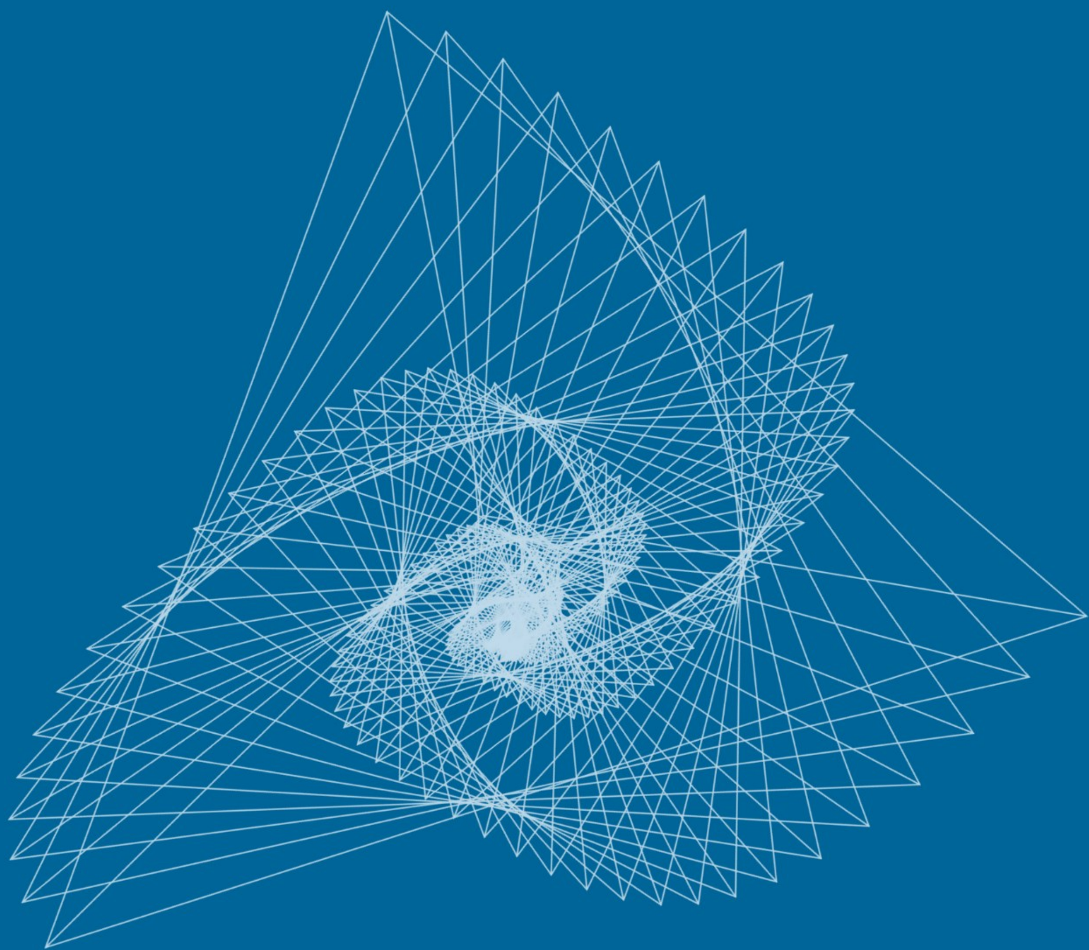
VOL. 9, NO. 2

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



2023

UMJURAN.RU

UMJ

ISSN 2414-3952
(Online)

Ural Mathematical Journal

Vol. 9, No. 2, 2023

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. 9, No. 2, 2023, pp. 4–27

DOI: 10.15826/umj.2023.2.001

APPROXIMATION OF DIFFERENTIATION OPERATORS BY BOUNDED LINEAR OPERATORS IN LEBESGUE SPACES ON THE AXIS AND RELATED PROBLEMS IN THE SPACES OF (p, q)-MULTIPLIERS AND THEIR PREDUAL SPACES[1]

Vitalii V. Arestov

Ural Federal University,
51 Lenin ave., Ekaterinburg, 620000, Russian Federation

vitalii.arestov@urfu.ru

Abstract: We consider a variant $E_{n,k}(N; r, r; p, p)$ of the four-parameter Stechkin problem $E_{n,k}(N; r, s; p, q)$ on the best approximation of differentiation operators of order k on the class of n times differentiable functions $(0 < k < n)$ in Lebesgue spaces on the real axis. We discuss the state of research in this problem and related problems in the spaces of multipliers of Lebesgue spaces and their predual spaces. We give two-sided estimates for $E_{n,k}(N; r, r; p, p)$. The paper is based on the author's talk at the S.B.Stechkin's International Workshop-Conference on Function Theory (Kyshtym, Chelyabinsk region, August 1–10, 2023).

Keywords: Differentiation operator, Stechkin's problem, Kolmogorov inequality, (p, q)-Multiplier, Predual space for the space of (p, q)-multipliers.

1. Introduction

1.1. Some notation

In this paper, we use the standard notation for the classical complex spaces of complex-valued measurable (in particular, continuous) functions of one variable on the real axis. Thus, $L_\gamma = L_\gamma(-\infty, \infty)$, $1 \le \gamma < \infty$, is the Lebesgue space of functions f measurable on the real axis $\mathbb{R} = (-\infty, \infty)$ such that the function $|f|^\gamma$ is integrable over the axis; the space L_γ is equipped with the norm

$$\|f\|_\gamma = \|f\|_{L_\gamma} = \left(\int |f(t)|^\gamma dt \right)^{1/\gamma};$$

hereinafter, we omit the integration set in integrals over the axis. The space $L_\infty = L_\infty(-\infty, \infty)$ consists of measurable essentially bounded functions on the axis; the space is equipped with the norm

$$\|f\|_\infty = \|f\|_{L_\infty} = \text{ess sup} \big\{ |f(t)| \colon t \in (-\infty, \infty) \big\}.$$

The space $L_\infty = L_\infty(-\infty, \infty)$ contains the space $C = C(-\infty, \infty)$ of bounded continuous functions on the axis with the uniform norm

$$\|f\|_C = \sup\{|f(t)| \colon t \in (-\infty, \infty)\}.$$

Let $C_0 = C_0(-\infty, \infty)$ be the subspace of $C = C(-\infty, \infty)$ of functions vanishing at infinity. Denote by V the space of (complex) bounded Borel measures on $(-\infty, \infty)$. We will identify this set with the set of (complex) functions μ of bounded variation on $(-\infty, \infty)$ such that values of their real

[1]This work was supported by the Russian Science Foundation, project no. 22-21-00526, https://rscf.ru/project/22-21-00526/ .

and imaginary parts at the discontinuity points are between the right-sided and left-sided limits. The norm in the space V is the total variation $\bigvee \mu = \bigvee_{-\infty}^{\infty} \mu$ of a measure (a function) $\mu \in V$.

These spaces and their norms are invariant under the group of translations $\{\tau_h,\ h \in \mathbb{R}\}$ defined by the formula $(\tau_h f)(t) = f(t-h)$, $t \in \mathbb{R}$, as well as under the family of operators $\{\sigma_h,\ h \in \mathbb{R}\}$ given by the formula $(\sigma_h f)(t) = f(h-t)$, $t \in \mathbb{R}$. The operators of these two families are related as follows: $\sigma_h = \tau_h \sigma_0$, where σ_0 is the operator of changing the sign of a function argument: $(\sigma_0 f)(t) = f(-t)$, $t \in \mathbb{R}$.

We define the direct and inverse Fourier transforms of functions (at least from the space $L = L_1(\mathbb{R})$) by the formulas

$$\widehat{f}\,(t) = \int e^{-2\pi t\eta i} f(\eta)\, d\eta, \quad \widecheck{f}\,(t) = \int e^{2\pi t\eta i} f(\eta)\, d\eta = \widehat{f}\,(-t), \tag{1.1}$$

respectively. Properties of the Fourier transform can be found, for example, in [44, Ch. I, Sects. 1, 2].

Let $\mathscr{S}$ be the space of rapidly decreasing, infinitely differentiable functions on the axis, and let $\mathscr{S}'$ be the corresponding dual space of generalized functions (see, for example, [41, 42, 44]). The value of a functional $\theta \in \mathscr{S}'$ on a function $\phi \in \mathscr{S}$ will be denoted by $\langle \theta, \phi \rangle$. The space $\mathscr{S}'$ contains the set $\mathscr{L} = \mathscr{L}(\mathbb{R})$ of functions f measurable and locally integrable on $\mathbb{R}$, and satisfying the condition

$$\int (1 + |t|)^d |f(t)| dt < \infty$$

with some exponent $d = d(f) \in \mathbb{R}$; functions $f \in \mathscr{L}$ are called slowly growing (classical) functions. A function $f \in \mathscr{L}$ is associated with a functional $f \in \mathscr{S}'$ by the formula

$$\langle f, \phi \rangle = \int f(t)\phi(t) dt, \quad \phi \in \mathscr{S}.$$

The convolution $\theta * \phi$ of an element $\theta \in \mathscr{S}'$ and a function $\phi \in \mathscr{S}$ is the function $y(\eta) = \langle \theta, \sigma_\eta \phi \rangle$. If $\theta \in \mathscr{L}$ is a classical function, then

$$(\theta * \phi)(\eta) = \int \theta(t)\phi(\eta - t)\, dt.$$

The Fourier transform $\widehat{\theta}$ of a functional $\theta \in \mathscr{S}'$ is a functional $\widehat{\theta} \in \mathscr{S}'$ acting by the formula $\langle \widehat{\theta}, \phi \rangle = \langle \theta, \widehat{\phi} \rangle$, $\phi \in \mathscr{S}$. If $\theta \in L_\gamma$, $1 \le \gamma \le 2$, then $\widehat{\theta} \in L_{\gamma'}$, $1/\gamma + 1/\gamma' = 1$; moreover, the Hausdorff–Young inequality $\|\widehat{\theta}\|_{\gamma'} \le \|\theta\|_\gamma$ holds (see, for example, [44, Ch. V, Sect. 1]).

1.2. Stechkin's problem on the best approximation of differentiation operators by bounded linear operators in Lebesgue spaces on the real axis

Let r, s, p, and q be parameters satisfying the constraints $1 \le r, s, p, q \le \infty$. Let us agree that, for $r = \infty$, by $L_\infty = L_\infty(-\infty, \infty)$, we mean the space $C_0 = C_0(-\infty, \infty)$ of continuous functions on the axis vanishing at infinity. For $p = \infty$, by $L_\infty = L_\infty(-\infty, \infty)$, we mean the classical space of essentially bounded functions on the axis. For $s = \infty$ and $q = \infty$, by $L_\infty = L_\infty(-\infty, \infty)$, we mean the space $C = C(-\infty, \infty)$ of bounded continuous functions on the axis or even the space $C_0 = C_0(-\infty, \infty)$ of continuous functions vanishing at infinity depending on the situation; these situations will be stipulated.

For an integer $n \ge 1$, we define the space $W^n_{r,p}$ of functions $f \in L_r$ that are $n-1$ times continuously differentiable on the axis, their derivatives $f^{(n-1)}$ of order $n-1$ are locally absolutely continuous, and $f^{(n)} \in L_p$. In the space $W^n_{r,p}$, consider the class

$$Q^n_{r,p} = \{f \in W^n_{r,p} \colon \|f^{(n)}\|_p \le 1\}.$$

Denote by $\mathfrak{B}(L_r, L_s)$ the set of all bounded linear operators from L_r to L_s, and let $\mathfrak{B}(N; L_r, L_s)$ for $N > 0$ be the set of operators $T \in \mathfrak{B}(L_r, L_s)$ with the norm $\|T\|_{L_r \to L_s} \le N$. Let $0 \le k < n$ be an integer and $k > 0$ if $r = s$. For an operator $T \in \mathfrak{B}(L_r, L_s)$, define

$$U(T) = \sup\big\{\|f^{(k)} - Tf\|_q \colon f \in Q^n_{r,p}\big\}.$$

If the difference $f^{(k)} - Tf$ does not belong to the space L_q, then we assume that $\|f^{(k)} - Tf\|_{L_q} = \infty$. For $N > 0$, the quantity

$$E(N) = E_{n,k}(N) = E_{n,k}(N; r, s; p, q) = \inf\big\{U(T) \colon T \in \mathfrak{B}(N; L_r, L_s)\big\} \tag{1.2}$$

is the best approximation (in the space L_q) of the differentiation operator of order k on the class $Q^n_{r,p}$ by the set of bounded linear operators $\mathfrak{B}(N; L_r, L_s)$. Stechkin's problem is to study quantity (1.2) and an extremal operator on which the infimum is attained in (1.2); we will call it problem (1.2), and sometimes the problem $E_{n,k}(N; r, s; p, q)$.

Problem (1.2) is a specific version of Stechkin's problem on the best approximation of an unbounded linear operator by bounded linear operators on a class of elements of a Banach space, which arose in his paper [46]. Problem (1.2) and its specific cases were studied by many mathematicians: S.B. Stechkin, L.V. Taikov, Yu.N. Subbotin, V.N. Gabushin, V.I. Berdyshev; V.M.Tikhomirov and his colleagues A.P. Buslaev and G.G. Magaril-Il'yaev; V.F. Babenko and his colleagues and students; V.V. Arestov, R.R. Akopyan, V.G. Timofeev, M.A. Filatova, E.E. Berdysheva, and a lot others; see, for example, the review papers [9, 15, 16] and the bibliography therein. Some specific results will be described in what follows.

Note some facts. Necessary and sufficient conditions for the finiteness of quantity (1.2) are known, see [26] ($s = q$) and [4] ($s \ne q$). Roughly speaking, these conditions are

$$s \ge r, \quad q \ge p. \tag{1.3}$$

More precisely, if conditions (1.3) are satisfied, then there exists $N_0 > 0$ such that $E(N) < \infty$ for $N \ge N_0$. If the problem parameters satisfy the constraints

$$k - \frac{1}{s} + \frac{1}{r} > 0, \quad n - k + \frac{1}{q} - \frac{1}{p} > 0, \tag{1.4}$$

then conditions (1.3) are necessary and sufficient for the quantity $E(N)$ to be finite for any $N > 0$. For a discussion of conditions (1.4), see [4].

Under conditions (1.3) and (1.4), we have the formula

$$E(N) = E(1)N^{-\gamma}, \quad \gamma = (n - k + 1/q - 1/p)/(k + 1/r - 1/s) > 0; \tag{1.5}$$

See [46] for the case $q = p = s = r = \infty$; in the general case, formula (1.5) is justified similarly, see, for example, [2].

The present paper considers problem (1.2) in the case $s = r$ and $q = p$, i.e., the variant $E_{n,k}(N; r, r; p, p)$. We review the results obtained so far in this version of the problem and related problems in multiplier spaces of Lebesgue spaces and predual spaces of multiplier spaces. In the last section of the paper, we give two-sided estimates for the value $E_{n,k}(N; r, r; p, p)$ in this variant of the problem.

1.3. Connection with the Kolmogorov inequality

Stechkin's problem (1.2) is related to several other extremal problems of function theory. Among them are the exact Kolmogorov inequalities for differentiable functions on the axis

$$\|f^{(k)}\|_{L_q} \le G\|f\|^{\alpha}_{L_r}\|f^{(n)}\|^{\beta}_{L_p}, \quad f \in W^n_{r,p}, \tag{1.6}$$

$$\alpha = (n-k-1/p+1/q)/(n-1/p+1/r), \quad \beta = 1-\alpha.$$

Such inequalities were studied by G.H. Hardy, J.E. Littlewood, E. Landau, J. Hadamard, B. Szőkefalvi-Nagy, A.N. Kolmogorov, S.B. Stechkin, L.V. Taikov, V.N. Gabushin, V.I. Berdyshev, N.P. Kuptsov, A.P. Buslaev, G.G. Magaryl-Il'yaev, V.M. Tikhomirov, V.F. Babenko, etc. (see the bibliography in [9, 16, 17, 48]). V.N. Gabushyn found necessary and sufficient conditions for the existence of inequality (1.6), more precisely, for the finiteness of the constant $G = G(n,k;r,p,q)$ in (1.6). Namely, he proved [24] (see also [27, 28]) that $G < \infty$ if and only if

$$\frac{n-k}{r} + \frac{k}{p} \geq \frac{n}{q}. \tag{1.7}$$

S.B. Stechkin made an important observation that, in the classical case $s = q$, the value (1.2) and the best constant in (1.6) are related by the inequality

$$E_{n,k}(N;r,q;p,q) \geq \beta\alpha^{\alpha/\beta}G^{1/\beta}N^{-\alpha/\beta}, \quad N > 0. \tag{1.8}$$

To avoid discussing specific degenerate values of the parameters, we will assume that conditions (1.3) and (1.4) hold. Inequality (1.8) is a specific version of a more general statement and more general considerations by S.B. Stechkin contained in [46, Sect. 2].

As follows from (1.3) and (1.7), the conditions for the finiteness of the value $E(N) = E_{n,k}(N;r,q;p,q)$ of Stechkin's problem and the best constant $G = G(n,k;r,p,q)$ in (1.6) are different. Consequently, there are cases when $E(N) = \infty$ and $G < \infty$; in this situation, inequality (1.8) is strict. In the case $E(N) < \infty$, depending on the values of the parameters, both possibilities are realized: inequality (1.8) can turn into equality, and inequality (1.8) can be strict. A more informative discussion of this issue can be found in [9, Sect. 4].

Inequality (1.8) is an important tool for studying both the three-parameter version of problem (1.2) (the case $s = q$) and inequality (1.6). Indeed, an arbitrary specific function $f^* \in W^n_{r,p}$ estimates the best constant G in inequality (1.6) from below. This, due to (1.8), gives a lower estimate for the quantity $E_{n,k}(N;r,q;p,q)$. A specific operator $T^* \in \mathfrak{B}(L_r, L_q)$ gives an upper estimate for the quantity $E_{n,k}(N;r,q;p,q)$: $E_{n,k}(N;r,q;p,q) \leq U(T^*)$ for $N = \|T^*\|$; in this case, it is not necessary to have the exact value of $U(T^*)$ but only again an upper estimate. If we managed to choose a function f^* and an operator T^* so that the obtained upper and lower estimates for the quantity $E(N)$ coincide, then we have a solution to both the problems. More precisely, we have exact values of $E(\|T^*\|)$ and the best constant G in (1.6). Moreover, the operator T^* is extremal in Stechkin's problem, and the function f^* is extremal in the Kolmogorov inequality. Along this path, a solution to both problems was found in several new cases; see [9, Sect. 4] and the references therein.

The considerations just outlined are not universal in the study of problem (1.2). Firstly, inequality (1.8) can be strict and, therefore, in this case, it is impossible to obtain an exact lower estimate for $E_{n,k}(N;r,q;p,q)$. Secondly, in the four-parameter case $s \neq q$, there is no analog of inequality (1.8), at least in Lebesgue spaces.

In the study of Stechkin's problem, the property of the translation invariance of problem (1.2) occurs useful. The norms of spaces, the class $Q^n_{r,p}$, and the approximated differentiation operator $D^k = d^k/dt^k$ are invariant under the translation group $\{\tau_h\}$; precisely in this sense, we say that problem (1.2) is translation invariant. Due to this property, in problem (1.2), we can restrict ourselves to approximating operators T that are also translation invariant; details can be found in [4–6, 8, 9]. This property makes it possible to solve Stechkin's problem in some cases (in particular, for $s \neq q$) and, which is no less essential, expands the environment of the problem. It is these issues that most of this paper is devoted to.

The property of invariance of approximating operators in Stechkin's problem and related problems in spaces of periodic functions was obtained and applied in the study of these problems by B.E. Klotz [33, 34].

2. Translation invariance of Stechkin's problem

In this section, we present some properties of spaces of bounded linear operators in Lebesgue spaces on the axis that are translation invariant; in particular, we describe their predual spaces.

2.1. The space of translation invariant bounded operators

For $1 \le p, q \le \infty$, denote by $\mathfrak{T}_{p,q} = \mathfrak{T}_{p,q}(\mathbb{R})$ the set of bounded linear operators from $L_p = L_p(\mathbb{R})$ to $L_q = L_q(\mathbb{R})$ that are invariant under (any) translation, i.e., such that $\tau_h T = T\tau_h$ on L_p for all $h \in \mathbb{R}$. Extensive research has been devoted to the properties of invariant bounded operators (see [32, 37, 44] and the references therein). It is known (see, for example, [32, Theorem 1.1]) that if $p > q$, then, for $p < \infty$, the set $\mathfrak{T}_{p,q}$ consists only of the operator $T \equiv 0$, and, for $p = \infty$, the restriction of an operator $T \in \mathfrak{T}_{\infty,q}$ to the set $(L_\infty)_0$ of functions from L_∞ having zero limit at infinity is the zero operator. In this regard, when discussing the properties of bounded invariant operators in what follows, we will assume that $1 \le p \le q \le \infty$.

In a joint paper [23], Figà-Talamanca and Gaudry (1967) proved that, for $1 \le p \le q < \infty$, the space $\mathfrak{T}_{p,q}(G)$ of bounded linear operators from $L_p(G)$ to $L_q(G)$ on a locally compact Abelian group G invariant under translation (more precisely, under the group operation) is the conjugate space for a function space $A_{p,q}(G)$ constructively described by them. More precisely, in [23], function spaces $A_{p,q}(G)$ were constructed such that the space $\mathfrak{T}_{p,q}(G)$ of invariant operators is isometrically isomorphic to the dual space $A^*_{p,q}(G)$, in short, $\mathfrak{T}_{p,q}(G) = A^*_{p,q}(G)$. Two years earlier (in 1965), Figa-Talamanca [22] obtained a similar result for the case $1 < q = p < \infty$.

Let X and Y be a pair of normed linear spaces such that Y is the conjugate space of X, i.e., $X^* = Y$. In this case, we say that X is the predual space of Y. In this terminology, the results of [22] and [23] mean that the spaces $A_{p,q}(G)$ (for $1 \le p \le q < \infty$) are predual of the spaces $\mathfrak{T}_{p,q}(G)$.

The results of [22] and [23] are valid, in particular, for the spaces $\mathfrak{T}_{p,q}(\mathbb{R})$ of bounded linear operators from the space $L_p(\mathbb{R})$ to the space $L_q(\mathbb{R})$ invariant under the group of translations τ_h, $h \in \mathbb{R}$. So, for $1 \le p \le q < \infty$, the spaces $\mathfrak{T}_{p,q}(\mathbb{R})$ are conjugate spaces of the spaces $A_{p,q} = A_{p,q}(\mathbb{R})$ constructed in [22] and [23]; i.e., $A_{p,q}$ are their predual.

In the author's papers (see [12, 14] and the references therein), a function space $F_{p,q} = F_{p,q}(\mathbb{R}) \subset L_r(\mathbb{R})$ was constructed which is the predual space of the space $\mathfrak{T}_{p,q} = \mathfrak{T}_{p,q}(\mathbb{R})$ of translation invariant bounded linear operators from $L_p(\mathbb{R})$ to $L_q(\mathbb{R})$. It is described in terms different from [22, 23], however, (for $1 \le p \le q < \infty$) it coincides, more precisely, is isometrically isomorphic to the space $A_{p,q}(\mathbb{R})$ of Figá-Talamanca and Gaudry [23]. The space $F_{p,q}$ will be described and used in what follows.

2.1.1. The space of (p, q)-multipliers

Let us discuss some properties of bounded linear operators from $L_p(\mathbb{R})$ to $L_q(\mathbb{R})$ that are invariant under (any) translation.

It is known (see [32, Theorem 1.2] or [44, Ch. I, Theorem 3.16]) that, if $q \ge p$, then an operator $T \in \mathfrak{T}_{p,q}$ on $\mathscr{S}$ has the form of the convolution with an element $\theta = \theta_T \in \mathscr{S}'$:

$$T\phi = \theta * \phi, \quad \phi \in \mathscr{S}.$$

The set $M_{p,q} = \{\theta_T \colon T \in \mathfrak{T}_{p,q}\} \subset \mathscr{S}'$ is a Banach space with respect to the norm

$$\|\theta_T\|_{M_{p,q}} = \|T\|_{L_p \to L_q}.$$

Elements $\theta \in M_{p,q}$, $1 \le p \le q \le \infty$ are often called (p,q)-multipliers.

In what follows, we always assume that $1 \le p \le q \le \infty$. Denote by ρ a parameter chosen from the condition

$$1/p - 1/q = 1 - 1/\rho; \tag{2.1}$$

we have $1 \le \rho \le \infty$. It is known that if $\theta \in L_\rho$ and $x \in L_p$, then $\theta * x \in L_q$ and the Young inequality holds (see, for example, [44, Ch. V, Sect. 1]):

$$\|\theta * x\|_q \le \|\theta\|_\rho \|x\|_p. \tag{2.2}$$

This fact and inequality (2.2) imply the embedding

$$L_\rho \subset M_{p,q}, \quad \frac{1}{\rho} = 1 - \left(\frac{1}{p} - \frac{1}{q}\right)$$

with the inequality $\|\theta\|_{M_{p,q}} \le \|\theta\|_\rho$, $\theta \in L_\rho$, for the norms of the elements.

Let us mention further known properties of the spaces $M_{p,q}$ (see, for example, [32, Sect. 1.2], [44, Ch. V, Sect. 1]). For two pairs of conjugate exponents (p,q) and (q',p'), the equality

$$M_{p,q} = M_{q',p'}$$

holds together with the equality of the norms of the elements: $\|\theta\|_{M_{p,q}} = \|\theta\|_{M_{q',p'}}$, $\theta \in M_{p,q}$. From this and the Riesz–Thorin interpolation theorem (see, for example, [21, Ch. VI, Sect. 10, Theorem 11] or [44, Ch. V, Sect. 1, Theorem 1.16]), it follows that if

$$\frac{1}{\alpha} = \frac{1-t}{p} + \frac{t}{q'}, \quad \frac{1}{\beta} = \frac{1-t}{q} + \frac{t}{p'}, \quad 0 \le t \le 1,$$

then we have the embedding

$$M_{p,q} \subset M_{\alpha,\beta}$$

and the inequality

$$\|\theta\|_{M_{\alpha,\beta}} \le \|\theta\|_{M_{p,q}}, \quad \theta \in M_{p,q}.$$

A constructive description of multipliers is known only in several cases. The structure of the spaces $M_{2,2}$ and $M_{p,\infty} = M_{1,p'}$ is known; namely (see, for example, [32, Sect. 1.2] and [44, Ch. 1, Sect. 3]), the following equalities are valid (together with the equalities of the norms of the elements):

$$M_{2,2} = \widehat{L}_\infty = \{\widehat{\theta} \colon \theta \in L_\infty\},$$

$$M_{p,\infty} = M_{1,p'} = L_{p'} \quad \text{for} \quad 1 \le p < \infty,$$

$$M_{\infty,\infty} = M_{1,1} = V;$$

here, $V = V(\mathbb{R})$ is the space of (complex) bounded Borel measures on $\mathbb{R}$.

2.1.2. The predual space of the space of (p, q)-multipliers

This section describes function spaces $F_{p,q}$ constructed by the author in [14] and some of their properties. These spaces are predual of the spaces of multipliers $M_{p,q}$: $F^*_{p,q} = M_{p,q}$. The spaces $F_{p,q}$ are described in different terms compared to $A_{p,q}$ in [23], although, in fact, they are isometrically isomorphic [14, Theorem 3.2]. Here, as before, $1 \le p \le q \le \infty$. Let γ be a parameter defined by the relation

$$1/\gamma = 1/p - 1/q; \tag{2.3}$$

for $\gamma = \infty$ (i.e., for $q = p$), we assume that $L_\gamma = C_0$. Comparing (2.3) with (2.1), we conclude that $\gamma = \rho'$.

On the set $\mathscr{S}$, we define the functional

$$\|\phi\|_{p,q} = \sup\{|\langle\theta, \phi\rangle| \colon \theta \in M_{p,q},\ \|\theta\|_{M_{p,q}} \le 1\}, \quad \phi \in \mathscr{S}. \tag{2.4}$$

Functional (2.4) on the set $\mathscr{S}$ is finite and is a norm [14, Lemma 2.1].

Let $F_{p,q} = F_{p,q}(\mathbb{R})$ be the completion of the space $\mathscr{S}$ with respect to the norm (2.4). For all $1 \le p \le q \le \infty$, the space $F_{p,q}$ is a function space; moreover, it is embedded in the space L_γ [14, Lemma 2.3]:

$$F_{p,q} \subset L_\gamma \quad \text{and} \quad \|f\|_\gamma \le \|f\|_{F_{p,q}}, \quad f \in F_{p,q}. \tag{2.5}$$

Hereinafter, we use the notation $\|f\|_{p,q}$ for the norms $\|f\|_{F_{p,q}}$ of functions $f \in F_{p,q}$.

For the convenience of reference, we formulate as a separate lemma the following statement from [14, Lemma 2.5].

Lemma 1. *For specific values of the parameters, the space $F_{p,q}$ has the following properties.*

(1) *For $q = \infty$,*

$$\begin{gathered} F_{p,\infty} = F_{1,p'} = L_p, \quad 1 \le p < \infty, \\ F_{\infty,\infty} = F_{1,1} = C_0. \end{gathered} \tag{2.6}$$

(2) *For $q = p = 2$,*

$$F_{2,2} = \check{L} = \{f \in C_0 \colon \widehat{f} \in L\}, \quad \|f\|_{2,2} = \|\widehat{f}\,\|_L, \quad f \in F_{2,2}. \tag{2.7}$$

(3) *Let $q = p$ and $\overline{p} = \max\{p, p'\}$. The spaces $F_{p,p}$ do not decrease in $\overline{p}$; more exactly, if $2 \le \overline{p}_1 \le \overline{p}_2 \le \infty$, then*

$$F_{p_1,p_1} \subset F_{p_2,p_2} \quad \textit{and} \quad \|f\|_{p_2,p_2} \le \|f\|_{p_1,p_1}, \quad f \in F_{p_1,p_1}; \tag{2.8}$$

in particular, for all $1 \le p \le \infty$,

$$F_{p,p} \subset C_0 \quad \textit{and} \quad \|f\|_{p,p} \ge \|f\|_{C_0}, \quad f \in F_{p,p},$$

$$F_{2,2} \subset F_{p,p} \quad \textit{and} \quad \|f\|_{p,p} \le \|f\|_{2,2} = \|\widehat{f}\,\|_L, \quad f \in F_{2,2}.$$

The spaces $F_{p,q}$ that are predual spaces of the spaces of (p, q)-multipliers will sometimes be briefly called the *predual spaces*.

2.2. Two extremal problems related to Stechkin's problem (1.2) in the spaces of multipliers and their predual spaces

Let r, s, p, and q be parameters satisfying the constraints $1 \le r \le s \le \infty$ and $1 \le p \le q \le \infty$. For integer $n \ge 1$, we define the space $\mathcal{W}^n_{r,s;p,q}$ of functions $f \in F_{r,s}$ that are $n-1$ times continuously differentiable on the axis, their derivatives $f^{(n-1)}$ of order $n-1$ are locally absolutely continuous, and $f^{(n)} \in F_{p,q}$. As a consequence of (2.5), we have the embedding $\mathcal{W}^n_{r,s;p,q} \subset W^n_{\gamma_1,\gamma_2}$, where $1/\gamma_1 = 1/r - 1/s$ and $1/\gamma_2 = 1/p - 1/q$.

In the space $\mathcal{W}^n_{r,s;p,q}$, consider the class

$$\mathcal{Q} = \mathcal{Q}^n_{r,s;p,q} = \{f \in \mathcal{W}^n_{r,s;p,q} \colon \|f^{(n)}\|_{p,q} \le 1\}.$$

On this class, consider a variant of Stechkin's problem on the best approximation of the functional $f^{(k)}(0)$ by the ball $M_{r,s}(N)$ of radius $N > 0$ in the space of multipliers $M_{r,s}$:

$$e(N) = e_{n,k}(N) = e_{n,k}(N; r, s; p, q) = \inf \{u(\theta) \colon \theta \in M_{r,s},\ \|\theta\|_{r,s} \le N\}, \tag{2.9}$$

where

$$u(\theta) = u_{n,k}(\theta) = \sup \{|f^{(k)}(0) - \langle \theta, f\rangle| \colon f \in \mathcal{Q}^n_{r,s;p,q}\}$$

is the deviation of a functional $\theta \in M_{r,s}$ from the functional $f^{(k)}(0)$ on the class $\mathcal{Q}$.

Problem (2.9) is associated with a multiplicative inequality of Kolmogorov type, but in the predual spaces:

$$\|f^{(k)}\|_C \le B_{n,k} \|f\|^{\alpha}_{r,s} \|f^{(n)}\|^{\beta}_{p,q}, \quad f \in \mathcal{W}^n_{r,s;p,q}, \tag{2.10}$$

$$\alpha = \frac{n - k + 1/q - 1/p}{n + 1/q - 1/p + 1/r - 1/s}, \quad \beta = 1 - \alpha = \frac{k + 1/r - 1/s}{n + 1/q - 1/p + 1/r - 1/s};$$

we assume that here $B_{n,k} = B_{n,k}(r, s; p, q)$ is the best (the smallest possible) constant (independent of the function f).

The following statement is contained in the author's paper [8, Theorem 3]; however, this result was preceded by several years of research by the author, see [4–6, 8] and [13, 14].

Theorem 1. *If $s \ge r \ge 1$, $q \ge p > 1$, and conditions (1.4) hold, then the following equality holds for any $N > 0$ for the values of problems (1.2) and (2.9) and the best constant B in (2.10):*

$$E_{n,k}(N) = e_{n,k}(N) = \beta \alpha^{\alpha/\beta} B^{1/\beta}_{n,k} N^{-\alpha/\beta}. \tag{2.11}$$

In addition, there is an extremal multiplier in problem (2.9); the convolution with this multiplier is an extremal operator of Stechkin's problem (1.2).

3. Stechkin's problem and related problems in the case $s = r$, $q = p$

In this section, we will discuss Stechkin's problem (1.2) and the corresponding problems (2.9) and (2.10) with the following relationship between the parameters:

$$1 \le s = r \le \infty, \quad 1 \le q = p \le \infty. \tag{3.1}$$

These restrictions and restrictions (1.4) imply that $k > 0$, so from now on $0 < k < n$. Let us agree further in all situations instead of the set of parameters $r, r; p, p$ write $r; p$; so instead of $\mathcal{W}^n_{r,r;p,p}$ the notation $\mathcal{W}^n_{r;p}$ will be used.

In several cases when (3.1) holds, the exact solutions to all three problems are known; a review of the corresponding results will be given here.

For the convenience of further references, we repeat the definitions of the problem and of the quantities in problem (1.2) under restrictions (3.1):

$$E(N) = E_{n,k}(N) = E_{n,k}(N;r;p) = \inf\big\{U(T)\colon T \in \mathfrak{B}(N;L_r,L_r)\big\}, \tag{3.2}$$

$$U(T) = \sup\big\{\|f^{(k)} - Tf\|_p\colon f \in Q^n_{r,p}\big\}. \tag{3.3}$$

In this case, inequality (2.10) has the form

$$\|f^{(k)}\|_C \le B_{n,k}\|f\|^{\alpha}_{r,r}\|f^{(n)}\|^{\beta}_{p,p}, \quad f \in \mathcal{W}^n_{r;p}, \tag{3.4}$$

$$\alpha = \frac{n-k}{n}, \quad \beta = 1-\alpha = \frac{k}{n}; \tag{3.5}$$

here, $B_{n,k} = B_{n,k}(r;p)$ is the best (the least possible) constant (independent of the function f). Note that indices (3.5) in inequality (3.4) are independent of the parameters r and p.

Restrictions (3.1) contain in particular the two sets of parameters $s = r = q = p = \infty$ and $s = r = q = p = 2$, which the study of Stechkin's problem (1.2) began with. As we will see below, these two cases are, in a sense, "extreme" in set (3.1). These two cases are discussed in the subsequent two sections.

3.1. The classical variant of Stechkin's problem

Problem (1.2) was first studied by Stechkin in the uniform norm on the axis and semi-axis, see [46] and an earlier paper [45].

We will denote by $E_{n,k}(N;C)$, along with $E_{n,k}(N;\infty;\infty)$, problem (1.2) and the value of this problem in the uniform norm on the axis; more exactly, for

$$s = r = q = p = \infty.$$

As already noted above in Section 1.3 (see inequality (1.8)), Stechkin found out that the problem $E_{n,k}(N;C)$ is related to the exact inequality

$$\|f^{(k)}\|_C \le C_{n,k}\|f\|_C^{(n-k)/n}\,\|f\|^{k/n}_{L_\infty}, \quad f \in W^n_{\infty,\infty}, \tag{3.6}$$

between the norms of derivatives of differentiable functions. Namely, Stechkin showed [46] that the smallest constant $C_{n,k}$ in (3.6) gives an estimate from below of the value $E_{n,k}(N;C)$ (see (1.8)). It turned out later that this estimate is in fact an equality:

$$E_{n,k}(N;C) = k\left(\frac{C_{n,k}}{n}\right)^{n/k}\left(\frac{N}{n-k}\right)^{-(n-k)/k}, \quad N > 0. \tag{3.7}$$

This fact is a consequence of Domar's result [20] and of a more general result by Gabushin [25] on the best approximation of unbounded functionals by bounded ones.

Inequality (3.6) with a certain finite constant was obtained and used by Hardy and Littlewood in 1912 [30]. The exact inequality (3.6), i.e., the inequality with the best constant was first obtained in 1914 by Hadamard [29] for $n = 2$ and $k = 1$; and by Shilov in 1937 [18] for $n = 3, 4$ for all $1 \le k < n$ and for $n = 5$ and $k = 2$. In 1939, Kolmogorov found [35] the exact constant in inequality (3.6) for all $1 \le k < n$ using an elegant comparison theorem. Kolmogorov's result is very striking and important in this topic; in this regard, inequality (3.6) and more general inequalities (1.6) on the axis and semi-axis are often called *Kolmogorov inequalities.*

The Favard–Akhiezer–Krein function

$$f_n(t) = \frac{4}{\pi} \sum_{\ell=0}^{\infty} \frac{\sin\left((2\ell+1)t - n\pi/2\right)}{(2\ell+1)^{n+1}} \tag{3.8}$$

is extremal in inequality (3.6) [35]. For the properties of this function, see, for example, [36, Ch. 5, Sect. 5.4]. The uniform norm of function (3.8) has the following value:

$$K_n = \|f_n\|_C = \frac{4}{\pi} \sum_{\ell=0}^{\infty} \frac{(-1)^{\ell(n+1)}}{(2\ell+1)^{n+1}}.$$

For all $1 \le k \le n$, we have the relation $f_n^{(k)} = f_{n-k}$; in particular, $f_n^{(n)}(t) = f_0(t) = \operatorname{sign} \sin t$. The extremal function (3.8) in inequality (3.6) and its properties listed imply the following formula for the best constant in (3.6):

$$C_{n,k} = K_{n-k}\,(K_n)^{-(n-k)/n}; \tag{3.9}$$

for this value, the estimates $1 < C_{n,k} < \pi/2$ hold [35, (3)].

Stechkin proved [45, 46] that the following classical (difference) operators $T_h^{n,k}$ are extremal in the problem $E_{n,k}(N;C)$ for $n = 2$ and 3 and $1 \le k < n$:

$$(T_{2,1}^h f)(t) = (T_{3,1}^h f)(t) = \frac{f(t+h) - f(t-h)}{2h}, \quad N = h^{-1}, \tag{3.10}$$

$$(T_{3,2}^h f)(t) = \frac{f(t+h) - 2f(t) + f(t-h)}{h^2}, \quad N = \frac{4}{h^2}.$$

For $n = 4$ and 5, the solution to this case of problem (1.2) was found (1967) by Arestov [1], and for an arbitrary $n \ge 6$ by Buslaev [19]. For $n \ge 4$, the extremal operators are infinite difference operators with uniform nodes. More precisely, for example, for $k = 1$, the extremal operator has the form

$$T_{n,1} f(t) = h^{-1} \sum_{\ell=0}^{\infty} \alpha_\ell (f(t + (2\ell+1)h) - f(t - (2\ell+1)h)).$$

The sequence $\{\alpha_\ell\}_{\ell \ge 0}$ is the sum of several geometric progressions. To prove the results, we used the lower estimate (3.7) and the exact Kolmogorov inequality (3.6).

According to the results of Stechkin [46], Arestov [1], and Buslaev [19], in the classical version of Stechkin's problem $E_{n,k}(N;C)$, there is an extremal operator $T_{n,k}^* = T_{n,k}^*(N)$, which is a finite difference operator for $n = 2$ and 3 and infinite difference with a uniform step for $n \ge 4$. The norm of this operator in the space C and the deviation value (3.3) have the following extremal values:

$$\|T_{n,k}^*\|_{C \to C} = N; \quad U_{n,k}(T_{n,k}^*; C) = E_{n,k}(N;C).$$

The operator $T_{n,k}^*$ is bounded linear in the spaces L_r for all $1 \le r < \infty$, and

$$\|T_{n,k}^*\|_{L_r \to L_r} \le N.$$

Let us discuss the corresponding inequality (3.4). According to (2.6), the space $\mathcal{W}_{\infty;\infty}^n = \mathcal{W}_{\infty,\infty;\infty,\infty}^n$ consists of functions $f \in C_0$ continuously differentiable n times on the axis, for which $f^{(n)} \in C_0$. Inequality (3.4) in this case coincides with inequality (3.6) on a narrower space $\mathcal{W}_{\infty;\infty}^n$; inequality (3.6) with constant (3.9) remains exact on $\mathcal{W}_{\infty;\infty}^n$.

3.2. Approximation of the differentiation operator in the space L_2 and related problems

3.2.1. Approximation of the differentiation operator in the space L_2

A version of the Stechkin problem on the best approximation of the differentiation operator in the space $L_2(-\infty,\infty)$ (i.e., problem (3.2) for $s=r=q=p=2$) was solved by Subbotin and Taikov [47] back in 1968. They proved the following formula for the best approximation value $E_{n,k}(N;L_2)$:

$$E_{n,k}(N_{n,k}(h);L_2)=\frac{k}{n}h^{n-k},\quad N_{n,k}(h)=\frac{n-k}{n}h^{-k},\quad h>0. \tag{3.11}$$

The extremal operator they constructed will be discussed below. The proof of (3.11) used Stechkin's lower estimate (1.8). The corresponding exact inequality (1.6) in this case has the form

$$\|f^{(k)}\|_{L_2}<\|f\|_{L_2}^{(n-k)/n}\|f^{(n)}\|_{L_2}^{k/n},\quad f\in W_{2,2}^n,\quad f\not\equiv 0. \tag{3.12}$$

A proof of inequality (3.12) for $n=2$ and $k=1$ see in [31, Ch. VII, Theorem 261]; the general case is proved similarly.

To prove (3.11), Subbotin and Taikov [47] constructed an extremal operator $T_{n,k}^h$, $h>0$. This operator is a convolution:

$$\widehat{T_{n,k}^h f}=\lambda\cdot\widehat{f},\quad f\in L^2,$$

in which the multiplier $\lambda=\lambda_h$ is defined by the formulas

$$\begin{gathered}\lambda(\eta)=i^k\left((2\pi\eta)^k-\frac{k}{n}h^{n-k}(2\pi\eta)^n\operatorname{sign}\eta^{n-k}\right),\quad |\eta|\le\frac{1}{2\pi h}\left(\frac{n}{k}\right)^{1/(n-k)},\\ \lambda(\eta)=0,\quad |\eta|>\frac{1}{2\pi h}\left(\frac{n}{k}\right)^{1/(n-k)}.\end{gathered} \tag{3.13}$$

Note that function (3.13) differs from the multiplier of [47] by a change of a variable; this is because the definition of the Fourier transform adopted here differs from that used in [47] by a factor of -2π in the exponent.

3.2.2. The space $\mathcal{W}_{2;2}^n$

Before considering inequality (2.10) and problem (2.9) in the case $s=r=q=p=2$, we discuss the properties of functions from the space $\mathcal{W}_{2;2}^n$.

Lemma 2. *The space $\mathcal{W}_{2;2}^n$ consists of functions $f\in C_0$ that can be represented in the form*

$$f(t)=\check{x}(t)=\int e^{2\pi t\eta i}x(\eta)\,d\eta, \tag{3.14}$$

where the function $x=\widehat{f}$ belongs to L and has the property

$$y(\eta)=(2\pi\eta i)^n x(\eta)\in L. \tag{3.15}$$

Moreover,

$$f^{(n)}(t)=\check{y}(t)=\int e^{2\pi t\eta i}(2\pi\eta i)^n x(\eta)\,d\eta.$$

P r o o f. The space $\mathcal{W}^n_{2;2}$ is formed by functions $f \in F_{2,2}$ such that $f^{(n)} \in F_{2,2}$. The derivative $f^{(n)}$ is understood in the sense of the theory of generalized functions, see, for example, [42, 44]. Namely, for a pair of functions $f, g \in \mathscr{L} = \mathscr{L}(\mathbb{R})$, it is assumed that $g = f^{(n)}$ if the following equality holds for all functions $\phi \in \mathscr{S}$:

$$\int g(t)\phi(t)dt = (-1)^n \int f(t)\phi^{(n)}(t)dt. \tag{3.16}$$

According to (2.7) and (1.1), a function $f \in F_{2,2}$ has the form (3.14). Its derivative $g = f^{(n)}$ has a similar form:

$$g\,(t) = \int e^{2\pi t\eta i} y(\eta)\, d\eta, \quad y \in L. \tag{3.17}$$

Substituting representations (3.14) and (3.17) into (3.16), we obtain

$$\int \phi(t) \int e^{2\pi t\eta i} y(\eta)\, d\eta dt = (-1)^n \int \phi^{(n)}(t) \int e^{2\pi t\eta i} x(\eta)\, d\eta dt.$$

We may change the orders of integration on both sides of this relation:

$$\int y(\eta) \int e^{2\pi t\eta i} \phi(t)\, dt\, d\eta = (-1)^n \int x(\eta) \int e^{2\pi t\eta i} \phi^{(n)}(t)\, dt\, d\eta. \tag{3.18}$$

Let us introduce the notation

$$\psi(\eta) = \check{\phi}(\eta) = \int e^{2\pi t\eta i} \phi(t)\, dt. \tag{3.19}$$

Together with the function ϕ, the function ψ also belongs to the space $\mathscr{S}$. Relation (3.19) implies that

$$\phi(\eta) = \widehat{\psi}(\eta) = \int e^{-2\pi t\eta i} \psi(t)\, dt. \tag{3.20}$$

Differentiate relation (3.20) n times:

$$\phi^{(n)}(\eta) = \int e^{-2\pi t\eta i} (-2\pi t i)^n \psi(t)\, dt.$$

Hence, we conclude that

$$\widecheck{\phi^{(n)}}(\eta) = \int e^{2\pi t\eta i} \phi^{(n)}(t)\, dt = (-2\pi\eta i)^n \psi(\eta). \tag{3.21}$$

Substituting (3.21) and (3.19) into (3.18), we obtain

$$\int y(\eta)\psi(\eta)\, d\eta = (-1)^n \int x(\eta)(-2\pi\eta i)^n \psi(\eta)\, d\eta$$

and

$$\int \left(y(\eta) - (2\pi\eta i)^n x(\eta)\right) \psi(\eta)\, d\eta = 0, \quad \psi \in \mathscr{S}.$$

The Fourier transform, and therefore the inverse Fourier transform (3.19), is a bijection of $\mathscr{S}$ onto itself, and therefore ψ in the last relation is an arbitrary function from $\mathscr{S}$. Hence,

$$y(\eta) - (2\pi\eta i)^n x(\eta) = 0, \qquad \text{a.e. on the axis.}$$

Property (3.15) is justified. Lemma 2 is proved. □

Consider now the corresponding inequality (3.4). It is convenient to study it in terms of Fourier transforms of functions $f \in \mathcal{W}^n_{2;2}$. Let us introduce the notation

$$Y^n = \widehat{\mathcal{W}^n_{2;2}} = \{x = \widehat{f} : f \in \mathcal{W}^n_{2;2}\} = \{x \in L\colon (2\pi ti)^n x \in L\}.$$

In terms of functions from the space Y^n, inequality (3.4) takes the following form in this case:

$$\|\check{x}^{(k)}\|_C \le B_{n,k}\, \|x\|_L^{(n-k)/n}\, \|(2\pi ti)^n x\|_L^{k/n}, \quad x \in Y^n. \tag{3.22}$$

Obviously, for every function $x \in Y^n$, the function $|x|$ also belongs to the space Y^n, and the function

$$\check{x}^{(k)}(t) = \int e^{2\pi t\eta i}(2\pi\eta i)^k x(\eta)\, d\eta$$

satisfies the relations

$$\|\check{x}^{(k)}\|_C \le \|\,\widecheck{|x|}^{(k)}\|_C = |\,\widecheck{|x|}^{(k)}(0)\,| = \|(2\pi ti)^k x\|_L.$$

Therefore, inequality (3.22) is equivalent to the inequality

$$\|(2\pi ti)^k x\|_L \le B_{n,k}\, \|x\|_L^{(n-k)/n}\, \|(2\pi ti)^n x\|_L^{k/n}, \quad x \in Y^n, \tag{3.23}$$

(with the same value of the best constant $B_{n,k}$).

3.2.3. Stechkin's problem in the space of multipliers $M_{2,2}$ and the corresponding inequality in the predual space

Consider now the corresponding variant of Stechkin's problem (2.9) on the best approximation of the functional

$$\check{x}^{(k)}(0) = \int (2\pi\eta i)^k x(\eta) d\eta$$

by the space of multipliers $M_{2,2}$. The class $Q^n_{2;2} \subset \mathcal{W}^n_{2;2}$ is correspond in Y^n to the class of functions

$$\Theta^n_2 = \widehat{Q^n_{2;2}} = \big\{x \in L\colon (2\pi ti)^n x \in L,\ \|(2\pi ti)^n x\|_L \le 1\big\}.$$

As a result, we have the problem

$$e_{n,k}(N) = e_{n,k}(N;2;2) = \inf\big\{u(\theta)\colon \theta \in L_\infty,\ \|\theta\|_{L_\infty} \le N\big\}, \tag{3.24}$$

where

$$u(\lambda) = u_{n,k}(\lambda) = \sup\Big\{\Big|\int (2\pi\eta i)^k x(\eta) d\eta - \int \lambda(\eta)x(\eta)\, d\eta\Big| \colon x \in \Theta^n_2\Big\}.$$

The best upper estimate for value (3.24) is given by multiplier (3.13). The relevant properties of this multiplier are summarized in the following lemma; all of them are available in [47].

Lemma 3. *The following two statements are valid for function* (3.13).

(1) *Function* (3.13) *is continuous and bounded on the axis, and*

$$\|\lambda\|_{C(-\infty,\infty)} = \big|\lambda(\pm(2\pi h)^{-1})\,\big| = h^{-k}\frac{n-k}{n}. \tag{3.25}$$

(2) *The function*

$$\Delta(\eta) = \frac{(2\pi\eta i)^k - \lambda(\eta)}{(2\pi\eta i)^n} \tag{3.26}$$

belongs to the space $L_\infty(-\infty,\infty)$, *and*

$$\|\Delta\|_{L_\infty(-\infty,\infty)} = \frac{k}{n} h^{n-k}. \tag{3.27}$$

P r o o f. The continuity, boundedness, and property (3.25) for function (3.13) are rather evident.

Let us now study function (3.26). For

$$0 < |\eta| \leq \frac{1}{2\pi h}\left(\frac{n}{k}\right)^{1/(n-k)},$$

we have

$$\Delta(\eta) = i^k \frac{k}{n}\frac{h^{n-k}(2\pi\eta)^n \operatorname{sign}\eta^{n-k}}{(2\pi\eta i)^n} = \frac{k}{n}h^{n-k}\operatorname{sign}\eta^{n-k} i^{k-n}.$$

In the case when

$$|\eta| \geq \frac{1}{2\pi h}\left(\frac{n}{k}\right)^{1/(n-k)},$$

we have

$$\Delta(\eta) = \frac{(2\pi\eta i)^k}{(2\pi\eta i)^n} = \frac{1}{(2\pi\eta i)^{n-k}};$$

hence,

$$|\Delta(\eta)| \leq \frac{k}{n}h^{n-k}, \quad |\eta| \geq \frac{1}{2\pi h}\left(\frac{n}{k}\right)^{1/(n-k)}.$$

This implies property (3.27) of function (3.26). Lemma 3 is proved. □

The following statement is contained in equality (3.11), inequality (3.12), Lemma 2, and Theorem 1. However, its proof will be given here. This proof largely repeats that of statement (3.11) in [47].

Theorem 2. *The following statements are valid for value* (3.24) *and the best constant* $B_{n,k}$ *in inequality* (3.23) *for* $0 < k < n$.

(1) *For all* $h > 0$,

$$e_{n,k}(N_{n,k}(h)) = \frac{k}{n}h^{n-k}, \quad N_{n,k}(h) = h^{-k}\frac{n-k}{n}, \quad h > 0; \tag{3.28}$$

and functional (3.13) *is extremal.*

(2) *The best constant in inequality* (3.23) *is one*:

$$B_{n,k} = 1. \tag{3.29}$$

P r o o f. (1) First, we obtain an upper estimate for the value $e_{n,k}(N)$. To do this, we use multiplier (3.13). Relations (3.25) and (3.27) imply the following upper estimate for $e_{n,k}(N)$:

$$e_{n,k}(N_{n,k}(h)) \leq \frac{k}{n}h^{n-k}, \quad N_{n,k}(h) = h^{-k}\frac{n-k}{n}, \quad h > 0. \tag{3.30}$$

(2) Let us now obtain a lower estimate for the best constant $B_{n,k}$ in inequality (3.23). We start with the function

$$f(t) = e^{2\pi t i} = \int e^{2\pi t\eta i}\, d\mu(\eta);$$

here μ is the measure on the axis, which can be written as $d\mu(\eta) = \delta(\eta - 1)\, d\eta$, where δ is the Dirac δ-function. For $\rho > 0$, we define a function x_ρ on the axis by the relation

$$x_\rho(\eta) = \begin{cases} \dfrac{1}{\rho}, & \eta \in [1, 1+\rho]\,; \\ 0, & \eta \notin [1, 1+\rho]\,. \end{cases}$$

For this function, we have $\|x_\rho\|_L = 1$ and

$$\|(2\pi ti)^k x_\rho\|_L \to (2\pi)^k, \quad \|(2\pi ti)^n x_\rho\|_L \to (2\pi)^n \quad \text{as } \rho \to +0.$$

Substitute the function x_ρ into inequality (3.23) and let $\rho \to +0$. As a result,

$$B_{n,k} \geq \frac{\|(2\pi ti)^k x_\rho\|_L}{\|x_\rho\|_L^{(n-k)/n} \|(2\pi ti)^n x_\rho\|_L^{k/n}} \to 1.$$

Thus, the following lower estimate holds for the best constant $B_{n,k}$ in inequality (3.23):

$$B_{n,k} \geq 1. \tag{3.31}$$

(3) Statement (2.11) and estimates (3.30) and (3.31) imply equalities (3.28) and (3.29). Theorem 2 is proved. □

Inequality (3.23) is an inequality of the Carlson type; the studies of V.I. Levin, F.I. Andrianov, and others were devoted to such inequalities in the middle of the last century, see [38], [31, Levin V.I., Stechkin S.B. Additions to the Russian edition], [3], and the references therein. For statements like Theorem 2 related to Carlson's inequalities, see [3].

In the previous two Sections 3.1 and 3.2, Stechkin's lower estimate (1.8) for the value of the best approximation of the differentiation operator in terms of the best constant in the corresponding Kolmogorov inequality was applied in the study of Stechkin's problem. At the time of studying Stechkin's problem, the exact constant in the corresponding inequalities (3.6) and (3.12) was known; moreover, inequality (1.8) gave an exact estimate for the value of the best approximation. In the next two Sections 3.3 and 3.4, Stechkin's problem will be discussed in situations where there is no corresponding inequality (1.8). A lower estimate for the best approximation will be based on the considerations of the translation invariance of Stechkin's problem; more precisely, the statements of Theorem 1 will be used.

3.3. Approximation in the uniform norm on the axis by operators bounded in the space L_r: the case $1 \leq s = r \leq \infty$, $p = q = \infty$

Here, we discuss Stechkin's problem (3.2) for values of the parameters

$$1 \leq s = r \leq \infty, \quad p = q = \infty. \tag{3.32}$$

For real r, $1 \leq r \leq \infty$, and integer $n \geq 1$, the space $W^n_{r,\infty}$ consists of functions $f \in L_r$ that are $n-1$ times continuously differentiable on the axis, their derivatives $f^{(n-1)}$ of order $n-1$ are locally absolutely continuous, and $f^{(n)} \in L_\infty$. In the space $W^n_{r,\infty}$, consider the class

$$Q^n_{r,\infty} = \{f \in W^n_{r,\infty} \colon \|f^{(n)}\|_{L_\infty} \leq 1\}.$$

Denote by $\mathfrak{B}(L_r)$ the set of all bounded linear operators in the space L_r. Let $\mathfrak{B}(N; L_r)$ for $N > 0$ be the set of operators $T \in \mathfrak{B}(L_r)$ with the norm $\|T\|_{L_r \to L_r} \leq N$. In this section, for $r = \infty$, we mean by L_∞ the space $C = C(-\infty, \infty)$.

We are interested in the best approximation (in the space C) of the differentiation operator $D^k = d^k/dt^k$ on the class $Q^n_{r,\infty}$ by the set of bounded linear operators $\mathfrak{B}(N; L_r)$:

$$E_{n,k}(N) = E_{n,k}(N; r; \infty) = \inf\{U(T) \colon T \in \mathfrak{B}(N; L_r)\}, \quad N > 0, \tag{3.33}$$

$$U(T) = U_{n,k}(T; r; \infty) = \sup\{\|f^{(k)} - Tf\|_C \colon f \in Q^n_{r,\infty}\}.$$

The problem is to calculate value (3.33) and an extremal operator on which the infimum in (3.33) is attained.

3.3.1. Case $n \geq 3$, $1 \leq r \leq \infty$

Recall that

$$K_n = \|f_n\|_C = \frac{4}{\pi} \sum_{\ell=0}^{\infty} \frac{(-1)^{\ell(n+1)}}{(2\ell+1)^{n+1}} \tag{3.34}$$

is the uniform norm of the Favard–Akhiezer–Krein function (3.8). Define

$$\overline{K}_n = \frac{4}{\pi} \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)^{n+1}}. \tag{3.35}$$

This is, in a sense, the "norm" of the same function f_n in the space $F_{2,2}$. Comparing (3.35) with (3.34), we see that $K_n \leq \overline{K}_n$; more exactly,

$$K_n = \overline{K}_n \text{ if } n \text{ is odd}; \quad K_n < \overline{K}_n \text{ if } n \text{ is even}.$$

The following two statements are valid [13] for the problem $E_{n,k}(N; L_r)$.

Theorem 3. *The following two-sided estimates for the value of problem* (3.33) *hold for all* $n \geq 2$, $1 \leq k < n$, *and* $1 \leq r \leq \infty$:

$$k \left(\frac{K_{n-k}}{n}\right)^{n/k} \left(\frac{N\,\overline{K}_n}{n-k}\right)^{-(n-k)/k} \leq E_{n,k}(N; r; \infty) \leq k \left(\frac{K_{n-k}}{n}\right)^{n/k} \left(\frac{N\,K_n}{n-k}\right)^{-(n-k)/k}. \tag{3.36}$$

Theorem 4. *The following statements hold in problem* (3.33) *for odd* $n \geq 3$ *and arbitrary* k, $1 \leq k < n$.

(1) *The following formula holds for value* (3.33) *independently of* r, $1 \leq r \leq \infty$:

$$E_{n,k}(N; r; \infty) = E_{n,k}(N, C) = k \left(\frac{K_{n-k}}{n}\right)^{n/k} \left(\frac{N\,K_n}{n-k}\right)^{-(n-k)/k}.$$

(2) *An operator* $T^*_{n,k} = T^*_{n,k}(N)$ *that is extremal in the problem* $E_{n,k}(N, C)$ *is also extremal in the problem* $E_{n,k}(N, L_r)$ *for all* r, $1 \leq r < \infty$.

3.3.2. Case $n = 2$, $r = 2$

For even $n \geq 2$ and $1 < r < \infty$, the statements of Theorem 4, generally speaking, no longer hold. The author's paper [11] provides a solution to problem (3.33) for

$$n = 2 \quad (k = 1); \quad r = s = 2, \quad p = q = \infty. \tag{3.37}$$

In this case, the first inequality in (3.36) is exact. More precisely, the following statement is true.

Theorem 5. *The following formula holds for values of the parameters* (3.37) *for all* $h > 0$:

$$E_{2,1}(N_{2,1}(h); 2; \infty) = \frac{\pi h}{4}, \quad N_{2,1}(h) = \frac{\pi^2}{2h} \left(4 \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)^3}\right)^{-1}. \tag{3.38}$$

An extremal operator in (3.38) *is the singular convolution operator on the space* L_2 *defined by the formula*

$$(\Theta_h f)(t) = A(h) \int_0^{\pi h} (f(t+u) - f(t-u))\, y\left(uh^{-1}\right) du,$$

where

$$y(u) = \frac{\pi - u}{4 \sin u}, \quad u \in (0, \pi); \quad A(h) = h^{-2} \left(4 \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)^3}\right)^{-1}.$$

For comparison, consider the result of Stechkin [46] for $n=2$, $k=1$, and $(q=p=)s=r=\infty$:

$$E_{2,1}(N_{2,1}(h);\infty;\infty)=\frac{\pi h}{4},\quad N_{2,1}(h)=h^{-1}.$$

An extremal operator is the difference operator (3.10):

$$(T^h_{2,1}f)(t)=\frac{f(t+h)-f(t-h)}{2h}.$$

3.3.3. Inequalities for values of the parameters (3.32) in predual spaces

In the case $1\le s=r\le\infty$ and $p=q=\infty$ under consideration, inequality (2.10) has the form

$$\|f^{(k)}\|_C\le B_{n,k}(r;\infty)\|f\|^{(n-k)/n}_{r,r}\big(\|f^{(n)}\|_\infty\big)^{k/n},\quad f\in\mathcal{W}^n_{r;\infty}. \tag{3.39}$$

For $s=r=\infty$, this is the classical variant (3.6) of the inequality between the uniform norms of derivatives studied by Kolmogorov. In the case $s=r=2$, inequality (3.39) takes the form

$$\|f^{(k)}\|_C\le B_{n,k}(2;\infty)\|\widehat{f}\|^{(n-k)/n}_{1}\big(\|f^{(n)}\|_\infty\big)^{k/n},\quad f\in\mathcal{W}^n_{2;\infty}. \tag{3.40}$$

The following inequality holds [11, 13] for the best constants in (3.39) and, in particular, in (3.40):

$$B_{n,k}(r;\infty)\le B_{n,k}(\infty;\infty)=C_{n,k},\quad 1\le r\le\infty; \tag{3.41}$$

recall that $C_{n,k}$ was defined in (3.9). For odd $n\ge 3$, we have the equality $B_{n,k}(r;\infty)=B_{n,k}(\infty)$, $1\le r\le\infty$, and the Favard–Akhiezer–Krein function f_n (3.8) is extremal for all r.

For even $n\ge 2$, this is, generally speaking, no longer the case. At least for $n=2$ $(k=1)$ and $r=2$, the best constant in inequality (3.40) has the following value [11]:

$$B_{2,1}(2;\infty)=\frac{\pi}{2}\left(\frac{4}{\pi}\sum_{\ell=0}^{\infty}\frac{1}{(2\ell+1)^3}\right)^{-1/2}, \tag{3.42}$$

and the Favard–Akhiezer–Krein function f_2 is extremal again. The following estimates hold for constant (3.42):

$$\sqrt{\frac{\pi}{2}}<B_{2,1}(2;\infty)<\sqrt{2} \tag{3.43}$$

(see details in [11]). According to Hadamard's result [29], the best constant in inequality (3.6) for $n=2$ and $k=1$ is $C_{2,1}=\sqrt{2}$. Consequently, the second inequality (3.43) means that $B_{2,1}(2;\infty)<B_{2,1}(\infty;\infty)=C_{2,1}$, so that inequality (3.41) is strict in this case.

Inequalities of type (3.40) containing the norms of intermediate and highest derivatives and the norm of the Fourier transform of functions, with norm parameters different from (3.40), also arose in the studies by Magaril-Il'yaev and Osipenko of extremal problems of recovering functions from information about their spectrum [39, 40].

3.4. Case $1\le s=r\le\infty$, $p=q=2$

Here we will discuss Stechkin's problem (3.2) studied in [5, 7, 10] for the parameter values

$$1\le s=r\le\infty,\quad p=q=2. \tag{3.44}$$

For real r, $1\le r\le\infty$, and integer $n\ge 1$, the space $W^n_{r,2}$ consists of functions $f\in L_r$ that are $n-1$ times continuously differentiable on the axis, their derivatives $f^{(n-1)}$ of order $n-1$ are locally absolutely continuous, and $f^{(n)} \in L_2$. Here, for $r = \infty$, we mean by L_∞ the space $C = C(-\infty, \infty)$. In the space $W^n_{r,2}$, consider the class $Q^n_{r,2} = \{f \in W^n_{r,2} \colon \|f^{(n)}\|_{L_2} \le 1\}$. As was said above, $\mathfrak{B}(L_r)$ denotes the set of all bounded linear operators in the space L_r, and $\mathfrak{B}(N; L_r)$ for $N > 0$ is the set of operators $T \in \mathfrak{B}(L_r)$ with the norm $\|T\|_{L_r \to L_r} \le N$.

We are interested in the best approximation in the space L_2 of the differentiation operator D^k on the class $Q^n_{r,2}$ by the set of bounded linear operators $\mathfrak{B}(N; L_r)$:

$$E_{n,k}(N) = E_{n,k}(N; r; 2) = \inf\left\{U_{n,k}(T; r; 2) \colon T \in \mathfrak{B}(N; L_r)\right\}, \quad N > 0,$$

$$U(T) = U_{n,k}(T; r; 2) = \sup\left\{\|f^{(k)} - Tf\|_{L_2} \colon f \in Q^n_{r,2}\right\}.$$

3.4.1. Case $n \ge 3,\ 1 \le r \le \infty$ [5]

Theorem 6. *For $n \ge 3,\ 1 \le k < n,\ 1 \le r \le \infty$, and all $h > 0$,*

$$E_{n,k}(N_{n,k}(h); r; 2) = \frac{k}{n} h^{n-k}, \tag{3.45}$$

where

$$N_{n,k}(h) = \frac{n-k}{n} h^{-k}. \tag{3.46}$$

For $r = 2$, statement (3.45)+(3.46) is statement (3.11) of Subbotin and Taikov [47]. To justify (3.45)+(3.46), the author used in [5] an operator that differs from the one in [47]; for more detailed discussion see Section 3.4.3.

3.4.2. Cases $n = 2,\ k = 1,\ r = \infty$ [7, 10], and $r = 2$ [47]

Theorem 7. *For all $h > 0$,*

$$E_{2,1}(N_{2,1}(h); r; 2) = \frac{1}{2} h,$$

where

$$N_{2,1}(h) = \frac{1}{2} h^{-1}$$

for $r = 2$ and

$$N_{2,1}(h) = \frac{16}{h\pi^3} \sum_{l=0}^{\infty} \frac{1}{(2l+1)^3}$$

for $r = \infty$.

3.4.3. Extremal operators

Case $n \ge 3$ and $k = 1$. Let us describe the construction of an extremal operator [5]. Let η be the 2π-periodic odd function defined on $[0, \pi]$ by the relations

$$\eta(t) = \begin{cases} t - \dfrac{1}{n}\left(\dfrac{2}{\pi}\right)^{n-1} t^n, & t \in \left[0, \dfrac{\pi}{2}\right], \\[2ex] \eta(\pi - t), & t \in \left[\dfrac{\pi}{2}, \pi\right]. \end{cases}$$

The Fourier series of this function has the form

$$\eta(t) = \sum_{l=0}^{\infty} c_l \sin(2l+1)t, \quad c_\ell = \frac{4}{\pi} \int_0^{\pi/2} \eta(t) \sin(2l+1)t dt.$$

The coefficients of this expansion for $n \geq 3$ have the following signs (see [5, proof of Theorem 4.1]):

$$(-1)^l c_l \geq 0, \quad l \geq 0. \tag{3.47}$$

For a number $h > 0$, we set $\nu = \nu(h) = \pi h/2$ and define an operator $T_{n,1}$ by the formula

$$(T_{n,1}f)(t) = \frac{1}{2\nu(h)} \sum_{l=0}^{\infty} c_l \big\{ f(t + (2l+1)\nu) - f(t - (2l+1)\nu) \big\}.$$

It is clear that $T_{n,1}$ is a bounded linear operator in the space L_r for all $1 \leq r \leq \infty$ and

$$\|T_{n,1}\|_{L_r \to L_r} = \frac{1}{\nu} \sum_{l=0}^{\infty} |c_l| = \frac{1}{\nu} \eta\left(\frac{\pi}{2}\right) = \frac{n-1}{nh}.$$

For this operator,

$$U_{n,k}(T_{n,1}; r; 2) = \frac{k}{n} h^{n-k}. \tag{3.48}$$

It is this operator that is extremal in (3.45) for $n \geq 3$, $k = 1$, and all $1 \leq r \leq \infty$. It is different from the operator constructed by Subbotin and Taikov [47] for $r = 2$.

Case $n = 2$, $k = 1$, and $r = \infty$. For $n = 2$, the property of signs (3.47) is violated. More precisely, we have

$$\eta(t) = \sum_{l=0}^{\infty} c_l \sin(2l+1)t, \quad c_l = \frac{8}{\pi^2} \frac{1}{(2l+1)^3}.$$

The operator $T_{2,1}$ defined by the formula

$$(T_{2,1}f)(t) = \frac{1}{2\nu(h)} \sum_{l=0}^{\infty} c_l \big\{ f(t + (2l+1)\nu(h)) - f(t - (2l+1)\nu(h)) \big\},$$

where $\nu = \nu(h) = \pi h/2$, is a bounded linear operator in C and

$$\|T_{2,1}\|_{C \to C} = \frac{1}{\nu} \sum_{l=0}^{\infty} c_l = \frac{16}{\pi^3 h} \sum_{l=0}^{\infty} \frac{1}{(2l+1)^3} = N_{2,1}(h).$$

The norm of the operator $T_{2,1}$ has a different expression in comparison with (3.46). The same formula (3.48) holds for the value of the deviation.

Paper [47] by Subbotin and Taikov contains the case $n = 2$, $k = 1$, and $r = 2$ as a special case.

3.4.4. Inequalities for cases (3.44) in predual spaces

Let us discuss now inequality (2.10) for the set of parameters (3.44) in the space

$$\mathcal{W}^n_{r;2} = \big\{ f \in F_{r,r} \colon \widehat{f^{(n)}} \in L \big\} = \big\{ f \in F_{r,r} \colon f^{(n)} = \check{z},\ z \in L \big\}, \tag{3.49}$$

and, in particular, in the space

$$\mathcal{W}^n_{\infty;2} = \big\{ f \in C_0 \colon \widehat{f^{(n)}} \in L \big\} = \big\{ f \in C_0 \colon f^{(n)} = \check{z},\ z \in L \big\}. \tag{3.50}$$

Theorem 8 [5]. *The following inequality holds for functions of space* (3.49) *for* $1 \le r \le \infty$, $n \ge 3$, *and* $1 \le k < n$:

$$\|f^{(k)}\|_C \le B_{n,k}(r;2)\|f\|_{r,r}^{(n-k)/n}\|\widehat{f^{(n)}}\|_L^{k/n}, \quad f \in \mathcal{W}^n_{r;2}, \tag{3.51}$$

with the smallest possible constant

$$B_{n,k}(r;2) = 1. \tag{3.52}$$

For all $n \ge 3$ *and* $1 \le k < n$, *an "ideal" extremal function is* sin.

For $n = 2$ $(k = 1)$, inequality (3.51) with constant (3.52) holds for $r = 2$ (see Theorem 9 below) and does not hold for $r = \infty$. The value of the constant $B_{2,1}(r;2)$ for other values of r is currently unknown. The following statement highlights the case $r = \infty$ of Theorem 8 and adds information about inequality (3.51) in the case $n = 2$ and $r = 2$.

Theorem 9 [7]. *The following inequality holds for functions of space* (3.50) *for* $n \ge 2$, $1 \le k < n$, *and* $r = \infty$:

$$\|f^{(k)}\|_C \le B_{n,k}(\infty;2)\|f\|_C^{(n-k)/n}\|\widehat{f^{(n)}}\|_L^{k/n}, \quad f \in \mathcal{W}^n_{\infty;2},$$

with the smallest possible constants

$$B_{2,1}(\infty;2) = \left\{\frac{32}{\pi^3}\sum_{l=0}^{\infty}\frac{1}{(2l+1)^3}\right\}^{1/2} > 1, \quad n = 2, \quad k = 1;$$

$$B_{n,k}(\infty;2) = 1, \quad n \ge 3, \quad 1 \le k < n.$$

For $n \ge 3$, *an "ideal" extremal function is* sin. *For* $n = 2$, *it is the entire function*

$$f(t) = \frac{1}{2}\int_0^\pi \frac{\pi - u}{\sin u}\sin 2\pi t u\, du.$$

4. Two-sided estimates for the value of Stechkin's problem (3.2)

For parameters $1 \le r, p \le \infty$, define $\overline{r} = \max\{r, r'\}$ and $\overline{p} = \max\{p, p'\}$. In statements of this section, we assume the following condition on two pairs of parameters r_1, r_2 and p_1, p_2:

$$\overline{r}_1 \le \overline{r}_2, \quad \overline{p}_1 \le \overline{p}_2. \tag{4.1}$$

Theorem 10. *The following two statements hold for the value* $E_{n,k}(N;r;p) = E_{n,k}(N;r,r;p,p)$ *for* $1 \le r \le \infty$, $1 < p \le \infty$, *and* $0 < k < n$.

(1) *For all* $N > 0$, *the value* $E_{n,k}(N;r;p)$ *of Stechkin's problem* (3.2) *does not decrease in the parameters* $\overline{r}$ *and* $\overline{p}$; *more exactly, if two pairs of parameters* r_1, r_2 *and* p_1, p_2 *satisfy conditions* (4.1), *then the following inequality holds*:

$$E_{n,k}(N;r_1;p_1) \le E_{n,k}(N;r_2;p_2).$$

(2) *For all* $N > 0$, *the following (exact) two-sided estimates hold for the values of Stechkin's problems* (3.2) *and* (2.9):

$$\beta\alpha^{\alpha/\beta}N^{-\alpha/\beta} \le E_{n,k}(N;r;p) = e_{n,k}(N;r;p) \le \beta\alpha^{\alpha/\beta}\left(C_{n,k}\right)^{1/\beta}N^{-\alpha/\beta},$$

where

$$\alpha = \frac{n-k}{n}, \quad \beta = \frac{k}{n}.$$

4.1. Auxiliary statement

Lemma 4. *The following two statements hold for the best constant $B_{n,k} = B_{n,k}(r;p)$ in inequality* (3.4) *for $1 \le r \le \infty$, $1 \le p \le \infty$, and $0 < k < n$.*

(1) *If two pairs of parameters r_1, r_2 and p_1, p_2 satisfy conditions* (4.1), *then the following inequality holds for the best constant in inequality* (3.4):

$$B_{n,k}(r_1;p_1) \le B_{n,k}(r_2;p_2). \tag{4.2}$$

(2) *The following (exact) two-sided estimates hold:*

$$1 \le B_{n,k}(r;p) \le C_{n,k} \left(< \frac{\pi}{2}\right). \tag{4.3}$$

P r o o f. The constant in inequality (3.4) can be represented in the form

$$B_{n,k}(r;p) = \sup\left\{ \frac{\|f^{(k)}\|_C}{\|f\|_{r,r}^{(n-k)/n} \|f^{(n)}\|_{p,p}^{k/n}} : f \in \mathcal{W}^n_{r;p},\ f \not\equiv 0 \right\}. \tag{4.4}$$

According to statement (2.8) of Lemma 1, under conditions (4.1), we have the embeddings

$$F_{r_1,r_1} \subset F_{r_2,r_2} \quad \text{and} \quad \|f\|_{r_2,r_2} \le \|f\|_{r_1,r_1}, \quad f \in F_{r_1,r_1},$$

$$F_{p_1,p_1} \subset F_{p_2,p_2} \quad \text{and} \quad \|g\|_{p_2,p_2} \le \|g\|_{p_1,p_1}, \quad g \in F_{p_1,p_1},$$

and hence the embeddings

$$\mathcal{W}^n_{r_1;p_1} \subset \mathcal{W}^n_{r_2;p_2}; \tag{4.5}$$

moreover, the following inequalities hold on $\mathcal{W}^n_{r_1;p_1}$:

$$\|f\|_{r_2,r_2} \le \|f\|_{r_1,r_1}, \quad \|f^{(n)}\|_{p_2,p_2} \le \|f^{(n)}\|_{p_1,p_1}, \quad f \in \mathcal{W}^n_{r_1;p_1}. \tag{4.6}$$

Representation (4.4), embedding (4.5), and inequality (4.6) imply property (4.2).

In particular, we have the inequalities

$$B_{n,k}(2;2) \le B_{n,k}(r;p) \le B_{n,k}(\infty;\infty).$$

According to the result (3.29) of Lemma 2, $B_{n,k}(2;2) = 1$. In the case $r = p = \infty$, the value $B_{n,k}(\infty;\infty)$ coincides with the best constant (3.9) in the Kolmogorov inequality (3.6): $B_{n,k}(\infty;\infty) = C_{n,k}$. Thus, statement (4.3) is verified. Lemma 4 is proved.

4.2. The proof of Theorem 10

Both statements of Theorem 10 follow from the corresponding statement of Lemma 4 and statement (2.11) of Theorem 1. Theorem 10 is proved.

5. Conclusions

As can be seen from the results described above, even in the case (3.1), the topics considered here are far from exhausted. One of the main reasons for the difficulties in studying Stechkin's problem is that the description of (p,q)-multipliers and the value of the norm of multipliers are known only in several exceptional cases (see Section 2.1.1).

For example, Stechkin's problem and the corresponding inequalities between the norms of derivatives in the case of equal exponents

$$1 \le s = r = q = p \le \infty$$

are of interest. Denote by $E_{n,k}(N)_p$ Stechkin's problem and its value for this case. This case is embedded in the assumptions and conclusions of Theorem 10. According to Theorem 10, the value $E_{n,k}(N)_p$ does not decrease in the parameter $\overline{p} = \{p, p'\}$ and the following estimates hold:

$$E_{n,k}(N)_2 \le E_{n,k}(N)_p \le E_{n,k}(N)_\infty.$$

The solution to Stechkin's problem $E_{n,k}(N)_p$ is known only in the cases $p = \infty$, 2, and 1. Of course, one of the reasons for this is that the description of the multipliers of the Lebesgue spaces $L^p(-\infty, \infty)$ is known only for these values of the parameter p.

Let $B_{n,k}(p)$ be the best constant in the corresponding inequality

$$\|f^{(k)}\|_p \le B_{n,k}(p)\, \|f\|_p^{(n-k)/n} \|f^{(n)}\|_p^{k/n}, \quad f \in W^n_{p,p},$$

between the p-norms of the derivatives. The following estimates are known for the best constant in this inequality:

$$B_{n,k}(2) = 1 \le B_{n,k}(p) \le B_{n,k}(\infty) = C_{n,k} < \frac{\pi}{2}. \tag{5.7}$$

The second inequality in (5.7) is Stein's result [43]. To justify the first inequality, one should substitute the (appropriately smoothed) function sin into (5.7). No results regarding monotonicity of the value $B_{n,k}(p)$ in p are unknown to the author.

Acknowledgements

The author expresses his deep gratitude to the reviewer, who thoroughly read the paper and made several helpful comments.

REFERENCES

1. Arestov V. V. On the best approximation of differentiation operators. *Math. Notes*, 1967. Vol. 1, No. 2. P. 100–103. DOI: 10.1007/BF01268057
2. Arestov V. V. On the best uniform approximation of differentiation operators. *Math. Notes*, 1969. Vol. 5, No. 3. P. 167–173. DOI: 10.1007/BF01388621
3. Arestov V. V. Approximation of linear operators, and related extremal problems. *Proc. Steklov Inst. Math.*, 1977. Vol. 138. P. 31–44.
4. Arestov V. V. Approximation of operators that are invariant under a shift. *Proc. Steklov Inst. Math.*, 1977. Vol. 138. P. 45–74.
5. Arestov V. V. Approximation of operators of convolution type by linear bounded operators. *Proc. Steklov Inst. Math.*, 1981. Vol. 145. P. 1–18.
6. Arestov V. V. Approximation of invariant operators. *Math. Notes*, 1983. Vol. 34, No. 1. P. 489–499. DOI: 10.1007/BF01160861
7. Arestov V. V. On the best approximation of the differentiation operator. In: *Priblizhenie funkcii polinomami i splainami* [Approximation of Functions by Polynomials and Splines]: collection of papers. Sverdlovsk, 1985. P. 3–14. (in Russian)
8. Arestov V. V. Best approximation of translation invariant unbounded operators by bounded linear operators. *Proc. Steklov Inst. Math.*, 1994. Vol. 198. P. 1–16.
9. Arestov V. V. Approximation of unbounded operators by bounded operators and related extremal problems. *Russian Math. Surveys*, 1996. Vol. 51, No. 6. P. 1093–1126. DOI: 10.1070/rm1996v051n06abeh003001

10. Arestov V. V. On the best approximation of the differentiation operator. *Ural Math. J.*, 2015. Vol. 1, No. 1. P. 20–29. DOI: 10.15826/umj.2015.1.002
11. Arestov V. V. Best uniform approximation of the differentiation operator by operators bounded in the space L_2. *Proc. Steklov Inst. Math.*, 2020. Vol. 308, Suppl. 1. P. S9–S30. DOI: 10.1134/S0081543820020029
12. Arestov V. V. On the conjugacy of the space of multipliers. *Trudy Inst. Mat. Mekh. UrO RAN*, 2019. Vol. 25, No. 4. P. 5–14. DOI: 10.21538/0134-4889-2019-25-4-5-14 (in Russian)
13. Arestov V. V. Uniform approximation of differentiation operators by bounded linear operators in the space L_r. *Anal. Math.*, 2020. Vol. 46. P. 425–445. DOI: 10.1007/s10476-020-0040-z
14. Arestov V. V. Predual spaces for the space of (p, q)-multipliers and their application in Stechkin's problem on approximation of differentiation operators. *Anal. Math.*, 2023. Vol. 49, No. 1. P. 43–65. DOI: 10.1007/s10476-022-0184-0
15. Arestov V. V., Akopyan R. R. Stechkin's problem on the best approximation of an unbounded operator by bounded ones and related problems. *Trudy Inst. Mat. Mekh. UrO RAN*, 2020. Vol. 26, No. 4. P. 7–31. DOI: 10.21538/0134-4889-2020-26-4-7-31 (in Russian)
16. Arestov V. V., Gabushin V. N. Best approximation of unbounded operators by bounded operators. *Russian Math.* (*Iz. VUZ*), 1995. Vol. 39, No. 11. P. 38–63.
17. Babenko V. F., Korneichuk N. P., Kofanov V. A., Pichugov S. A. *Neravenstva dlya proizvodnykh i ikh prilozheniya* [Inequalities for Derivatives and Their Applications]. Kiev: Naukova Dumka, 2003. 591 p. (in Russian)
18. Bosse Yu. G. (Shilov G. E.) On inequalities between derivatives. In: *Collection of Works of Student Research Societies of Moscow State University*. Moscow: Izd. MGU, 1937. Vol. 1. P. 68–72. (in Russian)
19. Buslaev A. P. Approximation of a differentiation operator. *Math. Notes*, 1981. Vol. 29, No. 5. P. 372–378. DOI: 10.1007/BF01158361
20. Domar Y. An extremal problem related to Kolmogoroff's inequality for bounded functions. *Arkiv för Mat.*, 1968. Vol. 7, No. 5. P. 433–441.
21. Dunford N., Schwartz J. T. *Linear Operators. Part 1: General Theory.* New York: Interscience, 1958.
22. Figà-Talamanca A. Translation invariant operators in L_p. *Duke. Math. J.*, 1965. Vol. 32, No. 3. P. 495–502. DOI: 10.1215/S0012-7094-65-03250-3
23. Figà-Talamanca A., Gaudry G. I. Density and representation theorems for multipliers of type (p, q). *J. Australian Math. Soc.*, 1967. Vol. 7, No. 1. P. 1–6. DOI: 10.1017/S1446788700005012
24. Gabushin V. N. Inequalities for the norms of a function and its derivatives in metric L_p. *Math. Notes*, 1967. Vol. 1, No. 3. P. 194–198. DOI: 10.1007/BF01098882
25. Gabushin V. N. Best approximations of functionals on certain sets. *Math. Notes Acad. Sci. USSR*, 1970. Vol. 8, No. 5. P. 780–785. DOI: 10.1007/BF01146932
26. Gabushin V. N. The best approximation of the differentiation operator in the metric of L_p. *Math. Notes*, 1972. Vol. 12, No. 5. P. 756–760. DOI: 10.1007/BF01099059
27. Gabushin V. N. Inequalities between derivatives in L_p-metrics for $0 < p \le \infty$. *Math. USSR-Izv.*, 1976. Vol. 10, No. 4. P. 823–844.
28. Gabushin V. N. Inequalities for derivatives of solutions of ordinary differential equations in the metric of L_p $(0 < p \le \infty)$. *Differ. Equ.*, 1988. Vol. 24, No. 10. P. 1088–1094.
29. Hadamard J. Sur le module maximum d'une fonction et de ses dérivées. *Soc. math. France, Comptes rendus des Séances*, 1914. Vol. 41. P. 68–72. (in French)
30. Hardy G. H., Littlewood J. E. Contribution to the arithmetic theory of series. *Proc. London Math. Soc.* 2, 1912. Vol. 11. P. 411–478.
31. Hardy G. H., Littlewood J. E., Pólya G. *Inequalities.* Cambridge: Cambridge Univ. Press, 1934.
32. Hörmander L. Estimates for translation invariant operators in L^p spaces. *Acta Math.*, 1960. Vol. 104, No. 1–2. P. 93–140. DOI: 10.1007/BF02547187
33. Klots B. E. Linear deviations of the classes $\widetilde{W}_p^{\alpha}$ and approximations in spaces of multipliers. *Math. Notes*, 1975. Vol. 18, No. 1. P. 640–646. DOI: 10.1007/BF01461146
34. Klots B. E. Approximation of differentiable functions by functions of large smoothness. *Math. Notes*, 1977. Vol. 21, No. 1. P. 12–19. DOI: 10.1007/BF02317028

35. Kolmogorov A. N. On inequalities between upper estimates of consecutive derivatives of an arbitrary function defined on an infinite interval. In: *Selected Works: Mathematics and Mechanics.* Moscow: Nauka, 1985. P. 252–263. (in Russian)
36. Korneichuk N. P. *Ekstremal'nye zadachi teorii funkcii* [Extremal Problems of Approximation Theory]. Moscow: Nauka, 1976. 320 p. (in Russian)
37. Larsen R. *An Introduction to the Theory of Multipliers.* Berlin etc.: Springer, 1971. 282 p.
38. Levin V. I. Exact constants in inequalities of the Carlson type. *Doklady Akad. Nauk. SSSR*, 1948. Vol. 59. P. 635–638. (in Russian)
39. Magaril-Il'yaev G. G., Osipenko K. Yu. Optimal recovery of functions and their derivatives from inaccurate information about the spectrum and inequalities for derivatives. *Funct. Anal. Appl.*, 2003. Vol. 37. P. 203–214. DOI: 10.1023/A:1026084617039
40. Magaril-Il'yaev G. G., Osipenko K. Yu. Optimal recovery of values of functions and their derivatives from inaccurate data on the Fourier transform. *Sbornik Math.*, 2004. Vol. 195, No. 10. P. 1461–1476. DOI: 10.1070/SM2004v195n10ABEH000853
41. Rudin W. *Functional Analysis.* New York: McGraw-Hill, 1973. 397 p.
42. Shilov G. E. *Matematicheskii analiz. Vtoroi spetsial'nyi kurs* [Mathematical Analysis: The Second Special Course]. Moscow: Nauka, 1965. 328 p. (in Russian)
43. Stein E. M. Functions of exponential type. *Ann. Math.* (2), 1957. Vol. 65. P. 582–592.
44. Stein E. M., Weiss G. *Introduction to Fourier Analysis on Euclidean Spaces.* Princeton: Princeton Univ. Press, 1971. 297 p.
45. Stechkin S. B. Inequalities between norms of derivatives of an arbitrary function. *Acta Sci. Math.*, 1965. Vol. 26, No. 3–4. P. 225–230.
46. Stechkin S. B. Best approximation of linear operators. *Math. Notes*, 1967. Vol. 1, No. 2. P. 91–99. DOI: 10.1007/BF01268056
47. Subbotin Yu. N., Taikov L. V. Best approximation of a differentiation operator in L_2-space. *Math. Notes*, 1968. Vol. 3, No. 2. P. 100–105. DOI: 10.1007/BF01094328
48. Tikhomirov V. M., Magaril-Il'yaev G. G. Inequalities for derivatives. In: *Kolmogorov A. N. Izbrannye trudy. Matematika i mekhanika* [Selected Works. Mathematics and Mechanics]: collection of papers. Moscow: Nauka, 1985. P. 387–390. (in Russian)

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 28–35
DOI: 10.15826/umj.2023.2.002

COUNTABLE COMPACTNESS MODULO AN IDEAL OF NATURAL NUMBERS

Prasenjit Bal[†], **Debjani Rakshit**[††] **and Susmita Sarkar**[†††]

Department of Mathematics, ICFAI University Tripura,
Kamalghat, India-799210

[†]balprasenjit177@gmail.com, [††]debjanirakshit88@gmail.com,
[†††]susmitamsc94@gmail.com

Abstract: In this article, we introduce the idea of I-compactness as a covering property through ideals of $\mathbb{N}$ and regardless of the I-convergent sequences of points. The frameworks of s-compactness, compactness and sequential compactness are compared to the structure of I-compact space. We began our research by looking at some fundamental characteristics, such as the nature of a subspace of an I-compact space, then investigated its attributes in regular and separable space. Finally, various features resembling finite intersection property have been investigated, and a connection between I-compactness and sequential I-compactness has been established.

Keywords: Ideal, Open cover, Compact space, I-convergence.

1. Introduction

The concept of statistical convergence ([17], also [21]) depends on the idea of natural density of subsets of the set of natural number $\mathbb{N}$. The density of a set $S \subseteq \mathbb{N}$ is denoted by $\delta(S)$ and defined as

$$\delta(S) = \lim_{n\to\infty} \frac{1}{n}\big|\big\{k \le n : \ k \in S\big\}\big|.$$

Later on, Maio and Kočinak [14] redefine the statistical convergence for a topological space. On the recent days, one of the most significant study area in pure mathematics is the ideal convergence which is an extension of statistical convergence and other convergence concepts. To define the ideal convergence Kostyrko et al. [19] used the notion of an ideal which is defined as $I \subseteq \mathcal{P}(\mathbb{N})$ having the following properties:

(i) $\varnothing \in I$;
(ii) $A \cup B \in I$, for each $A, B \in I$;
(iii) for each $A \in I$ and $B \subseteq A \Rightarrow B \in I$.

In a topological space (X, τ), a sequence $x = (x_n)$ is said to be I-convergent if there exists a ℓ such that for every open neighbourhood U of ℓ, the set [20]

$$\{n \in \mathbb{N} : x_n \notin U\} \in I.$$

This idea is being studied and used broadly by many researchers [1, 10, 12, 13].

On the other hand, the study of different covering properties (some recent works [7, 9, 11]) has received a lot of attention in topology. A family $\mathcal{U}$ of open subsets of topological space X is called an open cover of X if $\cup\mathcal{U} = X$. A topological space X is called compact if every open cover of X has a finite subcover. More specifically, it has became very essential to explore the structure of compactness and its generalized versions for topological spaces.

The star operator was introduce by E.K. van Douwen in 1991 [15] as

$$\mathrm{St}\,(A,\mathcal{U}) = \bigcup\{U \in \mathcal{U} : U \cap A \neq \varnothing\}$$

where A is a subset of space X and $\mathcal{U}$ is a family of subsets of X. Using star operator the concept of compactness has been generalized in many ways and has been studied by many authors extensively [2–6]. In our study we make an attempt to expand this region with the help of Ideal.

In recent days sequential compactness via ideal has been introduced by Singha and Roy [22] under the name of I-compactness, and I-compactness module via an ideal defined on X has also been studied by Gupta and Kaur [18] under the same name.

The purpose of our study is to explore the concept of compactness via ideal of natural number $\mathbb{N}$. We also establish a relation between sequential I-compactness and I–compactness.

2. Preliminaries

Throughout the paper a space X means a topological space with the corresponding topology τ, '$\because$' stands for 'therefore' and for other symbols and notions we follow [16].

Definition 1 [16]. *A topological space X is called a compact space if every open cover of X has a finite subcover, i.e., if for every open cover $\{U_s\}_{s\in\mathbb{S}}$ of the space X there exists a finite set $\{s_1, s_2, ..., s_k\} \subset \mathbb{S}$ such that $X = U_{s_1} \cup U_{s_2} \cup ... \cup U_{s_k}$.*

Definition 2 [16]. *A topological space X is called a Lindelöf space if every open cover of X has a countable subcover.*

It is known that every compact space is Lindelöf but the converse is not true.

Definition 3 [16]. *A topological space (X, τ) is called a countably compact space if every countable open cover of X has a finite sub-cover.*

Every compact space is countably compact. But the space W_0 of all countable ordinals is countably compact but not compact [16]. The space $\mathbb{N}$ of all natural numbers equipped with discrete topology is Lindelöf but it is not countably compact.

Definition 4 [16]. *A topological space (X, τ) is said to be sequentially compact space if every sequence in X has a convergent subsequence.*

Definition 5 [15]. *A topological space (X, τ) will be called a star compact space (in short* St*-compact) if for every open cover $\mathcal{U}$ of X, there exists a finite subset $\mathcal{U}' = \{U_k : k = 1, 2, 3, \dots, m\}$ such that*

$$\mathrm{St}\Big(\bigcup_{n=1}^{m} U_n, \mathcal{U}\Big) = X.$$

Definition 6 [8]. *A topological space (X, τ) will be called a statistical compact (in short s-compact) space if for every countable open cover $\mathcal{U} = \{U_n : n \in \mathbb{N}\}$ of X, there exists a sub-cover $\mathcal{V} = \{U_{m_k} : k \in \mathbb{N}\}$ of $\mathcal{U}$ such that $\delta\big(\{m_k : U_{m_k} \in \mathcal{V}\}\big) = 0$.*

3. Compactness via ideal

Definition 7. *Let I be a non-trivial ideal defined on $\mathbb{N}$. A topological space (X, τ) will be called an I-compact space if for every countable open cover $\mathcal{U} = \{A_n : n \in \mathbb{N}\}$ of X, there exists a sub-cover $\mathcal{V} = \{A_{n_k} : k \in \mathbb{N}\}$ such that $\{n_k : A_{n_k} \in \mathcal{V}\} \in I$.*

Remark 1. Countable compactness is equivalent to I_{fin}-compactness where I_{fin} indicates the ideal of all finite subsets of $\mathbb{N}$.

Proposition 1. *Every I_{fin}-compact space is a s-compact space.*

P r o o f. Let (X, τ) be an I_{fin}-compact space and $\mathcal{U} = \{A_n : n \in \mathbb{N}\}$ be a countable open cover of X. Therefore there exists a sub-cover $\mathcal{V} = \{A_{n_k} : k \in \mathbb{N}\}$ of $\mathcal{U}$ with $\{n_k : n_k \in \mathcal{V}\} \in I_{fin}$. i.e. $\mathcal{V} = \{A_{n_k} : k \in \mathbb{N}\}$ is a finite sub-cover of X. But the finite set of indices $\{n_k : A_{n_k} \in \mathcal{V}\}$ has natural density zero, i.e. $\delta(\{n_k : A_{n_k} \in \mathcal{V}\}) = 0$, $\therefore X$ is a s-compact space. □

Example 1. Converse of Proposition 1 may not be true. Indeed there exists a s-compact space which is not I_{fin}-compact.

Let $X = (-1, 1)$ and $\tau = \{(-\alpha, \alpha) : \alpha \in [0, 1]\}$. Clearly (X, τ) is a topological space. Consider a countable open cover $\mathcal{U} = \{U_n : n \in \mathbb{N}\}$ of X. If $X \in \mathcal{U}$, then $X = U_p$ for some $p \in \mathbb{N}$ and $\mathcal{V} = \{U_p\}$ is a sub-cover of $\mathcal{U}$ with $\delta(\{k : U_k \subset \mathcal{V}\}) = \delta(\{p\}) = 0$ and we are done.

Now let $X \notin \mathcal{U}$ and $\mathcal{U} = \{U_n : n \in \mathbb{N}\}$ is a non-trivial countable open cover of X. We consider the sub-cover $\mathcal{U}' = \{U_{n_k} : k \in \mathbb{N}\}$, where

$$U_{n_k} = \begin{cases} U_k, & k = 1, \\ \bigcup_{n \leq n_k} U_n, & \text{when } \bigcup_{n \leq n_k} U_n \text{ becomes a superset of } \quad U_{n_{k-1}} \text{ for } \quad k > 1. \end{cases}$$

Now, $\{U_{n_k} : k \in \mathbb{N}\}$ is an increasing sequence of open sets by means of inclusion ($\subseteq$) and is an open cover of X. It also has a sub-cover $\mathcal{V} = \{U_{n_{k^2}} : k \in \mathbb{N}\}$ with $\delta(\{n_{k^2} : U_{n_{k^2}} \in \mathcal{V}\}) = 0$. Moreover $\mathcal{V}$ is a subset of $\mathcal{U}$, $\therefore X$ is a s-compact space.

Again suppose that (X, τ) is I_{fin}-compact and consider the countable open cover

$$\mathcal{W} = \Big\{W_n = \Big(-1 + \frac{1}{n}, 1 - \frac{1}{n}\Big) : \ n \in \mathbb{N}\Big\}.$$

Since X is I_{fin}-compact there exists a sub-cover of $\mathcal{W}$, say $\mathcal{W}' = \{W_{n_k} : k = 1, 2, ..., q\}$ with $\{n_k : W_{n_k} \in \mathcal{W}'\} \in I_{fin}$.

Suppose $n_{k_{max}} = \max\{n_k : n_k \in \mathcal{W}'\}$ then we have

$$\therefore \bigcup \mathcal{W}' = \Big(-1 + \frac{1}{n_{k_{max}}}, 1 - \frac{1}{n_{k_{max}}}\Big) \neq X,$$

which is a contradiction. So X is not I_{fin}-compact.

Corollary 1. *Every Lindelöf I_{fin}-compact space is a compact space.*

P r o o f. By Lindelöfness, every open cover has a countable sub-cover. By I_{fin}-compactness, that countable sub-cover will have a finite sub-cover. Hence it will be a compact space. □

Theorem 1. *Every closed subspace of an I-compact space is an I-compact.*

P r o o f. Let (A, τ_A) be an arbitrary closed subspace of a I-compact space (X, τ) and $\mathcal{U} = \{U_n : n \in \mathbb{N}\}$ be a countably infinite cover of (A, τ_A). Then there exists a countable sequence $\mathcal{V} = \{V_n : n \in \mathbb{N}\} \in \tau$ such that $U_n = V_n \cap A$.

Now consider a countably infinite sequence $\mathcal{W} = \{(X \setminus A) \cup V_n : n \in \mathbb{N}\}$, which is a cover of X. Since (X, τ) is I-compact space, $\therefore \exists$ a sub-cover $\mathcal{W}' = \{(X \setminus A) \cup V_{n_k} : k \in \mathbb{N}\}$ of $\mathcal{W}$ with $\{n_k : (X \setminus A) \cup V_{n_k} \in \mathcal{W}'\} \in I$. Again, $\cup V_{n_k} \supseteq A$ then

$$A \cap (\cup V_{n_k}) = A \implies \cup(A \cap V_{n_k}) = A \implies \cup(U_{n_k}) = A, \quad \therefore \cup U_{n_k} \subseteq \cup U_n = \mathcal{U}$$

and $\mathcal{U}$ is a countably infinite cover of (A, τ_A). So $\{\cup U_{n_k} : k \in \mathbb{N}\}$ is a countably sub-cover of (A, τ_A) with $\{n_k : U_{n_k} \in \mathcal{U}\} \in I$.

Therefore (A, τ_A) is the I-compact space. □

Theorem 2. *Let (X, τ) be a I-compact space and (Y, σ) be a topological space. If $f : (X, \tau) \to (Y, \sigma)$ is the open continuous surjection mapping, then (Y, σ) is also the I-compact space.*

P r o o f. Let $f : (X, \tau) \to (Y, \sigma)$ be an open continuous surjection mapping and (X, τ) is I-compact space.

Let $\{U_n : n \in \mathbb{N}\}$ be a countable open cover of Y. So,

$$\cup\{U_n : n \in \mathbb{N}\} = Y \implies f^{-1}\big\{ \cup \{U_n : n \in \mathbb{N}\}\big\} = f^{-1}(Y) \implies \cup\{f^{-1}(U_n) : n \in \mathbb{N}\} = X.$$

Since f is a continuous surjection mapping and U_n is a countable open cover of Y then

$$\{f^{-1}(U_n) : n \in \mathbb{N}\} = \mathcal{V}$$

is a countable open cover of X. Again (X, τ) is I-compact space then there exists a sub-cover, $\{f^{-1}(U_{n_1}), f^{-1}(U_{n_2}), ...\}$ of $\mathcal{V}$ where $\{n_k : k \in \mathbb{N}\} \in I$,

$$\cup\{f^{-1}(U_{n_k}) : k \in \mathbb{N}\} = X,$$
$$f\big[\cup \{f^{-1}(U_{n_k}) : k \in \mathbb{N}\}\big] = f(X) = Y,$$
$$\cup\{U_{n_k} : k \in \mathbb{N}\} = Y, \quad \because f[f^{-1}(U_{n_k})] = U_{n_k},$$

$\therefore \{U_{n_k} : k \in \mathbb{N}\}$ is a countable sub-cover of $\{U_n : n \in \mathbb{N}\}$ where $\{n_k : k \in \mathbb{N}\} \in I$, (y, σ) is also I-compact space. □

Definition 8. *Let (X, τ) be a topological space and I be a ideal on $\mathbb{N}$. A subset $A \subseteq X$ will be called I-compact subset of X if for every countable cover $\{U_n : n \in \mathbb{N}\}$ of A by elements of τ there exists a $S \in I$ such that $A \subseteq \bigcup_{n \in S} U_n$.*

Theorem 3. *In a regular space (X, τ), if A is countable I_{fin}-compact subset of X, then for every closed set B disjoint from A there exists $U, V \in \tau$ such that $A \subseteq U$, $B \subseteq V$ and $U \cap V = \varnothing$.*

P r o o f. Let $A = \{x_n : n \in \mathbb{N}\}$ be a countable I_{fin}-compact subset of a regular space (X, τ) and B be an arbitrary closed set disjoint from A, $\therefore$ for every $x_n \in A$, $x_n \notin B$. But X is a regular space. Therefore there exists $U_n, V_n \in \tau$ such that $x_n \in U_n$, $B \subseteq V_n$ and $U_n \cap V_n = \varnothing$ $\forall n \in \mathbb{N}$.

It is obvious that $\{U_n : n \in \mathbb{N}\}$ is a countable open cover of A by the elements of τ. But A is an I_{fin}-compact subset of X, $\therefore$ there exist $S \in I_{fin}$ such that $A \subseteq \bigcup_{n \in S} U_n$.

Now $S \in I_{fin}$ is a finite set, $\therefore U = \bigcup_{n\in S} U_n \in \tau$ and $V = \bigcap_{n\in S} V_n \in \tau$ and $A \subseteq U$ and $B \subseteq V$. So we have to show that $U \cap V = \varnothing$. On the contrary suppose $U \cap V \neq \varnothing$ and $p \in U \cap V$

$$\begin{aligned}
&\implies p \in (\bigcup_{n\in S} U_n) \cap (\bigcap_{n\in S} V_n) \implies p \in \bigcup_{n\in S} U_n \quad \text{and} \quad p \in \bigcap_{n\in S} V_n \\
&\implies p \in U_k \quad \text{for some} \quad k \in S \quad \text{and} \quad p \in V_k \quad \forall k \in S \\
&\implies p \in U_k \quad \text{and} \quad p \in V_k \quad \text{for some} \quad k \in S \\
&\implies p \in U_k \cap V_k \quad \text{for some} \quad k \in S \subseteq \mathbb{N},
\end{aligned}$$

which is a contradiction to the fact that

$$U_n \cap V_n = \varnothing \quad \forall n \in \mathbb{N}, \quad \therefore U \cap V = \varnothing.$$

Hence the theorem is proven. □

Corollary 2. *If A is a countable I_{fin}-compact subset of a Hausdörff space X, then for every $x \notin A$ there exist $U, V \in \tau$ such that $A \subseteq U, x \in V$ and $U \cap V = \varnothing$.*

P r o o f. In a Hausdörff space, every singleton set $\{x\}$ is a closed set. So by Theorem 3 the result follows directly. □

Definition 9. *A topological space (X, τ) will be called sequentially I-compact if every sequence of elements of X has a I-convergent subsequence.*

Theorem 4. *A separable I_{fin}-compact space is a st-compact space*

P r o o f. Let (X, τ) be a separable I_{fin}-compact space. Therefore there exists a countable dense subset $A = \{x_n : n \in \mathbb{N}\}$ of X and $\mathcal{U}$ being an arbitrary open cover of X. Using the elements of $\mathcal{U}$ we construct a sequence of open sets $\{U_n : n \in \mathbb{N}\}$ where $U_n = \bigcup\{U \in \mathcal{U} : x_n \in U\}$ for all $n \in \mathbb{N}$. But A is a dense subset of X, $\therefore A \cap U \neq \varnothing \quad \forall U \in \mathcal{U}$.

$\therefore \mathcal{U}' = \{U_n : n \in \mathbb{N}\}$ is a countable open cover of X. But X is I_{fin}-compact, $\therefore \exists S \in I_{fin}$ such that $\bigcup_{n_k \in S} U_{n_k} = X$. But $S = \{U_{n_1}, U_{n_2}, ..., U_{n_p}\}$ is a finite subset of $\mathbb{N}$. Therefore

$$\bigcup_{k=1}^{n} U_{n_k} = X.$$

But $x_{n_k} \in U_{n_k} \quad \forall k = 1, 2, ..., p$, therefore $F = \{x_{n_k} : k = 1, 2, ..., p\}$ is a finite subset of X and

$$\begin{gathered}
\mathrm{St}\,(F, \mathcal{U}') \supseteq \bigcup_{k=1}^{n} U_{n_k} = X, \\
\mathrm{St}\,(F, \mathcal{U}') \supseteq \bigcup_{k=1}^{n} \Big\{ \bigcup\{U \in \mathcal{U} : x_{n_k} \in U\} \Big\} = X, \\
\mathrm{St}\,(F, \mathcal{U}) \supseteq \mathrm{St}\,(F, \mathcal{U}') \supseteq X, \\
\mathrm{St}\,(F, \mathcal{U}) = X,
\end{gathered}$$

$\therefore X$ is a St-compact space. □

Definition 10. *Let I be a ideal on $\mathbb{N}$. A family $\mathcal{F} = \{F_n : n \in \mathbb{N}\}$ of subsets of a space X is said to have I-intersection property if $\mathcal{F} \neq \varnothing$ and $\bigcap_{n\in S} F_n \neq \varnothing$ for all $S \in I$.*

Theorem 5. *For a topological space (X,τ) and for a non trivial ideal I the following statements are equivalent*:

(1) *For a family $\mathcal{G} = \{G_n : n \in \mathbb{N}\}$ of open sets of X, if for every $S \in I$, $\quad \mathcal{G}_S = \{G_{n_\alpha} : n_\alpha \in S\}$ fails to cover X, then $\mathcal{G}$ can not cover X.*
(2) *X is an I-compact space.*
(3) *Every family of countable closed subsets of X with I-intersection property has non-empty intersection.*
(4) *For a family $\mathcal{F} = \{F_n : n \in \mathbb{N}\}$ of closed subsets of X, if $\bigcap \mathcal{F} = \varnothing$, then there exists at least one $S \in I$ such that $\bigcap_{n\in S} F_n = \varnothing$.*

P r o o f. (1) $\Leftrightarrow$ (2): The statement (1) is the contrapositive statement of the definition of I-compact space, $\therefore$ statement (1) and (2) are equivalent.

(2) $\Leftrightarrow$ (3): Let X be an I-compact space, $\mathcal{H} = \{H_n : n \in \mathbb{N}\}$ be a arbitrary family of closed subsets of X having I-intersection property and suppose that $\bigcap \mathcal{H} = \varnothing$.

Let
$$\mathcal{G} = \{G_n = X \setminus H_n : n \in \mathbb{N}\},$$
then
$$\bigcup \mathcal{G} = \bigcup_{n\in\mathbb{N}} (G_n = X \setminus H_n) = X \setminus \bigcap_{n\in\mathbb{N}} H_n = X \setminus \varnothing = X,$$
$$\therefore \mathcal{G} = \{G_n : n \in \mathbb{N}\}$$
is an countable open cover of X. But X is I-compact. Therefore, there exists a $S \in I$ such that
$$\bigcup_{n\in S} G_n = X,$$
$$\implies \bigcup_{n\in S} \{X \setminus H_n\} = X \quad [\because G_n = X \setminus H_n \quad \forall n \in \mathbb{N}] \implies X \setminus \bigcap_{n\in S} H_n = X \implies \bigcap_{n\in S} H_n = \varnothing,$$
which is a contradiction to the fact that $\mathcal{H}$ has I-intersection property, $\therefore \bigcap \mathcal{H} \neq \varnothing$.

Conversely let every countable family of closed subsets with I-intersection property has non-empty intersection.

Let $\mathcal{U} = \{B_n : n \in \mathbb{N}\}$ is an arbitrary countable open cover of X, $\therefore \mathcal{F} = \{F_n = X \setminus B_n : n \in \mathbb{N}\}$ is a family of closed subsets of X and
$$\bigcap \mathcal{F} = \bigcap_{n\in\mathbb{N}} X \setminus B_n = X \setminus \bigcup_{n\in\mathbb{N}} B_n = X \setminus X = \varnothing.$$
Thus the countable family $\mathcal{F}$ of closed subsets of X has empty intersection. So it can not have I-intersection property by our assumption. Therefore, there exists a $S \in I$ such that
$$\bigcap_{n\in S} F_n = \varnothing \implies \bigcap_{n\in S} X \setminus B_n = \varnothing \implies X \setminus \bigcup_{n\in S} B_n = \varnothing \implies \bigcup_{n\in S} B_n = \varnothing,$$
$\therefore \mathcal{V} = \{B_n : n \in S\}$ is a sub-cover of $\mathcal{U}$ and $S \in I$. Therefore X is an I-compact space.

(3) $\Leftrightarrow$ (4): Statement (3) and statement (4) are contrapositive to each other. Therefore, statement (3) and statement (4) are equivalent. $\square$

Theorem 6. *Sequentially I_{fin}-compactness implies I_{fin}-compactness.*

P r o o f. Let (X,τ) be an I_{fin}-compact space. Then for every sequence $\{x_n : n \in \mathbb{N}\}$ of elements of X, there exists is a subsequence $\{x_{n_k} : n \in \mathbb{N}\}$ which is I-convergent. Let

$$I - \lim_{n\to\infty} x_n = \epsilon \in X.$$

On the other hand let $\mathcal{U} = \{U_n : n \in \mathbb{N}\}$ be a countable open cover of X. So $\exists U_m \in \mathcal{U}$ such that $\epsilon \in U_m$. Also $\{n_k \in \mathbb{N} : x_{n_k} \notin U_m\} \in I_{fin}$, suppose $\{n_{k_1}, n_{k_2}, ..., n_{k_p}\} = \{n_k \in \mathbb{N} : x_{n_k} \notin U_m\}$.

But $\mathcal{U}$ is a open cover of X. Therefore there exist

$$\begin{gathered} x_{n_{k_1}} \in U_{q_1} \in \mathcal{U}, \\ x_{n_{k_2}} \in U_{q_2} \in \mathcal{U}, \\ ... \\ x_{n_{k_p}} \in U_{q_p} \in \mathcal{U}. \end{gathered}$$

Now, the collection $\{U_m, U_{q_1}, U_{q_2}, ..., U_{q_p}\}$ is a sub-cover of $\mathcal{U}$ and $\{m, q_1, q_2, ..., q_p\}$ is a finite subset of $\mathbb{N}$, i.e. $\{m, q_1, q_2, ..., q_p\} \in I_{fin}$, $\therefore$ (X,τ) is a I_{fin}-compact space. □

4. Conclusion

The paper reveals that I_{fin}-compactness is stronger covering property than statistical compactness, closed subspace of an I-compact space is I-compact, open continuous surjection of an I-compact space is I-compact, a separable I-compact space is a star-compact space. A topological space is I-compact if and only if every family of countable closed subsets of the space which has I-intersection property has non-empty intersection. This study can further be extended for the covering properties like Mengerness and Rothbergerness in the context of modulo an ideal of natural numbers.

Acknowledgements

Authors are thankful to the reviewers for their insightful comments and suggestions which drastically improved the quality and presentation of the paper.

REFERENCES

1. Bal P. On the class of I-γ-open cover and I-St-γ-open cover. *Hacet. J. Math. Stat.*, 2023. Vol. 52, No. 3. P. 630–639.
2. Bal P., Bhowmik S. Star-selection principle: another new direction. *J. Indian Math. Soc.*, 2017. Vol. 84, No. 1–2. P. 1–6. DOI: 10.18311/jims/2017/6124
3. Bal P., Bhowmik S. On R-star-Lindelöf spaces. *Palest. J. Math.*, 2017. Vol. 6, No. 2. P. 480–486.
4. Bal P., Bhowmik S. Some new star-selection principles in topology. *Filomat*, 2017. Vol. 31, No. 13. P. 4041–4050. DOI: 10.2298/FIL1713041B
5. Bal P., Bhowmik S., Gauld D. On Selectively star-Lindelöf properties. *J. Indian Math. Soc.*, 2018. Vol. 85, No. 3–4. P. 291–304. DOI: 10.18311/jims/2018/20145
6. Bal P., Kočinac L. D. R. On selectively star-ccc spaces. *Topology Appl.*, 2020. Vol. 281. Art. no. 107184. DOI: 10.1016/j.topol.2020.107184
7. 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
8. Bal P., Rakshit D., Sarkar S. On statistical compactness. *Iran. J. Math. Sci. Inform.*, 2023. (accepted).
9. Bal P., De R. On strongly star semi-compactness of topological spaces. *Khayyam J. Math.*, 2023. Vol. 9, No. 1. P. 54–60. DOI: 10.22034/KJM.2022.342863.2546

10. Balcerzak M., Filipczak M. L. Ideal convergence of sequences and some of its applications. *Folia Math.*, 2017. Vol. 19, No. 1. P. 3–8.
11. Çakallı H. Sequential definitions of compactness. *Appl. Math. Lett.*, 2008. Vol. 21, No. 6. P. 594–598. DOI: 10.1016/j.aml.2007.07.011
12. Das P. Certain types of open covers and selection principles using ideals. *Houston J. Math.*, 2013, Vol. 39, No. 2. P. 637–650.
13. Debnath S., Rakshit D. On I-statistical convergence. *Iran. J. Math. Sci. Inform.*, 2018. Vol. 13, No. 2. P. 101–109.
14. Di Maio G., Kočinak 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
15. van Douwen E. K., Reed G. M., Roscoe A. W., Tree I. J. Star covering properties. *Topology Appl.*, 1991. Vol. 39, No. 1. P. 71–103. DOI: 10.1016/0166-8641(91)90077-Y
16. Engelking R. *General Topology.* Sigma Ser. Pure Math. Rev. compl. ed. Berlin: Heldermann, 1989. 529 p.
17. Fast H. Sur la convergence statistique. *Colloq. Math.*, 1951. Vol. 2, No. 3–4. P. 241–244. (in French)
18. Gupta A., Kaur R. Compact spaces with respect to an ideal. *Int. J. Pure Appl. Math.*, 2014. Vol. 92, No. 3. P. 443–448.
19. Kostyrko P., Salat T., Wilczynski W. I-convergence. *Real Anal. Exchange*, 2000. Vol. 26, No. 2. P. 669–686.
20. 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
21. 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
22. Singha M., Roy S. *Compactness with Ideals.* arXiv:2107.00050 [math.GN]

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 36–45
DOI: 10.15826/umj.2023.2.003

ON SEQUENCES OF ELEMENTARY TRANSFORMATIONS IN THE INTEGER PARTITIONS LATTICE

Vitaly A. Baransky[†], **Tatiana A. Senchonok**[††]

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

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

Abstract: An *integer partition*, or simply, a *partition* is a nonincreasing sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of nonnegative integers that contains only a finite number of nonzero components. The *length* $\ell(\lambda)$ of a partition λ is the number of its nonzero components. For convenience, a partition λ will often be written in the form $\lambda = (\lambda_1, \dots, \lambda_t)$, where $t \geq \ell(\lambda)$; i.e., we will omit the zeros, starting from some zero component, not forgetting that the sequence is infinite. Let there be natural numbers $i, j \in \{1, \dots, \ell(\lambda)+1\}$ such that (1) $\lambda_i - 1 \geq \lambda_{i+1}$; (2) $\lambda_{j-1} \geq \lambda_j + 1$; (3) $\lambda_i = \lambda_j + \delta$, where $\delta \geq 2$. We will say that the partition $\eta = (\lambda_1, \dots, \lambda_i - 1, \dots, \lambda_j + 1, \dots, \lambda_n)$ is obtained from a partition $\lambda = (\lambda_1, \dots, \lambda_i, \dots, \lambda_j, \dots, \lambda_n)$ by an *elementary transformation of the first type*. Let $\lambda_i - 1 \geq \lambda_{i+1}$, where $i \leq \ell(\lambda)$. A transformation that replaces λ by $\eta = (\lambda_1, \dots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \dots)$ will be called an *elementary transformation of the second type*. The authors showed earlier that a partition μ dominates a partition λ if and only if λ can be obtained from μ by a finite number (possibly a zero one) of elementary transformations of the pointed types. Let λ and μ be two arbitrary partitions such that μ dominates λ. This work aims to study the shortest sequences of elementary transformations from μ to λ. As a result, we have built an algorithm that finds all the shortest sequences of this type.

Keywords: Integer partition, Ferrers diagram, Integer partitions lattice, Elementary transformation.

1. Introduction

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

An *integer partition*, or simply a *partition*, is a nonincreasing sequence $\lambda = (\lambda_1, \lambda_2, \dots)$ of nonnegative integers that contains only a finite number of nonzero components (see [1]). Let $\operatorname{sum}\lambda$ denote the sum of all components of a partition λ and call it the *weight* of the partition λ. It is often said that a partition λ is a partition of the nonnegative integer $n = \operatorname{sum}\lambda$. The *length* $\ell(\lambda)$ of a partition λ is the number of its nonzero components. For convenience, a partition λ will often be written in the form $\lambda = (\lambda_1, \dots, \lambda_t)$, where $t \geq \ell(\lambda)$; i.e., we will omit the zeros, starting from some zero component, not forgetting that the sequence is infinite.

Let IPL denote the lattice of all (integer) partitions of all nonnegative integers, and let $IPL(m)$ denote the lattice of all partitions of a given nonnegative integer m. On the lattices IPL and $IPL(m)$, where $m \in \mathbb{N}$, the well-known domination relation is considered [7].

We define two types of elementary transformations (see [2, 3]) of the partition

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_t, 0, 0, \dots),$$

where $t = \ell(\lambda) + 1$.

Let there be natural numbers $i, j \in \{1, \dots, t\}$ such that $1 \leq i < j \leq \ell(\lambda) + 1$ and

(1) $\lambda_i - 1 \geq \lambda_{i+1}$ (or, equivalently, $\lambda_i > \lambda_{i+1}$);
(2) $\lambda_{j-1} \geq \lambda_j + 1$ (or, equivalently, $\lambda_{j-1} > \lambda_j$);
(3) $\lambda_i = \lambda_j + \delta$, where $\delta \geq 2$.

We will say that the partition $\eta = (\lambda_1, \dots, \lambda_i - 1, \dots, \lambda_j + 1, \dots, \lambda_n)$ is obtained from a partition $\lambda = (\lambda_1, \dots, \lambda_i, \dots, \lambda_j, \dots, \lambda_n)$ by an *elementary transformation of the first type* (or a *box movement*). Conditions (1), (2), and (3) guarantee that a partition again will be obtained. Note that η differs from λ by precisely two components with numbers i and j. For the Ferrers diagram, this transformation means moving the top box of i-column to the right to the top of the j-column. We will use Cartesian notation for the Ferrers diagram: each k-column consists of λ_k boxes (see [6]).

It should be noted that a box can also be thrown to the zero component with the number $\ell(\lambda) + 1$. The fact that η is obtained from λ by moving the box will be briefly written in the form $\lambda \rightharpoondown \eta$. Note that an elementary transformation of the first type preserves the weight of the partition, while the length of the partition can be preserved or lifted by 1.

We now define elementary transformations of the second type for the partition $\lambda = (\lambda_1, \lambda_2, \dots)$.

Let $\lambda_i - 1 \ge \lambda_{i+1}$ (or, equivalently, $\lambda_i > \lambda_{i+1}$), where $i \le \ell(\lambda)$. A transformation that replaces λ by $\eta = (\lambda_1, \dots, \lambda_{i-1}, \lambda_i - 1, \lambda_{i+1}, \dots)$ will be called an *elementary transformation of the second type* (or a *box removement*). As in the previous case, we will briefly write $\lambda \rightharpoondown \eta$. It should be noted that box removal reduces the weight of the partition exactly by 1, while the length of the partition can be preserved or lowered by 1.

It was shown in [2, 3] that a partition μ dominates a partition λ if and only if λ can be obtained from μ by sequentially applying a finite number (possibly a zero one) of elementary transformations of the pointed types.

Let λ and μ be two arbitrary partitions and $\lambda \le \mu$. The height (μ, λ) of a partition μ over a partition λ is the number of transformations in a shortest sequence of elementary transformations transforming μ into λ.

The following theorem was proved in [4].

Theorem 1 [4, Theorem 1]. *Let $\mu \ge \lambda$ in IPL and $C = \operatorname{sum}\mu - \operatorname{sum}\lambda$. Then*

$$\operatorname{height}(\mu, \lambda) = \sum_{j=1,\, \mu_j > \lambda_j}^{\infty} (\mu_j - \lambda_j) = \frac{1}{2}C + \frac{1}{2}\sum_{j=1}^{\infty} |\mu_j - \lambda_j|.$$

Let μ and λ be some fixed partitions such that $\mu > \lambda$. Consider sequences of elementary transformations from μ to λ (both types of elementary transformations are admissible):

$$\mu = \xi_{(0)} \rightharpoondown \xi_{(1)} \rightharpoondown \cdots \rightharpoondown \xi_{(s)} = \lambda.$$

This paper aims to describe an algorithm (Algorithm 1) for constructing all possible shortest sequences of this kind. Algorithm 1 generalizes an algorithm constructed in [4] (see Algorithm 2).

2. Main results

Let $\mu = (\mu_1, \dots, \mu_t)$ and $\lambda = (\lambda_1, \dots, \lambda_t)$ be two nonzero partitions, where t is the maximum length of μ and λ.

Note that if $\mu \ge \lambda$, then $\operatorname{sum}\mu \ge \operatorname{sum}\lambda$; i.e., the integer $C = \operatorname{sum}\mu - \operatorname{sum}\lambda$ is nonnegative.

By definition of the dominance relation, a condition $\mu \ge \lambda$ is equivalent to the system of inequalities

$$\mu_1 + \cdots + \mu_k \ge \lambda_1 + \cdots + \lambda_k \quad (k = 1, \dots, t).$$

We write this system of inequalities in the following equivalent form:

$$\sum_{j=1,\, \mu_j > \lambda_j}^{k} (\mu_j - \lambda_j) \ge \sum_{j=1,\, \mu_j < \lambda_j}^{k} (\lambda_j - \mu_j) \quad (k = 1, \dots, t). \tag{2.1}$$

Here and below, in the case when no index satisfies a summation condition, we will assume that the corresponding sum is equal to 0.

For some integer $C \geq 0$, a condition $\mu \geq \lambda$ is equivalent to the system

$$\begin{cases} \mu_1 + \cdots + \mu_t & = \lambda_1 + \cdots + \lambda_t + C, \\ \mu_1 + \cdots + \mu_{k-1} & \geq \lambda_1 + \cdots + \lambda_{k-1} \quad (k = 2, \ldots, t). \end{cases} \tag{2.2}$$

Since the following equalities are true:

$$\mu_1 + \cdots + \mu_{k-1} + \mu_k + \cdots + \mu_t = \lambda_1 + \cdots + \lambda_{k-1} + \lambda_k + \cdots + \lambda_t + C \quad (k = 2, \ldots, t),$$

system (2.2) is equivalent to the system

$$\begin{cases} \mu_1 + \cdots + \mu_t & = \lambda_1 + \cdots + \lambda_t + C, \\ \mu_k + \cdots + \mu_t & \leq \lambda_k + \cdots + \lambda_t + C \quad (k = 2, \ldots, t). \end{cases} \tag{2.3}$$

Let us rewrite system (2.3) in the equivalent form:

$$\begin{cases} \displaystyle\sum_{j=1,\, \mu_j > \lambda_j}^{t} (\mu_j - \lambda_j) & = \displaystyle\sum_{j=1,\, \mu_j < \lambda_j}^{t} (\lambda_j - \mu_j) + C, \\ \displaystyle\sum_{j=k,\, \mu_j > \lambda_j}^{t} (\mu_j - \lambda_j) & \leq \displaystyle\sum_{j=k,\, \mu_j < \lambda_j}^{t} (\lambda_j - \mu_j) + C \quad (k = 2, \ldots, t). \end{cases} \tag{2.4}$$

For $i = 1, 2, \ldots, t$, we say that an i-component of a partition μ has an *i-hill* (or simply a *hill*) with respect to the partition λ if $\mu_i > \lambda_i$. In the case when the condition $\mu_i > \lambda_i$ is satisfied, we assume that upper $\mu_i - \lambda_i$ boxes of the i-column of the Ferrers diagram of the partition μ form the i-hill of height $\mu_i - \lambda_i$.

For $j = 1, 2, \ldots, t$, we say that a j-component of a partition μ has a *j-pit* (or simply a *pit*) with respect to the partition λ if $\mu_j < \lambda_j$. In the case when the condition $\mu_j < \lambda_j$ is satisfied, we assume that there is the j-pit of depth $\lambda_j - \mu_j$ over the j-column of the Ferrers diagram of the partition μ.

Let us reformulate conditions (2.1) as follows.

For any $k = 1, \ldots, t$, the sum of the heights of all i-hills such that $1 \leq i \leq k$ is greater than or equal to the sum of the depths of all j-pits such that $1 \leq j \leq k$.

Respectively, system (2.4) is equivalent to the following statement.

For some nonnegative integer C:

- the sum of the heights of all hills is equal to the integer C plus the sum of the depths of all pits;
- for any $k = 2, \ldots, t$, the sum of the heights of all i-hills such that $i \geq k$ does not exceed the integer C plus the sum of the depths of all j-pits such that $j \geq k$.

In what follows, we will assume that $\mu \geq \lambda$.

We will say that a j-pit is *admissible* if $\mu_{j-1} > \mu_j$. Note that the admissibility of a j-pit is a necessary condition for the possibility of moving the box to the j-column of the partition μ from some column with a number less than j.

We will call an i-hill of a partition μ *open* if $\mu_i > \mu_{i+1}$. Note that the openness of i-hill is a necessary condition for the possibility of moving the box from i-column of the partition μ to some column with a greater number than i or for removal of the box from i-column of the partition μ.

For $k = 1, 2, \ldots, t$, we say that a number k is a *separator* for the partition μ with respect to the partition λ if

$$\mu_1 + \cdots + \mu_k = \lambda_1 + \cdots + \lambda_k.$$

This means that
$$\sum_{j=1,\,\mu_j>\lambda_j}^{k} (\mu_j - \lambda_j) = \sum_{j=1\,\mu_j<\lambda_j}^{k} (\lambda_j - \mu_j).$$

This condition is equivalent to the following statement.

The sum of the heights of all i-hills such that $1 \le i \le k$ is equal to the sum of the depths of all j-pits such that $1 \le j \le k$.

Let us make a simple remark that, for any i-hill, the number i is not a separator for the partition μ with respect to λ.

Indeed, by the condition $\mu \ge \lambda$, we have
$$\mu_1 + \cdots + \mu_{i-1} \ge \lambda_1 + \cdots + \lambda_{i-1}.$$

Since $\mu_i > \lambda_i$, it follows $\mu_1 + \cdots + \mu_i > \lambda_1 + \cdots + \lambda_i$. Consequently, the number i is not a separator.

Lemma 1. *Assume that $\mu \ge \lambda$, the partition μ has an i-hill with respect to the partition λ, where $1 \le i \le t$, and a partition μ' is obtained from the partition μ by applying an elementary transformation of the second type, which consists in removing the top box from the i-hill. Then*

(1) *if, for the partition μ, there is a separator k such that $i < k \le t$, then the condition $\mu' \ge \lambda$ is not satisfied*;

(2) *if, for the partition μ, there are no separators k such that $i < k \le t$, then the condition $\mu' \ge \lambda$ is satisfied.*

P r o o f. 1. A separator k satisfies the condition
$$\mu_1 + \cdots + \mu_i + \cdots + \mu_k = \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k.$$

For the partition μ', we have $\mu'_i = \mu_i - 1$ and $\mu'_p = \mu_p$ for $p \ne i$ and $p = 1, \ldots, k$. Hence, we have the inequality
$$\mu'_1 + \cdots + \mu'_i + \cdots + \mu'_k < \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k.$$

Therefore, μ' does not dominate λ, i.e., the condition $\mu' \ge \lambda$ is not satisfied.

2. For any number k such that $i \le k \le t$, since it is not a separator for the partition μ, the condition
$$\mu_1 + \cdots + \mu_i + \cdots + \mu_k > \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k$$
is true. Since, for the partition μ', we have $\mu'_i = \mu_i - 1$ and $\mu'_p = \mu_p$ for $p \ne i$ and $p = 1, \ldots, k$, it follows that
$$\mu'_1 + \cdots + \mu'_i + \cdots + \mu'_k \ge \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k.$$

In addition, for any $k = 1, \ldots, i-1$, the condition
$$\mu'_1 + \cdots + \mu'_k \ge \lambda_1 + \cdots + \lambda_k$$
is true.

Therefore, μ' dominates λ, i.e., the condition $\mu' \ge \lambda$ is true. $\square$

Assume that $\mu \ge \lambda$, the partition μ has an i-hill with respect to the partition λ, where $1 \le i \le t$, and the partition μ' is obtained from the partition μ by applying an elementary transformation of the second type, which consists in removing the top box from the i-hill. We will call such an elementary transformation of the second type *proper* for a partition μ with respect to a partition λ if, for a partition μ, there are no separators k such that $i < k \le t$. Lemma 1 states that $\mu' \ge \lambda$ holds if and only if the corresponding elementary transformation of the second type is proper.

Lemma 2. *Assume that $\mu \geq \lambda$, the partition μ has an i-hill and a j-pit with respect to the partition λ, where $1 \leq i < j \leq t$, and a partition μ' is obtained from the partition μ by applying an elementary transformation of the first type, which consists in moving the upper box from the i-hill to the j-pit. Then*

(1) *if the partition μ has a separator k such that $i < k < j$, then the condition $\mu' \geq \lambda$ is not satisfied;*
(2) *if, for the partition μ, there are no a separators k such that $i < k < j$, then the condition $\mu' \geq \lambda$ is satisfied.*

P r o o f. 1. It can be proven in exactly the same way as (1) in Lemma 1.

2. Due to the remark made before Lemma 1 and the conditions of this lemma, the numbers k such that $i \leq k < j$ are not separators for the partition μ with respect to the partition λ, hence, for such numbers k, we have

$$\mu_1 + \cdots + \mu_i + \cdots + \mu_k > \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k.$$

For the partition μ', we have $\mu'_i = \mu_i - 1$ and $\mu'_p = \mu_p$ for $p \neq i$ and $p = 1, \ldots, k$. Consequently, we get

$$\mu'_1 + \cdots + \mu'_i + \cdots + \mu'_k \geq \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_k.$$

Further, we note that the following condition is true for any $k = 1, \ldots, i-1$:

$$\mu'_1 + \cdots + \mu'_k \geq \lambda_1 + \cdots + \lambda_k.$$

For any $k \geq j$, the following condition holds:

$$\mu_1 + \cdots + \mu_i + \cdots + \mu_j + \cdots + \mu_k \geq \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_j + \cdots + \lambda_k.$$

Consequently,

$$\mu_1 + \cdots + (\mu_i - 1) + \cdots + (\mu_j + 1) + \cdots + \mu_k \geq \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_j + \cdots + \lambda_k.$$

Hence, we have

$$\mu'_1 + \cdots + \mu'_i + \cdots + \mu'_j + \cdots + \mu'_k \geq \lambda_1 + \cdots + \lambda_i + \cdots + \lambda_j + \cdots + \lambda_k.$$

Therefore, the condition $\mu' \geq \lambda$ is satisfied. □

Assume that $\mu \geq \lambda$, the partition μ has an i-hill and has a j-pit with respect to the partition λ, where $1 \leq i < j \leq t$, and a partition μ' is obtained from the partition μ by applying an elementary transformation of the first type, which consists in moving the upper box from i-hill to the j-pit. Such an elementary transformation of the first type will be called *proper* for the partition μ with respect to the partition λ if, for μ, there are no separators k such that $i < k < j$. Lemma 2 states that $\mu' \geq \lambda$ holds if and only if the corresponding elementary transformation of the first type is proper.

Lemma 3. *Let $\mu \geq \lambda$. Then*

(1) *for every pit of the partition μ, there is a hill to the left of that pit;*
(2) *if there is a pit of the partition μ, then also there is an admissible pit.*

P r o o f. Let the partition μ have a j-pit with respect to the partition λ for some $j \in \{1, \dots, t\}$.

Since $\mu \geq \lambda$, by conditions (2.1), there exists an i-hill such that $1 \leq i < j$. We assume that the i-hill is the nearest hill to the left from the j-pit, where $1 \leq i < j$. Then there are no hills between that i-hill and the j-pit; i.e., there are no s-hills such that $i < s < j$.

Then a pit closest to that i-hill on the right is admissible. Indeed, let such a pit be located in the k-column, i.e., it is a k-pit, where $i < k \leq j$. Then $\mu_{k-1} \geq \lambda_{k-1} \geq \lambda_k > \mu_k$, i.e., $\mu_{k-1} > \mu_k$. □

Lemma 4. *Let $\mu \geq \lambda$, and let an i-hill be the nearest hill to the left for an admissible j-pit, where $1 \leq j \leq t$. Then this i-hill is open, $\mu_i \geq 2 + \mu_j$, and there are no separators k for the partition μ such that $i < k < j$.*

P r o o f. Note first that $\mu_i > \lambda_i \geq \lambda_{i+1} \geq \mu_{i+1}$, so $\mu_i > \mu_{i+1}$, i.e., the i-hill is open. Moreover, $\mu_i > \lambda_i \geq \lambda_j > \mu_j$, hence $\mu_i \geq 2 + \mu_j$.

Let us show that there are no separators k for the partition μ such that $i < k < j$.

Consider a number k such that $i < k < j$. Since $\lambda_j > \mu_j$ and there are no s-hills such that $k < s < j$, we successively obtain

$$\sum_{p=1,\, \mu_p > \lambda_p}^{k} (\mu_p - \lambda_p) = \sum_{p=1,\, \mu_p > \lambda_p}^{j} (\mu_p - \lambda_p) \geq \sum_{p=1,\, \mu_p < \lambda_p}^{j} (\lambda_p - \mu_p) > \sum_{p=1,\, \mu_p < \lambda_p}^{k} (\lambda_p - \mu_p).$$

Therefore, the condition with the number k from (2.1) is a strict inequality, i.e., the number k is not a separator for the partition μ. □

Let $\mu \geq \lambda$, and let an i-hill be the nearest hill to the left for an admissible j-pit, where $1 \leq j \leq t$. Then, by Lemma 4, the following conditions are true:

(1) $\mu_i - 1 \geq \mu_{i+1}$;
(2) $\mu_{j-1} \geq \mu_j + 1$;
(3) $\mu_i \geq 2 + \mu_i$.

This is a necessary condition for a possibility of applying the box movement from i-column to the j-column of the partition μ.

The corresponding elementary transformation of the first type will be call a *moving the upper box into an admissible pit from the hill closest to it on the left.* By Lemma 2, such a transformation is proper.

Corollary 1. *Let μ' be a partition obtained from a partition μ by moving the upper box into an admissible pit from the hill closest to it on the left. Then $\mu' \geq \lambda$.*

Lemma 5. *Assume that $\mu \geq \lambda$, the last hill of the partition μ with respect to the partition λ has a number i, and μ' is a partition obtained from the partition μ by removing the upper box from i-hill. If $C = \operatorname{sum} \mu - \operatorname{sum} \lambda > 0$, then the last i-hill of the partition μ is open and $\mu' \geq \lambda$.*

P r o o f. Note that $\mu_i > \lambda_i \geq \lambda_{i+1} \geq \mu_{i+1}$; i.e., $\mu_i > \mu_{i+1}$, i.e., the i-column of the Ferrers diagram of the partition μ is open, and, therefore, an elementary transformation of the second type is applicable, which consists in removing of the upper box from i-hill. Let $C > 0$.

When passing from μ to μ' and replacing C by $C - 1$, the first condition of system (2.4) is preserved, since the sums decrease by 1 on the left and on the right.

When passing from μ to μ' and replacing C with $C - 1$, the second condition of system (2.4) for $k \leq i$ is preserved for the same reason.

Let $k > i$. Then the partition μ satisfies

$$0 = \sum_{j=k,\, \mu_j > \lambda_j}^{t} (\mu_j - \lambda_j) \leq \sum_{j=k,\, \mu_j < \lambda_j}^{t} (\lambda_j - \mu_j) + C.$$

Since $C > 0$ and $\mu_j = \mu'_j$ for $j \geq k$, we get

$$0 = \sum_{j=k,\, \mu'_j > \lambda_j}^{t} (\mu'_j - \lambda_j) \leq \sum_{j=k,\, \mu'_j < \lambda_j}^{t} (\lambda_j - \mu'_j) + (C - 1).$$

□

Lemmas 1 and 5 imply

Corollary 2. *Let μ' be a partition obtained from a partition μ by removing the top box from the last hill with number i. Then, for the partition μ, there are no separators k such that $i < k \leq t$.*

We now fix partitions μ and λ such that $\mu > \lambda$. Consider sequences of elementary transformations from μ to λ (both types of elementary transformations are admissible):

$$\mu = \xi_{(0)} \rightharpoondown \xi_{(1)} \rightharpoondown \cdots \rightharpoondown \xi_{(s)} = \lambda. \tag{2.5}$$

Let us construct now an algorithm for finding all possible shortest sequences of this kind.

Let $C = \operatorname{sum} \mu - \operatorname{sum} \lambda$. Since $\mu > \lambda$, we have $C \geq 0$. Obviously, in the sequence (2.5), there are exactly C elementary transformations of the second type, since elementary transformations of the second type reduce the weight of the partition by 1, and elementary transformations of the first type preserve the weight of the partition.

On the other hand, if some sequence of elementary transformations transforms μ into λ, then each box contained in any of the hills must be removed or moved. Therefore, the number of elementary transformations in such a sequence is not less than the sum of the heights of all hills. By (2.4), the sum of the heights of all hills is equal to C plus the sum of the depths of all pits. It is clear that all pits must be eliminated when passing from μ to λ in accordance with (2.5). Therefore, in sequence (2.5), there are at least p movements of the boxes to pits, where p is equal to the total depth of all pits. This implies that $s \geq C + p$.

The following algorithm constructs all shortest sequences of length $C + p$ of elementary transformations from μ to λ.

Algorithm 1. Let $\mu > \lambda$ and $C = \operatorname{sum} \mu - \operatorname{sum} \lambda$.

1. We set $\eta = \mu$ and $C' = C$.

2. To a current partition η and a number C', we apply any of the following proper elementary transformations for the partition η with respect to the partition λ:

   • if η has a pit, then we replace η with the partition obtained from η by moving the upper box from some open i-hill to some admissible j-pit for which there are no separators k such that $i < k < j$;

   • if $C' > 0$, then we replace C' with $C' - 1$ and replace the partition η with the partition obtained from η by removing the top box from some i-hill for which there are no separators k such that $i < k \leq t$.

3. Do step 2 as long as possible. The process will definitely end. The performed transformations will form the shortest sequence of elementary transformations from μ to λ. Its length is equal to the sum of the heights of all hills of the partition μ with respect to the partition λ and it is equal to

$$C + p = \frac{1}{2}C + \frac{1}{2}\sum_{j=1}^{\infty} |\mu_j - \lambda_j|.$$

Theorem 2. *Algorithm 1 is correct. Every shortest sequence of elementary transformations of the form (2.5) can be obtain by appropriate application of this algorithm.*

P r o o f. By Lemmas 1 and 2, all the constructed partitions η satisfy the condition $\eta \geq \lambda$.

If the current partition η satisfies $\eta > \lambda$, then, by Lemma 4 and Corollary 2, step 2 of the algorithm can be continued. It is clear that the algorithm will complete its work and the shortest sequence of the form (2.5) will be constructed. Its length is equal to the sum of the heights of all the hills of the partition μ with respect to the partition λ, i.e., it is equal to $C + p$. It is not difficult to see that

$$C + p = \frac{1}{2}C + \frac{1}{2}\sum_{j=1}^{\infty} |\mu_j - \lambda_j|.$$

Let sequence (2.5) be the shortest sequence of length $C + p$. There are C removals of boxes in this sequence. It must be eliminated all hills with the total height $C + p$, where p is equal to the total depth of all pits. Therefore, that sequence (2.5) consists of exactly C removals of boxes from hills and exactly p moves of boxes from hills. Because it must be eliminated all pits with the total depth p, every transformation $\xi_{(k-1)} \rightharpoondown \xi_{(k)}$ $(k = 1, \ldots, s)$ consists in removing the box from a hill or consists in moving a box from a hill to a pit. Due to the conditions $\xi_{(k-1)} \geq \xi_{(k)} \geq \lambda$ and by Lemmas 1 and 2, every elementary transformation $\xi_{(k-1)} \rightharpoondown \xi_{(k)}$ is a proper elementary transformation with respect to the partition λ.

Therefore, a simple execution of Algorithm 1 along the shortest sequence (2.5) is correct. $\square$

For applications the following special case of Algorithm 1 is useful which generally speaking does not construct all shortest sequences of the form (2.5).

Algorithm 2. [4] Let $\mu > \lambda$ and $C = \operatorname{sum}\mu - \operatorname{sum}\lambda$.

1. We set $\eta = \mu$ and $C' = C$.

2. Apply any of the following possible elementary transformations to a current partition η and an integer C':

   - if η has a pit, then we replace η with the partition obtained from η by moving the upper box into some admissible j-pit from the hill closest to it on the left;

   - if $C' > 0$, then we replace C' with $C' - 1$ and replace the partition η with the partition obtained from η by removing the top box from the last hill.

3. Do step 2 as long as possible. The process will definitely end. The performed transformations will form the shortest sequence of elementary transformations from μ to λ. Its length is equal to the sum of the heights of all hills of the partition μ with respect to the partition λ and is equal to

$$C + p = \frac{1}{2}C + \frac{1}{2}\sum_{j=1}^{\infty} |\mu_j - \lambda_j|.$$

Example 1. Let $\mu = (12, 8, 8, 4, 3, 2, 2, 2, 0, 0)$ and $\lambda = (10, 10, 6, 3, 3, 2, 1, 1, 1, 1)$. We have $\operatorname{sum}\mu = 41$, $\operatorname{sum}\lambda = 38$, and $C = 3$.

Consider the component-wise difference between μ and λ:

$$\begin{array}{rllllllllll} \mu = & (12, & 8, & 8, & 4, & 3, & 2, & 2, & 2, & 0, & 0)^3, \\ \lambda = & (10, & 10, & 6, & 3, & 3, & 2, & 1, & 1, & 1, & 1), \\ \delta = \mu - \lambda = & (+2, & -2, & +2, & +1, & 0, & 0, & +1, & +1, & -1, & -1), \\ \Delta = & (2, & \underline{0}, & 2, & 3, & 3, & 3, & 4, & 5, & 4, & 3). \end{array}$$

Here, we have five hills and three pits, $t = 10$, and the height(μ, λ) is equal to the sum of the heights of all hills, i.e., it is equal to 7. At the end of the notation of μ at the top, the number 3 is indicated, which is equal to the number of boxes to be removed when working Algorithm 1.

Here, the sequence $\Delta = (\Delta_1, \Delta_2, \dots, \Delta_{10})$ is given by the condition

$$\Delta_k = (\mu_1 + \dots + \mu_k) - (\lambda_1 + \dots + \lambda_k),$$

i.e., $\Delta_k = \Delta_{k-1} + \delta_k$. It is clear that the condition $\mu \geq \lambda$ is equivalent to the fact that $\Delta_k \geq 0$ for any $k = 1, 2, \dots, t$. The sequence Δ has only one zero $\Delta_2 = 0$. This zero underlined below in Δ. Hence, there is exactly one separator (the integer 2) for μ with respect to λ.

Note that the 1-hill is open. Using Algorithm 1, we cannot remove boxes from the 1-hill ("move them across the separator"). We can move boxes from the 1-hill only to the admissible 2-pit.

The 9-pit is admissible but the 10-pit is not. We can move a box from any hills with numbers 3, 4, and 8 to the 9-pit since they are open and there are no separators between these hills and the 9-pit. Note that the 7-hill is not open.

Let us move the box from the open 3-hill to the admissible 9-pit:

$$\begin{array}{rllllllllll} \mu = & (12, & 8, & 7, & 4, & 3, & 2, & 2, & 2, & 1, & 0)^3, \\ \lambda = & (10, & 10, & 6, & 3, & 3, & 2, & 1, & 1, & 1, & 1), \\ \delta = \mu - \lambda = & (+2, & -2, & +1, & +1, & 0, & 0, & +1, & +1, & 0, & -1), \\ \Delta = & (2, & \underline{0}, & 1, & 2, & 2, & 2, & 3, & 4, & 4, & 3). \end{array}$$

Note that, after such a transformation, the 10-pit became admissible.

Now let us remove the box from the 3-hill, which remained open:

$$\begin{array}{rllllllllll} \mu = & (12, & 8, & 6, & 4, & 3, & 2, & 2, & 2, & 1, & 0)^2, \\ \lambda = & (10, & 10, & 6, & 3, & 3, & 2, & 1, & 1, & 1, & 1), \\ \delta = \mu - \lambda = & (+2, & -2, & 0, & +1, & 0, & 0, & +1, & +1, & 0, & -1), \\ \Delta = & (2, & \underline{0}, & \underline{0}, & 1, & 1, & 1, & 2, & 3, & 3, & 2). \end{array}$$

Note that another separator has appeared — the integer 3. In addition, we have replaced the counter value of the number of boxes to be deleted by 2.

Continuing to apply elementary transformations in the same spirit according to Algorithm 1, we will find some shortest sequence of elementary transformations of length 7 that transforms μ into λ.

Note that we can remove a box from an open hill at any step of the algorithm execution if the value of the counter of deleted boxes is greater than zero; it is only important that there is no any separator to the right of the hill.

An example of an operation of Algorithm 2 see in [4].

3. Conclusion

We note that finding the shortest chains of elementary transformations is an important problem in studying the properties of graphic partitions. The use of Algorithm 2 allowed us to obtain several interesting properties of graphic partitions (see, for example, [5, 6]). Using Algorithm 1 opens up more possibilities.

REFERENCES

1. Andrews G. E. *The Theory of Partitions.* Cambridge: Cambridge University Press, 1984. 255 p. DOI: 10.1017/CBO9780511608650
2. Baransky V. A., Koroleva T. A., Senchonok T. A. On the partition lattice of an integer. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2015. Vol. 21, No. 3. P. 30–36. (in Russian)
3. Baransky V. A., Koroleva T. A., Senchonok T. A. On the partition lattice of all integers. *Sib. Electron. Mat. Izv.*, 2016. Vol. 13. P. 744–753. DOI: 10.17377/semi.2016.13.060 (in Russian)
4. Baransky V. A., Senchonok T. A. On the shortest sequences of elementary transformations in the partition lattice. *Sib. Electron. Mat. Izv.*, 2018. Vol. 15. P. 844–852. DOI: 10.17377/semi.2018.15.072 (in Russian)
5. Baransky V. A., Senchonok T. A. Bipartite-threshold graphs and lifting rotations of edges in bipartite graphs. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2023. Vol. 29, No. 1. P. 24–35. DOI: 10.21538/0134-4889-2023-29-1-24-35 (in Russian)
6. 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
7. Brylawski T. The lattice of integer partitions. *Discrete Math.*, 1973. Vol. 6, No. 3. P. 201–219. DOI: 10.1016/0012-365X(73)90094-0

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 46–59

DOI: 10.15826/umj.2023.2.004

$\mathcal{I}^{\mathcal{K}}$-SEQUENTIAL TOPOLOGY

H. S. Behmanush[†], **M. Küçükaslan**[††]

Mersin University, Science Faculty, Department of Mathematics,
33100 Yenişehir, Mersin, Turkey

[†]hsbehmanush1989@gmail.com [††]mkucukaslan@mersin.edu.tr

Abstract: In the literature, $\mathcal{I}$-convergence (or convergence in $\mathcal{I}$) was first introduced in [11]. Later related notions of $\mathcal{I}$-sequential topological space and $\mathcal{I}^*$-sequential topological space were introduced and studied. From the definitions it is clear that $\mathcal{I}^*$-sequential topological space is larger(finer) than $\mathcal{I}$-sequential topological space. This rises a question: is there any topology (different from discrete topology) on the topological space $\mathcal{X}$ which is finer than $\mathcal{I}^*$-topological space? In this paper, we tried to find the answer to the question. We define $\mathcal{I}^{\mathcal{K}}$-sequential topology for any ideals $\mathcal{I}$, $\mathcal{K}$ and study main properties of it. First of all, some fundamental results about $\mathcal{I}^{\mathcal{K}}$-convergence of a sequence in a topological space $(\mathcal{X}, \mathcal{T})$ are derived. After that, $\mathcal{I}^{\mathcal{K}}$-continuity and the subspace of the $\mathcal{I}^{\mathcal{K}}$-sequential topological space are investigated.

Keywords: Ideal convergence, $\mathcal{I}^{\mathcal{K}}$-convergence, Sequential topology, $\mathcal{I}^{\mathcal{K}}$-sequential topology.

1. Introduction

The notion of convergence of real or complex valued sequences was generalized using asymptotic density and was called statistical convergence by Fast [7] and Steinhause [20] in the same year 1951, independently. After some years P. Kostyrko, T. Šalát, W. Wilczyńki [11] gave a generalization of statistical convergence and called it as ideal convergence (or converges in ideal). Various fundamental properties (convergence in $\mathcal{I}$ and $\mathcal{I}^*$) were investigated. Later B.K. Lahiri and P. Das in [12] discussed convergence in $\mathcal{I}$ and in $\mathcal{I}^*$ and investigate some additional results related to mentioned concepts [4, 8–10, 15–17].

The concept of $\mathcal{I}^*$-convergence of functions was extended to $\mathcal{I}^{\mathcal{K}}$-convergence by M. Mačaj and M. Sleziak in [13] in 2011. The authors of [2, 3, 5, 6, 14] gave further properties and results about $\mathcal{I}^{\mathcal{K}}$-convergence.

In first part of this paper we introduce $\mathcal{I}^{\mathcal{K}}$-sequential topological (seq.-top.) space, which is a natural generalization of $\mathcal{I}^*$-seq.-top. space. Later we discuss the $\mathcal{I}^{\mathcal{K}}$-continuity of the function and in last two section we write about $\mathcal{I}^{\mathcal{K}}$-subspace and $\mathcal{I}^{\mathcal{K}}$-connectedness. We will use further the abbreviation T.S. for a topological space.

2. Definition and preliminaries

In this part, we give some known definitions and necessary results.

Definition 1 [7, 20]. *Let $\mathcal{A} \subset \mathbb{N}$, and for $m \in \mathbb{N}$ let the set*

$$\mathcal{A}_m := \{x \in \mathcal{A} : x < m\}$$

and $|\mathcal{A}_m|$ stand for the cardinality of $\mathcal{A}_m$. Natural density of $\mathcal{A}$ is defined by

$$\beta(\mathcal{A}) := \lim_{m\to\infty} \frac{|\mathcal{A}_m|}{m}$$

whenever the limit exists. A real sequence $\tilde{x} = (x_i)$ *is said to statistically converges to* x_0 *if for any* $\varepsilon > 0$,

$$\beta\big(\{n : |x_i - x_0| > \varepsilon\}\big) = 0$$

holds.

Definition 2 [11]. *Let* $\mathcal{I}$ *be any subfamily of* $\mathcal{P}(\mathbb{N})$, *with* $\mathcal{P}(\mathbb{N})$ *being the family of all subsets of* $\mathbb{N}$. *Then,* $\mathcal{I}$ *is called an ideal on* $\mathbb{N}$ *if the following requirements hold*:

(i) *finite union of sets in* $\mathcal{I}$ *is again in* $\mathcal{I}$;
(ii) *any subset of a set in* $\mathcal{I}$ *is in* $\mathcal{I}$.

$\mathcal{I}$ is admissible if all singleton subsets of $\mathbb{N}$ belong to $\mathcal{I}$. The ideal $\mathcal{I}$ is non-trivial if $\mathcal{I} \neq \varnothing$ and $\mathcal{I} \neq \mathcal{P}(\mathbb{N})$. A non-trivial ideal $\mathcal{I}$ is called proper if $\mathbb{N}$ is not in $\mathcal{I}$.

The family of finite subsets of the $\mathbb{N}$ is an admissible non-trivial ideal denoted by $\mathcal{F}in$ and the family of the subsets of $\mathbb{N}$ with natural density zero is also an admissible non-trivial ideal denoted by $\mathcal{I}_\beta$. The set of all non-trivial admissible ideals will be denoted as NA throughout the study.

Example 1. [11] Consider the decomposition of $\mathbb{N}$ as $\mathbb{N} = \bigcup_{j=1}^{\infty} \beta_j$ where all β_j are infinite subsets of $\mathbb{N}$ and are mutually disjoint. Take the family

$$\mathcal{I} = \{N \subset \mathbb{N} : N \text{ intersect only finite number of } \beta_j's\}.$$

Then, $\mathcal{I}$ belongs to NA.

Definition 3 [19]. *Assume* $\mathcal{F} \subset \mathcal{P}(\mathbb{N})$. *The collection* $\mathcal{F}$ *is a filter on* $\mathbb{N}$ *if*

(i) *a finite intersection of elements of* $\mathcal{F}$ *is in* $\mathcal{F}$ *and*
(ii) *if* $\mathcal{C} \in \mathcal{F} \wedge \mathcal{C} \subseteq \mathcal{D}$, *then* $\mathcal{D} \in \mathcal{F}$.

If empty set is not in $\mathcal{F}$ then $\mathcal{F}$ is proper. If $\mathcal{I} \in NA$ then the collection

$$\mathcal{F} = \{N \subset \mathbb{N} : N^C \in \mathcal{I}\}$$

is a filter on $\mathbb{N}$. It is known as the $\mathcal{I}$-associated filter.

Definition 4 [21]. *In a T.S.* $(\mathcal{X}, \mathcal{T})$ *a sequence* $\tilde{x} = (x_i) \subset \mathcal{X}$ *is called to converging in* $\mathcal{I}$ *to a point* $x \in \mathcal{X}$ *if*

$$\{i \in \mathbb{N} : \ x_i \in \upsilon\} \in \mathcal{F}(\mathcal{I})$$

holds for each neighborhood υ *of* x. *The point* x *is referred to as the ideal limit of the sequence* $\tilde{x} = (x_i)$ *and it is represented by* $x_i \xrightarrow{\mathcal{I}} x$ (*or* $\mathcal{I} - \lim x_i = x$).

Remark 1.

(i) Statistical and $\mathcal{I}_\beta-$ convergence are coincide.
(ii) Classical convergence and $\mathcal{F}in-$convergence are coincide.

Lemma 1 [1]. *Assume that* $\mathcal{I}, \mathcal{I}_1$ *and* $\mathcal{I}_2$ *be ideals on the set* $\mathbb{N}$ *and consider a T.S.* $(\mathcal{X}, \mathcal{T})$, *then*

1. *If* $\mathcal{I} \in NA$, *then every convergent sequence is* $\mathcal{I}$*-convergent sequence which converges to same point.*
2. *If* $\mathcal{I}_1 \subseteq \mathcal{I}_2$ *and* $(x_i) \subseteq \mathcal{X}$ *is a sequence which* $x_i \xrightarrow{\mathcal{I}_1} x$, *then* $x_i \xrightarrow{\mathcal{I}_2} x$.
3. *If* $\mathcal{X}$ *the Hausdorff space, then the limit of every convergent sequence is unique.*

3. $\mathcal{I}^{\mathcal{K}}$-convergence of sequence

In this part we will investigate some results related to $\mathcal{I}^{\mathcal{K}}$-convergence of sequences which is a generalized form of $\mathcal{I}^*$-convergence of sequences. If we consider $\mathcal{F}in$ instead of $\mathcal{K}$, then we will have $\mathcal{I}^*$-convergence.

Definition 5 [6]. *In a T.S. $(\mathcal{X}, \mathcal{T})$ a sequence $\tilde{x} = (x_i) \subset \mathcal{X}$ is called to be $\mathcal{I}^*$-converging to $x_0 \in \mathcal{X}$ if $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the sequence*

$$y_i := \begin{cases} x_i, & i \in M, \\ x, & i \notin M \end{cases}$$

is $\mathcal{F}in$ convergent to x.

That is, for each neighborhood υ of x,

$$\{i \in \mathbb{N} : \; y_i \in \upsilon\} \in \mathcal{F}(\mathcal{F}in),$$

or

$$\{i \in M : \; y_i \notin \upsilon\} \cup \{i \in M^C : \; y_i \notin \upsilon\} \in \mathcal{F}in.$$

So,

$$\{i \in M : \; x_i \notin \upsilon\} \cup \{i \in M^C : \; x \notin \upsilon\} \in \mathcal{F}in.$$

This implies that

$$\{i \in M : \; y_i \notin \upsilon\} \in \mathcal{F}in.$$

Therefore,

$$\{i \in M : \; y_i \in \upsilon\} \in \mathcal{F}(\mathcal{F}in).$$

It is clear that this definition is the same as the definition given in [6]. In the definition of $\mathcal{I}^*$-convergence of sequence if we consider an arbitrary ideal $\mathcal{K}$ instead of the ideal $\mathcal{F}in$ then it yields the definition of $\mathcal{I}^{\mathcal{K}}$-convergence of a sequence. That is, $\mathcal{I}^{\mathcal{K}}$-convergence is the generalized form of $\mathcal{I}^*$-convergence.

Definition 6 [13]. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider a T.S. $(\mathcal{X}, \mathcal{T})$. The sequence $\tilde{x} = (x_i) \subset \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-convergent to a point $x \in \mathcal{X}$ if $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the sequence*

$$y_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M, \end{cases}$$

$\mathcal{K}$-converges to x. We represent it as $\mathcal{I}^{\mathcal{K}} - \lim(x_i) = x$ or $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$.

Definition 7. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represent a T.S. Consider the sequences $\tilde{x} = (x_i) \subset \mathcal{X}$ and $\tilde{y} = (y_i) \subset \mathcal{X}$. Define a relation $\sim_{\mathcal{I}}$ as*

$$\tilde{x} \sim_{\mathcal{I}} \tilde{y} \Leftrightarrow \{i : x_i \neq y_i\} \in \mathcal{I}.$$

The relation $\sim_{\mathcal{I}}$ is an equivalence relation. That is,

1. $\forall \; \tilde{x} = (x_i) \subset \mathcal{X}$, $\{i : x_i \neq x_i\} = \varnothing \in \mathcal{I} \Rightarrow \tilde{x} \sim_{\mathcal{I}} \tilde{x}$.
2. Let $\tilde{x} \sim_{\mathcal{I}} \tilde{y}$. Since $\{i : y_i \neq x_i\} = \{i : x_i \neq y_i\} \in \mathcal{I}$, then $\tilde{y} \sim_{\mathcal{I}} \tilde{x}$.
3. Let $\tilde{x} \sim_{\mathcal{I}} \tilde{y}$ and $\tilde{y} \sim_{\mathcal{I}} \tilde{z}$. Then, $A := \{i : x_i = y_i\} \in \mathcal{F}(\mathcal{I})$ and $B := \{i : y_i = z_i\} \in \mathcal{F}(\mathcal{I})$. So, $\{i : x_i = z_i\} = A \cap B \in \mathcal{F}(\mathcal{I})$. Hence, $\tilde{x} \sim_{\mathcal{I}} \tilde{z}$ holds.

Lemma 2. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$ and the sequences $\tilde{x} = (x_i) \subseteq \mathcal{X}$. Assume $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ for any $x \in \mathcal{X}$ and $\tilde{t} = (t_i) \subseteq \mathcal{X}$ is a sequence s.t. $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$. Then, the sequence $t_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$.*

P r o o f. Let $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$, then $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the following sequence

$$y_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to x. Since $(x_i) \sim_{\mathcal{I}} (t_i)$. So $\forall i \in M$, $x_i = t_i$. Therefore, the following sequence

$$y_i = \begin{cases} t_i, & i \in M, \\ x, & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to x which shows that $t_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ holds.

□

The Definition 7 gives the possibility that the definition of $\mathcal{I}^{\mathcal{K}}$-convergence of a sequence can be rewritten as follows:

Definition 8. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$. A sequence $\tilde{x} = (x_i) \subset X$ is $\mathcal{I}^{\mathcal{K}}$-convergent to the point $x \in \mathcal{X}$ if there exist a sequence $\tilde{t} = (t_i) \subset X$ s.t. $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$ and $t_i \xrightarrow{\mathcal{K}} x$ holds.*

In the following lemma we demonstrate that Definition 6 and Definition 8 are equivalent for any ideals $\mathcal{I}$ and $\mathcal{K}$ and for any T.S. $(\mathcal{X}, \mathcal{T})$.

Lemma 3. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$ and $\tilde{x} = (x_i) \subset \mathcal{X}$ be a sequence. Then, $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ iff $\exists\, \tilde{t} = (t_i) \subset \mathcal{X}$ s.t. $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$ and $t_i \xrightarrow{\mathcal{K}} x$ hold.*

P r o o f. Let $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ holds. Then, $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the following sequence

$$y_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to x. Let us chose $(t_i) = (y_i)$ $\forall i \in \mathbb{N}$. Then, the proof will complete if we show that $\tilde{x} \sim_{\mathcal{I}} \tilde{y}$.

Consider the fact $\{i \in \mathbb{N} : x_i = y_i\} = \{i \in M : x_i = y_i\} \in \mathcal{F}(\mathcal{I})$. Hence, $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$.

Conversely, let $\tilde{x} = (x_i)$ and $\tilde{t} = (t_i)$ be sequences s.t. $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$ and $t_i \xrightarrow{\mathcal{K}} x$ hold. Since $\tilde{x} \sim_{\mathcal{I}} \tilde{t}$, then

$$M = \{i \in \mathbb{N} : x_i = t_i\} \in \mathcal{F}(\mathcal{I})$$

holds. Define a sequence

$$y_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M. \end{cases}$$

Since $x_i = t_i$ hold $\forall i \in M$, then we can write

$$t_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M. \end{cases}$$

Because $\tilde{t} = (t_i)$ is $\mathcal{K}$-convergent to x, the sequence $\tilde{y} = (y_i)$ is also $\mathcal{K}$-convergent to x. Hence, the sequence $\tilde{x} = (x_i)$ is $\mathcal{I}^{\mathcal{K}}$-convergent to the point x and this completes the proof. □

4. $\mathcal{I}^{\mathcal{K}}$-seq.-top. space

In this section, we are going to define a new topology on the $\mathcal{X}$ using the ideal $\mathcal{I}$ and $\mathcal{K}$ and investigate some properties of the new T.S. This topology will be an extended version of the $\mathcal{I}^*$-seq.-top. space which was discussed in [18]. If we take $\mathcal{I} = \mathcal{F}in$, then $\mathcal{I}^{\mathcal{K}}$-seq.-top. space is coincide with $\mathcal{I}^*$-T.S.

Definition 9. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$. Then*

1. *A set $F \subseteq \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-closed, if for each $(x_i) \subseteq F$ with $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$, then $x \in F$.*
2. *A set $V \subset \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-open, if its complement V^C is $\mathcal{I}^{\mathcal{K}}$-closed.*

Remark 2. Consider the T.S. $(\mathcal{X}, \mathcal{T})$. An $O \subset \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-open iff each sequence in $\mathcal{X} - O$ has $\mathcal{I}^{\mathcal{K}}$-limit in $\mathcal{X} - O$.

P r o o f. The proof is evident from Definition 9. Therefore, it is omitted here. □

Definition 10. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$. For any subset $A \subseteq \mathcal{X}$ define a set $\overline{A}^{\mathcal{I}^{\mathcal{K}}}$ (it is called $\mathcal{I}^{\mathcal{K}}$-closure of A) by*

$$\overline{A}^{\mathcal{I}^{\mathcal{K}}} := \{x \in \mathcal{X} : \exists (x_i) \subseteq A,\ x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x\}.$$

It is clear that $\overline{\varnothing}^{\mathcal{I}^{\mathcal{K}}} = \varnothing$, $\overline{\mathcal{X}}^{\mathcal{I}^{\mathcal{K}}} = \mathcal{X}$, and $A \subseteq \overline{A}^{\mathcal{I}^{\mathcal{K}}}$ holds $\forall A \subseteq \mathcal{X}$.

Remark 3. A subset C of the T.S. $\mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$ closed set iff $\overline{C}^{\mathcal{I}^{\mathcal{K}}} = C$.

P r o o f. Proof is obvious from the Definition 10. So, it is omitted here. □

Lemma 4. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and let $(\mathcal{X}, \mathcal{T})$ represent a T.S. For any subset $A \subseteq \mathcal{X}$, $\mathcal{I}^{\mathcal{K}}$-closure of A is $\mathcal{I}^{\mathcal{K}}$-closed.*

P r o o f. We must show that

$$\overline{\left(\overline{A}^{\mathcal{I}^{\mathcal{K}}}\right)}^{\mathcal{I}^{\mathcal{K}}} = \overline{A}^{\mathcal{I}^{\mathcal{K}}}.$$

It is clear that

$$\overline{A}^{\mathcal{I}^{\mathcal{K}}} \subset \overline{\left(\overline{A}^{\mathcal{I}^{\mathcal{K}}}\right)}^{\mathcal{I}^{\mathcal{K}}}.$$

Let $x \in \overline{\left(\overline{A}^{\mathcal{I}^{\mathcal{K}}}\right)}^{\mathcal{I}^{\mathcal{K}}}$. Then, there exist a sequence $(x_i) \subset \overline{A}^{\mathcal{I}^{\mathcal{K}}}$ s.t. $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ holds. Since $(x_i) \subset \overline{A}^{\mathcal{I}^{\mathcal{K}}}$, then there exist sequences $(x_i^n) \subset A$ s.t. $x_i^n \xrightarrow{\mathcal{I}^{\mathcal{K}}} x_i$. Therefore there exist the sets $M_n \in \mathcal{F}(\mathcal{I})$ s.t.

$$\{i \in M_n : x_i^n \notin \upsilon^n\} \in \mathcal{K}$$

for each neighborhood υ^n of x_i. Choose m_1 the i where x_i^1 is belonging to neighborhood υ^1 of x_1, similarly m_2 the i where x_i^2 is belonging to neighborhood υ^2 of x_2. If we continue this process and take m_p the i where x_i^p is belonging to neighborhood υ^n of x_p. The obtained sequence (x_{m_p}) belongs to A. The theorem will be proved if we show that $x_{m_p} \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$. Since $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$, so $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the sequence

$$y_i = \begin{cases} x_i, & i \in M, \\ x, & i \notin M, \end{cases} \quad y_i \xrightarrow{\mathcal{K}} x.$$

So,

$$\{i \in M : x_i \notin \upsilon\} \in \mathcal{K}$$

for each neighborhood υ of x. Now,

$$\{i \in M : \upsilon^n \not\subset \upsilon\} \subseteq \{i \in M : x_i \notin \upsilon\} \in \mathcal{K}.$$

Therefore,

$$\{i \in M : \upsilon^n \not\subset \upsilon\} \in \mathcal{K}$$

and

$$\{i \in M : x_{m_p} \notin \upsilon\} \subset \{i \in M : \upsilon^n \not\subseteq U\} \in \mathcal{K}$$

hold. So, $x_{m_p} \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$ and $x \in \overline{A}^{\mathcal{I}^{\mathcal{K}}}$. □

Definition 11. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represent a T.S. Then, for $A \subset \mathcal{X}$, $\mathcal{I}^{\mathcal{K}}$-interior of A is defined as*

$$A^{\circ \mathcal{I}^{\mathcal{K}}} := A - (\overline{\mathcal{X} - A}^{\mathcal{I}^{\mathcal{K}}}).$$

Proposition 1. *Let $\mathcal{V}$ be a subset of T.S. $\mathcal{X}$, then $\mathcal{V}$ is $\mathcal{I}^{\mathcal{K}}$-open iff $\mathcal{V}^{\circ \mathcal{I}^{\mathcal{K}}} = \mathcal{V}$.*

P r o o f. Let $\mathcal{V}$ be an $\mathcal{I}^{\mathcal{K}}$-open set. Then, $\mathcal{X} - \mathcal{V}$ is $\mathcal{I}^{\mathcal{K}}$-closed set and

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - \mathcal{V}) = \mathcal{X} - \mathcal{V}$$

holds. So, we have

$$\mathcal{V}^{\circ \mathcal{I}^{\mathcal{K}}} = \mathcal{V} - (\mathcal{X} - \mathcal{V}) = \mathcal{V}.$$

Conversely assume that

$$\mathcal{V}^{\circ \mathcal{I}^{\mathcal{K}}} = \mathcal{V}$$

holds. From the definition of $\mathcal{I}^{\mathcal{K}}$-interior of $\mathcal{V}$ we have

$$\mathcal{V} = \mathcal{V} - (\overline{\mathcal{X} - \mathcal{V}}^{\mathcal{I}^{\mathcal{K}}}).$$

Hence,

$$\mathcal{V} \cap \overline{\mathcal{X} - \mathcal{V}}^{\mathcal{I}^{\mathcal{K}}} = \varnothing.$$

Consequently

$$\overline{\mathcal{X} - \mathcal{V}}^{\mathcal{I}^{\mathcal{K}}} \subset \mathcal{X} - \mathcal{V}.$$

Thus,

$$\overline{\mathcal{X} - \mathcal{V}}^{\mathcal{I}^{\mathcal{K}}} = \mathcal{X} - \mathcal{V}$$

is satisfied. Therefore, $\mathcal{X} - \mathcal{V}$ is $\mathcal{I}^{\mathcal{K}}$-closed and $\mathcal{V}$ is $\mathcal{I}^{\mathcal{K}}$-open. □

Definition 12 [21]. *A sequence (x_i) in a T.S. $\mathcal{X}$ is $\mathcal{I}$-eventually in a subset A of $\mathcal{X}$ if*

$$\{i \in \mathbb{N} : x_i \in A\} \in \mathcal{F}(\mathcal{I}).$$

Definition 13. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X},\mathcal{T})$. A sequence $\tilde{x}=(x_i)\subseteq\mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-eventually in a subset $\mathcal{V}$ of $\mathcal{X}$. If there exist a sequence $\tilde{y}=(y_i)\subseteq\mathcal{X}$ s.t. $\tilde{y}\sim_I\tilde{x}$ and $\tilde{y}$ is $\mathcal{K}$-eventually in $\mathcal{V}$.*

In the next theorem, we will provide a sequence characterization of $\mathcal{I}^{\mathcal{K}}-$ open set.

Theorem 1. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X},\mathcal{T})$. A subset υ of $\mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-open iff each $\mathcal{I}^{\mathcal{K}}$-convergent sequence to $x_0\in\upsilon$ is $\mathcal{I}^{\mathcal{K}}$-eventually in υ.*

P r o o f. Let υ is $\mathcal{I}^{\mathcal{K}}$-open. Then, $\mathcal{X}-\upsilon$ is $\mathcal{I}^{\mathcal{K}}$-closed and $\overline{\mathcal{X}-\upsilon}^{\mathcal{I}^{\mathcal{K}}}=\mathcal{X}-\upsilon$ holds. Let $\tilde{x}=(x_i)\subset\mathcal{X}$ be a sequence s.t. $x_i\xrightarrow{\mathcal{I}^{\mathcal{K}}}x$ and $x\in\upsilon$. Then, $\exists M\in\mathcal{F}(\mathcal{I})$ s.t. the sequence

$$t_i=\begin{cases}x_i, & i\in M,\\ x, & i\notin M\end{cases}$$

is $\mathcal{K}$-convergent to x. Since υ is a neighborhood of x, then we have

$$H=\{i\in\mathbb{N}:x_i\notin\upsilon\}\in\mathcal{K}.$$

If we choose $y_i=t_i$, then

$$\{i\in\mathbb{N}:y_i=x_i\}=\{i\in\mathbb{N}:t_i=x_i\}=M\in\mathcal{F}(\mathcal{I})$$

holds. So, $(y_i)\sim_{\mathcal{I}}(x_i)$ holds and (y_i) is eventually in υ.

Conversely, let $\tilde{x}=(x_i)\subset\mathcal{X}$ is a sequence which is $\mathcal{I}^{\mathcal{K}}$-convergent sequence to a point $x\in\upsilon$ and it is $\mathcal{I}^{\mathcal{K}}$-eventually in υ. Assume that υ is not $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. So there exists $x_0\in\overline{\mathcal{X}-\upsilon}^{\mathcal{I}^{\mathcal{K}}}$ which $x_0\notin\mathcal{X}-\upsilon$. This means that there exists a sequence $(x_i)\subset\mathcal{X}-\upsilon$ which is $\mathcal{I}^{\mathcal{K}}$-convergence to $x_0\in\upsilon$. So, (x_i) is $\mathcal{I}^{\mathcal{K}}$-eventually in υ.

Therefore, $\exists\tilde{y}=(y_i)\subset\mathcal{X}$ which $\tilde{x}\sim_{\mathcal{I}}\tilde{y}$ and $\tilde{y}$ is $\mathcal{K}$-eventually in υ. This implies that $\tilde{y}$ is $\mathcal{K}$-eventually in υ which is not in case. □

Theorem 2. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X},\mathcal{T})$. A subset $\mathcal{C}\subset\mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-closed iff*

$$\mathcal{C}=\cap\{\mathcal{A}:\mathcal{A}\textit{ is }\mathcal{I}^{\mathcal{K}}\textit{— closed and }\mathcal{C}\subset\mathcal{A}\}.$$

P r o o f. Let

$$\mathcal{C}=\cap\{\mathcal{A}:\mathcal{A}\text{ is }\mathcal{I}^{\mathcal{K}}\text{— closed and }\mathcal{C}\subset\mathcal{A}\}.$$

Let x be any element of $\mathcal{I}^{\mathcal{K}}$-closure of $\mathcal{C}$. Then there exists $(x_i)\subset\mathcal{C}$ s.t. $x_i\xrightarrow{\mathcal{I}^{\mathcal{K}}}x$. Let $x\notin\mathcal{C}$ so

$$x\notin\cap\{\mathcal{A}:\mathcal{A}\text{ is }\mathcal{I}^{\mathcal{K}}\text{— closed and }\mathcal{C}\subset\mathcal{A}\}.$$

This implies that $\exists\,\mathcal{I}^{\mathcal{K}}$-closed subset F of $\mathcal{X}$ s.t. $x\notin\mathcal{A}$, but $\mathcal{C}$ is $\mathcal{I}^{\mathcal{K}}$-closed and it is a subset of $\mathcal{A}$, which is a contradiction.

The converse is obvious. □

Theorem 3. *Let $\mathcal{I}$ and $\mathcal{K}$ be ideals of $\mathbb{N}$ and $(\mathcal{X},\mathcal{T})$ be a T.S. A function $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}:\mathcal{P}(\mathcal{X})\to\mathcal{P}(\mathcal{X})$ defined as $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)=\overline{A}^{\mathcal{I}^{\mathcal{K}}}$ is satisfying Kuratowski closure axioms*

$(K1)$ $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\varnothing)=\varnothing$ *and* $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X})=\mathcal{X}$,

$(K2)$ $A \subseteq \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \quad \forall A \subseteq \mathcal{X}$,
$(K3)$ $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) = \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \quad \forall A \subseteq \mathcal{X}$,
$(K4)$ $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A \cup B) = \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \cup \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B) \quad \forall A, B \subseteq \mathcal{X}$.

P r o o f. $(K1)$ and $(K2)$ are clear from the definition of $\mathcal{I}^{\mathcal{K}}$-closure function. By Lemma 4, $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$ is closed. So, $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)) = \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$. Therefore, $(K3)$ holds.

To prove $(K4)$, let $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \cup \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$. Then, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$ or $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$. Without lost of generality assume that $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$. So, $\exists (x_i) \subset A$ s.t. $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$. Therefore, $\exists (x_i) \subset A \cup B$ s.t. $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$. So, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \cup \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$.

Conversely, let $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A \cup B)$. Then, there exist a sequence $(x_i) \subset (A \cup B)$ s.t. $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$. Assume that $x \notin \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$ and $x \notin \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$. So, neither set A nor set B contains a sequence s.t. $\mathcal{I}^{\mathcal{K}}$-converges to the point x. Consequently, there is not any sequence in the $A \cup B$ which is convergent to x. But $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A \cup B)$ which is a contradiction. Hence,

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A \cup B) = \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) \cup \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$$

holds. □

Corollary 1. *A subset A of $\mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-closed iff $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A) = A$ and a subset $O \subset \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-open iff $\mathcal{X} - O$ is $\mathcal{I}^{\mathcal{K}}$-closed.*

Theorem 4. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and consider the T.S. $(\mathcal{X}, \mathcal{T})$. Then,*

$$\mathcal{T}_{\mathcal{I}^{\mathcal{K}}} := \{A \subset X : \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A) = \mathcal{X} - A\}$$

is a topology over the set $\mathcal{X}$.

P r o o f. By $(K1)$, it is clear that $\mathcal{X} \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ and $\varnothing \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ hold. Let $A, B \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ be arbitrary sets. To prove $A \cup B \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ we must to prove that

$$\mathcal{X} - A \cup B = \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A \cup B)$$

holds. By $(K2)$, we have

$$\mathcal{X} - A \cup B \subset \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A \cup B).$$

Now, let $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A \cup B)$ be an arbitrarily element. Then, $\exists (x_i) \subset \mathcal{X} - (A \cup B)$ s.t. it is $\mathcal{I}^{\mathcal{K}}$-convergent to x. This implies that (x_i) is not subset of $A \cup B$. So, (x_i) is neither subset of A nor subset of B. Therefore, $(x_i) \subset \mathcal{X} - A$ or $(x_i) \subset \mathcal{X} - B$ which $\mathcal{I}^{\mathcal{K}}$-converges to point x. So, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A)$ or $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - B)$. Since $\mathcal{X} - A$ and $\mathcal{X} - B$ are closed sets, then

$$x \in (\mathcal{X} - A) \cup (\mathcal{X} - B) = \mathcal{X} - A \cup B$$

holds.

Let $\{A_i\}$ be a collection of $\mathcal{I}^{\mathcal{K}}$-open subsets of $\mathcal{X}$. Then, $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - A_i) = \mathcal{X} - A_i$ $\forall i \in \mathbb{N}$. By considering $(K2)$, we have

$$\cap_{i \in \mathbb{N}}(\mathcal{X} - A_i) \subseteq \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}\big(\cap_{n \in \mathbb{N}} (\mathcal{X} - A_i)\big).$$

Let $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}} \cap_{n \in \mathbb{N}} (\mathcal{X} - A_i)$ be an arbitrary element. Then, $\exists (x_i) \subset \cap_{n \in \mathbb{N}}(\mathcal{X} - A_i)$ which is $\mathcal{I}^{\mathcal{K}}$-convergent to x. Then, $(x_i) \subset (\mathcal{X} - A_i)$ $\forall i \in \mathbb{N}$. Since $\mathcal{X} - A_i$ are closed sets, then $x \in \mathcal{X} - A_i$ $\forall i \in \mathbb{N}$. Therefore,

$$x \in \cap_{i \in \mathbb{N}}(\mathcal{X} - A_i).$$

Hence, the set $\mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ is a topology and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ is a T.S. □

Definition 14. *The T.S. $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ is called as $\mathcal{I}^{\mathcal{K}}$-sequential T.S. For abbreviation we will show it by $\mathcal{I}^{\mathcal{K}}$-seq.-top. An $\mathcal{I}^{\mathcal{K}}$-seq.-top. $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ is said to be $\mathcal{I}^{\mathcal{K}}$-discrete space if $\mathcal{T}_{\mathcal{I}^{\mathcal{K}}} = \mathcal{P}(\mathcal{X})$.*

Theorem 5. *Let $\mathcal{I}$, $\mathcal{K}$, $\mathcal{I}_1$, $\mathcal{K}_1$, $\mathcal{I}_2$ and $\mathcal{K}_2$ stand for ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represents a T.S. Let $\mathcal{I}_1 \subset \mathcal{I}_2$ and $\mathcal{K}_1 \subset \mathcal{K}_2$. Then,*
1. $\mathcal{T}_{\mathcal{I}^{\mathcal{K}_2}} \prec \mathcal{T}_{\mathcal{I}^{\mathcal{K}_1}}$,
2. $\mathcal{T}_{\mathcal{I}_2^{\mathcal{K}}} \prec \mathcal{T}_{\mathcal{I}_1^{\mathcal{K}}}$.

P r o o f. Let υ be any $\mathcal{I}^{\mathcal{K}_2}$-open subset of $\mathcal{X}$. Then, $\mathcal{X}-\upsilon$ is $\mathcal{I}^{\mathcal{K}_2}$-closed and $\mathrm{cl}_{\mathcal{I}^{\mathcal{K}_2}}(\mathcal{X}-\upsilon) = \mathcal{X}-\upsilon$ hold. To prove υ is $\mathcal{I}^{\mathcal{K}_1}$-open subset of$\mathcal{X}$, we will show that

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}_1}}(\mathcal{X} - \upsilon) \subset \mathcal{X} - \upsilon.$$

Let $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}_1}}(\mathcal{X} - \upsilon)$ be any point. Then, there exists $(x_i) \subset \mathcal{X} - \upsilon$ s.t. $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}_1}} x$. Since $\mathcal{K}_1 \subset \mathcal{K}_2$, then by Proposition 3.6 in [13], $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}_2}} x$. So, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}_2}}(\mathcal{X} - \upsilon)$. Therefore, $x \in \mathcal{X} - \upsilon$. Hence $\mathcal{X} - \upsilon$ is $\mathcal{I}^{\mathcal{K}_2}$-closed set and υ is $\mathcal{I}^{\mathcal{K}_2}$-open subset of $\mathcal{X}$.

The second one can be proved by using the fact that if $\mathcal{I}_1 \subset \mathcal{I}_2$, then, $x_i \xrightarrow{\mathcal{I}_1^{\mathcal{K}}} x$ implies $x_i \xrightarrow{\mathcal{I}_2^{\mathcal{K}}} x$, it easily can be proved. □

Theorem 6. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represent a T.S. Then, every $\mathcal{I}^*$-open set is $\mathcal{I}^{\mathcal{K}}$-open set.*

P r o o f. If we take $\mathcal{K} = \mathcal{F}in$ then $\mathcal{I}^*$-open set will be $\mathcal{I}^{\mathcal{K}}$-open set. □

Theorem 7. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represent a T.S. Then, every $\mathcal{I}^{\mathcal{K}}$-open set is $\mathcal{K}$-open set.*

P r o o f. Let υ be an arbitrary $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. Then, $\mathcal{X} - \upsilon$ is $\mathcal{I}^{\mathcal{K}}$-closed and

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - \upsilon) = \mathcal{X} - \upsilon.$$

To prove υ is $\mathcal{K}$ open, it is sufficient to show that $\mathcal{X} - \upsilon$ is $\mathcal{K}$-closed, i.e,

$$\mathcal{X} - \upsilon = \overline{\mathcal{X} - \upsilon}^{\mathcal{K}}.$$

It is clear that $\mathcal{X} - \upsilon \subset \overline{\mathcal{X} - \upsilon}^{\mathcal{K}}$. Let $x \in \overline{\mathcal{X} - \upsilon}^{\mathcal{K}}$ be an arbitrary element s.t. $\exists (x_i) \subset \mathcal{X} - \upsilon$ satisfying $x_i \xrightarrow{\mathcal{K}} x$.

Then, by Lemma 3.5 in [13] we have $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$. So, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{X} - \upsilon) = \mathcal{X} - \upsilon$. Hence, the theorem proved. □

Proposition 2. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T})$ represent a T.S. Then, the following statements are true:*

1. *If $\mathcal{K} \subset \mathcal{I}$, then, each $\mathcal{I}$-open set is $\mathcal{I}^{\mathcal{K}}$-open set.*
2. *If the space $\mathcal{X}$ is a first countable space and the ideal $\mathcal{I}$ has additive property with respect to $\mathcal{K}$ (see Definition 3.10 in [13]), then, each $\mathcal{I}^{\mathcal{K}}$-open set is $\mathcal{I}$-open set.*
3. *If $\mathcal{I} \subset \mathcal{K}$, then every $\mathcal{K}$-open set is $\mathcal{I}^{\mathcal{K}}$-open.*

P r o o f. The proof is obvious from Proposition 3.7 and Theorem 3.11 of [13]. □

5. $\mathcal{I}^{\mathcal{K}}$-continuity of functions

In this section we will define $\mathcal{I}^{\mathcal{K}}-$continuous and sequential $\mathcal{I}^{\mathcal{K}}$-continuous functions. We will prove that in any $\mathcal{I}^{\mathcal{K}}$-sequential T.S. these two concepts coincide. Also, we will state some theorems that give the definition of $\mathcal{I}^{\mathcal{K}}$-continuous function in different words and ways. At the end of this section we will see that the combination of $\mathcal{I}^{\mathcal{K}}$-continuous functions is $\mathcal{I}^{\mathcal{K}}$-continuous.

Definition 15. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces. A function f, from $\mathcal{X}$ to $\mathcal{Y}$ is said to be*

(*i*) *$\mathcal{I}^{\mathcal{K}}$-continuous which provides that inverse image of any $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-open in $\mathcal{X}$.*

(*ii*) *Sequentially $\mathcal{I}^{\mathcal{K}}$-continuous which provides that $f(x_i) \overset{\mathcal{I}^{\mathcal{K}}}{\to} f(x)$ $\forall (x_i) \subset \mathcal{X}$ with $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$.*

Theorem 8. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces; and f, from $\mathcal{X}$ to $\mathcal{Y}$ be a function. Then, f is $\mathcal{I}^{\mathcal{K}}$-continuous iff it is sequentially $\mathcal{I}^{\mathcal{K}}$-continuous.*

P r o o f. Let f be an $\mathcal{I}^{\mathcal{K}}$-continuous function. Then, inverse image of any $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-open subset in $\mathcal{X}$. Let $(x_i) \subset \mathcal{X}$ be a sequence with $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$. Then, there exists $M \in \mathcal{F}(\mathcal{I})$ s.t. the following sequence

$$t_i := \begin{cases} x_i, & i \in M, \\ x, & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to x. That is, for each neighborhood υ of x we have

$$\{i \in \mathbb{N} : \ t_i \in \upsilon\} \in \mathcal{F}(\mathcal{K}).$$

Let $\mathcal{V}$ be any $\mathcal{I}^{\mathcal{K}}$-open neighborhood of $f(x)$. Then, $f^{-1}(\mathcal{V})$ is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$ which contains the point x. So, it is a neighborhood of x. Therefore,

$$\{i \in \mathbb{N} : \ t_i \in f^{-1}(\mathcal{V})\} \in \mathcal{F}(\mathcal{K}),$$

implies that $\{i \in \mathbb{N} : \ f(t_i) \in \mathcal{V}\} \in \mathcal{F}(\mathcal{K})$. Hence, the sequence

$$f(t_i) := \begin{cases} f(x_i), & i \in M, \\ f(x), & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to $f(x)$. So, $f(x_i) \overset{\mathcal{I}^{\mathcal{K}}}{\to} f(x)$. Hence, f is sequentially $\mathcal{I}^{\mathcal{K}}$-continuous function.

Conversely, let the function f be sequentially $\mathcal{I}^{\mathcal{K}}$-continuous and υ is any $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$. Assume that $f^{-1}(\upsilon)$ is not $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. Then, $\mathcal{X} - f^{-1}(\upsilon)$ is not $\mathcal{I}^{\mathcal{K}}$-closed subset of $\mathcal{X}$. So,

$$\exists (x_i) \subset \mathcal{X} - f^{-1}(\upsilon) \quad s.t. \quad x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x \quad \text{and} \quad x \notin \mathcal{X} - f^{-1}(\upsilon),$$

i.e. $x_i \notin f^{-1}(\upsilon)$ $\forall n$ and $x_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} x$ which means $x \in f^{-1}(\upsilon)$. Since f is $\mathcal{I}^{\mathcal{K}}$-sequentially continuous function then $f(x_i) \overset{\mathcal{I}^{\mathcal{K}}}{\to} f(x)$. So, $f(x) \in \upsilon$ and $f(x_i) \notin \upsilon$ $\forall n$. This is a contradiction. □

Lemma 5. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces and f, from $\mathcal{X}$ to $\mathcal{Y}$ be an $\mathcal{I}^{\mathcal{K}}$-continuous function. If $(y_i) \subset \mathcal{Y}$ be a sequence s.t. $y_i \overset{\mathcal{I}^{\mathcal{K}}}{\to} y$, then $f^{-1}(y_i) \overset{\mathcal{I}^{\mathcal{K}}}{\to} f^{-1}(y)$.*

P r o o f. Let f be an $\mathcal{I}^{\mathcal{K}}$-continuous function. Let $y_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} y$ then $\exists M \in \mathcal{F}(\mathcal{I})$ s.t. the sequence

$$s_n = \begin{cases} y_i, & i \in M, \\ y, & i \notin M \end{cases}$$

is $\mathcal{K}$-convergent to y. So, for each neighborhood υ of $\mathcal{Y}$,

$$\{i \in \mathbb{N} : y_i \in \upsilon\} \in \mathcal{F}(\mathcal{K}).$$

Since f is $\mathcal{I}^{\mathcal{K}}$-continuous function, then inverse image of any $\mathcal{I}^{\mathcal{K}}-$ open set in $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-open in $\mathcal{X}$, $f^{-1}(\upsilon)$ is open neighborhood of x in $\mathcal{X}$. Then

$$\{i \in \mathbb{N} : f^{-1}(y_i) \in f^{-1}(\upsilon)\} \in \mathcal{F}(\mathcal{K}).$$

Therefore,

$$f^{-1}(s_n) = \begin{cases} f^{-1}(y_i), & i \in M, \\ f^{-1}(y), & i \notin M, \end{cases}$$

is $\mathcal{K}$-convergent to $f^{-1}(y)$ and hence $f^{-1}(y_i) \xrightarrow{\mathcal{I}^{\mathcal{K}}} f^{-1}(y)$. □

Theorem 9. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces. Then the function f, from $\mathcal{X}$ to $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous iff*

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B) = f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$$

holds $\forall B \subset \mathcal{Y}$.

P r o o f. Assume that function f, from $\mathcal{X}$ to $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous function. Let

$$x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B)).$$

Then, $\exists (x_i) \subset f^{-1}(B)$ s.t. $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$. Since f is $\mathcal{I}^{\mathcal{K}}$-continuous so,

$$f(x_i) \xrightarrow{\mathcal{I}^{\mathcal{K}}} f(x).$$

In another hand $(x_i) \subset B$, so $f(x) \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$ and $x \in f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B))$.

Now, let $x \in f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B))$, i.e. $f(x) \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$. Therefore, $\exists (y_i) \subset B$ s.t. $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$. Then, by Lemma 5 there exists $(x_i) = (f^{-1}(y_i) \subset f^{-1}(B)$ s.t. $x_i \xrightarrow{\mathcal{I}^{\mathcal{K}}} x$, where $x = f^{-1}(y)$ holds. So, $x \in \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B))$. Hence,

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B) = f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B).$$

Conversely, let

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B) = f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(B), \quad \forall B \in \mathcal{P}(\mathcal{Y}).$$

Let υ be $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$ then

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{Y} - B) = \mathcal{Y} - B.$$

Let $B = \mathcal{Y} - \upsilon$, then

$$\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(\mathcal{Y} - \upsilon)) = f^{-1}(\mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(\mathcal{Y} - \upsilon)) = f^{-1}(\mathcal{Y} - \upsilon).$$

This shows that $f^{-1}(\mathcal{Y} - \upsilon)$ is $\mathcal{I}^{\mathcal{K}}$-closed. Hence, the following equality

$$f^{-1}(\mathcal{Y} - \upsilon) = \mathcal{X} - f^{-1}(\upsilon)$$

implies that $\mathcal{X} - f^{-1}(\upsilon)$ is $\mathcal{I}^{\mathcal{K}}$-closed. Therefore $f^{-1}(\upsilon)$ is $\mathcal{I}^{\mathcal{K}}$-open set. □

Corollary 2. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces. A function f, from $\mathcal{X}$ to $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous iff*

$$\operatorname{int}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B) = f^{-1}(\operatorname{int}_{\mathcal{I}^{\mathcal{K}}}(B) \quad \forall B \subset \mathcal{Y}.$$

Definition 16. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces and f, from $\mathcal{X}$ to $\mathcal{Y}$ be a function. The function f is $\mathcal{I}^{\mathcal{K}}$-continuous at a point $x \in \mathcal{X}$ if inverse image of any neighborhood of $f(x)$ is a neighborhood of x in $\mathcal{X}$.*

Corollary 3. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces. Then, the function f, from $\mathcal{X}$ to $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous iff it is $\mathcal{I}^{\mathcal{K}}$-continuous at every point $x \in \mathcal{X}$.*

Definition 17. *Let $\mathcal{I}$ and $\mathcal{K}$ stand for the ideals of $\mathbb{N}$ and $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ represent $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces and f, from $\mathcal{X}$ to $\mathcal{Y}$ be a function, f is said to be $\mathcal{I}^{\mathcal{K}}$-closure preserving if*

$$f(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(A)) = \operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f(A) \quad \forall A \subset \mathcal{X}.$$

Theorem 10. *The function f, from $\mathcal{X}$ to $\mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous iff it is $\mathcal{I}^{\mathcal{K}}$-closure preserving.*

P r o o f. Let $f : \mathcal{X} \to \mathcal{Y}$ be an $\mathcal{I}^{\mathcal{K}}$-continuous function. Then, for any subset B of $\mathcal{Y}$

$$\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(B) = f^{-1}(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(B)$$

holds. Consider a set $A \subset \mathcal{X}$ s.t. $f(A)$ is subset of $\mathcal{Y}$. So,

$$\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(f(A)) = f^{-1}(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f(A))$$

holds and it implies that $f(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(A)) = \operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f(A))$ $\forall A \subset \mathcal{X}$ holds.

Conversely, let f be $\mathcal{I}^{\mathcal{K}}$-closure preserving function, then

$$f(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(A)) = \operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f(A)) \quad \forall A \subset \mathcal{X}.$$

Let υ be any subset of $\mathcal{Y}$, then $f^{-1}(\upsilon)$ is subset of $\mathcal{X}$ and

$$f(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(\upsilon))) = \operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f(f^{-1}(\upsilon) = \operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(\upsilon)$$

holds. So

$$\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(f^{-1}(\upsilon) = f^{-1}(\operatorname{cl}_{\mathcal{I}^{\mathcal{K}}}(\upsilon)$$

and by Theorem 9 the function f is $\mathcal{I}^{\mathcal{K}}$-continuous. □

Theorem 11. *Let $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$ be $\mathcal{I}^{\mathcal{K}}$-seq.-top. spaces. Let f, from $\mathcal{X}$ to $\mathcal{Y}$ and g, from $\mathcal{Y}$ to $\mathcal{Z}$ be $\mathcal{I}^{\mathcal{K}}$-continuous functions. Then $g \circ f : \mathcal{X} \to \mathcal{Z}$ is $\mathcal{I}^{\mathcal{K}}$-continuous functions.*

P r o o f. Let υ be any $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Z}$. Since g is $\mathcal{I}^{\mathcal{K}}$-continuous function then $g^{-1}(\upsilon)$ is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$ and because f is $\mathcal{I}^{\mathcal{K}}$-continuous function therefore $f^{-1}(g^{-1}(\upsilon))$ is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$ hence $(g \circ f)^{-1}(\upsilon)$ is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. □

6. Subspace of $\mathcal{I}^{\mathcal{K}}$-seq.-top. space

In this section subspaces of the $\mathcal{I}^{\mathcal{K}}$-seq.-top. space and its properties under an $\mathcal{I}^{\mathcal{K}}$-continuous function will be discussed.

Definition 18. *Let $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ be an $\mathcal{I}^{\mathcal{K}}$-seq.-top. space and $\mathcal{Y} \subset \mathcal{X}$. Then*

$$C_Y : \mathcal{P}(\mathcal{Y}) \to \mathcal{P}(\mathcal{Y}), \quad C_Y(A) = \mathcal{Y} \cap \mathrm{cl}_{\mathcal{I}^{\mathcal{K}}}(A)$$

is a Kuratowsky operator. Define a T.S. as $(\mathcal{Y}, \mathcal{T}^{\mathcal{Y}}_{\mathcal{I}^{\mathcal{K}}})$, where

$$\mathcal{T}^{\mathcal{Y}}_{\mathcal{I}^{\mathcal{K}}} = \{U \cap \mathcal{Y}, Y \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}\} \subset \mathcal{P}(\mathcal{Y}).$$

This T.S. is called $\mathcal{I}^{\mathcal{K}}$-subspace of $\mathcal{X}$.

Lemma 6. *Let $\mathcal{Y}$ be an $\mathcal{I}^{\mathcal{K}}$-subspace of $\mathcal{I}^{\mathcal{K}}$-seq.-top. space $\mathcal{X}$. If set A is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$ and $\mathcal{Y}$ is an $\mathcal{I}^{\mathcal{K}}$-subset of $\mathcal{X}$. Then A is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$.*

P r o o f. Let A be $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$. Then $\exists U \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$ s.t. $A = \mathcal{Y} \cap U$. Since $\mathcal{Y}$ is an $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. Then $A \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$. □

Proposition 3. *Let $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ and $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ be $\mathcal{I}^{\mathcal{K}}$-sequential spaces, $f : \mathcal{X} \to \mathcal{Y}$ be $\mathcal{I}^{\mathcal{K}}$-continuous function and $A \subset \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-subspace of $\mathcal{X}$. Then $f_{/A} : A \to \mathcal{Y}$, the restriction f over A is $\mathcal{I}^{\mathcal{K}}$-continuous function.*

P r o o f. Let U be an $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{Y}$. Since f is $\mathcal{I}^{\mathcal{K}}$-continuous function then $f^{-1}(U)$ is $\mathcal{I}^{\mathcal{K}}$-open subset of $\mathcal{X}$. That is $f^{-1}(U) \in \mathcal{T}_{\mathcal{I}^{\mathcal{K}}}$.

In other hand $f^{-1}_{/A}(U) = A \cap f^{-1}(U)$. So $f^{-1}_{/A}(U)$ is $\mathcal{I}^{\mathcal{K}}$-open subset of subspace A. Hence $f_{/A}$ is $\mathcal{I}^{\mathcal{K}}$-continuous function. □

Lemma 7. *If A is $\mathcal{I}^{\mathcal{K}}$-subspace of $\mathcal{I}^{\mathcal{K}}$-sequential T.S. $\mathcal{X}$. Then the inclusion map $j : A \to \mathcal{X}$ is $\mathcal{I}^{\mathcal{K}}$-continuous.*

P r o o f. If U is $\mathcal{I}^{\mathcal{K}}$-open in $\mathcal{X}$ then $j^{-1}(U) = U \cap A$ is $\mathcal{I}^{\mathcal{K}}$-open in subspace $\mathcal{Y}$ hence j is $\mathcal{I}^{\mathcal{K}}$-continuous. □

Proposition 4. *Let $(\mathcal{X}, \mathcal{T}_{\mathcal{I}^{\mathcal{K}}})$ and $(\mathcal{Y}, \mathcal{T}'_{\mathcal{I}^{\mathcal{K}}})$ be $\mathcal{I}^{\mathcal{K}}$-sequential spaces, $B \subset \mathcal{Y}$ be subspace of $\mathcal{Y}$ and $f : \mathcal{X} \to B$ be $\mathcal{I}^{\mathcal{K}}$-continuous function. Then, $h : \mathcal{X} \to \mathcal{Y}$ obtained by expanding the range of f is $\mathcal{I}^{\mathcal{K}}$-continuous.*

P r o o f. To show $h : \mathcal{X} \to \mathcal{Y}$ is $\mathcal{I}^{\mathcal{K}}$-continuous function, if B as subspace of $\mathcal{Y}$ then note that h is the composition of the map $f : \mathcal{X} \to B$ and $j : B \to \mathcal{Y}$. □

7. Conclusion

In this article we defined the notion of $\mathcal{I}^{\mathcal{K}}$-closed (resp. $\mathcal{I}^{\mathcal{K}}$-open) set in a T.S. $(\mathcal{X},\mathcal{T})$ and established some important results concerning this notion. Furthermore, we defined the $\mathcal{I}^{\mathcal{K}}$-seq.-top., which is a generalized form of the $\mathcal{I}^*$-sequential space. We also talked about $\mathcal{I}^{\mathcal{K}}$-continuity of functions and saw that in $\mathcal{I}^{\mathcal{K}}$-seq.-top. space the notion of continuity and sequential continuity are the same. And in the last section of the paper, subspace of $\mathcal{I}^{\mathcal{K}}$-sequential space have been studied and some important results established.

REFERENCES

1. Blali A., El Amrani A., Hasani R. A., Razouki A. On the uniqueness of $\mathcal{I}$-limits of sequences. *Siberian Electron. Math. Rep.*, 2021. Vol. 8, No. 2. P. 744–757. DOI: 10.33048/semi.2021.18.055
2. Banerjee A. K., Paul M. Strong $\mathcal{I}^{\mathcal{K}}$-convergence in probabilistic metric spaces. *Iranian J. Math. Sci. Inform.*, 2022. Vol. 17, No. 2. P. 273–288. DOI: 10.52547/ijmsi.17.2.273
3. Banerjee A. K., Paul M. *A note on I^K and I^{K^*}-Convergence in Topological Spaces.* 2018. 10 p. arXiv:1807.11772v1 [math.GN]
4. Das P. Some further results on ideal convergence in topological spaces. *Topology Appl.*, 2012. Vol. 159, No. 10–11. P. 2621–2626. DOI: 10.1016/j.topol.2012.04.007
5. Das P., Sleziak M., Toma V. $\mathcal{I}^{\mathcal{K}}$-Cauchy functions. *Topology Appl.*, 2014. Vol. 173. P. 9–27. DOI: 10.1016/j.topol.2014.05.008
6. Das P., Sengupta S., Šupina J. $\mathcal{I}^{\mathcal{K}}$-convergence of sequence of functions. *Math. Slovaca*, 2019. Vol. 69, No. 5. P. 1137–1148. DOI: 10.1515/ms-2017-0296
7. Fast H. Sur la convergence statistique. *Colloq. Math.*, 1951. Vol. 2, No. 3–4. P. 241–244. URL: https://eudml.org/doc/209960 (in French)
8. Georgiou D., Iliadis S., Megaritis A., Prinos G. Ideal-convergence classes. *Topology Appl.*, 2017. Vol. 222. P. 217–226. DOI: 10.1016/j.topol.2017.02.045
9. Jasinski J., Recław I. Ideal convergence of continuous functions. *Topology Appl.*, 2006. Vol. 153, No. 18. P. 3511–3518. DOI: 10.1016/j.topol.2006.03.007
10. Kostyrko P., Mačaj M., Šalát T., Sleziak M. $\mathcal{I}$-convergence and extremal $\mathcal{I}$-limit points. *Math. Slovaca*, 2005. Vol. 55, No. 4. P. 443–464.
11. Kostyrko P., Šalát T., Wilczyńki W. $\mathcal{I}$-convergence. *Real Anal. Exchange*, 2000/2001. Vol. 26, No. 2. P. 669–685.
12. Lahiri B. K., Das P. I and I^*-convergence in topological space. *Math. Bohem.*, 2005. Vol. 130, No. 2. P. 153–160. DOI: 10.21136/MB.2005.134133
13. Mačaj M., Sleziak M. $\mathcal{I}^{\mathcal{K}}$-convergence. *Real Anal. Exchange*, 2010–2011. Vol. 36, No. 1. P. 177–194.
14. Mursaleen M., Debnath S., Rakshit D. I-statistical limit superior and I-statistical limit inferior. *Filomat*, 2017. Vol. 31, No. 7. P. 2103–2108. DOI: 10.2298/FIL1707103M
15. Öztürk F., Şençimen C., Pehlivan S. Strong Γ-ideal convergence in a probabilistic normed space. *Topology Appl.*, 2016. Vol. 201. P. 171–180. DOI: 10.1016/j.topol.2015.12.035
16. Pal S. K. $\mathcal{I}$-sequential topological space. *Appl. Math. E-Notes*, 2014. Vol. 14. P. 236–241.
17. Renukadevi V., Prakash B. $\mathcal{I}$: Fréchet-Urysohn spaces. *Math. Morav.*, 2016. Vol. 20, No. 2. P. 87–97. DOI: 10.5937/MatMor1602087R
18. Sabor Behmanush H., Küçükaslan M. *Some Properties of $\mathcal{I}^*$-Sequential Topological Space.* 2023. 13 p. arXiv:2305.19647 [math.GN]
19. Šalát T., Tripathy B. C., Ziman M. On some properties of $\mathcal{I}$-convergence. *Tatra Mt. Math. Publ.*, 2004. Vol. 28, No. 2. P. 279–286.
20. Steinhaus H. Sur la convergence ordinaire et la convergence asymptotique. *Colloq. Math.*, 1951. Vol. 2. P. 73–74. (in French)
21. Zhou X., Liu L., Lin S. On topological space defined by $\mathcal{I}$-convergence. *Bull. Iran. Math. Soc.*, 2020. Vol. 46, No. 3. P. 675–692. DOI: 10.1007/s41980-019-00284-6

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 60–76
DOI: 10.15826/umj.2023.2.005

HETEROGENEOUS SERVER RETRIAL QUEUEING MODEL WITH FEEDBACK AND WORKING VACATION USING ARTIFICIAL BEE COLONY OPTIMIZATION ALGORITHM

Divya Kothandaraman[†], **Indhira Kandaiyan**[††]

Department of Mathematics, School of Advanced Sciences,
Vellore Institute of Technology, Vellore-632014, Tamil Nadu, India

[†]divya.k2020@vitstudent.ac.in [††]kindhira@vit.ac.in

Abstract: This research delves into the dynamics of a retrial queueing system featuring heterogeneous servers with intermittent availability, incorporating feedback and working vacation mechanisms. Employing a matrix geometric approach, this study establishes the steady-state probability distribution for the queue size in this complex heterogeneous service model. Additionally, a range of system performance metrics is developed, alongside the formulation of a cost function to evaluate decision variable optimization within the service system. The Artificial Bee Colony (ABC) optimization algorithm is harnessed to determine service rates that minimize the overall cost. This work includes numerical examples and sensitivity analyses to validate the model's effectiveness. Also, a comparison between the numerical findings and the neuro-fuzzy results has been examined by the adaptive neuro fuzzy interface system (ANFIS).

Keywords: Retrial queue, Working vacation, MGA, ANFIS, ABC Optimization.

1. Introduction

In our modern, fast-paced society, it is crucial to prioritize the optimization of service systems due to the ever-changing needs and demands of diverse customers. Although traditional queueing models are valuable, they often fail to address the complexities of modern service environments. This research presents a queueing model that effectively addresses these challenges. This model fundamentally recognizes that service tasks can vary in nature and importance. It acknowledges the significance of selecting the appropriate service provider for a particular task. The concept of 'heterogeneous servers' is relevant here. Some servers specialize in handling routine requests, while others are particularly skilled at addressing complex issues.

Furthermore, we recognize that servers are not able to be available around the clock. Instead, they alternate between performing routine tasks and dedicating their attention to more specialized, secondary jobs. Intermittent availability optimizes resource allocation by ensuring that highly skilled servers are readily available when they are most needed. Finally, we have implemented a "working vacation" feature to guarantee uninterrupted service during periods of downtime. This feature ensures that customers are not left unattended, minimizing disruptions in service even when servers are on a break. This research goes beyond being a mere theoretical innovation; it tackles the actual challenges that service industries encounter in the real world. Our model provides a practical and competitive advantage in the dynamic landscape of modern service provision by enhancing efficiency, boosting customer satisfaction, and optimizing resource utilization.

The novelty of this work lies in its pioneering approach to designing service systems that can adapt to the multifaceted demands of contemporary industries. Unlike traditional queueing models, which rely on uniform servers and predictable service patterns, our model introduces heterogeneity, recognizing that not all service tasks are equal, which is shown in Fig. 1.

The remaining sections of this research paper: Section 2 outlines the model's development and the quasi-birth-death process framework. Moving on to Section 3, we delve into the matrix geometric approach, demonstrating the process of compute steady-state probabilities. In Section 4, we explore various performance metrics derived from the model and their practical implications. Section 5 is dedicated to a comprehensive discussion of sensitivity analysis and a cost assessment for the considered paradigm. Section 6 offers graphical representations of ANFIS and presents the numerical outcomes. The subsequent Section 7, delves into cost optimization strategies. Finally, Section 8 serves as the conclusion, where we summarize our investigation by highlighting the notable characteristics and real-world applications of our study.

1.1. Survey of literature

Our model incorporates two types of servers: one that handles routine tasks (server 1) and another that periodically shifts (server 2) its focus to secondary, more specialized tasks. The presence of heterogeneity enables a service delivery that is more customized and effective. In this heterogeneous queueing model, the servers provide service at a different rate. Morse [14] was the first to propose the notion of service heterogeneity. A queueing model with two classes and two servers is being discussed. A non-preemptive priority structure that is heterogeneous has been studied by Leemans [12]. According to [3], a heterogeneous two-server queueing system with feedback, reverse balking, and reneging and retaining renege customers can be analyzed. Markovian queueing model with discouraged arrivals, reneging customers, and retention of reneged customers was studied by [11] based on two heterogeneous servers finite capacities. A study presents an investigation of the heterogeneous queueing system $M/M/2$ with two types of server failures and catastrophes, along with their respective restoration processes, as conducted by the [16]. A queueing model with MAP arrivals and heterogeneous phase-type group services was researched by [4].

Agarwal [1] initially introduced the concept of a server with intermittent availability, where server 1 is consistently accessible while server 2 is periodically accessible. In this scenario, server 2 is responsible for executing a range of peculiar and unconventional tasks. Service interruptions may occur for a variable duration, but they are limited to instances when the ongoing task has been completed. This particular service is referred to as an intermittently available service. Sharda [19] investigated a queuing issue involving a server that is intermittently accessible, with entries and exits occurring in batches of varying sizes.

In recent times, queueing systems featuring server vacations have become increasingly popular. These "vacations" can arise from server outages or when the server is tasked with other responsibilities. Our model acknowledges the critical importance of maintaining continuous service, even during these working vacation periods. It is designed to ensure that customers are never left unattended, thus minimizing any disruptions in the quality of service provided. A recent trend in vacation queues has been working vacation, where service is provided at a lower rate during vacation periods than it is normally provided; i.e., while on vacation, the server provides service at a slower rate instead of ceasing completely. An initial proposal for a working vacation model has been made by Servi and Finn [13]. Madhu Jain [7] conducted a study on a single server working vacation queueing model that incorporates multiple types of server breakdowns. Sudhesh et al. [21] investigated the time-dependent dynamics of a single server queueing model featuring slow service. The researchers examined the effects of both single and multiple working vacations, as well as customers' impatience during periods of slow service. Krishnamoorthy et al. [9] discussed a queueing system with two heterogeneous servers. One server is always accessible, while the other takes vacations when no users are waiting. Laxmi et al. [23] examined a queuing system with several working vacations, incorporating elements of renewal input, balking, reneging, and heterogeneous servers. Two types of Working Vacations (WVs) and impatient clients were handled with in a multi-server queueing system by Yohapriyadharsini et al. [25]. Kumar et al. [10] examined a unreliable Markovian queueing model with two stage service, incorporating hybrid vacation.

The Matrix Geometric Method is a key technique in queueing theory. It simplifies the analysis of systems with varying service rates and transitions by using matrices, providing efficient solutions for steady-state probabilities and performance metrics. Initially introduced by Neuts [8], matrix-geometric models are the basis for stochastic computations. Recently, Divya [5] conducts an investigation on a Markovian queueing model that incorporates heterogeneous, intermittently available servers with feedback, operating under a hybrid vacation policy. ANFIS is an Adaptive Neuro Fuzzy Inference System. This ANFIS computer model uses fuzzy logic and neural networks to analyze and make decisions. ANFIS is famous for tackling complicated issues in numerous industries because it provides a foundation for building hybrid systems that can learn and adapt from data. ANFIS was established in the early 1980s by Professor Lotfi A. Zadeh [26]. Ahuja et al. [2] presented a comprehensive analysis of a single server queueing model with multiple stage service and functioning vacation, focusing on transient behavior. Additionally, they employed ANFIS computing techniques to enhance their analysis. Sethi et al. [17] conducted a study on the application of ANFIS in analyzing the performance of an unreliable $M/M/1$ queueing system. The study specifically focused on the impact of customers' impatience under N-policy. The ANFIS concept has garnered attention from a multitude of researchers in diverse fields of study [6], [18], [20], [22]. Wu and Yang [24] conducted optimization of a bi-objective queueing model that incorporates a two-phase heterogeneous service.

2. Model description and assumptions

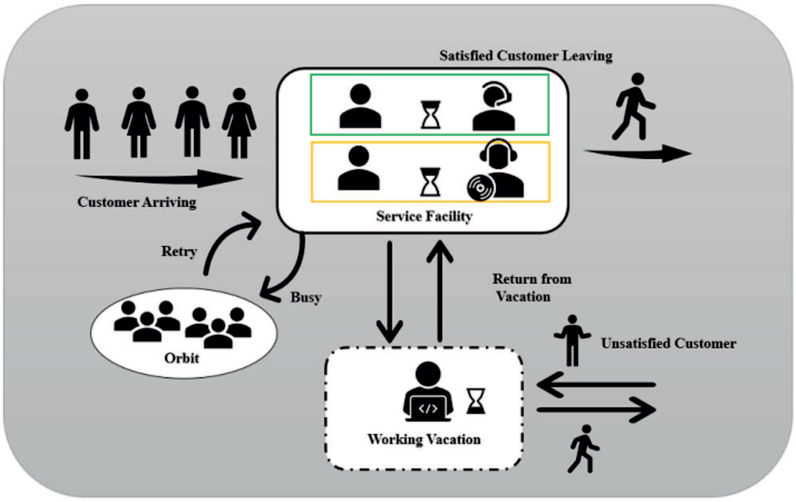


Figure 1. Model diagram.

1. **Arrival process.** In the RQ system, customers arrive according to a Poisson process with a rate of ω.

2. **Service process.** Server 1 is always obtainable, server 2 is intermittently obtainable. The servers provide service to customers with service rates of γ_1 and γ_2, respectively. The retrieval capacity time on server 2 follows an exponential distribution with a rate of β.

3. **Retrial process.** The retrial queuing mechanism enables customers to opt to orbit in the event that the servers are occupied upon their arrival. After a constant retrial rate ϕ, individuals may make another attempt to receive service following an exponentially distributed time.

4. **Vacation process.** A queuing system with working vacation is analyzed, If there are no customers in the orbits when the server 2 finishes servicing, it goes on a working vacation, where the server 2 serves customers with a reduced service rate γ_v during such periods, which follow an exponential distribution. When a vacation ends and there are customers waiting for service, the server 2 switches to regular service with retrieval rate τ. If there are no customers waiting, the server 2 retains in the same working vacation.

5. **Feedback rule.** During WV there are two possible outcomes for customers receiving service during working vacation: they may receive satisfactory service with probability p or unsatisfactory service with complementary probability $\overline{p}(1-p)$. In the event of unsatisfactory service, customers must undergo supplementary service, which follows an exponential distribution.

All stochastic processes in the system are independent of one another. The structure of the models transition diagram is depicted in the below Fig. 2. At time t, let $\chi(t)$ be the state of the server, which is defined as

$$\chi(t) = \begin{cases} 0, & \text{the server 2 is in WV \& it's free,} \\ 1, & \text{the server 2 is in WV \& it's busy,} \\ 2, & \text{the server 2 is in busy,} \\ 3, & \text{the server 2 is in intermittently obtainable} \end{cases}$$

and $\phi(t)$ be the number of customers in the system. The bi-variate process $\{(\phi(t), \chi(t)), t \geq 0\}$ that operates on a state space of $\{0, 1, 2, \dots\} \times \{0, 1, 2, 3\}$.

$$\Upsilon(t) = \{(l, m) | l \geq 0, \ m = 0, 1, 2, 3\}.$$

The state space of a Markov process is arranged in a lexicographical manner, as described below.

$$\Omega = \{(0,0) \bigcup \{(l,m)|l \geq 0, m = 0, 1, 2, 3\}.$$

2.1. Steady-state equation

To solve this problem and obtain effective and mathematically accurate model solutions, we employ the matrix-geometric method described in the following section. The matrix-geometric method is an effective method for obtaining steady-state probabilities when the state-space expands very quickly.

$$\overline{p}\gamma_v \pi_{0,0} = \omega \pi_{0,1}, \tag{2.1}$$

$$(\phi + \overline{p}\gamma_v)\pi_{l,0} = p\gamma_v \pi_{l-1,1} + \omega \pi_{l,1}, \quad l = 1, 2, 3 \dots, \tag{2.2}$$

$$(2\omega + \xi + p\gamma_v)\pi_{0,1} = \overline{p}\gamma_v \pi_{0,0} + \phi \pi_{1,0} + \gamma_1 \pi_{1,1}, \tag{2.3}$$

$$(2\omega + \gamma_1 + \tau + p\gamma_v)\pi_{l,1} = \omega \pi_{l-1,1} \overline{p}\gamma_v \pi_{l,0} + \phi \pi_{l+1,0} + \gamma_1 \pi_{l+1,1}, \quad l = 1, 2, 3 \dots, \tag{2.4}$$

$$\omega \pi_{0,2} = (\gamma_1 + \gamma_2)\pi_{1,2} + \beta \pi_{0,3} + \tau \pi_{0,1}, \tag{2.5}$$

$$\omega \pi_{l,2} = \omega \pi_{l-1,2} + \tau \pi_{l,1} + (\gamma_1 + \gamma_2)\pi_{l+1,2} + \beta \pi_{l,3}, \quad l = 1, 2, 3 \dots, \tag{2.6}$$

$$(\beta + \omega)\pi_{0,3} = \gamma_1 \pi_{1,3}, \tag{2.7}$$

$$(\gamma_1 + \beta + \omega)\pi_{l,3} = \omega \pi_{l-1,3} + \gamma_1 \pi_{l+1,3}, \quad l = 1, 2, 3. \dots \tag{2.8}$$

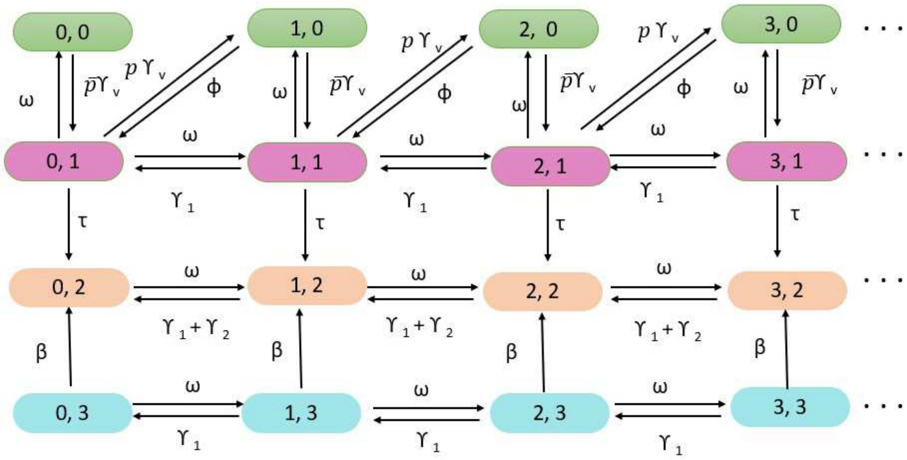


Figure 2. Transition diagram of the model.

3. Matrix-geometric solution

To calculate the steady-state probabilities of the model using the matrix-geometric approach, we utilize a system of equations denoted as (2.1) to (2.8). These equations help determine the probabilities at a steady state. The transition rate matrix Q, which represents the Markov chain in this model, is structured as a block tridiagonal matrix. The matrix Q is subdivided into sub matrices.

$$Q = \begin{bmatrix} S_0 & T_0 & & & & & \\ U_0 & V_0 & T_0 & & & & \\ & U_0 & V_0 & T_0 & & & \\ & & U_0 & V_0 & T_0 & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \end{bmatrix}$$

$$S_0 = \begin{bmatrix} -\omega & \omega & 0 & 0 \\ \overline{p}\gamma_v & -(\omega+\gamma_v+\tau) & \tau & 0 \\ 0 & 0 & -\omega & 0 \\ 0 & 0 & \beta & -(\omega+\beta) \end{bmatrix};$$

$$T_0 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ p\gamma_v & \omega & 0 & 0 \\ 0 & 0 & \omega & 0 \\ 0 & 0 & 0 & \omega \end{bmatrix}; \quad U_0 = \begin{bmatrix} 0 & \phi & 0 & 0 \\ 0 & \gamma_1 & 0 & 0 \\ 0 & 0 & \gamma_1+\gamma_2 & 0 \\ 0 & 0 & 0 & \gamma_1 \end{bmatrix};$$

$$V_0 = \begin{bmatrix} -(\omega+\phi) & \omega & 0 & 0 \\ \overline{p}\gamma_v & -(\omega+\gamma_1+\tau+\gamma_v) & \tau & 0 \\ 0 & 0 & -(\omega+\gamma_1+\gamma_2) & 0 \\ 0 & 0 & \beta & -(\omega+\beta+\gamma_1) \end{bmatrix}.$$

The steady-state probability vector Π for Q is partitioned as $\Pi = (\Pi_0, \Pi_1, \Pi_2, \ldots)$, where the sub-vectors $\Pi_l = \{\pi_{l,0}, \pi_{l,1}, \pi_{l,2}, \pi_{l,3}\},\ l \geq 0$.

3.1. Stability criteria

Theorem 1. *The inequality*

$$\rho = \frac{\gamma_1 + \gamma_2}{\omega} < 1$$

is the necessary and sufficient condition for the system to be stable.

P r o o f. Let us define the matrix $\mathcal{E} = T_0 + V_0 + U_0$ given by

$$\mathcal{E} = \begin{bmatrix} -\xi_1 & \xi_1 & 0 & 0 \\ \gamma_v & -\xi_2 & \tau & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \beta & -\beta \end{bmatrix}.$$

Where $\xi_1 = (\omega + \phi)$; $\xi_2 = (\tau + \gamma_v)$. There exists a stationary probability $\Pi = [\Pi_0, \Pi_1, \Pi_2, \Pi_3]$ of $\mathcal{E}$ such that

$$\Pi\mathcal{E} = 0, \quad \Pi e = 1, \tag{3.1}$$

where $e = [1, 1, 1, 1]^T$. Using Theorem 3.1.1 of Netus [8], the necessary and sufficient condition for the stability of the system is as follows:

$$\Pi T_0 e < \Pi U_0 e. \tag{3.2}$$

Solving (3.1) and (3.2), we get

$$\frac{\gamma_1 + \gamma_2}{\omega} < 1. \tag{3.3}$$

□

3.2. Stationary probability distribution

Let Π_{lm} be the steady-state probability that the process is in state (l, m), which is defined as follows:

$$\Pi_{lm} = \lim_{t \to 0} Pr\big[\phi(t) = l,\ \chi(t) = m\big], \quad l = 0, 1, 2, 3 \dots \quad \text{and} \quad m = 0, 1, 2, 3.$$

We denote the steady state probability vector of Q by $\Pi = (\Pi_0, \Pi_1, \Pi_2, \dots)$. Where $\Pi_l = \big(\pi_{l,1}, \pi_{l,1}, \pi_{l,2}, \pi_{l,3}\big)$ for $l \geq 0$. Under the stability condition (3.3). The steady-state equations can be expressed in matrix form as follows,

$$\Pi \mathrm{Q} = 0. \tag{3.4}$$

Equation (3.4) can be written as

$$\begin{gathered} \Pi_0 S_0 + \Pi_1 U_0 = 0, \\ \Pi_0 T_0 + \Pi_1 V_0 + \Pi_2 U_0 = 0, \\ \vdots \\ \Pi_{i-1} T_0 + \Pi_i V_0 + \Pi_{i+1} U_0 = 0, \quad i = 1, 2, 3, \dots . \end{gathered}$$

Based on the matrix-geometric method [8, 15], we obtain

$$\Pi_i = \Pi_0 \mathcal{R}^i \quad \text{for} \quad i \geq 1 \tag{3.5}$$

and Π_0 satisfies the set of equations

$$\Pi_0 \left(S_0 + \mathcal{R} U_0 \right) = 0. \tag{3.6}$$

Where $\mathcal{R}$ is referred to as the rate matrix which satisfies

$$T_0 + \mathcal{R} V_0 + \mathcal{R}^2 U_0 = 0,$$

and 0 denotes a zero squared matrix of an appropriate order.

The rate matrix can be approximate iteratively by considering one sequence with initialization $\mathcal{R}_0 = 0$ and calculating

$$\mathcal{R}_{i+1} = -\left(T_0 + \mathcal{R}_i^2 U_0 \right) V_0^{-1}, \quad i = 1, 2, \ldots.$$

Thus, $\lim_{i \to \infty} \mathcal{R}_i$ is an approximate solution of the rate matrix $\mathcal{R}$. From the normalization condition, we obtain the following:

$$\sum_{i=0}^{\infty} \Pi e = \sum_{i=0}^{\infty} \Pi_0 \mathcal{R}^i e = \Pi_0 \left(I - \mathcal{R} \right)^{-1} e = 1.$$

Combining (3.6) with the normalization condition yields

$$\Pi_0 = [\pi_{0,0}, \pi_{0,1} \ldots].$$

Once the steady-state probability vector Π_0 is available, then Π_i (i $\geq$ 1) can be determined using (3.5).

4. Performance measures

4.1. Performance measures

Based on the steady-state probabilities, we give numerous performance metrics for the model under evaluation.

- Prob that the servers are in idle

$$P_i = \pi_{0,0}.$$

- Prob that the server is in busy state

$$P_b = P[m = 2] = \sum_{l=1}^{\infty} \pi_{l,2}.$$

- Prob that the server 2 is in WV state

$$P_{wv} = P[m = 0] + P[m = 1] = \pi_{0,0} + \sum_{l=0}^{\infty} \pi_{l,1}.$$

- Prob that the server 2 is in IO state

$$P_{Io} = P[m = 3] = \sum_{l=0}^{\infty} \pi_{l,3}.$$

- Average system length

$$ASL = \sum_{l=0}^{\infty} l\pi_{l,1} + \sum_{l=0}^{\infty} l\pi_{l,0} + \sum_{l=0}^{\infty} l\pi_{l,2} + \sum_{l=0}^{\infty} l\pi_{l,3}.$$

- Average queue length

$$AQL = \sum_{l=0}^{\infty} l - 1\pi_{l,1} + \sum_{l=0}^{\infty} l - 1\pi_{l,0} + \sum_{l=0}^{\infty} l - 1\pi_{l,2} + \sum_{l=0}^{\infty} l - 1\pi_{l,3}.$$

4.2. Practical application

The proposed queueing model is applicable to a semi-attended self-checkout system in retail stores.

Server 1 – Always Available (Self-Checkout Machine): This server is designed to operate continuously and is accessible to customers at all times, similar to the self-checkout machines commonly found in retail shops. Customers have the freedom to use it at any time without any interruptions.

Server 2, also known as the Intermittently Obtainable (Human Billing Counter), functions similarly to a human cashier at a retail store. The service is available intermittently, which means that it serves customers but may take breaks or go on working vacations. A working vacation refers to planned breaks or vacations for server 2, during which it is temporarily unavailable to serve customers. For instance, a cashier may take a lunch break or have a scheduled time off during their shift. There are numerous systems similar to the queueing model, such as call centers, healthcare triage, banking services, online customer support, and restaurant service.

5. Sensitivity and cost analysis

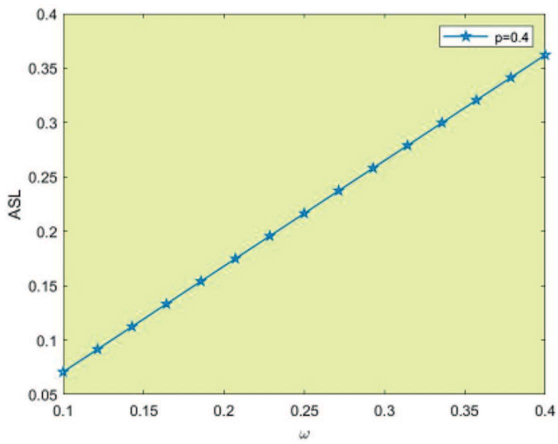
5.1. Sensitivity analysis

Here, we provide numerical examples to demonstrate the influence of various system settings on three distinct efficiency metrics (ω, γ_1, and γ_2) are applied to the following scenarios:
Case 1: $\gamma_1 = 0.5, \gamma_2 = 1$, and vary the value of ω from 0.1 to 0.4.
Case 2: $\omega = 0.05, \gamma_2 = 2$, and vary the value of γ_1 from 0.5 to 1.2.
Case 3: $\omega = 0.06, \gamma_1 = 0.5$, and vary the value of γ_2 from 1 to 3.

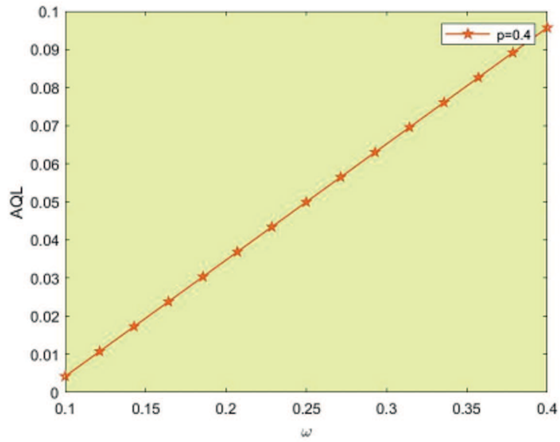
All values assigned to the system parameters in our numerical analysis satisfy the stability condition as described by (3.3). The curves depicting the performance measures against the parameters are illustrated in Fig. 3 to 5. Table 1 presents numerical results showing that increasing arrival rate (ω) corresponds to increasing system size (ASL), queue size (AQL) and server busy state probability (P_b). leads to, decreases the probabilities of other states. Similarly, as we increase the service rates (γ_1 and γ_2), the server busy probability (P_b), queue size (AQL), and system size (ASL) decreases, the probabilities of other states increase.

6. ANFIS implementation and results

An Adaptive Neuro-Fuzzy Inference System (ANFIS) is a computational model that combines the principles of neural networks and fuzzy logic to perform complex tasks, such as pattern recognition and system modeling. ANFIS utilizes a hybrid approach that blends the adaptability of neural

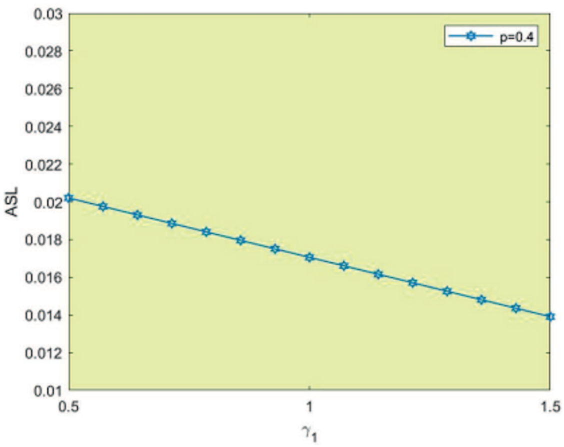


(a) ω $Vs.$ ASL

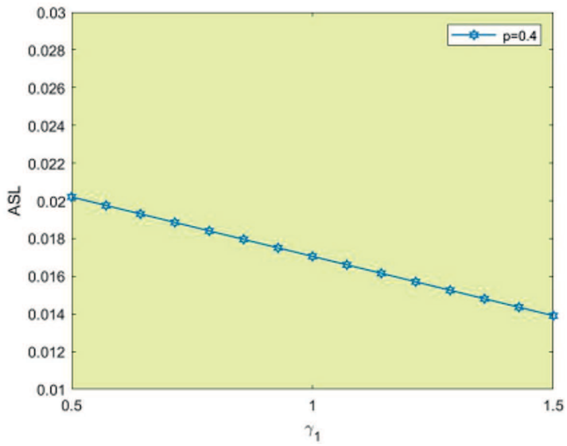


(b) ω $Vs.$ AQL

Figure 3. The curves depicting the performance measures against the parameter ω.

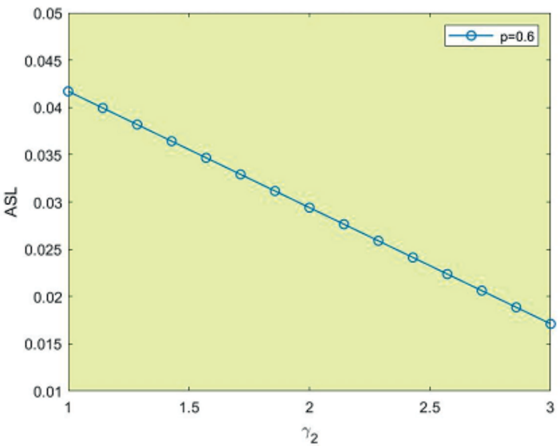


(a) γ_1 $Vs.$ ASL

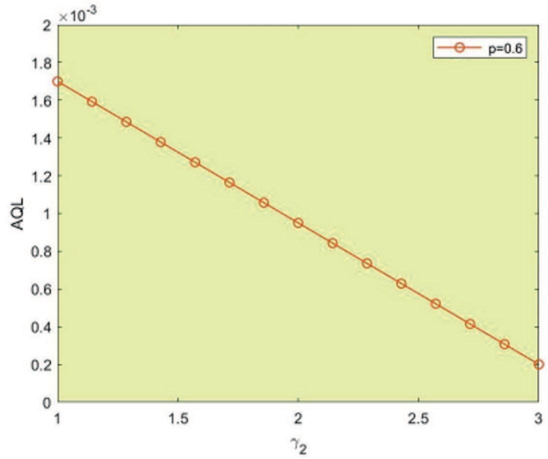


(b) γ_1 $Vs.$ AQL

Figure 4. The curves depicting the performance measures against the parameter γ_1.



(a) γ_2 $Vs.$ ASL



(b) γ_2 $Vs.$ AQL

Figure 5. The curves depicting the performance measures against the parameter γ_2.

Table 1. Various performance measures by varying ω, γ_1, γ_2, γ_v.

p	ω	γ_1	γ_2	γ_v	ASL	AQL	P_b
0.4	0.1	0.5	1	0.5	0.0707	0.0042	1.0007
	0.2				0.1535	0.0202	0.9995
	0.3				0.2500	0.0500	1.0000
	0.4				0.3622	0.0957	1.0000
0.6	0.1				0.0717	0.0049	1.0010
	0.2				0.1534	0.0202	0.9995
	0.3				0.2499	0.0499	1.0000
	0.4				0.3610	0.0947	0.9998
0.4	0.05				0.0202	0.0003	0.9999
		0.7			0.0187	0.0003	0.9981
		0.9			0.0174	0.0003	0.9954
0.6		0.5			0.0379	0.0004	1.0009
		0.7			0.0384	0.0013	0.9980
		0.9			0.0432	0.0040	0.7762
0.4	0.06	0.5			0.0417	0.0017	1.0000
			1.5		0.0307	0.0008	0.9999
			2		0.0244	0.0005	0.9999
			2.5		0.0204	0.0004	1.0000
			3		0.0171	0.0002	0.9952
0.6			1		0.0417	0.0017	1.0000
			1.5		0.0309	0.0009	1.0000
			2		0.0246	0.0006	1.0000
			2.5		0.0204	0.0004	1.0000
			3		0.0173	0.0003	0.9955

networks with the interpretability of fuzzy logic. It consists of a layered architecture where input data is passed through a series of nodes, each representing a fuzzy membership function. These nodes calculate membership values based on the input data's similarity to predefined linguistic terms. ANFIS learns and adjusts its parameters using a combination of gradient descent and least-squares methods. This enables it to fine-tune the strengths of its fuzzy rules and the connection weights between nodes to accurately model intricate relationships within the data. The model is particularly useful when dealing with non-linear and uncertain data, making it suitable for applications in various fields, including control systems, prediction, and optimization. By incorporating both neural networks and fuzzy logic, ANFIS provides a balance between the strengths of both approaches, offering a powerful tool for researchers and practitioners to tackle complex problems effectively.

Table 2. Values of the MF for the linguistics based on input parameters.

Input parameters	No. of membeship function	Linguistic Values
p, ω, γ_1, γ_2	5	Very Low, Low, Medium, High, Very High

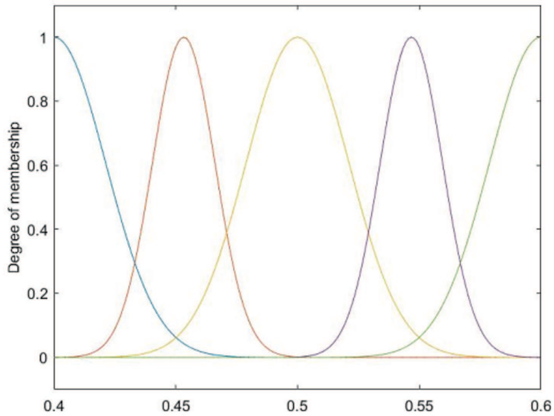
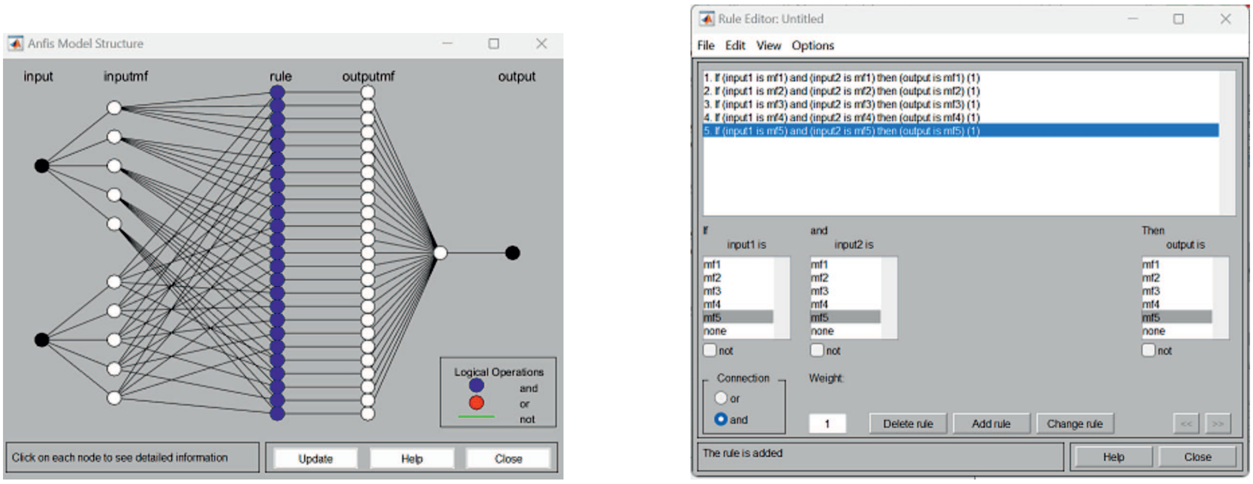


Figure 6. Membership function for ω.

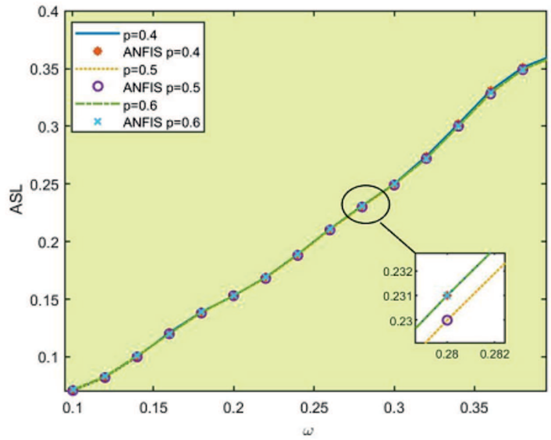


(a) (b)

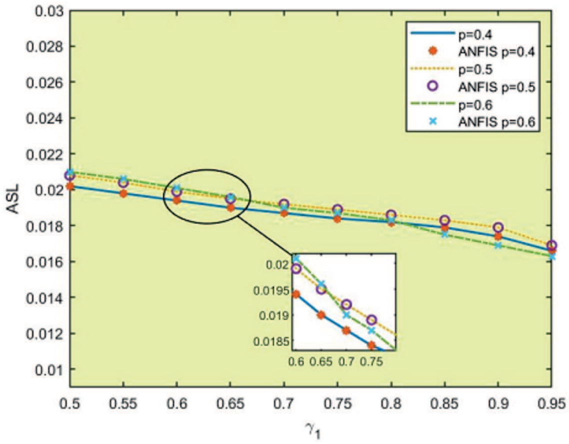
Figure 7. (a) ANFIS Structure and (b) Rules.

ANFIS was founded in the first few years of the 1980s by Professor Lotfi A. Zadeh [26]. The architecture of a two-input $(p, \omega, \gamma_1, \gamma_2)$, one-output (ASL) Adaptive Neuro-Fuzzy Inference System (ANFIS) model with five rules is illustrated in Fig. 7. Five fuzzy rules were created and the resulting Gaussian-shaped membership functions (MFs) of the inputs are displayed in Fig. 6. It is noteworthy to emphasize that each colored MF depicted in the curve represents a distinct cluster inside the input space. In ANFIS methodology, the parameters p, ω, γ_1, γ_2 and ASL are regarded as linguistic variables and subjected to training for a total of ten epochs. Three linguistic values have been utilized for the variables p, ω, γ_1, γ_2 and ASL, namely very low, low, medium, high, and very high. Gaussian membership functions were employed to represent the linguistic variables, as illustrated in Table 2.

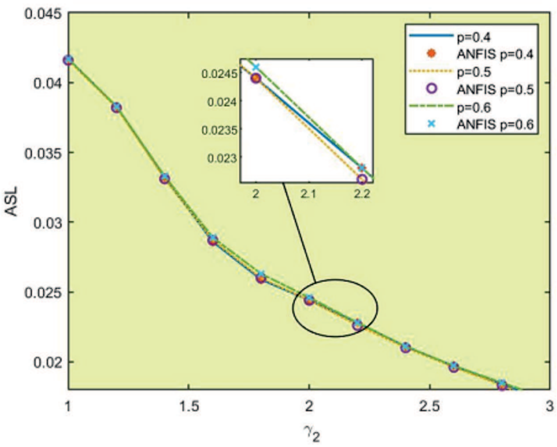
Fig. 8 displays the analytical results using continuous lines, while the results acquired using ANFIS for ω, γ_1, and γ_2 are shown by a dotted marker point. Furthermore, it was observed that the analytical and ANFIS outcomes had a high degree of concurrence, displaying a significant overlap in their respective trends. Based on the data presented, it can be noted that there is a positive correlation between the system length (ASL) and the arrival rate (ω) while considering different values of p. As the arrival rate increases, the system length also increases. Conversely, the system



(a) ω $Vs.$ ASL



(b) γ_1 $Vs.$ ASL



(c) γ_2 $Vs.$ ASL

Figure 8. Comparison of 2D numerical and ANFIS values for different values of p.

length lowers as the service rates (γ_1 and γ_2) decrease, while considering different values of p.

7. Cost model and optimization

The suggested queueing model has the potential to be implemented in both self-checkout systems and retail stores with human cashiers. In such scenarios, the primary objective of the manager is to minimize operational expenses. An essential consideration related to the self-checkout process pertains to the determination of the appropriate quantity of self-checkout machines for the service. Another crucial matter to consider is the necessity of upholding a satisfactory service rate in order to ensure the quality of service and customer satisfaction. It is feasible to improve the calibre of service rendered by the staff through comprehensive training. Moreover, it is not possible. To ascertain the customer's familiarity with the operation of a particular system. Self-checkout machines. Therefore, we proceed to formulate a cost function per unit of time in a manner that aligns with our expectations.

$$\mathcal{F} = F(\gamma_1, \gamma_2) = F_h \cdot ASL + F_b \cdot P_b + F_v \cdot P_{wv} + F_i \cdot P_{Io} + F_1 \cdot \gamma_1 + F_2 \cdot \gamma_2. \tag{7.1}$$

The variables in the equation are defined as follows: F_h represents the holding cost for each customer in the system, F_v represents the cost per unit of time when the server 2 in the working vacation service, F_b represents the cost per unit of time when server 2 is busy, F_i represents the cost per unit of time when server 2 is intermittently obtainable, F_1 represents the cost of providing a mean service rate γ_1 through server 1, and F_2 represents the cost of providing a mean service rate γ_2 through server 2. It is noteworthy to mention that (7.1) represents a mathematical function that depends on two continuous decision variables, denoted as γ_1 and γ_2. We set the cost elements as given in Table 3.

Table 3. Cost set values for various cost aspects.

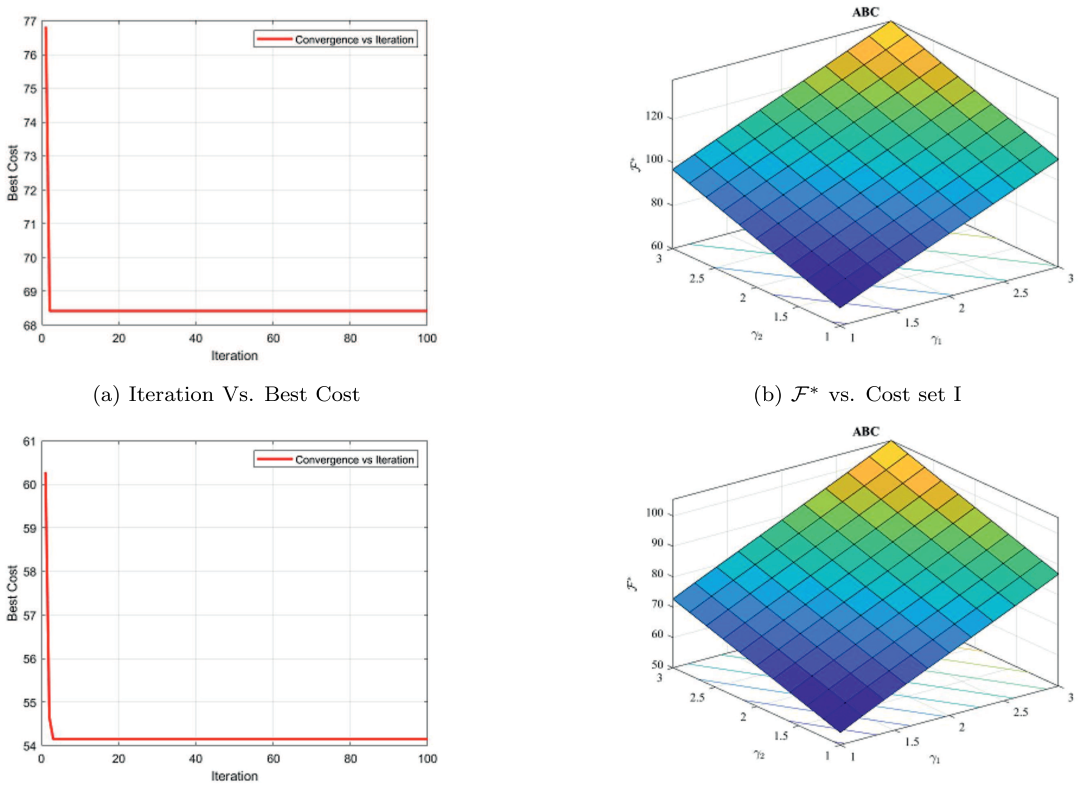
Cost set	F_h	F_b	F_v	F_i	F_1	F_2
I	60	50	40	30	25	15
II	50	45	35	20	20	10

7.1. Artificial bee colony optimization

ABC optimization, also known as Artificial Bee Colony optimization, is a meta-heuristic algorithm inspired by the foraging behavior of honey bees. It is a swarm-based optimization technique that can be applied to solve various optimization problems. The ABC optimization technique, initially introduced by Karaboga in 2005, has garnered significant recognition and acclaim in the field of optimization. The system employs worker bees, observer bees, and scout bees to explore the search space, exchange information, and discover improved solutions. The historical effect in ABC ensures that the algorithm strategically priorities regions of the search space that have demonstrated favorable outcomes, thus leveraging past successes. This enables ABC to efficiently converge towards optimal solutions and effectively address complex optimization problems.

In order to optimize ABC, the following default values are taken into account: $\omega = 3$, $\gamma_1 = 1.5$, $\gamma_2 = 0.6$, $\gamma_v = 0.5$, $\phi = 0.05$, $\tau = 0.5$, $\beta = 0.5$, $p = 0.4$, $p_1 = 0.6$ with a colony size of 100, a maximum of 100 iterations, an acceleration coefficient upper and lower bound are 1 and 5, and number of Onlooker bee 50, an abandonment limit parameter of 60.

Table 4 show the effect of cost elements F_h, F_b, F_v, F_i, F_1 and F_2 on the optimal service rates (γ_1^*, γ_2^*) and optimal total cost $(\mathcal{F}^*)$ for all two cost sets. The pseudo code of ABC algorithm is given in Table 1.



(a) Iteration Vs. Best Cost

(b) $\mathcal{F}^*$ vs. Cost set I

(c) Iteration Vs. Best Cost

(d) $\mathcal{F}^*$ vs. Cost set II

Figure 9. 2D and 3D visualization of ABC optimization.

7.2. Convergence

Convergence is a crucial aspect of meta-heuristic optimization algorithms. It signifies the process of gradually refining the candidate solutions toward an optimal or near-optimal solution. The convergence behavior of an algorithm is indicative of its ability to effectively search the solution space and approach the global optimum. In ABC, particles move towards the best-known solution, converging when their movements become limited and the best solution stabilizes. Fig. 9 (a) and (c) shows that ABC reach optimal cost convergence, Fig. 9 (b) and (d) shows the convexity and optimal of the cost function with respect to cost sets which are considered in the optimization analysis.

8. Conclusion

In this research has tackled the complexities of a retrial queueing system with heterogeneous servers, intermittent availability, feedback, and working vacation mechanisms. Employing a matrix geometric approach, we established the steady-state probability distribution and formulated

Table 4. The estimated optimal solutions (γ_1^*, γ_2^*) and their corresponding expected cost $\mathcal{F}^*$

$(\omega,\ \gamma_1,\ \gamma_2)$	γ_1^*	γ_2^*	$\mathcal{F}^*$	Iterations	CPU time(in Sec)
Cost set I					
(3,1.5,0.6)	1.1619	1.6258	80.7970	11	$12.50e^{-6}$
(3,1.5,1)	1.1640	1.0003	72.0522	15	$16.90e^{-6}$
(3,3.5,0.6)	1.0349	1.0190	69.4773	24	$20.60e^{-6}$
(3,3.5,1)	1.0131	1.2647	72.4869	14	$15.60e^{-6}$
(5,1.5,0.6)	1.0119	1.0079	67.7305	20	$21.00e^{-6}$
(5,1.5,1)	1.0533	2.0517	84.8465	14	$15.60e^{-6}$
(5,3.5,0.6)	2.6079	1.0745	108.0219	18	$19.70e^{-6}$
(5,3.5,1)	1.0526	1.0593	69.7106	30	$30.40e^{-6}$
Cost set II					
(3,1.5,0.6)	1.0116	1.4552	58.6088	13	$16.80e^{-6}$
(3,1.5,1)	1.0066	1.0265	54.5188	13	$17.60e^{-6}$
(3,3.5,0.6)	1.1555	1.1191	57.9891	12	$13.60e^{-6}$
(3,3.5,1)	1.0008	1.0089	54.2473	11	$13.50e^{-6}$
(5,1.5,0.6)	1.0102	1.0084	53.4592	10	$11.90e^{-6}$
(5,1.5,1)	1.3133	1.4153	64.3520	19	$22.30e^{-6}$
(5,3.5,0.6)	1.1714	1.0394	57.4689	12	$12.70e^{-6}$
(5,3.5,1)	1.0068	1.0198	53.5026	21	$23.40e^{-6}$

Algorithm 1 Artificial Bee Colony

```
Input: Objective function F(γ1, γ2), Maximum number of iterations
Output: The best solution found
Initialization;
Initialize employed bees with random solutions;
Evaluate the fitness of each solution;
Set the best solution as the solution with the best fitness;
while Termination condition not met do
    for each employed bee do
        Select a solution randomly from the population;
        Generate a new solution by modifying the selected solution;
        Evaluate the fitness of the new solution;
        If the new solution is better, replace the old solution;
    end for
    Update the best solution if a better solution is found;
end while
return The bestsolution found;
```

performance metrics and a cost function. Leveraging the Artificial Bee Colony optimization algorithm, we optimized service rates effectively. Furthermore, we compared our numerical findings with ANFIS results, highlighting the potential synergy between traditional methods and advanced machine learning approaches in queueing theory research. In future, it is possible to expand the proposed model to include additional factors such as different server vacations, server breakdowns, and customer impatience.

REFERENCES

1. Agarwal N. N. *Some Problems in the Theory of Reliability and Queues.* Ph.D. Thesis. Kurukshetra, India: Kurukshetra University, 1965.
2. Ahuja A., Jain A., Jain M. Transient analysis and ANFIS computing of unreliable single server queueing model with multiple stage service and functioning vacation. *Math. Comput. Simulation*, 2022. Vol. 192. P. 464–490. DOI: 10.1016/j.matcom.2021.09.011
3. Bouchentouf A. A., Kadi M., Rabhi A. Analysis of two heterogeneous server queueing model with balking, reneging and feedback. *Math. Sci. Appl E-Notes*, 2014. Vol. 2, No. 2. P. 10–21.
4. Chakravarthy S. R. Analysis of a queueing model with MAP arrivals and heterogeneous phase-type group services. *Mathematics*, 2022. Vol. 10, No. 19. Art. no. 3575. DOI: 10.3390/math10193575
5. Divya K., Indhira K. Analysis of a heterogeneous queuing model with intermittently obtainable servers under a hybrid vacation schedule. *Symmetry*, 2023. Vol. 15, No. 7. Art. no. 1304. DOI: 10.3390/sym15071304
6. Divya K., Indhira K. Performance analysis and ANFIS computing of an unreliable Markovian feedback queuing model under a hybrid vacation policy. *Math. Comput. Simulation*, 2024. Vol. 218. P. 403–419. DOI: 10.1016/j.matcom.2023.12.004
7. Jain M., Jain A. Working vacations queueing model with multiple types of server breakdowns. *Appl. Math. Model.*, 2010. Vol. 34, No. 1. P. 1–13. DOI: 10.1016/j.apm.2009.03.019
8. Neuts M. F. Matrix-Geometric solutions in stochastic models. In: *DGOR. Operations Research Proceedings, vol. 1983. Steckhan H. and al. (eds.).* Berlin, Heidelberg: Springer, 1984. 425 p. DOI: 10.1007/978-3-642-69546-9_91
9. Krishnamoorthy A., Sreenivasan C. An $M/M/2$ queueing system with heterogeneous servers including one with working vacation. *Int. J. Stoch. Anal.*, 2012. Vol. 2012. Art. no. 145867. DOI: 10.1155/2012/145867
10. Kumar A., Jain M. Cost optimization of an unreliable server queue with two stage service process under hybrid vacation policy. *Math. Comput. Simulation*, 2023. Vol. 204. P. 259–281. DOI: 10.1016/j.matcom.2022.08.007
11. Kumar R., Sharma S. K. Two heterogeneous server Markovian queueing model with discouraged arrivals, reneging and retention of reneged customers. *Int. J. Oper. Res.*, 2014. Vol. 11, No. 2. P. 64–68.
12. Leemans H. E. *The Two-Class Two-Server Queueing Model with Nonpreemptive Heterogeneous Priority Structures.* Ph.D. Thesis. Leuven: Katholieke Universiteit Leuven, 1998.
13. Servi L. D., Finn S. G. $M/M/1$ queues with working vacations ($M/M/1/WV$). *Perform. Eval.*, 2002. Vol. 50, No. 1. P. 41–52. DOI: 10.1016/S0166-5316(02)00057-3
14. Morse P. M. *Queues, Inventories and Maintenance: The Analysis of Operational Systems with Variable Demand and Supply.* Mineola, New York: Courier Corporation, 2004. 202 p.
15. Latouche G., Ramaswami V. *Introduction to Matrix Analytic Methods in Stochastic Modeling.* ASA-SIAM Series Stat. Appl. Math. Philadelphia, Pennsylvania: SIAM; Alexandria, Virginia: ASA, 1999. 334 p. DOI: 10.1137/1.9780898719734
16. Seenivasan M., Senthilkumar R., Subasri K. S. $M/M/2$ heterogeneous queueing system having unreliable server with catastrophes and restoration. *Mater. Today: Proc.*, 2022. Vol. 51, No. 8. P. 2332–2338. DOI: 10.1016/j.matpr.2021.11.567
17. Sethi R., Jain M., Meena R. K., Garg D. Cost optimization and ANFIS computing of an unreliable $M/M/1$ queueing system with customers' impatience under N-policy. *Int. J. Appl. Comput. Math.*, 2020. Vol. 6. Art. no. 51. P. 1–14. DOI: 10.1007/s40819-020-0802-0
18. Sanga S. S., Jain M. Cost optimization and ANFIS computing for admission control of $M/M/1/K$ queue with general retrial times and discouragement. *Appl. Math. Comput.*, 2019. Vol. 363. Art. no. 124624. DOI: 10.1016/j.amc.2019.124624
19. Sharda. A queuing problem with intermittently available server and arrivals and departures in batches of variable size. *ZAMM*, 1968. Vol. 48. P. 471–476.
20. Stojčić M., Pamučar D., Mahmutagić E., Stević Ž. Development of an ANFIS Model for the Optimization of a Queuing System in Warehouses. *Information*, 2018. Vol. 9, No. 10. Art. no. 240. 20 p. DOI: 10.3390/info9100240

21. Sudhesh R., Azhagappan A., Dharmaraja S. Transient analysis of $M/M/1$ queue with working vacation, heterogeneous service and customers' impatience. *RAIRO-Oper. Res.*, 2017. Vol. 51, No. 3. P. 591–606. DOI: 10.1051/ro/2016046
22. Thakur S., Jain A., Jain M. ANFIS and cost optimization for Markovian queue with operational vacation. *Int. J. Math. Eng. Management Sci.*, 2021. Vol. 6, No. 3. P. 894–910. DOI: 10.33889/IJMEMS.2021.6.3.053
23. Vijaya Laxmi P., Jyothsna K. Balking and reneging multiple working vacations queue with heterogeneous servers. *J. Math. Model. Algor.*, 2015. Vol. 14. P. 267–285. DOI: 10.1007/s10852-015-9271-6
24. Wu C.-H., Yang D.-Y. Bi-objective optimization of a queueing model with two-phase heterogeneous service. *Comput. Oper. Res.*, 2021. Vol. 130. Art. no. 105230. DOI: 10.1016/j.cor.2021.105230
25. Yohapriyadharsini R. S., Suvitha V. Multi-server Markovian heterogeneous arrivals queue with two kinds of working vacations and impatient customers. *Yugosl. J. Oper. Res.*, 2023. Vol. 33, No. 4. P. 643–666. DOI: 10.2298/YJOR221117011Y
26. Zadeh L. A. The concept of a linguistic variable and its application to approximate reasoning–I. *Inform. Sci.*, 1975. Vol. 8, No. 3. P. 199–249. DOI: 10.1016/0020-0255(75)90036-5

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 77–85
DOI: 10.15826/umj.2023.2.006

CANONICAL APPROXIMATIONS IN IMPULSE STABILIZATION FOR A SYSTEM WITH AFTEREFFECT[1]

Yurii F. Dolgii

Ural Federal University,
19 Mira str., Ekaterinburg, 620002, Russian Federation

jury.dolgy@urfu.ru

Abstract: For optimal stabilization of an autonomous linear system of differential equations with aftereffect and impulse controls, the formulation of the problem in the functional state space is used. For a system with aftereffect, approximating systems of ordinary differential equations proposed by S.N. Shimanov and J. Hale are used. A method for constructing approximations for optimal stabilizing control of an autonomous linear system with aftereffect and impulse controls is proposed. Matrix Riccati equations are used to find approximating controls.

Keywords: Differential equation with aftereffect, Canonical approximation, Optimal stabilization, Impulse control.

1. Introduction

The control object is described as an autonomous linear system of differential equations with aftereffect and impulse control

$$\frac{dx(t)}{dt} = \int_{-\tau}^{0} [d_s\eta(s)]x(t+s) + Bu. \tag{1.1}$$

Here, $t \in \mathbb{R}^+ = (0, +\infty)$, $x : [-\tau, +\infty) \to \mathbb{R}^n$, $\tau > 0$, B is a constant matrix of dimension $n \times r$, the matrix function η has bounded variation on $[-\tau, 0]$, and $\eta(0) = 0$. Impulse controls are generalized functions defined by the formulas

$$u(t) = \frac{dv(t)}{dt}, \quad t \in \mathbb{R}^+,$$

in which control impulses $v\colon [0, +\infty) \to \mathbb{R}^r$ have bounded variations on any finite interval and $v(0) = 0$.

For any initial function $\varphi \in \mathbb{H}$, there is a unique solution $x(t, \varphi)$, $t \geq -\tau$, to equation (1.1) satisfying the condition $x(t, \varphi) = \varphi(t)$, $-\tau \leq t \leq 0$, and the integral equation

$$x(t) = \varphi(0) + \int_0^t \left(\int_{-\tau}^{0} [d_\xi\eta(\xi)]x(s+\xi) \right) ds + B\left(v(t) - v(+0)\right), \quad t \in \mathbb{R}^+.$$

Here, $\mathbb{H} = \mathbb{L}_2([-\tau, 0), \mathbb{R}^n) \times \mathbb{R}^n$ is a Hilbert space of functions with the scalar product

$$\langle \varphi, \psi \rangle_H = \psi^\top(0)\varphi(0) + \int_{-\tau}^{0} \psi^\top(\vartheta)\varphi(\vartheta)d\vartheta.$$

[1]This work was supported by the Russian Science Foundation (project no. 22-21-00714).

Solutions to the integral equation are functions with bounded variations on any finite interval of the positive semi-axis $[0,+\infty)$. They define generalized solutions to the differential equation (1.1).

Need to find an impulse control formed according to the feedback principle, which ensures stable operation of system (1.1) and minimizes a given criterion for the quality of transient processes

$$J=\int\limits_0^{+\infty}\Big(x^\top(t)C_x x(t)+v^\top(t)C_v v(t)\Big)\,dt, \tag{1.2}$$

where C_x and C_v are positive definite matrices.

The problems of optimal stabilization of autonomous linear systems of differential equations with aftereffects for non-impulse controls have been studied quite well [5, 8, 10, 11]. For impulse controls, they were studied in [1, 6, 21]. Constructive procedures for constructing optimal stabilizing controls are associated with finite-dimensional approximations of differential equations with aftereffects. In control problems and the theory of differential games for finite-dimensional approximations of equations with aftereffects, systems of ordinary differential equations proposed by Krasovskii are widely used. Approximations of optimal nonimpulse controls are constructed [4, 8, 12, 15]. An estimate of the accuracy of these approximations in the optimal stabilization problem for differential equations with concentrated delay was obtained by Bykov and Dolgii [2]. In [7], for the problem of optimal impulse stabilization, finite-dimensional approximations to a differential equation with aftereffect proposed by Krasovskii were used.

Canonical approximations were used in the problem of optimal stabilization of systems of differential equations with aftereffect and non-impulse controls in the works of Krasovskii and Osipov [13, 17], Markushin and Shimanov [16], Pandolfi [18, 19], Bykov and Dolgii [3]. In this work, when constructing approximations for optimal impulse stabilizing control, we use canonical approximations to the differential equation with aftereffect.

2. Stabilization problem in a Hilbert state space

When solving the problem, it is convenient, following Krasovskii [14, p. 162], to move from a finite-dimensional to an infinite-dimensional formulation, introducing functional elements

$$\mathbf{x}_t(\vartheta)=x(t+\vartheta),\quad \vartheta\in[-\tau,0],\quad t\ge 0,$$

belonging to a separable Hilbert space $\mathbb{H}$ for solutions of system (1.1).

System (1.1) is associated with the differential equation

$$\frac{d\mathbf{x}_t}{dt}=\mathfrak{A}\mathbf{x}_t+\mathfrak{B}u,\quad t\in\mathbb{R}^+. \tag{2.1}$$

Here, $\mathfrak{A}:\mathbb{H}\to\mathbb{H}$ is an unbounded operator with the domain

$$D(\mathfrak{A})=\big\{\mathbf{x}\in\mathbb{H}:\mathbf{x}\in\mathbb{W}_2^1([-\tau,0],\mathbb{R}^n)\big\}$$

defined by the formulas

$$(\mathfrak{A}\mathbf{x})(\vartheta)=\frac{d\mathbf{x}(\vartheta)}{d\vartheta},\quad \vartheta\in[-\tau,0),\quad (\mathfrak{A}\mathbf{x})(0)=\int\limits_{-\tau}^0[d_s\eta(s)]\mathbf{x}(s).$$

A bounded operator $\mathfrak{B}:\mathbb{R}^r\to\mathbb{H}$ is defined by the formulas

$$(\mathfrak{B}\mathbf{u})(\vartheta)=0,\quad \vartheta\in[-\tau,0),\quad (\mathfrak{B}\mathbf{u})(0)=Bu.$$

The quality criterion for transient processes corresponding to (1.2) has the form

$$\mathbf{J} = \int\limits_{0}^{+\infty} \Big(\langle \mathbf{C}_x \mathbf{x}_t, \mathbf{x}_t \rangle_H + v^\top(t) C_v v(t) \Big)\, dt, \tag{2.2}$$

where a bounded self-adjoint nonnegative operator $\mathbf{C}_x : \mathbb{H} \to \mathbb{H}$ is defined by the formulas

$$(\mathbf{C}_x \mathbf{x})(\vartheta) = 0, \quad \vartheta \in [-\tau, 0), \quad (\mathbf{C}_x \mathbf{x})(0) = C_x \mathbf{x}(0).$$

Using the complexification of the space $\mathbb{H}$, we will consider the scalar product

$$\langle \mathbf{x}, \mathbf{y} \rangle_H = \mathbf{y}^*(0)\mathbf{x}(0) + \int\limits_{-\tau}^{0} \mathbf{y}^*(\vartheta)\mathbf{x}(\vartheta)\, d\vartheta.$$

The eigenvalues of the operator $\mathfrak{A}$ coincide with the roots of the characteristic equation

$$\delta(\lambda) = \det \Delta(\lambda) = 0, \quad \lambda \in \mathbb{C}, \tag{2.3}$$

where (see [14, p. 164])

$$\Delta(\lambda) = \lambda I_n - \int\limits_{-\tau}^{0} [d_s \eta(s)] \exp(\lambda s), \quad \lambda \in \mathbb{C}.$$

We will consider the nondegenerate case when the characteristic equation has a countable number of roots λ_k, $k \in \mathbb{N}$. To simplify further calculations, we will restrict ourselves to describing the canonical expansion procedure only for differential equations (2.1), all roots of the characteristic equations of which are simple. For any $\alpha \in \mathbb{R}$, a finite number of roots of equation (2.3) lie in the half-plane

$$\{\lambda \in \mathbb{C} : \mathrm{Re}\,(\lambda) > \alpha\}.$$

Consequently, they can be numbered in descending order of their real parts, and the numbers of complex conjugate roots must differ by one. The sequence of roots of the characteristic equation satisfies the condition $\mathrm{Re}\,(\lambda_n) \to -\infty$ as $n \to +\infty$. For the general case, the theory of canonical expansion is described in [9, 20].

Choose a positive integer N that satisfies requirement (A):

$$\mathrm{Re}\,(\lambda_n) < 0, \quad n > N.$$

Let $\mathbb{H}^N$ be the linear span of the eigenfunctions of the operator $\mathfrak{A}$ corresponding to its eigenvalues belonging to the set

$$\sigma_N = \{\lambda_1, \dots, \lambda_N\} \subset \sigma(\mathfrak{A}),$$

where $\lambda_k \in \mathbb{C}$, $k = \overline{1, N}$, and $\sigma(\mathfrak{A})$ is the set of eigenvalues of the operator $\mathfrak{A}$. The projector $\mathfrak{P}_N \left(\mathfrak{P}_N \mathbb{H} = \mathbb{H}^N\right)$ defines the canonical decomposition of the space $\mathbb{H}$ into a direct sum, in which an element $\mathbf{x} \in \mathbb{H}$ uniquely defines the elements $\mathbf{x}^N \in \mathbb{H}$ and $\mathbf{z}^N \in (I - \mathfrak{P}_N)\,\mathbb{H}$ such that $\mathbf{x} = \mathbf{x}^N + \mathbf{z}^N$.

When constructing canonical approximations to the stabilization problem, the projection method scheme is used. We use the complexification of state space elements $\mathbf{x} \in \mathbb{H}$ and controls $u \in \mathbb{C}^r$. Applying the projector $\mathfrak{P}_N$ to equation (2.1) and taking into account the equalities

$$\mathfrak{P}_N \mathfrak{A} = \mathfrak{A} \mathfrak{P}_N = \mathfrak{A} \mathfrak{P}_N^2, \quad \mathbf{x}^N = \mathfrak{P}_N \mathbf{x},$$

we obtain the approximating equation

$$\frac{d\mathbf{x}_t^N}{dt} = \mathfrak{A}_N \mathbf{x}_t^N + \mathfrak{B}_N u, \quad t \in \mathbb{R}^+, \tag{2.4}$$

where finite-dimensional operators $\mathfrak{A}_N : \mathbb{H}^N \to \mathbb{H}^N$ and $\mathfrak{B}_N : \mathbb{C}^r \to \mathbb{H}^N$ are defined by the formulas $\mathfrak{A}_N = \mathfrak{A}\mathfrak{P}_N$ and $\mathfrak{B}_N = \mathfrak{P}_N\mathfrak{B}$.

The new quality criterion corresponding to (2.2) has the form

$$\mathbf{J}_N = \int\limits_0^{+\infty} \left(\langle \mathbf{C}_x \mathbf{x}_t^N, \mathbf{x}_t^N \rangle_H + v^*(t) C_v v(t) \right) \, dt. \tag{2.5}$$

3. Finite-dimensional optimal stabilization problem

The subspace $\mathbb{H}^N$ is topologically equivalent to the finite-dimensional Hilbert space $\mathbb{C}^N$ with the inner product $z^* y$, where $y, z \in \mathbb{C}^N$. Let the topological isomorphism be given by the mapping

$$\pi_N : \mathbb{H}^N \to \mathbb{C}^N, \quad x^N = \pi_N \mathbf{x}^N, \quad \mathbf{x}^N \in \mathbb{H}^N, \quad x^N \in \mathbb{C}^N.$$

Using the mapping π_N, we replace equation (2.4) in the spaces $\mathbb{H}^N$ with an equivalent equation in the space $\mathbb{C}^N$

$$\frac{dx^N}{dt} = A_N x^N + B_N u, \quad t \in \mathbb{R}^+, \tag{3.1}$$

where finite-dimensional operators $A_N : \mathbb{C}^N \to \mathbb{C}^N$ and $B_N : \mathbb{C}^r \to \mathbb{C}^N$ are defined by the formulas

$$A_N = \pi_N \mathfrak{A}_N \pi_N^{-1}, \quad B_N = \pi_N \mathfrak{B}_N.$$

The equivalent quality criterion corresponding to (2.5) has the form

$$J_N = \int\limits_0^{+\infty} \left(x^{N*}(t) C_x^N x^N(t) + v^*(t) C_v v(t) \right) \, dt, \tag{3.2}$$

where a finite-dimensional operator $C_x^N : \mathbb{C}^N \to \mathbb{C}^N$ is defined by the formula

$$C_x^N = \pi_N^{-1*} \mathbf{C}_x \pi_N^{-1}.$$

Using the substitutions

$$u(t) = \frac{dv(t)}{dt}, \quad y^N(t) = x^N(t) - B_N v(t), \quad t \in \mathbb{R}^+, \tag{3.3}$$

we replace the finite-dimensional problem of optimal impulse stabilization (3.1), (3.2) with the finite-dimensional problem of optimal nonimpulse stabilization. It is posed for the system of differential equations

$$\frac{dy^N}{dt} = A_N y^N + A_N B_N v, \quad t \in \mathbb{R}^+, \tag{3.4}$$

with new nonimpulse controls v and quality criterion corresponding to (3.2) of the form

$$\hat{J}_N = \int\limits_0^{+\infty} \left(y^{N*}(t) C_{yy}^N y^N(t) + 2 y^{N*}(t) C_{yv}^{N*} v^N(t) + v^*(t) C_{vv}^N v(t) \right) \, dt, \tag{3.5}$$

where

$$C_{yy}^N = C_x^N, \quad C_{yv}^N = C_x^N B_N, \quad C_{vv}^N = C_v + B_N^* C_x^N B_N.$$

Assume that, for the problem of optimal non-impulse stabilization (3.4), (3.5) the matrix Riccati equation,

$$\begin{gathered} K^N A_N + A_N^* K^N + C_x^N - \left(K^N A_N + C_x^N\right) \tilde{C}_{vv}^N \left(A_N^* K^N + C_x^N\right) = 0, \\ \tilde{C}_{vv}^N = B_N \left(C_{vv}^N\right)^{-1} B_N^*, \end{gathered} \tag{3.6}$$

has a unique positive definite solution K^N. Then the optimal stabilizing control of problem (3.4), (3.5) is defined by the formula

$$v^{No}[y^N] = -\left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) y^N, \quad y^N \in \mathbb{C}^N. \tag{3.7}$$

Using formula (3.7), we find optimal stabilizing impulse controls of problem (3.1), (3.2).

Theorem 1. *Let the matrix Riccati equation* (3.6) *have a unique positive definite solution* K^N *and*

$$\det\left(C_v - B_N^* A_N^* K^N B_N\right) \neq 0.$$

Then the optimal stabilizing impulse control of problem (3.1), (3.2) *is defined by the formula*

$$u^{No}[t, x_0^N, x^N] = -\left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right)\left(x_0^N \delta(t) + A_N x^N\right), \quad x^N \in \mathbb{C}^N, \tag{3.8}$$

where $\delta(\cdot)$ *is the Dirac function.*

P r o o f. Using formulas (3.7) and (3.3), we obtain

$$v^N(t) = -\left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right)\left(x^N(t) - B_N v^N(t)\right), \quad t \in \mathbb{R}^+, \quad x^N \in \mathbb{C}^N,$$

or

$$\begin{gathered} \left(I_r - \left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) B_N\right) v^N(t) = \\ -\left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) x^N(t), \quad t \in \mathbb{R}^+, \quad x^N \in \mathbb{C}^N. \end{gathered}$$

Taking into account the equality

$$I_N - \left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) B_N = \left(C_{vv}^N\right)^{-1} \left(C_v - B_N^* A_N^* K^N B_N\right)$$

and the condition

$$\det\left(C_v - B_N^* A_N^* K^N B_N\right) \neq 0,$$

we get

$$\begin{gathered} v^N(t) = -\left(C_v - B_N^* A_N^* K^N B_N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) x^N(t), \\ t \in \mathbb{R}^+, \quad v^N(0) = 0, \quad x^N \in \mathbb{C}^N. \end{gathered}$$

The control v^N is differentiable on the positive semi-axis $\mathbb{R}^+$ and has a unique discontinuity point of the first kind $t = 0$ with a limit value

$$v^N(+0) = -\left(C_v - B_N^* A_N^* K^N B_N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) x_0^N.$$

As a result, the impulse control of problem (3.1), (3.2) is defined by the formula

$$u^N(t) = -\left(C_v - B_N^* A_N^* K^N B_N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right)\left(x_0^N \delta(t) + \frac{dx^N(t)}{dt}\right), \quad t \geq 0, \quad x^N \in \mathbb{C}^N.$$

Using (3.1), we obtain the equality

$$\begin{gathered} u^N(t) = -\left(C_v - B_N^* A_N^* K^N B_N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right)\left(x_0^N \delta(t) + A_N x^N(t) + B_N u^N(t)\right), \\ t \geq 0, \quad x^N \in \mathbb{C}^N. \end{gathered}$$

This explains the validity of formula (3.8), which completes the proof of the theorem. $\square$

4. Stabilizing impulse control of a system of differential equations with aftereffect

Using formula (3.8) and the connection between elements of the spaces $\mathbb{H}^N$ and $\mathbb{C}^N$, we find a stabilizing control for an autonomous linear system of differential equations with aftereffect.

Theorem 2. *Let requirement (A) and the conditions of Theorem 1 be satisfied. Then the control*

$$u^{No}[t,\varphi,\mathbf{x}_t] = -\left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right)\left(\pi_N\varphi\delta(t) + A_N\pi_N\mathbf{x}_t\right), \quad \varphi, \mathbf{x}_t \in \mathbb{H}, \quad t > 0, \tag{4.1}$$

is stabilizing for the system of differential equations with aftereffect (1.1).

P r o o f. For control (4.1), the differential equation (2.1) takes the form

$$\frac{d\mathbf{x}_t}{dt} = \left(\mathfrak{A} - \mathfrak{D}_N A_N \pi\right)\mathbf{x}_t - \mathfrak{D}_N \pi\varphi\delta(t), \quad t \in \mathbb{R}^+.$$

Here

$$\left(\mathfrak{D}_N v\right)(\vartheta) = 0, \quad \vartheta \in [-\tau, 0), \quad \left(\mathfrak{D}_N v\right)(0) = B_N \left(C_{vv}^N\right)^{-1} B_N^* \left(A_N^* K^N + C_x^N\right) v, \quad v \in \mathbb{C}^N.$$

Using the canonical expansion of the space $\mathbb{H}$, we obtain the system of differential equations

$$\frac{d\mathbf{x}_t^N}{dt} = \left(\mathfrak{A}P_N - P_N\mathfrak{D}_N A_N\pi\right)\mathbf{x}_t^N - P_N\mathfrak{D}_N\pi\varphi\delta(t),$$

$$\frac{d\mathbf{z}_t^N}{dt} = \mathfrak{A}\left(I - P_N\right)\mathbf{z}_t^N - \left(I - P_N\right)\mathfrak{D}_N A_N\pi\mathbf{x}_t^N - \left(I - P_N\right)\mathfrak{D}_N\pi\varphi\delta(t), \quad t \geq 0$$

with the initial conditions

$$\mathbf{x}_0^N = P_N\varphi, \quad \mathbf{z}_0^N = \left(I - P_N\right)\varphi.$$

The control used guarantees exponential boundedness of the solutions of the first subsystem with negative exponents. The evolutionary operator $T_N(t)$, $t \in \mathbb{R}^+$, of the homogeneous part of the first subsystem is exponentially bounded with a negative exponent, according to the chosen canonical expansion [9, p. 170].

The solution of the second subsystem is defined by the formula [9, p. 185]

$$\mathbf{z}_t^N = T_N(t)\left(I - P_N\right)\varphi - \int_0^t T_N(t-s)\left(I - P_N\right)\mathfrak{D}_N\left(\mathfrak{A}_N\pi\mathbf{x}_s^N - \pi\varphi\delta(s)\right)ds$$

$$= T_N(t)\left(I - P_N\right)\left(\varphi - \mathfrak{D}_N\pi\varphi\right) - \int_0^t T_N(t-s)\left(I - P_N\right)\mathfrak{D}_N\mathfrak{A}_N\pi\mathbf{x}_s^N ds, \quad t \in \mathbb{R}^+.$$

This implies that the solutions of the second subsystem with negative exponents are exponentially bounded, which completes the proof of the theorem. □

Let us consider the eigenfunctions φ^i, $i = \overline{1,N}$, corresponding to the eigenvalues λ_i, $i = \overline{1,N}$, of the operator $\mathfrak{A}$. Due to their linear independence, they define the basis of the subspace $\mathbb{H}^N$. The eigenfunctions of the operator $\mathfrak{A}$ are defined by the formulas

$$\varphi^k(\vartheta) = \exp(\lambda_k\vartheta)\hat{\varphi}^k, \quad \vartheta \in [-\tau, 0],$$

where $\hat{\varphi}^k$ are nontrivial solutions to the algebraic system

$$\left(\lambda_k I_N - \int_{-\tau}^{0} [d_s\eta(s)] \exp(\lambda_k s)\right)\hat{\varphi}^k = 0, \quad k = \overline{1, N}.$$

To find a coordinate representation of the projector $\mathfrak{P}_N$ in the selected basis, it is necessary to consider for it a biorthogonal system of functions $\{\psi^j\}_{j=1}^N$. The unbounded operator $\mathfrak{A}$ has a dense domain in the space $\mathbb{H}$. Therefore, there is an unbounded conjugate operator $\mathfrak{A}^* : \mathbb{H} \to \mathbb{H}$ with the domain

$$D(\mathfrak{A}^*) = \big\{\mathbf{y} \in \mathbb{H} : \tilde{\mathbf{y}} \in \mathbb{W}_2^1([-\tau, 0], \mathbb{C}^n),\ \tilde{\mathbf{y}}(\vartheta) = \mathbf{y}(\vartheta) - \eta^\top(\vartheta)\mathbf{y}(0),$$
$$\vartheta \in [-\tau, 0],\ \tilde{\mathbf{y}}(-\tau) + \eta^\top(-\tau)\mathbf{y}(0) = 0\big\}.$$

It is defined by the formulas

$$(\mathfrak{A}^*\mathbf{y})(\vartheta) = -\frac{d\tilde{\mathbf{y}}(\vartheta)}{d\vartheta}, \quad \vartheta \in [-\tau, 0), \quad (\mathfrak{A}^*\mathbf{y})(0) = \tilde{\mathbf{y}}(0).$$

The eigenfunctions of the operator $\mathfrak{A}^*$ corresponding to its eigenvalues $\bar{\lambda}_k$, $k \in \mathbb{N}$, are defined by the formulas

$$\psi^k(\vartheta) = \exp(-\bar{\lambda}_k\vartheta)\left(\bar{\lambda}_k I_N - \int_{\vartheta}^{0} [d_s\eta^\top(s)] \exp(\bar{\lambda}_k s)\right)\hat{\psi}^k,$$
$$\vartheta \in [-\tau, 0), \quad \psi^k(0) = \hat{\psi}^k,$$

where $\hat{\psi}^k$ are nontrivial solutions to the algebraic system

$$\left(\bar{\lambda}_k I_N - \int_{-\tau}^{0} [d_s\eta^\top(s)] \exp(\bar{\lambda}_k s)\right)\hat{\psi}^k = 0, \quad k = \overline{1, N}.$$

The requirement of simplicity of the eigenvalues of the operator $\mathfrak{A}$ imposed above generates the biorthogonality of the system of eigenfunctions $\{\psi^j\}_{j=1}^N$ of the operator $\mathfrak{A}^*$ with respect to the system of eigenfunctions $\{\varphi^i\}_{i=1}^N$ of the operator $\mathfrak{A}$. For the fulfilment of the conditions $\langle\varphi^i, \psi^j\rangle_H = \delta_{ij}$, where δ_{ij}, $i, j = \overline{1, N}$, is the Kronecker symbol, it is necessary that

$$1 = \langle\varphi^i, \psi^i\rangle_H = \hat{\psi}^{i*}\left(I_n - \int_{-\tau}^{0} [d_s\eta^\top(s)] s \exp(\lambda_i s)\right)\hat{\varphi}^i, \quad i = \overline{1, N}.$$

These normalization conditions can be ensured by freedom in choosing the vectors $\hat{\psi}^i$, $i = \overline{1, N}$.

Let us define a coordinate representation of the projector $\mathfrak{P}_N$ by the formulas

$$\mathfrak{P}_N\mathbf{x} = \sum_{k=1}^{N} y_k\varphi^k = \mathbf{x}^N = \sum_{k=1}^{N}\langle\mathbf{x}^N, \psi^k\rangle_H\varphi^k, \quad \mathbf{x} \in \mathbb{H}, \quad \mathbf{x}^N \in \mathbb{H}^N, \quad \{y_k\}_{k=1}^N = \mathbf{y}^N \in \mathbb{C}^N.$$

The topological isomorphism $\pi_N : \mathbb{H}^N \to \mathbb{C}^N$ is defined by the formulas

$$\pi_N\mathbf{x}^N = \{\langle\mathbf{x}^N, \psi^k\rangle_H\}_{k=1}^N = \mathbf{y}^N, \quad \pi_N^{-1}\mathbf{y}^N = \sum_{k=1}^{N} y_k\varphi^k = \mathbf{x}^N, \quad \mathbf{x} \in \mathbb{H}, \quad \mathbf{x}^N \in \mathbb{H}^N, \quad \mathbf{y}^N \in \mathbb{C}^N.$$

We have the estimates
$$\|\pi_N\| \leq \left(\sum_{k=1}^{N} \|\psi^k\|^2\right)^{1/2}, \quad \|\pi_N^{-1}\| \leq \lambda_{max},$$
where λ_{max} is the spectral radius of the matrix $\{\langle\varphi^k, \varphi^m\rangle_H\}_{k,m=1}^N$.

Theorem 3. *If the conditions of Theorem 2 hold, then the stabilizing controls for the system of differential equations with aftereffect* (1.1) *are defined by the formulas*
$$u^{No}[t, \varphi, \mathbf{x}_t] = -\left(C_{vv}^N\right)^{-1} B^\top \sum_{i,j=1}^{N} \hat{\psi}^i \left(\bar{\lambda}_i K_{ij}^N + \hat{\varphi}^{i*} C_x \hat{\varphi}^j\right) \left(\langle\varphi, \psi^j\rangle_H \delta(t) + \lambda_j \langle\mathbf{x}_t, \psi^j\rangle_H\right), \tag{4.2}$$
$$\varphi, \mathbf{x}_t \in \mathbb{H}, \quad t > 0,$$
where
$$C_{vv}^N = C_v + B^\top \sum_{i,j=1}^{N} \hat{\psi}^i \hat{\varphi}^{i*} C_x \hat{\varphi}^j \hat{\psi}^{j*} B.$$

P r o o f. Using the coordinate representations of the projector $\mathfrak{P}_N$ and the topological isomorphism π_N, we find the following coordinate representations for the operators:
$$\mathfrak{A}_N \mathbf{x}^N = \mathfrak{A}\mathfrak{P}_N \mathbf{x}^N = \sum_{i=1}^{N} \langle\mathbf{x}^N, \psi^i\rangle_H \mathfrak{A}\varphi^i = \sum_{i=1}^{N} \lambda_i \langle\mathbf{x}^N, \psi^i\rangle_H \varphi^i, \quad \mathbf{x}^N \in \mathbb{H}^N,$$
$$A_N \mathbf{y}^N = \pi_N \mathfrak{A}_N \pi_N^{-1} \mathbf{y}^N = \pi_N \sum_{i=1}^{N} \lambda_i \langle \sum_{k=1}^{N} y_k \varphi^k, \psi^i\rangle_H \varphi^i = \sum_{i=1}^{N} \lambda_i y_i \pi_N \varphi^i$$
$$= \sum_{i=1}^{N} \lambda_i y_i \{\langle\varphi^i, \psi^k\rangle_H\}_{k=1}^N = \{\lambda_k y_k\}_{k=1}^N, \quad \mathbf{y}^N \in \mathbb{C}^N,$$
$$B_N u = \pi_N \mathfrak{B}_N u = \pi_N \mathfrak{P}_N \mathfrak{B} u = \pi_N \sum_{i=1}^{N} \hat{\psi}^{i*} B u \pi_N \varphi^i$$
$$= \sum_{i=1}^{N} \hat{\psi}^{i*} B u \{\langle\varphi^i, \psi^k\rangle_H\}_{k=1}^N = \{\hat{\psi}^{k*} B u\}_{k=1}^N, \quad u \in \mathbb{C}^r,$$
$$C_x^N \mathbf{y}^N = \pi_N^{-1*} \mathbf{C}_x \pi_N^{-1} \mathbf{y}^N = \left\{\langle \mathbf{C}_x \pi_N^{-1} \mathbf{y}^N, \varphi^i\rangle_H\right\}_{i=1}^N$$
$$= \left\{\hat{\varphi}^{i*} \left(C_x \pi_N^{-1} \mathbf{y}^N\right)(0)\right\}_{i=1}^N = \left\{\sum_{k=1}^{N} \hat{\varphi}^{i*} C_x \hat{\varphi}^k y_k\right\}_{i=1}^N, \quad \mathbf{y}^N \in \mathbb{C}^N.$$
Using these formulas, from (4.1) we obtain (4.2), which completes the proof of the theorem. □

As the positive integer N increases, the constructed stabilizing controls approximate the optimal impulse controls for the autonomous linear system of differential equations with aftereffect (1.1).

5. Conclusion

Approximations to an optimal impulse stabilizing control for an autonomous linear system of differential equations with aftereffect have been constructed. Evaluating the accuracy of approximations to an optimal impulse stabilizing control is a challenging problem.

REFERENCES

1. Andreeva I. Yu., Sesekin A. N. An impulse linear-quadratic optimization problem in systems with aftereffect. *Russian Math. (Iz. VUZ),* 1995. Vol. 39, No. 10. P. 8–12.
2. Bykov D. S., Dolgii Yu. F. Error estimate for approximations of an optimal stabilizing control in a delay system. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2012. Vol. 18, No. 2. P. 38–47. (in Russian)
3. Bykov D. S., Dolgii Yu. F. Canonical approximations in task of optimal stabilization of autonomous systems with aftereffect. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2011. Vol. 17, No. 2. P. 20–34. (in Russian)
4. Delfour M. C., McCalla C., Mitter S. K. Stability and the infinite-time quadratic cost problem for linear hereditary differential systems. *SIAM J. Control,* 1975. Vol. 13, No. 1. P. 48–88. DOI: 10.1137/0313004
5. Dolgii Yu. F. Stabilization of linear autonomous systems of differential equations with distributed delay. *Autom. Remote Control*, 2007. Vol. 68, No. 1. P. 1813–1825. DOI: 10.1134/S0005117907100098
6. Dolgii Yu. F., Sesekin A. N. Regularization analysis of a degenerate problem of impulsive stabilization for a system with time delay. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2022. Vol. 28, No. 1. P. 74–95. DOI: 10.21538/0134-4889-2022-28-1-74-95 (in Russian)
7. Dolgii Y. Approximation of the problem of optimal impulse stabilization for an autonomous linear system with delay. In: *Proc. 2022 16th Int. Conf. "Stability and Oscillations of Nonlinear Control Systems" (Pyatnitskiy's Conference 2022). V.N. Tkhai (ed.).* IEEE Xplore, 2022. P. 1–4. DOI: 10.1109/STAB54858.2022.9807487
8. Gibson J. S. Linear-quadratic optimal control of hereditary differential systems: infinite dimensional Riccati equations and numerical approximations. *SIAM J. Control Optim.,* 1983. Vol. 21, No. 1. P. 95–139. DOI: 10.1137/0321006
9. Hale J. K. *Theory of Functional Differential Equations.* NY: Springer-Verlag, 1977. 366 p. DOI: 10.1007/978-1-4612-9892-2
10. Kolmanovskiĭ V. B., Shaĭkhet L. E. *Control of Systems with Aftereffect.* Ser. Monogr. Math., vol. 157. Providence, R.I.: American Mathematical Society, 1996. 336 p.
11. Krasovskii N. N. On the analytic construction of an optimal control in a system with time lags. *J. Appl. Mat. Mech.,* 1962. Vol. 26, No. 1. P. 50–67. DOI: 10.1016/0021-8928(62)90101-6
12. Krasovskii N. N. The approximation of a problem of analytic design of controls in a system with time-lag. *J. Appl. Math. Mech.*, 1964. Vol. 28, No. 4. P. 876–885. (in Russian)
13. Krasovskii N. N., Osipov Yu. S. On stabilization of a controlled object with a delay in a control system. *Izv. Akad. Nauk SSSR, Tekh. Kibern.*, 1963. No. 6. P. 3–15. (in Russian)
14. Krasovskii N. N. *Stability of Motions.* Stanford: University Press, 1963. 212 p.
15. Lukoyanov N. Yu., Plaksin A. R. Finite-dimensional modeling guides in time-delay systems. *Trudy Inst. Mat. i Mekh. UrO RAN*, 2013. Vol. 19, No. 1. P. 182–195. (in Russian)
16. Markushin E. M., Shimanov S. N. Approximate solution of the problem of an analytic construction of a regulator for an equation with retardation. *Differ. Uravn.*, 1966. Vol. 2, No. 8. P. 1018–1026. (in Russian)
17. Osipov Yu. S. On stabilization of control systems with delay. *Differ. Uravn.*, 1965. Vol. 1, No. 5. P. 605–618. (in Russian)
18. Pandolfi L. Stabilization of neutral functional differential equations. *J. Optim. Theory Appl.*, 1976. Vol. 20, No. 2. P. 191–204. DOI: 10.1007/BF01767451
19. Pandolfi L. Canonical realizations of systems with delays. *SIAM J. Control Optim.*, 1983. Vol. 21, No. 4. P. 598–613. DOI: 10.1137/0321036
20. Shimanov S. N. On the theory of linear differential equations wits after-effect. *Differ. Uravn.*, 1965. Vol. 1, No. 1. P. 102–116. (in Russian)
21. Zhelonkina N. I., Lozhnikov A. B., Sesekin A. N. On pulse optimal control of linear systems with aftereffect. *Autom. Remote Control,* 2013. Vol. 74, No. 11. P. 1802–1809. DOI: 10.1134/S0005117913110039

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 86–98

DOI: 10.15826/umj.2023.2.007

KERNEL DETERMINATION PROBLEM FOR ONE PARABOLIC EQUATION WITH MEMORY

Durdimurod K. Durdiev

Bukhara Branch of V.I. Romanovskiy Institute of Mathematics,
Academy of Sciences of Uzbekistan,
11 M. Ikbal St., Bukhara, 200100, Uzbekistan

durdiev65@mail.ru

Zhavlon Z. Nuriddinov

Bukhara State University,
11 M. Ikbal St., Bukhara, 200100, Uzbekistan

j.zafarovich@mail.ru

Abstract: This paper studies the inverse problem of determining a multidimensional kernel function of an integral term which depends on the time variable t and $(n-1)$-dimensional space variable $x'=(x_1,\ldots,x_{n-1})$ in the n-dimensional diffusion equation with a time-variable coefficient at the Laplacian of a direct problem solution. Given a known kernel function, a Cauchy problem is investigated as a direct problem. The integral term in the equation has convolution form: the kernel function is multiplied by a solution of the direct problem's elliptic operator. As an overdetermination condition, the result of the direct question on the hyperplane $x_n=0$ is used. An inverse question is replaced by an auxiliary one, which is more suitable for further investigation. After that, the last problem is reduced to an equivalent system of Volterra-type integral equations of the second order with respect to unknown functions. Applying the fixed point theorem to this system in Hölder spaces, we prove the main result of the paper, which is a local existence and uniqueness theorem.

Keywords: Inverse problem, Resolvent, Integral equation, Fixed point theorem, Existence, Uniqueness.

1. Introduction

The constitutive relations for a linear nonhomogeneous heat propagation and diffusion processes in a medium with memory contain a time- and space-dependent kernel in an integral term of time variable convolution type [11, 14–16, 19]. Often, in practical applications, these kernels are unknown functions, and it is required to determine them. Memory function determination problems in heat equations have been the object of study since the end of the last century. The nonlinear inverse source and linear inverse coefficient problems with different types of over-determination conditions can be mostly found in the literature (see, for example, [1–3, 8, 10, 12, 13, 17, 20, 21] and the references therein). The authors of these researches argued solutions by the special solvability and stability estimates as well as the numerical outlook for solving this type of problems.

Among works devoted to finding the kernel depending on one time variable (one-dimensional inverse problem), we note [4, 14, 16, 19]. Multidimensional inverse problems, when a kernel, in addition to the time variable, also depends on all or a part of spatial variables, are few studied. In this direction, we observe [4, 5, 7, 9, 16]. In [7], the problem of determining a kernel depending on a time variable t and an $(n-1)-$dimensional spatial variable $x'=(x_1,\ldots,x_{n-1})$ was investigated. The principal part of the integrodifferential equation in [7] is an n-dimensional heat conduction operator and the integral part has a form of time-convolution with respect to unknown functions:

the solutions of direct and inverse problems. However, in applications, the study of kernel determination problems is of great interest when the kernel in a convolution type integral is multiplied by an elliptic operator of a solution to the direct problem (see [12]). The present paper considers this kind of parabolic integrodifferential equations, for which the inverse problem will be studied.

Consider the problem of determining functions $u(x,t)$ and $k(x',t)$, $x = (x',x_n) = (x_1,\dots,x_{n-1},x_n)$, $t>0$, from the equations

$$u_t = a(t)\Delta u - \int_0^t k(x',t-\tau)a(\tau)\Delta u(x,\tau)d\tau, \quad (x,t)\in\mathbb{R}^n_T, \tag{1.1}$$

$$u(x,0)=\varphi(x), \quad x\in\mathbb{R}^n, \tag{1.2}$$

$$u(x',0,t)=f(x',t), \quad (x',t)\in\mathbb{R}^{n-1}_T, \quad f(x',0)=\varphi(x',0), \tag{1.3}$$

where Δ is the Laplace operator with respect to spatial variables $x=(x_1,\dots,x_n)$,

$$\mathbb{R}^n_T = \left\{(x,t)|\ x=(x',x_n)\in\mathbb{R}^n,\ 0<t<T\right\}$$

is a strip of thickness T, $T>0$ is an arbitrary fixed number, $a(t)\in C^2[0,T]$, $0<a_0\le a(t)\le a_1<\infty$, and a_0 and a_1 are given numbers.

Our investigations were devoted to the results of [4, 5, 7, 9] under the condition of the integrodifferential heat equation of parabolic type with a variable coefficient and a particular convolution integral.

In this paper, we use the Hölder space H^α with exponent α, where α is a positive integer, for functions depending only on spatial variables. We also use the space $H^{\alpha,\alpha/2}$ with exponents α and $\alpha/2$ for functions depending on both time and spatial variables.

Throughout this paper, we require that

$$\varphi(x)\in H^{l+8}\left(\mathbb{R}^n\right), \quad \varphi(x)\ge\varphi_0=\text{const}>0, \quad f(x',t)\in H^{l+6,(l+6)/2}\left(\bar{\mathbb{R}}^{n-1}_T\right),$$
$$\bar{\mathbb{R}}^{n-1}_T=\left\{(x',t)|x'\in\mathbb{R}^{n-1},\ 0\le t\le T\right\}.$$

The spaces $H^l(Q)$ and $H^{l,l/2}(Q_T)$ and their norms are defined in [6, p. 16–27]. In what follows, we denote by $|\cdot|^{l,l/2}_T$ the norm of functions in the space $H^{l,l/2}(Q_T)$ (in the particular cases $Q_T=\mathbb{R}^n_T$ or $Q_T=\mathbb{R}^{n-1}_T$) depending on time and spatial variables and by $|\cdot|^l$ the norms of functions depending only on spatial variables (for $Q=\mathbb{R}^n$ or $Q=\mathbb{R}^{n-1}$).

The paper is organized as follows. In Section 2, we reduce the inverse problem (1.1)–(1.3) to an auxiliary problem with the additional unknown k outside the integral. In Section 3, using the Poisson formula, we reduce the auxiliary problem to an equivalent system of integral equations with respect to unknown functions. In Section 4, we study the inverse problem as the problem of determining functions $k(t)$ from problem (1.1)–(1.3) using the contraction mapping principle.

2. Preliminaries. Auxiliary problem

Lemma 1. *Let $\{k(t),r(t)\}\in C[0,T]$, and let $k(t)$ and $r(t)$ satisfy the integral equation*

$$r(t)=k(t)+\int_0^t k(t-\tau)r(\tau)d\tau, \quad t\in[0,T].$$

Then a solution of the integral equation

$$\varphi(t)=\int_0^t k(t-\tau)\varphi(\tau)d\tau+f(t), \quad f(t)\in C[0,T],$$

is defined by the formula

$$\varphi(t) = \int\limits_0^t r(t-\tau)f(\tau)d\tau + f(t).$$

P r o o f. We can prove this assertion using a resolvent kernel method for linearly integral equations (see, for example [6]).

Let $u(x,t)$ be the classical solution to the Cauchy problem (1.1)–(1.2). We solve equation (1.1) with respect to $a(t)\Delta u$ and obtain

$$a(t)\Delta u = \int\limits_0^t k(x', t-\tau)a(\tau)\Delta u(x,\tau)d\tau + u_t. \tag{2.1}$$

Then, applying Lemma 1 to (2.1), we obtain for every fixed $x \in \mathbb{R}^n$

$$u_t - a(t)\Delta u = -\int\limits_0^t r(x', t-\tau)u_\tau(x,\tau)d\tau. \tag{2.2}$$

The function $r(x',t)$ in (2.2) is related to $k(x',t)$ as follows:

$$r(x',t) = k(x',t) + \int\limits_0^t k(x', t-\tau)r(x',t)d\tau, \quad (x,t) \in \mathbb{R}^n_T. \tag{2.3}$$

We study the question of finding functions $u(x,t)$ and $r(x',t)$ that satisfy equations (2.2), (1.2), and (1.3). To solve this problem, we first will find $k(x',t)$ from (2.3).

Consider a new function $\vartheta^{(1)}(x,t) = u_{x_n x_n}(x,t)$. Differentiating equations (2.2) and (1.2) twice with respect to x_n, we obtain the following relation for $\vartheta^{(1)}(x,t)$:

$$\vartheta^{(1)}_t - a(t)\Delta\vartheta^{(1)} = -\int\limits_0^t r(x', t-\tau)\vartheta^{(1)}_\tau(x,\tau)d\tau, \tag{2.4}$$

$$\vartheta^{(1)}(x,0) = \varphi_{x_n x_n}(x). \tag{2.5}$$

We obtain an overdetermination condition as follows. Introduce the term $a(t)u_{x_n x_n}$ into the expression $a(t)\Delta u$ of (2.2) and set $x_n = 0$. Then, taking into account that $a(t)u_{x_n x_n} = a(t)\vartheta^{(1)}$ and using (1.2), we get

$$\vartheta^{(1)}(x',0,t) = \frac{1}{a(t)}f_t(x',t) - \sum_{i=1}^{n-1} f_{x_i x_i}(x',t) + \frac{1}{a(t)}\int\limits_0^t r(x',t-\tau)f_\tau(x',\tau)d\tau. \tag{2.6}$$

For the continuity of the function $\vartheta^{(1)}(x,t)$ for $x_n = t = 0$, $x \in \mathbb{R}^{n-1}$, we require the following matching condition:

$$\varphi_{x_n x_n}(x',0) = \frac{1}{a(0)}f_t(x',0) - \sum_{i=1}^{n-1} f_{x_i x_i}(x',0). \tag{2.7}$$

We understand the values of the functions $a(t)$ and $f(x',t)$ and of their derivatives at $t = 0$ as the limit as $t \to +0$.

Consider another transformation of the question. Let $\vartheta^{(2)}(x,t)$ be the derivative of $\vartheta^{(1)}(x,t)$ with respect to t, i.e., let $\vartheta^{(2)}(x,t) := \vartheta^{(1)}_t(x,t)$, and let $h(x',t) := r_t(x',t)$. From (2.4)–(2.6), we get

$$\vartheta^{(2)}_t - a(t)\Delta\vartheta^{(2)} = a'(t)\Delta\vartheta^{(1)} - r(x',0)\vartheta^{(2)} - \int\limits_0^t h(x',t-\tau)\vartheta^{(2)}(x,\tau)d\tau, \tag{2.8}$$

$$\vartheta^{(2)}(x,0) = a(0)\Delta\varphi_{x_nx_n}(x), \tag{2.9}$$

$$\begin{gathered}\vartheta^{(2)}(x',0,t) = \frac{a'(t)}{a^2(t)}f_t(x',t) + \frac{1}{a(t)}f_{tt}(x',t) - \sum_{i=1}^{n-1} f_{tx_ix_i}(x',t)\\ -\frac{a'(t)}{a^2(t)}\int\limits_0^t r(x',t-\tau)f_\tau(x',\tau)d\tau + \frac{1}{a(t)}\int\limits_0^t h(x',\tau)f_\tau(x',t-\tau)d\tau + \frac{1}{a(t)}r(x',0)f_t(x',t).\end{gathered} \tag{2.10}$$

Here, we are obtained the initial condition (2.8) using (2.4) by setting $t=0$ and (2.5). The unknown function $r(x',0)$ is a term of equations (2.8) and (2.10). One can define this function as follows. Similarly to obtaining equality (2.7), we need the continuity of the function $\vartheta^{(2)}(x,t)$ for $x_n = t = 0$, $x \in \mathbb{R}^{n-1}$. Then, (2.9) and (2.10) give some equation, solving which with respect to $r(x',0)$ leads to

$$r(x',0) = \frac{1}{f_t(x',0)}\left[a^2(0)\Delta\varphi_{x_nx_n}(x',0) - \frac{a'(0)}{a(0)}f_t(x',0) - f_{tt}(x',0) + a(0)\sum_{i=1}^{n-1} f_{tx_ix_i}(x',0)\right]. \tag{2.11}$$

In the following calculations, we assume that $r(x',0)$ is known.

Let $\vartheta(x,t) := \vartheta^{(2)}_t(x,t)$. Then, we obtain the main problem of determining $\vartheta(x,t)$ and $h(x',t)$ satisfying the equations

$$\begin{gathered}\vartheta_t - a(t)\Delta\vartheta = 2a'(t)\Delta\vartheta^{(2)} + a''(t)\Delta\vartheta^{(1)}\\ -r(x',0)\vartheta - h(x',t)a(0)\Delta\varphi_{x_nx_n}(x) - \int\limits_0^t h(x',\tau)\vartheta(x,t-\tau)d\tau,\end{gathered} \tag{2.12}$$

$$\vartheta(x,0) = \Psi(x), \tag{2.13}$$

$$\begin{gathered}\vartheta(x',0,t) = F(x',t) + \left(2\frac{(a'(t))^2}{a^3(t)} - \frac{a''(t)}{a^2(t)}\right)\int\limits_0^t r(x',t-\tau)f_\tau(x',\tau)d\tau\\ -2\frac{a'(t)}{a^2(t)}\int\limits_0^t h(x',\tau)f_\tau(x',t-\tau)d\tau - \frac{1}{a(t)}\int\limits_0^t h(x',\tau)f_{tt}(x',t-\tau)d\tau + \frac{1}{a(t)}h(x',t)f_t(x',0),\end{gathered} \tag{2.14}$$

where

$$\Psi(x) = a^2(0)\Delta^2\varphi_{x_nx_n}(x) + a'(0)\Delta\varphi_{x_nx_n}(x) - r(x',0)a(0)\Delta\varphi_{x_nx_n}(x),$$

and therefore we get

$$\begin{gathered}F(x',t) = \left(\frac{a''(t)}{a^2(t)} - \frac{(a'(t))^2}{a^3(t)}\right)f_t(x',t) + \frac{1}{a(t)}f_{ttt}(x',t) - \sum_{i=1}^{n-1} f_{ttx_ix_i}(x',t)\\ -2\frac{a'(t)}{a^2(t)}r(x',0)f_t(x',t) + \frac{1}{a(t)}r(x',0)f_{tt}(x',t).\end{gathered}$$

Equation (2.12) contains $2a'(t)\Delta\vartheta^{(2)} + a''(t)\Delta\vartheta^{(1)}$ on the right-hand side. Taking into consideration $\vartheta_t^{(1)} = \vartheta^{(2)}$ and using (2.4), we replace it by $\vartheta^{(2)}$:

$$a''(t)\Delta\vartheta^{(1)} = \frac{a''(t)}{a(t)}\vartheta^{(2)} + \frac{a''(t)}{a(t)}\int\limits_0^t r(x',t-\tau)\vartheta^{(2)}(x,\tau)d\tau. \tag{2.15}$$

Similarly, from (2.4) and (2.8), we obtain

$$\begin{aligned} 2a'(t)\Delta\vartheta^{(2)} = 2\,(\ln a(t))'\Bigg[\vartheta - (\ln a(t))'\left(\vartheta^{(2)} + \int\limits_0^t r(x',t-\tau)\vartheta^{(2)}(x,\tau)d\tau\right) \\ -r(x',0)\vartheta^{(2)} - \int\limits_0^t h(x',t-\tau)\vartheta^{(2)}(x,\tau)d\tau\Bigg]. \end{aligned} \tag{2.16}$$

Further, we will deduce that the relation $2a'(t)\Delta\vartheta^{(2)}+a''(t)\Delta\vartheta^{(1)}$ in equation (2.12) is eliminated with the help of (2.15) and (2.16).

In case (2.7) and (2.11), it does not bring difficulties following out the inverse changes to derive the equations (1.1)–(1.3) from (2.8), (2.9), and (2.12)–(2.14) [7]. So, the inverse problem (1.1)–(1.3) is similar to problem (2.8), (2.9), and (2.12)–(2.14) of determining the functions $\vartheta^{(2)}(x,t)$, $\vartheta(x,t)$, $h(x',t)$, and $r(x,'t)$. □

3. Reduction of the auxiliary problem

The following statement is the main result of this section.

Lemma 2. *The auxiliary problems* (2.8)–(2.9), (2.12)–(2.13), *and the equality* $h(x',t) := r_t(x',t)$, *are equivalent to the problem of finding the functions* $\vartheta^{(2)}(x,t)$, $\vartheta(x,t)$, $h(x',t)$, *and* $r(x,'t)$ *from the following system of integral equations*:

$$\begin{aligned} \vartheta^{(2)}(x,t) = \int\limits_{\mathbb{R}^n} a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)G\left(x-\xi,\theta(t)\right)d\xi + \int\limits_0^{\theta(t)} \frac{d\tau}{a(\theta^{-1}(\tau))} \\ \times\int\limits_{\mathbb{R}^n}\Bigg[\left(\ln a(\theta^{-1}(\tau))\right)'\left(\vartheta^{(2)}\left(\xi,\theta^{-1}(\tau)\right) + \int\limits_0^{\theta^{-1}(\tau)} r(\xi',\theta^{-1}(\tau)-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha\right) \\ -r(\xi',0)\vartheta^{(2)}\left(\xi,\theta^{-1}(\tau)\right) - \int\limits_0^{\theta^{-1}(\tau)} h(\xi',\theta^{-1}(\tau)-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha\Bigg]G\left(x-\xi,\theta(t)-\tau\right)d\xi, \end{aligned} \tag{3.1}$$

$$\begin{aligned} \vartheta(x,t) = \int\limits_{\mathbb{R}^n}\Psi(\xi)G(x-\xi,\theta(t))d\xi + \int\limits_0^{\theta(t)}\frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}\Bigg[\Bigg(\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))} \\ -2((\ln a(\theta^{-1}(\tau)))')^2\Bigg)\vartheta^{(2)}(\xi,\theta^{-1}(\tau)) + \Big(2(\ln a(\theta^{-1}(\tau)))' - r(\xi',0)\Big)\vartheta(\xi,\theta^{-1}(\tau)) \end{aligned}$$

$$-\int\limits_0^{\theta^{-1}(\tau)} h(\xi',\alpha)\vartheta(\xi,\theta^{-1}(\tau)-\alpha)d\alpha+\Big(\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))}+2(\ln a(\theta^{-1}(\tau)))'$$

$$-2((\ln a(\theta^{-1}(\tau)))')^2\Big)\int\limits_0^{\theta^{-1}(\tau)} r(\xi',\theta^{-1}(\tau)-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha++2(\ln a(\theta^{-1}(\tau)))' \tag{3.2}$$

$$\times\int\limits_0^{\theta^{-1}(\tau)} h(\xi',\theta^{-1}(\tau)-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha-h(\xi'\theta^{-1}(\tau))a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)\Big]G(x-\xi,\theta(t)-\tau)d\xi,$$

$$h(x',t)=\frac{a(t)}{f_t(x',0)}\Big[\int\limits_{\mathbb{R}^n}\Psi(\xi)G(x'-\xi',\xi_n,\theta(t))d\xi-F(x',t)\Big]$$

$$+\frac{a(t)}{f_t(x',0)}\Big[\int\limits_0^{\theta(t)}\frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}\Big(\Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))}-2((\ln a(\theta^{-1}(\tau)))')^2\Big]\vartheta^{(2)}(\xi,\theta^{-1}(\tau))$$

$$+\Big[2(\ln a(\theta^{-1}(\tau)))'-r(\xi',0)\Big]\vartheta(\xi,\theta^{-1}(\tau))-\int\limits_0^{\theta^{-1}(\tau)} h(\xi',\alpha)\vartheta(\xi,\theta^{-1}(\tau)-\alpha)d\alpha$$

$$+\Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))}+2(\ln a(\theta^{-1}(\tau)))'-2((\ln a(\theta^{-1}(\tau)))')^2\Big]\int\limits_0^{\theta^{-1}(\tau)} r(\xi',\tau-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha \tag{3.3}$$

$$+2(\ln a(\theta^{-1}(\tau)))'\int\limits_0^{\theta^{-1}(\tau)} h(\xi',\theta^{-1}(\tau)-\alpha)\vartheta^{(2)}(\xi,\alpha)d\alpha-h(\xi',\theta^{-1}(\tau))a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)\Big)$$

$$\times G(x'-\xi',\xi_n,\theta(t)-\tau)d\xi\Big]-f_t(x',0)(2((\ln(a(t))')^2-\frac{a''(t)}{a(t)})\int\limits_0^t r(x',t-\tau)f_\tau(x',\tau)d\tau+$$

$$+2f_t(x',0)(\ln(a(t))'\int\limits_0^t h(x',\tau)f_\tau(x',t-\tau)d\tau+f_t(x',0)\int\limits_0^t h(x',\tau)f_{tt}(x',t-\tau)d\tau,$$

$$r(x,'t)=r(x',0)+\int\limits_0^t h(x',\tau)d\tau. \tag{3.4}$$

P r o o f. To prove Lemma 2, we use the formula [3]

$$p(x,t)=\int\limits_{\mathbb{R}^n}\varphi(\xi)G(x-\xi;\theta(t))d\xi+\int\limits_0^{\theta(t)}\frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}F(\xi,\theta^{-1}(\tau))G(x-\xi;\theta(t)-\tau)d\xi, \tag{3.5}$$

which provides a solution to the following Cauchy problem for the heat equation with a time-variable coefficient of thermal conductivity:

$$p_t-a(t)\Delta p=F(x,t),\quad x\in\mathbb{R}^n,\quad t>0,$$
$$p(x,0)=\varphi(x),\quad x\in\mathbb{R}.$$

In (3.5),

$$\theta(t) = \int\limits_0^t a(\tau)d\tau$$

and $\theta^{-1}(t)$ is the inverse function of $\theta(t)$;

$$G(x-\xi;\theta(t)-\tau) = \frac{1}{(2\sqrt{\pi(\theta(t)-\tau)})^n} e^{-|x-\xi|^2/4(\theta(t)-\tau)}$$

is the fundamental solution related to the operator of heat with the coefficient of thermal conductivity that depends on time:

$$\frac{\partial}{\partial t} - a(t)\Delta, \quad \xi = (\xi_1, \ldots, \xi_n), \quad \xi' = (\xi_1, \ldots, \xi_{n-1}), \quad d\xi = d\xi_1 \cdots d\xi_n, \quad |x|^2 = x_1^2 + \cdots + x_n^2.$$

Equations (3.1) and (3.2) follow from the Cauchy problems (2.8), (2.9) and (2.12), (2.13) with (3.5), independently. In (3.2), we set $x_n = 0$ and use another case of (2.14). After that, we get equation (3.3). Equality (3.4) is clear.

We add to the equations (3.1)–(3.4) the integral equation. It can be gained from relations (2.2) and (1.2). First, we use formula (3.5) after integrating by parts in the integral on the right-hand side of (2.2). In conclusion, we get the following equivalent integral equation for $u(x,t)$:

$$\begin{aligned} u(x,t) = \int\limits_{\mathbb{R}^n} \varphi(\xi) G(x-\xi;\theta(t))d\xi + \int\limits_0^{\theta(t)} \frac{d\tau}{a(\theta^{-1}(\tau))} \int\limits_{\mathbb{R}^n} \Big[r(\xi', \theta^{-1}(\tau))\varphi(\xi) \\ -r(\xi',0)u(\xi,\theta^{-1}(\tau)) - \int\limits_0^{\theta^{-1}(\tau)} h(\xi', \theta^{-1}(\tau)-\alpha)u(\xi,\alpha)d\alpha \Big] G(x-\xi;\theta(t)-\tau)d\xi. \end{aligned} \tag{3.6}$$

□

4. Existence and uniqueness

In this section, we show that a solution to the system of integral equations (3.1)–(3.4), (3.6) exists and is unique. To this end, we use the well-known Banach's principle [18, pp. 87–97]. Our goal is to set the integral equations like a system with a nonlinear operator for unknown functions $\vartheta^{(2)}(x,t)$, $\vartheta(x,t)$, $h(x',t)$, and $r(x',t)$, and show that an operator of this type is a contraction mapping operator. The uniqueness and existence then follow straight away.

Recall that F is a contraction mapping operator in a closed set Ω, which is a subset of a Banach space, if it satisfies the following two properties:

(1) if $y \in \Omega$, then $Fy \in \Omega$ (i.e., F maps Ω into itself);
(2) if $y, z \in \Omega$, then $\|Fy - Fz\| \le \rho \|y - z\|$ with $\rho < 1$ (ρ is a constant independent of y and z).

Right now, we introduce the primary result of this research.

Theorem 1. *Suppose that all cases of Section 1 on regard to the drawn functions $a(t)$, $\varphi(x)$, and $f(x',t)$ and the matching cases (1.3) and (2.7) are fulfilled except $|f_t(x',0)| > f_0 = \mathrm{const} > 0$, f_0 is a fixed number. Then there is a sufficiently small number $T > 0$ such that the unique answer to the inverse question (1.1)–(1.3) exists in the class of functions $u(x,t) \in H^{l+2,(l+2)/2}\left(\bar{\mathbb{R}}^n_T\right)$ and $k(x',t) \in H^{l,l/2}\left(\bar{\mathbb{R}}^{n-1}_T\right)$.*

P r o o f. The system of equations (3.1)–(3.4), (3.6) is a closed system of unknown functions $\vartheta^{(2)}(x,t)$, $\vartheta(x,t)$, $h(x',t)$, $r(x,'t)$, and $u(x,t)$ in $\mathbb{R}^n_T$. It can be written as a nonlinear operator equation

$$\psi = A\psi; \tag{4.1}$$

here $\psi = (\psi_1,\psi_2,\psi_3,\psi_4,\psi_5)^* = \big(\vartheta^{(2)}(x,t),\vartheta(x,t),h(x',t),r(x,'t),u(x,t)\big)^*$, where $*$ is the transposition symbol. According to equations (3.1)–(3.4) and (3.6), the operator $A\psi = [(A\psi)_1,(A\psi)_2,(A\psi)_3,(A\psi)_4,(A\psi)_5]$ has the form

$$\begin{aligned}(A\psi)_1 = \psi_{01}(x,t) + \int\limits_0^{\theta(t)} \frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}\Big[(\ln a(\theta^{-1}(\tau)))'\Big(\psi_1(\xi,\theta^{-1}(\tau))\\ + \int\limits_0^{\theta^{-1}(\tau)}\psi_4(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha\Big) - r(\xi',0)\psi_1(\xi,\alpha)\\ - \int\limits_0^{\theta^{-1}(\tau)}\psi_3(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha\Big]G(x-\xi,\theta(t)-\tau)d\xi,\end{aligned} \tag{4.2}$$

$$\begin{aligned}(A\psi)_2 = \psi_{02}(x,t) + \int\limits_0^{\theta(t)}\frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}\Big(\Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))} - 2((\ln a(\theta^{-1}(\tau)))')^2\Big]\psi_1(\xi,\theta^{-1}(\tau))\\ +\Big[2(\ln a(\theta^{-1}(\tau)))' - r(\xi',0)\Big]\psi_2(\xi,\theta^{-1}(\tau)) - \int\limits_0^{\theta^{-1}(\tau)}\psi_3(\xi',\alpha)\times\psi_2(\xi,\theta^{-1}(\tau)-\alpha)d\alpha\\ +\Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))} + 2(\ln a(\theta^{-1}(\tau)))' - 2((\ln a(\theta^{-1}(\tau)))')^2\Big]\int\limits_0^{\theta^{-1}(\tau)}\psi_4(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha\\ +2(\ln a(\theta^{-1}(\tau)))'\int\limits_0^{\theta^{-1}(\tau)}\psi_3(\xi',\theta^{-1}(\tau)-\alpha)\\ \times\psi_1(\xi,\alpha)d\alpha - \psi_3(\xi',\theta^{-1}(\tau))a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)\Big)G(x-\xi,\theta(t)-\tau)d\xi,\end{aligned} \tag{4.3}$$

$$\begin{aligned}(A\psi)_3 = \psi_{03}(x',t) + \frac{a(t)}{f_t(x',0)}\Big[\int\limits_0^{\theta(t)}\frac{d\tau}{a(\theta^{-1}(\tau))}\int\limits_{\mathbb{R}^n}\Big(\Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))}\\ -2((\ln a(\theta^{-1}(\tau)))')^2\Big]\psi_1(\xi,\theta^{-1}(\tau)) + \Big[2(\ln a(\theta^{-1}(\tau)))' - r(\xi',0)\Big]\psi_2(\xi,\theta^{-1}(\tau))\\ - \int\limits_0^{\theta^{-1}(\tau)}\psi_3(\xi',\alpha)\psi_2(\xi,\theta^{-1}(\tau)-\alpha)d\alpha + \Big[\frac{a''(\theta^{-1}(\tau))}{a(\theta^{-1}(\tau))} + 2(\ln a(\theta^{-1}(\tau)))'\\ -2((\ln a(\theta^{-1}(\tau)))')^2\Big]\int\limits_0^{\theta^{-1}(\tau)}\psi_1(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha + 2(\ln a(\theta^{-1}(\tau)))'\times\end{aligned}$$

$$\times \int_0^{\theta^{-1}(\tau)} \psi_3(\xi', \theta^{-1}(\tau) - \alpha)\psi_1(\xi, \alpha)d\alpha - \psi_3(\xi', \theta^{-1}(\tau))a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)\Big)$$

$$\times G(x' - \xi', \xi_n, \theta(t) - \tau)d\xi\Big] - f_t(x', 0)(2((\ln(a(t))')^2 - \frac{a''(t)}{a(t)}) \int_0^t \psi_4(x', t - \tau)f_\tau(x', \tau)d\tau \tag{4.4}$$

$$+2f_t(x', 0)(\ln(a(t))' \int_0^t \psi_3(x', \tau)f_\tau(x', t - \tau)d\tau + f_t(x', 0) \int_0^t \psi_3(x', \tau)f_{tt}(x', t - \tau)d\tau,$$

$$(A\psi)_4 = \psi_{04}(x', t) + \int_0^t \psi_3(x', \tau)d\tau, \tag{4.5}$$

$$\begin{gathered}(A\psi)_5 = \psi_{05}(x, t) + \int_0^{\theta(t)} \frac{d\tau}{a(\theta^{-1}(\tau))} \int_{\mathbb{R}^n} \Big[\psi_4(\xi', \theta^{-1}(\tau))\varphi(\xi) \\ -r(\xi', 0)\psi_5(\xi, \theta^{-1}(\tau)) - \int_0^{\theta^{-1}(\tau)} \psi_3(\xi', \theta^{-1}(\tau) - \alpha)\psi_5(\xi, \alpha)d\alpha\Big] G(x - \xi; \theta(t) - \tau)d\xi.\end{gathered} \tag{4.6}$$

In (4.2)–(4.6), we introduced the notation:

$$\psi_{01}(x, t) = \int_{\mathbb{R}^n} a(0)\Delta\varphi_{\xi_n\xi_n}(\xi)G(x - \xi, \theta(t))d\xi,$$

$$\psi_{02}(x, t) = \int_{\mathbb{R}^n} \Psi(\xi)G(x - \xi, \theta(t))d\xi,$$

$$\psi_{03}(x', t) = \frac{a(t)}{f_t(x', 0)}\Big[\int_{\mathbb{R}^n} \Psi(\xi)G(x' - \xi', \xi_n, \theta(t))d\xi - F(x', t)\Big],$$

$$\psi_{04}(x', t) = r(x', 0), \quad \psi_{05}(x, t) = \int_{\mathbb{R}^n} \varphi(\xi)G(x - \xi; \theta(t))d\xi.$$

Define

$$|\psi|_T^{l,l/2} = \max\big(|\psi_1|_T^{l,l/2},\ |\psi_2|_T^{l,l/2},\ |\psi_3|_T^{l,l/2}, |\psi_4|_T^{l,l/2},\ |\psi_5|_T^{l,l/2}\big),$$

fix T_0 such that $T_0 > T$, and consider in the space $H^{l,l/2}(\mathbb{R}^n_T)$ the set $S(T)$ of functions $\psi(x, t)$ satisfying the inequality

$$|\psi - \psi_0|_T^{l,l/2} \le |\psi_0|_{T_0}^{l,l/2}, \tag{4.7}$$

where $\psi_0 = (\psi_{01},\ \psi_{02},\ \psi_{03},\ \psi_{04},\ \psi_{05})$ and

$$|\psi_0|_{T_0}^{l,l/2} = \max\big(|\psi_{01}|_{T_0}^{l,l/2},\ |\psi_{02}|_{T_0}^{l,l/2},\ |\psi_{03}|_{T_0}^{l,l/2}\ |\psi_{04}|_{T_0}^{l,l/2},\ |\psi_{05}|_{T_0}^{l,l/2}\big).$$

For a sufficiently small T, the operator A is a contraction mapping operator in $S(T)$. Then the uniqueness and existence theorem follows right away from the contraction mapping principle.

First, it is seen that A has the first property of a contraction mapping operator. Let $\psi \in S(T)$, $T < T_0$. Then, from relation (4.7), we have

$$|\psi_i|_T^{l,l/2} \le 2\,|\psi_0|_{T_0}^{l,l/2}, \quad i = 1, 2, 3, 4, 5.$$

Define

$$a_1 := \|a\|_{C^2[0,T]}, \quad a_2 := \max_{t\in[0,T]} |(\ln a(t))'|$$
$$r_1 = |r(x',0)|^l, \quad f_1 := |f(x',t-\tau)|^{l+6,(l+6)/2}, \quad \varphi_1 := |\varphi(x)|^{l+6}.$$

It is not hard to see that

$$|(A\psi)_1 - \psi_{01}|_T^{l,l/2} = \Big| \int\limits_0^{\theta(t)} \frac{d\tau}{a(\theta^{-1}(\tau))} \int\limits_{\mathbb{R}^n} \Big[(\ln a(\theta^{-1}(\tau)))' \Big(\psi_1(\xi,\theta^{-1}(\tau))$$
$$+ \int\limits_0^{\theta^{-1}(\tau)} \psi_4(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha \Big) - r(\xi',0)\psi_1(\xi,\alpha)$$
$$- \int\limits_0^{\theta^{-1}(\tau)} \psi_3(\xi',\theta^{-1}(\tau)-\alpha)\psi_1(\xi,\alpha)d\alpha \Big] G(x-\xi,\theta(t)-\tau)d\xi \Big|_T^{l,l/2}$$
$$\le \int\limits_0^{\theta(t)} \frac{d\tau}{|a(\theta^{-1}(\tau))|_T} \int\limits_{\mathbb{R}^n} \Big[|(\ln a(\theta^{-1}(\tau)))'|_T$$
$$\times \Big(\psi_1(\xi,\theta^{-1}(\tau))|_T^{l,l/2} + \int\limits_0^{\theta^{-1}(\tau)} |\psi_4(\xi',\theta^{-1}(\tau)-\alpha)|_T^l |\psi_1(\xi,\alpha)|_T^l d\alpha \Big) + |r(\xi',0)|^l |\psi_1(\xi,\alpha)|_T^{l,l/2}$$
$$+ \int\limits_0^{\theta^{-1}(\tau)} |\psi_3(\xi',\theta^{-1}(\tau)-\alpha)|_T^{l,l/2} |\psi_1(\xi,\alpha)|_T^{l,l/2} d\alpha \Big] G(x-\xi,\theta(t)-\tau)d\xi$$
$$\le |\psi_0|_{T_0}^{l,l/2} \frac{2T^2}{a_0} \Big(a_2 + 2Ta_2|\psi_0|_{T_0}^{l,l/2} + r_1 + 2T|\psi_0|_{T_0}^{l,l/2} \Big) := |\psi_0|_{T_0}^{l,l/2}\beta_1.$$

In the same way, we obtain

$$|(A\psi)_2 - \psi_{02}|_T^{l,l/2} \le |\psi_0|_{T_0}^{l,l/2} \Big[\frac{2T^2}{a_0} \Big(\frac{a_1}{a_0} + 2a_2^2 + 2a_2 + r_1 + 2T|\psi_0|_{T_0}^{l,l/2} \Big)$$
$$+2T|\psi_0|_{T_0}^{l,l/2} \Big(\frac{a_2}{a_0} + 2a_2 + 2a_2^2 \Big) + 4Ta_2|\psi_0|_{T_0}^{l,l/2} + a_1\varphi_1 \Big] := |\psi_0|_{T_0}^{l,l/2}\beta_2,$$
$$|(A\psi)_3 - \psi_{03}|_T^{l,l/2} \le |\psi_0|_{T_0}^{l,l/2} \Big(2\frac{T^2}{f_1} \Big[\frac{a_1}{a_0} + 2a_2^2 + 2a_2 + r_1 + 2T|\psi_0|_{T_0}^{l,l/2}$$
$$+2T|\psi_0|_{T_0}^{l,l/2} \Big(\frac{a_1}{a_0} + 2a_2 + 2a_2^2 \Big) + 4T|\psi_0|_{T_0}^{l,l/2} a_2 + a_1\varphi_1 \Big] + Tf_1^2 \Big(\frac{a_1}{a_0} + 2a_2^2 + 2a_2 + 1 \Big) \Big) := |\psi_0|_{T_0}^{l,l/2}\beta_3$$
$$|(A\psi)_4 - \psi_{04}|_T^{l,l/2} \le 2T|\psi_0|_{T_0}^{l,l/2} := |\psi_0|_{T_0}^{l,l/2}\beta_4,$$
$$|(A\psi)_5 - \psi_{05}|_T^{l,l/2} \le |\psi_0|_{T_0}^{l,l/2} \cdot \frac{2T^2}{a_0} (\varphi_1 + r_1 + 2T|\psi_0|_{T_0}^{l,l/2}) := |\psi_0|_{T_0}^{l,l/2}\beta_5,$$

where $\beta_i(T) \to 0$ as $T \to 0$, $i = 1,2,3,4,5$. Accordingly, if we take T $(T < T_0)$ such that the following relation holds:

$$\beta := \max\{\beta_1,\ \beta_2,\ \beta_3,\ \beta_4, \beta_5\} < 1,$$

then the operator A has the first property of a contraction operator of mapping, i.e., $A\psi \in S(T)$. Next, let us think about the second property of a contraction mapping operator for A. Let

$$\psi^{(1)} = \left(\psi_1^{(1)}, \psi_2^{(1)}, \psi_3^{(1)}, \psi_4^{(1)}, \psi_5^{(1)}\right) \in S(T), \quad \psi^{(2)} = \left(\psi_1^{(2)}, \psi_2^{(2)}, \psi_3^{(2)}, \psi_4^{(2)}, \psi_5^{(2)}\right) \in S(T).$$

Based on the inequalities

$$\begin{aligned}\big|\psi_2^{(1)}\psi_1^{(1)} - \psi_2^{(2)}\psi_1^{(2)}\big|_T^{l,l/2} &= \Big|\big(\psi_2^{(1)} - \psi_2^{(2)}\big)\psi_1^{(1)} + \psi_2^{(2)}\big(\psi_1^{(1)} - \psi_1^{(2)}\big)\Big|_T^{l,l/2} \\ \le 2\big|\psi^{(1)} - \psi^{(2)}\big|_T^{l,l/2} &\max\left(\big|\psi_1^{(1)}\big|_T^{l,l/2},\ \big|\psi_2^{(2)}\big|_T^{l,l/2}\right) \le 4\,|\psi_0|_T^{l,l/2}\,\big|\psi^{(1)} - \psi^{(2)}\big|_T^{l,l/2},\end{aligned}$$

we evaluate the difference

$$\begin{aligned}|((A\psi)^{(1)} - A\psi)^{(2)})_1|_T^{l,l/2} \le \int\limits_0^{\theta(t)} \frac{d\tau}{|a(\theta^{-1}(\tau))|_T} \int\limits_{\mathbb{R}^n} \Big[&|(\ln a(\theta^{-1}(\tau)))'|_T|\Big((\psi_1^{(1)}(\xi,\theta^{-1}(\tau)) \\ -\psi_1^{(2)}(\xi,\theta^{-1}(\tau)))|_T^{l,l/2} + &\int\limits_0^{\theta^{-1}(\tau)} |\Big[\psi_4^{(1)}(\xi',\theta^{-1}(\tau)-\alpha)\psi_1^{(1)}(\xi,\alpha) \\ -\psi_4^{(2)}(\xi',\theta^{-1}(\tau)-\alpha)\psi_1^{(2)}(\xi,\alpha)\Big]|_T^{l,l/2}d\alpha\Big) &+ |r(\xi',0)|^l|(\psi_1^{(1)}(\xi,\alpha) - \psi_1^{(2)}(\alpha))|_T^{l,l/2} \\ + \int\limits_0^{\theta^{-1}(\tau)} |\Big[\psi_3^{(1)}(\xi',\theta^{-1}(\tau)-\alpha)&\psi_1^{(1)}(\xi,\alpha) - \psi_3^{(2)}(\xi',\theta^{-1}(\tau) \\ -\alpha)\psi_1^{(2)}(\xi,\alpha)\Big]|_T^{l,l/2}d\alpha&\Big]G(x-\xi,\theta(t)-\tau)d\xi \\ \le |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\frac{T^2}{a_0}(a_2 + 4Ta_2|\psi_0|_{T_0}^{l,l/2} + r_1 + 4T|\psi_0|_{T_0}^{l,l/2}) &:= |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\mu_1.\end{aligned}$$

For other components of A, we can write

$$\begin{aligned}|((A\psi)^{(1)} - A\psi)^{(2)})_2|_T^{l,l/2} &\le |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\bigg(\frac{T^2}{a_0}\Big(\frac{a_1}{a_0} + 2a_2^2 + 2a_2 + r_1 + 4T|\psi_0|_{T_0}^{l,l/2}\Big) \\ +4T|\psi_0|_{T_0}^{l,l/2}\Big(\frac{a_1}{a_0} + 2a_2 + 2a_2^2\Big) &+ 8Ta_2|\psi_0|_{T_0}^{l,l/2} + a_1\varphi_1\bigg) := |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\mu_2, \\ |((A\psi)^{(1)} - A\psi)^{(2)})_3|_T^{l,l/2} &\le |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\Big(\frac{T^2}{f_1}\Big[\frac{a_1}{a_0} + 2a_2^2 + 2a_2 + r_1 + 4T|\psi_0|_{T_0}^{l,l/2} \\ +4T|\psi_0|_{T_0}^{l,l/2}&\Big(\frac{a_1}{a_0} + 2a_2 + 2a_2^2\Big) + 8T|\psi_0|_{T_0}^{l,l/2}a_2 + a_1\varphi_1\Big] \\ +Tf_1^2\Big(\frac{a_1}{a_0} &+ 2a_2^2 + 2a_2 + 1\Big)\Big) := |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\mu_3, \\ |((A\psi)^{(1)} - A\psi)^{(2)})_4|_T^{l,l/2} &\le |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}T := |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\mu_4, \\ |((A\psi)^{(1)} - A\psi)^{(2)})_5|_T^{l,l/2} &\le |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\frac{T^2}{a_0}(\varphi_1 + r_1 + 4T|\psi_0|_{T_0}^{l,l/2}) := |\psi^{(1)} - \psi^{(2)}|_{T_0}^{l,l/2}\mu_5.\end{aligned}$$

Hence,

$$\Big|\Big(A\psi^{(1)} - A\psi^{(2)}\Big)\Big|_T^{l,l/2} < \mu\Big|\psi^{(1)} - \psi^{(2)}\Big|_T^{l,l/2}$$

if T satisfies the condition

$$\mu := \max\{\mu_1, \mu_2, \mu_3, \mu_4, \mu_5\} < 1.$$

It is not difficult to see that if we set $T_0 = \min(\beta, \mu)$, then, for any $T \in (0, T_0)$, the operator A has the two properties of a contraction mapping operator, i.e., A takes the set $S(T)$ onto itself. Therefore, by the Banach theorem (see, for example, [22, pp. 87–97]), there is a unique fixed point of A in $S(T)$; i.e., there exists only one solution to (4.1). $\square$

5. Conclusion

In this paper, we have considered the problem of finding the functions $u(x,t)$ and $k(x',t)$ from the (1.1)–(1.3). First, the above problem has been reduced to an auxiliary problem. The equivalence of the auxiliary problem to Volterra-type integral equations has been shown. The existence and uniqueness of a solution to the problem have been obtained using the fixed point principle.

REFERENCES

1. Colombo F. An inverse problem for a parabolic integro-differential model in the theory of combustion. *Phys. D.*, 2007. Vol. 236, No. 2. P. 81–89. DOI: 10.1016/j.physd.2007.07.012
2. Durdiev D. K., Jumaev J. J., Atoev D. D. Inverse problem on determining two kernels in integro-differential equation of heat flow. *Ufa Math. J.*, 2023. Vol. 15, No. 2. P. 120–135.
3. Durdiev D. K., Nuriddinov J. Z. On investigation of the inverse problem for a parabolic integro-differential equation with a variable coefficient of thermal conductivity. *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2020. Vol. 30, No. 4. P. 572–584. DOI: 10.35634/vm200403
4. Durdiev D. K., Nuriddinov Zh. Z. Problem of determining a multidimensional kernel in one parabolic integro-differential equation. *J. Siberian Federal Univ. Math. Phys.*, 2021. Vol. 14, No. 1. P. 117–127. DOI: 10.17516/1997-1397-2020-14-1-117-127
5. Durdiev D. K., Rashidov A. S. Inverse problem of determining the kernel in an integro-differential equation of parabolic type. *Diff. Equ.*, 2014. Vol. 50, No. 1. P. 110–116. DOI: 10.1134/S0012266114010145
6. Durdiev D., Shishkina E., Sitnik S. The Explicit formula for solution of anomalous diffusion equation in the multi-dimensional space. *Lobachevskii J. Math.*, 2021. Vol. 42, No. 6. P. 1264–1273. DOI: 10.1134/S199508022106007X
7. Durdiev D. K., Zhumaev Zh. Zh. Problem of determining a multidimensional thermal memory in a heat conductivity equation. *Methods Funct. Anal. Topology*, 2019. Vol. 25, No. 3. P. 219–226.
8. Durdiev D. K., Zhumaev Zh. Zh. Problem of determining the thermal memory of a conducting medium. *Diff. Equ.*, 2020. Vol. 56, No. 6. P. 785–796. DOI: 10.1134/S0012266120060117
9. Durdiev D. K., Zhumaev Zh. Zh. Memory kernel reconstruction problems in the integro-differential equation of rigid heat conductor. *Math. Methods Appl. Sci.*, 2020. Vol. 45, No. 14. P. 8374–8388. DOI: 10.1002/mma.7133
10. Dyatlov G. V. Determination for the memory kernel from boundary measurements on a finite time interval. *J. Inverse Ill-Posed Probl.*, 2003. Vol. 11, No. 1. P. 59–66. DOI: 10.1515/156939403322004937
11. Grasselli M. An identification problem for a linear integrodifferential equation occurring in heat flow. *Math. Meth. Appl. Sci.*, 1992. Vol. 15, No. 3. P. 167–186. DOI: 10.1002/mma.1670150304
12. Gurtin M. E., Pipkin A. C. A general theory of heat conduction with finite wave speeds. *Arch. Rational Mech. Anal.*, 1968. Vol. 31. P. 113–126. DOI: 10.1007/BF00281373
13. Hazanee A., Lesnic D., Ismailov M. I., Kerimov N. B. Inverse time-dependent source problems for the heat equation with nonlocal boundary conditions. *Appl. Math. Comput.*, 2019. Vol. 346. P. 800–815. DOI: 10.1016/j.amc.2018.10.059
14. Huntul M. J., Lesnic D., Hussein M. S. Reconstruction of time-dependent coefficients from heat moments. *Appl. Math. Comput.*, 2017. Vol. 301. P. 233–253. DOI: 10.1016/j.amc.2016.12.028
15. Ivanchov M. I., Saldina N. V. Inverse problem for a parabolic equation with strong power degeneration. *Ukr. Math. J.*, 2006. Vol. 58, No. 11. P. 1685–1703. DOI: 10.1007/s11253-006-0162-x
16. Janno J., Lorenzi A. Recovering memory kernels in parabolic transmission problems. *J. Inverse Ill-Posed Probl.*, 2008. Vol. 16, No. 3. P. 239–265. DOI: 10.1515/JIIP.2008.015
17. Janno J., Wolfersdorf L. Inverse problems for identification of memory kernels in heat flow. *J. Inverse Ill-Posed Probl.*, 1996, Vol. 4, No. 1. P. 39–66. DOI: 10.1515/jiip.1996.4.1.39

18. Ladyženskaja O., Solonnikov V. A., Ural'ceva N. N. *Linear and Quasi-linear Equations of Parabolic Type.* Transl. Math. Monographs, vol. 23. Providence, RI.: American Mathematical Soc., 1968. 648 p.
19. Lorenzi A. Severely ill-posed linear parabolic integro-differential problems. *J. Inverse Ill-Posed Probl.*, 2013. Vol. 21, No. 2. P. 223–264. DOI: 10.1515/jip-2012-0099
20. Lorenzi A., Lorenzi L., Yamamoto M. Continuous dependence and uniqueness for lateral Cauchy problems for linear integro-differential parabolic equations. *J. Inverse Ill-Posed Probl.*, 2017. Vol. 25, No. 5. P. 617–631. DOI: 10.1515/jiip-2016-0066
21. Lorenzi A., Messina F. Unique continuation and continuous dependence results for a severely ill-posed integrodifferential parabolic problem with a memory term in the principal part of the differential operator. *J. Inverse Ill-Posed Probl.*, 2013. Vol. 21, No. 2. P. 281–309. DOI: 10.1515/jip-2012-0047
22. Kolmogorov A., Fomin S. *Elementy teorii funkcij i funkcional'nogo analiza* [Elements of Function Theory and Functional Analysis]. Moscow: Nauka, 1972. 544 p. (in Russian)

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 99–108
DOI: 10.15826/umj.2023.2.008

COMPUTING THE REACHABLE SET BOUNDARY FOR AN ABSTRACT CONTROL SYSTEM: REVISITED

Mikhail I. Gusev

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

gmi@imm.uran.ru

Abstract: A control system can be treated as a mapping that maps a control to a trajectory (output) of the system. From this point of view, the reachable set, which consists of the ends of all trajectories at a given time, can be considered an image of the set of admissible controls into the state space under a nonlinear mapping. The paper discusses some properties of such abstract reachable sets. The principal attention is paid to the description of the set boundary.

Keywords: Reachable set, Nonlinear mapping, Control system, Extremal problem, Maximum principle.

1. Introduction

The paper explores the issue of describing the boundary of the reachable set of a nonlinear control system. A reachable set consists of all state vectors that can be reached along trajectories generated by admissible controls. For a system with geometric (point-wise) constraints, it is known that control steering the trajectory to the boundary of the set satisfies Pontryagin's maximum principle [13, 16]. Many algorithms for computing reachable sets are established based on solving optimal control problems and (or) use of the maximum principle [2, 5, 12, 14, 17]. For systems with integral constraints, some properties of reachable sets and algorithms for their construction are given in [6, 7, 15].

For integral quadratic constraints, it was shown in [8, 10] that any admissible control leading to the reachable set boundary provides a local extremum in some optimal control problem. Therefore, this control satisfies the maximum principle. This result was generalized in [11] for several mixed integral constraints in which the integrands depend on both control and state variables. In [9] (see, also [1]), we proposed to consider the reachability problem in terms of nonlinear mappings of Banach spaces. With this approach, the reachable set is treated as the image of the set of all admissible controls under the action of a nonlinear mapping. In the present paper, we extend the results of [9] to a broader class of abstract control systems. These systems are determined by differentiable maps of Banach spaces with different types of constraints on controls. The paper weakens the conditions of [9], which makes it possible to consider the problem with constraints specified by nonsmooth functionals. The use of nonsmooth analysis constructions allowed us to consider problems with multiple constraints within the framework of a unified scheme.

2. Single constraint control systems

Let us consider the system

$$\dot{x}(t) = f_1(t, x(t)) + f_2(t, x(t))u(t), \quad x(t_0) = x_0, \quad u(\cdot) \in U, \tag{2.1}$$

on a time interval $[t_0, t_1]$. Here, $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^r$, and U is a given set in the space $\mathbb{L}_p$, $p > 1$.

Functions $f_2 : \mathbb{R}^{n+1} \to \mathbb{R}^{n\times r}$ are considered to have continuous Fréchet derivatives in x and satisfying the conditions:

$$\|f_1(t,x)\| \leq l_1(t)(1+\|x\|), \quad \|f_2(t,x)\|_{n\times r} \leq l_2(t), \quad t_0 \leq t \leq t_1, \quad x \in \mathbb{R}^n.$$

Here, $l_1(\cdot) \in \mathbb{L}_1$ and $l_2(\cdot) \in \mathbb{L}_2$, where $\mathbb{L}_1$ and $\mathbb{L}_2$ denote the spaces of summable and square summable functions, respectively.

For any $u(\cdot) \in \mathbb{L}_1$, there is a unique absolutely continuous solution $x(t, u(\cdot))$ to system (2.1) such that $x(t_0) = x_0$.

A reachable set $G(t_1)$ of system (2.1) at time t_1 under the constraint $u(\cdot) \in U \subset \mathbb{L}_1$ is defined as follows:

$$G(t_1) = \big\{y \in \mathbb{R}^n : y = x(t_1, u(\cdot)),\ u(\cdot) \in U\big\}.$$

This definition of a reachable set fits into the framework of the following abstract construction. Let X and Y be real Banach spaces, and let $U \subset X$ be a given set. We will call a map $F : U \to Y$ an abstract control system. Here, $u \in U$ is called a control and the set U is called a constraint. The reachable set G of this system is

$$G = \big\{y \in Y : y = F(u),\ u \in U\big\}.$$

Thus, $G = F(U)$ is an image of the set U under the mapping F.

Further, we set

$$U = \big\{u \in X : \varphi(u) \leq \mu\big\},$$

so U is a level set of a continuous function $\varphi : X \to \mathbb{R}$; $\mu > 0$ is a given number. In control problems for system (2.1), one can take $X = \mathbb{L}_p$, $p > 1$, including $p = \infty$, as the space X and $Y = \mathbb{R}^n$.

The mapping F in this case is determined as

$$F(u) = F(u(\cdot)) = x(t_1, u(\cdot)). \tag{2.2}$$

With standard requirements on system (2.1) (see, for example, [10]), $F(u(\cdot))$ is a single-valued mapping having a continuous Fréchet derivative $F'(u(\cdot)) : \mathbb{L}_2 \to \mathbb{R}^n$:

$$F'_u(u(\cdot))\Delta u(\cdot) = \Delta x(t_1).$$

Here, $\Delta x(t)$ is a solution to system (2.1) linearized around $(x(t, u(\cdot)), u(t))$,

$$\begin{gathered} \dot{\Delta x}(t) = A(t)\Delta x(t) + B(t)\Delta u(t), \quad \Delta x(t_0) = 0, \\ A(t) = \frac{\partial f_1}{\partial x}(t, x(t)) + \frac{\partial}{\partial x}\big[f_2(t, x(t))u(t)\big], \quad B(t) = f_2(t, x(t)), \end{gathered} \tag{2.3}$$

corresponding to the control $\Delta u(t)$. If system (2.3) is controllable on $[t_0, t_1]$, then $\operatorname{Im} F'(u(\cdot)) = \mathbb{R}^n$.

Let us consider the geometric constraints on controls that are standard for control theory:

$$u(t) \in \Omega, \quad \text{a.e.} \quad t \in [t_0, t_1].$$

In many cases, the set Ω can be represented as

$$\Omega = \{v \in \mathbb{R}^r : \|Qv\| \leq 1\},$$

where Q is a matrix and $\|\cdot\|$ is some norm in $\mathbb{R}^m$. It is clear that we can take here $X = \mathbb{L}_\infty$ and

$$\varphi(u(\cdot)) = \operatorname*{ess\,sup}_{t_0 \leq t \leq t_1} \|Qu(t)\|.$$

Such a functional is obviously continuous in the space $\mathbb{L}_\infty$.

Another example of control constraints is an integral constraint. In this case, $X = \mathbb{L}_p$, $p > 1$, and

$$\varphi(u(\cdot)) = \int_{t_0}^{t_1} \|u(t)\|^p \, dt.$$

We call the joint constraints on both control and state variables of the form

$$\varphi(u(\cdot)) := \int_{t_0}^{t_1} \big(Q(t, x(t)) + u^\top(t) R(t, x(t)) u(t)\big) \, dt \le \mu, \quad u(\cdot) \in \mathbb{L}_2,$$

the isoperimetric constraints.

Let $B_X(x, r)$ and $B_Y(y, r)$ be the balls of radius r centered at $x \in X$ and $y \in Y$, respectively. Further analysis is based on a well-known Lyusternik's theorem.

Theorem 1 [4, Theorem 2]. *Let a mapping F from a Banach space X to a Banach space Y be continuously Fréchet differentiable at a point $\hat{u}$ and such that $\operatorname{Im} F'(\hat{u}) = Y$. Then there are a neighborhood V of the point $\hat{u}$ and a number $s > 0$ such that, for any $B_X(u, r) \subset V$,*

$$B_Y(F(u), sr) \subset F(B_X(u, r)).$$

The condition $\operatorname{Im} F'(\hat{u}) = Y$ is called the Lyusternik (regularity) condition. If this condition is met, F is said to be regular at the point $\hat{u}$.

Using this theorem we get the following statement.

Theorem 2. *Let W be some neighborhood of the set U, let $F : W \to Y$ be a mapping continuously Fréchet differentiable at a point $\hat{u} \in U$, and let $\operatorname{Im} F'(\hat{u}) = Y$. To $\hat{x} = F(\hat{u}) \in \partial G$, it is necessary that $\hat{u}$ be a local extremum in the problem*

$$\varphi(u) \to \min, \quad F(u) = \hat{x}, \tag{2.4}$$

and $\varphi(\hat{u}) = \mu$.

P r o o f. The proof is by contradiction. Assume that $\varphi(\hat{u}) < \mu$. Since $\varphi(u)$ is continuous at the point $\hat{u}$, there is a neighborhood V_1 of $\hat{u}$ such that $\varphi(u) < \mu \ \forall u \in V_1$. Let us choose a neighborhood V and a number s whose existence follows from Theorem 1. Then, for any ball $B_X(\hat{u}, r) \in V \bigcap V_1$, we have

$$B_X(\hat{u}, r) \subset U,$$
$$B_Y(\hat{x}, sr) = B_Y(F(\hat{u}), sr) \subset F(B_X(\hat{u}, r)) \subset F(U) = G,$$

which contradicts the condition $\hat{x} \in \partial G$. Hence, $\varphi(\hat{u}) = \mu$.

Let us again choose V and s from Theorem 1. Assume that $\hat{u}$ is not a local minimum in (2.4). Then there is $\bar{u} \in V$ such that $F(\bar{u}) = \hat{x}$ and $\varphi(\bar{u}) < \varphi(\hat{u}) = \mu$. Let us choose $r > 0$ such that $B_X(\bar{u}, r) \subset V$. Then, by Theorem 1,

$$B_Y(\hat{x}, sr) = B_Y(F(\bar{u}), sr) \subset F(B_X(\bar{u}, r)) \subset F(U) = G$$

contrary to the condition $\hat{x} \in \partial G$. This completes the proof. □

Let us write down the necessary extremum condition for problem (2.4), assuming that $\varphi(u)$ is continuously differentiable at $\hat{u}$. Since the constraint $F(u) = \hat{x}$ is regular at the point $\hat{u}$, there is a Lagrange multiplier $y^* \in Y^*$ such that

$$\varphi'(\hat{u}) + F'^*(\hat{u})y^* = 0. \tag{2.5}$$

Here, $F'^*(\hat{u})$ denotes the operator conjugate to the continuous linear operator $F'(\hat{u})$.

If $\varphi'(\hat{u}) \neq 0$, then equality (2.5) implies that $y^* \neq 0$. If we divide both sides of equality (2.5) by $\|y^*\|$, then it takes the form

$$F'^*(\hat{u})y^* + \lambda\varphi'(\hat{u}) = 0, \tag{2.6}$$

where $\|y^*\| = 1$ and $\lambda > 0$. Since $\varphi(\hat{u}) - \mu = 0$, we also have the equality

$$\lambda(\varphi(\hat{u}) - \mu) = 0. \tag{2.7}$$

It is easy to see that relations (2.6) and (2.7) also give the necessary optimality conditions for the problem

$$\langle y^*, F(u)\rangle \to \min, \quad \varphi(u) \leq \mu, \tag{2.8}$$

where $\langle \cdot, \cdot \rangle$ denotes a bilinear form establishing the duality of the spaces Y and Y^*. Here, equality (2.6) means that the derivative of the Lagrange function

$$L(u, \lambda) = \langle y^*, F(u)\rangle + \lambda(\varphi(u) - \mu)$$

in u is equal to zero, and equality (2.7) is a complementary slackness condition. Thus, the following statement is true.

Theorem 3. *Assume that $F(\hat{u}) = \hat{x} \in \partial G$, $u \in U$, $F(u)$ is regular, and $\varphi(u)$ is continuously differential at the point $\hat{u}$ and $\varphi(\hat{u}) \neq 0$. Then, there is $y^* \in Y^*$, $\|y^*\| = 1$, such that $\hat{u}$ satisfies the necessary extremum conditions* (2.6) *and* (2.7) *in problem* (2.8).

As it is easy to see, problem (2.8) can be rewritten in the equivalent form

$$\langle z^*, y\rangle \to \max, \quad y \in G.$$

where $z^* = -y^*$. The latter is the problem of calculating the support function of G. Recall that a support function $\psi_G(z^*)$ is defined on Y^* by the equality

$$\psi_G(z^*) = \sup_{y \in G} \langle z^*, y\rangle.$$

The point at which the supremum is reached is called the support point. Since the reachable set G in the nonlinear case is not necessarily convex, the boundary point $\hat{x}$ is not necessarily a support point. But it meets the necessary optimality conditions as if it would be a support point.

Next, we will consider the case when φ is not continuously differentiable but is Lipschitz continuous at the point $\hat{u}$. For simplicity, we will assume also that $Y = \mathbb{R}^n$.

Denote by $\partial_C f(u)$ the Clarke subdifferential of a function f at a point u. If f is Lipschitz continuous in some neighborhood of u, then $\partial_C f(u) \neq \varnothing$ is a convex weakly* compact set [3].

Let L be a Lagrange function

$$L(u, \lambda, y^*) = \lambda\varphi(u) + \langle y^*, F(u) - \hat{x}\rangle,$$

where $\lambda \geq 0$ and $y^* \in Y^* = \mathbb{R}^n$ are Lagrange multipliers.

Assume that $\hat{u}$ is a local solution to problem (2.4) and $\varphi(u)$ is Lipschitz continuous at the point $\hat{u}$. Then, there exist $\lambda \geq 0$ and $y^* \in \mathbb{R}^n$, $\lambda + \|y^*\| \neq 0$, such that

$$0 \in \partial_C L(\hat{u}, \lambda, y^*) = \lambda \partial_C \varphi(\hat{u}) + F'^*(\hat{u}) y^*, \tag{2.9}$$

where $\partial_C L$ is taken with respect to u (see, for example, [3, Theorem 6.1.1]). Let us show that $\lambda > 0$. Indeed, if $\lambda = 0$, then $\|y^*\| \neq 0$ and $F'^*(\hat{u})y^* = 0$. This contradicts the regularity of F at the point $\hat{u}$.

Without loss of generality, we set $\lambda = 1$. Suppose that $0 \notin \partial_C \varphi(\hat{u})$. Then $F'^*(\hat{u})y^* \neq 0$ and condition (2.9) takes the form

$$-F'^*(\hat{u})y^* \in \partial_C \varphi(\hat{u}). \tag{2.10}$$

Let us show that this inclusion is a necessary extremum condition in problem (2.8). Let

$$L(u, \alpha, \beta) = \alpha \langle y^*, F(u) \rangle + \beta(\varphi(u) - \mu)$$

be the Lagrange function for problem (2.8). If $\hat{u}$ is a local minimum point in problem (2.8), then there are $\alpha \geq 0$ and $\beta \geq 0$, $\alpha + \beta \neq 0$, such that

$$0 \in \partial_C L(\hat{u}, \alpha, \beta). \tag{2.11}$$

Note that if $0 \notin \partial_C \varphi(\hat{u})$, then $\alpha > 0$ and $\beta > 0$. Indeed, if $\alpha = 0$, then $\beta > 0$ and $0 \in \partial_C \varphi(\hat{u})$. If $\beta = 0$, then $\alpha F'^*(\hat{u})y^* = 0$ and $\alpha > 0$, which is impossible due to the regularity condition. Divide both sides of inclusion (2.11) by β and take $\alpha y^*/\beta$ as a new vector y^*. Then inclusion (2.11) takes the form (2.10).

As a result, we get the following statement.

Theorem 4. *Assume that $F(\hat{u}) = \hat{x} \in \partial G$, $\hat{u} \in U$, $F(u)$ is regular, and $\varphi(u)$ is Lipschitz continuous at the point $\hat{u}$ and $0 \notin \partial_C \varphi(\hat{u})$. Then there is $y^* \in Y^*$, $\|y^*\| = 1$, such that $\hat{u}$ satisfies the necessary extremum condition* (2.10) *in problem* (2.8).

Remark 1. If $\varphi(u)$ is convex, then $\partial_C \varphi(u) = \partial \varphi(u)$ is a subdifferential of a convex function. The condition $0 \notin \partial_C \varphi(\hat{u})$ in this case is equivalent to Slater's condition: there is $\bar{u}$ such that $\varphi(\bar{u}) < \varphi(\hat{u})$.

Remark 2. If a mapping F is defined by formula (2.2) and $\varphi(u(\cdot))$ is an integral quadratic in u functional, then Theorem 2 implies the necessary extremum conditions [10] in the form of Pontryagin's maximum principle.

Note that, under integral quadratic constraints, the relations of the maximum principle follow directly from the extremum conditions (2.10). Below we present its proof. Assume that $X = \mathbb{L}_2$, $Y = \mathbb{R}^n$, the mapping F is defined by formula (2.2), and $\varphi(u(\cdot)) = 1/2 \langle u(\cdot), u(\cdot) \rangle$ is an integral quadratic functional. In this case, $\partial \varphi(u(\cdot)) = \{\varphi'(u(\cdot))\} = \{u(\cdot)\}$ and the equality $\varphi(\hat{u}(\cdot)) = \mu$ implies that $\varphi'(\hat{u}(\cdot)) \neq 0$. Therefore, (2.10) takes the following equivalent form:

$$F'^*(\hat{u})z^* = \hat{u}, \quad z^* = -y^*, \quad z^* \neq 0.$$

Recall that $F'(u) = F'(u(\cdot))$ is defined by the equality $F'(u(\cdot))\Delta u(\cdot) = \Delta x(t_1)$, where $x(t)$ is the solution of (2.3). Let us represent this solution in the integral form

$$\Delta x(t_1) = \int_{t_0}^{t_1} X(t_1, \tau) B(\tau) \Delta u(\tau) \, d\tau,$$

where $X(t,\tau)$ is the Cauchy matrix. For any $z^* \in \mathbb{R}^n$, we have

$$\big(z^*, F'(u(\cdot))\Delta u(\cdot)\big) = \big\langle F'^*(u(\cdot))z^*, \Delta u(\cdot)\big\rangle = z^{*\top}\int_{t_0}^{t_1} X(t_1,\tau)B(\tau)\Delta u(\tau)d\tau$$
$$= \int_{t_0}^{t_1} p^\top(\tau)B(\tau)\Delta u(\tau)d\tau,$$

where $p(\tau) = X^\top(t_1,\tau)z^*$ satisfies the adjoint equation

$$\dot{p}(t) = -A^\top(t)p(t), \quad p(t_1) = z^*.$$

Thus, we have

$$F'^*(u(\cdot))z^* = B^\top(\cdot)\,p(\cdot) = \hat{u}(\cdot),$$

which implies that

$$\hat{u}(t) = B^\top(t)p(t), \quad t_0 \le t \le t_1.$$

Finally, we obtain a system of relations of the maximum principle for the boundary control $\hat{u}(t)$ (see [10])

$$\dot{x}(t) = f_1(t,x(t)) + f_2(t,x(t))B(t)p(t), \quad x(t_0) = x_0, \tag{2.12}$$
$$\dot{p}(t) = -A^\top(t)p(t), \quad p(t) \ne 0, \quad \hat{u}(t) = B(t)p(t), \tag{2.13}$$
$$A(t) = \frac{\partial f_1}{\partial x}(t,x(t)) + \frac{\partial}{\partial x}[f_2(t,x(t))\hat{u}(t)], \quad B(t) = f_2(t,x(t)).$$

Now suppose that the constraints have the form

$$\gamma(u(t)) \le \mu, \quad \text{a.e. in } [t_0,t_1],$$

where $\gamma(u)$ is a convex function in $\mathbb{R}^r$ (for example, a norm in $\mathbb{R}^r$). In this case, we can take $X = \mathbb{L}_\infty$ and

$$\varphi(u(\cdot)) = \operatorname*{ess\,sup}_{t_0 \le t \le t_1} \gamma(u(t)).$$

Such a functional is obviously convex and continuous in the space X. Assume that there is $\bar{u} \in \mathbb{R}^r$ such that $\gamma(\bar{u}) < \mu$. As before, we believe that $Y = \mathbb{R}^n$. Since $\varphi(u(\cdot))$ is convex, we can substitute $\partial_C\varphi(\hat{u}(\cdot))$ by a subdifferential of the convex function $\partial\varphi(\hat{u}(\cdot))$.

If $F(\hat{u}(\cdot)) \in \partial G$, then $\varphi(\hat{u}(\cdot)) = \mu$ and hence $0 \notin \partial\varphi(\hat{u}(\cdot))$. Thus,

$$F'^*(\hat{u}(\cdot))z^* \in \partial\varphi(\hat{u}(\cdot))$$

for some $z^* \in \mathbb{R}^n$, $z^* \ne 0$. Here, the point $F'^*(\hat{u}(\cdot))z^*$ belongs to the space $\mathbb{L}_\infty^*$. Similar to the previous case, it can be proven that $F'^*(\hat{u}(\cdot))z^* = B^\top(\cdot)p(\cdot)$, where $p(t) \ne 0$ is a solution to the adjoint system.

From the properties of $\partial\varphi(\hat{u}(\cdot))$, we get

$$\varphi(u(\cdot)) - \varphi(\hat{u}(\cdot)) \ge \langle F'^*(\hat{u}(\cdot))z^*, u(\cdot) - \hat{u}(\cdot)\rangle$$

for every $u(\cdot) \in \mathbb{L}_\infty$. From this inequality, for every $u(\cdot)$ such that $\varphi(u(\cdot)) \le \mu$, we have

$$0 \ge \int_{t_0}^{t_1} p^\top(\tau)B(\tau)(u(\tau) - \hat{u}(\tau))d\tau. \tag{2.14}$$

Choose a point $\tau \in (t_0, t_1)$ and a vector $v \in \mathbb{R}^r$ such that $\gamma(v) \le \mu$, and sufficiently small $\varepsilon > 0$. Let

$$u(t) = \begin{cases} \hat{u}(t), & t \notin [\tau, \tau+\varepsilon], \\ v, & t \in [\tau, \tau+\varepsilon]. \end{cases}$$

Then, (2.14) implies the inequality

$$\frac{1}{\varepsilon}\int_{\tau}^{\tau+\varepsilon} p^\top(t)B(t)\hat{u}(t)dt \ge \frac{1}{\varepsilon}\int_{\tau}^{\tau+\varepsilon} p^\top(t)B(t)vdt.$$

Passing here to the limit, we get

$$p^\top(\tau)B(\tau)\hat{u}(\tau) \ge p^\top(\tau)B(\tau)v$$

for almost every $\tau \in [t_0, t_1]$ and every v such that $\gamma(v) \le \mu$. So, we have

$$p^\top(\tau)B(\tau)\hat{u}(\tau) = \max_{\gamma(v)\le\mu} p^\top(\tau)B(\tau)v,$$
$$\dot{p}(\tau) = -A(\tau)p(\tau), \quad p(\cdot) \ne 0.$$

Introducing the Hamiltonian

$$H(t, x, p, u) = p^\top(f_1(t, x) + f_2(t, x)u),$$

we can write the last relations in the standard form of the maximum principle:

$$H(\tau, x(\tau), p(\tau), \hat{u}(\tau)) = \max_{\gamma(v)\le\mu} H(\tau, x(\tau), p(\tau), v), \quad \text{a.e.} \quad \tau \in [t_0, t_1], \tag{2.15}$$
$$\dot{p}(\tau) = -A(\tau)p(\tau) = -\frac{\partial H}{\partial x}(\tau, x(\tau), p(\tau), \hat{u}(\tau)), \quad \tau \in [t_0, t_1]. \tag{2.16}$$

3. Multiple constraints on the control

In this section, we consider constraints specified by the inequalities

$$\varphi_i(u) \le \mu_i, \quad i = 1, \dots, k. \tag{3.1}$$

Here, $\varphi_i : X \to \mathbb{R}$ are functionals and μ_i, $i = 1, \dots, k$, are given positive numbers.

One can assume without loss of generality that $\mu_i = 1$, $i = 1, \dots, k$. Then (3.1) can be replaced by the single constraint $\varphi(u) \le 1$ by setting

$$\varphi(u) = m(\varphi_1(u), \dots, \varphi_k(u)), \quad m(x) = m(x_1, \dots, x_k) = \max_{1\le i\le k} x_i.$$

Since $m(x)$ is a continuous function, the functional $\varphi(u)$ is obviously continuous at a point of continuity of all functionals $\varphi_i(u)$. Therefore, for describing the reachable set boundary, we can use Theorem 2, which leads to the following statement.

Corollary 1. *Let W be a neighborhood of the set U, and let $F : W \to Y$ be a mapping continuously Fréchet differentiable at the point $\hat{u} \in U$ such that $\operatorname{Im} F'(\hat{u}) = Y$. Assume that*

$$G = \{F(u) : \varphi_i(u) \le 1,\ i = 1, \dots, k\},$$

where $\varphi_i(u)$ are continuous at the point $\hat{u}$. To $\hat{x} = F(\hat{u}) \in \partial G$, it is necessary that $\hat{u}$ be a local extremum in the problem

$$\varphi(u) = m(\varphi_1(u), \dots, \varphi_k(u)) \to \min, \quad F(u) = \hat{x},$$

and $\varphi(\hat{u}) = 1$.

The derivation of extremum conditions in this problem is more complicated than before because the function $m(x)$ is not differentiable. However, the superposition $\varphi(u) = m(\varphi_1(u), \dots, \varphi_k(u))$ is locally Lipschitz at the point $\hat{u}$ if such are the functions $\varphi_i(u)$. Moreover, if each of the functions $\varphi_i(u)$ is either convex or continuously differentiable at the point $\hat{u}$, then

$$\partial_C \varphi(\hat{u}) = \operatorname{co} \bigcup_{i \in I(\hat{u})} \partial_C \varphi_i(\hat{u}), \tag{3.2}$$

where $I(\hat{u}) = \{i : \varphi_i(\hat{u}) = \varphi(\hat{u})\}$ and $\operatorname{co} A$ denotes a convex hull of A [3].

Let the conditions of Corollary 1 be satisfied. Let initially all functionals φ_i be continuously differentiable at $\hat{u}$. Then $\partial_C \varphi_i(\hat{u}) = \{\varphi_i'(\hat{u})\}$ and, taking into account (3.2), we get

$$\partial_C \varphi_i(\hat{u}) = \Big\{ \sum_{i \in I(\hat{u})} \alpha_i \varphi_i'(\hat{u}) : \sum_{i \in I(\hat{u})} \alpha_i = 1,\ \alpha_i \geq 0 \Big\}$$
$$= \Big\{ \sum_{1 \leq i \leq k} \alpha_i \varphi_i'(\hat{u}) : \sum_{1 \leq i \leq k} \alpha_i = 1,\ \alpha_i \geq 0,\ \alpha_i(\varphi_i(\hat{u}) - 1) = 0,\ i = 1, \dots, k \Big\}.$$

Here, the condition $0 \notin \partial_C \varphi_i(\hat{u})$ takes the form

$$\sum_{1 \leq i \leq k} \alpha_i = 1, \quad \alpha_i \geq 0, \quad \alpha_i(\varphi_i(\hat{u}) - 1) = 0, \quad i = 1, \dots, k \quad \Rightarrow \quad \sum_{1 \leq i \leq k} \alpha_i \varphi_i^{'}(\hat{u}) = 0.$$

In particular, it is satisfied if the vectors $\varphi_i(\hat{u})$ form a positive linear independent set. If this condition is met, we can write down the necessary condition for the inclusion $F(\hat{u}) \in \partial G$ as follows:

$$F'^*(\hat{u}) z^* = \sum_{1 \leq i \leq k} \alpha_i \varphi_i'(\hat{u}), \quad \sum_{1 \leq i \leq k} \alpha_i = 1, \quad \alpha_i \geq 0, \quad \alpha_i(\varphi_i(\hat{u}) - 1) = 0, \quad i = 1, \dots, k.$$

Using the previous scheme, we can also write this condition in the form of Pontryagin's maximum principle [16] (see also [11]).

Let us next consider a system with double control constraints. We will assume that one of the constraints is specified by a convex differentiable functional $\varphi_1(u)$ and the second by a convex functional $\varphi_2(u)$. An example of such a problem is system (2.1) with integral quadratic and geometric constraints. If $\varphi_2(\hat{u}) < \varphi_1(\hat{u})$, then $\partial_C \varphi(\hat{u}) = \{\varphi_1'(\hat{u})\}$; if $\varphi_1(\hat{u}) < \varphi_2(\hat{u})$, then $\partial_C \varphi(\hat{u}) = \{\partial \varphi_2(\hat{u})\}$; and, finally, if $\varphi_1(\hat{u}) = \varphi_2(\hat{u})$, then $\partial_C \varphi(\hat{u}) = \operatorname{co}(\{\varphi_1'(\hat{u})\} \cup \partial \varphi_2(\hat{u}))$.

Lemma 1. *Let $a \in X$, and let $B \subset X$ be a convex set. Then*

$$\operatorname{co}(\{a\} \cup B) = C := \bigcup_{0 \leq \lambda \leq 1} (\lambda a + (1 - \lambda) B).$$

P r o o f. Obviously, $C \subset \operatorname{co}(\{a\} \cup B)$. To prove the lemma, it suffices to prove the convexity of C. Let

$$c_1 = \lambda_1 a + (1 - \lambda_1) b_1, \quad c_2 = \lambda_2 a + (1 - \lambda_2) b_2, \quad b_1, b_2 \in B.$$

Let us choose $\alpha, \beta \geq 0$, $\alpha + \beta = 1$, and show that

$$c_3 = \alpha c_1 + \beta c_2 \in \lambda_3 a + (1 - \lambda_3) B$$

for some $\lambda_3 \in [0, 1]$. To this end, we try to find numbers $\alpha_1, \beta_1 \geq 0$, $\alpha_1 + \beta_1 = 1$, such that

$$\alpha c_1 + \beta c_2 = \alpha(\lambda_1 a + (1 - \lambda_1) b_1) + \beta(\lambda_2 a + (1 - \lambda_2) b_2) = \lambda_3 a + (1 - \lambda_3)(\alpha_1 b_1 + \beta_1 b_2).$$

Equating the coefficients at the vectors a, b_1, and b_2 on both sides of the equality, we obtain

$$\lambda_3 = \alpha\lambda_1 + \beta\lambda_2, \quad \alpha(1-\lambda_1) = \alpha_1(1-\lambda_3), \quad \beta(1-\lambda_2) = \beta_1(1-\lambda_3).$$

This implies the inequality $0 \le \lambda_3 \le 1$. For $0 \le \lambda_3 < 1$, we have

$$\alpha_1 = \frac{\alpha(1-\lambda_1)}{1-\lambda_3}, \quad \beta_1 = \frac{\beta(1-\lambda_2)}{1-\lambda_3};$$

so, $\alpha_1, \beta_1 \ge 0$ and $\alpha_1 + \beta_1 = 1$. If $\lambda_3 = 1$, then either $\alpha\lambda_1 = 1$ or $\beta\lambda_2 = 1$. In both of these cases, we get $c_3 = a$. This completes the proof. □

Let us further assume that Slater's condition is satisfied: there exists $\bar{u}$ such that $\varphi_i(\bar{u}) < 1$, $i = 1, 2$. Then the condition $0 \notin \partial_C\varphi(\hat{u})$ is satisfied. Indeed, suppose on the contrary that $0 \in \partial_C\varphi(\hat{u})$. Then, it follows from Lemma 1 that there is $\lambda \in [0, 1]$ such that

$$0 \in \lambda\varphi_1'(\hat{u}) + (1-\lambda)\partial\varphi_2(\hat{u}) = \partial(\lambda\varphi_1 + (1-\lambda)\varphi_2)(\hat{u}).$$

For the convex function $\lambda\varphi_1 + (1-\lambda)\varphi_2$, the last condition is necessary and sufficient for the minimum at $\hat{u}$. Thus,

$$(\lambda\varphi_1 + (1-\lambda)\varphi_2)(\hat{u}) \le (\lambda\varphi_1 + (1-\lambda)\varphi_2)(\bar{u}),$$

which contradicts Slater's condition.

Let further $X = \mathbb{L}_\infty$ and

$$\varphi_1(u(\cdot)) = c/2\langle u(\cdot), u(\cdot)\rangle = c/2\int_{t_0}^{t_1} u^\top(t)u(t)dt, \quad \varphi_2(u(\cdot)) = \operatorname*{ess\,sup}_{t_0 \le t \le t_1} \gamma(u(t)). \tag{3.3}$$

The constant $c > 0$ is chosen here such that to write down the constraints in the form $\varphi_i(u(\cdot)) \le 1$, $i = 1, 2$. Since $\varphi_1'(u(\cdot)) = cu(\cdot)$, the optimality conditions $F'^*(\hat{u}(\cdot))z^* \in \partial\varphi(\hat{u}(\cdot))$ take the form

$$F'^*(\hat{u}(\cdot))z^* - \lambda c\hat{u}(\cdot) \in (1-\lambda)\partial\varphi_2(\hat{u}(\cdot))$$

for some $\lambda \in [0, 1]$.

For $\lambda = 0$, we get a maximum principle of the form (2.15), (2.16).

For $\lambda = 1$, we get (2.12), (2.13).

Finally, for $0 < \lambda < 1$, we get

$$F'^*(\hat{u}(\cdot))w^* - \sigma c\hat{u}(\cdot) \in \partial\varphi_2(\hat{u}(\cdot)),$$

where $w^* = z^*/(1-\lambda)$ and $\sigma = \lambda/(1-\lambda)$. Introducing the Hamiltonian

$$H(t, x, p, \sigma, u) = -\sigma cu + p^\top(f_1(t, x) + f_2(t, x)u),$$

we can write these relations in the form of maximum principle:

$$H(\tau, x(\tau), p(\tau), \sigma, \hat{u}(\tau)) = \max_{\gamma(v)\le\mu} H(\tau, x(\tau), p(\tau), \sigma, v), \quad \text{a.e.} \quad \tau \in [t_0, t_1],$$

$$\dot{p}(\tau) = -A(\tau)p(\tau) = -\frac{\partial H}{\partial x}(\tau, x(\tau), p(\tau), \sigma, \hat{u}(\tau)), \quad \tau \in [t_0, t_1].$$

Thus, we arrive at the following statement.

Corollary 2. *Let functionals $\varphi_i(u(\cdot)) : \mathbb{L}_\infty \to \mathbb{R}$, $i = 1, 2$, be given by equalities (3.3), and let $F(u(\cdot)) = x(t_1)$, where $x(t)$ is a solution to system (2.1). Let*

$$G = \{F(u(\cdot)) : \varphi_i(u(\cdot)) \le 1, \ i = 1, 2\}.$$

If $F(\hat{u}(\cdot)) \in \partial G$ and system (2.1) linearized around $\hat{u}(\cdot)$ is controllable, then there exist a function $p(\cdot) \ne 0$ and a number $\sigma \ge 0$ such that the relations of maximum principle are satisfied.

4. Conclusion

The paper proposes a unified scheme for studying extremal properties of the reachable set boundary. Within the framework of this approach, the reachable set is treated as the image of the set of admissible controls under a nonlinear mapping of a Banach space. The proposed scheme is based on the results of nonlinear and nonsmooth analysis and is equally applicable to systems with integral and geometric control constraints, including multiple constraints.

REFERENCES

1. Ananyev B. I., Gusev M. I., Filippova T. F. *Upravlenie i ocenivanie sostoyanij dinamicheskikh sistem s neopredelennost'yu* [Control and Estimation of Dynamical Systems States with Uncertainty]. Novosibirsk: Izdatel'stvo SO RAN, 2018. 193 p. (in Russian)
2. Baier R., Gerdts M., Xausa I. Approximation of reachable sets using optimal control algorithms. *Numer. Algebra Control Optim.*, 2013. Vol. 3, No. 3. P. 519–548. DOI: 10.3934/naco.2013.3.519
3. Clarke F. H. *Optimization and Nonsmooth Analysis.* New York: J. Willey and Sons Inc., 1983. 308 p.
4. Dmitruk A. V., Milyutin A. A., Osmolovskii N. P. Lyusternik's theorem and the theory of extrema. *Russian Math. Surveys*, 1980. Vol. 35, No. 6. P. 11–51. DOI: 10.1070/RM1980v035n06ABEH001973
5. Gornov A. Yu., Finkel'shtein E. A. Algorithm for piecewise-linear approximation of the reachable set boundary. *Autom. Remote Control*, 2015. Vol. 76, No. 3. P. 385–393. DOI: 10.1134/S0005117915030030
6. Guseinov Kh. G. Approximation of the attainable sets of the nonlinear control systems with integral constraint on controls. *Nonlinear Anal. Theory, Methods Appl.*, 2009. Vol. 71, No. 1–2. P. 622–645. DOI: 10.1016/j.na.2008.10.097
7. Guseinov K.G., Ozer O., Akyar E., Ushakov V.N. The approximation of reachable sets of control systems with integral constraint on controls. *Nonlinear Differ. Equ. Appl.*, 2007. Vol. 14, No. 1–2. P. 57–73. DOI: 10.1007/s00030-006-4036-6
8. Gusev M. I. On reachability analysis of nonlinear systems with joint integral constraints. In: *Lecture Notes in Comput. Sci., vol. 10665: Large-Scale Scientific Computing (LSSC 2017). Lirkov I., Margenov S. (eds.).* Cham: Springer, 2018. P. 219–227. DOI: 10.1007/978-3-319-73441-5_23
9. Gusev M. I. Computing the reachable set boundary for an abstract control problem. *AIP Conf. Proc.*, 2018. Vol. 2025, No. 1. Art. no. 040009. DOI: 10.1063/1.5064893
10. Gusev M. I., Zykov I. V. On extremal properties of the boundary points of reachable sets for control systems with integral constraints. *Proc. Steklov Inst. Math.*, 2018. Vol. 300, Suppl. 1. P. 114–125. DOI: 10.1134/S0081543818020116
11. Gusev M. I., Zykov I. V. On the geometry of reachable sets for control systems with isoperimetric constraints. *Proc. Steklov Inst. Math.*, 2019. Vol. 304, Suppl. 1. P. S76–S87. DOI: 10.1134/S0081543819020093
12. Kurzhanski A. B., Varaiya P. *Dynamics and Control of Trajectory Tubes. Theory and Computation.* Systems Control Found. Appl., vol. 85. Basel: Birkhäuser, 2014. 445 p. DOI: 10.1007/978-3-319-10277-1
13. Lee E. B., Marcus L. *Foundations of Optimal Control Theory.* New York: J. Willey and Sons Inc., 1967. 576 p.
14. Patsko V. S., Pyatko S. G., Fedotov A. A. Three-dimensional reachability set for a nonlinear control system. *J. Comput. Syst. Sci. Int.,* 2003. Vol. 42. No. 3. P. 320–328.
15. Polyak B. T. Convexity of the reachable set of nonlinear systems under L_2 bounded controls. *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.*, 2004. Vol. 11. P. 255–267.
16. Pontryagin L. S., Boltyanskii V. G., Gamkrelidze R. V., Mishechenko E. F. *The Mathematical Theory of Optimal Processes.* New York/London: J. Willey and Sons Inc., 1962. 360 p.
17. Vdovin S. A., Taras'ev A. M., Ushakov V. N. Construction of an attainability set for the Brockett integrator. *J. Appl. Math. Mech.,* 2004. Vol. 68, No. 5. P. 631–646. DOI: 10.1016/j.jappmathmech.2004.09.001

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 109–120

DOI: 10.15826/umj.2023.2.009

A NEW CHARACTERIZATION OF SYMMETRIC DUNKL AND q-DUNKL-CLASSICAL ORTHOGONAL POLYNOMIALS

Yahia Habbachi

Faculty of Sciences, University of Gabes,
Erriadh City 6072, Zrig, Gabes, Tunisia

Yahiahabbachi@hotmail.com

Abstract: In this paper, we consider the following $\mathcal{L}$-difference equation

$$\Phi(x)\mathcal{L}P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0,$$

where Φ is a monic polynomial (even), $\deg \Phi \leq 2$, ξ_n, ϑ_n, λ_n, $n \geq 0$, are complex numbers and $\mathcal{L}$ is either the Dunkl operator T_μ or the the q-Dunkl operator $T_{(\theta,q)}$. We show that if $\mathcal{L} = T_\mu$, then the only symmetric orthogonal polynomials satisfying the previous equation are, up a dilation, the generalized Hermite polynomials and the generalized Gegenbauer polynomials and if $\mathcal{L} = T_{(\theta,q)}$, then the q^2-analogue of generalized Hermite and the q^2-analogue of generalized Gegenbauer polynomials are, up a dilation, the only orthogonal polynomials sequences satisfying the $\mathcal{L}$-difference equation.

Keywords: Orthogonal polynomials, Dunkl operator, q-Dunkl operator.

1. Introduction

The classical orthogonal polynomials (Hermite, Laguerre, Bessel, and Jacobi) have a lot of useful characterizations: they satisfy a Hahn's property, that the sequence of their monic derivatives is again orthogonal (see [1, 8, 14, 16]), they are characterized as the polynomial eigenfunctions of a second order homogeneous linear differential (or difference) hypergeometric operator with polynomial coefficients [4, 15, 16], their corresponding linear functionals satisfy a distribution equation of Pearson type (see [11, 13, 15]).

Another characterization was established by Al-Salam and Chihara in [1], in particular they showed that the sequences Hermite, Laguerre and Jacobi are the only monic orthogonal polynomial sequences $\{P_n\}_{n\geq 0}$ that satisfy an equation of the form:

$$\pi(x)P'_{n+1}(x) = (a_n x + b_n)P_{n+1} + c_n P_n(x), \quad n \geq 0, \tag{1.1}$$

where π is a monic polynomial, $\deg \pi \leq 2$.

Recently, Datta and J. Griffin [9] studied the q-analogue of (1.1). More precisely they studied a q-difference equation of the form:

$$\pi(x)D_q P_{n+1}(x) = (a_n x + b_n)P_{n+1} + c_n P_n(x), \quad n \geq 0, \tag{1.2}$$

where π is a monic polynomial, $\deg \pi \leq 2$ and D_q is the Hahn operator defined by

$$D_q f(x) = (f(qx) - f(x))/(q-1)x, \quad f \in \mathcal{P}.$$

In particular they showed that the only orthogonal polynomials satisfying (1.2) are the Al-Salam-Carlitz I, the little and big q-Laguerre, the little and big q-Jacobi and the q-Bessel polynomials. The aim of this paper is to study the equation of the form:

$$\Phi(x)\mathcal{L}P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0, \tag{1.3}$$

where Φ is a monic polynomial (even), $\deg\Phi \leq 2$ and $\mathcal{L} \in \{T_\mu, T_{(\theta,q)}\}$.

This paper is organized as follows. In Section 2, we introduce the basic background and some preliminary results that will be used in what follows. In Section 3, we show that the only symmetric orthogonal polynomials satisfying (1.3), are, up a dilation, the generalized Hermite polynomials and the generalized Gegenbauer polynomials if $\mathcal{L} = T_\mu$ and the q^2-analogue of generalized Hermite polynomials and the q^2-analogue of generalized Gegenbauer polynomials if $\mathcal{L} = T_{(\theta,q)}$.

2. Preliminaries and notations

Let $\mathcal{P}$ be the vector space of polynomials with coefficients in $\mathbb{C}$ and let $\mathcal{P}'$ be its dual. We denote by $\langle u, f\rangle$ the action of $u \in \mathcal{P}'$ on $f \in \mathcal{P}$. In particular, we denote by $(u)_n = \langle u, x^n\rangle$, $n \geq 0$, the moments of u. For any form u, any polynomial f and any $a \in \mathbb{C}\setminus\{0\}$, let fu and $h_a u$, be the forms defined by duality:

$$\langle fu, p\rangle = \langle u, fp\rangle, \quad \langle h_a u, p\rangle = \langle u, h_a p\rangle, \quad p \in \mathcal{P},$$

where $h_a p(x) = p(ax)$.

Let $\{P_n\}_{n\geq 0}$ be a sequence of monic polynomials (MPS, in short) with $\deg P_n = n$, $n \geq 0$. The dual sequence associated with $\{P_n\}_{n\geq 0}$ is the sequence $\{u_n\}_{n\geq 0}$, $u_n \in \mathcal{P}'$ such that $\langle u_n, P_m\rangle = \delta_{n,m}$, n, $m \geq 0$, where $\delta_{n,m}$ is the Kronecker symbol [14].

The linear functional u is called regular if there exists a MPS $\{P_n\}_{n\geq 0}$ such that (see [8, p. 7]):

$$\langle u, P_m P_n\rangle = r_n \delta_{n,m}, \quad n, m \geq 0, \quad r_n \neq 0, \quad n \geq 0.$$

Then the sequence $\{P_n\}_{n\geq 0}$ is said to be orthogonal with respect to u. In this case, we have

$$u_n = \left(\langle u_0, P_n^2\rangle\right)^{-1} P_n u_0, \quad n \geq 0.$$

Moreover, $u = \lambda u_0$, where $(u)_0 = \lambda \neq 0$ [17].

In what follows all regular linear functionals u will be taken normalized i.e., $(u)_0 = 1$. Therefore, $u = u_0$.

A polynomial set $\{P_n\}_{n\geq 0}$ is called symmetric if

$$P_n(-x) = (-1)^n P_n(x), \quad n \geq 0.$$

According to Favard's theorem [8], a sequence of monic orthogonal polynomials $\{P_n(x)\}_{n\geq 0}$ (MOPS, in short) satisfies a three-term recurrence relation

$$\begin{cases} P_0(x) = 1, \quad P_1(x) = x, \\ P_{n+2}(x) = (x - \beta_{n+1})P_{n+1}(x) - \gamma_{n+1}P_n(x), \quad n \geq 0, \quad \gamma_{n+1} \neq 0, \quad n \geq 0. \end{cases} \tag{2.1}$$

with

$$\beta_n = \frac{\langle u_0, xP_n^2\rangle}{\langle u_0, P_n^2\rangle}, \quad \gamma_{n+1} = \frac{\langle u_0, P_{n+1}^2\rangle}{\langle u_0, P_n^2\rangle}, \quad n \geq 0.$$

A dilatation preserves the property of orthogonality. Indeed, the sequence $\{\widetilde{P}_n(x)\}_{n\geq 0}$ defined by

$$\widetilde{P}_n(x) = a^{-n} P_n(ax), \quad n \geq 0, \quad a \in \mathbb{C}\setminus\{0\},$$

satisfies the recurrence relation [16]

$$\begin{cases} \widetilde{P}_0(x) = 1, \quad \widetilde{P}_1(x) = x - \widetilde{\beta}_0, \\ \widetilde{P}_{n+2}(x) = (x - \widetilde{\beta}_{n+1})\widetilde{P}_{n+1}(x) - \widetilde{\gamma}_{n+1}\widetilde{P}_n(x), \quad n \geq 0, \end{cases}$$

where

$$\widetilde{\beta}_n = \frac{\beta_n}{a}, \quad \widetilde{\gamma}_{n+1} = \frac{\gamma_{n+1}}{a^2}, \quad n \geq 0. \tag{2.2}$$

Moreover, if $\{P_n\}_{n\geq 0}$ is a MOPS with respect to the regular form u_0, then $\{\widetilde{P}_n\}_{n\geq 0}$ is a MOPS with respect to the regular form $\widetilde{u}_0 = h_{a^{-1}} u_0$.

Theorem 1 [8]. *Let $\{P_n\}_{n\geq 0}$ be a MOPS satisfying (2.1) and orthogonal with respect to a linear functional u. The following statements are equivalent*:

(*i*) *the sequence $\{P_n\}_{n\geq 0}$ is symmetric*;
(*ii*) $(u)_{2n+1} = 0,\ n \geq 0$;
(*iii*) $\beta_n = 0,\ n \geq 0$.

Next, we introduce the Dunkl operator T_μ defined on $\mathcal{P}$ by [10, 18]

$$(T_\mu f)(x) = f'(x) + \mu H_{-1} f(x), \quad \mu > -\frac{1}{2}, \quad f \in \mathcal{P},$$

where

$$(H_{-1} f)(x) = \frac{f(x) - f(-x)}{2x}.$$

For the Dunkl operator, we have the property [6]

$$T_\mu(fg)(x) = (T_\mu f)(x) g(x) + f(x)(T_\mu g)(x) - 4\mu x (H_{-1} f)(x)(H_{-1} g)(x), \quad f, g \in \mathcal{P}.$$

In particular,

$$T_\mu(x P_{n+1}) = \big(1 + 2\mu(-1)^{n+1}\big) P_{n+1}(x) + x (T_\mu P_{n+1})(x), \quad n \geq 0. \tag{2.3}$$

We define the operator T_μ from $\mathcal{P}'$ to $\mathcal{P}'$ as follows:

$$\langle T_\mu u, f\rangle = -\langle u, T_\mu f\rangle, \quad f \in \mathcal{P}, \quad u \in \mathcal{P}'.$$

In particular,

$$(T_\mu u)_n = -\mu_n (u)_{n-1}, \quad n \geq 0,$$

with the convention $(u)_{-1} = 0$, where

$$\mu_n = n + \mu(1 - (-1)^n), \quad n \geq 0.$$

We introduce also the q-Dunkl operator $T_{(\theta,q)}$ defined on $\mathcal{P}$ by [2, 5, 7]

$$(T_{(\theta,q)} f)(x) = \frac{f(qx) - f(x)}{(q-1)x} + \theta H_{-1} f(x), \quad f \in \mathcal{P}, \quad \theta \in \mathbb{C}.$$

Remark 1. Note that when $q \to 1$, we again meet the Dunkl operator.

From the last definition, it is easy to prove that

$$T_{(\theta,q)}(fg) = (T_{(\theta,q)} f) g + (h_q f)(T_{(\theta,q)} g) + \theta(h_{-1} f - h_q f) H_{-1} g, \quad f, g \in \mathcal{P}.$$

In particular,

$$T_\mu(x P_{n+1}) = \left(1 + \theta - \theta(q+1)\frac{1 - (-1)^{n+1}}{2}\right) P_{n+1}(x) + qx (T_{(\theta,q)} P_{n+1})(x), \quad n \geq 0. \tag{2.4}$$

We define the operator $T_{(\theta,q)}$ from $\mathcal{P}'$ to $\mathcal{P}'$ as follows:

$$\langle T_{(\theta,q)}u, f\rangle = -\langle u, T_{(\theta,q)}f\rangle, \quad f \in \mathcal{P}, \quad u \in \mathcal{P}'.$$

In particular,

$$(T_{(\theta,q)})_n = -\theta_{n,q}(u)_{n-1}, \quad n \geq 0,$$

where $(u)_{-1} = 0$ and

$$\theta_{n,q} = [n]_q + \theta\frac{1-(-1)^n}{2}, \quad n \geq 0, \tag{2.5}$$

here $[n]_q$, $n \geq 0$, denotes the basic q-number defined by

$$[n]_q = \frac{1-q^n}{1-q} = 1 + q + ... + q^{n-1}, \quad n \geq 1, \quad [0]_q = 0.$$

According to the definitions of T_μ and $T_{(\theta,q)}$, we have

$$T_\mu(x^n) = \mu_n x^{n-1}, \quad T_{(\theta,q)}(x^n) = \theta_{n,q}x^{n-1}.$$

3. The main results

In this section, we will look for all symmetric MOPS satisfying (1.3). We distinguish two cases. The first case is when $\mathcal{L} = T_\mu$ and the second one is when $\mathcal{L} = T_{(\theta,q)}$.

3.1. First case: when $\mathcal{L} = T_\mu$

Theorem 2. *The only symmetric MOPS satisfying a T_μ-difference equation of the form*

$$\Phi(x)T_\mu P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0, \tag{3.1}$$

where Φ is a monic polynomial (even), $\deg\Phi \leq 2$, are, up a dilation, the generalized Hermite polynomials and the generalized Gegenbauer polynomials.

P r o o f. Let $\{P_n\}_{n\geq 0}$ be a symmetric MOPS satisfying (3.1). Since Φ is a monic, even and $\deg\Phi \leq 2$, then we distinguish two cases: $\Phi(x) = 1$ and $\Phi(x) = x^2 + c$.

Case 1. $\Phi(x) = 1$, then (3.1) becomes

$$T_\mu P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0.$$

By comparing the degrees in the last equation (in x^{n+2} and x^{n+1}), we obtain $\xi_n = \vartheta_n = 0$, $n \geq 0$ and then

$$T_\mu P_{n+1}(x) = \lambda_n P_n(x), \quad n \geq 0. \tag{3.2}$$

Identifying coefficients in the monomials of degree n in the last equation, we obtain

$$\lambda_n = \mu_{n+1}, \quad n \geq 0. \tag{3.3}$$

On the other hand, applying the operator T_μ to (2.1) with $\beta_{n+1} = 0$ and using (2.3), we get

$$T_\mu P_{n+2}(x) = \big(1 + 2\mu(-1)^{n+1}\big)P_{n+1}(x) + x(T_\mu P_{n+1})(x) - \gamma_{n+1}(T_\mu P_n)(x), \quad n \geq 0.$$

Substituting (3.2) and (3.3) in the last equation and taking into account the fact that

$$1 + 2\mu(-1)^{n+1} = \mu_{n+2} - \mu_{n+1},$$

we get

$$\mu_{n+1}P_{n+1}(x) = \mu_{n+1}xP_n(x) - \mu_n\gamma_{n+1}P_{n-1}(x), \quad n \geq 0.$$

From (2.1), the last equation is equivalent to

$$\mu_{n+1}\gamma_n P_{n-1}(x) = \mu_n\gamma_{n+1}P_{n-1}(x), \quad n \geq 0,$$

hence,

$$\mu_{n+1}\gamma_n = \mu_n\gamma_{n+1}, \quad n \geq 1.$$

Therefore,

$$\gamma_{n+1} = \frac{\gamma_1}{\mu_1}\mu_{n+1}, \quad n \geq 1.$$

Since the last relation remains valid for $n = 0$, then we have

$$\gamma_{n+1} = \frac{\gamma_1}{\mu_1}\mu_{n+1}, \quad n \geq 0.$$

Using (2.2), where $a^2 = 2\gamma_1/\mu_1$, we obtain

$$\widetilde{\beta}_n = 0, \quad \widetilde{\gamma}_{n+1} = \frac{\mu_{n+1}}{2}, \quad n \geq 0.$$

So, we meet the recurrence coefficients for the generalized Hermite polynomial sequence (see [8]).

Case 2. $\Phi(x) = x^2 + c$, then (3.1) becomes

$$(x^2 + c)T_\mu P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0. \tag{3.4}$$

Identifying the coefficients of higher degree in both sides of (3.4), we obtain $\xi_n = \mu_{n+1}$, $n \geq 0$. Therefore, (3.4) becomes

$$(x^2 + c)T_\mu P_{n+1}(x) = (\mu_{n+1}x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0. \tag{3.5}$$

Applying the operator T_μ to (2.1) with $\beta_{n+1} = 0$ and using (2.3) and the fact that

$$1 + 2\mu(-1)^{n+1} = \mu_{n+2} - \mu_{n+1},$$

we get

$$T_\mu P_{n+2}(x) = (\mu_{n+2} - \mu_{n+1})P_{n+1}(x) + x(T_\mu P_{n+1})(x) - \gamma_{n+1}(T_\mu P_n)(x), \quad n \geq 0.$$

Multiplying the previous equation by $x^2 + c$ and using (3.5), we get

$$\begin{gathered}(\mu_{n+2}x + \vartheta_{n+1})P_{n+2}(x) + \lambda_{n+1}P_{n+1}(x) = (\mu_{n+2} - \mu_{n+1})(x^2 + c)P_{n+1}(x) + \\ (\mu_{n+1}x^2 + \vartheta_n x)P_{n+1}(x) + \lambda_n x P_n(x) - \gamma_{n+1}\big((\mu_n x + \vartheta_{n-1})P_n(x) + \lambda_{n-1}P_{n-1}(x)\big), \quad n \geq 1,\end{gathered}$$

or, equivalently,

$$\begin{gathered}(\vartheta_{n+1} - \vartheta_n)xP_{n+1}(x) - c(\mu_{n+2} - \mu_{n+1})P_{n+1} + \lambda_{n+1}P_{n+1}(x) \\ = \lambda_n x P_n(x) + \gamma_{n+1}\big((\mu_{n+2}x + \vartheta_{n+1})P_n(x) - (\mu_n x + \vartheta_{n-1})P_n(x) - \lambda_{n-1}P_{n-1}(x)\big), \quad n \geq 1.\end{gathered} \tag{3.6}$$

Comparing the degrees in the last equation, we obtain $\vartheta_{n+1} = \vartheta_n$, $n \geq 1$. But, from (3.5) and the fact that $\{P_n\}_{n\geq 0}$ is symmetric, where $n = 0$ and $n = 1$, we get, respectively,

$$\begin{aligned} v_0 &= 0, \quad \lambda_0 = c(1 + 2\mu), \\ v_1 &= 0, \quad \lambda_1 = 2(\gamma_1 + c). \end{aligned} \tag{3.7}$$

Thus,

$$\vartheta_n = 0, \quad n \geq 0.$$

Therefore, (3.6) becomes

$$\begin{aligned} &c(\mu_{n+1} - \mu_{n+2})P_{n+1}(x) + \lambda_{n+1}P_{n+1}(x) \\ &= \lambda_n x P_n(x) + \gamma_{n+1}\big((\mu_{n+2} - \mu_n)xP_n(x) - \lambda_{n-1}P_{n-1}(x)\big), \quad n \geq 1. \end{aligned}$$

Taking into account (2.1), we get

$$\begin{aligned} &\big(\lambda_{n+1} + c(\mu_{n+1} - \mu_{n+2})\big)xP_n(x) - \gamma_n\big(\lambda_{n+1} + c(\mu_{n+1} - \mu_{n+2})\big)P_{n-1}(x) \\ &= \big(\lambda_n + (\mu_{n+2} - \mu_n)\gamma_{n+1}\big)xP_n(x) - \lambda_{n-1}\gamma_{n+1}P_{n-1}(x), \quad n \geq 1. \end{aligned}$$

Then,

$$\lambda_{n+1} + c(\mu_{n+1} - \mu_{n+2}) = \lambda_n + (\mu_{n+2} - \mu_n)\gamma_{n+1}, \quad n \geq 1, \tag{3.8}$$

$$\big(\lambda_{n+1} + c(\mu_{n+1} - \mu_{n+2})\big)\gamma_n = \lambda_{n-1}\gamma_{n+1}, \quad n \geq 1. \tag{3.9}$$

Since $\mu_{n+2} - \mu_n = 2$, then, substitution of (3.8) in (3.9) gives

$$(\lambda_n + 2\gamma_{n+1})\gamma_n = \lambda_{n-1}\gamma_{n+1}, \quad n \geq 1.$$

Therefore,

$$\frac{\lambda_n}{\gamma_{n+1}} = \frac{\lambda_{n-1}}{\gamma_n} - 2, \quad n \geq 1.$$

So,

$$\lambda_n = \frac{\lambda_0 - 2n\gamma_1}{\gamma_1}\gamma_{n+1}, \quad n \geq 1. \tag{3.10}$$

It is clear that (3.10) remains valid for $n = 0$. Then, we have

$$\lambda_n = \frac{\lambda_0 - 2n\gamma_1}{\gamma_1}\gamma_{n+1}, \quad n \geq 0. \tag{3.11}$$

Substitution of (3.11) in (3.8) gives

$$\lambda_{n+1} = \frac{\lambda_0 - 2(n-1)\gamma_1}{\lambda_0 - 2n\gamma_1}\lambda_n + c(\mu_{n+2} - \mu_{n+1}), \quad n \geq 1.$$

By virtue of fourth equality in (3.7), we obtain that the previous equation remains valid for $n = 0$. Hence,

$$\lambda_{n+1} = \frac{\lambda_0 - 2(n-1)\gamma_1}{\lambda_0 - 2n\gamma_1}\lambda_n + c(\mu_{n+2} - \mu_{n+1}), \quad n \geq 0. \tag{3.12}$$

We will distinguish two situations: $c = 0$ and $c \neq 0$.

- If $c = 0$, then from (3.7) we have $\lambda_0 = 0$. Therefore, $\lambda_n = 0$, $n \geq 0$. Consequently, according to (3.11) and the fourth equality in (3.7), $\gamma_{n+1} = 0$, $n \geq 0$. This contradicts the orthogonality of $\{P_n\}_{n\geq 0}$.
- If $c \neq 0$, using a dilatation, we can take $c = -1$. Putting

$$\gamma_1 = \frac{1 + 2\mu}{3 + 2\mu + 2\alpha},$$

then (3.12) becomes

$$\lambda_{n+1} = \frac{2n+2\alpha+2\mu+1}{2n+2\alpha+2\mu+3}\lambda_n + \mu_{n+1} - \mu_{n+2}, \quad n \geq 0. \tag{3.13}$$

From (3.13), we can easily prove by induction that

$$\lambda_n = -\frac{\mu_{n+1}(\mu_{n+1}+2\alpha)}{2n+2\alpha+2\mu+1}, \quad n \geq 0.$$

Thus, (3.11) gives

$$\gamma_{n+1} = \frac{\mu_{n+1}(\mu_{n+1}+2\alpha)}{(2n+2\alpha+2\mu+1)(2n+2\alpha+2\mu+3)}, \quad n \geq 0.$$

So, we meet the recurrence coefficients for the generalized Gegenbauer polynomial (see [3, 8]). □

Remark 2. Notice that when $\mu = 0$ in (3.1), we again meet (1.1) for the symmetric case.

3.2. Second case: when $\mathcal{L} = T_{(\theta,q)}$

Theorem 3. *The only symmetric MOPS satisfying a $T_{(\theta,q)}$-difference equation of the form:*

$$\Phi(x)T_{(\theta,q)}P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0, \tag{3.14}$$

where Φ is a monic polynomial (even), $\deg \Phi \leq 2$, are, up a dilation, the q^2-analogue of generalized Hermite polynomials and the q^2-analogue of generalized Gegenbauer polynomials.

P r o o f. Let $\{P_n\}_{n\geq 0}$ be a symmetric MOPS satisfying (3.1). As in proof of Theorem 2, we distinguish two cases: $\Phi(x) = 1$ and $\Phi(x) = x^2 + c$.

Case 1. $\Phi(x) = 1$, then (3.14) becomes

$$T_{(\theta,q)}P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0. \tag{3.15}$$

By comparing the degrees in (3.15), we obtain $\xi_n = \vartheta_n = 0,\ n \geq 0$. Then,

$$T_{(\theta,q)}P_{n+1}(x) = \lambda_n P_n(x), \quad n \geq 0.$$

The comparison of the coefficients of x^n in the previous equation leads to $\lambda_n = \theta_{n+1,q},\ n \geq 0$. Therefore,

$$T_{(\theta,q)}P_{n+1}(x) = \theta_{n+1,q}P_n(x), \quad n \geq 0. \tag{3.16}$$

Now, applying $T_{(\theta,q)}$ to (2.1) with $\beta_{n+1} = 0$ and using (2.4), we get

$$\begin{aligned} T_{(\theta,q)}P_{n+2}(x) = &\left(1+\theta-\theta(q+1)\frac{1-(-1)^{n+1}}{2}\right)P_{n+1}(x) \\ &+q\,x(T_{(\theta,q)}P_{n+1})(x) - \gamma_{n+1}(T_{(\theta,q)}P_n)(x), \quad n \geq 0. \end{aligned}$$

Substituting (3.16) in the last equation, we get

$$\begin{aligned} \theta_{n+2,q}P_{n+1}(x) = &\left(1+\theta-\theta(q+1)\frac{1-(-1)^{n+1}}{2}\right)P_{n+1}(x) \\ &+q\theta_{n+1,q}\,xP_n(x) - \gamma_{n+1}\theta_{n,q}P_{n-1}(x), \quad n \geq 0. \end{aligned}$$

Using the fact that

$$xP_n = P_{n+1} + \gamma_n P_{n-1},$$

we obtain

$$\begin{aligned}\Big(\theta_{n+2,q} - 1 - \theta + \theta(q+1)\frac{1-(-1)^{n+1}}{2} - q\theta_{n+1,q}\Big)P_{n+1}(x)\\ = q\theta_{n+1,q}\gamma_n P_{n-1}(x) - \theta_{n,q}\gamma_{n+1}P_{n-1}(x), \quad n \geq 0.\end{aligned}$$

After easy calculations from (2.5), we have

$$\theta_{n+2,q} - 1 - \theta + \theta(q+1)\frac{1-(-1)^{n+1}}{2} - q\theta_{n+1,q} = 0, \quad n \geq 0. \tag{3.17}$$

Therefore,

$$(q\theta_{n+1,q}\gamma_n - \theta_{n,q}\gamma_{n+1})P_{n-1}(x) = 0, \quad n \geq 0.$$

Hence,

$$q\theta_{n+1,q}\gamma_n = \theta_{n,q}\gamma_{n+1}, \quad n \geq 1.$$

Then, we can deduce by induction that

$$\gamma_{n+1} = \frac{\gamma_1}{1+\theta}\, q^n\, \theta_{n+1,q}, \quad n \geq 1.$$

Moreover, the previous identity remains valid for $n = 0$, thus

$$\gamma_{n+1} = \frac{\gamma_1}{1+\theta}\, q^n\, \theta_{n+1,q}, \quad n \geq 0.$$

Then, according to (2.2), with the choice

$$a^2 = q(q+1)\frac{\gamma_1}{1+\theta}$$

and putting

$$\mu = \frac{1+\theta}{q(q+1)} - \frac{1}{2},$$

we obtain

$$\widetilde{\beta}_n = 0, \quad \widetilde{\gamma}_{n+1} = q^n\,\frac{\theta_{n+1,q}}{q(q+1)}, \quad n \geq 0,$$

which are the recurrence coefficients for the q^2-analogue of generalized Hermite polynomial $H_n^{(\mu,q^2)}$ [12], with

$$\mu = \frac{1+\theta}{q(q+1)} - \frac{1}{2}.$$

Case 2: $\Phi(x) = x^2 + c$, then in this case (3.14) becomes

$$(x^2+c)T_{(\theta,q)}P_{n+1}(x) = (\xi_n x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0. \tag{3.18}$$

By comparing terms of higher degree in the previous equation, we obtain

$$\xi_n = \theta_{n+1,q}, \quad n \geq 0.$$

Then, equation (3.18) becomes

$$(x^2+c)T_{(\theta,q)}P_{n+1}(x) = (\theta_{n+1,q}x + \vartheta_n)P_{n+1}(x) + \lambda_n P_n(x), \quad n \geq 0. \tag{3.19}$$

Applying the operator $T_{(\theta,q)}$ to (2.1) with $\beta_{n+1}=0$ and using (2.4), we get

$$\begin{aligned} T_{(\theta,q)}P_{n+2}(x) &= \left(1+\theta-\theta(q+1)\frac{1-(-1)^{n+1}}{2}\right)P_{n+1}(x)\\ &+qx(T_{(\theta,q)}P_{n+1})(x)-\gamma_{n+1}(T_{(\theta,q)}P_n)(x), \quad n\geq 0. \end{aligned}$$

By (3.17), the last equation becomes

$$\begin{aligned} T_{(\theta,q)}P_{n+2}(x) &= (\theta_{n+2,q}-q\theta_{n+1,q})P_{n+1}(x)\\ &+qx(T_{(\theta,q)}P_{n+1})(x)-\gamma_{n+1}(T_{(\theta,q)}P_n)(x), \quad n\geq 0. \end{aligned}$$

Multiplying the above equation by x^2+c and substituting (3.19) into the result, we get

$$\begin{aligned} (\theta_{n+2,q}x+\vartheta_{n+1})P_{n+2}(x)+\lambda_{n+1}P_{n+1}(x) &= (\theta_{n+2,q}-q\theta_{n+1,q})(x^2+c)P_{n+1}(x)\\ +q(\theta_{n+1,q}x^2+\vartheta_n x)P_{n+1}(x)+q\lambda_n xP_n(x) &-\gamma_{n+1}\big((\theta_{n,q}x+\vartheta_{n-1})P_n(x)+\lambda_{n-1}P_{n-1}(x)\big), \quad n\geq 1. \end{aligned}$$

Substituting of (2.1) in the previous equation, we get

$$\begin{aligned} (\vartheta_{n+1}-q\vartheta_n)xP_{n+1}(x)+\big(\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})\big)P_{n+1}(x) &=\\ q\lambda_n xP_n(x)+\gamma_{n+1}\big(((\theta_{n+2,q}-\theta_{n,q})x+\vartheta_{n+1}-\vartheta_{n-1})P_n(x)-\lambda_{n-1}P_{n-1}(x)\big), &\quad n\geq 1. \end{aligned}$$

The comparison of the coefficients of x^{n+2} in the previous equation gives $\vartheta_{n+1}=q\vartheta_n$, $n\geq 1$ and putting $n=0$ and $n=1$ in (3.19), we get respectively

$$\begin{aligned} v_0=0, &\quad \lambda_0=c(1+\theta),\\ v_1=0, &\quad \lambda_1=(1+q)(\gamma_1+c). \end{aligned} \tag{3.20}$$

Hence, $\vartheta_n=0,\ n\geq 0$.

Therefore, the last equation becomes

$$\begin{aligned} &\big(\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})\big)P_{n+1}(x)\\ &=q\lambda_n xP_n(x)+\gamma_{n+1}\big((\theta_{n+2,q}-\theta_{n,q})xP_n(x)-\lambda_{n-1}P_{n-1}(x)\big), \quad n\geq 1. \end{aligned}$$

Using the fact that $P_{n+1}=xP_n(x)-\gamma_nP_{n-1}$, the above equation is equivalent to

$$\begin{aligned} \big(\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})\big)xP_n(x)-\gamma_n\big(\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})\big)P_{n-1}(x)\\ =\big(q\lambda_n+(\theta_{n+2,q}-\theta_{n,q})\gamma_{n+1}\big)xP_n(x)-\lambda_{n-1}\gamma_{n+1}P_{n-1}(x), \quad n\geq 1. \end{aligned}$$

Then, we deduce

$$\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})=q\lambda_n+(\theta_{n+2,q}-\theta_{n,q})\gamma_{n+1}, \quad n\geq 1, \tag{3.21}$$

$$\Big(\lambda_{n+1}-c(\theta_{n+2,q}-q\theta_{n+1,q})\Big)\gamma_n=\lambda_{n-1}\gamma_{n+1}, \quad n\geq 1. \tag{3.22}$$

Since

$$\theta_{n+2,q}-\theta_{n,q}=(1+q)q^n,$$

then the substitution of (3.21) in (3.22) gives

$$(q\lambda_n+(1+q)q^n\gamma_{n+1})\gamma_n=\lambda_{n-1}\gamma_{n+1}, \quad n\geq 1,$$

therefore,

$$q\lambda_n=\left(\frac{\lambda_{n-1}}{\gamma_n}-(1+q)q^n\right)\gamma_{n+1}, \quad n\geq 1.$$

We can easily deduce by induction that

$$q^n\lambda_n = \left(\frac{\lambda_0}{\gamma_1} - q(q+1)[n]_{q^2}\right)\gamma_{n+1}, \quad n \geq 1.$$

It is clear that the previous identity remains valid for $n = 0$. Then, we have

$$q^n\lambda_n = \left(\frac{\lambda_0}{\gamma_1} - q(q+1)[n]_{q^2}\right)\gamma_{n+1}, \quad n \geq 0. \tag{3.23}$$

Now, we will determine λ_n. By (3.23), we have

$$\gamma_{n+1} = q^n \frac{\gamma_1}{\lambda_0 - q(q+1)\gamma_1[n]_{q^2}}\, \lambda_n, \quad n \geq 0. \tag{3.24}$$

Therefore, (3.21) becomes

$$\lambda_{n+1} = \frac{q\lambda_0 - (q+1)\gamma_1([n]_{q^2} - 1)}{\lambda_0 - q(q+1)\gamma_1[n]_{q^2}}\lambda_n + c(\theta_{n+2,q} - q\theta_{n+1,q}), \quad n \geq 1.$$

By virtue of (3.20), we obtain that the previous equation remains valid for $n = 0$.

Then,

$$\lambda_{n+1} = \frac{q\lambda_0 - (q+1)\gamma_1([n]_{q^2} - 1)}{\lambda_0 - q(q+1)\gamma_1[n]_{q^2}}\lambda_n + c(\theta_{n+2,q} - q\theta_{n+1,q}), \quad n \geq 0. \tag{3.25}$$

We will distinguish two situations: $c = 0$ and $c \neq 0$.

- If $c = 0$, then from (3.20) $\lambda_0 = 0$. Therefore, $\lambda_n = 0$, $n \geq 0$. Consequently, according to (3.24) and the fourth equality in (3.20), $\gamma_{n+1} = 0$, $n \geq 0$. This contradicts the orthogonality of $\{P_n\}_{n\geq 0}$.
- If $c \neq 0$, using a suitable dilatation, we can suppose that $c = -1$. Putting

$$\gamma_1 = \frac{1+\theta}{1+\theta+q(q+1)(\alpha+1)}. \tag{3.26}$$

Equation (3.25) becomes

$$\lambda_{n+1} = q\,\frac{q(q+1)(\alpha+1) + \theta_{2n-1,q}}{q(q+1)(\alpha+1) + \theta_{2n+1,q}}\lambda_n - (\theta_{n+2,q} - q\theta_{n+1,q}), \quad n \geq 0. \tag{3.27}$$

Therefore, from (3.27), we can prove by induction that

$$\lambda_n = -\frac{\theta_{n+1,q}\Big(q(q+1)(\alpha+1) + \theta_{n-1,q}(1+\theta(1-q)(1-(-1)^n)/2)\Big)}{q(q+1)(\alpha+1) + \theta_{2n-1,q}}, \quad n \geq 0. \tag{3.28}$$

By virtue of (3.24), (3.26) and (3.28), we get

$$\gamma_{n+1} = q^n \frac{\theta_{n+1,q}\Big(q(q+1)(\alpha+1) + \theta_{n-1,q}(1+\theta(1-q)(1-(-1)^n)/2)\Big)}{\Big(q(q+1)(\alpha+1) + \theta_{2n-1,q}\Big)\Big(q(q+1)(\alpha+1) + \theta_{2n+1,q}\Big)}, \quad n \geq 0.$$

So, we meet the recurrence coefficients for the q^2-anlogue of generalized Gegenbauer polynomial $S_n^{(\alpha,\beta,q^2)}$, with

$$\beta = \frac{1+\theta}{q(q+1)} - 1$$

(see [12]). □

Remark 3. Notice that when $q \to 1$, we recover the result in Theorem 2 and when $\theta = 0$ in (3.14), we again meet (1.2) for symmetric case.

4. Conclusion

To conclude this paper, we will present two tables in which we give the only symmetric MOPS verifying the $\mathcal{L}$-difference (1.3).

Polynomial	Φ	ξ_n	ϑ_n	λ_n
Generalized Hermite $H_n^{(\mu,q^2)}$	1	0	0	$\mu_{n+1}, \quad n \geq 0$
Generalized Gegenbauer $S_n^{(\alpha,\beta,q^2)}$	x^2-1	μ_{n+1}	0	$-\dfrac{\mu_{n+1}(\mu_{n+1}+2\alpha)}{2n+2\alpha+2\mu+1}, \quad n \geq 0$

Table 1: Case when $\mathcal{L} = T_\mu$

Polynomial	Φ	ξ_n	ϑ_n	λ_n
q^2-analogue of generalized Hermite $H_n^{(\mu,q^2)}$	1	0	0	$\theta_{n+1,q}, \quad n \geq 0$
q^2-analogue of generalized Gegenbauer $S_n^{(\alpha,\beta,q^2)}$	x^2-1	$\theta_{n+1,q}$	0	$-\frac{\theta_{n+1,q}\left(q(q+1)(\alpha+1)+\theta_{n-1,q}(1+\theta(1-q)(1-(-1)^n)/2)\right)}{q(q+1)(\alpha+1)+\theta_{2n-1,q}},$ $n \geq 0.$

Table 2: Case when $\mathcal{L} = T_{(\theta,q)}$

Remark 4. In this paper, we have studied only the symmetric case. The question for non-symmetric case remains open.

Acknowledgements

The author thanks Professor Francisco Marcellán for many enlightening discussions and the anonymous referees for their careful reading of the manuscript and corrections.

REFERENCES

1. Al-Salam W. A., Chihara T. S. Another characterization of the classical orthogonal polynomials. *SIAM J. Math. Anal.*, 1972. Vol. 3, No. 1. P. 65–70. DOI: 10.1137/0503007
2. Aloui B., Souissi J. Characterization of q-Dunkl-classical symmetric orthogonal q-polynomials. *Ramanujan J.*, 2022. Vol. 57, No. 2. P. 1355–1365. DOI: 10.1007/s11139-021-00425-8
3. Belmehdi S. Generalized Gegenbauer orthogonal polynomials. *J. Comput. Appl. Math.*, 2001. Vol. 133, No. 1–2. P. 195–205. DOI: 10.1016/S0377-0427(00)00643-9
4. Bochner S. Über Sturm–Liouvillesche polynomsysteme. *Math. Z.*, 1929. Vol. 29. P. 730–736. DOI: 10.1007/BF01180560 (in German)
5. Bouanani A., Khériji L., Ihsen Tounsi M. Characterization of q-Dunkl Appell symmetric orthogonal q-polynomials. *Expo. Math.*, 2010. Vol. 28, No. 4. P. 325–336. DOI: 10.1016/j.exmath.2010.03.003
6. Bouras B. Some characterizations of Dunkl-classical orthogonal polynomials. *J. Difference Equ. Appl.*, 2014. Vol. 20, No. 8. P. 1240–1257. DOI: 10.1080/10236198.2014.906590
7. Bouras B., Habbachi Y., Marcellán F. Characterizations of the symmetric $T_{(\theta,q)}$-classical orthogonal q-polynomials. *Mediterr. J. Math.*, 2022. Vol. 19, No. 2. Art. no. 66. 18 p. DOI: 10.1007/s00009-022-01986-8
8. Chihara T. S. *An Introduction to Orthogonal Polynomials.* New York: Gordon and Breach, Sci. Publ., Inc., 1978. 249 p.
9. Datta S., Griffin J. A characterization of some q-orthogonal polynomials. *Ramanujan J.*, 2006. Vol. 12. P. 425–437. DOI: 10.1007/s11139-006-0152-5

10. Dunkl C. F. Integral kernels reflection group invariance. *Canad. J. Math.*, 1991. Vol. 43, No. 6. P. 1213–1227. DOI: 10.4153/CJM-1991-069-8
11. Geronimus J. L. On polynomials orthogonal with respect to numerical sequences and on Hahn's theorem. *Izv. Akad. Nauk.*, 1940. Vol. 250. P. 215–228. (in Russian)
12. Ghressi A., Khériji L. The symmetrical H_q-semiclassical orthogonal polynomials of class one. *SIGMA Symmetry Integrability Geom. Methods Appl.*, 2009. Vol. 5. Art. no. 076. 22 p. DOI: 10.3842/SIGMA.2009.076
13. Hildebrandt E. H. Systems of polynomials connected with the Charlier expansions and the Pearson differential and difference equation. *Ann. Math. Statist.*, 1931. Vol. 2, No. 4. P. 379–439. DOI: 10.1214/aoms/1177732950
14. Marcellán F., Branquinho A., Petronilho J. Classical orthogonal polynomials: a functional approach. *Acta. Appl. Math.*, 1994. Vol. 34. P. 283–303. DOI: 10.1007/BF00998681
15. Maroni P. Fonctions eulériennes. Polynômes orthogonaux classiques. *Techniques de L'ingénieur*, 1994. Vol. 154. P. 1–30. DOI: 10.51257/a-v1-a154 (in French)
16. Maroni P. Une théorie algébrique des polynômes orthogonaux. Application aux polynômes orthogonaux semi-classiques. *IMACS Ann. Comput. Appl. Math.*, 1991. Vol. 9. P. 95–130. (in French)
17. Maroni P. Variations around classical orthogonal polynomials. Connected problems. *J. Comput. Appl. Math.*, 1993. Vol. 48, No. 1–2. P. 133–155. DOI: 10.1016/0377-0427(93)90319-7
18. Sghaier M. A note on Dunkl-classical orthogonal polynomials. *Integral Transforms Spec. Funct.*, 2012. Vol. 23, No. 10. P. 753–760. DOI: 10.1080/10652469.2011.631186

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 121–131

DOI: 10.15826/umj.2023.2.010

CONTROL PROBLEM FOR A PARABOLIC SYSTEM WITH UNCERTAINTIES AND A NON-CONVEX GOAL[1]

Igor' V. Izmest'ev[a,b,†], **Viktor I. Ukhobotov**[a,††]

[a]South Ural State University,
76 Lenina ave., Chelyabinsk, 454080, Russian Federation

[b]Chelyabinsk State University,
129 Brat'ev Kashirinykh str., Chelyabinsk, 454001, Russian Federation

[†]j748e8@gmail.com [††]ukh@csu.ru

Abstract: We consider the control problem for a parabolic system that describes the heating of a given number of rods. Control is carried out through heat sources that are located at the ends of the rods (only at one end or at both). The density functions of the internal heat sources and exact values of the temperature at the right ends of some rods are unknown, and only the segments of their change are given. The goal of choosing control is to ensure that at a fixed time moment the weighted sum of the average temperatures of the rods belongs to a non-convex terminal set for any admissible unknown functions. After a change of variables, this problem reduces to a one-dimensional differential game. Necessary and sufficient conditions for the game termination are found.

Keywords: Control, Uncertainty, Parabolic system.

1. Introduction

Mathematical modelling of controlled processes of thermal conductivity, diffusion, filtration leads to problems of control of parabolic equations [2, 4, 8, 11, 13]. In applications, problems often arise about heating a rod at the ends of which there are controlled heat sources. In a formalized form, these problems are reduced to the study of the heat equation, the boundary conditions of which depend on the control functions (see, for example, [1, 10]).

Control processes for real dynamic systems often occur in conditions where some of the system parameters and boundary conditions are not precisely specified, and there is also influence from uncontrolled disturbances [3, 5, 18, 19].

To study such problems, the method of optimization of guaranteed result [9] can be applied. This method is based on the theory of differential games (see, for example, [12, 14]). Uncertainties and disturbances affecting the system are taken as the second player – the opponent. In [12, 14] control is constructed within the framework of the theory of positional differential games.

This article continues the research begun in [6, 15]. The work [15] considers the problem of heating a rod by controlling the rate of temperature change at its left end. The temperature at the right end of the rod is determined by an unknown function limited in value. The density function of the internal heat sources of the rod is not precisely known, and only the boundaries of its possible values are given. The goal of the control is to bring the average temperature of the rod at a fixed time moment to a given segment for any unknown temperature at the right end of the rod and for any function of the density of internal heat sources. The average temperature value is calculated as the integral of the product of temperature and a given function. In [6] the problem of controlling a

[1]The research was supported by a grant from the Russian Science Foundation no. 23-21-00539, https://rscf.ru/project/23-21-00539/.

parabolic system describing the heating of a given number of rods using point heat sources located at the ends of the rods is considered. The goal of choosing a control is to ensure that at a fixed time moment the modulus of the linear function, determined using the average temperatures of the rods, does not exceed a given value.

In this work, a modification of problems [6, 15] is solved. A finite set of desired temperature values is given. The goal of the control is to bring the weighted sum of the average temperatures of the rods into the ε–neighbourhood of one of the desired values. After changing variables, taking unknown functions as a control of the second player, the original problem is reduced to a single-type one-dimensional differential game. For the resulting differential game, a solvability set and corresponding player controls are constructed.

2. Problem statement

The heat equation

$$\frac{\partial T_i(x,t)}{\partial t} = \frac{\partial^2 T_i(x,t)}{\partial x^2} + f_i(x,t), \quad 0 \le t \le p, \quad 0 \le x \le 1, \quad i = \overline{1,n}, \tag{2.1}$$

describes the temperature distribution $T_i(x,t)$ in i-th $(i = \overline{1,n})$ homogeneous rod of unit length as a function of time t. At the initial time moment $t = 0$, the temperature distributions $T_i(x,0) = g_i(x)$, $i = \overline{1,n}$, are given, where $g_i(x)$ are continuous functions.

We assume that the controlled temperature $T_i(0,t)$ at the left end of i–th rod varies according to equation

$$\frac{dT_i(0,t)}{dt} = a_i^{(1)}(t) + a_i^{(2)}(t) G_i^{(1)} \overline{\xi}(t). \tag{2.2}$$

Here, $a_i^{(\zeta)}(t)$, $i = \overline{1,n}$, $\zeta = 1,2$, are continuous functions for $0 \le t \le p$, and $a_i^{(2)}(t) > 0$. The vector-function $\overline{\xi}(t) = (\xi_1(t), \xi_2(t), \ldots, \xi_q(t))^* \in U$, where U is compact in $\mathbb{R}^q$, is a control. The symbol $*$ denotes the transposition operation. The choice of the corresponding one-dimensional controls $\xi_\iota(t)$ for the left end of each rod is given by the matrix $G^{(1)}$ of n by q dimension. $G_i^{(1)}$ denotes the i–th row of the corresponding matrix.

The temperature value $T_i(1,t)$ at the right end of the i-th rod is given as follows:

1. Determined by $\overline{\xi}$ control

$$\frac{dT_i(1,t)}{dt} = b_i^{(1)}(t) + b_i^{(2)}(t) G_i^{(2)} \overline{\xi}(t), \quad i = \overline{1,k}. \tag{2.3}$$

   Here, the functions $b_i^{(1)}(t)$ and $b_i^{(2)}(t)$, $i = \overline{1,k}$, are continuous for $0 \le t \le p$, and $b_i^{(2)}(t) > 0$. The choice of the corresponding one-dimensional controls $\xi_\iota(t)$ for the right end of the rods with indices $i = \overline{1,k}$ is given by the matrix $G^{(2)}$ of k by q dimension. $G_i^{(2)}$ denotes the i–th row of the corresponding matrix.
2. The temperature values $T_i(1,t)$, $i = \overline{k+1,l}$, which depend continuously on the time $t \in [0,p]$, are not exactly known, but the limits of their change are given

$$\beta_i^{(1)}(t) \le T_i(1,t) \le \beta_i^{(2)}(t), \quad 0 \le t \le p. \tag{2.4}$$

   Here $\beta_i^{(\zeta)}(t)$, $i = \overline{k+1,l}$, $\zeta = 1,2$, are continuous functions for $0 \le t \le p$.
3. $T_i(1,t)$, $i = \overline{l+1,n}$, are known continuous functions.

In addition, we know estimates of the continuous functions $f_i(x,t)$, which are the densities of internal heat sources of the rods:

$$f_i^{(1)}(x,t) \le f_i(x,t) \le f_i^{(2)}(x,t), \quad 0 \le t \le p, \quad 0 \le x \le 1, \quad i = \overline{1,n}. \tag{2.5}$$

Here functions $f_i^{(\zeta)}(x,t)$, $i = \overline{1,n}$, $\zeta = 1,2$, are continuous.

Assumption 1. *Each function $f_i : [0,1] \times [0,p] \to \mathbb{R}$, $i = \overline{1,n}$, is such that for any numbers $0 \le \tau < \nu$ and for any continuous functions $\varrho_i^{(\zeta)} : [\tau,\nu] \to \mathbb{R}$, $\zeta = 1,2$, $\mu_i : [0,1] \to \mathbb{R}$ such that the matching condition $\varrho_i^{(1)}(\tau) = \mu_i(0)$, $\varrho_i^{(2)}(\tau) = \mu_i(1)$ is satisfied, the first boundary value problem*

$$\frac{\partial Q_i(x,t)}{\partial t} = \frac{\partial^2 Q_i(x,t)}{\partial x^2} + f_i(x,t), \tag{2.6}$$

$$Q_i(0,t) = \varrho_i^{(1)}(t), \quad Q_i(1,t) = \varrho_i^{(2)}(t), \quad \tau \le t \le \nu; \tag{2.7}$$

$$Q_i(x,\tau) = \mu_i(x), \quad 0 \le x \le 1 \tag{2.8}$$

has a unique solution $Q_i(x,t)$ continuous for $0 \le x \le 1$, $\tau \le t \le \nu$.

Let numbers α_s, $s = \overline{1,r}$, and $\varepsilon \ge 0$ be such that $\alpha_{s+1} - \alpha_s = \Delta > 0$, $s = \overline{1, r-1}$ and $\Delta > 2\varepsilon$, and vector $\overline{\lambda} = (\lambda_1, \dots, \lambda_n)^* \in \mathbb{R}^n$ such that $\lambda_i > 0$, $i = \overline{1,n}$, be given. The goal of choosing control $\overline{\xi}(t)$ in (2.2), (2.3) is to implement the inclusion

$$\sum_{i=1}^{n} \lambda_i \int_0^1 T_i(x,p)\sigma_i(x)dx \in Z(\varepsilon) = \bigcup_{s=\overline{1,r}} [\alpha_s - \varepsilon, \alpha_s + \varepsilon] \tag{2.9}$$

for any continuous functions $T_i(1,t)$ (2.4), $i = \overline{k+1,l}$, and for any continuous functions $f_i(x,t)$ (2.5), $i = \overline{1,n}$, satisfying Assumption 1.

Here continuous functions $\sigma_i : [0,1] \to \mathbb{R}$, $i = \overline{1,n}$ are given and satisfy the conditions

$$\sigma_i(0) = \sigma_i(1) = 0. \tag{2.10}$$

3. Problem formalization

Let us describe an admissible rule for choosing control $\overline{\xi}(t)$. It means that for each time moment $0 \le \nu < p$ and for each admissible temperature distribution

$$\overline{T}(x,\nu) = (T_1(x,\nu), T_2(x,\nu), \dots, T_n(x,\nu))$$

at this time moment, a measurable vector-function $\overline{\xi}(t)$ such that $\overline{\xi} : [\nu,p] \to U$ is choosing. We will denote such a rule as

$$\overline{\xi}(t) = N(t, \overline{T}(\cdot,\nu)), \quad t \in [\nu,p]. \tag{3.1}$$

Fix a partition $\omega : 0 = t_0 < t_1 < \ldots < t_j < t_{j+1} < \ldots < t_{m+1} = p$ of the segment $[0,p]$ with diameter

$$d(\omega) = \max_{0 \le j \le m} (t_{j+1} - t_j).$$

Let the temperature distribution $\overline{T}^{(\omega)}(x,t_j)$, $0 \le x \le 1$ be realized at time moment t_j, $j = \overline{0,m}$. Denote $\overline{\xi}^{(j)}(t) = N(t, \overline{T}^{(\omega)}(\cdot,t_j))$, $t \in [t_j,p]$. Let continuous functions (2.4) $T_i(1,t) = \varrho_i^{(2)}(t)$ for $t_j \le t \le t_{j+1}$, $i = \overline{k+1,l}$, for which $\varrho_i^{(2)}(t_j) = T_i^{(\omega)}(1,t_j)$, and continuous functions $f_i(x,t)$ (2.5), $i = \overline{1,n}$, for $t_j \le t \le t_{j+1}$, $0 \le x \le 1$, be realized.

We denote by $T_i^{(\omega)}(x,t)$ for $0 \le x \le 1$, $t_j \le t \le t_{j+1}$ the solution $Q_i(x,t)$ of the problem (2.6)–(2.8) for $\tau = t_j$, $\nu = t_{j+1}$ and for the following initial and boundary conditions:

$$\beta_i(x) = T_i^{(\omega)}(x,t_j), \quad x \in [0,1]; \tag{3.2}$$

$$Q_i(0,t) = T_i^{(\omega)}(0,t_j) + \int_{t_j}^{t} (a_i^{(1)}(r) + a_i^{(2)}(r) G_i^{(1)} \overline{\xi}^{(j)}(r)) dr, \quad t \in [t_j, t_{j+1}]; \tag{3.3}$$

$$Q_i(1,t) = T_i^{(\omega)}(1,t_j) + \int_{t_j}^{t} (b_i^{(1)}(r) + b_i^{(2)}(r) G_i^{(2)} \overline{\xi}^{(j)}(r)) dr, \quad i = \overline{1,k}, \tag{3.4}$$

$$Q_i(1,t) = T_i(1,t), \quad i = \overline{k+1,n}, \quad t \in [t_j, t_{j+1}]. \tag{3.5}$$

Definition 1. *We say that control of the form (3.1) guarantees the fulfilment of the stated goal (2.9), if for any number $\gamma \in (\varepsilon, \Delta/2)$ there exists a number $\delta > 0$ such that for any partition ω with diameter $d(\omega) < \delta$, for any continuous functions $f_i(x,t)$ (2.5), $i = \overline{1,n}$, that satisfy Assumption 1, and for any continuous functions $T_i(1,t)$ (2.4), $i = \overline{k+1,l}$, the inclusion*

$$\sum_{i=1}^{n} \lambda_i \int_0^1 T_i^{(\omega)}(x,p) \sigma_i(x) dx \in Z(\gamma) = \bigcup_{s=\overline{1,r}} [\alpha_s - \gamma, \alpha_s + \gamma] \tag{3.6}$$

holds.

Note that when inequality $\gamma < \Delta/2$ is satisfied, segments $[\alpha_s - \gamma, \alpha_s + \gamma]$, $s = \overline{1,r}$, do not intersect.

4. Reduction to a one-dimensional problem

Let us denote by $\psi_i(x,\tau)$ for $0 \le x \le 1$, $0 \le \tau \le p$ solutions of the following first boundary value problems

$$\frac{\partial \psi_i(x,\tau)}{\partial \tau} = \frac{\partial^2 \psi_i(x,\tau)}{\partial x^2}, \quad \psi_i(x,0) = \sigma_i(x), \quad \psi_i(0,\tau) = \psi_i(1,\tau) = 0, \quad i = \overline{1,n}. \tag{4.1}$$

Equality (2.6) implies that the matching conditions at the ends of the segment in problems (2.9) are satisfied.

Using the conditions (2.4), it can be shown that [15]

$$\begin{gathered} \left\{ \int_0^1 f_i(x,t) \psi_i(x,p-t) dx : f_i^{(1)}(x,t) \le f_i(x,t) \le f_i^{(2)}(x,t) \right\} = \\ = \{ c_i^{(1)}(t) + c_i^{(2)}(t) s_i(t) : |s_i(t)| \le 1 \}, \quad i = \overline{1,n}, \end{gathered} \tag{4.2}$$

where

$$c_i^{(1)}(t) = \frac{1}{2} \int_0^1 (f_i^{(1)}(x,t) + f_i^{(2)}(x,t)) \psi_i(x,p-t) dx,$$

$$c_i^{(2)}(t) = \frac{1}{2} \int_0^1 (f_i^{(2)}(x,t) - f_i^{(1)}(x,t)) |\psi_i(x,p-t)| dx.$$

Note that the functions $c_i^{(1)}(t)$ and $c_i^{(2)}(t)$, $i = \overline{1,n}$, are continuous for $0 \le t \le p$ and $c_i^{(2)}(t) \ge 0$.

The inequalities (2.4) imply

$$\{T_i(1,t)\} = \left\{ \frac{\beta_i^{(1)}(t) + \beta_i^{(2)}(t)}{2} + \frac{\beta_i^{(2)}(t) - \beta_i^{(1)}(t)}{2} \widehat{\eta}_i(t) : \ |\widehat{\eta}_i(t)| \le 1 \right\} \tag{4.3}$$

for $i=\overline{k+1,l}$.

Introduce new variables

$$
\begin{aligned}
y_i(t) = \int_0^1 T_i(x,t)\psi_i(x,p-t)dx + T_i(0,t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr+\\
+\int_t^p \left(a_i^{(1)}(\tau)\int_\tau^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr + c_i^{(1)}(\tau)\right)d\tau - \theta_i(t), \quad i=\overline{1,n},
\end{aligned}
\tag{4.4}
$$

where

$$
\theta_i(t) = T_i(1,t)\int_t^p \frac{\partial\psi_i(1,p-r)}{\partial x}dr + \int_t^p b_i^{(1)}(\tau)\int_\tau^p \frac{\partial\psi_i(1,p-r)}{\partial x}drd\tau, \quad i=\overline{1,k},
$$

$$
\theta_i(t) = \int_t^p \left(\frac{\beta_i^{(1)}(\tau)+\beta_i^{(2)}(\tau)}{2}\frac{\partial\psi_i(1,p-\tau)}{\partial x}\right)d\tau, \quad i=\overline{k+1,l},
$$

$$
\theta_i(t) = \int_t^p \left(T_i(1,\tau)\frac{\partial\psi_i(1,p-\tau)}{\partial x}\right)d\tau, \quad i=\overline{l+1,n}.
$$

We fix a partition ω of the segment $[0,p]$ and a control (3.1). Let us substitute the realized functions $T_i^{(\omega)}(x,t)$, $i=\overline{1,n}$, into formula (4.4). Further, taking into account formulas (2.1), (3.2)–(3.5) and (4.1)–(4.3), we obtain

$$
\begin{aligned}
\dot{y}_i^{(\omega)}(t) = \left(a_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr\right)G_i^{(1)}\overline{\xi}^{(j)}(t)-\\
-\left(b_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(1,p-r)}{\partial x}dr\right)G_i^{(2)}\overline{\xi}^{(j)}(t) + c_i^{(2)}(t)s_i(t), \quad i=\overline{1,k},
\end{aligned}
\tag{4.5}
$$

$$
\begin{aligned}
\dot{y}_i^{(\omega)}(t) = \left(a_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr\right)G_i^{(1)}\overline{\xi}^{(j)}(t)-\\
-\left(\frac{\beta_i^{(2)}(t)-\beta_i^{(1)}(t)}{2}\frac{\partial\psi_i(1,p-t)}{\partial x}\right)\widehat{\eta}_i(t) + c_i^{(2)}(t)s_i(t), \quad i=\overline{k+1,l},
\end{aligned}
\tag{4.6}
$$

$$
\dot{y}_i^{(\omega)}(t) = \left(a_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr\right)G_i^{(1)}\overline{\xi}^{(j)}(t) + c_i^{(2)}(t)s_i(t), \quad i=\overline{l+1,n}.
\tag{4.7}
$$

Next, we rewrite (4.5)–(4.7) in the matrix form

$$
\dot{\overline{y}}^{(\omega)}(t) = -A(t)\overline{\xi}^{(j)}(t) + B(t)\overline{\eta}(t), \quad \overline{\xi}^{(j)}(t)\in U, \quad \overline{\eta}(t)\in\Pi(n).
\tag{4.8}
$$

Here

$$
\overline{y}^{(\omega)}(t) = \left(y_1^{(\omega)}(t), y_2^{(\omega)}(t),\dots,y_n^{(\omega)}(t)\right)^*;
$$

$$
\Pi(n) = \{\overline{s}=(s_1,s_2,\dots,s_n)^*\in\mathbb{R}^n : |s_i|\le 1, i=\overline{1,n}\};
$$

$$
A_i(t) = -\left(a_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr\right)G_i^{(1)} + \left(b_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(1,p-r)}{\partial x}dr\right)G_i^{(2)}
$$

for $i=\overline{1,k}$,

$$
A_i(t) = -\left(a_i^{(2)}(t)\int_t^p \frac{\partial\psi_i(0,p-r)}{\partial x}dr\right)G_i^{(1)}, \quad \text{for} \quad i=\overline{k+1,n};
$$

$$
\begin{aligned}
B(t) = \operatorname{diag}\Big\{c_1^{(2)}(t),\dots,c_k^{(2)}(t), c_{k+1}^{(2)}(t) + \frac{\beta_{k+1}^{(2)}(t)-\beta_{k+1}^{(1)}(t)}{2}\left|\frac{\partial\psi_{k+1}(1,p-t)}{\partial x}\right|,\\
\dots, c_l^{(2)}(t) + \frac{\beta_l^{(2)}(t)-\beta_l^{(1)}(t)}{2}\left|\frac{\partial\psi_l(1,p-t)}{\partial x}\right|, c_{l+1}^{(2)}(t),\dots,c_n^{(2)}(t)\Big\}.
\end{aligned}
$$

Denote by $\langle\cdot,\cdot\rangle$ the operation of the scalar product of two vectors. Define

$$a_-(t) = \min_{\overline{\xi}\in U}\langle\overline{\lambda}, A(t)\overline{\xi}\rangle, \quad a_+(t) = \max_{\overline{\xi}\in U}\langle\overline{\lambda}, A(t)\overline{\xi}\rangle, \quad b(t) = \max_{\overline{\eta}\in\Pi(n)}\langle\overline{\lambda}, B(t)\overline{\eta}\rangle.$$

Note that these functions are continuous.

Then the connectedness of the compact sets U, $\Pi(n)$ and the symmetry of $\Pi(n)$ imply

$$\langle\overline{\lambda}, A(t)\overline{\xi}\rangle = \frac{1}{2}(a_+(t)+a_-(t)) + a(t)u, \quad |u|\le 1, \quad a(t) = \frac{1}{2}(a_+(t)-a_-(t)) \ge 0; \tag{4.9}$$

$$\langle\overline{\lambda}, B(t)\overline{\eta}\rangle = b(t)v, \quad |v|\le 1. \tag{4.10}$$

We introduce a new one-dimensional variable

$$z = \langle\overline{\lambda}, \overline{y}\rangle. \tag{4.11}$$

Taking into account (4.11), we obtain a polygonal line $z^{(\omega)}(t)$, which satisfies the equality

$$z^{(\omega)}(p) = \sum_{i=1}^{n}\lambda_i\int_0^1 T_i^{(\omega)}(x,p)\sigma_i(x)dx.$$

It follows that inclusion (2.9) takes the form

$$z^{(\omega)}(p) \in Z(\gamma). \tag{4.12}$$

Differentiate z, taking into account formulas (4.8)–(4.10). Taking the uncertain function v as a control of the second player, we obtain the following one-dimensional differential game

$$\dot z^{(\omega)}(t) = -a(t)u + b(t)v, \quad |u|\le 1, \quad |v|\le 1, \quad z(p)\in Z(\varepsilon). \tag{4.13}$$

5. Termination conditions

Define function

$$g(t) = \int_t^p (a(r)-b(r))dr$$

for $t\le p$ and denote

$$q_1(\varepsilon) = \inf\left\{t<p : \varepsilon + g(\tau) < \Delta - \varepsilon - g(\tau) \text{ for all } t<\tau\le p\right\},$$
$$q_2(\varepsilon) = \inf\left\{t<p : 0 \le \varepsilon + g(\tau) \text{ for all } t<\tau\le p\right\},$$
$$q_3(\varepsilon) = \inf\left\{t<p : \alpha_1 - \varepsilon - g(\tau) \le \alpha_r + \varepsilon + g(\tau) \text{ for all } t<\tau\le p\right\}.$$

Let us define the set $W(t,\varepsilon)$ for $t\le p$ as follows:

$$W(t,\varepsilon) = \begin{cases} \bigcup\limits_{s=\overline{1,r}} [\alpha_s - \varepsilon - g(t), \alpha_s + \varepsilon + g(t)] & \text{for} \quad \max(q_1(\varepsilon), q_2(\varepsilon)) \le t \le p, \\ [\alpha_1 - \varepsilon - g(t), \alpha_r + \varepsilon + g(t)] & \text{for} \quad q_3(\varepsilon)\le t < q_1(\varepsilon), \quad q_2(\varepsilon) < q_1(\varepsilon), \\ \varnothing & \text{for} \quad \max(t, q_1(\varepsilon)) < q_2(\varepsilon) \text{ or } t < q_3(\varepsilon). \end{cases} \tag{5.1}$$

Here $\varnothing$ denotes the empty set.

Theorem 1. *Let the initial temperature distributions $T_i(x,0) = g_i(x)$ be such that the inclusion*

$$z(0) \in W(0,\varepsilon) \tag{5.2}$$

holds. Then there exists a control $\overline{\xi}$ that guarantees the fulfillment of the stated goal (2.9) for any unknown functions (2.4), (2.5).

P r o o f. *Case 1.* Let $\max(q_1(\varepsilon), q_2(\varepsilon)) \le 0 \le p$. Then, according to (5.1), inclusion (5.2) implies conditions

$$-g(\tau) \le \varepsilon \quad \text{for all} \quad 0 < \tau \le p, \quad z(0) \in [\alpha_s - \varepsilon - g(0), \alpha_s + \varepsilon + g(0)] \tag{5.3}$$

for some $s \in \overline{1, r}$.

Let's make a change of variable $z_* = z - \alpha_s$ and rewrite (5.3) as follows

$$F(z_*(0)) \le \varepsilon, \tag{5.4}$$

where

$$F(z) = \max\left(|z| - g(0), -\min_{0 \le \tau \le p} g(\tau)\right).$$

Define $\overline{\xi}_0(t) = N(t, \overline{T}(\cdot, \tau))$, $t \in [\tau, p]$ as the solution of problem

$$\langle \overline{\lambda}, A(t)\overline{\xi}(t)\rangle \text{sign } z_*(t) \to \max_{\overline{\xi}(t) \in U}.$$

Here and henceforth sign $0 = 1$.

Next, taking into account (4.9), we substitute the control $\xi_0(t)$ into (4.13) with $z = z_*$. We get that

$$\dot{z}_*^{(\omega)}(t) = -a(t)\text{sign } z_*(t_j) + b(t)v(t), \quad |v(t)| \le 1. \tag{5.5}$$

Here $v(t)$ satisfies the conditions: $|v(t)| \le 1$ if $b(t) = 0$, and

$$v(t) = \frac{\langle \overline{\lambda}, B(t)\overline{\eta}(t)\rangle}{b(t)} \quad \text{for} \quad b(t) > 0.$$

Each measurable function $v : [0, p] \to [-1, 1]$ with $z_*^{(\omega)}(0) = z_*(0)$ defines a polygonal line $z_*^{(\omega)}(t)$ satisfying equation (5.5). The family of these polygonal lines defined on the interval $[0, p]$ is uniformly bounded and equicontinuous [16, p. 46]. According to Arzel's theorem [7, p. 104] from any sequence of these polygonal lines we can select a subsequence uniformly converging on the segment $[0, p]$. The limit function $z_*(t)$ satisfies [16, Theorem 8.1] the inequality

$$|z_*(p)| \le F(z_*(0)). \tag{5.6}$$

Fix a number $\gamma \in (\varepsilon, \Delta/2)$. Let us show that there exists a number $\delta > 0$ such that inclusion (4.12) holds for any polygonal line $z^{(\omega)}(t)$ with partition diameter $d(\omega) < \delta$.

Indeed, let us assume the opposite. Then there exists a sequence of polygonal lines $z^{(\omega_k)}(t)$ with diameters $d(\omega_k) \to 0$ such that $z^{(\omega_k)}(p) \notin Z(\gamma)$ or what is the same

$$|z^{(\omega_k)}(p) - \alpha_s| > \gamma$$

for all $s \in \overline{1, r}$. We can assume that the functions $z^{(\omega_k)}(t)$ converge on the segment $[0, p]$ uniformly to the function $z(t)$ (otherwise we move on to a subsequence). Then

$$|z(p) - \alpha_s| \ge \gamma$$

for all $s \in \overline{1, r}$. This inequality contradicts inequalities (5.4) and (5.6).

Case 2. Let $q_3(\varepsilon) \le 0 < q_1(\varepsilon)$, $q_2(\varepsilon) < q_1(\varepsilon)$. Then, according to (5.1), inclusion (5.2) implies conditions

$$\alpha_1 - \varepsilon - g(\tau) \le \alpha_r + \varepsilon + g(\tau) \text{ for all } 0 < \tau \le p, \quad z(0) \in \big[\alpha_1 - \varepsilon - g(0), \alpha_r + \varepsilon + g(0)\big]. \tag{5.7}$$

Define $\overline{\xi}^0(t) = N(t, \overline{T}(\cdot, \tau))$, $t \in [\tau, p]$ as the solution of problem

$$\langle \overline{\lambda}, A(t)\overline{\xi}(t)\rangle \mathrm{sign}(z(t) - 0.5(\alpha_1 + \alpha_r)) \to \max_{\overline{\xi}(t)\in U}.$$

Taking into account (4.9), we substitute the control $\xi^0(t)$ into (4.13)

Next, reasoning by analogy with case 1 of the proof and relying on the results of work [17], it can be shown that when conditions (5.7) are satisfied, the limit function $z(t)$ satisfies the inclusion

$$z(q_1(\varepsilon)) \in \left[\alpha_1 - \varepsilon - g(q_1(\varepsilon)), \alpha_r + \varepsilon + g(q_1(\varepsilon))\right].$$

According to the definition of $q_1(\varepsilon)$, equality

$$\left[\alpha_1 - \varepsilon - g(q_1(\varepsilon)), \alpha_r + \varepsilon + g(q_1(\varepsilon))\right] = \bigcup_{s=\overline{1,r}} \left[\alpha_s - \varepsilon - g(q_1(\varepsilon)), \alpha_s + \varepsilon + g(q_1(\varepsilon))\right]$$

holds, and, therefore,

$$z(q_1(\varepsilon)) \in \left[\alpha_s - \varepsilon - g(q_1(\varepsilon)), \alpha_s + \varepsilon + g(q_1(\varepsilon))\right] \tag{5.8}$$

holds for some $s \in \overline{1, r}$.

Since $q_2(\varepsilon) < q_1(\varepsilon)$, then inequality $-g(\tau) \le \varepsilon$ holds for all $q_1(\varepsilon) < \tau \le p$. From here and from (5.8) we fall into the condition of case 1.

□

Now we consider the case when function

$$\overline{\eta}_*(t) = \mathrm{sign}(z^{(\omega)}(t_j) - 0.5(\alpha_1 + \alpha_r))(1, 1, \ldots, 1)^*$$

is realized in (4.8) for $t_j < t < t_{j+1}$.

Taking (4.10) into account, let us substitute this function $\overline{\eta}_*(t)$ into (4.13). We get that

$$\dot{z}^{(\omega)}(t) = -a(t)u_j(t) + b(t)\,\mathrm{sign}\,(z^{(\omega)}(t_j) - 0.5(\alpha_1 + \alpha_r)), \tag{5.9}$$

where

$$a(t)u_j(t) = \langle \overline{\lambda}, A(t)\overline{\xi}^{(j)}(t)\rangle.$$

Choosing arbitrary measurable functions $\xi^{(j)}(t) \in U$ and solving equation (5.9) with $z^{(\omega)}(0) = z(0)$, we obtain a family of polygonal lines $z^{(\omega)}(t)$.

Theorem 2. *Let at least one of the following inequalities be satisfied:*

$$z(0) < \alpha_1 - \gamma - g(0), \quad \alpha_r + \gamma + g(0) < z(0), \quad 0 < q_3(\gamma). \tag{5.10}$$

Then there exists a number $\delta > 0$ such that $z^{(\omega)}(p) \notin Z(\gamma)$ for any polygonal line $z^{(\omega)}(t)$ (5.9) with partition diameter $d(\omega) < \delta$.

P r o o f. Let's assume the opposite. Let's take a sequence of numbers $\delta_k \to 0$. Then there exists a sequence of polygonal lines $z^{(\omega_k)}(t)$ with diameter $d(\omega_k) < \delta_k$ and $z^{(\omega_k)}(t) \in Z(\gamma)$. The family of polygonal lines (5.9) with $z^{(\omega)}(0) = z(0)$ satisfies the conditions of Arzela's theorem. Passing, if necessary, to a subsequence, we can assume that the sequence of polygonal lines $z^{(\omega_k)}(t)$ converges to $z(t)$ uniformly.

Let's make a change of variables

$$\tilde{z} = z - 0.5(\alpha_1 + \alpha_r)$$

and rewrite conditions (5.10) in the following form:

$$|\tilde{z}(0)| > \gamma + g(0) + 0.5(\alpha_r - \alpha_1) \quad \text{or} \quad \gamma + 0.5(\alpha_r - \alpha_1) + \min_{0\le\tau\le p} g(\tau) < 0.$$

Hence,

$$0 \le \gamma < F(\tilde{z}(0)) - 0.5(\alpha_r - \alpha_1). \tag{5.11}$$

On the other hand, the limit function satisfies the inequality $|\tilde{z}(p)| \ge F(\tilde{z}(0))$ [16, Theorem 8.2]. From this and from (5.11) we obtain that

$$|\tilde{z}^{(\omega_k)}(p)| > \gamma + 0.5(\alpha_r - \alpha_1)$$

for any sufficiently large number k. After a reverse change of variables, we obtain one of the inequalities

$$z^{(\omega_k)}(p) < \alpha_1 - \gamma \quad \text{or} \quad z^{(\omega_k)}(p) > \alpha_r + \gamma.$$

Thus, we get a contradiction.

□

Next, consider the case when the function

$$\overline{\eta}^*(t) = -\mathrm{sign}\left(z^{(\omega)}(t_j) - 0.5(\alpha_s + \alpha_{s+1})\right)(1, 1, \ldots, 1)^*$$

is realized in (4.8) for $t_j < t < t_{j+1}$. Here, number $s \in \overline{1, r-1}$ can be calculated as the solution of the minimization problem

$$\min_{s\in\overline{1,r-1}} |z(0) - 0.5(\alpha_s + \alpha_{s+1})|.$$

Taking (4.10) into account, let us substitute this function $\overline{\eta}^*(t)$ into (4.13). We obtain

$$\dot{z}^{(\omega)}(t) = -a(t)u_j(t) - b(t)\,\mathrm{sign}\,(z^{(\omega)}(t_j) - 0.5(\alpha_s + \alpha_{s+1})). \tag{5.12}$$

Further, we define a family of polygonal lines $z^{(\omega)}(t)$ for equation (5.12) by analogy with (5.9).

Theorem 3. *Let the following inequalities be satisfied:*

$$\alpha_s + \gamma + g(0) < z(0) < \alpha_{s+1} - \gamma - g(0), \quad q_1(\gamma) < 0. \tag{5.13}$$

Then there exists a number $\delta > 0$ such that $z^{(\omega)}(p) \notin Z(\gamma)$ for any polygonal line $z^{(\omega)}(t)$ (5.12) with partition diameter $d(\omega) < \delta$.

P r o o f. Let's assume the opposite. By analogy with the proof of Theorem 2, we construct a sequence of polygonal lines $z^{(\omega_k)}(t)$ that converges to $z(t)$ uniformly.

Let's introduce the variable

$$\widehat{z} = z - 0.5(\alpha_s + \alpha_{s+1})$$

and write inequalities (5.13) as follows:

$$|\widehat{z}(0)| < 0.5\Delta - \gamma - g(0), \quad 0 < 0.5\Delta - \gamma - \max_{0\le\tau\le p} g(\tau).$$

From here we obtain

$$G(\widehat{z}(0)) = \max\left\{|\widehat{z}(0)| - \int_0^p (b(r) - a(r))dr, - \min_{0\le\tau\le p} \int_\tau^p (b(r) - a(r))dr\right\} < 0.5\Delta - \gamma \qquad (5.14)$$

On the other hand, applying [16, Theorem 8.1] from the point of view of the second player (in variables $\widehat{z}$ the roles of the players change, and the second player becomes the pursuer), we obtain that the limit function satisfies the inequality $|\widehat{z}(p)| \le G(\widehat{z}(0))$. From this and from (5.14) we obtain that

$$|\widehat{z}^{(\omega_k)}(p)| < 0.5\Delta - \gamma$$

for all sufficiently large numbers k. After a reverse change of variables, we obtain the inequalities

$$\alpha_s + \gamma < z^{(\omega_k)}(p) < \alpha_{s+1} - \gamma.$$

Thus, we get a contradiction. □

Remark 1. Let $q_3(\gamma) \le 0 < q_1(\gamma) < q_2(\gamma)$. Let's substitute an arbitrary function $v(t, z)$ ($|v(t, z)| \le 1$) into (4.13) and, by analogy with the proof of Theorem 2, define $z(t)$ as the uniform limit of a sequence of polygonal lines. Then there exists a time moment $t_* \in (q_1(\gamma), q_2(\gamma))$ such that $\gamma + g(\tau) < 0$ for all $\tau \in [t_*, q_2(\gamma))$. Then one of the following conditions is satisfied:

$$\alpha_s + \gamma + g(t_*) < z(t_*) < \alpha_{s+1} - \gamma - g(t_*) \quad \text{for some} \quad s \in \overline{1, r-1};$$
$$z(t_*) < \alpha_1 - \gamma - g(t_*); \quad \alpha_r + \gamma + g(t_*) < z(t_*).$$

From here we find ourselves in the conditions of Theorem 2 or 3 with the initial time moment t_*.

Corollary 1. *Theorems 1, 2, 3 and Remark 1 imply that the set $W(0, \varepsilon)$ determines the necessary and sufficient termination conditions in the differential game* (4.13).

6. Conclusion

This paper considers the problem of controlling a parabolic system that describes the heating of a given number of rods, with a non-convex one-dimensional terminal set, which is defined as the union of a finite number of disjoint segments of equal length. Necessary and sufficient conditions have been found under which there exists a control (3.1) that guarantees the achievement of the stated goal (2.9) for all continuous functions (2.4) and for all density functions of internal heat sources (2.5) that satisfy the Assumption 1.

In the future, it is planned to consider a version of this problem with an arbitrary n-dimensional non-convex terminal set. This will require the development of approximate algorithms for solving differential games: constructing a solvability set and restoring the corresponding control $\overline{\xi}$.

REFERENCES

1. Barseghyan V., Solodusha S. The problem of boundary control of the thermal process in a rod. *Mathematics,* 2023. Vol. 11, No. 13. Art. no. 2881. DOI: 10.3390/math11132881
2. Casas E., Yong J. Optimal control of a parabolic equation with memory. *ESAIM Control Optim. Calc. Var.*, 2023. Vol. 29. Art. no. 23. DOI: 10.1051/cocv/2023013
3. Dai J., Ren B. UDE-based robust boundary control of heat equation with unknown input disturbance. *IFAC-PapersOnLine,* 2017. Vol. 50, No. 1. P. 11403–11408. DOI: 10.1016/j.ifacol.2017.08.1801

4. Egorov A. I. *Optimal'noe upravlenie teplovymi i diffuzionnymi protsessami* [Optimal Control of Thermal and Diffusion Processes]. Moscow: Nauka, 1978. 464 p. (in Russian)
5. Feng H., Xu Ch.-Z., Yao P.-F. Observers and disturbance rejection control for a heat equation. *IEEE Trans. Automat. Control*, 2020. Vol. 65, No. 11. P. 4957–4964. DOI: 10.1109/TAC.2020.3022849
6. Izmest'ev I. V., Ukhobotov V. I. On one problem of controlling the heating of a rod system under uncertainty. *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2022. Vol. 32, Iss. 4. P. 546–556. DOI: 10.35634/vm220404 (in Russian)
7. Kolmogorov A. N., Fomin S. V. *Elementy teorii funktsii i funktsional'nogo analiza* [Elements of the Theory of Functions and Functional Analysis]. Moscow: Nauka, 1972. 496 p. (in Russian)
8. Korotkii A. I., Osipov Iu. S. Approximation in problems of position control of parabolic systems. *J. Appl. Math. Mech.*, 1978. Vol. 42, No. 4. P. 631–637. DOI: 10.1016/0021-8928(78)90004-7
9. Krasovskii N. N. *Upravlenie dinamicheskoi sistemoi* [Control of a Dynamical System]. Moscow: Nauka, 1985. 520 p. (in Russian)
10. Lohéac J. Nonnegative boundary control of 1D linear heat equations. *Vietnam J. Math.*, 2021. Vol. 49, No. 3. P. 845–870. DOI: 10.1007/s10013-021-00497-5
11. Maksimov V. I. On the reconstruction of an input disturbance in a reaction-diffusion system. *Comput. Math. Math. Phys.*, 2023. Vol. 63, No. 6. P. 990–1000. DOI: 10.1134/S0965542523060143
12. Okhezin S. P. Differential encounter-evasion game for a parabolic system under integral constraints on the player's controls. *J. Appl. Math. Mech.*, 1977. Vol. 41, No. 2. P. 194–201. DOI: 10.1016/0021-8928(77)90002-8
13. Osipov Iu. S. Position control in parabolic systems. *J. Appl. Math. Mech.*, 1977. Vol. 41, No. 2. P. 187–193. DOI: 10.1016/0021-8928(77)90001-6
14. Osipov Yu. S., Okhezin S. P. On the theory of differential games in parabolic systems. *Dokl. Akad. Nauk SSSR*, 1976. Vol. 226, No. 6. P. 1267–1270.
15. Ukhobotov V. I., Izmest'ev I. V. A control problem for a rod heating process with unknown temperature at the right end and unknown density of the heat source. *Trudy Inst. Mat. Mekh. UrO RAN,* 2019. Vol. 25, No. 1. P. 297–305. DOI: 10.21538/0134-4889-2019-25-1-297-305 (in Russian)
16. Ukhobotov V. I. *Metod odnomernogo proektirovaniya v lineinykh differentsial'nykh igrakh s integral'nymi ogranicheniyami* [Method of One-Dimensional Projecting in Linear Differential Games with Integral Constraints]. Chelyabinsk: Chelyabinsk State University, 2005. 124 p. (in Russian)
17. Ukhobotov V. I. One type differential games with convex goal. *Trudy Inst. Mat. Mekh. UrO RAN,* 2010. Vol. 16, No. 5. P. 196–204. (in Russian)
18. Wang S., Qi J., Diagne M. Adaptive boundary control of reaction-diffusion PDEs with unknown input delay. *Automatica,* 2021. Vol. 134. Art. no. 109909. DOI: 10.1016/j.automatica.2021.109909
19. Zheng G., Li J. Stabilization for the multi-dimensional heat equation with disturbance on the controller. *Automatica,* 2017. Vol. 82. P. 319–323. DOI: 10.1016/j.automatica.2017.04.011

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 132–140

DOI: 10.15826/umj.2023.2.011

ON TWO-SIDED UNIDIRECTIONAL MEAN VALUE INEQUALITY IN A FRÉCHET SMOOTH SPACE[1]

Dmitry V. Khlopin

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

khlopin@imm.uran.ru

Abstract: The paper is devoted to a new unidirectional mean value inequality for the Fréchet subdifferential of a continuous function. This mean value inequality finds an intermediate point and localizes its value both from above and from below; for this reason, the inequality is called two-sided. The inequality is considered for a continuous function defined on a Fréchet smooth space. This class of Banach spaces includes the case of a reflexive space and the case of a separable Asplund space. As some application of these inequalities, we give an upper estimate for the Fréchet subdifferential of the upper limit of continuous functions defined on a reflexive space.

Keywords: Smooth Banach space, Fréchet subdifferential, Unidirectional mean value inequality, Upper limit of continuous functions.

1. Introduction

Consider the following mean value inequalities.

Proposition 1 [12, Theorems 2.1 and 2.2]. *Let a scalar function f be defined and lower semicontinuous on a Fréchet smooth Banach space $\mathbb{X}$. Let points $\check{u}$ and $\check{v}$ in $\mathbb{X}$ be given. Then, for arbitrary numbers $\check{s} < f(\check{v}) - f(\check{u})$ and $\check{\varkappa} > 0$, there exist a point $z_- \in [\check{u};\check{v}] + \check{\varkappa}B$ and a Fréchet subgradient $\zeta_- \in \hat{\partial} f(z_-)$ such that*

$$\check{s} < \zeta_-(\check{v} - \check{u}) \quad \text{and} \quad f(z_-) < f(\check{u}) + \max(0, \check{s}) + \check{\varkappa}. \tag{1.1}$$

Furthermore, if f is continuous, there are a point $z_+ \in [\check{u};\check{v}] + \check{\varkappa}B$ and a Fréchet subgradient $\zeta_+ \in \hat{\partial} f(z_+)$ such that

$$\check{s} < \zeta_+(\check{v} - \check{u}) \quad \text{and} \quad f(z_+) > f(\check{v}) - \max(0, \check{s}) - \check{\varkappa}. \tag{1.2}$$

Note that inequalities (1.1) and (1.2) are similar. This suggests that, in the case of continuity of f, it is possible to get a common point $z_+ = z_-$ such that the value $f(z)$ is localized from both above and below. Proving the corresponding two-sided unidirectional mean inequality is the primary goal of this paper.

As part of the historical background, note that the existence of a pair (z_-, ζ_-) satisfying inequalities like (1.1) has been widely studied (see, for example, [13, Subsect. 3.4.8] and [14, Sect. 4.4]). Unlike different variants of Lagrange's mean value theorem for certain classes of Lipschitz continuous functions, they ensure an upper bound of $f(v) - f(u)$ through some subgradient ζ. These inequalities apply to any lower semicontinuous function. Furthermore, the corresponding to the

[1]This study was funded by the RFBR and DFG (project no. 21-51-12007).

Fréchet subdifferentials unidirectional mean value inequality is equivalent to the Asplund property of a Banach space [14, 17], and therefore is equivalent to several basic principles of variational analysis [1, 18], for example, to the inspired by [16] and [5] multidirectional mean value inequality [4, Subsection 3.6.1]; for more recent references, see [10] and [8]. However, the multidirectional mean value theorem as well as the previous unidirectional mean value inequality also localizes $f(\hat{z})$ on one side only.

The rest of the paper is organized as follows. In Section 3, we will prove the desired two-sided unidirectional mean value inequality for continuous functions. Then, applying this result, in Section 4, we will show an upper estimate for the Fréchet subdifferential of the upper limit of continuous functions. But first, we will recall several elementary definitions and notions.

2. Definitions and notation

We will use elementary notions from the set-valued and variational analyses [4, 13, 15].

For a nonempty set $\mathcal{X}$ of some real Banach space $\mathbb{X}$, denote by $\operatorname{cl}\mathcal{X}$ and $\operatorname{co}\mathcal{X}$ the closure and the convex hull of $\mathcal{X}$. For a point $x \in \mathcal{X}$, the contingent (Bouligand tangent) cone to $\mathcal{X}$ at x is the set $T(x;\mathcal{X})$ of all $v \in \mathbb{X}$ such that one finds a decreasing to 0 sequence of positive t_n and a converging to v sequence of $v_n \in \mathbb{X}$ such that $x + t_n v_n \in \mathcal{X}$ for all positive integers n. For a point $x \in \mathbb{X}$, we say that $\zeta \in \mathbb{X}^*$ is a Fréchet normal to $\mathcal{X}$ at x if one has $x \in \mathcal{X}$ and

$$\limsup_{n\to\infty} \frac{\zeta(z_n - x)}{\|z_n - x\|} = 0$$

for all converging to x sequences of $z_n \in \mathcal{X}$. Denote by $\hat{N}(x;\mathcal{X})$ the set of all Fréchet normals to $\mathcal{X}$ at x.

We call a Banach space $\mathbb{X}$ Fréchet smooth if this space has an equivalent norm that is C^1-smooth off the origin. Note that any reflexive Banach space and any separable Asplund space are Fréchet smooth [4, Theorem 6.1.6]. It is worth mentioning that each Fréchet smooth space has a C^1-smooth Lipschitz function with bounded nonempty support [3, Ex. 4.3.9].

Denote by B and B^* the unit closed balls in $\mathbb{X}$ and $\mathbb{X}^*$, respectively.

Given an extended-real-valued function $g : \mathbb{X} \to \mathbb{R} \cup \{-\infty, \infty\}$, define its lower semicontinuous envelope $\operatorname{lsc} g$ by the rule:

$$\operatorname{lsc} g(x) \overset{\triangle}{=} \liminf_{\varkappa\downarrow 0} \inf_{z\in x+\varkappa B} g(z) \text{ for all } x \in \mathbb{X}.$$

Note that this function is lower semicontinuous. In addition, a function g is lower semicontinuous iff its epigraph

$$\operatorname{epi} g \overset{\triangle}{=} \{(x,a) \in \mathbb{X} \times \mathbb{R} \mid a \geq g(x)\}$$

is closed. In the case of lower semicontinuous function g, define the Fréchet subdifferential of g at x as

$$\hat{\partial} g(x) \overset{\triangle}{=} \{\zeta \in \mathbb{X}^* \mid (\zeta, -1) \in \hat{N}(x, g(x); \operatorname{epi} g)\}$$

for a point $x \in \mathbb{X}$ with finite $g(x)$; let also $\hat{\partial} g(x) = \varnothing$ if $|g(x)| = \infty$.

3. Two-sided mean value inequality

Theorem 1. *Let $\mathbb{X}$ be a Fréchet smooth space. Let a continuous function $f : \mathbb{X} \to \mathbb{R}$ and some closed interval $[u; v]$ in $\mathbb{X}$ be given. Then, for a real number $s < f(v) - f(u)$ and positive ε, there exist some point $\hat{z} \in [u; v] + \varepsilon B$ and Fréchet subgradient $\hat{\zeta} \in \hat{\partial} f(\hat{z})$ such that*

$$s < \hat{\zeta}(v - u) \quad \textit{and} \quad |f(\hat{z}) - f(u)| \leq |s| + \varepsilon. \tag{3.1}$$

P r o o f. Without loss of generality, we can assume that $u = 0$ and $f(u) = 0$. Now, the initial inequality can be written as $s < f(v)$.

Case $s < 0$. Let s be negative. Choose a positive number $\varepsilon < \min(|s|, f(v) - s)$. Define $\bar{\varepsilon} = \varepsilon/4$ and $\bar{s} \triangleq s + 3\bar{\varepsilon}$. Since $\bar{s} < 0 = f(0) < |\bar{s}|$ and $\bar{s} < f(v)$, one finds a positive number $\hat{t} < 1$ such that

$$|\bar{s}| > f(\hat{t}v) > -\hat{t}|\bar{s}| > -|\bar{s}| = \bar{s} \tag{3.2}$$

because f is continuous on $[0; v]$. For the same reason, there is a positive $\varkappa < \bar{\varepsilon}$ such that

$$|f(z) - f(0)| < \bar{\varepsilon} \quad \text{for all} \quad z \in [0; v] \cap \varkappa B. \tag{3.3}$$

We claim that there exist some $\hat{z} \in [0; \hat{t}v] + \varkappa B$ and $\hat{\zeta} \in \hat{\partial} f(\hat{z})$ such that

$$s < -|\bar{s}| < \hat{\zeta}v \quad \text{and} \quad |f(\hat{z})| < |\bar{s}| + 2\bar{\varepsilon} < |s|. \tag{3.4}$$

To this end, consider the continuous map

$$[0; \hat{t}\,] \ni \tau \mapsto h(\tau) = f(\tau v) - \tau f(\hat{t}v)/\hat{t}.$$

Since $h(\hat{t}) = h(0) = 0$ holds, due to the intermediate value theorem, there exists positive $\hat{\tau} \leq \hat{t}$ that satisfies the equality $h(\hat{\tau}) = 0$ and at least one of the following conditions:

$$\text{(i)} \quad \hat{\tau} < \varkappa; \quad \text{(ii)} \quad h|_{[0,\hat{\tau}]} \text{ is nonpositive;} \quad \text{(iii)} \quad h|_{[0,\hat{\tau}]} \text{ is nonnegative.}$$

Now, the relations $0 < \hat{\tau} \leq \hat{t} < 1$, $h(\hat{\tau}) = 0$, and (3.2) yield the inequality

$$-|\bar{s}| \leq -|\bar{s}|\hat{\tau} \overset{(3.2)}{<} \hat{\tau} f(\hat{t}v)/\hat{t} = f(\hat{\tau}v) - f(0). \tag{3.5}$$

Let us apply Proposition 1 to this inequality with

$$\check{u}_- \triangleq 0, \quad \check{v}_+ \triangleq \hat{\tau}v, \quad \check{s} \triangleq -|\bar{s}|\hat{\tau}, \quad \text{and} \quad \varkappa_n \triangleq \varkappa/n$$

for all positive integers n. This gives some $r_-, r_+ \in [0; \hat{\tau}]$, $z_-, z_+ \in \mathbb{X}$, $\zeta_- \in \hat{\partial} f(z_-)$, and $\zeta_+ \in \hat{\partial} f(z_+)$ such that

$$\begin{aligned} -|\bar{s}|\hat{\tau} &< \hat{\tau}\zeta_- v, \quad \|r_- v - z_-\| < \varkappa_n, \quad f(z_-) - f(0) < \varkappa_n, \\ -|\bar{s}|\hat{\tau} &< \hat{\tau}\zeta_+ v, \quad \|r_+ v - z_+\| < \varkappa_n, \quad f(z_+) - f(\hat{\tau}v) > -\varkappa_n. \end{aligned}$$

Next, taking into account the inequalities $\hat{\tau} > 0$ and $f(0) + \varkappa_n = \varkappa_n < \bar{\varepsilon}$, we have

$$-|\bar{s}| < \zeta_\pm v, \quad z_\pm = r_\pm v + \varkappa_n B \subset [0; v] + \varkappa B,$$
$$f(z_-) < f(0) + \varkappa_n < \bar{\varepsilon}, \quad \text{and} \quad f(z_+) > -\varkappa_n + f(\hat{\tau}v) \overset{(3.5)}{\geq} -\varkappa_n - |\bar{s}| > -\bar{\varepsilon} - |\bar{s}|. \tag{3.6}$$

Now, in the case of $\hat{\tau} < \varkappa$ (condition (i)) and in the case of nonpositive $h|_{[0;\hat{\tau}]}$ (condition (ii)), let us set $\hat{z}_n \triangleq z_+$, $\hat{r}_n \triangleq r_+$, and $\hat{\zeta}_n \triangleq \zeta_+$ for all positive integers n; and in the case of nonnegative $h|_{[0;\hat{\tau}]}$ (condition (iii)), set $\hat{z}_n \triangleq z_-$, $\hat{r}_n \triangleq r_-$, and $\hat{\zeta}_n \triangleq \zeta_-$ for all positive integers n. Then, in all these cases and for all positive integers n, we have proved the first inequality in (3.4). So, it is required to check only

$$|f(\hat{z}_n)| \leq |\bar{s}| + 2\bar{\varepsilon}$$

for at least one positive integer n.

Note that all $\hat{r}_n v$ lie in the compact set $[0;\hat{\tau}v]$. Passing to a subsequence, we can assume that this sequence converges. By $\|\hat{z}_n - \hat{r}_n v\| \to 0$, the both sequences of $\hat{z}_n$ and $\hat{r}_n v$ has the common limit. The sequences of $f(\hat{z}_n)$ and $f(\hat{r}_n v)$ are the same by the continuity of f; in particular, one finds a positive integer N such that

$$|f(\hat{z}_N) - f(\hat{r}_N v)| < \bar{\varepsilon}. \tag{3.7}$$

So, it is required to check only the inequality

$$|f(\hat{r}_N v)| < |\bar{s}| + \bar{\varepsilon}.$$

Now, in the case of nonnegative $h|_{[0;\hat{\tau}]}$ (condition (iii)), by the choice of $\hat{r}_N = r_-$ and $\hat{z}_N = z_-$, we obtain

$$0 \leq h(r_-) = f(r_- v) - r_- f(\hat{t}v)/\hat{t} \leq f(r_- v) + |f(\hat{t}v)|$$

and

$$\bar{\varepsilon} \overset{(3.6)}{>} f(\hat{z}_N) = f(z_-) \overset{(3.7)}{\geq} f(r_- v) - \bar{\varepsilon} \geq -|f(\hat{t}v)| - \bar{\varepsilon} \overset{(3.2)}{\geq} -|s| - \bar{\varepsilon}.$$

In the case $h|_{[0;\hat{\tau}]} \leq 0$ (condition (ii)), one has

$$0 \geq h(r_+) \geq f(r_+ v) - |f(\hat{t}v)|$$

and

$$-\bar{\varepsilon} - |\bar{s}| \overset{(3.6)}{<} f(\hat{z}_N) = f(z_+) \overset{(3.7)}{\leq} f(r_+ v) + \bar{\varepsilon} \leq |f(\hat{t}v)| + \bar{\varepsilon} \overset{(3.2)}{\leq} |\bar{s}| + \bar{\varepsilon}.$$

Finally, in the case $\hat{\tau} < \varkappa$ (condition (i)), (3.3) and (3.7) yield

$$|f(r_- v)| < 2\bar{\varepsilon} < |\bar{s}| + \bar{\varepsilon}.$$

Inequalities (3.4) have been proved.

Case $s \geq 0$. Assume that s is nonnegative. Recall that $s < f(v)$. Choose a positive ε such that $s + \varepsilon < f(v)$. Define

$$\bar{s} \overset{\triangle}{=} s + \varepsilon/2.$$

This entails

$$0 < \bar{s} < s + \varepsilon < f(v),$$

and one can choose positive $\bar{\varepsilon}$ such that

$$\bar{\varepsilon} + (1 + \bar{\varepsilon})^3 \bar{s} < s + \varepsilon.$$

Consider the map $\bar{f} : \mathbb{X} \times \mathbb{R} \to \mathbb{R}$ defined as

$$\bar{f}(x, r) \overset{\triangle}{=} f(x) - (1 + \bar{\varepsilon}) r \bar{s} \quad \text{for all} \quad x \in \mathbb{X}, \quad r \in \mathbb{R}.$$

Then, we have $\bar{f}(0,0) = 0$,

$$\hat{\partial} \bar{f}(x, r) = \hat{\partial} f(x) \times \{-(1 + \bar{\varepsilon})\bar{s}\},$$
$$\bar{f}(v, 1) = \bar{f}(v, 1) - \bar{f}(0, 0) = f(v) - (1 + \bar{\varepsilon})\bar{s} > \bar{s} - (1 + \bar{\varepsilon})\bar{s} = -\bar{\varepsilon}\bar{s}.$$

Since $-\bar{\varepsilon}\bar{s} < 0$, we can apply the first case of our theorem to the inequality

$$-\bar{\varepsilon}\bar{s} < \bar{f}(v, 1) - \bar{f}(0, 0).$$

Then, there exist some number $r \in [0;1]$, point

$$\bar{z} = (\hat{z}, \hat{r}) \in (rv, r) + \bar{\varepsilon}B,$$

and subgradient

$$\bar{\zeta} = (\hat{\zeta}, -(1+\bar{\varepsilon})\bar{s}) \in \hat{\partial}\bar{f}(\hat{z}, \hat{r})$$

that satisfy (3.4); i.e.,

$$-\bar{\varepsilon}\bar{s} < \hat{\zeta}v - (1+\bar{\varepsilon})\bar{s} \quad \text{and} \quad |\bar{f}(\hat{z}, \hat{r})| \leq \bar{\varepsilon}|\bar{s}| + \bar{\varepsilon}.$$

Now, the first inequality leads to $s < \bar{s} < \hat{\zeta}v$ by $s < \bar{s}$; on the other hand, the second inequality entails

$$|f(\hat{z})| = \left|\bar{f}(\hat{z}, \hat{r}) + (1+\bar{\varepsilon})\hat{r}\bar{s}\right| < \bar{\varepsilon}\bar{s} + \bar{\varepsilon} + (1+\bar{\varepsilon})|\hat{r}|\bar{s} < \bar{\varepsilon}\bar{s} + \bar{\varepsilon} + (1+\bar{\varepsilon})^2\bar{s} < (1+\bar{\varepsilon})^3\bar{s} + \bar{\varepsilon} < s + \varepsilon$$

by $|\hat{r}| < |r| + \bar{\varepsilon} \leq 1 + \bar{\varepsilon}$ and the choice of $\bar{\varepsilon}$.

The theorem is proved. □

Remark 1. As [12, Example 2.1] has shown, (1.2) can be violated if $f : \mathbb{R} \to \mathbb{R}$ is only lower semicontinuous. Therefore, the assumption of the continuity of f is essential in this theorem as well.

Remark 2. In the case of Lipschitz continuous function f, for its G-subdifferential, there exists a variant of unidirectional mean value inequality that guaranties the inclusion $z \in [u; v]$ instead of $z \in [u; v] + \varepsilon B$ (see [9, Theorem 4.70]). However, this is not true for a Fréchet subdifferential. Indeed, for the Lipschitz continuous function

$$\mathbb{R}^2 \ni (x, y) \mapsto f(x, y) \stackrel{\triangle}{=} -|x|,$$

its Fréchet subdifferential is empty on the interval $[(0,0); (0,1)]$; in particular, no Fréchet subgradient ζ satisfies (3.1).

Remark 3. It may mistakenly seem that Theorem 1 does not essentially use the asymmetry of a Fréchet subdifferential and can be directly extended to the symmetric case. Indeed, Lebourg's mean value theorem [6, Theorem 2.4] for Clarke subdifferentials, the mean value theorem [2] for MP-subdifferentials, and the symmetric subdifferential mean value theorem [13, Theorem 3.47], [14, Theorem 4.11] give the corresponding gradient ζ of f at some $\hat{z} \in [u; v]$ that satisfies the symmetric bound

$$|f(v) - f(u)| = \left|\hat{\zeta}(v - u)\right|. \tag{3.8}$$

This bound is exactly the limit of bounds

$$s_+ < \hat{\zeta}(v-u) + \varepsilon \quad \text{and} \quad -s_- < (-\hat{\zeta})(v-u) + \varepsilon$$

as $s_+ \uparrow f(v) - f(u)$, $-s_- \uparrow (-f)(u) - (-f)(v)$, and $\varepsilon \downarrow 0$. Similarly, passing to the limit in $|f(\hat{z}) - f(u)| < |s| + \varepsilon$, we could hope for the eatimate

$$|f(\hat{z}) - f(u)| \leq \left|\hat{\zeta}(v - u)\right| \tag{3.9}$$

together with (3.8). However, in the case

$$f(x) \stackrel{\triangle}{=} x(x-2) \quad \text{and} \quad [u, v] \stackrel{\triangle}{=} [0, 2],$$

inequalities (3.8) and (3.9) should give $f'(\hat{z}) = \zeta = 0$ and $|f(\hat{z})| \leq 0$; i.e., $1 = \hat{z} \in \{0, 2\}$. This contradiction negates the hope of adding two-side estimate (3.9) to (3.8).

4. Subdifferentials of the upper limit of continuous functions

Let a family of continuous functions $f_\theta : \mathbb{X} \to \mathbb{R} \cup \{-\infty, \infty\}$, $\theta \in [0; \infty)$ be given. Define a function $f_{\sup} : \mathbb{X} \to \mathbb{R} \cup \{-\infty, \infty\}$ by the following rule:

$$f_{\sup}(x) \stackrel{\triangle}{=} \limsup_{\theta \uparrow \infty} f_\theta(x) \quad \text{for all} \quad x \in \mathbb{X}. \tag{4.1}$$

For every positive δ, denote by $Z_\delta(\check{x})$ the set of all $\zeta \in \mathbb{X}^*$ for which there exists a pair $(\theta, x) \in [0; \infty) \times \mathbb{X}$ such that $\zeta \in \hat{\partial} f_\theta(x)$,

$$\theta > 1/\delta, \quad x \in \check{x} + \delta B, \quad \text{and} \quad |f_\theta(x) - (\operatorname{lsc} f_{\sup})(\check{x})| < \delta. \tag{4.2}$$

The following estimate of the subdifferential of the upper limit function is the enlargement of [11, Lemma 6] on reflexive spaces as well as the refinement of [12, Theorem 6.1(a)] in the case of continuous functions; its proof is similar to that of [12, Theorem 6.1(a)].

Proposition 2. *Assume that $\mathbb{X}$ is a reflexive space, a family of scalar functions f_θ, $\theta \in [0; +\infty)$, continuous on $\mathbb{X}$ is given, and $f_{\sup}$ is defined by (4.1). For all $\check{x} \in \mathbb{X}$ and $\xi \in \hat{\partial} \operatorname{lsc} f_{\sup}(\check{x})$, for every positive δ, there exist some $N \in \mathbb{N}$, $\alpha_1, \alpha_2, \ldots, \alpha_N \in [0; 1]$, and $\zeta_1, \zeta_2, \ldots, \zeta_N \in Z_\delta\big(\check{x}, \operatorname{lsc} f_{\sup}(\check{x})\big)$ such that $\alpha_1 + \cdots + \alpha_N = 1$ and*

$$\xi \in \sum_{k=1}^{N} \alpha_k \zeta_k + \delta B^*. \tag{4.3}$$

P r o o f. **The special case: $\check{x} = 0$ is a local minimum of** $\operatorname{lsc} f_{\sup}$**.** Assume that $\check{x} \stackrel{\triangle}{=} 0$ and $\xi \stackrel{\triangle}{=} 0$; furthermore, assume that

$$(\operatorname{lsc} \limsup_{\theta \uparrow \infty} f_\theta)(0) = \inf_{x \in \delta_0 B} \limsup_{\theta \uparrow \infty} f_\theta(x) = 0$$

for some positive δ_0. Then $0 \in \hat{\partial} f_{\sup}(0) = \hat{\partial} \operatorname{lsc} f_{\sup}(0)$.

Note that $f_{\sup}(z) = \inf_{T>0} E(T, z)$ for all $z \in \mathbb{X}$; here $E : [0; +\infty) \times \mathbb{X} \to \mathbb{R} \cup \{-\infty, +\infty\}$ is defined as

$$E(T, x) \stackrel{\triangle}{=} \sup_{\theta \geq T} f_\theta(x) \quad \text{for all} \quad T > 0, \quad x \in \mathbb{X}.$$

Fix a vector $v \in B$ and a positive number $\delta < \min(\delta_0, 1/3)$. Define $t = \delta^2$. Since 0 is a local minimum of $\operatorname{lsc} f$, there exists a point $\check{z} \in tB$ such that

$$0 \leq \operatorname{lsc} f_{\sup}(\check{z}) \leq f_{\sup}(\check{z}) < \delta^4.$$

Then,

$$\|\check{z} + tv\| < 2\delta^2 < \delta < \delta_0 \quad \text{and} \quad f_{\sup}(\check{z} + tv) \geq \operatorname{lsc} f_{\sup}(\check{z} + tv) \geq 0.$$

So,

$$f_{\sup}(\check{z} + tv) - f_{\sup}(\check{z}) > -\delta^4 = -\delta^2 t.$$

Further, we can find positive numbers $\bar{T} \geq 1/\delta$ and $\hat{\theta} > \bar{T}$ such that

$$\delta^2 t > E(\bar{T}, \check{z}) - f_{\sup}(\check{z}) \quad \text{and} \quad \delta^2 t + f_{\hat{\theta}}(\check{z} + tv) > E(\bar{T}, \check{z} + tv). \tag{4.4}$$

By definition of E, we also have

$$0 \leq E(\bar{T}, \check{z} + tv) - f_{\sup}(\check{z} + tv) \quad \text{and} \quad f_{\hat{\theta}}(\check{z}) \leq E(\bar{T}, \check{z}). \tag{4.5}$$

Subtracting the sum of inequalities (4.5) from the sum of inequalities (4.4), we have

$$2\delta^2 t + f_{\hat\theta}(\check z + tv) - f_{\hat\theta}(\check z) > f_{\sup}(\check z + tv) - f_{\sup}(\check z).$$

From $f_{\sup}(\check z + tv) - f_{\sup}(\check z) > -\delta^2 t$ and $\delta < 1/3$, it follows that

$$f_{\hat\theta}(\check z + tv) - f_{\hat\theta}(\check z) > -\delta t.$$

Now, Theorem 1 for $f = f_{\hat\theta}$ with $u = \check z$, $v = \check z + tv$, $s = -\delta t$, and $\varepsilon = \delta(\delta - t)$ gives a number $r \in [0,t]$, a point $\hat z \in \mathbb{X}$, and a subgradient $\hat\zeta \in \hat\partial f_{\hat\theta}(\hat z)$ such that

$$-\delta t < t\hat\zeta v, \quad \|\hat z - \check z\| < \|\hat z - rv\| + r \le t + \delta(\delta - t) < 2\delta^2, \quad \text{and} \quad |f_{\hat\theta}(\hat z) - f_{\hat\theta}(\check z)| < \delta t + \delta(\delta - t) = \delta^2.$$

Then, by the choice of $\check z$, we obtain

$$\|\hat z\| \le \|\check z\| + 2\delta^2 < 3\delta^2 < \delta \quad \text{and} \quad |f_{\hat\theta}(\hat z)| \le |f_{\hat\theta}(\check z) - f_{\hat\theta}(u)| + \delta^2 \le 2\delta^2 < \delta.$$

So, we show (4.2) for $(\check x, \check y) = (0,0)$, $(\theta, x) = (\hat\theta, \hat z)$, therefore we obtain $\hat\zeta \in Z_\delta(0,0)$. Hence, for each $v \in B$, we have found $\hat\zeta \in Z_\delta(0,0)$ such that $\hat\zeta v > -\delta$. This entails

$$-\delta < \inf_{v\in B} \sup_{\zeta \in Z_\delta(0,0)} \zeta v \le \inf_{v \in B} \sup_{\zeta \in \operatorname{cl}\operatorname{co} Z_\delta(0,0)} \zeta v.$$

The set B is an weak compact subset of $\mathbb{X}^{**} = \mathbb{X}$ and, together with $\operatorname{cl}\operatorname{co} Z_\delta(0,0)$, is convex. In addition, the map $(\zeta, v) \mapsto \zeta v$ is continuous and linear in $(\zeta, v) \in \mathbb{X}^* \times \mathbb{X}^{**}$. Therefore, the nonsymmetrical Minimax Theorem [4, Theorem 3.6.14] ensures

$$-\delta < \inf_{v\in B} \sup_{\zeta \in \operatorname{cl}\operatorname{co} Z_\delta(0,0)} \zeta v = \sup_{\zeta \in \operatorname{cl}\operatorname{co} Z_\delta(0,0)} \inf_{v \in B} \zeta v.$$

Since there exists $\zeta \in \operatorname{cl}\operatorname{co} Z_\delta$ such that $\delta > -\zeta v$ for all $v \in B$, we obtain $\|\zeta\| \le \delta$. Therefore, (4.3) holds in the special case. The special case of this lemma is proved.

The general case. Let a point $\check x \in \mathbb{X}$ and a subgradient $\xi \in \hat\partial \operatorname{lsc} f_{\sup}(\check x)$ be given. Define $\check y = \operatorname{lsc} f_{\sup}(\check x)$. Choose a positive number $\delta < 1/3$.

Since $\mathbb{X}$ is a Fréchet smooth space, by [7, Theorem 4.6 (i)], there exist a C^1-smooth function g and a positive number $\delta_1 < \delta^2$ such that

$$\xi = g'(\check x), \quad \operatorname{lsc} f_{\sup}(\check x) = g(\check x), \quad \text{and} \quad \operatorname{lsc} f_{\sup}(\check x + tv) - \operatorname{lsc} f_{\sup}(\check x) \ge g(\check x + tv) - g(\check x)$$

if $x \in \check x + \delta_1 B$. Further, decreasing δ_1 if necessary, we can also ensure $\xi \in g'(x) + \delta^2 B^*$ and $g(x) \in g(\check x) + \delta^2 B$ for all $x \in \check x + \delta_1 B$. So,

$$\operatorname{lsc}(f_{\sup} - g)(0) \le (f_{\sup} - g)(x) \quad \text{for all} \quad x \in \delta_1 B.$$

Using the special case for the maps

$$\mathbb{X} \ni x \mapsto \bar f_\theta(x) = f_\theta(x - \check x) - g(x - \check x),$$

and a positive number δ^2, we find $\zeta \in \operatorname{cl}\operatorname{co} \bar Z_{\delta^2}(\check x, \check y) \cap \delta^2 B^*$. Then, in the account of the fuzzy sum rule, one finds a positive integer N, points $x_1, \dots, x_N \in \mathbb{X}$, subgradients $\bar\zeta_1 \in \hat\partial f_{\theta_1}(\bar x_1) - g'(\bar x_1), \dots, \zeta_N \in \hat\partial \bar f_{\theta_N}(\bar x_N) - g'(\bar x_N)$, and convex coefficients $\alpha_i \in [0,1]$ such that $\alpha_1 + \ldots + \alpha_N = 1$,

$$\bar x_i \in \check x + 2\delta^2 B, \quad \theta_i \ge \delta^{-2}, \quad |\bar f_{\theta_i}(\bar x_i) - g(\bar x_i) - \operatorname{lsc} f_{\sup}(\check x) + g(\check x)| \le 2\delta^2$$

for all i and
$$\sum_{k=1}^{N} \alpha_k \zeta_k \in 2\delta^2 B^*.$$
Define
$$\bar{\zeta}'_i \stackrel{\triangle}{=} \bar{\zeta}_i + g'(\bar{x}_1) \in \hat{\partial} f_{\theta_1}(\bar{x}_1).$$
By the choice of a positive number δ_1 and a smooth function g, we obtain
$$|\bar{f}_{\theta_i}(\bar{x}_i) - \operatorname{lsc} f_{\sup}(\check{x})| \le |g(\bar{x}_i) - g(\check{x})| + 2\delta^2 < \delta,$$
$$\Big\| \sum_{k=1}^{N} \alpha_k \zeta'_k - \sum_{k=1}^{N} \alpha_k \bar{\zeta}_k \Big\| \le \max_{i \in [1:N]} \big\| g'(\bar{x}_i) - \xi \big\| < 2\delta^2 < \delta, \quad \text{and} \quad \xi \in \sum_{k=1}^{N} \alpha_k \zeta'_k + \delta B^*.$$
So, the proposition is proved. □

Remark 4. If $\mathbb{X} \stackrel{\triangle}{=} \mathbb{R}^d$, by the famous Carathéodory theorem [15, Theorem 2.29], any finite convex sum of a (co)vectors can be represented by some finite convex sum of no more than $d+1$ of them. So, we can assume that $N \le d+1$.

Remark 5. If every f_θ is C^1-smooth, we conclude that every $\hat{\partial} f_\gamma(x)$ is a singleton; therefore, $\zeta_i = f'_{\gamma_i}(x_i)$ for all i.

REFERENCES

1. Aussel D., Corvellec J.-N., Lassonde M. Nonsmooth constrained optimization and multidirectional mean value inequalities. *SIAM J. Optim.,* 1999. Vol. 9, No. 3. P. 690–706. DOI: 10.1137/S105262349732339X
2. Borwein J. M., Fitzpatrick S. P., Giles J. R. The differentiability of real functions on normed linear space using generalized subgradients. *J. Math. Anal. Appl.,* 1987. Vol. 128, No. 2. P. 512–534. DOI: 10.1016/0022-247X(87)90203-4
3. Borwein J. M., Vanderwerff J. D. *Convex Functions: Constructions, Characterizations and Counterexamples.* Cambridge: Cambridge University Press, 2010. x+521 p.
4. Borwein J. M., Zhu Q. J. *Techniques of Variational Analysis.* Springer-Verlag: New York, 2005. vi+362 p. DOI: 10.1007/0-387-28271-8
5. Clarke F., Ledyaev Yu. S. New formulas for finite increments. *Dokl. Math.*, 1994. Vol. 48, No 1. P. 75–79.
6. Clarke F. H., Ledyaev Yu. S., Stern R. J., Wolenski P. R. *Nonsmooth Analysis and Control Theory.* New York: Springer-Verlag, 1998. xiv+276 p. DOI: 10.1007/b97650
7. Fabian M., Mordukhovich B. S. Nonsmooth characterizations of Asplund spaces and smooth variational principles. *Set-Valued Anal.*, 1998. Vol. 6. P. 381–406. DOI: 10.1023/A:1008799412427
8. Gomoyunov M. I., Plaksin A. R. Equivalence of minimax and viscosity solutions of path-dependent Hamilton–Jacobi equations. *J. Funct. Anal.*, 2023. Vol. 285, No. 11. Art. no. 110155. DOI: 10.1016/j.jfa.2023.110155
9. Ioffe A. D. *Variational Analysis of Regular Mappings: Theory and Applications.* Ser. Springer Monogr. Cham: Math. Springer, 2017. xxi+495 p. DOI: 10.1007/978-3-319-64277-2
10. Kipka R., Ledyaev Yu. S. A generalized multidirectional mean value inequality and dynamic optimization. *Optimization,* 2019. Vol. 68, No 7. P. 1365–1389. DOI: 10.1080/02331934.2019.1638388
11. Khlopin D. Necessary conditions in infinite-horizon control problems that need no asymptotic assumptions. *Set-Valued Var. Anal.*, 2023. Vol. 31, No. 1. Art. no. 8. DOI: 10.1007/s11228-023-00672-5
12. Ledyaev Yu. S., Treiman J. S. Sub- and supergradients of envelopes, semicontinuous closures, and limits of sequences of functions. *Russian Math. Surveys*, 2012. Vol. 67, No. 2. P. 345–373. DOI: 10.1070/RM2012v067n02ABEH004789
13. Mordukhovich B. S. *Variational Analysis and Generalized Differentiation I: Basic Theory.* Berlin: Springer-Verlag, 2006. xxii+579 p. DOI: 10.1007/3-540-31247-1

14. Mordukhovich B. S. *Variational Analysis and Applications.* Ser. Springer Monogr. Math. Cham: Springer, 2018. xix+622 p. DOI: 10.1007/978-3-319-92775-6
15. Rockafellar R. T., Wets R. J. B. *Variational Analysis.* Berlin: Springer-Verlag, 2009. vii+735 p. DOI: 10.1007/978-3-642-02431-3
16. Subbotin A. I. On a property of the subdifferential. *Math. USSR-Sb.*, 1993. Vol. 74, No. 1. P. 63–78. DOI: 10.1070/SM1993v074n01ABEH003335
17. Trang N. T. Q. A note on an approximate mean value theorem for Fréchet subgradients, *Nonlinear Anal.,* 2012. Vol. 75, No. 1. P. 380–383. DOI: 10.1016/j.na.2011.08.041
18. Zhu Q. J. The equivalence of several basic theorems for subdifferentials. *Set-Valued Anal.,* 1998. Vol. 6. P. 171–185. DOI: 10.1023/A:1008656207314

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 141–156

DOI: 10.15826/umj.2023.2.012

CONVEXITY OF REACHABLE SETS OF QUASILINEAR SYSTEMS[1]

Ivan O. Osipov

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

i.o.osipov@imm.uran.ru

Abstract: This paper investigates convexity of reachable sets for quasilinear systems under integral quadratic constraints. Drawing inspiration from B.T. Polyak's work on small Hilbert ball image under nonlinear mappings, the study extends the analysis to scenarios where a small nonlinearity exists on the system's right-hand side. At zero value of a small parameter, the quasilinear system turns into a linear system and its reachable set is convex. The investigation reveals that to maintain convexity of reachable sets of these systems, the nonlinear mapping's derivative must be Lipschitz continuous. The proof methodology follows a Polyak's scheme. The paper's structure encompasses problem formulation, exploration of parameter linear mapping and image transformation, application to quasilinear control systems, and concludes with illustrative examples.

Keywords: Quasilinear control system, Small parameter, Integral constraints, Reachable sets, Convexity.

1. Introduction

This paper focuses on studying the reachable sets of quasilinear systems with integral quadratic constraints.

The study is based on the work of B.T. Polyak [21, 22], wherein sufficient conditions were derived for establishing convexity of a nonlinear mapping applied to a small ball in Hilbert space. These conditions were further applied to problems in control theory, demonstrating that the reachable set of a nonlinear system exhibits convexity given a sufficient small control resource, provided that the linearized system is controllable. A series of papers [12–14, 19] used the convexity conditions of the small ball mapping to investigate the reachable sets of nonlinear systems under integral constraints over small time intervals. In this case, it is important to note that the controllability of the linearized system alone does not guarantee convexity of the reachable sets for the nonlinear system. Additional conditions related to the asymptotic behavior of the eigenvalues of the controllability Gramian of the linearized system need to be imposed. Once these conditions are fulfilled, the reachable sets of the nonlinear system not only exhibit convexity but are also asymptotically equivalent to the reachable sets of the linearized system.

Therefore, the study investigates the convexity of reachable sets of nonlinear systems with a small control resource and on a small time interval. This paper discusses a variant of convexity of reachable sets of systems with a small parameter, namely with a small nonlinearity on the right hand side.

Systems that have small nonlinearity on the right-hand side are commonly called quasilinear systems. The study of such systems in control theory dates back to the 1960s [16, 17, 23].

[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 no. 075-02-2023-913).

E.G. Albrecht solved several problems concerning quasilinear systems [3], including the optimal motion problem [1] and the game problem of quasilinear objects rendezvous [2]. Control problems for quasilinear systems are also addressed in the following works [6, 9, 15, 18]. In modern applications of control theory, quasilinear systems arise after feedback linearization [4] and stochastic linearization [5, 10].

The paper studies the convexity of the reachable sets of quasilinear systems under integral constraints. In line with researches [12–14, 19, 21, 22], the study is reduced to the analysis of a nonlinear mapping from the control space and the small parameter space to the space of the trajectories endpoints generated by these controls. In this case, the reachable set is the image of the ball under this mapping. The specific feature of the mapping defined by the solution of a quasilinear system is the fact that, at zero value of the small parameter, this mapping becomes linear in control. For the image of the ball to preserve its convexity for small values of the small parameter, it is necessary for the nonlinear part of the mapping to have a Lipschitz continuous derivative. The scheme for proving this statement is in many ways similar to the scheme for proving the mail theorem in [21].

The paper is organized as follows. The problem statement and some remarks are given in the second section. The third section contains the investigation of parameter linear mapping and image of a ball. In next section, we apply the results of the third section to the quasilinear control system. Finally, we provide two illustrative examples in the fifth section.

2. Problem statement and preparatory remarks

Further we use the following notation. By $A^\top$ we denote the transpose of a real matrix A, I is an identity matrix, 0 stands for a zero vector or a zero matrix of appropriate dimension. For a real $n \times n$ matrix A a spectral matrix norm induced by the Euclidean vector norm is denoted as $\|A\|$. The symbols $\mathbb{L}_1$ and $\mathbb{L}_2$ stand for the spaces of summable and square summable functions respectively. The norms in these spaces are denoted as $\|\cdot\|_{\mathbb{L}_1}$ and $\|\cdot\|_{\mathbb{L}_2}$. By $B_X(a,r)$ we will denote the closed ball of radius $r > 0$ centered at a,

$$B_X(a,r) = \{x \in X : \|x - a\|_X \leqslant r\}.$$

Here X is some linear space with a norm $\|\cdot\|_X$.

Let us consider the quasilinear control system

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + \varepsilon f\big(x(t), t\big), \quad t_0 \leqslant t \leqslant T, \quad x(t_0) = x_0, \tag{2.1}$$

where $x \in \mathbb{R}^n$ is a state vector, $u \in \mathbb{R}^r$ is a control vector, t_0 is a non-negative number, T is a positive number and ε is a small parameter such that $\varepsilon \in [0, \overline{\varepsilon}]$, $\overline{\varepsilon} > 0$. The matrix maps $A : [t_0, T] \to \mathbb{R}^{n \times n}$, $B : [t_0, T] \to \mathbb{R}^{n \times r}$ are assumed to be continuous. The function $f : \mathbb{R}^n \times [t_0, T] \to \mathbb{R}^n$ is assumed to be continuous in (x, t) and continuously differentiable in x.

The control $u(\cdot)$ will be chosen from $B_{\mathbb{L}_2}(0, \mu)$ with some $\mu > 0$.

For any control $u(\cdot) \in \mathbb{L}_2$ and any $\varepsilon \in [0, \overline{\varepsilon}]$ there exists the unique absolutely continuous solution $x(t, \varepsilon, u(\cdot))$ of the system (2.1), satisfying the initial condition $x(t_0, \varepsilon, u(\cdot)) = x_0$, and this solution is defined on some interval $[t_0, t_0 + \Delta]$, where $t_0 + \Delta < T$.

Further we will suppose that the conditions of the following assumption are satisfied.

Assumption 1. *There exists $\overline{\mu} > \mu$ such that for all $\varepsilon \in [0, \overline{\varepsilon}]$ all solutions $x(t, \varepsilon, u(\cdot))$ generated by controls $u(\cdot) \in B_{\mathbb{L}_2}(0, \overline{\mu})$ belong to some convex compact set $D \subset \mathbb{R}^n$. In addition, it* *is assumed that the function $f : \mathbb{R}^n \times [t_0, T] \to \mathbb{R}^n$ and its derivative on x satisfy the Lipschitz condition with constants L_f, l_f respectively*

$$\|f(x_1,t) - f(x_2,t)\| \leqslant L_f \|x_1 - x_2\|, \quad t \in [t_0, T], \quad x_1, x_2 \in D,$$
$$\left\| \frac{\partial f(x_1,t)}{\partial x} - \frac{\partial f(x_2,t)}{\partial x} \right\| \leqslant l_f \|x_1 - x_2\|, \quad t \in [t_0, T], \quad x_1, x_2 \in D.$$

In particular, first part of Assumption 1 holds if f satisfies one of the following conditions [8]:

$$\begin{aligned} &\big\|f\big(x,t\big)\big\| \leqslant l_1(t)(1 + \|x\|), \\ &x \cdot f(x,t) \leqslant a(t)\|x\|^2 + b(t), \end{aligned} \tag{2.2}$$

where $l_1(\cdot) \in \mathbb{L}_1[t_0, T]$ and $a(\cdot), b(\cdot)$ are continuous functions.

Definition 1. *The reachable set $G(T, \mu, \varepsilon)$ of the system (2.1) at time T is the set consisting of all possible states that can be reached by the system at time T while satisfying the given integral constraints on the control*

$$G(T, \mu, \varepsilon) = \big\{\widetilde{x} \in \mathbb{R}^n : \exists u(\cdot) \in B_{\mathbb{L}_2}(0, \mu),\ x(T, \varepsilon, u(\cdot)) = \widetilde{x}\big\}.$$

The question to be studied is under which conditions the reachable set will be convex for small ε.

3. Nonlinear mappings depending on a small parameter

In this section, x (including x_0, x_1 and others) is not related to the state vector of system (2.1). Here x is an element of the space X, ε is still a small non-negative parameter.

Consider the mapping

$$F(x, \varepsilon) = a_0 + A_0 x + \varepsilon A_1(x, \varepsilon) : B_X(0, r) \times \mathbb{R}_+ \to Y,$$

where X and Y are Hilbert spaces. Here $a_0 \in Y$ is a constant, it does not depend on either x or ε, A_0 is a linear continuous operator which we assume to be a surjective mapping from X to Y. The last implies, that there exists $\nu > 0$, such that

$$\|A_0^* y\| \geqslant \nu \|y\|, \quad \forall y \in Y. \tag{3.1}$$

Here A_0^* is a linear operator adjoint to A_0. Let $A_1 : B_X(0, r) \times \mathbb{R}_+ \to Y$ be a nonlinear operator, which is continuous in x and ε.

Assumption 2. *There exists $\overline{\varepsilon} > 0$, such that for all $x \in B_X(0, r)$, $\varepsilon \in [0, \overline{\varepsilon}]$ the mapping $A_1(x, \varepsilon)$ has a Frechet derivative*

$$\frac{\partial A_1(x, \varepsilon)}{\partial x} = A_1'(x, \varepsilon)$$

which is continuous in ε and Lipschitz continuous in x: there exists $L > 0$, such that

$$\|A_1'(x_1, \varepsilon) - A_1'(x_2, \varepsilon)\| \leqslant L\|x_1 - x_2\|, \quad x_1, x_2 \in B_X(0, r), \quad \varepsilon \in [0, \overline{\varepsilon}].$$

In order to justify this further, it is useful to quote the following result from [20, 21]. In the formulation of following lemma it is assumed that $f : X \to Y$ is a nonlinear Frechet differentiable map.

Lemma 1 [20, 21]. *Suppose there exist L, ρ, $\mu > 0$ and $y_0 \in Y$, such that*

$$\|f'(x) - f'(z)\| \leqslant L\|x - z\|, \quad \forall x, z \in B(x_0, \rho),$$
$$\|f'(x)^* y\| \geqslant \mu\|y\|, \quad \forall y \in Y, \quad \forall x \in B(x_0, \rho),$$
$$\|f(x_0) - y_0\| \leqslant \rho\mu,$$

then the equation $f(x) = y_0$ has a solution $x^* \in B(x_0, \rho)$ and

$$\|x^* - x_0\| \leqslant \frac{1}{\mu} \|f(x_0) - y_0\| .$$

The following theorem can now be formulated and proven.

Theorem 1. *Denote the image of the ball $B_X(0, r)$ under the map $F(x, \varepsilon)$ by $F\big(B_X(0, r), \varepsilon\big)$, i.e.*

$$F\big(B_X(0, r), \varepsilon\big) = \big\{F(x, \varepsilon) : x \in B_X(0, r)\big\}.$$

Suppose the condition of Assumption 2 to be fulfilled and $F\big(B_X(0, r), \varepsilon\big)$ is a closed set for each $\varepsilon \in [0, \overline{\varepsilon}]$. There exists $\varepsilon_0 \in (0, \overline{\varepsilon}]$, such that for all positive $\varepsilon < \varepsilon_0$ the image $F\big(B_X(0, r), \varepsilon\big)$ is a convex set in Y.

P r o o f. Note that the constant a_0 has no impact on the convexity of the image $F\big(B_X(0, r), \varepsilon\big)$. Therefore, for the proof, we will consider it as zero.

Let us consider two arbitrary points, x_1 and x_2, in $B_X(0, r)$. Let

$$x_0 = \frac{1}{2}(x_1 + x_2), \quad y(\varepsilon) = \frac{1}{2}\big(y_1(\varepsilon) + y_2(\varepsilon)\big),$$

where $y_1(\varepsilon) = F(x_1, \varepsilon)$ and $y_2(\varepsilon) = F(x_2, \varepsilon)$.

To prove that the set $F\big(B_X(0, r), \varepsilon\big)$ is convex we need to show $y(\varepsilon) \in F\big(B_X(0, r), \varepsilon\big)$ or, equivalently, there exists $x^* \in B_X(0, r)$, such that $F(x^*, \varepsilon) = y(\varepsilon)$. Let us write down the expression for $y(\varepsilon)$

$$\begin{aligned} y(\varepsilon) = \frac{1}{2}\big(F(x_1, \varepsilon) + F(x_2, \varepsilon)\big) &= \frac{1}{2}\big(A_0 x_1 + \varepsilon A_1(x_1, \varepsilon) + A_0 x_2 + \varepsilon A_1(x_2, \varepsilon)\big) \\ &= A_0 x_0 + \frac{1}{2}\varepsilon\big(A_1(x_1, \varepsilon) + A_1(x_2, \varepsilon)\big). \end{aligned} \tag{3.2}$$

Let $x \in X$ and $h \in X$ be chosen such that the inclusions $x \in B_X(0, r)$ and $x + h \in B_X(0, r)$ are valid. Under Assumption 2, we will expand A_1 in a series around the point x:

$$A_1(x + h, \varepsilon) = A_1(x, \varepsilon) + A_1'(x, \varepsilon)h + R(\varepsilon, x, h). \tag{3.3}$$

Multiplying both sides of this equality by $y^* \in Y^*$, $\|y^*\| \leqslant 1$ we get

$$\langle y^*, R(\varepsilon, x, h)\rangle = \langle y^*, A_1(x + h, \varepsilon)\rangle - \langle y^*, A_1(x, \varepsilon)\rangle - \langle y^*, A_1'(x, \varepsilon)h\rangle.$$

Apply mean value theorem to function $\langle y^*, A_1(x, \varepsilon)\rangle$ to obtain

$$\langle y^*, A_1(x + h, \varepsilon)\rangle - \langle y^*, A_1(x, \varepsilon)\rangle = \langle y^*, A_1'(x + \theta h, \varepsilon)h\rangle, \qquad 0 \leqslant \theta \leqslant 1.$$

The last two relations lead to the following estimates

$$\|\langle y^*, R(\varepsilon, x, h)\rangle\| \leqslant \|y^*\|\|A_1'(x + \theta h, \varepsilon) - A_1'(x, \varepsilon)\|\|h\| \leqslant L\theta\|h\|^2 \leqslant L\|h\|^2,$$
$$\|R(\varepsilon, x, h)\| \leqslant L\|h\|^2,$$

Substituting (3.3) into the expression (3.2), we obtain

$$\begin{aligned}y(\varepsilon) = A_0x_0 + \frac{1}{2}\varepsilon\Big(&A_1(x_0,\varepsilon) + A_1'(x_0,\varepsilon)(x_1-x_0) + R(\varepsilon,x_0,x_1-x_0)\\ &+A_1(x_0,\varepsilon) + A_1'(x_0,\varepsilon)(x_2-x_0) + R(\varepsilon,x_0,x_2-x_0)\Big)\\ &= A_0x_0 + \varepsilon A_1(x_0,\varepsilon) + \xi(\varepsilon,x_1,x_2),\end{aligned}$$

where the residual term has the form

$$\xi(\varepsilon,x_1,x_2) = \frac{1}{2}\varepsilon\big(R(\varepsilon,x_0,x_1-x_0) + R(\varepsilon,x_0,x_2-x_0)\big),$$

and it could be estimated as

$$\|\xi(\varepsilon,x_1,x_2)\| \leqslant \frac{1}{2}\varepsilon L\Big(\frac{1}{4}\|x_1-x_2\|^2 + \frac{1}{4}\|x_1-x_2\|^2\Big) \leqslant \frac{1}{4}L\varepsilon\|x_1-x_2\|^2.$$

As a result, we have

$$y(\varepsilon) = A_0x_0 + \varepsilon A_1(x_0,\varepsilon) + \xi(\varepsilon,x_1,x_2) = F(x_0,\varepsilon) + \xi(\varepsilon,x_1,x_2)$$

for all $x_1, x_2 \in B(0,r)$, $x_1 \neq x_2$, $\varepsilon \in [0,\overline{\varepsilon}]$, and we have

$$\|F(x_0,\varepsilon) - y(\varepsilon)\| = \|\xi(\varepsilon,x_1,x_2)\| \leqslant \frac{1}{4}L\varepsilon\|x_1-x_2\|^2.$$

Now let us study the derivative of the mapping $F(x_0,\varepsilon)$ in x_0 for a fixed ε,

$$F_x'(x_0,\varepsilon) = A_0 + \varepsilon A_1'(x_0,\varepsilon).$$

Using Assumption 2 we can estimate $\|A_1'(x_0,\varepsilon\|$ from above:

$$\begin{aligned}\|A_1'(x_0,\varepsilon) - A_1'(0,\varepsilon)\| &\leqslant L\|x_0\| \leqslant Lr,\\ \|A_1'(x_0,\varepsilon)\| &\leqslant \|A_1'(0,\varepsilon)\| + Lr.\end{aligned}$$

Since

$$\|A_1'(x_0,\varepsilon)\| = \|\big(A_1'(x_0,\varepsilon)\big)^*\|,$$

it follows

$$\|\big(A_1'(x_0,\varepsilon)\big)^*\| \leqslant \|A_1'(0,\varepsilon)\| + Lr.$$

From this and (3.1), we have

$$\big\|F_x'(x_0,\varepsilon)^*y\big\| = \big\|\big(A_0 + \varepsilon A_1'(x_0,\varepsilon)\big)^*y\big\| \geqslant \|A_0^*y\| - \varepsilon\big\|\big(A_1'(x_0,\varepsilon)\big)^*\big\|\,\|y\| \geqslant (\nu - k\varepsilon)\|y\|,$$

where

$$k = \max_{\varepsilon\in[0,\overline{\varepsilon}]}\|A_1'(0,\varepsilon)\| + Lr > 0.$$

For small ε, the following inequality is true

$$(\nu - k\varepsilon) \geqslant \frac{\nu}{2}$$

and we have

$$\big\|F_x'(x_0,\varepsilon)^*y\big\| \geqslant \frac{\nu}{2}\|y\|.$$

In order to use Lemma 1, we require that

$$\|F(x_0,\varepsilon)-y(\varepsilon)\| = \|\xi(\varepsilon,x_1,x_2)\| \leqslant \frac{1}{4}L\varepsilon\|x_1-x_2\|^2 \leqslant \frac{\nu}{2}\frac{\|x_1-x_2\|^2}{8r}.$$

To achieve this, it is necessary to choose a value of ε such that it satisfies the inequality

$$\varepsilon \leqslant \varepsilon_0 = \min\left\{\frac{\nu}{4Lr}, \frac{\nu}{2k}, \overline{\varepsilon}\right\}.$$

Then, from Lemma 1 with parameters

$$\mu = \frac{\nu}{2}, \qquad \rho = \frac{\|x_1-x_2\|^2}{8r},$$

it follows that there exists $x^* \in B(x_0,\rho)$ such that $F(x^*,\varepsilon) = y(\varepsilon)$.

Since $B_X(0,r)$ is Hilbert ball, it is strongly convex and the inclusion $B(x_0,\rho) \subset B_X(0,r)$ is true, therefore $x^* \in B_X(0,r)$. So, the point

$$y(\varepsilon) = \frac{1}{2}\big(F(x_1,\varepsilon)+F(x_2,\varepsilon)\big)$$

is contained within the image of the ball $F\big(B_X(0,r),\varepsilon\big)$ for all $\varepsilon \leqslant \varepsilon_0$ and $x_1, x_2 \in B_X(0,r)$. Due to the closeness, for all $\varepsilon \leqslant \varepsilon_0$, the image of the ball $F\big(B_X(0,r),\varepsilon\big)$ is convex. □

4. On the properties of the solutions of quasilinear systems

In this section we investigate the solutions of (2.1) to verify the applicability of the previous results, in particular Theorem 1.

By $X(t,\tau)$ we denote the Cauchy matrix of the linear system $\dot{x}(t) = A(t)x(t)$. This matrix is the solution of the following equation

$$\frac{\partial X(t,\tau)}{\partial t} = A(t)X(T,\tau), \qquad X(\tau,\tau) = I.$$

If $x(\cdot,\varepsilon,u(\cdot))$ is the solution of (2.1), produced by control $u(\cdot)$ and initial condition x_0, it satisfies the next integral equation

$$x\big(T,\varepsilon,u(\cdot)\big) = X(T,t_0)x_0 + \int_{t_0}^{T} X(T,\tau)\Big(Bu(\tau)+\varepsilon f\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)\Big)\,d\tau$$

$$= X(T,t_0)x_0 + \int_{t_0}^{T} X(T,\tau)B(t)u(\tau)\,d\tau + \varepsilon\int_{t_0}^{T} X(T,\tau)f\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)\,d\tau.$$

Let us define the mapping $F: B_{\mathbb{L}_2}(0,\overline{\mu})\times[0,\overline{\varepsilon}] \to \mathbb{R}^n$ by the equality $F(u(\cdot),\varepsilon) = x(T,\varepsilon,u(\cdot))$, where $x(T,\varepsilon,u(\cdot))$ is the solution of (2.1) at moment T generated by the control $u(\cdot)$ and the small parameter ε.

In order to use the results from the previous sections, we now rewrite the mapping F as

$$F(u(\cdot),\varepsilon) = a_0 + A_0u(\cdot) + \varepsilon A_1(u(\cdot),\varepsilon),$$

where $a_0 = X(T,0)x_0$, the linear map $A_0 : B_{\mathbb{L}_2}(0,\overline{\mu}) \mapsto \mathbb{R}^n$ is defined by

$$A_0 u(\cdot) = \int\limits_{t_0}^{T} X(T,\tau)B(t)u(\tau)\ d\tau$$

and nonlinear map $A_1 : B_{\mathbb{L}_2}(0,\overline{\mu}) \times [0,\overline{\varepsilon}] \to \mathbb{R}^n$ is defined by

$$A_1(u(\cdot),\varepsilon) = \int\limits_{t_0}^{T} X(T,\tau) f\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)\ d\tau. \tag{4.1}$$

Reachable set $G(T,\mu,\varepsilon)$ of the quasilinear system (2.1) is the image under mapping F of the ball $B_{\mathbb{L}_2}(0,\mu)$,

$$G(T,\mu,\varepsilon) = F(B_{\mathbb{L}_2}(0,\mu),\varepsilon).$$

Assertion 1. *Assume the Assumption 1 is fulfilled. Then, for all $\varepsilon \in [0,\overline{\varepsilon}]$, the reachable set $G(T,\mu,\varepsilon)$ is closed.*

P r o o f. The proof is based on the equicontinuity of trajectories, the uniform boundedness of the set of trajectories, and the weak compactness of the ball $B_{\mathbb{L}_2}(0,\mu)$ (see, for example [11]). □

To apply Theorem 1 to the mapping F, we must demonstrate that Assumption 2 holds for A_1, defined in equation (4.1).

Lemma 2. *Assume Assumption 1 to be fulfilled. Then, for all $\varepsilon \in [0,\overline{\varepsilon}]$, there exists a constant $L_x(\varepsilon) > 0$, such that for any $u_i(\cdot) \in B_{\mathbb{L}_2}(0,\mu)$, $i = 1,2$ and $t \in [t_0,T]$,*

$$\|x_1(t) - x_2(t)\| \leqslant L_x(\varepsilon)\|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2},$$

where $x_i(t) = x(t,\varepsilon,u_i(\cdot))$, $i = 1,2$. Furthermore, $L_x(\varepsilon) \leqslant L_x(\overline{\varepsilon})$.

P r o o f. Since $x_i(t) \in D$ for all $t \in [t_0,T]$, from Assumption 1, we have

$$\|f\big(x_1(t),t\big) - f\big(x_2(t),t\big)\| \leqslant L_f\|x_1(t) - x_2(t)\|.$$

From the integral identities

$$x_i(t) = x_0 + \int\limits_{t_0}^{t} A(\tau)x_i(\tau)\ d\tau + \int\limits_{t_0}^{t} B(\tau)u_i(\tau)\ d\tau + \varepsilon\int\limits_{t_0}^{t} f\big(x_i(\tau),\tau\big)\ d\tau \tag{4.2}$$

we get

$$\begin{aligned}\|x_1(t) - x_2(t)\| &\leqslant \left\|\int\limits_{t_0}^{t} A(\tau)\big(x_1(\tau) - x_2(\tau)\big)\ d\tau\right\| + \left\|\int\limits_{t_0}^{t} B(\tau)\big(u_1(\tau) - u_2(\tau)\big)\ d\tau\right\| \\ &\quad + \varepsilon\left\|\int\limits_{t_0}^{t}\Big(f\big(x_1(\tau),\tau\big) - f\big(x_2(\tau),\tau\big)\Big)\ d\tau\right\| \\ &\leqslant \int\limits_{t_0}^{t}\left(k_A + L_f\varepsilon\right)\|x_1(\tau) - x_2(\tau)\|\ d\tau + k_u\|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2}.\end{aligned}$$

Here,

$$k_u = \sqrt{(T - t_0) \max_{\tau \in [t_0, t]} \|B(\tau)\|}, \quad k_A = \max_{\tau \in [t_0, t]} \|A(\tau)\|.$$

From the Grownwall inequality we have

$$\|x_1(t) - x_2(t)\| \leqslant L_x(\varepsilon) \|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2},$$

where

$$L_x(\varepsilon) = k_u \exp\big((k_A + L_f \varepsilon)(T - t_0)\big).$$

Note, that $L_x(\varepsilon) \leqslant L_x(\overline{\varepsilon})$. □

Introduce the mapping $\overline{F} : [t_0, T] \times [0, \overline{\varepsilon}] \times B_{\mathbb{L}_2}(0, \overline{\mu}) \to \mathbb{R}^n$,

$$\overline{F}(\tau, \varepsilon, u(\cdot)) = x\big(\tau, \varepsilon, u(\cdot)\big),$$

where $x\big(\tau, \varepsilon, u(\cdot)\big)$ is a solution of (2.1) at moment τ generated by the control $u(\cdot)$ the small parameter ε. The derivative of $\overline{F}$ in $u(\cdot)$, $\overline{F}' : B_{\mathbb{L}_2}(0, \overline{\mu}) \to \mathbb{R}^n$ is the solution of the linearized system as it was shown in [11]

$$\overline{F}'(\tau, \varepsilon, u(\cdot))\delta u(\cdot) = \delta x(\tau),$$

where $\delta x(\tau)$ is a solution of the the system (2.1) linearized along $(u(\cdot), x(\cdot, \varepsilon, u(\cdot))$, corresponding to the control $\delta u(\cdot)$ and zero initial condition:

$$\delta \dot{x} = \overline{A}\big(t, \varepsilon, u(\cdot)\big)\delta x + B(t)\delta u(t), \quad 0 \leqslant t \leqslant \tau, \quad \delta x(0) = 0, \tag{4.3}$$

where

$$\overline{A}\big(t, \varepsilon, u(\cdot)\big) = A(t) + \varepsilon \frac{\partial f\big(x(t, \varepsilon, u(\cdot)), t\big)}{\partial x}.$$

Lemma 3. *Suppose Assumption 1 to be fulfilled. There exists a constant $L_u(\varepsilon) > 0$, such that for any $\varepsilon \in [0, \overline{\varepsilon}]$, $u_i(\cdot) \in B_{\mathbb{L}_2}(0, \mu)$ and $\tau \in [t_0, T]$,*

$$\|\overline{F}'(\tau, \varepsilon, u_1(\cdot)) - \overline{F}'(\tau, \varepsilon, u_2(\cdot))\| \leqslant L_u(\varepsilon) \|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2},$$

where $i = 1, 2$.

P r o o f. The solution of (4.3) has the form

$$\delta x\big(\tau, \varepsilon, u_i(\cdot), \delta u(\cdot)\big) = \int\limits_{t_0}^{\tau} \overline{X}(\tau, s, \varepsilon, u_i(\cdot)) B(s) \delta u(s)\, ds, \tag{4.4}$$

where $X(\tau, s, \varepsilon, u(\cdot))$ is fundamental matrix of system (4.3), and it satisfies the equation

$$\frac{\partial \overline{X}(\tau, s, \varepsilon, u(\cdot))}{\partial s} = -\overline{A}\big(s, \varepsilon, u(\cdot)\big)^{\top} \overline{X}(\tau, s, \varepsilon, u(\cdot)), \quad \overline{X}\big(\tau, \tau, \varepsilon, u(\cdot)\big) = I.$$

It is well-known (for example, it follows from the proof of Theorem 3 in [7]), that there exists $k_X > 0$ such that

$$\|\overline{X}(\tau, s, \varepsilon, u(\cdot))\| \leqslant k_X, \quad \tau \in [t_0, T], \quad s \in [t_0, T]$$

for all $u(\cdot) \in B_{\mathbb{L}_2}(0, \mu)$ and sufficiently small ε. For the sake of brevity, we use the notation $\overline{A}_i(t) = \overline{A}\big(t, \varepsilon, u_i(\cdot)\big)$ and $\overline{X}_i(t, s) = \overline{X}(t, s, \varepsilon, u_i(\cdot))$. Under Assumption 1 and using Lemma 2 we can obtain the estimate

$$\int\limits_{t_0}^{\tau} \|\overline{A}_1(s) - \overline{A}_2(s)\| ds \leqslant L_A \|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2}.$$

Here $L_A > 0$ does not depend on $u_1(\cdot)$, $u_2(\cdot)$, τ and ε. Since,

$$\frac{\partial}{\partial t}\big(\overline{X}_1(t,s) - \overline{X}_2(t,s)\big) = -\overline{A}_1^{\top}(t)\big(\overline{X}_1(t,s) - \overline{X}_2(t,s)\big) + (\overline{A}_2(t) - \overline{A}_1(t))^{\top}\overline{X}_2(t,s), \quad t \in [s, \tau]$$

we get the following formula

$$\overline{X}_1(\tau, s) - \overline{X}_2(\tau, s) = \int\limits_{s}^{\tau} Y(t,s)\big(\overline{A}_2(t) - \overline{A}_1(t)\big)^{\top} X_2(t,s) dt.$$

Here $Y(t, s)$ is a fundamental matrix of the system

$$\dot{y} = -\overline{A}_1(t) y.$$

Like $\overline{X}_i(\tau, s)$, this matrix is also bounded: there exists $k_Y > 0$ such that

$$\|Y(t,s)\| \leqslant k_Y, \quad t, s \in [t_0, \tau]$$

for all $u(\cdot) \in B(0, \mu)$. We get

$$\|\overline{X}_1(\tau, s) - \overline{X}_2(\tau, s)\| \leqslant L_X \|u_1(\cdot) - u_2(\cdot)\|_{\mathbb{L}_2},$$

where

$$L_X = k_Y L_A k_X (T - t_0).$$

Hence from (4.4) it follows the statement of the lemma and $L_u(\varepsilon) = L_X(\varepsilon)\tau$. □

Now we will claim Frechet differentiability of the mapping $A_1(u(\cdot), \varepsilon)$ in $u(\cdot)$. Let us choose arbitrary $u(\cdot) \in B_{\mathbb{L}_2}(0, \mu)$ and $\delta u(\cdot)$, such that $\|\delta u(\cdot)\|_{\mathbb{L}_2} \leqslant \overline{\mu} - \mu$ and consider

$$\begin{aligned} &A_1(u(\cdot) + \delta u(\cdot), \varepsilon) - A_1(u(\cdot), \varepsilon) \\ &= \int\limits_{t_0}^{T} X(T, \tau)\Big[f\Big(x\big(\tau, \varepsilon, u(\cdot) + \delta u(\cdot)\big), \tau\Big) - f\Big(x\big(\tau, \varepsilon, u(\cdot)\big), \tau\Big)\Big]\, d\tau. \end{aligned} \tag{4.5}$$

Here we should study the difference between solutions of (2.1), produced by $u(\cdot)$ and $u(\cdot)+\delta u(\cdot)$. From (4.2) it follows

$$\begin{aligned} x\big(t, \varepsilon, u(\cdot) + \delta u(\cdot)\big) - x\big(t, \varepsilon, u(\cdot)\big) &= \int\limits_{t_0}^{t} A(\tau)\left[x\big(\tau, \varepsilon, u(\cdot) + \delta u(\cdot)\big) - x\big(\tau, \varepsilon, u(\cdot)\big)\right]\, d\tau \\ + \int\limits_{t_0}^{t} B(\tau)\delta u(\tau)\, d\tau + \varepsilon \int\limits_{t_0}^{t} &\left[f\big(x(\tau, \varepsilon, u(\cdot) + \delta u(\cdot)), \tau\big) - f\big(x(\tau, \varepsilon, u(\cdot)), \tau\big)\right]\, d\tau. \end{aligned} \tag{4.6}$$

Let $y \in \mathbb{R}^n$ and $h \in \mathbb{R}^n$ be chosen such that the inclusions $y \in D$ and $y + h \in D$ are valid. Then, for all $\tau \in [t_0, T]$, using representation of the increment of a function through the integral over a parameter, we have

$$f(y+h,\tau) - f(y,\tau) = \Big(\int\limits_0^1 \frac{\partial f}{\partial x}(y + \xi h, \tau)\ d\xi \Big) h = \frac{\partial f}{\partial x}(y,\tau) h + \omega(y,h,\tau),$$

where

$$\omega(y,h,\tau) = \Big(\int\limits_0^1 \Big[\frac{\partial f}{\partial x}(y + \xi h, \tau) - \frac{\partial f}{\partial x}(y,\tau) \Big]\ d\xi \Big) h.$$

Since D is convex, $y + \xi h \in D$ for all $0 \leqslant \xi \leqslant 1$. Therefore, using Assumption 1, we can obtain the following estimate

$$\|\omega(y,h,\tau)\| \leqslant l_f \Big(\int\limits_0^1 \|\xi h\|\ \ d\xi \Big) h \leqslant \frac{l_f}{2} \|h\|^2.$$

When $y = x\big(\tau, \varepsilon, u(\cdot)\big)$ and

$$h = \Delta x\big(\tau, \varepsilon, \delta u(\cdot)\big) = x\big(\tau, \varepsilon, u(\cdot) + \delta u(\cdot)\big) - x\big(\tau, \varepsilon, u(\cdot)\big),$$

for all $\tau \in [t_0, T]$ we have

$$\begin{gathered} f\big(x(\tau,\varepsilon,u(\cdot)+\delta u(\cdot)),\tau\big) - f\big(x(\tau,\varepsilon,u(\cdot)),\tau\big) \\ = \frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u(\cdot)),\tau\big) \Delta x\big(\tau,\varepsilon,\delta u(\cdot)\big) + \omega\big(x(\tau,\varepsilon,u(\cdot)), \Delta x(\tau,\varepsilon,\delta u(\cdot)),\tau\big), \end{gathered} \tag{4.7}$$

where (see Lemma 2)

$$\big\|\omega\big(x\big(\tau,\varepsilon,u(\cdot)\big), \Delta x\big(\tau,\varepsilon,\delta u(\cdot)\big),\tau\big)\big\| \leqslant \frac{l_f}{2}\ \|\Delta x(\tau,\varepsilon,\delta u(\cdot))\|^2 \leqslant \frac{l_f}{2} L_x^2(\overline{\varepsilon}) \|\delta u(\cdot)\|^2_{\mathbb{L}_2}. \tag{4.8}$$

From (4.7) it follows, that $\omega(x(\tau,\varepsilon,u(\cdot)), \Delta x(\tau,\varepsilon,\delta u(\cdot)),\cdot)$ is measurable, as the sum of measurable functions. Substituting (4.7) to (4.6), we obtain

$$\begin{gathered} \Delta x\big(t,\varepsilon,\delta u(\cdot)\big) = \int\limits_{t_0}^{t} \overline{A}\big(\tau,\varepsilon,u(\cdot)\big) \Delta x\big(\tau,\varepsilon,\delta u(\cdot)\big)\ d\tau + \int\limits_{t_0}^{t} B(\tau)\delta u(\tau)\ d\tau \\ + \varepsilon \int\limits_{t_0}^{t} \omega\Big(x\big(\tau,\varepsilon,u(\cdot)\big), \Delta x\big(\tau,\varepsilon,\delta u(\cdot)\big),\tau\Big) d\tau = \delta x(t) + \Omega(t,\varepsilon,\delta u(\cdot)), \end{gathered}$$

where $\delta x(t)$ is the solution of system (4.3) and

$$\Omega(t,\varepsilon,\delta u(\cdot)) = \varepsilon \int\limits_{t_0}^{t} \omega\Big(x\big(\tau,\varepsilon,u(\cdot)\big), \Delta x\big(\tau,\varepsilon,\delta u(\cdot)\big),\tau\Big) d\tau.$$

Since (4.8) we can estimate $\Omega(t,\varepsilon,\delta u(\cdot))$ from above for all $t \in [t_0, T]$

$$\|\Omega(t,\varepsilon,\delta u(\cdot))\| \leqslant \frac{l_f}{2}\overline{\varepsilon}L_x^2(\overline{\varepsilon})(T-t_0)\|\delta u(\cdot)\|^2_{\mathbb{L}_2}.$$

Here we are going to rewrite (4.7),

$$\begin{gathered} f\big(x(\tau,\varepsilon,u(\cdot)+\delta u(\cdot)),\tau\big) - f\big(x(\tau,\varepsilon,u(\cdot)),\tau\big) \\ = \frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u(\cdot)),\tau\big)\delta x(t) + \frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u(\cdot)),\tau\big)\Omega(t,\varepsilon,\delta u(\cdot)) + \omega\big(x(\tau,\varepsilon,u(\cdot)),\Delta x(\tau,\varepsilon,\delta u(\cdot)),\tau\big). \end{gathered}$$

We can estimate the norm of residial term from above:

$$\left\|\frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u(\cdot)),\tau\big)\Omega(t,\varepsilon,\delta u(\cdot))\right\| \leqslant \frac{l_f}{2}\overline{\varepsilon}L_x^2(\overline{\varepsilon})(T-t_0)\max_{\substack{x\in D\\ \tau\in[t_0,T]}}\left\|\frac{\partial f}{\partial x}(x,\tau)\right\|\|\delta u(\cdot)\|^2_{\mathbb{L}_2},$$

Therefore, we are able to rewrite (4.5) in form

$$A_1(u(\cdot)+\delta u(\cdot),\varepsilon) - A_1(u(\cdot),\varepsilon) = \int\limits_{t_0}^{T} X(T,\tau)\frac{\partial f}{\partial x}\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)\delta x(t)\;d\tau + o(\|\delta u(\cdot)\|^2).$$

This implies, that the Frechet derivative $A_1'(u(\cdot),\varepsilon) : B_{\mathbb{L}_2}(0,\overline{\mu}) \to \mathbb{R}^n$ exists and could be defined by equality

$$A_1'(u(\cdot),\varepsilon)\delta u(\cdot) = \int\limits_{t_0}^{T} X(T,\tau)\frac{\partial f}{\partial x}\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)\delta x(t)\;d\tau \tag{4.9}$$

The Lipschitz continuity of $\delta x(\cdot)$ was proved in Lemma 3. The derivative

$$\frac{\partial f}{\partial x}\Big(x\big(\tau,\varepsilon,u(\cdot)\big),\tau\Big)$$

is Lipschitz continuous as a composition of Lipschitz continuous functions

$$\begin{gathered} \left\|\frac{\partial f\big(x(\tau,\varepsilon,u_1(\cdot)),\tau\big)}{\partial x} - \frac{\partial f\big(x(\tau,\varepsilon,u_2(\cdot)),\tau\big)}{\partial x}\right\| \leqslant l_f\left\|x\big(\tau,\varepsilon,u_1(\cdot)\big) - x\big(\tau,\varepsilon,u_2(\cdot)\big)\right\| \\ \leqslant l_f L_x(\varepsilon)\left\|u_1(\cdot)-u_2(\cdot)\right\|_{\mathbb{L}_2}, \quad \tau\in[t_0,T], \qquad u_1(\cdot),u_2(\cdot)\in B_{\mathbb{L}_2}(0,\mu). \end{gathered}$$

Then the integrand in (4.9) also fulfills the Lipschitz condition for all $\varepsilon\in[0,\overline{\varepsilon}]$ and $\tau\in[t_0,T]$,

$$\begin{gathered} \left\|\frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u_1(\cdot)),\tau\big)\overline{F}'(\tau,\varepsilon,u_1(\cdot))\delta u(\cdot) - \frac{\partial f}{\partial x}\big(x(\tau,\varepsilon,u_2(\cdot)),\tau\big)\overline{F}'(\tau,\varepsilon,u_2(\cdot))\delta u(\cdot)\right\| \\ \leqslant (\overline{\mu}-\mu)\Big(l_f L_x(\varepsilon)\max_{\substack{u(\cdot)\in B_{\mathbb{L}_2}(0,\mu)\\ \tau\in[t_0,T]}}\|\overline{F}'(\tau,\varepsilon,u(\cdot))\| + L_u(\varepsilon)\max_{\substack{x\in D,\\ \tau\in[t_0,T]}}\left\|\frac{\partial f}{\partial x}\Big(x,\tau\Big)\right\|\Big)\left\|u_1(\cdot)-u_2(\cdot)\right\|, \end{gathered}$$

and the whole derivative $A_1'(u(\cdot),\varepsilon)$ will be Lipschitz continuous in $u(\cdot)$.

In order to fulfill the condition of Assumption 2, it remains to show that this derivative will be continuous in ε. This is valid due to the facts that the right-hand side of system (2.1) is linear in the parameter ε and the matrix $\overline{A}\big(t,\varepsilon,u(\cdot)\big)$ of the linearized system (4.3) depends continuously on ε.

Thus, the mapping $A_1(u(\cdot),\varepsilon)$ defined in (4.1) fulfills the condition of Assumption 2 and we are able to formulate the main result of this paper in the following theorem

Theorem 2. *Assume the conditions of Assumption 1 are satisfied, then there exists a positive value ε_0 such that the reachable sets $G(T, \mu, \varepsilon)$ of the quasilinear system (2.1) are convex for all $\varepsilon < \varepsilon_0$.*

P r o o f. The statement's validity can be confirmed by applying Theorem 1 to the mapping F, given that Lipschitz continuity of A_1' and closeness of $G(T, \mu, \varepsilon)$ (Assertion 1) were previously established. □

Remark 1. In the article [2], E.G. Albrecht investigates the support functions of reachable sets for quasilinear systems with integral constraints. The paper defines conditions under which the support functions of reachable sets have continuous dependence on parameter. The author also noted that the continuous dependence of the reachable set on the parameter implies its convexity for small values of parameter. However, no proof of this fact was provided. Furthermore, continuity of reachable sets alone was not sufficient to prove it.

5. Examples

In this section, we present the results of numerical experiments that are intended to illustrate the application of the Theorems 1 and 2.

Example 1. First system under study is Duffing oscillator. We deal with equations

$$\dot{x_1} = x_2, \qquad \dot{x_2} = -x_1 - 10\varepsilon x_1^3 + u, \qquad 0 \leqslant t \leqslant 2 \tag{5.1}$$

describing the motion of a non-linear elastic spring under the influence of an external force u. The impact of the nonlinear elastic force term is determined by the small parameter $\varepsilon > 0$. The initial state is $x_1(0) = x_2(0) = 0$, and the control is bounded by

$$\int_0^2 u^2 dt \leqslant 1. \tag{5.2}$$

When $\varepsilon = 0$, the equations (5.1) describe a linear system with the matrices

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The nonlinear term comprises of a small parameter and the function $f(x) = [-10x^3; 0]$. Condition (2.2) is not fulfilled for this nonlinear term. However, we can use estimates obtained in paper [26] to show, that all the trajectories of the system (5.1) corresponding to admissible controls and zero initial state are lying in a compact set D.

We set

$$v_\varepsilon(t, x) = \frac{5}{2}\varepsilon x_1^4 + \frac{1}{2}x_1^2 + \frac{1}{2}x_2^2$$

and calculate the time derivative

$$\frac{d}{dt}v_\varepsilon(t, x(t)) = \nabla v_\varepsilon(t, x(t))\big(Ax(t) + Bu(t) + \varepsilon f(x(t))\big) = x_2(t)u(t). \tag{5.3}$$

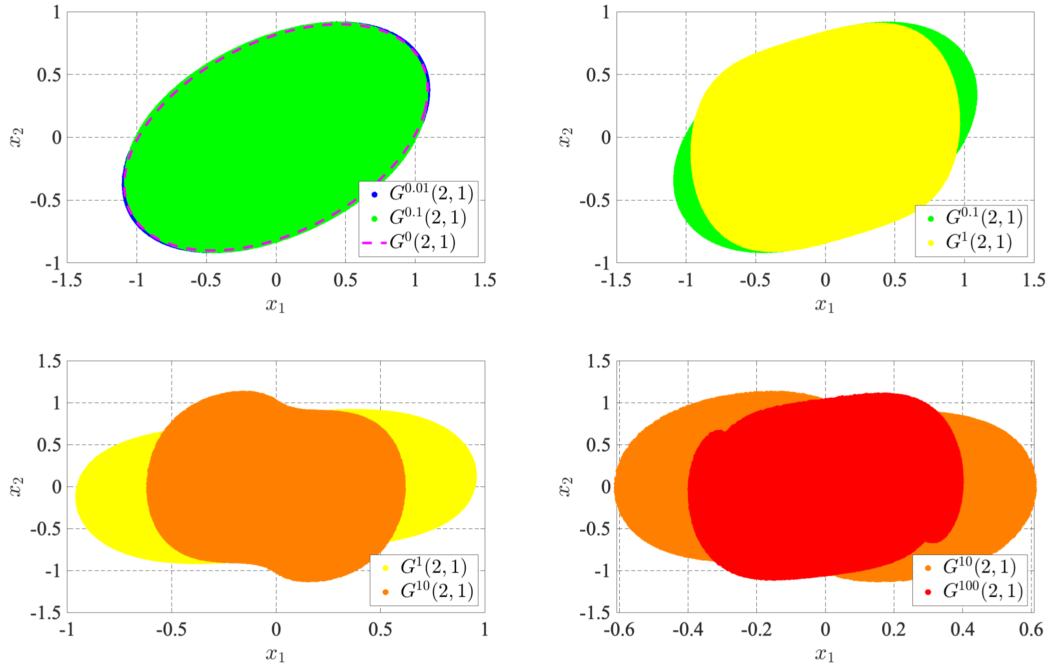


Figure 1. The reachable sets of Duffing oscillator.

For each $\varepsilon \geqslant 0$ and each control $u(\cdot)$ satisfied (5.2), there exists $\tau > 0$, such that the solution of (5.1) generated by this control $u(\cdot)$ and by zero initial state is defined on time interval $[0, \tau]$. Let us integrate (5.3) from 0 to τ. We have

$$v_\varepsilon(\tau, x(\tau)) = \int_0^\tau x_2(t)u(t)\ dt \leqslant \Big(\int_0^2 u^2(t)\ dt\Big)^{1/2} \Big(\int_0^\tau x_2^2(t)\ dt\Big)^{1/2} \leqslant \sqrt{2}\Big(\int_0^\tau v_\varepsilon(t, x(t))\ dt\Big)^{1/2}.$$

Applying comparison theorem to this inequality, one can obtain, that $v_\varepsilon(\tau, x(\tau)) \leqslant \tau$ and, therefore, $\|x(\tau)\|^2 \leqslant 2\tau$. Using well-known technique, we could conclude that any solution (5.1) generated by a control $u(\cdot) \in B_{\mathbb{L}_2}(0, 1)$ and zero initial state, could be continued to time interval $[0, 2]$ and it will belong to the convex set $D = B_{\mathbb{R}^n}(0, 2)$.

The Assumption 1 are fullfilled: the pair (A, B) is a constant; the function f is continuous and continuously differentiable; also, the function f and its derivative $\partial f/\partial x$ satisfy the Lipschitz condition on the set D.

Therefore, the requirements of Theorem 2 are fulfilled for system (5.1), and the corresponding reachable sets should be convex for small parameter values. This is evident in Fig. 1, which demonstrates the constructed reachable sets $G^\varepsilon(T, \mu)$ using numerical Monte–Carlo based technique [24, 25].

It can be seen that sets $G^{0.01}(1, 1)$ and $G^{0.1}(1, 1)$ are close to set $G^0(1, 1)$ constructed for the linear system. One can also see that the sets become non-convex as the parameter ε increases.

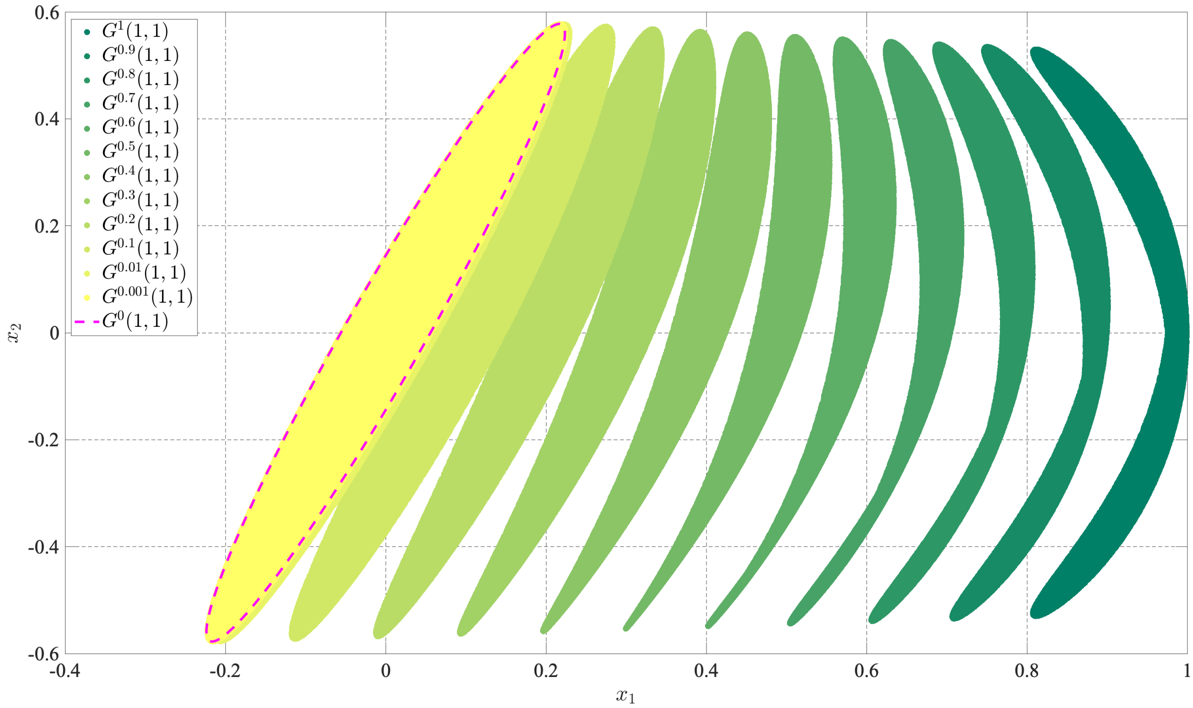


Figure 2. The reachable sets of system (5.4).

Example 2. Second system under study is

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \varepsilon \begin{pmatrix} \cos x_3 - x_2 \\ \sin x_3 - x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} u. \tag{5.4}$$

When $\varepsilon = 0$, the equations (5.4) describe a linear system with matrices

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

and when $\varepsilon = 1$, they describe a unicycle. The nonlinear term comprises of a small parameter and the function

$$f(x) = \begin{pmatrix} \cos x_3 - x_2 \\ \sin x_3 - x_3 \\ 0 \end{pmatrix}.$$

The initial state is zero $x_1(0) = x_2(0) = x_3(0)$, the constraints on the controls are the same as in the first example, but we will consider this system on the time interval $0 \leqslant t \leqslant 1$.

Similar to the previous example, the conditions of Assumption 1 are satisfied, allowing the application of Theorem 2. Fig. 2 displays the projections in the plane (x_1, x_2) of the numerically constructed reachable sets $G^\varepsilon(T, \mu)$ for the system (5.4).

It can be seen that projections of sets $G^{0.001}(1,1)$ and $G^{0.01}(1,1)$ are close to projection of set $G^0(1,1)$ constructed for the linear system. One can also see that the projections of sets become non-convex as the parameter ε increases.

Acknowledgements

The author is grateful to the anonymous reviewers for their careful scrutiny of the manuscript and their constructive feedback, which greatly contributed to its revision.

REFERENCES

1. Al'brekht E. H. The optimal control of the motion of quasilinear systems. *Differ. Uravn.*, 1969. Vol. 5, No. 3. P. 430–442. (in Russian)
2. Al'brekht E. H. The coming together of quasilinear objects in the regular case. *Differ. Uravn.*, 1971. Vol. 7, No. 7. P. 1171–1178. (in Russian)
3. Albrecht E. G. *Metod Lyapunova-Puankare v zadachah optimalnogo upravleniya*. Diss. dokt. fiz.-mat. nauk [Lyapunov-Poincare method in optimal control problems. Dr. phys. and math. sci. diss.]. Sverdlovsk, 1986. 280 p. (in Russian)
4. Calvet J.-P., Arkun Y. Design of P and PI stabilizing controllers for quasi-linear systems. *Comput. Chem. Eng.*, 1990. Vol. 14, No. 4–5. P. 415–426. DOI: 10.1016/0098-1354(90)87017-j
5. Ching Sh., Eun Yo., Gokcek C., Kabamba P. T., Meerkov S. M. *Quasilinear Control: Performance Analysis and Design of Feedback Systems with Nonlinear Sensors and Actuators*. Cambridge: Cambridge University Press, 2010. 282 p. DOI: 10.1017/CBO9780511976476
6. Dauer J. P. Nonlinear perturbations of quasi-linear control systems. *J. Math. Anal. Appl.*, 1976. Vol. 54, No. 3. P. 717–725. DOI: 10.1016/0022-247X(76)90191-8
7. Filippov A. F. *Differential Equations with Discontinuous Righthand Sides*. Dordrecht: Springer, 1988. 304 p. DOI: 10.1007/978-94-015-7793-9
8. Filippov A. F. *Vvedenie v teoriju differencial'nyh uravnenij* [Introduction to the theory of differential equations]. Moscow: Comkniga, 2007. 240 p. (in Russian)
9. Gabasov R. F., Kalinin A. I., Kirillova F. M., Lavrinovich L. I. On asymptotic optimization methods for quasilinear control systems. *Trudy Inst. Mat. Mekh. UrO RAN*, 2019. Vol. 25, No. 3. P. 62–72. DOI: 10.21538/0134-4889-2019-25-3-62-72
10. Guo Y., Kabamba P. T., Meerkov S. M., Ossareh H. R., Tang C. Y. Quasilinear control of wind farm power output. *IEEE Trans. Control Syst. Technol.*, 2015. Vol. 23, No. 4. P. 1555–1562. DOI: 10.1109/TCST.2014.2363431
11. Gusev M. I., Zykov I. V. On Extremal properties of the boundary points of reachable sets for control systems with integral constraints. *Proc. Steklov Inst. Math.*, 2018. Vol. 300, No. Suppl. 1. P. 114–125. DOI: 10.1134/S0081543818020116
12. Gusev M. I. The limits of applicability of the linearization method in calculating small-time reachable sets. *Ural Math. J.*, 2020. Vol. 6, No. 1. P. 71–83. DOI: 10.15826/umj.2020.1.006
13. Gusev M. I., Osipov I. O. Asymptotic behavior of reachable sets on small time intervals. *Proc. Steklov Inst. Math.*, 2020. Vol. 309, Suppl. 1. P. S52–S64. DOI: 10.1134/S0081543820040070
14. Gusev M. I., Osipov I. O. On a local synthesis problem for nonlinear systems with integral constraints. *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2022. Vol. 32, No. 2. P. 171–186. DOI: 10.35634/vm220202
15. Kalinin A. I., Lavrinovich L. I. Asymptotic minimization method of the integral quadratic functional on the trajectories of a quasilinear dynamical system. *Dokl. NAN Belarusi*, 2018. Vol. 62, No. 5. P. 519–524. DOI: 10.29235/1561-8323-2018-62-5-519-524 (in Russian)
16. Kiselev Yu. N. An asymptotic solution of the problem of time-optimal control systems which are close to linear ones. *Soviet Math. Dokl.*, 1968. Vol. 9, No. 5. P. 1093–1097.
17. Krasovskii N. N. *Teoriya upravleniya dvizheniem* [Theory of Control of Motion]. Moscow: Nauka, 1968. 476 p. (in Russian)
18. Kremlev A. G. Control of a quasilinear system under indeterminate initial conditions, *Differ. Uravn.*, 1980. Vol. 16, No. 11. P. 1967–1979. (in Russian)
19. Osipov I. O. On the convexity of the reachable set with respect to a part of coordinates at small time intervals. *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2021. Vol. 31, No. 2. P. 210–225. DOI: 10.35634/vm210204

20. Polyak B. T. Gradient methods for solving equations and inequalities. *USSR Comput. Math. Math. Phys.*, 1964. Vol. 4, No. 6. P. 17–32. DOI: 10.1016/0041-5553(64)90079-5
21. Polyak B. T. Convexity of nonlinear image of a small ball with applications to optimization. *Set-Valued Analysis*, 2001. Vol. 9. P. 159–168. DOI: 10.1023/A:1011287523150
22. Polyak B. T. Convexity of the reachable set of nonlinear systems under L_2 bounded controls. *Dynam. Contin. Discrete Impuls. Systems Ser. A Math. Anal.*, 2004. Vol. 11, Suppl. 2–3. P. 255–267.
23. Subbotin A. I. Control of motion of a quasilinear system. *Differ. Uravn.,* 1967. Vol. 3, No. 7. P. 1113–1118. (in Russian)
24. Zykov I. V., Osipov I. O. A program for constructing the reachable sets of nonlinear systems with integral control constraints by the Monte Carlo method. *Certificates of State Registration of a Computer Program or a Database*, 2020. No. 2020661557.
25. Zykov I. V. An algorithm for constructing reachable sets for systems with multiple integral constraints. In: *Mathematical Analysis with Applications: Int. Conf. CONCORD-90, Pinelas S., Kim A., Vlasov V. (eds.), Ekaterinburg, July 2018*. Springer Proc. Math. Stat., vol. 318. Cham: Springer, 2020. P. 51–60. DOI: 10.1007/978-3-030-42176-2_6
26. Zykov I. V. External estimates of reachable sets for control systems with integral constraints. *Itogi Nauki i Tekhniki. Ser. Sovrem. Mat. Pril. Temat. Obz.*, 2021. Vol. 190. P. 107–114. DOI: 10.36535/0233-6723-2021-190-107-114

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 157–164

DOI: 10.15826/umj.2023.2.013

POLYNOMIALS LEAST DEVIATING FROM ZERO IN $L^p(-1;1)$, $0 \le p \le \infty$, WITH A CONSTRAINT ON THE LOCATION OF THEIR ROOTS[1]

Alena E. Rokina

Ural Federal University,
51 Lenin ave., Ekaterinburg, 620000, Russian Federation

a.e.pestovskaya@mail.ru

We study Chebyshev's problem on polynomials that deviate least from zero with respect to L^p-means on the interval $[-1;1]$ with a constraint on the location of roots of polynomials. More precisely, we consider the problem on the set $\mathcal{P}_n(D_R)$ of polynomials of degree n that have unit leading coefficient and do not vanish in an open disk of radius $R \ge 1$. An exact solution is obtained for the geometric mean (for $p = 0$) for all $R \ge 1$; and for $0 < p < \infty$ for all $R \ge 1$ in the case of polynomials of even degree. For $0 < p < \infty$ and $R \ge 1$, we obtain two-sided estimates of the value of the least deviation.

Keywords: Algebraic polynomials, Chebyshev polynomials, Constraints on the roots of a polynomial.

1. Statement and discussion of the problem

Let

$$D_R := \{z \in \mathbb{C}\colon |z| < R\}$$

be an open disk with center at zero and radius $R > 0$. For $R = 1$, denote by D the unit open disk. Let I be the interval $[-1;1]$.

Denote by $\mathcal{P}_n$ the set of algebraic polynomials of (exact) degree n with complex coefficients and leading coefficient equal to one. A polynomial p_n from $\mathcal{P}_n$ is uniquely defined by its roots z_k, $k = \overline{1,n}$, by the equality

$$p_n(z) = \prod_{k=1}^{n}(z - z_k).$$

Denote by $\mathcal{P}_n(D_R)$ the set of algebraic polynomials from $\mathcal{P}_n$ that do not vanish in an open disk of radius $R > 0$:

$$\mathcal{P}_n(D_R) := \big\{p_n \in \mathcal{P}_n : p_n(z) \ne 0,\ |z| < R\big\}.$$

We use the following notation:

$$\|p_n\|_\infty = \|p_n\|_{C(I)} := \max\big\{|p_n(x)|\colon x \in [-1;1]\big\};$$

$$\|p_n\|_p := \Big(\frac{1}{2}\int_{-1}^{1}|p_n(x)|^p \mathrm{d}x\Big)^{1/p}, \quad 0 < p < \infty;$$

$$\|p_n\|_0 := \exp\Big(\frac{1}{2}\int_{-1}^{1}\ln|p_n(x)|\mathrm{d}x\Big).$$

[1]This work was supported by the Russian Science Foundation, project no. 22-21-00526, https://rscf.ru/project/22-21-00526/ .

For $p \geq 1$, this functional is a norm. It is known (see, e.g., [13]) that

$$\|p_n\|_0 = \lim_{p\to 0} \|p_n\|_p, \quad p_n \in \mathcal{P}_n.$$

In this paper, we study polynomials that deviate least from zero with respect to L^p-means on the interval $[-1;1]$ among all polynomials from the set $\mathcal{P}_n(D_R)$.

Define the *value of the least deviation from zero of polynomials from* $\mathcal{P}_n(D_R)$ *with respect to* L^p*-means on the interval* $[-1;1]$ by the equality

$$\tau_n(I, D_R)_p := \min\big\{\|p_n\|_p \colon p_n \in \mathcal{P}_n(D_R)\big\}. \tag{1.1}$$

The problem is to find quantity (1.1) and polynomials from $\mathcal{P}_n(D_R)$ least deviating from zero on the interval $[-1;1]$, that is, polynomials for which the minimum in (1.1) is attained. It will follow from the further reasoning that the minimum in (1.1) is attained.

The problem on polynomials that deviate least from zero is one of the important problems of approximation theory. In the uniform norm, the problem without constraints on the location of roots was posed and solved by Chebyshev in 1854 [5]. The Chebyshev polynomial of the first kind with unit leading coefficient is extremal in this problem. The polynomial that deviates least from zero in the space $L^1(-1;1)$ was found by E.I. Zolotarev and A.N. Korkin, Chebyshev's disciples, in 1873 (see, for example, [1]). The Chebyshev polynomial of the second kind with unit leading coefficient is extremal. The Legendre polynomials are extremal in the space $L^2(-1;1)$ (see, for example, [20]). Polynomials that deviate least from zero in the space $L^0(-1;1)$ were obtained by Glazyrina in 2005 [8]. Although an explicit form of the polynomials least deviating from zero in spaces $L^p(-1;1)$ for $p \neq 0, 1, 2, \infty$ is unknown, some of their general properties, which can be found in [14, Sects. 2.3–2.4], are useful in studying many important problems of approximation theory.

Note that (as will be seen below), unlike polynomials that deviate least from zero on the interval $[-1;1]$, an extremal polynomial in (1.1) is, generally speaking, not unique.

Studying extremal properties of algebraic polynomials with restrictions on the location of their roots began apparently in 1939 with paper [21] by Turán devoted to inequalities that give a lower estimate for the norm of the derivative of a polynomial in terms of the norm of the polynomial itself. A detailed history of studies of such inequalities can be found in [9, 10].

In 1947, Lax [15] proved the conjecture of P. Erdős. The statement is that, in the classical Bernstein inequality

$$\|p_n'\|_{C(\overline{D})} \leq n\,\|p_n\|_{C(\overline{D})}, \quad p_n \in \mathcal{P}_n,$$

considered on the set $\mathcal{P}_n(D)$ of polynomials that do not vanish in the unit disk, the exact (smallest) constant is half as large (is equal to $n/2$); i.e., the following inequality holds:

$$\|p_n'\|_{C(\overline{D})} \leq \frac{n}{2}\,\|p_n\|_{C(\overline{D})}, \quad p_n \in \mathcal{P}_n(D).$$

The inequality turns into an equality on an arbitrary polynomial having all its roots on the unit circle.

Akopyan [3, Theorem 2] found polynomials in $\mathcal{P}_n(D_R)$, $R > 0$, that deviate least from zero on the unit circle with respect to L^p-norms, $0 \leq p \leq \infty$ (L^p-means for $0 \leq p < 1$). These are polynomials of the form $z^n + \varepsilon R^n$, $|\varepsilon| = 1$.

The sharp Bernstein inequality on the set of polynomials $\mathcal{P}_n(D)$ with respect to L^p-norms on the unit circle was obtained by Lax [15] ($p = 2, \infty$), de Bruijn [6] ($1 \leq p < \infty$), and Rahman and Schmeisser [18] ($0 \leq p < 1$). Arestov obtained [4] a generalization of the Bernstein inequality on the set of polynomials $\mathcal{P}_n(D)$ for rather wide class of operators. The sharp Bernstein inequality on the set of polynomials $\mathcal{P}_n(D_R)$ in the case $p = \infty$ and $R > 1$ was obtained by Malik [16]. Several results for $p = 2$ can be found in Akopyan's paper [2].

Denote by $M_{n,m}(D_R)_p$ the exact (the smallest) constant in the Markov brothers inequality for polynomials from the class $\mathcal{P}_n(D_R)$ with respect to L^p-means on the interval $I = [-1, 1]$:

$$\|p_n^{(m)}\|_p \le M_{n,m}(D_R)_p \, \|p_n\|_p, \quad p_n \in \mathcal{P}_n(D_R), \quad 0 \le p \le \infty, \quad m = 0, 1, \dots, n. \tag{1.2}$$

It is clear that, for $m = n$, the inequality (1.2) is related to problem (1.1); more precisely, the following equality holds:

$$n! = M_{n,n}(D_R)_p \, \tau_n(I, D_R)_p.$$

For the results related to the Markov brothers inequality for $p = \infty$ with constraints on the location of the roots of polynomials, see [7, 12] and the references therein.

In the author's paper [17], the problem on polynomials that deviate least from zero on a compact set K of the complex plane $\mathbb{C}$ with respect to the uniform norm and with a constraint on the location of roots was studied. In particular, a solution to problem (1.1) was found for $p = \infty$ (see Theorem A below). Moreover, the existence of an extremal polynomial was proved, and the problem was reduced to polynomials with roots on the boundary of the domain which gives the constraints.

Similar statements are valid for the more general case of problem (1.1) for $0 \le p \le \infty$. In the following statement, we prove that an extremal polynomial exists for $0 \le p \le \infty$.

Assertion 1. *For $0 \le p \le \infty$, the minimum in problem (1.1) is attained.*

P r o o f of Assertion 1 is performed by the scheme of the proof of Theorem 1 from [17]. Let $q_{n,k}$, $k \in \mathbb{N}$, be an extremal sequence in (1.1), i.e., $\lim\limits_{k\to\infty} \|q_{n,k}\|_p = \tau_n(I, D_R)_p$. Using the different metrics inequality, we get

$$\|q_{n,k}\|_\infty \le c(n)_p \|q_{n,k}\|_p,$$

where the constant $c(n)_p$ is independent of k. The existence of $c(n)_p$ in the case $p \ge 1$ is a well-known fact (the equivalence of norms in finite-dimensional spaces). In the case $0 \le p < 1$, a finite constant also exists, see [9, Lemma 1] for $0 < p < 1$ and [8] for $p = 0$. Then the sequence $q_{n,k}$ is uniformly bounded on $[-1; 1]$. Hence, using the Lagrange interpolation formula, we get its uniformly boundedness on an arbitrary compact set from $\mathbb{C}$.

By the principle of compactness (condensation) in analytic function theory, there exists a subsequence that uniformly converges inside $\mathbb{C}$. It follows from the convergence of coefficients of polynomials of the subsequence that the limiting analytic function is a polynomial. Taking into account the continuity of roots of polynomials as functions of their coefficients and the closedness of $\overline{\mathbb{C}} \setminus D_R$, we conclude that zeros of the limiting polynomial do not belong to D_R. At the same time, the roots of polynomials of the extremal sequence do not tend to infinity, because we get $\tau_n(I, D_R)_p = \infty$ if even one root tends to infinity. Thus, we conclude that the limiting polynomial belongs to $\mathcal{P}_n(D_R)$. The assertion is proved. □

The following statement on the reduction of problem (1.1) to a similar problem for polynomials with roots on a circle is a consequence of a more general Theorem 2 from [17] (see Remark 1). In the particular case considered in the present paper, the proof is simplified. We will give it for the completeness.

Assertion 2. *For $0 \le p \le \infty$ and $R \ge 1$, every extremal polynomial in problem (1.1) has all n roots on the circle of radius R centered at the origin.*

P r o o f. Assume that at least one root of a polynomial $p_n \in \mathcal{P}_n(D_R)$ does not lie on the circle of radius R. Denote it by $z_0 = \rho e^{it}$, where $\rho > R$. Then the polynomial p_n can be represented in the form

$$p_n(x) = p_{n-1}(x)(x - z_0), \quad p_{n-1} \in \mathcal{P}_{n-1}(D_R).$$

Consider the polynomial $q_n(x) = p_{n-1}(x)(x - \widetilde{z}_0)$, where $\widetilde{z}_0 = Re^{it}$. It is clear that $q_n \in \mathcal{P}_n(D_R)$. Since $R < \rho$, we have $|x - \widetilde{z}_0| < |x - z_0|$ for all $x \in [-1;1]$, and hence the following pointwise inequality holds: $|q_n(x)| < |p_n(x)|$, $x \in [-1;1]$. Taking into account the monotonicity of L^p-means, we obtain the inequality $\|q_n\|_p < \|p_n\|_p$. Consequently, the polynomial p_n is not a polynomial from $\mathcal{P}_n(D_R)$ that deviates least from zero on the interval $[-1;1]$ with respect to L^p-means. The assertion is proved. $\square$

The further scheme of presentation in the paper is as follows. In the next two sections, we give a solution to the problem in the two extreme cases $p = \infty$ and $p = 0$. In the last section, we estimate quantity (1.1) from below and above for $0 < p < \infty$. These estimates coincide for polynomials of even degrees, which makes it possible to find an exact value of (1.1) and extremal polynomials.

2. Solution to problem (1.1) in the case $p = \infty$

Let ϱ_n be equal to $1/\sqrt{2}$ if $n = 2m$ and to the unique root of the equation

$$(\varrho^2 - 1)^{2m}(\varrho^2 + 1) = \varrho^{4m+2}$$

in the interval $(1/\sqrt{2}, 1/\sqrt[4]{2})$ if $n = 2m + 1$, $m \ge 1$.

Theorem A. [17, Theorem 3] *The following equality holds*:

$$\tau_n(I, D_R)_\infty = \begin{cases} \sqrt{1 + R^2}, & n = 1, \quad R \ge 0, \\ R^n, & n > 1, \quad R \ge \varrho_n. \end{cases} \tag{2.1}$$

The minimum in (1.1) *is attained on the polynomials*

$$p_n^*(x) = (x^2 - R^2)^m \quad \text{for} \quad n = 2m;$$
$$p_n^*(x) = (x^2 - R^2)^m (x \pm iR) \quad \text{for} \quad n = 2m + 1.$$

The polynomials from $\mathcal{P}_n(D_R)$, given in the theorem, that deviate least from zero on $[-1, 1]$ are not unique. For example, the polynomials

$$p_{2mk}^{**}(x) = (x^{2k} - R^{2k})^m, \quad k, m \in \mathbb{N},$$

are extremal for even n and $R \ge 1/\sqrt[2k]{2}$.

3. Solution to problem (1.1) in the case $p = 0$

In this section, we find an exact solution to problem (1.1) in the case $p = 0$ for $R \ge 1$.

Theorem 1. *The following equality holds for* $R \ge 1$:

$$\tau_n(I, D_R)_0 = \|x + R\|_0^n = \begin{cases} 2^n e^{-n}, & R = 1, \\ e^{-n} \left((R+1)^{(R+1)}/(R-1)^{(R-1)}\right)^{n/2}, & R > 1. \end{cases} \tag{3.1}$$

The polynomials

$$p_n^*(x) = (x - R)^k (x + R)^{n-k}, \quad 0 \le k \le n,$$

are unique extremal polynomials.

P r o o f. According to Assertion 2, it suffices to consider polynomials with roots on the circle. First, we study the case of polynomials of the first degree ($n = 1$). Consider the polynomials $p(x) = x - z_0$, where $|z_0| = |x_0 + iy_0| = R$. The following equalities hold:

$$\ln \|p\|_0 = \frac{1}{2}\int_{-1}^{1} \ln |p(x)| \mathrm{d}x = \frac{1}{2}\int_{0}^{1} (\ln |p(x)| + \ln |p(-x)|) \mathrm{d}x = \frac{1}{2}\int_{0}^{1} \ln((x^2 + R^2)^2 - 4x^2x_0^2) \mathrm{d}x.$$

It is clear that, under the condition $x_0 \in [-R; R]$, the quantity $\|p\|_0$ attains its minimal value only for $x_0 = R$ and $-R$. Thus, in the case $n = 1$, the polynomials $p_1^*(x) = x \pm R$ are extremal. It is not difficult to verify the equalities

$$\|p_1^*\|_0 = e^{-1}\left((R+1)^{(R+1)}/(R-1)^{(R-1)}\right)^{1/2} \quad \text{for} \quad R > 1,$$
$$\|p_1^*\|_0 = 2e^{-1} \quad \text{for} \quad R = 1.$$

Now, let $n > 1$. In view of the multiplicativity of L^0-means, for the polynomial

$$p_n(x) = \prod_{k=1}^{n} (x - z_k),$$

we have

$$\|p_n\|_0 = \exp\left(\frac{1}{2}\int_{-1}^{1} \sum_{k=1}^{n} \ln |x - z_k| \mathrm{d}x\right) = \prod_{k=1}^{n} \exp\left(\frac{1}{2}\int_{-1}^{1} \ln |x - z_k| \mathrm{d}x\right) = \prod_{k=1}^{n} \|x - z_k\|_0.$$

Then the following equality holds for the value of the least deviation:

$$\tau_n(I, D_R)_0 = \prod_{k=1}^{n} \tau_1(I, D_R)_0 = \|x + R\|_0^n.$$

The uniqueness of extremal polynomials of degree n follows from the uniqueness of polynomials for $n = 1$. This proves equality (3.1). □

4. Studying of problem (1.1) in the case $0 < p < \infty$

In this section, we find estimates of quantity (1.1) from below and above for $0 < p < \infty$. These estimates coincide for polynomials of even degrees; hence, we find an exact value of (1.1).

Lemma 1. *The following inequality holds for arbitrary* $0 < p < \infty$, $R \ge 1$, *and positive integer* n:

$$\tau_n(I, D_R)_p \ge \left(\int_0^1 (R^2 - x^2)^{np/2} \mathrm{d}x\right)^{1/p}. \tag{4.1}$$

P r o o f. The following chain of relations holds for an arbitrary polynomial $p_n \in \mathcal{P}_n(D_R)$:

$$\|p_n\|_p^p = \frac{1}{2}\int_{-1}^{1} |p_n(x)|^p \mathrm{d}x = \frac{1}{2}\int_{0}^{1} \left(|p_n(-x)|^p + |p_n(x)|^p\right) \mathrm{d}x \ge \frac{1}{2}\int_{0}^{1} \psi_n(x) \mathrm{d}x,$$

where

$$\psi_n(x) = \min\left\{(|p_n(-x)|^p + |p_n(x)|^p) \colon p_n \in \mathcal{P}_n(D_R)\right\}.$$

Using the inequality of means, we obtain the inequality

$$|p_n(-x)|^p + |p_n(x)|^p \geq 2|p_n(-x)\, p_n(x)|^{p/2}.$$

Consider the absolute value of the product:

$$|p_n(-x)\, p_n(x)| = \Big| \prod_{k=1}^{n} (x^2 - z_k^2) \Big| = |q_n(x^2)|,$$

where $q_n(x) \in \mathcal{P}_n(D_{R^2})$. It follows the inequality

$$\psi_n(x) \geq 2 \min\big\{ |q_n(x^2)|^{p/2} \colon q_n \in \mathcal{P}_n(D_{R^2}) \big\}.$$

The following equality holds for an arbitrary point $z_0 \in D_{R^2}$:

$$\min\big\{ |q_n(z_0)| \colon q_n \in \mathcal{P}_n(D_{R^2}) \big\} = \min\big\{ |z_0 - z|^n \colon |z| = R^2 \big\}.$$

Using this equality, we obtain

$$\psi_n(x) \geq 2(R^2 - x^2)^{np/2}.$$

Consequently, the inequality

$$\|p_n\|_p^p \geq \int_0^1 (R^2 - x^2)^{np/2} \mathrm{d}x$$

holds for an arbitrary polynomial $p_n \in \mathcal{P}_n(D_R)$. This implies estimate (4.1). The lemma is proved. □

Now, we pass to obtaining an upper estimate.

Lemma 2. *The following inequality holds for arbitrary* $0 < p < \infty$, $R \geq 1$, *and positive integer* n:

$$\tau_n(I, D_R)_p \leq \begin{cases} \Big(\dfrac{1}{2} \displaystyle\int_{-1}^{1} (R^2 - x^2)^{mp} \mathrm{d}x \Big)^{1/p}, & n = 2m, \\ \Big(\dfrac{1}{2} \displaystyle\int_{-1}^{1} (R^2 - x^2)^{mp} \cdot (R^2 + x^2)^{p/2} \mathrm{d}x \Big)^{1/p}, & n = 2m + 1. \end{cases} \tag{4.2}$$

P r o o f. We obtain an upper estimate directly from the definition of the value of the least deviation by means of the polynomials

$$p_n(x) = (R^2 - x^2)^m \quad \text{for} \quad n = 2m,$$
$$p_n(x) = (R^2 - x^2)^m (x + iR) \quad \text{for} \quad n = 2m + 1.$$

□

For polynomials of even degrees, the lower and upper estimates coincide, therefore, we obtain an exact solution to problem (1.1) for all $0 < p < \infty$.

Theorem 2. *The following equality holds for* $0 < p < \infty$ *and* $R \geq 1$ *in the case of even* $n = 2m$:

$$\tau_n(I, D_R)_p = \Big(\frac{1}{2} \int_{-1}^{1} (R^2 - x^2)^{mp} \mathrm{d}x \Big)^{1/p}. \tag{4.3}$$

The polynomials $p_{2m}^*(x) = (x^2 - R^2)^m$ *are extremal.*

In the case $p = 2$, we write out a solution for polynomials of small odd degrees.

Theorem 3. *The following equalities hold for $R \ge 1$:*

$$\tau_1(I, D_R)_2 = \Big(R^2 + \frac{1}{3}\Big)^{1/2}, \tag{4.4}$$

$$\tau_3(I, D_R)_2 = \Big(R^6 - \frac{R^4}{3} - \frac{R^2}{5} + \frac{1}{7}\Big)^{1/2}. \tag{4.5}$$

Any polynomial with a root on a circle of radius R, i.e., any polynomial of the form $(x + Re^{it})$, $t \in [0; 2\pi]$, is extremal for $n = 1$. Any polynomial of the form $(x^2 - R^2)(x + Re^{it})$, $t \in [0; 2\pi]$, is extremal for $n = 3$.

P r o o f. According to Assertion 2, we may consider polynomials with roots on the circle. In the case $n = 1$, all polynomials with roots on the circle have the same norm; this implies equality (4.4).

Consider polynomials of the third degree. Let

$$p_3(x) = \prod_{k=1}^{3}(x - z_k),$$

where $|z_k| = |x_k + iy_k| = R$, $k = 1, 2, 3$. Calculating the norm of p_3, we obtain the relation

$$\tau_3(I, D_R)_2^2 = \min_{x_k \in [-R;R]} \bigg(R^6 + R^4 + \frac{3R^2}{5} + \frac{1}{7} + \Big(\frac{4}{5} + \frac{4R^2}{3}\Big)(x_1x_2 + x_2x_3 + x_1x_3)\bigg).$$

Minimizing the function $\sigma(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_1x_3$ in $x_k \in [-R; R]$, we get equality (4.5). The theorem is proved. □

In conclusion, let us give explicit values of the quantity $\tau_{2m}(I, D_R)_p$ for $0 < p < \infty$.

(1) The following equality holds for $R = 1$ and $0 < p < \infty$:

$$\tau_{2m}(I, D_1)_p = \Big(\frac{1}{2}\int_{-1}^{1}(1 - x^2)^{mp}\mathrm{d}x\Big)^{1/p} = \Big(\frac{\sqrt{\pi}}{2}\Gamma\Big(\frac{np+3}{2}\Big)\Big)^{1/p}.$$

(2) The relation

$$\tau_{2m}(I, D_R)_p = \Big(\frac{1}{2}R^{2mp}\int_{-1}^{1}\Big(1 - \frac{x^2}{R^2}\Big)^{mp}\mathrm{d}x\Big)^{1/p} = \Big(R^{2mp} + \sum_{k=1}^{\infty}\frac{(-1)^k C_{mp}^k R^{2(mp-k)}}{2k+1}\Big)^{1/p},$$

where

$$C_{mp}^k = \prod_{l=1}^{k}\frac{mp - l + 1}{l},$$

holds for arbitrary $0 < p < \infty$ and $R > 1$.

Acknowledgements

The author is grateful to her scientific supervisor R.R. Akopyan for setting the problem and useful discussions.

REFERENCES

1. Akhiezer N. I. *Lekcii po teorii approksimacii* [Lectures on Approximation Theory]. Moscow: Nauka, 1965. 408 p. (in Russian)
2. Akopyan R. R. Turán's inequality in H_2 for algebraic polynomials with restrictions to their zeros. *East J. Approx.,* 2000. Vol. 6, No. 1. P. 103–124.
3. Akopyan R. R. Certain extremal problems for algebraic polynomials which do not vanish in a disk. *East J. Approx.*, 2003. Vol. 9, No. 2. P. 139–150.
4. Arestov V. V. Integral inequalities for algebraic polynomials with a restriction on their zeros. *Anal. Math.*, 1991. Vol. 17, No. 1. P. 11–20. DOI: 10.1007/BF02055084
5. Chebyshev P. L. Theory of the mechanisms known as parallelograms. In: *Chebyshev P.L. Collected works. Vol. II: Mathematical Analysis.* Moscow, Leningrad: Acad. Sci. USSR, 1947. P. 23–51. (in Russian)
6. De Bruijn N. G. Inequalities concerning polynomials in the complex domain. *Nederl. Akad. Watensh., Proc.*, 1947. Vol. 50. P. 1265–1272.
7. Erdélyi T. Markov-type inequalities for constrained polynomials with complex coefficients. *Illinois J. Math.*, 1998. Vol. 42, No. 4. P. 544–563. DOI: 10.1215/ijm/1255985460
8. Glazyrina P. Yu. The Markov brothers inequality in L_0-space on an interval. *Math. Notes*, 2005. Vol. 78, No. 1. P. 53–58. DOI: 10.1007/s11006-005-0098-8
9. Glazyrina P. Yu., Révész Sz. Gy. Turán type oscillation inequalities in L^q norm on the boundary of convex domains. *Math. Inequal. Appl.*, 2017. Vol. 20, No. 1. P. 149–180. DOI: 10.7153/mia-20-11
10. Glazyrina P. Yu., Révész Sz. Gy. Turán–Erőd type converse Markov inequalities on general convex domains of the plane in the boundary L^q norm. *Proc. Steklov Inst. Math.*, 2018. Vol. 303. P. 78–104. DOI: 10.1134/S0081543818080084
11. Goluzin G. M. *Geometric Theory of Functions of a Complex Variable.* Transl. Math. Monogr., vol. 26. Providence, R. I.: American Math. Soc., 1969. 676 p.
12. Halász G. Markov-type inequalities for polynomials with restricted zeros. *J. Approx. Theory*, 1999. Vol. 101. P. 148–155.
13. Hardy G. H., Littlewood J. E. Pólya G. *Inequalities.* Cambridge: Cambridge Univ. Press, 1934.
14. Korneychuk N. P., Babenko V. F., Ligun A. A. *Ekstremal'nye svojstva polinomov i splajnov* [Extreme Properties of Polynomials and Splines]. Kyiv: Scientific Opinion, 1992. 304 p. (in Russian)
15. Lax P. D. Proof of the conjecture of P. Erdős on the derivative of a polynomial. *Bull. Amer. Math. Soc.*, 1947. Vol. 50. P. 509–513. DOI: 10.1090/S0002-9904-1944-08177-9
16. Malik M. A. On the derivative of a polynomial. *J. London Math. Soc.*, 1969. Vol. s2-1, No. 1. P. 57–60. DOI: 10.1112/jlms/s2-1.1.57
17. Pestovskaya A. E. Polynomials least deviating from zero with a constraint on the location of roots. *Trudy Inst. Mat. Mekh. UrO RAN*, 2022. Vol. 28, No. 3. P. 166–175. DOI: 10.21538/0134-4889-2022-28-3-166-175
18. Rahman Q. I., Schmeisser G. L^p inequalities for polynomials. *J. Approx. Theory*, 1988. Vol. 53, No. 1. P. 26–33. DOI: 10.1016/0021-9045(88)90073-1
19. Smirnov V. I., Lebedev N. A. *Functions of a Complex Variable (Constructive Theory).* London: Iliffe Books Ltd., 1968. 488 p.
20. Suetin P. K. *Klassicheskie ortogonal'nye mnogochleny* [The Classical Orthogonal Polynomials]. Moscow: Nauka, 1976. 327 p. (in Russian)
21. Turán P. Über die Ableitung von Polynomen. *Compos. Math.*, 1939. Vol. 7. P. 89–95. (in German)

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 165–174

DOI: 10.15826/umj.2023.2.0141

SOME INEQUALITIES BETWEEN THE BEST SIMULTANEOUS APPROXIMATION AND THE MODULUS OF CONTINUITY IN A WEIGHTED BERGMAN SPACE

Muqim S. Saidusainov

University of Central Asia,
155 Qimatsho Imatshoev, Khorog, GBAO, Republic of Tajikistan

muqim.saidusainov@ucentralasia.org

Abstract: Some inequalities between the best simultaneous approximation of functions and their intermediate derivatives, and the modulus of continuity in a weighted Bergman space are obtained. When the weight function is $\gamma(\rho)=\rho^{\alpha},\ \alpha>0$, some sharp inequalities between the best simultaneous approximation and an mth order modulus of continuity averaged with the given weight are proved. For a specific class of functions, the upper bound of the best simultaneous approximation in the space $B_{2,\gamma_1},\ \gamma_1(\rho)=\rho^{\alpha},\ \alpha>0$, is found. Exact values of several n-widths are calculated for the classes of functions $W_p^{(r)}(\omega_m,q)$.

Keywords: The best simultaneous approximation, Modulus of continuity, Upper bound, n-widths.

1. Introduction

Extremal problems of polynomial approximation of functions in a Bergman space were studied, for example, in [8, 13–15]. Here, we will continue our research in this direction and study the simultaneous approximation of functions and their intermediate derivatives in a weighted Bergman space based on the works [4–6, 10]. Note that the problem of simultaneous approximation of periodic functions and their intermediate derivatives by trigonometric polynomials in the uniform metric was studied by Garkavi [1]. In the case of entire functions, this problem was studied by Timan [12].

To solve the problem, we first will prove an analog of Ligun's inequality [2].

Let us introduce the necessary definitions and notation to formulate our results. Let

$$U:=\{z\in\mathbb{C}:|z|<1\}$$

be the unit disk in $\mathbb{C}$, and let $\mathcal{A}(U)$ be the set of functions analytic in the disk U. Denote by $B_{2,\gamma}$ the weighted Bergman space of analytic functions $f\in\mathcal{A}(U)$ such that [8]

$$\|f\|_{2,\gamma}:=\left(\frac{1}{2\pi}\iint_{(U)}|f(z)|^2\gamma(|z|)d\sigma\right)^{1/2}<\infty, \tag{1.1}$$

$d\sigma$ is an area element, $\gamma:=\gamma(|z|)$ is a nonnegative measurable function that is not identically zero, and the integral is understood in the Lebesgue sense. It is obvious, that the norm (1.1) can be written in the form

$$\|f\|_{2,\gamma}=\left(\frac{1}{2\pi}\int_0^1\int_0^{2\pi}\rho\gamma(\rho)|f(\rho e^{it})|^2d\rho dt\right)^{1/2}.$$

In the particular case of $\gamma \equiv 1$, $B_q := B_{q,1}$ is the usual Bergman space. The mth order modulus of continuity in $B_{2,\gamma}$ is defined as

$$\omega_m(f,t)_{2,\gamma} = \sup\left\{\|\Delta_m(f,\cdot,\cdot,h)\|_{2,\gamma} \colon |h| \le t\right\} =$$
$$= \sup\left\{\left(\frac{1}{2\pi}\int_0^1\int_0^{2\pi}\rho\gamma(\rho)|\Delta_m(f;\rho,u,h)|^2 d\rho du\right)^{1/2} \colon |h| \le t\right\},$$

where

$$\Delta_m(f;\rho,u,h) = \sum_{k=0}^{m}(-1)^k C_m^k f\left(\rho e^{i(u+kh)}\right).$$

Let $\mathcal{P}_n$ be the set of complex polynomials of order at most n. Consider the best approximation of functions $f \in B_{2,\gamma}$:

$$E_{n-1}(f)_{2,\gamma} = \inf\left\{\|f - p_{n-1}\|_{2,\gamma} \colon p_{n-1} \in \mathcal{P}_{n-1}\right\}$$

Denote by $\mathscr{B}_{2,\gamma}^{(r)}$ and $\mathscr{B}_2^{(r)}$, $r \in \mathbb{N}$ the class of functions $f \in \mathcal{A}(U)$ whose rth order derivatives

$$f^{(r)}(z) = d^r f/dz^r$$

belong to the spaces $B_{2,\gamma}$ and B_2, respectively. Define

$$\alpha_{n,r} = n(n-1)\cdots(n-r+1), \quad n > r.$$

It is well known [7, 8] that the best approximation of functions

$$f = \sum_{k=0}^{\infty} c_k(f) z^k \in B_{2,\gamma}$$

is equal to

$$\begin{aligned} E_{n-1}(f)_{2,\gamma} &= \left(\sum_{k=n}^{\infty}|c_k(f)|^2\int_0^1\rho^{2k+1}\gamma(\rho)d\rho\right)^{1/2}, \\ E_{n-s-1}\left(f^{(s)}\right)_{2,\gamma} &= \left(\sum_{k=n}^{\infty}|c_k(f)|^2\,\alpha_{k,s}^2\int_0^1\rho^{2(k-s)+1}\gamma(\rho)d\rho\right)^{1/2}, \end{aligned} \tag{1.2}$$

and the modulus of continuity of $f \in B_{2,\gamma}$ is

$$\omega_m\left(f^{(r)},t\right)_{2,\gamma} = 2^{m/2}\sup_{|h|\le t}\left\{\sum_{k=r}^{\infty}\alpha_{k,r}^2|c_k(f)|^2(1-\cos(k-r)h)^m\int_0^1\rho^{2(k-r)+1}\gamma(\rho)d\rho\right\}^{1/2}. \tag{1.3}$$

Denote by

$$\mu_s(\gamma) = \int_0^1\gamma(\rho)\rho^s d\rho, \quad s = 0,1,2,\ldots \tag{1.4}$$

the moments of order s of the weight function $\gamma(\rho)$ on $[0,1]$. According to notation (1.4), we write equalities (1.2) and (1.3) in compact form:

$$\begin{aligned} E_{n-1}(f)_{2,\gamma} &= \left(\sum_{k=n}^{\infty}|c_k(f)|^2\mu_{2k+1}(\gamma)\right)^{1/2}, \\ E_{n-s-1}\left(f^{(s)}\right)_{2,\gamma} &= \left(\sum_{k=n}^{\infty}|c_k(f)|^2\,\alpha_{k,s}^2\,\mu_{2(k-s)+1}(\gamma)\right)^{1/2}, \\ \omega_m\left(f^{(r)},t\right)_{2,\gamma} &= 2^{m/2}\sup_{|h|\le t}\left\{\sum_{k=r}^{\infty}\alpha_{k,r}^2|c_k(f)|^2(1-\cos(k-r)h)^m\,\mu_{2(k-r)+1}(\gamma)\right\}^{1/2}. \end{aligned} \tag{1.5}$$

2. Analog of Ligun's inequality

For compact statement of the results, we introduce the following extremal characteristic:

$$\mathscr{K}_{m,n,r,s,p}(q,\gamma,h)=\sup_{f\in\mathscr{B}_{2,\gamma}^{(r)}}\frac{2^{m/2}E_{n-s-1}\left(f^{(s)}\right)_{2,\gamma}}{\left(\int_0^h\omega_m^p(f^{(r)},t)_{2,\gamma}q(t)dt\right)^{1/p}},$$

where $m,n\in\mathbb{N}$, $r\in\mathbb{Z}_+$, $n>r\geq s$, $0<p<2$, $0<h\leq\pi/(n-r)$, and $q(t)$ is a real, nonnegative, measurable weight function that is not identically zero on $[0,h]$.

Theorem 1. *Let $k,m,n\in\mathbb{N}$, $r,s\in\mathbb{Z}_+$, $k>n>r\geq s$, $0<p<2$, $0<h\leq\pi/(n-r)$, and let $q(t)$ be a nonnegative, measurable function that is not identically zero on $[0,h]$. Then*

$$\frac{1}{\mathscr{L}_{n,r,s,p}(q,\gamma,h)}\leq\mathscr{K}_{m,n,r,s,p}(q,\gamma,h)\leq\frac{1}{\inf\limits_{n\leq k<\infty}\mathscr{L}_{k,r,s,p}(q,\gamma,h)},\tag{2.1}$$

where

$$\mathscr{L}_{k,r,s,p}(q,\gamma,h)=\frac{\alpha_{k,r}}{\alpha_{k,s}}\left(\frac{\mu_{2(k-r)+1}(\gamma)}{\mu_{2(k-s)+1}(\gamma)}\right)^{1/2}\left(\int_0^h(1-\cos(k-r)t)^{mp/2}q(t)dt\right)^{1/p}.$$

P r o o f. Consider the simplified variant of Minkowski's inequality [3, p. 104]:

$$\left(\int_0^h\left(\sum_{k=n}^{\infty}|g_k(t)|^2\right)^{p/2}dt\right)^{1/p}\geq\left(\sum_{k=n}^{\infty}\left(\int_0^h|g_k(t)|^pdt\right)^{2/p}\right)^{1/2},\tag{2.2}$$

which is hold for all $0<p\leq 2$ and $h\in\mathbb{R}_+$. Setting

$$g_k=f_kq^{1/p}\quad(0<p\leq 2)$$

in (2.2), we get

$$\left(\int_0^h\left(\sum_{k=n}^{\infty}|f_k(t)|^2\right)^{p/2}q(t)dt\right)^{1/p}\geq\left(\sum_{k=n}^{\infty}\left(\int_0^h|f_k(t)|^pq(t)dt\right)^{2/p}\right)^{1/2}.\tag{2.3}$$

From (1.3) with respect to (2.3), we get

$$\begin{aligned}\left\{\int_0^h\omega_m^p\left(f^{(r)},t\right)_{2,\gamma}q(t)dt\right\}^{1/p}&=\left\{\int_0^h\left(\omega_m^2(f^{(r)},t)_{2,\gamma}\right)^{p/2}q(t)dt\right\}^{1/p}\\ \geq\left\{\int_0^h\left(2^m\sum_{k=n}^{\infty}\alpha_{k,r}^2|c_k(f)|^2(1-\cos(k-r)t)^m\mu_{2(k-r)+1}(\gamma)\right)^{p/2}q(t)dt\right\}^{1/p}&\\ \geq\left\{\sum_{k=n}^{\infty}\left[2^{mp/2}\alpha_{k,r}^p|c_k(f)|^p\int_0^h(1-\cos(k-r)t)^{mp/2}\left(\mu_{2(k-r)+1}(\gamma)\right)^{p/2}q(t)dt\right]^{2/p}\right\}^{1/2}&\\ =2^{m/2}\left\{\sum_{k=n}^{\infty}|c_k(f)|^2\mu_{2(k-r)+1}(\gamma)\left[\alpha_{k,r}^p\int_0^h(1-\cos(k-r)t)^{mp/2}q(t)dt\right]^{2/p}\right\}^{1/2}&\\ =2^{m/2}\left\{\sum_{k=n}^{\infty}|c_k(f)|^2\alpha_{k,s}^2\,\mu_{2(k-s)+1}(\gamma)\mu_{2(k-r)+1}(\gamma)\left(\mu_{2(k-s)+1}(\gamma)\right)^{-1}\right.&\end{aligned}$$

$$\left[\left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p\int_0^h(1-\cos(k-r)t)^{mp/2}q(t)dt\right]^{2/p}\Bigg\}^{1/2}$$

$$\geq 2^{m/2}\inf_{n\leq k<\infty}\left\{\frac{\alpha_{k,r}}{\alpha_{k,s}}\left(\frac{\mu_{2(k-r)+1}(\gamma)}{\mu_{2(k-s)+1}(\gamma)}\right)^{1/2}\left(\int_0^h(1-\cos(k-r)t)^{mp/2}q(t)dt\right)^{1/p}\right\}$$

$$\times\left(\sum_{k=n}^{\infty}|c_k(f)|^2\alpha_{k,s}^2\,\mu_{2(k-s)+1}(\gamma)\right)^{1/2}=2^{m/2}E_{n-s-1}(f^{(s)})_{2,\gamma}\inf_{n\leq k<\infty}\mathscr{L}_{k,r,s,p}(q,\gamma,h),$$

and this yields the inequality

$$\frac{2^{m/2}E_{n-s-1}\left(f^{(s)}\right)_{2,\gamma}}{\left(\int_0^h\omega_m^p(f^{(r)},t)_{2,\gamma}q(t)dt\right)^{1/p}}\leq\frac{1}{\inf\limits_{n\leq k<\infty}\mathscr{L}_{k,r,s,p}(q,\gamma,h)} \tag{2.4}$$

or

$$\mathscr{K}_{m,n,r,s,p}(q,\gamma,h)\leq\frac{1}{\inf\limits_{n\leq k<\infty}\mathscr{L}_{k,r,s,p}(q,\gamma,h)}. \tag{2.5}$$

To estimate the value in (2.1) from below, consider the function

$$f_0(z)=z^n\in\mathscr{B}_{2,\gamma}^{(r)}.$$

Simple calculation leads to the following relations:

$$E_{n-s-1}(f_0^{(s)})_{2,\gamma}=\alpha_{n,s}\left(\int_0^1\rho^{2(n-s)+1}\gamma(\rho)d\rho\right)^{1/2}=\alpha_{n,s}\left(\mu_{2(n-s)+1}(\gamma)\right)^{1/2},$$

$$\omega_m^2\big(f_0^{(r)},t\big)_{2,\gamma}=2^m\alpha_{n,r}^2\big(1-\cos(n-r)t\big)^m\int_0^1\rho^{2(n-r)+1}\gamma(\rho)d\rho$$
$$=2^m\alpha_{n,r}^2\big(1-\cos(n-r)t\big)^m\mu_{2(n-r)+1}(\gamma),$$

using which, we get the lower estimate

$$\mathscr{K}_{m,n,r,p}(q,\gamma,h)\geq\frac{2^{m/2}E_{n-s-1}\big(f_0^{(s)}\big)_{2,\gamma}}{\left(\int_0^h\omega_m^p(f_0^{(r)},t)_{2,\gamma}q(t)dt\right)^{1/p}}$$
$$=\frac{2^{m/2}\alpha_{n,s}\left(\mu_{2(n-s)+1}(\gamma)\right)^{1/2}}{\left(2^{mp/2}\alpha_{n,r}^p\left(\mu_{2(n-r)+1}(\gamma)\right)^{p/2}\int_0^h(1-\cos(n-r)t)^{mp/2}q(t)dt\right)^{1/p}}=\frac{1}{\mathscr{L}_{n,r,s,p}(q,\gamma,h)}. \tag{2.6}$$

Comparing the upper estimate (2.5) and the lower estimate (2.6), we obtain the required two-sided inequality (2.1). This completes the proof of Theorem 1. □

Corollary 1. *The following two-sided inequality holds for $\gamma_1(\rho)=\rho^\alpha$, $\alpha\geq 0$, in Theorem 1:*

$$\frac{1}{\mathscr{G}_{n,r,s,p,\alpha}(q,h)}\leq\mathscr{K}_{m,n,r,s,p}(q,\gamma_1,h)\leq\frac{1}{\inf\limits_{n\leq k<\infty}\mathscr{G}_{k,r,s,p,\alpha}(q,h)}, \tag{2.7}$$

where

$$\mathscr{G}_{k,r,s,p,\alpha}(q,h)=\frac{\alpha_{k,r}}{\alpha_{k,s}}\left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{1/2}\left(\int_0^h(1-\cos(k-r)t)^{mp/2}\,q(t)dt\right)^{1/p}. \tag{2.8}$$

The following problem naturally arises from (2.7): to find an exact upper bound for the extremal characteristic

$$\mathscr{K}_{m,n,r,s,p}(q,\gamma_1,h) = \sup_{f\in\mathscr{B}_{2,\gamma_1}^{(r)}} \frac{2^{m/2}E_{n-s-1}(f^{(s)})_{2,\gamma_1}}{\left(\int_0^h \omega_m^p(f^{(r)},t)_{2,\gamma_1}q(t)dt\right)^{1/p}},$$

where $m,n\in\mathbb{N}$, $r,s\in\mathbb{Z}_+$, $n>r\geq s$, $0<p<2$, $0<h\leq\pi/(n-r)$, $\gamma_1(\rho)=\rho^\alpha$, and $\alpha\geq 0$.

Theorem 2. *Let a weight function $q(t)$, $t\in[0,h]$, be continuous and differentiable on the interval. If the differential inequality*

$$\left(\sum_{l=s}^{r-1}\frac{p}{k-l} - \frac{2p(r-s)}{[2(k-r+1)+\alpha](2(k-s+1)+\alpha)} - \frac{1}{k-r}\right)q(t) - \frac{1}{k-r}tq'(t)\geq 0 \tag{2.9}$$

holds for all $k\in\mathbb{N}$, $r,s\in\mathbb{Z}_+$, $k>n>r\geq s$, $0<p\leq 2$, and $\alpha\geq 0$, then the following equality holds for all $m,n\in\mathbb{N}$ and $0<h\leq\pi/(n-r)$:

$$\mathscr{K}_{m,n,r,s,p}(q,\gamma_1,h) = \frac{\alpha_{n,s}}{\alpha_{n,r}}\left(\frac{2(n-r+1)+\alpha}{2(n-s+1)+\alpha}\right)^{1/2}\left(\int_0^h (1-\cos(n-r)t)^{mp/2}q(t)dt\right)^{1/p}. \tag{2.10}$$

P r o o f. To prove equality (2.10), it suffices to show that the following equality holds in (2.7):

$$\inf_{n\leq k<\infty}\mathscr{G}_{k,r,s,p,\alpha}(q,h) = \mathscr{G}_{n,r,s,p,\alpha}(q,h). \tag{2.11}$$

We should note that a similar problem of finding a lower bound in (2.11) for some specific weights for $p=2$ was considered in [2]. In the general case, this problem was studied in [9], where it was proved that, if the weight function $q\in C^{(1)}[0,h]$ for $1/r<p\leq 2$, $r\geq 1$, and $0<t\leq h$ satisfies the differential equation

$$(rp-1)q(t)-tq'(t)\geq 0,$$

then (2.11) holds.

Let us now show that, under all constrains on the parameters k, r, s, m, p, α, and h in Theorem 2, the function

$$\psi(k) = \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p\left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2}\int_0^h \left(1-\cos(k-r)t\right)^{mp/2}q(t)dt \tag{2.12}$$

increases for $n\leq k<\infty$. Indeed, differentiating (2.12) and using the identity

$$\frac{d}{dk}(1-\cos(k-r)t)^{mp/2} = \frac{t}{k-r}\frac{d}{dt}(1-\cos(k-r)t)^{mp/2},$$

we obtain

$$\psi'(k) = \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p\sum_{l=s}^{r-1}\frac{p}{k-l}\left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2}\int_0^h (1-\cos(k-r)t)^{mp/2}q(t)dt$$
$$+\left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p\frac{p}{2}\left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2-1}\frac{4s-4r}{[2(k-r+1)+\alpha]^2}\int_0^h (1-\cos(k-r)t)^{mp/2}q(t)dt$$
$$+\left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p\left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2}\int_0^h \frac{d}{dk}(1-\cos(k-r)t)^{mp/2}q(t)dt$$

$$= \int_0^h (1-\cos(k-r)t)^{mp/2} q(t)dt \bigg\{ \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p \sum_{l=s}^{r-1} \frac{p}{k-l} \left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2}$$

$$- \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p \frac{2p(r-s)}{[2(k-r+1)+\alpha](2(k-s+1)+\alpha)} \left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2} \bigg\}$$

$$+ \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p \left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2} \int_0^h \frac{t}{k-r} \frac{d}{dt}(1-\cos(k-r)t)^{mp/2} q(t)dt$$

$$= \left(\frac{\alpha_{k,r}}{\alpha_{k,s}}\right)^p \left(\frac{2(k-s+1)+\alpha}{2(k-r+1)+\alpha}\right)^{p/2} \bigg\{ \frac{h}{k-r}(1-\cos(k-r)h)^{mp/2} q(h) + \int_0^h (1-\cos(k-r)t)^{mp/2}$$

$$\times \bigg[\bigg(\sum_{l=s}^{r-1} \frac{p}{k-l} - \frac{2p(r-s)}{[2(k-r+1)+\alpha](2(k-s+1)+\alpha)} - \frac{1}{k-r} \bigg) q(t) - \frac{1}{k-r} t q'(t) \bigg] dt \bigg\}.$$

This relation and condition (2.9) imply that $\psi(k) > 0$, $k \geq n > r \geq s$, and we obtain equality (2.10). Theorem 2 is proved. □

Denote by $W_p^{(r)}(\omega_m, q)$ $(r \in \mathbb{Z}_+,\ 0 < p \leq 2)$ the set of functions $f \in \mathscr{B}_{2,\gamma_1}^{(r)}$ whose rth derivatives $f^{(r)}$ satisfy the following condition for all $0 < h \leq \pi/(n-r)$ and $n > r$:

$$\int_0^h \omega_m^p \big(f^{(r)}, t\big)_{2,\gamma_1} q(t)dt \leq 1.$$

Since, for $f \in \mathscr{B}_{2,\gamma_1}^{(r)}$, its intermediate derivatives $f^{(s)}$ $(1 \leq s \leq r-1)$ also belong to L_2, the behavior of the value $E_{n-s-1}\left(f^{(s)}\right)_2$ for some classes $\mathfrak{M}^{(r)} \subset \mathscr{B}_{2,\gamma_1}^{(r)}$, $n > r \geq s$, $n \in \mathbb{N}$, and $r, s \in \mathbb{Z}_+$, is of interest. More precisely, it is required to find the value

$$\mathscr{A}_{n,s}\big(\mathfrak{M}^{(r)}\big) := \sup \big\{ E_{n-s-1}(f^{(s)})_{2,\gamma_1} :\ f \in \mathfrak{M}^{(r)} \big\}.$$

Corollary 2. *The following equality holds for all $n \in \mathbb{N}$, $n > r \geq s$, $0 < p \leq 2$, and $0 < h \leq \pi/(n-r)$:*

$$\mathscr{A}_{n,s}\big(W_p^{(r)}(\omega_m, q)\big) := \sup \Big\{ E_{n-s-1}(f^{(s)})_{2,\gamma_1} :\ f \in W_p^{(r)}(\omega_m, q) \Big\} = \frac{1}{2^{m/2} \mathscr{G}_{n,r,s,p,\alpha}(q,h)}. \tag{2.13}$$

Moreover, there is a function $g_0 \in W_p^{(r)}(\omega_m, q)$ on which the upper bound in (2.13) is attained.

P r o o f. Assuming that $\gamma = \gamma_1(\rho) = \rho^\alpha$ in (2.4), with respect to (2.8), we can write

$$E_{n-s-1}(f^{(s)})_{2,\gamma_1} \leq \frac{\left(\int_0^h \omega_m^p(f^{(r)}, t)_{2,\gamma_1} q(t)dt\right)^{1/p}}{2^{m/2} \inf\limits_{n \leq k < \infty} \mathscr{L}_{k,r,s,p}(q, \gamma_1, h)} = \frac{\left(\int_0^h \omega_m^p(f^{(r)}, t)_{2,\gamma_1} q(t)dt\right)^{1/p}}{2^{m/2} \inf\limits_{n \leq k < \infty} \mathscr{G}_{k,r,s,p,\alpha}(q,h)}.$$

Using equality (2.11) and the definition of the class $W_p^{(r)}(\omega_m, q)$, we get

$$E_{n-s-1}(f^{(s)})_{2,\gamma_1} \leq \frac{1}{2^{m/2} \mathscr{G}_{n,r,s,p,\alpha}(q,h)}. \tag{2.14}$$

From (2.14), it follows the upper estimate of the value on the left-hand side of (2.13):

$$\mathscr{A}_{n,s}\left(W_p^{(r)}(\omega_m; q, \Phi)\right) \leq \frac{1}{2^{m/2} \mathscr{G}_{n,r,s,p,\alpha}(q,h)}. \tag{2.15}$$

To obtain the lower estimate for this value, consider the function

$$g_0(z) = \frac{\sqrt{2(n-r+1)+\alpha}}{2^{m/2}\alpha_{n,r}} \left(\int_0^h (1-\cos(n-r)t)^{mp/2}\, q(t)dt \right)^{-1/p} z^n$$

and show that g_0 belongs to $W_p^{(r)}(\omega_m, q)$. Differentiating this function r times, we obtain

$$g_0^{(r)}(z) = \sqrt{\frac{2(n-r+1)+\alpha}{2^m}} \left(\int_0^h (1-\cos(n-r)t)^{mp/2}\, q(t)dt \right)^{-1/p} z^{n-r}.$$

Using this equality and formulas (1.3), we get

$$\omega_m \left(g_0^{(r)}, t \right)_{2,\gamma_1} = \frac{[1-\cos(n-r)t]^{m/2}}{\left(\int_0^h (1-\cos(n-r)t)^{mp/2}\, q(t)dt \right)^{1/p}}.$$

Raising both sides of this inequality to a power p $(0 < p \le 2)$, multiplying them by the weight function $q(t)$, and integrating with respect to t from 0 to h, we obtain

$$\int_0^h \omega_m^p (g_0^{(r)}, t)_{2,\gamma_1} q(t)dt = 1$$

or, equivalently,

$$\left(\int_0^h \omega_m^p (g_0^{(r)}, t)_{2,\gamma_1} q(t)dt \right)^{1/p} = 1.$$

Thus, the inclusion $g_0 \in W_p^{(r)}(\omega_m, q)$ is proved.

Since the relation

$$g_0^{(s)}(z) = \sqrt{\frac{2(n-r+1)+\alpha}{2^m}} \frac{\alpha_{n,s}}{\alpha_{n,r}} \left(\int_0^h (1-\cos(n-r)t)^{mp/2}\, q(t)dt \right)^{-1/p} z^{n-s}$$

holds for all $0 \le s \le r < n$, $n \in \mathbb{N}$, and $r, s \in \mathbb{Z}_+$, according to (1.5), we have

$$\begin{aligned} E_{n-s-1}\left(g_0^{(s)}\right)_{2,\gamma_1} &= \frac{1}{2^{m/2}} \frac{\alpha_{n,s}}{\alpha_{n,r}} \sqrt{\frac{2(n-r+1)+\alpha}{2(n-s+1)+\alpha}} \left(\int_0^h [1-\cos(n-r)t]^{mp}\, q(t)dt \right)^{-1/p} \\ &= \frac{1}{2^{m/2}\, \mathscr{G}_{n,r,s,p,\alpha}(q,h)}. \end{aligned}$$

Using this equality, we obtain the lower estimate

$$\sup \left\{ E_{n-s-1}(f^{(s)})_{2,\gamma_1} :\ f \in W_p^{(r)}(\Omega_m, q) \right\} \ge E_{n-s-1}(g_0^{(s)})_{2,\gamma_1} = \frac{1}{2^{m/2}\, \mathscr{G}_{n,r,s,p,\alpha}(q,h)}. \tag{2.16}$$

Comparing the upper estimate (2.15) and the lower estimate (2.16), we obtain the required equality (2.13). $\square$

3. Exact values of n-widths for the classes $W_p^{(r)}(\omega_m, q)$ $(r \in \mathbb{Z}_+,\ 0 < p \le 2)$

Recall definitions and notation needed in what follows. Let X be a Banach space, let S be the unit ball in X, let $\Lambda_n \subset X$ be an n-dimensional subspace, let $\Lambda^n \subset X$ be a subspace of codimension n, let $\mathscr{L} : X \to \Lambda_n$ be a continuous linear operator, let $\mathscr{L}^\perp : X \to \Lambda_n$ be a continuous linear projection operator, and let $\mathfrak{M}$ be a convex centrally symmetric subset of X. The quantities

$$\begin{gathered}
b_n(\mathfrak{M}, X) = \sup\big\{ \sup\{\varepsilon > 0;\ \varepsilon S \cap \Lambda_{n+1} \subset \mathfrak{M}\} : \Lambda_{n+1} \subset X\big\},\\
d_n(\mathfrak{M}, X) = \inf\big\{ \sup\{\inf\{\|f-g\|_X : g \in \Lambda_n\} : f \in \mathfrak{M}\} : \Lambda_n \subset X\big\},\\
\delta_n(\mathfrak{M}, X) = \inf\big\{ \inf\{\sup\{\|f-\mathscr{L}f\|_X : f \in \mathfrak{M}\} : \mathscr{L}X \subset \Lambda_n\} : \Lambda_n \subset X\big\},\\
d^n(\mathfrak{M}, X) = \inf\big\{ \sup\{\|f\|_X : f \in \mathfrak{M} \cap \Lambda^n\} : \Lambda^n \subset X\big\},\\
\Pi_n(\mathfrak{M}, X) = \inf\big\{ \inf\{\sup\{\|f-\mathscr{L}^\perp f\|_X : f \in \mathfrak{M}\} : \mathscr{L}^\perp X \subset \Lambda_n\} : \Lambda_n \subset X\big\}
\end{gathered}$$

are called the *Bernstein, Kolmogorov, linear, Gelfand, and projection n-widths* of a subset $\mathfrak{M}$ in the space X, respectively. These n-widths are monotone in n and related as follows in a Hilbert space X (see, e.g., [3, 11]):

$$b_n(\mathfrak{M}, X) \le d^n(\mathfrak{M}, X) \le d_n(\mathfrak{M}, X) = \delta_n(\mathfrak{M}, X) = \Pi_n(\mathfrak{M}, X). \tag{3.1}$$

For an arbitrary subset $\mathfrak{M} \subset X$, we set

$$E_{n-1}(\mathfrak{M})_X := \sup\big\{E_{n-1}(f)_2 \colon f \in \mathfrak{M}\big\}.$$

Theorem 3. *The following equalities hold for all $m, n \in \mathbb{N}$, $r \in \mathbb{Z}_+$, $n > r$, and $0 \le h \le \pi/(n-r)$:*

$$\begin{aligned}
&\lambda_n(W_p^{(r)}(\omega_m, q), B_{2,\gamma_1}) = E_{n-1}(W_p^{(r)}(\omega_m, q), B_{2,\gamma_1})\\
&= \frac{1}{2^{m/2}\alpha_{n,r}}\sqrt{\frac{2(n-r+1)+\alpha}{2(n+1)+\alpha}}\left(\int_0^h [1-\cos(n-r)t]^{mp}\, q(t)dt\right)^{-1/p},
\end{aligned} \tag{3.2}$$

where $\lambda_n(\cdot)$ is any of the n-widths $b_n(\cdot)$, $d_n(\cdot)$, $d^n(\cdot)$, $\delta_n(\cdot)$, and $\Pi_n(\cdot)$.

P r o o f. We obtain the upper estimates of all n-widths for the class $W_p^{(r)}(\omega_m, q)$ with $s = 0$ from (2.14) since

$$\begin{aligned}
&E_{n-1}\big(W_p^{(r)}(\omega_m, q)\big)_{2,\gamma_1} = \sup\big\{E_{n-1}(f)_{2,\gamma_1} : f \in W_p^{(r)}(\omega_m, q)\big\}\\
&\le \frac{1}{2^{m/2}\alpha_{n,r}}\sqrt{\frac{2(n-r+1)+\alpha}{2(n+1)+\alpha}}\left(\int_0^h [1-\cos(n-r)t]^{mp}\, q(t)dt\right)^{-1/p}.
\end{aligned}$$

Using relations (3.1) between the n-widths, we obtain the upper estimate in (3.2):

$$\begin{aligned}
&\lambda_n\big(W_p^{(r)}(\omega_m, q)\big) \le E_{n-1}\big(W_p^{(r)}(\omega_m, q)\big)_{2,\gamma_1}\\
&\le \frac{1}{2^{m/2}\alpha_{n,r}}\sqrt{\frac{2(n-r+1)+\alpha}{2(n+1)+\alpha}}\left(\int_0^h [1-\cos(n-r)t]^{mp}\, q(t)dt\right)^{-1/p}.
\end{aligned} \tag{3.3}$$

To obtain the lower estimate on the right-hand side of (3.2) for all n-widths in the $(n+1)$-dimensional subspace of complex algebraic polynomials

$$\mathcal{P}_{n+1} = \Big\{p_n(z) : p_n(z) = \sum_{k=0}^{n} a_k z^k,\ a_k \in \mathbb{C}\Big\},$$

we introduce the ball

$$\mathbb{B}_{n+1}:=\left\{p_n(z)\in\mathcal{P}_n:\|p_n\|\le\frac{1}{2^{m/2}\alpha_{n,r}}\sqrt{\frac{2(n-r+1)+\alpha}{2(n+1)+\alpha}}\left(\int_0^h[1-\cos(n-r)t]^{mp}\,q(t)dt\right)^{-1/p}\right\},$$

where $n>r$, $n\in\mathbb{N}$, $r\in\mathbb{Z}_+$, and show that $\mathbb{B}_{n+1}\subset W_p^{(r)}(\omega_m,q)$. Indeed, for all $p_n(z)\in\mathbb{B}_{n+1}$, from (1.3), we write

$$\begin{aligned}\omega_m^2\left(p_n^{(r)},t\right)_{2,\gamma_1}&=2^m\sum_{k=r}^{\infty}\frac{\alpha_{k,r}^2|a_k(f)|^2}{2(k-r+1)+\alpha}(1-\cos(k-r)h)^m\\&\le 2^m\max_{r\le k\le n}\left\{\alpha_{k,r}^2(1-\cos(k-r)h)^m\right\}\sum_{k=r}^{\infty}\frac{|a_k(f)|^2}{2(k-r+1)+\alpha}.\end{aligned}\tag{3.4}$$

We have to prove that

$$\max_{r\le k\le n}\left\{\alpha_{k,r}^2(1-\cos(k-r)h)^m\right\}=\alpha_{n,r}^2(1-\cos(n-r)h)^m,\quad 0\le h\le\pi/(n-r).$$

Consider the function

$$\varphi(k)=\alpha_{k,r}^2(1-\cos(k-r)h)^m,\quad r\le k\le n,\quad 0\le h\le\pi/(n-r).$$

We will show that the function $\varphi(k)$ is monotone increasing for all accepted values k and h. To this end, it suffices to show that $\varphi'(k)>0$. In fact

$$\varphi'(k)=2\alpha_{k,r}^2\sum_{l=0}^{r-1}\frac{1}{k-l}(1-\cos(k-r)h)^m+mh\alpha_{k,r}^2\sin(k-r)h(1-\cos(k-r)h)^{m-1}\ge 0.$$

Hence, we can write (3.4) in the form

$$\begin{aligned}\omega_m^2\left(p^{(r)},t\right)_{2,\gamma_1}&\le 2^m\alpha_{n,r}^2(1-\cos(n-r)h)^m\sum_{k=r}^{\infty}\frac{|a_k(f)|^2}{2(k-r+1)+\alpha}\\\le 2^m\alpha_{n,r}^2(1-\cos(n-r)h)^m\sum_{k=0}^{\infty}\frac{|a_k(f)|^2}{2(k-r+1)+\alpha}&=2^m\alpha_{n,r}^2(1-\cos(n-r)h)^m\|p_n\|_{2,\gamma_1}^2.\end{aligned}\tag{3.5}$$

From (3.5), we have

$$\omega_m\left(p^{(r)},t\right)_{2,\gamma_1}\le 2^{m/2}\alpha_{n,r}(1-\cos(n-r)h)^{m/2}\|p_n\|_{2,\gamma_1}.$$

Raising both sides of this inequality to a power p $(0<p\le 2)$, multiplying them by the weight function $q(t)$, and integrating with respect to t from 0 to h, we obtain

$$\int_0^h\omega_m^p\left(p^{(r)},t\right)_{2,\gamma_1}q(t)dt\le 2^{mp/2}\alpha_{n,r}^p\|p_n\|_{2,\gamma_1}^p\int_0^h(1-\cos(n-r)h)^{mp/2}q(t)dt\le 1$$

for all $p_n\in\mathbb{B}_{n+1}$. It follows that $\mathbb{B}_{n+1}\subset W_p^{(r)}(\omega_m,q)$. Then, according to the definition of the Bernstein n-width and (3.1), we can write the following lower estimate for all above listed n-widths:

$$\begin{aligned}&\lambda_n(W_p^{(r)}(\omega_m,q),B_{2,\gamma_1})\ge b_n(W_p^{(r)}(\omega_m,q),B_{2,\gamma_1})\ge b_n(\mathbb{B}_{n+1},B_{2,\gamma_1})\\&\ge\frac{1}{2^{m/2}\alpha_{n,r}}\sqrt{\frac{2(n-r+1)+\alpha}{2(n+1)+\alpha}}\left(\int_0^h[1-\cos(n-r)t]^{mp}\,q(t)dt\right)^{-1/p}.\end{aligned}\tag{3.6}$$

Comparing the upper estimate (3.3) and the lower estimate in (3.6), we obtain the required equality (3.2). Theorem 3 is proved. □

4. Conclusion

Upper and lower estimates have been proven for extremal characteristics in a weighted Bergman space. In the case of a power function considered instead of a general weight, the values of n-widths have been calculated for a specific class of functions.

REFERENCES

1. Garkavi A. L. Simultaneous approximation to a periodic function and its derivatives by trigonometric polynomials. *Izv. Akad. Nauk SSSR Ser. Mat.*, 1960. Vol. 24, No. 1. P. 103–128. (in Russian)
2. Ligun A. A. Some inequalities between best approximations and moduli of continuity in an L_2 space. *Math Notes Acad. Sci. USSR*, 1978. Vol. 24, No. 6. P. 917–921. DOI: 10.1007/BF01140019
3. Pinkus A. *n-Widths in Approximation Theory.* Berlin, Heidelberg: Springer-Verlag, 1985. 294 p. DOI: 10.1007/978-3-642-69894-1
4. 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, No. 4. P. 656–668. DOI: 10.1134/S0001434618030343
5. 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)
6. Shabozov M. Sh., Saidusaynov M. S. Approximation of functions of a complex variable by Fourier sums in orthogonal systems in L_2. *Russian Math. (Iz. VUZ)*, 2020. Vol. 64, No. 6. P. 56–62. DOI: 10.3103/S1066369X20060080
7. 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)
8. Shabozov M. Sh., Shabozov O. Sh. On the best approximation of some classes of analytic functions in the weighted Bergman spaces. *Dokl. Math.*, 2007. Vol. 75. P. 97–100. DOI: 10.1134/S1064562407010279
9. Shabozov M. Sh., Yusupov G. A. Best polynomial approximations in L_2 of classes of 2π-periodic functions and exact values of their widths. *Math. Notes*, 2011. Vol. 90, No. 5. P. 748–757. DOI: 10.1134/S0001434611110125
10. Shabozov M. Sh., Yusupov G. A., Zargarov J. J. On the best simultaneous polynomial approximation of functions and their derivatives in Hardy spaces. *Trudy Inst. Mat. Mekh. UrO RAN*, 2021. Vol. 27, No. 4. P. 239–254. DOI: 10.21538/0134-4889-2021-27-4-239-254 (in Russian)
11. Tikhomirov V. M. *Nekotorie voprosi teorii priblizhenij* [Some Questions in Approximation Theory]. Moscow: Izdat. Moskov. Univ., 1976. 304 p. (in Russian)
12. Timan A. F. On the question of simultaneous approximation of functions and their derivatives on the whole real axis. *Izv. Akad. Nauk SSSR Ser. Mat.*, 1960. Vol. 24. No. 3. P. 421–430.
13. Vakarchuk S. B. Diameters of certain classes of functions analytic in the unit disc. I. *Ukr. Math. J.*, 1990. Vol. 42. P. 769–778. DOI: 10.1007/BF01062078
14. Vakarchyuk S. B. Best linear methods of approximation and widths of classes of analytic functions in a disk. *Math. Notes*, 1995. Vol. 57, No. 1–2. P. 21–27. DOI: 10.1007/BF02309390
15. Vakarchuk S. B., Shabozov M. Sh. The widths of classes of analytic functions in a disc. *Sb. Math.*, 2010. Vol. 201, No. 8. P. 1091–1110. DOI: 10.1070/SM2010v201n08ABEH004104

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 175–192

DOI: 10.15826/umj.2023.2.015

A PRESENTATION FOR A SUBMONOID OF THE SYMMETRIC INVERSE MONOID

Apatsara Sareeto

Institute of Mathematics, University of Potsdam,
Potsdam, 14476, Germany

channypooii@gmail.com

Jörg Koppitz

Institute of Mathematics and Informatics,
Bulgarian Academy of Sciences,
Sofia, 1113, Bulgaria

koppitz@math.bas.bg

Abstract: In the present paper, we study a submonoid of the symmetric inverse semigroup I_n. Specifically, we consider the monoid of all order-, fence-, and parity-preserving transformations of I_n. While the rank and a set of generators of minimal size for this monoid are already known, we will provide a presentation for this monoid.

Keywords: Symmetric inverse monoid, Order-preserving, Fence-preserving, Presentation.

1. Introduction

Let $\overline{n}$ be a finite chain with n elements, where n is a positive integer, denoted by $\overline{n} = \{1 < 2 < \cdots < n\}$. We denote by PT_n the monoid (under composition) of all partial transformations on $\overline{n}$. A partial transformation α on the set $\overline{n}$ is a mapping from a subset A of $\overline{n}$ into $\overline{n}$. The domain (respectively, image or range) of α is denoted by $dom(\alpha)$ (respectively, $im(\alpha)$). The empty transformation is denoted by ε. A transformation $\alpha \in PT_n$ is called order-preserving if $x < y$ implies $x\alpha \leq y\alpha$ for all $x, y \in dom(\alpha)$. It is worth noting that we write mappings on the right of their arguments and perform composition from left to right. Furthermore, an $\alpha \in PT_n$ is called a partial injection when α is injective. The set of all partial injections forms a monoid, the symmetric inverse semigroup I_n, as introduced by Wagner [17]. We denote by POI_n the submonoid of I_n, consisting of all order-preserving partial injections on $\overline{n}$. This monoid has already been well-studied (see e.g., [6]).

A non-linear order that is closed to a linear order in some sense is the so-called zig-zag order. The pair $(\overline{n}, \preceq)$ is called a zig-zag poset or fence if

$$1 \prec 2 \succ \cdots \prec n-1 \succ n \text{ if n is odd and } 1 \prec 2 \succ \cdots \succ n-1 \prec n \text{ if n is even, respectively.}$$

The definition of the partial order $\preceq$ is self-explanatory. A transformation $\alpha \in PT_n$ is referred to as fence-preserving if it preserves the partial order $\preceq$, meaning that for all $x, y \in dom(\alpha)$ with $x \prec y$, we have $x\alpha \preceq y\alpha$. The set of fence-preserving transformations on $\overline{n}$ was initially explored by Currie, Visentin, and Rutkowski. In [2, 14], the authors investigated the number of order-preserving maps of a finite fence. In particular, a formula for the number of order-preserving self-mappings of a fence was introduced. It is noteworthy that every element of a fence is either minimal or maximal. For all $x, y \in \overline{n}$ with $x \prec y$, we have $y \in \{x-1, x+1\}$. We denote by PFI_n the submonoid of I_n, consisting of all fence-preserving partial injections of $\overline{n}$. We denote by IF_n the inverse submonoid of PFI_n of all regular elements in PFI_n. It is easy to see that IF_n is the set of all $\alpha \in PFI_n$ with $\alpha^{-1} \in PFI_n$. It is worth mentioning that several properties of a variety of monoids of fence-preserving transformations were studied [3, 7, 9, 11, 12, 16].

In the present paper, we focus on a submonoid of $IOF_n = IF_n \bigcap POI_n$. Let $a \in dom(\alpha)$ for some $\alpha \in IOF_n$. If $a+1 \in dom(\alpha)$ or $a-1 \in dom(\alpha)$ then it is easy to verify that a and $a\alpha$ have the same parity. In other words, a is odd if and only if $a\alpha$ is odd. However, if $a-1$ and $a+1$ are not in $dom(\alpha)$, then a and $a\alpha$ can have different parity. In order to exclude this case, we require that the image of any $a \in dom(\alpha)$ has the same parity as $a\alpha$. In this context, we refer to α as parity-preserving. In our paper, we consider the monoid IOF_n^{par} of all parity-preserving transformations of IOF_n. Notably, for any $\alpha \in IOF_n^{par}$, the inverse partial injection α^{-1} exists and possesses order-preserving, fence-preserving, and parity-preserving. This observation implies that IOF_n^{par} is an inverse submonoid of I_n, as explained in [15]. Furthermore, in the same paper [15], the authors provided a characterization of the monoid IOF_n^{par}:

Proposition 1 [15]. *Let $p \le n$ and let*

$$\alpha = \begin{pmatrix} d_1 & < & d_2 & < & \cdots & < & d_p \\ m_1 & & m_2 & & \cdots & & m_p \end{pmatrix} \in I_n.$$

Then $\alpha \in IOF_n^{par}$ if and only if the following four conditions hold:

(i) $m_1 < m_2 < \ldots < m_p$;
(ii) d_1 *and* m_1 *have the same parity*;
(iii) $d_{i+1} - d_i = 1$ *if and only if* $m_{i+1} - m_i = 1$ *for all* $i \in \{1, \ldots, p-1\}$;
(iv) $d_{i+1} - d_i$ *is even if and only if* $m_{i+1} - m_i$ *is even for all* $i \in \{1, \ldots, p-1\}$.

Also in [15], a set of generators of IOF_n^{par} of minimal size is given. This leads to the question of a presentation of IOF_n^{par}. In this paper, we will exhibit a monoid presentation for IOF_n^{par}. A monoid presentation is represented as an ordered pair $\langle X \mid R \rangle$, where X is a set, referred to as the alphabet (a set whose elements are called letters), and R is a binary relation on the free monoid generated by X, denoted by X^*. A pair $(u, v) \in X^* \times X^*$ is represented by $u \approx v$ and is called relation. We state that $u \approx v$, for $u, v \in X^*$, is a consequence of R if $(u, v) \in \rho_R$, where ρ_R denotes the congruence on X^* generated by R. We say that the momoid IOF_n^{par} has (monoid) presentation $\langle X \mid R \rangle$ if IOF_n^{par} is isomorphic to the factor semigroup X^*/ρ_R. For a more comprehensive understanding of semigroups, presentations, and standard notation see [1, 8, 10, 13].

Given that IOF_n^{par} is a finite monoid, we can always exhibit a presentation for it. A usual method to establish a good presentations is the Guess and Prove Method, which is described by the following theorem, adapted to monoids from Ruškuc (1995, Proposition 3.2.2).

Theorem 1 [13]. *Let X be a generating set for IOF_n^{par}, let $R \subseteq X^* \times X^*$ be a set of relations and let $W \subseteq X^*$ that the following conditions are satisfied*:

1. *The generating set X of IOF_n^{par} satisfies all the relations from R*;
2. *For each word $w \in X^*$, there exists a word $w' \in W$ such that the relation $w \approx w'$ is a consequence of R*;
3. $|W| \le |IOF_n^{par}|$.

Then IOF_n^{par} is defined by the presentation $\langle X \mid R \rangle$.

In the next section, we introduce the alphabet (generating set) denoted as X_n and the binary relation R on X_n^*. Furthermore, we will demonstrate that X_n fulfills all the relations in R as outlined in Theorem 1, item 1. Following the guidance of item 2 in Theorem 1, we will establish a set of forms, denoted as P, in Section 3. Finally, in the last section, we will provide a proof for item 3 of Theorem 1.

2. The generator and relations

In this section, we will define the alphabet X_n and introduce a binary relation R on X_n^*. We will also demonstrate that the corresponding generating set satisfies all the relations in R. Let $\overline{v}_i$ be the partial identity with the domain $\overline{n}\backslash\{i\}$ for all $i \in \{1, ..., n\}$. Additionally, let us define

$$\overline{u}_i = \begin{pmatrix} 1 & \cdots & i & i+1 & i+2 & i+3 & i+4 & \cdots & n \\ 3 & \cdots & i+2 & - & - & - & i+4 & \cdots & n \end{pmatrix}$$

and $\overline{x}_i = (\overline{u}_i)^{-1}$ for all $i \in \{1, ..., n-2\}$. By Proposition 1, it is easy to verify that $\overline{u}_i$ as well as $\overline{x}_i$, $i \in \{1, ..., n-2\}$, belong to IOF_n^{par}. In [15], the authors have shown that $\{\overline{v}_1, \overline{v}_2, ..., \overline{v}_n, \overline{u}_1, \overline{u}_2, ..., \overline{u}_{n-2}, \overline{x}_1, \overline{x}_2, ..., \overline{x}_{n-2}\}$ is a generating set of IOF_n^{par}. In order to use Theorem 1, we define an alphabet

$$X_n = \{v_1, v_2, ..., v_n, u_1, u_2, ..., u_{n-2}, x_1, x_2, ..., x_{n-2}\},$$

which corresponds to the set of generators of IOF_n^{par}. For $w = w_1...w_m$ with $w_1, ..., w_m \in X_n$, where m being a positive integer, we write w^{-1} for the word $w^{-1} = w_m...w_1$.

We fix a particular sequence of letters as follows: $x_{i,j} = x_i x_{i+2} ... x_{i+2j-2}$ and $u_{i,j} = u_i u_{i+2} ... u_{i+2j-2}$ for $i \in \{1, ..., n-2\}$, $j \in \{1, ..., \lfloor (n-i)/2 \rfloor\}$ and obtain the following sets of words:

$$W_x = \left\{x_{i,j} : i \in \{1, ..., n-2\},\ j \in \left\{1, ..., \left\lfloor \frac{n-i}{2} \right\rfloor\right\}\right\},$$
$$W_x^{-1} = \{x_{i,j}^{-1} : x_{i,j} \in W_x\},$$
$$W_u = \left\{u_{i,j} : i \in \{1, ..., n-2\},\ j \in \left\{1, ..., \left\lfloor \frac{n-i}{2} \right\rfloor\right\}\right\}.$$

Let w be any word of the form $w = w_1...w_m$ with $w_1, ..., w_m \in W_x \cup W_u$ and m is a positive integer. For $k \in \{1, ..., m\}$, the word w_k is of the form

$$w_k = \begin{cases} u_{i_k, j_k} & \text{if} \quad w_k \in W_u; \\ x_{i_k, j_k} & \text{if} \quad w_k \in W_x \end{cases}$$

for some $i_k \in \{1, ..., n-2\}$, $j_k \in \{1, ..., \lfloor (n-i)/2 \rfloor\}$. We observe $j_k = |w_k|$, i.e. j_k is the length of the word w_k. We define two sequences $1_x, 2_x, ..., m_x$ and $1_u, 2_u, ..., m_u$ of indicators: for $k \in \{1, ..., m\}$ let

$$k_x = \begin{cases} i_k + 2|w_k| + 2|W_u^k| - 2|W_x^k| & \text{if} \quad w_k \in W_u; \\ i_k & \text{if} \quad w_k \in W_x \end{cases}$$

and

$$k_u = \begin{cases} i_k + 2|w_k| - 2|W_u^k| + 2|W_x^k| & \text{if} \quad w_k \in W_x; \\ i_k & \text{if} \quad w_k \in W_u, \end{cases}$$

where W_u^s (respectively, W_x^s) means the word $w_{s+1}...w_m$ without the letters in $\{x_1,...,x_{n-2}\}$ respectively, in $\{u_1,...,u_{n-2}\}$) for $s \in \{0,1,...,m-1\}$ and $W_u^m = W_x^m = \epsilon$, where ϵ is the empty word. Let Q_0 be the set of all words $w = w_1...w_m$ with $w_1,...,w_m \in W_x \cup W_u$ and m being a positive integer such that:

(1_q) If $w_k, w_l \in W_x$ then $i_k + 2j_k + 1 < i_l$ for $k < l \leq m$;
(2_q) If $w_k, w_l \in W_u$ then $i_k + 2j_k + 1 < i_l$ for $k < l \leq m$;
(3_q) If $w_k \in W_u$ then $i_k + 2j_k + 2 \leq (k+1)_u$ for $k \in \{1,...,m-1\}$ and $(k+1)_x - k_x \geq 2$;
(4_q) If $w_k \in W_x$ then $i_k + 2j_k + 2 \leq (k+1)_x$ for $k \in \{1,...,m-1\}$ and $(k+1)_u - k_u \geq 2$.

Let now $w = w_1...w_m \in Q_0$ and let $w^* = W_u^0(W_x^0)^{-1}$. Further, we define recursively a set A_w:

(5_q) If $m_u > m_x$ and $m_u + 2 \leq n$ then $A_m = \{m_u+2,...,n\}$,
if $m_u < m_x$ and $m_x + 2 \leq n$ then $A_m = \{m_x+2,...,n\}$,
otherwise $A_m = \emptyset$;
(6_q) If $w_k \in W_u$ then $A_k = A_{k+1} \cup \{i_k + 2j_k + 2, ..., (k+1)_u - 1\}$ for $k \in \{1,...,m-1\}$,
if $w_k \in W_x$ then $A_k = A_{k+1} \cup \{k_u + 2, ..., (k+1)_u - 1)\}$ for $k \in \{1,...,m-1\}$;
(7_q) If $1 \in \{1_x, 1_u\}$ then $A_w = A_1$,
if $1 < 1_u \leq 1_x$ then $A_w = A_1 \cup \{1,...,1_u - 1\}$,
if $1 < 1_x < 1_u$ then $A_w = A_1 \cup \{1_u - 1_x + 1, ..., 1_u - 1\}$.

For a set $A = \{i_1 < i_2 < \cdots < i_k\} \subseteq \overline{n}$, let $v_A = v_{i_1}v_{i_2}...v_{i_k}$ for some $k \in \{1,...,n\}$. Note that $v_\emptyset$ means the empty word ϵ. For convenience, we put $v_i = \epsilon$ for $i \geq n+1$. Let

$$W_n = \{v_A w^* : w \in Q_0, A \subseteq A_w\} \cup \{v_A : A \subseteq \overline{n}\}.$$

On the other hand, we will define now a set of relations. For this, let W_t be the set of all words of the form $u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}}$ with the following four properties:

(i) $l \in \{0,...,n-2\}$, and $m \in \{0,...,n-3\}$;
(ii) $i_0 < i_1 < \cdots < i_l \in \{1,...,n-2\}$;
(iii) $j_1 > j_2 > \cdots > j_m > j_{m+1} \in \{1,...,n-2\}$;
(iv) if $k \in \{i_0,...,i_{l-1}\}$ (respectively, $k \in \{j_2,...,j_{m+1}\}$) then $k+1, k+3 \notin \{i_1,...,i_l\}$ (respectively, $k+1, k+3 \notin \{j_1,...,j_m\}$) for all $k \in \{1,...,n-3\}$.

Then we define a sequence R of relations on X_n^* as follows: for $i,j \in \{1,...,n\}$ and $k = i+2j-2$, let

$$(E)\ x_iu_j \approx \begin{cases} v_1v_2v_{i+3}...v_{j+3}, & \text{if} \quad i<j, j-i=2,3; \\ v_1v_2v_{j+3}...v_{i+3}, & \text{if} \quad i>j, i-j=2,3; \\ v_1v_2v_{j+3}v_{j+4}, & \text{if} \quad i>j, i-j=1; \\ v_1v_2v_{j+2}v_{j+3}, & \text{if} \quad i<j, j-i=1; \\ v_1v_2v_{i+3}, & \text{if} \quad i=j; \\ v_1v_2u_jx_{i+2}, & \text{if} \quad i<j, j-i\geq 4; \\ v_1v_2u_{j+2}x_i, & \text{if} \quad i>j, i-j\geq 4; \end{cases}$$

($L1$) $u_2u_1 \approx u_1u_2 \approx x_1x_2 \approx x_2x_1 \approx u_2^2 \approx x_2^2 \approx v_1v_2v_3v_4v_5$;
($L2$) $u_3u_2 \approx x_2x_3 \approx v_1v_2v_3v_4v_5v_6$;
($L3$) $u_iu_1 \approx v_1v_2u_i$ and $x_1x_i \approx v_3v_4x_i, i \geq 3$;
($L4$) $u_iu_2 \approx v_1v_2v_3u_i$ and $x_2x_i \approx v_3v_4v_5x_i, i \geq 4$;
($L5$) $u_iu_{i-1} \approx v_{i+3}u_{i-3}u_{i-1}$ and $x_{i-1}x_i \approx v_{i+3}x_{i-1}x_{i-3}, i \geq 4$;
($L6$) $u_iu_j \approx u_{j-2}u_i$ and $x_jx_i \approx x_ix_{j-2}, i > j \geq 3, i-j \geq 2$;

$(R1)$ $v_i^2 \approx v_i$, $i \in \{1, ..., n\}$;
$(R2)$ $v_i v_j \approx v_j v_i$, $i, j \in \{1, ..., n\}, i \neq j$;
$(R3)$ $v_i u_j \approx u_j v_i$ and $v_i x_j \approx x_j v_i$, $i \in \{j+4, ..., n\}$;
$(R4)$ $v_i u_j \approx u_j v_{i+2}$ and $v_{i+2} x_j \approx x_j v_i$, $1 \leq i \leq j$;
$(R5)$ $v_i u_j \approx u_j$ and $x_j v_i \approx x_j$, $i \in \{j+1, j+2, j+3\}$;
$(R6)$ $u_j v_i \approx u_j$ and $v_i x_j \approx x_j$, $i \in \{1, 2, j+3\}$;
$(R7)$ $u_1^2 \approx x_1^2 \approx v_1 ... v_4$;
$(R8)$ $u_i^2 \approx u_{i-2} u_i$ and $x_i^2 \approx x_i x_{i-2}$, $i \geq 3$;
$(R9)$ $u_i u_{i+1} \approx u_{i-1} u_{i+1}$ and $x_{i+1} x_i \approx x_{i+1} x_{i-1}$, $i \in \{2, ..., n-5\}$;
$(R10)$ $u_i u_{i+3} \approx v_{i+6} u_i u_{i+2}$ and $x_{i+3} x_i \approx v_{i+6} x_{i+2} x_i$, $i \leq n-5$;
$(R11)$ $w \approx v_{i_0+1} v_{i_0+2} v_{i_0+3} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m}$, $w = u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}} \in W_t$ with $j_{m+1} = i_0 + 2l - 2m$;
$(R12)$ $w \approx v_{i_0} v_{i_0+1} v_{i_0+2} v_{i_0+3} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m}$, $w = u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}} \in W_t$ with $j_{m+1} = i_0 + 2l - 2m - 1$;
$(R13)$ $w \approx v_{i_0+1} v_{i_0+2} v_{i_0+3} v_{i_0+4} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m}$, $w = u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}} \in W_t$ with $j_{m+1} = i_0 + 2l - 2m + 1$;
$(R14)$ $w \approx u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m}$, $w = u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}} \in W_t$ with $j_{m+1} < 2l - 2m$;
$(R15)$ $w \approx u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}}$, $w = u_{i_0} u_{i_1} ... u_{i_l} x_{j_1} ... x_{j_m} x_{j_{m+1}} \in W_t$ with $i_0 < 2m - 2l$;
$(R16)$ $v_1 ... v_i u_{i,j} \approx v_1 ... v_{k+3}$, $i \in \{1, ..., n-2\}$;
$(R17)$ $v_{k-i+3} ... v_{k+2} x_{i,j}^{-1} \approx v_1 ... v_{k+3}$, $i \in \{1, ..., n-2\}$;
$(R18)$ $v_i u_{i,j} \approx v_{k+3} u_{i-1,j}$, $i \in \{2, ..., n-2\}$;
$(R19)$ $v_{k+2} x_{i,j}^{-1} \approx v_{k+3} x_{i-1,j}^{-1}$, $i \in \{2, ..., n-2\}$.

Lemma 1. *The relations from R hold as equations in IOF_n^{par}, when the letters are replaced by the corresponding transformations.*

P r o o f. We show the statement diagrammatically. This method was also used in [4, 5]. We give an example calculation for the relation $(R10)$ $u_i u_{i+3} \approx v_{i+6} u_i u_{i+2}$, $i \leq n-5$, in Figures 1 and 2 below. Note we can show $x_{i+3} x_i \approx v_{i+6} x_{i+2} x_i$ in a similar way. $\square$

By Figures 1 and 2, we have that $\overline{u}_i \overline{u}_{i+3} = \overline{v}_{i+6} \overline{u}_i \overline{u}_{i+2}$.

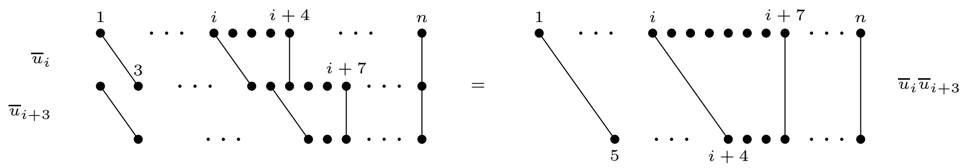


Figure 1. $\overline{u}_i \overline{u}_{i+3}$.

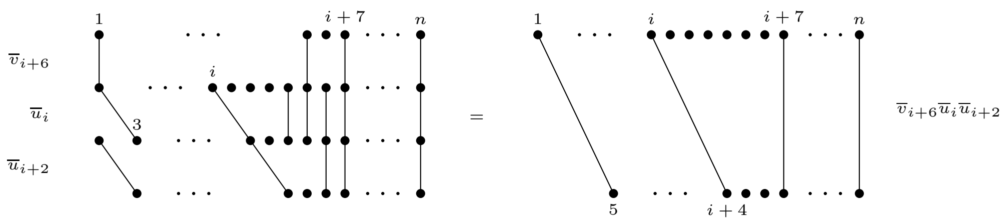


Figure 2. $\overline{v}_{i+6} \overline{u}_i \overline{u}_{i+2}$.

Next, we will verify consequences of R, which are important by technical reasons.

Lemma 2. *(i) For* $w = u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}} \in W_t$ *with* $j_{m+1} = 2l - 2m$, *we have* $w \approx v_1u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}$.
(ii) For $w = u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}} \in W_t$ *with* $i_0 = 2m - 2l$, *we have* $w \approx v_{i_0+3}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}}$.

P r o o f. (i) We have

$$u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}} \overset{(R14)}{\approx} u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}-1}x_{j_{m+1}}.$$

Suppose $j_{m+1} = 2l - 2m \geq 4$. Then

$$\begin{aligned} u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}-1}x_{j_{m+1}} &\overset{(L5)}{\approx} u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}v_{j_{m+1}+3}x_{j_{m+1}-1}x_{j_{m+1}-3} \\ &\overset{(R4)}{\approx} v_1u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}-1}x_{j_{m+1}-3} \overset{(R14)}{\approx} v_1u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}. \end{aligned}$$

Suppose $j_{m+1} = 2l - 2m < 4$, i.e. $j_{m+1} = 2$. We prove that

$$u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}x_{j_{m+1}} \approx v_1u_{i_0}u_{i_1}...u_{i_l}x_{j_1}...x_{j_m}$$

by using $(L1)$ and $(R4)$–$(R6)$ in a similar way.

(ii) The proof is similar to (i), by using $(R15)$ and $(L5)$ if $i_0 \geq 4$ and $(R15)$, $(L1)$, and $(R4)$–$(R6)$ if $i_0 = 2$. □

3. Set of forms

In this section, we introduce an algorithm, which transforms any word $w \in X_n^*$ to a word in W_n using R, with other words, we show that for all $w \in X_n^*$, there is $w' \in W_n$ such that $w \approx w'$ is a consequence of R. First, the algorithm transforms each $w \in X_n^*$ to a "new" word w'. All these "new" words will be collected in a set. Later, we show that this set belongs to W_n. Let $w \in X_n^* \backslash \{\epsilon\}$.

- Using $(R1)$–$(R6)$, we can move any v_i for $i \in \{1, 2, ..., n\}$, at the beginning of the word or we can cancel it. So we obtain $w \approx \tilde{v}\tilde{w}$, where $\tilde{v} \in \{v_1, ..., v_n\}^*$ and $\tilde{w} \in \{u_1, u_2, ..., u_{n-2}, x_1, x_2, ..., x_{n-2}\}^*$.

- Moreover, we separate the u_i*'s* and x_i*'s* for $i \in \{1, ..., n-2\}$ by (E) and $(R1)$–$(R6)$. Then $\tilde{v}\tilde{w} \approx \overline{v}\overline{BC}$, where $\overline{v} \in \{v_1, ..., v_n\}^*$, $\overline{B} \in \{u_1, u_2, ..., u_{n-2}\}^*$, and $\overline{C} \in \{x_1, x_2, ..., x_{n-2}\}^*$.

- By $(L1)$–$(L6)$ and $(R1)$–$(R6)$, we get $\overline{v}\overline{BC} \approx v'B'C'$, where $v' \in \{v_1, ..., v_n\}^*$, $B' \in \{u_1, u_2, ..., u_{n-2}\}^*$, and $C' \in \{x_1, x_2, ..., x_{n-2}\}^*$ such that the indices of the letters in the word B' are ascending and in the word C' are descending (reading from the left to the right).

- By $(L1)$, $(R7)$–$(R10)$, and $(R1)$–$(R6)$, we replace subwords of $B'C'$ of the form $x_{i+3}x_i, x_{i+1}x_i, x_i^2, u_i^2, u_iu_{i+3}$, and u_iu_{i+1} until $v'B'C' \approx v''w_1...w_p$ with $v'' \in \{v_1, ..., v_n\}^*$ and $w_1, ..., w_p \in W_x^{-1} \cup W_u$ such that

  if $u_i \in var(w_1...w_p)$ (respectively, $x_i \in var(w_1...w_p)$) then $u_{i+1}, u_{i+3} \notin var(w_1...w_p)$ (respectively, $x_{i+1}, x_{i+3} \notin var(w_1...w_p)$) for all $i \in \{1, ..., n-2\}$ and each letter in $w_1...w_p$ is unique. (∗)

  Note that this is possible since each of the relations $(L1)$, $(R7)$–$(R10)$, and $(R1)$–$(R6)$ does not increase the index of any letter in $\{u_1, u_2, ..., u_{n-2}, x_1, x_2, ..., x_{n-2}\}$ in the "new" word.

- Using $(R11)$–$(R15)$, Lemmas 2, and $(R1)$–$(R6)$, we remove letters x_i and u_i, respectively, until one can not more remove a letter x_i or u_i for $i \in \{1, 2, ..., n-2\}$. We obtain $v''w_1...w_p \approx v'''w'_1...w'_{p'}$, where $v''' \in \{v_1, ..., v_n\}^*$ and $w'_1, ..., w'_{p'} \in W_x^{-1} \cup W_u$. Note that is possible since each of the relations $(R11)$–$(R15)$ as well as Lemmas 2 only removes letters (and add letters in $\{v_1, ..., v_n\}$, respectively).

- We decrease the indices of the letters in $\{u_1, u_2, ..., u_{n-2}, x_1, x_2, ..., x_{n-2}\}$ (if possible) by $(R16)$–$(R19)$ as well as $(R1)$–$(R6)$ and obtain $v'''w'_1...w'_{p'} \approx v^*B^*C^*$ with $v^* \in \{v_1, ..., v_n\}^*$, $B^* \in \{u_1, u_2, ..., u_{n-2}\}^*$, and $C^* \in \{x_1, x_2, ..., x_{n-2}\}^*$. Note that the indices of the letters in B^* (respectively, in C^*) are ascending (respectively, are descending).

We repeat all steps. The procedure terminates if the word will not change more in all steps. We obtain $v^*B^*C^* \approx v_A\hat{w}_1...\hat{w}_{\hat{p}}$, where $\hat{w}_1, ..., \hat{w}_{\hat{p}} \in W_x^{-1} \cup W_u$ and $A \subseteq \overline{n}$ such that no v_j $(j \in A)$ can be canceled by using $(R1)$–$(R6)$. This case has to happen since the number of the letters from $\{u_1, u_2, ..., u_{n-2}, x_1, x_2, ..., x_{n-2}, v_1, ..., v_n\}$ decreases or is kept and the indices of the u_i's and x_i's decrease or are kept in each step.

We denote by P the set of all words obtained from $w \in X_n^*$ by that algorithm.
By $(*)$, we obtain immediately from the algorithm.

Remark 1. Let $\hat{w} = v_A\hat{w}_1...\hat{w}_m \in P$ and let $1 \leq k < k' \leq m$.
If $\hat{w}_k, \hat{w}_{k'} \in W_u$ then $i_k + 2|\hat{w}_k| + 2 \leq i_{k'}$.
If $\hat{w}_k, \hat{w}_{k'} \in W_x$ then $i_{k'} + 2|\hat{w}_{k'}| + 2 \leq i_k$.

Let fix a word $\hat{w} = v_A\hat{w}_1...\hat{w}_m \in P$. There are $a, b \in \{0, ..., n\}$ with $a + b = m$, $t_1, ..., t_{a+b} \in \{1, ..., m\}$, $w_{t_1}, ..., w_{t_a} \in W_u$ and $w_{t_{a+1}}, ..., w_{t_{a+b}} \in W_x$ such that

$$\hat{w} = v_A\hat{w}_1...\hat{w}_m = v_A w_{t_1}...w_{t_a} w_{t_{a+1}}^{-1}...w_{t_{a+b}}^{-1},$$

where $\{w_{t_1}, ..., w_{t_a}\} = \emptyset$ or $\{w_{t_{a+1}}, ..., w_{t_{a+b}}\} = \emptyset$ (i.e. $a = 0$ or $b = 0$) is possible. We observe that $\{\hat{w}_1, ..., \hat{w}_m\} = \{w_{t_1}, ..., w_{t_a}, w_{t_{a+1}}^{-1}, ..., w_{t_{a+b}}^{-1}\}$ and $\{t_1, ..., t_a, t_{a+1}, ..., t_{a+b}\} = \{1, ..., m\}$. We define an order on $\{t_1, ..., t_a, t_{a+1}, ..., t_{a+b}\}$ by $t_1 < \cdots < t_a$ and $t_{a+b} < \cdots < t_{a+1}$. If $a, b \geq 1$, the order between $t_1, ..., t_a$ and $t_{a+1}, ..., t_{a+b}$ is given by the following rule:

Let $k \in \{1, ..., a\}$ and $l \in \{1, ..., b\}$
if $i_{t_k} + 2|w_{t_k}| - 2 + 2|w_{t_{k+1}}...w_{t_a}| - 2|w_{t_{a+1}}^{-1}...w_{t_{a+l-1}}^{-1}| < i_{t_{a+l}} + 2|w_{t_{a+l}}^{-1}| - 2$ then $t_k < t_{a+l}$ and

if $i_{t_k} + 2|w_{t_k}| - 2 + 2|w_{t_{k+1}}...w_{t_a}| - 2|w_{t_{a+1}}^{-1}...w_{t_{a+l-1}}^{-1}| > i_{t_{a+l}} + 2|w_{t_{a+l}}^{-1}| - 2$ then $t_k > t_{a+l}$.
The case

$$i_{t_k} + 2|w_{t_k}| - 2 + 2|w_{t_{k+1}}...w_{t_a}| - 2|w_{t_{a+1}}^{-1}...w_{t_{a+l-1}}^{-1}| = i_{t_{a+l}} + 2|w_{t_{a+l}}^{-1}| - 2$$

is not possible, since otherwise we can cancel $u_{i_{t_k}+2|w_{t_k}|-2}$ and $x_{i_{t_{a+l}}+2|w_{t_{a+l}}^{-1}|-2}$ in $\hat{w}$ by $(R11)$. Our next aim is to describe the relationships between $k_u, (k+1)_u$ and $k_x, (k+1)_x$ for all $k \in \{1, ..., m-1\}$ for the word $w = w_1...w_m$.

Lemma 3. *For all $k \in \{1, ..., m-1\}$, we have $k_u < (k+1)_u$ and $k_x < (k+1)_x$.*

P r o o f. Let $k \in \{1, ..., m-1\}$. Suppose $w_k, w_{k+1} \in W_u$. We obtain $k_u < (k+1)_u$ and

$$k_x = i_k + 2|w_k| + 2|W_u^k| - 2|W_x^k|,$$
$$(k+1)_x = i_{k+1} + 2|w_{k+1}| + 2|W_u^{k+1}| - 2|W_x^{k+1}|.$$

By Remark 1, we have $i_k + 2|w_k| + 2 \leq i_{k+1}$. This gives

$$i_k + 2|w_k| + 2|W_u^k| - 2|W_x^k| < i_{k+1} + 2|W_u^k| - 2|W_x^k| = i_{k+1} + 2|w_{k+1}| + 2|W_u^{k+1}| - 2|W_x^{k+1}|$$

(since $w_{k+1} \in W_u$ implies $2|W_x^k| = 2|W_x^{k+1}|$). Then $k_x < (k+1)_x$. For the case $w_k, w_{k+1} \in W_x$, we can show that $k_u < (k+1)_u$ and $k_x < (k+1)_x$ in a similar way.

Suppose $w_k \in W_u$ and $w_{k+1} \in W_x$. First, we will show $k_u < (k+1)_u$. We have $k_u = i_k$ and

$$(k+1)_u = i_{k+1} + 2|w_{k+1}| + 2|W_x^{k+1}| - 2|W_u^{k+1}|.$$

Since $k \in \{t_1, ..., t_a\}$ and $k+1 \in \{t_{a+1}, ..., t_{a+b}\}$, we obtain

$$i_k + 2|w_k| - 2 + 2|W_u^k| - 2|W_x^{k+1}| < i_{k+1} + 2|w_{k+1}| - 2.$$

Then

$$i_k < i_k + 2|w_k| < i_{k+1} + 2|w_{k+1}| + 2|W_x^{k+1}| - 2|W_u^{k+1}|$$

(since $w_{k+1} \in W_x$ implies $|W_u^k| = |W_u^{k+1}|$). Then $k_u < (k+1)_u$. Moreover, we prove $k_x < (k+1)_x$ similarly. The case $w_k \in W_x$ and $w_{k+1} \in W_u$ can be shown in a similar way as above. □

Of course, the next goal should be the proof of $w = w_1...w_m \in Q_0$, i.e. we will show that w satisfies (1_q)–(4_q).

Lemma 4. *We have $w = w_1...w_m \in Q_0$.*

P r o o f. Exactly, w satisfies (1_q) and (2_q). This is trivially checked by Remark 1.

Let $k \in \{1, ..., m-1\}$ and let $w_k \in W_u, w_{k+1} \in W_x$. This provides $k \in \{t_1, ..., t_a\}$, $k+1 \in \{t_{a+1}, ..., t_{a+b}\}$. We have

$$i_k + 2|w_k| - 2 + 2|W_u^k| - 2|W_x^{k+1}| < i_{k+1} + 2|w_{k+1}| - 2.$$

Since $w_{k+1} \in W_x$, we have

$$2|W_u^k| = 2|W_u^{k+1}|.$$

So

$$i_k + 2|w_k| - 2 + 2|W_u^{k+1}| - 2|W_x^{k+1}| < i_{k+1} + 2|w_{k+1}| - 2.$$

We observe that

$$i_k + 2|w_k| - 2 + 2|W_u^{k+1}| - 2|W_x^{k+1}| + 1 \leq i_{k+1} + 2|w_{k+1}| - 2.$$

If

$$i_k + 2|w_k| - 2 + 2|W_u^{k+1}| - 2|W_x^{k+1}| + 1 = i_{k+1} + 2|w_{k+1}| - 2,$$

we can cancel $u_{i_k+2|w_k|-2}, x_{i_{k+1}+2|w_{k+1}|-2}$ by $(R13)$ in $\hat{w}$. This contradicts $\hat{w} \in P$. Then

$$i_k + 2|w_k| - 2 + 2|W_u^{k+1}| - 2|W_x^{k+1}| + 2 \leq i_{k+1} + 2|w_{k+1}| - 2,$$

i.e.

$$i_k + 2|w_k| + 2 \leq i_{k+1} + 2|w_{k+1}| - 2|W_u^{k+1}| + 2|W_x^{k+1}| = (k+1)_u.$$

Next, to show that $(k+1)_x - k_x \geq 2$. Lemma 3 gives $(k+1)_x - k_x \geq 1$.

If $(k+1)_x - k_x = 1$ then

$$i_{k+1} - i_k - 2|w_k| - 2|W_u^k| + 2|W_x^k| = 1.$$

This implies

$$i_{k+1} + 2|w_{k+1}| - 2 = i_k + 2|w_k| - 2 + 2|W_u^k| - 2|W_x^{k+1}| + 1$$

since

$$2|W_x^k| = 2|W_{k+1}| + 2|W_x^{k+1}|.$$

We can cancel $u_{i_k+2|w_k|-2}, x_{i_{k+1}+2|w_{k+1}|-2}$ in $\hat{w}$ by $(R13)$. This contradicts $\hat{w} \in P$. Thus, $(k+1)_x - k_x \geq 2$. In case $w_k, w_{k+1} \in W_u$, by using Remark 1, we easily get

$$i_k + 2|w_k| + 2 \leq (k+1)_u.$$

To show $(k+1)_x - k_x \geq 2$, it is routine to calculate directly. Together with Remark 1, we will get that $(k+1)_x - k_x \geq 2$. Altogether, w satisfies (3_q). We prove that w satisfies (4_q) in a similar way. Therefore, $w \in Q_0$. □

We have shown $w \in Q_0$. This leads us to the next step, showing that $A \subseteq A_w$. First, we point out subsets of $\overline{n}$, which do not contain any element of A.

Lemma 5. *Let $q \in \{1, ..., a\}$ and let*

$$\rho \in \{i_{t_q} + 1, ..., i_{t_q} + 2|w_{t_q}| + 1\} \cap \overline{n}.$$

Then $\rho \notin A$.

P r o o f. Assume $\rho \in A$. Then

$$v_\rho w_{t_1} ... w_{t_q} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \overset{(R3)}{\approx} w_{t_1} ... v_\rho w_{t_q} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1}.$$

If $\rho \in \{i_{t_q} + 1, i_{t_q} + 2, i_{t_q} + 3\} \cap \overline{n}$ then

$$v_\rho u_{i_{t_q}} \overset{(R5)}{\approx} u_{i_{t_q}}.$$

If $\rho = i_{t_q} + h + t$ for some $h \in \{2, 4, ..., 2|w_{t_q}| - 2\}$ and $t \in \{2, 3\}$ then

$$\begin{aligned} w_{t_1} ... v_\rho w_{t_q} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} &= w_{t_1} ... v_\rho u_{i_{t_q}} u_{i_{t_q}+2} ... u_{i_{t_q}+2|w_{t_q}|-2} w_{t_{q+1}} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R3)}{\approx} w_{t_1} ... u_{i_{t_q}} ... v_{(i_{t_q}+h+t)} u_{i_{t_q}+h} ... u_{i_{t_q}+2|w_{t_q}|-2} w_{t_{q+1}} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R5)}{\approx} w_{t_1} ... u_{i_{t_q}} ... u_{i_{t_q}+h} ... u_{i_{t_q}+2|w_{t_q}|-2} w_{t_{q+1}} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1}, \end{aligned}$$

i.e. we can cancel v_ρ in $\hat{w}$ using $(R3)$ and $(R5)$, a contradiction. □

Lemma 6. *Let $\rho \in A$ and let $q \in \{1, ..., a\}$ such that $t_q \neq m$. If $\rho \in \{(t_q)_u + 1, ..., (t_q + 1)_u - 1\}$ then*

$$\rho \in \left\{(t_q)_u + 2|w_{t_q}| + 2, ..., (t_q + 1)_u - 1\right\} \subseteq A_w.$$

P r o o f. We have $(t_q)_u = i_{t_q}$. It is a consequence of Lemma 5 that

$$\rho \in \left\{i_{t_q} + 2|w_{t_q}| + 2, ..., (t_q + 1)_u - 1\right\}$$

and by (6_q), we have

$$\left\{i_{t_q} + 2|w_{t_q}| + 2, ..., (t_q + 1)_u - 1\right\} \subseteq A_w.$$

□

Lemma 7. *Let $\rho \in A$, if $t_a = m$ and $\rho \in \{i_m + 1, ..., n\}$ then $\rho \in \{m_x + 2, ..., n\} \subseteq A_w$.*

P r o o f. Assume $\rho \in \{i_m + 1, ..., m_x + 1\}$. We have $m_x + 1 = i_{t_a} + 2|w_{t_a}| + 1$. Then $\rho \in \{i_{t_a} + 1, ..., i_{t_a} + 2|w_{t_a}| + 1\}$. By Lemma 5, we have $\rho \notin A$. Therefore, $\rho \in \{m_x + 2, ..., n\} \subseteq A_w$ by (5_q). $\square$

Lemma 8. *Let $\rho \in A$, then $\rho \neq (t_{a+l})_u + 1$ for all $l \in \{1, ..., b\}$.*

P r o o f. Let $l \in \{1, ..., b\}$. Assume $\rho = (t_{a+l})_u + 1$. Suppose that there exists $q \in \{1, ..., a\}$ with $t_q > t_{a+l}$. Then

$$\begin{aligned} v_\rho w_{t_1} ... w_{t_q} ... w_{t_a} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}} &\overset{(R3)}{\approx} w_{t_1} ... v_\rho w_{t_q} ... w_{t_a} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}} \\ &\overset{(R4)}{\approx} w_{t_1} ... w_{t_q} ... w_{t_a} v_{\rho + 2|w_{t_q} ... w_{t_a}|} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}}. \end{aligned}$$

Since

$$(t_{a+l})_u + 1 = i_{t_{a+l}} + 2|w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+l}}| - 2|w_{t_q} ... w_{t_a}| + 1,$$

we have

$$\rho + 2|w_{t_q} ... w_{t_a}| = i_{t_{a+l}} + 2|w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+l}}| + 1.$$

Suppose $t_q < t_{a+l}$ for all $q \in \{1, ..., a\}$. Then we have

$$(t_{a+l})_u + 1 = i_{t_{a+l}} + 2|w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+l}}| + 1,$$

i.e.

$$v_\rho w_{t_1} ... w_{t_q} ... w_{t_a} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}} \overset{(R3)}{\approx} w_{t_1} ... w_{t_q} ... w_{t_a} v_\rho w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}}.$$

Both cases imply

$$\begin{aligned} & w_{t_1} ... w_{t_q} ... w_{t_a} v_{i_{t_{a+l}} + 2|w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+l}}| + 1} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+b}} \\ \overset{(R4)}{\approx} & w_{t_1} ... w_{t_q} ... w_{t_a} w^{-1}_{t_{a+1}} ... v_{i_{t_{a+l}} + 2|w^{-1}_{t_{a+l}}| + 1} w^{-1}_{t_{a+l}} ... w^{-1}_{t_{a+b}} \overset{(R6)}{\approx} w_{t_1} ... w_{t_q} ... w_{t_a} w^{-1}_{t_{a+1}} ... w^{-1}_{t_{a+l}} ... w^{-1}_{t_{a+b}}, \end{aligned}$$

i.e. we can cancel v_ρ in $\hat{w}$ using $(R3)$, $(R4)$, and $(R6)$, a contradiction. $\square$

Lemma 9. *Let $\rho \in A$ and let $l \in \{1, ..., b\}$ such that $t_{a+l} \neq m$. If $\rho \in \{(t_{a+l})_u + 1, ..., (t_{a+l} + 1)_u - 1\}$ then*

$$\rho \in \{(t_{a+l})_u + 2, ..., (t_{a+l} + 1)_u - 1\} \subseteq A_w.$$

P r o o f. It is a consequence of Lemma 8 that $\rho \in \{(t_{a+l})_u + 2, ..., (t_{a+l} + 1)_u - 1\}$ and by (6_q), we have $\{(t_{a+l})_u + 2, ..., (t_{a+l} + 1)_u - 1\} \subseteq A_w$. $\square$

Lemma 10. *Let $\rho \in A$. If $t_{a+1} = m$ and $\rho \in \{m_u + 1, ..., n\}$ then $\rho \in \{m_u + 2, ..., n\} \subseteq A_w$.*

P r o o f. Suppose $\rho = m_u + 1 = (t_{a+1})_u + 1$. By Lemma 8, we have $\rho \notin A$. Therefore, $\rho \in \{m_u + 2, ..., n\} \subseteq A_w$ by (5_q). $\square$

Lemma 11. *If* $1 < 1_x < 1_u$ *then* $\rho \notin A$ *for all* $\rho \in \{1, ..., 1_u - 1_x\}$.

P r o o f. Let $\rho \in \{1, ..., 1_u - 1_x\}$. Assume $\rho \in A$. We observe that

$$1_u - 1_x = 2|w_{t_{a+b}}^{-1}...w_{t_{a+1}}^{-1}| - 2|w_{t_1}...w_{t_a}| = 2k$$

for some positive integer k. We put $\mathcal{U} = w_{t_1}...w_{t_a}$ and $\mathcal{X} = w_{t_{a+b}}^{-1}...w_{t_{a+1}}^{-1}$, i.e. $2k = 2|\mathcal{X}| - 2|\mathcal{U}|$ and $|\mathcal{X}| = |\mathcal{U}| + k$. Let

$$w_{t_{a+1}}^{-1}...w_{t_{a+b}}^{-1} = y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k},$$

where $y_1, ..., y_{|\mathcal{U}|+k} \in \{x_1, ..., x_{n-2}\}$. Then

$$v_\rho w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k} \overset{(R4)}{\approx} w_{t_1}...w_{t_a}v_{\rho+2|w_{t_1}...w_{t_a}|}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k}.$$

Using Remark 1, it is routine to calculate that

$$2|w_{t_{a+b}}^{-1}...w_{t_{a+1}}^{-1}| < i_{t_{a+1}} + 2|w_{t_{a+1}}^{-1}|,$$

i.e.

$$(1_u - 1_x) + 2|w_{t_1}...w_{t_a}| = 2|w_{t_{a+b}}^{-1}...w_{t_{a+1}}^{-1}| < i_{t_{a+1}} + 2|w_{t_{a+1}}^{-1}|.$$

This implies

$$\rho + 2|w_{t_1}...w_{t_a}| \leq i_{t_{a+1}} + 2|w_{t_{a+1}}^{-1}|.$$

Then

$$w_{t_1}...w_{t_a}v_{\rho+2|w_{t_1}...w_{t_a}|}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k} \overset{(R4)}{\approx} w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}v_\rho y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k}.$$

Note that $1_u - 1_x$ is even and there is $i \in \{2, 4, ..., 1_u - 1_x\}$ such that $\rho \in \{i-1, i\}$. If $\rho = i - 1$ then

$$\rho - 2|y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+i/2-1}| = 1.$$

If $\rho = i$ then

$$\rho - 2|y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+i/2-1}| = 2.$$

Thus,

$$\begin{aligned} & w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}v_\rho y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+k} \\ \overset{(R4)}{\approx}\ & w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...v_{\rho-2|y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+i/2-1}|}y_{|\mathcal{U}|+i/2}...y_{|\mathcal{U}|+(1_u-1_x)/2} \\ =\ & w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...v_{\hat{\rho}}y_{|\mathcal{U}|+i/2}...y_{|\mathcal{U}|+(1_u-1_x)/2} \end{aligned}$$

(where $\hat{\rho} \in \{1, 2\}$)

$$\overset{(R6)}{\approx} w_{t_1}...w_{t_a}y_1...y_{|\mathcal{U}|}y_{|\mathcal{U}|+1}...y_{|\mathcal{U}|+i/2}...y_{|\mathcal{U}|+(1_u-1_x)/2},$$

i.e. we can cancel v_ρ in $\hat{w}$ using (R4) and (R6), a contradiction. □

Lemma 12. *Let* $\rho \in A$ *with* $\rho \in \{1, ..., 1_u - 1\}$. *If* $1 < 1_u \leq 1_x$ *then* $\rho \in \{1, ..., 1_u - 1\} \subseteq A_w$ *and if* $1 < 1_x < 1_u$ *then* $\rho \in \{1_u - 1_x + 1, ..., 1_u - 1\} \subseteq A_w$.

P r o o f. If $1 < 1_u \leq 1_x$ then $\{1, ..., 1_u - 1\} \subseteq A_w$ by (7_q). If $1 < 1_x < 1_u$, it is a consequence of Lemma 11 that $\rho \in \{1_u - 1_x + 1, ..., 1_u - 1\}$ and by (7_q), we have $\{1_u - 1_x + 1, ..., 1_u - 1\} \subseteq A_w$. □

Lemma 13. *We have $(t_q)_u \notin A$ for all $q \in \{1, ..., a\}$.*

P r o o f. Let $q \in \{1, ..., a\}$. We have

$$w_{t_q} = u_{i_{t_q}} u_{i_{t_q}+2} ... u_{i_{t_q}+2|w_{t_q}|-2}$$

and $(t_q)_u = i_{t_q}$. Assume $(t_q)_u \in A$. If $i_{t_q} \geq 2$ then

$$\begin{aligned} v_{i_{t_q}} w_{t_1} ... w_{t_q} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} &\overset{(R3)}{\approx} w_{t_1} ... v_{i_{t_q}} u_{i_{t_q}} u_{i_{t_q}+2} ... u_{i_{t_q}+2|w_{t_q}|-2} w_{t_{q+1}} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R18)}{\approx} w_{t_1} ... v_{i_{t_q}+2|w_{t_q}|+1} u_{i_{t_q}-1} u_{i_{t_q}+1} ... u_{i_{t_q}+2|w_{t_q}|-3} w_{t_{q+1}} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1}. \end{aligned}$$

If $i_{t_q} = 1$ then $q = 1$ and

$$\begin{aligned} v_{i_{t_1}} w_{t_1} w_{t_2} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} &= v_1 u_1 u_3 ... u_{1+2|w_{t_1}|-2} w_{t_2} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R16)}{\approx} v_1 v_2 ... v_{1+2|w_{t_1}|+1} w_{t_2} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1}. \end{aligned}$$

We observe that we can replace several letters in $\hat{w}$ by letters with decreasing index by $(R18)$ and the letters $u_1, u_3, ..., u_{1+2|w_{t_1}|-2}$ were canceled in $\hat{w}$ by $(R16)$, respectively, a contradiction. □

Lemma 14. *We have $(t_{a+l})_u \notin A$ for all $l \in \{1, ..., b\}$.*

P r o o f. Let $l \in \{1, ..., b\}$. Now assume that $(t_{a+l})_u \in A$. We will have the following two cases. In the first case, we suppose that there exists $q \in \{1, ..., a\}$ with $t_q > t_{a+l}$ and, of course, for the trivial second case is supposed $t_q < t_{a+l}$ for all $q \in \{1, ..., a\}$. Using $(R3)$ and $(R4)$ in the first case and $(R4)$ in the second case, together with a few tedious calculations, both cases imply

$$v_{(t_{a+l})_u} w_{t_1} ... w_{t_q} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \approx w_{t_1} ... w_{t_a} v_{i_{t_{a+l}}+2|w_{t_{a+1}}^{-1} ... w_{t_{a+l}}^{-1}|} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1}.$$

It is routine to calculate that

$$w_{t_1} ... w_{t_a} v_{i_{t_{a+l}}+2|w_{t_{a+1}}^{-1} ... w_{t_{a+l}}^{-1}|} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \overset{(R4)}{\approx} w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... v_{i_{t_{a+l}}+2|w_{t_{a+l}}^{-1}|} w_{t_{a+l}}^{-1} ... w_{t_{a+b}}^{-1}.$$

If $i_{t_{a+l}} + 2|w_{t_{a+l}}^{-1}| > 3$ then

$$\begin{aligned} & w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... v_{i_{t_{a+l}}+2|w_{t_{a+l}}^{-1}|} w_{t_{a+l}}^{-1} ... w_{t_{a+b}}^{-1} \\ &= w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... v_{i_{t_{a+l}}+2|w_{t_{a+l}}^{-1}|} x_{i_{t_{a+l}}+2|w_{t_{a+l}}|-2} x_{i_{t_{a+l}}+2|w_{t_{a+l}}|-4} ... x_{i_{t_{a+l}}} w_{t_{a+l+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R19)}{\approx} w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... v_{i_{t_{a+l}}+2|w_{t_{a+l}}^{-1}|+1} x_{i_{t_{a+l}}+2|w_{t_{a+l}}|-3} x_{i_{t_{a+l}}+2|w_{t_{a+l}}|-5} ... x_{i_{t_{a+l}}-1} w_{t_{a+l+1}}^{-1} ... w_{t_{a+b}}^{-1}. \end{aligned}$$

If $i_{t_{a+l}} + 2|w_{t_{a+l}}^{-1}| = 3$ then $w_{t_{a+b}}^{-1} = x_1$. Thus,

$$\begin{aligned} & w_{t_1} ... w_{t_a} v_{i_{t_{a+l}}+2|w_{t_{a+1}}^{-1} ... w_{t_{a+l}}^{-1}|} w_{t_{a+1}}^{-1} ... w_{t_{a+b}}^{-1} \\ &\overset{(R4)}{\approx} w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b-1}}^{-1} v_3 x_1 \overset{(R17)}{\approx} w_{t_1} ... w_{t_a} w_{t_{a+1}}^{-1} ... w_{t_{a+b-1}}^{-1} v_1 v_2 v_3 v_4. \end{aligned}$$

We observe that we can replace several letters in $\hat{w}$ by letters with decreasing index by $(R19)$ and the letter x_1 can be canceled in $\hat{w}$ by $(R17)$, respectively, a contradiction. □

If we summarize the previous lemmas, then we obtain:

Lemma 15. *We have $A \subseteq A_w$.*

P r o o f. Let $\rho \in A$. Then it is easy to verify that $\rho \in \{1, ..., 1_u\}$ or $\rho \in \{k_u + 1, ..., (k+1)_u\}$ for some $k \in \{1, ..., m-1\}$ or $\rho \in \{m_u + 1, ..., n\}$. Suppose that $\rho \in \{k_u + 1, ..., (k+1)_u - 1\}$ for some $k \in \{1, ..., m-1\}$. Lemmas 13 and 14 show that $k_u \notin A$. Then we can conclude that $\rho \in A_w$ by Lemmas 6 and 9. Suppose $\rho \in \{m_u + 1, ..., n\}$. Then we can conclude that $\rho \in A_w$ by Lemmas 7 and 10. Finally, we suppose that $\rho \in \{1, ..., 1_u - 1\}$. Then we can conclude that $\rho \in A_w$ by Lemma 12. Eventually, we have $\rho \in A_w$ for all $\rho \in A$. Therefore, $A \subseteq A_w$. □

Lemmas 4 and 15 prove that $\hat{w} = v_A \hat{w}_1 ... \hat{w}_m \in W_n$. Consequently, we have:

Proposition 2. $P \subseteq W_n$.

By the definition of the set P and Proposition 2, it is proved:

Corollary 1. *Let $w \in X_n^*$. Then there is $w' \in P \subseteq W_n$ with $w \approx w'$.*

4. A presentation for IOF_n^{par}

In this section, we exhibit a presentation for IOF_n^{par}. Concerning the results from the previous sections, it remains to show that $|W_n| \leq |IOF_n^{par}|$. For this, we construct a word w_α, for all $\alpha \in IOF_n^{par}$, in the following way.

Let

$$\alpha = \begin{pmatrix} d_1 & < & d_2 & < & \cdots & < & d_p \\ m_1 & & m_2 & & \cdots & & m_p \end{pmatrix} \in IOF_n^{par} \backslash \{\varepsilon\}$$

for a positive integer $p \leq n$. There are a unique $l \in \{0, 1, ..., p-1\}$ and a unique set $\{r_1, ..., r_l\} \subseteq \{1, ..., p-1\}$ such that (i)–(iii) are satisfied:

(i) $r_1 < ... < r_l$;
(ii) $d_{r_i+1} - d_{r_i} \neq m_{r_i+1} - m_{r_i}$ for $i \in \{1, ..., l\}$;
(iii) $d_{i+1} - d_i = m_{i+1} - m_i$ for $i \in \{1, ..., p-1\} \backslash \{r_1, ..., r_l\}$.

Note that $l = 0$ means $\{r_1, ..., r_l\} = \emptyset$. Further, we put $r_{l+1} = p$. For $i \in \{1, ..., l\}$, we define

$$w_i = \begin{cases} x_{m_{r_i}, ((m_{r_i+1} - m_{r_i}) - (d_{r_i+1} - d_{r_i}))/2} & \text{if} \quad m_{r_i+1} - m_{r_i} > d_{r_i+1} - d_{r_i}; \\ u_{d_{r_i}, ((d_{r_i+1} - d_{r_i}) - (m_{r_i+1} - m_{r_i}))/2} & \text{if} \quad m_{r_i+1} - m_{r_i} < d_{r_i+1} - d_{r_i}. \end{cases}$$

Obviously, we have $w_i \in W_x \cup W_u$ for all $i \in \{1, ..., l\}$. If $m_p = d_p$ then we put $w_{l+1} = \epsilon$. If $m_p \neq d_p$, we define additionally

$$w_{l+1} = \begin{cases} x_{m_p, (d_p - m_p)/2} & \text{if} \quad d_p > m_p; \\ u_{d_p, (m_p - d_p)/2} & \text{if} \quad d_p < m_p. \end{cases}$$

Clearly, $w_{l+1} \in W_x \cup W_u$. We consider the word

$$w = w_1 ... w_{l+1}.$$

From this word, we construct a new word w_α^* by arranging the subwords $s \in W_x$ in reverse order at the end, replacing s by s^{-1}. In other words, we consider the word

$$w_\alpha^* = w_{s_1} ... w_{s_a} w_{s_{a+1}}^{-1} ... w_{s_{a+b}}^{-1}$$

such that $w_{s_1}, ..., w_{s_a} \in W_u$, $w_{s_{a+1}}, ..., w_{s_{a+b}} \in W_x$ and

$$\{w_{s_1}, ..., w_{s_a}, w_{s_{a+1}}, ..., w_{s_{a+b}}\} = \{w_1, ..., w_{a+b}\},$$

where $s_1 < ... < s_a, s_{a+b} < ... < s_{a+1}$, and $a, b \in \overline{n} \cup \{0\}$ with

$$a + b = \begin{cases} l & \text{if} \quad d_p = m_p; \\ l+1 & \text{if} \quad d_p \neq m_p. \end{cases}$$

For convenience, $a = 0$ means $w_\alpha^* = w_{s_{a+1}}^{-1} ... w_{s_{a+b}}^{-1}$ and $b = 0$ means $w_\alpha^* = w_{s_1} ... w_{s_a}$. Now, we add recursively letters from the set $\{v_1, ..., v_n\} \subseteq X_n$ to the word w_α^*, obtaining new words $\lambda_0, \lambda_1, ..., \lambda_p$.

(1) For $d_p \leq n-2$:

(1.1) if $m_p < d_p$ then $\lambda_0 = v_{d_p+2} ... v_n w_\alpha^*$;
(1.2) if $n-1 > m_p > d_p$ then $\lambda_0 = v_{m_p+2} ... v_n w_\alpha^*$;
(1.3) if $m_p = d_p$ then $\lambda_0 = v_{m_p+1} ... v_n w_\alpha^*$;
otherwise $\lambda_0 = w_\alpha^*$.

(2) If $d_p = m_p = n-1$ then $\lambda_0 = v_n w_\alpha^*$. Otherwise $\lambda_0 = w_\alpha^*$.

(3) For $k \in \{2, ..., p\}$:

(3.1) if $2 \leq m_k - m_{k-1} = d_k - d_{k-1}$ then $\lambda_{p-k+1} = v_{d_{k-1}+1} ... v_{d_k-1} \lambda_{p-k}$;
(3.2) if $2 < m_k - m_{k-1} < d_k - d_{k-1}$ then $\lambda_{p-k+1} = v_{d_k-(m_k-m_{k-1}-2)} ... v_{d_k-1} \lambda_{p-k}$;
(3.3) if $m_k - m_{k-1} > d_k - d_{k-1} > 2$ then $\lambda_{p-k+1} = v_{d_{k-1}+2} ... v_{d_k-1} \lambda_p$;
otherwise $\lambda_{p-k+1} = \lambda_{p-k}$.

(4) If $d_1 = 1$ or $m_1 = 1$ then $\lambda_p = \lambda_{p-1}$.

(5) If $1 < d_1 \leq m_1$ then $\lambda_p = v_1 ... v_{d_1-1} \lambda_{p-1}$.

(6) If $1 < m_1 < d_1$ then $\lambda_p = v_{d_1-m_1+1} ... v_{d_1-1} \lambda_{p-1}$.

The word λ_p induces a set $A = \{a \in \overline{n} : v_a \text{ is a letter in } \lambda_p\}$ and it is easy to verify that $\rho \notin A$ for all $\rho \in dom(\alpha)$. We put $w_\alpha = \lambda_p$. The word w_α has the form $w_\alpha = v_A w_\alpha^*$.

Our next aim is to present the relationship between cardinality of W_n and IOF_n^{par}. This leads us to assume the existence of a map $f : IOF_n^{par} \backslash \{\varepsilon\} \to W_n \backslash \{v_{\overline{n}}\}$, where $f(\alpha) = w_\alpha$ for all $\alpha \in IOF_n^{par} \backslash \{\varepsilon\}$. We start by constructing the transformation $\alpha_{v_A w^*}$ for any $v_A w^* \in W_n$, different from $v_{\overline{n}}$. Let $v_A w^* \in W_n \backslash \{v_{\overline{n}}\}$. We have $w \in Q_0, A \subseteq A_w$, and there are $w_1, ..., w_m \in W_u \cup W_x$ such that $w = w_1 ... w_m$ for some positive integer m. For $k \in \{1, ..., m\}$, we define $a_k = k_u + 2$ and $b_k = i_k + 2j_k + 2$, whenever $w_k \in W_x$. On the other hand, we define $a_k = i_k + 2j_k + 2$ and $b_k = k_x + 2$, whenever $w_k \in W_u$. It is easy to verify that $a_m = b_m$. We put

$$\alpha_{v_A w^*} = \overline{v}_A \begin{pmatrix} 1 + 1_u - \min\{1_u, 1_x\} ... 1_u & a_1 ... 2_u & \cdots & a_{m-1} ... m_u & a_m ... n \\ 1 + 1_x - \min\{1_u, 1_x\} ... 1_x & b_1 ... 2_x & \cdots & b_{m-1} ... m_x & b_m ... n \end{pmatrix}.$$

For convenience, we also give

$$\alpha_{v_A w^*} = \begin{pmatrix} d_1 & d_2 & \cdots & d_p \\ m_1 & m_2 & \cdots & m_p \end{pmatrix}$$

for some positive integer $p \leq n$. In the following, we show that $\alpha_{v_A w^*}$ is well-defined in the sense that the construction of $\alpha_{v_A w^*}$ gives a transformation.

Lemma 16. *$\alpha_{v_A w^*}$ is well-defined.*

P r o o f. Let $k \in \{1, ..., m-1\}$. Suppose $w_k, w_{k+1} \in W_u$. We have

$$k_u = i_k, \quad k_x = i_k + 2|w_k| + 2|W_u^k| - 2|W_x^k|,$$
$$(k+1)_u = i_{k+1}, (k+1)_x = i_{k+1} + 2|w_{k+1}| + 2|W_u^{k+1}| - 2|W_x^{k+1}|,$$

and $a_k = i_k + 2j_k + 2, b_k = k_x + 2$. Then

$$(k+1)_u - a_k = i_{k+1} - (i_k + 2j_k + 2),$$
$$(k+1)_x - b_k = i_{k+1} + 2|w_{k+1}| + 2|W_u^{k+1}| - 2|W_x^{k+1}| - k_x - 2$$
$$= i_{k+1} + 2|w_{k+1}| + 2|W_u^{k+1}| - 2|W_x^{k+1}| - i_k - 2|w_k| - 2|W_u^k| + 2|W_x^k| - 2$$
$$= i_{k+1} - i_k - 2j_k - 2 = i_{k+1} - (i_k + 2j_k + 2).$$

Therefore, $(k+1)_u - a_k = (k+1)_x - b_k$.

For the rest cases ($w_k \in W_u$ and $w_{k+1} \in W_x$, $w_k \in W_x$ and $w_{k+1} \in W_u$ as well as $w_k, w_{k+1} \in W_x$), a proof similar as above will eventually show that $(k+1)_u - a_k = (k+1)_x - b_k$. Furthermore, suppose $d_p = m_p$. Let $k \in \{1, ..., m\}$ and $w_k \in W_u$. We have

$$a_k - k_u = i_k + 2j_k + 2 - k_u = i_k + 2j_k + 2 - i_k = 2j_k + 2,$$
$$b_k - k_x = k_x + 2 - k_x = 2.$$

Thus, $a_k - k_u \neq b_k - k_x$.

For the case $w_k \in W_x$, we can show $a_k - k_u \neq b_k - k_x$ in the same way.

Continuously, suppose $d_p \neq m_p$. By the previous part of the proof, we have $a_k - k_u \neq b_k - k_x$ for all $k \in \{1, ..., m-1\}$. Moreover, we observe that $d_p \notin \{a_m, ..., n\}$ and $m_p \notin \{b_m, ..., n\}$ because $n - a_m = n - b_m$. This implies $d_p = m_u$ and $m_p = m_x$. By any of the above, we can conclude that $\alpha_{v_A w^*}$ is well-defined. □

The proof of Lemma 16 shows $(k+1)_u - a_k = (k+1)_x - b_k$ for all $k \in \{1, ..., m-1\}$. Then $a_k - k_u \neq b_k - k_x$ for all $k \in \{1, ..., m\}$, whenever $d_p = m_p$, and $a_k - k_u \neq b_k - k_x$ for all $k \in \{1, ..., m-1\}$ and $d_p = m_u, m_p = m_x$, whenever $d_p \neq m_p$. Furthermore, observing by trivial calculation, $a_k - k_u \geq 2$ and $b_k - k_x \geq 2$. Therefore, if there exists $i \in \{1, ..., p-1\}$, where $d_{i+1} - d_i \neq m_{i+1} - m_i$, then $d_i \in \{1_u, ..., (m-1)_u\}(\cup\{m_u\})$, $m_i \in \{1_x, ..., (m-1)_x\}(\cup\{m_x\})$ and we put $k_u = d_{r_k}, k_x = m_{r_k}$ for all $k \in \{1, ..., m-1\}(\cup\{m\})$ (we put $r_m = p$, whenever $d_p \neq m_p$). This gives the unique set $\{r_1, ..., r_m\}$ as required by the definition of $w_{\alpha_{v_A w^*}}$. Moreover, we need to show that $\alpha_{v_A w^*} \in IOF_n^{par} \backslash \{\varepsilon\}$ by checking (i)-(iv) of Proposition 1. We will now show that $\alpha_{v_A w^*} \in IOF_n^{par}$ as well as $w_{\alpha_{v_A w^*}} = v_A w^*$. This gives the tools to calculate that $|W_n| \leq |IOF_n^{par}|$.

Lemma 17. $\alpha_{v_A w^*} \in IOF_n^{par} \backslash \{\varepsilon\}$.

P r o o f. Clearly, $\alpha_{v_A w^*} \neq \varepsilon$. We will prove that $\alpha_{v_A w^*}$ satisfies the conditions (i)–(iv) in Proposition 1. We observe that $d_1 < d_2 < \cdots < d_p$ and $m_1 < m_2 < \cdots < m_p$ by definition of $\alpha_{v_A w^*}$. We have $1_u - d_1 = 1_x - m_1$, i.e. $1_u - 1_x = d_1 - m_1$. By the definition of k_u and k_x, for $k \in \{1, ..., m\}$, we observe that $1_u - 1_x$ is even, i.e. $d_1 - m_1$ is even. Thus, d_1 and m_1 have the same parity.

Let $d_{i+1} - d_i = 1$ for some $i \in \{1, ..., p-1\}$. Then $d_i \in dom(\alpha) \backslash \{1_u, ..., m_u\}$ implies $m_{i+1} - m_i = d_{i+1} - d_i = 1$.

Let $m_{i+1} - m_i = 1$ for some $i \in \{1, ..., p-1\}$. Then $m_i \in im(\alpha) \backslash \{1_x, ..., m_x\}$ implies $d_{i+1} - d_i = m_{i+1} - m_i = 1$.

Let $d_{i+1} - d_i$ is even. Suppose $d_{i+1} - d_i \neq m_{i+1} - m_i$. This gives $d_i = k_u$ and $m_i = k_x$ for some $k \in \{1, ..., m-1\}$. By the definition of k_u and k_x, we observe that $k_u - k_x$ is even.

Moreover, $(k+1)_u - d_{i+1} = (k+1)_x - m_{i+1}$ since $(k+1)_u - (k+1)_x$ is even, we have $d_{i+1} - m_{i+1}$ is even. Then d_{i+1}, d_i and d_i, m_i as well as d_{i+1}, m_{i+1} have the same parity. This implies that m_{i+1}, m_i have the same parity, i.e. $m_{i+1} - m_i$ is even. Conversely, we can prove similarly that, if $m_{i+1} - m_i$ is even then $d_{i+1} - d_i$ is even. By Proposition 1, we get $\alpha_{v_A w^*} \in IOF_n^{par}$. □

We can construct $f(\alpha_{v_A w^*}) = w_{\alpha_{v_A w^*}}$, where $w_{\alpha_{v_A w^*}} = v_{\tilde{A}} \hat{w}^*_{\alpha_{v_A w^*}}$ with $\hat{w} = \hat{w}_1 ... \hat{w}_m$ for $\hat{w}_1, ..., \hat{w}_m \in W_u \cup W_x$ and $\tilde{A} \subseteq \overline{n}$. We will prove that f is surjective in the next lemma.

Lemma 18. *Let $v_A w^* \in W_n \backslash \{v_{\overline{n}}\}$. Then there is $\alpha \in IOF_n^{par} \backslash \{\varepsilon\}$ with $v_A w^* = w_\alpha$.*

P r o o f. We have $w_{\alpha_{v_A w^*}} = v_{\tilde{A}} \hat{w}^*_{\alpha_{v_A w^*}}$, where $\hat{w} = \hat{w}_1 ... \hat{w}_m$ with $\hat{w}_1, ..., \hat{w}_m \in W_u \cup W_x$ and $\tilde{A} \subseteq \overline{n}$. First, our goal is to show that $\hat{w} = w$. Suppose $d_p = m_p$ and let $k \in \{1, ..., m\}$ such that $b_k - k_x > a_k - k_u$. By the definition of $\hat{w}_k$, we have $\hat{w}_k = x_{k_x, ((b_k - k_x) - (a_k - k_u))/2}$ and $k_x = i_k$. Then

$$\frac{(b_k - k_x) - (a_k - k_u)}{2} = \frac{i_k + 2j_k + 2 - i_k - k_u - 2 + k_u}{2} = j_k,$$

i.e. $\hat{w}_k = x_{i_k, j_k} = w_k$. For the case $b_k - k_x < a_k - k_u$, we can prove that $\hat{w}_k = w_k$ in a similar way. This gives $\hat{w}_1 ... \hat{w}_m = w_1 ... w_m$.

Suppose $d_p \neq m_p$. We have $a_k - k_u \neq b_k - k_x$ for all $k \in \{1, ..., m-1\}$ and by a similar proof as above, we have $\hat{w}_1 ... \hat{w}_{m-1} = w_1 ... w_{m-1}$. If $m_p < d_p$ then $\hat{w}_m = x_{m_p, (d_p - m_p)/2}$ and $m_p = m_x = i_m$. Then

$$\frac{d_p - m_p}{2} = \frac{m_u - m_x}{2} = \frac{i_m + 2j_m - i_m}{2} = j_m,$$

i.e. $\hat{w}_m = x_{i_m, j_m} = w_m$. For the case $m_p > d_p$, we can prove $\hat{w}_m = w_m$ in a similar way. Thus, $\hat{w}_1 ... \hat{w}_{m-1} \hat{w}_m = w_1 ... w_{m-1} w_m$. Then $w = \hat{w}$, i.e. $w^* = \hat{w}^*_{\alpha_{v_A w^*}}$. The next goal is to show that $A = \tilde{A}$.

1) To show that $A \subseteq \tilde{A}$: let $a \in A$. We have $A \subseteq A_w$ since $v_A w^* \in W_n$. Therefore, we have the following cases: $a \in \{a_m, ..., n\} = A_1$ or $a \in \{a_k, ..., (k+1)_u - 1\} = A_2$ for some $k \in \{1, ..., m-1\}$ or

$$a \in \{1 + 1_u - \min\{1_u, 1_x\}, ..., 1_u - 1\} = A_3.$$

If $a \in A_1$ and $m_p \neq d_p$ then $a \in \tilde{A}$ since (1.1) and (1.2), respectively. If $a \in A_1$ and $a \in \{d_p + 1, ..., n\}$ with $m_p = d_p$ then $a \in \tilde{A}$ since (1.3) and (2), respectively.

Suppose $a \in A_2$ with $a \in \{a_k, ..., d_{r_k+1} - 1\}$. If $2 < d_{r_k+1} - d_{r_k} < m_{r_k+1} - m_{r_k}$ then $w_k \in W_x$. Note that $a_k = k_u + 2 = d_{r_k} + 2$. Thus, $a \in \tilde{A}$ since (3.3). If $2 < m_{r_k+1} - m_{r_k} < d_{r_k+1} - d_{r_k}$ then $w_k \in W_u$.

Note

$$d_{r_k+1} - a_k = m_{r_k+1} - b_k, \quad b_k = k_x + 2,$$
$$a_k = a_k - b_k + b_k = d_{r_k+1} - m_{r_k+1} + k_x + 2 = d_{r_k+1} - m_{r_k+1} + m_{r_k} + 2.$$

Thus, $a \in \tilde{A}$ since (3.2).

Suppose $a \in A_3$. If $1 < d_1 \leq m_1$ and $a \in \{1, ..., d_1 - 1\}$ then $a \in \tilde{A}$ since (5). If $1 < m_1 < d_1$ and $a \in \{d_1 - m_1 + 1, ..., 1_u - 1\}$ then $a \in \tilde{A}$ since (6) (note that $1_u - 1_x = d_1 - m_1$).

Suppose $a \in A_1 \cup A_2 \cup A_3$ and there exists $s \in \{2, ..., p\}$ such that $d_s - d_{s-1} = m_s - m_{s-1} \geq 2$ with $a \in \{d_{s-1} + 1, ..., d_s - 1\}$. Then $a \in \tilde{A}$ since (3.1). By any of the above, we have $A \subseteq \tilde{A}$.

2) To show that $\tilde{A} \subseteq A$: let

$$A_1 = \{1 + 1_u - \min\{1_u, 1_x\}, ..., 1_u - 1\},$$
$$A_2 = \{a_1, ..., 2_u - 1\} \cup \{a_2, ..., 3_u - 1\} \cup ... \cup \{a_{m-1}, ..., m_u - 1\},$$
$$A_3 = \{a_m, ..., n\}.$$

Because $A \subseteq A_w$, we have $A \subseteq A_1 \cup A_2 \cup A_3$ and $A \cap \{d_1, ..., d_p\} = \emptyset$. This implies $A \subseteq A_1 \cup A_2 \cup A_3 \backslash \{d_1, ..., d_p\}$. Conversely, we have $A_1 \cup A_2 \cup A_3 \backslash \{d_1, ..., d_p\} \subseteq A$ by the definition of $\alpha_{v_A w^*}$. Thus, $A = A_1 \cup A_2 \cup A_3 \backslash \{d_1, ..., d_p\}$.

Let $a \in \tilde{A}$. By the definition of $\tilde{A}$, we can observe that $a \neq d_i$ for all $i \in \{1, ..., p\}$.

Suppose a is given by (1.1) or (1.2) or (1.3) or (2). Then $a \in A_3 \backslash \{d_1, ..., d_p\}$.

Suppose a is given by (3.1). Then $a \in A_1 \cup A_2 \cup A_3 \backslash \{d_1, ..., d_p\}$.

Suppose a is given by (3.2), i.e. $a \in \{d_s - m_s + m_{s-1} + 2, ..., d_s - 1\}$ for some $s \in \{2, ..., p\}$. We have already shown that there is $k \in \{1, ..., m-1\}$ such that $d_s - m_s + m_{s-1} + 2 = a_k$. Then $a \in A_2 \backslash \{d_1, ..., d_p\}$.

Suppose a is given by (3.3). Then $a \in A_2 \backslash \{d_1, ..., d_p\}$.

Suppose a is given by (5). Then $a \in A_1 \backslash \{d_1, ..., d_p\}$.

Suppose a is given by (6). Then $a \in A_1 \backslash \{d_1, ..., d_p\}$ (note that $d_1 - m_1 = 1_u - 1_x$). Therefore, we have $a \in A$, i.e. $\tilde{A} \subseteq A$.

By 1) and 2), we get $A = \tilde{A}$. This implies $v_A w^* = v_{\tilde{A}} \hat{w}^* = w_{\alpha_{v_A w^*}}$. □

Lemma 18 establishes that f is surjective, which implies $|W_n| \leq |IOF_n^{par}|$. We will now adjust our alphabet and relations to meet the requirements of Theorem 1. As mentioned previously, $\overline{X}_n = \{\overline{s} : s \in X_n\}$ is a generating set for the monoid IOF_n^{par}. Building on the insights from Lemma 1, we can conclude that $\overline{X}_n$ satisfies all the relations from $\overline{R} = \{\overline{s}_1 \approx \overline{s}_2 : s_1 \approx s_2 \in R\}$.

Corollary 1 further shows that for any $w \in \overline{X}_n^*$, there exists a corresponding $w' \in \overline{W}_n$, for which $w \approx w'$ is a consequence of $\overline{R}$. This implies that $\overline{R} \subseteq \overline{X}_n^* \times \overline{X}_n^*$ and that $\overline{W}_n \subseteq \overline{X}_n^*$ meet the conditions 1–3 in Theorem 1. We now possess all the necessary items to conclude our main result.

Theorem 2. *$\langle \overline{X}_n \,|\, \overline{R} \rangle$ is a monoid presentation for IOF_n^{par}.*

REFERENCES

1. Clifford A. H., Preston G. B. *The Algebraic Theory of Semigroups.* Math. Surveys Monogr., vol. 7, part 1. Providence. R.I.: Amer. Math. Soc. Surveys, 1961. DOI: 10.1090/surv/007.1; Math. Surveys Monogr., vol. 7, part 2. Providence. R.I.: Amer. Math. Soc. Surveys, 1967. DOI: 10.1090/surv/007.2
2. Currie J. D., Visentin T. I. The number of order-preserving maps of fences and crowns. *Order*, 1991. Vol. 8. P. 133–142.
3. Dimitrova I., Koppitz J., Lohapan L. Generating sets of semigroups of partial transformations preserving a zig-zag order on ℕ. *Int. J. Pure Appl. Math.*, 2017. Vol. 117, No. 2. P. 279–289. DOI: 10.12732/ijpam.v117i2.4
4. East J. A presentation of the singular part of the symmetric inverse monoid. *Comm. Algebra*, 2006. Vol. 34, No. 5. P. 1671–1689. DOI: 10.1080/00927870500542689
5. East J. A presentation for the singular part of the full transformation semigroup. *Semigroup Forum*, 2010. Vol 81. P. 357–379. DOI: 10.1007/s00233-010-9250-1
6. Fernandes V. H. The monoid of all injective order preserving partial transformations on a finite chain. *Semigroup Forum*, 2001. Vol. 62. P. 178–204. DOI: 10.1007/s002330010056
7. Fernandes V. H., Koppitz J., Musunthia T. The rank of the semigroup of all order-preserving transformations on a finite fence. *Bull. Malays. Math. Sci. Soc.*, 2019. Vol. 42. P. 2191–2211. DOI: 10.1007/s40840-017-0598-1
8. Howie J. M. *Fundamentals of Semigroup Theory.* Oxford: Clarendon Press, 1995. 351 p.

9. Jendana K., Srithus R. Coregularity of order-preserving self-mapping semigroups of fences. *Commun. Korean Math. Soc.*, 2015. Vol. 30, No. 4. P. 349–361. DOI: 10.4134/CKMS.2015.30.4.349
10. Lallement G. *Semigroups and Combinatorial Applications*. New York: Wiley, 1979. 376 p.
11. Lohapan L., Koppitz J. Regular semigroups of partial transformations preserving a fence ℕ. *Novi Sad J. Math.*, 2017. Vol. 47, No. 2. P. 77–91. DOI: 10.30755/NSJOM.05333
12. Lohapan L., Koppitz J., Worawiset S. Congruences on infinite semigroups of transformations preserving a zig-zag order. *J. Algebra Appl.*, 2021. Vol. 20, No. 09. Art. no. 2150167. DOI: 10.1142/S021949882150167X
13. Ruškuc N. *Semigroup Presentations*. Ph. D. Thesis. St Andrews: University of St Andrews, 1995.
14. Rutkowski A. The formula for the number of order-preserving selfmappings of a fence. *Order*, 1992. Vol. 9. P. 127–137. DOI: 10.1007/BF00814405
15. Sareeto A., Koppitz J. The rank of the semigroup of order-, fence- and parity-preserving partial injections on a finite set. *Asian-Eur. J. Math.*, 2023. DOI: 10.1142/S1793557123502236
16. Srithus R., Chinram R., Khongthat C. Regularity in the semigroup of transformations preserving a zig-zag order. *Bull. Malays. Math. Sci. Soc.*, 2020. Vol. 43, No. 2. P. 1761–1773. DOI: 10.1007/s40840-019-00772-2
17. Wagner V. V. Generalized groups. *Dokl. Akad. Nauk SSSR*, 1952. Vol. 84, No. 6. P. 1119–1122. (in Russian)

URAL MATHEMATICAL JOURNAL, Vol. 9, No. 2, 2023, pp. 193–208

DOI: 10.15826/umj.2023.2.016

GRACEFUL CHROMATIC NUMBER OF SOME CARTESIAN PRODUCT GRAPHS[1]

I Nengah Suparta[a,†], **Mathiyazhagan Venkathacalam**[b,††], **I Gede Aris Gunadi**[a,†††], **Putu Andi Cipta Pratama**[a,††††]

[a]Department of Mathematics, Universitas Pendidikan Ganesha, Jl. Udayana 11, Singaraja-Bali 81117, Indonesia

[b]Department of Mathematics, Kongunadu Arts and Science College, Coimbatore–641029, Tamil Nadu, India

[†]nengah.suparta@undiksha.ac.id [††]venkatmaths@kongunaducollege.ac.in
[†††]igedearisgunadi@undiksha.ac.id [††††]andicipta25@gmail.com

Abstract: A graph $G(V,E)$ is a system consisting of a finite non empty set of vertices $V(G)$ and a set of edges $E(G)$. A (*proper*) *vertex colouring* of G is a function $f: V(G) \to \{1,2,\ldots,k\}$, for some positive integer k such that $f(u) \neq f(v)$ for every edge $uv \in E(G)$. Moreover, if $|f(u)-f(v)| \neq |f(v)-f(w)|$ for every adjacent edges $uv, vw \in E(G)$, then the function f is called *graceful colouring* for G. The minimum number k such that f is a graceful colouring for G is called the *graceful chromatic number* of G. The purpose of this research is to determine graceful chromatic number of Cartesian product graphs $C_m \times P_n$ for integers $m \geq 3$ and $n \geq 2$, and $C_m \times C_n$ for integers $m, n \geq 3$. Here, C_m and P_m are cycle and path with m vertices, respectively. We found some exact values and bounds for graceful chromatic number of these mentioned Cartesian product graphs.

Keywords: Graceful colouring, Graceful chromatic number, Cartesian product.

1. Introduction

A graph $G(V,E)$ is a system consisting of a finite non empty set of vertices $V(G)$ and a set of edges $E(G)$. Let G and H be two disjoint graphs. The *Cartesian product* of G and H, denoted by $G \times H$, is the graph with vertex set $V(G) \times V(H)$, and edges $xy, uv \in V(G) \times V(H)$ are adjacent in $G \times H$, if $x = u$ and $yv \in E(H)$ or $y = v$ and $xu \in E(G)$. A (*proper*) *vertex colouring* of G is a way of colouring vertices in G such that each adjacent vertices are assigned to different colours.

If for a vertex colouring of G we have that every adjacent edges in G have different induced colours, then the vertex colouring is called *graceful*. We may think a graceful colouring of G as a function $f: V(G) \to \{1,2,\ldots,k\}$, for some positive integer k, such that for every edge $uv \in E(G)$ we have $f(u) \neq f(v)$, and for any vertex $u \in V(G)$ we have $|f(u)-f(v)| \neq |f(u)-f(w)|$ for every vertices $v, w \in V(G)$ which are adjacent to u. The absolute value $|f(u)-f(v)|$ for every $uv \in E(G)$, is the *induced label* of the edge uv ∈E(G). In this sense, the terms colour and label are interchangeable. The smallest value of k for which the function f is a graceful vertex colouring of G is called the *graceful chromatic number* of G. The graceful colouring is a variation of graceful labeling which was introduced by Alexander Rosa in 1967 (see Gallian in [5]). Whereas, the notion of graceful colouring was introduced by Gary Chartrand in 2015, as a variant of the proper vertex k-colouring problem (see [3]). Since then, researches on graceful colouring numbers started to be celebrated.

Byers in [3] derived exact values for the graceful chromatic number of some graphs: path, cycle, wheel, and caterpillar; and introduced some bounds for certain connected regular graphs.

[1]This work was supported by LP2M of Universitas Pendidikan Ganesha.

Moreover, English, et al. in [4] invented graceful chromatic number of some classes of trees, and gave a lower bound for the graceful chromatic number of connected graphs with certain minimum degree. Mincu et al. in [6] derived graceful chromatic number of some well-known graph classes, such as diamond graph, Petersen graph, Moser spindle graph, Goldner-Harary graph, friendship graphs, and fan graphs. Graceful chromatic number of some particular unicyclic class graphs were presented by Alfarisi et al. (2019) in [1].

Furthermore, in 2022, Asy'ari et al. in [2] presented graceful chromatic numbers of several types of graphs, including star graphs, diamond graphs, book graphs. In addition, Asy'ari, et al. also stated some open problems. One of the problems is to determine the graceful chromatic number of some Cartesian product of certain graphs. Here we derive graceful chromatic number of Cartesian product graph $C_m \times P_n$, $m \geq 3, n \geq 2$, where C_m is the cycle with m vertices and P_n is the path with n vertices. The Cartesian product graph $C_m \times P_n$ is known as *prism* for $n = 2$ and as *generalized prism* for $n \geq 3$. We also introduce bounds for Cartesian product graph $C_m \times C_n$, $m, n \geq 3$.

To proceed with the main results, we need to introduce some introductory facts which will be beneficial for our further discussion.

Let G be a graph and x be a vertex of G. All vertex which are adjacent to x are called *the neighbors of* x, and denoted by $N(x)$. The degree of the vertex x, denoted by $\deg(x)$, is equal to the cardinality of $N(x)$, $\deg(x) = |N(x)|$. We will start with the following lemma.

Lemma 1. *Let G be a graph and u be a vertex in G with degree $d \geq 1$. Let f be a graceful colouring for G. If $f(u) = a$, $1 \leq a \leq d$, then there is a vertex $v \in N(u)$ with colour $f(v) \geq d + a$.*

P r o o f. Let $f(u) = a$ with $1 \leq a \leq d$. If $a = 1$, the smaller possible colours we can assign for the all d neighbors $v \in N(u)$ of u, are $2, 3, \ldots, d$ and the colour $d + 1$. This means that, there is a vertex $v \in N(u)$ with $f(v) \geq d + 1 = d + a$. We are done for the case $a = 1$.

Now, assume $f(u) = a$, $1 < a \leq d$. Note that the colours k and $2a-k$, for every $k, 1 \leq k \leq a-1$, can not be assigned simultaneously for the vertices in $N(u)$, since they give the same difference from the colour a. Therefore, the maximum number of colours we may assign from the first $2(a-1)$ smallest colours $\{k, 2a-k : 1 \leq k \leq a-1\}$ is equal to $a-1$. It implies that the remaining vertices in $N(u)$ which are not coloured yet, is at least $d-(a-1)$ vertices. The colours we need for these vertices are started from a colour $\geq 2a$. This means that the next $d-(a-1)$ smallest colours we should assign are $2a, 2a+1, \ldots, 2a+(d-(a-1)-1)$. So, there is a vertex $v \in N(u)$ such that its colour $f(v) \geq 2a + (d-(a-1)-1) = d + a$. □

In a specific case, the colour of a vertex u is equal to the degree of u, $f(u) = deg(u)$, we have the following corollary.

Corollary 1. *In a graph G with graceful colouring f, if the vertex u has degree $d \geq 1$ and colour d, then there is a vertex $v \in N(u)$ with colour $f(v) \geq 2d$.*

P r o o f. Let G be a graph and u be a vertex of G with $deg(u) = d$. Let f be a graceful colouring for G where $f(u) = d$. By Lemma 1, we found a neighbor v of u such that $f(v) \geq d + d = 2d$.□

The following result was introduced by Byers (2018) in [3].

Lemma 2 (Byers in [3]). *The graceful chromatic number of cycle C_n on $n \geq 3$ vertices is*

$$\chi_g(C_n) = \begin{cases} 4, & \text{if} \quad n \neq 5, \\ 5, & \text{if} \quad n = 5. \end{cases} \tag{1.1}$$

Then, we will introduce some terminologies related with certain ladder graphs.

A *ladder* of $2m$ vertices, $m \geq 2$, denoted by L_m, is the Cartesian product graph of the path on m vertices and the path on two vertices. The ladder L_2 is the cycle graph of four vertices. Assume that the vertices of L_m are $v_1, v_2, \ldots, v_m, w_1, w_2, \ldots, w_m$ such that its edges are v_iv_{i+1}, w_iw_{i+1} : $1 \leq i \leq m-1, v_iw_i : 1 \leq i \leq m$. For $m \geq 4$, if the vertices v_1 and v_m, and the vertices w_1 and w_m are identified, then we obtain a prism $C_{m-1} \times P_2$. In this resulting $C_{m-1} \times P_2$, $v_1 = v_m$, $w_1 = w_m$, and edge $v_1w_1 = v_mw_m$. Due to this, we may call the ladder L_m as the *open* graph of $C_{m-1} \times P_2$ about the edge v_1w_1.

On the other side, let $C_m \times P_2, m \geq 3$, be a prism. This prism has vertex set $\{v_1, v_2, \ldots, v_m, w_1, w_2, \ldots, w_m\}$ and edge set

$$\{v_iv_{i+1}, w_iw_{i+1} : 1 \leq i \leq m-1\} \cup \{v_1v_m, w_1w_m,\} \cup \{v_iw_i : 1 \leq i \leq m\}.$$

After opening $C_m \times P_2$ about the edge v_1w_1 into the ladder L_{m+1}, the vertices v_1 and w_1 copy themselves into two copies each; the first copy of v_1(resp. w_1) is adjacent with v_2(resp. w_2), and the second copy of v_1(resp. w_1) is adjacent with v_m(resp. w_m). These last vertex copies in the ladder L_{m+1} are named as v_{m+1} and w_{m+1}, respectively. Therefore, if f a colouring for the prism $C_m \times P_2$, then in the ladder L_{m+1} we have $f(v_1) = f(v_{m+1}$ as well as $f(w_1) = f(w_{m+1})$. In this case, we may also call $C_m \times P_2$ as the *closed* graph of L_{m+1} about the edges v_1w_1 and v_mw_m.

In the following lemma we will show that a ladder of $2m$ vertices, with $m \not\equiv 0 \pmod 4$, can not be gracefully coloured using 4 colours.

Lemma 3. *Using four different colours, the graph $C_m \times P_2$, with $m \geq 3$, $m \not\equiv 0 \pmod 4$, can not be gracefully coloured.*

P r o o f. Let a, b, c and d be four different colours, and let $m = 4k + r$, $1 \leq r \leq 3$. Consider the ladder L_{m+1} as the opened graph of $C_m \times P_2$. Let the vertex and edge sets of the ladder L_{m+1} be $\{v_i, w_i : 1 \leq i \leq m+1\}$ and $\{v_iv_{i+1}, w_iw_{i+1} : 1 \leq i \leq m, v_iw_i : 1 \leq i \leq m+1\}$, respectively. Observe that the colour of v_j (resp. w_j) must be the same with the colour of w_{j+2} (resp. v_{j+2}) or of w_{j-2} (resp. v_{j-2}) for realizable integer j (realizable means in the range of discussion). Without loss of generality, let the colour of v_1 is a. Therefore, the colour of w_{4s+3} and of v_{4t+1} is a, for some realizable non-negative integers s, t. Now let us see cases: $r = 1$, $r = 2$, and $r = 3$. Suppose that f is a graceful colouring for $C_m \times P_2$.

Case $r = 1$. If we take $t = k$, then we have $f(v_1) = a = f(v_{4k+1}) = f(v_m)$. Note that $v_{m+1} = v_{4k+2}$ is adjacent with v_m. Thus, $f(v_{m+1})$ can not be a to maintain proper colouring property. But, in $C_m \times P_2$, vertices v_1 and v_{m+1} are identical which insist $f(v_{m+1}) = f(v_1) = a$. This implies a contradiction. So, for $r = 1$ the graph $C_m \times P_2$ can not be gracefully coloured.

Case $r = 2$. Applying a similar argument, by assuming the colour of v_1 is a, we have that $f(w_{m+1}) = f(w_{4k+3}) = f(v_1) = a$. In graph $C_m \times P_2$, vertices w_1 and w_{m+1} are identical. On the other side, w_1 is adjacent with v_1, so that they can not get the same colour. Thus, a contradiction occurs.

Case $r = 3$. Again by using a similar reason, we have that $f(w_m) = f(w_{4k+3}) = f(v_1) = a$. We know that w_{m+1} in $C_m \times P_2$ is identified with w_1, and therefore is adjacent with both w_m and v_1. This implies that the induced edge colours of $v_1w_1 (= v_1w_{m+1})$ and w_1w_m are the same which then contradicts the gracefulness property.

In any case we have proven that $C_m \times P_2$, $m \not\equiv 0 \pmod 4$, can not be gracefully coloured using only 4 colours. □

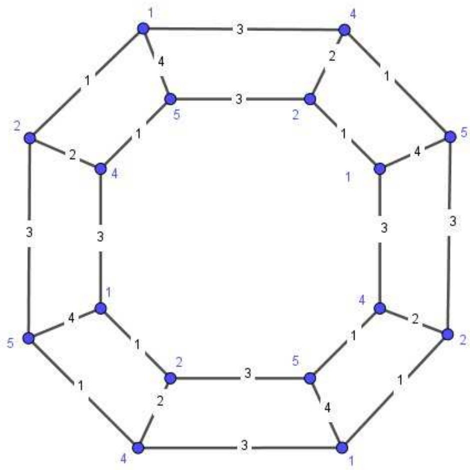


Figure 1. A graceful colouring of $C_8 \times P_2$.

2. Results on prism and generalized prism graphs

In this section, we will be dealing with the graceful chromatic number of prism $C_m \times P_2$ first, $m \geq 3$, and then with the graceful chromatic number of generalized prism graphs $C_m \times P_n$, $m, n \geq 3$. As for some consequences, we also derive some bounds for graceful chromatic number of graph $C_m \times C_n$, $m, n \geq 3$, for some specific values of m and n.

Our main discussion will be separated into two subsections: For $C_m \times P_2, m \geq 3$ and for $C_m \times P_n$, with $m, n \geq 3$.

2.1. Prism graph $C_m \times P_2$ for $m \geq 3$.

Theorem 1. *If* $m \equiv 0 \pmod 4$*, then the graceful chromatic number of graph* $C_m \times P_2$ *is equal to* 5.

P r o o f. Note that the graph $C_m \times P_2$ contains subgraph C_4. Based on Lemma 2, we may conclude that $\chi_g(C_m \times P_2) \geq 4$. Since all vertices of $C_m \times P_2$ has degree 3, if the colour 3 is used, then by Corollay 1, the colour greater than 6 should occur. Therefore, the four colours we will use are $1, 2, 4$, and 5. Now we will prove that using these four colours, we are able to colour $C_m \times P_2$ gracefully. To confirm this, we will do by introducing the following graceful colouring technique for $C_m \times P_2$ using only labels $1, 2, 4$, and 5.

Let the vertices of $C_m \times P_2$ is the set

$$\{v_{1+i}, v_{2+i}, v_{3+i}, v_{4+i}, w_{1+i}, w_{2+i}, w_{3+i}, w_{4+i} : \ i = 4k, \ k = 0, 1, 2, \ldots, m/4 - 1\}$$

and its edge set is

$$\{v_1 v_m, w_1 w_m, v_m w_m, v_i v_{i+1}, w_i w_{i+1}, v_i w_i : \ i = 1, 2, \ldots, m-1\}.$$

Define a colouring f for $C_m \times P_2$ as follows.

$$f(v_i) = \begin{cases} 1, & \text{if} \quad i \equiv 1 \pmod 4, \\ 4, & \text{if} \quad i \equiv 2 \pmod 4, \\ 5, & \text{if} \quad i \equiv 3 \pmod 4, \\ 2, & \text{if} \quad i \equiv 0 \pmod 4, \end{cases} \qquad f(w_i) = \begin{cases} 5, & \text{if} \quad i \equiv 1 \pmod 4, \\ 2, & \text{if} \quad i \equiv 2 \pmod 4, \\ 1, & \text{if} \quad i \equiv 3 \pmod 4, \\ 4, & \text{if} \quad i \equiv 0 \pmod 4. \end{cases} \tag{2.1}$$

Based on the above function f, it is clear that for every adjacent vertices u and v we have $f(u) \neq f(v)$. We can immediately observe that for any adjacent edges uw and wv in C_m we have

$$\{|f(u) - f(w)|,\ |f(w) - f(v)|\} = \{1, 3\}.$$

Furthermore, we also have

$$\{|f(v_i) - f(w_i)| :\ 1 \leq i \leq m\} = \{2, 4\}.$$

Remember that each vertex u in $C_m \times P_2$ has degree 3; say x_1, x_2, and x_3 are the vertices adjacent to u. From the function f we can immediately conclude that the set

$$\{|f(u) - f(x_1)|,\ |f(u) - f(x_2)|,\ |f(u) - f(x_3)|\}$$

is equal to $\{1, 2, 3\}$ or to $\{1, 3, 4\}$. Thus, the function f satisfies the property to become graceful colouring for $C_m \times P_2$. Therefore, $\chi_g(C_m \times P_2) = 5$. □

Theorem 2. *If* $m \not\equiv 0 \pmod 4$*, then the graceful chromatic number of graph* $C_m \times P_2$ *is equal to* 6.

P r o o f. The proof of Theorem 2 will make use of the result described in the proof of Theorem 1.

For some positive integer $k \geq 1$, consider $C_{4k} \times P_2$ which is coloured as in (2.1). Let the ladder L_{4k+1} be the open graph of $C_{4k} \times P_2$ about v_1w_1. Since $C_m \times P_2$ contains subgraph C_4, to colour it gracefully, one needs at least 4 colours. But, when $m \equiv 1, 2$ or $3 \pmod 4$, based on Lemma 3, we can not colour the graph $C_{4k} \times P_2$ gracefully using only 4 colours. Therefore, we have to use at least 5 colours. The smallest five colours are $1, 2, 3, 4$, and 5. But, based on Corollary 1, whenever we apply 3 for a vertex colour, the colour 6 or greater colour must occur. Thus, the graceful chromatic number of $C_m \times P_2$ is at least 6. To conclude that $\chi_g(C_m \times P_2) = 6$, we will proceed by showing that a graceful colouring exist with maximum colour 6, as follows.

Case 1: $m \equiv 1 \pmod 4$. First, consider $C_5 \times P_2$ with vertex set $\{a_1, a_2, a_3, a_4, a_5, b_1, b_2, b_3, b_4, b_5\}$ and with edge set $\{a_1a_5, b_1b_5, a_ia_{i+1}, b_ib_{i+1} : i = 1 \leq i \leq 4\} \cup \{a_ib_i : 1 \leq i \leq 5\}$. Now, we colour vertices using the following function f:

$$f(a_i) = \begin{cases} 1, & \text{if} \quad i = 1, \\ 4, & \text{if} \quad i = 2, \\ 3, & \text{if} \quad i = 3, \\ 5, & \text{if} \quad i = 4, \\ 2, & \text{if} \quad i = 5, \end{cases} \qquad f(b_i) = \begin{cases} 5, & \text{if} \quad i = 1, \\ 2, & \text{if} \quad i = 2, \\ 6, & \text{if} \quad i = 3, \\ 1, & \text{if} \quad i = 4, \\ 4, & \text{if} \quad i = 5. \end{cases}$$

The coloured $C_5 \times P_2$ will be used as the seed of our general construction for Case 1, and its diagram is depicted in Fig. 2.

Consider the opened ladder L_6 from the coloured $C_5 \times P_2$ above about a_1b_1. In L_6, the colours of a_1, a_2, a_3, a_4, a_5, and a_6 are $1, 2, 5, 3, 4$, and 1, while the colours of b_1, b_2, b_3, b_4, b_5, and b_6 are $5, 4, 1, 6, 2$, and 5.

Then, consider the open ladder L_{4k+1}, for some positive integer $k \geq 1$, from the coloured $C_{4k} \times P_2$ in Theorem 1 about v_1w_1. Here, the colours of v_1 and w_1 are also 1 and 5, respectively. The same colours are also for v_{4k+1} which is 1, and for w_{4k+1} which is 5. Based on (2.1), we have $f(v_{4k}) = 2$, and $f(w_{4k}) = 4$. By identifying v_{4k+1} with a_6 and w_{4k+1} with b_6, and maintaining the

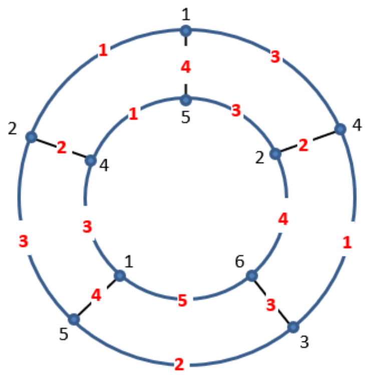


Figure 2. A graceful colouring of $C_5 \times P_2$.

other vertex colours, then we get a new ladder on $4(k+1)+2$ vertices, $L_{4(k+1)+2}$, with graceful colouring.

Furthermore, we know that $f(v_2) = 4$, $f(w_2) = 2$, $f(a_2) = 2$, and $f(b_2) = 4$. Thus by identifying v_1 with a_1 and w_1 with b_1 in the ladder $L_{4(k+1)+2}$, we obtain $C_{4(k+1)+1} \times P_2$ with a graceful colouring.

From here, we may infer that the graceful chromatic number of the graph $C_m \times P_2$, for $m \equiv 1 \pmod 4$ is equal to 6.

Case 2: $m \equiv 2 \pmod 4$. First, consider $C_6 \times P_2$ with vertex set

$$\{a_1, a_2, a_3, a_4, a_5, a_6, b_1, b_2, b_3, b_4, b_5, b_6\},$$

and with edge set

$$\big\{a_1a_6, b_1b_6, a_ia_{i+1}, b_ib_{i+1} : i = 1 \le i \le 5, \quad a_ib_i : 1 \le i \le 6\big\}.$$

As a seed graph, we define the following colouring for $C_6 \times P_2$ as follows.

$$f(a_i) = \begin{cases} 1, & \text{if} \quad i = 1, \\ 3, & \text{if} \quad i = 2, \\ 4, & \text{if} \quad i = 3, \\ 1, & \text{if} \quad i = 4, \\ 3, & \text{if} \quad i = 5, \\ 4, & \text{if} \quad i = 6, \end{cases} \qquad f(b_i) = \begin{cases} 5, & \text{if} \quad i = 1, \\ 6, & \text{if} \quad i = 2, \\ 2, & \text{if} \quad i = 3, \\ 5, & \text{if} \quad i = 4, \\ 6, & \text{if} \quad i = 5, \\ 2, & \text{if} \quad i = 6. \end{cases}$$

By inspection we can verify that the above colouring for $C_6 \times P_2$ is graceful. The diagram of the coloured graph is shown in Fig. 3.

Let the ladder of 7 vertices, L_7, is the open graph from the $C_6 \times P_2$ above about v_1w_1. We emphasize here that in this ladder L_7, vertices a_7 and b_7 have colours 1 and 5, respectively; the same as the colours of a_1 and b_1, respectively.

We use again the same ladder L_{4k+1}, $k \ge 1$, as in Case 1. Now we identify v_{4k+1} with a_7 and w_{4k+1} with b_7, and maintaining the other vertex colours. Then we get a new ladder on $4(k+1)+3$ vertices, $L_{4(k+1)+3}$, with graceful colouring.

Furthermore, we identify v_1 with a_1 and w_1 with b_1 in the ladder $L_{4(k+1)+3}$. Based on the previous colours, we know that the colours of $v_2, w_2, a_2, b_2, v_1 = a_1, w_1 = b_1$, are $4, 2, 3, 6, 1, 5$, respectively. This means that after the last identification, the gracefulness colouring of $C_{4(k+1)+2}$ are maintained. Thus, we may conclude that $C_{4(k+1)+2} \times P_2$ is with graceful colouring.

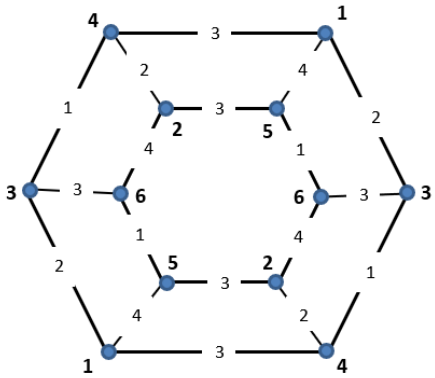


Figure 3. A graceful colouring of $C_6 \times P_2$.

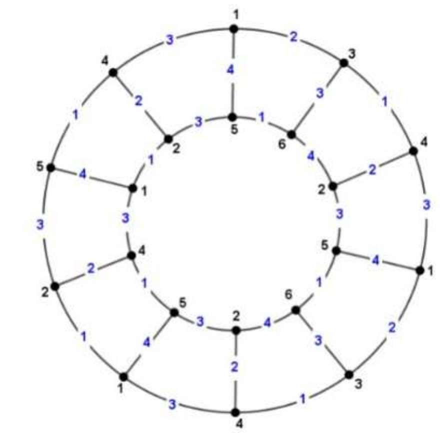


Figure 4. A graceful colouring of $C_{10} \times P_2$.

A graceful labeled $C_{10} \times P_2$ which is constructed using this method is depicted in Fig. 4.

From here, we may infer that the graceful chromatic number of the graph $C_m \times P_2$, for $m \equiv 2 \pmod 4$ is equal to 6.

Case 3: $m \equiv 3 \pmod 4$. Here we will introduce a construction for graceful colouring of $C_m \times P_2$ with $m \equiv 3 \pmod 4$. We start with $C_3 \times P_2$ with vertex set $\{a_1, a_2, a_3, b_1, b_2, b_3\}$ and edge set $\{a_3a_1, a_1a_2, a_2a_3, b_3b_1, b_1b_2, b_2b_3, a_1b_1, a_2b_2, a_3b_3\}$. Then we colour $C_3 \times P_2$ using the following colouring f.

$$f(a_i) = \begin{cases} 1, & \text{if} \quad i = 1, \\ 3, & \text{if} \quad i = 2, \\ 4, & \text{if} \quad i = 3, \end{cases} \qquad f(b_i) = \begin{cases} 5, & \text{if} \quad i = 1, \\ 6, & \text{if} \quad i = 2, \\ 2, & \text{if} \quad i = 3. \end{cases}$$

We can immediately check that this colouring f is graceful. The diagram of the gracefully coloured graph $C_3 \times P_2$ is shown in Fig. 5. We can verify that the graceful chromatic number of this graph is 6.

We should mention again that this above colouring of $C_3 \times P_2$ is graceful. As we did for Case 1 and Case 2, first we will observe the open ladder L_4 from $C_3 \times P_2$ about a_1b_1. In this L_4, the

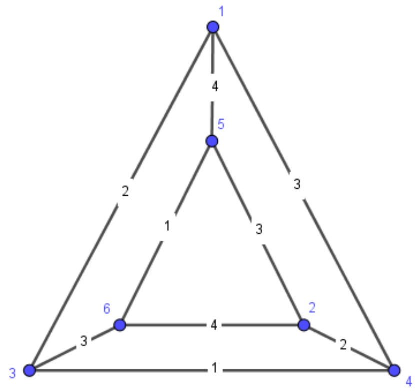


Figure 5. A graceful colouring of $C_3 \times P_2$.

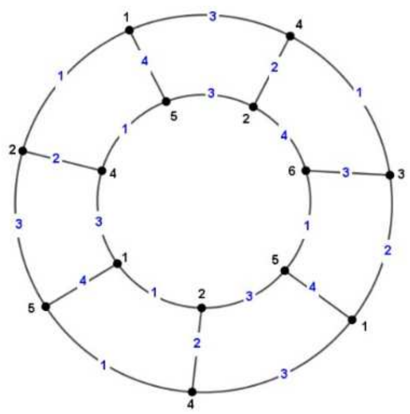


Figure 6. A graceful colouring of $C_7 \times P_2$.

colour of vertices $a_4 = a_1 = 1$ and $b_4 = b_1 = 5$. Observe back the open ladder L_{4k+1} in Case 1 (and Case 2).

Now we identify v_{4k+1} with a_4 and w_{4k+1} with b_4 to obtain a graceful colouring ladder L_{4k+4}. Let us denote the colouring as α. We can easily see that in this ladder we have $\alpha(a_1) = \alpha(v_1) = 1$ and $\alpha(b_1) = \alpha(w_1) = 5$. Moreover, we have also $\alpha(a_2) = f(a_2) = 3$, $\alpha(b_2) = f(b_2) = 6$, $\alpha(v_2) = 4$, and $\alpha(w_2) = 2$. Thus, by identifying v_1 with a_1 and w_1 with b_1, we get a graceful colouring $C_{4k+3} \times P_2$, with graceful chromatic number is 6. See the labeled graph $C_7 \times P_2$ in Fig. 6 as an example of the graph resulted from the construction.

Therefore, we may conclude that the graceful chromatic number of the graph $C_m \times P_2$, with $m \equiv 3 \pmod 4$ is also 6.

Since in all cases of m we proved that $C_m \times P_2$ has graceful chromatic number 6, we may conclude that $\chi_g(C_m \times P_2) = 6$. $\square$

2.2. Results on generalized prism graphs $C_m \times P_n$, $m, n \geq 3$.

For a graph G, let f be a graceful colouring for G. It is obvious that for a vertex $u \in V(G)$, if $v, w \in N(u)$, then $f(v) \neq f(w)$. Therefore, we can immediately observe that the graph $P_3 \times P_3$ can not be coloured by only four different colours. This observation gives

$$\chi_g(P_3 \times P_3) \geq 5.$$

But, if we use only five colours $1, 2, 3, 4$ and 5, the center vertex of $P_3 \times P_3$ must be 1 or 5. Then, by inspection we can show that using only five colours, we can not colour $P_3 \times P_3$ gracefully. This gives the following lemma.

Lemma 4. *The graceful chromatic number of the graph $P_3 \times P_3$, $\chi_g(P_3 \times P_3) \geq 6$.*

The following Lemma 5 will be an important tool for the proofs of our main results encountered in this section.

Lemma 5. *The graceful chromatic number of the graph $P_5 \times P_5$, $\chi_g(P_5 \times P_5) \geq 7$.*

P r o o f. Let the vertices of $P_5 \times P_5$ be $V(P_5 \times P_5) = \{v_{ij} : i, j = 0, 1, 2, 3, 4\}$ and $E(P_5 \times P_5) = \{v_{ij}v_{i(j+1)}, v_{ij}v_{(i+1)j} : i, j = 0, 1, 2, 3\}$. Now, observe the subgraph $P_3 \times P_3$ with $V(P_3 \times P_3) = \{v_{ij} : i, j = 1, 2, 3\}$ and

$$E(P_3 \times P_3) = \{v_{ij}v_{(i+1)j}, v_{ij}v_{(i)(j+1)} : i, j = 1, 2\}.$$

In $P_5 \times P_5$, every vertex of the subgraph $P_3 \times P_3$ has degree 4. Based on Lemma 4, for gracefully colouring $P_3 \times P_3$, we need at least five colours. If the colour 3 or 4 is assigned for a vertex of $P_3 \times P_3$, then based on Lemma 1 the colour greater than or equal to $4 + 3 = 7$ must appear in $P_5 \times P_5$. If the colors 3 and 4 both are not assigned for any vertex of $P_3 \times P_3$, then, since we need at least five colours, we need some color greater than or equal to 7 for gracefully colouring $P_5 \times P_5$. $\square$

Now, observe the graph $P_4 \times P_3$. We will make use of this observation for facilitating the result which will be formulated in Lemma 6. Let $V(P_4 \times P_3) = \{v_{ij} : i = 0, 1, 2, 3; j = 0, 1, 2\}$, and $E(P_4 \times P_3) = \{v_{ij}v_{i(j+1)} : i = 0, 1, 2, 3; j = 0, 1\} \cup \{v_{ij}v_{(i+1)j} : i = 0, 1, 2; j = 0, 1, 2\}$. The picture in Fig. 7 is the diagram of graph $P_4 \times P_3$ with vertex names.

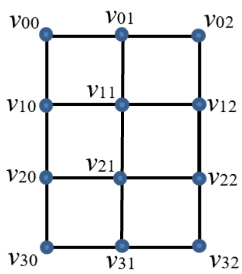


Figure 7. The graph $P_4 \times P_3$ with vertex names.

In here, we will restrict a vertex colouring α for $P_4 \times P_3$ as $\alpha(v_{0j}) = \alpha(v_{3j}), \forall j = 0, 1, 2$. We will show that under this restriction, using only six colours, the vertex colouring α can not be graceful.

Let the six colours be $1, 2, 3, 4, 5$ and 6. Based on Lemma 1, since the degree of vertices v_{11} and v_{21} each is four, the colours 3 and 4 both can not be used for these two vertices. So, there are four colours: $1, 2, 5$, and 6 that can be assigned for the vertices v_{11} and v_{21}. In total, there are six different combinations for colouring these two vertices: $\{\alpha(v_{11}), \alpha(v_{21})\} = \{a, b\}$, $a, b \in \{1, 2, 5, 6\}$, with $a \neq b$. We can check by inspection that any one of these combinations results in the colouring α is not graceful. But, for the space consideration, we will only describe the detail process for combination $\{\alpha(v_{11}), \alpha(v_{21})\} = \{1, 2\}$ as in Fig. 8. Note that the case $\alpha(v_{11}) = a$ and $\alpha(v_{21}) = b$ is similar to the case $\alpha(v_{11}) = b$ and $\alpha(v_{21}) = a$.

The explanation of the colouring process in Fig. 8 is the following:

1) The colours $\alpha(v_{11}) = 1$ and $\alpha(v_{21}) = 2$ are fixed as the initial combination.
2) The next vertex colouring follows the following vertices order: $v_{20}, v_{10}, v_{00}, v_{01}, v_{02}, v_{12}, v_{22}$. Note that $\alpha(v_{3j}) := \alpha(v_{0j}), \forall j = 0, 1, 2$, based on the restriction imposed for α.
3) For some colours x, y and z, a notation $x/\mathbf{y}/z$ means that we assign the colour y(indicated with bold face) for the related vertex among the possible colours x, y and z.
4) The colour which stands alone (written in red bold face), indicates that the colour is the only possible colour for the related vertex.
5) The red cross sign **X** informs that the colouring process is discontinue at the related vertex, since there is no possible choice of colours to colour the vertex. The appearance of **X** indicates that the colouring fails to be graceful.

From Fig. 8 we can see that each colouring process ends to be not graceful which is indicated by the appearance of the sign **X**. Thus, we may conclude that under the restriction $\alpha(v_{1j}) = \alpha(v_{4j}), j = 0, 1, 2$, using exactly six different colours, we can not colour the graph $P_4 \times P_3$ gracefully.

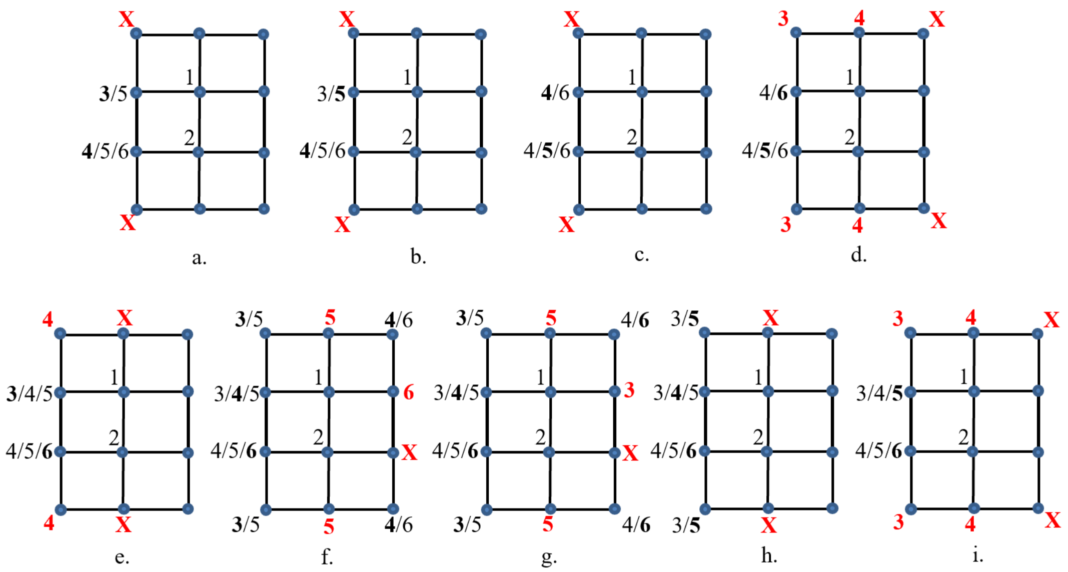


Figure 8. The colouring process for $P_4 \times P_3$ with $\alpha(v_{11}) = 1$ and $\alpha(v_{21}) = 2$.

If we extend this last observation to graph $P_4 \times P_n$, $n \geq 3$, with

$$V(P_4 \times P_n) = \{v_{ij} : i = 0, 1, 2, 3; j = 0, 1, \ldots, n-1\},$$

and

$$E(P_m \times P_n) = \{v_{ij}v_{i(j+1)} : i = 0, 1, 2, 3; j = 0, 1, \ldots, n-2\} \cup \{v_{ij}v_{(i+1)j} : i = 0, 1, 2; j = 0, 1, \ldots, n-1\},$$

under restriction that $\alpha(v_{0j}) = \alpha(v_{3j}), j = 0, 1, \ldots, n-1$, we may also conclude that we need at least seven colours to maintain the colouring α becomes graceful for $P_4 \times P_n$.
From this last observation we can formulate the following result.

Lemma 6. *For $n \geq 3$, the graceful chromatic number of the graph $C_3 \times P_n$, $\chi_g(C_3 \times P_n) \geq 7$.*

P r o o f. The generalized prism graph $C_3 \times P_n, n \geq 3$, can be obtained by identifying vertices v_{0j} and v_{3j} for every $j = 0, 1, 2, \ldots, n-1$ as it is in the last observation. By considering a graceful colouring α for the graph $P_4 \times P_n$ under the above mentioned restriction, we are done. $\square$

For facilitating the discussion of our main results in this section, we need the following definition, as we defined a ladder as an open graph of $C_m \times P_2$ in the previous section. Here we will define a similarone as an open graph from the graph $C_m \times P_n$, $m, n \geq 3$. Let the vertex set of graph $C_m \times P_n$, $m, n \geq 3$, be

$$\{v_{ij}, 0 \leq i \leq m-1,\ 0 \leq j \leq n-1\},$$

and its edge set be

$$\{v_{ij}v_{kl},\ \text{if } i = k \quad \text{and} \quad |j-l| = 1 \quad \text{or} \quad j = l \quad \text{and} \quad |i-k| \equiv 1 \pmod{m}\}.$$

Consider the open graph of $C_m \times P_n$, $m, n \geq 3$, about the path P which has end vertices v_{00} and v_{0n}, and has vertex set and edge set $\{v_{0j}, j = 0, 1, \ldots, n-1\}$ and $\{v_{0j}v_{0(j+1)}, j = 0, 1, \ldots, n-2\}$, respectively. Denote this open graph by $\mathcal{L}_{m+1,n}$. This graph is a grid graph having $(m+1) \times n$ vertices which involves two copies of path P. These two copies of path P, each has vertices $v_{0j}, j = 0, 1, \ldots, n-1$ and edges $v_{0j}v_{0(j+1)}, j = 0, 1, \ldots, n-2$. In the open graph $\mathcal{L}_{m+1,n}$, the vertices and edges of the second copy of P will be denoted by $v_{mj}, j = 0, 1, \ldots, n-1$, and $v_{mj}v_{(m)(j+1)}$, $j = 0, 1, \ldots, n-1$, respectively. It is clear that the vertex v_{mj} is adjacent with $v_{(m-1)j}$ for every $j = 0, 1, \ldots, n-2$. In this case, $C_m \times P_n$ can be reconstructed from $\mathcal{L}_{m+1,n}$ by identifying vertex v_{0j} and v_{mj} for every $j = 0, 1, \ldots, n-1$.

Theorem 3. *For any positive integers $m, n \geq 3$, with $m \equiv 0 \pmod 3$, $\chi_g(C_m \times P_n) = 7$.*

P r o o f. From Lemma 4 we know that the graceful chromatic number of $C_m \times P_n$ is at least seven. Now we will show that a graceful colouring exists for $C_m \times P_n$ such that it uses only seven different colours, and therefore $\chi_g(C_m \times P_n) = 7$.

Let the vertex set of $C_m \times P_n$ is $\{v_{ij} | 0 \leq i \leq m-1;\ 0 \leq j \leq n-1\}$, and edge set

$$\big\{v_{ij}v_{rs} | i = r \quad \text{and} \quad |s-j| \equiv 1 \pmod{n} \quad \text{or} \quad j = s \quad \text{and} \quad |i-r| \equiv 1 \pmod{m}\big\}.$$

To this end, here we define a colouring function f for $C_m \times P_n$ as follows.

$$f(v_{ij}) = \begin{cases} 1, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 0 \pmod 6, \\ 5, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 1 \pmod 6, \\ 6, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 2 \pmod 6, \\ 2, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 3 \pmod 6, \\ 3, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 4 \pmod 6, \\ 7, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 5 \pmod 6, \\ 3, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 0 \pmod 6, \\ 7, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 1 \pmod 6, \\ 1, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 2 \pmod 6, \\ 5, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 3 \pmod 6, \\ 6, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 4 \pmod 6, \\ 2, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 5 \pmod 6, \\ 6, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 0 \pmod 6, \\ 2, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 1 \pmod 6, \\ 3, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 2 \pmod 6, \\ 7, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 3 \pmod 6, \\ 1, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 4 \pmod 6, \\ 5, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 5 \pmod 6. \end{cases} \tag{2.2}$$

An example of a graceful coloured graph $C_6 \times P_n$ using (2.2) is shown in Fig. 9. In this figure we may also see the related open graph $\mathcal{L}_{7,n}$ of $C_6 \times P_n$.

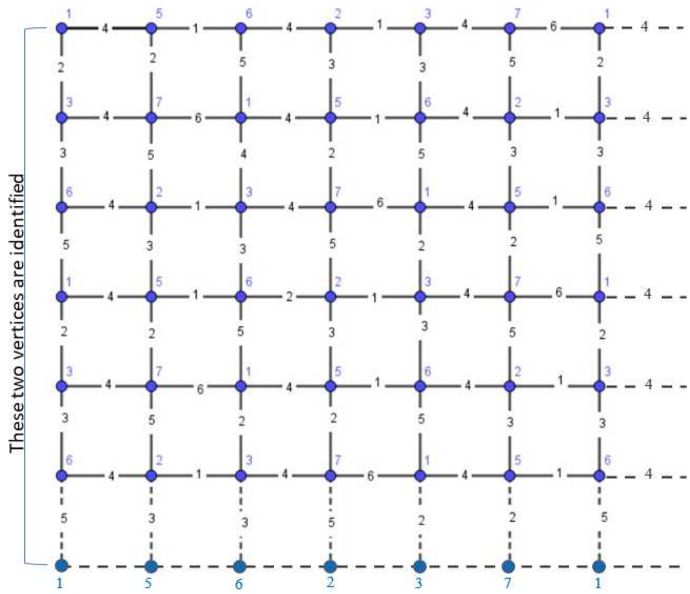


Figure 9. A graceful colouring of $C_6 \times P_n$, $n \geq 3$.

Fig. 9 also helps us to be able to check by inspection that f is a graceful colouring for the graph $C_m \times P_n$, with $m \equiv 0 \pmod 3$. Therefore, we may conclude that this graph has chromatic number 7. □

Furthermore, based on (2.2) we see that for every i, $0 \leq i \leq m-1$, we have $f(v_{ij}) = f(v_{ik})$ provided $|j-k| \equiv 0 \pmod 6$.

Corollary 2. *For any positive integers* $m, n \geq 3$, *with* $m \equiv 0 \pmod 3$ *and with* $n \equiv 0 \pmod 6$, $\chi_g(C_m \times C_n) = 7$.

P r o o f. The proof of this corollary may be derived from (2.2). From Theorem 3 we conclude that $\chi_g(C_m \times P_n) = 7$, if $m \equiv 0 \pmod 3$, and $n \geq 3$. From (2.2) we know that $f(v_{ij}) = f(v_{ik})$ whenever $|j-k| \equiv 0 \pmod 6$. Thus, if $n \equiv 0 \pmod 6$, then if we identify vertex v_{i0} and v_{in} for every $i, 0 \leq i \leq m-1$ in $C_m \times P_n$, then we get a graceful coloured graph $C_m \times C_n$, $m \equiv 0 \pmod 3$ and $n \equiv 0 \pmod 6$. Therefore, we may conclude that $\chi_g(C_m \times C_n) = 7$ where $m \equiv 0 \pmod 3$ and $n \equiv 0 \pmod 6$. □

In the remaining part of this section we will see the graceful colouring number for $C_m \times P_n$, with $m \not\equiv 0 \pmod 3$, $n \geq 3$. We start to observe the case $m \equiv 1 \pmod 3$ as we formulate in the following theorem.

Theorem 4. *If* $m \equiv 1 \pmod 3$, *then* $7 \leq \chi_g(C_m \times P_n) \leq 8$.

P r o o f. We will make use of prism graph $C_4 \times P_n$ as the seed of our graceful colouring construction. We first introduce a colouring for the graph $C_4 \times P_n$, $n \geq 3$.

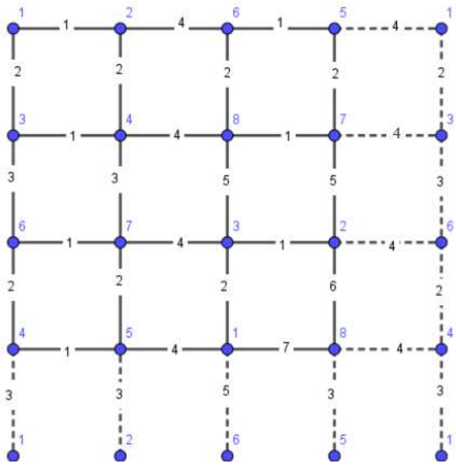


Figure 10. A graceful colouring of $C_4 \times P_n$.

Let the vertex set of $C_4 \times P_n$ is

$$\{v_{ij} | 0 \leq i \leq 3;\ 0 \leq j \leq n-1\},$$

and edge set

$$\{v_{ij}v_{rs} |\ i = r \quad \text{and} \quad |s-j| \equiv 1 \pmod{n} \quad \text{or} \quad j = s \quad \text{and} \quad |i-r| \equiv 1 \pmod{4}\}.$$

To this end, we define a colouring function f as follows.

$$f(v_{ij}) = \begin{cases} 1, & \text{if} \quad i \equiv 0 \pmod{4}, \quad j \equiv 0 \pmod{4}, \\ 2, & \text{if} \quad i \equiv 0 \pmod{4}, \quad j \equiv 1 \pmod{4}, \\ 6, & \text{if} \quad i \equiv 0 \pmod{4}, \quad j \equiv 2 \pmod{4}, \\ 5, & \text{if} \quad i \equiv 0 \pmod{4}, \quad j \equiv 3 \pmod{4}, \\ 3, & \text{if} \quad i \equiv 1 \pmod{4}, \quad j \equiv 0 \pmod{4}, \\ 4, & \text{if} \quad i \equiv 1 \pmod{4}, \quad j \equiv 1 \pmod{4}, \\ 8, & \text{if} \quad i \equiv 1 \pmod{4}, \quad j \equiv 2 \pmod{4}, \\ 7, & \text{if} \quad i \equiv 1 \pmod{4}, \quad j \equiv 3 \pmod{4}, \\ 6, & \text{if} \quad i \equiv 2 \pmod{4}, \quad j \equiv 0 \pmod{4}, \\ 7, & \text{if} \quad i \equiv 2 \pmod{4}, \quad j \equiv 1 \pmod{4}, \\ 3, & \text{if} \quad i \equiv 2 \pmod{4}, \quad j \equiv 2 \pmod{4}, \\ 2, & \text{if} \quad i \equiv 2 \pmod{4}, \quad j \equiv 3 \pmod{4}, \\ 4, & \text{if} \quad i \equiv 3 \pmod{4}, \quad j \equiv 0 \pmod{4}, \\ 5, & \text{if} \quad i \equiv 3 \pmod{4}, \quad j \equiv 1 \pmod{4}, \\ 1, & \text{if} \quad i \equiv 3 \pmod{4}, \quad j \equiv 2 \pmod{4}, \\ 8, & \text{if} \quad i \equiv 3 \pmod{4}, \quad j \equiv 3 \pmod{4}. \end{cases} \tag{2.3}$$

For an illustration one can see in Fig. 10

Fig. 10 helps us to see that (2.3) gives a graceful colouring for $C_4 \times P_n$ for every $n \geq 3$ with $\chi_g(C_4 \times P_n) \leq 8$. Therefore, based on Lemma 4, we may conclude that $7 \leq \chi_g(C_4 \times P_n) \leq 8$.

Furthermore, the graceful colouring of $C_m \times P_n$, with $m \equiv 1 \pmod{3}$ and $n \geq 3$ in general, is obtained by extending graceful coloured graph $C_4 \times P_n$ using the prism graph $C_3 \times P_n$ which has colouring as we will show below.

Let the vertex set of $C_3 \times P_n$ is

$$\{v_{ij} |\ 0 \leq i \leq 2;\ 0 \leq j \leq n-1\},$$

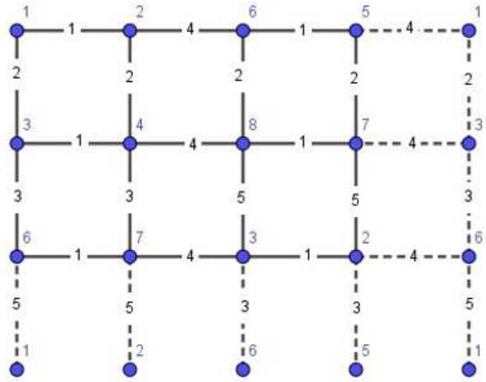


Figure 11. A graceful colouring of $C_3 \times P_n$.

and edge set

$$\{v_{ij}v_{rs}|\ i = r \quad \text{and} \quad |s - j| \equiv 1 \pmod n \quad \text{or} j = s \text{ and } |i - r| \equiv 1 \pmod 3\}.$$

To this end, we define a colouring function f as follows.

$$f(v_{ij}) = \begin{cases} 1, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 0 \pmod 4, \\ 2, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 1 \pmod 4, \\ 6, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 2 \pmod 4, \\ 5, & \text{if} \quad i \equiv 0 \pmod 3, \quad j \equiv 3 \pmod 4, \\ 3, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 0 \pmod 4, \\ 4, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 1 \pmod 4, \\ 8, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 2 \pmod 4, \\ 7, & \text{if} \quad i \equiv 1 \pmod 3, \quad j \equiv 3 \pmod 4, \\ 6, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 0 \pmod 4, \\ 7, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 1 \pmod 4, \\ 3, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 2 \pmod 4, \\ 2, & \text{if} \quad i \equiv 2 \pmod 3, \quad j \equiv 3 \pmod 4. \end{cases} \tag{2.4}$$

The diagram of coloured graph $\mathcal{L}_{4,n}$ from $C_3 \times P_n$ is depicted in Fig. 11. The coloured graph $C_3 \times P_n$ is obtained by identifying v_{0j} and v_{3j} for all $j, 0 \leq j \leq n-1$. We can immediately observe that (2.4) gives a graceful colouring for the prism graph $C_3 \times P_n$ with $\chi_g(C_3 \times P_n) \leq 8$. Again based on Lemma 4, we conclude that $7 \leq \chi_g(C_3 \times P_n) \leq 8$.

For producing a graceful colouring for $C_m \times P_n$, $m \equiv 1 \pmod 3$ we use $\mathcal{L}_{5,n}$ from $C_4 \times P_n$ and $\mathcal{L}_{4,n}$ from $C_3 \times P_n$, by identifying v_{5j} of $\mathcal{L}_{5,n}$ and v_{0j} of $\mathcal{L}_{4,n}$ for all j, $0 \leq j \leq n-1$. This identification results in a graceful coloured grid graph $\mathcal{L}_{8,n}$. Then, if we identify v_{8j} of $\mathcal{L}_{8,n}$ and v_{0j} of $\mathcal{L}_{4,n}$ for all $j, 0 \leq j \leq n-1$, we get a graceful coloured grid graph $\mathcal{L}_{11,n}$. Continuing the same procedure, then we get a graceful coloured grid graph $\mathcal{L}_{(m+1),n}$. Then by identifying vertex v_{0j} and v_{mj} from $\mathcal{L}_{(m+1),n}$ we obtain $C_m \times P_n$ with $m \equiv 1 \pmod 3$ and $n \geq 3$. □

As one consequence, as we formulated Corollary 2 based on Theorem 3, we also formulate a corollary based on Theorem 4 as the following.

Corollary 3. *If* $m \equiv 1 \pmod 3$ *and* $n \equiv 0 \pmod 4$, *then* $7 \leq \chi_g(C_m \times C_n) \leq 8$.

Now we go to the next case $m \equiv 2 \pmod 3$. The result is formulated in the following theorem.

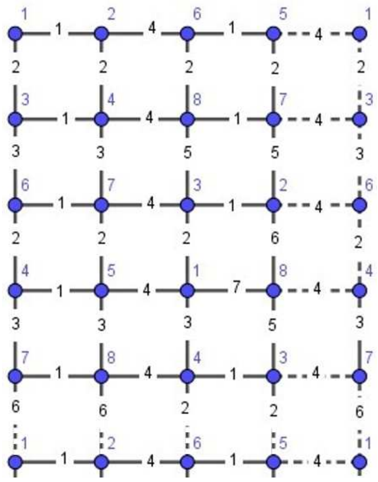


Figure 12. A graceful colouring of grid graph $\mathcal{L}_{6,n}$ of $C_5 \times P_n$.

Theorem 5. *If* $m \equiv 2 \pmod 3$, *then* $7 \leq \chi_g(C_m \times P_n) \leq 8$.

P r o o f. To proof this theorem, we will start with a graceful colouring for $C_5 \times P_n$, $m \equiv 2 \pmod 3$, $n \geq 3$. We introduce the following colouring for the graph $C_5 \times P_n$, $n \geq 3$.

$$f(v_{ij}) = \begin{cases} 1, & \text{if} \quad i \equiv 0 \pmod 5, \quad j \equiv 0 \pmod 4, \\ 2, & \text{if} \quad i \equiv 0 \pmod 5, \quad j \equiv 1 \pmod 4, \\ 6, & \text{if} \quad i \equiv 0 \pmod 5, \quad j \equiv 2 \pmod 4, \\ 5, & \text{if} \quad i \equiv 0 \pmod 5, \quad j \equiv 3 \pmod 4, \\ 3, & \text{if} \quad i \equiv 1 \pmod 5, \quad j \equiv 0 \pmod 4, \\ 4, & \text{if} \quad i \equiv 1 \pmod 5, \quad j \equiv 1 \pmod 4, \\ 8, & \text{if} \quad i \equiv 1 \pmod 5, \quad j \equiv 2 \pmod 4, \\ 7, & \text{if} \quad i \equiv 1 \pmod 5, \quad j \equiv 3 \pmod 4, \\ 6, & \text{if} \quad i \equiv 2 \pmod 5, \quad j \equiv 0 \pmod 4, \\ 7, & \text{if} \quad i \equiv 2 \pmod 5, \quad j \equiv 1 \pmod 4, \\ 3, & \text{if} \quad i \equiv 2 \pmod 5, \quad j \equiv 2 \pmod 4, \\ 2, & \text{if} \quad i \equiv 2 \pmod 5, \quad j \equiv 3 \pmod 4, \\ 4, & \text{if} \quad i \equiv 3 \pmod 5, \quad j \equiv 0 \pmod 4, \\ 5, & \text{if} \quad i \equiv 3 \pmod 5, \quad j \equiv 1 \pmod 4, \\ 1, & \text{if} \quad i \equiv 3 \pmod 5, \quad j \equiv 2 \pmod 4, \\ 8, & \text{if} \quad i \equiv 3 \pmod 5, \quad j \equiv 3 \pmod 4, \\ 7, & \text{if} \quad i \equiv 4 \pmod 5, \quad j \equiv 0 \pmod 4, \\ 8, & \text{if} \quad i \equiv 4 \pmod 5, \quad j \equiv 1 \pmod 4, \\ 4, & \text{if} \quad i \equiv 4 \pmod 5, \quad j \equiv 2 \pmod 4, \\ 3, & \text{if} \quad i \equiv 4 \pmod 5, \quad j \equiv 3 \pmod 4. \end{cases} \tag{2.5}$$

The diagram for coloured grid graph $\mathcal{L}_{6,n}$ of $C_5 \times P_n$, which is derived from (2.5), can be seen in Fig. 12. Using this diagram we may conclude that the colouring is graceful. It is clear that $\chi_g(C_5 \times P_n) \leq 8$.

The process of expanding to get coloured graph $C_m \times P_n$, $m \equiv 2 \pmod 3$, $n \geq 3$, is similar to the previous process as was described in the proof of Theorem 4. Here we use graceful coloured grid graph $\mathcal{L}_{6,n}$ from graceful coloured graph $C_5 \times P_n$, and graceful coloured grid graph $\mathcal{L}_{4,n}$ from graceful coloured graph $C_3 \times P_n$. Again by considering Lemma 4,we then conclude that

$7 \le \chi_g(C_5 \times P_n) \le 8.$ □

Similar to the previous corollaries, here we formulate the following corollary as a consequence of Theorem 5.

Corollary 4. *If* $m \equiv 2 \pmod 3$ *and* $n \equiv 0 \pmod 4$*, then* $7 \le \chi_g(C_m \times C_n) \le 8$.

3. Conclusion

In the discussion above, it has been proven that prism graph $C_m \times P_2$ has a chromatic number equal to 5 when $m \equiv 0 \pmod 4$, and equal to 6 when $m \not\equiv 0 \pmod 4$. While for generalized prism $C_m \times P_n$ we found that its chromatic number is equal to 7 while $m \equiv 0 \pmod 3$. Whereas for $m \not\equiv 0 \pmod 3$, we got that $7 \le \chi_g(C_m \times P_n) \le 8$. Based on these results, we could also derive some exact and bound values of graceful chromatics number of $C_m \times C_n$ for certain $m, n \ge 3$. Regarding this last observation, we propose the following open problem and conjecture.

Conjecture. If $m \not\equiv 0 \pmod 3$ and $n \ge 3$, then $\chi_g(C_m \times P_n) = 8$.

Open problem. What is $\chi_g(C_m \times C_n)$, if $m \not\equiv 0 \pmod 3$, $m, n \ge 3$?

Acknowledgements

The authors express their deep gratitude to referees for their invaluable corrections and suggestions. The first author thanks to the Research and Social Service Board of Universitas Pendidikan Ganesha for the provided support to conduct this research.

REFERENCES

1. Alfarisi R., Dafik, Prihandini R. M., Adawiyah R., Albirri E. R., Agustin I. H. Graceful Chromatic Number of Unicyclic Graphs. *J. Phys.: Conf. Ser.*, 2019. Vol. 1306: The 2nd Int. Conf. Math.: Education, Theory, and Application (30–31 October 2018, Sukoharjo, Indonesia). Art. no. 012039. DOI: 10.1088/1742-6596/1306/1/012039
2. Asyari M. L., Agustin I. H., Nisviasari R., Adawiyah R. On graceful chromatic number of some graphs. *J. Phys.: Conf. Ser.*, 2022. Vol. 2157: The 5th Int. Conf. Combinatorics, Graph Theory, and Network Topology, ICCGANT 2021. (21–22 August 2021, Jember, Indonesia). Art. no. 012013. DOI: 10.1088/1742-6596/2157/1/012013
3. Byers A. D. *Graceful Colorings and Connection in Graphs.* Diss. Doct. at Western Michigan University, 2018. Dissertations. Vol. 3308. URL: https://scholarworks.wmich.edu/dissertations/3308
4. English E., Zhang P., Kalamazoo. On graceful colourings of trees. *Math. Bohem.*, 2016. Vol. 142, No. 1. P. 57–73. URL: http://dml.cz/dmlcz/146009
5. Gallian J. A. A dynamic survey of graph labeling. *Electron. J. Combin.*, 2021. Vol. 24, #DS6. 623 p. URL: https://www.combinatorics.org/files/Surveys/ds6/ds6v25-2022.pdf
6. Mincu R., Obreja C., Popa A. The graceful chromatic number for some particular classes of graphs. In: *Proc. 21st Int. Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC).* IEEE Xplore, 2019. Art. no. 19492926. DOI: 10.1109/SYNASC49474.2019.00024

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