schmidt R. Über divergente Folgen und lineare Mittelbildungen. *Math. Z.*, 1925. Vol. 22. P. 89–152. DOI: 10.1007/BF01479600 (in German)
26. Szász O. On a Tauberian theorem for Abel summability. *Pacific J. Math.*, 1951. Vol. 1, No. 1. P. 117–125.
27. Tauber A. Ein Satz aus der Theorie der unendlichen Reihen. *Monatsh. Math.*, 1897. Vol. 8. P. 273–277. DOI: 10.1007/BF01696278 (in German)
28. Totur Ü., Dik M. One-sided Tauberian conditions for a general summability method. *Math. Comput. Model.*, 2011. Vol. 54, No. 11–12. P. 2639–2644. DOI: 10.1016/j.mcm.2011.06.041
29. Tietz H. Schmidtsche Umkehrbedingungen für Potenzreihenverfahren. *Acta Sci. Math.*, 1990. Vol. 54. P. 355–365. (in German)
30. Tietz H., Trautner R. Tauber Sätze für Potenzreihenverfahren. *Arch. Math.*, 1988. Vol. 50. P. 164–174. DOI: 10.1007/BF01194575 (in German)
31. Tietz H., Zeller K. Tauber-Bedingungen für Verfahren mit Abschnittskonvergenz. *Acta Math. Hungar.*, 1998. Vol. 81. P. 241–247. DOI: 10.1023/A:1006598501040 (in German)

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 92–106

DOI: 10.15826/umj.2024.2.009

TRAJECTORIES OF DYNAMIC EQUILIBRIUM AND REPLICATOR DYNAMICS IN COORDINATION GAMES

Nikolay A. Krasovskii[a,†], **Alexander M. Tarasyev**[a,b,††]

[a]Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences,
16 S. Kovalevskaya Str., Ekaterinburg, 620108, Russian Federation

[b]Ural Federal University,
19 Mira Str., Ekaterinburg, 620002, Russian Federation

[†]n.a.krasovskii@imm.uran.ru, n.a.krasovskii@gmail.com
[††]tam@imm.uran.ru, a.m.tarasyev@urfu.ru

Abstract: The paper analyzes average integral payoff indices for trajectories of the dynamic equilibrium and replicator dynamics in bimatrix coordination games. In such games, players receive large payoffs when choosing the same type of behavior. A special feature of a 2×2 coordination game is the presence of three static Nash equilibria. In the dynamic formulation, the trajectories of coordination games are estimated by the average integral payoffs for a wide range of models arising in economics and biology. In optimal control problems and dynamic games, average integral payoffs are used to synthesize guaranteed strategies, which are involved, among other things, in the constructions of the dynamic Nash equilibrium. In addition, average integral payoffs are a natural tool for assessing the quality of trajectories of replicator dynamics. In the paper, we compare values of average integral indices for trajectories of replicator dynamics and trajectories generated by guaranteed strategies in constructing the dynamic Nash equilibrium. An analysis is provided for trajectories of mixed dynamics when the first player plays a guaranteed strategy, and the behavior of replicator dynamics guides the second player.

Keywords: Dynamic bimatrix games, Coordination games, Average integral payoffs, Guaranteed strategies, Replicator dynamics, Dynamic Nash equilibrium.

1. Introduction

The paper is devoted to analyzing the behavior of equilibrium trajectories in dynamic bimatrix coordination games with average integral indices of players' payoffs. In such games, players obtain better payoffs when choosing the same type of behavior. A feature of a 2×2 coordination game is the presence of three static Nash equilibria. Players' benefits on each time interval are determined as mathematical expectations of payoffs. On the infinite time interval players' payoff functionals are defined as average integral indices (time average values), methods for whose analysis in control theory were studied in papers [1, 14].

In the first step, we consider a solution of the dynamic bimatrix game using approaches of the theory of differential–evolutionary games, ideas of N.N. Krasovskii guaranteed strategies [5, 7, 8], and constructions of L.S. Pontryagin maximum principle [12]. Based on the proposed approach we elaborate an algorithm for constructing positional strategies and equilibrium trajectories of dynamic Nash equilibrium [4, 6]. Equilibrium trajectories generated by guaranteed strategies provide payoff results not worse than those of the static Nash equilibrium [15] located inside the square of the game. In this sense, guaranteed strategies allow one to shift game solutions toward Pareto maximum points generated by cooperative constructions [9, 11].

In the second step, we consider an analysis of constructions for replicator dynamics which is widely used in the theory of evolutionary games and applications [2, 3, 10, 13, 16]. Trajectories of the replicator dynamics in coordination games converge to static Nash equilibria located at vertices of the game square and demonstrate the bifurcation behavior depending on chosen initial positions.

In the third step, we consider the so-called "mixed" dynamics, when the first player uses guaranteed strategies and equilibrium trajectories of the second player are generated by the replicator dynamics. Values of players' payoff functionals at the attraction points of the motion for equilibrium trajectories of "mixed" dynamics majorate values of payoffs at the point of the static Nash equilibrium.

A model is considered for a dynamic coordination game of two coalitions of players called "hawks" and "doves." We construct equilibrium trajectories for guaranteed strategies, replicator dynamics, and "mixed" constructions for such a game. A comparison is carried out for equilibrium trajectories of all three types of dynamics.

2. Game dynamics. Players' payoff functionals

To describe the behavior of two players, we consider the system of differential equations

$$\begin{cases} \dot{\xi}(t) = -\xi(t) + u(t), & \xi(t_0) = \xi_0, \\ \dot{\eta}(t) = -\eta(t) + v(t), & \eta(t_0) = \eta_0, \end{cases} \tag{2.1}$$

where the parameters $\xi = \xi(t)$, $0 \leq \xi \leq 1$, and $\eta = \eta(t)$, $0 \leq \eta \leq 1$ determine the probabilities of choosing strategies by players. For example, the parameter ξ stands for the probability that the first player holds to the first strategy (respectively, $(1 - \xi)$ is the probability that he holds to the second strategy). The parameter η stands for the probability of choosing the first strategy by the second player (respectively, $(1 - \eta)$ means the probability that he holds to the second strategy). The control parameters $u = u(t)$ and $v = v(t)$ satisfy the conditions $0 \leq u \leq 1$ and $0 \leq v \leq 1$ and are signals recommending players to change their strategy. For example, the value $u = 0$ $(v = 0)$ corresponds to the signal: "change the first strategy to the second". The value $u = 1$ $(v = 1)$ corresponds to the signal: "change the second strategy to the first". The value $u = \xi$ $(v = \eta)$ corresponds to the signal: "keep the previous strategy"

The square, $(\xi, \eta) \in [0, 1] \times [0, 1]$, of the game is a strongly invariant set due to the properties of the dynamics (2.1). So, any trajectory of the dynamics (2.1), that starts in the square, survives in it on the infinite horizon of time.

Matrices A and B describe players' payoffs

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Terminal quality functionals are defined as the mathematical expectations of payoffs given by corresponding matrices A and B in a bimatrix game and can be interpreted as "local" interests of players

$$g_A(\xi(T), \eta(T)) = C_A \xi(T)\eta(T) - \alpha_1 \xi(T) - \alpha_2 \eta(T) + a_{22}$$

at given time T. Here, the parameters C_A, α_1, and α_2 are defined according to the classic theory of bimatrix games [15]

$$C_A = a_{11} - a_{12} - a_{21} + a_{22},$$
$$\alpha_1 = a_{22} - a_{12}, \quad \alpha_2 = a_{22} - a_{21}.$$

The quality functional g_B of the second player and the parameters C_B, β_1, and β_2 are defined analogously by the coefficients of the matrix B.

The “global” interests J_A^∞ of the first player are determined as limit relations for quality functionals on an infinite planning horizon

$$
\begin{aligned}
JI_A^\infty &= [JI_A^-, JI_A^+],\\
JI_A^- = JI_A^-(\xi(\cdot),\eta(\cdot)) &= \liminf_{T\to\infty} \frac{1}{(T-t_0)} \int_{t_0}^{T} g_A(\xi(t),\eta(t))dt,\\
JI_A^+ = JI_A^+(\xi(\cdot),\eta(\cdot)) &= \limsup_{T\to\infty} \frac{1}{(T-t_0)} \int_{t_0}^{T} g_A(\xi(t),\eta(t))dt,
\end{aligned}
\tag{2.2}
$$

calculated for the trajectories $(\xi(\cdot),\eta(\cdot))$ of system (2.1). For the second player, the “global” interests J_B^∞ are determined symmetrically.

The average integral functionals (2.2) are widely used for the problems of evolution in economics and biology. In optimal control problems, such functionals were studied in the papers [1, 14] and called time average values. Unlike the payoff functionals optimized at each time, average integral payoffs allow potential losses in some periods to win in others. Thus, they obtain the best integral result over all periods on the infinite horizon. Such property guarantees another character of switching lines in optimal closed-loop control strategies compared with the problems where payoffs are optimized at the terminal time. This construction allows the system to stay longer in advantageous domains where the local payoffs of coalitions are strictly better than the payoffs at the points of the static Nash equilibrium.

3. Dynamic Nash equilibrium

3.1. Definition of dynamic Nash equilibrium

The solution of the dynamic game is considered based on the optimal control theory [12] and differential game theory [5, 8]. Following [4], we present the definition of the dynamic Nash equilibrium in the class of positional strategies (feedbacks) $U = u(t,\xi,\eta,\varepsilon)$ and $V = v(t,\xi,\eta,\varepsilon)$.

Definition 1. *Let $\varepsilon > 0$ and $(\xi_0,\eta_0) \in [0,1]\times[0,1]$. The pair of feedbacks $U^0 = u^0(t,\xi,\eta,\varepsilon)$ and $V^0 = v^0(t,\xi,\eta,\varepsilon)$ is called the Nash equilibrium at the initial point (ξ_0,η_0) if the following conditions hold for any other feedbacks $U = u(t,\xi,\eta,\varepsilon)$ and $V = v(t,\xi,\eta,\varepsilon)$: the inequalities*

$$
\begin{aligned}
J_A^-(\xi^0(\cdot),\eta^0(\cdot)) &\geq J_A^+(\xi_1(\cdot),\eta_1(\cdot)) - \varepsilon,\\
J_B^-(\xi^0(\cdot),\eta^0(\cdot)) &\geq J_B^+(\xi_2(\cdot),\eta_2(\cdot)) - \varepsilon
\end{aligned}
$$

are true for any trajectories

$$
\begin{aligned}
(\xi^0(\cdot),\eta^0(\cdot)) &\in X(\xi_0,\eta_0,U^0,V^0),\\
(\xi_1(\cdot),\eta_1(\cdot)) &\in X(\xi_0,\eta_0,U,V^0),\\
(\xi_2(\cdot),\eta_2(\cdot)) &\in X(\xi_0,\eta_0,U^0,V).
\end{aligned}
$$

Here, the symbol X stands for the set of trajectories that start from the initial point (ξ_0,η_0) and are generated by the corresponding strategies (U^0,V^0), (U,V^0), and (U^0,V) (see [8]).

3.2. Auxiliary zero-sum games

We employ the results of [4] for constructing the desired equilibrium feedbacks U^0 and V^0. Based on this approach, one can develop the notion of equilibrium using optimal feedbacks for differential games $\Gamma_A = \Gamma_A^- \cup \Gamma_A^+$ and $\Gamma_B = \Gamma_B^- \cup \Gamma_B^+$ with payoffs J_A^∞ and J_B^∞. Let us note that, in the game Γ_A, the first player aims to maximize the functional $J_A^-(\xi(\cdot),\eta(\cdot))$ with the guarantee, using the feedback $U = u(t,\xi,\eta,\varepsilon)$, and the second player, as an antagonist, intends to minimize the functional $J_A^+(\xi(\cdot),\eta(\cdot))$, using the feedback $V = v(t,\xi,\eta,\varepsilon)$. In parallel, in the game Γ_B, the second player tries to maximize the functional $J_B^-(\xi(\cdot),\eta(\cdot))$ with the guarantee, and the first player, as an opponent, wishes to minimize the functional $J_B^+(\xi(\cdot),\eta(\cdot))$.

For a description of the dynamic equilibrium, we need the following notation. Let us denote by symbols $u_A^0 = u_A^0(t,\xi,\eta,\varepsilon)$ and $v_B^0 = v_B^0(t,\xi,\eta,\varepsilon)$ the feedbacks, which solve, respectively, the problem of guaranteed maximization of the payoff functionals J_A^- and J_B^-. It is worth noting, that such feedbacks are oriented of the guaranteed maximization of players' payoffs in the long run, and can be called "positive" feedbacks. In addition, we use the symbols $u_B^0 = u_B^0(t,\xi,\eta,\varepsilon)$ and $v_A^0 = v_A^0(t,\xi,\eta,\varepsilon)$ for denoting feedbacks that work most unfavorably for the opposing players. These feedbacks aim to minimize the payoff functionals J_B^+ and J_A^+ of the opponents. We call these strategies the "penalizing" feedbacks.

According to [4], the dynamic Nash equilibrium is formed by sticking together "positive" feedbacks u_A^0, v_B^0 and "penalizing" feedbacks u_B^0 and v_A^0 by the relations

$$U^0 = \begin{cases} u_A^0, & \text{if } \|(\xi,\eta) - (\xi_\varepsilon(t),\eta_\varepsilon(t))\| < \varepsilon, \\ u_B^0, & \text{otherwise,} \end{cases}$$

$$V^0 = \begin{cases} v_B^0, & \text{if } \|(\xi,\eta) - (\xi_\varepsilon(t),\eta_\varepsilon(t))\| < \varepsilon, \\ v_A^0, & \text{otherwise.} \end{cases}$$

In the next sections, we build "positive" feedbacks u_A^0 and v_B^0 for generating trajectories $(\xi^0(\cdot),\eta^0(\cdot))$ that lead the system to more favorable positions than static Nash equilibrium located in the interior of the game square by both the criteria

$$J_A^\infty(\xi^0(\cdot),\eta^0(\cdot)) \geq v_A, \quad J_B^\infty(\xi^0(\cdot),\eta^0(\cdot)) \geq v_B.$$

4. Optimal control problems for players

To construct "positive" feedbacks $u_A^0 = u_A^0(\xi,\eta)$ and $v_B^0 = v_B^0(\xi,\eta)$, we consider in this section an auxiliary two-step optimal control problem with average integral payoff functional for the first player in the case when actions of the second player are most unfavorable. For that, we analyze an optimal control problem for the dynamic system (2.1)

$$\begin{cases} \dot{\xi} = -\xi + u, & \xi(0) = \xi_0, \\ \dot{\eta} = -\eta + v, & \eta(0) = \eta_0 \end{cases} \tag{4.1}$$

with the payoff functional

$$J_A^f = \int_0^{T_f} g_A(\xi(t),\eta(t))dt.$$

Here, without loss of generality, we assume that $t_0 = 0$, $T = T_f$, and the terminal time $T_f = T_f(\xi_0,\eta_0)$ is determined by the condition of reaching the target set.

One can assume that the value of the static game equals to zero and the following conditions holds:

$$w_A = \frac{D_A}{C_A} = 0, \quad C_A > 0, \quad 0 < \xi_A = \frac{\alpha_2}{C_A} < 1, \quad 0 < \eta_A = \frac{\alpha_1}{C_A} < 1. \tag{4.2}$$

Let us consider the case when the initial conditions (ξ_0,η_0) of system (4.1) satisfy the following relations:

$$\xi_0 = \xi_A, \quad \eta_0 > \eta_A. \tag{4.3}$$

We suppose that the actions of the second player are mostly unfavorable to the first player. For trajectories of system (4.1) that start from the initial positions (ξ_0, η_0) (4.3), these actions are determined by the relation

$$v_A^0 = 0.$$

In this situation, the optimal actions u_A^0 of the first player according to the payoff functionals J_A^f can be presented as a two-step impulse control: it equals to unit from the initial time $t_0 = 0$ till the moment of optimal switch s and then equals to zero till the terminal time T_f:

$$u_A^0(t) = \begin{cases} 1 & \text{if} \quad t_0 \le t < s, \\ 0 & \text{if} \quad s \le t < T_f. \end{cases}$$

Here, the value s is the parameter of optimization. The terminal time T_f is determined from the following condition. The trajectory $(\xi(\cdot),\ \eta(\cdot))$ of system (4.1) that starts from the line on which $\xi(t_0) = \xi_A$ returns to this line when $\xi(T_f) = \xi_A$, which can be considered as the target set.

Let us consider two aggregates of characteristics. The first one is described by the system of differential equations with the value of the control parameter $u = 1$

$$\begin{cases} \dot{\xi} = -\xi + 1, \\ \dot{\eta} = -\eta, \end{cases} \tag{4.4}$$

solutions of which are determined by the Cauchy formula

$$\begin{cases} \xi(t) = (\xi_0 - 1)e^{-t} + 1, \\ \eta(t) = \eta_0 e^{-t}. \end{cases} \tag{4.5}$$

Here, the initial positions (ξ_0, η_0) satisfy conditions (4.3), and the time parameter t satisfies the inequality $0 \le t < s$.

The second aggregate of characteristics is given by the system of differential equations with the value of the control parameter $u = 0$:

$$\begin{cases} \dot{\xi} = -\xi, \\ \dot{\eta} = -\eta, \end{cases} \tag{4.6}$$

solutions of which are determined by the Cauchy formula

$$\begin{cases} \xi(t) = \xi_1 e^{-t}, \\ \eta(t) = \eta_1 e^{-t}. \end{cases} \tag{4.7}$$

Here, the initial positions $(\xi_1, \eta_1) = (\xi_1(s), \eta_1(s))$ are determined by the relations

$$\begin{cases} \xi_1 = \xi_1(s) = (\xi_0 - 1)e^{-s} + 1, \\ \eta_1 = \eta_1(s) = \eta_0 e^{-s}, \end{cases} \tag{4.8}$$

and the time parameter t satisfies the inequality $0 \le t < p$. Here, the terminal time $p = p(s)$ and the final position $(\xi_2, \eta_2) = (\xi_2(s), \eta_2(s))$ of the whole trajectory $(\xi(\cdot), \eta(\cdot))$ are given by the formulas

$$\xi_1 e^{-p} = \xi_A, \quad p = p(s) = \ln \frac{\xi_1(s)}{\xi_A}, \quad \xi_2 = \xi_A, \quad \eta_2 = \eta_1 e^{-p}. \tag{4.9}$$

The optimal control problem is to find such time s and the corresponding switching point $(\xi_1, \eta_1) = (\xi_1(s), \eta_1(s))$ on the trajectory $(\xi(\cdot), \eta(\cdot))$, where the integral $I = I(s)$ reaches its maximum,

$$\begin{aligned} I(s) &= I_1(s) + I_2(s), \\ I_1(s) &= \int_0^s \big(C_A((\xi_0 - 1)e^{-t} + 1)\eta_0 e^{-t} - \alpha_1((\xi_0 - 1)e^{-t} + 1) - \alpha_2 \eta_0 e^{-t} + a_{22}\big)dt, \\ I_2(s) &= \int_0^{p(s)} \big(C_A \xi_1(s)\eta_1(s)e^{-2t} - \alpha_1 \xi_1(s)e^{-t} - \alpha_2 \eta_1(s)e^{-t} + a_{22}\big)dt. \end{aligned} \tag{4.10}$$

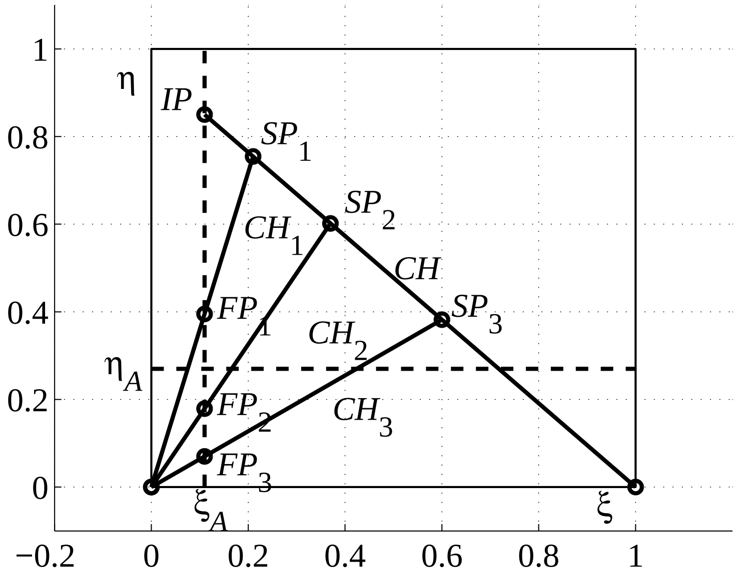


Figure 1. Characteristics of the Hamilton–Jacobi equation and the switching points.

Figure 1 shows the initial position IP chosen on the line $\xi = \xi_A$ with $\eta > \eta_A$, the characteristics CH oriented toward the vertex $(1, 0)$, the characteristics CH_1, CH_2, and CH_3 oriented toward the vertex $(0, 0)$, the switching points SP_1, SP_2, and SP_3 of the motion along the characteristics, and the endpoints FP_1, FP_2, and FP_3 of the motion located on the target line $\xi = \xi_A$.

5. Construction of switching lines

To solve the optimal control problem (4.4)–(4.10), we are based on the following algorithm. We use the necessary optimality conditions and calculate the derivative dI/ds, derive it as the function of optimal switching points $(\xi, \eta) = (\xi_1, \eta_1)$, equate this derivative to zero $dI/ds = 0$, and obtain the equation $F(\xi, \eta) = 0$ for the curve that consists of optimal switching points (ξ, η). This curve is called the switching line.

In the first stage, let us calculate the integrals I_1 and I_2:

$$
\begin{aligned}
I_1 = I_1(s) &= C_A(\xi_0 - 1)\eta_0 \frac{(1 - e^{-2s})}{2} + C_A\eta_0(1 - e^{-s}) - \alpha_1((\xi_0 - 1)(1 - e^{-s}) + s) \\
&\quad -\alpha_2\eta_0(1 - e^{-s}) + a_{22}s, \\
I_2 = I_2(s) &= C_A\xi_1(s)\eta_1(s)\frac{(1 - e^{-2p(s)})}{2} - \alpha_1\xi_1(s)(1 - e^{-p(s)}) - \alpha_2\eta_1(s)(1 - e^{-p(s)}) + a_{22}p(s).
\end{aligned}
$$

Next, we calculate the derivatives dI_1/ds and dI_2/ds and represent them as functions of optimal switching points $(\xi, \eta) = (\xi_1, \eta_1)$

$$\begin{aligned}
\frac{dI_1}{ds} &= C_A(\xi_0 - 1)\eta_0 e^{-2s} + C_A\eta_0 e^{-s} - \alpha_1\big((\xi_0 - 1)e^{-s} + 1\big) - \alpha_2\eta_0 e^{-s} + a_{22} \\
&= C_A\xi\eta - \alpha_1\xi - \alpha_2\eta + a_{22}, \\
\frac{dI_2}{ds} &= C_A\Big(\frac{d\xi}{ds}\eta\frac{(1-e^{-2p})}{2} + \xi\frac{d\eta}{ds}\frac{(1-e^{-2p})}{2} + \xi\eta e^{-2p}\frac{dp}{ds}\Big) - \alpha_1\frac{d\xi}{ds}(1-e^{-p}) \\
&\quad -\alpha_1\xi e^{-p}\frac{dp}{ds} - \alpha_2\frac{d\eta}{ds}(1-e^{-p}) - \alpha_2\eta e^{-p}\frac{dp}{ds} + a_{22}\frac{dp}{ds} \\
&= (C_A^2\xi^2\eta - \alpha_2^2\eta - 2C_A^2\xi^3\eta - 2\alpha_1 C_A\xi^2 + 2\alpha_1 C_A\xi^3 + 2\alpha_2 C_A\xi^2\eta + 2C_A a_{22}\xi - 2C_A a_{22}\xi^2)/(2C_A\xi^2).
\end{aligned}$$

In the latter equation, we use the following expressions for the derivatives $d\xi/ds$, $d\eta/ds$, and dp/ds and the exponents e^{-p}, e^{-2p}, $(1 - e^{-p})$, and $(1 - e^{-2p})$ as functions of the variables (ξ, η):

$$\frac{d\xi}{ds} = 1 - \xi, \quad \frac{d\eta}{ds} = -\eta, \quad \frac{dp}{ds} = \frac{1-\xi}{\xi},$$
$$e^{-p} = \frac{\alpha_2}{C_A\xi}, \quad e^{-2p} = \frac{\alpha_2^2}{C_A^2\xi^2}, \quad 1 - e^{-p} = \frac{C_A\xi - \alpha_2}{C_A\xi}, \quad 1 - e^{-2p} = \frac{C_A^2\xi^2 - \alpha_2^2}{C_A^2\xi^2}.$$

Transforming the derivatives dI_1/ds and dI_2/ds, we obtain the following equation for the switching line:

$$\frac{C_A^2\xi^2\eta - 2\alpha_1 C_A\xi^2 - \alpha_2^2\eta + 2C_A a_{22}\xi}{2C_A\xi^2} = 0.$$

Using the assumption that $w_A = 0$ (see (4.2)), we get the final expression for the switching line M_A^1:

$$\eta = \frac{2\alpha_1\xi}{C_A\xi + \alpha_2}.$$

The curve M_A^1 is a hyperbola that passes through the points $(0, 0)$, (ξ_A, η_A) and possesses the horizontal asymptote

$$\eta = \frac{2\alpha_1}{C_A}.$$

To complete the construction of the switching line M_A in the case when $C_A > 0$, we add a similar line M_A^2 to the line M_A^1 in the domain when $\eta \le \eta_A$:

$$\begin{aligned}
M_A &= M_A^1 \cup M_A^2, \qquad (5.11)\\
M_A^1 &= \Big\{(\xi, \eta) \in [0,1] \times [0,1] \colon \eta = \frac{2\alpha_1\xi}{C_A\xi + \alpha_2},\ \eta \ge \frac{\alpha_1}{C_A}\Big\}, \\
M_A^2 &= \Big\{(\xi, \eta) \in [0,1] \times [0,1] \colon \eta = -\frac{2(C_A - \alpha_1)(1-\xi)}{C_A(1-\xi) + (C_A - \alpha_2)} + 1,\ \eta \le \frac{\alpha_1}{C_A}\Big\}.
\end{aligned}$$

Let us note that, in the case when $C_A < 0$, the lines M_A, M_A^1, and M_A^2 are described by the formulas

$$\begin{aligned}
M_A &= M_A^1 \cup M_A^2, \qquad (5.12)\\
M_A^1 &= \Big\{(\xi, \eta) \in [0,1] \times [0,1] \colon \eta = \frac{2\alpha_1(1-\xi)}{C_A(1-\xi) + (C_A - \alpha_2)},\ \eta \ge \frac{\alpha_1}{C_A}\Big\}, \\
M_A^2 &= \Big\{(\xi, \eta) \in [0,1] \times [0,1] \colon \eta = -\frac{2(C_A - \alpha_1)\xi}{C_A\xi + \alpha_2} + 1,\ \eta \le \frac{\alpha_1}{C_A}\Big\}.
\end{aligned}$$

One can see that the line M_A divides the unit square $[0,1] \times [0,1]$ into two parts: the upper part

$$D_A^u \supset \big\{(\xi,\eta)\colon \xi = \xi_A,\ \eta > \eta_A\big\}$$

and the lower part

$$D_A^l \supset \big\{(\xi,\eta)\colon \xi = \xi_A,\ \eta < \eta_A\big\}.$$

The "positive" feedback u_A^0 has the following structure:

$$u_A^0 \quad = \quad u_A^0(\xi,\eta) = \begin{cases} \max\{0, -\operatorname{sgn}(C_A)\} & \text{if} \quad (\xi,\eta) \in D_A^u, \\ \max\{0, \operatorname{sgn}(C_A)\} & \text{if} \quad (\xi,\eta) \in D_A^l, \\ [0,1] & \text{if} \quad (\xi,\eta) \in M_A. \end{cases} \tag{5.13}$$

One can obtain the similar switching lines M_B for the second player whose profit is oriented on the payoff matrix B. For example, in the case when $C_B > 0$, the switching line M_B is presented by the relations

$$M_B = M_B^1 \cup M_B^2, \tag{5.14}$$

$$M_B^1 = \Big\{(\xi,\eta) \in [0,1] \times [0,1]\colon \eta = \frac{\beta_1 \xi}{2\beta_2 - C_B \xi},\ \xi \geq \frac{\beta_2}{C_B}\Big\},$$

$$M_B^2 = \Big\{(\xi,\eta) \in [0,1] \times [0,1]\colon \eta = -\frac{(C_B - \beta_1)(1-\xi)}{2(C_B - \beta_2) - C_B(1-\xi)} + 1,\ \xi \leq \frac{\beta_2}{C_B}\Big\}.$$

When the parameter C_B is negative, $C_B < 0$, the lines M_B, M_B^1, and M_B^2 are constructed by the formulas

$$M_B = M_B^1 \cup M_B^2, \tag{5.15}$$

$$M_B^1 = \Big\{(\xi,\eta) \in [0,1] \times [0,1]\colon \eta = -\frac{(C_B - \beta_1)\xi}{2\beta_2 - C_B \xi} + 1,\ \xi \geq \frac{\beta_2}{C_B}\Big\},$$

$$M_B^2 = \Big\{(\xi,\eta) \in [0,1] \times [0,1]\colon \eta = \frac{\beta_1(1-\xi)}{2(C_B - \beta_2) - C_B(1-\xi)},\ \xi \leq \frac{\beta_2}{C_B}\Big\}.$$

Similarly, the line M_B divides the unit square $[0,1] \times [0,1]$ into two parts: the left part

$$D_B^l \supset \big\{(\xi,\eta)\colon \xi < \xi_B,\ \eta = \eta_B\big\}$$

and the right part

$$D_B^r \supset \big\{(\xi,\eta)\colon \xi > \xi_B,\ \eta = \eta_B\big\}.$$

The "positive" feedback v_B^0 has the following structure:

$$v_B^0 = v_B^0(\xi,\eta) = \begin{cases} \max\{0, -\operatorname{sgn}(C_B)\} & \text{if} \quad (\xi,\eta) \in D_B^l, \\ \max\{0, \operatorname{sgn}(C_B)\} & \text{if} \quad (\xi,\eta) \in D_B^r, \\ [0,1] & \text{if} \quad (\xi,\eta) \in M_B. \end{cases} \tag{5.16}$$

6. Models of coordination games

Let us consider two different examples of coordination games.

The first example is the following. Two individuals (two species) compete for territory or a useful resource. Each player can choose one of the strategies: "hawk" or "dove" (see, for example, [16]). The names of the strategies are conditional, denoting only two types of behavior: enter into an aggressive conflict or retreat. In the asymmetric form of the game, we will consider the damage to the players to be different if they choose different strategies.

Let the first player be “the owner” and the second be “the invader” in the game of competing for territory. If both choose aggressive behavior, the damage will be considered the same and equal to 1, if both retreated — 0. In the event of an attack by “an invader”, the damage is equal to 4 and 3, respectively. In the aggressive behavior of “the owner”, the damage equals 3 and 5, respectively.

The matrix A reflects the damage of the first player, and the matrix B stands for the damage of the second player:

$$A = \begin{pmatrix} 1 & 3 \\ 4 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 5 \\ 3 & 0 \end{pmatrix}. \tag{6.1}$$

Let us present the main “game” parameters with the matrices A and B [15]:

$$\begin{gathered} C_A = a_{11} - a_{12} - a_{21} + a_{22} = -6, \\ \alpha_1 = a_{22} - a_{12} = -3, \quad \alpha_2 = a_{22} - a_{21} = -4, \\ \xi_A = \frac{\alpha_2}{C_A} = 0.67, \quad \eta_A = \frac{\alpha_1}{C_A} = 0.5, \end{gathered} \tag{6.2}$$

$$\begin{gathered} C_B = b_{11} - b_{12} - b_{21} + b_{22} = -7, \\ \beta_1 = b_{22} - b_{12} = -5, \quad \beta_2 = b_{22} - b_{21} = -3, \\ \xi_B = \frac{\beta_2}{C_B} = 0.43, \quad \eta_B = \frac{\beta_1}{C_B} = 0.71. \end{gathered} \tag{6.3}$$

In parallel, we consider the second example where we construct a modification of the previous coordination game “hawk” and “dove” with the following payoff matrices:

$$A = \begin{pmatrix} 10 & 0 \\ 7 & 23 \end{pmatrix}, \quad B = \begin{pmatrix} 19 & 0 \\ 4 & 11 \end{pmatrix}, \tag{6.4}$$

$$\begin{gathered} C_A = a_{11} - a_{12} - a_{21} + a_{22} = 26, \\ \alpha_1 = a_{22} - a_{12} = 23, \quad \alpha_2 = a_{22} - a_{21} = 16, \\ \xi_A = \frac{\alpha_2}{C_A} = 0.62, \quad \eta_A = \frac{\alpha_1}{C_A} = 0.88, \end{gathered} \tag{6.5}$$

$$\begin{gathered} C_B = b_{11} - b_{12} - b_{21} + b_{22} = 26, \\ \beta_1 = b_{22} - b_{12} = 11, \quad \beta_2 = b_{22} - b_{21} = 7, \\ \xi_B = \frac{\beta_2}{C_B} = 0.27, \quad \eta_B = \frac{\beta_1}{C_B} = 0.42. \end{gathered} \tag{6.6}$$

7. Feedback strategies and equilibrium trajectories

In this section, we provide feedback strategies and equilibrium trajectories for the given examples of the “hawk”–“dove” game based on the solution constructions given in formulas (5.11)–(5.16).

The structure of the dynamic Nash equilibrium of the first example (6.1)–(6.3) is presented in Figure 2. Here we depict the saddle points S_A and S_B of the static game, points of the static Nash equilibria NE_1, NE_2, and NE_3, and the switching lines $M_A = M_A^1 \cup M_A^2$ and $M_B = M_B^1 \cup M_B^2$. The equilibrium trajectories start from the initial points IP_1, IP_2, and IP_3, then move along characteristics of the Hamilton–Jacobi equations, meet the switching lines where they change orientation, and converge to the final points FP_1, FP_2, and FP_3.

The values of players' payoff functionals at the final points of the motion of the equilibrium trajectories are the following: $g_A(FP_1) = g_A(FP_2) = 3$, $g_B(FP_1) = g_B(FP_2) = 5$, $g_A(FP_3) = 4$,

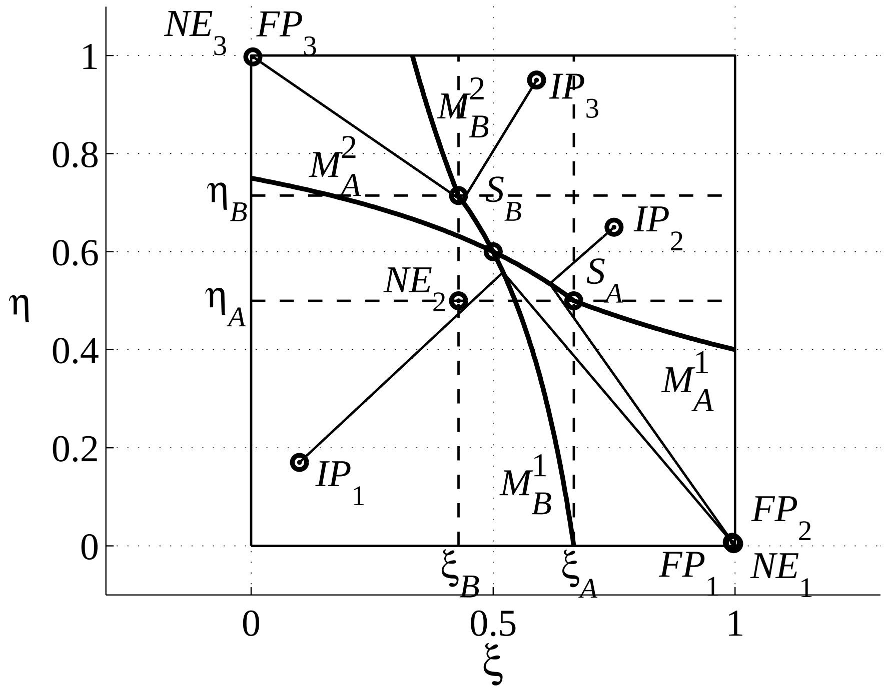


Figure 2. Equilibrium trajectories in the game with average integral payoffs (Example 1).

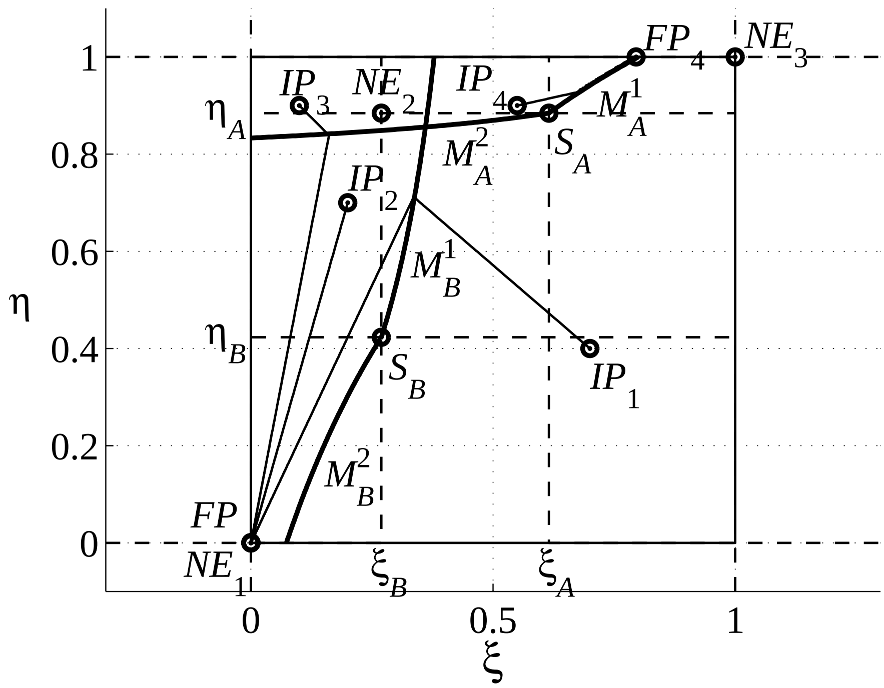


Figure 3. Equilibrium trajectories in the game with average integral payoffs (Example 2).

and $g_B(FP_3) = 3$. Let us note that these values majorate the payoffs at the point of the static Nash equilibrium NE_2: $g_A(NE_2) = 2$ and $g_B(NE_2) = 2.14$.

The structure of the dynamic Nash equilibrium of the second example (6.4)–(6.6) is presented in Figure 3. Here we depict the saddle points S_A and S_B of the static game, points of the static Nash equilibria NE_1, NE_2, and NE_3, and the switching lines $M_A = M_A^1 \cup M_A^2$ and $M_B = M_B^1 \cup M_A^B$. Equilibrium trajectories start from the initial points IP_1, IP_2, IP_3, and IP_4, then move along characteristics of the Hamilton–Jacobi equations, meet the switching lines where they change orientation, and converge to the final points FP and FP_4. Let us note that the final point FP_4 does not coincide with the Nash equilibrium NE_3.

The values of players' payoff functionals at the final points of the motion of the equilibrium trajectories are the following: $g_A(FP) = 23, g_B(FP) = 11$, $g_A(FP_4) = 9.39$, and $g_B(FP_4) = 15.93$. Let us note that these values majorate the payoffs at the point of the static Nash equilibrium NE_2: $g_A(NE_2) = 6.91$ and $g_B(NE_2) = 7.11$. At the point FP_4 located on the boundary of the game square, guaranteed strategies provide a result that brings closer the interests of the players.

8. Replicator dynamics

In this section, we present the structure of the replicator dynamics.

The general view of replicator dynamics for the dynamic bimatrix game can be presented as follows (see, for example, [2, 3, 16]):

$$\begin{cases} \dot{u_i} = u_i\big((\mathbf{Av})_i - (\mathbf{u}, \mathbf{Av})\big), \\ \dot{v_j} = v_j\big((\mathbf{Bu})_j - (\mathbf{v}, \mathbf{Bu})\big), \quad 1 \le i \le n, \quad 1 \le j \le n. \end{cases} \tag{8.1}$$

Here, the vectors $\mathbf{u} = (u_1, \dots, u_n)$ and $\mathbf{v} = (v_1, \dots, v_n)$ describe the system state. The symbols $(\mathbf{Av})_i$ and $(\mathbf{Bu})_j$ stand for the fitness of the corresponding type. An average fitness is defined as follows:

$$(\mathbf{u}, \mathbf{Av}) = \sum_{i=1}^{n} u_i(\mathbf{Av})_i, \quad (\mathbf{v}, \mathbf{Bu}) = \sum_{i=1}^{n} v_i(\mathbf{Bu})_i.$$

System (8.1) is consistent with one of the basic principles of Darwinism: the reproductive success of an individual or a group depends on the advantage of one's fitness over the population's average fitness.

Let us present the main characteristics of the replicator systems of the type (8.1).

The Jacobi matrix at the stationary point (the static Nash equilibrium) in the general case has the form [3]:

$$\mathbf{J} = \begin{bmatrix} \mathbf{0} & \mathbf{C} \\ \mathbf{D} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{0}$ is the zero submatrix of size $(n-1) \times (n-1)$, and $\mathbf{C}$ and $\mathbf{D}$ are submatrices formed by some constant coefficients.

The characteristic polynomial of the system has the form

$$p(\lambda) = \det(\lambda^2 \mathbf{I} - \mathbf{DC}).$$

From the structure of the characteristic polynomial, it follows that, in the two-dimensional case, the system cannot have a stationary point of the focus or node type.

For the dynamic bimatrix 2×2 game, the replicator dynamics can be written in the form of the system of differential equations of the second order:

$$\begin{cases} \dot{\xi}(t) = \xi(t)\big(1 - \xi(t)\big)\big(C_A\eta(t) - \alpha_1\big), \quad \xi(t_0) = \xi_0, \\ \dot{\eta}(t) = \eta(t)\big(1 - \eta(t)\big)\big(C_B\xi(t) - \beta_2\big), \quad \eta(t_0) = \eta_0. \end{cases} \tag{8.2}$$

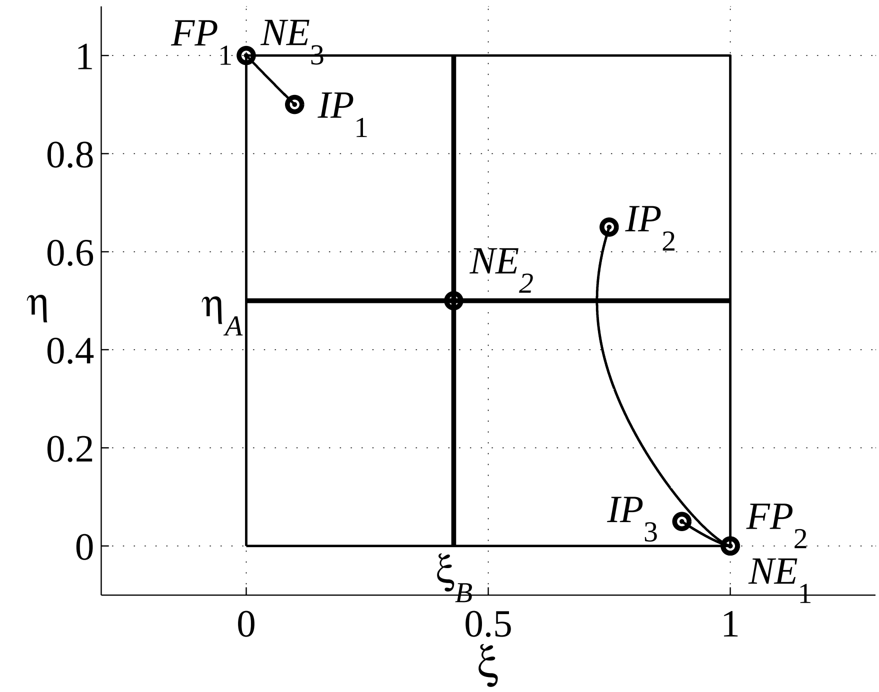


Figure 4. Trajectories of replicator dynamics (Example 1).

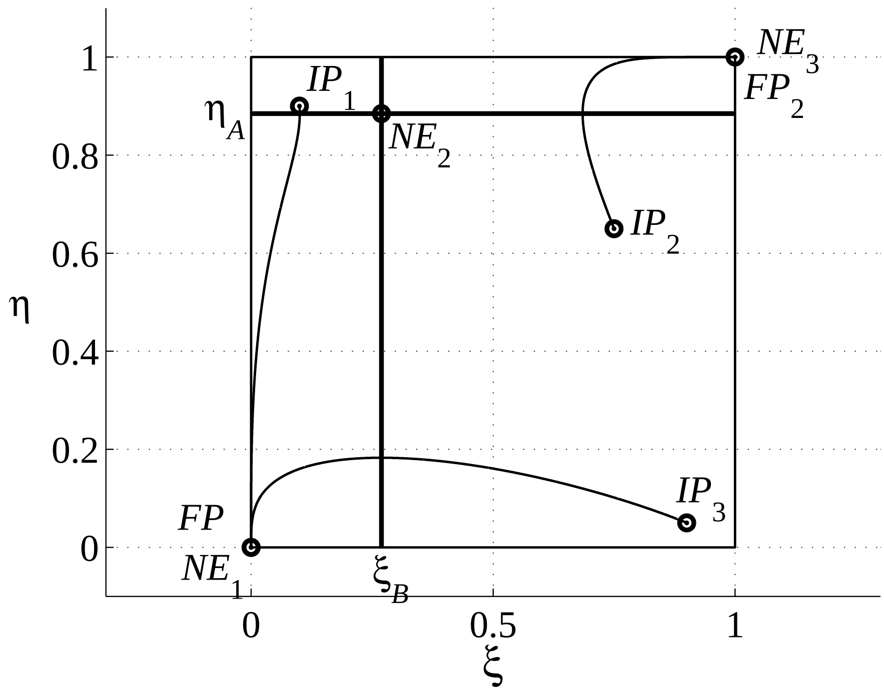


Figure 5. Trajectories of replicator dynamics (Example 2).

Figure 4 presents the trajectories of the replicator dynamics for the first example. They start from the initial points IP_1, IP_2, and IP_3 and tend to the final points FP_1 and FP_2, which coincide with the Nash equilibria NE_1 and NE_3.

Figure 5 presents the trajectories of the replicator dynamics for the second example. They start from the initial points IP_1, IP_2, and IP_3 and terminate their motion at the points FP and FP_2 matching with the Nash equilibria NE_1 and NE_3.

9. Mixed dynamics

In this section, we consider mixed dynamics when the first player uses the guaranteed strategy with switching line M_A (5.11), (5.12) that has the form $u_A^0 = u_A^0(\xi(t), \eta(t))$ (5.13), and the strategy of the second player is formed by the replicator dynamics (8.2):

$$\begin{cases} \dot{\xi}(t) = -\xi(t) + u_A^0(\xi(t), \eta(t)), & \xi(t_0) = \xi_0, \\ \dot{\eta}(t) = \eta(t)\big(1 - \eta(t)\big)\big(C_B\xi(t) - \beta_2\big), & \eta(t_0) = \eta_0. \end{cases}$$

Figure 6 presents the mixed dynamics for the first example. Here we show the switching line $M_A = M_A^1 \cup M_A^2$ for the control of the first player and the switching line $\xi = \xi_B$ for the control of the second player related to the replicator dynamics. The trajectories of the mixed dynamics start from the initial points IP_1, IP_2, and IP_3, switch control on the line M_A, and converge to the final points FP_1 and FP_2.

Figure 7 presents the mixed dynamics for the second example. Here we show the switching line $M_A = M_A^1 \cup M_A^2$ for the control of the first player and the switching line $\xi = \xi_B$ for the control of the second player formed by the replicator dynamics. The trajectories of the mixed dynamics start from the initial points IP_1, IP_2, IP_3, and IP_4, have a control switch on the line M_A, and converge to the final points FP_1 and FP_2.

In the second example, the mixed dynamics demonstrate that guaranteed strategies can provide convergence to the final points, for instance, to the final point FP_1, which differs from the Nash equilibrium NE_3 and gives the payoff results with closer interests of the players.

10. Conclusion

An analysis of the behavior of equilibrium trajectories is provided for the 2×2 dynamic bimatrix coordination game. First, trajectories of the dynamic Nash equilibrium are constructed within the approach of guaranteed strategies in the sense of N.N. Krasovskii in combination with the L.S. Pontryagin maximum principle. Second, an analysis is provided for the replicator dynamics whose trajectories converge to the static Nash equilibrium points located in the vertices of the game square. Third, computational experiments are carried out for the mixed dynamics in which we couple the strategies of the considered dynamics: the strategies of the dynamic Nash equilibrium and the replicator dynamics. Finally, the comparison results are presented for equilibrium trajectories of the considered dynamics.

REFERENCES

1. Arnold V. I. Optimization in mean and phase transitions in controlled dynamical systems. *Funct. Anal. Appl.*, 2002. Vol. 36, No. 2. P. 83–92. DOI: 10.1023/A:1015655005114
2. Bratus A. S., Novozhilov A. S., Platonov A. P. *Dinamicheskiye systemy i modeli biologii* [Dynamic Systems and Models of Biology]. Moscow: Fizmatlit, 2010. 400 p. (in Russian)

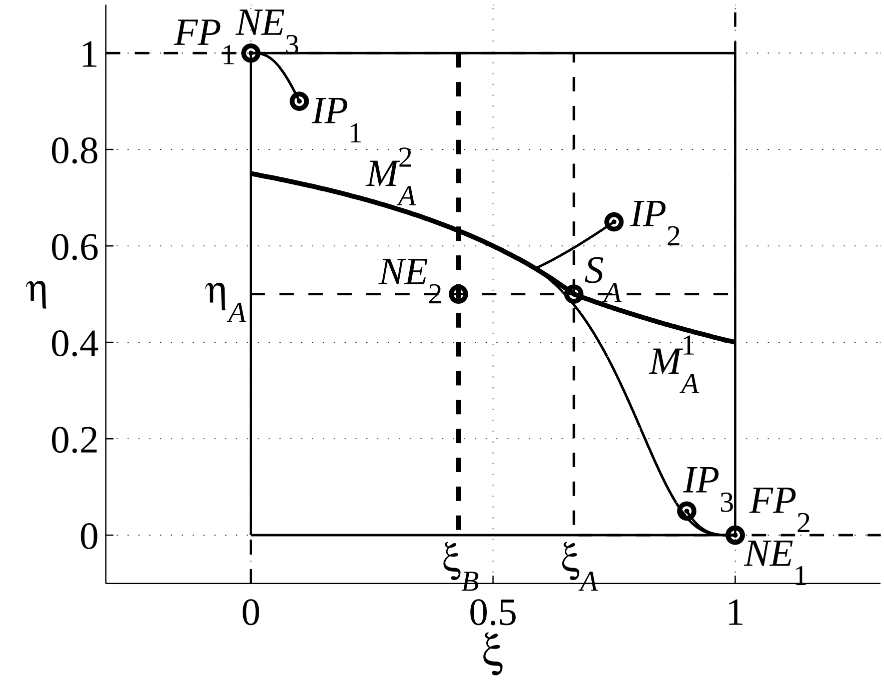


Figure 6. Equilibrium trajectories of mixed dynamics (Example 1).

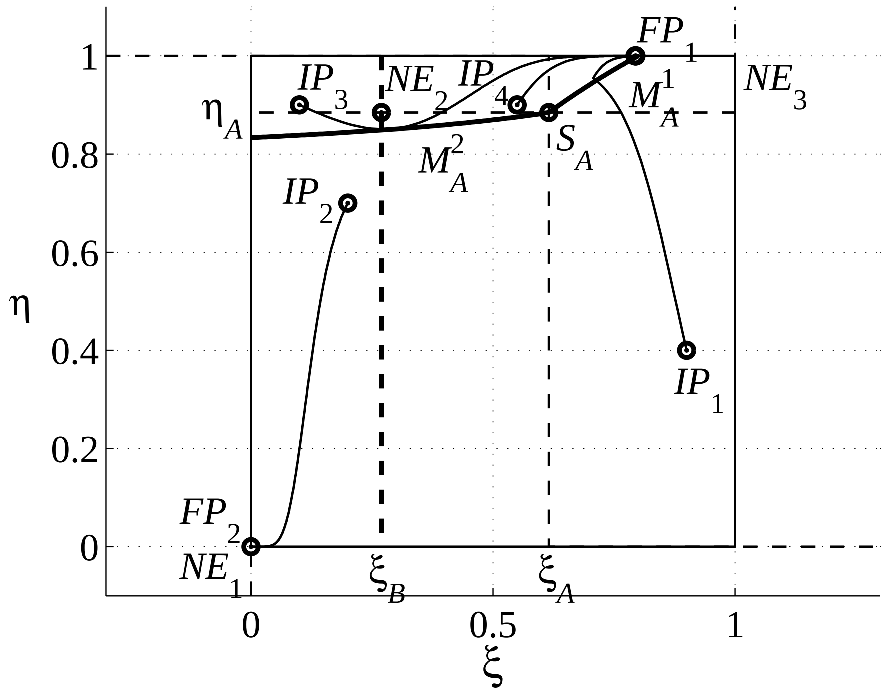


Figure 7. Equilibrium trajectories of mixed dynamics (Example 2).

3. Hofbauer J., Sigmund K. *The Theory of Evolution and Dynamical Systems*. Cambridge etc.: Cambridge Univ. Press, 1988. 341 p. DOI: 10.1002/zamm.19900700210
4. Kleimenov A. F. *Neantagonisticheskiye pozitsionniye differentsial'niye igry* [Non-antagonistic Positional Differential Games]. Yekaterinburg: Nauka, 1993. 185 p. (in Russian)
5. Krasovskii A. N., Krasovskii N. N. *Control Under Lack of Information*. Boston: Burkhäuser, 1995. 322 p. DOI: 10.1007/978-1-4612-2568-3
6. Krasovskii N. A., Tarasyev A. M. Equilibrium trajectories in dynamical bimatrix games with average integral payoff functionals. *Autom. Remote Control*, 2018. Vol. 79, No. 6. P. 1148–1167. DOI: 10.1134/S0005117918060139
7. Krasovskii N. N. *Upravleniye dinamicheskoy sistemoy* [Control of Dynamic System]. Moscow: Nauka, 1985. 520 p. (in Russian)
8. Krasovskii N. N., Subbotin A. I. *Game–Theoretical Control Problems*. New-York: Springer–Verlag, 1988. 517 p.
9. Mazalov V. V., Rettieva A. N. Application of bargaining schemes for equilibrium determination in dynamic games. *Mat. Teor. Igr Pril.*, 2023. Vol. 15., No. 2. P. 75–88. (in Russian)
10. Mertens J.-F., Sorin S., Zamir S. *Repeated Games*. Cambridge: Cambridge University Press, 2015. 567 p.
11. Petrosjan L. A., Zenkevich N. A. Conditions for sustainable cooperation. *Autom. Remote Control*, 2015. Vol. 76. P. 1894–1904. DOI: 10.1134/S0005117915100148
12. Pontryagin L. S., Boltyanskii V. G., Gamkrelidze R. V., Mischenko E. F. *The Mathematical Theory of Optimal Processes*. New-York: Wiley Interscience, 1962. 360 p.
13. Sorin S. Replicator dynamics: old and new. *J. Dynam. Games*, 2020. Vol. 7, No. 4. P. 365–386. DOI: 10.3934/jdg.2020028
14. Vinnikov E. V., Davydov A. A., Tunitskiy D. V. Existence of maximum of time averaged harvesting in the KPP-model on sphere with permanent and impulse harvesting. *Dokl. Math.*, 2023. Vol. 108, No. 3. P. 472–476. DOI: 10.1134/S1064562423701387
15. Vorobyev N. N. *Teoriya igr dlya ekonomistov–kibernetikov* [The Theory of Games for Economists–Cyberneticians]. Moscow: Nauka, 1985. 272 p. (in Russian)
16. Yakushkina T. S. A distributed replicator system corresponding to a bimatrix game. *Moscow Univ. Comput. Math. Cybernet.*, 2016. Vol. 40, No. 1. P. 19–27. DOI: 10.3103/S0278641916010064

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 107–120

DOI: 10.15826/umj.2024.2.010

INTERPOLATION WITH MINIMUM VALUE OF L_2-NORM OF DIFFERENTIAL OPERATOR

Sergey I. Novikov

Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences,
16 S. Kovalevskaya Str., Ekaterinburg, 620108, Russian Federation

Sergey.Novikov@imm.uran.ru

Abstract: For the class of bounded in l_2-norm interpolated data, we consider a problem of interpolation on a finite interval $[a,b]\subset\mathbb{R}$ with minimal value of the L_2-norm of a differential operator applied to interpolants. Interpolation is performed at knots of an arbitrary N-point mesh $\Delta_N:\ a\le x_1<x_2<\cdots<x_N\le b$. The extremal function is the interpolating natural $\mathcal{L}$-spline for an arbitrary fixed set of interpolated data. For some differential operators with constant real coefficients, it is proved that on the class of bounded in l_2-norm interpolated data, the minimal value of the L_2-norm of the differential operator on the interpolants is represented through the largest eigenvalue of the matrix of a certain quadratic form.

Keywords: Interpolation, Natural $\mathcal{L}$-spline, Differential operator, Reproducing kernel, Quadratic form.

1. Introduction

Let N be any positive integer, $1\le q<\infty$, and

$$\mathfrak{M}_{N,q}=\Big\{z:\ z=\{z_j\}_{j=1}^N,\Big(\sum_{j=1}^N|z_j|^q\Big)^{1/q}\le 1\Big\}$$

be a class of interpolated values that is the unit ball in the space l_q^N.

Let $[a,b]\subset\mathbb{R}$ be an arbitrary finite interval and $W_q^m[a,b]$ be the standard Sobolev space equipped with the norm

$$\|f\|_{W_q^m[a,b]}=\|f\|_q+\sum_{j=1}^m\|f^{(j)}\|_q, \tag{1.1}$$

where $\|f\|_q$ is the usual L_q-norm of a function f on $[a,b]$.

Let $D=d/dx$ be the operator of differentiation, I be the identical operator, and

$$\mathcal{L}_m(D)=D^m+a_{m-1}D^{m-1}+\ldots+a_1D+a_0I$$

be a linear differential operator of order m with constant real coefficients. Denote by $p_m:=p_m(x)$ the characteristic polynomial of the differential operator $\mathcal{L}_m(D)$:

$$p_m(x)=x^m+a_{m-1}x^{m-1}+\cdots+a_1x+a_0.$$

We restrict our attention to the case when $p_m(x)$ has only real roots $\{\beta_j\}_{j=1}^m$. This means that the differential operator $\mathcal{L}_m(D)$ has the factorization into a product of differential operators of the first order, i.e.,

$$\mathcal{L}_m(D)=(D-\beta_1I)(D-\beta_2I)\cdots(D-\beta_mI). \tag{1.2}$$

We interpolate at the knots of an arbitrary fixed mesh of N points from the interval $[a, b]$

$$\Delta_N : \ a \le x_1 < x_2 < \ldots < x_N \le b, \quad (a > -\infty, \ b < +\infty).$$

For an arbitrary fixed $z \in \mathfrak{M}_{N,q}$, we introduce the quantity

$$K_{N,q}(z) = \inf_{\substack{f \in W_q^m[a,b] \\ f(x_k) = z_k, \ k=\overline{1,N}}} \|\mathcal{L}_m(D) f\|_q. \tag{1.3}$$

The problem of finding quantity (1.3) is known as the Favard type interpolation problem (see [2, 4, 20], and the references therein).

In the paper, we study an analog of quantity (1.3) for the class $\mathfrak{M}_{N,q}$ of interpolated data, namely

$$\mathfrak{B}^q_{\mathcal{L}_m}(\Delta_N) = \sup_{z \in \mathfrak{M}_{N,q}} K_{N,q}(z). \tag{1.4}$$

Problem (1.4) can also be interpreted as the Favard type interpolation problem, but considered for the entire class of interpolated data. For the differential operator $\mathcal{L}_m(D) = D^m$, quantity (1.4) was found by the author [9] in the case of $q = 2$.

Problem (1.4) is close to extremal interpolation problems (see [15–18] and the references therein). However, the set of interpolated data in our setting (1.4) is given by the constraint imposed on the interpolated values $z = (z_1, z_2, \ldots, z_N)$, but not on their finite or divided differences.

In the present paper, we consider problems (1.3) and (1.4) only for $q = 2$. For this reason, index 2 in $\mathfrak{M}_{N,2}$, $K_{N,2}(z)$, and $\mathfrak{B}^2_{\mathcal{L}_m}(\Delta_N)$ will be omitted.

The main result of the paper is Theorem 1, in which we give the exact value of the quantity $\mathfrak{B}_{\mathcal{L}_m}(\Delta_N)$. This exact value is expressed in terms of the largest eigenvalue of the matrix of a quadratic form of interpolated data.

For $q = 2$, the extremal function in (1.3) is known. This function is a specific spline, which is called an interpolation natural $\mathcal{L}$-spline. This fact is a particular case of results of the variation spline theory.

The paper is organized as follows. Section 2 is devoted to the Favard type interpolation problem (1.3) considered from the point of view of general approaches. In Section 3, we write two representations of interpolation natural $\mathcal{L}$-splines. In Section 4, we prove the lemmas needed to prove the main result. In Section 5, we formulate and prove the main result of the paper. Section 6 is devoted to discussions and some comments.

2. On Favard-type interpolation problems

Consider problem (1.3) in the case of $q = 2$. As shown in [20], (1.3) is one of convex programming problems. In [20, p.87], it is proved that there exists a solution to (1.3), and for $N > m$, the solution is unique. As noted above, an extremal function in (1.3) is known. To write this function, we need some known results of the variation theory of splines (see for example, [1, 14], and the references therein).

We first introduce some notation. Let X be a real Hilbert space of functions with a norm $\|\cdot\|$, let $T : X \to X$ be a bounded linear operator, and let $\ker T$ be its null-space, i.e., the set of functions $\varphi \in X$ such that $T\varphi = 0$. By X^*, we denote the conjugate space of X. Let $\varphi_i \in X^*$ $(i = 1, 2, \ldots, N)$, i.e., let φ_i be bounded linear functionals on X. For every $\tau \in X$, we set $A\tau = (\varphi_1(\tau), \varphi_2(\tau), \ldots, \varphi_N(\tau))$. This means that A is a linear operator that maps a function $\tau \in X$ onto an N-dimensional vector consisting of values of the functionals $\{\varphi_i\}_{i=1}^N$ on this function.

As such functionals, we take values at the points of the mesh $\Delta_N = \{x_i\}_{i=1}^N$, i.e., we set $\varphi_i(\tau) = \tau(x_i)$, $(i = 1, 2, \dots, N)$. Thus, we have

$$A\tau = (\tau(x_1), \tau(x_2), \dots, \tau(x_N)).$$

The abstract Favard-type interpolation problem is to find the quantity

$$\mathcal{B}(T, z) = \inf_{\substack{Af=z \\ f\in X}} \|Tf\|_X^2. \tag{2.1}$$

Following [14, p. 77], a real Hilbert space X with a norm $\|\cdot\|$ and a seminorm $\rho(\,\cdot\,)$ is called an S-space if the following conditions hold:

(i) the seminorm ρ is bounded in X; i.e., for every $\tau \in X$, the inequality $\rho(\tau) \le C\|\tau\|$ is true with some constant $C > 0$ independent of τ;

(ii) X is complete with respect to the seminorm ρ.

It is proved (see [1] and [14, Sect. 5.15]) that if the space X is an S-space and is continuously embedded into the space of continuous functions, then the extremal function of (2.1) has the following form:

$$\sigma(x) = q(x) + \sum_{j=1}^N \lambda_j G_m(x, x_j). \tag{2.2}$$

Here $q \in \ker T$, $G_m(x,\cdot)$ is a reproducing kernel of the S-space X (see [1, 14]), and the scalars $\{\lambda_j\}_{j=1}^N$ are determined from the condition

$$\sum_{j=1}^N \lambda_j u(x_j) = 0 \qquad \forall u \in \ker T. \tag{2.3}$$

Problem (1.3) is a particular case (up to root-squaring) of (2.1) when $X = W_2^m[a,b]$ with norm (1.1), $T = \mathcal{L}_m(D): \; W_2^m[a,b] \to L_2[a,b]$, and $\tau = f$. In our case, the seminorm $\rho(\cdot)$ is defined as

$$\rho(f) = \left(\int_a^b |\mathcal{L}_m(D)f(t)|^2 dt\right)^{1/2}.$$

Since the seminorm $\rho(\cdot)$ is estimated through the coefficients of the differential operator $\mathcal{L}_m(D)$ as

$$\rho(f) = \|\mathcal{L}_m(D)f\|_2 \le \max\left\{1, |a_0|, |a_1|, \dots, |a_{m-1}|\right\} \|f\|_{W_2^m[a,b]},$$

the seminorm $\rho(f)$ is bounded.

The operator $\mathcal{L}_m(D)$ acts "onto" $L_2[a,b]$, and the space $L_2[a,b]$ is complete. Therefore, $W_2^m[a,b]$ is complete with respect to the seminorm ρ. Thus, $W_2^m[a,b]$ is an S-space. In addition, the space $W_2^m[a,b]$ is continuously embedded into the space $C[a,b]$ of continuous functions (the Sobolev embedding theorem).

For finding the reproducing kernel of S-space $W_2^m[a,b]$, we introduce two subspaces

$$U_m = \left\{f \in W_2^m[a,b]: \quad f^{(i)}(a) = 0, \; i = 0, 1, \dots, m-1\right\}$$

and

$$V_m = \left\{f \in W_2^m[a,b]: \quad f^{(i)}(b) = 0, \; i = 0, 1, \dots, m-1\right\}.$$

Each of the subspaces U_m and V_m has codimension m. Also, it is easy to see that $\big(\ker \mathcal{L}_m(D)\big) \cap U_m = \{0\}$ and $\big(\ker \mathcal{L}_m(D)\big) \cap V_m = \{0\}$. From these simple facts, it follows that $W_2^m[a,b] = (\ker \mathcal{L}_m(D)) \cup U_m$ and $W_2^m[a,b] = (\ker \mathcal{L}_m(D)) \cup V_m$.

Let $\mathcal{L}_m^*(D)$ be a linear differential operator that is formal adjoint to the operator $\mathcal{L}_m(D)$, i.e., $\mathcal{L}_m^*(D) = \mathcal{L}_m(-D)$. Now, we introduce the differential operator $\mathcal{L}_{2m}(D)$ of order $2m$ as follows

$$\mathcal{L}_{2m}(D) = \mathcal{L}_m(D)\ \mathcal{L}_m^*(D).$$

From (1.2), we have

$$\mathcal{L}_{2m}(D) = (-1)^m(D^2 - \beta_1^2 I)(D^2 - \beta_2^2 I)\cdots(D^2 - \beta_m^2 I).$$

For the S-space $W_2^m[a,b]$, the reproducing kernel is coordinated with the subspaces U_m and V_m and is built through a fundamental solution of the differential operator $\mathcal{L}_{2m}(D)$ (see, e.g., [14, Ch. 5]).

As is known (see, e.g., [21, Ch. III]), the fundamental solution of a differential operator $\mathcal{L}(D)$ is a distribution $\mathcal{E}$ satisfying $\mathcal{L}(D)\mathcal{E} = \delta$, where δ is the Dirac δ-function (or δ-distribution). The fundamental solution is defined up to a summand that is an arbitrary solution of the equation $\mathcal{L}(D)y(t) = 0$. We will assume that this summand is identically equal to zero. Distributions are understood as linear continuous functionals in the space of infinitely differentiable functions with compact supports.

The following result is known.

Lemma 1 (see, e.g., [21, p. 114]). *Let $\mathcal{L}_r(D)$ be an arbitrary linear differential operator of order $r \geq 2$ with constant real coefficients. Then the fundamental solution of this operator has the form*

$$\mathcal{E}_r(t) = \theta(t)\ Z_r(t),$$

where $Z_r(t)$ is a unique solution to the initial value problem

$$\begin{cases} \mathcal{L}_r(D)Z_r(t) = 0, \\ Z_r(0) = Z_r'(0) = \cdots = Z_r^{(r-2)}(0) = 0, \\ Z_r^{(r-1)}(0) = 1 \end{cases}$$

and $\theta(t)$ is the Heaviside function

$$\theta(t) = \begin{cases} 1, & t > 0, \\ 0, & t \leq 0. \end{cases}$$

Now, we set $r = 2m$ and apply Lemma 1 to the differential operator $\mathcal{L}_{2m}(D)$. Since it has the leading coefficient $(-1)^m$, its fundamental solution is

$$\mathcal{E}_{2m}(t) = (-1)^m\ \theta(t)\ Z_{2m}(t), \tag{2.4}$$

where

$$\begin{cases} (D^2 - \beta_1^2 I)(D^2 - \beta_2^2 I)\cdots(D^2 - \beta_m^2 I)Z_{2m}(t) = 0, \\ Z_{2m}(0) = Z_{2m}'(0) = \cdots = Z_{2m}^{(2m-2)}(0) = 0, \\ Z_{2m}^{(2m-1)}(0) = 1. \end{cases}$$

Based on [14, Sect. 5.13], we will prove that the function $\mathcal{E}_{2m}(x-t)$ is the reproducing kernel of the S-space $W_2^m[a,b]$.

Lemma 2. $\mathcal{E}_{2m}(x-t) \in U_m$ *for any fixed $t \in [a,b]$.*

P r o o f. From (2.4), we have

$$\mathcal{E}_{2m}^{(i)}(u) = (-1)^m \ \theta(\tau) \ Z_{2m}^{(i)}(u) \quad (i = 0, 1, \ldots, 2m-1). \tag{2.5}$$

Now, we set $u = x - t$. From the definition of the Heaviside function, we see that

$$\mathcal{E}_{2m}^{(i)}(a-t) = 0, \qquad (i = 0, 1, \ldots, 2m-1),$$

i.e., $\mathcal{E}_{2m}(x-t) \in U_m$ for any fixed $t \in [a, b]$. □

Lemma 3. *The following equality holds for any fixed $t \in [a, b]$ and any function $f \in V_m$:*

$$\int_a^b \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big) \ \mathcal{L}_m(D) f(x) dx = f(t).$$

P r o o f. Let $a < t < b$. We write the integral on the left-hand side as the sum of two integrals

$$\int_a^b \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big) \ \mathcal{L}_m(D) f(x) \ dx = I_1 + I_2,$$

where

$$I_1 = \int_a^t \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big) \ \mathcal{L}_m(D) f(x) dx, \quad I_2 = \int_t^b \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big) \ \mathcal{L}_m(D) f(x) dx.$$

Changing the variable $u = x - t$ and noting that $\mathcal{E}_{2m}(u) = 0$ for all $u \le 0$, we have

$$I_1 = \int_{a-t}^0 \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(u)\big) \ \mathcal{L}_m(D) f(u+t) \ du = 0.$$

Integrating I_2 by parts, we obtain

$$I_2 = w(b) - w(t) + \int_t^b \big(\mathcal{L}_{2m}(D)\mathcal{E}_{2m}(x-t)\big) f(x) dx,$$

where

$$w(x) = f^{(m-1)}(x) \ \mathcal{L}_m(D)\mathcal{E}_{2m}(x-t) + f^{(m-2)}(x) \ \Big[a_{m-1} \ \mathcal{L}_m(D)\mathcal{E}_{2m}(x-t) - \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big)'_x\Big]$$
$$+ \cdots + f'(x) \ \sum_{\nu=0}^{m-2} (-1)^\nu a_{\nu+2} \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big)_x^{(\nu)} + f(x) \ \sum_{\nu=0}^{m-1} (-1)^\nu a_{\nu+1} \big(\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)\big)_x^{(\nu)},$$

and $\{a_\nu\}_{\nu=0}^m$, $a_m = 1$, are the constant real coefficients in the standard representation of the differential operator $\mathcal{L}_m(D)$.

By the definition of the set V_m, we have $w(b) = 0$.

From (2.5) for $i = 0, 1, \ldots, m-1$, it follows that

$$\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)|_{x=t} = 0.$$

From (2.5) for $i = m, m+1, \ldots, 2m-1$, we conclude that all derivatives of $\mathcal{L}_m(D)\mathcal{E}_{2m}(x-t)$ in $w(x)$ are equal to zero when $x = t$. Therefore, $w(t) = 0$.

Using the definition of the fundamental solution and one of the known properties of the Dirac δ-function (see, for example, [21, p. 134]), we finally have

$$I_2 = \int_t^b \big(\mathcal{L}_{2m}(D)\mathcal{E}_{2m}(x-t)\big) f(x) dx = \int_t^b \delta(x-t) f(x) dx = f(t).$$

The cases $t = a$ and $t = b$ are easily checked. Lemma 3 is proved. □

From [14, Sect. 5.13] and Lemmas 2 and 3, we obtain the following statement.

Lemma 4. $G_m(x,t) = \mathcal{E}_{2m}(x-t)$.

Lemma 4 together with (2.2) and (2.3) give the following expression of the extremal function $\sigma(x)$ in problem (1.3):

$$\sigma(x) = q(x) + \sum_{j=1}^{N} \lambda_j \mathcal{E}_{2m}(x - x_j), \tag{2.6}$$

where $q \in \ker \mathcal{L}_m(D)$ and there are additional conditions

$$\sum_{j=1}^{N} \lambda_j g_\nu(x_j) = 0 \quad (\nu = 1, 2, \ldots, m), \tag{2.7}$$

for finding parameters $\{\lambda_j\}_{j=1}^N$. Here, the set of functions $\{g_\nu(x)\}_{\nu=1}^m$ is a basis in $\ker \mathcal{L}_m(D)$.

From (2.4), it follows that

$$\mathcal{E}_{2m}(x - x_j) \in C^{2m-2}(\mathbb{R}) \quad (j = 1, 2, \ldots, N).$$

Therefore, the function $\sigma(x)$ has the following properties:

(1) $\sigma \in C^{2m-2}(\mathbb{R})$;
(2) $\mathcal{L}_{2m}(D)\sigma(x) = 0$ for all $x \in (x_j, x_{j+1})$ $(j = 1, 2, \ldots, N-1)$;
(3) $\sigma(x_i) = z_i$ $(i = 1, 2, \ldots, N)$;
(4) $\sigma \in \ker \mathcal{L}_m(D)$ for $x \leq x_1$ and $x \geq x_N$.

Properties (1)–(3) mean that $\sigma(x)$ is an interpolating $\mathcal{L}$-spline corresponding to the differential operator $\mathcal{L}_{2m}(D)$ with knots at the points of the mesh Δ_N and has the minimal defect. Due to property (4), the $\mathcal{L}$-spline $\sigma(x)$ can be extended beyond the interval $[x_1, x_N]$ with maintaining of the smoothness. By analogy with polynomial splines, such splines are called *natural* $\mathcal{L}$-splines.

Thus, $\sigma(x)$ is a natural $\mathcal{L}$-spline corresponding to the differential operator

$$\mathcal{L}_{2m}(D) = \mathcal{L}_m(D)\ \mathcal{L}_m^*(D)$$

with knots at the points of the mesh $\Delta_N = \{x_j\}_{j=1}^N$. If $N > m$, then this solution is unique [14, Sect. 5.21].

Remark 1. If $\mathcal{L}_m(D) = D^m$, $m \geq 2$, then it is to see from Lemma 1 that the fundamental solution of the operator $\mathcal{L}_{2m}(D) = (-1)^m D^{2m}$ is

$$\mathcal{E}_{2m}(x-t) = \frac{(-1)^m (x-t)_+^{2m-1}}{(2m-1)!},$$

where $(x-t)_+ = \max\{x-t, 0\}$ is the truncated function, which is traditionally widely used in the spline theory. By choosing $g_\nu(x) = x^\nu$ $(\nu = 0, 1, \ldots, m)$, from (2.6) and (2.7), we arrive at the well-known polynomial natural splines. More information about these splines can be found, for example, in [5, 8, 14].

3. On natural $\mathcal{L}$-splines

First, we get an explicit expression of the fundamental solution for the differential operator $\mathcal{L}_{2m}(D) = \mathcal{L}_m(D)\mathcal{L}_m^*(D)$, where $\mathcal{L}_m(D)$ is given in (1.2). For simplicity, we impose additional restrictions on the roots of the characteristic polynomial of the differential operator (1.2). We will assume that $\beta_j \in \mathbb{R}\backslash\{0\}$ and $\beta_i \neq \pm\beta_j$, $i \neq j$, for all $i, j = 1, 2, \ldots, r$.

Lemma 5. *Let $\mathcal{L}_r(D)$ be a linear differential operator of order $r \geq 2$ of the form*

$$\mathcal{L}_r(D) = (D - b_1 I)(D - b_2 I) \cdots (D - b_r I).$$

If $b_j \in \mathbb{R} \backslash \{0\}$ and $b_i \neq \pm b_j$, $i \neq j$, for all $i, j = 1, 2, \ldots, r$. Then

$$\mathcal{E}_r(t) = (-1)^{r-1} \theta(t) \sum_{s=1}^{r} \frac{e^{b_s t}}{\prod_{\nu=1,\, \nu \neq s}^{r} (b_\nu - b_s)},$$

where $\theta(t)$ is the Heaviside function.

P r o o f. Find a solution to the initial value problem from Lemma 1. The assumptions about the numbers $\{b_j\}_{j=1}^m$ mean that all roots of the characteristic polynomial of the differential operator $\mathcal{L}_r(D)$ are simple and nonzero.

The general solution to the differential equation $\mathcal{L}_r(D) Z_r(t) = 0$ is written as

$$Z_r(t) = C_1 e^{b_1 t} + C_2 e^{b_2 t} + \ldots + C_r e^{b_r t},$$

where $\{C_s : s = 1, 2, \ldots, r\}$ are some real numbers. To find the function $Z_r(t)$, we have initial conditions, which lead to a system of linear algebraic equations with respect to $\{C_s\}_{s=1}^r$. The system has the Vandermonde matrix. We solve the system using Cramer's rule and obtain

$$C_s = \frac{(-1)^{r-1}}{\prod_{\nu=1,\, \nu \neq s}^{r} (b_\nu - b_s)}, \quad s = 1, 2, \ldots, r.$$

It remains to use Lemma 1. □

Now, we apply Lemma 5 to the differential operator

$$\mathcal{L}_{2m}(D) = \mathcal{L}_m(D) \mathcal{L}_m^*(D) = (-1)^m (D^2 - \beta_1^2 I)(D^2 - \beta_2^2 I) \ldots (D^2 - \beta_m^2 I)$$

with the restrictions $\beta_j \in \mathbb{R} \backslash \{0\}$ and $\beta_i \neq \pm \beta_j$, $i \neq j$, for all $i, j = 1, 2, \ldots, m$.

By B_{2m}, we denote the set of roots of the characteristic polynomial of the operator $\mathcal{L}_{2m}(D)$:

$$B_{2m} = \{b_1, b_2, \ldots, b_{2m}\}.$$

Each of the numbers b_j $(j = 1, 2, \ldots, 2m)$ is either a root of the characteristic polynomial p_m of the differential operator $\mathcal{L}_m(D)$ or a root of the characteristic polynomial p_m^* of the formal adjoint operator $\mathcal{L}_m^*(D)$. All numbers b_j $(j = 1, 2, \ldots, 2m)$ are different. Therefore, Lemma 5 gives

$$\mathcal{E}_{2m}(t) = (-1)^{m-1}\ \theta(t) \sum_{s=1}^{2m} \frac{e^{b_s t}}{\prod_{\nu=1,\, \nu \neq s}^{2m} (b_\nu - b_s)}. \tag{3.1}$$

Under imposed constraints on $\{\beta_j\}_{j=1}^m$, the system of functions $\{e^{\beta_1 x}, e^{\beta_2 x}, \ldots, e^{\beta_m x}\}$ is a basis in $\ker \mathcal{L}_m(D)$. Therefore, from (2.6) and (2.7) for an arbitrary set of interpolated data $z = \{z_i\}_{i=1}^N$, we have the system of $N + m$ linear equations with respect to $N + m$ unknowns $\{\lambda_j\}_{j=1}^N$ and $\{c_k\}_{k=1}^m$:

$$\begin{cases} \sum_{j=1}^N \lambda_j \mathcal{E}_{2m}(x_i - x_j) + \sum_{k=1}^m c_k e^{\beta_k x_i} = z_i \quad (i = 1, 2, \ldots, N), \\ \sum_{j=1}^N \lambda_j e^{\beta_1 x_j} = 0, \ldots, \sum_{j=1}^N \lambda_j e^{\beta_m x_j} = 0. \end{cases}$$

If $N > m$, then the system has a unique solution.

Taking into account the properties of the function $\mathcal{E}_{2m}(x)$, we study the matrix $\mathcal{A}$ of the system. Since $x_1 < x_2 < \cdots < x_{N-1} < x_N$, we have $\mathcal{E}_{2m}(x_i - x_j) = 0$ for $i \le j$. Therefore, the matrix $\mathcal{A}$ consists of four blocks

$$\mathcal{A} = \left(\begin{array}{c|c} E_{N\times N} & B_{N\times m} \\ \hline V_{m\times N} & O_{m\times m} \end{array}\right),$$

where

$$B_{N\times m} = \begin{pmatrix} e^{\beta_1 x_1} & e^{\beta_2 x_1} & \dots & e^{\beta_m x_1} \\ e^{\beta_1 x_2} & e^{\beta_2 x_2} & \dots & e^{\beta_m x_2} \\ \dots & \dots & \dots & \dots \\ e^{\beta_1 x_N} & e^{\beta_2 x_N} & \dots & e^{\beta_m x_N} \end{pmatrix},$$

$$E_{N\times N} = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 & 0 \\ \mathcal{E}_{2m}(x_2 - x_1) & 0 & \cdots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathcal{E}_{2m}(x_{N-1} - x_1) & \mathcal{E}_{2m}(x_{N-1} - x_2) & \cdots & \mathcal{E}_{2m}(x_{N-1} - x_{N-2}) & 0 & 0 \\ \mathcal{E}_{2m}(x_N - x_1) & \mathcal{E}_{2m}(x_N - x_2) & \cdots & \mathcal{E}_{2m}(x_N - x_{N-2}) & \mathcal{E}_{2m}(x_N - x_{N-1}) & 0 \end{pmatrix},$$

$O_{m\times m}$ is the zero block of size $m \times m$, and $V_{m\times N} = B^T_{N\times m}$. Here, the upper index T means the transpose of the matrix. The first row of the matrix $\mathcal{A}$ has N zeros, the second row has $(N-1)$ zeros, etc.; and finally, the $(N-1)$th row has two zeros, and the Nth row has only one zero.

From the existence and uniqueness of the natural $\mathcal{L}$-spline, it follows that $\det \mathcal{A} \neq 0$.

Now, we will find a representation of the natural $\mathcal{L}$-spline by fundamental natural $\mathcal{L}$-spline interpolants. Let

$$\{\mathcal{E}_{2m}(x - x_1), \dots, \mathcal{E}_{2m}(x - x_N), e^{\beta_1 x}, e^{\beta_2 x}, \dots, e^{\beta_m x}\}.$$

One by one, we replace the kth $(k = 1, 2, \dots, N)$ row of the matrix $\mathcal{A}$ with this row. The obtained matrices are denoted by $\mathcal{A}_k(x)$. Now, we set

$$Q_k(x) = \frac{\det \mathcal{A}_k(x)}{\det \mathcal{A}}, \quad (k = 1, 2, \dots, N).$$

The functions $\{Q_k(x)\}_{k=1}^N$ are the fundamental natural $\mathcal{L}$-spline interpolants, since it is easy to see that $\det \mathcal{A}_k(x_j) = \delta_{kj} \det \mathcal{A}$ $(k, j = 1, 2, \dots, N)$, where δ_{kj} is the Kronecker symbol.

The natural $\mathcal{L}$-spline $\sigma(x)$ (see (2.6)) is a linear combination of the fundamental natural $\mathcal{L}$-splines $\{Q_k(x)\}_{k=1}^N$; i.e.,

$$\sigma(x) = \sum_{k=1}^{N} z_k Q_k(x), \tag{3.2}$$

where $\{z_k\}_{k=1}^N$ are interpolated data.

Thus, we have two representations (2.6) and (3.2) for the extremal function of the Favard-type interpolation problem.

4. Lemmas

In this section, we establish several lemmas needed to prove our main result.

Lemma 6. *Let $N > m$, and let $\sigma(x)$ be the natural $\mathcal{L}$-spline that is the extremal function in the Favard-type interpolation problem (1.3). Let $z = \{z_k\}_{k=1}^N$ be an arbitrary set of interpolated data. Then*

$$\mathcal{L}_m(D)\sigma(x) = \frac{1}{\det \mathcal{A}} \left(\sum_{\mu=1}^{N} \left(\sum_{k=1}^{N} z_k \alpha_{\mu k} \right) \mathcal{L}_m(D) \mathcal{E}_{2m}(x - x_\mu) \right),$$

where $\{\alpha_{\mu k}\}$ are certain values independent of x.

P r o o f. Applying the differential operator $\mathcal{L}_m(D)$ to (3.2), we obtain

$$\mathcal{L}_m(D)\sigma(x) = \sum_{k=1}^{N} z_k \mathcal{L}_m(D) Q_k(x) = \frac{1}{\det \mathcal{A}} \sum_{k=1}^{N} z_k \, \big(\mathcal{L}_m(D) \det \mathcal{A}_k(x)\big).$$

Using the well-known rule of differentiation of determinants and taking into account that only one row in $\mathcal{A}_k(x)$ depends on the variable x, we find that $\mathcal{L}_m(D) \det \mathcal{A}_k(x)$ is the determinant in which the kth row has the form

$$\mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_1) \dots \mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_N) \underbrace{0 \dots 0}_{m \text{ times}}.$$

We expand each determinant $\mathcal{L}_m(D) \det \mathcal{A}_k(x)$ $(k = 1, 2, \dots, N)$ according to the elements of kth row and obtain the required equality, in which $\alpha_{\mu k}$ are the corresponding minors taken with their signs. Lemma 6 is proved. $\square$

Lemma 7. *Let x_μ be an arbitrary point of the mesh $\Delta_N : x_i < x_2 < \dots < x_N$, and let the differential operator $\mathcal{L}_m(D)$ be such that*

$$\mathcal{L}_m(D) = (D - \beta_1 I)(D - \beta_2 I) \dots (D - \beta_m I)$$

with $\beta_j \in \mathbb{R}\backslash\{0\}$ and $\beta_i \neq \pm\beta_j$, $i \neq j$, for all $i, j = 1, 2, \dots, m$. Then

$$\mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_\mu) = (-1)^m \sum_{i=1}^{m} \frac{e^{-\beta_i(x - x_\mu)}}{\prod_{\nu=1,\, \nu\neq i}^{m} (\beta_\nu - \beta_i)} \tag{4.1}$$

for $x > x_\mu$ and $\mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_\mu) = 0$ for $x \le x_\mu$.

P r o o f. We write (3.1) as

$$\mathcal{E}_{2m}(x - x_\mu) = (-1)^{m-1} \, \theta(x - x_\mu) \sum_{s=1}^{2m} \omega_s e^{b_s(x - x_\mu)}, \tag{4.2}$$

where $b_s = \beta_s$, $b_{s+m} = -\beta_s$ $(s = 1, 2, \dots, m)$, and

$$\omega_s = \Big(\prod_{\nu=1,\, \nu\neq s}^{2m} (b_\nu - b_s) \Big)^{-1} \quad (s = 1, 2, \dots, 2m).$$

Hence, $\mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_\mu) = 0$ for $x \le x_\mu$.

Let $x > x_\mu$. From (4.2), we have

$$\mathcal{L}_m(D)\mathcal{E}_{2m}(x - x_\mu) = (-1)^{m-1} \Bigg(\sum_{s:\ b_s = \beta_s} \omega_s \, \mathcal{L}_m(D) e^{b_s(x - x_\mu)} + \sum_{s:\ b_s = -\beta_s} \omega_s \, \mathcal{L}_m(D) e^{b_s(x - x_\mu)} \Bigg).$$

The former sum on the right-hand side vanishes because

$$\mathcal{L}_m(D) e^{\beta_s x} = 0 \quad (s = 1, 2, \dots, m).$$

By simple calculation, we obtain

$$\mathcal{L}_m(D) e^{-\beta x} = e^{-\beta x} \, p_m^*(\beta),$$

where p_m^* is the characteristic polynomial of the formal adjoint operator $\mathcal{L}_m^*(D)$. By using this relation, we arrive at the equality

$$\mathcal{L}_m(D)\mathcal{E}_{2m}(x-x_\mu) = (-1)^{m-1} \sum_{i=1}^{m} \omega_i \, e^{-\beta_i(x-x_\mu)} \, p_m^*(\beta_i).$$

Since

$$\omega_i^{-1} = 2(-1)^{m-1}\beta_i \prod_{\nu=1,\,\nu\neq i}^{m} (\beta_\nu^2-\beta_i^2)$$

and

$$p_m^*(\beta_i) = 2(-1)^{m-1}\beta_i \prod_{\nu=1,\,\nu\neq i}^{m} (\beta_\nu+\beta_i),$$

we come to (4.1). Lemma 7 is proved. □

5. The main result

Finally, we directly turn to the problem of finding quantity (1.4).

Theorem 1. *Assume that*

$$\mathcal{L}_m(D) = (D-\beta_1 I)(D-\beta_2 I)\dots(D-\beta_m I)$$

is a linear differential operator of order $m \geq 2$ such that $\beta_j \in \mathbb{R}\backslash\{0\}$ and $\beta_i \neq \pm\beta_j$, $i \neq j$ $(i, j = 1, 2, \dots, m)$.

If $N > m$, then

$$\mathfrak{B}_{\mathcal{L}_m}(\Delta_N) = \frac{1}{|\det \mathcal{A}|} \sqrt{\lambda_{max}}\,,$$

where λ_{max} is the largest eigenvalue of the matrix $Q = (a_{ij})_{i,j=1}^N$ whose entries $\{a_{ij}\}$ are such that

$$a_{kk} = \sum_{\mu=1}^{N} R_{\mu\mu}\alpha_{\mu k}^2 + 2\sum_{\mu>n}^{N} R_{\mu n}\alpha_{\mu k}\alpha_{nk},$$

$$a_{kj} = \sum_{\mu=1}^{N} R_{\mu\mu}\alpha_{\mu k}\alpha_{\mu j} + 2\sum_{\mu>n}^{N} R_{\mu n}\alpha_{\mu k}\alpha_{nj} \quad (k \neq j),$$

$$R_{\mu n} = \sum_{i,\nu=1}^{m} \omega_i\omega_\nu \, \frac{e^{-\beta_\nu(x_\mu-x_n)} - e^{-\beta_i(x_N-x_\mu)}e^{-\beta_\nu(x_N-x_n)}}{\beta_i+\beta_\nu} \quad (\mu \geq n),$$

$$\omega_i = \Big(\prod_{s=1,\,s\neq i}^{m} (\beta_s-\beta_i)\Big)^{-1} \quad (i = 1, 2, \dots, m),$$

and $\{\alpha_{s\mu}\}_{s,\mu=1}^N$ are the minors taken with their signs in the decompositions of the determinants $\mathcal{L}_m(D)\det\mathcal{A}_k(x)$ $(k = 1, 2, \dots, N)$ according to the elements of the kth row.

P r o o f. Since $N > m$, the natural $\mathcal{L}$-spline $\sigma(x)$ is the unique extremal function in (1.3). After applying the differential operator $\mathcal{L}_m(D)$ to (2.6), we have

$$\mathcal{L}_m(D)\sigma(x) = \sum_{\mu=1}^{N} \lambda_\mu \, \mathcal{L}_m(D)\mathcal{E}_{2m}(x-x_\mu).$$

Comparing this result with Lemma 6 yields

$$\lambda_\mu = \frac{1}{\det \mathcal{A}} \sum_{k=1}^{N} z_k \alpha_{\mu k} \quad (\mu = 1, 2, \dots, N). \tag{5.1}$$

Using Lemma 7, we obtain

$$\mathcal{L}_m(D)\sigma(x) = (-1)^m \sum_{\mu=1}^{N} \lambda_\mu F_\mu(x),$$

where

$$F_\mu(x) = \sum_{i=1}^{m} \omega_i \begin{cases} e^{-\beta_i(x-x_\mu)}, & x > x_\mu, \\ 0, & x \le x_\mu, \end{cases}$$

and

$$\omega_i = \Big(\prod_{\nu=1,\, \nu\neq i}^{m} (\beta_\nu - \beta_i) \Big)^{-1} \quad (i = 1, 2, \dots, m).$$

Define

$$R_{\mu n} = \int\limits_{x_1}^{x_N} F_\mu(x) F_n(x) dx \quad (\mu, n = 1, 2, \dots, N, \;\; \mu \ge n). \tag{5.2}$$

For the natural $\mathcal{L}$-spline $\sigma(x)$, we have

$$\int_a^b \big|\mathcal{L}_m(D)\sigma(x)\big|^2 dx = \int_{x_1}^{x_N} \big|\mathcal{L}_m(D)\sigma(x)\big|^2 dx = \int_{x_1}^{x_N} \Big|\sum_{\mu=1}^{N} \lambda_\mu F_\mu(x)\Big|^2 dx$$

$$= \int_{x_1}^{x_N} \Big(\sum_{\mu=1}^{N} \lambda_\mu^2 F_\mu^2(x) + 2 \sum_{\mu>n}^{N} \lambda_\mu \lambda_n F_\mu(x) F_n(x) \Big) dx = \sum_{\mu=1}^{N} \lambda_\mu^2 R_{\mu\mu} + 2 \sum_{\mu>n}^{N} \lambda_\mu \lambda_n R_{\mu n}.$$

Now, we substitute λ_μ from (5.1) into the latter expression. After simple transformations, we obtain

$$\big(K_N(z)\big)^2 = \Big(\frac{1}{\det \mathcal{A}}\Big)^2 \Big\{ \sum_{\mu=1}^{N} \Big(\sum_{k=1}^{N} z_k \alpha_{\mu k} \Big)^2 R_{\mu\mu} + 2 \sum_{\mu>n}^{N} R_{\mu n} \Big(\sum_{k=1}^{N} z_k \alpha_{\mu k} \Big) \Big(\sum_{j=1}^{N} z_j \alpha_{nj} \Big) \Big\}$$

$$= \frac{1}{(\det \mathcal{A})^2} \Big(\sum_{k=1}^{N} a_{kk} z_k^2 + \sum_{k,j=1,\; k\neq j}^{N} a_{kj} z_k z_j \Big),$$

where

$$a_{kk} = \sum_{\mu=1}^{N} R_{\mu\mu} \alpha_{\mu k}^2 + 2 \sum_{\mu>n}^{N} R_{\mu n} \alpha_{\mu k} \alpha_{nk},$$

$$a_{kj} = \sum_{\mu=1}^{N} R_{\mu\mu} \alpha_{\mu k} \alpha_{\mu j} + 2 \sum_{\mu>n}^{N} R_{\mu n} \alpha_{\mu k} \alpha_{nj} \quad (k \neq j).$$

It remains to calculate the integrals (5.2). Suppose that $\mu \ge n$.Then $x_\mu \ge x_n$, and we have

$$R_{\mu n} = \int_{x_\mu}^{x_N} \Big(\sum_{i=1}^{m} \omega_i e^{-\beta_i(x-x_\mu)} \Big) \Big(\sum_{\nu=1}^{m} \omega_\nu e^{-\beta_\nu(x-x_n)} \Big) dx = \sum_{i,\nu=1}^{m} \omega_i \omega_\nu \Big(\int_{x_\mu}^{x_N} e^{-(\beta_i+\beta_\nu)x} dx \Big) e^{\beta_i x_\mu + \beta_\nu x_n}$$

$$= \sum_{i,\nu=1}^{m} \omega_i \omega_\nu \, \frac{e^{-\beta_\nu(x_\mu - x_n)} - e^{-\beta_i(x_N - x_\mu)} e^{-\beta_\nu(x_N - x_n)}}{\beta_i + \beta_\nu}.$$

In our calculations, we used the fact that the functions $F_\mu(x)$ are equal to zero for $x \le x_\mu$. Note that all the denominators do not vanish due to the restrictions imposed on the differential operator $\mathcal{L}_m(D)$.

Thus, to find quantity (1.4), it is sufficient to solve the following extremal problem:

$$\begin{cases} V(z_1, z_2, \ldots, z_N) = \sum\limits_{k=1}^{N} a_{kk} z_k^2 \; + \sum\limits_{k,j=1,\; k\neq j}^{N} a_{kj} z_k z_j \; \to \; \max, \\ \sum\limits_{k=1}^{N} |z_k|^2 \le 1. \end{cases}$$

As in [9, p. 224–225], it is proved that the maximum is attained at the boundary of the set $\mathfrak{M}_N$, i.e., at some points $z = (z_1, z_2, \ldots, z_N)$ with $\sum_{k=1}^{N} |z_k|^2 = 1$. As a result, we arrive at the problem of maximizing the quadratic form on the unit sphere in the space l_2^N.

It is known (see, for example, [3, p. 476–477]) that the unique solution to this problem is the largest eigenvalue $\lambda = \lambda_{max}$ of the matrix $Q = (a_{ij})_{i,j=1}^N$ of the quadratic form $V(z_1, z_2, \ldots, z_N)$, i.e., the maximum root of the equation

$$\det(Q - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1N} \\ a_{12} & a_{22} - \lambda & \cdots & a_{2N} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1N} & a_{2N} & \cdots & a_{NN} - \lambda \end{vmatrix} = 0.$$

The matrix Q is symmetric. Therefore, all its eigenvalues are real. It is also not difficult to see that $\lambda_{max} > 0$. Theorem 1 is proved. □

6. Discussion and comments

(1) The constraints imposed on the roots of the characteristic polynomial $p_m(t)$ of the differential operator $\mathcal{L}_m(D)$ in Theorem 1 can be partially weakened. In particular, one can exclude the constraints $\beta_i \neq 0$ and $\beta_i \neq \pm\beta_j$, $i \neq j$ $(i, j = 1, 2, \ldots, m)$. In this case, the characteristic polynomials of the differential operators $\mathcal{L}_m(D)$ and $\mathcal{L}_m^*(D)$ will have common roots. The approach used in the paper can be implemented with minor modifications for such differential operators. In so doing, the explicit expressions for $\mathcal{E}_{2m}(t)$ (see Lemma 5), the entries of the matrix $\mathcal{A}$, and the numbers $R_{\mu n}$ (see Theorem 1) will be different from those given in the paper. However, in the framework of the applied approach, it is impossible to discard the constraint $\beta_i \in \mathbb{R}$, $i = 1, 2, \ldots, m$. The point is that if the polynomial p_m has nonreal complex roots (two or more), then for an arbitrary interval $[a, b]$, it fails to prove that a solution to the Favard-type interpolation problem is a natural $\mathcal{L}$-spline. Apparently, in this case, we need to introduce some additional restrictions that would associate the segment $[a, b]$ with oscillation properties of the differential operator $\mathcal{L}_m(D)$.

(2) Another approach to the study of problems (1.3) and (1.4) is known. This approach is based on the concept of Chebyshevian splines (see e.g. [6, Chap. 10, Sect. 3] or [7, Chap. 11]). Since $\ker \mathcal{L}_m(D) = \operatorname{span}\{e^{\beta_1 x}, e^{\beta_2 x}, \ldots, e^{\beta_m x}\}$, this set of functions is an ECT-system on any finite interval $[a, b]$. Due to this, Theorem 1 can be proved by using of this approach.

(3) For any given N and prescribed knots of the mesh Δ_N, all coefficients of the quadratic form $V(z_1, z_2, \ldots, z_N)$ are calculated. This can be made directly or by using symbolic computation systems (Maple and others). Numerical algorithms allow one to find the largest eigenvalue of

the symmetric matrix Q approximately. One can also estimate λ_{max} from above and below (see, for example, [12, 19] and many other publications). Specifically, from [19], we can write the estimate

$$\frac{|\mathrm{tr}\ Q|}{N} \le |\lambda_{max}| \le \frac{1}{N}\Big(|\mathrm{tr}\ Q| + \sqrt{N-1}\ \big(N\ \mathrm{tr}\ Q^2 - (\mathrm{tr}\ Q)^2\big)^{1/2}\Big),$$

where tr Q is the sum of elements located on the main diagonal of the matrix Q.

(4) The periodical analogs of quantities (1.3) and (1.4) for the differential operator $\mathcal{L}_m(D) = D^m$ were studied in the author's paper [10]. Periodicity requirements for both interpolated values $z = \{z_k\}_{k=0}^{2N-1}$ and interpolating functions have been added to problem statements. However, the results in [10] were obtained only for the mesh with equidistant knots $\Delta_{2N} = \{j\pi/N\}_{j=0}^{2N-1}$ that was 2π-periodically extended into $\mathbb{R}$. For the class of interpolated values

$$\widetilde{\mathfrak{M}_{2N}} = \Big\{z :\ z = \{z_j\}_{j=0}^{2N-1}, \Big(\sum_{j=0}^{2N-1} |z_j|^2\Big)^{1/2} \le 1\Big\},$$

it was proved that if $N > m,\ m \ge 2$, then

$$\widetilde{\mathfrak{B}_m}(\Delta_{2N}) = \sup_{z\in\widetilde{\mathfrak{M}_{2N}}} \inf_{\substack{f^{(m-1)}\in\widetilde{AC} \\ f(j\pi/N)=z_j}} \Big(\int_0^{2\pi} |f^{(m)}(t)|^2 dt\Big)^{1/2} = \Big(\frac{\pi}{N}\Big)^{1/2} \Bigg(\sum_{l\in\mathbb{Z}} \frac{1}{(2Nl+(N-1))^{2m}}\Bigg)^{-1/2},$$

where $\widetilde{AC}$ is the class of 2π-periodic absolutely continuous functions.

Note that the series on the right hand side of the last equality is convergent. It is easy to see using the comparison test for number series.

While proving this result, the largest eigenvalue of an analog of the matrix Q also arose. However, due to the specificity of the periodic case and the uniform grid, the analog of the matrix Q is the circulant, and its eigenvalues are known in an explicit form.

(5) Along with our settings of problems (1.3) and (1.4), one can consider the corresponding multivariate settings. For the case of two variables, analogs of natural splines are known [13]. Problem (1.4) was considered in the case of the Laplace operator. This was done in the author's recent paper [11].

7. Conclusion

We have considered the problem of finding quantity (1.4). In a certain sense, this quantity is the value of the L_2-norm of the differential operator applied to the "best" interpolant under interpolating the "worst" data from the given class. In this paper, we found the exact value of the studied quantity.

REFERENCES

1. Bezhaev A. Yu., Vasilenko V. A. *Variation Spline Theory.* Ser. Numer. Anal. Bulletin No. 3. Novosibirsk: NCC Publisher, 1993. 258 p.
2. Favard J. Sur l'interpolation. *J. Math. Pures Appl.*, 1940. Vol. 19, No. 1–4. P. 281–306. (in French)
3. Fichtenholz G.M. *Kurs differencial'nogo i integral'nogo ischisleniya* [Course of Differential and Integral Calculus]. Vol. 1. Moscow: Nauka Publ., 1970. 608 p. (in Russian)

4. Fisher S. D., Jerome J. W. *Minimum Norm Extremals in Function Spaces: With Applications to Classical and Modern Analysis.* Lecture Notes in Math., vol. 479. 1975. 214 p. DOI: 10.1007/BFb0097059
5. Jerome J. W., Schumaker L. L. A note on obtaining natural spline functions by the abstract approach of Atteia and Laurent. *SIAM J. Numer. Anal.,* 1968. Vol. 5, No. 4. P. 657–663. DOI: 10.1137/0705052
6. Karlin S. *Total Positivity.* Vol. 1. London: Stanford Univer. Press, 1968. 576 p.
7. Karlin S., Studden W J. *Tchebycheff Systems: With Applications in Analysis and Statistics.* Pure and Appl. Math., vol. 15. New York: Interscience, 1966. 586 p.
8. Malozyomov V.N., Pevny A.B. *Polinomial'nye splajny* [Polynomial Splines]. Leningrad: Leningrad State University, 1986. 120 p. (in Russian)
9. Novikov S.I. Optimal interpolation on an interval with the smallest mean-square norm of the r-th derivative. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2023. Vol. 29, No. 4. P. 217–228. DOI: 10.21538/0134-4889-2023-29-4-217-228 (in Russian)
10. Novikov S. I. Periodic interpolation with minimal norm of m-th derivative. *Sib. Zh. Vychisl. Math.,* 2006. Vol. 9, No. 2. P. 165–172. (in Russian)
11. Novikov S. I. On an interpolation problem with the smallest L_2-norm of the Laplace operator. *Proc. Steklov Inst. Math.*, 2022. Vol. 319, Suppl. 1. P. S193–S203. DOI: 10.1134/S0081543822060177
12. Parodi M. *La Localisation des Valeurs Caractéristiques des Matrices et ses Applications.* Paris: Gauthier Villars, 1959. 172 p. (in French)
13. Pevnyi A. B. Natural splines of two and three variables. *Metody Vychisl.*, 1985. No. 14. P. 160–170. (in Russian)
14. Rozhenko A. I. *Abstraktnaya teoriya splajnov* [Abstract Theory of Splines]. Novosibirsk: Publ. Center of Novosibirsk State University, 1999. 177 p. (in Russian)
15. Shevaldin V. T. Subbotin's splines in the problem of extremal interpolation in the space $L_p(\mathbb{R})$ for second-order linear differential operators. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2021. Vol. 27, No. 4. P. 255–262. DOI: 10.21538/0134-4889-2021-27-4-255-262 (in Russian)
16. Shevaldin V. T. Extremal interpolation with the least value of the norm of the second derivative in $L_p(\mathbb{R})$. *Izv. Math.*, 2022. Vol. 86, No. 1. P. 203–219. DOI: 10.1070/IM9125
17. Subbotin Yu. N. Functional interpolation in the mean with smallest n derivative. *Proc. Steklov Inst. Math.*, 1967. Vol. 88. P. 31–63.
18. Subbotin Yu. N., Novikov S. I., Shevaldin V. T. Extremal function interpolation and splines. *Trudy Inst. Mat. i Mekh. UrO RAN,* 2018. Vol. 24, No. 3. P. 200–225. DOI: 10.21538/0134-4889-2018-24-3-200-225 (in Russian)
19. Tarazaga P. Eigenvalue estimates for symmetric matrices. *Linear Algebra Appl.*, 1990. Vol. 135, No. 1. P. 171–179. DOI: 10.1016/0024-3795(90)90120-2
20. Tikhomirov V. M., Bojanov B. D. On some convex problems of approximation theory. *Serdica*, 1979. Vol. 5. P. 83–96. (in Russian)
21. Vladimirov V. S. *Uravneniya matematicheskoj fiziki* [Equations of Mathematical Physics]. Moscow: Nauka Publ., 1971. 512 p. (in Russian)

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 121–130

DOI: 10.15826/umj.2024.2.011

ON WIDTHS OF SOME CLASSES OF ANALYTIC FUNCTIONS IN A CIRCLE

Mirgand Sh. Shabozov

Tajik National University,
17 Rudaky Ave., Dushanbe, 734025, Republic of Tajikistan

shabozov@mail.ru

Muqim S. Saidusainov

University of Central Asia,
155 Qimatsho Imatshoev, Khorog, GBAO, Republic of Tajikistan

muqim.saidusainov@ucentralasia.org

Abstract: We calculate exact values of some n-widths of the class $W_q^{(r)}(\Phi)$, $r \in \mathbb{Z}_+$, in the Banach spaces $\mathscr{L}_{q,\gamma}$ and $B_{q,\gamma}$, $1 \le q \le \infty$, with a weight γ. These classes consist of functions f analytic in the unit circle, their rth order derivatives $f^{(r)}$ belong to the Hardy space H_q, $1 \le q \le \infty$, and the averaged moduli of smoothness of boundary values of $f^{(r)}$ are bounded by a given majorant Φ at the system of points $\{\pi/(2k)\}_{k\in\mathbb{N}}$; more precisely,

$$\frac{k}{\pi-2}\int_0^{\pi/(2k)} \omega_2(f^{(r)}, 2t)_{H_{q,\rho}}\, dt \le \Phi\left(\frac{\pi}{2k}\right)$$

for all $k \in \mathbb{N}$, $k > r$.

Keywords: Modulus of smoothness, The best approximation, n-widths, The best linear method of approximation.

1. Introduction

There are many studies devoted to calculating exact values of various n-widths of classes of functions analytic in the unit circle both in the Hardy space H_q $(1 \le q \le \infty)$ and in the Bergman space B_q $(1 \le q \le \infty)$ (see, e.g., [1–36]). The present paper aims to obtain new results related to calculating exact values of various n-widths of some classes of functions analytic in the unit circle.

First, we introduce some notation and concepts. Define

$$U_\rho := \{z \in \mathbb{C} : |z| < \rho\}, \quad 0 < \rho \le 1,$$

Let $U := U_1$, let $\mathscr{A}(U_\rho)$ be the set of functions analytic in a circle U_ρ, and let H_q $(1 \le q \le \infty)$ be the Hardy space of functions $f \in \mathscr{A}(U)$ such that

$$\|f\|_{H_q} = \lim_{\rho\to 1-0} M_q(f,\rho),$$

where

$$M_q(f,\rho) := \begin{cases} \left(\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} |f(\rho e^{it})|^q dt\right)^{1/q}, & 1 \le q < \infty, \\ \max\left\{|f(\rho e^{it})| : 0 < t \le 2\pi\right\}, & q = \infty; \end{cases}$$

the integral is understood in the Lebesgue sense.

It is known [26] that the norm $\|f\|_{H_q}$ is attained on angular boundary values $f(e^{it})$ of functions $f \in H_q$, which exist almost everywhere on $[0, 2\pi)$. We set

$$H_{q,\rho} := \left\{ f \in \mathscr{A}(U_\rho) : \|f(\cdot)\|_{H_{q,\rho}} := \|f(\rho\cdot)\|_{H_q} < \infty \right\}$$

and, for $r \in \mathbb{Z}_+$,

$$H_q^{(r)} := \left\{ f \in \mathscr{A}(U) : f^{(r)} \in H_q \right\} \quad (H_q^{(0)} \equiv H_q),$$

where

$$f^{(r)}(z) = \sum_{k=r}^{\infty} \alpha_{k,r} c_k(f) z^k,$$

$$\alpha_{k,r} = k(k-1)\cdots(k-r+1), \quad k \geq r, \quad k \in \mathbb{Z}_+, \quad \alpha_{k,0} \equiv 1,$$

and $c_k(f)$ are coefficients of the Taylor series

$$f(z) = \sum_{k=0}^{\infty} c_k(f) z^k.$$

Denote by

$$\mathscr{L}_q := \mathscr{L}_q(U) \quad (1 \leq q < \infty)$$

the Banach space of complex-valued functions f on U with finite norms

$$\|f\|_{\mathscr{L}_q} = \left(\frac{1}{2\pi} \iint_{(U)} |f(z)|^q dxdy \right)^{1/q} = \left(\frac{1}{2\pi} \int_0^1 \int_0^{2\pi} \rho |f(\rho e^{it})|^q dt d\rho \right)^{1/q},$$

where the integral is understood in the Lebesgue sense.

Let $\gamma(|z|)$ be a nonnegative measurable function not equivalent to zero and summable on the set U. Denote by

$$\mathscr{L}_{q,\gamma} := \mathscr{L}(U, \gamma) \quad (1 \leq q < \infty)$$

the set of complex-valued functions f on U such that

$$\gamma^{1/q} f \in \mathscr{L}_q(U), \quad \|f\|_{\mathscr{L}_{q,\gamma}} := \|\gamma^{1/q} f\|_{\mathscr{L}_q}.$$

By $B_{q,\gamma}$ $(1 \leq q < \infty)$, we mean the Banach space of functions $f \in \mathscr{A}(U)$ such that $f \in \mathscr{L}_{q,\gamma}$. In this case,

$$\|f\|_{B_{q,\gamma}} = \left(\int_0^1 \rho \gamma(\rho) M_q^q(f, \rho) d\rho \right)^{1/q}.$$

In the particular case of $\gamma \equiv 1$, the space $B_q := B_{q,1}$ is the well-known Bergman space.

Assume that X is a Banach space, $\mathbb{B}$ is the unit ball in this space, $\mathfrak{M}$ is a convex centrally symmetric subset of X, $L_n \subset X$ is an n-dimensional linear subspace, $L^n \subset X$ is a subspace of codimension n, and $\Lambda : X \to L_n$ is a continuous linear operator from X into L_n. Define the best approximation to an element $f \in X$ by elements of the subspace $L_n \subset X$ as

$$E_n(f)_X := E(f, L_n)_X = \inf \left\{ \|f - \varphi\|_X : \varphi \in L_n \right\}.$$

The approximation to the fixed set $\mathfrak{M} \subset X$ by the fixed subspace $L_n \subset X$ is defined by

$$E(\mathfrak{M}, L_n)_X := \sup \left\{ E(f, L_n)_X : f \in \mathfrak{M} \right\}.$$

If the approximation is performed with a linear operator A then, we will study the sharp upper bound

$$\sup\left\{\|f-A(f)\|_X:\ f\in\mathfrak{M}\right\},$$

and the quantity

$$\mathscr{E}(\mathfrak{M},L_n)_X=\inf\left\{\sup\left\{\|f-A(f)\|_X:\ f\in\mathfrak{M}\right\}:AX\subset L_n\right\},\tag{1.1}$$

which characterizes the best linear approximation of the set $\mathfrak{M}$ by elements of $L_n\subset X$. If there exists a linear operator A^*, $A^*X\subset L_n$ realizing the infimum in (1.1), i.e., an operator such that

$$\mathscr{E}(\mathfrak{M},L_n)_X=\sup\left\{\|f-A^*(f)\|_X:\ f\in\mathfrak{M}\right\},$$

then A^* is called the best linear method of approximation to $\mathfrak{M}$.

The quantities

$$\begin{gathered}b_n(\mathfrak{M},X):=\sup\big\{\sup\{\varepsilon>0;\ \varepsilon\mathbb{B}\cap L_{n+1}\subset\mathfrak{M}\}:L_{n+1}\subset X\big\},\\ d_n(\mathfrak{M},X):=\inf\{E(\mathfrak{M},L_n)_X:L_n\subset X\},\\ d^n(\mathfrak{M},X):=\inf\big\{\sup\{\|f\|_X:f\in\mathfrak{M}\cap L^n\}:L^n\subset X\big\},\\ \delta_n(\mathfrak{M},X):=\inf\left\{\mathscr{E}(\mathfrak{M},L_n)_X:L_n\subset X\right\},\end{gathered}\tag{1.2}$$

are called *Bernstein*, *Kolmogorov*, *Gelfand*, and *linear n-widths*, respectively (see, for example, [8, Ch. II], [30, Ch. III]).

If there exists a subspace $\bar{L}_{n+1}\subset X$, $\dim\bar{L}_{n+1}=n+1$, for which

$$b_n(\mathfrak{M},X):=\sup\left\{\varepsilon>0:\ \varepsilon\mathbb{B}\cap L_{n+1}\subset\mathfrak{M}\right\},$$

then it is an extremal subspace for $b_n(\mathfrak{M},X)$. A subspace $L_n^*\subset X$, $\dim L_n^*=n$, on which the infimum in (1.2) is attained, i.e., $d_n(\mathfrak{M},X)=E(\mathfrak{M},L_n^*)$ is called an extremal subspace for the Kolmogorov n-width $d_n(\mathfrak{M},X)$. If there exist a subspace $L_*^n\subset X$ of codimention n such that

$$d^n(\mathfrak{M},X):=\sup\left\{\|f\|_X:f\in\mathfrak{M}\cap L_*^n\right\},$$

then L_*^n is said to be extremal for $d^n(\mathfrak{M},X)$. A subspace $\tilde{L}_n\subset X$, $\dim\tilde{L}_n=n$ such that

$$\delta_n(\mathfrak{M},X)=\mathscr{E}(\mathfrak{M},\tilde{L}_n),$$

if it exists, is called extremal for $\delta_n(\mathfrak{M},X)$. Finding extremal subspaces $\hat{L}_n\subset X$, $\dim\hat{L}_n=n$, such that

$$E(\mathfrak{M},\hat{L}_n)_X=\mathscr{E}(\mathfrak{M},\hat{L}_n)_X=d_n(\mathfrak{M},X)=\delta_n(\mathfrak{M},X)$$

is of special interest. The n-widths mentioned above satisfy the relations [8, 30]

$$b_n(\mathfrak{M},X)\le\begin{array}{c}d_n(\mathfrak{M},X)\\ d^n(\mathfrak{M},X)\end{array}=\delta_n(\mathfrak{M},X).\tag{1.3}$$

2. Main theorem

Following [28, p. 652] and [14, p. 284], for an arbitrary function $f\in H_q$ $(1\le q\le\infty)$, we consider the modulus of smoothness

$$\omega_2(f,2x)_{H_q}:=\sup_{|t|\le x}\left\|f(e^{i(\cdot+t)})-2f(e^{i(\cdot)})+f(e^{i(\cdot-t)})\right\|_{L_q[0,2\pi]},$$

where the $L_q[0,2\pi]$-norm is defined by

$$\|f\|_{L_q[0,2\pi]} = \begin{cases} \left(\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi} |f(e^{it})|^q dt\right)^{1/q}, & 1 \le q < \infty, \\ \operatorname*{ess\,sup}_{0\le t\le 2\pi} |f(e^{it})|, & q = \infty. \end{cases}$$

Let $\Phi(t)$, $t \ge 0$, be a continuous increasing function such that $\Phi(0) = 0$. Using Φ as a majorazing function, we consider the class of functions studied by Taikov [28]:

$$W_q^{(r)}(\Phi) := \left\{ f \in \mathscr{A}(U) : \ f^{(r)} \in H_q, \ \frac{k}{\pi - 2}\int_0^{\pi/(2k)} \omega_2(f^{(r)}, 2t)_{H_q} dt \le \Phi\left(\frac{\pi}{2k}\right), \, k \in \mathbb{N} \right\},$$

where $r \in \mathbb{Z}_+$ and $1 \le q \le \infty$.

In [28, Theorem 4], it is proved that, if the majorant $\Phi(t)$ for $0 < t \le \pi/2$ satisfies the inequality

$$\frac{\Phi(\lambda t)}{\Phi(t)} \ge \frac{\pi}{\pi - 2} \begin{cases} 1 - \dfrac{2}{\lambda\pi}\sin\dfrac{\lambda\pi}{2}, & \text{if} \quad 0 < \lambda \le 2, \\ 2\left(1 - \dfrac{1}{\lambda}\right), & \text{if} \quad \lambda \ge 2, \end{cases} \tag{2.1}$$

then the following equality holds for all $n \in \mathbb{N}$, $r \in \mathbb{Z}_+$, $n > r$, and $1 \le q \le \infty$:

$$d_n\left(W_q^{(r)}(\Phi), H_q\right) = \frac{1}{\alpha_{n,r}}\Phi\left(\frac{\pi}{2(n-r)}\right).$$

It is also proved that the function $\Phi_*(t) = t^{2/(\pi-2)}$ satisfies constraint (2.1).

It is also of interest to calculate the exact values of the above n-widths for the classes $W_q^{(r)}(\Phi)$ in the spaces $\mathscr{L}_{q,\gamma}$ and $B_{q,\gamma}$, $r \in \mathbb{Z}_+$, $1 \le q < \infty$.

For this purpose, we specify the extremal subspaces L_n^*, L_*^n, and $\bar{L}_{n+1}$ and the best linear approximation method Λ_{n-1}^* already mentioned in the first section.

We set

$$\begin{aligned} L_n^* &:= \operatorname{span}\left\{ \{z^k\}_{k=0}^{r-1}, \left[\left\{1 + \frac{\rho^{2(n-k)}\alpha_{k,r}}{\alpha_{2n-k,r}}\left[\beta_{k,r}\left(1 - \left(\frac{k-r}{2n-k-r}\right)^2\right) - 1\right]\right\} z^k\right]_{k=r}^{n-1} \right\}, \\ \Lambda_{n-1}^* &:= \sum_{k=0}^{r-1} c_k(f) z^k + \sum_{k=r}^{n-1}\left\{1 + \frac{\rho^{2(n-k)}\alpha_{k,r}}{\alpha_{2n-k,r}}\left[\beta_{k,r}\left(1 - \left(\frac{k-r}{2n-k-r}\right)^2\right) - 1\right]\right\} c_k(f) z^k, \end{aligned} \tag{2.2}$$

where

$$\beta_{k,r} := \frac{2(n-r)}{\pi - 2}\int_0^{\pi/2(n-r)} (1 - \sin(n-r)x)\cos(k-r)x dx, \quad k \ge n > r, \quad k, n \in \mathbb{N}, \quad r \in \mathbb{Z}_+.$$

Theorem 1. *Let $r \in \mathbb{Z}_+$, $1 \le q < \infty$, and let the majorant Φ satisfies condition (2.1). Then, the following equalities hold for all $n \in \mathbb{N}$, $n > r$:*

$$\begin{aligned} &b_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) = b_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) = d^n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) = d^n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right), \\ &d_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) = d_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) = E\left(W_q^{(r)}(\Phi); L_n^*\right)_{\mathscr{L}_{q,\gamma}} = \mathscr{E}\left(W_q^{(r)}(\Phi); L_n^*\right)_{\mathscr{L}_{q,\gamma}} \\ &= \sup\left\{\|f - \Lambda_{n-1}^*(f)\|_{\mathscr{L}_{q,\gamma}} : f \in W_q^{(r)}(\Phi)\right\} = \frac{1}{\alpha_{n,r}}\Phi\left(\frac{\pi}{2(n-r)}\right)\left(\int_0^1 \rho^{nq+1}\gamma(\rho) d\rho\right)^{1/q}. \end{aligned} \tag{2.3}$$

Moreover,

(1) *the subspace* L_n^* *is extremal in the case of* n*-widths* $d_n\big(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\big)$ *and* $\delta_n\big(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\big)$;
(2) *the continuous linear operator* Λ_{n-1}^* *is the best linear approximation method for* $W_q^{(r)}(\Phi)$ *in* $\mathscr{L}_{q,\gamma}$;
(3) *the subspace* L_*^n *is extremal for the* n*-width* $d^n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right)$;
(4) *the subspace* $\bar{L}_{n+1}$ *is extremal for the* n*-width* $b_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right)$.

To prove the theorem, we need the following lemma.

Lemma 1. *The following inequality holds for an arbitrary function* $f \in H_q^{(r)}$ $(r \in \mathbb{Z}_+,$ $1 \le q < \infty)$:

$$E_{n-1}(f)_{\mathscr{L}_{q,\gamma}} \le \frac{1}{\alpha_{n,r}} \left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho \right)^{1/q} E_{n-r-1}(f^{(r)})_{H_q}. \tag{2.4}$$

Inequality (2.4) *turns into an equality at the function* $f_0(z) = z^n$, $n > r$.

P r o o f. Relation (2.3) from [14] with $s = 0$ implies that, for an arbitrary function $f \in H_q^{(r)}$ $(r \in \mathbb{Z}_+,\ 1 \le q < \infty)$, there exists a polynomial $p_{n-1} \in \mathscr{P}_{n-1}$ satisfying the following inequality for $n \in \mathbb{N}$, $n > r$, and $0 < \rho \le 1$:

$$\big\| f(\rho e^{it}) - p_{n-1}(\rho e^{it}) \big\|_{H_q} \le \frac{\rho^n}{\alpha_{n,r}} E_{n-r-1}(f^{(r)})_{H_q}. \tag{2.5}$$

We raise both sides of (2.5) to the power q $(1 \le q < \infty)$, multiply both sides by $\rho\gamma(\rho)$, integrate with respect to ρ over $[0,1]$, and raise the obtained result to the power $1/q$ $(1 \le q < \infty)$. Finally, we have

$$\|f - p_{n-1}\|_{\mathscr{L}_{q,\gamma}} \le \frac{1}{\alpha_{n,r}} \left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho \right)^{1/q} E_{n-r-1}(f^{(r)})_{H_q}.$$

This implies inequality (2.4). The equality in (2.4) for the function $f_0(z) = z^n$ is verified by direct calculation. The proof of lemma is complete. $\square$

P r o o f of Theorem 1. Taikov proved [28, p. 288] the following inequality for an arbitrary function $f \in H_q$ $(1 \le q \le \infty)$:

$$E_{n-1}(f)_{H_q} \le \frac{n}{\pi - 2} \int_0^{\pi/(2n)} \omega_2(f, 2t)_{H_q} dt; \tag{2.6}$$

and the equality in (2.6) for the function $f_0(z) = z^n$, $n \in \mathbb{N}$.

Replacing in (2.6) the number n with $n - r$ and the function f with $f^{(r)} \in H_q$, we obtain the following inequality for any function $f \in H_q^{(r)}$:

$$E_{n-r-1}(f^{(r)})_{H_q} \le \frac{n-r}{\pi - 2} \int_0^{\pi/2(n-r)} \omega_2(f^{(r)}, 2t)_{H_q} dt. \tag{2.7}$$

In view of (2.7), we can write inequality (2.4) in the form

$$E_{n-1}(f)_{\mathscr{L}_{q,\gamma}} \le \frac{1}{\alpha_{n,r}} \left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho \right)^{1/q} \frac{n-r}{\pi - 2} \int_0^{\pi/2(n-r)} \omega_2(f^{(r)}, 2t)_{H_q} dt. \tag{2.8}$$

From (2.8), assuming that $f \in W_q^{(r)}(\Phi)$, we obtain

$$E_{n-1}(f)_{\mathscr{L}_{q,\gamma}} \le \frac{1}{\alpha_{n,r}} \left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho \right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)} \right).$$

Hence, by relations (1.3), we write upper estimates for the Bernstein and Kolmogorov n-widths:

$$\begin{aligned} b_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) \le d_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) \le E_{n-1}\left(W_q^{(r)}(\Phi)\right)_{\mathscr{L}_{q,\gamma}} \\ \le \frac{1}{\alpha_{n,r}}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)}\right). \end{aligned} \tag{2.9}$$

To obtain a similar upper estimate for the linear n-width, we will use a result of Vakarchuk [36, p. 324]. He proved the following inequality for an arbitrary function $f \in W_q^{(r)}(\Phi)$ ($r \in \mathbb{Z}_+$, $1 \le q \le \infty$) for all $n \in \mathbb{N}$ and $0 < \rho \le 1$:

$$\left\| f(\rho e^{i(\cdot)}) - \Lambda_{n-1}^*(f, \rho e^{i(\cdot)}) \right\|_{H_q} \le \frac{\rho^n}{\alpha_{n,r}} \Phi\left(\frac{\pi}{2(n-r)}\right),$$

Hence, we obtain an upper estimate for the linear n-widths:

$$\begin{aligned} \delta_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) \le \mathscr{E}_{n-1}\left(W_q^{(r)}(\Phi)\right)_{\mathscr{L}_{q,\gamma}} \\ = \sup\left\{\|f - \Lambda_{n-1}^*(f)\|_{\mathscr{L}_{q,\gamma}} : f \in W_q^{(r)}(\Phi)\right\} \le \frac{1}{\alpha_{n,r}}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)}\right). \end{aligned} \tag{2.10}$$

Relations (2.9) and (2.10) imply the following upper estimates for the n-widths $b_n(\cdot)$, $d_n(\cdot)$, and $\delta_n(\cdot)$:

$$\begin{aligned} \lambda_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) \le E_{n-1}\left(W_q^{(r)}(\Phi); L_n^*\right)_{\mathscr{L}_{q,\gamma}} \le \mathscr{E}_{n-1}\left(W_q^{(r)}(\Phi); L_n^*\right)_{\mathscr{L}_{q,\gamma}} \\ \le \frac{1}{\alpha_{n,r}}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)}\right), \end{aligned} \tag{2.11}$$

where $\lambda(\cdot)$ is any of the n-widths $b_n(\cdot)$, $d_n(\cdot)$, or $\delta_n(\cdot)$.

It is known [8, Ch. II, Sect. 3] that, if X and Y are linear normed spaces and X is the subspace of Y ($X \subset Y$), then $d^n(\mathfrak{N}, X) = d^n(\mathfrak{N}, Y)$, where $\mathfrak{N} \subset X$. Consequently, we can write

$$d^n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) = d^n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right).$$

By definition of the Bernstein n-width, we write

$$b_n\left(W_q^{(r)}(\Phi), \mathscr{L}_{q,\gamma}\right) \ge b_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right).$$

In view of relation (1.3), to complete the proof of Theorem 1, it remains to obtain the inequality

$$b_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) \ge \frac{1}{\alpha_{n,r}}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)}\right).$$

To this end, let us introduce the $(n+1)$-dimensional ball of polynomials

$$\mathbb{B}_{n+1} := \left\{p_n \in \mathscr{P}_n : \|p_n\|_{B_{q,\gamma}} \le \frac{1}{\alpha_{n,r}}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \Phi\left(\frac{\pi}{2(n-r)}\right)\right\}$$

and prove the possibility of the embedding $\mathbb{B}_{n+1} \subset W_q^{(r)}(\Phi)$.

We also introduce the notation

$$(1-\cos nx)_* := \left\{1-\cos nx, \text{ if } 0 < nx \le \pi;\ 2, \text{ if } nx > \pi\right\}.$$

The following inequality was proved in [27] for an arbitrary polynomial $p_n \in \mathscr{P}_n$:

$$\|p_n^{(r)}\|_{H_q} \le \alpha_{n,r}\|p_n\|_{H_q}, \quad n > r, \quad n \in \mathbb{N}, \quad r \in \mathbb{Z}_+.$$

We also need the inequality

$$\rho^{nq}\|p_n\|_{H_q}^q \le M_q^q(p_n,\rho) \quad (n \in \mathbb{N}, \quad 1 \le q \le \infty, \quad 0 < \rho \le 1), \tag{2.12}$$

which follows from the inequality

$$\int_{|z|=1} |p_n(z)|^q|dz| \le \rho^{-(nq+1)} \int_{|z|=\rho} |p_n(z)|^q|dz|$$

established by Hille, Szegő, and Tamarkin (see, for example, [25]). Multiplying both sides of (2.12) by $\rho\gamma(\rho)$ and integrating with respect to ρ over $[0,1]$, we obtain

$$\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q} \|p_n\|_{H_q} \le \|p_n\|_{B_{q,\gamma}}$$

and hence

$$\|p_n\|_{H_q} \le \left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{-1/q} \|p_n\|_{B_{q,\gamma}}. \tag{2.13}$$

To prove that the ball $\mathbb{B}_{n+1}$ belongs to the class $W_q^{(r)}(\Phi)$, we will use the inequality

$$\omega_2(p_n^{(r)},2t)_{H_q} \le 2\alpha_{n,r}(1-\cos(n-r)t)_*\|p_n\|_{H_q} \tag{2.14}$$

obtained from one of Taikov's result [28].

Consider two cases: $2k \ge n-r$ and $2k < n-r$.

Let $2k \ge n-r$. By (2.13) and (2.14), for an arbitrary polynomial $p_n \in \mathbb{B}_{n+1}$, we have

$$\begin{aligned} &\frac{k}{\pi-2}\int_0^{\pi/(2k)} \omega_2(p_n^{(r)},2t)_{H_q}dt \le 2\alpha_{n,r}\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{-1/q}\|p_n\|_{B_{q,\gamma}} \\ &\times\frac{k}{\pi-2}\int_0^{\pi/(2k)}(1-\cos(n-r)t)dt \le \frac{\pi}{\pi-2}\left(1-\frac{2k}{\pi(n-r)}\sin\frac{\pi(n-r)}{2k}\right)\Phi\left(\frac{\pi}{2(n-r)}\right). \end{aligned} \tag{2.15}$$

Using (2.15) and the first inequality from (2.1) with

$$t = \frac{\pi}{2(n-r)}, \quad \lambda = \frac{n-r}{k}, \quad \lambda t = \frac{\pi}{2k}, \tag{2.16}$$

we obtain

$$\frac{k}{\pi-2}\int_0^{\pi/(2k)} \omega_2(p_n^{(r)},2t)_{H_q}dt \le \Phi\left(\frac{\pi}{2k}\right). \tag{2.17}$$

Let $2k < n-r$. By (2.14) and (2.13), for an arbitrary polynomial $p_n \in \mathbb{B}_{n+1}$, we have

$$\frac{k}{\pi-2}\int_0^{\pi/(2k)} \omega_2(p_n^{(r)},2t)_{H_q}dt$$

$$\begin{aligned} &\le \Phi\left(\frac{\pi}{2(n-r)}\right)\frac{k}{\pi-2}\left(\int_0^{\pi/(n-r)} 2(1-\cos(n-r)t)dt + \int_{\pi/(n-r)}^{\pi/(2k)} 4dt\right) \\ &= \frac{2\pi}{\pi-2}\left(1-\frac{k}{n-r}\right)\Phi\left(\frac{\pi}{2(n-r)}\right). \end{aligned} \tag{2.18}$$

Using (2.16) and the second inequality from (2.1) with (2.18), we obtain equality (2.17). The inclusion $\mathbb{B}_{n+1} \subset W_q^{(r)}(\Phi)$ is proved. Then, by definition of the Bernstein n-width, we obtain

$$b_n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) \ge b_n(\mathbb{B}_{n+1}, B_{q,\gamma}) \ge \frac{1}{\alpha_{n,r}}\,\Phi\left(\frac{\pi}{2(n-r)}\right)\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q}. \tag{2.19}$$

Comparing relations (2.11) and (2.19), we obtain the required equality (2.3).

It follows from the proof of Theorem 1 that the subspace L_n^* is extremal for the class $W_q^{(r)}(\Phi)$ in the space $\mathscr{L}_{q,\gamma}$ in the case of exact values of the Kolmogorov n-width $d_n(\cdot)$ and the linear n-width $\delta_n(\cdot)$. The subspace $\bar{L}_{n+1}$ is extremal for the Bernstein n-width $b_n(\cdot)$. The linear continuous operator Λ_{n-1}^* defined by equality (2.2) is the best linear approximation method for the class $W_q^{(r)}(\Phi)$ in $\mathscr{L}_{q,\gamma}$. By definition of the Gelfand n-width, the last inequality in (2.10) particularly implies the following inequality for an arbitrary function $f \in W_q^{(r)}(\Phi)$ in the case $c_k(f) = 0$, $k = \overline{0, n-1}$:

$$\begin{aligned} d^n\left(W_q^{(r)}(\Phi), B_{q,\gamma}\right) &\le \sup\left\{\|f\|_{B_{q,\gamma}} : f \in W_q^{(r)}(\Phi) \cup L_*^n\right\} \\ &\le \frac{1}{\alpha_{n,r}}\,\Phi\left(\frac{\pi}{2(n-r)}\right)\left(\int_0^1 \rho^{nq+1}\gamma(\rho)d\rho\right)^{1/q}. \end{aligned} \tag{2.20}$$

Comparing inequalities (2.19) and (2.20) and taking into account relation (1.3), we see that the subspace L_*^n of codimension n is extremal for the Gelfand n-widths $d^n(\cdot)$. Theorem 1 is proved. □

3. Conclusion

In the Banach spaces $\mathscr{L}_{q,\gamma}$ and $B_{q,\gamma}$, $1 \le q \le \infty$, with a weight γ, exact values of some n-widths of the classes $W_q^{(r)}(\Phi)$, $r \in \mathbb{Z}_+$, have been calculated. It was proved that the subspace L_n^* is extremal for the Kolmogorov and linear n-widths in the class $W_q^{(r)}(\Phi)$, the continuous linear operator Λ_{n-1}^* is the best linear approximation method for $W_q^{(r)}(\Phi)$ in $\mathscr{L}_{q,\gamma}$, and the subspace L_*^n is extremal for the n-width $d^n(W_q^{(r)}(\Phi), B_{q,\gamma})$. The subspace $\bar{L}_{n+1}$ is extremal for the n-width $b_n(W_q^{(r)}(\Phi), B_{q,\gamma})$.

REFERENCES

1. Ainulloev N., Taikov L. V. Best approximation in the sense of Kolmogorov of classes of functions analytic in the unit disc. *Math. Notes*, 1986. Vol. 40, No. 3. P. 699–705. DOI: 10.1007/BF01142473
2. Babenko K. I. Best approximations for a class of analytic functions. *Izv. Akad. Nauk SSSR. Ser. Mat.*, 1958. Vol. 22, No. 5. P. 631–640. (in Russian)
3. Dveĭrin M. Z. Widths and ε-entropy of classes of analytic functions in the unit disc. *Teor. Funkts., Funkts. Anal. Prilozh.*, 1975. Vol. 23. P. 32–46. (in Russian)
4. Dvejrin M. Z., Chebanenko I. V. On the polynomial approximation in Banach spaces of analytic functions. In: *Theory of mappings and approximation of functions. Collect. Sci. Works.* Kiev: Naukova Dumka, 1983. P. 62–73. (in Russian)
5. Farkov Yu. A. Widths of Hardy classes and Bergman classes on the ball in $\mathbb{C}^n$. *Russian Math. Surveys*, 1990. Vol. 45, No. 5. P. 229–231. DOI: 10.1070/RM1990v045n05ABEH002677
6. Farkov Yu. A. n-Widths, Faber expansion, and computation of analytic functions. *J. Complexity*, 1996. Vol. 12, No. 1. P. 58–79. DOI: 10.1006/jcom.1996.0007
7. Fisher S. D., Stessin M. I. The n-width of the unit ball of H^q. *J. Approx. Theory*, 1991. Vol. 67, No. 3. P. 347–356. DOI: 10.1016/0021-9045(91)90009-Y
8. Pinkus A. *n-Widths in Approximation Theory.* Heidelberg: Springer-Verlag, 1985. 252 p. DOI: 10.1007/978-3-642-69894-1

9. Saidusainov M. S. On the best linear method of approximation of some classes analytic functions in the weighted Bergman space. *Chebyshevskii Sb.*, 2016. Vol. 17, No. 1. P. 240–253. (in Russian)
10. Saidusainov M. S. $\mathscr{K}$-functionals and exact values of n-widths in the Bergman space. *Ural Math. J.*, 2017. Vol. 3, No. 2. P. 74–81. DOI: 10.15826/umj.2017.2.010
11. Saidusainov M. S. Analysis of a theorem on the Jackson–Stechkin inequality in the Bergman space B_2. *Trudy Inst. Mat. Mekh. UrO RAN,* 2018. Vol. 24, No. 4. P. 217–224. DOI: 10.21538/0134-4889-2018-24-4-217-224 (in Russian)
12. Saidusainov M. S. Some inequalities between the best simultaneous approximation and the modulus of continuity in a weighted Bergman space. *Ural Math. J.*, 2023. Vol. 9, No. 2. P. 165–174. DOI: 10.15826/umj.2023.2.014
13. Shabozov M. Sh. Widths of some classes of analytic functions in the Bergman space. *Dokl. Math.*, 2002. Vol. 65, No. 2. P. 194–197.
14. Shabozov M. Sh. On the best simultaneous approximation of functions in the Hardy space. *Trudy Inst. Mat. Mekh. UrO RAN,* 2023. Vol. 29, No. 4. P. 283–291. DOI: 10.21538/0134-4889-2023-29-4-283-291 (in Russian)
15. Shabozov M. Sh. On the best simultaneous approximation in the Bergman space B_2. *Math. Notes*, 2023. Vol. 114. P. 377–386. DOI: 10.1134/S0001434623090080
16. Shabozov M. Sh., Saidusaynov M. S. Mean-square approximation of complex variable functions by Fourier series in the weighted Bergman space. *Vladikavkazskii Mat. Zh.*, 2018. Vol. 20, No. 1. P. 86–97. DOI: 10.23671/VNC.2018.1.11400 (in Russian)
17. Shabozov M. Sh., Saidusaynov M. S. Upper bounds for the approximation of certain classes of functions of a complex variable by Fourier series in the space L_2 and n-widths. *Math. Notes*, 2018. Vol. 103. P. 656–668. DOI: 10.1134/S0001434618030343
18. Shabozov M. Sh., Saidusainov M. S. Mean-square approximation of functions of a complex variable by Fourier sums in orthogonal systems. *Trudy Inst. Mat. Mekh. UrO RAN,* 2019. Vol. 25, No. 2. P. 258–272. DOI: 10.21538/0134-4889-2019-25-2-258-272 (in Russian)
19. Shabozov M. Sh., Saidusaynov M. S. approximation of functions of a complex variable by Fourier sums in orthogonal systems in L_2. *Russ. Math.*, 2020. Vol. 64. P. 56–62. DOI: 10.3103/S1066369X20060080
20. Shabozov M. Sh., Saidusainov M. S. Mean-squared approximation of some classes of complex variable functions by Fourier series in the weighted Bergman space $B_{2,\gamma}$ *Chebyshevskii Sb.*, 2022. Vol. 23, No. 1. P. 167–182. DOI: 10.22405/2226-8383-2022-23-1-167-182 (in Russian)
21. Shabozov M. Sh., Shabozov O. Sh. Widths of some classes of analytic functions in the Hardy space H_2. *Math. Notes*, 2000. Vol. 68, No. 5–6. P. 675–679. DOI: 10.1023/A:1026692112651
22. Shabozov M. Sh., Shabozov O. Sh. On the best approximation of some classes of analytic functions in weighted Bergman spaces. *Dokl. Math.*, 2007. Vol. 75, No. 1. P. 97–100. DOI: 10.1134/S1064562407010279
23. Shabozov M. Sh., Yusupov G. Y. Best approximation and width of some classes of analytic functions. *Dokl. Math.*, 2002. Vol. 65, No. 1. P. 111–113.
24. Shabozov M. Sh., Yusupov G. A. On the best polynomial approximation of functions in the Hardy space $H_{q,R}\,(1 \leq q \leq \infty, R \geq 1)$. *Chebyshevskii Sb.*, 2023. Vol. 24, No. 1. P. 182–193. DOI: 10.22405/2226-8383-2023-24-1-182-193 (in Russian)
25. Shikhalev N. I. An inequality of Bernstein-Markov kind for analytic functions. *Dokl. Akad. Nauk Azerb. SSR*, 1975. Vol. 31, No. 8. P. 9–14. (in Russian)
26. Smirnov V. I., Lebedev N. A. *A Constructive Theory of Functions of a Complex Variable*. Moscow-Leningrad: Nauka. 1964. 440 p.
27. Taikov L. V. On the best approximation in the mean of certain classes of analytic functions. *Math. Notes*, 1967. Vol. 1. P. 104–109. DOI: 10.1007/BF01268058
28. Taikov L. V. Diameters of certain classes of analytic functions. *Math. Notes*, 1977. Vol. 22. P. 650–656. DOI: 10.1007/BF01780976
29. Tikhomirov V. M. Diameters of sets in function spaces and the theory of best approximations. *Russian Math. Surveys*, 1960. Vol. 15, No. 3. P. 75–111. DOI: 10.1070/rm1960v015n03abeh004093
30. Tikhomirov V. M. Approximation Theory. In: *Analysis II. Convex Analysis and Approximation Theory, Gamkrelidze R. V. eds.* Encyclopaedia Math. Sci., vol. 14. Berlin, Heidelberg: Springer, 1990. P. 93–243. DOI: 10.1007/978-3-642-61267-1_2

31. Vakarchuk S. B. Widths of certain classes of analytic functions in the Hardy space H_2. *Ukr. Math. J.*, 1989. Vol. 41. No. 6. P. 686–689. DOI: 10.1007/BF01060570
32. Vakarchyuk S. B. Best linear methods of approximation and widths of classes of analytic functions in a disk. *Math. Notes*, 1995. Vol. 57. P. 21–27. DOI: 10.1007/BF02309390
33. Vakarchuk S. B. Exact values of widths for certain functional classes. *Ukr. Math. J.*, 1996. Vol. 48. P. 151–153. DOI: 10.1007/BF02390993
34. Vakarchyuk S. B. Exact values of widths of classes of analytic functions on the disk and best linear approximation methods. *Math. Notes*, 2002. Vol. 72. P. 615–619. DOI: 10.1023/A:1021496620022
35. Vakarchuk S. B. On some extremal problems of approximation theory in the complex plane. *Ukr. Math. J.*, 2004. Vol. 56. P. 1371–1390. DOI: 10.1007/s11253-005-0122-x
36. Vakarchuk S. B., Zabutnaya V. I. Best linear approximation methods for functions of Taikov classes in the Hardy spaces $H_{q,\rho}$, $q \geq 1$, $0 < \rho \leq 1$. *Math. Notes*, 2009. Vol. 85. P. 322–327. DOI: 10.1134/S000143460903002X

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 131–143
DOI: 10.15826/umj.2024.2.012

INTEGRAL ANALOGUE OF TURÁN-TYPE INEQUALITIES CONCERNING THE POLAR DERIVATIVE OF A POLYNOMIAL

Mayanglambam Singhajit Singh†, **Barchand Chanam**††

Department of Mathematics, National Institute of Technology Manipur,
Imphal, 795004, India

†msinghasingho@gmail.com ††barchand_2004@yahoo.co.in

Abstract: If $w(\zeta)$ is a polynomial of degree n with all its zeros in $|\zeta| \leq \Delta$, $\Delta \geq 1$ and any real $\gamma \geq 1$, Aziz proved the integral inequality [1]

$$\left\{\int_0^{2\pi} \left|1+\Delta^n e^{i\theta}\right|^{\gamma} d\theta\right\}^{1/\gamma} \max_{|\zeta|=1} |w'(\zeta)| \geq n \left\{\int_0^{2\pi} \left|w\left(e^{i\theta}\right)\right|^{\gamma} d\theta\right\}^{1/\gamma}.$$

In this article, we establish a refined extension of the above integral inequality by using the polar derivative instead of the ordinary derivative consisting of the leading coefficient and the constant term of the polynomial. Besides, our result also yields other intriguing inequalities as special cases.

Keywords: Polar derivative, Turán-type inequalities, Integral inequalities.

1. Introduction

In the late nineteenth century, renowned chemist Mendeleev became interested in the subject of the extremal properties of polynomials while searching for an upper bound of a quadratic polynomial. More specifically, he [14] established that, if $w(r)$ is a quadratic polynomial of real variable r with real coefficients, then for $-1 \leq w(r) \leq 1$ and $-1 \leq r \leq 1$,

$$\max_{-1\leq r\leq 1} |w'(r)| \leq 4.$$

While working on a problem in Approximation Theory, Bernstein needed an upper bound estimate of the maximum modulus $|w'(\zeta)|$ of a complex polynomial in terms of the maximum modulus of $|w(\zeta)|$, where $|\zeta| = 1$, which is an analogue of above Mendeleev's problem in the complex domain. He [5] proved his famous inequality which states that, if $w(\zeta)$ is a n degree polynomial, then

$$\max_{|\zeta|=1} |w'(\zeta)| \leq n \max_{|\zeta|=1} |w(\zeta)|. \tag{1.1}$$

This inequality is sharp if and only if $w(\zeta) = \delta\zeta^n$, where

$$|\delta| = \max_{|\zeta|=1} |w(\zeta)|.$$

Inequality (1.1) is an immediate consequence of an inequality concerning trigonometric polynomials proved by him.

Paul Turán [21] was the first to estimate the maximum modulus for the derivative of a polynomial through a lower bound in terms of the maximum modulus of the polynomial. He established, in particular, that if $w(\zeta)$ is a n degree polynomial and all of its zeros lie in $|\zeta| \leq 1$, then

$$\max_{|\zeta|=1} w'(\zeta) \geq \frac{n}{2} \max_{|\zeta|=1} |w(\zeta)|. \tag{1.2}$$

Equality in (1.2) attains for $w(\zeta) = \delta\zeta^n + \beta$, where $|\delta| = |\beta|$. If $w(\zeta)$ is a n degree polynomial over the complex numbers $\mathbb{C}$, and for a real number $\gamma > 0$, the integral mean of $w(\zeta)$ is defined by

$$\|w\|_\gamma = \left\{ \frac{1}{2\pi} \int_0^{2\pi} \left| w(e^{i\theta}) \right|^\gamma d\theta \right\}^{1/\gamma}.$$

Taking limit as $\gamma \to \infty$ and using the fact from the analysis [18, 20] that

$$\lim_{\gamma \to \infty} \left\{ \frac{1}{2\pi} \int_0^{2\pi} \left| w(e^{i\theta}) \right|^\gamma d\theta \right\}^{1/\gamma} = \max_{|\zeta|=1} |w(\zeta)|,$$

we can legitimately denote

$$\|w\|_\infty = \max_{|\zeta|=1} |w(\zeta)|.$$

Aziz and Dawood [2] improved (1.2) into the form

$$\|w'\|_\infty \geq \frac{n}{2} \Big\{ \|w\|_\infty + \min_{|\zeta|=1} |w(\zeta)| \Big\}. \tag{1.3}$$

Throughout this paper, $\mathbb{P}_{n,s,\Delta}$ represents the class of all polynomials

$$w(\zeta) = \zeta^s \sum_{j=0}^{n=s} \alpha_j \zeta^j, \quad 0 \leq s \leq n,$$

with zero of multiplicity s at the origin having all its zeros in $|\zeta| \leq \Delta,\ \Delta \geq 1$ and $\mathbb{P}_{n,\Delta}$, the class of all polynomials

$$w(\zeta) = \sum_{j=0}^{n} \alpha_j \zeta^j$$

with all their zeros in $|\zeta| \leq \Delta,\ \Delta \geq 1$.

Applications and interest in inequality (1.2) have been substantial. Thus, it would be very interesting to determine its generalisation for polynomials whose zeros are all in $|\zeta| \leq \Delta,\ \Delta > 0$. For $0 < \Delta \leq 1$, Malik [13] proved

$$\|w'\|_\infty \geq \frac{n}{1+\Delta} \|w\|_\infty. \tag{1.4}$$

For $\Delta \geq 1$, Govil [9] found

$$\|w'\|_\infty \geq \frac{n}{1+\Delta^n} \|w\|_\infty. \tag{1.5}$$

Equality in (1.5) holds for $w(\zeta) = \zeta^n + \Delta^n, \Delta \geq 1$.

Govil [10] refined inequality (1.4) by proving that

$$\|w'\|_\infty \geq \frac{n}{1+\Delta} \Big(\|w\|_\infty + \frac{1}{\Delta^{n-1}} \min_{|\zeta|=\Delta} |w(\zeta)| \Big). \tag{1.6}$$

Equality in (1.6) holds for $w(\zeta) = (\zeta + \Delta)^n$.

For the polynomials which have all their zeros in $|\zeta| \leq \Delta$, $\Delta \leq 1$ with zero of multiplicity s at the origin, Aziz and Shah [4] obtained the following generalization of (1.4) that

$$\|w'\|_\infty \geq \frac{n + s\Delta}{1+\Delta} \|w\|_\infty.$$

The above inequality is sharp with the extremal polynomial being $w(\zeta) = \zeta^s (\zeta + \Delta)^{n-s}$, $0 \leq s \leq n$.

Using the same assumption, Govil [10] was able to improve (1.5) as

$$\|w'\|_\infty \geq \frac{n}{1+\Delta^n}\left\{\|w\|_\infty + \min_{|\zeta|=\Delta}|w(\zeta)|\right\}. \tag{1.7}$$

Inequality (1.7) attains equality for

$$w(\zeta) = \zeta^n + \Delta^n, \quad \Delta \geq 1.$$

Malik [12] extended inequality (1.2) for the first time in 1984 into its integral analogue by establishing that if $w(\zeta)$ is a n degree polynomial with all its zeros in $|\zeta| \leq 1$, then for $\gamma > 0$,

$$\|1+\zeta\|_\gamma \|w'\|_\infty \geq n\|w\|_\gamma.$$

The result is best possible for $w(\zeta) = (\zeta+1)^n$.

In 1988, Aziz [1] extended to integral form of (1.5) by establishing

Theorem 1. *If $w(\zeta) \in \mathbb{P}_{n,\Delta}$, then for $\gamma \geq 1$,*

$$\|1+\Delta^n\zeta\|_\gamma \|w'\|_\infty \geq n\|w\|_\gamma. \tag{1.8}$$

Equality in (1.8) *holds for*

$$w(\zeta) = \delta\zeta^n + \beta\Delta^n, \quad |\delta| = |\beta|.$$

For a n degree polynomial $w(\zeta)$ and any $\delta \in \mathbb{C}$, we define the polar derivative of the polynomial $w(\zeta)$ with regard to δ by

$$D_\delta w(\zeta) = nw(\zeta) + (\delta - \zeta)w'(\zeta).$$

Note that $D_\delta w(\zeta)$ has atmost $n-1$ degree, and it is a generalization of the ordinary derivative as

$$\lim_{\delta\to\infty} \frac{D_\delta w(\zeta)}{\delta} = w'(\zeta),$$

uniformly with respect to ζ for $|\zeta| \leq R, R > 0$.

Inequality (1.4) was first extended to the polar derivative by Aziz and Rather [3]. They obtained that if $w(\zeta)$ is a n degree polynomial with all its zeros in $|\zeta| \leq \Delta, \Delta \leq 1$, then for $\delta \in \mathbb{C}$, $|\delta| \geq \Delta$,

$$\|D_\delta w\|_\infty \geq n\left(\frac{|\delta|-\Delta}{1+\Delta}\right)\|w\|_\infty.$$

Besides, in the same article [3], they could extend (1.5) to polar derivative by proving that

$$\|D_\delta w\|_\infty \geq n\left(\frac{|\delta|-\Delta}{1+\Delta^n}\right)\|w\|_\infty, \tag{1.9}$$

where $\delta \in \mathbb{C}$ with $|\delta| \geq \Delta$.

Dewan et al. [7] obtained the polar derivative version of (1.7), which also sharpens (1.9) by proving that if $w(\zeta) \in \mathbb{P}_{n,\Delta}$, then for $\delta \in \mathbb{C}$ with $|\delta| \geq \Delta$,

$$\|D_\delta w\|_\infty \geq \frac{n}{1+\Delta^n}\Big\{(|\delta|-\Delta)\|w\|_\infty + \Big(|\delta| + \frac{1}{\Delta^{n-1}}\Big)\min_{|\zeta|=\Delta}|w(\zeta)|\Big\}. \tag{1.10}$$

The following generalization and improvement of (1.9) consisting of the polynomial's constant term and leading coefficient was recently established by Singh and Chanam [19].

Theorem 2. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then $\delta \in \mathbb{C}$ with $|\delta| \geq \Delta$,*

$$\|D_\delta w\|_\infty \geq \frac{|\delta| - \Delta}{1 + \Delta^n} \left\{ n + s + \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}|} - \sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}|}} \right\} \|w\|_\infty . \tag{1.11}$$

Milovanovic et al. [15] proved the following improvement and generalization of (1.9), (1.10) and (1.11).

Theorem 3. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then for $\delta \in \mathbb{C}$, $|\delta| \geq \Delta$,*

$$\begin{aligned} \|D_\delta w\|_\infty &\geq \frac{n}{1 + \Delta^n} \left\{ \left(|\delta| - \Delta\right) \|w\|_\infty + \left(|\delta| + \frac{1}{\Delta^{n-1}}\right) m \right\} \\ &+ \frac{|\delta| - \Delta}{1 + \Delta^n} \left\{ s + \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}| - m} - \sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}| - m}} \right\} \left(\|w\|_\infty - \frac{m}{\Delta^n} \right), \end{aligned} \tag{1.12}$$

where $m = \min\limits_{|\zeta|=\Delta} |w(\zeta)|$.

2. Main result

Below we derive the generalized integral extension of Theorem 2, which further improves Theorem 3 and also gives many other interesting results as special cases. In particular, we prove

Theorem 4. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then for $\delta \in \mathbb{C}$, $|\delta| \geq \Delta$ and $\lambda \in \mathbb{C}$, $|\lambda| < 1$ and $\gamma > 0$,*

$$\left\| D_\delta \left\{ w(e^{i\theta}) - \frac{m}{\Delta^n} \lambda e^{in\theta} \right\} \right\|_\gamma \geq \frac{|\delta| - \Delta}{2E_\gamma} A \left\| w(e^{i\theta}) - \frac{m}{\Delta^n} \lambda e^{in\theta} \right\|_\gamma , \tag{2.1}$$

where

$$m = \min_{|\zeta|=\Delta} |w(\zeta)|, \quad A = \left\{ n + s + \frac{\sqrt{\Delta^n|\alpha_{n-s}| - |\lambda| m} - \sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}| - |\lambda| m}} \right\}$$

and

$$E_\gamma = \frac{\left\{ \int_0^{2\pi} \left|1 + \Delta^{n-s} e^{i\theta}\right|^\gamma d\theta \right\}^{1/\gamma}}{\left\{ \int_0^{2\pi} \left|1 + e^{i\theta}\right|^\gamma d\theta \right\}^{1/\gamma}}.$$

Remark 1. Suppose $w(\zeta)$ has all its zeros in $|\zeta| \leq \Delta$, $\Delta \geq 1$. Now, for $|\zeta| = \Delta$

$$m = \min_{|\zeta|=\Delta} |w(\zeta)| \leq |w(\zeta)|. \tag{2.2}$$

As a consequence of Maximum Modulus Principle, we have

$$\max_{|\zeta|=\Delta} |w(\zeta)| \leq \Delta^n \max_{|\zeta|=1} |w(\zeta)|. \tag{2.3}$$

Using (2.3) to (2.2), we get

$$m \leq \Delta^n \max_{|\zeta|=1} |w(\zeta)|,$$

i.e.

$$\frac{m}{\Delta^n} \leq \max_{|\zeta|=1} |w(\zeta)|. \tag{2.4}$$

For arbitrary $\lambda \in \mathbb{C}$, $|\lambda| < 1$, we have

$$\frac{|\lambda| m}{\Delta^n} < \max_{|\zeta|=1} |w(\zeta)|. \tag{2.5}$$

Remark 2. Suppose $\gamma \to \infty$ in (2.1) and knowing the simple fact that

$$E_\gamma \to \frac{1+\Delta^n}{2} \quad \text{as} \quad \gamma \to \infty,$$

we get

$$\max_{|\zeta|=1} \left| D_\delta \left\{ w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right\} \right| \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right|,$$

i.e.

$$\max_{|\zeta|=1} \left| D_\delta w(\zeta) - \frac{|\delta| m n \lambda}{\Delta^n} \zeta^{n-1} \right| \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right|. \tag{2.6}$$

Let ζ_0 on $|\zeta| = 1$ be such that

$$\max_{|\zeta|=1} \left| D_\delta w(\zeta) - \frac{|\delta| m n \lambda}{\Delta^n} \zeta^{n-1} \right| = \left| D_\delta w(\zeta_0) - \frac{|\delta| m n \lambda}{\Delta^n} \zeta_0^{n-1} \right|. \tag{2.7}$$

In the right side of (2.7), we can choose the argument of λ with

$$\left| D_\delta w(\zeta_0) - \frac{|\delta| m n \lambda}{\Delta^n} \zeta_0^{n-1} \right| = |D_\delta w(\zeta_0)| - \frac{n|\delta||\lambda|}{\Delta^n} m. \tag{2.8}$$

From (2.7) and (2.8), (2.6) becomes

$$|D_\delta w(\zeta_0)| - \frac{n|\delta||\lambda|}{\Delta^n} m \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right|. \tag{2.9}$$

Since

$$|D_\delta w(\zeta_0)| \leq \max_{|\zeta|=1} |D_\delta w(\zeta)|,$$

(2.9) gives

$$\max_{|\zeta|=1} |D_\delta w(\zeta)| - \frac{n|\delta||\lambda|}{\Delta^n} m \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right|. \tag{2.10}$$

Let ζ_1 on $|\zeta| = 1$ be such that $\max\limits_{|\zeta|=1} |w(\zeta)| = |w(\zeta_1)|$. Then

$$\max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right| \geq \left| w(\zeta_1) - \frac{m\lambda}{\Delta^n} \zeta^n \right| \geq \left| |w(\zeta_1)| - \frac{m|\lambda|}{\Delta^n} \right|. \tag{2.11}$$

Using (2.5) to (2.11), we get

$$\max_{|\zeta|=1} \left| w(\zeta) - \frac{m\lambda}{\Delta^n} \zeta^n \right| \geq \max_{|\zeta|=1} |w(\zeta)| - \frac{m|\lambda|}{\Delta^n}. \tag{2.12}$$

Using (2.12), (2.10) gives

$$\max_{|\zeta|=1} |D_\delta w(\zeta)| - \frac{n|\delta||\lambda|}{\Delta^n} m \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \Big(\max_{|\zeta|=1} |w(\zeta)| - \frac{|\lambda|}{\Delta^n} m \Big). \tag{2.13}$$

When $|\lambda| \to l$ in (2.13), we have

$$\max_{|\zeta|=1} |D_\delta w(\zeta)| - \frac{n|\delta| l}{\Delta^n} m \geq \frac{|\delta| - \Delta}{1+\Delta^n} A \Big(\max_{|\zeta|=1} |w(\zeta)| - \frac{l}{\Delta^n} m \Big),$$

which becomes the following result on simply taking limit $l \to 1$.

Corollary 1. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then for $\delta \in \mathbb{C}$, $|\delta| \geq \Delta$,*

$$\begin{aligned}\|D_\delta w\|_\infty &\geq \frac{n}{1+\Delta^n}\Big\{(|\delta|-\Delta)\|w\|_\infty + \Big(|\delta| + \frac{1}{\Delta^{n-1}}\Big)m\Big\}\\ &+\frac{|\delta|-\Delta}{1+\Delta^n}\left(s + \frac{\sqrt{\Delta^n|\alpha_{n-s}|-m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-m}}\right)\left(\|w\|_\infty - \frac{m}{\Delta^n}\right),\end{aligned} \tag{2.14}$$

where $m = \min_{|\zeta|=\Delta}|w(\zeta)|$.

Remark 3. Using the three facts (2.4), (4.1) and (4.3) in (2.14), it is obvious that Corollary 1 improves (1.10).

Remark 4. Also, the function

$$f(x) = \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}|-x}-\sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}|-x}}$$

is non-increasing for x. Therefore, for $\Delta \geq 1$

$$f\left(\frac{m}{\Delta^s}\right) \geq f(m),$$

that is,

$$\frac{\sqrt{\Delta^n|\alpha_{n-s}|-m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-m}} \geq \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}|-m}-\sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}|-m}}.$$

This shows that Corollary 1 is an improvement of (1.12).

Remark 5. If we divide both sides of (2.14) by $|\delta|$ and let $|\delta| \to \infty$, the next result which improves (1.7), is obtained.

Corollary 2. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then*

$$\|w'\|_\infty \geq \frac{n}{1+\Delta^n}(\|w\|_\infty + m) + \frac{1}{1+\Delta^n}\left(s + \frac{\sqrt{\Delta^n|\alpha_{n-s}|-m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-m}}\right)\left(\|w\|_\infty - \frac{m}{\Delta^n}\right), \tag{2.15}$$

where $m = \min_{|\zeta|=\Delta}|w(\zeta)|$.

Remark 6. If we divide both sides of (2.1) of Theorem 4 by $|\delta|$ and let $|\delta| \to \infty$, the following generalized integral extension of Corollary 2 is obtained.

Corollary 3. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then for each $\lambda \in \mathbb{C}$, $|\lambda| < 1$ and $\gamma > 0$,*

$$\left\|w'(e^{i\theta}) - \frac{mn}{\Delta^n}\lambda e^{i(n-1)\theta}\right\|_\gamma \geq \frac{A}{2E_\gamma}\left\|w(e^{i\theta}) - \frac{m}{\Delta^n}\lambda e^{in\theta}\right\|_\gamma,$$

where m, A and E_r are defined in Theorem 4.

Remark 7. When $\lambda = 0$ in (2.1) of Theorem 4, the below integral extension of Theorem 2 yields an improved and generalised integral analogue for polar derivative of Theorem 1.

Corollary 4. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then for $\delta \in \mathbb{C}$, $|\delta| \geq \Delta$ and $\gamma > 0$,*

$$\|D_\delta w\|_\gamma \geq \frac{|\delta| - \Delta}{2E_\gamma}\left(n + s + \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}|} - \sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}|}}\right)\|w\|_\gamma\,, \tag{2.16}$$

where E_γ is defined in Theorem 4.

Remark 8. In case $r \to \infty$ in (2.16), Corollary 4, in particular, becomes Theorem 2 and dividing both sides by $|\delta|$ and making $|\delta| \to \infty$, we have an improved form of (1.5).

Corollary 5. *If $w(\zeta) \in \mathbb{P}_{n,s,\Delta}$, then*

$$\|w'\|_\infty \geq \frac{1}{1 + \Delta^n}\left(n + s + \frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}|} - \sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}|}}\right)\|w\|_\infty. \tag{2.17}$$

Remark 9. If degree n of polynomial $w(\zeta)$ is greater than or equal to 1, the leading coefficient α_n is different from zero, and using the fact (4.1), it follows obviously that inequality (2.17) always provides better bounds than that of (1.5). When $\Delta = 1$, (2.15) and (2.17) sharpen (1.3) and (1.2) respectively.

3. Example with numerical illustration

Example. Consider $w(\zeta) = \zeta(\zeta + 1)$ with all zeros $0, -1$. Now, all the zeros lie in the closed disk $|\zeta| \leq 1$. On the unit circle $|\zeta| = 1$,

$$\left|w\left(e^{i\theta}\right)\right| = \sqrt{2 + 2\cos\theta}.$$

Since the non-negative function

$$f(\theta) = 2 + 2\cos\theta, \quad 0 \leq \theta < 2\pi,$$

attains its maximum at $\theta = 0$,

$$\max_{|\zeta|=1}\left|w(\zeta)\right| = 2.$$

For each fixed $\Delta = \Delta_0$,

$$|w(\Delta_0 e^{i\theta})| = \Delta_0\sqrt{\Delta_0^2 + 2\Delta_0\cos\theta + 1}.$$

Since the function

$$g(\theta) = \Delta_0^2 + 2\Delta_0\cos\theta + 1, \quad 0 \leq \theta < 2\pi,$$

attains its minimum at $\theta = \pi$,

$$m = \min_{|\zeta|=\Delta_0}|w(\zeta)| = \Delta_0(\Delta_0 - 1).$$

If we take $\Delta_0 = 1.95$ and $|\delta| = 10$, then by using Theorem 2, we have

$$\|D_{10}w\|_\infty \geq \frac{10 - 1.95}{1 + 1.95^2}\left\{2 + 1 + \frac{\sqrt{1.95 \times 1} - \sqrt{1}}{\sqrt{1.95 \times 1}}\right\} \times 2 \approx 11.009,$$

while by Theorem 3,

$$\|D_{10}w\|_\infty \geq \frac{2}{1+1.95^2}\Big\{(10-1.95)2+\Big(10+\frac{1}{1.95}\Big)1.95(1.95-1)\Big\}$$
$$+\frac{10-1.95}{1+1.95^2}\Bigg\{1+\frac{\sqrt{1.95\times 1-1.95(1.95-1)}-\sqrt{1}}{\sqrt{1.95\times 1-1.95(1.95-1)}}\Bigg\}\Bigg(2-\frac{1.95(1.95-1)}{1.95^2}\Bigg)\approx 11.7657$$

Meanwhile, if we use Corollary 1, we get

$$\|D_{10}w\|_\infty \geq \frac{2}{1+1.95^2}\Big\{(10-1.95)2+\Big(10+\frac{1}{1.95}\Big)1.95(1.95-1)\Big\}$$
$$+\frac{10-1.95}{1+1.95^2}\Bigg\{1+\frac{\sqrt{1.95^2\times 1-1.95(1.95-1)}-\sqrt{1.95\times 1}}{\sqrt{1.95^2\times 1-1.95(1.95-1)}}\Bigg\}\Bigg\{2-\frac{1.95(1.95-1)}{1.95^2}\Bigg\}\approx 17.351,$$

which is larger than the bounds obtained by using Theorems 2 and 3. In other words, the bound of Corollary 1 improves over those of Theorems 2 and 3 respectively due to Singh and Chanam [19] and Milovanovic et al. [15] by about 57.61% and 47.47%. From this, it is easy to see that by making appropriate choices of the polynomial $w(\zeta)$, and the parameters Δ and δ, this improvement can be scaled up.

4. Lemmas

We need the following auxiliary results to prove the theorem and its corollaries. For a n degree polynomial $w(\zeta)$, we will use

$$q(\zeta)=\zeta^n\overline{w\left(1/\bar{\zeta}\right)}.$$

Lemma 1 [13]. *If $w(\zeta)$ is a n degree polynomial with all its zeros in $|\zeta|\leq\Delta$, $\Delta\leq 1$, then for $|\zeta|=1$,*

$$|q'(\zeta)|\leq\Delta|w'(\zeta)|.$$

Lemma 2. *If $w(\zeta)$ is a n degree polynomial, then for $R\geq 1$ and $\gamma>0$,*

$$\left\{\int_0^{2\pi}\left|w\left(Re^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}\leq R^n\left\{\int_0^{2\pi}\left|w\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}.$$

It is difficult to trace the origin of Lemma 2. However, it could be followed from a famous result of Hardy [11], by which for any function $f(\zeta)$ analytic in $|\zeta|<t_0$, and for each $\gamma>0$,

$$\left\{\int_0^{2\pi}\left|f\left(xe^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}$$

is non-decreasing for $x\in(0,t_0)$. If $w(\zeta)$ is a n degree polynomial, then

$$f(\zeta)=\zeta^n\overline{w\left(1/\bar{\zeta}\right)}$$

is a polynomial of degree at most n and is an entire function, and by Hardy's result for $\gamma>0$,

$$\left\{\int_0^{2\pi}\left|f\left(xe^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}\leq\left\{\int_0^{2\pi}\left|f\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma},$$

for $x=1/R<1$, and hence Lemma 2.

Lemma 3 [19]. *If $w(\zeta) \in P_{n,s,1}$, then for $|\zeta| = 1$,*

$$\left|w'(\zeta)\right| \geq \frac{1}{2}\left\{n + s + \frac{\sqrt{|\alpha_{n-s}|} - \sqrt{|\alpha_0|}}{\sqrt{|\alpha_{n-s}|}}\right\}|w(\zeta)|.$$

Lemma 4 [6, 16]. *If $w(\zeta)$ is a n degree polynomial and $w(\zeta) \neq 0$ in $|\zeta| < 1$, then for $R \geq 1$ and $\gamma > 0$,*

$$\left\{\int_0^{2\pi} \left|w\left(Re^{i\theta}\right)\right|^{\gamma} d\theta\right\}^{1/\gamma} \leq B_{\gamma}\left\{\int_0^{2\pi} \left|w\left(e^{i\theta}\right)\right|^{\gamma} d\theta\right\}^{1/\gamma},$$

where

$$B_{\gamma} = \frac{\left\{\int_0^{2\pi} |1 + R^n e^{i\theta}|^{\gamma} d\theta\right\}^{1/\gamma}}{\left\{\int_0^{2\pi} |1 + e^{i\theta}|^{\gamma} d\theta\right\}^{1/\gamma}}.$$

This is due to Boas and Rahman [6] for $\gamma \geq 1$. Later, Rahman and Schmeisser [16] verified validity for $0 < \gamma < 1$.

Lemma 5 [8]. *If $w(\zeta)$ is a n degree polynomial and $w(\zeta) \neq 0$ in $|\zeta| < \Delta$, $\Delta > 0$, then for $|\zeta| < \Delta$*

$$|w(\zeta)| > m,$$

where $m = \min_{|\zeta|=\Delta} |w(\zeta)|$.

Lemma 6. *If*

$$w(\zeta) = \zeta^s\left(\sum_{j=0}^{n-s} \alpha_j \zeta^j\right), \quad 0 \leq s \leq n,$$

is a polynomial with all its zeros in $|\zeta| \leq \Delta$, $\Delta > 0$, then for $\lambda \in \mathbb{C}$, $|\lambda| < 1$

$$\sqrt{\Delta^{n-s}|\alpha_{n-s}| - |\lambda|\frac{m}{\Delta^s}} - \sqrt{|\alpha_0|} \geq 0, \tag{4.1}$$

where $m = \min_{|\zeta|=\Delta} |w(\zeta)|$.

P r o o f. By hypothesis,

$$w(\zeta) = \zeta^s h(\zeta) = \zeta^s\left(\sum_{j=0}^{n-s} \alpha_j \zeta^j\right), \quad 0 \leq s \leq n,$$

has all its zeros in $|\zeta| \leq \Delta$, $\Delta > 0$. Then, the polynomial $W(\zeta) = e^{-i\arg \alpha_{n-s}} h(\zeta)$ has the same zeros as $h(\zeta)$.
Now,

$$\begin{aligned} W(\zeta) &= e^{-i\arg \alpha_{n-s}}\left\{\alpha_0 + \alpha_1\zeta + \cdots + \alpha_{n-s-1}\zeta^{n-s-1} + |\alpha_{n-s}|e^{i\arg \alpha_{n-s}}\zeta^{n-s}\right\} \\ &= e^{-i\arg \alpha_{n-s}}\left\{\alpha_0 + \alpha_1\zeta + \cdots + \alpha_{n-s-1}\zeta^{n-s-1}\right\} + |\alpha_{n-s}|\zeta^{n-s}. \end{aligned}$$

Now, on $|\zeta| = \Delta$ for $\lambda \in \mathbb{C}$, $|\lambda| < 1$ and $m = \min_{|\zeta|=\Delta} w(\zeta) \neq 0$, we have

$$\left|\frac{m\lambda}{\Delta^n}\zeta^{n-s}\right| < \frac{m}{\Delta^s} = \min_{|\zeta|=\Delta} |h(\zeta)| = \min_{|\zeta|=\Delta} |W(\zeta)| \leq |W(\zeta)|.$$

Then by Rouche's theorem,

$$R(\zeta) = W(\zeta) - \frac{m|\lambda|}{\Delta^n}\zeta^{n-s}$$

has all its zeros in $|\zeta| < \Delta$. By Vieta's formula applied to $R(\zeta)$, we get

$$\frac{|\alpha_0|}{\big||\alpha_{n-s}| - m|\lambda|/\Delta^n\big|} < \Delta^{n-s},$$

that is,

$$\left(\frac{|\alpha_0|}{\big||\alpha_{n-s}| - m|\lambda|/\Delta^n\big|}\right)^{1/2} < \Delta^{(n-s)/2}. \tag{4.2}$$

Since $W(\zeta)$ is a polynomial of degree $n-s$ with all its zeros in $|\zeta| \le \Delta$, then

$$Q(\zeta) = \zeta^{n-s}\overline{W\left(1/\overline{\zeta}\right)}$$

is a polynomial having atmost $n-s$ degree having no zero in $|\zeta| < 1/\Delta$. Using Lemma 5 to $Q(\zeta)$, we obtain

$$|\alpha_{n-s}| = |Q(0)| > \min_{|\zeta|=1/\Delta} |Q(\zeta)| = \frac{1}{\Delta^{n-s}} \min_{|\zeta|=\Delta} |W(\zeta)| = \frac{m}{\Delta^n},$$

i.e.

$$|\alpha_{n-s}| > \frac{m}{\Delta^n}. \tag{4.3}$$

Using (4.3) to (4.2), we have

$$\sqrt{\Delta^{n-s}|\alpha_{n-s}| - |\lambda|\frac{m}{\Delta^s}} - \sqrt{|\alpha_0|} > 0.$$

For $m = \min_{|\zeta|=\Delta} |w(\zeta)| = 0$, the result becomes trivial, simply by the similar reasoning of inequality (4.2) to

$$h(\zeta) = \sum_{j=0}^{n-s} \alpha_j \zeta^j,$$

i.e.

$$\sqrt{\Delta^{n-s}|\alpha_{n-s}|} - \sqrt{|\alpha_0|} \ge 0.$$

□

5. Proof of Theorem 4

By assumption, $w(\zeta)$ has all its zeros in $|\zeta| \le \Delta$, $\Delta \ge 1$. For $m = \min_{|\zeta|=\Delta} |w(\zeta)| \ne 0$, consider

$$R(\zeta) = w(\zeta) - \frac{m}{\Delta^n}\lambda\zeta^n,$$

where $\lambda \in \mathbb{C}$, $|\lambda| < 1$. Now, on $|\zeta| = \Delta$

$$\left|\frac{m}{\Delta^n}\lambda\zeta^n\right| < \frac{m}{\Delta^n}\Delta^n \le |w(\zeta)|.$$

Consequently, from Rouche's theorem, $R(\zeta)$ has all its zeros in $|\zeta| < \Delta$. When $m = 0$, $R(\zeta) = w(\zeta)$. Therefore, $R(\zeta)$ has all its zeros in $|\zeta| \le \Delta$ in any case. Then, all the zeros of $W(\zeta) = R(\Delta\zeta)$ are in $|\zeta| \le 1$. It is a simple fact that for $|\zeta| = 1$

$$|Q'(\zeta)| = \left|nW(\zeta) - \zeta W'(\zeta)\right|, \tag{5.1}$$

where
$$Q(\zeta)=\zeta^n\overline{W\left(1/\bar{\zeta}\right)}.$$
Using Lemma 1 to $W(\zeta)$, we have for $|\zeta|=1$
$$\left|Q'(\zeta)\right|\le\left|W'(\zeta)\right|. \tag{5.2}$$
Using (5.1) and (5.2), we have for $|\delta/\Delta|\ge 1$ and $|\zeta|=1$
$$\begin{aligned}\left|D_{\delta/\Delta}W(\zeta)\right|&=\left|nW(\zeta)+\Big(\frac{\delta}{\Delta}-\zeta\Big)W'(\zeta)\right|\ge\left|\frac{\delta}{\Delta}\right|\left|W'(\zeta)\right|-\left|nW\left(\zeta\right)-\zeta W'\left(\zeta\right)\right|\\&=\left|\frac{\delta}{\Delta}\right||W'(\zeta)|-|Q'\left(\zeta\right)|\ge\Big(\left|\frac{\delta}{\Delta}\right|-1\Big)|W'(\zeta)|.\end{aligned} \tag{5.3}$$
Applying Lemma 3 to $W(\zeta)$, we have for $|\zeta|=1$
$$\left|W'(\zeta)\right|\ge\frac{1}{2}\Big\{n+s+\frac{\sqrt{\Delta^{n-s}|\alpha_{n-s}-(m/\Delta^n)\lambda|}-\sqrt{|\alpha_0|}}{\sqrt{\Delta^{n-s}|\alpha_{n-s}-(m/\Delta^n)\lambda|}}\Big\}|W(\zeta)|.$$
Since $f(x)=(x-|a|)/x$ is non-decreasing and in view of (4.3), we get
$$\left|W'(\zeta)\right|\ge\frac{1}{2}\Big\{n+s+\frac{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}}\Big\}|W(\zeta)|. \tag{5.4}$$
Combining (5.4) and (5.3), we get
$$\left|D_{\delta/\Delta}W\left(\zeta\right)\right|\ge\frac{|\delta|-\Delta}{2\Delta}\Big\{n+s+\frac{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}}\Big\}|W(\zeta)|.$$
Replacing $W(\zeta)$ by $R(\Delta\zeta)$, this inequality gives
$$\left|nR(\Delta\zeta)+\Big(\frac{\delta}{\Delta}-\zeta\Big)\Delta R'(\Delta\zeta)\right|\ge\frac{|\delta|-\Delta}{2\Delta}A|R(\Delta\zeta)|, \tag{5.5}$$
where
$$A=\{n+s+\frac{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}-\sqrt{\Delta^s|\alpha_0|}}{\sqrt{\Delta^n|\alpha_{n-s}|-|\lambda|m}}\}.$$
Inequality (5.5) becomes
$$\left|nR(\Delta\zeta)+\left(\delta-\Delta\zeta\right)R'(\Delta\zeta)\right|\ge\frac{|\delta|-\Delta}{2\Delta}A|R(\Delta\zeta)|,$$
therefore for any $\gamma>0$, we have
$$\left|D_\delta R\big(\Delta e^{i\theta}\big)\right|^\gamma\ge\Big(\frac{|\delta|-\Delta}{2\Delta}A\Big)^\gamma\left|R\big(\Delta e^{i\theta}\big)\right|^\gamma,\quad 0\le\theta<2\pi.$$
Equivalently,
$$\Big\{\int_0^{2\pi}\left|D_\delta R\big(\Delta e^{i\theta}\big)\right|^\gamma d\theta\Big\}^{1/\gamma}\ge\frac{|\delta|-\Delta}{2\Delta}A\Big\{\int_0^{2\pi}\left|R\big(\Delta e^{i\theta}\big)\right|^\gamma d\theta\Big\}^{1/\gamma}. \tag{5.6}$$
We have,
$$W(\zeta)=R(\Delta\zeta)=\alpha_0\Delta^s\zeta^s+\alpha_1\Delta^{s+1}\zeta^{s+1}+...+(\alpha_{n-s}\Delta^n-m\lambda)\zeta^n,$$

and

$$Q(\zeta) = \zeta^n \overline{W\left(1/\bar{\zeta}\right)}. \tag{5.7}$$

Applying Lemma 4 to $Q(\zeta)$, we get

$$\left\{\int_0^{2\pi} \left|Q\left(\Delta e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma} \leq E_\gamma \left\{\int_0^{2\pi} \left|Q\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}. \tag{5.8}$$

Now, it follows readily that $\left|Q(\Delta e^{i\theta})\right| = \Delta^n \left|R\left(e^{i\theta}\right)\right|$ and $\left|Q\left(e^{i\theta}\right)\right| = \left|R\left(\Delta e^{i\theta}\right)\right|$. Using the two relations, (5.8) gives

$$\Delta^n \left\{\int_0^{2\pi} \left|R\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma} \leq E_\gamma \left\{\int_0^{2\pi} \left|R\left(\Delta e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}. \tag{5.9}$$

Since $D_\delta R(\zeta)$ is a polynomial of degree at most $(n-1)$, by Lemma 2 to $D_\delta R(\zeta)$, $R = \Delta \geq 1$, we have

$$\frac{1}{\Delta^{n-1}} \left\{\int_0^{2\pi} \left|D_\delta R\left(\Delta e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma} \leq \left\{\int_0^{2\pi} \left|D_\delta R\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma}. \tag{5.10}$$

Using (5.10) to (5.6), we get

$$\Delta^{n-1} \left\{\int_0^{2\pi} \left|D_\delta R\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma} \geq \frac{|\delta| - \Delta}{2\Delta} A \left\{\int_0^{2\pi} \left|R(\Delta e^{i\theta})\right|^\gamma d\theta\right\}^{1/\gamma}. \tag{5.11}$$

Combining (5.9) and (5.11), we have

$$\left\{\int_0^{2\pi} \left|D_\delta R\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma} \geq \frac{|\delta| - \Delta}{2E_\gamma} A \left\{\int_0^{2\pi} \left|R\left(e^{i\theta}\right)\right|^\gamma d\theta\right\}^{1/\gamma},$$

which is equivalent to

$$\left\{\int_0^{2\pi} \left|D_\delta \left\{w\left(e^{i\theta}\right) - \frac{m}{\Delta^n} \lambda e^{in\theta}\right\}\right|^\gamma d\theta\right\}^{1/\gamma} \geq \frac{|\delta| - \Delta}{2E_\gamma} A \left\{\int_0^{2\pi} \left|w\left(e^{i\theta}\right) - \frac{m}{\Delta^n} \lambda e^{in\theta}\right|^\gamma d\theta\right\}^{1/\gamma}.$$

This proves Theorem 4. □

6. Conclusion

For the set of n degree polynomials with all their zeros in $|\zeta| \leq \Delta$, $\Delta \geq 1$, there has been no integral analogue of Turán-type inequalities for about 19 years until 2017 that Rather and Bhat [17] had extended inequality (1.9) to integral mean setting. In this paper, we provide an integral mean version of Theorem 2 by using some techniques different from those followed by Rather and Bhat [17]. Our result also implicates various existing known results in the literature.

Acknowledgements

The authors are very thankful to the referees for their valuable suggestions.

REFERENCES

1. Aziz A. Integral mean estimates for polynomials with restricted zeros. *J. Approx. Theory*, 1988. Vol. 55, No. 2. P. 232–239. DOI: 10.1016/0021-9045(88)90089-5
2. Aziz A., Dawood Q. M. Inequalities for a polynomial and its derivatives. *J. Approx. Theory*, 1988. Vol. 54, No. 3. P. 306–313. DOI: 10.1016/0021-9045(88)90006-8
3. Aziz A., Rather N. A. A refinement of a theorem of Paul Turán concerning polynomials. *Math. Inequal. Appl.*, 1998. Vol. 1, No. 2. P. 231–238. DOI: 10.7153/mia-01-21
4. Aziz A., Shah W. M. Inequalities for a polynomial and its derivative. *Math. Inequal. Appl.*, 2004. Vol. 7. No. 3. P. 379–391. DOI: 10.7153/mia-07-39
5. Bernstein S. *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réele*. Paris: Gauthier-Villars, 1926. 207 p. (in French)
6. Boas Jr R. P., Rahman Q. I. L^p inequalities for polynomials and entire functions. *Arch. Rational Mech. Anal.*, 1962. Vol. 11. P. 34–39. DOI: 10.1007/BF00253927
7. Dewan K. K., Singh N., Mir A., Bhat A. Some inequalities for the polar derivative of a polynomial. *Southeast Asian Bull. Math.*, 2010. Vol. 34. P. 69–77.
8. Gardner R. B., Govil N. K., Musukula S. R. Rate of growth of polynomials not vanishing inside a circle. *J. Inequal. Pure Appl. Math.*, 2005. Vol. 6. No. 2. Art. no. 53.
9. Govil N. K. On the derivative of a polynomials. *Proc. Amer. Math. Soc.*, 1973. Vol. 41, No. 2. P. 543–546. DOI: 10.2307/2039130
10. Govil N. K. Some inequalities for the derivative of a polynomial. *J. Approx. Theory*, 1991. Vol. 66, No. 1. P. 29–35. DOI: 10.1016/0021-9045(91)90052-C
11. Hardy G. H. The mean value of the modulus of an analytic function. *Proc. Lond. Math. Soc.*, 1915. Vol. s2_14, No. 1. P. 269–277. DOI: 10.1112/plms/s2_14.1.269
12. Malik M. A. An integral mean estimates for polynomials. *Proc. Amer. Math. Soc.*, 1984. Vol. 91, No. 2. P. 281–284. DOI: 10.2307/2044642
13. Malik M. A. On the derivative of a polynomial. *J. Lond. Math. Soc.*, 1969. Vol. s2-1, No. 1. P. 57–60. DOI: 10.1112/jlms/s2-1.1.57
14. Mendeleev D. *Issledovanie vodnykh rastvorov po udel'nomu vesu.* [Investigations of Aqueous Solutions Based on Specific Gravity]. St. Petersburg: Tip. V. Demakova, 1887. 521 p. (in Russain)
15. Milovanovic G. V., Mir A., Malik A. Estimates for the polar derivative of a constrained polynomial on a disk. *Cubo*, 2022. Vol. 24. No. 3. P. 541–554. DOI: 10.56754/0719-0646.2403.0541
16. Rahman Q. I., Schmeisser G. L^p inequalities for polynomials. *J. Approx. Theory*, 1988. Vol. 53, No. 1. P. 26–32. DOI: 10.1016/0021-9045(88)90073-1
17. Rather N. A., Bhat F. A. Inequalities for the polar derivative of a polynomial. *Appl. Math. E-Notes*, 2017. Vol. 17. P. 231–241.
18. Rudin W. *Real and Complex Analysis*. McGraw-Hill Publ. Comp., 1977. 416 p.
19. Singh T. B., Chanam B. Generalizations and sharpenings of certain Bernstein and Turán types of inequalities for the polar derivative of polynomial. *J. Math. Inequal.*, 2021. Vol. 15, No. 4. P. 1663–1675. DOI: 10.7153/jmi-2021-15-114
20. Taylor A. E. *Introduction to Functional Analysis*. New York: John Wiley & Sons Inc., 1958.
21. Turan P. Über die ableitung von polynomen. *Compos. Math.*, 1939. Vol. 7. P. 89–95. (in German)

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 144–156
DOI: 10.15826/umj.2024.2.013

ON THE MODULAR SEQUENCE SPACES GENERATED BY THE CESÀRO MEAN

Sukhdev Singh[1,3,†] **and Toseef Ahmed Malik**[2,3,††]

[1]Agam Tutorials, Adampur Doaba-144102, Jalandhar, Punjab, India

[2]Department of Mathematics, Govt. Boys Higher Secondary School, Darhal-185135, Jammu and Kashmir, India

[3]Department of Mathematics, School of Chemical Engineering and Physical Sciences, Lovely Professional University, Phagwara-144411, Punjab, India

[†]singh.sukhdev01@gmail.com, [††]tsfmlk5@gmail.com

Abstract: In this paper, the seminormed Cesàro difference sequence space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is defined by using the generalized Orlicz function. Some algebraic and topological properties of the space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ are investigated. Various inclusion relations for this sequence space are also studied.

Keywords: Difference sequences, Orlicz function, Modular sequence, AK-space and BK-space.

1. Preliminaries and introduction

The notation $\omega(\mathcal{X})$ represents the spaces of all $\mathcal{X}$-valued sequence spaces, and $(\mathcal{X}, g)$ is a seminormed space. By ℓ_∞, c, and c_0, we indicate the spaces of all bounded, convergent, and null convergent sequences, respectively. Also, we denote the set of natural numbers including zero by $\mathbb{N}$ and the zero sequence by θ.

In [9], Kızmaz introduced the notion of difference sequence spaces $\lambda(\Delta)$, where λ denotes any one of the classical sequence spaces ℓ_∞, c, and c_0. Çolak and Et [5] further generalized the notion of difference sequence space $\lambda(\Delta^m)$ for $\lambda \in \{\ell_\infty, c, c_0\}$. Following [14], for $t, s \in \mathbb{N}$ and $\lambda = \ell_\infty, c, c_0$, we have

$$\lambda(\Delta^t_{(s)}) = \big\{x \in \omega : (\Delta^t_{(s)} x_i) \in \lambda\big\},$$

where

$$\Delta^t_{(s)} x_i = \Delta^{t-1}_{(s)} x_i - \Delta^{t-1}_{(s)} x_{i+s}, \quad \Delta^0_{(s)} x_i = x_i \quad \forall i \in \mathbb{N},$$

which has the following binomial expression:

$$\Delta^t_{(s)} x_i = \sum_{k=0}^{t} (-1)^k \binom{t}{k} x_{i+sk}.$$

For $s = t = 1$, we obtain the spaces $\ell_\infty(\Delta)$, $c(\Delta)$, and $c_0(\Delta)$.

A linear metric space $\mathcal{X}$ over $\mathbb{C}$ is said to be a *paranormed space* if there is a subadditive function $q : \mathcal{X} \to \mathbb{C}$ such that $q(0) = 0$, $q(x) = q(-x)$, and scalar multiplication is continuous; that is, $|\alpha_n - \alpha| \to 0$ and $q(x_i - x) \to 0$ imply $q(\alpha_i x_i - \alpha x) \to 0$ as $i \to \infty$ $\forall \alpha \in \mathbb{C}$ and $x \in \mathcal{X}$.

A paranorm q is called *total* if $q(x) = 0$ implies $x = 0$. The pair $(\mathcal{X}, q)$ is called a total paranormed space.

A convex function $M : \mathbb{R} \to \mathbb{R}$ such that $M(0) = 0$ and $M(x) > 0$ for all $x > 0$ is called an *Orlicz function*. Let X_M be the set of all sequences (x_n) such that $\sum_n M(|x_n|/p) < \infty$ for some $p > 0$; X_M is a Banach space with the norm

$$\|x_n\|_M = \inf\Big\{p > 0 : \sum_{n=1}^{\infty} M\Big(\frac{|x_n|}{p}\Big) \le 1\Big\},$$

and $(X_M, \|\cdot\|)$ is called an *Orlicz sequence space*. An Orlicz function $\mathcal{F} : [0,\infty) \to [0,\infty)$ is called a *modulus function* if

$$\mathcal{F}(x+y) \le \mathcal{F}(x) + \mathcal{F}(y) \quad \forall x, y \in [0,\infty).$$

An Orlicz function $\mathcal{F}$ is said to satisfy Δ_2-condition for all values of $u \ge 0$ if there exists $K > 0$ such that

$$\mathcal{F}(2u) \le K\mathcal{F}(u).$$

This is equivalent to satisfying the inequality

$$\mathcal{F}(ru) \le Kr\mathcal{F}(u)$$

for $r > 1$ and all values of $u \ge 0$. The Δ_2-condition implies

$$\mathcal{F}(ru) \le Kr^{\log_2 K}\mathcal{F}(u)$$

for all values of $u \ge 0$ and for $r > 1$.

Two Orlicz functions M and N are said to be *equivalent* if there exist $\alpha, \beta > 0$, $0 < K \le L$, and $a > 0$ such that $KM(\alpha x) \le N(x) \le LM(\beta x)$ for each $x \in [0, a]$. A BK-space is a Banach space of complex sequences with continuous coordinate maps. A sequence $x = (x_i) \in \nu$ is called sectionally convergent if

$$x^{[n]} = \sum_{i=1}^{n} x_i e_i \to x$$

as $n \to \infty$, where $e_i = (\delta_{ij})$ is the Kronecker symbol, that is, $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \ne j$. A space ν is called an AK-space if and only if each element is sectionally convergent.

Orlicz [13] studied the Orlicz functions and introduced the sequence space $\ell_{\mathcal{F}}$. Orlicz spaces have many applications in various fields including the theory of nonlinear integral equations. Also, Orlicz sequence spaces generalize ℓ_p-spaces, and ℓ_p-spaces are enveloped in Orlicz spaces. Many researchers have studied different sequence spaces using the Orlicz functions. For a more detailed study of the Orlicz functions, one can refer to [**?**, **?**, **?**, 2–4, 6, 8, 15, 17].

For a sequence $(\mathcal{F}_i)$ of Orlicz functions, the vector space $\ell(\mathcal{F}_i)$ defined by

$$\ell(\mathcal{F}_i) = \Big\{x = (x_i) \in w : \sum_{i=1}^{\infty} \mathcal{F}_i\Big(\frac{|x_i|}{\rho}\Big) < \infty \text{ for some } \rho > 0\Big\}$$

is a Banach space with the norm

$$\|x\| = \inf\Big\{\rho > 0 : \sum_{i=1}^{\infty} \mathcal{F}_i\Big(\frac{|x_i|}{\rho}\Big) \le 1\Big\}$$

and is called a *modular sequence space*. Furthermore, the space $\ell(\mathcal{F}_i)$ generalizes the notion of modular sequence space studied by Nakano [12] who introduced the space $\ell(\mathcal{F}_i)$ for $\mathcal{F}_i(x) = x^{\alpha_i}$, where $1 \le \alpha_i < \infty$ for $i \in \{1, 2, \ldots\}$. In [10], Lindenstrauss and Tzafriri proved that every Orlicz sequence space contains a subspace isomorphic to ℓ_p for some $1 \le p < \infty$. They also proved that every subspace of a separable Orlicz sequence space is isomorphic to ℓ_p for some $1 \le p < \infty$. Woo [?] extended these findings to the separable modular sequence spaces.

In this paper, we define and study the seminormed Cesàro difference sequence space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ using the concept of the generalized Orlicz function. Throughout the paper, we use a well-known inequality which is explained as follows [1]:
let (q_j) be a sequence of positive real numbers with

$$0 \le q_j \le \sup_j q_j = H, \quad K = \max(1, 2^{H-1}),$$

then

$$|a_i + b_i|^{q_j} \le K\,|a_i|^{q_j} + K\,|b_i|^{q_j}$$

for any two complex numbers a_i and b_i, for each $i \in \mathbb{N}$.

2. Seminormed difference sequence space and Orlicz functions

Let $\mathcal{F} = (\mathcal{F}_j)$ be a sequence of Orlicz functions, let $(\mathcal{X}, g)$ be a seminormed space, and let (q_j) be a strictly bounded sequence of positive real numbers. Let $\mathcal{C}$ be the Cesàro matrix of order 1. Then, for a nonnegative real number r and a sequence of positive real numbers $\mu = (\mu_i)$, we define a difference sequence space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ as follows:

$$\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) = \left\{ x \in \omega(\mathcal{X}) : \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) \right]^{q_j} < \infty,\ \rho > 0 \right\}.$$

Theorem 1. *The sequence space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is a linear space over the field of complex numbers $\mathbb{C}$.*

P r o o f. Let $x = (x_i)$ and $y = (y_i)$ belong to $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. Let a and b be two nonzero complex numbers. To establish the result, we need to find some $\rho_3 > 0$ such that

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} (a x_i + b y_i)}{\rho_3(j+1)} \right) \right) \right]^{q_j} < \infty.$$

For $x, y \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$, there exist $\rho_1, \rho_2 > 0$ such that

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho_1(j+1)} \right) \right) \right]^{q_j} < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} \frac{1}{j^r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} y_i}{\rho_2(j+1)} \right) \right) \right]^{q_j} < \infty.$$

Consider

$$\frac{1}{\rho_3} = \min \left\{ \frac{1}{|a|\,\rho_1},\ \frac{1}{|b|\,\rho_2} \right\}.$$

Since $\mathcal{F} = (\mathcal{F}_j)$ is nondecreasing, g is a seminorm, and $\Delta^t_{(s)}$ is linear, we obtain

$$\begin{aligned} &\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j \circ g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} (a x_i + b y_i)}{\rho_3(j+1)} \right) \right]^{q_j} \\ &\le \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} a x_i}{\rho_3(j+1)} \right) + g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} b y_i}{\rho_3(j+1)} \right) \right) \right]^{q_j} \end{aligned}$$

$$\leq \sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho_1(j+1)}\right)+g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t y_i}{\rho_2(j+1)}\right)\right)\right]^{q_j}$$

$$\leq K\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho_1(j+1)}\right)\right)\right]^{q_j}+K\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t y_i}{\rho_2(j+1)}\right)\right)\right]^{q_j}<\infty.$$

Hence, $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta_{(s)}^t, \mathcal{C})$ is a linear space. □

Theorem 2. *The space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta_{(s)}^t, \mathcal{C})$ is a paranormed space (not necessarily total paranormed) with the paranorm $\mathfrak{H}$ given by*

$$\mathfrak{H}_\Delta(x)=\inf\left\{\rho^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho(j+1)}\right)\right)\right]\leq 1,\ \rho>0,\ t\in\mathbb{N}\right\},$$

where $G=\max\left\{1, H=\sup_{j\in\mathbb{N}} q_j\right\}$.

P r o o f. Trivially, $\mathfrak{H}_\Delta(x)=\mathfrak{H}_\Delta(-x)$. Since $\mathcal{F}_j(\theta)=0$ for all $j\in\mathbb{N}$, we obtain $\inf\{\rho^{q_n/G}\}=0$ for $x=\theta$.

Let $x=(x_i)$ and $y=(y_i)$ be two arbitrary sequences in $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta_{(s)}^t, \mathcal{C})$. Then, for some $\rho_1,\rho_2>0$, we have

$$\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho_1(j+1)}\right)\right)\right]\leq 1 \quad\text{and}\quad \sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t y_i}{\rho_2(j+1)}\right)\right)\right]\leq 1.$$

For $\rho=\rho_1+\rho_2$, we obtain

$$\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t (x_i+y_i)}{\rho(j+1)}\right)\right)\right]$$

$$\leq\left(\frac{\rho_1}{\rho_1+\rho_2}\right)\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho_1(j+1)}\right)\right)\right]+\left(\frac{\rho_2}{\rho_1+\rho_2}\right)\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t y_i}{\rho_2(j+1)}\right)\right)\right]<1.$$

Thus,

$$\mathfrak{H}_\Delta(x+y)=\inf\left\{(\rho_1+\rho_2)^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t (x_i+y_i)}{(\rho_1+\rho_2)(j+1)}\right)\right)\right]\leq 1,\ \rho_1>0,\ \rho_2>0\right\}$$

$$\leq\inf\left\{(\rho_1)^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{\rho_1(j+1)}\right)\right)\right]\leq 1,\rho_1>0,t\in\mathbb{N}\right\}$$

$$+\inf\left\{(\rho_2)^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t y_i}{\rho_2(j+1)}\right)\right)\right]\leq 1,\rho_2>0,t\in\mathbb{N}\right\}\leq\mathfrak{H}_\Delta(x)+\mathfrak{H}_\Delta(y).$$

Finally, for any scalar $\gamma\neq 0$ and $r=\rho/|\gamma|$, we have

$$\mathfrak{H}_\Delta(\gamma x)=\inf\left\{\rho^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t (\gamma x_i)}{\rho(j+1)}\right)\right)\right]\leq 1,\ \rho>0,\ t\in\mathbb{N}\right\}$$

$$=\inf\left\{(|\gamma|r)^{q_t/G}:\sum_{j=0}^{\infty} j^{-r}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta_{(s)}^t x_i}{r(j+1)}\right)\right)\right]\leq 1,\ r>0,\ t\in\mathbb{N}\right\}.$$

Hence, $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is a paranormed sequence space. $\square$

Theorem 3. *Let $\mathcal{F} = (\mathcal{F}_j)$ and $\mathcal{T} = (\mathcal{T}_j)$ be two sequences of Orlicz functions. Then*

$$\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \cap \ell(\mathcal{T}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j + \mathcal{T}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}).$$

P r o o f. Let

$$x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \cap \ell(\mathcal{T}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}).$$

Then there exist $\rho_1, \rho_2 > 0$ such that

$$\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_1(j+1)}\Big)\Big)\Big]^{q_j} < \infty \quad \text{and} \quad \sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{T}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_2(j+1)}\Big)\Big)\Big]^{q_j} < \infty.$$

Taking $1/\rho = \min\{1/\rho_1, 1/\rho_2\}$, we obtain

$$\sum_{j=0}^{\infty} j^{-r}\Big[(\mathcal{F}_j + \mathcal{T}_j)\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j}$$

$$\le K\Big[\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_1(j+1)}\Big)\Big)\Big]^{q_j}\Big] + K\Big[\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{T}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_2(j+1)}\Big)\Big)\Big]^{q_j}\Big] < \infty.$$

Therefore,

$$\sum_{j=0}^{\infty} j^{-r}\Big[(\mathcal{F}_j + \mathcal{T}_j)\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j} < \infty.$$

Hence, $x \in \ell(\mathcal{F}_j + \mathcal{T}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. $\square$

Theorem 4. *For $t \ge 1$, the inclusion $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^{t-1}_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is strict.*

P r o o f. Let $x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^{t-1}_{(s)}, \mathcal{C})$. Then, we have

$$\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^{t-1}_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j} < \infty \quad \text{for some} \quad \rho > 0.$$

Since $\mathcal{F} = (\mathcal{F}_j)$ is nondecreasing and g is a seminorm, we obtain

$$\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j} \le \sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i(\Delta^{t-1}_{(s)}x_i - \Delta^{t-1}_{(s)}x_{i+1})}{\rho(j+1)}\Big)\Big)\Big]^{q_j}$$

$$\le K\Big[\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^{t-1}_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j}\Big] + K\Big[\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^{t-1}_{(s)}x_{i+1}}{\rho(j+1)}\Big)\Big)\Big]^{q_j}\Big] < \infty.$$

Therefore,

$$\sum_{j=0}^{\infty} j^{-r}\Big[\mathcal{F}_j\Big(g\Big(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\Big)\Big)\Big]^{q_j} < \infty.$$

Hence, $x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. □

In general, $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^i_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ for $i = 1, 2, \ldots, t-1$, and the inclusion is strict.

Theorem 5. *Let (q_j) be a sequence of positive real numbers. Then*

(a) $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ *for* $0 < \inf_j q_j \le q_j \le 1$.

(b) $\ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ *for* $1 \le q_j \le \sup_j q_j < \infty$.

P r o o f. (a) Let $x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. Then

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right]^{q_j} < \infty \quad \text{for some} \quad \rho > 0.$$

Since $0 < \inf_j q_j \le q_j \le 1$, we have

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right] \le \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right]^{q_j} < \infty.$$

This implies that $x \in \ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. Hence, $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$.

(b) Let $q_j \ge 1$ for all j, $\sup_j q_j < \infty$, and $x \in \ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$.
Then

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right] < \infty \quad \text{for some} \quad \rho > 0. \tag{2.1}$$

Since $1 \le q_j \le \sup_j q_j < \infty$, we have

$$\sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right]^{q_j} \le \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right] < \infty.$$

Thus, $x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. Hence, $\ell(\mathcal{F}_j, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. □

Theorem 6. *Let $(\mathcal{F}_j)$ and $(\mathcal{T}_j)$ be two sequences of Orlicz functions satisfying the Δ_2-condition, and let $r > 1$. Then $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subset \ell(\mathcal{T}_j \circ \mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$.*

P r o o f. Let $x \in \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ and $\varepsilon > 0$. Choose $0 < \delta < 1$ such that $\mathcal{F}_j(v) < \varepsilon$ for $0 \le v \le \delta$. Write

$$y_j = \mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right) \quad \text{for each} \quad j \in \mathbb{N}.$$

Consider the equality

$$\sup_j \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j} = \sup_j \sum_{1} j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j} + \sup_j \sum_{2} j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j},$$

where $y_j \leq \delta$ for the first summation and $y_j > \delta$ for the second summation. Thus, for $r > 1$, we have

$$\sup_j \sum_1 j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j} < \max(1, \varepsilon^H) \sum j^{-r} < \infty.$$

For $y_j > \delta$, we get $y_j < y_j/\delta \leq 1 + y_j/\delta$.

Since each $\mathcal{T}_j$ is nondecreasing, convex, and satisfies the Δ_2-condition, it follows that

$$\begin{aligned}\mathcal{T}_j(y_j) < \mathcal{T}\left(1 + \frac{y_j}{\delta}\right) < \frac{1}{2}\mathcal{T}_j(2) + \frac{1}{2}\mathcal{T}_j\left(\frac{2y_j}{\delta}\right)\\ < \frac{1}{2}K\frac{y_j}{\delta}\mathcal{T}_j(2) + \frac{1}{2}K\frac{y_j}{\delta}\mathcal{T}_j(2) < Ky_j\delta^{-1}\mathcal{T}_j(2) \quad \text{for each} \quad j \in \mathbb{N}.\end{aligned}$$

Therefore,

$$\sup_j \sum_2 j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j} < \max(1, (K\delta^{-1}\mathcal{F}(2))^H) \sum_2 j^{-r} (y_j)^{q_j} < \infty.$$

Thus, (2.1) yields

$$\sup_j \sum_{j=0}^{\infty} j^{-r} \left[\mathcal{T}_j(y_j)\right]^{q_j} \leq \max(1, \varepsilon^j) \sum_{j=1}^{\infty} j^{-r} + \max(1, (K\delta^{-1}\mathcal{F}(2))^H) \sum_{j=2}^{\infty} j^{-r} (y_j)^{q_j} < \infty.$$

Hence, $x \in \ell(\mathcal{T}_j \circ \mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$. □

Corollary 1. *Let $(\mathcal{F}_j)$ be any sequence of Orlicz functions satisfying the Δ_2-condition, and let $r > 1$. If $\mathcal{F}_j(x) = x$ for all $x \in [0, \infty)$ and for all $\in \mathbb{N}$, then $\ell(q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) \subseteq \ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$.*

Corollary 2. *If $\mathcal{F}_j$ and $\mathcal{T}_j$ are Orlicz functions that are equivalent for each $j \in \mathbb{N}$, then $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C}) = \ell(\mathcal{T}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$.*

For $r = 0$, the space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ reduces to a sequence space as follows:

$$\ell(\mathcal{F}_j, q, g, \mu, \Delta^t_{(s)}, \mathcal{C}) = \left\{x \in \omega(\mathcal{X}) : \sum_{j=0}^{\infty} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right]^{q_j} < \infty \text{ for some } \rho > 0\right\}.$$

Theorem 7. *Let $(\mathcal{F}_j)$ be a sequence of Orlicz functions, let $q_j \in \ell_\infty$, and let $(\mathcal{X}, g)$ be a complete seminormed space. Then $\ell(\mathcal{F}_j, q, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ is a complete paranormed endowed with the paranorm $\mathfrak{H}_\Delta$ defined by*

$$\mathfrak{H}_\Delta(x) = \inf\left\{\rho^{q_t/K} : \sum_{j=0}^{\infty} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)}\right)\right)\right] \leq 1,\ \rho > 0,\ t \in \mathbb{N}\right\},$$

where $K = \max\{1, H = \sup_{j \in \mathbb{N}} q_j\}$.

P r o o f. Let (x_i) be a Cauchy sequence in $\ell(\mathcal{F}_j, q, g, \mu, \Delta^t_{(s)}, \mathcal{C})$. Let $\delta > 0$ be fixed, and let $s > 0$ be such that, for given $0 < \varepsilon < 1$, $\varepsilon/s\delta > 0$ and $s\delta \geq 1$. Then, there exists a positive integer n_0 such that

$$h(x^m - x^n) < \frac{\varepsilon}{s\delta} \quad \forall\, m, n \geq n_0.$$

Thus,

$$\inf\left\{\rho^{q_t/K} : \sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^n)}{\rho(j+1)}\right)\right)\right] \leq 1, t \in \mathbb{N}\right\} < \frac{\varepsilon}{s\delta} \quad \forall\, m, n \geq n_0.$$

It implies that

$$\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^n)}{h(x^m - x^n)(j+1)}\right)\right)\right] \leq 1 \quad \forall\, m, n \geq n_0.$$

Therefore,

$$\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^t)}{h(x^m - x^n)(j+1)}\right)\right) \leq 1 \quad \forall\, m, n \geq n_0 \quad \text{and} \quad j \in \mathbb{N}.$$

For $s > 0$ with $\mathcal{F}_j(s\delta/2) \geq 1$, we obtain

$$\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^n)}{h(x^m - x^n)(j+1)}\right)\right) \leq \mathcal{F}_j\left(\frac{s\delta}{2}\right) \quad \forall\, m, n \geq n_0 \quad \text{and} \quad j \in \mathbb{N}.$$

Since $\mathcal{F}_j$ is nondecreasing for each $j \in \mathbb{N}$, we have

$$g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^n)}{j+1}\right) \leq \frac{s\delta}{2} \times \frac{\varepsilon}{s\delta} = \frac{\varepsilon}{2}.$$

Hence, $(\Delta^t_{(s)}x_i^m)$ is a Cauchy sequence in $(\mathcal{X}, g)$ for each $i \in \mathbb{N}$. However, $(\mathcal{X}, g)$ is complete and so $(\Delta^t_{(s)}x_i^m)$ is convergent in $(\mathcal{X}, g)$ for all $i \in \mathbb{N}$.

Let $\lim\limits_{m\to\infty} \mu_i\Delta^t_{(s)}x_i^m = x_i$ exists for each $i \geq 1$. For $i = 1$, we obtain

$$\lim_{m\to\infty} \mu_1\Delta^t_{(s)}x_1^m = \lim_{m\to\infty} \mu_1 \sum_{k=0}^{t}(-1)^k\binom{t}{k}x_{1+sk} = \lim_{m\to\infty} \mu_1 x_1^m = x_1. \tag{2.2}$$

Similarly,

$$\lim_{m\to\infty} \mu_i\Delta^t_{(s)}x_i^m = \lim_{m\to\infty} \mu_i x_i^m = x_i \quad \text{for} \quad i = 1, \dots, ts. \tag{2.3}$$

From (2.2) and (2.3), it follows that $\lim\limits_{m\to\infty} \mu_i x_{1+ts}^m$ exists.

Let $\lim\limits_{m\to\infty} \mu_i x_{1+ts}^m = \mu_i\, x_{1+ts}$. Then, by induction, $\lim\limits_{m\to\infty} \mu_i x_i^m = x_i$ for all $i \in \mathbb{N}$.

Now, for each $m, n \geq n_0$, we have

$$\inf\left\{\rho^{q_t/K} : \sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x_i^m - \Delta^t_{(s)}x_i^n)}{\rho(j+1)}\right)\right)\right] \leq 1,\ t \in \mathbb{N}\right\} < \varepsilon.$$

Thus,

$$\lim_{n\to\infty}\left\{\inf\left\{\rho^{q_t/K}:\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x^m_i-\Delta^t_{(s)}x^n_i)}{\rho(j+1)}\right)\right)\right]\le 1,\ t\in\mathbb{N}\right\}\right\}<\varepsilon\quad\forall m,n\ge n_0.$$

Using the continuity of Orlicz functions, we obtain

$$\inf\left\{\rho^{q_t/K}:\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x^m_i-\Delta^t_{(s)}\lim_{n\to\infty}x^n_i)}{\rho(j+1)}\right)\right)\right]\le 1,\ t\in\mathbb{N}\right\}<\varepsilon\quad\forall m\ge n_0.$$

This implies that

$$\inf\left\{\rho^{q_t/K}:\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i(\Delta^t_{(s)}x^m_i-\Delta^t_{(s)}x_i)}{\rho(j+1)}\right)\right)\right]\le 1,\ t\in\mathbb{N}\right\}<\varepsilon\quad\forall n\ge n_0.$$

Hence, $(x^m-x)\in\ell(\mathcal{F}_j,q,g,\mu,\Delta^t_{(s)},\mathcal{C})$, and then $x=x^m-(x^m-x)\in\ell(\mathcal{F}_j,q,g,\mu,\Delta^t_{(s)},\mathcal{C})$. □

For $r=0$, $q_j=q$, a constant, the space $\ell(\mathcal{F}_j,q,g,r,\mu,\Delta^t_{(s)},\mathcal{C})$ reduces to a sequence space as follows:

$$\ell(\mathcal{F}_j,g,\mu,\Delta^t_{(s)},\mathcal{C})=\left\{x\in\omega(\mathcal{X}):\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right)\right]<\infty\text{ for some }\rho>0\right\}.$$

Theorem 8. *Let $(\mathcal{X},g)$ be a complete normed space. Then, $\ell(\mathcal{F}_j,g,\mu,\Delta^t_{(s)},\mathcal{C})$ is a Banach space with a norm $\|\cdot\|$ defined by*

$$\|x\|=\inf\left\{\rho:\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right)\right]\le 1\right\}.\tag{2.4}$$

P r o o f. To prove that $\|\cdot\|$ is a norm in $\ell(\mathcal{F}_j,g,\mu,\Delta^t_{(s)},\mathcal{C})$, we can verify the completeness of $\ell(\mathcal{F}_j,g,\mu,\Delta^t_{(s)},\mathcal{C})$ as in the proof of Theorem 7.

If $x=\theta$, then clearly $\|x\|=0$.

Conversely, suppose that $\|x\|=0$. Then,

$$\inf\left\{\rho:\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right)\right]\le 1\right\}=0.$$

Thus, for given $\varepsilon>0$, there exists ρ_ε $(0<\rho_\varepsilon<\varepsilon)$ such that

$$\sum_{j=0}^{\infty}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_\varepsilon(j+1)}\right)\right)\right]\le 1.$$

This implies that

$$\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho_\varepsilon(j+1)}\right)\right)\le 1\quad\forall j\in\mathbb{N}.$$

Therefore, we have

$$\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta_{(s)}^t x_i}{\varepsilon(j+1)}\right)\right) \le \mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta_{(s)}^t x_i}{\rho_\varepsilon(j+1)}\right)\right) \le 1 \quad \forall j \in \mathbb{N}.$$

Suppose that

$$\frac{\sum_{i=0}^{n_j} \mu_i \Delta_{(s)}^t x_i}{(n_j+1)} \ne 0$$

for some n_j. Then,

$$\frac{\sum_{i=0}^{n_j} \mu_i \Delta_{(s)}^t x_i}{\varepsilon(n_j+1)} \to \infty$$

as $\varepsilon \to 0$. This implies that

$$\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta_{(s)}^t x_i}{\epsilon(j+1)}\right)\right) \to \infty \quad \text{as} \quad \varepsilon \to 0 \quad \text{for some} \quad n_j \in \mathbb{N},$$

which leads to a contradiction. Therefore,

$$\frac{\sum_{i=0}^{j} \mu_i \Delta_{(s)}^t x_i}{(j+1)} = 0 \quad \forall j \in \mathbb{N}.$$

If $j = 0$, then $\mu_0 \Delta_{(s)}^t x_0 = 0$ and $\mu_0 x_0 = 0$; $\mu_1 x_1 = 0$ for $j = 1$.

Similarly, $x_j = 0$ for all $j \ge 1$. Hence, $x = \theta$.

Further, the properties $\|x+y\| \le \|x\| + \|y\|$ and $\|\alpha x\| = |\alpha| \, \|x\|$ for any scalar α can be proved as in the proof of Theorem 2. □

The above proof makes it easy to prove that $\|x^n\| \to 0$ implies that $x_i^n \to 0$ for each $n \ge 1$. Now, we state the following result.

Proposition 1. *The space $\ell(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C})$ is a BK-space.*

To prove the AK-property of the space $\ell(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C})$, we give the following definition and prove some related results.

Definition 1. *Let $\mathcal{F} = (\mathcal{F}_j)$ be any sequence of Orlicz functions. Define*

$$\ell'(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C}) = \left\{x \in \omega(\mathcal{X}) : \sum_{j=0}^{\infty} \left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j} \mu_i \Delta_{(s)}^t x_i}{\rho(j+1)}\right)\right)\right] < \infty \quad \textit{for every} \quad \rho > 0\right\}.$$

Evidently, $\ell'(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C})$ is a subspace of $\ell(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C})$, and its topology is inherited from $\|\cdot\|$.

Theorem 9. *Let $(\mathcal{F}_j)$ be a sequence of Orlicz functions satisfying the Δ_2-condition. Then $\ell(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C}) = \ell'(\mathcal{F}_j, g, \mu, \Delta_{(s)}^t, \mathcal{C})$.*

P r o o f. Let $x \in \ell(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$. Then, for some $\rho > 0$, we have

$$\sum_{j=0}^{\infty} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) \right] < \infty.$$

Consider any arbitrary $\eta > 0$. If $\rho \leq \eta$, then

$$\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\eta(j+1)} \right) \right) < \mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) < \infty \quad \text{for each} \quad j \in \mathbb{N}.$$

Let $\eta < \rho$. Since $\mathcal{F}_j$ satisfies the Δ_2-condition, there exists a constant $K_j > 0$ such that

$$\mathcal{F}_j \left(g \left(\frac{\sum\limits_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\eta(j+1)} \right) \right) \leq K_j \left(\frac{\rho}{\eta} \right)^{\log_2 K_j} \mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) \quad \text{for each} \quad j \in \mathbb{N}.$$

Now, we can find $R_j > 0$ such that

$$R_j = \sup_j K_j \left(\frac{\rho}{\eta} \right)^{\log_2 K_j}.$$

Then, for fixed $\eta > 0$ and for each $j \in \mathbb{N}$, we have

$$\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\eta(j+1)} \right) \right) \leq R_j \mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) < \infty.$$

It follows the result. □

Theorem 10. *Let $(\mathcal{X}, g)$ be a complete normed space. Then $\ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ is an AK-space.*

P r o o f. Let $x \in \ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$. Then, for each ε $(0 < \varepsilon < 1)$, we can find r_0 such that

$$\sum_{j \geq r_0} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\epsilon(j+1)} \right) \right) \right] \leq 1.$$

Therefore, for $r \geq r_0$, we have

$$\begin{aligned} \|x - x^{[r]}\| &= \inf \left\{ \rho : \sum_{j \geq r+1}^{\infty} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) \right] \leq 1 \right\} \\ &\leq \inf \left\{ \rho : \sum_{j \geq r} \left[\mathcal{F}_j \left(g \left(\frac{\sum_{i=0}^{j} \mu_i \Delta^t_{(s)} x_i}{\rho(j+1)} \right) \right) \right] \leq 1 \right\} < \varepsilon. \end{aligned}$$

Hence, $\ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(\mu s)}, \mathcal{C})$ is an AK-space. □

Now, using Proposition 1 and Theorem 9, we establish the following result.

Corollary 3. *Let* $(\mathcal{F}_j)$ *be a sequence of Orlicz functions satisfying the* Δ_2*-condition. Then* $\ell(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ *is an AK-space.*

Theorem 11. *The space* $\ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ *is a closed subspace of* $\ell(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$*.*

P r o o f. Let (x^r) be a sequence in $\ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ such that $\|x^r - x\| \to 0$.
It suffices to show that $x \in \ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$, i.e.,

$$\sum_{j\geq 0}\left[\mathcal{F}_j\left(g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right)\right] < \infty \quad \text{for every} \quad \rho > 0.$$

For $\rho > 0$, there exists m such that $\|x^m - x\| \leq \rho/2$. Since $\mathcal{F}_j$ is a convex function for each $j \in \mathbb{N}$, we have

$$\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right]$$

$$= \sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left(\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i^m\right| - \left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i^m\right| + \left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i\right|\right)}{\rho(j+1)}\right)\right]$$
$$\leq \frac{1}{2}\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i^m\right|}{\rho(j+1)}\right)\right] + \frac{1}{2}\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}(x_i^m - x_i)\right|}{\rho(j+1)}\right)\right]$$
$$\leq \frac{1}{2}\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i^m\right|}{\rho(j+1)}\right)\right] + \frac{1}{2}\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}(x_i^m - x_i)\right|}{\|x^m - x\|(j+1)}\right)\right].$$

From (2.4), we get

$$\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{2\left|\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}(x_i^m - x_i)\right|}{\|x^m - x\|(j+1)}\right)\right] \leq 1.$$

Thus,

$$\sum_{j\geq 0}\mathcal{F}_j\left[g\left(\frac{\sum_{i=0}^{j}\mu_i\Delta^t_{(s)}x_i}{\rho(j+1)}\right)\right] < \infty \quad \text{for every} \quad \rho > 0.$$

Hence, $x \in \ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$. □

Corollary 4. *The space* $\ell'(\mathcal{F}_j, g, \mu, \Delta^t_{(s)}, \mathcal{C})$ *is a BK-space.*

3. Conclusion

We have investigated the convergence of the difference sequence for the Cesàro mean of order 1, along with the generalized Orlicz function, using the technique of seminorm. In our study, we established that the newly defined sequence space $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is a paranormed sequence space. We examined both the algebraic and topological properties of this sequence space. Additionally, we verified that $\ell(\mathcal{F}_j, q, g, r, \mu, \Delta^t_{(s)}, \mathcal{C})$ is indeed a separable sequence space. In our upcoming research, we aim to extend this concept to the case of statistical convergence and the Cesàro mean of higher order.

Acknowledgements

The authors are very grateful to all the referees for their careful reading and insightful comments.

REFERENCES

1. Altay B., Başar F. Generalization of the sequence space $\ell(p)$ derived by weighted means. *J. of Math. Anal. Appl.*, 2007. Vol. 330, No. 1. P. 174–185. DOI: 10.1016/j.jmaa.2006.07.050
2. Bakery A. A., Kalthum Om, Mohamed S. K. Orlicz generalized difference sequence space and the linked pre-quasi operator ideal. *J. of Math.*, 2020. Art. no. 6664996. P. 1–9. DOI: 10.1155/2020/6664996
3. Dutta H., Basar F. A generalization of Orlicz sequence spaces by Cesàro mean of order one. *Acta Math. Univ. Comenian*, 2011. Vol. 80, No. 2. P. 185–200.
4. Esi A., Bipin H., Esi A. New type of lacunary Orlicz difference sequence spaces generated by infinite matrices. *Filomat*, 2016. Vol. 30, No. 12. P. 3195–3208. DOI: 10.2298/FIL1612195E
5. Et M., Çolak R. On some generalized sequence spaces. *Soochow J. Math.*, 1995. Vol. 21, No. 4. P. 377–386.
6. Et M., Lee P. Y., Tripathy B. C. Strongly almost $(V,\lambda)(\Lambda^r)$-summable sequences defined by Orlicz function. *Hokkaido Math. J.*, 2006. Vol. 35. P. 197–213.
7. Gupta M., Bhar A. Generalized Orlicz-Lorentz sequence spaces and corresponding operator ideals. *Math. Slovaca*, 2014. Vol. 64, No. 6. P. 1474–1496. DOI: 10.2478/s12175-014-0287-6
8. Jebril I. H. A generalization of strongly Cesàro and strongly lacunary summable spaces. *Acta Univ. Apalensis*, 2010. Vol. 23. P. 49–61.
9. Kizmaz H. On certain sequence spaces. *Canad. Math. Bull.*, 1981. Vol. 24, No. 2. P. 169–176. DOI: 10.4153/CMB-1981-027-5
10. Lindenstrauss J., Tzafriri L. On Orlicz sequence spaces. *Israel J. Math.*, 1971. Vol. 10. P. 379–390. DOI: 10.1007/BF02771656
11. Mursaleen M., Khan A. M., Qamaruddin. Difference sequence spaces defined by Orlicz functions. *Demonstratio Math.*, 1999. Vol. 32, No. 1. P. 145–150. DOI: 10.1515/dema-1999-0115
12. Nakano H. Modulared sequence spaces. *Proc. Japan Acad.*, 1951. Vol. 27, No. 9. P. 508–512. DOI: 10.3792/pja/1195571225
13. Orlicz W. Über Räume (L^M). *Bull. Int. Acad. Polon. Sci.*, 1936. P. 93–107. (in German)
14. Tripathy B. C., Esi A. A new type of difference sequence spaces. *Inter. J. Sci. Tech.* 2006. Vol. 1, No. 1. P. 11–14.
15. Tripathy B. C., Choudhary B., Sarma B. Some difference double sequence spaces defined by Orlicz function. *Kyungpook Math. J.*, 2008. Vol. 48, No. 4. P. 613–622. DOI: 10.5666/KMJ.2008.48.4.613
16. Tripathy B. C., Dutta H. Some difference paranormed sequence spaces defined by Orlicz functions. *Fasc. Math.*, 2009. Vol. 42. P. 121–131.
17. Tripathy B. C., Dutta H. On some lacunary difference sequence spaces defined by a sequence of Orlicz functions and q-lacunary Δ_m^n-statistical convergence. *An. Ştiint. Univ. Ovidius Constanţa. Ser. Mat.*, 2012. Vol. 20. No. 1. P. 417–430.
18. Woo J. On modular sequence spaces. *Studia Math.*, 1973. Vol. 48. No. 3. P. 271–289.

URAL MATHEMATICAL JOURNAL, Vol. 10, No. 2, 2024, pp. 157–173

DOI: 10.15826/umj.2024.2.014

PROPERTIES OF SOLUTIONS IN THE DUBINS CAR CONTROL PROBLEM[1]

Artem A. Zimovets

Krasovskii Institute of Mathematics and Mechanics,
Ural Branch of the Russian Academy of Sciences,
16 S. Kovalevskaya Str., Ekaterinburg, 620108, Russian Federation

aazimovets@gmail.com

Abstract: This paper addresses the time-optimal control problem of the Dubins car, which is closely related to the problem of constructing the shortest curve with bounded curvature between two points in a plane. This connection allows researchers to apply both geometric methods and control theory techniques during their investigations. It is established that the time-optimal control for the Dubins car is a piecewise constant function with no more than two switchings. This characteristic enables the categorization of all such controls into several types, facilitating the examination of the solutions to the control problem for each type individually. The paper derives explicit formulas for determining the switching times of the control signal. In each case, necessary and sufficient conditions for the existence of solutions are obtained. For certain control types, the uniqueness of optimal solutions is established. Additionally, the dependence of the movement time on the initial and terminal conditions is studied.

Keywords: Dubins car, Dubins problem, Time-optimal control, Curve with bounded curvature.

1. Introduction

The Dubins car is a simple mathematical model of a car-like vehicle that moves in a plane at a constant speed and is capable of making left and right turns with a bounded turning radius. The time-optimal control problem of the Dubins car is closely related to the problem of constructing the shortest curve with bounded curvature between two points in a plane. One of the first studies on this subject was by Markov [11], in which he considered the shortest curves with a prescribed tangent at one of the endpoints. The problem of constructing the shortest curve with a constraint on average curvature and with prescribed tangents at both endpoints was later investigated by Dubins [7]. This problem later became known as the Markov–Dubins problem or simply as the Dubins problem, and the solution to this problem was referred to as the Dubins path. Moreover, as demonstrated in [8], not only geometric methods but also control theory methods can be used to study plane curves. Significant results in this direction were obtained in [5, 9, 19].

The Markov–Dubins problem and its variations have been extensively studied over the past several decades. We mention in particular the construction of the shortest bounded-curvature paths in 3-dimensional space [18], the investigation of homotopy classes of bounded-curvature paths [1], and the description of the reachable sets for the Dubins car [13, 14]. Reeds and Shepp [16] notably extended Dubins' original work by considering a vehicle capable of both forward and reverse motion, resulting in the formulation of the Reeds–Shepp car model. Numerous other extended models can also be found in [3, 4, 12]. The practical applications of the Markov–Dubins problem are widespread, impacting fields such as railroad construction [11], air traffic control [15], robotics [2], and many other domains.

[1]This research was supported by the Russian Science Foundation (grant no. 24-21-00424, https://rscf.ru/en/project/24-21-00424/).

It was shown in [7] that the shortest path with bounded curvature between two points in a plane consists of no more than three segments, each of which is either an arc of a circle or a straight line segment. The same result was obtained for the trajectories of the Dubins car [5, 19]. As a consequence, the Markov–Dubins problem can be reduced to finding the shortest path among several candidate paths. In [15], the parameters of the candidate paths were found for a fixed terminal position. The case of a moving target was investigated in [6]. Paper [17] considers the case where the starting and ending points are far apart, and provides a decision table for finding the shortest path. In [10], the endpoints of the curve segments were found by a geometric approach.

In this paper, we investigate the properties of the candidate paths and the corresponding controls in the time-optimal control problem of the Dubins car. The paper is organized as follows. Section 2 outlines the time-optimal control problem for the Dubins car and categorizes the control set into three distinct types. In Section 3, we derive formulas for calculating the switching times associated with each type of control. Section 4 identifies key properties for each control type. Finally, Section 5 illustrates how these properties can be used to solve the time-optimal control problem for the Dubins car.

2. Problem statement

Consider a vehicle that moves in a horizontal plane at a constant speed, capable of making left and right turns. The motion of the vehicle is governed by the system of ordinary differential equations

$$\begin{cases} \dot{x} = v\cos\varphi, \\ \dot{y} = v\sin\varphi, \\ \dot{\varphi} = u, \end{cases} \tag{2.1}$$

where x and y are the Cartesian coordinates of the vehicle in the xy-plane, φ is the orientation of the velocity vector, v is the speed, and u is the control variable. It is assumed that the angle φ is measured counterclockwise from the positive x axis and can take any real values. An admissible control is a Lebesgue measurable function $u(t)$ that satisfies the constraint $|u(t)| \leqslant u_m$, $u_m > 0$, on any finite time interval. The mathematical model described by (2.1) is called the "Dubins car".

In this model, any two orientation angles φ_* and φ^* such that $\varphi_* - \varphi^* = 2\pi k$, $k \in \mathbb{Z}$, are considered equivalent. $\mathbb{Z}$ denotes the set of integers. We should note that the coefficient k will be used in this paper to define various sets. In this regard, all these coefficients should be treated independently of each other.

The time-optimal control problem of the Dubins car can be described as follows. Suppose we are given a vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$, where (x_0, y_0) and (x_f, y_f) are the initial and terminal positions of the vehicle, and φ_0 and φ_f are the initial and terminal orientations, respectively. It is required to find an admissible control that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ in the minimum possible time. Since system (2.1) is time-invariant, the initial time t_0 can be chosen arbitrarily.

In [5, 13, 19], it is shown that the time-optimal control for the Dubins car is a piecewise constant function having no more than three segments with lengths Δt_1, Δt_2, and Δt_3 and values u_1, u_2, and u_3, respectively, where $u_1 \in \{-u_m, u_m\}$, $u_2 \in \{-u_m, 0, u_m\}$, and $u_3 \in \{-u_m, u_m\}$. Let $u_* \in \{-u_m, u_m\}$. Then all such controls can be divided into the following types:

1. Control of the type $(u_*, -u_*, u_*)$, where $u_1 = u_*$, $u_2 = -u_*$, and $u_3 = u_*$.
2. Control of the type $(u_*, 0, u_*)$, where $u_1 = u_*$, $u_2 = 0$, and $u_3 = u_*$.
3. Control of the type $(u_*, 0, -u_*)$, where $u_1 = u_*$, $u_2 = 0$, and $u_3 = -u_*$.

Thus, the time-optimal control problem of the Dubins car can be solved by testing the optimal controls of several specific types. Accordingly, the purpose of this paper is to study the properties of each of these types of control. Specifically, we aim to investigate the necessary and sufficient conditions for the existence of solutions, the uniqueness of optimal solutions, and the dependence of the movement time on the initial and terminal conditions.

3. Switching times

This section provides a solution to the time-optimal control problem of the Dubins car for controls of the types $(u_*, -u_*, u_*)$, $(u_*, 0, u_*)$, and $(u_*, 0, -u_*)$. In each case, we derive explicit formulas for determining the optimal time intervals Δt_1, Δt_2, and Δt_3. The lengths of the intervals are assumed to be nonnegative. This condition, called a nonnegativity condition, can be written as

$$\begin{cases} \Delta t_1 \geqslant 0, \\ \Delta t_2 \geqslant 0, \\ \Delta t_3 \geqslant 0. \end{cases} \tag{3.1}$$

Note that the intervals are allowed to be degenerate. Knowing Δt_1, Δt_2, and Δt_3, we can find the switching times t_1 and t_2 and the terminal time t_f by the simple relations

$$t_1 = t_0 + \Delta t_1, \quad t_2 = t_1 + \Delta t_2, \quad t_f = t_2 + \Delta t_3.$$

Before proceeding to specific types of control, we introduce some notation and definitions.

Definition 1. *Denote by* $\operatorname{sgn}(x)$ *the function of a real variable* x *defined by*

$$\operatorname{sgn}(x) = \begin{cases} 1, & x \geqslant 0, \\ -1, & x < 0. \end{cases}$$

If $x \neq 0$, the function $\operatorname{sgn}(x)$ can be written as $\operatorname{sgn}(x) = |x|/x$.

Definition 2. *The modulo operation* $a \bmod b$ *is the binary operation that associates with each pair of real numbers* a *and* $b \neq 0$ *the nonnegative remainder after dividing* a *by* b, *that is, a number* $r \in [0, |b|)$ *such that* $a = qb + r$, *where* $q \in \mathbb{Z}$.

Definition 3. *By a multivalued function* $F\colon X \to \mathcal{P}(Y)$ *we mean a function that maps elements of* X *to subsets of* Y.

We extend standard binary operations that take two single-valued arguments to binary operations that take one single-valued argument and one multivalued argument as follows.

Definition 4. *Let* $*\colon X \times X \to X$ *be a binary operation. For each* $x \in X$ *and* $\sigma \subset X$, *define* $x * \sigma = \{x * y \mid y \in \sigma\}$ *and* $\sigma * x = \{y * x \mid y \in \sigma\}$.

Next, we proceed to prove a preliminary lemma.

Lemma 1. *Let* F *be a multivalued real function of the form* $F(x) = f(x) + G$, *where* f *is a continuous single-valued function,* $G = \{ka \mid k \in \mathbb{Z}\}$, *and* a *is a positive constant. Let a multivalued function* H *be defined as* $H(x) = F(x) \bmod a$, *and let a single-valued function* h *be defined as*

$$h(x) = \min\{y \in F(x) \mid y \geqslant 0\}.$$

Then $H(x) = \{h(x)\}$. *Moreover, if* $f(x_*) \neq ma$, $m \in \mathbb{Z}$, *then* h *is continuous at* x_*.

P r o o f. Let x^* be an arbitrary point in $\operatorname{dom} H$. We first show that $H(x^*)$ cannot contain two different elements. Suppose there are $h_1 \in H(x^*)$ and $h_2 \in H(x^*)$ such that $h_1 \neq h_2$. Then h_1 and h_2 must satisfy the system

$$\begin{cases} f(x^*) + k_1 a = q_1 a + h_1, & h_1 \in [0, a), \quad k_1 \in \mathbb{Z}, \quad q_1 \in \mathbb{Z}, \\ f(x^*) + k_2 a = q_2 a + h_2, & h_2 \in [0, a), \quad k_2 \in \mathbb{Z}, \quad q_2 \in \mathbb{Z}. \end{cases} \tag{3.2}$$

Subtracting the second equation from the first and collecting the terms that involve a factor of a on the left-hand side, we obtain

$$(k_1 - k_2 - q_1 + q_2)a = h_1 - h_2. \tag{3.3}$$

Since $h_1 \in [0, a)$ and $h_2 \in [0, a)$, we have $-a < h_1 - h_2 < a$. Hence, equality (3.3) holds only when $k_1 - k_2 - q_1 + q_2 = 0$ and $h_1 = h_2$. This is a contradiction.

From (3.2), it follows that $h_1 \in F(x^*)$. We claim that $h(x^*) = h_1$. Suppose there is $h_3 \in F(x^*)$ such that $0 \leqslant h_3 < h_1$. By definition, $h_3 = f(x^*) + k_3 a$, where $k_3 \in \mathbb{Z}$. Then

$$h_1 - h_3 = (k_1 - q_1 - k_3)a.$$

However, $0 < h_1 - h_3 < a$. This results in a contradiction. Hence, $h(x^*) = h_1$.

Next, we show that the function h is continuous at x_* if $f(x_*) \neq ma$, $m \in \mathbb{Z}$. Pick any $\varepsilon > 0$. Since f is continuous, it follows that

$$\exists \delta > 0\colon \|x - x_*\| < \delta \Rightarrow |f(x) - f(x_*)| < \varepsilon.$$

Let

$$\varepsilon_m = \min_{m \in \mathbb{Z}} |f(x_*) - ma|.$$

Then we get

$$\exists \delta_m > 0\colon \|x - x_*\| < \delta_m \Rightarrow |f(x) - f(x_*)| < \varepsilon_m.$$

Let $\delta_* = \min\{\delta, \delta_m\}$. Then we find that

$$\|x - x_*\| < \delta_* \Rightarrow |f(x) - f(x_*)| < \varepsilon.$$

Furthermore, if $\|x - x_*\| < \delta_*$, then there exists $p \in \mathbb{Z}$ such that

$$f(x) = pa + h(x), \quad f(x_*) = pa + h(x_*).$$

Finally, we obtain

$$\|x - x_*\| < \delta_* \Rightarrow |h(x) - h(x_*)| = |f(x) - f(x_*)| < \varepsilon.$$

Since ε was chosen arbitrarily, we conclude that the function h is continuous at x_*. □

3.1. Controls of the type $(u_*, -u_*, u_*)$

Given a vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$, the goal is to find a control of the type $(u_*, -u_*, u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ in minimum time.

For this type of control, the function $\varphi(t)$ on the time interval $[t_0, t_f]$ can be expressed as

$$\varphi(t) = \begin{cases} \varphi_0 + u_*(t - t_0), & t \in [t_0, t_1), \\ \varphi_0 + u_* \Delta t_1 - u_*(t - t_1), & t \in [t_1, t_2), \\ \varphi_0 + u_* \Delta t_1 - u_* \Delta t_2 + u_*(t - t_2), & t \in [t_2, t_f]. \end{cases} \tag{3.4}$$

Substituting (3.4) into (2.1) gives

$$
\begin{aligned}
x(t_f) &= x_0 + \int_0^{\Delta t_1} v\cos(\varphi_0 + u_*\tau)\,d\tau + \int_0^{\Delta t_2} v\cos(\varphi_0 + u_*\Delta t_1 - u_*\tau)\,d\tau \\
&\quad + \int_0^{\Delta t_3} v\cos(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2 + u_*\tau)\,d\tau \\
&= x_0 + \frac{v}{u_*}\big(2\sin(\varphi_0 + u_*\Delta t_1) - 2\sin(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2) + \sin(\varphi(t_f)) - \sin(\varphi_0)\big),
\end{aligned}
\tag{3.5}
$$

$$
\begin{aligned}
y(t_f) &= y_0 + \int_0^{\Delta t_1} v\sin(\varphi_0 + u_*\tau)\,d\tau + \int_0^{\Delta t_2} v\sin(\varphi_0 + u_*\Delta t_1 - u_*\tau)\,d\tau \\
&\quad + \int_0^{\Delta t_3} v\sin(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2 + u_*\tau)\,d\tau \\
&= y_0 - \frac{v}{u_*}\big(2\cos(\varphi_0 + u_*\Delta t_1) - 2\cos(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2) + \cos(\varphi(t_f)) - \cos(\varphi_0)\big).
\end{aligned}
\tag{3.6}
$$

Combining (3.4), (3.5), and (3.6) with the terminal condition, we obtain the system

$$
\begin{cases}
x_0 + \dfrac{v}{u_*}\big(2\sin(\varphi_0 + u_*\Delta t_1) - 2\sin(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2) + \sin(\varphi_f) - \sin(\varphi_0)\big) = x_f, \\
y_0 - \dfrac{v}{u_*}\big(2\cos(\varphi_0 + u_*\Delta t_1) - 2\cos(\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2) + \cos(\varphi_f) - \cos(\varphi_0)\big) = y_f, \\
\varphi_0 + u_*\Delta t_1 - u_*\Delta t_2 + u_*\Delta t_3 = \varphi_f + 2\pi k, \quad k \in \mathbb{Z},
\end{cases}
\tag{3.7}
$$

where Δt_1, Δt_2, and Δt_3 are unknowns.

Thus, the problem can be formulated as follows: find a solution to system (3.7) that satisfies the nonnegativity condition (3.1) and minimizes the performance index

$$T_1 = \Delta t_1 + \Delta t_2 + \Delta t_3.$$

Introduce the notation

$$\alpha = \varphi_0 + u_*\Delta t_1, \quad \beta = u_*\Delta t_2, \quad \gamma = u_*\Delta t_3, \tag{3.8}$$

$$a_1 = \frac{u_*}{v}(x_f - x_0) - \sin(\varphi_f) + \sin(\varphi_0), \tag{3.9}$$

$$b_1 = \frac{u_*}{v}(y_f - y_0) + \cos(\varphi_f) - \cos(\varphi_0). \tag{3.10}$$

With this notation, system (3.7) may be written as

$$
\begin{cases}
2\sin(\alpha) - 2\sin(\alpha - \beta) = a_1, \\
-2\cos(\alpha) + 2\cos(\alpha - \beta) = b_1, \\
\alpha - \beta + \gamma = \varphi_f + 2\pi k, \quad k \in \mathbb{Z}.
\end{cases}
\tag{3.11}
$$

1. Suppose that $a_1^2 + b_1^2 = 0$. Then the set of solutions to (3.7) can be expressed as

$$
\begin{cases}
\Delta t_1 + \Delta t_3 = \dfrac{1}{u_*}(\varphi_f - \varphi_0 + 2\pi k), \quad k \in \mathbb{Z}, \\
\Delta t_2 = \dfrac{1}{u_*}2\pi k, \quad k \in \mathbb{Z}.
\end{cases}
\tag{3.12}
$$

Applying the modulo operation to (3.12) and resolving the ambiguity, we obtain

$$
\begin{cases}
\Delta t_1 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)(\varphi_f - \varphi_0)) \mod 2\pi\big), \\
\Delta t_2 = 0, \\
\Delta t_3 = 0.
\end{cases}
\tag{3.13}
$$

Thus, solution (3.13) defines the constant control $u(t) \equiv u_*$ over the entire time interval $[t_0, t_f]$. The trajectory of the vehicle in this case is just an arc of a circle.

2. Suppose now that $a_1^2 + b_1^2 \neq 0$. Squaring both sides of the first and second equations of system (3.11) and adding the resulting equations together, we get

$$\begin{gathered} 8 - 8\cos(\alpha)\cos(\alpha - \beta) - 8\sin(\alpha)\sin(\alpha - \beta) = a_1^2 + b_1^2, \\ 8 - 8\cos(\beta) = a_1^2 + b_1^2, \\ \cos(\beta) = -\frac{a_1^2 + b_1^2 - 8}{8}. \end{gathered} \tag{3.14}$$

Let us assume that a solution to (3.14) exists. Then we can write this solution in the form

$$\beta_1 = \arccos\left(-\frac{a_1^2 + b_1^2 - 8}{8}\right) + 2\pi k, \quad k \in \mathbb{Z}, \tag{3.15}$$

$$\beta_2 = -\arccos\left(-\frac{a_1^2 + b_1^2 - 8}{8}\right) + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.16}$$

So, we have expressed β in terms of a_1 and b_1. Now let us express α in terms of β. Applying the appropriate trigonometric identities to (3.11), we have

$$\begin{cases} 2\sin(\alpha) - 2\sin(\alpha)\cos(\beta) + 2\cos(\alpha)\sin(\beta) = a_1, \\ -2\cos(\alpha) + 2\cos(\alpha)\cos(\beta) + 2\sin(\alpha)\sin(\beta) = b_1; \end{cases}$$

$$\begin{cases} 2\big(1 - \cos(\beta)\big)\sin(\alpha) + 2\sin(\beta)\cos(\alpha) = a_1, \\ 2\sin(\beta)\sin(\alpha) - 2\big(1 - \cos(\beta)\big)\cos(\alpha) = b_1. \end{cases} \tag{3.17}$$

Since $a_1^2 + b_1^2 \neq 0$, then $\cos(\beta) \neq 1$ and $\sin(\beta) \neq 0$. Solving (3.17) for $\cos(\alpha)$ and $\sin(\alpha)$ yields

$$\begin{cases} \cos(\alpha) = \dfrac{a_1\sin(\beta) - b_1\big(1 - \cos(\beta)\big)}{2\big(1 - \cos(\beta)\big)^2 + 2\sin^2(\beta)}, \\ \sin(\alpha) = \dfrac{a_1\big(1 - \cos(\beta)\big) + b_1\sin(\beta)}{2\big(1 - \cos(\beta)\big)^2 + 2\sin^2(\beta)}. \end{cases} \tag{3.18}$$

Let us assume that a solution to (3.18) exists. Then we can write this solution in the form

$$\alpha = \operatorname{sgn}\left(\frac{a_1\big(1 - \cos(\beta)\big) + b_1\sin(\beta)}{2\big(1 - \cos(\beta)\big)^2 + 2\sin^2(\beta)}\right)\arccos\left(\frac{a_1\sin(\beta) - b_1\big(1 - \cos(\beta)\big)}{2\big(1 - \cos(\beta)\big)^2 + 2\sin^2(\beta)}\right) + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.19}$$

Next, let us express γ in terms of α and β. From the third equation of system (3.11), we find

$$\gamma = \varphi_f - \alpha + \beta + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.20}$$

Solving (3.8) for Δt_1, Δt_2, and Δt_3, and then using the modulo operation, we obtain the following result:

$$\begin{cases} \Delta t_1 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)(\alpha - \varphi_0)) \mod 2\pi\big), \\ \Delta t_2 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)\beta) \mod 2\pi\big), \\ \Delta t_3 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)\gamma) \mod 2\pi\big), \end{cases} \tag{3.21}$$

where α, β, and γ are defined by (3.19), (3.15), (3.16), and (3.20), respectively. Formula (3.21) gives us two solutions: the first solution corresponds to $\beta = \beta_1$, and the second solution corresponds to $\beta = \beta_2$. It is clear that both solutions satisfy the nonnegativity condition (3.1). Therefore, to find the minimum of the performance index, it remains necessary to compare these two solutions.

Remark 1. When substituting (3.19), (3.15), (3.16), and (3.20) into (3.21), one can assume $k = 0$ in all these formulas.

Remark 2. From [5, Theorem 12], it follows that if the time-optimal control of the Dubins car is of the type $(u_*, -u_*, u_*)$ with nondegenerate Δt_1, Δt_2, and Δt_3, then $\Delta t_2 > \pi/|u_*|$. This condition implies that

$$\beta = -\operatorname{sgn}(u_*) \arccos\left(-\frac{a_1^2 + b_1^2 - 8}{8}\right) + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.22}$$

When substituting (3.22) into (3.21), one can assume $k = 0$.

3.2. Controls of the type $(u_*, 0, u_*)$

Given a vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$, the goal is to find a control of the type $(u_*, 0, u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ in minimum time.

For this type of control, the function $\varphi(t)$ on the time interval $[t_0, t_f]$ can be expressed as

$$\varphi(t) = \begin{cases} \varphi_0 + u_*(t - t_0), & t \in [t_0, t_1), \\ \varphi_0 + u_*\Delta t_1, & t \in [t_1, t_2), \\ \varphi_0 + u_*\Delta t_1 + u_*(t - t_2), & t \in [t_2, t_f]. \end{cases} \tag{3.23}$$

Substituting (3.23) into (2.1) gives

$$\begin{aligned} x(t_f) = x_0 &+ \int_0^{\Delta t_1} v\cos\bigl(\varphi_0 + u_*\tau\bigr)\,d\tau + \int_0^{\Delta t_2} v\cos\bigl(\varphi_0 + u_*\Delta t_1\bigr)\,d\tau \\ &+ \int_0^{\Delta t_3} v\cos\bigl(\varphi_0 + u_*\Delta t_1 + u_*\tau\bigr)\,d\tau \\ = x_0 &+ \frac{v}{u_*}\Bigl(\sin\bigl(\varphi(t_f)\bigr) - \sin\bigl(\varphi_0\bigr)\Bigr) + v\Delta t_2\cos\bigl(\varphi_0 + u_*\Delta t_1\bigr), \end{aligned} \tag{3.24}$$

$$\begin{aligned} y(t_f) = y_0 &+ \int_0^{\Delta t_1} v\sin\bigl(\varphi_0 + u_*\tau\bigr)\,d\tau + \int_0^{\Delta t_2} v\sin\bigl(\varphi_0 + u_*\Delta t_1\bigr)\,d\tau \\ &+ \int_0^{\Delta t_3} v\sin\bigl(\varphi_0 + u_*\Delta t_1 + u_*\tau\bigr)\,d\tau \\ = y_0 &- \frac{v}{u_*}\Bigl(\cos\bigl(\varphi(t_f)\bigr) - \cos\bigl(\varphi_0\bigr)\Bigr) + v\Delta t_2\sin\bigl(\varphi_0 + u_*\Delta t_1\bigr). \end{aligned} \tag{3.25}$$

Combining (3.23), (3.24), and (3.25) with the terminal condition, we obtain the system

$$\begin{cases} x_0 + \dfrac{v}{u_*}(\sin(\varphi_f) - \sin(\varphi_0)) + v\Delta t_2\cos(\varphi_0 + u_*\Delta t_1) = x_f, \\ y_0 - \dfrac{v}{u_*}(\cos(\varphi_f) - \cos(\varphi_0)) + v\Delta t_2\sin(\varphi_0 + u_*\Delta t_1) = y_f, \\ \varphi_0 + u_*\Delta t_1 + u_*\Delta t_3 = \varphi_f + 2\pi k, \quad k \in \mathbb{Z}, \end{cases} \tag{3.26}$$

where Δt_1, Δt_2, and Δt_3 are unknowns.

Thus, the problem can be formulated as follows: find a solution to system (3.26) that satisfies the nonnegativity condition (3.1) and minimizes the performance index

$$T_2 = \Delta t_1 + \Delta t_2 + \Delta t_3.$$

Introduce the notation

$$\begin{aligned} \alpha &= \varphi_0 + u_*\Delta t_1, \quad \gamma = u_*\Delta t_3, \\ a_1 &= \frac{u_*}{v}(x_f - x_0) - \sin(\varphi_f) + \sin(\varphi_0), \\ b_1 &= \frac{u_*}{v}(y_f - y_0) + \cos(\varphi_f) - \cos(\varphi_0). \end{aligned} \tag{3.27}$$

With this notation, system (3.26) may be written as

$$\begin{cases} u_*\Delta t_2 \cos(\alpha) = a_1, \\ u_*\Delta t_2 \sin(\alpha) = b_1, \\ \alpha + \gamma = \varphi_f + 2\pi k, \quad k \in \mathbb{Z}. \end{cases} \tag{3.28}$$

1. Suppose that $a_1^2 + b_1^2 = 0$. It is easy to show that, in this case, the solution is (3.13).

2. Suppose now that $a_1^2 + b_1^2 \neq 0$. Squaring both sides of the first and second equations of system (3.28) and adding the resulting equations together, we get

$$\begin{aligned} u_*^2\Delta t_2^2 &= a_1^2 + b_1^2, \\ \Delta t_2 &= \frac{1}{|u_*|}\sqrt{a_1^2 + b_1^2}. \end{aligned} \tag{3.29}$$

So, we have expressed Δt_2 in terms of a_1 and b_1. Now we can find α by substituting (3.29) into (3.28). We have

$$\alpha = \operatorname{sgn}\left(\frac{|u_*|b_1}{u_*\sqrt{a_1^2 + b_1^2}}\right) \arccos\left(\frac{|u_*|a_1}{u_*\sqrt{a_1^2 + b_1^2}}\right) + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.30}$$

Next, let us express γ in terms of α. From the third equation of system (3.28), we find

$$\gamma = \varphi_f - \alpha + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.31}$$

Solving (3.27) for Δt_1 and Δt_3, and then using the modulo operation, we obtain the following result:

$$\begin{cases} \Delta t_1 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)(\alpha - \varphi_0)) \mod 2\pi\big), \\ \Delta t_2 = \dfrac{1}{|u_*|}\sqrt{a_1^2 + b_1^2}, \\ \Delta t_3 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)\gamma) \mod 2\pi\big), \end{cases} \tag{3.32}$$

where α and γ are defined by (3.30) and (3.31), respectively.

Remark 3. When substituting (3.30) and (3.31) into (3.32), one can assume $k = 0$ in all these formulas.

3.3. Controls of the type $(u_*, 0, -u_*)$

Given a vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$, the goal is to find a control of the type $(u_*, 0, -u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ in minimum time.

For this type of control, the function $\varphi(t)$ on the time interval $[t_0, t_f]$ can be expressed as

$$\varphi(t) = \begin{cases} \varphi_0 + u_*(t - t_0), & t \in [t_0, t_1), \\ \varphi_0 + u_*\Delta t_1, & t \in [t_1, t_2), \\ \varphi_0 + u_*\Delta t_1 - u_*(t - t_2), & t \in [t_2, t_f]. \end{cases} \tag{3.33}$$

Substituting (3.33) into (2.1) gives

$$\begin{aligned} x(t_f) = x_0 + \int_0^{\Delta t_1} v\cos(\varphi_0 + u_*\tau)\, d\tau + \int_0^{\Delta t_2} v\cos(\varphi_0 + u_*\Delta t_1)\, d\tau \\ + \int_0^{\Delta t_3} v\cos(\varphi_0 + u_*\Delta t_1 - u_*\tau)\, d\tau \\ = x_0 + \frac{v}{u_*}\big(2\sin(\varphi_0 + u_*\Delta t_1) - \sin(\varphi(t_f)) - \sin(\varphi_0)\big) + v\Delta t_2\cos(\varphi_0 + u_*\Delta t_1), \end{aligned} \tag{3.34}$$

$$\begin{aligned} y(t_f) = y_0 + \int_0^{\Delta t_1} v\sin(\varphi_0 + u_*\tau)\, d\tau + \int_0^{\Delta t_2} v\sin(\varphi_0 + u_*\Delta t_1)\, d\tau \\ + \int_0^{\Delta t_3} v\sin(\varphi_0 + u_*\Delta t_1 - u_*\tau)\, d\tau \\ = y_0 - \frac{v}{u_*}\big(2\cos(\varphi_0 + u_*\Delta t_1) - \cos(\varphi(t_f)) - \cos(\varphi_0)\big) + v\Delta t_2\sin(\varphi_0 + u_*\Delta t_1). \end{aligned} \tag{3.35}$$

Combining (3.33), (3.34), and (3.35) with the terminal condition, we obtain the system

$$\begin{cases} x_0 + \dfrac{v}{u_*}\big(2\sin(\varphi_0 + u_*\Delta t_1) - \sin(\varphi_f) - \sin(\varphi_0)\big) + v\Delta t_2\cos(\varphi_0 + u_*\Delta t_1) = x_f, \\ y_0 - \dfrac{v}{u_*}\big(2\cos(\varphi_0 + u_*\Delta t_1) - \cos(\varphi_f) - \cos(\varphi_0)\big) + v\Delta t_2\sin(\varphi_0 + u_*\Delta t_1) = y_f, \\ \varphi_0 + u_*\Delta t_1 - u_*\Delta t_3 = \varphi_f + 2\pi k, \quad k \in \mathbb{Z}, \end{cases} \tag{3.36}$$

where Δt_1, Δt_2, and Δt_3 are unknowns.

Thus, the problem can be formulated as follows: find a solution to system (3.36) that satisfies the nonnegativity condition (3.1) and minimizes the performance index

$$T_3 = \Delta t_1 + \Delta t_2 + \Delta t_3.$$

Introduce the notation

$$\begin{gathered} \alpha = \varphi_0 + u_*\Delta t_1, \quad \gamma = u_*\Delta t_3, \\ a_2 = \frac{u_*}{v}(x_f - x_0) + \sin(\varphi_f) + \sin(\varphi_0), \\ b_2 = \frac{u_*}{v}(y_f - y_0) - \cos(\varphi_f) - \cos(\varphi_0). \end{gathered} \tag{3.37}$$

With this notation, system (3.7) may be written as

$$\begin{cases} 2\sin(\alpha) + u_*\Delta t_2\cos(\alpha) = a_2, \\ -2\cos(\alpha) + u_*\Delta t_2\sin(\alpha) = b_2, \\ \alpha - \gamma = \varphi_f + 2\pi k, \quad k \in \mathbb{Z}. \end{cases} \tag{3.38}$$

Squaring both sides of the first and second equations of system (3.38) and adding the resulting equations together, we get

$$4 + u_*^2 \Delta t_2^2 = a_2^2 + b_2^2. \tag{3.39}$$

Let us assume that a solution to (3.39) exists. Then, we have

$$\Delta t_2 = \frac{1}{|u_*|} \sqrt{a_2^2 + b_2^2 - 4}. \tag{3.40}$$

So, we have expressed Δt_2 in terms of a_2 and b_2. Now let us express α in terms of Δt_2. Solving the first two equations of system (3.38) for $\cos(\alpha)$ and $\sin(\alpha)$ yields

$$\begin{cases} \cos(\alpha) = \dfrac{a_2 u_* \Delta t_2 - 2b_2}{4 + u_*^2 \Delta t_2^2}, \\ \sin(\alpha) = \dfrac{b_2 u_* \Delta t_2 + 2a_2}{4 + u_*^2 \Delta t_2^2}. \end{cases} \tag{3.41}$$

Let us assume that a solution to (3.41) exists. Then, after substituting (3.40) into (3.41), we can write this solution in the form

$$\begin{aligned} \alpha = \operatorname{sgn} &\left(\frac{(b_2 u_*/|u_*|)\sqrt{a_2^2 + b_2^2 - 4} + 2a_2}{a_2^2 + b_2^2} \right) \\ \times \arccos &\left(\frac{(a_2 u_*/|u_*|)\sqrt{a_2^2 + b_2^2 - 4} - 2b_2}{a_2^2 + b_2^2} \right) + 2\pi k, \quad k \in \mathbb{Z}, \end{aligned} \tag{3.42}$$

where (3.40) guarantees that $a_2^2 + b_2^2 \neq 0$.

Next, let us express γ in terms of α. From the third equation of system (3.38), we find

$$\gamma = \alpha - \varphi_f + 2\pi k, \quad k \in \mathbb{Z}. \tag{3.43}$$

Solving (3.37) for Δt_1 and Δt_3, and then using the modulo operation, we obtain the following result:

$$\begin{cases} \Delta t_1 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)(\alpha - \varphi_0) r) \mod 2\pi\big), \\ \Delta t_2 = \dfrac{1}{|u_*|}\sqrt{a_2^2 + b_2^2 - 4}, \\ \Delta t_3 = \dfrac{1}{|u_*|}\big((\operatorname{sgn}(u_*)\gamma) \mod 2\pi\big), \end{cases} \tag{3.44}$$

where α and γ are defined by (3.42) and (3.43), respectively.

Remark 4. When substituting (3.42) and (3.43) into (3.44), one can assume $k = 0$ in all these formulas.

4. Analysis of solutions

Let us introduce some additional definitions.

Definition 5. *An open (closed) disc of radius r and center (x_*, y_*) is the set of points (x, y) such that*

$$(x - x_*)^2 + (y - y_*)^2 < r^2 \quad ((x - x_*)^2 + (y - y_*)^2 \leqslant r^2).$$

Definition 6. *We say that a vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$ is feasible for controls of the type (u_1, u_2, u_3) if there exists a control of the type (u_1, u_2, u_3) that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states*

$$\big\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\big\}.$$

We now proceed to investigate the properties of solutions.

4.1. Controls of the type $(u_*, -u_*, u_*)$

Let the notation be as in Section 3.1.

We first obtain necessary and sufficient conditions for the existence of a solution to the time-optimal control problem of the Dubins car for controls of the type $(u_*, -u_*, u_*)$. To do this, we prove the following proposition.

Proposition 1. *System (2.1) can be transferred from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, -u_*, u_*)$ if and only if the point (x_f, y_f) belongs to a closed disc $\mathbb{B}_1$ of radius $4v/|u_*|$ centered at the point (x_*, y_*) defined by*

$$(x_*, y_*) = \Big(x_0 + \frac{v}{u_*}\sin(\varphi_f) - \frac{v}{u_*}\sin(\varphi_0),\ y_0 - \frac{v}{u_*}\cos(\varphi_f) + \frac{v}{u_*}\cos(\varphi_0)\Big).$$

P r o o f. **1.** First, we show that if $(x_f, y_f) \in \mathbb{B}_1$, then there exists a control of the type $(u_*, -u_*, u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$.

It is easy to see that (x_f, y_f) is the center of the closed disc $\mathbb{B}_1$ if and only if $a_1^2 + b_1^2 = 0$. This follows immediately from (3.9) and (3.10). In this case, system (3.7) has solution (3.13).

Let the point (x_f, y_f) belong to the closed disc $\mathbb{B}_1$, but it is not the center of this disc. In this case, equation (3.14) has a solution if and only if

$$-1 \leqslant -\frac{a_1^2 + b_1^2 - 8}{8} < 1. \tag{4.1}$$

We write (4.1) as

$$0 < a_1^2 + b_1^2 \leqslant 16. \tag{4.2}$$

Multiplying all parts of (4.2) by $(v/u_*)^2$ gives

$$0 < \left(\frac{v}{u_*}a_1\right)^2 + \left(\frac{v}{u_*}b_1\right)^2 \leqslant \left(4\frac{v}{u_*}\right)^2. \tag{4.3}$$

Thinking of x_f and y_f as variables, it is easy to see that expression (4.3) defines all the points of the closed disc $\mathbb{B}_1$ except for the center point. Since we assumed that (x_f, y_f) belongs to the closed disc $\mathbb{B}_1$, but it is not the center of this disc, conditions (4.1)–(4.3) are met, which implies that a solution to equation (3.14) exists. Let us check that a solution to system (3.18) also exists. For this, we find the sum of the squares of the right-hand sides of the equations of this system. Substituting (3.14) into (3.18), we have

$$\begin{aligned}
&\frac{a_1^2\sin^2(\beta) - 2a_1b_1\sin(\beta)(1-\cos(\beta)) + b_1^2(1-\cos(\beta))^2}{4\big((1-\cos(\beta))^2 + \sin^2(\beta)\big)^2}\\
&+\frac{a_1^2(1-\cos(\beta))^2 + 2a_1b_1\sin(\beta)(1-\cos(\beta)) + b_1^2\sin^2(\beta)}{4\big((1-\cos(\beta))^2 + \sin^2(\beta)\big)^2} = \frac{(a_1^2+b_1^2)\big((1-\cos(\beta))^2 + \sin^2(\beta)\big)}{4\big((1-\cos(\beta))^2 + \sin^2(\beta)\big)^2}\\
&= \frac{a_1^2 + b_1^2}{4 - 8\cos(\beta) + 4\cos^2(\beta) + 4\sin^2(\beta)} = \frac{a_1^2 + b_1^2}{8 - 8\cos(\beta)} = 1.
\end{aligned}$$

Thus, we see that, for any β satisfying (3.14), under the condition $a_1^2 + b_1^2 \neq 0$, the equations of system (3.18) indeed represent the sine and cosine of some angle α. Consequently, system (3.7) has solution (3.21).

2. Now we show that if there exists a control of the type $(u_*, -u_*, u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$, then $(x_f, y_f) \in \mathbb{B}_1$.

Suppose that $(x_f, y_f) \notin \mathbb{B}_1$. Then (4.1)–(4.3) are not met. Hence, equation (3.14) has no solution, and therefore system (3.7) also has no solution. This is a contradiction. □

Corollary 1. *System (2.1) can be transferred from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, -u_*, u_*)$ if and only if the point (x_0, y_0) belongs to a closed disc $\mathbb{B}_1^*$ of radius $4v/|u_*|$ centered at the point (x^*, y^*) defined by*

$$(x^*, y^*) = \Big(x_f - \frac{v}{u_*}\sin(\varphi_f) + \frac{v}{u_*}\sin(\varphi_0),\ y_f + \frac{v}{u_*}\cos(\varphi_f) - \frac{v}{u_*}\cos(\varphi_0)\Big).$$

Corollary 2. *If the point (x_f, y_f) belongs to the closed disc $\mathbb{B}_1$, but it is not the center of this disc, then solutions to equations (3.14) and (3.18) exist. In this case, system (3.7) will have solution (3.21).*

Corollary 3. *If the point (x_f, y_f) is the center of the closed disc $\mathbb{B}_1$, then system (3.7) will have solution (3.13).*

Next, we turn to the question of the uniqueness of the time-optimal control.

Proposition 2. *Let W_1 be the set of all feasible vectors of boundary conditions for controls of the type $(u_*, -u_*, u_*)$. For any $\mathbf{w} \in W_1$, there are at most two different time-optimal controls of the type $(u_*, -u_*, u_*)$.*

P r o o f. **1.** Assume that $a_1^2 + b_1^2 = 0$. Then the set of all solutions to system (3.7) will be determined by expression (3.12). The right-hand sides of the equations of (3.12) can be expressed in the form $\lambda\big(f(x) + G\big)$, where λ is a positive real number, $f(x)$ is a constant function, and $G = \{2\pi k \mid k \in \mathbb{Z}\}$. By the first part of Lemma 1, it follows that the modulo operation allows us to extract the smallest nonnegative value from this sets of values. After doing this, we can see that the middle segment of the optimal control is degenerate, and since $u_1 = u_3$, we infer that all optimal solutions to system (3.7) generate the same optimal control. So, in this case, the optimal control is unique.

2. Assume that $a_1^2 + b_1^2 \neq 0$. Using Lemma 1, we see that expression (3.21) determines at most two different solutions to system (3.7). The first of them corresponds to the case $\beta = \beta_1$, and the second corresponds to the case $\beta = \beta_2$, where β_1 and β_2 are defined by (3.15) and (3.16). It also follows from Lemma 1 that one of these solutions will be optimal. If the values of the performance index are the same for both the solutions, then both the solutions will be optimal. □

Corollary 4. *Under condition (3.22), expression (3.21) defines the unique control that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states*

$$\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}.$$

Finally, we study the dependence of the movement time on the initial and terminal conditions.

Proposition 3. *Let W_1 be the set of all feasible vectors of boundary conditions for controls of the type $(u_*, -u_*, u_*)$, and let T_1^{opt} be a function that assigns to each $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$ in W_1 the minimum time required to transfer system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, -u_*, u_*)$ under condition (3.22). If $\mathbf{w}_*$ is a point of discontinuity of T_1^{opt}, then at least one of the following conditions holds at $\mathbf{w}_*$:*

1. $a_1^2 + b_1^2 = 0$;
2. $\Delta t_1 = 0$;
3. $\Delta t_3 = 0$.

P r o o f. We will prove this proposition by contradiction. Suppose that none of conditions 1–3 holds at $\mathbf{w}_*$. Observe that $T_1^{opt}(\mathbf{w})$, $\mathbf{w} \in W_1$, represents the optimal value of the performance index T_1 for $\mathbf{w}$ under condition (3.22). Therefore, according to Corollaries 2 and 3, the value of $T_1^{opt}(\mathbf{w})$, $\mathbf{w} \in W_1$, is determined by either (3.13) or (3.21). Since we assumed that condition 1 does not hold at $\mathbf{w}_*$, (3.13) can be ruled out. So, it remains to consider only (3.21). By Corollary 4, under condition (3.22), the value of $T_1^{opt}(\mathbf{w}_*)$ is unique. We need to prove that T_1^{opt} is continuous at $\mathbf{w}_*$. To do this, we will consider Δt_1, Δt_2, Δt_3, α, β, γ, a_1, and b_1 as functions of the vector of boundary conditions $\mathbf{w}$.

It is obvious that a_1 and b_1 are continuous on W_1.

Let us consider expression (3.22). We see that the function β is of the form $\beta(\mathbf{w}) = f_\beta(\mathbf{w}) + G$, where f_β is a continuous single-valued function and $G = \{2\pi k \mid k \in \mathbb{Z}\}$. Since we assumed that condition 1 does not hold at $\mathbf{w}_*$, we have $f_\beta(\mathbf{w}_*) \neq 2\pi k$, $k \in \mathbb{Z}$. Hence, by Lemma 1, the function Δt_2 is continuous at $\mathbf{w}_*$.

Let us consider expression (3.19). This expression is a solution of system (3.18). The values of the numerators and denominators of the fractions in this expression continuously depend on $\mathbf{w}$, and the denominators cannot vanish unless $a_1^2(\mathbf{w}) + b_1^2(\mathbf{w}) \neq 0$. Consequently, $\cos\big(\alpha(\mathbf{w})\big)$ and $\sin\big(\alpha(\mathbf{w})\big)$ are continuous at $\mathbf{w}_*$. It can be shown that, in a neighborhood of $\mathbf{w}_*$, the function α is of the form $\alpha(\mathbf{w}) = f_\alpha(\mathbf{w}) + G$, where f_α is a continuous single-valued function. Therefore, the expression $\operatorname{sgn}(u_*)(\alpha - \varphi_0)$ in the first equation of formula (3.21) can also be represented in the same form. Since we assumed that condition 2 does not hold at $\mathbf{w}_*$, we have

$$\operatorname{sgn}(u_*)\big(\alpha(\mathbf{w}_*) - \varphi_0\big) \neq 2\pi k, \quad k \in \mathbb{Z}.$$

Hence, by Lemma 1, the function Δt_1 is continuous at $\mathbf{w}_*$.

Let us consider expression (3.20). It can be shown that, in a neighborhood of $\mathbf{w}_*$, the function γ is of the form $\gamma(\mathbf{w}) = f_\gamma(\mathbf{w}) + G$, where f_γ is a continuous single-valued function. Since we assumed that condition 3 does not hold at $\mathbf{w}_*$, we have $f_\gamma(\mathbf{w}_*) \neq 2\pi k$, $k \in \mathbb{Z}$. Hence, by Lemma 1, the function Δt_3 is continuous at $\mathbf{w}_*$.

Thus, we have shown that each of the functions Δt_1, Δt_2, and Δt_3 is continuous at $\mathbf{w}_*$. So, T_1^{opt} is also continuous at $\mathbf{w}_*$. □

4.2. Controls of the type $(u_*, 0, u_*)$

Let the notation be as in Section 3.2.

We first obtain necessary and sufficient conditions for the existence of a solution to the time-optimal control problem of the Dubins car for controls of the type $(u_*, 0, u_*)$. To do this, we prove the following proposition.

Proposition 4. *For any vector of boundary conditions $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$, system (2.1) can be transferred from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, 0, u_*)$.*

P r o o f. It is obvious that system (3.28) is solvable for any a_1, b_1, and φ_f; so, the switching times can be easily found by using (3.13) and (3.32). $\square$

Corollary 5. *If the point (x_f, y_f) is not the center of the closed disc $\mathbb{B}_1$ defined in Proposition 1, then system (3.26) will have solution (3.32).*

Corollary 6. *If the point (x_f, y_f) is the center of the closed disc $\mathbb{B}_1$ defined in Proposition 1, then system (3.26) will have solution (3.13).*

Next, we turn to the question of the uniqueness of the time-optimal control.

Proposition 5. *For any vector of boundary conditions, the time-optimal control of the type $(u_*, 0, u_*)$ is unique.*

P r o o f is similar to that of Proposition 2. $\square$

Finally, we study the dependence of the movement time on the initial and terminal conditions.

Proposition 6. *Let T_2^{opt} be a function that assigns to each $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$ in $\mathbb{R}^6$ the minimum time required to transfer system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, 0, u_*)$. If $\mathbf{w}_*$ is a point of discontinuity of T_2^{opt}, then at least one of the following conditions holds at $\mathbf{w}_*$:*

1. $a_1^2 + b_1^2 = 0$;
2. $\Delta t_1 = 0$;
3. $\Delta t_3 = 0$.

P r o o f is similar to that of Proposition 3. $\square$

4.3. Controls of the type $(u_*, 0, -u_*)$

Let the notation be as in Section 3.3.

We first obtain necessary and sufficient conditions for the existence of a solution to the time-optimal control problem of the Dubins car for controls of the type $(u_*, 0, -u_*)$. To do this, we prove the following proposition.

Proposition 7. *System (2.1) can be transferred from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, 0, -u_*)$ if and only if the point (x_f, y_f) does not belong to an open disc $\mathbb{B}_2$ of radius $2v/|u_*|$ centered at the point (x_*, y_*) defined by*

$$(x_*, y_*) = \Big(x_0 - \frac{v}{u_*}\sin(\varphi_f) - \frac{v}{u_*}\sin(\varphi_0),\ y_0 + \frac{v}{u_*}\cos(\varphi_f) + \frac{v}{u_*}\cos(\varphi_0)\Big).$$

P r o o f. **1.** First, we show that if $(x_f, y_f) \notin \mathbb{B}_2$, then there exists a control of the type $(u_*, 0, -u_*)$ that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$.

Observe that equation (3.39) has a solution if and only if

$$a_2^2 + b_2^2 \geqslant 4. \tag{4.4}$$

Multiplying both sides of (4.4) by $(v/u_*)^2$ gives

$$\left(\frac{v}{u_*}a_2\right)^2 + \left(\frac{v}{u_*}b_2\right)^2 \geqslant \left(2\frac{v}{u_*}\right)^2. \tag{4.5}$$

Thinking of x_f and y_f as variables, it is easy to see that expression (4.5) defines the points that do not belong to the open disc $\mathbb{B}_2$. Since we assumed that (x_f, y_f) does not belong to the open disc $\mathbb{B}_2$, conditions (4.4) and (4.5) are met, which implies that a solution to equation (3.39) exists. Let us check that a solution to system (3.41) also exists. For this, we find the sum of the squares of the right-hand sides of the equations of this system. Substituting (3.39) into (3.41), we have

$$\frac{a_2^2u_*^2\Delta t_2^2 - 4a_2b_2u_*\Delta t_2 + 4b_2^2}{\left(4+u_*^2\Delta t_2^2\right)^2} + \frac{b_2^2u_*^2\Delta t_2^2 + 4a_2b_2u_*\Delta t_2 + 4a_2^2}{\left(4+u_*^2\Delta t_2^2\right)^2}$$
$$= \frac{a_2^2u_*^2\Delta t_2^2 + 4b_2^2 + b_2^2u_*^2\Delta t_2^2 + 4a_2^2}{\left(4+u_*^2\Delta t_2^2\right)^2} = \frac{\left(a_2^2+b_2^2\right)\left(4+u_*^2\Delta t_2^2\right)}{\left(4+u_*^2\Delta t_2^2\right)^2} = \frac{a_2^2+b_2^2}{4+u_*^2\Delta t_2^2} = 1.$$

Thus, we see that, for any Δt_2 satisfying (3.39), the equations of system (3.41) indeed represent the sine and cosine of some angle α. Consequently, system (3.36) has solution (3.44).

2. Now we show that if there exists a control of the type $(u_*, 0, -u_*)$ that transfers the system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$, then $(x_f, y_f) \notin \mathbb{B}_2$.

Suppose that $(x_f, y_f) \in \mathbb{B}_2$. Then (4.4), (4.5) are not met. Hence, equation (3.39) has no solution, and therefore system (3.36) also has no solution. This is a contradiction. □

Corollary 7. *System (2.1) can be transferred from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, 0, -u_*)$ if and only if the point (x_0, y_0) does not belong to an open disc $\mathbb{B}_2^*$ of radius $2v/|u_*|$ centered at the point (x^*, y^*) defined by*

$$(x^*, y^*) = \Big(x_f + \frac{v}{u_*}\sin(\varphi_f) + \frac{v}{u_*}\sin(\varphi_0),\ y_f - \frac{v}{u_*}\cos(\varphi_f) - \frac{v}{u_*}\cos(\varphi_0)\Big).$$

Corollary 8. *If the point (x_f, y_f) does not belong to the open disc $\mathbb{B}_2$, then solutions to equations (3.39) and (3.41) exist. In this case, system (3.36) will have solution (3.44).*

Next, we turn to the question of the uniqueness of the time-optimal control.

Proposition 8. *Let W_3 be the set of all feasible vectors of boundary conditions for controls of the type $(u_*, 0, -u_*)$. For any $\mathbf{w} \in W_3$, the time-optimal control of the type $(u_*, 0, -u_*)$ is unique.*

P r o o f is similar to that of Proposition 2. □

Finally, we study the dependence of the movement time on the initial and terminal conditions.

Proposition 9. *Let W_3 be the set of all feasible vectors of boundary conditions for controls of the type $(u_*, 0, -u_*)$, and let T_3^{opt} be a function that assigns to each $\mathbf{w} = (x_0, y_0, \varphi_0, x_f, y_f, \varphi_f)$ in W_3 the minimum time required to transfer system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ by a control of the type $(u_*, 0, -u_*)$. If $\mathbf{w}_*$ is a point of discontinuity of T_3^{opt}, then at least one of the following conditions holds at $\mathbf{w}_*$:*
1. $\Delta t_1 = 0$;
2. $\Delta t_3 = 0$.

P r o o f is similar to that of Proposition 3. □

5. Example

As an example, we will demonstrate how the properties of solutions deduced in Section 4 can reduce the computational effort required to solve the time-optimal control problem of the Dubins car. This issue was previously addressed in [17] for scenarios where the starting and ending points are far apart. However, the results from Section 4 are applicable to any configuration of the points.

Suppose that

$$x_0 = 0, \quad y_0 = 0, \quad \varphi_0 = \pi/2, \quad x_f = 3, \quad y_f = 0, \quad \varphi_f = 3\pi/2, \quad u_m = 1, \quad v = 1,$$

and it is desired to find a control that transfers system (2.1) from the initial state (x_0, y_0, φ_0) to one of the terminal states $\{(x_f, y_f, \varphi_f + 2\pi k) \mid k \in \mathbb{Z}\}$ in minimum time.

To solve this problem, we must find the values of Δt_1, Δt_2, and Δt_3 for controls of the types $(1,-1,1)$, $(-1,1,-1)$, $(1,0,1)$, $(-1,0,-1)$, $(1,0,-1)$, $(-1,0,1)$ using the corresponding formulas from Section 3, and then choose the one of these controls that transfers system (2.1) from the initial state to the terminal state in minimum time. Observe that Propositions 1 and 7 allow us to rule out some cases. Namely, we can exclude controls of the type $(1,-1,1)$ since the point $(3,0)$ does not belong to a closed disc of radius 4 centered at $(-2,0)$. Let us calculate the values of Δt_1, Δt_2, and Δt_3 for the remaining types of controls. Calculations for controls of the type $(-1,1,-1)$ will be carried out taking into account Remark 2. The results of these calculations are given in Table 1, where $T = \Delta t_1 + \Delta t_2 + \Delta t_3$.

Table 1. Time intervals for different control types.

Control types	Δt_1	Δt_2	Δt_3	T
$(-1,1,-1)$	4.46	5.78	4.46	14.7
$(1,0,1)$	4.71	5.0	4.71	14.42
$(-1,0,-1)$	1.57	1.0	1.57	4.14
$(1,0,-1)$	5.44	2.24	2.3	9.98
$(-1,0,1)$	2.3	2.24	5.44	9.98

Comparing the total movement times T of each type of control, we see that the time-optimal control is a control of the type $(-1,0,-1)$. Figure 1 shows the trajectory of the vehicle in the xy-plane, generated by this control. The arrow indicates the direction of the movement.

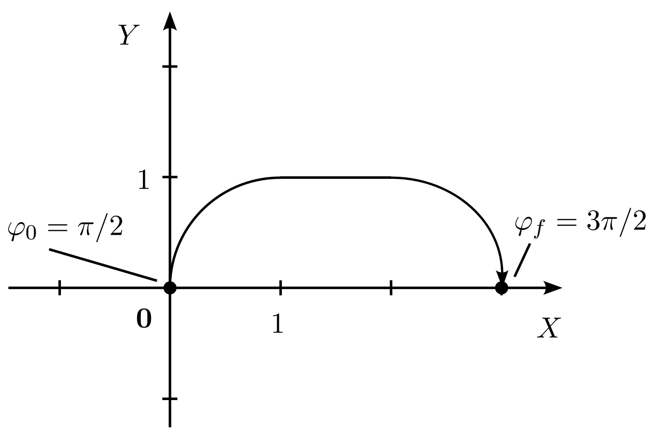


Figure 1. The optimal trajectory of the vehicle in the xy-plane.

6. Conclusion

In this paper, we have developed several fundamental properties for each type of controls in the time-optimal control problem of the Dubins car. The necessary and sufficient conditions for the existence of solutions determine the shape of the regions in the plane to which the vehicle can be driven by a control of the corresponding type. Since the regions are circular in shape, checking whether points belong to these regions can be done quite simply, and so this reduces the computational effort in solving the Dubins car control problem.

REFERENCES

1. Ayala J., Rubinstein H. The classification of homotopy classes of bounded curvature paths. *Isr. J. Math.*, 2016. Vol. 213. P. 79–107. DOI: 10.1007/s11856-016-1321-x
2. Balkcom D., Furtuna A., Wang W. The Dubins car and other arm-like mobile robots. In: *Proc. 2018 IEEE Int. Conf. on Robot. and Aut. (ICRA), Brisbane, QLD, Australia*, 2018. P. 380–386. DOI: 10.1109/ICRA.2018.8461017
3. Berdyshev Iu. I. Time-optimal control synthesis for a fourth-order nonlinear system. *J. Appl. Math. Mech.*, 1975. Vol. 39, No. 6. P. 948–956. DOI: 10.1016/0021-8928(76)90081-2
4. Boissonnat J.-D., Cerezo A., Leblond J. A note on shortest paths in the plane subject to a constraint on the derivative of the curvature. In: *INRIA Research Report.* RR-2160. 1994. URL: inria-00074512
5. Boissonnat J.-D., Cérézo A., Leblond J. Shortest paths of bounded curvature in the plane. *J. Intell. Robot. Syst.*, 1994. Vol. 11. P. 5–20. DOI: 10.1007/BF01258291
6. Buzikov M. E., Galyaev A. A. Minimum-time lateral interception of a moving target by a Dubins car. *Automatica*, 2022. Vol. 135. Art. no. 109968. DOI: 10.1016/j.automatica.2021.109968
7. Dubins L. E. On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents. *Amer. J. Math.*, 1957. Vol. 79, No. 3. P. 497–516. DOI: 10.2307/2372560
8. Johnson H. H. An application of the maximum principle to the geometry of plane curves. *Proc. Amer. Math. Soc.*, 1974. Vol. 44, No. 2. P. 432–435. DOI: 10.1090/S0002-9939-1974-0348631-6
9. Kaya C. Y. Markov–Dubins path via optimal control theory. *Comput. Optim. Appl.*, 2017. Vol. 68. P. 719–747. DOI: 10.1007/s10589-017-9923-8
10. Khabarov S. P., Shilkina M. L. Geometric approach to the solution of the Dubins car problem in the formation of program trajectories. *Nauch.-Tehn. Vest. Inf. Tekh., Mekh. i Opt.*, 2021. Vol. 21, No. 5. P. 653–663. DOI: 10.17586/2226-1494-2021-21-5-653-663 (in Russian)
11. Markov A. A. A few examples of solving special problems on the largest and smallest values. *Soobshch. Kharkov. Mat. Obshch.*, 1889. Ser. 2, vol. 1, No. 2. P. 250–276. (in Russian)
12. McGee T. G., Hedrick J. K. Optimal path planning with a kinematic airplane model. *J. Guid. Cont. Dyn.*, 2007. Vol. 30, No. 2. P. 629–633. DOI: 10.2514/1.25042
13. Patsko V. S., Pyatko S. G., Fedotov A. A. Three-dimensional reachability set for a nonlinear control system. *J. Comp. Syst. Sci. Int.*, 2003. Vol. 42, No. 3. P. 320–328.
14. Patsko V. S., Fedotov A. A. Three-dimensional reachable set for the Dubins car: Foundation of analytical description. *Commun. Optim. Theory*, 2022. Vol. 2022. P. 1–42. DOI: 10.23952/cot.2022.23
15. Pecsvaradi T. Optimal horizontal guidance law for aircraft in the terminal area. *IEEE Trans. Automatic Control*, 1972. Vol. 17, No. 6. P. 763–772. DOI: 10.1109/TAC.1972.1100160
16. Reeds J. A., Shepp L. A. Optimal paths for a car that goes both forwards and backwards. *Pacific J. Math.*, 1990. Vol. 145, No. 2. P. 367–393. DOI: 10.2140/pjm.1990.145.367
17. Shkel A. M., Lumelsky V. J. Classification of the Dubins set. *Robotics Auton. Syst.*, 2001. Vol. 34, No. 4. P. 179–202. DOI: 10.1016/S0921-8890(00)00127-5
18. Sussmann H. J. Shortest 3-dimensional paths with a prescribed curvature bound. In: *Proc. of 1995 34th IEEE Conf. on Decision and Control*, 1995. Vol. 4. P. 3306–3312. DOI: 10.1109/CDC.1995.478997
19. Sussmann H. J., Tang G. Shortest paths for the Reeds–Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control. In: *Rutgers Center for Syst. and Cont. Technical Report 91–10*, 1991. URL: SYCON-91-10

Editor: Tatiana F. Filippova

Managing Editor: Oxana G. Matviychuk

Design: Alexander R. Matviychuk

Contact Information

16 S. Kovalevskaya str., Ekaterinburg, Russia, 620990

Phone: +7 (343) 375-34-73

Fax: +7 (343) 374-25-81

Email: secretary@umjuran.ru

Web-site: https://umjuran.ru

N.N.Krasovskii Institute of Mathematics and Mechanics
of the Ural Branch of Russian Academy of Sciences

Ural Federal University named after the first President of Russia B.N.Yeltsin

Distributed for free