The 4th National Congress on Mathematics and Statistics

Gonbad Kavous University, 29 Feb 2024

Conference Papers (English)

Scientific Chairman

Dr. Esmaeil Shirazi

Executive Chairman

Dr. Abdolsaeed Toomaj

Dr. Mojtaba Kashani

Prepared By

Dr. Mehdi Shahini

Contents

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–112, pp. 1–16

A certain coupled system of q-FDEs on two consecutive intervals under Dirichlet conditions via Krasnoselskii's theorem

Mohammad Esmael Samei[a,*], Mohammad Izadi[b], Mohammed K. A. Kaabar[c]

[a] Department of Mathematics, Faculty of Basic Science, Bu-Ali Sina University, Hamedan 65178-38695, Iran
[b] Department of Applied Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran
[c] Institute of Mathematical Sciences, Faculty of Science, University of Malaya, Kuala Lumpur 50603, Malaysia

Article Info

Keywords:
nonlinear fractional equation
fractional $q-$differential equation
Dirichlet boundary conditions
Riemann–Liouville $q-$integral

2020 MSC:
34A08
39A12
39A13

Abstract

In this study, we use certain mathematical tools to analyze the solutions of a system of fractional q-differential equation ${}^{C}\mathbb{D}_q^{\sigma_i}[\wp](\mathfrak{t}) = \mathfrak{w}_i(\mathfrak{t}, \wp(\mathfrak{t}), {}^{C}\mathbb{D}_q^{i^{\nu_j}}[\wp](\mathfrak{t}), \mathbb{I}_q^{i^{\nu_j}}[\wp](\mathfrak{t}))$, $i = 1$ whenever $\mathfrak{t} \in [0, \mathfrak{t}_0]$, and $i = 2$ whenever $\mathfrak{t} \in [\mathfrak{t}_0, 1]$, for $j = 1, 2$, such as fixed point theorem of Krasnoselskii and Banach contraction principle, under simultaneous Dirichlet boundary conditions. Here, we use standard definitions of the Liouville-Caputo fractional type $q-$derivative and Riemann-Liouville $q-$integral. Some illustrative examples with numerical results are discussed, too.

1. Introduction: Problem's formulation

Mathematical subjects in the analysis of the problems of today's world are really welcomed by researchers. Among others, the study of new mathematical models have been a growing filed of study due to its importance and applications in diverse discipline of science and engineering. In this respect, using factional or non-integer derivatives provides more insight for the description of natural phenomena in the language of mathematical modeling [16, 17]. Many interesting real-life models with fractional derivatives have been proposed and analyzed mathematically. Among others, we can refer to the ppopulation models [11, 13], the blood ethanol system [18], the viscolastic models [9], the Layla and Mojnun's love story [14], the HBV, HIV and SEIR infection models [10, 20, 29], and the human liver model [25], to name a few. For definitions of fractional derivatives and integrals and some related special functions we refer to the recently published papers on the subject [27, 28]. In the meantime, fractional differential and q-differential equations (FDE, FqDE) are significant, see [1, 3, 5, 15, 21, 30, 31].

*Talker

Email addresses: mesamei@basu.ac.ir, mesamei@gmail.com (Mohammad Esmael Samei), izadi@uk.ac.ir (Mohammad Izadi), mohammed.kaabar@wsu.edu (Mohammed K. A. Kaabar)

The following FDE investigated by Ahmad *et al.* in 2014 as

$$\begin{cases} {}^cD_q^\beta({}^cD_q^\gamma+\lambda)u(t)=pf(t,u(t))+kI_q^\xi g(t,u(t)), & 0<\beta,\gamma\leq 1,\\ \alpha_1 u(0)-\beta_1(t^{(1-\gamma)}D_q u(0))\big|_{t=0}=\sigma_1 u(\eta_1), \quad \alpha_2 u(1)+\beta_2 D_q u(1)=\sigma_2 u(\eta_2), \end{cases}$$

where $t,q\in[0,1]$ and ${}^cD_q^\beta$ is the fractional Liouville-Caputo q-derivative. Moreover, the symbol $I_q^\xi(.)$ stands for the Riemann-Liouville integral for $\xi\in(0,1)$ and the functions f and g are two continuous functions. Finally, the parameters λ, p,k are real numbers. Similarly, we have $\alpha_\ell,\beta_\ell,\sigma_\ell\in\mathbb{R}$ and $\eta_\ell\in(0,1)$ for $\ell=1,2$ ([4]). Furthermore, Abdeljawad *et al.*, considered (with proof) a novel discrete $q-$fractional version of the well-known Grönwall inequality: $({}_qC_a^\alpha f)(t)=T\,(t,f(t))$ and $f(a)=\gamma$ in a way that $\alpha\in(0,1]$, $a\in\mathbb{T}_q=\{q^n:n\in\mathbb{Z}\}$, t belongs to $\mathbb{T}_a=[0,\infty)_q=\{q^{-\ell}a:\ell=0,1,2,\ldots\}$. Here, the notation ${}_qC_a^\alpha$ shows the Liouville-Caputo fractional difference of order α and a Lipschitz condition for the function $T(t,x)$ holds for all t and x ([1]). Later in 2017, Zhou *et al.* provided the existence criteria for the solutions of p-Laplacian Langevin FDE $D_{0+}^\beta\phi_p[(D_{0+}^\alpha+\lambda)x(t)]=f(t,x(t),D_{0+}^\alpha x(t))$, ${}_qD_{0+}^\beta\phi_p[(D_{0+}^\alpha+\lambda)x(t)]=g(t,x(t),{}_qD_{0+}^\alpha x(t))$ under anti-periodic boundary conditions $x(0)=-x(1)$, ${}_qD_{0+}^\alpha x(0)=-{}_qD_{0+}^\alpha x(1)$, in the whole domain $0\leq t\leq 1$. Here, $\phi_p(s)=|s|^{p-2}s$, with $p\in(1,2]$. Also, we have $0<\alpha,\beta\leq 1$, $0\leq\lambda$, $1<\alpha+\beta<2$, and $q\in(0,1)$ [31]. For more instance, see [7, 22–24]

In this work, some basic and fundamental results related to q-calculus are recalled in Sec. 2. Motivated by these achievements, in Sec. 3, we examine the positive solutions of FqDE in two consecutive segments

$${}^c\mathcal{D}_q^\sigma[k](t)=\begin{cases} f\left(t,k(t),{}^c\mathcal{D}_q^{\alpha_1}[k](t),\mathcal{I}_q^{\beta_1}[k](t)\right), & t\in[0,t_0),\\ g\left(t,k(t),{}^c\mathcal{D}_q^{\alpha_2}[k](t),\mathcal{I}_q^{\beta_2}[k](t)\right), & t\in[t_0,1], \end{cases} \tag{1}$$

under simultaneous Dirichlet boundary conditions

$$k(0)=h_1\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_3}[k](t_0),\mathcal{I}_q^{\beta_3}[k](t_0)\right), \qquad k(1)=h_2\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_4}[k](t_0),\mathcal{I}_q^{\beta_4}[k](t_0)\right), \tag{2}$$

where $\mathcal{I}_q^\beta$ and ${}^c\mathcal{D}_q^\sigma$ stand for the Riemann-Liouville $q-$integral and the Liouville-Caputo fractional $q-$derivative of order β and $1<\sigma\leq 2$ respectively, $t\in\overline{J}=[0,1]$, $t_0\in J=(0,1)$, $0<\alpha_\ell<1$ with $\ell=1,2,3,4$, $\beta_\ell>0$ with $\ell=1,2,3,4$, and the functions f,g,h_1 and h_2 map $J\times\mathbb{R}^3$ to $\mathbb{R}$ with $f(t_0,\cdot,\cdot,\cdot)=g(t_0,\cdot,\cdot,\cdot)$. Finally in Sec. 4 we consider two illustrated examples associated to the obtained results for the above model problems are provided in detail.

2. Essential preliminaries

Throughout the context, we shall apply the notations of time scales calculus [8]. Let us assume that $\mathrm{t}_0\in\mathbb{R}$ and $q\in\mathbb{I}$. Next, we define the time scale $\mathbb{T}_{\mathrm{t}_0}=\{0\}\cup\{\mathrm{t}:\mathrm{t}=\mathrm{t}_0q^n,\ \forall n\in\mathbb{N}\}$. However, for simplicity we sometimes drop the subscript t_0 and denote $\mathbb{T}_{\mathrm{t}_0}$ by $\mathbb{T}$ if there is no confusion about t_0. For a given $s\in\mathbb{R}$, let us define the symbol $[s]_q=(1-q^s)/(1-q)$ [15]. The next aim is to define the notation $(y-z)_q^{(n)}$ for the q-factorial function. It is given by $(y-z)_q^{(n)}=\prod_{k=0}^{n-1}(y-zq^k)$, $n\in\mathbb{N}_0:=\{0\}\cup\mathbb{N}$, $y,z\in\mathbb{R}$, and with $(y-z)_q^{(0)}=1$, see ([2]). One can also show that $(y-z)_q^{(\sigma)}=y^\sigma\prod_{k=0}^\infty\frac{y-zq^k}{y-zq^{\sigma+k}}$, $\sigma\in\mathbb{R}$, $s\neq 0$. It should be stressed that for $z=0$, we have obviously $y^{(\sigma)}=y^\sigma$. The next symbol is used for the q-Gamma function. It has the following definition $\Gamma_q(y)=(1-q)^{1-y}(1-q)_q^{(y-1)}$, $(y\in\mathbb{R}\backslash\{\cdots,-2,-1,0\})$ [15]. To proceed, let us σ and ν be two positive numbers. Let a function $y:\mathbb{T}\to\mathbb{R}$ is given. We define the q-derivative of y in the form $\mathbb{D}_q[y](\mathrm{t})=\left(\frac{\mathrm{d}}{\mathrm{dt}}\right)_q y(\mathrm{t})=\frac{y(q\mathrm{t})-y(\mathrm{t})}{\mathrm{t}(1-q)}$, for all $\mathrm{t}\in\mathbb{T}\backslash\{0\}$ and for $\mathrm{t}=0$ we have $\mathbb{D}_q[y](0)=\lim_{\mathrm{t}\to 0}\mathbb{D}_q[y](\mathrm{t})$ [2]. One can also define the higher-order q-derivative of y recursivly through the relation $\mathbb{D}_q^n[y](\mathrm{t})=\mathbb{D}_q[\mathbb{D}_q^{n-1}[y]](\mathrm{t})$ for all $n\geq 1$. Here, for $n=0$ we get $\mathbb{D}_q^0[y](\mathrm{t})=y(\mathrm{t})$ [2].

$$\mathbb{I}_q[y](\mathrm{t})=\int_0^{\mathrm{t}}y(\xi)\,\mathrm{d}_q\xi=\mathrm{t}(1-q)\sum_{k=0}^\infty q^k y(\mathrm{t}q^k), \tag{3}$$

for $0 \leq \mathfrak{t} \leq b$ and under condition that the involved series is absolutely convergent, see [2]. From this we can conclude the next identity for s in $[0, b]$ as

$$\int_s^b y(\xi)\, \mathrm{d}_q\xi = \mathbb{I}_q[y](b) - \mathbb{I}_q[y](s) = (1-q)\sum_{k=0}^{\infty} q^k \left[by(bq^k) - sy(sq^k)\right],$$

based upon the existence of the series. Suppose that $y \in C([0, b])$. For $n = 0$, the integral operator $\mathbb{I}_q^n$ is defined as $\mathbb{I}_q^0[y](\mathfrak{t}) = y(\mathfrak{t})$ and for for $n \geq 1$ we have $\mathbb{I}_q^n[y](\mathfrak{t}) = \mathbb{I}_q[\mathbb{I}_q^{n-1}[y]](\mathfrak{t})$, see [2]. If the function y be continuous at $\mathfrak{t} = 0$ one can assert that $\mathbb{D}_q[\mathbb{I}_q[y]](\mathfrak{t}) = y(\mathfrak{t})$ and $\mathbb{I}_q[\mathbb{D}_q[y]](\mathfrak{t}) = y(\mathfrak{t}) - y(0)$ [2]. For the function y, the next is the definition of fractional Riemann–Liouville type q-integral in the form

$$\mathbb{I}_q^{\sigma}[y](\mathfrak{t}) = \int_0^{\mathfrak{t}} (\mathfrak{t}-\xi)_q^{(\sigma-1)} \frac{y(\xi)}{\Gamma_q(\sigma)}\, \mathrm{d}_q\xi, \qquad \mathbb{I}_q^0[y](\mathfrak{t}) = y(\mathfrak{t}), \tag{4}$$

where $\sigma > 0$ and for all $\mathfrak{t} \in [0, 1]$ [6, 12]. Similarly for this function, the concept of Liouville-Caputo fractional q-derivative is given next as

$${}^{C}\mathbb{D}_q^{\sigma}[y](\mathfrak{t}) = \mathbb{I}_q^{[\sigma]-\sigma}\left[\mathbb{D}_q^{[\sigma]}[y]\right](\mathfrak{t}) = \int_0^{\mathfrak{t}} (\mathfrak{t}-\xi)_q^{([\sigma]-\sigma-1)} \frac{\mathbb{D}_q^{[\sigma]}[y](\xi)}{\Gamma_q([\sigma]-\sigma)}\, \mathrm{d}_q\xi, \tag{5}$$

where $\sigma > 0$ and for all $\mathfrak{t} \in [0, 1]$ [12, 19]. For $\sigma, \nu \geq 0$, we can prove that $\mathbb{I}_q^{\nu}[\mathbb{I}_q^{\sigma}[y]](\mathfrak{t}) = \mathbb{I}_q^{\sigma+\nu}[y](\mathfrak{t})$, and ${}^{C}\mathbb{D}_q^{\sigma}[\mathbb{I}_q^{\sigma}[y]](\mathfrak{t}) = y(\mathfrak{t})$, see [12].

Lemma 2.1 ([17])**.** *Let $y \in AC^n[\mathfrak{t}_1, \mathfrak{t}_2]$. Then for $n-1 < \sigma \leq n$, $n \in \mathbb{N}$ one has $\mathbb{I}^{\sigma}[{}^{C}\mathbb{D}_q^{\sigma}[y]](\mathfrak{t}) = y(\mathfrak{t}) + \sum_{i=0}^{n-1} c_i(\mathfrak{t}-\mathfrak{t}_1)^i$, $c_0, c_1, \ldots, c_{n-1} \in \mathbb{R}$.*

Lemma 2.2 ([17])**.** *Let suppose that $\sigma \in (0, 1)$. Then for each $y \in AC[0, 1]$ we have $\mathbb{I}^{\sigma}[\mathbb{D}^{\sigma}[y]](\mathfrak{t}) = y(\mathfrak{t})$ for a.e. $\mathfrak{t} \in [0, 1]$. Here, we have $\mathbb{D}^{\sigma}[y](\mathfrak{t}) = \frac{\mathrm{d}}{\mathrm{d}\mathfrak{t}} \int_0^{\mathfrak{t}} (\mathfrak{t}-\xi)^{-\sigma} \frac{y(\xi)}{\Gamma(1-\sigma)}\, \mathrm{d}\xi$.*

Theorem 2.3 ([26] Banach contraction principle)**.** *Let assume that the space $\mathcal{X}$ is a Banach space and let $A : \mathcal{X} \to \mathcal{X}$ be a contraction map. Then, there exists an $x \in \mathcal{X}$ such that $Ax = x$.*

Theorem 2.4 ([26] Krasnoselskii's fixed point theorem)**.** *Consider a nonempty subset S of a Banach space $\mathcal{X}$ such that be a closed and convex and two maps A and B of S into $\mathcal{X}$ such that $A[k] + B[l] \in S$ for $k, l \in S$. Let suppose that B is a contraction map and let A is also compact and continuous map. Then, there exists a $k \in S$ such that $k = A[k] + B[k]$.*

3. Main and basic results

The main aim of this section is to investigate the existence of the solutions for the FqDE (1)-(2) by considering the fixed point theorems. We consider the set $\mathcal{X} = C^1(\overline{J}, \mathbb{R})$ endowed with the norm $\|k\|_* = \sup_{t\in\overline{J}} |k(t)| + \sup_{t\in\overline{J}} |k'(t)|$.

Lemma 3.1. *Assume that we have $v \in L^1(\overline{J}, \mathbb{R})$. Assume further that the FqDE ${}^{c}\mathcal{D}_q^{\sigma}[k](t) = v(t)$ under the conditions*

$$k(0) = h_1\left(t_0, k(t_0), {}^{c}\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right), \qquad k(1) = h_2\left(t_0, k(t_0), {}^{c}\mathcal{D}_q^{\alpha_4}[x](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right),$$

is given. Then, the unique solution is obtained as

$$\begin{aligned} k(t) &= \mathcal{I}_q^{\sigma}[v](t) + h_1\left(t_0, k(t_0), {}^{c}\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\ &\quad + t\big[h_2\left(t_0, k(t_0), {}^{c}\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - \mathcal{I}_q^{\sigma}[v](1) - h_1\left(t_0, k(t_0), {}^{c}\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\big]. \end{aligned} \tag{6}$$

Proof. We assume that $k(t)$ satisfies in the equation ${}^c\mathcal{D}_q^{\sigma}[k](t) = v(t)$. Lemma 2.2 implies that $k(t) = \mathcal{I}_q^{\sigma}[v](t) + c_0 + c_1 t$, where $c_0, c_1 \in \mathbb{R}$. Considering the boundary conditions, we conclude that $c_0 = h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)$ and

$$c_1 = h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - \mathcal{I}_q^{\sigma}[v](1) - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right).$$

Clearly Eq. (6) satisfies on the boundary conditions

$$k(0) = h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \qquad k(1) = h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), I_q^{\beta_4}[k](t_0)\right).$$

On the other hand, Lemmas 2.1 and 2.2 imply that

$${}^c\mathcal{D}_q^{\sigma}[k](t) = \mathcal{I}_q^{2-\sigma} k''(t) = \mathcal{I}_q^{2-\sigma}\left[\mathcal{I}_q^{-2+\sigma}[v]\right](t) = \mathcal{I}_q^{2-\sigma}\left[{}^c\mathcal{D}_q^{2-\sigma}[v]\right](t) = v(t).$$

Now, our proof is complete. □

Corollary 3.2. *A given function $k \in \mathcal{X}$ is called a solution of FqDE (1)-(2) iff*

$$\begin{aligned} k(t) &= \mathcal{I}_q^{\sigma} f\left(t, k(t), {}^c\mathcal{D}_q^{\alpha_1}[k](t), \mathcal{I}_q^{\beta_1}[k](t)\right) + h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\ &\quad + \Bigg[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\ &\quad - \int_{t_0}^1 \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\Bigg] t, \end{aligned}$$

whenever $0 \le t \le t_0$, and

$$\begin{aligned} k(t) &= \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1})[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \mathrm{d}_q s \\ &\quad + h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \left[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right)\right. \\ &\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\beta_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\ &\quad \left. - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\right] t, \end{aligned}$$

whenever $t_0 \le t \le 1$.

Theorem 3.3. *Let suppose that there exist $\ell \in (0, \sigma-1)$ and $L_1, L_2 \in L^{\frac{1}{\ell}}(\overline{J}, (0,\infty))$ and $L_3, L_4 \in C(\overline{J}, (0,\infty))$ s.t $\left|f(t, k_1, k_2, k_3) - f(t, k_1', k_2', k_3')\right| \le L_1(t) \sum_{i=1}^3 |k_i - k_i'|$, $\left|g(t, k_1, k_2, k_3) - g(t, k_1', k_2', k_3')\right| \le L_2(t) \sum_{i=1}^3 |k_i - k_i'|$, $\left|h_1(t, k_1, k_2, k_3) - h_1(t, k_1', k_2', k_3')\right| \le L_3(t) \sum_{i=1}^3 |k_i - k_i'|$, $\left|h_2(t, k_1, k_2, k_3) - h_2(t, k_1', k_2', k_3')\right| \le L_4(t) \sum_{i=1}^3 |k_i - k_i'|$, for all $t \in \overline{J}$ and k_ℓ, k_ℓ', with $\ell = 1, 2, 3$. Then FqDE (1)-(2) has a unique solution if*

$$\begin{aligned} \Lambda_1 &= \tfrac{3\|L_1\|_{\frac{1}{\ell}} \eta_1}{\Gamma_q(\sigma)}\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right] + \tfrac{3\|L_2\|_{\frac{1}{\ell}} \eta_1}{\Gamma_q(\sigma)}\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right] \\ &\quad + 3\|L_3\|\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right] + 2\|L_4\|\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right] \\ &\quad + \tfrac{\|L_1\|_{\frac{1}{\ell}} \eta_2}{\Gamma_q(\sigma-1)}\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right] + \tfrac{\|L_2\|_{\frac{1}{\ell}} \eta_2}{\Gamma_q(\sigma-1)}\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right] < 1, \end{aligned} \tag{7}$$

where $\eta_1 = \left(\frac{1-\ell}{\sigma-\ell}\right)^{1-\ell}$, $\eta_2 = \left(\frac{1-\ell}{\sigma-\ell+1}\right)^{1-\ell}$, and $\|L\|_p = \left(\int_0^1 |L(s)|^p \,\mathrm{d}s\right)^{\frac{1}{p}}$, for each $L \in L^p(\overline{J}, \mathbb{R})$.

Proof. Define the operator $\mathcal{T}$ on $\mathcal{X}$ by

$$\begin{aligned} \mathcal{T}[k](t) &= \mathcal{I}_q^{\sigma} f\left(t, k(t), {}^c\mathcal{D}_q^{\alpha_1}[k](t), \mathcal{I}_q^{\beta_1}[k](t)\right) + h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\ &\quad + \Bigg[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\ &\quad - \int_{t_0}^1 \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2} k(s), \mathcal{I}_q^{\beta_2} k(s)\right) \mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\Bigg] t, \end{aligned}$$

for $0 \leq t \leq t_0$, and

$$\begin{aligned}
\mathcal{T}[k](t) = &\int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\alpha)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s \\
&+ h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \Big[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\
&- \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\
&- \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\Big] t,
\end{aligned}$$

for $t_0 \leq t \leq 1$. Clearly, FqDE (1)-(2) admits a solution iff the relation $\mathcal{T}[k] = k$ has a fixed point. Let $k, l \in \mathcal{X}$. We then for $0 \leq t \leq t_0$ have

$$\begin{aligned}
|\mathcal{T}[k](t) - \mathcal{T}[l](t)| = &\big|\mathcal{I}_q^{\sigma} f\left(t, k(t), {}^c\mathcal{D}_q^{\alpha_1}[k](t), \mathcal{I}_q^{\beta_1}[k](t)\right) + h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\
&+ \big[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\
&- \int_{t_0}^{1} \tfrac{(1-qs)^{(\alpha-1)}}{\Gamma_q(\alpha)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s \\
&-h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\big] t - \mathcal{I}_q^{\alpha} f\left(t, l(t), {}^c\mathcal{D}_q^{\alpha_1}[l](t), \mathcal{I}_q^{\beta_1} l(t)\right) \\
&- h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right) - \big[h_2\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_4}[l](t_0), \mathcal{I}_q^{\beta_4}[l](t_0)\right) \\
&- \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \,\mathrm{d}_q s \\
&- \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \,\mathrm{d}_q s \; -h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right)\big] t\big| \\
\leq &\, \mathcal{I}_q^{\sigma}\big|f\left(t, k(t), {}^c\mathcal{D}_q^{\alpha_1}[k](t), \mathcal{I}_q^{\beta_1}[k](t)\right) - f\left(t, l(t), {}^c\mathcal{D}_q^{\alpha_1}[l](t), \mathcal{I}_q^{\beta_1}[l](t)\right)\big| \\
&+ 2\big|h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) - h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right)\big| \\
&+ \big|h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) - h_2\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_4}[l](t_0), \mathcal{I}_q^{\beta_4}[l](t_0)\right)\big| \\
&+ \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \big|f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) - f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right)\big| \,\mathrm{d}_q s \\
&+ \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \big|g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) - g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right)\big| \,\mathrm{d}_q s \\
\leq &\, \mathcal{I}_q^{\sigma} L_1(t)\left(|k(t) - l(t)| + |{}^c\mathcal{D}_q^{\alpha_1}[k](t) - {}^c\mathcal{D}_q^{\alpha_1}[l](s)| + |\mathcal{I}_q^{\beta_1}[k](t) - \mathcal{I}_q^{\beta_1}[l](t)|\right) \\
&+ 2L_3(t_0)\left(|k(t_0) - l(t_0)| + |{}^c\mathcal{D}_q^{\alpha_3}[k](t_0) - {}^c\mathcal{D}_q^{\alpha_3}[l](t_0)| + |\mathcal{I}_q^{\beta_3}[k](t_0) - \mathcal{I}^{\beta_3}[l](t_0)|\right) \\
&+ L_4(t_0)\left(|k(t_0) - l(t_0)| + |{}^c\mathcal{D}_q^{\alpha_4}[k](t_0) - {}^c\mathcal{D}_q^{\alpha_4}[l](t_0)| + |\mathcal{I}_q^{\beta_4}[k](t_0) - \mathcal{I}_q^{\beta_4}[l](t_0)|\right) \\
&+ \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} L_1(s)\,(|k(s) - l(s)| \\
&+|{}^c\mathcal{D}_q^{\alpha_1}[k](s) - {}^c\mathcal{D}_q^{\alpha_1}[l](s)| + |\mathcal{I}_q^{\beta_1}[k](s) - \mathcal{I}_q^{\beta_1}[l](s)|\big) \,\mathrm{d}_q s \\
&+ \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} L_2(s)\left(|k(s) - l(s)| + |{}^c\mathcal{D}_q^{\alpha_2}[k](s) - {}^c\mathcal{D}_q^{\alpha_2}[l](s)| + |\mathcal{I}_q^{\beta_1}[k](s) - \mathcal{I}_q^{\beta_2}[l](s)|\right) \,\mathrm{d}_q s \\
\leq &\, \mathcal{I}_q^{\sigma} L_1(t)\left(|k(t) - l(t)| + \mathcal{I}_q^{1-\alpha_1}|k'(s) - l'(w)| + \mathcal{I}_q^{\beta_1}|k(s) - l(s)|\right) \\
&+ 2L_3(t_0)\left(|k(t_0) - l(t_0)| + \mathcal{I}_q^{1-\alpha_3}|k'(t_0) - l'(t_0)| + \mathcal{I}_q^{\beta_3}|k(t_0) - l(t_0)|\right) \\
&+ L_4(t_0)\left(|k(t_0) - l(t_0)| + \mathcal{I}_q^{1-\alpha_4}|k'(t_0) - l'(t_0)| + \mathcal{I}_q^{\beta_4}|k(t_0) - l(t_0)|\right)
\end{aligned}$$

$$
\begin{aligned}
&+\int_0^{t_0}(1-qs)^{(\sigma-1)}\tfrac{L_1(s)}{\Gamma_q(\sigma)}\Big(|k(s)-l(s)|+\mathcal{I}_q^{1-\alpha_1}|k'(s)-l'(s)|+\mathcal{I}_q^{\beta_1}|k(s)-l(s)|\Big)\\
&+\int_{t_0}^{1}(1-qs)^{(\sigma-1)}\tfrac{L_2(s)}{\Gamma_q(\sigma)}\left(|k(s)-l(s)|+\int_0^s(s-u)^{-\alpha_2}\tfrac{|k'(u)-l'(u)|}{\Gamma_q(1-\alpha_2)}\,\mathrm{d}u\right.\\
&\left.+\int_0^s(s-u)^{\beta_2-1}\tfrac{|k(u)-l(u)|}{\Gamma_q(\beta_2)}\,\mathrm{d}u\right)\mathrm{d}s\\
\le\ &\mathcal{I}_q^{\sigma}L_1(t)\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)\|k-l\|_*+2L_3(t_0)\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)\|k-l\|_*\\
&+L_4(t_0)\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)\|k-l\|_*\\
&+\int_0^{t_0}\tfrac{(1-s)^{\sigma-1}}{\Gamma_q(\sigma)}L_1(s)\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)\|k-l\|_*ds\\
&+\int_{t_0}^{1}\tfrac{(1-s)^{\sigma-1}}{\Gamma_q(\sigma)}L_2(s)\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)\|k-l\|_*\,\mathrm{d}s\\
\le\ &\tfrac{\|k-l\|_*}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)\left[\int_0^t\left((t-s)^{\sigma-1}\right)^{\frac{1}{1-k}}\mathrm{d}s\right]^{1-k}\left[\int_0^t\Big(L_1(s)\Big)^{\frac{1}{k}}\mathrm{d}s\right]^{\ell}\\
&+\left[2\|L_3\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)+\|L_4\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)\right]\|k-l\|_*\\
&+\tfrac{\|k-l\|_*}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)\left[\int_0^{t_0}\left((1-s)^{\sigma-1}\right)^{\frac{1}{1-\ell}}\mathrm{d}s\right]^{1-\ell}\left[\int_0^{t_0}\Big(L_1(s)\Big)^{\frac{1}{\ell}}\mathrm{d}s\right]^{\ell}\\
&+\tfrac{\|k-l\|_*}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)\left[\int_{t_0}^{1}\left((1-s)^{\sigma-1}\right)^{\frac{1}{1-\ell}}\mathrm{d}s\right]^{1-\ell}\left[\int_{t_0}^{1}\Big(L_2(s)\Big)^{\frac{1}{\ell}}\mathrm{d}s\right]^{\ell}\\
\le\ &\left[\tfrac{2\|L_1\|_{\frac{1}{\ell}}}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)\left[\tfrac{1-\ell}{\sigma-\ell}\right]^{1-\ell}+\tfrac{\|L_2\|_{\frac{1}{\ell}}}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)\left[\tfrac{1-\ell}{\sigma-\ell}\right]^{1-\ell}\right.\\
&\left.+2\|L_3\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)+\|L_4\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)\right]\|k-l\|_*,
\end{aligned}\tag{8}
$$

and

$$
\begin{aligned}
|(\mathcal{T}[k])'(t)-(\mathcal{T}[l])'(t)|=\ &\Big|\mathcal{I}_q^{\sigma-1}f\left(t,k(t),{}^c\mathcal{D}_q^{\alpha_1}[k](t),\mathcal{I}_q^{\beta_1}[k](t)\right)+h_2\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_4}[k](t_0),\mathcal{I}^{\beta_4}[k](t_0)\right)\\
&-\int_0^{t_0}\tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}f\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_1}[k](s),\mathcal{I}_q^{\beta_1}[k](s)\right)\mathrm{d}_qs\\
&-\int_{t_0}^{1}\tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}g\left(s,k(s),{}^c\mathcal{D}_q^{(\alpha_2)}[k](s),\mathcal{I}_q^{\beta_2}[k](s)\right)\mathrm{d}_qs\\
&-h_1\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_3}[k](t_0),\mathcal{I}_q^{\beta_3}[k](t_0)\right)-\mathcal{I}_q^{\sigma-1}f\left(t,l(t),{}^c\mathcal{D}_q^{\alpha_1}l(t),\mathcal{I}_q^{\beta_1}[l](t)\right)\\
&-h_2\left(t_0,l(t_0),{}^c\mathcal{D}_q^{\alpha_4}[l](t_0),\mathcal{I}_q^{\beta_4}[l](t_0)\right)+\int_0^{t_0}\tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}f\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_1}[l](s),\mathcal{I}_q^{\beta_1}[l](s)\right)\mathrm{d}_qs\\
&+\int_{t_0}^{1}\tfrac{(1-qs)^{\sigma-1}}{\Gamma_q(\sigma)}g\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_2}[l](s),\mathcal{I}_q^{\beta_2}[l](s)\right)\mathrm{d}_qs+h_1\left(t_0,l(t_0),{}^c\mathcal{D}_q^{\alpha_3}[l](t_0),\mathcal{I}_q^{\beta_3}[l](t_0)\right)\Big|\\
\le\ &\left[\tfrac{\|L_1\|_{\frac{1}{\ell}}\eta_2}{\Gamma_q(\sigma-1)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+\tfrac{\|L_1\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+\tfrac{\|L_2\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}\right.\right.\\
&\left.\left.+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)+\|L_3\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)+\|L_4\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)\right]\|k-l\|_*,
\end{aligned}\tag{9}
$$

where $\eta_1 = \left(\frac{1-\ell}{\sigma-\ell}\right)^{1-\ell}$ and $\eta_2 = \left(\frac{1-\ell}{\sigma-\ell+1}\right)^{1-\ell}$. Similarly, for $t_0 \leq t \leq 1$ we get

$$\begin{aligned}
|\mathcal{T}[k](t) - \mathcal{T}[l](t)| = \Bigg| & \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\
& + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \mathrm{d}_q s \\
& + h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \left[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right)\right. \\
& - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\
& - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \mathrm{d}_q s \\
& \left. -h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3} x)(t_0)\right)\right] t - \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \mathrm{d}_q s \\
& - \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \mathrm{d}_q s \\
& - h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right) - \left[h_2\big(t_0, l(t_0), {}^c\mathcal{D}^{\alpha_4}[l](t_0), \mathcal{I}_q^{\beta_4} l[y](t_0)\big)\right. \\
& - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \mathrm{d}_q s \\
& \left. - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \mathrm{d}_q s \; -h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right)\right] t \Bigg| \\
\leq & \left[\tfrac{2\|L_1\|_{\frac{1}{\ell}} \eta_1}{\Gamma_q(\sigma)} \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + \tfrac{2\|L_2\|_{\frac{1}{\ell}} \eta_1}{\Gamma_q(\sigma)} \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) \right. \\
& \left. +2\|L_3\| \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right) + \|L_4\| \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) \right] \|k - l\|_*,
\end{aligned} \tag{10}$$

and

$$\begin{aligned}
|(\mathcal{T}[k])'(t) - (\mathcal{T}[l])'(t)| = \Bigg| & \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\
& + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_g}[k](s), \mathcal{I}_q^{\beta_g}[k](s)\right) \mathrm{d}_q s + h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\
& - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \mathrm{d}_q s \\
& - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\
& - \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \mathrm{d}_q s \\
& - \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_g}[l](s), \mathcal{I}_q^{\beta_g}[l](s)\right) \mathrm{d}_q s - h_2\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_4}[l](t_0), \mathcal{I}_q^{\beta_4}[l](t_0)\right) \\
& + \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \mathrm{d}_q s \\
& + \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \mathrm{d}_q s + h_1\left(t_0, l(t_0), {}^c\mathcal{D}_q^{\alpha_3}[l](t_0), \mathcal{I}_q^{\beta_3}[l](t_0)\right) \Bigg| \\
\leq & \left[\tfrac{\|L_1\|_{\frac{1}{k}} k_2}{\Gamma_q(\sigma-1)} \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + \tfrac{\|L_2\|_{\frac{1}{\ell}} \eta_2}{\Gamma_q(\sigma-1)} \left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) \right.
\end{aligned}$$

$$+\frac{\|L_1\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{\|L_2\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)$$
$$+\|L_3\|\left(1+\frac{1}{\Gamma_q(2-\alpha_3)}+\frac{1}{\Gamma_q(1+\beta_3)}\right)+\|L_4\|\left(1+\frac{1}{\Gamma_q(2-\alpha_4)}+\frac{1}{\Gamma_q(1+\beta_4)}\right)\Bigg]\|k-l\|_*, \tag{11}$$

where $\eta_1=\left(\frac{1-\ell}{\sigma-\ell}\right)^{1-\ell}$ and $\eta_2=\left(\frac{1-\ell}{\sigma-\ell+1}\right)^{1-\ell}$. By utilizing relations (8), (9), (10), and (11) we have

$$\begin{aligned}\|\mathcal{T}[k]-\mathcal{T}[l]\|_* &= \|\mathcal{T}[k]-\mathcal{T}[l]\|+\|(\mathcal{T}[k])'-(\mathcal{T}[l])'\|\\
&\le \Bigg[\frac{3\|L_1\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{3\|L_2\|_{\frac{1}{\ell}}\eta_1}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)\\
&\quad+3\|L_3\|\left(1+\frac{1}{\Gamma_q(2-\alpha_3)}+\frac{1}{\Gamma_q(1+\beta_3)}\right)+2\|L_4\|\left(1+\frac{1}{\Gamma_q(2-\alpha_4)}+\frac{1}{\Gamma_q(1+\beta_4)}\right)\\
&\quad+\frac{\|L_1\|_{\frac{1}{\ell}}\eta_2}{\Gamma_q(\sigma-1)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{\|L_2\|_{\frac{1}{\ell}}\eta_2}{\Gamma_q(\sigma-1)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)\Bigg]\|k-l\|_*=\Lambda_1\|k-l\|_*.\end{aligned}$$

Thus $\mathcal{T}$ is a contraction mapping due to the fact that $\Lambda_1<1$. Therefore, by using the Banach contraction principle we conclude that $\mathcal{T}$ has a unique fixed point. This fixed point is the unique solution of the model problem (1)-(2) by using Corollary 3.2. □

Corollary 3.4. *Assume that there exist L_1, L_2, L_3 and $L_4\in\mathbb{R}^+$ such that*

$$\left|f(t,k_1,k_2,k_3)-f(t,k_1',k_2',k_3')\right|\le L_1(t)\sum_{i=1}^{3}|k_i-k_i'|,\qquad \left|g(t,k_1,k_2,k_3)-g(t,k_1',k_2',k_3')\right|\le L_2(t)\sum_{i=1}^{3}|k_i-k_i'|,$$
$$\left|h_1(t,k_1,k_2,k_3)-h_1(t,k_1',k_2',k_3')\right|\le L_3(t)\sum_{i=1}^{3}|k_i-k_i'|,\quad \left|h_2(t,k_1,k_2,k_3)-h_2(t,k_1',k_2',k_3')\right|\le L_4(t)\sum_{i=1}^{3}|k_i-k_i'|,$$

for all $t\in\overline{J}$ and k_ℓ, k_ℓ' with $\ell=1,2,3$. Then the FqDE (1)-(2) has a unique solution whenever

$$\frac{3L_1}{\Gamma_q(\sigma+1)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{3L_2}{\Gamma_q(\sigma+1)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)+3L_3\left(1+\frac{1}{\Gamma_q(2-\alpha_3)}+\frac{1}{\Gamma_q(1+\beta_3)}\right)$$
$$+2L_4\left(1+\frac{1}{\Gamma_q(2-\alpha_4)}+\frac{1}{\Gamma_q(1+\beta_4)}\right)+\frac{L_1}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{L_2}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)<1.$$

By the aid of the Krasnoselskii's fixed-point theorem, we state our next existence result.

Theorem 3.5. *Suppose that there exist L_1, L_2, μ_1 and $\mu_2\in C(\overline{J},[0,\infty))$ and two nondecreasing self-functions ψ_1 and ψ_2 defined on $\mathbb{R}^+$ s.t $|f(t,k_1,k_2,k_3)-f(t,k_1',k_2',k_3')|$ and $|g(t,k_1,k_2,k_3)-g(t,k_1',k_2',k_3')|$, are less than or equal to $L_1(t)\sum_{i=1}^3|k_i-k_i'|$, $L_2(t)\sum_{i=1}^3|k_i-k_i'|$ respectively, and $|h_1(t,k_1,k_2,k_3)|\le\mu_1(t)\psi_1\sum_{i=1}^3|k_i|$, and $|h_2(t,k_1,k_2,k_3)|\le\mu_2(t)\psi_1\sum_{i=1}^3|k_i|$, for each $t\in\overline{J}$ and k_ℓ, k_ℓ' with $\ell=1,2,3$. If*

$$\Lambda_2=\left[\frac{\|L_1\|}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{\|L_2\|}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)\right]\left(\frac{1}{\sigma}+1\right)<1, \tag{12}$$

then, the model problem (1) admits a solution on $\overline{J}$.

Proof. Consider $S=\{k\in\mathcal{X}:\|k\|\le r\}$ where

$$3\|\mu_1\|\psi_1\left(\left(1+\frac{1}{\Gamma_q(2-\alpha_3)}+\frac{1}{\Gamma_q(1+\beta_3)}\right)r\right)+2\|\mu_2\|\psi_2\left(\left(1+\frac{1}{\Gamma_q(2-\alpha_4)}+\frac{1}{\Gamma_q(1+\beta_4)}\right)r\right)+\frac{r}{\Gamma_q(\sigma)}\left(\frac{2}{\sigma}+\sigma+1\right)\Bigg[\|L_1\|\Bigg(1$$
$$+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\Bigg)+F_0\Bigg]+\frac{r}{\Gamma_q(\sigma)}\left(\frac{2}{\sigma}+\sigma+1\right)\left[\|L_2\|\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)+G_0\right]\le r.$$

Clearly the set S is an nonempty subset of the Banach space $\mathcal{X}$, closed and convex. Now, we define two operators A and B on S by

$$\begin{aligned} A[k](t) &= h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \Big[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\ &\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\ &\quad - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\Big] t, \end{aligned}$$

for each $0 \le t \le 1$, and $B[k](t) = \mathcal{I}_q^{\sigma} f)\left(t, k(t), {}^c\mathcal{D}_q^{\alpha_1}[k](t), \mathcal{I}_q^{\beta_1}[k](t)\right)$, whenever $0 \le t \le t_0$, and

$$B[k](t) = \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s,$$

whenever $t_0 \le t \le 1$. Let $k, l \in S$. On the interval $0 \le t \le t_0$ we get

$$\begin{aligned} |A[k](t) + B[l](t)| &= \big|h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \big[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\ &\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\ &\quad - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s \\ &\quad -h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right)\big] t + \mathcal{I}_q^{\sigma} f\left(t, l(t), {}^c\mathcal{D}_q^{\alpha_1}[l](t), \mathcal{I}_q^{\beta_1}[l](t)\right)\big| \\ &\le 2\mu_1(t_0)\psi_1\left(|k(t_0)| + |{}^c\mathcal{D}_q^{\alpha_3}[k](t_0)| + |\mathcal{I}_q^{\beta_3}[k](t_0)|\right) + \mu_2(t_0)\psi_2\left(|k(t_0)| + |{}^c\mathcal{D}_q^{\alpha_4}[k](t_0)| + |\mathcal{I}_q^{\beta_4}[k](t_0)|\right) \\ &\quad + \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_1(s)|k(s) + {}^c\mathcal{D}_q^{\alpha_1}[k](s) + \mathcal{I}_q^{\beta_1}[k](s)| + F_0\right) \,\mathrm{d}_q s \\ &\quad + \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_2(s)|k(s) + {}^c\mathcal{D}_q^{\alpha_2}[k](s) + \mathcal{I}_q^{\beta_2}[k](s)| + G_0\right) \,\mathrm{d}_q s \\ &\quad + \int_0^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_1(s)|l(s) + {}^c\mathcal{D}_q^{\alpha_1}[l](s) + \mathcal{I}_q^{\beta_1}[l](s)| + F_0\right) \,\mathrm{d}_q s \\ &\le 2\|\mu_1\|\psi_1\left(\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right] r\right) + \|\mu_2\|\psi_2\left(\left[1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right] r\right) \\ &\quad + \tfrac{r}{\Gamma_q(\sigma+1)}\left[2\|L_1\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + 2F_0 + \|L_2\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) + G_0\right], \end{aligned}$$

$$\begin{aligned} |(A[k])'(t) + (B[l])'(t)| &= \Big|h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\ &\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\ &\quad - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\ &\quad + I_q^{\sigma-1} f\left(t, l(t), {}^c\mathcal{D}_q^{\alpha_1}[l](t), \mathcal{I}_q^{\beta_1}[l](t)\right)\Big| \\ &\le \|\mu_2\|\psi_2\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) r\right) + \|\mu_1\|\psi_1\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right) r\right) \\ &\quad + \tfrac{r(\sigma+1)}{\Gamma_q(\sigma)}\left[\|L_1\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + F_0\right] + \tfrac{r}{\Gamma_q(\sigma+1)}\left[\|L_2\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) + G_0\right]. \end{aligned}$$

Similarly, on the interval $t_0 \le t \le 1$ one gets

$$\begin{aligned}
|A[k](t)+B[l](t)| &= \Big| h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) + \Big[h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\
&\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\
&\quad - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \Big] t \\
&\quad + \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \,\mathrm{d}_q s \\
&\quad + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \,\mathrm{d}_q s \Big| \\
&\le 2\mu_1(t_0)\psi_1\left(|k(t_0)| + |{}^c\mathcal{D}_q^{\alpha_3}[k](t_0)| + |\mathcal{I}_q^{\beta_3}[k](t_0)|\right) + \mu_2(t_0)\psi_2\left(|k(t_0)| + |{}^c\mathcal{D}_q^{\alpha_4}[k](t_0)| + |\mathcal{I}_q^{\beta_4}[k](t_0)|\right) \\
&\quad + \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_1(s)|k(s) + {}^c\mathcal{D}_q^{\alpha_1}[k](s) + \mathcal{I}_q^{\beta_1}[k](s)| + F_0\right) \mathrm{d}_q s \\
&\quad + \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_2(s)|k(s) + {}^c\mathcal{D}_q^{\alpha_2}[k](s) + \mathcal{I}_q^{\beta_2}[k](s)| + G_0\right) \mathrm{d}_q s \\
&\quad + \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_1(s)|l(s) + {}^c\mathcal{D}_q^{\alpha_1}[l](s) + \mathcal{I}_q^{\beta_1}[l](s)| + F_0\right) \mathrm{d}_q s \\
&\quad + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} \left(L_2(s)|l(s) + {}^c\mathcal{D}_q^{\alpha_2}[l](s) + \mathcal{I}_q^{\beta_2}[l](s)| + G_0\right) \mathrm{d}_q s \\
&\le 2\|\mu_1\|\psi_1\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right) r\right) + \|\mu_2\|\psi_2\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) r\right) \\
&\quad + \tfrac{2r}{\Gamma_q(\sigma+1)}\left[\|L_1\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + F_0 + \|L_2\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) + G_0\right],
\end{aligned}$$

$$\begin{aligned}
|(A[k])'(t) + (B[l])'(t)| &= \Big| h_2\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_4}[k](t_0), \mathcal{I}_q^{\beta_4}[k](t_0)\right) \\
&\quad - \int_0^{t_0} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} f\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_1}[k](s), \mathcal{I}_q^{\beta_1}[k](s)\right) \,\mathrm{d}_q s \\
&\quad - \int_{t_0}^{1} \tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)} g\left(s, k(s), {}^c\mathcal{D}_q^{\alpha_2}[k](s), \mathcal{I}_q^{\beta_2}[k](s)\right) \,\mathrm{d}_q s - h_1\left(t_0, k(t_0), {}^c\mathcal{D}_q^{\alpha_3}[k](t_0), \mathcal{I}_q^{\beta_3}[k](t_0)\right) \\
&\quad + \int_0^{t_0} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} f\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_1}[l](s), \mathcal{I}_q^{\beta_1}[l](s)\right) \,\mathrm{d}_q s \\
&\quad + \int_{t_0}^{t} \tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)} g\left(s, l(s), {}^c\mathcal{D}_q^{\alpha_2}[l](s), \mathcal{I}_q^{\beta_2}[l](s)\right) \,\mathrm{d}_q s \Big| \\
&\le \|\mu_2\|\psi_2\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) r\right) + \|\mu_1\|\psi_1\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right) r\right) \\
&\quad + \tfrac{r(\sigma+1)}{\Gamma_q(\sigma)}\left[\|L_1\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + F_0\right] + \tfrac{r(\sigma+1)}{\Gamma_q(\sigma)}\left[\|L_2\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) + G_0\right],
\end{aligned}$$

where $F_0 = \sup_{t\in\overline{J}} |f(t,0,0,0)|$ and $G_0 = \sup_{t\in\overline{J}} |g(t,0,0,0)|$. Thus

$$\begin{aligned}
\|A[k] + B[l]\|_* = \|A[k] + B[l]\| + \|(A[k])' + (B[l])'\| &\le 2\|\mu_1\|\psi_1\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_3)} + \tfrac{1}{\Gamma_q(1+\beta_3)}\right) r\right) \\
&\quad + \|\mu_2\|\psi_2\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) r\right) + \tfrac{2r}{\Gamma_q(\sigma+1)}\Big[\|L_1\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_1)} + \tfrac{1}{\Gamma_q(1+\beta_1)}\right) + F_0 \\
&\quad + \|L_2\|\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_2)} + \tfrac{1}{\Gamma_q(1+\beta_2)}\right) + G_0\Big] + \|\mu_2\|\psi_2\left(\left(1 + \tfrac{1}{\Gamma_q(2-\alpha_4)} + \tfrac{1}{\Gamma_q(1+\beta_4)}\right) r\right)
\end{aligned}$$

$$
\begin{aligned}
&+\|\mu_1\|\psi_1\left(\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)r\right)+\tfrac{r(\sigma+1)}{\Gamma_q(\sigma)}\left[\|L_1\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+F_0\right]\\
&+\tfrac{r(\sigma+1)}{\Gamma_q(\sigma)}\left[\|L_2\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)+G_0\right]\\
=\ &3\|\mu_1\|\psi_1\left(\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)r\right)+2\|\mu_2\|\psi_2\left(\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)r\right)\\
&+\tfrac{r}{\Gamma_q(\sigma)}\left(\tfrac{2}{\sigma}+\sigma+1\right)\left[\|L_1\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+F_0\right]\\
&+\tfrac{r}{\Gamma_q(\sigma)}\left(\tfrac{2}{\sigma}+\sigma+1\right)\left[\|L_2\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)+G_0\right]\leq r.
\end{aligned}
$$

Hence for each k and $l\in S$, $A[k]+B[l]\in S$. For each $k\in S$, we have

$$
\begin{aligned}
\|A[k]\|_*\leq\ &3\|\mu_1\|\psi_1\left(\left(1+\tfrac{1}{\Gamma_q(2-\alpha_3)}+\tfrac{1}{\Gamma_q(1+\beta_3)}\right)r\right)+2\|\mu_2\|\psi_2\left(\left(1+\tfrac{1}{\Gamma_q(2-\alpha_4)}+\tfrac{1}{\Gamma_q(1+\beta_4)}\right)r\right)\\
&+\tfrac{2r}{\Gamma_q(\sigma+1)}\left[\|L_1\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+F_0+\|L_2\|\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)+G_0\right].
\end{aligned}
$$

Thus, we conclude the uniformly boundedness of the operator A on S. For any $k\in S$ and $t<\tau\in\overline{J}$, we also have

$$
\begin{aligned}
|A[k](\tau)-A[k](t)|=(\tau-t)\Big[&h_2\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_4}[k](t_0),\mathcal{I}_q^{\beta_4}[k](t_0)\right)-\int_0^{t_0}\tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}f\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_1}[k](s),\mathcal{I}_q^{\beta_1}[k](s)\right)\,\mathrm{d}_qs\\
&-\int_{t_0}^1\tfrac{(1-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}g\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_2}[k](s),\mathcal{I}_q^{\beta_2}[k](s)\right)\,\mathrm{d}_qs-h_1\left(t_0,k(t_0),{}^c\mathcal{D}_q^{\alpha_3}[k](t_0),\mathcal{I}_q^{\beta_3}[k](t_0)\right)\Big],
\end{aligned}
$$

which is not dependent on k and approaches to 0 as $t\to\tau$. Indeed, the operator A is equicontinuous. Consequently, by invoking the Arzelá-Ascoli theorem we asserted that the operator A is compact on S. We now consider two elements k and l belonging to S. Then, we get

$$
\begin{aligned}
|B[k](t)-B[l](t)|&=\left|\mathcal{I}_q^{\sigma}f\left(t,k(t),{}^c\mathcal{D}_q^{\alpha_1}[k](t),\mathcal{I}_q^{\beta_1}[k](t)\right)-\mathcal{I}_q^{\sigma}f\left(t,l(t),{}^c\mathcal{D}_q^{\alpha_1}[l](t),\mathcal{I}_q^{\beta_1}[l](t)\right)\right|\\
&\leq\tfrac{\|L_1\|}{\Gamma_q(\sigma+1)}\left[1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right]\|k-l\|_*,\\
|(B[k])'(t)-(B[l])'(t)|&=\left|\mathcal{I}_q^{\sigma-1}f\left(t,k(t),{}^c\mathcal{D}_q^{\alpha_1}[k](t),\mathcal{I}_q^{\beta_1}[k](t)\right)-\mathcal{I}_q^{\sigma-1}f\left(t,l(t),{}^c\mathcal{D}_q^{\alpha_1}[l](t),\mathcal{I}_q^{\beta_1}[l](t)\right)\right|\\
&\leq\tfrac{\|L_1\|}{\Gamma_q(\sigma)}\left[1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right]\|k-l\|_*,
\end{aligned}
$$

whenever $0\leq t\leq t_0$. Also, we have

$$
\begin{aligned}
|B[k](t)-B[l](t)|=\Big|&\int_0^{t_0}\tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}f\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_1}[k](s),\mathcal{I}_q^{\beta_1}[k](s)\right)\,\mathrm{d}_qs\\
&+\int_{t_0}^t\tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}g\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_2}[k](s),\mathcal{I}_q^{\beta_2}[k](s)\right)\,\mathrm{d}_qs-\int_0^{t_0}\tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}f\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_1}[l](s),\mathcal{I}_q^{\beta_1}[l](s)\right)\,\mathrm{d}_qs\\
&-\int_{t_0}^t\tfrac{(t-qs)^{(\sigma-1)}}{\Gamma_q(\sigma)}g\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_2}[l](s),\mathcal{I}_q^{\beta_2}[l](s)\right)\,\mathrm{d}_qs\Big|\\
\leq\ &\|x-y\|_*\left[\tfrac{\|L_1\|}{\Gamma_q(\sigma+1)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+\tfrac{\|L_2\|}{\Gamma_q(\sigma+1)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)\right],\\
|(B[k])'(t)-(B[l])'(t)|=\Big|&\tfrac{1}{\Gamma_q(\sigma-1)}\int_0^{t_0}(t-qs)^{(\sigma-2)}f\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_1}[k](s),\mathcal{I}_q^{\beta_1}[k](s)\right)\,\mathrm{d}_qs\\
&+\int_{t_0}^t\tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)}g\left(s,k(s),{}^c\mathcal{D}_q^{\alpha_2}[k](s),\mathcal{I}_q^{\beta_2}[k](s)\right)\,\mathrm{d}_qs-\int_0^{t_0}\tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)}f\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_1}[l](s),\mathcal{I}_q^{\beta_1}[l](s)\right)\,\mathrm{d}_qs\\
&-\int_{t_0}^t\tfrac{(t-qs)^{(\sigma-2)}}{\Gamma_q(\sigma-1)}g\left(s,l(s),{}^c\mathcal{D}_q^{\alpha_2}[l](s),\mathcal{I}_q^{\beta_2}[l](s)\right)\,\mathrm{d}_qs\Big|\\
\leq\ &\|k-l\|_*\left[\tfrac{\|L_1\|}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_1)}+\tfrac{1}{\Gamma_q(1+\beta_1)}\right)+\tfrac{\|L_2\|}{\Gamma_q(\sigma)}\left(1+\tfrac{1}{\Gamma_q(2-\alpha_2)}+\tfrac{1}{\Gamma_q(1+\beta_2)}\right)\right],
\end{aligned}
$$

whenever $t_0 \leq t \leq 1$. Therefore,

$$\|B[k]-B[l]\|_* \leq \left[\frac{\|L_1\|}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_1)}+\frac{1}{\Gamma_q(1+\beta_1)}\right)+\frac{\|L_2\|}{\Gamma_q(\sigma)}\left(1+\frac{1}{\Gamma_q(2-\alpha_2)}+\frac{1}{\Gamma_q(1+\beta_2)}\right)\right]\left(\frac{1}{\sigma}+1\right)\|k-l\|_* \leq \Lambda_2\|k-l\|_*.$$

Since $\Lambda_2 < 1$. Indeed, operator B is contraction. In the other words, all conditions of Theorem 2.3 are fulfilled. This indicates that there exists a k belonging to S such that $A[k]+B[k]=k$. So, we conclude that equation (1) has a solution, which is in $\overline{J}$. This finishes the validation. □

4. Applications

At present, we provide two examples for illustrating our main results which their numerical results are presented using the required algorithms. In this way, some computational results are carried out to check the feasibility of the theoretical findings for FqDE (1)-(2).

Example 4.1. As the first test example, let us pay attention to the boundary value differential $q-$fractional problem in the form

$$ {}^c\mathcal{D}_q^{\frac{3}{2}}[k](t) = \begin{cases} \frac{1}{100}\left(t^2+\frac{1}{2}t-\frac{1}{2}\right)\left[k(t)+\tan^{-1}\left({}^c\mathcal{D}_q^{\frac{1}{3}}[k](t)\right)+\sin\left(\mathcal{I}_q^{\sqrt{2}}[k](t)\right)\right], & 0 \leq t \leq \frac{3}{7}, \\ \frac{1}{100}\left(t^2+\frac{(\sqrt{2}-1)}{2}t-\frac{\sqrt{2}}{4}\right)\left[\frac{|k(t)|}{1+|k(t)|}+\frac{\left|{}^c\mathcal{D}_q^{\frac{1}{4}}[k](t)+\mathcal{I}_q^{\sqrt{3}}[k](t)\right|}{1+\left|{}^c\mathcal{D}_q^{\frac{1}{4}}[k](t)+\mathcal{I}_q^{\sqrt{3}}[k](t)\right|}\right], & \frac{3}{7} \leq t \leq 1. \end{cases} \tag{13}$$

The prescribed boundary conditions are

$$k(0) = \frac{e^{\frac{3}{7}}}{100}\left[\frac{\left|k\left(\frac{3}{7}\right)+{}^c\mathcal{D}_q^{\frac{1}{5}}[k]\left(\frac{3}{7}\right)+\mathcal{I}_q^{\sqrt{5}}[k]\left(\frac{3}{7}\right)\right|}{1+\left|k\left(\frac{3}{7}\right)+{}^c\mathcal{D}_q^{\frac{1}{5}}[k]\left(\frac{3}{7}\right)+\mathcal{I}_q^{\sqrt{5}}[k]\left(\frac{3}{7}\right)\right|}\right], \tag{14}$$

and

$$k(1) = \frac{1}{100}\sin\left(\frac{3}{7}\right)\left[\cos\left(k\left(\frac{3}{7}\right)\right)+\sin\left(\left({}^c\mathcal{D}_q^{\frac{1}{6}}k\right)\left(\frac{3}{7}\right)\right)+\tan^{-1}\left(\left(\mathcal{I}_q^{\sqrt{6}}k\right)\left(\frac{3}{7}\right)\right)\right]. \tag{15}$$

Here, $\sigma=\frac{3}{2}$, $\alpha_1=\frac{1}{3}$, $\alpha_2=\frac{1}{4}$, $\alpha_3=\frac{1}{5}$, $\alpha_4=\frac{1}{6}$, $\beta_1=\sqrt{2}$, $\beta_2=\sqrt{3}$, $\beta_3=\sqrt{5}$, $\beta_4=\sqrt{6}$, $t_0=\frac{1}{2}$,

$$f(t,k_1,k_2,k_3) = \frac{1}{100}\left[t^2+\frac{1}{2}t-\frac{1}{2}\right]\left(k_1+\tan^{-1}(k_2)+\sin(k_3)\right),$$
$$g(t,k_1,k_2,k_3) = \frac{1}{100}\left[t^2+\frac{(\sqrt{2}-1)}{2}t-\frac{\sqrt{2}}{4}\right]\left(\frac{|k_1|}{1+|k_1|}+\frac{|k_2+k_3|}{1+|k_2+k_3|}\right),$$

and $h_1(t,k_1,k_2,k_3)=\frac{e^t}{100}\left(\frac{|k_1+k_2+k_3|}{1+|k_1+k_2+k_3|}\right)$, $h_2(t,k_1,k_2,k_3)=\frac{1}{100}\sin(t)\left[\cos(k_1)+\sin(k_2)+\tan^{-1}(k_3)\right]$. Clearly

$$|f(t,k_1,k_2,k_3)-f(t,k_1',k_2',k_3')| \leq \frac{1}{100}(|k_1-k_1'|+|k_2-k_2'|+|k_3-k_3'|),$$
$$|g(t,k_1,k_2,k_3)-g(t,k_1',k_2',k_3')| \leq \frac{2+\sqrt{2}}{400}(|k_1-k_1'|+|k_2-k_2'|+|k_3-k_3'|),$$
$$|h_1(t,k_1,k_2,k_3)-h_1(t,k_1',k_2',k_3')| \leq \frac{1}{100}e(|k_1-k_1'|+|k_2-k_2'|+|k_3-k_3'|),$$

and $|h_2(t,k_1,k_2,k_3)-h_2(t,k_1',k_2',k_3')| \leq \frac{1}{100}\sin(1)(|k_1-k_1'|+|k_2-k_2'|+|k_3-k_3'|)$, for $t \in \overline{J}$ and $k_1,k_1',k_2,k_2',k_3,k_3' \in \mathbb{R}$. Hence, $L_1=\frac{1}{100}$, $L_2=\frac{2+\sqrt{2}}{400}$, $L_3=\frac{1}{100}e$, $L_4=\frac{1}{100}$, and by using Eq. (7), we obtain $\Lambda_1 \approx 0.35919$, 0.32314, 0.30295 for $q=\frac{1}{10},\frac{1}{2},\frac{8}{9}$, respectively. These results show in Tables 1 such that they emphasize with underline. Hence, all conditions of Corollary 3.4 are hold. This indicates that the differential $q-$fractional equation (13) has an unique solution under the Dirichlet boundary conditions (14) and (15), here the unique solution is in $\overline{J}$. We also note that, L_1, L_2, L_3, and L_4 are maximum of functions f, g, h_1, and h_2, respectively.

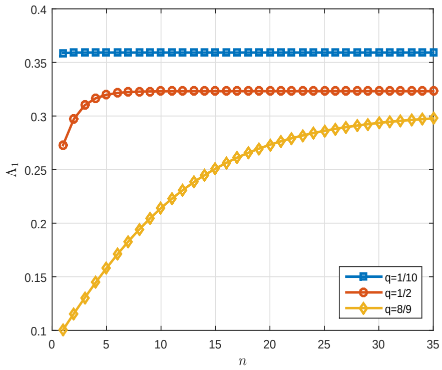


Fig. 1. Numerical evaluations of Λ_1 for various $q = \frac{1}{10}, \frac{1}{2}, \frac{8}{9}$ in Example 4.1

Table 1. Numerical evaluations of Λ_1 for various $q = \frac{1}{10}, \frac{1}{2}$, and $\frac{8}{9}$ in Example 4.1.

	Λ_1		
n	$q = \frac{1}{10}$	$q = \frac{1}{2}$	$q = \frac{8}{9}$
1	0.35874	0.27282	0.10052
2	0.35915	0.29729	0.11569
3	<u>0.35919</u>	0.31004	0.13043
4	0.35919	0.31655	0.14462
⋮	⋮	⋮	⋮
12	0.35919	0.32312	0.23091
13	0.35919	0.32313	0.23821
14	0.35919	<u>0.32314</u>	0.24485
15	0.35919	0.32314	0.25087
⋮	⋮	⋮	⋮
75	0.35919	0.32314	0.30293
76	0.35919	0.32314	0.30293
77	0.35919	0.32314	0.30294
78	0.35919	0.32314	0.30294
79	0.35919	0.32314	0.30294
80	0.35919	0.32314	<u>0.30295</u>
81	0.35919	0.32314	0.30295
82	0.35919	0.32314	0.30295

Example 4.2. The second test example devoted to the following boundary value differential $q-$fractional problem

$$
{}^{c}\mathcal{D}_q^{\frac{4}{3}}[k](t) = \begin{cases} \frac{\ln\left(t+\frac{7}{8}\right)}{3t+\pi^2+3}\left[\frac{\left|k(t)+{}^{c}\mathcal{D}_q^{\frac{1}{6}}[k](t)+\mathcal{I}_q^{\frac{4}{5}}[k](t)\right|}{1+\left|k(t)+{}^{c}\mathcal{D}_q^{\frac{1}{6}}[k](t)+\mathcal{I}_q^{\frac{4}{5}}[k](t)\right|}\right], & 0 \le t \le \frac{2}{5}, \\ \frac{1}{e^3+1}\left[t-\frac{2}{6}\right]^2\left[k(t)+\cos\left({}^{c}\mathcal{D}_q^{\frac{3}{7}}[k](t)\right)+\sin\left(\mathcal{I}_q^{\frac{3}{4}}[k](t)\right)\right], & \frac{2}{5} \le t \le 1, \end{cases} \tag{16}
$$

subjected to the boundary conditions

$$
k(0) = e^{\frac{2}{5}}\left[k\left(\tfrac{2}{5}\right) + {}^{c}\mathcal{D}_q^{\frac{3}{8}}[k]\left(\tfrac{2}{5}\right) + \mathcal{I}_q^{\frac{9}{7}}[k]\left(\tfrac{2}{5}\right)\right] \tag{17}
$$

and $k(1) = \sin\left(\frac{2}{5}\right)\left[k\left(\frac{2}{5}\right) + {}^{c}\mathcal{D}_q^{\frac{5}{8}}[k]\left(\frac{2}{5}\right) + \mathcal{I}_q^{\frac{10}{7}}[k]\left(\frac{2}{5}\right)\right]^{\frac{1}{2}}$. Here, $\sigma = \frac{4}{3}$, $\alpha_1 = \frac{1}{6}$, $\alpha_2 = \frac{3}{7}$, $\alpha_3 = \frac{3}{8}$, $\alpha_4 = \frac{5}{8}$, $\beta_1 = \frac{4}{5}$, $\beta_2 = \frac{3}{4}$, $\beta_3 = \frac{9}{7}$, $\beta_4 = \frac{10}{7}$, $t_0 = \frac{2}{5}$,

$$
f(t, k+1, k_2, k_3) = \frac{\ln\left(t+\frac{7}{8}\right)}{3t+\pi^2+3}\left[\frac{|k_1+k_2+k_3|}{1+|k_1+k_2+k_3|}\right], \qquad g(t, k_1, k_2, k_3) = \frac{1}{e^3+1}\left(t-\frac{2}{5}\right)^2(k_1+k_2+k_3),
$$

and $h_1(t, k_1, k_2, k_3) = e^t \,(k_1 + k_2 + k_3),\ h_2(t, k_1, k_2, k_3) = \sin(t)\,(k_1 + k_2 + k_3)^{\frac{1}{2}}$.

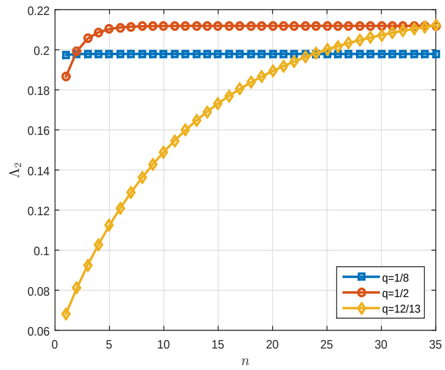


Fig. 2. Numerical evaluations of Λ_2 for various $q = \frac{1}{8}, \frac{1}{2}, \frac{12}{13}$ in Example 4.2

Clearly

$$\begin{aligned} |f(t, k_1, k_2, k_3) - f(t, k_1', k_2', k_3')| &\leq \frac{\ln\left(t + \frac{7}{8}\right)}{3t + \pi^2 + 3}\left(|k_1 - k_1'| + |k_2 - k_2'| + |k_3 - k_3'|\right), \\ |g(t, k_1, k_2, k_3) - g(t, k_1', k_2', k_3')| &\leq \frac{1}{e^3 + 1}\left(t - \frac{2}{5}\right)^2\left(|k_1 - k_1'| + |k_2 - k_2'| + |k_3 - k_3'|\right), \end{aligned}$$

and $|h_1(t, k_1, k_2, k_3)| \leq e^t(|k_1| + |k_2| + |k_3|)$, $|h_2(t, k_1, k_2, k_3)| \leq \sin(t)(|k_1| + |k_2| + |k_3|)^{\frac{1}{2}}$, for all $t \in \overline{J}$ and $k_1, k_2, k_3,\ k_1', k_2'$ and $k_3' \in \mathbb{R}$. Choose $L_1(t) = \frac{1}{3t+\pi^2+3}\ln(t + \frac{7}{8})$, $L_2(t) = \frac{1}{e^3+1}(t - \frac{2}{5})^2$, $\mu_1(t) = e^t$, $\mu_2(t) = \sin(t)$, $\psi_1(t) = t$, and $\psi_2(t) = t^{\frac{1}{2}}$. Eq. (12) yields $\Lambda_2 \approx 0.19790$, 0.21190 and $0.0.22194$ for $q = \frac{1}{8}, \frac{1}{2}$ and $\frac{12}{13}$, respectively. These values show in Tables 2 such that they emphasize with underline. Consequently, all the assumptions of Theorem 3.5 hold. This implies that the given fractional $q-$differential equation (16) admits at least one solution under the given Dirichlet boundary conditions.

Table 2. Numerical evaluations of Λ_2 for various $q = \frac{1}{8}, \frac{1}{2}$, and $\frac{12}{13}$ in Example 4.2.

	Λ_2		
n	$q = \frac{1}{8}$	$q = \frac{1}{2}$	$q = \frac{12}{13}$
1	0.19745	0.18656	0.06811
2	0.19785	0.19924	0.081
3	<u>0.1979</u>	0.20557	0.0925
4	0.1979	0.20874	0.10289
5	0.1979	0.21032	0.11235
⋮	⋮	⋮	⋮
11	0.1979	0.21188	0.15468
12	0.1979	0.21189	0.1599
13	0.1979	<u>0.2119</u>	0.1647
14	0.1979	0.2119	0.16913
⋮	⋮	⋮	⋮
112	0.1979	0.2119	0.22193
113	0.1979	0.2119	0.22193
114	0.1979	0.2119	0.22193
115	0.1979	0.2119	<u>0.22194</u>
116	0.1979	0.2119	0.22194
117	0.1979	0.2119	0.22194

5. Conclusion

In this paper, we first gave some properties of the fractional q–derivative and integral, and then using the proposed properties we have established the existence of solutions for the single and multi-dimensional fractional neutral functional q–differential equation (1) with Dirichlet boundary conditions (2) on a time scale. By numerical evaluations we confirmed our theoretical finding for the underlying model problem. Compared to existing published outcomes in the literature, this results of the current work are new form point of theoretical and numerical computational point of views.

References

[1] T. Abdeljawad, J. Alzabut, and D. Baleanu. A generalized q-fractional gronwall inequality and its applications to nonlinear delay q-fractional difference systems. *Journal of Inequalities and Applications*, 216:240, 2016. doi: 10.1186/s13660-016-1181-2.

[2] C. R. Adams. The general theory of a class of linear partial q-difference equations. *Transactions of the American Mathematical Society*, 26: 283–312, 1924.

[3] R. P. Agarwal, D. Baleanu, V. Hedayati, and S. Rezapour. Two fractional derivative inclusion problems via integral boundary condition. *Appl. Math. Comput.*, 257:205–212, 2015. doi: 10.1016/j.amc.2014.10.082.

[4] B. Ahmad, J. J. Nieto, A. Alsaedi, and H. Al-Hutami. Existence of solutions for nonlinear fractional q-difference integral equations with two fractional orders and nonlocal four-point boundary conditions. *J. Franklin Inst.*, 351:2890–2909, 2014.

[5] B. Ahmad, M. Alghanmi, A. Alsaedi, H. M. Srivastava, and S. K. Ntouyas. The Langevin equation in terms of generalized Liouville-Caputo derivatives with nonlocal boundary conditions involving a generalized fractional integral. *Mathematics*, 7(6):533, 2019.

[6] M. H. Annaby and Z. S. Mansour. *q-Fractional Calculus and Equations*. Springer Heidelberg, Cambridge, 2012. doi: 10.1007/978-3-642-30898-7.

[7] Z. Baitiche, C. Derbazi, J. Alzabut, M. E. Samei, M. K. A. Kaabar, and Z. Siri. Monotone iterative method for Langevin equation in terms of ψ-Caputo fractional derivative and nonlinear boundary conditions. *Fractal Fractional*, 5(2):81, 2021. doi: 10.3390/fractalfract5030081.

[8] M. Bohner and A. Peterson. *Dynamic Equations on Time Scales*. Birkhäuser, Boston, 2001.

[9] A. Bonfanti, J. L. Kaplan, G. Charras, and A. Kabla. Fractional viscoelastic models for power-law materials. *Soft Matter.*, 16:6002–6020, 2020.

[10] J. Danane, K. Allali, and Z. Hammouch. Mathematical analysis of a fractional differential model of HBV infection with antibody immune response. *Chaos Solit. Fract.*, 136:109787, 2020.

[11] S. Djilali, B. Ghanbari, and W. Adel. The influence of an infectious disease on a prey-predator model equipped with a fractional-order derivative. *Adv. Differ. Equ.*, 2021:20, 2021.

[12] R. A. C. Ferreira. Nontrivials solutions for fractional q-difference boundary value problems. *Electronic Journal of Qualitative Theory of Differential Equations*, 70:1–101, 2010.

[13] M. Izadi, S. Yuzbasi, and W. Adel. Accurate and efficient matrix techniques for solving the fractional Lotka-Volterra population model. *Physica A*, 600:127558, 2022.

[14] M. Izadi, N. Sene, W. Adel, and A. El-Mesady. The Layla and Majnun mathematical model of fractional order: Stability analysis and numerical study. *Results Phys.*, 51:106650, 2023.

[15] F. H. Jackson. q-difference equations. *American Journal of Mathematics*, 32:305–314, 1910. doi: 10.2307/2370183.

[16] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo. *Theory and Application of Fractional Differential Equations*. Elsevier, Amsterdam, The Netherlands, 2004.

[17] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego, 1999.

[18] S. Qureshi, A. Yusuf, A. A. Shaikh, M. Inc, and D. Baleanu. Fractional modeling of blood ethanol concentration system with real data application. *Chaos*, 29:013143, 2019.

[19] P. M. Rajković, S. D. Marinković, and M. S. Stanković. Fractional integrals and derivatives in q-calculus. *Applicable Analysis and Discrete Mathematics*, 1:311–323, 2007.

[20] S. Rezapour and H. Mohammadi. SEIR epidemic model for COVID-19 transmission by Caputo derivative of fractional order. *Adv. Differ. Equ.*, 2020:490, 2022.

[21] S. Rezapour and M. E. Samei. On the existence of solutions for a multi-singular pointwise defined fractional q-integro-differential equation. *Boundary Value Problems*, 2020:38, 2020. doi: 10.1186/s13661-020-01342-3.

[22] M. E. Samei and S. Rezapour. On a system of fractional q-differential inclusions via sum of two multi-term functions on a time scale. *Boundary Value Problems*, 2020:135, 2020. doi: 10.1186/s13661-020-01433-1.

[23] M. E. Samei and W. Yang. Existence of solutions for k-dimensional system of multi-term fractional q-integro-differential equations under anti-periodic boundary conditions via quantum calculus. *Mathematical Modelling in Applied Sciences*, 43(7):4360–4382, 2020. doi: 10.1002/mma.6198.

[24] M. E. Samei, R. Ghaffari, S. W. Yao, M. K. A. Kaabar, F. Martínez, and M. Inc. Existence of solutions for a singular fractional q-differential equations under Riemann-Liouville integral boundary condition. *Symmetry*, 13:135, 2021. doi: 10.3390/sym13071235.

[25] M. Shabibi, Z. Zeinalabedini Charandabi, H. Mohammadi, and S. Rezapour. Investigation of mathematical model of human liver by Caputo fractional derivative approach. *J. Adv. Math. Model.*, 11:750–760, 2021.

[26] D. R. Smart. *Fixed Point Theorems*. Cambridge University Press, New York, 1980.

[27] H. M. Srivastava. Some parametric and argument variations of the operators of fractional calculus and related special functions and integral transformations. *J. Nonlinear Convex Anal.*, 22:1501–1520, 2021.

[28] H. M. Srivastava. An introductory overview of fractional-calculus operators based upon the Fox-Wright and related higher transcendental functions. *J. Adv. Engrg. Comput.*, 5:135–166, 2021.

[29] S. Yuzbasi and M. Izadi. Bessel-quasilinearization technique to solve the fractional-order HIV-1 infection of CD4+T-cells considering the impact of antiviral drug treatment. *Appl. Math. Comput.*, 431:127319, 2022.

[30] Y. Zhao, H. Chen, and Q. Zhang. Existence results for fractional q-difference equations with nonlocal q–integral boundary conditions. *Adv. Differ. Equ.*, 2013:48, 2013. doi: 10.1186/1687-1847-2013-48.

[31] H. Zhou, J. Alzabut, and L. Yang. On fractional Langevin differential equations with anti-periodic boundary conditions. *The European Physical Journal Special Topics*, 226:3577–3590, 2017. doi: 10.1140/epjst/e2018-00082-0.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–113, pp. 17–21

Existence of solutions for an Ecological Model involving nonlocal operators

S. Shakeri[a]

[a]*Department of Mathematics, Ayatollah Amoli Branch, Islamic Azad University, Amol, P. O. Box 678, Iran*

Article Info

Keywords:
systems
Indefinite weight
key3
Grazing and constant yield harvesting
Sub-supersolution method

2020 MSC:
msc1
msc2

Abstract

This study concerns the existence of a positive solution for the following nonlinear boundary value problem

$$\begin{cases} -M\left(\int_\Omega |\nabla u|^p\,dx\right)\Delta_p u = am(x)u^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K & \text{in } \Omega,\\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

Here, $\Delta_p u := \operatorname{div}(|\nabla u|^{p-2}\nabla u)$ is the p-Laplacian operator, $p > 1$, a, b, c, γ, K are positive constants with $\gamma \geq 2$, $M : \mathbb{R}_0^+ \to \mathbb{R}^+$ is a continuous and increasing function and Ω is a smooth bounded region with $\partial\Omega$ belonging to C^2. The weight function $m(x)$ satisfies $m(x) \in C(\Omega)$ and $m(x) \geq m_0 > 0$ for $x \in \Omega$, also $\|m\|_\infty = l < \infty$. We prove the existence of a positive solution under certain conditions.

1. Introduction

In this paper, we consider the following reaction-diffusion equation

$$\begin{cases} -M\left(\int_\Omega |\nabla u|^p\,dx\right)\Delta_p u = am(x)u^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K & \text{in } \Omega,\\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where $\Delta_p u := \operatorname{div}(|\nabla u|^{p-2}\nabla u)$ is the p-Laplacian operator, $p > 1$, a, b, c, γ, K are positive constants with $\gamma \geq 2$, $M : \mathbb{R}_0^+ \to \mathbb{R}^+$ is a continuous and increasing function and Ω is a smooth bounded region with $\partial\Omega$ belonging to C^2. The weight function $m(x)$ satisfies $m(x) \in C(\Omega)$ and $m(x) \geq m_0 > 0$ for $x \in \Omega$, also $\|m\|_\infty = l < \infty$. We denote by λ_1 the first eigenvalue of

$$\begin{cases} -\Delta_p\phi = \lambda m(x)|\phi|^{p-2}\phi & x \in \Omega,\\ \phi = 0 & x \in \partial\Omega, \end{cases} \tag{2}$$

with positive principal eigenfunction ϕ_1 satisfying $\|\phi_1\|_\infty = 1$ (see [10]).
Here u is the population density and $am(x)u^{p-1} - bu^2$ represents logistics growth. This model describes grazing of a fixed number of grazers on a logistically growing species (see [12, 14]). The herbivore density is assumed to be a

*Talker
Email address: Saleh.Shakeri@iau.ac.ir (S. Shakeri)

constant which is a valid assumption for managed grazing systems and the rate of grazing is given by $\frac{cu^\gamma}{1+u^\gamma}$. At high levels of vegetation density this term saturates to c as the grazing population is a constant. This model has also been applied to describe the dynamics of fish populations (see [12]). The diffusive logistic equation with constant yield harvesting, in the absence of grazing was studied in [15]. Recently, in the case when $m(x) = 1$ and $p = 2$, problem (1) has been studied by Shivaji et al. (see [8]).
The purpose of this paper is to extend this study to the p-Laplacian case with the weight function . Our result in this note improves the previous one [4] in which $M(t) \equiv 1$. In recent years, problems involving Kirchhoff type operators have been studied in many papers, we refer to [1–3, 5, 6] in which the authors have used variational method and topological method to get the existence of solutions for (1).
We shall establish our existence results via the method of sub- and supersolutions. The concepts of sub- and supersolution were introduced by Nagumo [13] in 1937 who proved, using also the shooting method, the existence of at least one solution for a class of nonlinear Sturm-Liouville problems. In fact, the premises of the sub- and super-solution method can be traced back to Picard. He applied, in the early 1880s, the method of successive approximations to argue the existence of solutions for nonlinear elliptic equations that are suitable perturbations of uniquely solvable linear problems. This is the starting point of the use of sub- and supersolutions in connection with monotone methods. Picard's techniques were applied later by Poincaré [16] in connection with problems arising in astrophysics. We refer to [17].
Here and in what follows, $W_0^{1,p}(\Omega)$, $p > 1$, denotes the usual Sobolev space.

Definition 1.1. we say that ψ (*resp.* z) in $W_0^{1,p}(\Omega) \cap C(\overline{\Omega})$ is called a subsolution (*resp.* supersolution) of (1), if ψ (*resp.* z) satisfies

$$\begin{cases} M\left(\int_\Omega |\nabla\psi|^p\,dx\right)\int_\Omega |\nabla\psi|^{p-2}\nabla\psi.\nabla w dx \le \int_\Omega \left(am(x)\psi^{p-1} - b\psi^2 - c\frac{\psi^\gamma}{\psi^\gamma+1} - K\right)w dx, \\ \psi \le 0 \end{cases} \tag{3}$$

$$\left(\textit{resp.} \begin{cases} M\left(\int_\Omega |\nabla z|^p\,dx\right)\int_\Omega |\nabla z|^{p-2}\nabla z\nabla w dx \ge \int_\Omega \left(am(x)z^{p-1} - bz^2 - c\frac{z^\gamma}{z^\gamma+1} - K\right)w dx, \\ z \ge 0 \end{cases}\right) \tag{4}$$

for all non-negative test functions $w \in W_0^{1,p}(\Omega)$.

Then, the following lemma holds (see [7]).

Lemma 1.2 (see [7])**.** *If there exist sub-supersolutions ψ and z, respectively, such that $\psi \le z$ on Ω, then (1) has a positive solution u such that $\psi \le u \le z$ in Ω.*

Proposition 1.3. *If $a \le \frac{\lambda_1}{M_0}$, then (1) has no positive solution.*

Proof. Suppose not, i.e., assume that there exists a positive solution u of (1), then u satisfies

$$M\left(\int_\Omega |\nabla u|^p\,dx\right)\int_\Omega |\nabla u|^p dx \ge M_0\int_\Omega |\nabla u|^p dx = M_0\int_\Omega \left[am(x)u^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K\right]u dx.$$

But

$$M_0\int_\Omega |\nabla u|^p dx \ge \lambda_1\int_\Omega am(x)u^p dx.$$

Thus, we have

$$M_0\int_\Omega [am(x)u^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K]u dx \ge \lambda_1\int_\Omega am(x)u^p dx,$$

and hence

$$(a - \frac{\lambda_1}{M_0})\int_\Omega m(x)u^p dx \ge \int_\Omega \left[bu^2 + c\frac{u^\gamma}{u^\gamma+1} + K\right]u dx \ge 0.$$

Since $u > 0$, $m(x) \ge m_0 > 0$, this requires $a > \frac{\lambda_1}{M_0}$, which is a contradiction. Hence (1) has no positive solution. $\square$

2. Existence of solution

In this section we prove the existence of solution for problem (1) by comparison method (see [11]). It is easy to see that any subsolution of

$$\begin{cases} -M\left(\int_\Omega |\nabla u|^p\,dx\right)\Delta_p u = am_0u^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K & \text{in } \Omega, \\ u=0 & \text{on } \partial\Omega, \end{cases} \tag{5}$$

is a subsolution of (1). Also any supersolution of

$$\begin{cases} -M\left(\int_\Omega |\nabla u|^p\,dx\right)\Delta_p u = alu^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K & \text{in } \Omega, \\ u=0 & \text{on } \partial\Omega, \end{cases} \tag{6}$$

is a supersolution of (1), where l is defined as above.
We denote by λ_1' the first eigenvalue of

$$\begin{cases} -\Delta_p\phi = \lambda'|\phi|^{p-2}\phi & x\in\Omega, \\ \phi = 0 & x\in\partial\Omega, \end{cases} \tag{7}$$

with positive principal eigenfunction ϕ_1' satisfying $\|\phi_1'\|_\infty = 1$.

Theorem 2.1. *Assume that $M:\mathbb{R}_0^+\to\mathbb{R}^+$ is a continuous and increasing function satisfying*

$$M(t)\geq M_0 \textit{ for all } t\in\mathbb{R}_0^+, \tag{8}$$

where M_0 is a positive constant. If $a>\frac{\lambda_1' M_0}{m_0}, b>0, c>0$, then there exists a $K_0(a,b,c,m_0,\gamma)$ such that for $K<K_0$, (1) has a positive solution.

Proof. We use the method of sub-supersolutions. We recall the anti-maximum principle (see [9]) in the following form. Let λ' is defined as above, then there exists $\sigma(\Omega)>0$ and a solution $z_{\lambda'}$ of

$$\begin{cases} -\Delta_p z - \lambda' z^{p-1} = -1 & \text{in } \Omega, \\ z=0 & \text{on } \partial\Omega, \end{cases} \tag{9}$$

for $\lambda'\in(\lambda_1',\lambda_1'+\sigma)$, that $z_{\lambda'}$ is positive in Ω and is such that $\frac{\partial z_{\lambda'}}{\partial\nu}<0$ on $\partial\Omega$, where ν is outward normal vector at $\partial\Omega$. Fix

$$'\lambda_*'\in\left(\lambda_1', \min\{\lambda_1'+\sigma, m_0\alpha\}\right).$$

Let $z_{\lambda_*'}>0$ be the solution of (9) when $\lambda'=\lambda_*'$ and $\alpha=\|z_{\lambda_*'}\|_\infty$. Define

$$\psi := \mu K^{\frac{1}{p-1}} z_{\lambda_*'},$$

where $\mu\geq 1$ is to be determined later. We will choose μ and $K>0$ properly so that ψ is a subsolution. Let $w\in W_0^{1,p}(\Omega)$. Then

$$\begin{aligned} -M\left(\int_\Omega |\nabla\psi|^p\,dx\right)&\int_\Omega |\nabla\psi|^{p-2}\nabla\psi\nabla w dx + \int_\Omega \left[am_0\psi^{p-1} - b\psi^2 - c\frac{\psi^\gamma}{\psi^\gamma+1} - K\right]wdx \\ &\geq M_0\int_\Omega |\nabla\psi|^{p-2}\nabla\psi\nabla wdx + \int_\Omega\left[am_0\psi^{p-1} - b\psi^2 - c\frac{\psi^\gamma}{\psi^\gamma+1} - K\right]wdx \\ &= \int_\Omega\Big[-KM_0\mu^{p-1}(\lambda_*' z_{\lambda_*'}^{p-1}-1) + am_0K(\mu z_{\lambda_*'})^{p-1} - b(\mu K^{\frac{1}{p-1}}z_{\lambda_*'})^2 \\ &- c\frac{(\mu K^{\frac{1}{p-1}}z_{\lambda_*'})^\gamma}{(\mu K^{\frac{1}{p-1}}z_{\lambda_*'})^\gamma+1} - K\Big]wdx \\ &\geq \int_\Omega\Big[(am_0-\lambda_*'M_0)(\mu z_{\lambda_*'})^{p-1} - bK^{\frac{3-p}{p-1}}(\mu z_{\lambda_*'})^2 - c(\mu K574^{\frac{1}{p-1}}z_{\lambda_*'})^\gamma \\ &+ (\mu^{p-1}-1)\Big]Kwdx. \end{aligned}$$

Define

$$H(x) := (am_0 - \lambda'_* M_0)x^{p-1} - bK^{\frac{3-p}{p-1}}x^2 - cK^{\frac{\gamma-p+1}{p-1}}x^\gamma + (\mu^{p-1} - 1).$$

So, if we can find K and μ such that $H(x) \geq 0$ for all $x \in [0, \mu\alpha]$, then ψ will be a subsolution.
Notice that $H(0) = (\mu^{p-1} - 1) \geq 0$ since $\mu \geq 1$, and

$$H'(x) = x^{p-2}\left[(p-1)(am_0 - \lambda'_* M_0) - 2bK^{\frac{3-p}{p-1}}x^{3-p} - c\gamma K^{\frac{\gamma-p+1}{p-1}}x^{p-\gamma+3}\right].$$

This means that $H(x) \geq 0$ if $H(\mu\alpha) \geq 0$, i.e.,

$$(am_0 - \lambda'_* M_0)(\mu\alpha)^{p-1} - bK^{\frac{3-p}{p-1}}(\mu\alpha)^2 - cK^{\frac{\gamma-p+1}{p-1}}(\mu\alpha)^\gamma + \mu^{p-1} - 1 \geq 0.$$

Let

$$G(K) := (am_0 - \lambda'_* M_0)(\mu\alpha)^{p-1} - bK^{\frac{3-p}{p-1}}(\mu\alpha)^2 - cK^{\frac{\gamma-p+1}{p-1}}(\mu\alpha)^\gamma + \mu^{p-1} - 1.$$

Then $G(0) = (am_0 - \lambda'_* M_0)(\mu\alpha)^{p-1} + (\mu^{p-1} - 1) > 0$, since $\mu \geq 1$ and $am_0 > \lambda'_*$. Also we have

$$G'(K) = -b\Big(\frac{3-p}{p-1}\Big)(\mu\alpha)^2 K^{\frac{4-2p}{p-1}} - c\Big(\frac{\gamma-p+1}{p-1}\Big)K^{\frac{\gamma-2p+2}{p-1}}(\mu\alpha)^\gamma < 0.$$

Hence given μ and γ there exists a unique $K^* = K^*(a, b, c, \mu, \gamma, m_0) > 0$ with $G(K^*) = 0$. Since $G(K) \leq (am_0 - \lambda'_* M_0)(\mu\alpha)^{p-1} - b(\mu\alpha)^2 K^{\frac{3-p}{p-1}} + (\mu^{p-1} - 1) = \widetilde{G}(K)$, we see that

$$K^* \leq \left[\frac{(am_0 - \lambda'_* M_0)(\mu\alpha)^{p-1} + (\mu^{p-1} - 1)}{b\mu^2\alpha^2}\right]^{\frac{3-p}{p-1}} := K_1(a, b, \mu, m_0).$$

Note that $K_1(a, b, \mu, m_0)$ is bounded for $\mu \in [1, \infty)$. Hence K^* is bounded for $\mu \in [1, \infty)$. Let $K_0(a, b, c, m_0, \gamma) = \sup_{\mu\geq 1} K^*(a, b, c, \mu, m_0, \gamma) > 0$. Now let $\widetilde{K} < K_0(a, b, c, m_0, \gamma)$. By definition there will exist a $\widetilde{\mu} \geq 1$ such that $\widetilde{K} < K^*(a, b, c, \widetilde{\mu}, m_0, \gamma) < K_0(a, b, c, m_0, \gamma)$. Choose $\psi = \widetilde{\mu}\widetilde{K}^{\frac{1}{p-1}} z_{\lambda'_*}$ with $\mu = \widetilde{\mu}$. We have $G(\widetilde{K}) \geq 0$ and

$$(am_0 - \lambda'_*)(\widetilde{\mu}\alpha)^{p-1} - b\widetilde{K}^{\frac{3-p}{p-1}}(\widetilde{\mu}\alpha)^2 - c\widetilde{K}^{\frac{\gamma-p+1}{p-1}}(\widetilde{\mu}\alpha)^\gamma + (\widetilde{\mu}^{p-1} - 1) \geq 0.$$

Hence ψ turns out to be a sub-solution to (1).
We next construct the supersolution z for (1) such that $z \geq \psi$. Let $z = \overline{M}e_p$, where $\overline{M} > 0$ is such that $\frac{alu^{p-1} - bu^2 - c\frac{u^\gamma}{u^\gamma+1} - K}{M_0} \leq \overline{M}$ for all $u \geq 0$ and e_p is the unique positive solution of

$$\begin{cases} -\Delta_p e_p = 1 & \text{in } \Omega, \\ e_p = 0 & \text{on } \partial\Omega. \end{cases}$$

Then for $w \in W_0^{1,p}(\Omega)$,

$$\begin{aligned} M\left(\int_\Omega |\nabla z|^p\, dx\right)\int_\Omega |\nabla z|^{p-2}\nabla z \nabla w dx &= M\left(\int_\Omega |\nabla z|^p\, dx\right)\int_\Omega \overline{M} w dx \\ &\geq \int_\Omega \Big[alz^{p-1} - bz^2 - c\frac{z^\gamma}{z^\gamma + 1} - K\Big] w dx. \end{aligned}$$

Thus z is a supersolution of (6). Therefore z is a supersolution of (1), since we can choose $\overline{M} \gg 1$ so that $z \geq \psi$. Hence, by Lemma 1.2, problem (1) has a positive solution for all $K < K_0(a, b, c, m_0, \gamma)$. □

Acknowledgments

The authors are extremely grateful to the referees for their helpful suggestions for the improvement of the paper.

References
References

[1] G.A. Afrouzi, N.T. Chung and S. Shakeri, Positive solutions for a infinite semipositone problem involving nonlocal operator, *Rend. Semin. Mat. Univ. Padova*, **Vol. 2014**(132) (2014), 25-32.

[2] A. Bensedik and M. Bouchekif, On an elliptic equation of Kirchhoff-type with a potential asymptotically linear at infinity, *Math. Comput. Modelling*, **49** (2009), 1089-1096.

[3] F.J.S.A. Corrêa and G.M. Figueiredo, On an elliptic equation of p-Kirchhoff type via variational methods, *Bull. Aust. Math. Soc.*, **74** (2006), 263-277.

[4] G.A.Afrouzi ,S. Shakeri and A. Hadjian, On the existence of positive solutions for an ecological model with indefinite weight, *Arab Journal of Mathematical Sciences.*, **22** (2016), 132-137.

[5] J.J. Sun and C.L. Tang, Existence and multiplicity of solutions for Kirchhoff type equations, *Nonlinear Anal.*, **74** (2011), 1212-1222.

[6] M.H. Yang and Z.Q. Han, Existence and multiplicity results for Kirchhoff type problems with four-superlinear potentials, *Appl. Anal.*, **91**(11) (2012), 2045-2055.

[7] H. Amann, Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces, *SIAM Rev.*, **18** (1976), 620-709.

[8] R. Causey, S. Sasi and R. Shivaji, An ecological model with grazing and constant yield harvesting, *Bull. Belg. Math. Soc. Simon Stevin*, **17** (2010), 833-839.

[9] Ph. Clément and L.A. Peletier, An anti-maximum principle for second-order elliptic operators, *J. Differential Equations*, **34** (1979), 218-229.

[10] P. Drábek and J. Hernández, Existence and uniqueness of positive solutions for some quasilinear elliptic problems, *Nonlinear Anal.*, **44** (2001), 189-204.

[11] P. Drábek, P. Kerjcǐ and P. Takáč, *Nonlinear Differential Equations*, CRC Research Notes Math. 404, Chapman & Hall/CRC, Boca Raton, FL-London-New York-Washington, DC, 1999.

[12] R.M. May, Thresholds and breakpoints in ecosystems with a multiplicity of stable states, *Nature*, **269** (1977), 471-477.

[13] M. Nagumo, Über die Differentialgleichung $y'' = f(x, y, y')$, *Proceedings of the Physico-Mathematical Society of Japan* **19** (1937), 861-866.

[14] I. Noy-Meir, Stability of grazing systems: An application of predator-prey graphs, *J. Ecol.*, **63** (1975), 459-481.

[15] S. Oruganti, J. Shi and R. Shivaji, Diffusive logistic equation with constant yield harvesting, I: steady states, *Trans. Amer. Math. Soc.*, **354** (2002), 3601-3619.

[16] H. Poincaré, Les fonctions fuchsiennes et l'équation $\Delta u = e^u$, *J. Math. Pures Appl.* **4** (1898), 137-230.

[17] V. Rădulescu, *Qualitative Analysis of Nonlinear Elliptic Partial Differential Equations: Monotonicity, Analytic, and Variational Methods*, Contemporary Mathematics and Its Applications, vol. 6, Hindawi Publishing Corporation, New York, 2008.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–114, pp. 22–34

On a class of operators and its applications for existence of solution a system of integral equations

Shahram Banaei*

[a]Department of Mathematics, Bonab Branch, Islamic Azad University, Bonab, Iran.

Article Info

Keywords:
Measure of noncompactness
Integral equations
Fixed point theorem

2020 MSC:
47H08
47H10

Abstract

In this article, we extend Darbo's fixed point theorem via the concept of the class of operators $A(f;.)$ in Banach space and prove some Prešić type fixed point theorems. We apply the technique of measure of noncompactness in the presentation of our proofs. We handle our obtained results to inquiry existence of solution for a system of integral equations.

1. Introduction and preliminaries

Fixed point theory is an essential tool applied in analysis to investigate the solvability for a system of integral equations. The notion of measure of noncompactness, for shortly (MNC), was introduced by Kuratowski [10] which plays a fundamental role in the study of a system of integral equations. Darbo's fixed point theorem [8] is an important application of this measure, which it generalizes both Banach contraction principle and Schauder fixed point theorem. Up to now, many authors and researchers such as in [4, 5, 9, 12, 14] studied solvability of integral and differential equations. The aim of this paper is to present some Darbo type fixed point theorems associated with a (MNC) via the concept of operators $A(f;.)$. Moreover, in order to show the applicability of our results, we discuss the existence of solution for a system of integral equations. Now, we remember some concepts and definitions that are used in this article.

Suppose $\mathcal{E}$ is a real Banach space with norm $||.||$ and X be a nonempty subset of $\mathcal{E}$. Let $\overline{X}$ and $Conv(X)$ the closure and the closed convex hull of X, respectively. We mark by $\mathcal{M}(\mathcal{E})$ the family of all nonempty and bounded subsets of $\mathcal{E}$ and $\mathcal{N}(\mathcal{E})$ be the collection of all relatively compact subsets of $\mathcal{E}$. Moreover, let $\mathcal{R}$ indicates the set of all real numbers and $\mathcal{R}^+ = [0, +\infty)$. In addition, let $\overline{B}(\epsilon, r)$ be the closed ball with center ϵ and radius r. Also, let $\overline{B}_r$ denotes the ball $\overline{B}(0, r)$.

Definition 1.1. [6] We say that a mapping $\alpha : \mathcal{M}(\mathcal{E}) \longrightarrow \mathcal{R}^+$ is a (MNC) in the Banach space $\mathcal{E}$ if:

1° The family $ker\alpha = \{X \in \mathcal{M}(\mathcal{E}) : \alpha(X) = 0\}$ is nonempty and $ker\alpha \subset \mathcal{N}(\mathcal{E})$;

*Talker
Email address: math.sh.banaei@gmail.com (Shahram Banaei)

2° $X \subset Y \Longrightarrow \alpha(X) \leq \alpha(Y)$;

3° $\alpha(\overline{X}) = \alpha(X)$;

4° $\alpha(ConvX) = \alpha(X)$;

5° $\alpha(\lambda X + (1-\lambda)Y) \leq \lambda\alpha(X) + (1-\lambda)\alpha(Y)$ for all $\lambda \in [0, 1]$;

6° If $\{X_n\}$ is a sequence of closed sets from $\mathcal{M}(\mathcal{E})$ such that $X_{n+1} \subset X_n$ for $n = 1, 2, \cdots$, and if $\lim\limits_{n\to\infty} \alpha(X_n) = 0$, then $X_\infty = \cap_{n=1}^{\infty} X_n \neq \emptyset$.

Wardowski [15] introduced a significant extension of the Banach contraction principle. He presented a new class of control functions $\mathcal{K}$ which provide many contractions. Similar [11, 15], Σ indicates the set of all functions $V : (0, \infty) \to (1, \infty)$ such that
V_1. V is increasing and continuous,
V_2. $\lim\limits_{n\to\infty} \alpha_n = 0$ if and only if $\lim\limits_{n\to\infty} V(\alpha_n) = 1$, for all sequence $\{\alpha_n\} \subseteq (0, \infty)$.
Suppose Φ be set of all functions $\varphi : (1, \infty) \to (1, \infty)$
so that

φ_1. φ is a continuous and increasing function;

φ_2. $\lim\limits_{n\to\infty} \varphi^n(t) = 1$ for all $t \in (1, \infty)$.

The following concept of $A(f; .)$ was given by Altun and Turkoglu [3].
Let $F([0, \infty))$ be the class of all functions $f : [0, \infty) \to [0, \infty)$ and let $\mathcal{F}$ be the class of all operators
$A(\bullet; \cdot) : F([0, \infty)) \to F([0, \infty)), \ \ f \to A(f; \cdot)$
satisfying the following conditions:
(1) $A(f; p) > 0$ for $p > 0$ and $A(f; 0) = 0$.
(2) $A(f; p) \leq A(f; q)$ for $p \leq q$.
(3) $\lim_{n\to\infty} A(f; p_n) = A(f; \lim_{n\to\infty} p_n)$.
(4) $A(f; max\{p, q\}) = max\{A(f; p), A(f; q)\}$ for some $f \in F([0, \infty))$.

Now we remember some theorems used in the schedule.

Theorem 1.2. $(Schauder)$ *[2] Let C be a nonempty, bounded, closed and convex subset of a Banach space $\mathcal{E}$. Then each continuous and compact mapping $T : C \to C$ has at least one fixed point in the set C.*

Theorem 1.3. *[8] Let C be a nonempty, bounded, closed and convex subset of a Banach space $\mathcal{E}$ and let $T : C \to C$ be a continuous mapping. Assume that there exists a constant $\kappa \in [0, 1)$ such that $\mu(T\Lambda) \leq \kappa\mu(\Lambda)$ for any nonempty subset Λ of C, where μ is a MNC defined in $\mathcal{E}$. Then T has at least a fixed point in C.*

Theorem 1.4. *[1] Suppose that $\alpha_1, \alpha_2, \ldots, \alpha_n$ are (MNC) in Banach spaces $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$, respectively, the function $\omega : [0, \infty)^n \longrightarrow [0, \infty)$ is a convex function and $\omega(p_1, \cdots, p_n) = 0$ if and only if $p_i = 0$ for all $i = 1, 2, \cdots, n$. Then*

$$\tilde{\alpha}(X) = \omega(\alpha_1(X_1), \alpha_2(X_2), \ldots, \alpha_n(X_n)),$$

is a (MNC) in $\mathcal{E}_1 \times \mathcal{E}_2 \times \ldots \times \mathcal{E}_n$, where X_i denote the natural projection of X into $\mathcal{E}_i$, for all $i = 1, 2, \ldots, n$.

2. Main Results

In this section, we want to extend the Darbo's fixed point theorem [8] by applying the concept of the class of operators $A(f; .)$.

Theorem 2.1. *Let C be a nonempty, closed, bounded and convex $(NCBC)$ subset of a Banach space $\mathcal{E}$ and let $\Upsilon : C \to C$ be a continuous mapping such that*

$$A\big(f;\ \alpha(\Upsilon X)\big) \leq \mathcal{V}\big(A(f;\ \alpha(X))\big) - \mathcal{V}\big(A(f;\ \alpha(\Upsilon X))\big), \tag{1}$$

for all $X \subseteq C$, where $\mathcal{V} : [0, \infty) \to [0, 1]$ is such that $\lim_{u \to 0^+} \mathcal{V}(u) = 0$, $\mathcal{V}(0) = 0$, $A(\bullet; \cdot) \in \mathcal{F}$ and α is an arbitrary MNC. Then Υ has at least one fixed point.

Proof. Define a sequence $\{C_n\}$ as follows,
$C_0 = C$ and $C_{n+1} = \overline{Conv}(\Upsilon(C_n))$ for all $n \in \mathbb{N}$.
If there exists an integer $n \in \mathbb{N}$ such that $\alpha(C_n) = 0$, then C_n is relatively compact and by applying Theorem 1.2, we get Υ has a fixed point. So, we can assume that $\alpha(C_n) > 0$ for each $n \in \mathbb{N}$.
Obviously $\{C_n\}_{n \in \mathbb{N}}$ is a $(NCBC)$ sequence such that

$$C_0 \supseteq C_1 \supseteq \cdots \supseteq C_n \supseteq C_{n+1}.$$

Therefore, the sequence $\{\mathcal{V}(\alpha(C_n))\}$ is decreasing. On the other hand $\mathcal{V}$ is bounded below. Thus, there exists $r \in \mathbb{R}^+$ such that $\lim_{n \to \infty} \mathcal{V}(\alpha(C_n)) = r$. Moreover, we get,

$$\begin{aligned} A(f;\ \alpha(C_{n+1})) &= A(f;\ \alpha(\Upsilon C_n)) \\ &\leq \mathcal{V}(A(f;\ \alpha(C_n))) - \mathcal{V}(A(f;\ \alpha(\Upsilon C_n))) \\ &= \mathcal{V}(A(f;\ \alpha(C_n))) - \mathcal{V}(A(f;\ \alpha(C_{n+1}))). \end{aligned} \tag{2}$$

Using the property of $A(\bullet; \cdot)$ and taking $limsup$ in inequality 2, we get

$$\limsup_{n \to \infty} A\big(f;\ \alpha(C_{n+1})\big) \leq \limsup_{n \to \infty} \mathcal{V}\Big(A\big(f;\ \alpha(C_n)\big)\Big) - \liminf_{n \to \infty} \mathcal{V}\Big(A\big(f;\ \alpha(C_{n+1})\big)\Big).$$

Thus,

$$\lim_{n \to \infty} A\big(f;\ \alpha(C_{n+1})\big) = 0$$

In view of properties $A(\bullet; \cdot)$, we have

$$A\big(f;\ \lim_{n \to \infty} \alpha(C_{n+1})\big) = 0.$$

Property (1) of $A(\bullet; \cdot)$ implies that

$$\lim_{n \to \infty} \alpha(C_n) = 0.$$

we can deduce that $\alpha(C_n) \to 0$ as $n \to \infty$. Since C_n is a nested sequence, in view of (6°) of Definition 1.1 we infer that the set $C_\infty = \bigcap_{n=1}^{\infty} C_n$ is a nonempty, closed and convex set and belongs to $ker\alpha$. Then, according the Schauder fixed point theorem we deduce that Υ has a fixed point. □

Remark 2.2. Theorem 2.1 is an extension of Darbo's fixed point theorem.

Suppose $\Upsilon : X \to X$ be a Darbo mapping, then there exists a constant $k \in [0, 1)$ such that

$$\alpha(\Upsilon X) \leq k\alpha(X)$$

for all $X \subseteq C$. Thus,

$$\alpha(\Upsilon X) \leq k\alpha(X) \leq \frac{k}{1 - \sqrt{k} + k}\alpha(X)$$

Therefore,

$$k\alpha(\Upsilon X) + (1 - \sqrt{k})\alpha(\Upsilon X) \leq k\alpha(X),$$

and

$$(1-\sqrt{k})\alpha(\Upsilon X) \leq k\alpha(X) - k\alpha(\Upsilon X)$$

Thus,

$$\alpha(\Upsilon X) \leq \frac{k}{1-\sqrt{k}}\alpha(X) - \frac{k}{1-\sqrt{k}}\alpha(\Upsilon X)$$

and

$$A\Big(f;\ \alpha(\Upsilon X)\Big) \leq \frac{k}{1-\sqrt{k}}A\Big(f;\ \alpha(X)\Big) - \frac{k}{1-\sqrt{k}}A\Big(f;\ \alpha(\Upsilon X)\Big)$$

By taking $\mathcal{V}(x) = \frac{k}{1-\sqrt{k}}x$, we get

$$A\Big(f;\ \alpha(\Upsilon X)\Big) \leq \mathcal{V}\bigg(A\Big(f;\ \alpha(X)\Big)\bigg) - \mathcal{V}\bigg(A\Big(f;\ \alpha(\Upsilon X)\Big)\bigg)$$

Therefore, the inequality of Darbo's fixed point Theorem [8] is a special case of Theorem 2.1.

Definition 2.3. [7] We say that $(p, q, r) \in \mathcal{E}^3$ is a tripled fixed point of a mapping $\Upsilon : \mathcal{E}\times\mathcal{E}\times\mathcal{E} \to \mathcal{E}$ if $\Upsilon(p, q, r) = p$, $\Upsilon(q, p, r) = q$ and $\Upsilon(r, q, p) = r$

Theorem 2.4. *Let C be a nonempty, closed, bounded and convex $(NCBC)$ subset of a Banach space $\mathcal{E}$ and let $\Upsilon : C \times C \times C \to C$ be a continuous mapping such that*

$$\begin{aligned} A\big(f;\ \alpha\big(\Upsilon(X_1\times X_2\times X_3)\big)\big) \leq \frac{1}{3}[\mathcal{V}(\big(A(f;\ \alpha(X_1)+\alpha(X_2)+\alpha(X_3))\big))] \\ -\mathcal{V}\Big(A(f;\ \alpha(\Upsilon(X_1\times X_2\times X_3)))\Big) \end{aligned} \tag{3}$$

for any X_1, X_2 and X_3 of C, where $\mathcal{V}$ is a subadditive mapping, $A(\bullet;\cdot) \in \mathcal{F}$, α and $\mathcal{V}$ are as in Theorem 2.1. Then Υ has tripled fixed point.

Proof. Define the mapping $\widetilde{\Upsilon} : C^3 \to C^3$ by

$$\widetilde{\Upsilon}(p, q, r) = (\Upsilon(p, q, r), \Upsilon(q, p, r), \Upsilon(r, p, q)).$$

Obviously, $\widetilde{\Upsilon}$ is continuous. We prove that $\widetilde{\Upsilon}$ satisfies all the conditions of Theorem 2.1. Let $X \subset C^3$ be a nonempty subset. We know that $\widetilde{\alpha}(X) = \alpha(X_1) + \alpha(X_2) + \alpha(X_3)$ is a (MNC) [1] , where X_1, X_2 and X_3 denote the natural projections of X into $\mathcal{E}$. From (5) we get

$$
\begin{aligned}
\mathcal{V}\big(A(f;\ \widetilde{\alpha}(\widetilde{\Upsilon}(X)))\big) &= \mathcal{V}\big(A(f;\ \widetilde{\alpha}(\Upsilon(X_1 \times X_2 \times X_3) \times \Upsilon(X_2 \times X_1 \times X_3) \\
&\qquad\qquad \times \Upsilon(X_3 \times X_1 \times X_2)\big) \\
&= \mathcal{V}\big(A(f;\ \alpha(\Upsilon(X_1 \times X_2 \times X_3))) + \alpha(\Upsilon(X_2 \times X_1 \times X_3)) \\
&\qquad\qquad\qquad + \alpha(\Upsilon(X_3 \times X_1 \times X_2))\big) \\
&\leq \frac{1}{3}[\mathcal{V}(\big(A(f;\ \alpha(X_1) + \alpha(X_2) + \alpha(X_3))))] \\
&\qquad\qquad - \mathcal{V}\big(A(f;\ \alpha(\Upsilon X_1 \times X_2 \times X_3))\big) \\
&\quad + \frac{1}{3}[\mathcal{V}(\big(A(f;\ \alpha(X_2) + \alpha(X_1) + \alpha(X_3))))] \\
&\qquad\qquad - \mathcal{V}\big(A(f;\ \alpha(\Upsilon X_2 \times X_1 \times X_3))\big) \\
&\quad + \frac{1}{3}[\mathcal{V}(\big(A(f;\ \alpha(X_3) + \alpha(X_2) + \alpha(X_2))))] \\
&\qquad\qquad - \mathcal{V}\big(A(f;\ \alpha(\Upsilon X_3 \times X_1 \times X_2))\big) \\
&\leq \mathcal{V}(\big(A(f;\ \alpha(X_1) + \alpha(X_2) + \alpha(X_3)))\big) \\
&\qquad - \mathcal{V}\big(A(f;\ \widetilde{\alpha}(\Upsilon(X_1 \times X_2 \times X_3)))\big) \\
&= \mathcal{V}(A(f;\ \widetilde{\alpha}(X))) - \mathcal{V}\big(A(f;\ \widetilde{\alpha}(\widetilde{\Upsilon}(X)))\big). \qquad (4)
\end{aligned}
$$

From Theorem 2.1 we infer that $\widetilde{\Upsilon}$ has at least a fixed point which implies that Υ has a tripled fixed point. □

By taking $A(f;\ p) = p$, $f = I$ (identity map) and $\mathcal{V} = 1$ in Theorem 2.4, we infer the following corollary.

Corollary 2.5. *Let C be a nonempty, closed, bounded and convex $(NCBC)$ subset of a Banach space $\mathcal{E}$ and let $\Upsilon : C \times C \times C \to C$ be a continuous mapping such that*

$$\alpha\Big(\Upsilon(X_1 \times X_2 \times X_3)\Big) \leq \frac{1}{6}\Big(\alpha(X_1) + \alpha(X_2) + \alpha(X_3)\Big)$$

for any X_1, X_2 and X_3 of C, where α is an arbitrary MNC . Then Υ has tripled fixed point.

Theorem 2.6. *Let C be a $(NCBC)$ subset of a Banach space $\mathcal{E}$ and let $\Upsilon : C \to C$ be a continuous mapping such that*

$$V\big(A(f;\ \alpha(\Upsilon X))\big) \leq \varphi\big(V(A(f;\ \alpha(X)))\big), \qquad (5)$$

for all $X \subseteq C$, where $V \in \Sigma$, $\varphi \in \Phi$, $A(\bullet;\cdot) \in \mathcal{F}$ and α is an arbitrary MNC. Then Υ has at least one fixed point.

Proof. Let sequence $\{C_n\}$ be such that
$C_0 = C$ and $C_{n+1} = \overline{Conv}(\Upsilon(C_n))$ for all $n \in \mathbb{N}$.
If for an integer $n \in \mathbb{N}$ one has $\alpha(C_n) = 0$, then C_n is relatively compact and so Schauder Theorem 1.2 guarantees a fixed point for Υ. So, we can assume that $\alpha(C_n) > 0$ for each $n \in \mathbb{N}$.
Evidently, $\{C_n\}_{n\in\mathbb{N}}$ is a $(NCBC)$ sequence such that

$$C_0 \supseteq C_1 \supseteq \cdots \supseteq C_n \supseteq C_{n+1}.$$

On the other hand

$$
\begin{aligned}
V\big(A(f;\ \alpha(C_{n+1}))\big) &= V\big(A(f;\ \alpha(\Upsilon C_n))\big) \\
&\leq \varphi\big(V(A(f;\ \alpha(C_n)))\big) \\
&\ \vdots \\
&\leq \varphi^{n+1}\big(V(A(f;\ \alpha(C_0)))\big). \qquad (6)
\end{aligned}
$$

Therefore, $\alpha(C_{n+1})$ is a convergent sequence. Suppose that $\lim_{n\to\infty} \alpha(C_{n+1}) = r$. Now, we show that $r = 0$.
Using the property of $A(\bullet;\cdot)$ and Passing limit through in inequality 6, we get

$$\lim_{n\to\infty} V\big(A(f;\ \alpha(C_{n+1}))\big) = 1.$$

Thus, we have

$$\lim_{n\to\infty} \big(A(f;\ \alpha(C_{n+1}))\big) = 0.$$

In view of properties $A(\bullet;\cdot)$, we have

$$A\big(f;\ \lim_{n\to\infty} \alpha(C_{n+1})\big) = 0.$$

Property (i) of $A(\bullet;\cdot)$ implies that

$$\lim_{n\to\infty} \alpha(C_n) = 0.$$

we can deduce that $\alpha(C_n) \to 0$ as $n \to \infty$. Since C_n is a nested sequence, in view of $(6°)$ of Definition 1.1 we infer that the set $C_\infty = \bigcap_{n=1}^{\infty} C_n$ is a nonempty, closed and convex set and belongs to $ker\mu$. Then, according the Schauder fixed point theorem we deduce that Υ has a fixed point. □

Taking $A(f;p) = p,\ \ f = I$ (identity map) and $\varphi(x) = x^k\ k \in (0,1)$ in Theorem 2.6 we get the following Corollary.

Corollary 2.7. *Let C be a nonempty, bounded, closed and convex subset of a Banach space $\mathcal{E}$ and let $\Upsilon : C \to C$ be a continuous operator such that*

$$V\big(\alpha(\Upsilon X)\big) \le V\big(\alpha(X)\big)^k \tag{7}$$

for all $X \subseteq C$, where $V \in \Sigma$ and α is an arbitrary MNC. Then Υ has at least one fixed point in C.

Remark 2.8. If we take $V(x) = e^x$ in the above Corollary, then we get the Darbo's fixed point theorem.

3. Prešić type fixed point

In this part, we state and prove Prešić type fixed point theorem.
Prešić in [13] presented the following theorem as an extension of Banach contraction principle.

Theorem 3.1. *Suppose (X,d) be a complete metric space and let $\Upsilon : X^k \to X$ (k is a positive integer). Let,*

$$d\big(\Upsilon(\xi_1,\dots,\xi_k),\Upsilon(\xi_2,\dots,\xi_{k+1})\big) \le \sum_{i=1}^{k} \mu_i d(\xi_i,\xi_{i+1})$$

for all $\xi_1,\dots,\xi_{k+1}$ in X, where $\mu_i \ge 0$ and $\sum_{i=1}^{k} \mu_i \in [0,1)$. Then, Υ has a unique fixed point ξ^* (that is $\Upsilon(\xi^*,\dots,\xi^*) = \xi^*$ and now we say Prešić type fixed point).

Theorem 3.2. *Let $C \subseteq \mathcal{E}$ be a nonempty,closed, bounded and convex subset and let $\Upsilon : C^n \to C$ be a continuous function such that*

$$V\big(A(f;\ \alpha\Upsilon(X_1\times\ldots\times X_n))\big) \le \frac{1}{n}\varphi\big([V(A(f;\ \alpha(X_1)+\ldots+\alpha(X_n)))]\big) \tag{8}$$

for all $X_1,\dots,X_n \subseteq C$, where $V \in \Sigma$ is a subadditive mapping, $\varphi \in \Phi$ and α is an arbitrary MNC, then Υ has a Prešić type fixed point.

Proof. Define the mapping $\widetilde{\Upsilon} : C^n \to C^n$ by

$$\widetilde{\Upsilon}(\xi_1, \dots, \xi_n) = \big(\Upsilon(\xi_1, \dots, \xi_n), \dots (\xi_1, \dots, \xi_n)\big).$$

Obviously, $\widetilde{\Upsilon}$ is continuous. We prove that $\widetilde{\Upsilon}$ satisfies all the conditions of Theorem 2.6. Suppose $X \subset C^n$ be a nonempty subset. We know that $\widetilde{\alpha}(X) = \alpha(X_1) + \alpha(X_2) + \dots + \alpha(X_n)$ is a (MNC) [1] , where $X_1, X_2 \dots X_n$ denote the natural projections of X into $\mathcal{E}$. From (8) we get

$$\begin{aligned} V\big(A(f;\ \widetilde{\alpha}(\widetilde{\Upsilon}(X)))\big) &= V\big(A(f;\ \widetilde{\alpha}(\Upsilon(X_1 \times \dots \times X_n) \times \dots \times \Upsilon(X_1 \times \dots \times X_n)\big) \\ &= V\big(A(f;\ n\alpha(\Upsilon(X_1 \times \dots \times X_n)\big) \\ &\leq n\, V\big(A(f;\ \alpha(\Upsilon(X_1 \times \dots \times X_n)\big) \\ &\leq \varphi\big([V(A(f;\ \alpha(X_1) + \dots + \alpha(X_n)))]\big) \\ &= \varphi\big(V(A(f;\ \widetilde{\alpha}(X)))\big). \end{aligned} \tag{9}$$

From Theorem 2.6 we infer that $\widetilde{\Upsilon}$ has at least a fixed point which implies that there exists $\xi_1, \dots, \xi_n$ such that $\Upsilon(\xi_1, \dots, \xi_n) = \xi_1 = \dots = \xi_n$ that is, Υ has a Prešić type fixed point. $\square$

4. Application

In this section of the paper we investigate the existence of solutions for the following system of equations:

$$\begin{cases} u_1(x) = h(x, u_1(\varrho(x)), u_2(\varrho(x)), u_3(\varrho(x)), \\ \qquad\qquad \displaystyle\int_0^{\rho(x)} k\,(x, y, u_1(\varrho(y)), u_2(\varrho(y)), u_3(\varrho(y)))dy) \\ u_2(x) = h(x, u_2(\varrho(x)), u_1(\varrho(x)), u_3(\varrho(x)), \\ \qquad\qquad \displaystyle\int_0^{\rho(x)} k\,(x, y, u_2(\varrho(y)), u_1(\varrho(y)), u_3(\varrho(y)))\,dy) \\ u_3(x) = h(x, u_3(\varrho(x)), u_2(\varrho(x)), u_1(\varrho(x)), \\ \qquad\qquad \displaystyle\int_0^{\rho(x)} k\,(x, y, u_3(\varrho(y)), u_2(\varrho(y)), u_1(\varrho(y)))\,dy.) \end{cases} \tag{10}$$

where $x \in [0, T]$.
Suppose $C[0, T]$ be the space of all real functions which are bounded and continuous on the interval $[0, T]$ with the norm

$$\|p\| = \sup\{|p(r)| : r \in [0, T]\}.$$

The modulus of continuity of a function $p \in C[0, T]$ is as

$$\omega(p, \epsilon) = \sup\{|p(r) - p(s)| : r, s \in [0, T], |r - s| \leq \epsilon\}.$$

Let $\omega(X, \varepsilon) = \sup\{\omega(p, \varepsilon) : p \in X\}$. The Hausdorff (MNC) for all bounded sets X of $C[0, T]$ is as follows:

$$\mu(X) = \omega(X) = \lim_{\epsilon \to 0} \Big\{ \sup_{p \in X} \omega(p, \epsilon) \Big\}.$$

(For more details see [2]).

Theorem 4.1. *Suppose that the following conditions are satisfies.*

(i) *Let* $\varrho, \rho : [0,T] \longrightarrow [0,T]$ *are continuous functions.*

(ii) *The function* $h : [0,T] \times \mathbb{R}^4 \longrightarrow \mathbb{R}$ *is continuous and there exists function* γ *so that*

$$A\Big(f;\ h\big(x, u_1, u_2, u_3, p\big) - h\big(x, v_1, v_2, v_3, q\big)\Big) \leq \frac{1}{6}\gamma\Big(A(f;\ \Big|u_1 - v_1\Big| + \Big|u_2 - v_2\Big| + \Big|u_3 - v_3\Big|\Big) + \Big|p - q\Big|, \tag{11}$$

where $\gamma : [0,\infty) \to [0,\infty)$ *is a subadditive increasing mapping such that* $\lim_{n\to\infty} \gamma^n(t) = 0$ *and* $\gamma(0) = 0$*, for* $t > 0$.

(iii) $M := \sup\{A(f;\ \Big|h(x,0,0,0,0)\Big| < \infty,\ x \in [0,T]\}$ *and* $A(f,\epsilon) < \epsilon$.

(iv) $k : [0,T] \times [0,T] \times \mathbb{R}^3 \longrightarrow \mathbb{R}$ *is continuous and*
$\Gamma := \sup\Big\{ \int_0^{\rho(x)} k(x,y,u_1,u_2,u_3)dy\ : x, y \in [0,T], u_1, u_2, u_3 \in C([0,T])\Big\}.$

(v) *The inequality*

$$\frac{1}{6}\Big(A(f;\ 3r_0)\Big) + \Gamma + M \leq r_0$$

has a positive solution r_0.

Then the system of integral equations (10) *has at least one solution in the space* $(C[0,T])^3$.

Proof. Suppose that $\Upsilon : C[0,T] \times C[0,T] \times C[0,T] \longrightarrow C[0,T]$ defined by

$$\Upsilon(u_1,u_2,u_3)(x) = h\big(x, u_1(\varrho(x)), u_2(\varrho(x)), u_3(\varrho(x)) \\ , \int_0^{\rho(x)} k\left(x, y, u_1(\varrho(y)), u_2(\varrho(y)), u_3(\varrho(y))\right) dy\big).$$

In view of given assumptions, we deduce that the function $\Upsilon(u_1,u_2,u_3)$ is continuous for arbitrarily $u_1, u_2, u_3 \in C[0,T]$. Also, we obtain that

$$\begin{aligned}
&A\Big(f;\ \Upsilon(u_1,u_2,u_3)(x)\Big) \\
&= A\Big(f;\ h\big(x, u_1(\varrho(x)), u_2(\varrho(x)), u_3(\varrho(x)) \\
&\qquad\qquad , \int_0^{\rho(x)} k\left(x, y, u_1(\varrho(y)), u_2(\varrho(y)), u_3(\varrho(y))\right) dy\big)\Big) \\
&\leq A\Big(f;\ h\big(x, u_1(\varrho(x)), u_2(\varrho(x)), u_3(\varrho(x)), \\
&\quad \int_0^{\rho(x)} k\left(x, y, u_1(\varrho(y)), u_2(\varrho(y)), u_3(\varrho(y))\right) dy) - h(x,0,0,0,0)\Big) \\
&\qquad\qquad + A\Big(f; (x,0,0,0,0)\Big) \\
&\leq \frac{1}{6}\gamma\Big(A(f;\ \Big|u_1(\varrho(x))\Big| + \Big|u_2(\varrho(x))\Big| + \Big|u_3(\varrho(x))\Big|\Big) \\
&\qquad + \Big|\int_0^{\rho(x)} k\left(x, y, u_1(\varrho(y)), u_2(\varrho(y)), u_3(\varrho(y))\right) dy\Big| \\
&\qquad + A\Big(f; (x,0,0,0,0)\Big) \\
&\leq \frac{1}{6}\gamma\Big(A(f;\ \Big|u_1(\varrho(x))\Big| + \Big|u_2(\varrho(x))\Big| + \Big|u_3(\varrho(x))\Big|\Big) + \Gamma + M
\end{aligned}$$

Thus, we have

$$A\Big(f;\ |\Upsilon(u_1,u_2,u_3)|\Big) \leq \frac{1}{6}\gamma\Big(A(f;\ \|u_1\|+\|u_2\|+\|u_3\|)\Big| + \Gamma + M \tag{12}$$

Therefore,
In view of assumption (v) and inequality (12), we conclude that the function Υ maps $(\bar{B}_{r_0})^3$ into $\bar{B}_{r_0}$.
In this stage , we prove the continuity of map Υ on $(\bar{B}_{r_0})^3$.
Fix $\varepsilon > 0$ and choose (u_1,u_2,u_3) and $(v_1,v_2,v_3) \in (\bar{B}_{r_0})^3$ arbitrarily such that $\max\{\|u_i - v_i\|\} \leq \varepsilon$ for all $i = 1,2,3$. Then, for all $x \in [0,T]$, we have

$$\begin{aligned}
&A\Big(f;\ \Upsilon(u_1,u_2,u_3)(x) - \Upsilon(v_1,v_2,v_3)(x)\Big) = \\
&A\Big(f;\ h\big(x,u_1,u_2,u_3,\int_0^{\rho(x)} k\,(x,y,u_1,u_2,u_3)\,dy\big) \\
&\quad - h\big(x,v_1,v_2,v_3,\int_0^{\rho(x)} k\,(x,y,v_1,v_2,v_3)\,dy\big)\Big) \\
&\leq \frac{1}{6}\gamma\Big(A(f;\ \Big|u_1(\varrho(x)) - v_1(\varrho(x))\Big| + \Big|u_2(\varrho(x)) - v_2(\varrho(x))\Big| \\
&\qquad + \Big|u_3(\varrho(x)) - v_3(\varrho(x))\Big|)\Big) \\
&\qquad + \Big|\int_0^{\rho(x)} k\,(x,y,u_1,u_2,u_3) - k\,(x,y,v_1,v_2,v_3)\,dy\Big| \\
&\leq \frac{1}{6}\gamma\Big(A(f;\ 3\varepsilon)\Big) + \rho(T)\omega^T(k,\varepsilon),
\end{aligned}$$

where

$$\begin{aligned}
\omega^T(k,\varepsilon) = \sup\{&|k(x,y,u_1,u_2,u_3) - k(x,y,v_1,v_2,v_3)| : x,y \in [0,\rho(T)], \\
&u_i, v_i \in [-r_0,r_0], i = 1,2,3, \max|u_i - v_i| \leq \varepsilon\}.
\end{aligned}$$

$$\rho(T) = sup\{\rho(x) : x \in [0,T]\}.$$

From the continuity of k on the compact set $[0,T] \times [0,\rho(T)] \times [-r_0,r_0]^3$, we infer $\omega^T(k,\varepsilon) \longrightarrow 0$ as $\varepsilon \longrightarrow 0$. Therefore, from the above calculation, we get Υ is a continuous map on $(\bar{B}_{r_0})^3$. Now, we check that Υ satisfies all the assumptions of Theorem 2.4. Suppose X_1, X_2 and X_3 be nonempty and bounded subsets of $\bar{B}_{r_0}$. Let $\varepsilon > 0$ is an arbitrary constant. Also, we choose $p_1, p_2 \in [0,T]$, with $|p_2 - p_1| \leq \varepsilon$, $|\varrho(p_2) - \varrho(p_1)| \leq \varepsilon$ and $u_j \in X_j$ for all $j = 1,2,3$. Then, we have

$$
\begin{aligned}
&A(f;\ \Big|\Upsilon(u_1,u_2,u_3)(p_1)-\Upsilon(u_1,u_2,u_3)(p_2)\Big|)\\
&\le A(f;\ \Big|h(p_1,u_1(\varrho(p_1)),u_2(\varrho(p_1)),u_3(\varrho(p_1)))\\
&\qquad\qquad ,\int_0^{\rho(p_1)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\\
&\qquad -h(p_2,u_1(\varrho(p_2)),u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\varrho(p_2)}k(p_2,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\Big|)\\
&\quad +A(f;\ \Big|h(p_2,u_1(\varrho(p_1)),u_2(\varrho(p_1)),u_3(\varrho(p_1)))\\
&\qquad\qquad ,\int_0^{\rho(p_1)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\\
&\qquad -h(p_2,u_1(\varrho(p_2)),u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\rho(p_1)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\Big|)\\
&\quad +A(f;\ \Big|h(p_2,u_1(\varrho(p_2))u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\rho(p_1)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\\
&\qquad -h(p_2,u_1(\varrho(p_2)),u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\rho(p_2)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\Big|)\\
&\quad +A(f;\ \Big|h(p_2,u_1(\varrho(p_2)),u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\rho(p_2)}k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y))dy)\\
&\qquad -h(p_2,u_1(\varrho(p_2)),u_2(\varrho(p_2)),u_3(\varrho(p_2)))\\
&\qquad\qquad ,\int_0^{\rho(p_2)}k(p_2,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy)\Big|).
\end{aligned}
$$

Using condition (11) we have

$$
\begin{aligned}
& A\Big(f;\ \Big|\Upsilon(u_1,u_2,u_3)(p_1)-\Upsilon(v_1,v_2,v_3)(p_2)\Big|\Big)\\
&\le \omega_{r_0}(h,\varepsilon)\\
&\ +\frac{1}{6}A(f;\ \{|u_1(\varrho(p_1))-u_1(\varrho(p_2))|+|u_2(\varrho(p_1))-u_2(\varrho(p_2))|\\
&\qquad\qquad +|u_3(\varrho(p_1))-u_3(\varrho(p_2))|\})\\
&\ +A(f;\ \Big|\int_0^{\rho(p_1)} k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy)\\
&\qquad\quad -\int_0^{\rho(p_2)} k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy\Big|\\
&\ +A(f;\ \Big|\int_0^{\rho(p_2)} k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy\\
&\qquad\quad -\int_0^{\rho(p_2)} k(p_2,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy\Big|)\\
&\le \omega_{r_0}(h,\varepsilon)\\
&\qquad +\frac{1}{6}A(f;\ \{\omega(u_1,\varepsilon)+\omega(u_2,\varepsilon)+\omega(u_3,\varepsilon)\})\\
&\qquad +A(f;\ \Big|\int_{\rho(p_1)}^{\rho(p_2)} k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy\Big|)\\
&\qquad +A(f;\ \int_0^{\rho(p_2)} \Big|k(p_1,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))\\
&\qquad\qquad\qquad -k(p_2,y,u_1(\varrho(y)),u_2(\varrho(y)),u_3(\varrho(y)))dy\Big|)\\
&\le \omega_{r_0}(h,\varepsilon)+\frac{1}{6}A(f;\ \{\omega(u_1,\varepsilon)+\omega(u_2,\varepsilon)+\omega(u_3,\varepsilon)\})\\
&\qquad +A(f;\ \omega(\rho,\varepsilon)U_{r_0})+A(f;\ T\omega_{r_0}(k,\varepsilon)\big)
\end{aligned}
\tag{13}
$$

where

$$
\begin{aligned}
\omega_{r_0}(h,\varepsilon)&=\sup\{|h(p_1,s,t,z,w)-h(p_2,s,t,z,w)|:p_1,p_2\in[0,T],\\
&\qquad |p_2-p_1|\le\varepsilon, ||s||,||t||,||z||\le r_0,\ |w|\le\Gamma\},\\
\omega_{r_0}(k,\varepsilon)&=\sup\{|k(p_1,q,s,t,z)-k(p_2,q,s,t,z)|:p_1,p_2,q\in[0,T],\\
&\qquad |p_2-p_1|\le\varepsilon, ||s||,||t||,||z||\le r_0\},\\
U_{r_0}&=\sup\{|k(p,q,s,t,z)|:p,q\in[0,T]\ s,t,z\in[-r_0,r_0]\}.
\end{aligned}
$$

Since in (13), u_i was an arbitrary element of X_i for $i=1,2,3$ we conclude that

$$
\begin{aligned}
A\Big(f;\ \alpha(\Upsilon(X_1\times X_2\times X_3),\varepsilon)\Big) &\le A\Big(f;\ \omega_{r_0}(h,\varepsilon)\Big)\\
&\quad +A(f;\ \{\alpha(X_1,\varepsilon)+\alpha(X_2,\varepsilon)+\alpha(X_3,\varepsilon)\})\\
&\quad +A(f;\ \omega(\rho,\varepsilon)U_{r_0})\\
&\quad +A(f;\ T\omega_{r_0}(k,\varepsilon)).
\end{aligned}
$$

The uniform continuity of h, ρ and g on the compact sets $[0,T]\times[-r_0,r_0]^3\times[-M_{r_0},M_{r_0}]$, $[0,T]$ and $[0,T]^2\times[-r_0,r_0]^2$ respectively, yields that $\omega_{r_0}(h,\varepsilon)\longrightarrow 0$, $\omega(\rho,\varepsilon)\longrightarrow 0$ and $\omega_{r_0}(k,\varepsilon)\longrightarrow 0$ as $\varepsilon\longrightarrow 0$. Therefore, by taking $A(f;t)=t,\ f=I$, we have

$$\alpha(\Upsilon(X_1 \times X_2 \times X_3)) \leq \frac{1}{6}\big(\alpha(X_1) + \alpha(X_2) + \alpha(X_3)\big)$$

Therefore, Corollary 2.5 infers that the operator Υ has a tripled fixed point. Therefore, the system of functional integral equations (10) has at least one solution in $(C[0,T])^3$. $\square$

5. Example

Example 5.1. Suppose that the following system of integral equations is given:

$$\begin{cases} p(t) = \frac{1}{3}e^{-t^2} + \frac{\tan^{-1}p(t)+\sinh^{-1}q(t)+\tan^{-1}r(t)}{6\pi+t^{10}} \\ \qquad +\dfrac{1}{6}\displaystyle\int_0^t \frac{s(|cosp(s)| + \sqrt{(1+p^2(s))(1+sin^2q(s))(1+cos^2r(s))})}{e^t(1+p^2(s))(1+sin^2q(s))(1+cos^2r(s))}ds \\ q(t) = \frac{1}{3}e^{-t^2} + \frac{\tan^{-1}q(t)+\sinh^{-1}p(t)+\tan^{-1}r(t)}{6\pi+t^{10}} \\ \qquad +\dfrac{1}{6}\displaystyle\int_0^t \frac{s(|\cos q(s)| + \sqrt{(1+q^2(s))(1+sin^2p(s))(1+cos^2r(s))})}{e^t(1+q^2(s))(1+sin^2p(s))(1+cos^2r(s))}ds. \\ r(t) = \frac{1}{3}e^{-t^2} + \frac{\tan^{-1}r(t)+\sinh^{-1}q(t)+\tan^{-1}p(t)}{6\pi+t^{10}} \\ \qquad +\dfrac{1}{6}\displaystyle\int_0^t \frac{s(|cosr(s)| + \sqrt{(1+r^2(s))(1+sin^2q(s))(1+cos^2p(s))})}{e^t(1+r^2(s))(1+sin^2q(s))(1+cos^2p(s))}ds \end{cases} \tag{14}$$

It is clear that this system of integral equations (14) is a special case of the system (10) with

$$\varrho(t) = \rho(t) = t, \qquad t \in [0,1],$$

$$h(t,p,q,r,\lambda) = \frac{1}{3}e^{-t^2} + \frac{\tan^{-1}p + \sinh^{-1}q + \tan^{-1}r}{6\pi + t^{10}} + \frac{\lambda}{6},$$

and

$$k(t,s,p,q,r) = \frac{s(|cosp| + \sqrt{(1+p^2)(1+sin^2q)(1+cos^2r)})}{e^t(1+p^2)(1+sin^2q)(1+cos^2r)}.$$

To show that the above system has a solution, we should verify the conditions (i)-(iv) of Theorem 4.1.

We observe that condition (i) is evident. We define $\gamma(t) = t$ and $A(f;x) = x$,
Now, we have

$$\begin{aligned} &A\Big(f; \Big|h(t,p,q,r,m) - h(t,u,v,w,n)\Big|\Big) \\ &\leq \frac{|\tan^{-1}p - \tan^{-1}u| + |\sinh^{-1}q - \sinh^{-1}v| + |\tan^{-1}r - \tan^{-1}w|}{6\pi + t^{10}} + \frac{|m-n|}{6} \\ &\leq \frac{\tan^{-1}|p-u|}{6\pi} + \frac{\sinh^{-1}|q-v|}{6\pi} + \frac{\tan^{-1}|r-w|}{6\pi} + \frac{|m-n|}{6} \\ &\leq \frac{1}{6}\{|p-u| + |q-v| + |r-w|\} + |m-n|. \end{aligned}$$

Therefore, f satisfies condition (ii) of Theorem 4.1. Also,

$$M = \sup\left\{A(f;|h(p,0,0,0,0)|) : p \in [0,T]\right\} = \sup\{\frac{1}{3}e^{-t^2} : t \in [0,1]\} \leq 0.5$$

Clearly, condition (iii) of Theorem 4.1 is satisfy, that is, k is continuous on $[0,T] \times [0,T] \times \mathbb{R}^3$, and

$$\begin{aligned}
\Gamma &= \sup\left\{|\int_0^{\rho(p)} k(p,q,u_1,u_2,u_3)dy|\right\} \\
&= \sup\left\{|\int_0^t \frac{s(|cosp(s)| + \sqrt{(1+p^2(s))(1+sin^2q(s))(1+cos^2r(s))})}{e^t(1+p^2(s))(1+sin^2q(s))(1+cos^2r(s))}ds| \right. \\
&\qquad\qquad \left. : t,s \in [0,1], p,q \in [-r,r]\right\} \\
&\leq \sup \frac{t^2}{e^t} \leq 1.
\end{aligned}$$

Moreover, obviously every $r \geq 3$ satisfies inequality in condition (iv), i.e.,

$$\frac{1}{6}\Big(A(f,3r)\Big) + \Gamma + M \leq \frac{1}{6}\Big(3r\Big) + 0.5 + 1 \leq r.$$

Therefore, all of the assumptions Theorem 4.1 are fulfilled. Consequently, the above system of integral equations has at least one solution in $\{C[0,T]\}^3$.

References

[1] A. Aghajani and N. Sabzali. Existence of coupled fixed points via measure of noncompactness and applications. *Nonlinear Convex Anal*, 14: 941–952, 2014.

[2] R.R. Akhmerov, M.I. Kamenskii, A.S. Potapov, A.E. Rodkina, and B.N. Sadovskii. *Measures of Noncompactness and Condensing Operators*. Birkhauser, Basel, 1992.

[3] I. Altun and D. Turkoglu. Existence of coupled fixed points via measure of noncompactness and applications. *Journal of Computational Analysis and Applications*, 9:9–14, 2007.

[4] S. Banaei. On solvability of infinite system of integral equations of volterra together with hammerstein type in the frechet spaces. *Filomat*, 35(6):2055–2069, 2021.

[5] S. Banaei, V. Parvaneh, and M. Mursaleen. Measures of noncompactness and infinite systems of integral equations of urysohn type in $l^\infty(g)$. *Carpathian J. Math*, 37:407–416, 2021.

[6] J. Banás and K. Goebel. *Measures of noncompactness in Banach spaces*, volume 60. Lecture Notes in Pure and Applied Mathematics, New York, 1980.

[7] T. G. Bhaskar and V. Lakshmikantham. Fixed point theorems in partially ordered metric spaces and applications. *Nonlinear Anal*, 65: 1379–1393, 2006.

[8] G. Darbo. Punti uniti in trasformazioni a codominio non compatto. *Rendiconti del Seminario Matematico della Universita di Padova*, 24: 84–92, 1955.

[9] H. Işık, S. Banaei, F. Golkarmanesh, V. Parvaneh, C. Parkl, and M. Khorshidi. On new extensions of darbo's fixed point theorem with applications. *Symmetry*, 12(3), 2020. doi: 10.3390/sym12030424.

[10] C. Kuratowski. Sur les espaces complets. *Fundamenta Mathematicae*, 15:301–309, 1930.

[11] V. Parvaneh, N. Hussain, A. Mukheimer, and H. Aydi. On fixed point results formodifed js-contractions with applications. *Axioms*, 15(8), 2019. doi: 10.3390/axioms8030084.

[12] V. Parvaneh, M. Khorshidi, M. De La Sen, H. Işik, and M. Mursaleen. Measure of noncompactness and a generalized darbo's fixed point theorem and its applications to a system of integral equations. *Adv Differ Equ*, 243, 2020. doi: 10.1186/s13662-020-02703-z.

[13] S. B. Prešić. Sur une classe dinequations aux differences finies et sur la convergence de certaines suites. *Pub. de. lInstitut Math. Belgrade.*, 5:75–78, 1965.

[14] M. Rabbani, R. Arab, and B. Hazarika. Solvability of nonlinear quadratic integral equation by using simulation type condensing operator and measure of noncom pactness. *Applied Mathematics and Computation*, 349:102–117, 2019.

[15] D. Wardowski. fixed points of a new type of contractive mappings in complete metric spaces. *fixed Point Theory Appl*, 94, 2012. doi: 10.1186/1687-1812-2012-94.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–117, pp. 35–39

New common fixed point results on generalized metric spaces

Hamid Faraji*[a,*]

[a] *Department of Mathematics, Islamic Azad University, South Tehran Branch,Tehran, Iran*

Article Info

Keywords:
Fixed point
Vector valued metric
Weakly compatible

2020 MSC:
47H10
54H25

Abstract

In this paper, we give some common fixed point results for self-mappings on a complete generalized metric apace. To show the usability of our results, we present some examples.

1. First Section

Fixed point theory plays an important role for solving problems in various branches of mathematical, such as nonlinear analysis, integral and differential equations (see [3, 12, 21, 23]). There exist many interesting generalizations of metric spaces (see for example [2, 4, 5, 9, 15, 19]). In 1964, Perov [16] introduced the concept of of vector-valued metric space and gave a generalization of the Banach contraction principle.(see e.g. [6–8, 10, 11, 13, 14, 16–18, 20, 22] and references therein). We give below some definitions and results which will help in proving our main results.
Let X be a nonempty set. A function $D : X \times X \to \mathbb{R}^m_+$ is said to be a vector-valued metric if for all $x, y, z \in X$ the following properties are satisfied:
1) $D(x, y) \succ \theta$ for all $x, y \in X$; $D(x, y) = \theta$ if and only if $x = y$;
2) $D(x, y) = D(y, x)$ for all $x, y \in X$;
3) $D(x, y) \preceq D(x, z) + D(z, y)$ for all $x, y, z \in X$.
A nonempty set X endowed with a vector-valued metric D is called a generalized metric space and it will be denoted by (X, D).

Remark 1.1. If $\alpha, \beta \in \mathbb{R}^m$ with $\alpha = (\alpha_1, ..., \alpha_m), \beta = (\beta_1, ..., \beta_m)$, then by $\alpha \preceq \beta$ (respectively $\alpha \prec \beta$), we mean that $\alpha_i \leq \beta_i$ (respectively $\alpha_i < \beta_i$), for all $i = 1, ..., m$.

Throughout this paper we denote by $M_{m\times m}(R_+)$ the set of all $m \times m$ matrixs with positive elements, by 0 the zero $m \times m$ matrix and by I the identity $m \times m$ matrix. A matrix $A \in M_{m\times m}(R_+)$ is said to be matrix convergent to zero if $A^n \to 0$ as $n \to +\infty$.

*Talker
Email address: hf.math.ac@gmail.com (Hamid Faraji*)

Theorem 1.2. *[6] Let $A \in M_{m\times m}(R_+)$. Then following assertions are equivalent:*
1. A is convergent towards zero.
2. $A^n \to 0$ as $n \to +\infty$.
3.The matrix $(I-A)$ is non-singular and

$$(I-A)^{-1} = I + A + A^2 + ... + A^n + ...$$

4.The matrix $(I-A)$ is non-singular and $(I-A)^{-1}$ has non-negative elements.
5. $A^n q \to 0$ and $qA^n \to 0$ as $n \to +\infty$, for each $q \in \mathbb{R}^m$.

Example 1.3. [6] The given below matrices are convergent to a zero matrix.
1. Any matrix $A = \begin{bmatrix} a & a \\ b & b \end{bmatrix}$, where $a, b \in \mathbb{R}_+$ and $a + b < 1$.
2. Any matrix $A = \begin{bmatrix} a & b \\ a & b \end{bmatrix}$, where $a, b \in \mathbb{R}_+$ and $a + b < 1$.
3. $A = \begin{bmatrix} \frac{1}{4} & \frac{1}{2} \\ 0 & \frac{1}{4} \end{bmatrix}$.

In 1964 Perov [16] proved the following result.

Theorem 1.4. *Let (X, D) be a complete generalized metric space and $T : X \to X$ be an self-mapping. Suppose there exists a matrix $A \in M_{m\times m}(\mathbb{R}_+)$ convergent to zero such that*

$$D(T(x), T(y)) \preceq AD(x, y).$$

Then the following statements hold:
1. T has a unique fixed point x^*.
2. The Picard iterative sequence $x_n = T^n(x_0)$, $n \in \mathbb{N}$ converges to x^* for all $x_0 \in X$.
3. $D(x_n, x^*) \preceq A^n(I-A)^{-1}d(x_0, x_1)$, for all $n \in \mathbb{N}$. where $A \in M_{m\times m}(R+)$ is a matrix convergent to zero.

Definition 1.5. [1] Let X be a nonempty set and $T, S : X \to X$. Two self-mappings T and S are said to be weakly compatible if they commute at their coincidence points; i.e., if $T(x) = S(x)$ for some $x \in X$, then $T(S(x)) = S(T(x))$.

Proposition 1.6. *[1] Let T and S be weakly compatible self maps of a set X. If T and S have a unique point of coincidence $w = Tx = Sx$, then w is the unique common fixed point of T and S.*

2. Second Section

In this section of the paper, we will give some common fixed point results of Perov type contractive mappings in generalized metric spaces.

Theorem 2.1. *Let (X, D) be a generalized metric apace and let $T, S : X \to X$ be self-mappings which satisfy,*

$$D(Sx, Sy) \preceq B(D(Sy, Sx) + D(Sy, Ty)), \tag{1}$$

for all $x, y \in X$, where $B \in M_{m\times m}(R_+)$ and B be a nonzero matrix convergent to zero. Moreover if T and S are weakly compatible, $S(X) \subseteq T(X)$ and $T(X)$ be a complete subspace of X, then T and S have a unique common fixed point in X.

Proof. We choose elements $x_0, x_1 \in X$ such that $Tx_1 = Sx_0$. Since $S(X) \subseteq T(X)$, we can define a sequence $\{x_n\}$ such that $Tx_n = Sx_{n-1}$ for each $n \in \mathbb{N}$. From (1), we have

$$\begin{aligned} D(Tx_{n+1}, Tx_n) &= D(Sx_n, Sx_{n-1}) \\ &\preceq B(D(Sx_{n-1}, Tx_n) + D(Sx_{n-1}, Tx_{n-1})) \\ &= B(D(Tx_n, Tx_n) + D(Tx_n, Tx_{n-1})), \end{aligned}$$

for all $n \in \mathbb{N}$. Inductively, we get

$$D(Tx_{n+1}, Tx_n) \preceq B^n D(Tx_1, Tx_0), \tag{2}$$

for all $n \in \mathbb{N}$. Suppose that $mleqn$. Using (2), we can write

$$\begin{aligned} D(Tx_m, Tx_n) &\preceq D(Tx_m, Tx_{m-1}) + ... + D(Tx_{n+1}, Tx_n) \\ &\preceq (B^{m-1} + ... + B^n) D(Tx_1, Tx_0) \\ &= B^n (I + B + ...B^{m-n-1}) D(Tx_1, Tx_0) \\ &\preceq B^n (I - B)^{-1} D(Tx_1, Tx_0). \end{aligned}$$

Then, $\lim_{n\to+\infty} D(Tx_m, Tx_n) = \theta$. Hence, $\{Tx_n\}$ is a Cauchy sequence in $T(X)$. Since $T(X)$ is complete, there exists $x^* \in X$ such that $D(Tx_n, Tx^*) \to \theta$ as $n \to +\infty$. We show that $Sx^* = Tx^*$. Using (1), we have

$$\begin{aligned} D(Sx^*, Tx^*) &\preceq D(Sx^*, Tx_n) + D(Tx_n, Tx^*) \\ &\preceq D(Sx^*, Sx_{n-1}) + D(Tx_n, Tx^*)) \\ &\preceq B(D(Sx_{n-1}, Tx^*) + D(Tx_n, Tx_{n-1})) + D(Tx_n, Tx^*) \\ &= B(D(Tx_n, Tx^*) + D(Tx_n, Tx_{n-1})) + D(Tx_n, Tx^*) \to \theta, \text{as } n \to +\infty. \end{aligned}$$

Then, $D(Sx^*, Tx^*) = \theta$, i.e $Tx^* = Sx^*$. Now we show that T and S have a unique point of coincidence. For this, assume that there exists another point z^* in X such that $Tz^* = Sz^*$. Using (1), we obtain

$$\begin{aligned} D(Sx^*, Sz^*) &\preceq B(D(Sz^*, Tx^*) + D(Sz^*, Tz^*)) \\ &= BD(Sz^*, Sx^*). \end{aligned}$$

Thuse, $(I - B)D(Sz^*, Sx^*) \preceq \theta$. Since $I - B \neq \theta$, we have $D(Sz^*, Sx^*) = \theta$, i.e $Sz^* = Sx^*$. Therefore, T and S have a unique point of coincidence. Then by Proposition 1.6, T and S have a unique common fixed point. ☐

Corollary 2.2. *Let (X, D) be a complete generalized metric apace and let $S : X \to X$ be a self-mapping satisfying:*

$$D(Sx, Sy) \preceq B(D(Sy, x) + D(Sy, y)),$$

for all $x, y \in X$, where $B \in M_{m\times m}(R_+)$ and B be a nonzero matrix convergent to zero. Then S has a unique fixed point in X.

Example 2.3. Let $X = [0, +\infty)^2$ and vector valued-metric $D : X \times X \to \mathbb{R}^2$ be defined as follows:

$$D((x_1, x_2), (y_1, y_2)) = (|x_1 - y_1|, |x_2 - y_2|).$$

Define $T, S : X \to X$ by $T(x_1, x_2) = (4x_1, x_2)$ and $S(x_1, x_2) = (\frac{x_1}{3}, 0)$. Suppose that $B = \begin{bmatrix} \frac{3}{4} & \frac{1}{2} \\ 0 & \frac{3}{4} \end{bmatrix}$. Thus, Theorem 2.1 implies that T and S have a unique common fixed point in X. Note that $(0, 0)$ is common fixed point of T and S.

Theorem 2.4. *Let (X, D) be a complete generalized metric apace and let $T, S : X \to X$ be self-mappings which satisfy,*

$$D(Sx, Ty) \preceq AD(x, Sx) + BD(x, y), \tag{3}$$

for all $x, y \in X$, where $A, B \in M_{m\times m}(R_+)$ and $A + B$ be a nonzero matrix convergent to zero. Then T and S have a unique common fixed point in X.

Proof. Let x_0 be an arbitrary point in X. We can define a sequence $\{x_n\}$ in X by $x_{2k+1} = Sx_{2k}$ and $x_{2k+2} = Tx_{2k+1}$ for each $k \in \mathbb{N} \cup \{0\}$. If $D(x_{2k}, x_{2k+1}) = \theta$ for some $k_0 \in \mathbb{N}$, we have $x_{2k_0} = x_{2k_0+1} = Sx_{2k_0}$. Then x_{2k_0} is a point of S. From (3), we obtain

$$\begin{aligned} D(x_{2k_0}, Tx_{2k_0}) &= D(Sx_{2k_0}, Tx_{2k_0}) \\ &\preceq AD(x_{2k_0}, Sx_{2k_0}) + BD(x_{2k_0}, x_{2k_0}) = \theta. \end{aligned}$$

Consequently, $x_{2k_0} = Tx_{2k_0} = Sx_{2k_0}$, then x_{2k_0} is a common fixed point of T and S and the proof is finished. Assume that $x_{2k} \neq x_{2k+1}$ for all $k \in \mathbb{N} \cup \{0\}$. From (3), we have

$$\begin{aligned} D(x_{2k+1}, x_{2k+2}) &= D(Sx_{2k}, Tx_{2k+1}) \\ &\preceq AD(x_{2k}, Sx_{2k}) + BD(x_{2k}, x_{2k+1}) \\ &= (A+B)D(x_{2k}, x_{2k+1}), \end{aligned}$$

for all $k \in \mathbb{N} \cup \{0\}$. Then, we obtain

$$D(x_{2k+1}, x_{2k+2}) \preceq (A+B)^{2k+1} D(x_1, x_0). \tag{4}$$

By a similar method, we can show that

$$D(x_{2k+2}, x_{2k+3}) \preceq (A+B)^{2k+2} D(x_1, x_0). \tag{5}$$

From (4) and (5), we get

$$D(x_n, x_{n+1}) \preceq (A+B)^n D(x_1, x_0). \tag{6}$$

for all $n \in \mathbb{N}$. Suppose that $m \geq n$. Using (6), we can write

$$\begin{aligned} D(x_m, x_n) &\preceq D(x_m, x_{m-1}) + \ldots + D(x_{n+1}, x_n) \\ &\preceq ((A+B)^{m-1} + \ldots + (A+B)^n) D(x_1, x_0) \\ &= (A+B)^n (I + (A+B) + (A+B)^2 + \ldots + (A+B)^{m-n-1}) D(x_1, x_0) \\ &\preceq (A+B)^n (I - A - B)^{-1} D(x_1, x_0). \end{aligned}$$

Then, $D(x_m, x_n) \to \theta$ as $n \to +\infty$. Hence, $\{x_n\}$ is a Cauchy sequence in X. Since X is complete, there exists $x^* \in X$ such that

$$\lim_{n \to +\infty} x_n = x^*. \tag{7}$$

From (3), we have

$$\begin{aligned} D(Sx^*, x^*) &\preceq D(Sx^*, Tx_{2n+1}) + D(Tx_{2n+1}, x^*) \\ &\preceq AD(x^*, Sx^*) + BD(x^*, x_{2n+1}) + D(x_{2n+2}, x^*), \end{aligned}$$

for all $n \in \mathbb{N}$. From (7), we have $(I - A)D(Sx^*, x^*) \preceq \theta$. Since $I \neq A$, we have $Sx^* = x^*$. Now we show that x^* is a common fixed point of T and S. From (3), we obtain

$$D(Sx^*, Tx^*) \preceq AD(x^*, Sx^*) + BD(x^*, x^*).$$

Thuse $D(Sx^*, Tx^*) = \theta$ i.e $Sx^* = Tx^* = x^*$. Now, we show that T and S have a unique common fixed point. For this, assume that there exists another common fixed point z^* in X such that $Sz^* = Tz^* = z^*$. From (3), we have

$$\begin{aligned} D(z^*, x^*) &= D(Sz^*, Tx^*) \\ &\preceq AD(z^*, Sz^*) + BD(z^*, x^*). \end{aligned}$$

Then, $(I - B)D(z^*, x^*) \preceq \theta$. Since $I \neq B$, we have $D(z^*, x^*) = \theta$ i.e $z^* = x^*$. Then, T and S have a unique common fixed point. □

Corollary 2.5. *In the case that $T = S$ and $A = 0$, we obtain Perov Theorem* **??**.

Corollary 2.6. *Let (X, D) be a complete generalized metric apace and let $T : X \to X$ be self-mapping which satisfy,*

$$D(Tx, Ty) \preceq AD(x, Tx) + BD(x, y),$$

for all $x, y \in X$, where $A, B \in M_{m\times m}(R_+)$ and $A+B$ be a nonzero matrix convergent to zero. Then T has a unique fixed point in X.

Example 2.7. Let $X = [0, +\infty)^2$ and vector valued- metric $D : X \times X \to \mathbb{R}^2$ be defined as follows:

$$D((x_1, x_2), (y_1, y_2)) = (|x_1 - y_1|, |x_2 - y_2|).$$

Define $T, S : X \to X$ by $T(x_1, x_2) = (Arccot\frac{x_1,}{2}, \frac{x_2}{10})$ and $S(y_1, y_2) = (Arccot\frac{y_1}{2}, \frac{y_2}{20})$. Suppose that $B = \begin{bmatrix} \frac{3}{4} & \frac{\frac{3}{4}}{\frac{3}{4}} \\ 0 & \end{bmatrix}$ and $B = \begin{bmatrix} \frac{3}{4} & 0 \\ 0 & \frac{3}{4} \end{bmatrix}$ Thus, Theorem 2.4 implies that T and S have a unique common fixed point in X. Note that $(0, 0)$ is common fixed point of T and S.

References

[1] M. Abbas, G. Jungck, *Common fixed point results for noncommuting mappings without continuity in cone metric spaces*, J. Math. Anal. Appl. 341(1), 416-420, 2008.

[2] R. P. Agarwal, M. A. El-Gebeily and D. O'Regan, *Generalized contractions in partially ordered metric spaces*, Appl. Anal. 87, no.1, 109-116, 2008.

[3] S. Banach, *Sur les operations dans les ensembles abstraits et leurs applications aux equations integrales*, Fund. Math., 3, 133–181, 1922.

[4] I. Beg, A.R. Butt, *Fixed point for set-valued mappings satisfying an implicit relation in partially ordered metric spaces*, Nonlinear Anal. 71, (9), 3699–3704, 2009.

[5] S. Czerwik, *Nonlinear set-valued contraction mappings in b-metric spaces*, Atti Sem. Mat. Fis. Univ. Modena 46, no. 2, 263-276, 1998.

[6] A. D. Flip, A. Petrusel, *Fixed point theorems on space endowed with vector valued metrics*, Fixed Point Theory Appl. 2010, Art. ID 281381, 15 pp, 2010.

[7] R. George, R. Rajagopalan, H. A. Nabwey, S. Radenović, *Dislocated Cone Metric Spaces over Banach Algebra and Fixed Points of $\alpha-$Quasi Contraction Mappings of Perov type*, Fixed Point Theory Appl., 2017:24, 2017.

[8] L. Guran, M.-F.Bota, A.Naseem, Z. D. Mitrović, M. de la Sen, S. Radenović, *On some new multivalued results in the metric spaces of Perov's type*, Mathematics, 8(3), pages 1-12, 2020.

[9] L.G. Huang, X. Zhang, *Cone metric spaces and fixed point theorems of contractive mappings*, J. Math. Anal. Appl. 332 (2), 1468–1476, 2007.

[10] J. Jachymski, J. Klima, *Around Perov's fixed point theorem for mappings on generalized metric spaces*, Fixed Point Theory 17, no. 2, 367-380, 2016.

[11] Z. Kadelburg, N. Fabiano, N. Mirkov, S. Radenović, *On ordered topological vector groups -new results, Journal of Nonlinear and Convex Analysis*, vol 23, no 6, 1231-1254, 2022.

[12] G. N. V. Kishore, R. P. Agarwal, B. S. Rao, R. V. N. S. Rao, *Caristi type cyclic contraction and common fixed point theorems in bipolar metric spaces with applications*, Fixed Point Theory Appl., 2018, 1–13, 2018

[13] N. Mirković, S. Radenović, S. Radojević, *Some new observations for F-contractions in vector-valued metric spaces of Perov's type*, Axioms, 10(2), 127, 2021.

[14] N. Mirković, N. Fabiano, S. Radenović, *Critical remarks on "A new fixed point result of Perov type and its application to a semilinear operator system*, TWMS J. Pure Appl. Math. V 14, No 1, 141-145, 2023

[15] Z. Mustafa, H. Obiedat, *A fixed point theorem of Reich in G-metric spaces*, Cubo 12 (1), 83–93, 2010.

[16] A. I. Perov, *On the Cauchy problem for a system of ordinary differential equations*, Pviblizhen. Met. Reshen. Differ. Uvavn., no. 2, 115-134, 1964.

[17] S. Radenović, F. Vetro, *Some remarks on Perov type mappings in cone metric spaces*, Mediterr. J. Math. 14, no. 6, paper no. 240, 15 pp, 2017.

[18] A. Savić, N. Fabiano, N. Mirkov, A. Sretenović , S. Radenović, *Some significant remarks on multivalued Perov type contractions on cone metric spaces with a directed graph*, AIMS Math. 7 , no. 1, 187-198, 2022.

[19] T. Suzuki, *A generalized Banach contraction principle that characterizes metric completeness*, Proc. Amer. Math. Soc. 136, no. 5, 1861-1869, 2008.

[20] F.Vetro, S. Radenović, *Some results of Perov type on rectangular cone metric spaces*, Journal of Fixed Point Theory and Applications, vol. 20, no. 1, article 41, 2018.

[21] D. Wardowski, *Fixed points of a new type of contractive mappings in complete metric spaces*, Fixed Point Theory Appl.2012:94, 6 pp, 2012.

[22] S. Xu, Y. Han, S. Aleksić, S. Radenović, *Fixed Point Results for Nonlinear Contractions of Perov Type in Abstract Metric Spaces with Applications*, AIMS Math. 7, no. 8, 14895-14921, 2022.

[23] K. Zoto, Z. D. Mitrović, S. N. Radenović, *Unified setting of generalized contractions by extending simulation mappings in b-metric-like spaces*, Acta Math. Univ. Comenianae, 9, 247–258, 2022.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–118, pp. 40–44

Block fully indecomposable tensors

Malihe Nobakht-Kooshkghazi[a,*], Hamidreza Afshin[b]

[a] *Vali-e-Asr University of Rafsanjan*
[b] *Vali-e-Asr University of Rafsanjan*

Article Info

Keywords:
Fully indecomposable tensors
Partly decomposable tensors
Permanent
Tensors.

2020 MSC:
15A69

Abstract

In this paper, we generalize the definitions of partly decomposable and fully indecomposable matrices to tensors. We obtain a relation between partly decomposable and reducible tensors. Furthermore, we present a sufficient condition for block tensors to be fully indecomposable. We study the relations between fully indecomposable and irreducible tensors.

1. Introduction

A *tensor* $\mathscr{A} = (a_{i_1 \dots i_d})_{n_1 \times \dots \times n_d}$ of order d and dimension $(n_1, \dots, n_d)$ is a multi-array of entries $a_{i_1 \dots i_d} \in \mathbb{F}$, where $i_j = 1, \dots, n_j$ for $j = 1, \dots, d$ and $\mathbb{F}$ is a field. When $n_1 = n_2 = \dots = n_d = n$, we say that $\mathscr{A}$ is a square tensor of order d and dimension n. A tensor $\mathscr{A}$ is called non-negative if all its entries are non-negative [3]. Many combinatorial properties of non-negative matrices have been studied by using partly decomposable, fully indecomposable, nearly decomposable, and nearly reducible matrices [6]. Fully indecomposable matrices have played an interesting role in various topics of research. For example, using fully indecomposable matrices, one may find a necessary condition for a matrix to have a positive inverse [5]. In recent years, many researchers such as Lim, Chang, Pearson, and Zhang generalized the concepts of non-negative matrices to tensors [4],[2]. In this paper, we generalize the concepts partly decomposable and fully indecomposable from non-negative matrices to tensors.
This paper is organized as follows. In Section 2, we review some definitions that will use in other sections. In section 3, we give some theorems about non-negative tensors. We conclude the paper with a brief conclusion in Section 4.

2. Preliminary

In this section, we first give some definitions that will be used in the proofs of our theorems.

*Talker
Email addresses: m.nobakht88@gmail.com (Malihe Nobakht-Kooshkghazi), afshin@vru.ac.ir (Hamidreza Afshin)

Definition 2.1 ([9]). Let $\mathscr{A} = (a_{i_1\dots i_d})$ be a tensor of order d and dimension $n_1 \times n_2 \times \cdots \times n_d$. The *permanent* of $\mathscr{A}$ is defined by

$$per(\mathscr{A}) := \sum_{\sigma_k} \prod_{i=1}^{n_1} a_{i\sigma_2(i)\dots\sigma_d(i)},$$

where the summation runs over all one-to-one functions σ_k from $\{1,\dots,n_1\}$ to $\{1,\dots,n_k\}$ and $k = 2,\dots,d$, with $per(\mathscr{A}) = 0$ if $n_1 > n_k$ for some k.

Definition 2.2 ([3]). Let $\mathscr{A}$ be a tensor of order d and dimension n. We say that $\mathscr{A}$ is reducible if there exists a nonempty proper index subset $J \subset [n]$ such that

$$a_{i_1\dots i_d} = 0, \qquad \forall i_1 \in J, \qquad \forall i_2,\dots,i_d \notin J.$$

The following definition generalizes block matrices to the setting of tensors.

Definition 2.3 ([7]). A block tensor is a tensor whose entries themselves are tensors.

Definition 2.4 ([1]). Let $\mathscr{A}$ be a tensor of order d and dimension n and, $n_1, n_2, \dots, n_r$ $(r \geq 2)$ be positive integers with $n = n_1 + n_2 + \cdots + n_r$. Write $S_0 = 0$ and

$$S_j = n_1 + n_2 + \cdots + n_j, \qquad I_j = \{S_{j-1}+1,\dots,S_j\}, \quad j = 1,\dots,r.$$

If for any $j \in \{2,\dots,r\}$,

$$a_{ii_2\dots i_d} = 0, \qquad \forall i \in I_j, \quad and \quad \max\{i_2,\dots,i_d\} \leq S_{j-1},$$

then we say that $\mathscr{A}$ is a third-type $(n_1,\dots,n_r)$-upper triangular block tensor (or a 3rdUTB tensor).

Definition 2.5 ([8]). Let $\mathscr{A}$ (respectively, $\mathscr{B}$) be a tensor of order $d \geq 2$ (respectively, $k \geq 1$) and dimension n. Define the product $\mathscr{A}\mathscr{B}$ to be the following tensor $\mathscr{C}$ of order $(d-1)(k-1)+1$ and dimension n.

$$c_{i\alpha_1\dots\alpha_{d-1}} = \sum_{i_2,\dots,i_d=1}^{n} a_{ii_2\dots i_d} b_{i_2\alpha_1} \dots b_{i_d\alpha_{d-1}}, \quad (i \in [n], \alpha_1,\dots,\alpha_{d-1} \in [n]^{k-1}).$$

Definition 2.6 ([8]). Let $\mathscr{A}$ be a tensor of order d and dimension n. If P and Q are $n \times n$ matrices, then

$$(P\mathscr{A}Q)_{i_1\dots i_d} = \sum_{j_1,\dots,j_d=1}^{n} a_{j_1\dots j_d} p_{i_1j_1} q_{j_2i_2} \dots q_{j_di_d}.$$

Definition 2.7 ([8]). Let $\mathscr{A}$ and $\mathscr{B}$ be tensors of order d and dimension n. If there exists a permutation matrix P such that $\mathscr{B} = P\mathscr{A}P^T$, then we say that $\mathscr{A}$ and $\mathscr{B}$ are permutation-similar.

Remark 2.8 ([1]). A tensor $\mathscr{A}$ of order d and dimension n is reducible if and only if there exists an integer k with $1 \leq k \leq n-1$ such that $\mathscr{A}$ is permutation-similar to some $(k, n-k)$-3rdUTB tensor.

3. Partly decomposable and fully indecomposable tensors

In this section, by according to Definition 2.7 and Remark 2.8, we define *permutation-equivalent* and *partly decomposable* tensors. Also, we compare these concepts to the matrix case.

Definition 3.1. Let $\mathscr{A}$ and $\mathscr{B}$ be tensors of order d and dimension n. If there exist $n \times n$ permutation matrices P and Q such that $\mathscr{B} = P\mathscr{A}Q$, then we say that $\mathscr{A}$ and $\mathscr{B}$ are permutation-equivalent.

Definition 3.2. Let $\mathscr{A}$ be a tensor of order d and dimension n. We say that $\mathscr{A}$ is partly decomposable if there exists an integer k with $1 \leq k \leq n-1$ such that $\mathscr{A}$ is permutation-equivalent to some $(k, n-k)$-3rdUTB tensor. A fully indecomposable tensor is one which is not partly decomposable.

Example 3.3. Let A be the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}.$$

It is clear that A is permutation-equivalent to a $(1,1)$-3rdUTB matrix with the following P and Q.

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$PAQ = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

Since PAQ is a $(1,1)$-3rdUTB tensor (tensor with $d = 2$), A is a partly decomposable matrix [6].

The following theorem present a necessary and sufficient condition for a tensor to be partly decomposable.

Theorem 3.4. *Let $\mathscr{A}$ be a tensor of order d and dimension n. The tensor $\mathscr{A}$ is partly decomposable if and only if there exist sequences $\lambda \in Q_{n-k,n}$ and $\mu \in Q_{k,n}$, where $Q_{k,n}$ is the set of non-decreasing subsequences with cardinality k of $\{1, 2, 3, \ldots, n\}$, such that $\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu]$ is a zero subtensor of order d and dimension $n - k \times k \times k \times \cdots \times k$ (or $\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu] = 0$).*

Proof. Let $\mathscr{A}$ be a partly decomposable tensor. Then, there exists an integer k with $1 \leq k \leq n-1$ such that $\mathscr{A}$ is permutation-equivalent to some $(k, n-k)$-3rdUTB tensor. Let $\mathscr{G} = P\mathscr{A}Q$, where P and Q are permutation matrices of order n, and $\mathscr{G}$ is a $(k, n-k)$-3rdUTB tensor. Define two sequences $\lambda_1 = \{k+1, k+2, \ldots, n\}$ and $\mu_1 = \{1, 2, \ldots, k\}$. Let τ and τ' be the permutation functions corresponding to the permutation matrices P and Q, respectively. Define

$$\lambda = \tau(\lambda_1) = \{\tau(k+1), \tau(k+2), \tau(k+3), \ldots, \tau(n)\}$$

and

$$\mu = \tau'(\mu_1) = \{\tau'(1), \tau'(2), \tau'(3), \ldots, \tau'(k)\}.$$

It is clear that $\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu]$ is a tensor of order d and dimension $n - k \times k \times k \times \cdots \times k$. Clearly,

$$\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu] = \mathscr{A}[\tau(k+1)\ldots\tau(n)|\tau'(1)\ldots\tau'(k)|\tau'(1)\ldots\tau'(k)|\ldots|\tau'(1)\ldots\tau'(k)].$$

Let $i \in \{k+1, k+2, k+3, \ldots, n\}$ and $j \in \{1, 2, \ldots, k\}$. In what follows, we calculate the entry of the tensor $\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu]$ with index $(i, j, j, \ldots, j)$.

$$\begin{aligned}
\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu](i, j, j, \ldots, j) &= \mathscr{A}(\tau(i), \tau'(j), \tau'(j), \ldots, \tau'(j)) \\
&= P^T \mathscr{G} Q^T (\tau(i), \tau'(j), \tau'(j), \ldots, \tau'(j)) \\
&= \sum_{j_1, j_2, \ldots, j_d = 1}^{n} \mathscr{G}_{j_1 j_2 \ldots j_d} P^T_{\tau(i) j_1} Q^T_{j_2 \tau'(j)} Q^T_{j_3 \tau'(j)} \cdots Q^T_{j_d \tau'(j)} \\
&= \mathscr{G}_{ijj\ldots j} P^T_{\tau(i) i} Q^T_{j\tau'(j)} Q^T_{j\tau'(j)} \cdots Q^T_{j\tau'(j)} \\
&= \mathscr{G}_{ijj\ldots j} \\
&= 0.
\end{aligned}$$

This completes the proof.
Conversely, suppose that λ, μ are subsets of $\{1, 2, \ldots, n\}$ such that $\lambda \in Q_{n-k,n}$, $\mu \in Q_{k,n}$ and $\mathscr{A}[\lambda|\mu|\mu|\ldots|\mu] = 0$. Let

$$\lambda = \{i_{k+1}, i_{k+2}, \ldots, i_n\}, \qquad \mu = \{i_1, i_2, \ldots, i_k\},$$

where the subsets λ and μ are not necessarily disjoint. Let τ be a permutation function such that $\tau(i_l) = l$ for all $l = 1, 2, \ldots, k$. Let τ' be a permutation function such that $\tau'(j) = i_j$ for all $j = k+1, k+2, \ldots, n$. Also, let Q and P be the corresponding permutation matrices of τ and τ', respectively. If $k+1 \le s \le n$ and $1 \le t \le k$, then

$$\begin{aligned}(P\mathscr{A}Q)_{stt\ldots t} &= \sum_{j_1,j_2,\ldots,j_d=1}^{n} a_{j_1j_2\ldots j_d} P_{sj_1} Q_{j_2t} Q_{j_3t} \ldots Q_{j_dt} \\ &= a_{\tau'(s)\tau^{-1}(t)\ldots\tau^{-1}(t)} \\ &= a_{i_si_ti_t\ldots i_t} \\ &= 0.\end{aligned}$$

Thus, $P\mathscr{A}Q = \mathscr{G}$, where $\mathscr{G}$ is a $(k, n-k)$3rdUTB tensor. □

Example 3.5. Let $\mathscr{A}$ be a non-negative tensor of order 3 and dimension 3 with positive entries

$$a_{111}, a_{311}, a_{122}, a_{222}, a_{233}, a_{333}$$

whose all other entries are equal to 0. It follows from Theorem 3.4 that $\mathscr{A}$ is a fully indecomposable tensor.

Clearly, the zero tensor of order d and dimension 1 is partly decomposable, while a nonzero tensor of order d and dimension 1 is fully indecomposable by Theorem 3.4.

Proposition 3.6. *Let $\mathscr{A}$ be a non-negative tensor of order d and dimension n. If $\mathscr{A}$ is reducible, then it is partly decomposable.*

Proof. This follows from Remark 2.8 and Definition 3.2. □

We show in the following remark that it is not true that a non-negative fully indecomposable tensor satisfies in the below property and converse

$$per(\mathscr{A}(i_1|\ldots|i_d)), \quad \forall i_1, \ldots, i_d = 1, \ldots, n. \tag{1}$$

Remark 3.7. Let $\mathscr{A}$ be a tensor of order 3 and dimension 2 with positive entries $a_{111}, a_{211}, a_{212}, a_{121}, a_{122}, a_{221}, a_{222}$ and $a_{112} = 0$. Clearly, $\mathscr{A}$ is a fully indecomposable tensor. But, since $per(\mathscr{A}(2|2|1)) = a_{112} = 0$, it does not satisfy (1). Also, let $\mathscr{A}$ be a non-negative tensor of order 3 and dimension 3 whose all entries, except of a_{211} and a_{311}, are positive numbers (that is, $a_{211} = a_{311} = 0$). It is clear that $\mathscr{A}$ satisfies (1), but it is a partly decomposable tensor. We know that in the case $d = 2$, A is a fully indecomposable matrix if and only if $per(A(i|j)) > 0$ for all $i, j = 1, \ldots, n$ [6].

Theorem 3.8. *Let $\mathscr{A}$ be a non-negative block tensor of order d and dimension n. Let $n = n_1 + \cdots + n_r$, where $n_i > 0$ and $r \ge 2$. Suppose that $1 : n$ has exactly one partition of the form below for the index $i_j = 1, \ldots, n$, for each $j = 1, 2, \ldots, d$.*

$$1 : n = [1, \ldots, n_1 | n_1 + 1, \ldots, n_1 + n_2 | \ldots\ldots | n_1 + \cdots + n_{r-1} + 1, \ldots, n_1 + \cdots + n_r].$$

For each $i \in \{1, \ldots, r\}$, we denote the block $\mathscr{A}_{i\ldots i}$ of $\mathscr{A}$ by $\mathscr{A}_i$. Also, for each $i \in \{1, \ldots, r-1\}$, we denote the block $\mathscr{A}_{ii+1\ldots i+1}$ of $\mathscr{A}$ by $\mathscr{B}_i$. Finally, we use $\mathscr{B}_r$ to denote the block $\mathscr{A}_{r1\ldots 1}$ of $\mathscr{A}$. Let $\mathscr{A}_i$ be a tensor of order d and dimension n_i, $\mathscr{B}_i$ be a tensor of order d and dimension $n_i \times n_{i+1} \times n_{i+1} \times \cdots \times n_{i+1}$, where $i = 1, \ldots, r-1$, and $\mathscr{B}_r$ be a tensor of order d and dimension $n_r \times n_1 \times n_1 \times \cdots \times n_1$. In other words,

$$\begin{aligned}\mathscr{A}_i &= \mathscr{A}_{\underbrace{i\ldots i}_{d}}, & i = 1, \ldots, r, \\ \mathscr{B}_i &= \mathscr{A}_{i\underbrace{i+1\ldots i+1}_{d-1}}, & i = 1, \ldots, r-1, \\ \mathscr{B}_r &= \mathscr{A}_{r\underbrace{1\ldots 1}_{d-1}}.\end{aligned}$$

If $\mathscr{A}_i$ is a fully indecomposable tensor and $\mathscr{B}_i \neq 0$ for all $i = 1, \ldots, r$, then $\mathscr{A}$ is a fully indecomposable tensor.

Proof. Let $\mathscr{A}$ be a partly decomposable tensor. Then, there exists an integer k with $1 \leq k \leq n-1$ such that $\mathscr{A}$ is permutation-equivalent to some $(k, n-k)$-3rdUTB tensor. Thus, $\mathscr{A}[\alpha|\beta|\ldots|\beta] = 0$ for $s+t=n$, where $\alpha \in Q_{s,n}$ and $\beta \in Q_{t,n}$. Suppose that s_j entries of α and t_j entries of β intersect the subtensor $\mathscr{A}_j$, where $j = 1, \ldots, r$ (that is, $\mathscr{A}_j[\alpha_{s_1}, \ldots, \alpha_{s_j}|\beta_{t_1}, \ldots, \beta_{t_j}|\ldots|\beta_{t_1}, \ldots, \beta_{t_j}] = 0$). Since $\mathscr{A}_j$ is fully indecomposable, we must have $s_j + t_j \leq n_j$, where $s_j + t_j = n_j$ if $s_j = 0$ or $t_j = 0$. Now,

$$\begin{aligned} n &= s+t \\ &= \sum_{j=1}^{r} s_j + t_j \\ &\leq \sum_{j=1}^{r} n_j \\ &= n. \end{aligned}$$

Thus, $s_j + t_j = n_j$ for $j = 1, \ldots, r$. Since $\mathscr{A}_j$ is fully indecomposable, $s_j = 0$ or $t_j = 0$ for all $j = 1, \ldots, r$. It is clear that at least one of the s_j's and at least one of the t_j's must be positive, because $s_1 + \ldots, s_r = s \geq 1$ and $t_1 + \ldots, t_r = t \geq 1$. Thus, there exists at least one integer k such that $s_k = n_k$ and $t_{k+1} = n_{k+1}$ (Since otherwise, $t_1 = \cdots = t_r = 0$ and hence $t = 0$, which contradicts $t \geq 1$.) Thus, there exists a zero subtensor that contains the subtensor $\mathscr{B}_k$. Hence, $\mathscr{B}_k = 0$ and this contradicts our hypothesis. □

Theorem 3.9. *A non-negative, irreducible tensor with a positive main diagonal is fully indecomposable tensor.*

Proof. Let $\mathscr{A}$ be a non-negative, irreducible tensor with a positive main diagonal. If $\mathscr{A}$ is partly decomposable, then $\mathscr{A}[\alpha|\beta|\ldots|\beta] = 0$, where $\alpha \in Q_{n-k,n}$ and $\beta \in Q_{k,n}$. Since the entries of the main diagonal of $\mathscr{A}$ are positive, the sets α and β should be disjoint. Hence, $\mathscr{A}$ is a non-negative reducible tensor. But, this contradicts our hypothesis. □

It is not true that a non-negative, fully indecomposable tensor is permutation-equivalent to an irreducible tensor with a positive main diagonal.

Remark 3.10. Let $\mathscr{A}$ be a non-negative tensor with positive entries

$$a_{111}, a_{211}, a_{312}, a_{313}, a_{122}, a_{222}, a_{133}, a_{233}, a_{332}$$

whose all other entries are equal to 0. Then, $\mathscr{A}$ is a non-negative, fully indecomposable tensor which is not permutation-equivalent to an irreducible tensor with a positive main diagonal.

4. Conclusion

We defined partly decomposable and fully indecomposable tensors, and we proposed a necessary and sufficient condition for a tensor to be partly decomposable. We obtained a sufficient condition for block tensors to be fully indecomposable, and we found some relations between fully indecomposable and irreducible tensors. We compared the relations between fully indecomposable and irreducible tensors with the case of matrices using a counterexample.

References

[1] L. You J. Shao. On some properties of three different types of triangular blocked tensors. *Linear Algebra Appl.*, 511:110–140, 2016.
[2] T. Zhang K.C. Chang, K. Pearson. Perron-frobenius theorem for nonnegative tensors. *Commun. Math. Sci.*, 6(2):507–520, 2008.
[3] Z. Luo L. Qi. Tensor analysis: Spectral theory and special tensors, society for industrial and applied mathematics. *Philadelphia*, 7, 2017.
[4] L.H. Lim. Singular values and eigenvalues of tensors: a variational approach. *In 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, pages 129–132, 2005.
[5] H. Minc. On matrices with positive inverses. *Zastos. Mat.*, 7:165–169, 1963.
[6] H. Minc. Nonnegative matrices. *A Wiley-Interscience, New York*, pages 129–132, 1978.
[7] C.F. Van Loan S. Ragnarsson. Block tensor unfoldings. *IAM J. Matrix Anal. Appl.*, 33(1):149–169, 2012.
[8] J. Y. Shao. A general product of tensors with applications. *Linear Algebra Appl.*, 439(8):2350–2366, 2013.
[9] P.M. Gibson S.J. Dow. Permanents of d-dimensional matrices. *Linear Algebra Appl.*, 90:133–145, 1987.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–120, pp. 45–47

Quasi-multipliers on certain class of operators

Marjan Adib[a]

[a] *Department of Mathematics, Payamenoor University (PNU), Tehran, Iran.*

Article Info

Keywords:
Quasi-multiplier
multiplier
Banach algebra
second dual
Arens regularity

2020 MSC:
47B48
46H25

Abstract

In this paper we extend the notion of quasi-multipliers to the dual of a Banach algebra A which second dual has a mixed identity. We consider algebras satisfying weaker condition than Arens regularity. Among others we prove that for an Arens regular Banach algebra with a b.a.i., $QM(A^*)$ is isometrically isomorphic to A^{**} and apply our results to the space of trace class operators and the group algebra of a compact group.

1. Introduction

2. Main results

Definition 2.1. A bilinear map $m: A^* \times A^{**} \to A^*$ is a *right quasi-multiplier of* A^* if

$$m(F \cdot \xi, G) = F \cdot m(\xi, G) \qquad \text{and} \qquad m(\xi, G \circ F) = m(\xi, G) \cdot F \tag{1}$$

hold for arbitrary $\xi \in A^*$ and $F,\ G \in A^{**}$.
Similarly, a bilinear map $m': A^{**} \times A^* \to A^*$ is a *left quasi-multiplier of* A^* if

$$m'(F \circ G, \xi) = F \cdot m'(G, \xi) \qquad \text{and} \qquad m'(G, \xi \cdot F) = m'(G, \xi) \cdot F$$

hold for arbitrary $\xi \in A^*$ and $F,\ G \in A^{**}$.

Although in our investigation we do not assume that Arens regularity, we usually have to assume that the given algebra satisfies the following weaker condition. A Banach algebra A satisfies condition (K) if

$$(F \cdot \xi) \cdot G = F \cdot (\xi \cdot G) \qquad (F,\ G \in A^{**},\ \xi \in A^*).$$

Email address: m.adib.pnu@gmail.com (Marjan Adib)

Let $QM_r(A^*)$ be the set of all separately continuous right quasi-multipliers of A^*. It is obvious that $QM_r(A^*)$ is a linear space. Moreover, it is a Banach space with respect to the norm

$$||m|| = \sup\{||m(\xi, F)||; \quad \xi \in A^*,\ F \in A^{**},\ ||\xi|| \leq 1,\ ||F|| \leq 1\}.$$

Of course, the same holds for $QM_l(A^*)$, the set of all separately continuous left quasi-multipliers of A^*.
Let A be a Banach algebra. Recall that a map $T : A^* \to A^*$ is called a right multiplier of A^* if

$$T(F \cdot \xi) = F \cdot T(\xi),$$

for all $\xi \in A^*, F \in A^{**}$. With $M_r(A^*)$ we denote the space of all bounded linear right multipliers on A^*.

Theorem 2.2. *Assume that A is a Banach algebra satisfying condition (K) and A^{**} has a mixed identity. Then*

$$\rho_T(\xi, F) = (T\xi) \cdot F \qquad (T \in M_r(A^*),\ \xi \in A^*,\ F \in A^{**})$$

defines an injective linear map $\rho :\ M_r(A^*) \to QM_r(A^*)$ with norm $\|\rho\| \leq 1$. Moreover, ρ is onto if A^{**} has an identity. If A^{**} has a mixed identity with norm one, then ρ is an isometry.

Proof. Let $T \in M_r(A^*)$ be arbitrary. It is obvious that ρ_T is a bilinear map from $A^* \times A^{**}$ to A^* and that it is bounded with $\|T\|$. For $a \in A$, $\xi \in A^*$, and $F,\ G \in A^{**}$, we have

$$\rho_T(F \cdot \xi, G) = T(F \cdot \xi) \cdot G = (F \cdot T\xi) \cdot G = F \cdot (T\xi \cdot G) = F \cdot \rho_T(\xi, G)$$

and

$$\rho_T(\xi, G \circ F) = (T\xi) \cdot (G \circ F) = (T\xi \cdot G) \cdot F = \rho_T(\xi, G) \cdot F.$$

Thus, $\rho_T \in QM_r(A^*)$. It follows from the definition that $\rho :\ M_r(A^*) \to QM_r(A^*)$ is linear. Obviously, $\|\rho_T\| \leq \|T\|$, which gives $\|\rho\| \leq 1$. Let $E \in A^{**}$ be a mixed identity. If $\rho_T = 0$, then we have $(T\xi) \cdot E = 0$ for every $\xi \in A^*$ and consequently $T = 0$. Assume that E is an identity for A^{**}. Let $m \in QM_r(A^*)$ be arbitrary. It is easily seen that $T\xi = m(\xi, E)$ $(\xi \in A^*)$ defines a bounded right multiplier of A^*. Since equalities $\rho_T(\xi, F) = (T\xi) \cdot F = m(\xi, E) \cdot F = m(\xi, E \circ F) = m(\xi, F)$ hold for all $\xi \in A^*$ and $F \in A^{**}$ we conclude that ρ is onto.
At the end assume that E is mixed identity for A^{**} of norm one. Let $T \in M_r(A^*)$ and $\varepsilon > 0$ be arbitrary. If $\xi \in A^*$ is such that $\|\xi\| \leq 1$ and $\|T\| - \varepsilon < \|T\xi\|$, then

$$\|\rho_T\| \geq \|\rho_T(\xi, E)\| = \|T\xi\| > \|T\| - \varepsilon.$$

Thus, ρ is an isometry. □

Corollary 2.3. *If A is a C^*-algebra, then ρ is an isometrical isomorphism from $M_r(A^*)$ onto $QM_r(A^*)$.*

Proof. It is well known that every C^*-algebra is Arens regular and has b.a.i. Thus, A satisfies condition (K) and its dual A^{**} is unital. □

If A is a Banach algebra satisfying condition (K) and A^{**} has an identity, then Theorem 2.2 allows a natural definition of multiplication in $QM_r(A^*)$. Namely, for arbitrary $m_1,\ m_2 \in QM_r(A^*)$, let $T_1,\ T_2 \in M_r(A^*)$ be uniquely determined multipliers satisfying $m_1 = \rho(T_1)$ and $m_2 = \rho(T_2)$. Then

$$m_1 \circ_\rho m_2 = \rho(T_1) \circ_\rho \rho(T_2) := \rho(T_2 T_1)$$

gives a well defined multiplication. It is easy to see that $QM_r(A^*)$ is a unital Banach algebra.
Note that $QM_l(A^*)$ as well has a natural multiplication if A is a Banach algebra satisfying condition (K) and A^{**} has a mixed identity. Indeed, let $M_l(A^*)$ be the space of all bounded left multipliers on A^*, i.e., bounded linear operators T on A^* satisfying $T(\xi \cdot F) = T\xi \cdot F$, for all $\xi \in A^*$ and $F \in A^{**}$. A similar reasoning as in Theorem 2.2 shows that the mapping $\lambda :\ M_l(A^*) \to QM_l(A^*)$, which is defined by

$$\lambda_S(F, \xi) = F \cdot S\xi \qquad (S \in M_l(A^*),\ \xi \in A^*,\ F \in A^{**}),$$

is a linear bijection. Thus, a natural multiplication on $QM_l(A^*)$ is given by $\lambda(S_1) \circ_\lambda \lambda(S_2) := \lambda(S_1 S_2)$.

Theorem 2.4. *Let A be a Banach algebra such that A^{**} has an identity E and hypothesis (K) is valid in A. Assume A^* factors on the right. Then there exist an isomorphism of A^{**} onto $QM_r(A^*)$.*

Proof. Define a map $\psi : A^{**} \to QM_r(A^*)$ by $\psi(H) = \rho(R_H)$, where R_H is the right multiplication operator on A^* determined by $H \in A^{**}$. Then, for arbitrary $\xi \in A^*, F \in A^{**}$,

$$\psi(H)(\xi, F) := (\xi \cdot H) \circ F.$$

It is evident that ψ is linear and continuous. We check the multiplicativity of ψ. Let $H_1, H_2 \in A^{**}$. By Theorem 2.2, there exist $T_1, T_2 \in M_r(A^*)$ such that $\psi(H_1) = \rho(T_1)$ and $\psi(H_2) = \rho(T_2)$. Hence, for arbitrary $\xi \in A^*, F \in A^{**}$, we have

$$T_1(\xi) \cdot F = (\xi \cdot H_1) \circ F \qquad \text{and} \qquad T_2(\xi) \cdot F = (\xi \cdot H_2) \circ F).$$

It follows

$$\begin{aligned}(\psi(H_1) \circ_\rho \psi(H_2))(\xi, F) &= (\rho(T_1) \circ_\rho \rho(T_2))(\xi, F) = \rho(T_2 T_1)(\xi, F)\\ &= T_2(T_1(\xi)) \circ F = T_1\xi \cdot (H_2 \circ F)\\ &= \xi \cdot (H_1 \circ H_2 \circ F) = \psi(H_1 \circ H_2)(\xi, F),\end{aligned}$$

which means ψ is a homomorphism.
Assume that $\psi(H) = 0$ for $H \in A^{**}$. Since the mapping ρ is one to one $R_H = 0$. Hence, for each $\xi \in A^*$, one has $\xi \circ H = 0$. By the assumption, A^* factors on the right, which implies $H = 0$. Thus, ψ is one to one. Homomorphism ψ is onto as well. If $m \in QM_r(A^*)$, then there exist $T \in M_r(A^*)$ such that $m = \rho(T) = \rho(R_{T^*(E)})$. This means that ψ is onto. ☐

Corollary 2.5. *Let H be a Hilbert space and let $A = K(H)$, the algebra of all compact operators on H. The dual of the space of compact operators is the space of trace-class operators, $C_1(H)$. The second dual of A is $B(H)$. Since $K(H)$ is a C^*-algebra we have $QM_r(C_1(H)) \cong B(H)$.*

At the end we consider the group algebra of a compact group G. By [6], $L_1(G)$ is Arens regular if and only if G is finite. However, since $L_1(G)$ is a two-sided ideal in its second dual ([5]), it satisfies condition (K). Note that the dual $L_1(G)^*$ can be identified with $L_\infty(G)$.
Let $M(G)$ be the convolution algebra of all bounded regular measures on G. Recall that the convolution product of $f \in L_1(G)$ and $\mu \in M(G)$ is given by

$$f * \mu(x) = \int_G f(xy^{-1})\, d\mu(y).$$

Of course, $L_\infty(G)$ is a Banach $L_1(G)^{**}$-bimodule. However, the space $L_\infty(G)$ has also a natural structure of a Banach $M(G)$-bimodule. The same holds for $L_\infty(G)^* = L_1(G)^{**}$. We will denote all these module multiplications by $*$.

Proposition 2.6. *Let G be a compact group and $A = L_1(G)$. Then the equation*

$$(\theta_\mu(\xi, F) := (\xi * \mu) * F \qquad (\mu \in M(G), \xi \in L_\infty(G),\ F \in L_1(G)^{**}).$$

defines a linear isomorphism between $M(G)$ and a subspace of $QM_r(A^*)$.

Proof. Note that by definition of module action $(\xi * \mu) * F = \xi * (\mu * F)$. From this and condition (K) we conclude that $\theta_\mu \in QM_r(L_1(G)^*)$. Of course, $\theta : M(G) \to QM_r(L_1(G)^*)$ is a bounded linear map. We claim that θ is injective. Indeed, suppose that $\theta_\mu = 0$. Then $(\xi * \mu) * F = 0$ for all $\xi \in L_\infty(G)$ and $F \in (L_\infty(G))^*$. Since $L_1(G)$ has a b.a.i. it follows $\xi \circ \mu = 0$. ☐

References

[1] C. A. Akemann and G. K. Pedersen, *Complications of semicontinuity in C^*-algebra theory*, Duke Math. J. **40** (1973), 785–795.
[2] M. Kaneda, *Quasi-multipliers and algebrizations of an operator space*, J. Func. Anal **251** (2007), 346–359.
[3] G. Köthe, *Topological Vector Space I*, I. New York-Heidelberg-Berlin: Springer, (1969).
[4] M. McKennon, *Quasi-multipliers*, Trans. Amer. Math. Soc. **233** (1977), 105–123.
[5] S. Watanabe, *A Banach algebra which is an ideal in the second dual space*, Sci. Rep. Niigata Univ. Ser. **A 11** (1974), 95–101.
[6] N. Young, *The irregularity of multiplication in group algebras*, Quart. J. Math. Oxford **24** (1973), 59–62.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–124, pp. 48–55

The Hyers-Ulam-Rassias Stability of the Cauchy Linear Functional Equation in Random Normed Spaces

Mohammd. Hassn. Golmohammadi[a,*], Abbas. Zareei[a]

[a] *Address: Department of Mathematics, Payame Noor University, P. O. Bax 19395 - 4697 Tehran, Iran*

Article Info

Keywords:
Hyers-Ulam
Random Normed
fixed point
direct method

2020 MSC:
msc1
msc2

Abstract

Using the fixed point and direct method, we prove the generalized Hyers-Ulam stability in Random Normed Spaces of the following Cauchy linear functional equation

$$f(x+y+a) = f(x) + f(y)$$

where $f : X \longrightarrow Y$ and a is an arbitrary element in X.

1. First Section

2. Introduction

The stability problem of functional equations originated from a quationn of Ulam [18] concerning the stability of group homomorphisms.Hyers [8] gave a first affirmative answer to the question of Ulam for Banach spaces. Aoki[1] and Th.M.Rassias [14] proved a generalization of the Hyer's theorem for additive and linear mapping, respectively, by allowing the Cauchy di erence to be un- bounded. P.Gavruta [6] proved a further generalization of the Th.M.Rassias' theorem by using a general control function.The Cauchy linear functional equation.
The functional equation

$$f(x+y+a) = f(x) + f(y), x, y \in X \tag{1}$$

A particular case of this linear functional equation is

$$f(x+y) = f(x) + f(y), x, y \in X \tag{2}$$

If f is a solution of (1.2) it is said to be additive or satis es the Cauchy equation, In [17] S.M.Ulam posed the question of the stability of Cauchy equation: If a function f approximately statis es Cauchy's functional equation (1.2) when dose

*Talker
Email addresses: first.golmohamadi@pnu.ac.ir (Mohammd. Hassn. Golmohammadi), second.Zarei@pnu.ac.ir (Abbas. Zareei)

there exists an exact solution of (1.2) which f approximates. The priblem has been consider for many di erent types of writers including D. H. Hyers[9,10,11], Th.M.Rassias [11,14], Z. Gajda [5] and P. Gaveruta [7]. The interested reader should refer to the book by D. H. Hyers G. Isac and Th. M. Rassias[11]. B.Belaid and E.Elhoucien [3], introduced the Cauchy linear functional equation (1.1) and they established the general solution and provided a proof of functional stability in the spirit of Hyers-Ulam, Th.M.Rassias and P.Gavruta. In the section 2, we use direct method to prove the generalized Hyers-Ulam stability of general functional equations (1.1) in random normed spaces and in the section 3, we use Fixed point method to prove the generalized Hyers- Ulam stability of general functional equations (1.1) random normed spaces. Throughout this paper, the spaces of all probability distribution function is denoted by ∆+. Elements of ∆+ are functions F : R □[1; +1] ! [0; 1], such that F is left continuous and nondecreasing on R and F (0) = 0; F (+1) = 1. It's clear that the subset. Throughout this paper, the spaces of all probability distribution function is denoted by Δ^+. Elements of Δ^+ are functions $F : \mathbb{R}\bigcup[-\infty,+\infty] \to [0,1]$, such that F is left continuous and nondecreasing on $\mathbb{R}$ and $F(0) = 0, F(+\infty) = 1$. It's clear that the subset

$$D^+ = \{F \in \Delta^+ : l^- F(-\infty) = 1\},$$

where $l^- f(x) = \lim_{t\to x^-} f(t)$. is a subset of Δ^+ . The spaces Δ^+ is partially ordered by the usual pointwise ordering of functions, that is for all $t \in \mathbb{R}, F \leq G$ if and onaly if $F(t) \leq G(t)$.For every $a \geq 0$, $H_a(t)$ is the element of D^+ defined by

$$H_a(t) = \begin{cases} 0 & t \leq a \\ 1 & \text{t > a} \end{cases}$$

On can easily show that the maximal element for Δ^+ in this order is the distribution functional $H_0(t)$.

Definition 2.1. A faction $T : [0,1] \times [0,1] \to [0,1]$ is a continuous triangular norm (briefly a t-norm) if T satisfies the following conditions:

(1) T is commutative and associative;
(2) T is continuous;
(3) $T(x,1) = x$ for all $x \in [0,1]$;
(4) $T(x,y) \leq T(z,w)$ whenever $x \leq z$ and $y \leq w$ for all $x,y,z,w \in [0,1]$.

Three typical examples of continuous t-norms are $T(x,y) = xy, T(x,y) = max\{a+b-1\}$ and $T(x,y) = min(a,b)$. Recall that, if T is a t-norm and $\{x_n\}$ is a given of numbers in $[0,1]$, $T^n_{i=1}x_i$ is defined recursively by $T^1_{i=1}x_1$ and $T^n_{i=1}x_i = T(T^{n-1}_{i=1}x_i, x_n)$ for $n \geq 2$.

Definition 2.2. A random normed spaces(briefly RN-space) is a triple (X,Φ,T), where X is a vector space, T is a continuous t-norm and $\Phi : X \to D^+$ is a maping such that the following conditions hold:

(1) $\Phi_x(t) = H_0(t)$ for all $t > 0$ if and only if $x = 0$.
(2) $\Phi_{\alpha x}(t) = \Phi_x(\frac{t}{|\alpha|})$ for all $\alpha \in \mathbb{R}, \alpha \neq 0, x \in X$ and $t \geq 0$.
(3) $\Phi_{x+y}(t+s) \geq T(\Phi_x(t), \Phi_y(s))$, for all $x,y \in X$ and $t,s \geq 0$.

Example 2.3. Every normed spaces $(X, \|.\|)$ defines a random normed space (X,Φ,T_M) where for every $t > 0$,

$$\Phi_u(t) = \frac{t}{t + \|u\|}$$

and T_M is the minimum t-norm. This space is called the induced random normed space.

Example 2.4. Let $(X, \|.\|)$ be a normed linear space and $\alpha, \beta > 0$, and

$$F_x(t) = \begin{cases} \frac{\alpha t}{\alpha t + \beta\|x\|} & x \in X, t > 0 \\ 0 & \text{x} \in \text{X.t} \leq 0 \end{cases}$$

Then (X,F,T_M) is a random normed space.

Definition 2.5. Let (X, Φ, T) be an RN-space.
(1) A sequence $x_n \in X$ is said to be convergent to $x \in X$ if for all $t > 0$, $lim_{n\to\infty}\Phi_{x_n - x}(t) = 1$.
(2) A sequence $\{x_n\}$ in X is said to be Cauchy sequence in X if for all $t > 0$,$lim_{n\to\infty}\Phi_{x_n - x}(t) = 1$.
(3) The RN-space (X, Φ, T) is said to be complete if every Cauchy sequence in X is convergent.

Theorem 2.6. *If (X, Φ, T) is RN-space and $\{x_n\}$ is a sequence such that $x_n \to x$, then $lim_{n\to\infty}\Phi_{x_n}(t) = \Phi_x(t)$.*

Definition 2.7. Let X be a set. A function $d : X \times X \to [0, \infty]$ is called a generalized metric on X if d satisfies
(1) $d(x, y) = 0$ if and only if $x = y$,
(2) $d(x, y) = d(y, x)$ for all $x, y \in X$,
(3) $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

Theorem 2.8. *Let (X, d) be a complete generalized metric space and let $J : X \to X$ be a strictly contractive mapping with Lipschitz constant $L < 1$. Then for each given element $x \in X$, either*

$$d(J^n x, J^{n+1}x) = \infty$$

for all nonnegative integers n or there exists a positive integer n_0 such that
(1) $d(J^n x, J^{n+1}x) < \infty$ for all $n \geq n_0$;
(2) the sequence $\{J^n x\}$converges to a fixed point y^* of J;
(3) y^* is the uniqe fixed point of J in the set $Y = \{y \in X|d(J^{n_0}x, y) < \infty\}$;
(4) $d(y, y^*) \leq \frac{1}{1-L}d(y, Jy)$ for all $y \in Y$.

3. RN-stability of the functional equation (1): direct method

In this section, we use direct method to prove the generalized Hyers-Ulam stability of general quadratic functional equations (1).

Theorem 3.1. *Let X be a real linear space, (Z, Ψ, min) be an RN-space, $\varphi : X^2 \to Z$ be a function such that for some $0 < \alpha < 2$,*

$$\Psi_{\varphi(2x+a,2y+a)}(t) \geq \Psi_{\alpha\varphi(x,y)}(t) \qquad \forall x \in X, t > 0, \tag{3}$$

for all $x, y \in X$ and $t > 0$

$$\lim_{n\to\infty} \Psi_{\varphi(2^n x+(2^n-1)a,2^n y+(2^n-1)a)}(2^n t) = 1.$$

Let (Y, μ, min) be a complete RN-space. If $f : X \to Y$ is a mapping such that for all $x, y \in X$ and $t > 0$,

$$\mu_{f(x+y+a)-f(x)-f(y)}(t) \geq \Psi_{\varphi(x,y)}(t), \tag{4}$$

then there is a unique additive mapping $R : X \to Y$ such that

$$R(x) := \lim_{n\to\infty} \frac{1}{2^n} f(2^n x + (2^n - 1)a),$$

and

$$\mu_{f(x)-A(x)}(t) \geq \Psi_{\varphi(x,x)}\left((2 - \alpha)t\right). \tag{5}$$

Proof. Puting $y = x$ in(4) we see that for all $x \in X$ and $t > 0$,

$$\mu_{\frac{f(2x+a)}{2} - f(x)}(t) \geq \Psi_{\varphi(x,x)}(2t). \tag{6}$$

Replacing x by $2^n x + (2^n - 1)a$ in (6) and using (3), we obtain

$$\begin{aligned} &\mu_{\frac{1}{2^{n}+1}} f(2^{n+1}x + (2^{n+1} - 1)a) - \frac{1}{2^n} f(2^n x + (2^n - 1)a)(t) \\ &\geq \Psi_{\varphi(2^n x+(2^n-1)a),(2^n x+(2^n-1)a))}(2 \times 2^n t) \geq \Psi_{\varphi(x,x)}(\frac{2 \times r^n}{\alpha^n} t). \end{aligned} \tag{7}$$

so

$$\begin{aligned}
\mu_{\frac{f(r^n x)}{r^{2n}}-f(x)}\left(\sum_{k=0}^{n-1}\frac{t\alpha^k}{r^2\times r^{2k}}\right) &= \mu_{\sum_{k=0}^{n-1}\frac{f(r^{k+1}x)}{r^{2k+2}}-\frac{f(r^k x)}{r^{2k}}}\left(\sum_{k=0}^{n-1}\frac{t\alpha^k}{r^2\times r^{2k}}\right) \qquad (8)\\
&\geq T_{k=0}^{n-1}\mu_{\frac{f(r^{k+1}x)}{r^{2k+2}}-\frac{f(r^k x)}{r^{2k}}}\left(\frac{t\alpha^k}{r^2\times r^{2k}}\right)\\
&\geq T_{k=0}^{n-1}\left(\Psi_{\varphi(x,0)}(t)\right)\\
&= \Psi_{\varphi(x,0)}(t).
\end{aligned} \qquad (9)$$

This implies that

$$\mu_{\frac{f(r^n x)}{r^{2n}}-f(x)}(t)\geq\Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=0}^{n-1}\frac{\alpha^k}{r^{2k+2}}}\right). \qquad (10)$$

Replacing x by $r^p x$ in (8), we obtain

$$\mu_{\frac{f(r^{n+p}x)}{r^{2n+2p}}-\frac{f(r^p x)}{r^{2p}}}(t) \qquad (11)$$

$$nonumber \geq \Psi_{\varphi(r^p x,0)}\left(\frac{t}{\sum_{k=0}^{n-1}\frac{\alpha^k}{r^2\times r^{2k+2p}}}\right) \qquad (12)$$

$$\begin{aligned}
nonumber \qquad &\geq \Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=0}^{n-1}\frac{\alpha^{k+p}}{r^2\times r^{2k+2p}}}\right) \qquad (13)\\
&= \Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=p}^{n+p-1}\frac{\alpha^k}{r^2\times r^{2k+2p}}}\right).
\end{aligned}$$

As

$$\lim_{p,n\to\infty}\Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=p}^{n+p-1}\frac{\alpha^k}{r^{2k+2}}}\right)=1,$$

then $\{\frac{f(r^n x)}{r^{2n}}\}_{n=1}^{\infty}$ is a Cauchy sequence in complete RN-space (Y,μ,min),so there exist some point $R(x)\in Y$ such that

$$\lim_{n\to\infty}r^{-2n}f(r^n x)=R(x).$$

Fix $x\in X$ and put $p=0$ in (11). Then we obtain

$$\mu_{\frac{f(r^n x)}{r^{2n}}-f(x)}(t)\geq\Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=0}^{n-1}\frac{\alpha^k}{r^{2k+2}}}\right), \qquad (14)$$

and so, for every $\epsilon>0$, we have

$$\begin{aligned}
\mu_{f(x)-R(x)}(t+\epsilon) &\geq T\left(\mu_{R(x)-\frac{f(r^n x)}{r^{2n}}}(\epsilon),\mu_{\frac{f(r^n x)}{r^{2n}}-f(x)}(t)\right) \qquad (15)\\
&\geq T\left(\mu_{R(x)-\frac{f(r^n x)}{r^{2n}}}(\epsilon),\Psi_{\varphi(x,0)}\left(\frac{t}{\sum_{k=0}^{n-1}\frac{\alpha^k}{r^{2k+2}}}\right)\right).
\end{aligned}$$

Taking the Limit as $n\to\infty$ and using (15), we get

$$\mu_{R(x)-f(x)}(t+\epsilon)\geq\Psi_{\varphi(x)}((r^2-\alpha)t). \qquad (16)$$

Since ϵ was arbitrary by taking $\epsilon\to 0$ in (17), we get

$$\mu_{R(x)-f(x)}(t)\geq\Psi_{\varphi(x,0)}((r^2-\alpha)t). \qquad (17)$$

Replacing x and y by $r^n x$ and $r^n y$, respectively, in (4), we get for all $x, y \in X$ and for all $t > 0$,

$$\mu_{\frac{1}{r^{2n}}\left[f(r^{n+1}x+r^n sy)-r^2 f(r^n x)-s^2 f(r^n y)-\frac{rs}{2}[f(r^n(x+y))-f(r^n(x-y))]\right]}(t) \geq \Psi_{\varphi(r^n x, r^n y)}(r^{2n}t), \tag{18}$$

Since $\lim_{n\to\infty} \Psi_{\varphi(r^n x, r^n y)}(r^n t) = 1.$, we conclude that

$$f(rx+sy) = r^2 f(x) + s^2 f(y) + \frac{rs}{2}[f(x+y) - f(x-y)]$$

To prove the uniqueness of the quadratic mapping R, assume that there exist another quadratic mapping $S : X \to Y$ which satisfies (5). By induction one can easily show that for all $n \in \mathbb{N}$ and every $x \in X$, $R(r^n x) = r^{2n} R(x)$ and $S(r^n x) = r^{2n} S(x)$. So

$$\begin{aligned} \mu_{R(x)-S(x)}(t) &= \lim_{n\to\infty} \mu_{\frac{R(r^n x)}{r^{2n}} - \frac{S(r^n x)}{r^{2n}}}(t) \\ &\geq \lim_{n\to\infty} min\left\{\mu_{\frac{R(r^n x)}{r^{2n}} - \frac{f(r^n x)}{r^{2n}}}\left(\frac{t}{2}\right), \mu_{\frac{S(r^n x)}{r^{2n}} - \frac{f(r^n x)}{r^{2n}}}\left(\frac{t}{2}\right)\right\} \\ &\geq \lim_{n\to\infty} \Psi_{\varphi(r^n x, 0)}\left(\frac{r^{2n}(r^2-\alpha)t}{2r\alpha^n}\right). \end{aligned} \tag{19}$$

Since $\lim_{n\to\infty} \frac{r^{2n}(r^2-\alpha)t}{2r\alpha^n} = \infty$, we get

$$\lim_{n\to\infty} \Psi_{\varphi(x,0)}\left(\frac{r^{2n}(r^2-\alpha)t}{2r\alpha^n}\right) = 1.$$

Therefore, it follows that for all $t > 0$, $\mu_{R(x)-S(x)}(t) = 1$ and so $R(x) = S(x)$. This complete the proof. $\square$

$\square$

Corollary 3.2. *Let X be real linear space, (Z, Ψ, min) be an RN-space, and (Y, μ, min) a complete RN-space. Let $0 < p < 1$ and $y_0 \in Z$. If $f : X \to Y$ is a mapping that for all $x, y \in X$ and $t > 0$,*

$$\mu_{f(rx+sy)-r^2 f(x)-s^2 f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \Psi_{\|x\|^p x_0}(t), \tag{20}$$

then there is a unique additive mapping $R : X \to Y$ such that

$$R(x) = \lim_{n\to\infty} r^{-2n} f(r^n x) \tag{21}$$

and

$$\mu_{f(x)-R(x)}(t) \geq \Psi_{\|x\|^p x_0}(r - r^{2p-1}). \tag{22}$$

Proof. Let $\alpha = r^{2p}$ and $\varphi : X^2 \to Z$ be defined by $\varphi(x, y) = \|x\|^p x_0$. $\square$

Corollary 3.3. *Let X be real linear space, (Z, Ψ, min) be an RN-space, and (Y, μ, min) a complete RN-space. Let $0 < p < 1$ and $y_0 \in Z$. If $f : X \to Y$ is a mapping that for all $x, y \in X$ and $t > 0$,*

$$\mu_{f(rx+sy)-r^2 f(x)-s^2 f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \Psi_{(\|x\|^p+\|y\|^p)y_0}(t), \tag{23}$$

then there is a unique additive mapping $R : X \to Y$ such that

$$R(x) = \lim_{n\to\infty} r^{-2n} f(r^n x) \tag{24}$$

and

$$\mu_{f(x)-R(x)}(t) \geq \Psi_{\|x\|^p y_0}(r^{1-p} - r^{p-1}). \tag{25}$$

Proof. Let $\alpha = r^{2p}$ and $\varphi : X^2 \to Z$ be defined by $\varphi(x, y) = (\|x\|^p + \|y\|^p)y_0$. $\square$

4. RN-stability of functional equation(1): fixed point method

Theorem 4.1. *Let X be linear space, (Y, μ, T_M) be a complete RN-space and Φ be a mapping from X^2 to D^+ ($\Phi(x, y)$ is denoted by $\Phi_{x,y}$) such that, for some $0 < \alpha < \frac{1}{r^2}$,*

$$\Phi_{rx,ry}(t) \leq \Phi_{x,y}(\alpha t) \tag{26}$$

for all $x, y \in X$ and all $t > 0$. Let $f : X \to Y$ be a function with $f(0) = 0$, such that

$$\mu_{f(rx+sy)-r^2f(x)-s^2f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \Phi_{x,y}(t) \tag{27}$$

for all $x, y \in X$ and all $t > 0$. Then

$$A(x) := \lim_{n\to\infty} r^{2n} f\left(\frac{x}{r^n}\right) \tag{28}$$

exists for each $x \in X$ and defines a unique additive $A : X \to Y$ such that

$$\mu_{f(x)-A(x)}(t) \geq \Phi_{(x,0)}\left(\frac{(1-r^2\alpha)}{\alpha}t\right). \tag{29}$$

for all $x, y \in X$ and $t > 0$.

Proof. Putting $y = 0$ (27), we have

$$\mu_{f(rx)-r^2f(x)}(t) \geq \Phi_{(x,0)}(t). \tag{30}$$

for all $x \in X$ and all $t > 0$. Consider the set $S := \{g : X \to Y, g(0) = 0\}$ and introduce the generalized metric on S:

$$d(f, g) = inf\{u \in \mathbb{R}^+ : \mu_{g(x)-h(x)}(ut) \geq \Phi_{(x,0)}(t), \forall x \in X, \forall t > 0\} \tag{31}$$

where, as usual, $inf\emptyset = +\infty$. It is easy to show that (S, d) is complete. Now we consider the linear mapping $J : S \to S$ such that

$$Jh(x) := r^2 h\left(\frac{x}{r}\right) \tag{32}$$

for all $x \in X$ and we prove that J is a strictly contractive mapping with the Lipschitz constant $r^2\alpha$. Let $g, h \in S$ be given such that $d(g, h) < \epsilon$. Then

$$\mu_{g(x)-h(x)}(\epsilon t) \geq \Phi_{(x,0)}(t) \tag{33}$$

for all $x \in X$ and all $t > 0$. Hence

$$\begin{aligned} \mu_{Jg(x)-Jh(x)}(r^2\alpha\epsilon t) &= \mu_{r^2g(\frac{x}{r})-r^2h(\frac{x}{r})}(r^2\alpha\epsilon t) \\ &= \mu_{g(\frac{x}{r})-h(\frac{x}{r})}(\alpha\epsilon t) \\ &\geq \Phi_{\frac{x}{r}}(\alpha t) \\ &\geq \Phi_{(x,0)}(t) \end{aligned} \tag{34}$$

for all $x \in X$ and all $t > 0$. So $d(g, h) < \epsilon$ implies that $d(Jg, Jh) < r^2\alpha\epsilon$. This means that

$$d(Jg, Jh) \leq r^2\alpha d(g, h) \tag{35}$$

for all $g, h \in S$. It follows from (30) that

$$\mu_{f(x)-r^2f(\frac{x}{r})}(\alpha t) \geq \Phi_{(x,0)}(t) \tag{36}$$

for all $x \in X$ and all $t > 0$. So

$$d(g, Jg) \leq \alpha < 1. \tag{37}$$

By Theorem (**??**), there exists a mapping $A : X \to Y$ satisfying the following:
(1) A is a fixed point of J, the is

$$A\left(\frac{x}{r}\right) = \frac{1}{r^2}A(x) \tag{38}$$

for all $x \in X$. The mapping A is a unique fixed point of J in the set

$$\Omega = \{h \in S : d(g, h) < \infty\}. \tag{39}$$

This implies that A is a unique mapping satisfying (38) such that there exists a $u \in (0, \infty)$ satisfying

$$\mu_{g(x)-A(x)}(ut) \geq \Phi_{x,0}(t) \tag{40}$$

for all $x \in X$ and all $t > 0$;
(2) $d(J^n g, A) \to 0$ as $n \to \infty$. This implies the equality

$$\lim_{n\to\infty} r^{2n} f\left(\frac{x}{r^n}\right) = A(x) \tag{41}$$

for all $x \in X$.
(3) $d(f, A) \leq \frac{d(g.Jg)}{1-r^2\alpha}$ with $g \in \Omega$, which implies the inequality

$$d(g, A) \leq \frac{\alpha}{1-2\alpha} \tag{42}$$

from which it follows

$$\mu_{f(x)-A(x)}(t) \geq \Phi_{x,0}\left(\frac{\alpha t}{1-r^2\alpha}\right) \tag{43}$$

for all $x \in X$ and all $t > 0$. This implies that the inequality (29) holds. Replacing x and y by $r^n x$ and $r^n y$, respectively in (27), we obtain

$$\mu_{r^{2n}\left[f(\frac{rx}{r^n}+\frac{sy}{r^n})-r^2 f(\frac{x}{r^n})-s^2 f(\frac{y}{r^n})-\frac{rs}{2}[f(\frac{(x+y)}{r^n})-f(\frac{(x-y)}{r^n})]\right]}(t) \geq \Phi_{x,y}\left((\frac{1}{\alpha r^2})^n t\right) \tag{44}$$

for all $x, y \in X$, all $t > 0$ and all $n \in \mathbb{N}$. Since $\lim_{n\to\infty} \Phi_x\left(\left(\frac{1}{r^2\alpha}\right)^n t\right) = 1$ for all $x, y \in X$, all $t > 0$, then we deduce that

$$\mu_{A(rx+sy)-r^2A(x)-s^2A(y)-\frac{rs}{2}[A(x+y)-A(x-y)]}(t) = 1, \tag{45}$$

for all $x, y \in X$, all $t > 0$. Thus the mapping $A : X \to Y$ is additive. To prove the uniqueness of the additive mapping A is as Theorem (3.1). □

Corollary 4.2. *Let $\theta \geq$ and let p be a real number with $p > 1$. Let X be a normed vector space with norm $\|.\|$. Let $A : X \to Y$ be a mapping satisfying*

$$\mu_{f(rx+sy)-r^2f(x)-s^2f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \frac{t}{t+\theta(\|x\|^p+\|y\|^p)} \tag{46}$$

for all $x, y \in X$ and $t > 0$. Then

$$A(x) := \lim_{n\to\infty} r^{2n} f\left(\frac{x}{r^n})\right) \tag{47}$$

exists for each $x \in X$ and defines a unique additive mapping $A : X \to Y$ such that

$$\mu_{f(x)-A(x)}(t) \geq \frac{(r^p-r^2)t}{(r^p-r^2)t+\theta\|x\|^p} \tag{48}$$

Proof. The proof follows from Theorem (4.1) by takig

$$\Phi_{x,y}(t) = \frac{t}{t+\theta(\|x\|^p+\|y\|^p)} \tag{49}$$

for all $x, y \in X$ and all $t > 0$. Then we can choose $\alpha = r^{-p}$ and we get the desired result.

Theorem 4.3. *Let X be linear space, (Y, μ, T_M) be a complete RN-space and Φ be a mapping from X^2 to D^+ ($\Phi(x,y)$ is denoted by $\Phi_{x,y}$) such that, for some $0 < \alpha < r^2$,*

$$\Phi_{\frac{x}{r},\frac{y}{r}}(t) \leq \Phi_{x,y}(\alpha t) \tag{50}$$

for all $x, y \in X$ and all $t > 0$. Let $f : X \to Y$ be a function with $f(0) = 0$, such that

$$\mu_{f(rx+sy)-r^2 f(x)-s^2 f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \Phi_{x,y}(t) \tag{51}$$

for all $x, y \in X$ and all $t > 0$. Then

$$A(x) := \lim_{n\to\infty} \frac{f(r^n x)}{r^{2n}} \tag{52}$$

exists for each $x \in X$ and defines a unique additive $A : X \to Y$ such that

$$\mu_{f(x)-A(x)}(t) \geq \Phi_{(x,0)}\left(\frac{(r^2-\alpha)t}{r}\right). \tag{53}$$

for all $x, y \in X$ and $t > 0$.

Corollary 4.4. *Let $\theta \geq 0$ and let p be a real number with $0 < p < 1$. Let X be a normed vector space with norm $\|.\|$. Let $f : X \to Y$ be a mapping satisfying*

$$\mu_{f(rx+sy)-r^2 f(x)-s^2 f(y)-\frac{rs}{2}[f(x+y)-f(x-y)]}(t) \geq \frac{t}{t+\theta(\|x\|^p + \|y\|^p)} \tag{54}$$

for all $x, y \in X$ and $t > 0$. Then

$$A(x) := \lim_{n\to\infty} \frac{f(r^n x)}{r^{2n}} \tag{55}$$

exists for each $x \in X$ and defines a unique additive mapping $A : X \to Y$ such that

$$\mu_{f(x)-A(x)}(t) \geq \frac{(r^2-r^{2p})t}{(r^2-r^{2p})t+\theta\|x\|^p} \tag{56}$$

Proof. The proof follows from Theorem (4.3) by takig

$$\Phi_{x,y}(t) = \frac{t}{t+\theta(\|x\|^p + \|y\|^p)} \tag{57}$$

for all $x, y \in X$ and all $t > 0$. Then we can choose $\alpha = r^{2p}$ and we get the desired result.

References

[1] *T. Aoki, On the stability of the linear transformation in Banach spaces,* J. Math. Soc. Japan 2(1950),64-66

[2] *H. Azadi ,Kh. Shafaat ,M. Shafei and G.Takbiri, Hyers-Ulam-Rassias stability of the apollonius type quadratic mapping in RN-spaces*. The Journal of nonlinear sciences and applications.4(2011).no.1,110-119.

[3] *P.W. Cholewa, Remarks on the stability of functional equations,* Aequationes Math.27(1984),no. 1-2,76-86.1

[4] *P. Ga'vruta, A generalization of the Hyers-Ulam-Rassias stability of approximately additive mappings,* J. Math. Anal. Appl. 184(1994),no.3, 431-436.1

[5] *D. H. Hyers, On the stability of the linear functional equation,* Proc. Nat. Acad. Sci. U. S.A. 27(1941),222-224.1

[6] *D. Mihet and V. Radu, On the stability of the additive cauchy functional equation in random normed spaces,* J. Math. Anal. Appl. 343(2008), no. 1, 567-572.

[7] *A. Najati, Fuzzy stability of a generalized quadratic functional equation,* Commun. Korean Math. Soc. 25(2010),no, 3, pp. 405-412

[8] *C. Park, Fuzzy stability of a functional equation associated with inner product spaces,* Fuzzy Sets and Systems 160(2009), no. 11,1632-1642.

[9] *Th.M. Rassias, On the stability of the linear mapping in Banach spaces,* Proc. Amer. Math. Soc. 72(1978),no. 2,297-300.1,1.

[10] *B. Schewizer and A. Sklar,* Probabilistic Metric Spaces,North-Holland Series in Probability and Applied Mathematics,North-Holland, New York, NY, USA, 1983.1.

[11] *F. Skof, Local properties and approximation of operators,* Rend. Sem. Mat. Fis. Milano 53(1983),113-129.1.

[12] *S. M Ulam, A collecton of mathematical problems,* Interscience Tracts in Pure and Applied Mathematics, no. 8 Interscience Publishers, New York-London, 1960.1.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–135, pp. 56–60

On the difference of two perspectives for accretive matrices

Somayeh Malekinejad[a,*], Hamideh Mohammadzadehkan[b]

[a] *Department of Mathematics*
Payame Noor University, P.O. Box 19395-3697
Tehran, Iran.
[b] *Department of Mathematics*
Payame Noor University, P.O. Box 19395-3697
Tehran, Iran.

Article Info

Keywords:
Accretive matrix
Difference of two perspectives
mean of accretive matrices

2020 MSC:
47A63
47A60

Abstract

In this paper, we state and prove an extension of the inequalities obtained by Moradi et al. for sector matrices and perspective of two means. Among them Let $A, B \in \prod_n^\alpha$ and $f, g \in \mathbf{m}$. Then

$$\begin{aligned}&\mathcal{D}_{f,g}(\Re A|\Re B) + (\cos^2\alpha - 1)\Re(A\sigma_g B)\\ &\leq \Re\mathcal{D}_{f,g}(A|B)\\ &\leq \mathcal{D}_{f,g}(\Re A|\Re B) + (1-\cos^2\alpha)\Re(A\sigma_f B).\end{aligned}$$

1. Introduction

Let $\mathbb{M}_n$ be the class of all $n \times n$ complex matrices. Inequalities among elements of $\mathbb{M}_n$ has been an active research area due to its applications in various fields, not to mention its role in understanding the algebra $\mathbb{M}_n$. However, the order among elements in $\mathbb{M}_n$ is restricted to the so-called Hermitian matrices. A matrix $A \in \mathbb{M}_n$ is said to be Hermitian if $A^* = A$, where A^* is the conjugate transpose of A. Recall that a matrix $A \in \mathbb{M}_n$ is said to be accretive if $\Re A > 0$, where $\Re A$ is the real part of A defined by $\Re A = \frac{A+A*}{2}$ and $\Im A = \frac{A-A*}{2i}$. The class of accretive matrices in $\mathbb{M}_n$ will be denoted by $\prod_n$. It is clear that $\mathcal{P}_n \subset \prod_n$. Since elements of $\prod_n$ are not Hermitian, the predefined order does not apply to $\prod_n$. This is why inequalities among accretive matrices are usually stated in terms of their real parts. We must introduce sectorial matrices to deal with inequalities in $\prod_n$. If $0 \leq \alpha < \frac{\pi}{2}$, and if $A \in \mathbb{M}_n$ is such that

$$\{\langle Ax, x\rangle : x \in \mathbb{C}^n, \|x\| = 1\} \subset \{z \in \mathbb{C} : \Re z > 0, |\Im z| \leq (tan\alpha)\Re z\},$$

then A will be called a sectorial matrix and we simply write $A \in \prod_n$, where $\Im z$ denotes the imaginary part of z. We emphasize here that whenever we use the notation $\prod_n$ in this paper, we implicitly understand that $0 \leq \alpha < \frac{\pi}{2}$. We

*Talker
Email addresses: maleki60313@pnu.ac.ir; maleki60313@gmail.com (Somayeh Malekinejad), h_mohamadzadeh@pnu.ac.ir; mohammadzadeh83@gmail.com (Hamideh Mohammadzadehkan)

also remark that a matrix is accretive if and only if it is sectorial [4]. The study of accretive matrices differs from that of Hermitian matrices because a partial order among members of $\prod_n$ is not as well established as that in $\mathcal{H}_n$. So, in studying inequalities among members of $\prod_n$, we usually refer to the real parts of these elements, noting that the real part of any matrix is in $\mathcal{H}_n$. Our target in this paper is to further study possible inequalities among matrices in $\prod_n$, where we extend some of the well-established inequalities in $\mathcal{P}_n$ or $\mathbb{M}_n^+$ to the class $\prod_n$. For simplicity, we will use the notation:
$\mathbf{m} = \{f : (0,\infty) \to (0,\infty);\ f$ is a matrix monotone function with $f(1) = 1\}$. We will use the following lemma in our main results proof:

Lemma 1.1. *([2]), Let $A, B \in \prod_n^\alpha$. Then $A\sigma B \in \prod_n^\alpha$ and*

$$\Re A\sigma\Re B \le \Re(A\sigma B) \le \sec^2(\alpha)\,(\Re A\sigma\Re B).$$

Malekinejad et al. [10] were able to prove the general condition of the above inequalities. For example, they gave the following inequality.

Lemma 1.2. *([10]), If σ_1 and σ_2 are two means with $\sigma_1 \le \sigma_2$ and $A, B \in \prod_n^\alpha$, then*

$$\Re(A\sigma_1 B) \le \sec^2\alpha\,\Re\,(A\sigma_2 B)\,.$$

In particular

$$\cos^2\alpha\,\Re(A!_t B) \le \Re(A\sharp_t B) \le \sec^2\alpha\,\Re(A\nabla_t B). \tag{1}$$

We refer the reader to [2–6, 8–11, 14] for an almost comprehensive overview of the progress that has been made in studying inequalities in $\prod_n$. If $A,\ B$ are two strictly positive operators, and $0 \le t \le 1$ is a real number, then the relative operator entropy $S(A|B)$ is defined by

$$S(A|B) = A^{\frac{1}{2}}\log(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}$$

Moreover, the Tsallis relative operator entropy is defined by

$$\mathcal{T}_t(A|B) := \frac{A\sharp_t B - A}{t}.$$

It is known that [13, Theorem 5.18]

$$\lim_{t\to 0}\mathcal{T}_t(A|B) = S(A|B).$$

For more details, we refer the reader to[15].

Definition 1.3. Let σ_f be a matrix mean related to matrix monotone function $f \in \mathbf{m}$. Then $A\sigma_f B := A^{\frac{1}{2}}f(A^{-\frac{1}{2}}BA^{-\frac{1}{2}})A^{\frac{1}{2}}$ is often sometimes called a perspective [1]. We may define a difference between two perspectives as $\mathcal{D}_{f,g}(A|B) = A\sigma_f B - A\sigma_g B$, for $f, g \in \mathbf{m}$.

For two operator means σ_f and σ_g with representing functions f and g, we write $\sigma \le \tau$ if $A\sigma B \le A\tau B$ for every two positive operators A and B or equivalently if $f(t) \le g(t)$ for all positive $t \in \mathbb{R}$. Moradi et al. [12] proved the following relation between $\Re\,\mathcal{D}_{f,g}(A|B)$ and $\mathcal{D}_{f,g}(\Re A|\Re B)$.

Lemma 1.4. *[12] Let $A, B \in \prod_n^\alpha$ and $f, g \in \mathbf{m}$. Then*

$$\begin{aligned}&\mathcal{D}_{f,g}(\Re A|\Re B) + (1-\sec^2\alpha)(\Re A\sigma_g\Re B)\\ &\le \Re\,\mathcal{D}_{f,g}(A|B)\\ &\le \mathcal{D}_{f,g}(\Re A|\Re B) + (\sec^2\alpha - 1)(\Re A\sigma_g\Re B).\end{aligned} \tag{2}$$

Finally, they have the following double inequality, which bounds $\mathcal{D}_{f,g}(A|B)$ between certain differences between the harmonic mean $!_t$ and the arithmetic mean ∇_t.

Lemma 1.5. *[12] Let $A, B \in \prod_n^\alpha$ and $f, g \in$ **m** be such that $f'(1) = g'(1) = t$. Then*

$$\cos^2\alpha\, \Re(A!_tB) - \sec^2\alpha\, \Re(A\nabla_tB) \le \Re\left(\mathcal{D}_{f,g}(A|B)\right) \le \sec^2\alpha\, \Re(A\nabla_tB) - \cos^2\alpha\, \Re(A!_tB). \tag{3}$$

If we take $f(x) := \frac{x^t-1}{t} + 1,\ (0 < t \le 1)$ and $g(x) := 1$ in (2), then we have

$$\mathcal{T}_t(\Re A|\Re B) + (1 - \sec^2\alpha)\Re A \le \Re\left(\mathcal{T}_t(A|B)\right) \le \sec^2\alpha \mathcal{T}_t(\Re A|\Re B) + (\sec^2\alpha - 1)\Re A.$$

Since it is known the relation $\mathcal{T}_t(\Re A|\Re B) \le \Re(\mathcal{T}_t(A|B))$ for accretive matrices $A,\ B$ and $0 < t < 1$ in [14], the lower bound of $\Re(\mathcal{T}_t(A|B))$ in the inequalities above does not give a refined bound. However, we obtain the upper bound of $\Re(\mathcal{T}_t(A|B))$.
Ghazanfari and Malekinejad [7, Theorem 2.1] obtained the upper bound of $\Re(\mathcal{T}_t(A|B))$ as follows:

Lemma 1.6. *Let $A, B \in \prod_n^\alpha$ and $0 < t \le 1$. If there exists a $m > 1$ such that $m\Re A \le \Re B$, then there is a $\beta > 1 + \frac{\tan^2\alpha}{t}$ such that*

$$\frac{(\beta - 1)t}{(\beta - 1)t - \tan^2\alpha} \le m,$$

and

$$\Re(\mathcal{T}_t(A \mid B)) \le \beta \mathcal{T}_t(\Re A \mid \Re B),$$

consequently

$$\Re(\mathcal{S}(A \mid B)) \le \beta \mathcal{S}(\Re A \mid \Re B).$$

2. Main Result

In this section, we present our results. We study $\mathcal{D}_{f,g}(A|B)$ for accretive matrices A, B; as a new track in this research field. The following theorem is an extension of [12, Theorem 3.2].

Theorem 2.1. *Let $A, B \in \prod_n^\alpha$ and $f, g \in$ **m**. Then*

$$\begin{aligned} &\mathcal{D}_{f,g}(\Re A|\Re B) + (\cos^2\alpha - 1)\Re(A\sigma_g B) \\ &\le \Re\mathcal{D}_{f,g}(A|B) \\ &\le \mathcal{D}_{f,g}(\Re A|\Re B) + (1 - \cos^2\alpha)\Re(A\sigma_f B). \end{aligned} \tag{4}$$

Proof. By Lemma 1.1, we have

$$\begin{aligned} \mathcal{D}_{f,g}(\Re A|\Re B) &= \Re A\sigma_f \Re B - \Re A\sigma_g \Re B \\ &\le \Re(A\sigma_f B) - \cos^2\alpha\, \Re(A\sigma_g B) \\ &= \Re(A\sigma_f B) - \Re(A\sigma_g B) + \Re(A\sigma_g B) - \cos^2\alpha\, \Re(A\sigma_g B) \\ &= \Re(A\sigma_f B) - \Re(A\sigma_g B) + (1 - \cos^2\alpha)\Re(A\sigma_g B) \\ &= \Re\mathcal{D}_{f,g}(A|B) + (1 - \cos^2\alpha)\Re(A\sigma_g B), \end{aligned}$$

and

$$\begin{aligned} \mathcal{D}_{f,g}(\Re A|\Re B) &= \Re A\sigma_f \Re B - \Re A\sigma_g B \\ &\ge \cos^2\alpha\, \Re(A\sigma_f B) - \Re(A\sigma_g B) \\ &= \Re(A\sigma_f B) - \Re(A\sigma_f B) + \cos^2\alpha\, \Re(A\sigma_f B) - \Re(A\sigma_g B) \\ &= \Re\mathcal{D}_{f,g}(A|B) + (\cos^2\alpha - 1)\Re(A\sigma_f B). \end{aligned}$$

□

Remark 2.2. By Lemma 1.1 inequality (2) equivalent to

$$\begin{aligned}&\mathcal{D}_{f,g}(\Re A|\Re B) + (1-\sec^2\alpha)\Re(A\sigma_g B)\\ &\leq \Re\mathcal{D}_{f,g}(A|B)\\ &\leq \mathcal{D}_{f,g}(\Re A|\Re B) + (\sec^2\alpha - 1)\Re(A\sigma_f B).\end{aligned}$$

Since $\sec^2\alpha - 1 \geq 1-\cos^2\alpha$, therefore Inequalities (4) are stronger than Inequalities (2).

The following theorem is an extension of [12, Theorem 3.3].

Theorem 2.3. *Let $A, B \in \prod_n^\alpha$ and $f, g \in \boldsymbol{m}$ such that $\sigma_f \leq \sigma_g$. Then*

$$\Re\mathcal{D}_{f,g}(A|B) \leq (1-\cos^2\alpha)\Re(A\sigma_f B) \leq (\sec^2\alpha - 1)\Re(A\sigma_g B). \tag{5}$$

Proof. By Lemmas 1.1 and 1.2, we have

$$\begin{aligned}\Re\mathcal{D}_{f,g}(A|B) &= \Re(A\sigma_f B) - \Re(A\sigma_g B)\\ &\leq \Re(A\sigma_f B) - \cos^2\alpha\,\Re(A\sigma_f B)\\ &= (1-\cos^2\alpha)\Re(A\sigma_f B)\\ &\leq \sec^2\alpha(1-\cos^2\alpha)\Re(A\sigma_g B)\\ &= (\sec^2\alpha - 1)\Re(A\sigma_g B).\end{aligned}$$

□

Remark 2.4. If we take $f(x) := (1-t)+tx$ and $g(x) := x^t$ in Lemma 1.5, then we obtain

$$(1-\sec^2\alpha)\,\Re(A\nabla_t B) + \cos^2\alpha\,\Re(A!_t B) \leq \Re(A\sharp_t B) \leq (1+\sec^2\alpha)\Re(A\nabla_t B) - \cos^2\alpha\,\Re(A!_t B), \tag{6}$$

and if we take $f(x) := x^t$ and $g(x) := \left((1-t)+tx^{-1}\right)^{-1}$ in Lemma 1.5, then we obtain

$$(1+\cos^2\alpha)\,\Re(A!_t B) - \sec^2\alpha\,\Re(A\nabla_t B) \leq \Re(A\sharp_t B) \leq \sec^2\alpha\,\Re(A\nabla_t B) + (1-\cos^2\alpha)\,\Re(A!_t B). \tag{7}$$

However, we find from the inequality $\cos^2\alpha\,\Re(A!_t B) \leq \Re(A\nabla_t B)$ that both inequalities above do not improve the known inequalities (1).
On the other hand if we take $f(x) := \left((1-t)+tx^{-1}\right)^{-1}$ and $g(x) := x^t$ in Theorem 2.3, then we obtain

$$\cos^2\alpha\,\Re(A!_t B) \leq \Re(A\sharp_t B)$$

and if we take $f(x) := x^t$ and $g(x) := (1-t)+tx$ in Theorem 2.3, then we obtain

$$\Re(A\sharp_t B) \leq \sec^2\alpha\,\Re(A\nabla_t B).$$

Therefore inequalities (5) are stronger than inequalities (3).

References

[1] Kubo, F. and Ando, T. (1980), "Means of positive linear operators," Math. Ann., 246 , 205–224.
[2] Bedrani, Y. Kittaneh, F. and Sababeh, M. (2021), "From positive to accretive matrices," Positivity.,https:// doi.org /10.1007/ s 11117-021-00831-8.
[3] Y. Bedrani, F. Kittaneh, and M. Sababheh, (2021), "Numerical radii of accretive matrices," Linear Multilinear Algebra., 69(5) , 957–970.
[4] Bedrani, Y. Kittaneh, F. and Sababheh, M. (2022), "matrices and matrix convex functions," Results Math., 77, 52 . https://doi.org/10.1007/s00025-021-01590-4.
[5] Lin, M. (2014), "Singular value inequalities for matrices with numberical ranges in a sector," J. Oper. Matrices., 8 , 1143-1148.
[6] Ghazanfari, A. and Malekinejad, S. (2021), " Heron means and Pólya inequality for sector matrices," J. Bull. Math. Soc. Sci. Math. Roumanie. Tome., 64 , 329-339.

[7] Ghazanfari, A. and Malekinejad, S. (2021), “New inequalities for entropy and Tsallis entropy of two accretive operators,” in 52nd Annual Iranian Mathematics Conference, Kerman, IR.Iran.
[8] Lin, M.(2016), “Some inequalities for sector matrices,” J. Oper. Matrices., 10, 915-921.
[9] Liu, J. T. and Wang, Q. W. (2017) “ More inequalities for sector matrices,” J. Bull. Iran. Math. Soc., https://doi.org/10.1007/s41980 -018- 0069-y.
[10] Malekinejad, S. Khosravi, M. and Sheikhhosseini, A. (2022), “Mean inequalities for sector matrices involving positive linear maps, ” Positivity., 26 ,1-17.
[11] Mao, Y. (2020), “Inequalities for the Hienz Mean of Sector Matrices,” Bull. Iran. Math. Soc., https://doi.org /10.1007/s 41980-020-00357-x.
[12] Moradi, H. R. Furuichi, S. and Sababheh, M. “Further properties of accretive matrices,” https://arxiv.org/abs/2210 .08678.
[13] Pečarić, J. E. Furuta, T. Mićić Hot, J. and Seo, Y. (2005), “ Mond-pečarić Method in operator inequalities,” Element, Zagreb.
[14] Raissouli, M. Moslehian, M. S. and Furuichi, S. (2017), “Relative entropy and Tsallis entropy of two accretive operators,” C. R. Math. Acad. Sci. Paris, Ser, I, 355, 687-693.
[15] Soleimani, S. and Ghazanfari, A. G. (2019),“New refinements of some inequalities involving relative operator entropy,” Complex Analysis and Operator Theory 13:3337-3345.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–138, pp. 61–64

Notes about the direct sum of subspace-diskcyclic operators

Mansooreh Moosapoor*

[a] *Associate Professor, Department of Mathematics Education, Farhangian University, Tehran, Iran.*

Article Info

Keywords:
Subspace-diskcyclic operators
Diskcyclic operators
Direct sum

2020 MSC:
47A16
47B37

Abstract

In this paper, we investigate the properties of the direct sum of subspace-diskcyclic operators. We prove that subspace-diskcyclicity of two operators implies subspace-diskcyclicity of any of them. Also, we show that the direct sum of two subspace-diskcyclic operators are subspace-diskcyclic. Especially, the direct sum of a subspace-diskcyclic operator with itself is subspace-diskcyclic.

1. Introduction

Let X be a complex Banach space and let $B(X)$ be the set of linear continuous operators from X to X. Suppose $T \in B(X)$. By $orb(T, x)$ we mean:

$$orb(T, x) = \{x, Tx, T^2x, ...\}.$$

An operator T is hypercyclic if there exists some vector x such that $orb(T, x)$ be dense in X [3]. If M is a closed subspace of X, T is called M-hypercyclic if there is $x \in X$ such that $orb(T, x) \cap M$ is dense in M [4].
Some other types of operators are also defined. For example diskcyclic operators are defined in [6] as follows.

Definition 1.1. An operator $T \in B(X)$ is called diskcyclic if there is a vector $x \in X$ such that the disk orbit

$$\mathbb{D}orb(T, x) = \{\lambda T^n(x); \lambda \in \mathbb{C}, |\lambda| \leq 1, n \in \mathbb{N}_0\},$$

is dense in X and X is called a diskcyclic vector for T.

Recall that $\mathbb{D} = \{x \in \mathbb{C} : |x| \leq 1\}$ is the closed unit disk.
By the definition, that the set of hypercyclic operators is a subset of the set of diskcyclic operators. Authors in [2] charactrized diskcyclic forward weighted shifts. Recall that a forward weighted shift with respect to the canonical basis $\{e_n : n \in \mathbb{Z}\}$ if $T(e_n) = w_n e_{n+1}$, where the weight sequence $\{w_n : n \in \mathbb{Z}\}$ is a bounded subset of $\mathbb{C} \setminus \{0\}$. Also, they stated a diskcyclicity criterion as follows.

*Talker
Email address: m.mosapour@cfu.ac.ir; mosapor110@gmail.com (Mansooreh Moosapoor)

Theorem 1.2. *([2]) Let $T \in B(X)$. If there exists an increasing sequence of integers $\{n_k\} \in \mathbb{N}$ and two dense sets $D_1, D_2 \in X$ such that:*

(a) For each $y \in D_2$, there exists a sequence $\{x_k\}$ in X such that $x_k \to 0$, and $T^{n_k}(x_k) \to y$.

(b) $||T^{n_k}(x)||||x_k|| \to 0$ for all $x \in D_1$.

The concept of subspace-diskcyclicity is defined in [1] as follows.

Definition 1.3. Let $T \in B(X)$ and let M be a closed subspace of X. Then T is called an M-diskcyclic operator if there exists a vector $x \in X$ such that $\mathbb{D}orb(T, x) \cap M$ is dense in M. Such a vector x is called an M-diskcyclic vector for T.

Authors stated in [1] that there are subspace-diskcyclic operators that are not subspace-hypercyclic. Also, they stated some sufficient conditions for subspace-diskcyclicity. One can see more about subspace-diskcyclic operators in [5]. It is proved in [5] that any diskcyclic operator is subspace-diskcyclic.

Theorem 1.4. *If $T \in B(X)$ is a diskcyclic operator, then there is a non-trivial closed subspace M of X such that T is M-diskcyclic.*

In this paper, we investigate properties of the subspace-diskcyclicity operators and their direct sum. We show that subspace-diskcyclicity of any power of an operator implies subspace-diskcyclicity of the operator. Also, we prove that subspace-diskcyclicity of two operators implies subspace-diskcyclicity of any of them. Moreover, the direct sum of two subspace-diskcyclic operators leads to subspace-diskcyclicity of any of them.

2. Main results

We begin with proving a primarily property of subspace-diskcyclic operators.

Theorem 2.1. *Let $T \in B(X)$ and M be a closed subspace of X. If T^n is M-diskcyclic for some $n \in \mathbb{N}$, then T is M-diskcyclic.*

Proof. Suppose T^n is M-diskcyclic for some $n \in \mathbb{N}$. Hence, there is $x \in X$ such that $\mathbb{D}orb(T^n, x) \cap M$ is dense in M. But

$$\mathbb{D}orb(T^n, x) \cap M \subseteq \mathbb{D}orb(T, x) \cap M \subseteq M.$$

Hence, $orb(T, x) \cap M$ is dense in M. So, T is M-diskcyclic. □

In the next theorem, we prove that subspace-diskcyclicity of an operator implies some multiples of it.

Theorem 2.2. *Let $T \in B(X)$ and M be a closed subspace of X. If T is M-diskcyclic, then μT is M-diskcyclic for any $\mu \in \mathbb{C}$ with $|\mu| = 1$.*

Proof. Let $\mu \in \mathbb{C}$ with $|\mu| = 1$. T is M-diskcyclic. Hence, there is $x \in X$ such that $\mathbb{D}orb(T, x) \cap M$ is dense in M. Let U be a nonempty open set. So, there are $n \in \mathbb{N}$ and $\lambda \in \mathbb{D}$ such that $\lambda T^n x \in U$. But

$$\lambda T^n x = (\frac{\lambda}{\mu^n})\mu^n T^n x$$

and $|\frac{\lambda}{\mu^n}| = \frac{|\lambda|}{|\mu^n|} \leq 1$. So, for any nonempty open set U there are $\kappa \in \mathbb{C}$ and $n \in \mathbb{N}$ such that $|\kappa| \leq 1$ and $\kappa(\mu^n T^n)x \in U$. Hence, μT is M-diskcyclic. □

Now, we begin to investigate the properties of the direct sum of the subspace-diskcyclic operators. First, we prove that subspace-diskcyclicity of the direct sum of two operators, implies subspace-diskcyclicity of any of them. In the following Y denotes a Banach space.

Theorem 2.3. *Let $T \in B(X)$ and $S \in B(Y)$. Let M be a closed subspace of X and N be a closed subspace of Y. If $T \oplus S$ is $M \oplus N$-diskcyclic, then T is M-diskcyclic and S is N-diskcyclic.*

Proof. $T \oplus S$ is $M \oplus N$-diskcyclic. So, there is $x \in X$ and $y \in Y$ such that $\mathbb{D}orb(T \oplus S, x \oplus y) \cap (M \oplus N)$ is dense in $M \oplus N$. Let $m \in M$ and $n \in N$. Hence, $m \oplus n \in M \oplus N$. By $M \oplus N$-diskcyclicity of $T \oplus S$, there are $\lambda \in \mathbb{C}$ with $|\lambda| \leq 1$ and $k \in \mathbb{N}$ such that

$$\|\lambda(T \oplus S)^k(x \oplus y) - (m \oplus n)\|_{M \oplus N} \leq \varepsilon.$$

On the oter hand

$$\|\lambda T^k x - m\|_M \leq \|\lambda(T \oplus S)^k(x \oplus y) - (m \oplus n)\|_{M \oplus N},$$
$$\|\lambda S^k y - n\|_N \leq \|\lambda(T \oplus S)^k(x \oplus y) - (m \oplus n)\|_{M \oplus N}.$$

So,

$$\|\lambda T^k x - m\|_M \leq \varepsilon \quad \text{and} \quad \|\lambda S^k y - n\|_N \leq \varepsilon.$$

Hence, we can conclude that $\mathbb{D}orb(T, x) \cap M$ is dense in M and $\mathbb{D}orb(S, y) \cap N$ is dense in N. □

By Theorem 2.1, Theorem 2.2 and Theorem 2.3 we can state the following corollaries.

Corollary 2.4. *Let $T \in B(X)$ and $S \in B(Y)$. Let M be a closed subspace of X and N be a closed subspace of Y. If $T^n \oplus S^m$ is $M \oplus N$-diskcyclic for some $m, n \in \mathbb{N}$, then μT is M-diskcyclic and λS is N-diskcyclic for any $\mu, \lambda \in \mathbb{C}$ with $|\mu| = |\lambda| = 1$.*

Proof. By Theorem 2.3, T^n is M-diskcyclic and S^m is N-diskcyclic. Now, Theorem 2.1 asserts that T is M-diskcyclic and S is N-diskcyclic.
Hence, for any $\mu, \lambda \in \mathbb{C}$ with $|\mu| = |\lambda| = 1$, μT is M-diskcyclic and λS is N-diskcyclic by Theorem 2.2. □

Corollary 2.5. *Let $T \in B(X)$. Let M be a closed subspace of X. If $T \oplus T$ is $M \oplus M$-diskcyclic, then T is M-diskcyclic.*

Proof. It is sufficient to consider $S := T$ and $N := M$ in Theorem 2.3. □

We also can generalize Theorem 2.3 to a finite number of operators as follows. the proof is similar to the proof of Theorem 2.3

Lemma 2.6. *Suppose X_i, $1 \leq i \leq n$, are Banach spaces. Let M_i, $1 \leq i \leq n$, be a closed subspace of X_i. Let $T_i \in B(X_i)$. If $T_1 \oplus T_2 \oplus ... \oplus T_n$ is $M_1 \oplus M_2 \oplus ... \oplus M_n$-diskcyclic then for any $1 \leq i \leq n$, T_i is an M_i-diskcyclic operator.*

Now, we prove that subspace-diskcyclicity of two operators implies subspace-diskcyclicity of their direct sum.

Theorem 2.7. *Let $T \in B(X)$ and $S \in B(Y)$. Let M be a closed subspace of X and N be a closed subspace of Y. If T is M-diskcyclic and S is N-diskcyclic, then $T \oplus S$ is $M \oplus \{0\}$-diskcyclic and $\{0\} \oplus N$-diskcyclic.*

Proof. By hypothesis T is M-diskcyclic. So, there is $x \in X$ such that $\mathbb{D}orb(T, x) \cap M$ is dense in M.
Suppose $m \in M$. Hence, there are $\lambda \in \mathbb{D}$ and $k \in \mathbb{N}$ such that

$$\|\lambda T^k x - m\|_M \leq \varepsilon.$$

Hence,

$$\|\lambda(T \oplus S)^k(x \oplus 0) - (m \oplus 0)\|_{M \oplus \{0\}} \leq \varepsilon.$$

That means $T \oplus S$ is $M \oplus \{0\}$-diskcyclic.
Similarly, S is N-diskcyclic. So, there is $y \in Y$ such that $\mathbb{D}orb(T, y) \cap N$ is dense in N. Suppose $n \in N$. Hence, there are $\mu \in \mathbb{D}$ and $p \in \mathbb{N}$ such that

$$\|\mu S^p y - n\|_N \leq \varepsilon.$$

Hence,

$$\|\mu(T \oplus S)^p(0 \oplus y) - (0 \oplus n)\|_{\{0\}\oplus N} \leq \varepsilon.$$

That means $T \oplus S$ is $\{0\} \oplus N$-diskcyclic. □

We show in Theorem 2.7 that if T and S are subspace-diskcyclic, then $T \oplus S$ is subspace-diskcyclic with respect to at least two subspaces. Now, this question arises that can we conclude $M \oplus N$-diskcyclicity of $T \oplus S$, when T is M-diskcyclic and S is N-diskcyclic?

References

[1] N. Bamerni, A. Kilicman, *On subspace-diskcyclicity*, Arab. J. Math. Sci., (23)(2), (2017) 133-140.

[2] N. Bamerni, A. Kilicman, M. S. Md Noorani, *A review of some works in the theory of diskcyclic operators*, Bull. Malays. Math. Sci. Soc., (39)(2), (2016) 723-739.

[3] K.G. Grosse-Erdmann and A. Peris Manguillot, *Linear Chaos*, Universitext, Springer, London, 2011.

[4] B. F. Madore, R. A. Martinez-Avendano, *Subspace hypercyclicity*, J. Math. Anal. Appl., (373)(2) (2011), 502-511.

[5] M. Moosapoor, *On diskcyclic and subspace-diskcyclic operators*, The extended abstracts of the 9^{th} seminar on nonlinear analysis and its applications, Qazvin, Iran, (2018) 23-23.

[6] J. Zeana, *Cyclic phenomena of operators on Hilbert space (Ph.D. thesis)*, University of Baghdad, 2002.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–142, pp. 65–68

Multi Subspace-supercyclic operators

Mansooreh Moosapoor[a,*]

[a] *Associate Professor, Department of Mathematics Education, Farhangian University, Tehran, Iran.*

Article Info

Keywords:
Subspace-supercyclic operators
Multi-supercyclic operators
Supercyclic operators

2020 MSC:
47A16
47B37

Abstract

In this paper, we introduce and investigate multi subspace-supercyclic operators. We prove that any subspace-hypercyclic operator and multi-supercyclic operator are multi subspace-supercyclic. Also, we prove that an operator is multi subsace-supercyclic if and only if any powers of it is multi subspace-supercyclic.

1. Introduction and Preliminaries

Let H be an infinite-dimensional and separable Hilbert space. We denote the set of all linear continuous operators on H by $B(H)$. An operator $T \in B(H)$ is called hypercyclic, if there is $h \in H$ such that $orb(T, h)$ is dense in H, where

$$orb(T, h) = \{h, Th, T^2h, ...\}.$$

If for some $h \in H$,

$$\mathbb{C}orb(T, h) = \{\lambda T^n h, \lambda \in \mathbb{C} \text{ and } n \in \mathbb{N}_0\}$$

is dense in H, then T is called a supercyclic operator [2]. Hypercyclicity, supercyclicity and related topics are considered for decades. One can see more about them in [2].
We say that T is multi-supercyclic, if there is $\{x_1, x_2, ..., x_n\} \subseteq H$ such that $\bigcup_{i=1}^n \mathbb{C}orb(T, x_i)$ is dense in H [5].
Subspace-hypercyclic operators were defined by Madore and Martinez-Avendano in [3]. We say an operator $T \in B(H)$ is subspace-hypercyclic with respect to a closed subspace M of H if there is $x \in H$ such that $orb(T, x) \cap M$ is dense in M. Bamerni, Kadets and Kilicman in [1] answered to a question that is mentioned in [3] by stating the following theorem.

Theorem 1.1. *([1]) Let A be a dense subset of a Hilbert space H. Then there exists a non-trivial closed subspace M of H such that $A \cap M$ is dense in M.*

*Talker
Email address: m.mosapour@cfu.ac.ir; mosapor110@gmail.com (Mansooreh Moosapoor)

By Theorem 1.1, they stated in [1] that any hypercyclic operator is subspace-hypercyclic.
Subspace-supercyclic operators are introduced in [7] as follows.

Definition 1.2. An operator $T \in B(H)$ is called subspace-supercyclic with respect to a closed subspace M of H if there is $x \in H$ such that $\mathbb{C}orb(T,x) \cap M$ is dense in M.

Authors in [7] constructed some examples of this type of operators. They also proved several theorems about them. Zhang and Zhou in [6] stated a subspace-supercyclicity criterion and some criteria equivalent to it.
In [4] one can see some examples of operators that are subspace-supercyclic but they are not subspace-hypercyclic. Also, it is proved that subspace-supercyclic operators exists on finite-dimensional spaces. Moreover, we have the following theorem in [4]

Theorem 1.3. *Let $T \in B(H)$ be an invertible operator. If T is subspace-supercyclic, then T^n and T^{-n} are subspace-supercyclic for any $n \in \mathbb{N}$.*

Now it is natural to define multi subspace-supercyclic operators as follows.

Definition 1.4. Let $T \in B(H)$ and let M be a closed and non-zero subspace of H. We say that T is multi subspace-supercyclic with respect to M or multi M-supercyclic if there exists $\{x_1, x_2, ..., x_n\} \subseteq H$ such that

$$\overline{\bigcup_{i=1}^{n}(\mathbb{C}orb(T,x_i) \cap M)} = M.$$

We say $F = \{x_1, x_2, ..., x_n\} \subseteq H$ is a minimal set for multi M-supercyclicity for $T \in B(H)$ if for any $E \subset F$ we have

$$\overline{\bigcup_{x_i \in E}(\mathbb{C}orb(T,x_i) \cap M)} \neq M.$$

It is clear by the definition that subspace-supercyclic operators are multi subspace-supercyclic.
In this paper, we investigate multi subspace-supercyclic operators. We prove that any multi-supercyclic operator is multi subspace-supercyclic. Also, we prove that multi subsace-supercyclicity of an operator implies multi subspace-supercyclicity of any powers of it and vice versa.

2. Properties of Multi Subspace-supercyclic Operators

We start this section by proving that multi-supercyclic operators are multi subspace-supercyclic.

Theorem 2.1. *Suppose $T \in B(H)$ is a multi-supercyclic operator. Then there is a closed and non-trivial subspace M of H such that T is multi subspace-supercyclic.*

Proof. Let $x_1, x_2, ..., x_n \in H$ such that $\overline{\bigcup_{1 \leq i \leq n} \mathbb{C}orb(T,x_i)} = H$. By Theorem 1.1, there exists a closed and non-trivial subspace M of H such that $(\bigcup_{1 \leq i \leq n} \mathbb{C}orb(T,x_i)) \cap M$ is dense in M. On the other hand,

$$\overline{\bigcup_{1 \leq i \leq n}(\mathbb{C}orb(T,x_i) \cap M)} = \overline{(\bigcup_{1 \leq i \leq n} \mathbb{C}orb(T,x_i)) \cap M}.$$

So, T is multi subspace-supercyclic with respect to M. □

In the next theorem, we prove that multi subspace-supercyclicity of T^n can be concluded from multi subspace-supercyclicity of T for any $n \in \mathbb{N}$.

Theorem 2.2. *Let $T \in B(H)$. If T is multi subspace-supercyclic with respect to M, then for every $n \in \mathbb{N}$, T^n is a multi subspace-supercyclic operator with respect to M .*

Proof. For $n = 1$ the proof is clear. Suppose $n \geq 2$. By multi M-supercyclicity of T, there are $x_1, x_2, ..., x_m$ in H such that $\bigcup_{1\leq i\leq m}(\mathbb{C}orb(T, x_i) \cap M)$ is dense in M. Let $y_{i,j} = T^j x_i$, where $1 \leq i \leq m$ and $1 \leq j \leq n-1$. Consider that

$$\bigcup_{1\leq i\leq m} (\mathbb{C}orb(T, x_i) \cap M) = \bigcup_{\substack{1\leq i\leq m \\ 0\leq j\leq n-1}} (\mathbb{C}orb(T^n, y_{i,j}) \cap M). \tag{1}$$

The left side of (1) is dense in M. So, the right is dense in M too. Therefore, T^n is multi subspace-supercyclic with respect to M. □

Now by Theorem 2.1 and Theorem 2.2, we can conclude the following corollary.

Corollary 2.3. *Suppose $T \in B(H)$ is a multi-supercyclic operator. Then there is a closed and non-trivial subspace M of H such that T^n is multi subspace-supercyclic for any $n \in \mathbb{N}$.*

Also, we can state the following theorem.

Theorem 2.4. *Let $T \in B(H)$. Suppose that there exists $n \in \mathbb{N}$ such that T^n is multi subspace-supercyclic with respect to M. Then T is multi subspace-supercyclic with respect to M.*

Proof. Let n be a positive integer greater than or equal to 2 such that T^n is multi subspace-supercyclic with respect to M. So there are $x_1, x_2, ..., x_m \in H$ such that

$$\overline{\bigcup_{1\leq i\leq m} (\mathbb{C}orb(T^n, x_i) \cap M)} = M. \tag{2}$$

But

$$\mathbb{C}orb(T^n, x_i) \cap M \subseteq \mathbb{C}orb(T, x_i) \cap M. \tag{3}$$

Now by (2) and (3), we conclude that $\bigcup_{1\leq i\leq m}(\mathbb{C}orb(T, x_i) \cap M)$ is dense in M. Therefore T is multi M-supercyclic. □

By Theorem 2.2 and Theorem 2.4 we can conclude the following corollary.

Corollary 2.5. *Let $T \in B(H)$. Then T is multi subspace-supercyclic with respect to M if and only if T^n is multi subspace-supercyclic with respect to M for any $n \in \mathbb{N}$.*

In the next theorem, we show that if T is subspace-supercyclic, then any power of it is multi subspace-supercyclic.

Theorem 2.6. *Let $T \in B(H)$ be an M-supercyclic operator. Then T^n is multi M-supercyclic for any $n \in \mathbb{N}$.*

Proof. It is clear when $n = 1$. Now let $n \geq 2$. Let y be an M-supercyclic vector for T. So, $\overline{\mathbb{C}orb(T, y) \cap M} = M$. Let $x_1 := x, x_2 := Tx, ..., x_n := T^{n-1}x$. Hence,

$$\begin{aligned}
&\bigcup_{j=1}^{n} (\mathbb{C}orb(T^n, T^{j-1}y) \cap M) \\
&= (\mathbb{C}orb(T^n, y) \cup \mathbb{C}orb(T^n, Ty) \cup ... \cup \mathbb{C}orb(T^n, T^{n-1}y)) \cap M \\
&= \mathbb{C}\{y, Ty, ..., T^{n-1}y, T^n y, T^{n+1}y, ...\} \cap M \\
&= \mathbb{C}orb(T, y) \cap M.
\end{aligned}$$

Therefore T^n is multi M-supercyclic. □

By using Theorem 2.6, we can make the following example.

Example 2.7. Let T be a supercyclic operator on a Hilbert space H. If we consider $T \oplus I : H \oplus H \to H \oplus H$, then $T \oplus I$ is subspace-supercyclic with respect to $M := H \oplus \{0\}$. Now by Theorem 2.6, we can deduce that $(T \oplus I)^n = T^n \oplus I$ is multi M-supercyclic for any $n \in \mathbb{N}$.

Corollary 2.8. *Let $T \in B(H)$, where H is an infinite-dimensional Hilbert space. If $T^n = I$ for some $n \in \mathbb{N}$, then T can not be subspace-supercyclic.*

Proof. Consider that $T^n = I$ for some $n \in \mathbb{N}$. Without loss of generality, we can assume that $n \geq 2$. Suppose on contrary that T is subspace-supercyclic. By Theorem 2.6, T^n must be multi subspace-supercyclic. But this is impossible since the identity operator can not be multi subspace-supercyclic. $\square$

References

[1] N. Bamerni, V. Kadets and A. Kilicman, *Hypercyclic operators are subspace-hypercyclic*, J. Math. Anal. Appl. **435** (2016), no. 2, 1812–1815.
[2] K.-G. Grosse-Erdmann and A.Peris Manguillot, *Linear chaos*, Springer, 2011.
[3] B. F. Madore and R. A. Martiınez-Avendano, *Subspace hypercyclicity*, J. Math. Anal. Appl. **373** (2011), 502–511.
[4] M. Moosapoor, *On subspace-supercyclic operators*,Aust. J. Math. Anal. Appl. **17** (2020), Art. 4, 8 pp.
[5] A. Peris, *Multi-hypercyclic operators are hypercyclic*, Math. Z. **236** (2001), 779–786.
[6] L. Zhang, Z. H. Zhou, *Notes about subspace-supercyclic operators*, Ann. Funct. Anal. **6** (2015), 60–68
[7] X. F. Zhao, , Y. L. Shu and Y. H. Zhou, *Subspace-supercyclicity and common subspace-supercyclic vectors*, J. East China Norm. Univ. (Natural Science), **1** (2012), 106.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–166, pp. 69–78

Nonexistence and Multiplicity results for binonlocal Leray-Lions type problems

Z. Naghizadeh[a,*]

[a] *Faculty of Sciences, Department of Mathematics , University of Science and Technology of Mazandaran, P.O. Box 48518-78195, Behshahr, Iran*

Article Info

Keywords:
$r(x)$-Kirchhoff type problems
Leray-Lions type operators
Variational techniques

2020 MSC:
35D30
35J35
35J60

Abstract

The Leray-Lions operators accept many attentions because they are flexible enough to be specified to different elliptic operators. The goal of this paper is to obtain the existence of at least three distinct weak solutions for a Leray-Lions problem of $r(x)$-Kirchhoff type and nonexistence result in the exponent constant case. The technique is constructed on variational methods.

1. Introduction

The study of Leray-Lions type operator is a new subject for investigation, because they happen in some field, like as electrorheological fluids [12], image processing [14] and etc. Recently, some fourth order Leray-Lions type problems have been investigated. For instance, in [10] by using critical point theorem of [2], the authors ensured multiplicity of weak solutions for a nonlocal biharmonic system including Hardy potential and Leray-Lions operator. In [9], a multiplicity theorem for a fourth-order Leray Lions equation including indefinite weights, was established.
Relatively speaking, biharmonic $r(x)-$Kirchhoff type problems consisting of Leray-Lions operators have rarely been considered. In [13], by applying critical point theory and variational approach, some multiplicity results for a Leray-Lions $r(x)$-Kirchhoff type problem was obtained. Besides, the study of bi-nonlocal problems including Kirchhoff type operator together with the external force term with a nonlocal coefficient can better qualify multiple biological and physical systems (see [5]).
Here, we consider the following form of binonlocal $r(x)-$Kirchhoff type problems including Leray-Lions operator:

$$\begin{cases} M_1\Big(H_{r(x)}(u)\Big)\Delta\left(a(x,\Delta u)+|u|^{r(x)-2}u\right)=\lambda M_2\Big(K(u)\Big)f(x,u(x)) & \text{in } \Omega, \\ u=\Delta u=0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

*Talker
Email address: z.naghizadeh@mazust.ac.ir (Z. Naghizadeh)

in which
$$H_{r(x)}(u) = \int_\Omega \Big[A(x,\Delta u) + \frac{1}{r(x)}|u|^{r(x)}\Big]dx,\;\; K(u) = \int_\Omega F(x,u(x))dx, \tag{2}$$
and $\Omega \subset \mathbb{R}^N$ $(N \geq 2)$ represents a bounded domain with smooth boundary, $F(x,t) = \int_0^t f(x,\gamma)d\gamma$, $a : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ denotes a Carathéodory function obeying the subsequent assumptions:

(**A1**) $a(x,0) = 0,$ for $a.e.$ $x \in \overline{\Omega}$.

(**A2**) a verifies the growth condition
$$|a(x,t)| \leq c\Big(|t|^{r(x)-1} + g(x)\Big),\;\; \forall t \in \mathbb{R},\; a.e.\; x \in \Omega,$$
in which $c > 0$ denotes some constant, $g \in L^{\frac{r(x)}{r(x)-1}}(\Omega)$, denotes a nonnegative function and $r \in C(\overline{\Omega})$ denotes a Log-Holder continuous function obeying the relationship
$$1 < r^- := \inf_{x\in\Omega} r(x) \leq r(x) \leq r^+ := \sup_{x\in\Omega} r(x) < \frac{N}{2}. \tag{3}$$

(**A3**) For every $t, s \in \mathbb{R}$,
$$\Big(a(x,s) - a(x,t)\Big)(s-t) \geq 0 \text{ for a.e. } x \in \Omega.$$

(**A4**) There is $0 < \bar{c} < 3\min\{c,1\}$, obeying the following relationship
$$\bar{c}|t|^{r(x)} \leq \min\{r(x)A(x,t), a(x,t)t\}\; \forall t \in \mathbb{R},\; a.e.\; x \in \Omega,$$
in which $A : \overline{\Omega} \times \mathbb{R} \to \mathbb{R}$ denotes the primitive function of a, in other words,
$$A(x,t) = \int_0^t a(x,\gamma)d\gamma,$$

the operators $\Delta\Big(a(x,\Delta u)\Big)$ represents the fourth order Leray-Lions. If we consider
$$a(x,t) = \varrho(x)|t|^{r(x)-2}t, \tag{4}$$
in which $r \in C_+(\overline{\Omega}), r^+ < +\infty$, and select $\varrho \in L^\infty(\Omega)$ obeying the relationship
$$\exists \varrho_0 > 0\;;\; \varrho(x) \geq \varrho_0 > 0,\; \text{for a.e. } x \in \Omega,$$
so, (4) satisfies conditions (**A1**) $-$ (**A4**) and we achieve the following operator
$$\varrho(.)\Delta\left(|\Delta u|^{r(.)-2}\Delta u\right).$$
Whenever $\varrho \equiv 1$, the upper operator becomes the well-known $r(x)-$biharmonic operator $\Delta^2_{r(.)}$, [8].
Now, we give the hypothesis concerning the functions M_1, M_2 and f:

(**M**) $M_1 : \mathbb{R}^+ \to \mathbb{R}^+$ and $M_2 : \mathbb{R}^+ \to \mathbb{R}^+$ denote continuous functions and there exist three constants $m_1, m_2, m_2' > 0$ with $0 < m_2 \leq m_2'$ and two constants $\beta, \alpha > 1$ obeying the relationship
$$M_1(t) \geq m_1 t^{\alpha-1}, \qquad m_2 t^{\beta-1} \leq M_2(t) \leq m_2' t^{\beta-1},\;\; \forall t \geq 0.$$
$f : \Omega \times \mathbb{R} \to \mathbb{R}$ denotes a function described as
$$f(x,t) = \begin{cases} f_1(x,t) & |t| \geq 1, \\ f_2(x,t) & |t| < 1, \end{cases} \tag{5}$$
and verifies the following assumption:

(**F**) There are $o_i \in C(\overline{\Omega})$ and $c_i > 0, i = 1, 2$, obeying the relationship

$$|f_i(x,t)| \leq c_i|t|^{o_i(x)-1}, \quad 1 < o_1^- \leq o_1(x) \leq o_1^+ < \frac{\alpha r^-}{\beta} < \frac{\alpha r^+}{\beta} < o_2^- \leq o_2(x) < r_2^*(x), \tag{6}$$

in which $r_2^*(x)$ will be defined by (8) with $m = 2$.

The paper contains four sections as follows: Section 2 gives some background and notations related to the function space. The multiplicity and nonexistence theorems are presented in Section 3, whereas the proofs of this theorems are stated in subsections 4.1 and 4.2, respectively.

2. Background

We begin by recalling some essential notions of the generalized Sobolev spaces that will be applied in the next sections. For any $r \in C(\overline{\Omega})$ obeying the relationship (3), we describe the Lebesgue variable exponent space as

$$L^{r(x)}(\Omega) := \left\{ u : \Omega \to \mathbb{R} \text{ measurable} : \int_\Omega |u(x)|^{r(x)} \mathrm{d}x < +\infty \right\},$$

including the norm

$$|u|_{r(x)} := \inf \left\{ \nu > 0 : \int_\Omega \left| \frac{u(x)}{\nu} \right|^{r(x)} \mathrm{d}x \leq 1 \right\},$$

and the Holder-type inequality

$$\Big| \int_\Omega uv dx \Big| \leq \Big(\frac{1}{r^-} + \frac{1}{(\frac{r}{r-1})^-} \Big) |u|_{r(x)} |v|_{\frac{r(x)}{r(x)-1}}, \ \forall u \in L^{r(x)}(\Omega), v \in L^{\frac{r(x)}{r(x)-1}}(\Omega), \tag{7}$$

keeps true.

Proposition 2.1 (see [7, Proposition 2.7]). *If* $u \in L^{r(x)}(\Omega)$*, then*

$$\min \left\{ |u|_{r(x)}^{r^-}, |u|_{r(x)}^{r^+} \right\} \leq \int_\Omega |u(x)|^{r(x)} dx \leq \max \left\{ |u|_{r(x)}^{r^-}, |u|_{r(x)}^{r^+} \right\}.$$

For $m = 1, 2$, the variable exponent Sobolev space is described as

$$W^{m,r(x)}(\Omega) := \left\{ u \in L^{r(x)}(\Omega) : \ D^\delta u \in L^{r(x)}(\Omega), |\delta| \leq m \right\},$$

in which $\delta = (\delta_1, \ldots, \delta_N)$ denotes a multi-index, $|\delta| = \sum\limits_{j=1}^{N} \delta_j$ and $D^\delta u = \frac{\partial^{|\delta|}}{\partial x_1^{\delta_1} \ldots \partial x_N^{\delta_N}} u$.
The norm of this space is characterized by

$$\|u\|_{m,r(x)} := \inf \left\{ \nu > 0 : \ B_{r(x)}(\frac{u}{\nu}) \leq 1 \right\},$$

in which the modular $B_{r(x)} : W^{m,r(x)}(\Omega) \to \mathbb{R}$, is described as

$$B_{r(x)}(u) = \int_\Omega (|\Delta u(x)|^{r(x)} + |u|^{r(x)}) dx.$$

Let $W_0^{1,r(x)}(\Omega) = \overline{C_0^\infty(\Omega)}^{\text{in } W^{1,r(x)}(\Omega)}$. Our workspace Z is described as

$$Z := W^{2,r(x)}(\Omega) \cap W_0^{1,r(x)}(\Omega),$$

with

$$\|u\|_Z = \|u\|_{m,r(x)},$$

which displays separable and reflexive Banach space (see [4, 11]). Furthermore, the next embedding proposition take place.

Proposition 2.2 (see [6, Theorem 2.3]). *If $h \in C(\overline{\Omega})$ obeying the relationship $1 < h^- \leq h^+ \leq \infty$ and $h(x) \leq r_m^*(x), \forall x \in \overline{\Omega}$, in which*

$$r_m^*(x) = \begin{cases} \frac{Nr(x)}{N-mr(x)} & r(x) < \frac{N}{m}, \\ +\infty & r(x) \geq \frac{N}{m}. \end{cases} \tag{8}$$

Then the embedding $Z \hookrightarrow L^{h(x)}(\Omega)$ is continuous. If $h(x) < r_m^*(x)$ for each $x \in \overline{\Omega}$, the embedding becomes compact.

By Proposition 2.1, We arrive at the subsequent Proposition.

Proposition 2.3. *For every $u \in Z$, we get*

(i) $\|u\|_Z < 1(> 1; = 1) \Leftrightarrow B_{r(x)}(u) < 1((> 1; = 1);$

(ii) $\min\left\{\|u\|_Z^{r^-}, \|u\|_Z^{r^+}\right\} \leq B_{r(x)}(u) \leq \max\left\{\|u\|_Z^{r^-}, \|u\|_Z^{r^+}\right\}.$

Remark 2.4. From conditions $(\mathbf{A1}) - (\mathbf{A4})$, we conclude that the function $A(x,t)$ is C^1-Carathéodory and there is $\tilde{c} > 0$, obeying the following relation

$$\frac{\bar{c}}{r(x)}|t|^{r(x)} \leq |A(x,t)| \leq \tilde{c}\Big(|t|^{r(x)} + g(x)|t|\Big), \ \forall t \in \mathbf{R} \text{ and } a.e.\ x \in \Omega, \tag{9}$$

in which the constants $\bar{c}$, is as in condition $(\mathbf{A4})$.

Note that if Z^* represents the dual space of Z, then a mapping $G : Z \to Z^*$ is of $(S+)$ type if $u_j \rightharpoonup u$ and $\limsup_{j\to\infty}\langle G(u_j), u_j - u\rangle \leq 0$, imply $u_j \to u$ over Z.
A functional $B : Z \to \mathbb{R}$ is sequentially weakly lower semicontinous if $u_j \rightharpoonup u$ over Z implies $B(u) \leq \liminf_{j\to\infty} B(u_j)$.

Proposition 2.5 (see [3]). *Suppose that $(\mathbf{A1}) - (\mathbf{A4})$ are fulfilled and the functional $H_{r(x)} : Z \to \mathbb{R}$ characterized by (2). Then we get*

(i) *$H_{r(x)} \in C^1(Z, \mathbb{R})$ with derivative denoted by*

$$\langle H'_{r(x)}(u), \eta\rangle = \int_\Omega a(x, \Delta u)\Delta\eta dx + \int_\Omega |u|^{r(x)-2}u\eta dx, \tag{10}$$

for all $\eta \in Z$.

(ii) *$H_{r(x)}$ is sequentially weakly lower semicontinuous.*

(iii) *$H'_{r(x)} : Z \to Z'$ denotes a mapping of (S_+) type .*

Proposition 2.6. *Suppose that $(\mathbf{F})$ is fulfilled. Consider the functional $K : Z \to \mathbb{R}$ by (2). Then we get*

(i) *$K \in C^1(Z, \mathbb{R})$ together with derivative denoted by*

$$\langle K'(u), \eta\rangle = \int_\Omega f(x, u)\eta dx, \tag{11}$$

for all $\eta \in Z$.

(**ii**) *K is sequentially weakly continuous over Z, that is, $u_j \rightharpoonup u$ implies that $K(u_j) \to K(u)$.*

Proof. By (**F**), the proof of statement (**i**) is immediate. Now let $\{u_j\}$ be any sequence with $u_j \rightharpoonup u$ over Z. By using (**F**) and (7), we arrive at

$$\begin{aligned}|K(u_j)-K(u)| &\leq c_1\int_\Omega |u+\mu_j(u_j-u)|^{o_1(x)-1}|u_j-u|dx + c_2\int_\Omega |u+\mu_j(u_j-u)|^{o_2(x)-1}|u_j-u|dx\\ &\leq 2c_1\Big||u+\mu_j(u_j-u)|^{o_1(x)-1}\Big|_{\frac{o_1(x)}{o_1(x)-1}}|u_j-u|_{o_1(x)}\\ &\quad+2c_2\Big||u+\mu_j(u_j-u)|^{o_2(x)-1}\Big|_{\frac{o_2(x)}{o_2(x)-1}}|u_j-u|_{o_2(x)}\end{aligned} \tag{12}$$

in which $\forall x \in \Omega;\ 0 \leq \mu_j(x) \leq 1$. Besides, since $Z \hookrightarrow L^{o_i(x)}(\Omega)$ is compact for $i = 1, 2$, $u_j \to u$ in $L^{o_i(x)}(\Omega)$. So, it follows from (12) that (**ii**) holds true . □

Now, we will express the subsequent theorem that will be indispensable to prove the multiplicity result of this paper.

Proposition 2.7 (see [1, Theorem 2.1])**.** *Assume that $J, I : Z \to \mathbb{R}$ denote two continuously Gâteaux differentiable functionals over reflexive and separable real Banach space Z. If $I(z) \geq 0$ for each $z \in Z$ and there is $z_0 \in Z$ together with $I(z_0) = J(z_0) = 0$ and there are $\eta_0 > 0, z_1 \in Z$ so that*

(**i**) $\eta_0 < I(z_1)$;

(**ii**) $\sup_{I(z)<\eta_0} J(z) < \eta_0 \frac{J(z_1)}{I(z_1)}$.
Moreover, put

$$\xi = \frac{\sigma\eta_0}{\eta_0\frac{J(z_1)}{I(z_1)} - \sup_{I(z)<\eta_0} J(z)},$$

with $\sigma > 1$, and if $I - \lambda J$ denotes a sequentially weakly lower semicontinuous functional, verifies the (PS) condition and

(**iii**) $\lim_{\|z\|\to+\infty}(I(z) - \lambda J(z)) = +\infty$ *for all $\lambda \in [0, \xi]$.*

Then there is a number $\mu > 0$ and an open interval $L \subset [0, \xi]$, so that for all $\lambda \in L$, the equation

$$I'(z) - \lambda J'(z) = 0,$$

possesses at least 3 distinct solutions over Z whose norms are smaller than μ.

3. Main results

We consider two functionals $J, I : Z \to \mathbb{R}$ as

$$I(u) = \widehat{M_1}\Big(H_{r(x)}(u)\Big), \quad J(u) = \widehat{M_2}\Big(K(u)\Big) \quad \forall u \in Z, \tag{13}$$

in which $\widehat{M_i}(t) = \int_0^t M_i(\gamma)d\gamma$ for $i = 1, 2$ and $H_{r(x)}(u)$ and $K(u)$ are defined as (2).

By Propositions 2.5 and 2.6, $J, I \in C^1(Z, \mathbb{R})$ and

$$\langle I'(u), \eta\rangle = M_1\Big(H_{r(x)}(u)\Big)\langle H'_{r(x)}(u), \eta\rangle, \qquad \langle J'(u), \eta\rangle = M_2\Big(K(u)\Big)\langle K'(u), \eta\rangle, \quad ,$$

for all $\eta, u \in Z$, in which $H'_{r(x)}$ and K' as defined in (10) and (11), respectively.
Any function $u \in Z$ is named a weak solution of problem (1), if the following relationship is verified:

$$\langle I'(u), \eta\rangle - \lambda\langle J'(u), \eta\rangle = 0, \ \ \forall \eta \in Z. \tag{14}$$

The multiplicity result can be described by the following theorems.

Theorem 3.1. *(Multiplicity result) Suppose that* $(\mathbf{A1})-(\mathbf{A4})$, $(\mathbf{M})$ *and* $(\mathbf{F})$ *are fulfilled. Furthermore, there exist* $\bar{t}>0$ *with* $F(x,\bar{t})>0$ *for all* $x\in\Omega$. *Then, there is an open interval* $L\subset[0,\xi]$ *and a number* $\mu>0$ *so that for each* $\lambda\in L$ *problem* (1) *possesses at least 3 weak solutions whose norms are smaller than* μ, *in which* ξ *will given later one.*

In the special case, when $r(x)\equiv r$ be a constant, problem (1) reduces to the following r-Kirchhoff type problem

$$\begin{cases} M_1\Big(H_r(u)\Big)\Delta\left(a(x,\Delta u)+|u|^{r-2}u\right)=\lambda M_2\Big(K(u)\Big)f(x,u(x)) & \text{in } \Omega,\\ u=\Delta u=0 & \text{on } \partial\Omega, \end{cases} \tag{15}$$

and the nonexistence result are stated by the subsequent theorem.

Theorem 3.2. *(Nonexistence result in the exponent constant case) Suppose that* $(\mathbf{A4})$ *is fulfilled. In case that* $r(x)\equiv r$, *the conditions* $(\mathbf{M})$ *and* $(\mathbf{F})$ *change as follow, respectively:*

$(\mathbf{M}')$ $M_1,M_2:\mathbb{R}^+\to\mathbb{R}^+$ *denote two continuous functions and there are two numbers* $m_1,m_2'>0$ *and two constant* $\beta,\alpha>1$ *verifying*

$$M_1(t)\geq m_1t^{\alpha-1},\quad M_2(t)\leq m_2't^{\beta-1}.$$

$(\mathbf{F}')$ *There is* $c_3>0$ *so that*

$$|f(x,t)|\leq c_3|t|^{\frac{\alpha r}{\beta}-1},\quad 1<\frac{\alpha r}{\beta}<r^*,\quad \forall t\in\mathbb{R} \text{ and } \forall \text{ a.e. } x\in\Omega.$$

Then there is $\lambda_0>0$ *so that, problem* (15) *hasn't any nontrivial weak solution over* Z *for any* $\lambda<\lambda_0$.

4. Proof the of the main results

4.1. Proof of Theorem 3.1

We will use Proposition 2.7 to prove Theorem 3.1. So, it is necessary to check all conditions of Proposition 2.7.

Lemma 4.1. *The functional* $I-\lambda J$ *is sequentially weakly lower semicontinuous over* Z *for each* $\lambda>0$.

Proof. Assume that $\{u_j\}$ is a sequence that $u_j\rightharpoonup u$ over Z. By (ii) in Proposition 2.5, we get

$$\liminf_{j\to\infty}H_{r(x)}(u_j)\geq H_{r(x)}(u).$$

Besides, the function $t\to\hat{M}_1(t)$ is monotone. So, we deduce that

$$\liminf_{j\to\infty}I(u_j)=\liminf_{j\to\infty}\widehat{M_1}\Big(H_{r(x)}(u_j)\Big)\geq\widehat{M_1}\Big(\liminf_{j\to\infty}H_{r(x)}(u_j)\Big)\geq\widehat{M_1}\Big(H_{r(x)}(u)\Big)=I(u).$$

So, I is sequentially weakly lower semicontinuous over Z. By (ii) in Proposition 2.6, we arrive at

$$\lim_{j\to\infty}K(u_j)=K(u).$$

Besides, since the function $t\to\hat{M}_2(t)$ is continuous, we arrive at

$$\lim_{j\to\infty}J(u_j)=\lim_{j\to\infty}\widehat{M_2}\Big(K(u_j)\Big)=\widehat{M_2}\Big(\lim_{j\to\infty}K(u_j)\Big)=\widehat{M_2}\Big(K(u)\Big)=J(u).$$

Hence, J is sequentially weakly continuous and hence $I-\lambda J$ is sequentially weakly lower semicontinuous. □

Lemma 4.2. $I-\lambda J$ *represents a coercive functional, that is,* $\lim_{\|u\|\to+\infty}[I(u)-\lambda J(u)]=+\infty$.

Proof. Let $u \in Z$ with $\|u\|_Z > 1$. Put

$$\Omega_1 = \{x \in \Omega;\ |u(x)| \geq 1\},\ \ \Omega_2 = \{x \in \Omega;\ |u(x)| < 1\}.$$

By (**M**), (**F**) and (9), we arrive at

$$(I - \lambda J)(u) \geq \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\Big(B_{r(x)}(u)\Big)^\alpha - \lambda\frac{m_2'}{\beta}\Big(\frac{c_1}{o_1^-}\int_{\Omega_1}|u|^{o_1(x)}dx + \frac{c_2}{o_2^-}\int_{\Omega_2}|u|^{o_2(x)}dx\Big)^\beta.$$

By (6), we infer that

$$\begin{aligned}(I - \lambda J)(u) &\geq \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\Big(B_{r(x)}(u)\Big)^\alpha - \lambda\frac{m_2'}{\beta}\Big(\frac{c_1}{o_1^-}\int_{\Omega_1}|u|^{o_1(x)}dx + \frac{c_2}{o_2^-}\int_{\Omega_2}|u|^{o_1(x)}dx\Big)^\beta \\ &\geq \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\Big(B_{r(x)}(u)\Big)^\alpha - \lambda\frac{m_2'}{\beta o_1^-}\Big(\max\{c_1,c_2\}\Big)^\beta\Big(\int\limits_\Omega |u(x)|^{o_1(x)}dx\Big)^\beta.\end{aligned}$$

Now, since the embedding $Z \hookrightarrow L^{o_1(x)}(\Omega)$ is continuous, we infer that

$$\exists\, e_1 > 0;\ \ e_1|u|_{o_1(x)} \leq \|u\|_Z.$$

Since $\|u\|_Z > 1$, by Proposition 2.1, we arrive at

$$\int_\Omega |u|^{o_1(x)}dx \leq \max\Big\{\frac{1}{e_1^{o_1^+}}, \frac{1}{e_1^{o_1^-}}\Big\}\|u\|_Z^{o_1^+}.$$

So, by using (ii) in proposition 2.3, we infer that

$$(I - \lambda J)(u) \geq \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\|u\|_Z^{\alpha r^-} - \lambda\frac{m_2'}{\beta o_1^-}\Big(\max\{c_1,c_2\}\Big)^\beta\Big(\max\Big\{\frac{1}{e_1^{o_1^+}}, \frac{1}{e_1^{o_1^-}}\Big\}\Big)^\beta\|u\|_Z^{\beta o_1^+}.$$

By (6), we have $\alpha r^- > \beta o_1^+$ and so $I - \lambda J$ is coercive. $\square$

Lemma 4.3. *For each $\lambda > 0$, $I - \lambda J$ verifies the (PS) condition, that is, each sequence $\{u_j\}$ obeying the following condition*

$$\Big|(I - \lambda J)(u_j)\Big| \leq c,\ \ (I' - \lambda J')(u_j) \to 0 \text{ over } Z^* \text{ as } j \to \infty, \tag{16}$$

admits a convergent subsequence in Z.

Proof. Let $\{u_j\}$ be a (PS) sequence for $I - \lambda J$. By Lemma 4.2, $I - \lambda J$ is coercive on Z, so by the first relation in (16), the sequence $\{u_j\}$ is bounded over Z. Via the reflexivity of Z, there is a subsequence indicated by $\{u_j\}$ and some $u \in Z$ with $u_j \rightharpoonup u$. We will prove that $\{u_j\} \to u$. Indeed, since $u_j \rightharpoonup u$ over Z, we deduce that

$$M_1(H_{r(x)}(u_j))\langle H'_{r(x)}(u_j),(u_j - u)\rangle - \lambda M_2(K(u_j))\langle K'(u_j),(u_j - u)\rangle \to 0. \tag{17}$$

By (**F**), we infer that

$$\Big|\langle K'(u_j), u_j - u\rangle\Big| \leq c_1\int_{\Omega_1}|u_j|^{o_1(x)-1}|u_j - u|dx + c_2\int_{\Omega_2}|u_j|^{o_2(x)-1}|u_j - u|dx.$$

Since the embedding $Z \hookrightarrow L^{o_1(x)}$ is compact, by (6) and (7), we deduce that

$$\begin{aligned}\Big|\langle K'(u_j), u_j - u\rangle\Big| &\leq c_1\int_{\Omega_1}|u_j|^{o_1(x)-1}|u_j - u|dx + c_2\int_{\Omega_2}|u_j|^{o_1(x)-1}|u_j - u|dx \\ &\leq \max\{c_1,c_2\}\int_\Omega |u_j|^{o_1(x)-1}|u_j - u|dx \\ &\leq 2\max\{c_1,c_2\}\Big||u_j|^{o_1(x)-1}\Big|_{\frac{o_1(x)}{o_1(x)-1}}|u_j - u|_{o_1(x)} \to 0.\end{aligned} \tag{18}$$

Combining (18) with the continuity of the function M_2 and (ii) in Proposition 2.6, we arrive at

$$M_2\Big(K(u_j)\Big)\langle K'(u_j),(u_j-u)\rangle \to 0. \tag{19}$$

Hence by (17) and (19), we arrive at

$$M_1(H_{r(x)}(u_j))\langle H'_{r(x)}(u_j),(u_j-u)\rangle \to 0.$$

Besides, by (ii) in Proposition 2.3 and $(\mathbf{M})$, we arrive at

$$\begin{aligned}\Big|M_1\Big(H_{r(x)}(u_j)\Big)\langle H'_{r(x)}(u_j),(u_j-u)\rangle\Big| &\geq \Big|\frac{m_1(\min\{1,\bar{c}\})^{\alpha-1}}{(r^+)^{\alpha-1}}\Big(B_{r(x)}(u)\Big)^{\alpha-1}\Big| \times \Big|\langle H'_{r(x)}(u_j),(u_j-u)\rangle\Big| \\ &\geq \Big|\frac{m_1(\min\{1,\bar{c}\})^{\alpha-1}}{(r^+)^{\alpha-1}}\min\Big\{\|u_j\|_Z^{(\alpha-1)r^-},\|u_j\|_Z^{(\alpha-1)r^+}\Big\}\Big| \\ &\quad\times\Big|\langle H'_{r(x)}(u_j),(u_j-u)\rangle\Big| \geq 0.\end{aligned}$$

If $\|u_j\|_Z \to 0$ then $u_j \to 0$ over Z. Otherwise, $\|u_j\|_Z$ is bounded in Z. So

$$\lim_{j\to\infty}\langle H'_{r(x)}(u_j),(u_j-u)\rangle = 0.$$

By (iii) in Proposition 2.5, we obtain $u_j \to u$ over Z and this ends the proof. □

Here, we are prepare to prove Theorem 3.1.
Proof of Theorem 3.1 Obviously,

$$I(u)\geq 0 \ \ \forall u\in Z,\ \ I(0)=J(0)=0 \tag{20}$$

Given $\bar{t}$ as in Theorem 3.1. If $\tilde{\Omega}\subset\Omega$ denotes a sufficiently large compact subset and $\bar{u}\in C_0^\infty(\Omega)$, so that $0\leq\bar{u}(x)\leq\bar{t}$ over $\Omega\backslash\tilde{\Omega}$, $\bar{u}(x)=\bar{t}$ over $\tilde{\Omega}$. Then by $(\mathbf{M})$ and $(\mathbf{F})$, we arrive at

$$\begin{aligned}J(\bar{u}) &\geq \frac{m_2}{\beta}(K(\bar{u}))^\beta \geq \frac{m_2}{\beta}\Big(\int_{\tilde{\Omega}}F(x,\bar{t})dx - \frac{c_1}{o_1^-}\int_{(\Omega\backslash\tilde{\Omega})\cap\Omega_1}|\bar{u}|^{o_1(x)}dx - \frac{c_2}{o_2^-}\int_{(\Omega\backslash\tilde{\Omega})\cap\Omega_2}|\bar{u}|^{o_2(x)}dx\Big)^\beta \\ &\geq \frac{m_2}{\beta}\Big(\int_{\tilde{\Omega}}F(x,\bar{t})dx - \frac{c_1}{o_1^-}\max\Big\{|\bar{t}|^{o_1^+},|\bar{t}|^{o_1^-}\Big\}\Big|(\Omega\backslash\tilde{\Omega})\cap\Omega_1\Big| - \frac{c_2}{o_2^-}\max\Big\{|\bar{t}|^{o_2^+},|\bar{t}|^{o_2^-}\Big\}\Big|(\Omega\backslash\tilde{\Omega})\cap\Omega_2\Big|\Big)^\beta \\ &> 0,\end{aligned}$$

while $|\Omega\backslash\tilde{\Omega}|$ is small enough. Besides, we have

$$I(\bar{u}) = \widehat{M_1}\Big(H_{r(x)}(\bar{u})\Big) \geq \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\min\Big\{|\bar{t}|^{\alpha r^+},|\bar{t}|^{\alpha r^-}\Big\}|\tilde{\Omega}|^\alpha > 0.$$

Hence

$$0 < \frac{J(\bar{u})}{I(\bar{u})}. \tag{21}$$

Now, choose $0<\eta_0<\min\Big\{\frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}, I(\bar{u})\Big\}$. So, (i) in Proposition 2.7 is achieved with $z_1=\bar{u}$. Whenever $I(u)<\eta_0$, we arrive at

$$\frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}\Big(B_{r(x)}(u)\Big)^\alpha \leq I(u) < \eta_0 < \frac{m_1(\min\{1,\bar{c}\})^\alpha}{\alpha(r^+)^\alpha}. \tag{22}$$

So

$$B_{r(x)}(u) < 1,$$

and by (i) in Proposition 2.3, $\|u\|_Z < 1$ which implies by (22) that

$$\|u\|_Z^{r^+} = \min\{\|u\|_Z^{r^+}, \|u\|_Z^{r^-}\} \leq B_{r(x)}(u) \leq \Big(\frac{\eta_0 \alpha (r^+)^\alpha}{m_1(\min\{1,\bar{c}\})^\alpha}\Big)^{\frac{1}{\alpha}} = (\eta_0 c_4)^{\frac{1}{\alpha}}.$$

So, we get

$$\|u\|_Z \leq (\eta_0 c_4)^{\frac{1}{\alpha r^+}}. \tag{23}$$

By $(\mathbf{F})$ and (6), we arrive at

$$\begin{aligned} J(u) &\leq \frac{m_2'}{\beta}(K(u))^\beta \leq \frac{m_2'}{\beta}\Big(\frac{c_1}{o_1^-}\int_{\Omega_1}|u|^{o_1(x)}dx + \frac{c_2}{o_2^-}\int_{\Omega_2}|u|^{o_2(x)}dx\Big)^\beta \\ &\leq \frac{m_2'}{\beta}\Big(\frac{c_1}{o_1^-}\int_{\Omega_1}|u|^{o_2(x)}dx + \frac{c_2}{o_2^-}\int_{\Omega_2}|u|^{o_2(x)}dx\Big)^\beta \\ &\leq \frac{m_2'}{\beta o_1^-}\Big(\max\{c_1,c_2\}\Big)^\beta\Big(\int_\Omega |u|^{o_2(x)}dx\Big)^\beta, \end{aligned}$$

Now, by the continuous embedding $Z \hookrightarrow L^{o_2(x)}(\Omega)$, we infer that

$$\exists\, e_2 > 0; \quad e_2|u|_{o_2(x)} \leq \|u\|_Z.$$

Since $\|u\|_Z < 1$, by Proposition 2.1, we have

$$\int_\Omega |u|^{o_2(x)}dx \leq \max\Big\{\frac{1}{e_2^{o_2^+}}, \frac{1}{e_2^{o_2^-}}\Big\}\|u\|_Z^{o_2^-}.$$

By (23), we arrive at

$$\begin{aligned} J(u) &\leq \frac{m_2'}{\beta o_1^-}\Big(\max\{c_1,c_2\}\Big)^\beta\Big(\max\Big\{\frac{1}{e_2^{o_2^+}}, \frac{1}{e_2^{o_2^-}}\Big\}\Big)^\beta \|u\|_Z^{\beta o_2^-} \\ &\leq \frac{m_2'}{\beta o_1^-}\Big(\max\{c_1,c_2\}\Big)^\beta\Big(\max\Big\{\frac{1}{e_2^{o_2^+}}, \frac{1}{e_2^{o_2^-}}\Big\}\Big)^\beta (\eta_0 c_4)^{\frac{\beta o_2^-}{\alpha r^+}} = c_5 \eta_0^{\frac{\beta o_2^-}{\alpha r^+}}. \end{aligned}$$

Note that by (6), we obtain $\alpha r^+ < \beta o_2^-$. Therefore, relation (21) permits us to select η_0 small enough so that

$$\frac{\sup_{I(u)<\eta_0} J(u)}{\eta_0} < \frac{J(\bar{u})}{I(\bar{u})}, \tag{24}$$

and (ii) in Proposition 2.7 is achieved. Lemmas 4.1-4.3 and relations (20),(21) and (24) permits us to apply Proposition 2.7 with $z_0 = 0, z_1 = \bar{u}$. Thus there exists a number $\mu > 0$ and an open interval $L \subset [0,\xi]$, in which

$$\xi = \frac{\sigma \eta_0}{\eta_0 \frac{J(\hat{u})}{I(\hat{u})} - \sup_{I(u)<\eta_0} J(u)},$$

with $\sigma > 1$, so that for all $\lambda \in L$ problem (1) possesses at the minimum 3 distinct solutions over Z whose norms are smaller than μ.

4.2. Proof of Theorem 3.2

Note that we consider the exponent constant case $r(x) \equiv r$ be a constant. So, our work space will be $Z = W^{2,r} \cap W_0^{1,r}(\Omega)$ with

$$\|u\|_{Z_0} = \Big(B_r(u)\Big)^{\frac{1}{r}} = \Big(\int_\Omega (|\Delta u|^r + |u|^r)dx\Big)^{\frac{1}{r}}.$$

Since $1 < \frac{\alpha r}{\beta} < r^*$, the embedding $Z \hookrightarrow L^{\frac{\alpha r}{\beta}}(\Omega)$ is continuous. Let $e_3 > 0$ be the best Sobolev constants for that embedding, that is,

$$e_3 = \inf_{0 \neq u \in Z} \frac{\|u\|_Z}{|u|_{\frac{\alpha r}{\beta}}}.$$

Now, if u denotes any weak solution for the problem (1). By $(\mathbf{M}')$, $(\mathbf{A4})$ and $(\mathbf{F}')$, there is a number $c_6 > 0$ so that

$$\frac{m_1(\min\{1, \bar{c}\})^{\alpha}}{(r)^{\alpha-1}} \|u\|_{Z_0}^{\alpha r} \leq M_1\Big(H_{r(x)}(u)\Big)\langle H'_{r(x)}(u), u\rangle = \lambda M_2\Big(K(u)\Big)\langle K'(u), u\rangle \leq \lambda c_6 m'_2 |u|^{\alpha r}_{\frac{\alpha r}{\beta}}.$$

By taking $\lambda_0 = \frac{m_1(\min\{1,\bar{c}\})^{\alpha}}{(r)^{\alpha-1}} \frac{e_3^{\alpha r}}{m'_2 c_6}$, Theorem 3.2 is proved.

References

[1] G. Bonanno. Some remarks on a three critical points theorem. *Nonlinear Analysis: Theory, Methods & Applications*, 54(4):651–665, 2003.

[2] G. Bonanno and S. A. Marano. On the structure of the critical set of non-differentiable functions with a weak compactness condition. *Appl. Anal.*, 89(1):1–10, 2010.

[3] M.-M. Boureanu. Fourth-order problems with Leray-Lions type operators in variable exponent spaces. *screte Contin. Dyn. Syst. Ser. S*, 12(2), 2019.

[4] L. Diening, P. Harjulehto, P. Hästö, and M. Ruzicka. *Lebesgue and Sobolev Spaces with Variable Exponents*, volume 2017. Springer Science & Business Media, 2011.

[5] X. Fan. On nonlocal $p(x)$-Laplacian Dirichlet problems. *Nonlinear Analysis: Theory, Methods & Applications*, 72(7-8):3314–3323, 2010.

[6] X. Fan and D. Zhao. On the spaces $L^{p(x)}(\Omega)$ and $W^{m,p(x)}(\Omega)$. *J. Math. Anal. Appl.*, 263(2):424–446, 2001.

[7] Y. Karagiorgos and N. Yannakakis. A Neumann problem involving the $p(x)$-Laplacian with $p = \infty$ in a subdomain. *Adv. Calc. Var.*, 9(1): 65–76, 2016.

[8] K. Kefi and V. Rădulescu. On a $p(x)$-biharmonic problem with singular weights. *Z. Angew. Math. Phys.*, 68(4):1–13, 2017.

[9] K. Kefi, N. Irzi, and M. M. Al-shomrani. Existence of three weak solutions for fourth-order Leray–Lions problem with indefinite weights. *Complex Var. Elliptic Equ.*, pages 1–12, 2022.

[10] Z. Musbah and A. Razani. A class of biharmonic nonlocal quasilinear systems consisting of Leray–Lions type operators with Hardy potentials. *Bound. Value Probl.*, 2022(1):1–14, 2022.

[11] V. D. Radulescu and D. D. Repovs. *Partial differential equations with variable exponents: variational methods and qualitative analysis*, volume 9. CRC Press, Boca Raton, 2015.

[12] M. Ruzicka. Electrorheological fluids: modeling and mathematical theory. *Lecture Notes in Math.*, 1748:16–38, 2000.

[13] K. Soualhine, M. Filali, M. Talbi, and N. Tsouli. Three weak solutions for a class of fourth order $p(x)$-Kirchhoff type problem with Leray-Lions operators. *Bol. Soc. Paran. Mat*, 1799:1–13, 2022. doi: http://dx.doi.org/10.5269/bspm.62936.

[14] C. Yun-rnei, S. Levine, and R. Hurali. Variable exponent, linear growth functionals in image processing. i. *SIAM J. Appl. Math.*, 66(4): 1383–1406, 2006.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–168, pp. 79–82

Multiplicative Linear Mappings in Continuous Inverse Fréchet Algebras

Ali Naziri-Kordkandi*

Department of Mathematics, Payame Noor University, Tehran, Iran.

Article Info

Keywords:
Continuous inverse Fréchet algebra
Multiplicative linear functional
Spectrum

2020 MSC:
46H05
46H20

Abstract

A characterization of multiplicative linear functionals in Banach algebras was given by Gleason-Kahane-Zelazko. A version of the Gleason-Kahane-Zelazko theorem was also proved for certain algebras. In this paper, we investigate a characterization of multiplicative linear mappings, in particular, multiplicative linear functionals in continuous inverse Fréchet algebras.

1. Introduction

Throughout this paper, all algebras are complex and unital. The identity element of an algebra A is denoted by e. Multiplicative linear mappings, in particular, multiplicative linear functionals play a crucial role in functional analysis and topological algebras. One of the famous results in this direction is the classical Gleason-Kahane-Zelazko theorem [3], which states that every unital, invertibility-preserving, linear functional on a Banach algebra is necessarily multiplicative. Kowalski and Slodkowski [5] gave a characterization of multiplicative linear functionals on a Banach algera A without the linearity assumption. The characterization is:
Let $f : A \to \mathbb{C}$ satisfying $f(0) = 0$ and $f(x) - f(y) \in \mathrm{sp_A}(x - y)$ for each $x, y \in A$. Then f is multiplicative and linear. In fact, they generalized the classical Gleason-Kahane-Zelazko theorem on Banach algeras for not necessary linear functionals. In this paper, We try to generalize the Kowalski-Slodkowski theorem for continuous inverse Fréchet algebras.

2. Preliminaries

In this section, we present a collection of definitions and known results.

*Talker
Email address: ali_ naziri@pnu.ac.ir (Ali Naziri-Kordkandi)

Definition 2.1. For an algebra A, the spectrum $sp_A(x)$ of an element $x \in A$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda e - x$ is not invertible in A. The spectral radius $r_A(x)$ of an element $x \in A$ is defined by $r_A(x) = \sup\{|\lambda| : \lambda \in sp_A(x)\}$. Also, the set of all invertible elements of A is denoted by $Inv(A)$.

Definition 2.2. Definition 1.1. Let X and Y be metric spaces with metrics d_X and d_X respectively. A mapping $f : X \to Y$ is called Lipschitz iff there exists an $0 \leq M$ such that

$$d_Y(f(x), f(y)) \leq M d_X(x, y) \quad \text{for all } \ x, y \in X.$$

Definition 2.3. A topological algebra A is called a Q-algebra if the set of all invertible elements of A is open.

Definition 2.4. [2] An F-algebra is a complete metrizable topological algebra. A locally m-convex F-algebra is called a Frechet algebra.

Definition 2.5. [2] A continuous inverse algebra is a locally convex algebra in which the set of invertible elements is a neighbourhood of e and inversion is continuous at e.

Proposition 2.6. *[2, 1.2] Let A be a continuous inverse algebra. Then $Inv(A)$ is an open subset of A, and inversion is a continuous map from $Inv(A)$ into itself.*

Theorem 2.7. *[7, 4.1] Let A be a continuous inverse algebra and $\varphi : A \longrightarrow \mathbb{C}$ be a linear functional on A. If $\varphi(x) \in sp_A(x)$, for all $x \in A$, then φ is multiplicative.*

Definition 2.8. (Gateaux derivative). Suppose that X and Y are locally convex topological vector spaces, $U \subset X$ is open, and $f : X \to Y$ is a function. The Gateaux differential of f at $a \in U$ in the direction $x \in X$, denoted by $(Df)_a(x)$, is defined as

$$(Df)_a(x) = \lim_{r \to 0} \frac{f(a + rx) - f(a)}{r}.$$

If the limit exists for every $x \in X$, then the function f is called Gateaux differentiable at a [1].

Definition 2.9. [5] Let X be a vector space. We say that a mapping $\varphi : X \to \mathbb{C}$ is complex (real) linear (shortly C-linear or R-linear) if it is additive and homogeneous with respect to complex (real) scalars.

Definition 2.10. [5] We say that a mapping in a Frechet space has a real differential at a point if it has a Gateaux differential with respect to real scalars which, in addition, is continuous.

Lemma 2.11. *[6, 3.1] Let X be a real or complex topological vector space. If a functional $f : X \to \mathbb{C}$ is additive and continuous, then it is $\mathbb{R}$-linear.*

3. Multiplicative linear mappings

A characterization of multiplicative linear functionals in Banach algebras was given by Gleason-Kahane-Zelazko [4]. A version of the Gleason-Kahane-Zelazko theorem is also generalized for continuous inverse algebras [7]. Now, we weaken the assumption of this theorem and prove it for continuous inverse Fréchet algebras without linearity. Before giving the main theorem, we need the following lemmas.

Lemma 3.1. *Let $f : \mathbb{C} \to \mathbb{C}$ be an entire function. If f is a Lipschitz function, then it is an affine mapping.*

Proof. Let f be holomorphic in an convex open set U. If f is Lipschitz, then for every distinc z_1 and z_2 in U, there exists a constant $m > 0$ such that

$$\left| \frac{f(z_2) - f(z_1)}{z_2 - z_1} \right| \leq m.$$

Keeping z_1 fixed and letting $z_2 \to z_1$, we have $|\, f'(z_1)| \leq m$. Since z_1 is arbitrary, f' is bounded in U. Now, if f is entire and Lipschitz, then f' is bounded in $U = \mathbb{C}$. Since derivatives of entire functions are themselves entire, it follows from Liouville's theorem that f' is constant. This implies that f is an affine mapping, i.e., for all $z \in \mathbb{C}$, f(z)=mz+n , where m, n are complex constants.

□

Lemma 3.2. *Let A be a continuous inverse algebra and φ be $\mathbb{R}$-linear on A such that $\varphi(x) \in \mathrm{sp}_{\mathrm{A}}(x)$ for each $x \in A$, then φ is $\mathbb{C}$-linear.*

Proof. The proof of this lemma follows by the same reasoning as in [5, 2.1]. □

Now, we prove the main result.

Theorem 3.3. *Let $(A, (p_n))$ be a continuous inverse Fréchet algebra and let $f : A \to \mathbb{C}$ satisfy $f(0) = 0$ and $f(x) - f(y) \in \mathrm{sp}_{\mathrm{A}}(x - y)$ for each $x, y \in A$. Then f is multiplicative and linear.*

Proof. We use the same way as in the proof of Theorem1.2 in [5]. First we may assume that A is separable. Suppose that f has a differential at a point $a \in A$. We have

$$\frac{f(a + rx) - f(a)}{r} \in \frac{\mathrm{sp}_{\mathrm{A}}(a + rx - a)}{r} = \mathrm{sp}_{\mathrm{A}}(x), \quad r \in R,\ r \neq 0,\ x \in A.$$

So,

$$(Df)_a(x) = \lim_{r\to 0} \frac{f(a + rx) - f(a)}{r} \in \mathrm{sp}_{\mathrm{A}}(x).$$

Thus, by Lemma 2.11, the differential is $\mathbb{R}$-linear, and by Lemma 3.2, it is $\mathbb{C}$-linear. On the other hand, since A is a Q-algebra, by [4, 6.18], we have $r_A(x) \leq p_m(x)$, for some $m \in \mathbb{N}$ and for all $x \in A$. Hence, we get

$$|f(x) - f(y)| \in |\mathrm{sp}_{\mathrm{A}}(x - y)| \leq p_m(x - y).$$

Thus, f is a Lipschitz function. By [5, 2.3] and [5, 2.4], we obtain that f is an entire function. We define the function $f_{a,b} : \mathbb{C} \to \mathbb{C}$ by $f_{a,b}(z) = f(az + b)$, for $a, b \in A$. Therefore, $f_{a,b}$ is Lipschitz and entire. By Lemma 3.1, it is affine. By the same reasoning as in [5, 1.2], and the Gleason-Kahane-Zelazko theorem for continuous inverse algebras [7, 4.1], we conclude that f is linear and multiplicative, when A is separable. Now, we consider the general case. Let $a, b \in A$. Clearly, $[e,a,b]$ (the subalgebra of A generated by e,a,b in A) is continuous inverse Fréchet subalgebra of A. The function f of Theorem 3.3, restricted to subalgebra $[e,a,b]$ of A satisfies conditions of the theorem. As $[e,a,b]$ is separable, from the preceding part of the proof, it follows that $f\big|_{[e,a,b]}$ is multiplicative and linear. Since a and b is arbitrary, we deduce that f is multiplicative and linear in the whole of A. □

It is well known that every Banach algebra is a continuous inverse Fréchet algebra, but the following example shows that the converse may be false in general.

Example 3.4. [7, 3.1] Consider the algebra $A = C^\infty[0, 1]$ of all C^∞- functions on $[0, 1]$ with topology τ defined by the algebra seminorms

$$p_n(f) = \sup_{0\leq t\leq 1} [\sum_{k=0}^{n} \frac{|f^{(k)}(t)|}{k!}].$$

Then (A, τ) is a Frechet algebra whose the set of invertible elements is open. So A is a continuous inverse algebra, but not a Banach algebra.

In the following, we present several results of Theorem 3.3.

Theorem 3.5. *Let A be a continuous inverse Fréchet algebra and let B be a semisimple and commutative continuous inverse Fréchet algebra. Suppose that f is a mapping from A into B such that*

$$\mathrm{sp}_{\mathrm{B}}(f(x) + f(y)) \subseteq \mathrm{sp}_{\mathrm{A}}(x + y), \quad \textit{for all } \ x, y \in A.$$

Then f is linear and multiplicative.

Theorem 3.6. *Let A be a continuous inverse Fréchet algebra and let $f : A \to \mathbb{C}$ be a functional such that*

$$f(a) \in \mathrm{sp}_A(a), \qquad \text{for all } \ a \in A.$$

If the Real differentials of f are constant, then f is linear and multiplicative.

Theorem 3.7. *Let A and B be a continuous inverse Fréchet algebras such that B is commutative and semisimple. Let $f : A \to \mathbb{C}$ be a functional such that*

$$\mathrm{sp}_B(f(a)) \subseteq \mathrm{sp}_A(a), \qquad \text{for all } \ a \in A.$$

If the Real differentials of f are constant, then f is linear and multiplicative.

Theorem 3.8. *Let A and B be continuous inverse Fréchet algebras and let B be commutative and semisimple. If $f : A \to B$ is a linear mapping such that $f(e_A) = e_B$ and $f(a) \in Inv(B)$ for $a \in Inv(A)$, then f is multiplicative.*

Theorem 3.9. *Let A be a continuous inverse Fréchet algebra and let B be a semisimple and commutative continuous inverse Fréchet algebra and $p(u, v) = \lambda u + \mu v \ \ (\lambda\mu \neq 0)$ be two-variable polynomials. Suppose that T is a mapping from A into B such that*

$$\mathrm{sp_B}(p(Tf, Tg)) \subseteq \mathrm{sp_A}(p(f, g)), \quad \text{for all } \ f, g \in A.$$

If $\lambda + \mu \neq 0$, then T is linear and multiplicative.

Corollary 3.10. *If $\lambda + \mu = 0$, in the above theorem, then $T - T(0)$ is linear and multiplicative.*

4. Conclusion

In this paper, we investigated some charactrizations of multiplicative linear mappings, in particular, multiplicative linear functionals in continuous inverse Fréchet algebras.

Acknowledgment. The authors would like to thank the referee for helpful suggestions.

References

[1] S. M. Abdolkhaleghzade, S. Effati and S. A. Rakhshan, *An efficient design for solving discrete optimal control problem with time-varying multi-delays*, Iran. J. Numer. Anal. Optim. **12** (2022), 719 − 738.

[2] H. Biller, *Continuous inverse algebras with involution*, Forum Mathematicum. **22** (2010), 1033 − 1059.

[3] F. F. Bonsal and J. Doncan, *Complete Normed Algebras*, Speringer-Verlag, Berlin, 1973.

[4] M. Fragoulopoulou, *Topological Algebras with Involution*, Elsevier, Amsterdam, 2005.

[5] S. Kowalski and Z. Slodkowaski, *A characterization of multiplicative linear functionals in Banach algebras*, Studia Math. **67** (1980), 215–223.

[6] M. Luczak, *A characterization of linear-multiplicative functionals in topological algebras*, Commentationes Mathematicae. **46** (2006), 79–91.

[7] A. Naziri-Kordkandi, *Topics on continuous inverse algebras*, J. Algebr. Syst. **9** (2022), 219 − 227.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–170, pp. 83–85

Some existence theorems for variational relation problems

Maryam Lotfipour[a,*]

[a]*Department of Mathematics, Faculty of Science, Fasa University, Fasa, Fars, Iran*

Article Info

Keywords:
variational relation
generalized KKM theorem
nonempty intersection problem

2020 MSC:
47H04
49J53

Abstract

In this paper, first we present some nonempty intersection theorems without the compactness of the domain of the set-valued maps. Then, by using these results, some variational relation problems are investigated under new conditions.

1. Introduction and preliminaries

There are some different problems in nonlinear analysis such as optimization, equilibrium problem and variational inclusion problem which have a similarity in their structure and it is useful to have a unified model to investigate them. In [6], Luc introduced a general framework for these problems as variational relation problem.
Let A, B, Y are nonempty sets. Consider nonempty-valued set-valued maps $S_1 : A \rightrightarrows A, S_2 : A \rightrightarrows B$ and $T : A \times B \rightrightarrows Y$ and let $R(a, b, y)$ be a relation for $a \in A, b \in B$ and $y \in Y$. Then the variational relation problem was defined in [6] as follows:
Find $\bar{a} \in A$ such that $\bar{a} \in S_1(\bar{a})$ and $R(\bar{a}, b, y)$ holds for each $b \in S_2(\bar{a}), y \in T(\bar{a}, b)$.
After that, many studies have been done to solve this problem in different spaces and different versions of variational relation problems were introduced; see for example [1, 2].
Let X, Y, Z be nonempty sets, $P : Y \rightrightarrows Z$ be a nonempty-valued map and $R(x, y, z)$ be a relation for $x \in X, y \in Y$ and $z \in Z$. In this paper, we study the following variational relation problem [1]:

$$\text{(VRP) Find } \bar{y} \in Y \text{ such that for each } x \in X \text{ there exists } z \in P(\bar{y}) \text{ such that } R(x, \bar{y}, z) \text{ holds.}$$

Here, first by using a KKM result, we obtain an intersection theorem for generalized KKM maps and then the variational relation problem (VRP) will be investigated.
Through this paper, all nonempty finite subsets of a set K is denoted by $\langle K \rangle$ and for each $A \in \langle K \rangle$, the convex hull of A is denoted by conv$A$.

---

*Talker

Email address: lotfipour@fasau.ac.ir (Maryam Lotfipour)

Let X, Y be topological spaces. A set-valued map $F : X \rightrightarrows Y$ is said to be upper semicontinuous, if for any closed subset B of Y the set $\{x \in X : F(x) \cap B \neq \emptyset\}$ is closed. Let X be a subset of a vector space E, then the set-valued map $F : X \rightrightarrows E$ is called to be a KKM map if

$$\mathrm{conv}A \subseteq \bigcup_{x \in A} F(x), \quad \text{for each } A \in \langle X \rangle.$$

Let $T, S : X \rightrightarrows Y$ be set-valued maps. The set-valued map S is said to be a generalized KKM map with respect to (w.r.t.) T if for each $A \in \langle X \rangle$,

$$T(\mathrm{conv}A) \subseteq S(A)$$

where $S(A) = \cup_{x \in A} S(x)$. The set-valued map F is said to be topological pseudomonotone [3] if

$$\mathrm{cl}(\bigcap_{u \in [a,b]} F(u)) \cap [a, b] = \bigcap_{u \in [a,b]} F(u) \cap [a, b], \ \forall a, b \in X$$

and F is said to be intersectionally closed on $A \subseteq X$, if

$$\bigcap_{x \in A} \mathrm{cl}(F(x)) = \mathrm{cl}(\bigcap_{x \in A} F(x)).$$

2. Main results

In [4], Ky Fan extended the well-known KKM Lemma [5] to infinite dimensional topological vector spaces and showed some applications of this result. Since then, it has become a useful tool for solving many other problems such as game theory and equilibrium problems. In [3], several new KKM-type theorems under new coercivity and closedness conditions were obtained. Here, in view of this theorem and studying some intersection theorems, we investigate a variational relation problem which includes many other problems such as variational and equilibrium problems.

Theorem 2.1. *([3])Let K be a nonempty and convex subset of a Hausdorff topological vector space E and $F : K \rightrightarrows K$. Suppose that the following conditions hold:*

(A1) F is a KKM map;

(A2) for each $A \in \langle K \rangle$, the set-valued map $F \cap \mathrm{conv}A$ is intersectionally closed on $\mathrm{conv}A$*;*

(A3) F is topological pseudomonotone;

(A4) there exist a nonempty subset B of K and a nonempty compact subset D of K such that $\mathrm{conv}(A \cup B)$ *is compact, for any $A \in \langle K \rangle$, and for each $y \in K \setminus D$ there exists $x \in \mathrm{conv}(B \cup \{y\})$ such that $y \notin F(x)$.*

Then, $\bigcap_{x \in K} F(x) \neq \emptyset$.

Lemma 2.2. *Suppose that $S, T : X \rightrightarrows Z$ be set-valued maps with nonempty values such that S is generalized KKM map w.r.t. T. Then the map $H : X \rightrightarrows X$ such that $H(x) := \{y \in X : T(y) \cap S(x) \neq \emptyset\}$ is a KKM map.*

Theorem 2.3. *Let X be a nonempty convex subset of a Hausdorff topological vector space E, Z be a nonempty set and $S, T : X \rightrightarrows Z$ be set-valued maps with nonempty values such that*

(I) S is generalized KKM map w.r.t. T;

(II) for each $x \in X$ and for each $A \in \langle X \rangle$, the set $\{y \in X : T(y) \cap S(x) \neq \emptyset\} \cap \mathrm{conv}A$ is intersectionally closed on $\mathrm{conv}A$*;*

(III) for each $x \in X$, the set $\{y \in Y : T(y) \cap S(x) \neq \emptyset\}$ is topological pseudomonotone;

(IV) there exist a nonempty subset B of X and a nonempty compact subset D of X such that $\mathrm{conv}(A \cup B)$ *is compact, for any $A \in \langle X \rangle$, and for each $y \in X \setminus D$ there exists $x \in \mathrm{conv}(B \cup \{y\})$ such that $T(y) \cap S(x) = \emptyset$.*

Then, there exists a point $\bar{x} \in X$ such that for each $x \in X$,

$$T(\bar{x}) \cap S(x) \neq \emptyset.$$

Corollary 2.4. *Let X be a nonempty convex subset of a Hausdorff topological vector space E, Z be a nonempty set and $S, T : X \rightrightarrows Z$ be set-valued maps with nonempty values such that*

(I) *S is generalized KKM map w.r.t. T;*

(II) *for each $x \in X$, the set $\{y \in Y : T(y) \cap S(x) \neq \emptyset\}$ is closed;*

(III) *there exist a nonempty subset B of X and a nonempty compact subset D of X such that $\mathrm{conv}(A \cup B)$ is compact, for any $A \in \langle X \rangle$, and for each $y \in X \setminus D$ there exists $x \in \mathrm{conv}(B \cup \{y\})$ such that $T(y) \cap S(x) = \emptyset$.*

Then, there exists a point $\bar{x} \in X$ such that for each $x \in X$,

$$T(\bar{x}) \cap S(x) \neq \emptyset.$$

Theorem 2.5. *Let X be a nonempty convex subset of a Hausdorff topological vector space E, Z be a topological space. Suppose that $P : X \rightrightarrows Z$ is an upper semicontinuous set-valued map with nonempty values and $R(x, y, z)$ is a relation such that*

(i) *for each $x \in X$, the set $\{(y, z) \in X \times Z : R(x, y, z) \text{ holds}\}$ is closed in $X \times Z$;*

(ii) *for each $K \in \langle X \rangle, y \in \mathrm{conv}K$ and $z \in P(y)$, we have $R(x, y, z)$ holds for some $x \in K$;*

(iii) *there exist a nonempty subset B of X and a nonempty compact subset D of X such that $\mathrm{conv}(A \cup B)$ is compact, for any $A \in \langle X \rangle$, and for each $y \in X \setminus D$ there exists $x \in \mathrm{conv}(B \cup \{y\})$ such that $R(x, y, z)$ does not hold, for each $z \in P(y)$.*

Then, the variational relation problem (VRP) has a solution.

References

[1] Balaj, M.: Intersection theorems for generalized weak KKM set-valued mappings with applications in optimization, Math. Nach. **294**, 1262–1276 (2021)

[2] Balaj, M.: Three types of variational relation problems, Taiwanese J. Math. **17** no. 1, 47—61 (2013)

[3] Fakhar, M., Lotfipour, M., Zafarani, J.: On the Brézis Nirenberg Stampacchia-type theorems and their applications. J. Glob. Optim. **55**(4), 751–770 (2013)

[4] Fan, K.: A genaralization of Tychonoff's fixed point theorem. Math. Ann. **142**, 305-310 (1961)

[5] Knaster, B., Kuratowski, C., Mazurkiewicz, S.: Ein beweis des fixpunktsatzes für n-dimensionale simplexe. Fund. Math. **14**, 132–137 (1929)

[6] Luc, D. T.: An abstract problem in variational analysis, J. Optim. Theory Appl. **138**, no. 1, 65–76 (2008)

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–181, pp. 86–88

A note on the local spectrum preserving maps

Rohollah Parvinianzadeh[a,*], Morad Ali Peyvand[b]

[a] Department of Mathematics, University of Yasouj, Yasouj 75918, Iran
[b] Department of Mathematics, University of Yasouj, Yasouj 75918, Iran

Article Info

Keywords:
Local spectrum
Rank-one operators
the single-valued extension property

2020 MSC:
47A11
47A15

Abstract

Let $B(H)$ denote the algebra of all bounded linear operators on a Hilbert spaces H. For a fix nonzero vector $h_0 \in H$, we show that if two surjective maps φ_1 and φ_2 from $B(H)$ into itself satisfies

$$\sigma_{\varphi_1(A)\varphi_2(B)^*}(h_0) = \sigma_{AB^*}(h_0)$$

for all $A, B \in B(H)$, then $\varphi_2(I)$ is invertible and there exists a bijective linear map $P : H \to H$ such that $\varphi_1(A) = PA(\varphi_2(I)^*P)^{-1}$ and $\varphi_2(A) = P^{-1}AP^*\varphi_2(I)$ for all $A \in B(H)$.

1. Introduction

The problem of characterizing linear or additive maps on $B(X)$, the algebra of all bounded linear operators on a complex Banach space X, preserving local spectra was initiated by A. Bourhim and T. Ransford in [5], and continued by several authors; see for instance [3] and the references therein. In [7], J. Bračič and V. Müller characterized surjective linear and continuous mappings on $B(X)$ preserving the local spectrum and local spectral radius at a fixed nonzero vector x_0 of X, and thus extending the main results of [6, 8] to infinite-dimensional Banach spaces. In [4], A. Bourhim and J. Mashreghi showed that surjective map φ on $B(X)$ preserving the local spectrum of product of operators at a fixed nonzero vector $x_0 \in X$ if and only if φ is a positive or negative multiple automorphism and x_0 is an eigenvector of the intertwining operator. Abdelali et al. [2] characterized maps φ that preserve the local spectrum at fixed nonzero vector of the skew-product operators. In this paper, we investigate the form of all maps φ_1 and φ_2 of $B(H)$ onto $B(H)$ such that, for every A and B in $B(H)$, the local spectrum of AB^* and $\varphi_1(A)\varphi_2(B)^*$ are the same at a nonzero fixed vector.

2. Preliminaries

The first lemma summarizes some known basic properties of the local spectrum.

*Talker
Email addresses: r.parvinian@yu.ac.ir (Rohollah Parvinianzadeh), peyvand@yu.ac.ir (Morad Ali Peyvand)

Lemma 2.1. *[1] Let X be a Banach space and $T \in B(X)$. For every $x, y \in X$ and a scalar $\alpha \in \mathbb{C}$ the following statements hold.*
(a) $\sigma_T(\alpha x) = \sigma_T(x)$ if $\alpha \neq 0$, and $\sigma_{\alpha T}(x) = \alpha\sigma_T(x)$.
(b) If $Tx = \lambda x$ for some $\lambda \in \mathbb{C}$, then $\sigma_T(x) \subseteq \{\lambda\}$. Further, if $x \neq 0$ and T has SVEP, then $\sigma_T(x) = \{\lambda\}$.

For a nonzero $h \in H$ and $T \in B(H)$, we use a useful notation defined by A. Bourhim and J. Mashreghi in [4] by

$$\sigma_T^*(h) := \begin{cases} \{0\} & if\ \sigma_T(h) = \{0\}, \\ \sigma_T(h) \setminus \{0\} & if\ \sigma_T(h) \neq \{0\}. \end{cases}$$

For any $x, y \in H$, let $x \otimes y$ denote the operator of rank at most one on H defined by

$$(x \otimes y)z = \langle z, y \rangle\, x, \qquad \forall\, z \in H.$$

The following lemma is an elementary observation that gives the nonzero local spectrum of any rank one operator.

Lemma 2.2. *(See [4, Lemma 2.2]) Let h_0 be a nonzero vector in H. For every vectors $x, y \in H$, the following statements hold.*
(a)

$$\sigma_{x\otimes y}^*(h_0) := \begin{cases} \{0\} & if\ \langle h_0, y\rangle = 0, \\ \langle x, y\rangle & if\ \langle h_0, y\rangle \neq 0. \end{cases}$$

(b) For all rank one operators $R \in B(H)$ and all $T, S \in B(H)$, we have

$$\sigma_{(T+S)R}^*(h_0) = \sigma_{TR}^*(h_0) + \sigma_{SR}^*(h_0)$$

The following theorem, which may be of independent interest, gives a spectral characterization of rank one operators in term of local spectrum.

Theorem 2.3. *(See [4, Theorem 4.1]) For a nonzero vector $h \in H$ and a nonzero operator $R \in B(H)$, the following statements are equivalent.*
(a) R has rank one.
(b) $\sigma_{RT}^*(h)$ contains at most one element for all $T \in B(H)$.
(c) $\sigma_{RT}^*(h)$ contains at most one element for all $T \in F_2(H)$.

The following result characterizes in term of the local spectrum when two operators are the same.

Lemma 2.4. *(See [4, Theorem 3.2]) For a nonzero vector h in H and two operators A and B in $B(H)$, the following statements are equivalent.*
(a) $A = B$.
(b) $\sigma_{AT}(h) = \sigma_{BT}(h)$ for all operators $T \in B(H)$.
(c) $\sigma_{AT}(h) = \sigma_{BT}(h)$ for all rank one operators $T \in B(H)$.
(d) $\sigma_{AT}^*(h) = \sigma_{BT}^*(h)$ for all rank one operators $T \in B(H)$.

3. Main Results

Throughout this paper, H is an infinite-dimensional complex Hilbert spaces and $B(H)$ denote the algebra of all bounded linear operators on a Hilbert space H, and its unit will be denoted by I. For an operator $T \in B(H, K)$, let T^* denote as usual its adjoint. The local resolvent set, $\rho_T(x)$, of an operator $T \in B(H)$ at a point $x \in H$ is the union of all open subsets U of the complex plane $\mathbb{C}$ for which there is an analytic function $f : U \longrightarrow H$ such that $(\mu I - T)f(\mu) = x$ for all $\mu \in U$. The complement of local resolvent set is called the local spectrum of T at x, denoted by $\sigma_T(x)$, and is obviously a closed subset (possibly empty) of $\sigma(T)$, the spectrum of T. Recall that an operator

$T \in B(H)$ is said to have the single-valued extension property (henceforth abbreviated to SVEP) if, for every open subset U of $\mathbb{C}$, there exists no nonzero analytic solution, $f : U \longrightarrow H$, of the equation

$$(\mu I - T)f(\mu) = 0, \quad \forall\, \mu \in U.$$

For more information about these notions one can see the books [1].

We begin with the following identity principal.

Lemma 3.1. *For a nonzero vector h in H and two operators A and B in $B(H)$, the following statements are equivalent.*
(a) $A = B$.
(b) $\sigma_{TA}(h) = \sigma_{TB}(h)$ for all operators $T \in B(H)$.
(c) $\sigma_{TA}(h) = \sigma_{TB}(h)$ for all rank one operators $T \in B(H)$.
(d) $\sigma^*_{TA}(h) = \sigma^*_{TB}(h)$ for all rank one operators $T \in B(H)$.

The followig theorem is the main result of this paper.

Theorem 3.2. *Let $h_0 \in H$ be a nonzero vector. Suppose that φ_1 and φ_2 be surjective maps from $B(H)$ into $B(H)$ which satisfy*

$$\sigma_{\varphi_1(A)\varphi_2(B)^*}(h_0) = \sigma_{AB^*}(h_0), \ \ (A, B \in B(H)). \tag{1}$$

Then $\varphi_2(I)$ is invertible and there exists a bijective linear map $P : H \to H$ such that $\varphi_1(A) = PA(\varphi_2(I)^*P)^{-1}$ and $\varphi_2(A) = P^{-1}AP^*\varphi_2(I)$ for all $A \in B(H)$.

Proof. The proof is long and we break it into several claims.

Claim 1. φ_1 and φ_2 are injective, and $\varphi_1(0) = 0, \varphi_2(0) = 0$.

Claim 2. φ_1 preserves rank one operators in both directions.

Claim 3. φ_1 is linear.

Claim 4. There exist bijective linear operators $C : H \to H$ and $D : H \to H$ such that $\varphi_1(x \otimes y) = Cx \otimes Dy$ for all $x, y \in H$.

Claim 5. $\varphi_2(I)$ is invertible.

Claim 6. The result in the theorem holds.

□

References

[1] P. Aiena, *Fredholm and local spectral theory, with applications to multipliers*, Kluwer, Dordrecht, 2004.
[2] Z. Abdelali, A. Achchi and R. Marzouki, *Maps preserving the local spectrum of skew-product of operators*, Linear Algebra Appl., **485** (2015), No. 3, pp. 58–71.
[3] A. Bourhim, J. Mashreghi, *A survey on preservers of spectra and local spectra*, Contemp Math. 45 (2015) 45-98.
[4] A. Bourhim, J. Mashreghi, *Maps preserving the local spectrum of product of operators*, Glasgow Math. J. 57 (2015) 709-718.
[5] A. Bourhim , T. Ransford, *Additive maps preserving local spectrum*. Integral Equations Operator Theory. 5 (2006) 377-385.
[6] A. Bourhim and V. G. Miller, Linear maps on $M_n(\mathbb{C})$ preserving the local spectral radius, *Studia Math.*, **188** (2008), pp. 67-75.
[7] J. Bracic and V. Muller,*Local spectrum and local spectral radius of an operator at a fixed vector*, Studia Math., **194** (2009), No. 2, pp. 155-162.
[8] M. Gonzalez and M. Mbekhta, Linear maps on $M_n(\mathbb{C})$ preserving the local spectrum, *Linear Algebra Appl.*, **427** (2007), pp. 176-182.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–184, pp. 89–92

Derivations on certain Banach algebras related to locally compact groups

Seyedeh Somayeh Jafari*

Department of mathematics, Payame Noor University, Tehran, Iran.

Article Info

Keywords:
Banach algebra
derivation
locally compact group
vanishing at infinity

2020 MSC:
43A15
46H20

Abstract

For a locally compact group $\mathcal{G}$, let $L_0^\infty(\mathcal{G})$ be the Banach space of all essentially bounded measurable functions on $\mathcal{G}$ vanishing at infinity. Here, we deal with a derivation problem for the Banach algebra $L_0^\infty(\mathcal{G})^*$ equipped with a multiplication of Arens type; in particular, our interest to us here are some identifications on this subject in term of abelian groups. For instance, we show that G is abelian if and only if every weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ is zero. Also, it's known the Singer-Wermer conjecture for $L_0^\infty(G)^*$ is true if G is abelian. We show that the Singer-Wermer conjecture for $L_0^\infty(G)^*$ is valid only this case.

1. Introduction and Preliminaries

Let $\mathcal{G}$ be a locally compact group with a fixed left Haar measure dt, and let $L^1(\mathcal{G})$ be the usual group algebra; i.e., the set of all measurable functions on $\mathcal{G}$ equipped with the convolution product $*$ and the norm $\|.\|_1$. Let also $L^\infty(\mathcal{G})$ denote the Lebesgue space, consisting of all locally essentially bounded measurable functions on $\mathcal{G}$ equipped with the essential supremum norm $\|.\|_\infty$. Then $L^\infty(\mathcal{G})$ is the dual of $L^1(\mathcal{G})$ for the pairing

$$\langle f, \phi\rangle = \int_{\mathcal{G}} f(t)\ \phi(t)\ dt.$$

for all $f \in L^\infty(\mathcal{G})$ and $\phi \in L^1(\mathcal{G})$; see for example [2]. Suppose that $L_0^\infty(\mathcal{G})$ is the subspace of $L^\infty(\mathcal{G})$ consisting of all elements $f \in L^\infty(\mathcal{G})$ that vanish at infinity; it means that for each $\varepsilon > 0$, there is a compact subset K of $\mathcal{G}$ for which $\|f\ \chi_{\mathcal{G}\setminus K}\|_\infty < \varepsilon$, where $\chi_{\mathcal{G}\setminus K}$ denotes characteristic function of $\mathcal{G}\setminus K$ on $\mathcal{G}$. For every $n \in L_0^\infty(\mathcal{G})^*$ and $g \in L_0^\infty(\mathcal{G})$, we denote by ng the function in $L^\infty(\mathcal{G})$ defined by

$$\langle ng, \phi\rangle = \langle n,\ g\cdot\phi\rangle \quad (\phi \in L^1(\mathcal{G})),$$

*Talker
Email address: ss.jafari@pnu.ac.ir, ss.jafari@alumni.iut.ac.ir (Seyedeh Somayeh Jafari)

where

$$g \cdot \phi = \frac{1}{\Delta}\tilde{\phi} * g,$$

$\widetilde{\phi}(s) = \phi(s^{-1})$ and Δ denotes the modular function of $\mathcal{G}$. The space $L_0^\infty(\mathcal{G})$ is left introverted in $L^\infty(\mathcal{G})$; i.e., for each $n \in L_0^\infty(\mathcal{G})^*$ and $g \in L_0^\infty(\mathcal{G})$, we have $ng \in L_0^\infty(\mathcal{G})$. This lets us to endow $L_0^\infty(\mathcal{G})^*$ with the *first Arens product* $\diamond$ defined by

$$\langle m \diamond n, g\rangle = \langle m, ng\rangle$$

for all $m, n \in L_0^\infty(\mathcal{G})^*$ and $g \in L_0^\infty(\mathcal{G})$. Then $L_0^\infty(\mathcal{G})^*$ with this product is a Banach algebra which is in relevance to the group algebra $L^1(\mathcal{G})$ of $\mathcal{G}$. For each $\phi \in L^1(\mathcal{G})$, let ϕ also denote the functional in $L_0^\infty(\mathcal{G})^*$ defined by

$$\langle \phi, g\rangle := \int_{\mathcal{G}} \phi(s)\ g(s)\ d(s)$$

for all $g \in L_0^\infty(\mathcal{G})$. Note that this duality defines a linear isometric embedding of $L^1(\mathcal{G})$ into $L_0^\infty(\mathcal{G})^*$, and that $L^1(\mathcal{G}) = L_0^\infty(\mathcal{G})^*$ if and only if $\mathcal{G}$ is discrete. Moreover, observe that $\phi \diamond \psi = \phi * \psi$ for all $\phi, \psi \in L^1(\mathcal{G})$, and that $L^1(\mathcal{G})$ is a closed ideal in $L_0^\infty(\mathcal{G})^*$; see Lau and Pym [4] as the survey article for details.

Let us recall that an element $u \in L_0^\infty(\mathcal{G})^*$ is called a mixed identity if $\phi \diamond u = u \diamond \phi = \phi$ for all $\phi \in L^1(\mathcal{G})$. Denote by $\Lambda_0(\mathcal{G})$ the nonempty set of all mixed identities u with norm one in $L_0^\infty(\mathcal{G})^*$, and recall from Ghahramani, Lau and Losert [1] that $u \in \Lambda_0(\mathcal{G})$ if and only if it is a weak*-cluster point of an approximate identity in $L^1(\mathcal{G})$ bounded by one; or equivalently, it is a right identity of $L_0^\infty(\mathcal{G})^*$ with norm one. Moreover, $L_0^\infty(\mathcal{G})^*$ has an identity if and only if $\mathcal{G}$ is discrete.

Suppose that A is a Banach algebra. By a *derivation* $D : A \to A$, we shall mean a linear map satisfying $D(ab) = D(a)b + aD(b)$ for all $a, b \in A$. Moreover, a derivation D is called *inner* if there is $x \in A$ such that $D = \mathrm{ad}_x$, where the derivation $\mathrm{ad}_x : A \longrightarrow A$ is defined by $\mathrm{ad}_x(a) = ax - xa$ for all $a \in A$.

The study of the range of derivations on Banach algebras was initiated by Singer and Wermer [9], since 1955. Having shown that the range of a continuous derivation on a commutative Banach algebra is contained within the radical of algebra, they conjectured that continuity could be ignored. More than thirty years later, Thomas [10] has proved it. So far, there have been several generalizations of Singer-Wermer conjecture presented to non-commutative Banach algebras. Conditions have been investigated under which every derivation on a Banach algebra maps into the radical. Our aim in this paper is to study derivations on $L_0^\infty(\mathcal{G})^*$. From this point of view, Theorem 2.1 states some characterizations for abelian groups. Also, we show that the Singer-Wermer conjecture is valid only on the case where $\mathcal{G}$ is abelian.

2. The results

Mehdipour and Saeedi [8] have recently studied derivations on $L_0^\infty(\mathcal{G})^*$ for locally compact groups $\mathcal{G}$. They have proved that the zero map is the only weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ when $\mathcal{G}$ is an abelian locally compact group. Our first result is in fact a more general result for all locally compact groups.

Before stating, let us recall that the measure algebra $M(\mathcal{G})$ of $\mathcal{G}$ is the Banach algebra of all complex Radon measures on $\mathcal{G}$ endowed with convolution product $*$ and total variation norm as defined in [2].

Theorem 2.1. *Let $\mathcal{G}$ be a locally compact group. Then every weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ is inner. Also, $\mathcal{G}$ is abelian if and only if either of the following statements holds.*
(a) *Every weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ is zero.*
(b) *Every derivation from $L_0^\infty(\mathcal{G})^*$ into $L^1(\mathcal{G})$ is zero.*

Proof. Suppose that $D : L_0^\infty(\mathcal{G})^* \longrightarrow L_0^\infty(\mathcal{G})^*$ is a weak*-weak*-continuous derivation. On the one hand, since $L^1(\mathcal{G}) * L^1(\mathcal{G}) = L^1(\mathcal{G})$ and

$$D(\phi * \psi) = D(\phi) \diamond \psi + \phi \diamond D(\psi) \in L^1(\mathcal{G}) \quad (\phi, \psi \in L^1(\mathcal{G})),$$

the range of the derivation $D|_{L^1(\mathcal{G})}$ is containing in $L^1(\mathcal{G})$. So, there exists a measure $\mu \in M(\mathcal{G})$ such that $D|_{L^1(\mathcal{G})} = \mathrm{ad}_\mu$ by Corollary 1.2 of [5].
On the other hand, the restriction map defines a continuous epimorphism $\tau : L_0^\infty(\mathcal{G})^* \longrightarrow M(\mathcal{G})$ and so, there exists an element $m \in L_0^\infty(\mathcal{G})^*$ such that $\tau(m) = \mu$. Now, note that

$$\phi \diamond n = \phi * \tau(n), \qquad n \diamond \phi = \tau(n) * \phi,$$

for all $n \in L_0^\infty(\mathcal{G})^*$ and, $\phi \in L^1(\mathcal{G})$. So, we can regard the $M(\mathcal{G})$-bimodule $L^1(\mathcal{G})$ is even an $L_0^\infty(\mathcal{G})^*$-bimodule. Hence,

$$D(\phi) = \phi \diamond m - m \diamond \phi \quad (\phi \in L^1(\mathcal{G})).$$

So, the result will follow when $\mathcal{G}$ is discrete; otherwise,

$$M(\mathcal{G})^* = L^1(\mathcal{G})^* \oplus M_d(\mathcal{G})^* \oplus M_s(\mathcal{G})^*$$

and so, for each $f \in M(\mathcal{G})^*$, we have $f = f_a + f_d + f_s$, where $f_a \in L^1(\mathcal{G})^*$, $f_d \in M_d(\mathcal{G})^*$ and $f_s \in M_s(\mathcal{G})^*$. Now, let $k \in L_0^\infty(\mathcal{G})^*$. Then there exists a net $(\phi_i) \subseteq L^1(\mathcal{G})$ such that $\phi_i \longrightarrow k$ with respect to the weak*-topology in $L^1(\mathcal{G})^{**}$. It follows that

$$\begin{aligned} \lim_i \langle \phi_i,\, f\rangle &= \lim_i \langle \phi_i,\, f_a + f_d + f_s\rangle \\ &= \lim_i \langle \phi_i,\, f_a\rangle \\ &\longrightarrow \langle k,\, f_a\rangle \\ &= \langle k,\, f\rangle. \end{aligned}$$

Therefore, $\phi_i \longrightarrow k$ with respect to the weak*-topology in $M(\mathcal{G})^{**}$. On the other hand, $Z_t(M(\mathcal{G})^{**}) = M(\mathcal{G})$ by [6]. It follows that

$$m \diamond \phi_i = \mu * \phi_i \longrightarrow m \diamond k$$

with respect to weak*-topology in $M(\mathcal{G})^{**}$ and so, $L_0^\infty(\mathcal{G})^*$. Whence,

$$\begin{aligned} k \diamond m - m \diamond k &= \text{weak}^* - \lim_i\, (\phi_i \diamond m - m \diamond \phi_i) \\ &= \text{weak}^* - \lim_i\, D(\phi_i) \\ &= D(k). \end{aligned}$$

That is, $D = \mathrm{ad}_m$. So, every weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ is inner.
As noted earlier, every weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ is zero if $\mathcal{G}$ is abelian; see Corollary 2 of [8]. For converse, suppose that $\mathcal{G}$ is not abelian. Then ad_ϕ is a non-zero weak*-weak*-continuous derivation on $L_0^\infty(\mathcal{G})^*$ for some $\phi \in L^1(\mathcal{G}) \setminus Z(L^1(\mathcal{G}))$.
To complete the proof, let $D : L_0^\infty(\mathcal{G})^* \longrightarrow L^1(\mathcal{G})$ be a derivation. Then $D|_{L^1(\mathcal{G})} = ad_\mu$, for some $\mu \in M(\mathcal{G})$. So, if $\mathcal{G}$ is abelian, then $D|_{L^1(\mathcal{G})} = 0$. Moreover, for each $n \in L_0^\infty(\mathcal{G})^*$, we have

$$D(n) = \lim_\alpha e_\alpha * D(n) = \lim_\alpha [D(e_\alpha \diamond n) - D(e_\alpha) \diamond n] = 0,$$

where (e_α) is a bounded approximate identity of $L^1(\mathcal{G})$. The converse is clear. □

Mehdipour and Saeedi [8] have shown that the Singer-Wermer conjecture is true for $L_0^\infty(\mathcal{G})^*$ if $\mathcal{G}$ is abelian; here, we show that this is an "if and only if" statement.

Proposition 2.2. *Let $\mathcal{G}$ be a locally compact group. Then the image of every derivation on $L_0^\infty(\mathcal{G})^*$ is contained in the radical of $L_0^\infty(\mathcal{G})^*$ if and only if $\mathcal{G}$ is abelian.*

Proof. Suppose that the image of every derivation on $L_0^\infty(\mathcal{G})^*$ is contained in the radical of $L_0^\infty(\mathcal{G})^*$. Then for each $\phi \in L^1(\mathcal{G})$ and $n \in L_0^\infty(\mathcal{G})^*$, we have

$$\mathrm{ad}_\phi(n) \in \mathrm{rad}(L_0^\infty(\mathcal{G})^*) \cap L^1(\mathcal{G}).$$

But the right annihilator of $L_0^\infty(\mathcal{G})^*$ and its radical coincide with $C_0(\mathcal{G})^\perp$; refer to Theorem 2.1 and Corollary 2.2 of [7]. So,

$$\mathrm{ad}_\phi(n) \in \mathrm{ran}(L_0^\infty(\mathcal{G})^*) \cap L^1(\mathcal{G}) = \{0\}.$$

It follows that $L^1(\mathcal{G})$ is commutative, whence $\mathcal{G}$ is abelian. As we have mentioned above the converse is Corollary 1 of [8]. □

The end result of this note follows immediately from Proposition 2.2.

Corollary 2.3. *Let $\mathcal{G}$ be a locally compact group. Then every derivation on $L_0^\infty(\mathcal{G})^*$ is zero if and only if $\mathcal{G}$ is discrete and abelian.*

References

[1] F. Ghahramani, A. T. Lau, and V. Losert. *Isometric isomorphisms between Banach algebras related to locally compact groups*, Trans. Amer. Math. Soc. 321 (1990), 273-283.
[2] E. Hewitt and K. Ross, *Abstract Harmonic Analysis*, Springer-Verlag, New York, 1970.
[3] N. Isik, J. Pym and A. Ulger, *The second dual of the group algebra of a compact group*, J. London Math. Soc. **35**, 135-148 (1987)
[4] A. T. Lau and J. Pym, *Concerning the second dual of the group algebra of a locally compact group*, J. London Math. Soc. 41 (1990), 445-460.
[5] V. Losert, *The derivation problem for group algebras*, Annals Math. 168 (2008), 221-246.
[6] V. Losert, M. Neufang, J. Pachl and J. Steprans, *Proof of the Ghahramani–Lauconjecture*, Adv. Math. 2908 (2016), 709-73.
[7] S. Maghsoudi and R. Nasr-Isfahani, *On the Maximal and Minimal Left Ideals of Certain Banach Algebras on Locally CompactGroups*, Results Math. 62, (2012) 157-165.
[8] M. J. Mehdipour and Z. Saeedi, *Derivations on group algebras of a locally compact abelian group*, Monatsh. Math. 180 (3) (2016), 595-605.
[9] I.M. Singer and J. Wermer, *Derivations on commutative normed algebras*, Math. Ann. 129 (1955), 260-264.
[10] M. Thomas, *The image of a derivation is contained in the radical*, Ann. Math. 128, (1988) 435-460.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–187, pp. 93–96

Amenable locally compact hypergroups

Ali Ghaffari

Department of Mathematics, Semnan University, Semnan, Iran.

Article Info

Keywords:
Amenable hypergroup
Hypergroup algebra
Topological left invariant mean

2020 MSC:
22A20
43A60

Abstract

In this talk, among the other things, we prove a variety of characterizations of amenable hypergroups. We give sufficient conditions and some necessary conditions for H to have a left invariant mean. We know that every topologically left invariant mean on $L^\infty(H)$ is also left invariant mean. Sufficient conditions on a left invariant mean to be a topologically left invariant mean are given.

1. Preliminaries and notations

The theory of hypergroups was initiated by Dunkl, Jewett and Spector and has received a good deal of attention from harmonic analysts. A hypergroup is a locally compact space with a convolution product mapping each pair of points to a probability measure with compact support [9]. Hypergroups are a generalization of locally compact groups wherein the convolution of two points corresponds to the point evaluation measure at their product, for more details see [1], [2] and [3]. There are a lot of results in abstract harmonic analysis on amenability of a locally compact group. A good deal of attention was paid to the study of amenable hypergroups. The study of amenable hypergroups was initiated by Skantharajah [11] and pursued by Wilson [13], see also [7] and [12].
Throughout, H will denote a hypergroup with a left Haar measure λ. It is still unknown if an arbitrary hypergroup admits a left Haar measure, but all the known examples such as commutative hypergroups and central hypergroups do, for more information see [4] and [14]. In [8] Lau introduced and studied a class of Banach algebras which include $L^1(H)$. He called such algebras F-algebras. He extended several important characterizations of amenable locally compact groups to left amenable F-algebras. The F-algebra $L^1(H)$ is amenable if and only if H is amenable [8].
For the sake of convenience for the readers we bring in our notations here. Let H be a locally compact hypergroup with a left Haar measure λ. All of the Lebesgue spaces $L^p(H)$ $1 \leq p \leq \infty$, are taken with respect to left Haar measure λ on H. If H is compact, we normalize λ; so that $\lambda(H) = 1$. Duality between Banach spaces is denoted by $\langle \ \rangle$; thus for $f \in L^\infty(H)$ and $\varphi \in L^1(H)$, we have $\langle f, \varphi\rangle = \int f(x)\varphi(x)dx$. Let $C_b(H)$ and $C_c(H)$ denote the Banach space of all continuous bounded (complex-valued) functions on H, and the subspace of all members of $C_b(H)$ with compact supports, respectively. We denote by $C_0(H)$ the uniform closure of $C_c(H)$ in $C_b(H)$. Thus the dual

*Talker
Email address: aghaffari@semnan.ac.ir (Ali Ghaffari)

of $C_0(H)$ may be identified with the convolution measure algebra $M(H)$ of H. For each $f \in L^\infty(H)$ and $x \in H$, define $L_x f \in L^\infty(H)$ by setting

$$L_x f(y) = f(x * y) = \int f(t) d\delta_x * \delta_y(t)$$

where δ_x denote the unit point-measure at x. Thus we have $\langle f, \delta_x * \varphi\rangle = \langle L_x f, \varphi\rangle$ for $f \in L^\infty(H)$ and $\varphi \in L^1(H)$ [4]. A function $f \in C_b(H)$ is called left uniformly continuous if the map $x \mapsto L_x f$ from H to $C_b(H)$ is continuous. Let $LUC(H)$ denote the subspace of $C_b(H)$ consisting of all left uniformly continuous functions. Skantharajah [11] showed that for hypergroups with left Haar measure, $LUC(H) = L^1(H) * L^\infty(H)$. As far as possible, we follow [4] in our notation and refer to [10] for basic functional analysis and to [5] for basic harmonic analysis.

2. Existence of invariant means

We start by recalling the following definition.

Definition 2.1. Let H be a hypergroup with a left Haar measure λ, and let X be one of the spaces $LUC(H)$ or $L^\infty(H)$. A linear functional M on X is called a mean if:

(i) $\langle M, \overline{f}\rangle = \overline{\langle M, f\rangle}$ for all $f \in X$;

(ii) $f \geq 0$ implies $\langle M, f\rangle \geq 0$ and $\langle M, 1\rangle = 1$.

It is easy to see that a linear functional M on X is a mean if and only if $\langle M, 1\rangle = \|M\| = 1$ and thus the set $\mathcal{M}(X)$ of all means on X is a non-empty weak* compact convex set in X^* [5]. A mean M on X is called a left invariant mean if $\langle M, L_x f\rangle = \langle M, f\rangle$ for all $f \in X$, $x \in H$. The convex set of left invariant means on X is denoted by $LIM(X)$. A hypergroup H is called *amenable* if there is a left invariant mean on $L^\infty(H)$ [6]. The amenability of H can be characterized by the existence of nets of positive, norm one functions in $L^1(H)$ which tend to left invariance in any of several ways [13]. Skantharajah [11] showed that if H is a hypergroup which admits a left Haar measure, then the function spaces of $LUC(H)$, $C_b(H)$, and $L^\infty(H)$ all either admit a left invariant mean (if H is amenable) or all do not. Let

$$P^1(H) = \{\varphi \in L^1(H); \varphi \geq 0 \text{ and } \|\varphi\|_1 = 1\}.$$

A mean M on X is said to be a topological left invariant mean if $\langle M, \varphi * f\rangle = \langle M, f\rangle$ for all $\varphi \in P^1(H)$ and $f \in X$. Every topological left invariant mean on X is a left invariant mean [6]. If $X = LUC(H)$, then every left invariant mean on X is also a topological left invariant mean.

Let H be a non compact amenable locally compact hypergroup, and let $TLIM(L^\infty(H))$ be the set of all topological left invariant means on $L^\infty(H)$. Skantharajah proved in [11] that $|TLIM(L^\infty(H))| = 2^{2^d}$, where d is the smallest cardinality of a cover of H by compact sets.

Theorem 2.2. *A necessary and sufficient condition for the amenability of a locally compact hypergroup H is given by each of the following properties:*

(i) For every $f \in L^\infty(H)$, there exists a mean M_f on $L^\infty(H)$ such that $\langle M_f, \varphi * f\rangle = \langle M_f, \psi * f\rangle$ whenever $\varphi, \psi \in P^1(H)$;

(ii) There exists a net $\{\varphi_\alpha\}$ in $P^1(H)$ such that, for every weakly compact subset S of $P^1(H)$,

$$\lim_\alpha \|\varphi * \varphi_\alpha - \varphi_\alpha\|_1 = 0$$

uniformly for every $\varphi \in S$.

Proposition 2.3. *If H is a locally compact hypergroup, the following properties are equivalent:*

(i) H is amenable,

(ii) For every $\psi \in L^1(H)$;

$$D_\psi := \inf\{\|\psi * \varphi\|_1; \varphi \in P^1(H)\} \leq \Big| \int \psi(xdx \Big|.$$

Let X be a locally convex Hausdorff topological vector space and let C be a compact convex subset of X. The pair $(L^1(H), C)$ is called a semiflow, if;

(1) There exists a map $(\varphi, x) \to \varphi.x$ from $L^1(K) \times X$ into X such that for every $x \in C$, the map $\varphi \to \varphi.x$ is continuous from $L^1(H)$ into X and linear where $L^1(H)$ has the weak topology;

(2) $P^1(H).C \subseteq C$;

(3) For any $\varphi, \psi \in L^1(H)$ and $x \in X$, $\varphi.(\psi.x) = (\varphi * \psi).x$.

Theorem 2.4. *If H is a locally compact hypergroup. The following statements are equivalent:*

(i) H is amenable;

(ii) for any semiflow $(L^1(H), C)$, there is some $x \in C$ such that $\varphi.x = x$ for all $\varphi \in P^1(H)$.

Theorem 2.5. *Let H be a hypergroup with a left Haar measure λ. Then the following statements are equivalent:*

(i) H is amenable;

(ii) for all $f \in LUC(H)$, $n \in \mathbb{N}$ and $\mu_1, ..., \mu_n \in M(H)$,

$$\inf\{\sup\{|\langle f, \mu_i * \varphi\rangle|; 1 \leq i \leq n\}; \varphi \in P^1(H)\} \leq \sup\{|\mu_i(H)|; 1 \leq i \leq n\}\|f\|.$$

3. Relation between left invariant means and topological left invariant means

Recall that $L^1(H)^{**}$, the second conjugate space of $L^1(H)$, is a Banach algebra with the first Arens product [5]. More specifically, let $F, G \in L^1(H)^{**}$, $f \in L^1(H)^*$ and $\varphi, \psi \in L^1(H)$; we define$f\varphi, Ff \in L^1(H)^*$, $GF \in L^1(H)^{**}$ by the equations

$$\langle f\varphi, \psi\rangle = \langle f, \varphi * \psi\rangle, \langle Ff, \varphi\rangle = \langle F, f\varphi\rangle \text{ and } \langle GF, f\rangle = \langle G, Ff\rangle.$$

Theorem 3.1. *Let H be an amenable hypergroup. The following conditions are equivalent:*

(i) $LIM(L^\infty(G)) \cap P^1(H) \neq \emptyset$.

(ii) H is compact.

(iii) $TLIM(L^\infty(H)) \subseteq P^1(H)$.

For each φ in $L^1(H)$ define $\rho_\varphi : L^\infty(H) \to [0, \infty)$ by $\rho_\varphi(f) = \|f\varphi\|$. Then ρ_φ is a seminorm. If $f \neq 0$, then select some $\varphi \in L^1(H)$ with $\langle f, \varphi\rangle \neq 0$. For a bounded approximate identity $\{\varphi_\alpha\}$ in $P^1(H)$ [4], we have

$$0 \neq |\langle f, \varphi\rangle| = \lim_\alpha |\langle f, \varphi * \varphi_\alpha\rangle| = |\langle f\varphi, \varphi_\alpha\rangle| \leq \|f\varphi\|.$$

This implies that $\{\rho_\varphi; \varphi \in L^1(H)\}$ separates the points of $L^1(H)^*$ and makes $L^1(H)^*$ into a locally convex space. The topology defined by these seminorms is denoted by τ.

Proposition 3.2. *Let H be an amenable hypergroup. A left invariant mean M on $L^\infty(H)$ is a topologically left invariant mean if and only if M is τ-continuous.*

Proposition 3.3. *Let H be an amenable hypergroup. If $M \in LIM(L^\infty(H))$ and there exists $\varphi \in P^1(H)$ such that $\langle M, \varphi * f\rangle = \langle M, f\rangle$ whenever $f \in L^\infty(H)$, then also $M \in TLIM(L^\infty(H))$.*

We close this section with the following remark. Let H be a hypergroup with a left Haar measure. Suppose that $\{x\} * \{y\}$ is finite for all $x, y \in H$. Let H have a nondiscrete normal subgroup of finite index. If H_d is amenable, then $\tau \subsetneqq \tau_{\|.\|}$. Indeed, Skantharajah [11] has shown that there is an $M \in LIM(L^\infty(H)) \setminus TLIM(L^\infty(H))$. By Theorem 3.1, M is not τ-contniuous and so $\tau \subsetneqq \tau_{\|.\|}$.

References

[1] M. Amini and C.H. Chu, *Harmonic functions on hypergroups,* Journal of Functional Analysis, **261** (2011), 1835–1864.
[2] M. Amini, H. Nikpey and S.M. Tabatabaie, *Crossed product of C^*-algebras by hypergroups,* Mathematische Nachrichten, **292** (2019), 1897–1910.
[3] A. Bagheri Salec, V. Kumar and S. Mohammad Tabatabaie, *Convolution properties of Orlicz spaces on hypergroups*, Proceeding of the American Mathematical Society, **150** (2022), 1685–1696.
[4] W.R. Bloom and H. Heyer, *Harmonic analysis of probability measures on hypergroups*, vol. 20, de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, 1995.
[5] H.G. Dales, *Banach algebra and automatic continuity*, London Math. Soc. Monogr. Ser. Clarendon Press, 2000.
[6] A. Ghaffari and M.B. Sahabi, *Characterizations of amenable hypergroups*, Wavelets and Linear Algebra, **4** (2017), 1– 9.
[7] A. Ghaffari, T. Haddadi, S. Javadi and M. Sheybani, *On the structure of hypergroups with respect to the induced topology*, Rocky Mountain Journal of Mathematics, **52** (2022), 519–533.
[8] A.T. Lau, *Analysis on a class of Banach algebras with applications to harmonic analysis on locally compact groups,* Fundamenta Mathematicae, **118** (1983), 161–175.
[9] S. Mohammad Tabatabaie, M. Amini and A.A. Amjadi, *Property A for hypergroups*, Semigroup Forum, **104** (2022), 464–479.
[10] W. Rudin, *Functional analysis*, McGraw Hill, New York, 1991.
[11] M. Skantharajah, *Amenable hypergroups,* Illinois Journal of Mathematics, **36** (1992), 15–46.
[12] N. Tahmasebi, *Fixed point properties, invariant means and invariant projections related to hypergroups,* Journal of Mathematical Analysis and Applications, **437** (2016), 526–544.
[13] B. Willson, *Configurations and invariant nets for amenable hypergroups and related algebras,* Transactions of the American Mathematical Society, **366** (2014), 5087–5112.
[14] B. Willson, *A fixed point theorem and the existence of a Haar measure for hypergroups satisfying conditions related to amenability,* Canadian Mathematical Bulletin, **58** (2015), 415–422.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–192, pp. 97–99

Fixed point results for noncyclic φ-contractions in metric spaces equipped with a transitive relation $\mathcal{R}$

Akram Safari-Hafshejani

Department of Pure Mathematics, Payame Noor University (PNU), P. O. Box: 19395-3697, Tehran, Iran.

Article Info

Keywords:
Fixed point
Noncyclic φ-contraction
Metric space
Transitive relation

2020 MSC:
47H10
54H25

Abstract

In this work, we first introduce a new class of maps, called noncyclic φ-contraction in metric spaces equipped with a transitive relation $\mathcal{R}$. Then, we study the existence, uniqueness and convergence of fixed points for such mappings. Also, iterative algorithms are furnished to determine fixed points. Presented results extend and improve some recent results in the literature.

1. Introduction and Preliminaries

Let A and B be two nonempty subsets of a metric space (X, d). If self mapping $T : A \cup B \to A \cup B$ be a noncyclic map , i.e., $T(A) \subseteq A$ and $T(B) \subseteq B$; then $x^* \in A \cup B$ is called a fixed point of T provided that $Tx^* = x^*$. We say that $(x^*, y^*) \in A \times B$ is an optimal pair of fixed points of the noncyclic mapping T provided that

$$Tx^* = x^*, \qquad Ty^* = y^* \qquad \text{and} \qquad d(x^*, y^*) = d(A, B),$$

where $d(A, B) = \inf\{d(a, b) : \ a \in A, \ b \in B\}$.
In 2013, the class of noncyclic contractions was first introduced by Espínola and Gabeleh [3]. For these mappings, the authors presented the following existence theorem.

Theorem 1.1. *Let A and B be nonempty convex subsets of a uniformly convex Banach space X such that A is closed and let $T : A \cup B \to A \cup B$ be a noncyclic contraction map that is, there exists $\lambda \in [0, 1)$ such that*

$$d(Tx, Ty) \leq \lambda d(x, y) + (1 - \lambda) d(A, B), \tag{1}$$

for all $x \in A$ and $y \in B$. For $x_0 \in A$, define $x_{n+1} := Tx_n$ for each $n \geq 0$. Then there exists a unique fixed point $x \in A$ such that $x_n \to x$.

---

*Talker

Email address: asafari@pnu.ac.ir (Akram Safari-Hafshejani)

Since then, the problems of the existence of a fixed pointof noncyclic mappings, have been extensively studied by many authors; see for instance [1, 4–6] and references therein.
In this article, we want to achieve more general results from [3] by using a function $\varphi : [0, +\infty) \to [0, +\infty)$ instead of the constant value λ in relation (1) and equipping the metric space (X, d) with a transitive relation $\mathcal{R}$. For this purpose, we introduce the concept of noncyclic φ-contraction. We study the existence, uniqueness and convergence of fixed points for such mappings in metric spaces equipped with a transitive relation $\mathcal{R}$. Also, iterative algorithms are furnished to determine such fixed points.
Here, we recall the definitions of UC and WUC property.

Definition 1.2. [2, 7] Let A and B be nonempty subsets of the metric space (X, d). Then (A, B) is said to satisfies

(i) property UC, if $\{x_n\}$ and $\{x'_n\}$ are sequences in A and $\{y_n\}$ is a sequence in B such that $\lim_n d(x_n, y_n) = \lim_n d(x'_n, y_n) = d(A, B)$, then $\lim_n d(x_n, x'_n) = 0$;

(ii) property WUC, if for any $\{x_n\} \subseteq A$ such that for every $\epsilon > 0$ there exists $y \in B$ satisfying that $d(x_n, y) \leq d(A, B) + \epsilon$ for $n \geq n_0$, then it is the case that $\{x_n\}$ is Cauchy.

It was announced in [7] that if A and B are nonempty subsets of a uniformly convex Banach space X such that A is convex, then (A, B) has the property UC. Moreover, if A and B are nonempty subsets of a metric space (X, d) such that A is complete and the pair (A, B) has the property UC, then (A, B) has the property WUC (see [2]).

2. Noncyclic φ-contractions

We begin our main conclusions of this section with the following lemma.

Lemma 2.1. *Let $\varphi : [0, +\infty) \to [0, +\infty)$ be a strictly increasing function and I be an identity function defined on $[0, +\infty)$. If $I - \varphi$ is a strictly increasing function, then*

(a) $\varphi(t) > 0$*, for all* $t > 0$*;*

(b) $(I - \varphi)(t) < t$ *for all* $t > 0$*;*

(c) φ *is continuous.*

From now on for the strictly increasing function $\varphi : [0, +\infty) \to [0, +\infty)$ we assume that $I - \varphi$ is also strictly increasing. Also, for given a nonempty subsets A and B of a metric space (X, d) we set

$$d^*(x, y) := d(x, y) - d(A, B), \quad \forall (x, y) \in A \times B.$$

Definition 2.2. Let A and B be nonempty subsets of a metric space (X, d) and ”$\mathcal{R}$” be a transitive relation on A. Let T be a noncyclic mapping on $A \cup B$

(i) we say that T is $\mathcal{R}$-continuous at $x \in A$ if for every sequence $\{x_n\}$ in A with $x_n \to x$ and $x_n \,\mathcal{R}\, x_{n+1}$ for all $n \in \mathbb{N}$, we have $Tx_n \to Tx$;

(ii) we say that T preserves ”$\mathcal{R}$” on A if $Tu \,\mathcal{R}\, Tv$ for every $u, v \in A$ with $u \,\mathcal{R}\, v$;

(iii) we say that ”$\mathcal{R}$” has property $(*)$ on A, if for any sequence $\{x_n\}$ in A with $x_n \to x \in A$ and $x_n \,\mathcal{R}\, x_{n+1}$ for all $n \in \mathbb{N}$, we have $x_n \,\mathcal{R}\, x$ for all $n \in \mathbb{N}$.

Now, with these prerequisites and inspired by the findings of the study about noncyclic Fisher quasi-contractions [6], we introduce the concept of a noncyclic φ-enriched quasi-contraction in the metric space (X, d) equipped with a transitive relation ”$\mathcal{R}$” as follows:

Definition 2.3. Let A and B be nonempty subsets of the metric space (X, d) equipped with a transitive relation ”$\mathcal{R}$”. Let T be a noncyclic mapping on $A \cup B$, then, T is said to be a noncyclic φ-contraction if

$$d^*(Tx, Ty) \leq (I - \varphi)(d^*(x, y)), \tag{2}$$

for all $x \in A$ and $y \in B$ that are comparable with respect to ”$\mathcal{R}$”.

Example 2.4. Let A and B be nonempty subsets of a metric space (X, d) and $T : A \cup B \to A \cup B$ be a noncyclic contraction, Then T is a noncyclic φ-contraction with $\mathcal{R} := X \times X$ and $\varphi(t) := (1 - \lambda)t$ for $t \geq 0$ and $\lambda \in [0, 1)$.

Lemma 2.5. *Let A and B be nonempty subsets of the metric space (X, d) equipped with a transitive relation "$\mathcal{R}$". Let (A, B) has the property WUC and T be a noncyclic φ–contraction mapping on $A \cup B$. Let $x_0 \in A$ and $y_0 \in B$ be such that $x_0 \,\mathcal{R}\, y_0 \,\mathcal{R}\, Tx_0$. Define $x_{n+1} := Tx_n$ for each $n \geq 0$, then $\{x_n\}$ is Cauchy.*

The next theorem is our main result in this section which is an extension of Theorem 2.4 in [6].

Theorem 2.6. *Let A and B be nonempty and complete subsets of the metric space (X, d) equipped with a transitive relation "$\mathcal{R}$". Let T is a noncyclic φ-contraction mapping on $A \cup B$ such that T preserves "$\mathcal{R}$" on $A \cup B$, then*

(i) *if the pair (A, B) satisfies the property WUC, $x_0 \,\mathcal{R}\, y_0 \,\mathcal{R}\, Tx_0$ and $T\,|_A \colon A \to A$ is $\mathcal{R}$-continuous on A, then there exists $x^* \in A$ such that $Tx^* = x^*$;*

(ii) *if the pair (B, A) satisfies the property WUC, $x_0 \,\mathcal{R}\, y_0 \,\mathcal{R}\, Tx_0$ and $T\,|_B \colon B \to B$ is $\mathcal{R}$-continuous on B, then there exists $y^* \in B$ such that $Ty^* = y^*$;*

(iii) *$d(x^*, y^*) = d(A, B)$;*

(iv) *if both pairs (A, B) and (B, A) have property WUC and every pair of elements $x \in A$ and $y \in B$ are comparable with respect to "$\mathcal{R}$", then the optimal pair of fixed points of T obtained in (i) and (ii) is unique.*

From Theorem 2.6, we obtain the following common fixed point result, immediately.

Corollary 2.7. *Let (X, d) be a complete metric space and let $T : X \to X$ and $S : X \to X$ be two continuous mappings satisfying*

$$d(Tx, Sy) \leq (I - \varphi)\,(d(x, y))\,,$$

for all $x, y \in X$. Then S and T have a unique common fixed point in X.

References

[1] A. Anthony Eldred, W. A. Kirk and P. Veeramani,*Proximal normal structure and relatively nonexpansive mappings*, Studia Math., 171(3) (2005), pp. 283-293.

[2] R. Espínola and A. Fernández-León, *On best proximity points in metric and Banach space*, Canad. J. math., 63(3) (2011), pp. 533-550.

[3] R. Espínola and M. Gabeleh, *On the structure of minimal sets of relatively nonexpansive mappings*, Numer. Func. Anal. Optim., 34(8) (2013), pp. 845-860.

[4] A. Fernández-León and M. Gabeleh, *Best proximity pair theorems for noncyclic mappings in Banach and metric spaces*, Fixed Point Theory, 17(1) (2016), pp. 63-84.

[5] A. Safari-Hafshejani, *The existence of best proximity points for generalized cyclic quasi-contractions in metric spaces with the UC and ultrametric properties*, Fixed Point Theory, 23 (2022), pp. 507-518.

[6] A. Safari-Hafshejani, A. Amini-Harandi and M. Fakhar, *Best proximity points and fixed points results for noncyclic and cyclic Fisher quasi-contractions*, Numer. Funct. Anal. Optim., 40(5) (2019), pp. 603-619.

[7] T. Suzuki, M. Kikawa and C. Vetro, *The existence of best proximity points in metric spaces with the property UC*, Nonlinear Anal., 71 (2009), pp. 2918-2926.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–205, pp. 100–105

Banach fixed point theorem in exponential metric spaces

Zynab Izadi[a,*], Vahid Parvaneh[b]

[a] *Department of Mathematics, Faculty of Sciences, Payame Noor University, P. O. Box 19395-3697, Tehran, Iran*
[b] *Department of Mathematics, Gilan-E-Gharb Branch, Islamic Azad University, Gilan-E-Gharb, Iran*

Article Info

Keywords:
fixed point
α-admissible mapping
exponential metric space.

2020 MSC:
47H10
54H25.

Abstract

One of the most fundamental theorems in fixed point theory is Banach's fixed point theorem, which has been generalized by many researchers in different metric type spaces. In this article, a generalization of the Banach contraction principle is presented, which will be of interest to researchers in the field of fixed point theory and numerical analysis and other branches related to the theory of fixed point theory. To clarify the benefits of this theorem, we will present some corollaries which doubles the interest of the subject.

1. Introduction and Preliminaries

Banach contractive principle or Banach fixed point theorem is the most celebrated result in fixed point theory which illustrates that in a complete metric space, each contractive mapping has a unique fixed point. There is a great number of generalizations of this principle by using different forms of contractive conditions in various spaces. Recently, Wardowski [6] introduced a new contraction called F-contraction and proved a fixed point result as a generalization of the Banach contraction principle. Abbas *et al.* [4], as well as Wardowski and Van Dung [11] generalized the concept of F-contraction and proved certain fixed and common fixed point results.
One of the interesting results which also generalizes the Banach contraction principle was given by Samet *et al.* [9] by defining α-ψ-contractive and α-admissible mappings.

Definition 1.1. [1] Let T be a self-mapping on a set X and let $\alpha : X \times X \to [0, \infty)$ be a function. We say that T is an α-admissible mapping if

$$x, y \in X, \quad \alpha(x, y) \geq 1 \quad \Longrightarrow \quad \alpha(Tx, Ty) \geq 1.$$

Hussain *et al.* [7], as well as Karapınar *et al.* [10] extended this result as follows (see also [4, 8]).

Definition 1.2. [11] Let (X, d) be a metric space. Let $\alpha : X \times X \to [0, \infty)$ and $T : X \to X$ be mappings. We say that T is an α-continuous mapping on (X, d), if, for given $x \in X$ and sequence $\{x_n\}$,

$$x_n \to x \text{ as } n \to \infty \text{ and } \alpha(x_n, x_{n+1}) \geq 1 \text{ for all } n \in \mathbb{N} \implies Tx_n \to Tx.$$

*Talker
Email addresses: zynab- izadi@pnu.ac.ir (Zynab Izadi), zam.dalahoo@gmail.com (Vahid Parvaneh)

Definition 1.3. [4] Let $f : X \to X$ and $\alpha : X \times X \to [0, +\infty)$. We say that f is a triangular α-admissible mapping if

(T1) $\alpha(x, y) \geq 1 \quad \text{implies} \quad \alpha(fx, fy) \geq 1, \quad x, y \in X;$

(T2) $\begin{cases} \alpha(x, z) \geq 1 \\ \alpha(z, y) \geq 1 \end{cases} \quad \text{implies} \quad \alpha(x, y) \geq 1, \quad x, y, z \in X.$

Lemma 1.4. [4] *Let f be a triangular α-admissible mapping. Assume that there exists $x_0 \in X$ such that $\alpha(x_0, fx_0) \geq 1$. Define sequence $\{x_n\}$ by $x_n = f^n x_0$. Then*

$$\alpha(x_m, x_n) \geq 1 \text{ for all } m, n \in \mathbb{N} \text{ with } m < n.$$

Bakhtin in [1] and Czerwik in [2, 3] introduced the concept of a b-metric space. Since then, several papers dealt with fixed point theory for single-valued and multi-valued operators in b-metric spaces (see, *e.g.*, [5, 7]).
Let $s \geq 1$ be a fixed real number. We will consider the following class of functions.
Δ will denote the set of all functions $\Gamma : [0, \infty) \to [1, \infty)$ such that

(Δ_1) Γ is continuous and strictly increasing;

(Δ_2) for each sequence $\{t_n\} \subseteq R^+$, $\lim_{n\to\infty} t_n = 0$ if and only if $\lim_{n\to\infty} F(t_n) = 1$.

Example 1.5. If $\Gamma_1(t) = \cosh t$ and $\Gamma_2(t) = e^t$, then $\Gamma_1, \Gamma_2 \in \Delta$.

Definition 1.6. [13] Let X be a (nonempty) set and $s \geq 1$ be a given real number. A function $d : X \times X \to \mathbb{R}^+$ is an exponential metric if, for all $x, y, z \in X$, the following conditions are satisfied:

(b_1) $d(x, y) = 0$ iff $x = y$,

(b_2) $d(x, y) = d(y, x)$,

(b_3) $\Gamma[d(x, z)] \leq [\Gamma[d(x, y)]\Gamma[d(y, z)]]^s$.

In this case, the pair (X, d) is called an exponential metric space.

Definition 1.7. [8] Let (X, d) be an exponential metric space.
(a) A sequence $\{x_n\}$ in X is called convergent if there exists $x \in X$ such that $d(x_n, x) \to 0$, as $n \to \infty$. In this case, we write $\lim_{n\to\infty} x_n = x$.
(b) $\{x_n\}$ in X is said to be Cauchy if $d(x_n, x_m) \to 0$, as $n, m \to \infty$.
(c) The exponential metric space (X, d) is complete if every Cauchy sequence in X be convergent.

Note that an exponential metric need not be to a continuous function. It is also an interesting generalization of the concept of b-metric.

Lemma 1.8. *Let (X, d) be an exponential metric space with parameter $s \geq 1$, and suppose that $\{x_n\}$ and $\{y_n\}$ are convergent to x and y, respectively. Then we have:*

$$\sqrt[s^2]{\Gamma[d(x, y)]} \leq \Gamma[\liminf_{n\to\infty} d(x_n, y_n)] \leq \Gamma[\limsup_{n\to\infty} d(x_n, y_n)] \leq \Gamma[d(x, y)]^{s^2}.$$

In particular, if $x = y$, then we have $\lim_{n\to\infty} d(x_n, y_n) = 0$. Moreover, for each $z \in X$, we have,

$$\sqrt[s]{\Gamma[d(x, y)]} \leq \Gamma[\liminf_{n\to\infty} d(x_n, y_n)] \leq \Gamma[\limsup_{n\to\infty} d(x_n, y_n)] \leq \Gamma[d(x, y)]^s.$$

Proof. The reader can follow the following inequalities:

$$\begin{aligned}\Gamma[d(x_n, y_n)] &\leq [\Gamma[d(x_n, x)]\Gamma[d(x, y_n)]]^s \\ &\leq \left[\Gamma[d(x_n, x)]\left[\Gamma[d(x, y)]\Gamma[d(y, y_n)]\right]^s\right]^s,\end{aligned}$$

and

$$\begin{aligned}\Gamma[d(x, y)] &\leq [\Gamma[d(x, x_n)]\Gamma[d(x_n, y)]]^s \\ &\leq \left[\Gamma[d(x, x_n)]\left[\Gamma[d(x_n, y_n)]\Gamma[d(y_n, y)]\right]^s\right]^s.\end{aligned}$$

Also,

$$\Gamma[d(x_n, y)] \leq [\Gamma[d(x_n, x)]\Gamma[d(x, y)]]^s,$$

and

$$\Gamma[d(x, y)] \leq [\Gamma[d(x, x_n)]\Gamma[d(x_n, y)]]^s.$$

□

In this paper, we introduce the concept of α-Γ-contraction and obtain some fixed point results in exponential metric spaces.

2. Fixed point results for α-admissible-Γ-contractions

Definition 2.1. Let (X, d) be an exponential metric space with parameter $s \geq 1$ and let T be a self-mapping on X. Also, suppose that $\alpha : X \times X \to [0, \infty)$ is a function. We say that T is an α-Γ-contraction if for all $x, y \in X$ with $1 \leq \alpha(x, y)$ and $d(Tx, Ty) > 0$ we have

$$[\Gamma[d(Tx, Ty)]]^s \leq K \cdot M_s(x, y), \tag{1}$$

where $\Gamma \in \Delta$, $0 < K < 1$ and

$$M_s(x, y) = \max\left\{d(x, y), d(x, Tx), d(y, Ty), \Gamma^{-1}[\sqrt[2s]{\Gamma[d(x, Ty)]\Gamma[d(y, Tx)]}]\right\}. \tag{2}$$

Now we state and prove our main result of this section.

Theorem 2.2. *Let (X, d) be a complete exponential metric space with parameter $s \geq 1$. Let $T : X \to X$ be a self-mapping satisfying the following assertions:*

(i) T is a triangular α-admissible mapping;

(ii) T is an α-Γ-contraction;

(iii) there exists $x_0 \in X$ such that $\alpha(x_0, Tx_0) \geq 1$;

(iv) T is α-continuous.

Then T has a fixed point. Moreover, T has a unique fixed point if $\alpha(x, y) \geq 1$ for all $x, y \in Fix(T)$.

Proof. Let $x_0 \in X$ be such that $\alpha(x_0, Tx_0) \geq 1$. We define the sequence $\{x_n\}$ by $x_n = T^n x_0 = Tx_{n-1}$. Now since, T is an α-admissible mapping then $\alpha(x_0, x_1) = \alpha(x_0, Tx_0) \geq 1$. By continuing this process we have

$$\alpha(x_{n-1}, x_n) \geq 1$$

for all $n \in \mathbb{N}$. Also, let there exists $n_0 \in \mathbb{N}$ such that $x_{n_0} = x_{n_0+1}$. Then x_{n_0} is a fixed point of T and we have nothing to prove. Hence, we assume that $x_n \neq x_{n+1}$, i.e., $d(Tx_{n-1}, Tx_n) > 0$ for all $n \in \mathbb{N} \cup \{0\}$. Since, T is an α-Γ-contraction, so we derive

$$[\Gamma[d(Tx_{n-1}, Tx_n)]]^s = [\Gamma[d(x_n, x_{n+1})]]^s \leq K \cdot M_s(x_{n-1}, x_n),$$

where

$$\begin{aligned}
&M_s(x_{n-1}, x_n)\\
&\quad = \max\left\{d(x_{n-1}, x_n), d(x_{n-1}, Tx_{n-1}), d(x_n, Tx_n), \Gamma^{-1}[\sqrt[2s]{\Gamma[d(x_{n-1}, Tx_n)]\Gamma[d(x_n, Tx_{n-1})]}]\right\}\\
&\quad = \max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1}), \Gamma^{-1}[\sqrt[2s]{\Gamma[d(x_{n-1}, Tx_n)]}]\right\}\\
&\quad \leq \max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1}), \Gamma^{-1}[\sqrt{\Gamma[d(x_{n-1}, x_n)]\Gamma[d(x_n, x_{n+1})]}]\right\}\\
&\quad = \max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1})\right\},
\end{aligned}$$

which implies

$$[\Gamma(d(x_n, x_{n+1}))]^s \leq K \cdot \max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1})\right\}.$$

If

$$\max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1})\right\} = d(x_n, x_{n+1}),$$

then we have

$$[\Gamma(d(x_n, x_{n+1}))]^s \leq K \cdot d(x_n, x_{n+1}),$$

a contradiction. Therefore, $\max\left\{d(x_{n-1}, x_n), d(x_n, x_{n+1})\right\} = d(x_{n-1}, x_n)$, and we have

$$d(x_n, x_{n+1}) \leq K \cdot d(x_{n-1}, x_n).$$

Consequently, we deduce that

$$d(x_n, x_{n+1}) \leq K^n \cdot d(x_0, x_1).$$

By taking the limit as $n \to \infty$ in (**??**), we have

$$\lim_{n\to\infty} d(x_n, x_{n+1}) = 0. \tag{3}$$

Next, we show that $\{x_n\}$ is a Cauchy sequence in X.
Suppose to the contrary, that is, $\{x_n\}$ is not a Cauchy sequence. Then there exists $\varepsilon > 0$ for which we can find two subsequences $\{x_{m_i}\}$ and $\{x_{n_i}\}$ of $\{x_n\}$ such that n_i is the smallest index for which

$$n_i > m_i > i \text{ and } d(x_{m_i}, x_{n_i}) \geq \varepsilon. \tag{4}$$

This means that

$$d(x_{m_i}, x_{n_i-1}) < \varepsilon. \tag{5}$$

From (4) and using the triangular inequality, we get

$$\Gamma(\varepsilon) \leq \Gamma(d(x_{m_i}, x_{n_i})) \leq \Gamma[d(x_{m_i}, x_{m_i+1})]^s \Gamma[d(x_{m_i+1}, x_{n_i})]^s.$$

Taking the upper limit as $i \to \infty$, we get

$$\Gamma^{-1}(\sqrt[s]{\Gamma(\varepsilon)}) \leq \limsup_{i\to\infty} d(x_{m_i+1}, x_{n_i}). \tag{6}$$

Also, from (5),

$$\limsup_{i\to\infty} d(x_{m_i}, x_{n_i-1}) \leq \varepsilon. \tag{7}$$

On the other hand, we have

$$\Gamma[d(x_{m_i}, x_{n_i})] \leq \Gamma[d(x_{m_i}, x_{n_i-1})]^s \Gamma[d(x_{n_i-1}, x_{n_i})]^s.$$

Using (3) and (5) and taking the upper limit as $i \to \infty$, we get

$$\Gamma[\limsup_{i\to\infty} d(x_{m_i}, x_{n_i})] \leq \Gamma(\varepsilon)^s. \tag{8}$$

Again, using the triangular inequality, we have

$$\Gamma[d(x_{m_i+1}, x_{n_i-1})] \leq \Gamma[d(x_{m_i+1}, x_{m_i})]^s \Gamma[d(x_{m_i}, x_{n_i-1})]^s.$$

Taking the upper limit as $i \to \infty$ in the above inequality and using (3) and (5), we get

$$\Gamma[\limsup_{i\to\infty} d(x_{m_i+1}, x_{n_i-1})] \leq \Gamma(\varepsilon)^s. \tag{9}$$

Since T is a triangular α-admissible mapping, we have that $\alpha(x_{m_i}, x_{n_i}) \geq 1$, and so we can apply (1) to conclude that

$$\begin{aligned}[\Gamma\big(d(x_{m_i+1}, x_{n_i})\big)]^s &= [\Gamma\big(d(Tx_{m_i}, Tx_{n_i-1})\big)]^s \\ &\leq K.M_s(x_{m_i}, x_{n_i-1}),\end{aligned} \tag{10}$$

where,

$$\begin{aligned}&M_s(x_{m_i}, x_{n_i-1}) \\ &= \max\left\{d(x_{m_i}, x_{n_i-1}), d(x_{m_i}, Tx_{m_i}), d(x_{n_i-1}, Tx_{n_i-1}), \Gamma^{-1}[\sqrt[2s]{\Gamma[d(x_{m_i}, Tx_{n_i-1})]\Gamma[d(Tx_{m_i}, x_{n_i-1})]}]\right\} \\ &= \max\left\{d(x_{m_i}, x_{n_i-1}), d(x_{m_i}, x_{m_i+1}), d(x_{n_i-1}, x_{n_i}), \Gamma^{-1}[\sqrt[2s]{\Gamma[d(x_{m_i}, x_{n_i})]\Gamma[d(x_{m_i+1}, x_{n_i-1})]}]\right\}.\end{aligned} \tag{11}$$

Now, taking the upper limit as $i \to \infty$ in (10) and using (6), (7), (9) and (11), we have

$$\begin{aligned}\Gamma(\varepsilon) = \left[\Gamma(\Gamma^{-1}(\sqrt[s]{\Gamma(\varepsilon)}))\right]^s &\leq [\Gamma[\limsup_{i\to\infty} d(x_{m_i+1}, x_{n_i})]]^s \\ &\leq K.\limsup_{i\to\infty} M_s(x_{m_i}, x_{n_i-1}) \leq K.\max\{\varepsilon, \varepsilon\} \\ &= K.\varepsilon,\end{aligned}$$

which is impossible.

Thus, we have proved that $\{x_n\}$ is a Cauchy sequence in the exponential metric space (X, d). Since (X, d) is complete, the sequence $\{x_n\}$ converges to some $z \in X$, that is, $\lim_{n\to\infty} d(x_n, z) = 0$.

Suppose that $z \neq Tz$. Then, from Lemma **??**, as T is α-continuous,

$$\sqrt[s]{\Gamma[d(z, Tz)]} \leq \Gamma(\liminf_{n\to\infty} d(x_n, Tx_n)) \leq \Gamma(\limsup_{n\to\infty} d(x_n, Tx_n)) = \limsup_{n\to\infty} d(x_n, x_{n+1}) = 0.$$

Hence, we have $d(Tz, z) = 0$ and so $Tz = z$. Thus, z is a fixed point of T.

Let $x, y \in Fix(T)$ where $x \neq y$. Then from

$$d(Tx, Ty) \leq [\Gamma\big(d(Tx, Ty)\big)]^s \leq K.M_s(x, y),$$

where

$$\begin{aligned} M_s(x,y) &= \max\left\{d(x,y), d(x,fx), d(y,fy), \Gamma^{-1}\sqrt[2s]{\Gamma[d(x,fy)]\Gamma[d(fx,y)]}\right\} \\ &= d(x,y). \end{aligned}$$

So, we get

$$d(x, y) \leq K.d(x, y),$$

so $d(x, y) = 0$, a contradiction. Hence, $x = y$.
Therefore, T has a unique fixed point. □

Taking $\Gamma(t) = e^t$ in the above theorem, we obtain the Banach contraction principle in the setup of b-metric spaces. Taking $s = 1$, we obtain the Banach contraction principle in a metric space.

References

[1] IA Bakhtin. The contraction principle in quasimetric spaces, func. an., unianowsk, gos. ped. ins. 30 (1989), 26–37.

[2] Stefan Czerwik. Contraction mappings in b-metric spaces. *Acta mathematica et informatica universitatis ostraviensis*, 1(1):5–11, 1993.

[3] Stefan Czerwik. Nonlinear set-valued contraction mappings in b-metric spaces. *Atti Sem. Mat. Fis. Univ. Modena*, 46:263–276, 1998.

[4] Nawab Hussain, Vahid Parvaneh, Jamal Rezaei Roshan, and Zoran Kadelburg. Fixed points of cyclic weakly (ψ, φ, l, a, b)-contractive mappings in ordered b-metric spaces with applications. *Fixed Point Theory and Applications*, 2013:1–18, 2013.

[5] Mirko Jovanović, Zoran Kadelburg, and Stojan Radenović. Common fixed point results in metric-type spaces. *Fixed point theory and applications*, 2010:1–15, 2010.

[6] E Karapinar, P Kumam, and P Salimi. On–meir-keeler contractive mappings. *Fixed Point Theory Appl*, 94:2013, 2013.

[7] MA Khamsi and Nawab Hussain. Kkm mappings in metric type spaces. *Nonlinear Analysis: Theory, Methods & Applications*, 73(9): 3123–3129, 2010.

[8] Peyman Salimi, Abdul Latif, and Nawab Hussain. Modified α-ψ-contractive mappings with applications. *Fixed Point Theory and Applications*, 2013:1–19, 2013.

[9] Bessem Samet, Calogero Vetro, and Pasquale Vetro. Fixed point theorems for α–ψ-contractive type mappings. *Nonlinear analysis: theory, methods & applications*, 75(4):2154–2165, 2012.

[10] Dariusz Wardowski. Fixed points of a new type of contractive mappings in complete metric spaces. *Fixed point theory and applications*, 2012:1–6, 2012.

[11] Dariusz Wardowski and N Van Dung. Fixed points of f-weak contractions on complete metric spaces. *Demonstratio Mathematica*, 47(1): 146–155, 2014.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–208, pp. 106–113

On properly efficient solution of multiobjective convex generalized semi-infinite optimization problems

Hamed Soroush[a,*]

[a] *Department of Mathematics, Payame Noor University (PNU), Tehran, Iran*

Article Info

Keywords:
optimality conditions
generalized semi-infinite programming
Slater constraint qualification3
Basic constraint qualification

2020 MSC:
90C34
90C40
49J52

Abstract

In this paper, for a non-differentiable multiobjective generalized semi-infinite programming problem, where the objective and constraint functions are convex, the lower-level Basic and Slater constraint qualifications are given. Then, some necessary optimality conditions are derived at properly efficient solutions of the considered problem, under these constraint qualifications and using convex subdifferential.

1. Introduction

Generalized semi-infinite programming problem (GSIP, in brief) is a field of active research, because of not only its surprising structural aspects, but also its numerous applications. Stein in [7] lists a number of problems from engineering and economics which give rise to GSIP models, including reverse Chebyshev approximation, minimax problems, robust optimization, design centering, and disjunctive programming. These problems have the form

$$\text{GSIP}: \qquad \inf \varphi(x) \quad \text{s.t.} \;\; \psi(x,y) \geq 0, \;\; \forall y \in A(x),$$

where $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a given function and the upper-level index set $A(x)$ is defined as

$$A(x) := \{y \in \mathbb{R}^m \mid \phi_i(x,y) \leq 0, \quad i \in P\},$$

in which the lower-level index set P is finite, and the functions ψ and ϕ_i as $i \in P$ are real-valued on $\mathbb{R}^{n+m}$.
In almost all existing literature on GSIP theory, the continuously differentiable (smoothness) assumption on the emerging functions φ, ψ, and ϕ_i as $i \in P$, is principle and restrictive. In order to establish optimality conditions for smooth GSIP, several kinds of lower-level constraint qualifications are studied. Extensive references to optimality conditions

*Talker
Email address: h_soroush2011@pnu.ac.ir (Hamed Soroush)

and constraint qualifications for smooth GSIPs, and their historical notes, can be found in [7, 9] and their references. Kanzi and his coauthors extend these optimality conditions to nonsmooth GSIPs with DC (difference of convex) functions [4] and locally Lipschitz functions [5].
Recently, GSIP is considered when $\varphi := (\varphi_1, \ldots, \varphi_q) : \mathbb{R}^n \to \mathbb{R}^q$ is a vector-valued function, and this problem is named multiobjective generalized semi-infinite programming problem (MGSIP in short). In the case when all appearing functions of MGSIP are continuously differentiable, some necessary first-order optimality conditions have been given in [8]. More recently, some optimality conditions for MGSIPs with nondifferentiable convex functions are presented by Soroush [6] and by Edalat *et. al.* [2]. In these articles, the results are stated on upper-level Mangasarian-Fromovitz constraint qualification.
Since in all of the references cited above (except for [4]), the uniform boundedness on the set-valued map $x \mapsto A(x)$ is a standard assumption, and this assumption is very restrictive, one of the goal of this article is to free the problem from this condition. Another goal of this article is to introduce two lower-level constraint qualification in Slater and Basic types for convex MGSIPs. In order to these aims, in this paper we consider the following MGSIP:

$$\text{GMGSIP}: \qquad \inf F(x) := \big(f_1(x), \ldots, f_q(x)\big), \quad \text{s.t.} \quad x \in S,$$

with the feasible set

$$S := \{x \in \mathbb{R}^n \mid g(x, y) \geq 0, \;\; y \in Y(x)\},$$

and the index set,

$$Y(x) := \{y \in \mathbb{R}^m \mid h_i(x, y) \leq 0, \; i \in P\},$$

where $P := \{1, \ldots, p\}$, and all the appearing functions f, g and h_i (as $i \in P$) are real-valued convex on $\mathbb{R}^n$ and $\mathbb{R}^{n+m}$, respectively.
The organization of the paper is as follows. In Section 2, basic notations and results of convex analysis are reviewed. Section 3 contains the main results of the paper, which includes the introduction of constraint qualifications and proof of main theorems.

2. Notations and Preliminaries

In this section we describe our notation and present preliminary results. Throughout the paper, the inner product of two vectors u and v in the n-dimensional al space $\mathbb{R}^n$ will be denoted by $\langle u, v\rangle$, and the null vector in $\mathbb{R}^n$ will be denoted by 0_n. The closure of $A \subseteq \mathbb{R}^n$ is denoted by $\overline{A}$.
• For a set $A \subseteq \mathbb{R}^n$ we shall denote

1. the convex hull of A by $conv(A)$, that is defined by

$$conv(A) := \Big\{ \sum_{j=1}^{k} \lambda_j a_j \mid a_j \in A, \; k \in \mathbb{N}, \; \lambda_j \in [0, 1], \; \sum_{j=1}^{k} \lambda_j = 1 \Big\}.$$

2. the convex cone of A by $cone(A)$, that is defined by

$$cone(A) := \bigcup_{\alpha \geq 0} \alpha conv(A).$$

3. the negative polar cone of A by $A^{\ominus}$, that is defined by

$$A^{\ominus} := \{x \in \mathbb{R}^n \mid \langle x, a\rangle \leq 0, \; \forall\, a \in A\}.$$

4. the strictly negative polar of A by A^s, that is defined by

$$A^s := \{x \in \mathbb{R}^n \mid \langle x, a\rangle < 0, \; \forall\, a \in A\}.$$

It is easy to check $A^s = \big(conv(A)\big)^s$ for all $A \subseteq \mathbb{R}^n$. Also, if $A^s \neq \emptyset$, we can see $A^{\ominus} = \overline{A^s}$.

Theorem 2.1. *([3]) Suppose that $A \subseteq \mathbb{R}^n$ is given. Then,*

- *$conv(A)$ is closed provided A is closed.*
- *$cone(A)$ is closed provided A is compact and $0_n \notin conv(A)$.*
- *$\left(A^{\ominus}\right)^{\ominus} = \overline{cone(A)}$.*

Theorem 2.2. *([3]) Let $C_1, \dots, C_k$ be nonempty convex sets in $\mathbb{R}^n$. Then, every non-zero vector of $conv\left(\bigcup_{i=1}^{k} C_i\right)$ can be expressed as a convex combination of vectors, each belonging to a different C_i.*

• For a convex set $C \subseteq \mathbb{R}^n$, we shall denote

1. the tangent cone of C at $\hat{c} \in \overline{C}$ by $T_C(\hat{c})$, that is defined by
$$T_C(\hat{c}) := \overline{cone(C - \{\hat{x}\})}.$$
2. the normal cone of C at $\hat{c} \in \overline{C}$ by $N_C(\hat{c})$, that is defined by
$$N_C(\hat{c}) := \left\{x \in \mathbb{R}^n \mid \left\langle x, c - \hat{c}\right\rangle \leq 0 \quad \text{for all } c \in C\right\}.$$

We can see the negative polar cone of $N_C(\hat{c})$ is $T_C(\hat{c})$, and the negative polar cone of $T_C(\hat{c})$ is $N_C(\hat{c})$.
• Let $\varphi : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz function.

1. The generalized Clarke directional derivative of φ at $\hat{x} \in \mathbb{R}^n$ in the direction $d \in \mathbb{R}^n$ is defined by
$$\varphi^0(\hat{x}; d) := \limsup_{y \to \hat{x},\ t \downarrow 0} \frac{\varphi(y + td) - \varphi(y)}{t}.$$
2. The Clarke subdifferential of φ at $\hat{x}$ is defined by
$$\partial^c \varphi(\hat{x}) := \left\{\xi \in \mathbb{R}^n \mid \varphi^0(\hat{x}; d) \geq \langle \xi, d\rangle, \quad \forall d \in \mathbb{R}^n\right\}.$$

It is known ([1]) that the Clarke subdifferential of a locally Lipschitz function at each point of its domain is always a non-empty convex compact set. Moreover, $\partial^c \varphi(\hat{x}) = \{\nabla \varphi(\hat{x})\}$ when the function $\varphi(\cdot)$ is continuously differentiable at $\hat{x}$ (as always, $\nabla \varphi(\hat{x})$ denotes the gradient of φ at $\hat{x}$). Also, if $\varphi : \mathbb{R}^n \to \mathbb{R}$ is a convex function, we have
$$\partial^c \varphi(\hat{x}) = \partial \varphi(\hat{x}) := \left\{\xi \in \mathbb{R}^n \mid \varphi(x) - \varphi(\hat{x}) \geq \langle \xi, x - \hat{x}\rangle, \quad \forall x \in \mathbb{R}^n\right\},$$
and $\varphi^0(\hat{x}; d) = \varphi'(\hat{x}; d)$, where the standard directional derivative of φ at $\hat{x}$ in the direction d is denoted by $\varphi'(\hat{x}; d)$. We will use the following important relations for two locally Lipschitz functions $\varphi_1, \varphi_2 : \mathbb{R}^n \to \mathbb{R}$ hereafter:
$$\varphi_i^0(\hat{x}; d) = \max\{\langle \xi, d\rangle \mid \xi \in \partial^c \varphi(\hat{x})\}, \tag{1}$$
$$\partial^c\big(\max\{\varphi_1, \varphi_2\}\big)(\hat{x}) \subseteq conv\big(\partial^c \varphi_1(\hat{x}) \cup \partial^c \varphi_2(\hat{x})\big), \tag{2}$$
$$\partial^c\big(\alpha_1 \varphi_1 + \alpha_2 \varphi_2\big)(\hat{x}) \subseteq \alpha_1 \partial^c \varphi_1(\hat{x}) + \alpha_2 \partial^c \varphi_2(\hat{x}), \qquad \forall \alpha_1, \alpha_2 \in \mathbb{R}. \tag{3}$$
The last inclusion increases to equality for convex functions φ_1 and φ_2 and non-negative scalars α_1 and α_2.

Theorem 2.3. *([1]) If the locally Lipschitz function $\varphi : \mathbb{R}^n \to \mathbb{R}$ attains its minimum on convex set $C \subseteq \mathbb{R}^n$ at $\hat{c} \in C$, we have*
$$0_n \in \partial^c \varphi(\hat{c}) + N_C(\hat{c}).$$

For a locally Lipschitz function $\phi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and a point $(\hat{x}, \hat{y}) \in \mathbb{R}^n \times \mathbb{R}^m$, let $\partial_x^c \phi(\hat{x}, \hat{y})$ and $\partial_y^c \phi(\hat{x}, \hat{y})$ denote the partial Clarke subdifferential of $\phi(\cdot, \cdot)$ at $(\hat{x}, \hat{y})$, which are defined as $\partial^c \phi(\cdot, \hat{y})(\hat{x})$ and $\partial^c \phi(\hat{x}, \cdot)(\hat{y})$, respectively.

3. Main Results

We start this section by recalling the following definition.

Definition 3.1. A feasible point $x_0 \in S$ is called a properly efficient solution to MGSIP when there exist some positive scalars $\gamma_1, \ldots, \gamma_q$ such that

$$\sum_{j=1}^{q} \gamma_j f_j(x_0) \leq \sum_{j=1}^{q} \gamma_j f_j(x), \qquad \forall x \in S.$$

The lower level problem of MGSIP at $\hat{x} \in S$ (which depends on the parameter $\hat{x}$) is defined as

$$\inf g(\hat{x}, y) \quad \text{s.t.} \quad y \in Y(x). \tag{4}$$

We associate with MGSIP the Lagrangian

$$\$: \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}^p \longrightarrow \mathbb{R}$$
$$\$(x, y, \alpha, \beta) = \alpha g(x, y) + \sum_{i=1}^{p} \beta_i h_i(x, y).$$

Given $\hat{x} \in S$, set of active constraints at $\hat{x}$ is denoted by $Y_0(\hat{x})$, i.e.,

$$Y_0(\hat{x}) := \big\{ y \in Y(\hat{x}) \mid g(\hat{x}, y) = 0 \big\}.$$

Note that $Y_0(\hat{x})$ is just the set of minimizers of the lower-level problem (4). For each $\hat{y} \in Y_0(\hat{x})$, we define the set (maybe empty) of Karush-Kahn-Tucker (KKT) multiplier of the lower-level problem (4) at $\hat{y}$ as

$$K(\hat{x}, \hat{y}) := \big\{ \beta \in \mathbb{R}^p_+ \mid 0_m \in \partial_y \$\hat{x}, \hat{y}, 1, \beta), \;\; \beta_i h_i(\hat{x}, \hat{y}) = 0, \;\; \forall\, i \in P \big\}.$$

Definition 3.2. We say that

(i): the Basic constraint qualification (BCQ) satisfies at $(x_0, y_0) \in \Lambda$ if

$$N_\Lambda(x_0, y_0) \subseteq cone\Big(\bigcup_{i=1}^{p} \partial h_i(x_0, y_0) \Big),$$

where

$$\Lambda := \big\{ (x, y) \in \mathbb{R}^n \times \mathbb{R}^m \mid h_i(x, y) \leq 0 \;\; \forall i \in P \big\}.$$

(ii): the Slater constraint qualification (SCQ) satisfies if there exists a $(x^*, y^*) \in \Lambda$ such that $h_i(x^*, y^*) < 0$ for all $i \in P$.

Now, we can formulate our first main result.

Theorem 3.3. *Suppose that $\hat{x}$ is a properly efficient solution for convex MGSIP.*

(i): *If $Y(x_0) = \emptyset$, there exist some positive coefficients $\lambda_1, \ldots, \lambda_q > 0$ such that*

$$0_n \in \sum_{j=1}^{q} \lambda_j \partial f_j(\hat{x}).$$

(ii): *If $Y(x_0) \neq \emptyset$ and for all $\hat{y} \in Y_0(\hat{x})$ the BCQ holds at $(\hat{x}, \hat{y})$, there exist some non-negative coefficients $\lambda_1, \ldots, \lambda_q \geq 0$, as well as some non-negative scalars $\mu \geq 0$ and $\hat{\beta} \in K(\hat{x}, \hat{y})$, such that*

$$0_n \in \sum_{j=1}^{q} \lambda_j \partial f_j(\hat{x}) - \mu \partial_x \$(\hat{x}, \hat{y}, 1, \hat{\beta}) \qquad \textit{and} \qquad \mu + \sum_{j=1}^{q} \lambda_j > 0.$$

Proof. Since $\hat{x}$ is a properly efficient solution for MGSIP, there exist some positive numbers $\lambda_1, \ldots, \lambda_q > 0$ such that

$$\sum_{j=1}^{q} \gamma_j f_j(x) - \sum_{j=1}^{q} \gamma_j f_j(\hat{x}) \geq 0, \qquad \forall x \in S. \tag{5}$$

The value function of lower-level problem (4) is defined as

$$\vartheta(x_0) := \begin{cases} \inf\{g(x_0, y) \mid y \in Y(x_0)\}, & \text{if } Y(x_0) \neq \emptyset, \\ +\infty, & \text{if } Y(x_0) = \emptyset. \end{cases} \tag{6}$$

Consider the following function

$$\theta(x) := \max\Big\{ \sum_{j=1}^{q} \gamma_j f_j(x) - \sum_{j=1}^{q} \gamma_j f_j(\hat{x}) \,,\, -\mu(x) \Big\}.$$

Note that θ is a locally Lipschitz function because that it is maximum of a convex function and a concave function. If $x \in S$, then $\theta(x) \geq 0$ by (5). Also, if $x \notin S$, there is a $y_0 \in Y(x)$ with $g(\hat{x}, y_0) < 0$ and so $-\vartheta(x) > 0$ by (6). Hence, $\theta(x) \geq 0$ for all $x \in \mathbb{R}^n$, and since $\theta(\hat{x}) = 0$, we conclude that $\hat{x}$ is a minimizer of θ on $\mathbb{R}^n$. Now, Theorem 2.3 implies that

$$0_n \in \partial^c \theta(\hat{x}) + N_{\mathbb{R}^n}(\hat{x}) = \partial^c \theta(\hat{x}). \tag{7}$$

If $Y(x_0) = \emptyset$, then $\theta(x) = \sum_{j=1}^{q} \gamma_j f_j(x) - \sum_{j=1}^{q} \gamma_j f_j(\hat{x})$, and so

$$0_n \in \sum_{j=1}^{q} \gamma_j \partial f_j(\hat{x}),$$

by (7). Thus, (i) is proved. From now we suppose that $Y(x_0) \neq \emptyset$. Employing (7), we have

$$0_n \in \partial^c \theta(\hat{x}) \subseteq conv\Big(\sum_{j=1}^{q} \gamma_j \partial f_j(\hat{x}) \,\cup\, \big(-\partial\vartheta(\hat{x})\big) \Big),$$

and hence, there exists some $\tau \in [0, 1]$ such that

$$0_n \in \tau \sum_{j=1}^{q} \gamma_j \partial f_j(\hat{x}) - (1 - \tau)\partial\vartheta(\hat{x}). \tag{8}$$

Given $\hat{y} \in Y_0(\hat{x})$ and $\xi \in \partial\vartheta(\hat{x})$, we have $\vartheta(\hat{x}) = g(\hat{x}, \hat{y})$ and

$$\vartheta(x) - \vartheta(\hat{x}) \geq \langle \xi, x - \hat{x} \rangle, \qquad \forall\, x \in \mathbb{R}^n,$$

by the definition of the convex subdifferential. This inequality and

$$\vartheta(x) \leq g(x, y), \qquad \forall\, (x, y) \in S \times Y(x),$$

conclude that

$$g(x, y) - g(\hat{x}, \hat{y}) \geq \langle \xi, x \rangle - \langle \xi, \hat{x} \rangle, \quad \forall\, (x, y) \in S \times Y(x).$$

This means that $(\hat{x}, \hat{y})$ is a solution to the following convex optimization problem:

$$\begin{aligned} &\min g(x, y) - \langle \xi, x \rangle \\ &\text{s.t. } \; h_i(x, y) \leq 0, \qquad i = 1, 2, \ldots, p. \end{aligned}$$

Since the BCQ holds at $(\hat{x}, \hat{y})$ for the above problem, by KKT necessary condition [3, VII Prop. 2.2.1] there is a $\beta := (\beta_1, \beta_2, \dots, \beta_p) \in \mathbb{R}^p_+$, such that

$$\begin{aligned}(0_n, 0_m) &\in \partial\Big(g(x,y) - \langle \xi . x\rangle\Big)(\hat{x}, \hat{y}) + \sum_{i=1}^{p} \beta_i \partial h_i(\hat{x}, \hat{y}) \\ &= \partial g(\hat{x}, \hat{y}) - (\xi, 0_m) + \sum_{i=1}^{p} \beta_i \partial h_i(\hat{x}, \hat{y}), \qquad (9)\end{aligned}$$

and

$$\beta_i h_i(\hat{x}, \hat{y}) = 0, \quad \forall\, i = 1, 2, \dots, p.$$

Now, we use the following important relationship between the full and partial subdifferentials of convex functions $\Psi(x, y)$ that holds by, e.g., [1, Prop. 2.3.15]:

$$\partial\Psi(x, y) \subseteq \partial_x \Psi(x, y) \times \partial_y \Psi(x, y). \qquad (10)$$

According to (10) and (9), we deduce that

$$\begin{cases} 0_n \in \partial_x g(\hat{x}, \hat{y}) - \xi + \sum_{i=1}^{p} \beta_i \partial_x h_i(\hat{x}, \hat{y}), \\ 0_m \in \partial_y g(\hat{x}, \hat{y}) + \sum_{i=1}^{p} \beta_i \partial_y h_i(\hat{x}, \hat{y}), \\ \beta_i h_i(\hat{x}, \hat{y}) = 0, \quad \forall\, i \in P. \end{cases} \iff \begin{cases} \xi \in \partial_x g(\hat{x}, \hat{y}) + \sum_{i=1}^{p} \beta_i \partial_x h_i(\hat{x}, \hat{y}), \\ 0_m \in \partial_y g(\hat{x}, \hat{y}) + \sum_{i=1}^{p} \beta_i \partial_y h_i(\hat{x}, \hat{y}), \\ \beta_i h_i(\hat{x}, \hat{y}) = 0, \quad \forall\, i \in P. \end{cases}$$

$$\iff \begin{cases} \xi \in \partial_x \Big(g + \sum_{i=1}^{p} \beta_i h_i\Big)(\hat{x}, \hat{y}), \\ 0_m \in \partial_y \Big(g + \sum_{i=1}^{p} \beta_i h_i\Big)(\hat{x}, \hat{y}), \\ \beta_i h_i(\hat{x}, \hat{y}) = 0, \quad \forall\, i \in P. \end{cases} \iff \begin{cases} \xi \in \partial_x \$(\hat{x}, \hat{y}, 1, \beta), \\ 0_m \in \partial_y \$(\hat{x}, \hat{y}, 1, \beta), \\ \beta_i h_i(\hat{x}, \hat{y}) = 0, \quad \forall\, i \in P. \end{cases}$$

$$\iff \xi \in \bigcup_{\beta \in K(\hat{x}, \hat{y})} \partial_x \$(\hat{x}, \hat{y}, 1, \beta).$$

Since ξ was an arbitrary element of $\vartheta(\hat{x})$, we thus proved

$$\partial\vartheta(\hat{x}) \subseteq \bigcup_{\beta \in K(\hat{x}, \hat{y})} \partial_x \$(\hat{x}, \hat{y}, 1, \beta). \qquad (11)$$

Owning to (8) and (11), we find some $\hat{\beta} \in K(\hat{x}, \hat{y})$ such that

$$0_n \in \sum_{j=1}^{q} \tau\gamma_j \partial f_j(\hat{x}) - (1-\tau)\partial_x \$(\hat{x}, \hat{y}, 1, \hat{\beta}).$$

Taking $\lambda_j := \tau\gamma_j$ as $j = 1, \dots, q$ and $\mu := 1 - \tau$ in above inclusion, the result is proved. □

It is known that if the h_i functions as $i \in P$ are affine, then BCQ holds at each $(\hat{x}, \hat{y}) \in \Lambda$ (see [3]). Therefore, the following corollary is immediate by Theorem 3.3.

Corollary 3.4. *Let $\hat{x}$ be a properly efficient solution for MGSIP and $Y_0(\hat{x}) \neq \emptyset$. Suppose that the functions f and g are convex, and that h_i as $i \in P$ are sffine on $\mathbb{R}^n \times \mathbb{R}^m$. Then, for all $\hat{y} \in Y_0(\hat{x})$ there exist some non-negative coefficients $\lambda_1, \ldots, \lambda_q > 0$, as well as some non-negative scalars $\mu \geq 0$ and $\hat{\beta} \in K(\hat{x}, \hat{y})$, such that*

$$0_n \in \sum_{j=1}^{q} \lambda_j \partial f_j(\hat{x}) - \mu \partial_x \$(\hat{x}, \hat{y}, 1, \hat{\beta}) \qquad \textit{and} \qquad \mu + \sum_{j=1}^{q} \lambda_j > 0.$$

Note that the above necessary optimality condition for MGSIPs with linear lower-level constraint is very appealing since no uniform boundedness assumption and no constraint qualification are required.
Observe that the necessary optimality condition in Theorem 3.3 can be stated for one $y_0 \in Y_0(x_0)$ only. Since the Slater condition does not depend on (x_0, y_0), we have the following necessary optimality condition for all $(x_0, y_0) \in gphY_0$.

Theorem 3.5. *Suppose that $\hat{x}$ is a properly efficient solution for convex MGSIP with $Y_0(\hat{x}) \neq \emptyset$. If SCQ holds, then for each $\hat{y} \in Y_0(\hat{x})$, there exist some non-negative coefficients $\lambda_1, \ldots, \lambda_q \geq 0$, as well as some non-negative scalars $\mu \geq 0$ and $\hat{\beta} \in K(\hat{x}, \hat{y})$, such that*

$$0_n \in \sum_{j=1}^{q} \lambda_j \partial f_j(\hat{x}) - \mu \partial_x \$(\hat{x}, \hat{y}, 1, \hat{\beta}) \qquad \textit{and} \qquad \mu + \sum_{j=1}^{q} \lambda_j > 0.$$

Proof. By the definition of SCQ, there is an $(x^*, y^*) \in \Lambda$ such that $h_i(x^*, y^*) < 0$ for all $i \in P$. Let $\hat{y} \in Y_0(\hat{x})$ and $i \in P$ are given, arbitrarily. For each $\xi \in \partial h_i(\hat{x}, \hat{y})$, we have

$$\langle \xi, (x^*, y^*) - (\hat{x}, \hat{y}) \rangle \leq \overbrace{h_i(x^*, y^*)}^{<0} - \overbrace{h_i(\hat{x}, \hat{y})}^{=0} < 0.$$

Hence, $(x^*, y^*) - (\hat{x}, \hat{y}) \in \big(\partial h_i(\hat{x}, \hat{y})\big)^s$ for all $i \in P$, and so

$$\Big(conv\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)\Big)^s = \Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)^s = \bigcap_{i \in P} \big(\partial h_i(\hat{x}, \hat{y})\big)^s \neq \emptyset. \tag{12}$$

Thus, $0_{n+m} \notin \Big(conv\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)\Big)^s$, and since $\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})$ is a compact set, Theorem 2.1 concludes that $cone\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)$ is a closed set.
On the other hand, for each

$$\zeta \in \Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)^s = \Big(conv\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)\Big)^s. \tag{13}$$

Taking

$$G(x, y) := \max\{h_i(x, y) \mid i \in P\}, \quad \forall (x, y) \in \Lambda,$$

we conclude by (**??**) that

$$\partial^c G(\hat{x}, \hat{y}) \subseteq conv\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big) \implies \Big(conv\Big(\bigcup_{i \in P} \partial h_i(\hat{x}, \hat{y})\Big)\Big)^s \subseteq \big(\partial^c G(\hat{x}, \hat{y})\big)^s.$$

This inclusion and (1) deduce that $G\big((\hat{x}, \hat{y}); \zeta\big) < 0$ for all $\zeta \in \mathbb{R}^n \times \mathbb{R}^m$. So, the definition of Clarke directional derivative implies that there exists a $\varepsilon > 0$ such that

$$G\big((\hat{x}, \hat{y}) + \delta\zeta\big) - \overbrace{G(\hat{x}, \hat{y})}^{\leq 0} < 0, \qquad \forall\, \delta \in (0, \varepsilon].$$

Consequently, $G\big((\hat{x},\hat{y})+\delta\zeta\big)<0$, and so $h_i\big((\hat{x},\hat{y})+\delta\zeta\big)<0$ for all $i\in P$ and all $\delta\in(0,\varepsilon]$. This means that $(\hat{x},\hat{y})+\delta\zeta\in\Lambda$ for all $\delta\in(0,\varepsilon]$, which concludes $subdifferential\in T_\Lambda(\hat{x},\hat{y})$. Since ζ was chosen arbitrarily, we proved that

$$\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)^s\subseteq T_\Lambda(\hat{x},\hat{y}),$$

and so

$$\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)^\ominus=\overline{\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)^s}\subseteq\overline{T_\Lambda(\hat{x},\hat{y})}=T_\Lambda(\hat{x},\hat{y}).$$

Consequently,

$$N_\Lambda(\hat{x},\hat{y})=\Big(T_\Lambda(\hat{x},\hat{y})\Big)^\ominus\subseteq\bigg(\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)^\ominus\bigg)^\ominus=\overline{cone\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)}.$$

Since we proved that $cone\Big(\bigcup_{i\in P}\partial h_i(\hat{x},\hat{y})\Big)$ is a closed set, the above inclusion implies that BCQ holds at $(\hat{x},\hat{y})$. Now, Employing Theorem 3.3, the proof is complete, □

Note that SCQ is weaker than the Slater condition, introduced in [7], which requires the existence of $y^*\in Y_0(\hat{x})$ such that $h_i(\hat{x},y^*)<0$ as $i\in P$. Thus, Theorem 3.5 is a generalization of [7, Theorem 4.3.5], that the required Slater condition is weaker and the uniform boundedness assumption is not needed.

References

[1] F.H. Clarke, *Optimization and nonsmooth analysis*, Wiley-Interscience 1983.

[2] G. Edalat, N. Kanzi and M. Adib, Upper Level Qualifications and Optimality in Multiobjective Convex Generalized Semi-Infinite Problems. Journal of Mathematical Extension, 7, 11-23 (2023).

[3] J. B. Hiriart- Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms, I & II*, Springer, Berlin, Heidelberg 1991.

[4] N. Kanzi, Lagrange multiplier rules for non-differentiable DC generalized semi-infinite programming problems. Journal of Global Optimization 56, 417-430 (2013).

[5] N. Kanzi and S. Nobakhtian, Necessary optimality conditions for nonsmooth generalized semi-infinite programming problems. European Journal of Operational Research 205, 253- 261 (2010).

[6] H. Soroush, Necessary Conditions in Generalized Semi-Infinite Optimization with Nondifferentiable Convex Data. Journal of Mathematical Extension, 16, 1-14 (2022).

[7] O. Stein, *Bi-level strategies in semi-infinite programming*, Kluwer, Boston 2003.

[8] F. G. Vazquez, L. A. H. Rebollar and J. J. R¨uckmann, On Vector Generalizrd Semi-infinite Programming. Revista Investigacion Operacional, 39, 341-352 (2018).

[9] J. J. Ye and S. Y. Wu, *First order optimality conditions for generalized semi-infinite programming problems*, J. Optim. Theory Appl. **137**, (2008), 419-434.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–209, pp. 114–119

The Generalized Orthogonally Additive $\rho-$Functional Equation between Orthogonally Banach Algebras

Vahid Keshavarz[a,*], Mohammad Taghi Heydari[a]

[a] *Department of Mathematics, College of Sciences, Yasouj University, Yasouj, 75918 Iran*

Article Info

Keywords:
generalized orthogonally additive $\rho-$functional equation
Hyers-Ulam-Rassias stability
Orthogonally fixed point

2020 MSC:
47H10
17B40

Abstract

In this paper, first, we define and solve the generalized orthogonally additive $\rho-$functional equation between orthogonally Banach algebras and we investigate it is an orthogonally additive mapping. After that, explain generalized orthogonally additive $\rho-$functional equation can be $\mathbb{C}-$linear mapping. In the following, we prove the Hyers-Ulam stability and Hyers-Ulam-Rassias stability of generalized orthogonally additive $\rho-$functional equation between orthogonally Banach algebras. Finally, by using orthogonally fixed point, we investigate generalized orthogonally additive $\rho-$functional equation can be hyperstable.

1. Introduction and Preliminaries

The concept of stability of functional equations is of great importance in mathematics and has wide-ranging implications. It was first introduced by Ulam [23] in 1940, when he raised the fundamental question of when a function that approximately satisfies a functional equation must be close to an exact solution of the equation. This question was further developed by Hyers [12] in the context of Banach spaces the following year, representing a significant advancement in the analysis of functional equation stability.In 1978, Rassias [21] introduced generalized Hyers-Ulam stability, which incorporated a new stability concept involving a control function $\varepsilon(\|a\|^r + \|b\|^r)$, where $\varepsilon > 0$ and $r < 1$, for additive mappings. Găvruta [10] later replaced Rassias's control function with $\varphi(a, b)$ and demonstrated its stability in 1994.

Hyers-Ulam stability is characterized by the property of a functional equation where a small approximation in the equation's solutions results in only a minor approximation in the solutions themselves. This concept finds applications in various branches of mathematics, including functional analysis, dynamic equations, and other pure and applied mathematical areas. Such that in 1993, M. Obloza [19] extended the concept of stability to linear differential equations by focusing on the concept of Hyers-Ulam stability, generating substantial interest in stability for fractional differential equations. This interest led to extensive research, with researchers extending the concept of Hyers-Ulam stability to

*Talker
Email addresses: v.keshavarz68@yahoo.com (Vahid Keshavarz), heydari@yu.ac.ir (Mohammad Taghi Heydari)

fractional differential equations and fractional calculus. These developments have significantly expanded the understanding and application of stability concepts in various mathematical contexts. For more information see [17, 22, 24].

After that, Miura et al. [18] expanded the reach and significance of Hyers-Ulam stability by establishing its applicability to stability for linear operators. This extension represents a significant development in the field, broadening the understanding of stability concepts beyond the realm of functional equations. Additionally, researchers have delved into the exploration of this concept within a variety of Hilbert spaces, including Hardy spaces, weighted Hardy spaces, and Fock spaces, indicating the versatility and wide-ranging implications of the stability framework. The investigation of stability in the context of linear operators and various Hilbert spaces offers valuable insights into the behavior and properties of mathematical structures in these settings, contributing to a deeper understanding of stability phenomena in diverse mathematical contexts. For further in-depth information on these advancements, interested readers are encouraged to consult the sources [3, 11, 13], which provide comprehensive details and analysis on the topic.

The concept of hyperstability was initially introduced in [6], focusing on specific ring homomorphisms. Hyperstability refers to a functional equation being considered hyperstable when every approximate solution of the equation is also an exact solution. This notion represents a significant advancement in the study of stability in functional equations. In other words, hyperstability pertains to functional equations where every approximate solution is also an exact solution, representing a significant advancement in the study of stability. Many researchers worked in this field further details are available in [4, 8, 16].

There are several notions of orthogonality in a real normed space, including Birkhoff-James, Boussouis, (semi–)inner product, Singer, Carlsson, unitary–Boussouis, Roberts, Pythagorean, and Diminnie (see, [1, 2]). The concept of orthogonal sets and related notions was introduced by Eshaghi Gordji and colleagues. They provide definitions and properties related to these sets such that, these definitions and concepts lay the groundwork for studying metric spaces and mappings in the context of orthogonality, providing a broader framework for understanding convergence, continuity, and contraction properties in these spaces [14, 15]. In 2017, the set of orthogonal on normed spaces was introduced by M. Eshaghi Gordji et al. [9] as follows.

Definition 1.1. [9] (i) Suppose $X \neq \emptyset$ and $\perp \subseteq X \times X$ is a binary relation. If $\perp$ satisfies the following condition

$$\exists x_0; (\forall y; y \perp x_0) \ \ or \ \ (\forall y; x_0 \perp y),$$

then $(X, \perp)$ is termed an orthogonal set (abbreviated as O-set).
(ii) Given an O-set $(X, \perp)$, a sequence $\{x_n\}_{n\in\mathbb{N}}$ is referred to as an orthogonal sequence (abbreviated as O-sequence) if

$$(\forall n; x_n \perp x_{n+1}) \ \ or \ \ (\forall n; x_{n+1} \perp x_n).$$

(iii) If $(X, \perp)$ is an O-set and (X, d) is a metric space, then $(X, \perp, d)$ forms an orthogonally metric space. A mapping $f : X \to X$ is considered $\perp-$continuous at $x \in X$ if for each O-sequence $\{x_n\}_{n\in\mathbb{N}}$ in X with $x_n \to x$, $f(x_n) \to f(x)$. It is evident that every continuous mapping is also $\perp-$continuous.
(iv) A Cauchy sequence $\{x_n\}$ in X is called a Cauchy orthogonally sequence (abbreviated as Cauchy O-sequence) if for all $n \in A$, $x_n \perp x_{n+1}$ or $x_{n+1} \perp x_n$. An orthogonally metric space $(X, \perp, d)$ is orthogonally complete (abbreviated as O-complete) if every Cauchy O-sequence converges.
(v) : Let $(X, \perp, d)$ be an orthogonally metric space and $0 < \lambda < 1$. A mapping $f : X \to X$ is considered an orthogonality contraction with Lipschitz constant λ if for any x, y with $x \perp y$

$$d(f(x), f(y)) \leq \lambda d(x, y).$$

By using the concept of orthogonal sets, Bahraini et al. as follows proved the fixed point theorem of Diaz-Margolis for these sets. They showed states the conditions under which a function on an O-complete generalized metric space has a unique fixed point.

Theorem 1.2. *[5] Assume that $(X, d, \perp)$ is an O-complete generalized metric space. Let $T : X \to X$ be a $\perp$-preserving, $\perp$-continuous, and $\perp$-λ-contraction. Let $x_0 \in X$ satisfy for all $y \in X$, $x_0 \perp y$ or for all $y \in X$,*

$y \perp x_0$, and consider the O-sequence of successive approximations with initial element x_0; x_0, $T(x_0)$, $T^2(x_0)$, ..., $T^n(x_0)$, ... Then, either $d(T^n(x_0), T^{n+1}(x_0)) = \infty$ for all $n \geq 0$, or there exists a positive integer n_0 such that $d(T^n(x_0), T^{n+1}(x_0)) < \infty$ for all $n > n_0$. If the second alternative holds, then
(i): the O-sequence of $\{T^n(x_0)\}$ is convergent to a fixed point x^* of T,
(ii): x^* is the unique fixed point of T in

$$X^* = \{y \in X : d(T^n(x_0), y) < \infty\}.$$

(iii): if $y \in X$, then

$$d(y, x^*) \leq \frac{1}{1-\lambda} d(y, T(y)).$$

In 2015, Park [20], introduced additive ρ-functional inequalities and demonstrated the Hyers-Ulam stability of these inequalities in both Banach spaces and non-Archimedean Banach spaces. In the following, we mention the definition of orthogonally generalized additive $\rho-$functional equation between orthogonally normed spaces, a more expand of the study of additive ρ-functional equations in the context of orthogonality and normed spaces.
Let X and Y are orthogonally normed spaces. The mapping $T : X \to Y$, called orthogonally additive mapping if it satisfies $T(x+y) = T(x) + T(y)$, for all $x, y \in X$ with $x \perp y$. By attention to concept of orthogonally additive mapping, in the next, we explain about orthogonally generalized additive $\rho-$functional equation where $\rho \neq 0, \pm 1$. The mapping T called orthogonally generalized additive $\rho-$functional equation on orthogonally normed spaces for all $x, y, z \in X$, with $x \perp y,\ x \perp z$ and $y \perp z$, if satisfies

$$T(x+y+z) + T(x) - T(x+y) - T(x+z) = \rho\Big(T(x-y) + T(y-z) - T(x-z)\Big). \tag{1}$$

In this paper, we solve orthogonally generalized additive $\rho-$functional equation such that we show it is an orthogonally additive mappings. After that, we can investigate it is an orthogonally $\mathbb{C}$-linearity. In the next, by using orthogonally fixed point, we prove the Hyers-Ulam stability of orthogonally generalized additive $\rho-$functional equation using control functions of Găvruta and Rassias. Finally, we can show of orthogonally generalized additive $\rho-$functional equation can be hyperstable with both of control functions of Găvruta and Rassias.

2. Stability and Hyperstability

In this section, let $\rho \neq 0, \pm 1$, and X and Y be orthogonally normed spaces, and $\mathfrak{A}$ and $\mathfrak{B}$ be orthogonally Banach algebras. Firstly, we will prove a lemma for the orthogonally generalized additive ρ-functional equation, demonstrating that it is an additive mapping.

Lemma 2.1. *Let X and Y are orthogonally normed spaces. If the mapping $T : X \to Y$ satisfies*

$$T(x+y+z) + T(x) - T(x+y) - T(x+z) = \rho\Big(T(x-y) + T(y-z) - T(x-z)\Big) \tag{2}$$

for all $x, y, z \in X$, with $x \perp y,\ x \perp z$ and $y \perp z$, then T is an orthogonally additive mapping.

Proof. The mapping T satisfies the equation (2). By setting $x = y = z = 0$ in (2), we find that $T(0) = 0$. Then, setting $x = z = 0$ in (2), we obtain the following equation

$$T(-y) = -T(y). \tag{3}$$

In the following step, we use equation (3) and substitute $x = 0$ and $y = z$ into equation (2), resulting in the following expression

$$T(2y) = 2T(y). \tag{4}$$

Lastly, once more substituting $x = 0$ and utilizing equations (3) and (4) in equation (2), we obtain

$$T(x+y) = T(x) + T(y).$$

This statement means that the function T from X to Y preserves orthogonally addition. □

In the continue of section, we define α as a member of the set $T^1_{\frac{1}{n_0}} := \{e^{i\theta};\quad 0 \le \theta \le \frac{2\pi}{n_0}\}$. This set consists of complex numbers that lie on the unit circle in the complex plane, with their angles ranging from 0 to $\frac{2\pi}{n_0}$ radians.
In the next lemma, we incorporate the element α into the orthogonally generalized additive ρ-functional equation. By doing so, we proceed to verify that T, the function in question, is a $\mathbb{C}$-linear mapping. This means that T satisfies in above lemma and it is homogeneous over the complex numbers (i.e., for any complex number α and any element $x \in X$,

$$T(\alpha u) = \alpha T(u).$$

Lemma 2.2. *If X and Y are two orthogonally normed spaces, and the mapping $T : X \to Y$ satisfies the equation*

$$\alpha T(x+y+z) + \alpha T(x) - T(\alpha x + \alpha y) - T(\alpha x + \alpha z) = \rho\Big(T(\alpha x - \alpha y) + \alpha T(y-z) - T(\alpha x - \alpha z)\Big) \tag{5}$$

then T is an orthogonally $\mathbb{C}$-linear mapping.

Proof. Using lemma 2.1, we conclude that T is a additive mapping. Setting $x = z = 0$ in (2), we obtain the following

$$T(\alpha x) = \alpha T(x) \tag{6}$$

The mapping T is an orthogonally $\mathbb{C}-$linear mapping for complex numbers, based on the reasoning used in the proof of Theorem 3.2 in [7]. □

In the next theorem, we will use Theorem 1.2 to establish the Hyers-Ulam stability of the orthogonally generalized additive ρ-functional equation with the control function of Găvruta.

Theorem 2.3. *Let the mappings $\sigma : \mathfrak{A}^3 \to [0,\infty)$, satisfy the inequality*

$$\sigma\left(\frac{x}{2}, \frac{y}{2}, \frac{z}{2}\right) \le \frac{\beta}{2}\sigma(x,y,z). \tag{7}$$

for some $0 < \beta < 1$. And if $f : \mathfrak{A} \to \mathfrak{B}$ satisfies a following inequality

$$\left\|\alpha f(x+y+z) + \alpha f(x) - f(\alpha x+\alpha y) - f(\alpha x+\alpha z) - \rho\Big(f(\alpha x-\alpha y) + \alpha f(y-z) - f(\alpha x-\alpha z)\Big)\right\| \le \sigma(x,y,z), \tag{8}$$

then there exists an orthogonally unique $\mathbb{C}-$linear mapping $T : \mathfrak{A} \to \mathfrak{B}$ such that

$$\|f(x) - T(x)\| \le \frac{\beta}{1-\beta}\sigma(0,x,x)$$

for all $x \in \mathfrak{A}$.

In the next theorem, the goal is to investigate whether the orthogonally generalized additive ρ-functional equation can exhibit hyperstability with Găvruta's control function.

Theorem 2.4. *Suppose there is a function $\sigma : \mathfrak{A}^3 \to [0,\infty)$ such that:*

$$\lim_{n\to\infty} \frac{1}{2^n}\sigma\left(0, 2^n y, 2^n z\right) = 0,$$

Additionally, if an orthogonally mapping $f : \mathfrak{A} \to \mathfrak{B}$ satisfies a following inequality

$$\left\|\alpha f(x+y+z) + \alpha f(x) - f(\alpha x+\alpha y) - f(\alpha x+\alpha z) - \rho\Big(f(\alpha x-\alpha y) + \alpha f(y-z) - f(\alpha x-\alpha z)\Big)\right\| \le \sigma(0,y,z). \tag{9}$$

Then $f : \mathfrak{A} \to \mathfrak{B}$ is an orthogonally $\mathbb{C}-$linear mapping.

Based on the theorems 2.3 and 2.4, the following corollaries, we investigate the stability and hyperstability of the orthogonally generalized additive ρ-functional equation between orthogonally algebras, considering the control function of Rassias.

Corollary 2.5. *Let $t < 1$ and θ be nonegative real numbers. If an orthogonally mapping $f : \mathfrak{A} \to \mathfrak{B}$ satisfies*

$$\left\| \alpha f(x+y+z)+\alpha f(x)-f(\alpha x+\alpha y)-f(\alpha x+\alpha z)-\rho\Big(f(\alpha x-\alpha y)+\alpha f(y-z)-f(\alpha x-\alpha z)\Big) \right\| \leq \theta(\|x\|^t+\|y\|^t+\|z\|^t). \tag{10}$$

Then there is an unique orthogonally $\mathbb{C}-$linear mapping $T : \mathfrak{A} \to \mathfrak{B}$ such that

$$\|f(x) - T(x)\| \leq \frac{2\alpha}{2-2^t}\|x\|^t.$$

Now, we investigate the hyperstability for orthogonally generalized additive ρ-functional equation.

Corollary 2.6. *Let $t < 1$ and θ be nonegative real numbers. If an orthogonally mapping $f : \mathfrak{A} \to \mathfrak{B}$ satisfies*

$$\left\| \alpha f(x+y+z)+\alpha f(x)-f(\alpha x+\alpha y)-f(\alpha x+\alpha z)-\rho\Big(f(\alpha x-\alpha y)+\alpha f(y-z)-f(\alpha x-\alpha z)\Big) \right\| \leq \theta(\|y\|^t+\|z\|^t).$$

Then $f : \mathfrak{A} \to \mathfrak{B}$ is an orthogonally $\mathbb{C}-$linear mapping.

3. Conclusions

First, a new orthogonally generalized additive ρ-functional equation is introduced, with the condition that ρ is not equal to 0, ± 1, and this equation is defined on orthogonally normed spaces. Next, the equation is solved within a class of additive functions. Then, we demonstrated that the solution satisfies the condition of being a $\mathbb{C}$-linear mapping. Finally, by utilizing the orthogonally fixed point theorem, we investigated is conducted to determine whether the orthogonally generalized additive ρ-functional equation, where ρ is not equal to 0, ± 1, can be both Hyers-Ulam stable and hyper stable using the control functions of G$\breve{a}$vruta and Rassias.

References

[1] J. Alonso, C. Benítez, Carlos orthogonality in normed linear spaces: A survey II. Relations between main orthogonalities, Extracta Math. **4** (1989), 121–131.

[2] J. Alonso, C. Benítez, Orthogonality in normed linear spaces: A survey I. Main properties, Extracta Math. **3** (1988), 1–15.

[3] D. R. Anderson and M. Onitsuka, Best constant for Hyers–Ulam stability of two step sizes linear difference equations, Journal of Mathematical Analysis and Applications, 496 (2021), https://doi.org/10.1016/j.jmaa.2020.124807

[4] M. Almahalebi, S. Kabbaj and G. H. Kim, On the hyperstability of Jensen functional equation in 2-Banach spaces, Fixed Point Theory, **20**(2) (2019), 417-430. https://doi.org/10.24193/fpt-ro.2019.2.27

[5] A. Bahraini, G. Askari, M. Eshaghi Gordji and R. Gholami, Stability and hyperstability of orthogonally $*$-m-homomorphisms in orthogonally Lie C^*-algebras: A fixed point approach, J. Fixed Point Theory Appl., (2018). doi.org/10.1007/s11784-018-0571-0

[6] D. G. Bourgin, Approximately isometric and multiplicative transformations on continuous function rings, Duke Math. J., 16(1949), 385–397.

[7] M. Eshaghi Gordji and A. Fazeli, Stability and superstability of homomorphisms on C*-ternary algebras, An. St. Univ. Ovidius Constanta, 20 (2012), 173–188.

[8] M. Eshaghi Gordji and V. Keshavarz, Hyperstability of ternary Jordan homomorphisms on unital ternary C*-algebras, In: Analysis, Geometry, Nonlinear Optimization and Applications. World Scientific Publishing Company, Editors: P. M. Pardalos and Th. M. Rassias. https://doi.org/10.1142/9789811261572_0011

[9] M. Eshaghi Gordji, M. Ramezani, M. De La Sen, Y.J. Cho, On orthogonal sets and Banach fixed point theorems, Fixed Point Theory, **18** (2017), 569–578.

[10] P. G$\breve{a}$vruta, A generalization of the Hyers–Ulam–Rassias stability of approximately additive mappings, J. Math. Anal. Appl., **184** (1994), 431–436.

[11] M. Haji Shaabani and M. Farzi Haromi, Hyers-Ulam stability of some linear operators on a Hilbert space, J. Math. Ext., 16 (2022).

[12] D. H. Hyers, On the stability of the linear functional equation, Proceedings of the National Academy of Sciences of the United States of America, **27** (1941), 222–224.

[13] V. Keshavarz, M. T. Heydari and D. R. Anderson, Hyers–Ulam stabilities for mth differential operators on H^2_β, Chaos, Solitons & Fractals, **179** (2024). https://doi.org/10.1016/j.chaos.2023.114443

[14] V. Keshavarz and S. Jahedi, Hyperstability of orthogonally 3-lie homomorphism: an orthogonally fixed point approach, Approximation and Computation in Science and Engineering, In: Daras, N.J., Rassias, T.M. (eds) Approximation and Computation in Science and Engineering. Springer Optimization and Its Applications, **180**. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-84122-5_25

[15] V. Keshavarz, S. Jahedi, M. Eshaghi Gordji and S. Bazeghi, Hyperstability of orthogonally pexider lie functional equation: an orthogonally fixed point approach, Thai Journal of Mathematics, **20** (2022), 235-243.

[16] S. O. Kim and A. Bodaghi, Hyperstability of Jordan triple derivations on banach algebras, Acta Mathematica Vietnamica, **41** (2016), 583–594. https://doi.org/10.1007/s40306-015-0156-6

[17] M. Li, X. Yang, Q. Song and X. Chen, Iterative sequential approximate solutions method to Hyers–Ulam stability of time-varying delayed fractional-order neural networks, Neurocomputing, **557** (2023). https://doi.org/10.1016/j.neucom.2023.126727

[18] T. Miura, S. Miyajima, S.E. Takahasi, Hyers-Ulam stability of linear differential operator with constant coefficients, Math. Nachr., **258** (2003), 90–-96.

[19] M. Obloza, Hyers stability of the linear differential equation, Rocznik Nauk. Dydakt. Prace Mat., **13** (1993), 259–270.

[20] C. Park, Additive ρ-functional inequalities in non-Archimedean normed spaces, J. Math. Inequal., 9(2015), 397–407.

[21] Th. M. Rassias, On the stability of the linear mapping in Banach spaces, Proc. Amer.Math. Soc., **72** (1978), 297–300.

[22] H. M. Srivastava, A. K. Nai and R.K. Vats, A theoretical study of the fractional-order p-Laplacian nonlinear Hadamard type turbulent flow models having the Ulam–Hyers stability, Revista de la Real Academia de Ciencias Exactas, F$\acute{i}$sicas y Naturales. Serie A. Matem$\acute{a}$ticas, **117** (160) (2023). https://doi.org/10.1007/s13398-023-01488-6

[23] S. M. Ulam , Problems in Modern Mathematics, Chapter IV, Science edn. Wiley, New York, 1960.

[24] S. Wang, The Ulam stability of fractional differential equation with the Caputo-Fabrizio derivative, Journal of Function Spaces, **2022**, Article ID 7268518, 9 pages, (2022). https://doi.org/10.1155/2022/7268518

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–214, pp. 120–122

On an open problem regarding character injectivity

Mohammad Fozouni[a,*]

[a] *Gonbad Kavous University, Faculty of Basic Science and Engineering*

Article Info

Keywords:
Injectivity
ϕ-injectivity
ϕ-amenability
semigroup algebras

2020 MSC:
46M10
46H25

Abstract

S. S. Renani and R. Nasr Isfahani in [7] introduced the notion of ϕ-injectivity. Essmaili et al. in [5] continued the investigations of this outstanding work and obtain some results on semigroup algebras. In this paper, we aim to draw your attention to an open problem that arose when the authors of [5] were working on this topic.

1. Introduction and Basic Definitions

The aim of this work is not to present some new findings. Here, we just want to repeat some definitions and ground breaking works in the world of Horological algebras. Then, give the definitions of some new adoption of injectivity which is related to characters. Finally, we present an open problem which we believe plays a fundamental role in the Harmonic Analysis world. So, let's get started.

Suppose that A is a Banach algebra. We denote by **A-mod** and **mod-A** the categories of Banach left A-modules and Banach right A-modules, respectively. In the case where A is unital, we also denote by **A-unmod** the categories of unital Banach left A-modules. For each $E, F \in$ **A-mod**, let ${}_AB(E,F)$ be the closed subspace of $B(E,F)$ consisting of the left A-module morphisms. An operator $T \in B(E,F)$ is called *admissible* if ker$T$ and ImT are closed complemented subspaces of E and F, respectively. It is easy to verify that T is admissible if and only if there exists $S \in B(F,E)$ such that $T \circ S \circ T = T$.

A Banach left A-module E is called *injective* if for each $F, K \in$ **A-mod** and admissible monomorphism $T \in_A B(F,K)$, the induced map ${}_AB(K,E) \longrightarrow {}_AB(F,E)$ is onto. We also say $E \in$ **mod-A** is *flat* if the dual module of $E^* \in$ **A-mod** is injective with the following left module action:

$$(a \cdot f)(x) = f(x \cdot a) \qquad (a \in A, x \in E).$$

The notions of injectivity and flatness of Banach algebras were introduced by A. Ya. Helemskii. These notions have been studied for various classes of Banach modules; see [3], [4], [9] and [10] for more details. Recently, Ramsden in [9]

* Talker
Email address: fozouni@gonbad.ac.ir (Mohammad Fozouni)

studied injectivity and flatness of Banach modules over semigroup algebras. It is well known that if A is amenable, then every Banach A-modules is flat but the converse is a long standing open problem. We recall that the answer is positive for some classes of Banach algebras associated with locally compact groups such as, the class of group algebras and measure algebras; see [3] and [8].

Kaniuth, Lau and Pym introduced and studied in [1] and [2] the notion of ϕ-amenability for Banach algebras, where $\phi : A \longrightarrow \mathbb{C}$ is a character, i.e., a non-zero homomorphism on A. Afterwards, Monfared introduced and studied in [6] the notion of character amenability for Banach algebras. Let $\Delta(A)$ be the set of all characters of the Banach algebra A, and let $\phi \in \Delta(A)$. The Banach algebra A is called *left ϕ-amenable* if for all Banach A-bimodules E for which the left module action is given by

$$a \cdot x = \phi(a)x \qquad (a \in A, x \in E),$$

every derivation $D : A \longrightarrow E^*$ is inner. It is clear that amenability of A implies ϕ-amenability for all $\phi \in \Delta(A)$.

Recently, Nasr-Isfahani and Soltani Renani in [7] introduced and studied the notion of ϕ-injectivity and ϕ-flatness for Banach modules (see Definition 2.1). As an important result, it is shown in [7, Proposition 3.1] that the Banach algebra A is left ϕ-amenable if and only if every Banach left A-modules E is ϕ-flat. Indeed, this result gives a positive answer to the above open problem arises by A. Ya. Helemskii in this homology setting based on character ϕ. Furthermore, they obtained some necessary and sufficient conditions for ϕ-injectivity and characterized ϕ-injectivity of Banach modules in terms of a coretraction problem; see [7, Theorem 2.4].

2. Definition of Character Injectivity

First, we recall some standard notations that we shall use and define the notions of ϕ-injectivity and ϕ-flatness of Banach modules.

Let A be a Banach algebra and $E \in$ **A-mod**. Throughout the paper, we regard E as a Banach left $A^\sharp$-module (the unitization of A) with the following left module action:

$$(a, \lambda) \cdot x = a \cdot x + \lambda x \qquad (a \in A, \lambda \in \mathbb{C}, x \in E).$$

Moreover, the space $B(A, E)$ is a Banach A-bimodule with the following module actions:

$$(a \cdot T)(b) = T(ba), \qquad (T \cdot a)(b) = T(ab) \quad (T \in B(A, E), a, b \in A).$$

Suppose that A is a Banach algebra and $\phi \in \Delta(A)$. For each $E \in$ **A-mod** we define,

$$I(\phi, E) = \text{span}\{a \cdot x - \phi(a)x : a \in A, x \in E\}.$$

Following [7], we also consider

$$_\phi B(A^\sharp, E) = \{T \in B(A^\sharp, E) : T(ab - \phi(b)a) = a \cdot T(b - \phi(b)e^\sharp) \ \text{ for all } \ a, b \in A\},$$

where $e^\sharp = (0, 1)$ denotes the unite of $A^\sharp$. It is straightforward to check that $_\phi B(A^\sharp, E)$ is a closed A-submodule of $B(A^\sharp, E)$. Moreover, we define *the canonical morphism* $_\phi\Pi^\sharp : E \longrightarrow_\phi B(A^\sharp, E)$ as follows:

$$_\phi\Pi^\sharp(x)(a) = a \cdot x \qquad (x \in E, a \in A^\sharp).$$

Definition 2.1. Let A be a Banach algebra, $\phi \in \Delta(A)$ and $E \in$ **A-mod**. We say that E is *ϕ-injective* if, for each $F, K \in$ **A-mod** and admissible monomorphism $T : F \longrightarrow K$ with $I(\phi, K) \subseteq \text{Im}(T)$, the induced map $T_E : {_A B(K, E)} \longrightarrow {_A B(F, E)}$ defined by $T_E(R) = R \circ T$ is onto.

The following theorem gives a characterization of ϕ-injectivity in terms of a coretraction problem.

Theorem 2.2. *([7, Theorem 2.4]) Let A be a Banach algebra and $\phi \in \Delta(A)$. For $E \in$ **A-mod** the following statements are equivalent.*

(i) E is ϕ-injective.

(ii) $_\phi\Pi^\sharp \in {_A B(E, {_\phi B(A^\sharp, E)})}$ is a coretraction, (that is there exists $_\phi\rho^\sharp \in {_A B(_\phi B(A^\sharp, E), E)}$ such that is a left inverse for $_\phi\Pi^\sharp$).

A Banach right (left) A-module E is ϕ-flat if E^* is ϕ-injective as a left (right) A-module. It is shown that Banach algebra A is left ϕ-amenable if and only if each Banach left A-module E is ϕ-flat [7, Proposition 3.1].

3. Main Point of This Paper

Fasten your seat belt and put on your perfectionism glasses. Now, we want to take a deep dive in the mysterious world of Groups in Harmonic Analysis.
Let G be a locally compact group. It is shown in [3, Theorem 4.9] that $L^1(G) \in \mathbf{L^1(G)}$**-mod** is injective if and only if G is amenable and discrete. Now, suppose that $L^1(G) \in \mathbf{L^1(G)}$**-mod** is character injective.

Open problem: Are the following statements equivalent?

(i) $L^1(G) \in \mathbf{L^1(G)}$**-mod** is character injective.

(ii) G is amenable and discrete.

Up to our knowledge this problem has not been solved yet. This question is one of the fundamental piece in the realm of Harmonic Analysis. If you can give a partial answer to it, you can open a new and fundamental door for other mathematicians. We take this great opportunity of conference to bring into your precious attention this question and hope you can solve it.

References

[1] A. T. Lau E. Kaniuth and J. S. Pym. On φ-amenability of banach algebras. *Math. Proc. Cambridge philos. Soc.*, 144:85–96, 2008.

[2] A. T. Lau E. Kaniuth and J. S. Pym. On character amenability of banach algebras. *J. Math. Anal. Appl.*, 344:942–955, 2008.

[3] M. E. Polyakov H. G. Dales. Homological properties of modules over group algebras. *Proc. London Math. Soc.*, 89:390–426, 2004.

[4] A. Ya. Helemskii. A certain class of flat banach modules and its applications. *Vestnik. Moskov. Univ. Ser. Mat. Mekh.*, 27:29–36, 1972.

[5] M. Fozouni M. Essmaili and J. Laali. Hereditary properties of character injectivity with applications to semigroup algebras. *Ann. Funct. Anal.*, 6(2):162–172, 2015.

[6] M. S. Monfared. Character amenability of banach algebras. *Math. Proc. Cambridge Philos. Soc.*, 144:697–706, 2008.

[7] S. Soltani Renani R. Nasr-Isfahani. Character injectivity and projectivity of banach modules. *Quart. J. Math.*, 2013. doi: 10.1093/qmath/hat015.

[8] P. Ramsden. *Homological properties of semigroup algebras*. PhD thesis, University of Leeds, 2008.

[9] P. Ramsden. Homological properties of modules over semigroup algebras. *J. Funct. Anal.*, 258:3988–4009, 2010.

[10] M. C. White. Injective modules for uniform algebras. *Proc. London Math. Soc.*, 73(3):155–184, 1996.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–215, pp. 123–128

The general 3D Jensen $\rho-$functional equation and 3-Lie homomorphism-derivation

Zohreh Kefayati[a,*], Eskandar Naraghirad[b]

[a] *Department of Mathematics, College of Sciences, Yasouj University, Yasouj, 75918 Iran*
[b] *Department of Mathematics, College of Sciences, Yasouj University, Yasouj, 75918 Iran*

Article Info

Keywords:
generalized 3D Jensen $\rho-$functional equation
Hyers-Ulam stability
3-Lie homomorphism-derivation

2020 MSC:
47B47
47H10, 17B40, 16W25

Abstract

In this work, first, we define generalized 3D Jensen $\rho-$functional equation on 3-Lie algebras and solve it belongs to a class of additive mappings, accompanied by its $\mathbb{C}$-linearity. In the next, by the fixed point theorem, we will inquire into the stability of the generalized 3D Jensen $\rho-$functional equation with Găvruta's control function within the framework of Hyers-Ulam stability. After that, the confirmation of the stability of relation (1), we will investigate the concurrent Hyers-Ulam stability of 3-Lie derivations and 3-Lie homomorphisms. Finally, we will discuss the results of the theorem of Rassias for the generalized 3D Jensen $\rho-$functional equation and 3-Lie derivations and 3-Lie homomorphisms on 3-Lie algebras.

1. Introduction and Preliminaries

In the 1800s, Cayley played a pioneering role in introducing 3-ary operations involving cubic matrices. This breakthrough paved the way for the emergence of n-ary algebraic structures spanning various fields. N-ary algebras, with a focus on ternary algebraic structures, have been extensively researched. In 2008, J. Bagger and N. Lambert conducted a study on gauge symmetry, proposing a supersymmetric theory for multiple M2-branes. This theory involves an algebra incorporating ternary operations, now recognized as Bagger-Lambert algebras. This idea transcends the realm of mathematics, as n-ary and ternary algebras represent generalizations of algebraic structures that surpass conventional binary operations. These structures adhere to properties like associativity and linearity, enabling the definition of algebraic systems that extend beyond the confines of binary operations. For example, Both ternary algebra and n-ary algebra have applications in various areas of quantum mechanics, coding theory, physics, computer science, and the domain of Nambu mechanics. For more information see [1, 5, 6, 15].

In the year 2008, the concept of ternary algebras was introduced by M. Amyari and M. S. Moslehian [2]. They defined a ternary algebra $\mathfrak{A}$ as a complex space equipped with a ternary product $(x_1, x_2, x_3) \to [x_1, x_2, x_3]$ mapping from $\mathfrak{A}^3$ to $\mathfrak{A}$. This product is stipulated to be $\mathbb{C}$-linear in the outer variables, conjugate $\mathbb{C}$-linear in the middle variable, and associative. Furthermore, it must satisfy specific norm properties, including $\|[x_1, x_2, x_3]\| \leq \|x_1\|.\|x_2\|.\|x_3\|$ and

*Talker
Email addresses: zohrekefayati68@gmail.com (Zohreh Kefayati), esnaraghirad@yu.ac.ir (Eskandar Naraghirad)

$\|[x,x,x]\| = \|x\|^3$. When a ternary algebra $\mathfrak{A}$ is also a Banach space, it earns the designation of a ternary Banach algebra.

The stability problem of functional equations holds a fundamental place in mathematics, bearing significant implications for their study. In 1940, Ulam [22] introduced this concept, posing the question, "Under what conditions does a function that approximately satisfies a functional equation have to be close to an exact solution of the equation?" The following year, Hyers [10] addressed this query in the context of Banach spaces, representing a crucial advancement in the analysis of functional equation stability. Subsequently, in 1978, Rassias [21] introduced generalized Hyers-Ulam stability, incorporating a novel stability concept with a control function $\varepsilon(\|x_1\|^r + \|x_2\|^r)$, where $\varepsilon > 0$ and $r < 1$, for additive mappings. Later in 1994, Găvruta [8] replaced Rassias's control function with $\varphi(x_1, x_2)$ and demonstrated its stability. Hyers-Ulam stability essentially denotes the property of a functional equation, where a slight perturbation in the equation's solutions results in only a minor perturbation in the solutions themselves. This concept finds applications in various branches of mathematics, including functional analysis, dynamic equations, and other pure and applied mathematical areas. Focusing on the concept of Hyers-Ulam stability, M. Obloza introduced stability for linear differential equations in 1994 [16, 17], generating substantial interest in stability for fractional differential equations. This interest led to extensive research, with researchers extending the concept of Hyers-Ulam stability to fractional differential equations and fractional calculus [4, 7, 11, 18].

Expanding the scope of Hyers-Ulam stability and its significance, Miura et al. established stability for linear operators. Additionally, researchers have explored this concept within various Hilbert spaces, including Hardy spaces, weighted Hardy spaces, and Fock spaces. Further information can be found in [3, 9, 12].

Lie algebras are attributed to Professor Sophus Lie, and in 1893, Scheffers [19] compiled much of the original formulation from Lie's lecture notes in Leipzig. An Banach algebra $\mathcal{X}$ equipped with the product $[x_1, x_2] := \frac{(x_1x_2 - x_2x_1)}{2}$ for all $x_1, x_2 \in \mathcal{X}$ forms an Lie algebra and similarly for 3-Lie algebra $\mathcal{X}$ endowed with the bracket $\Big[[x_1, x_2], x_3\Big] := \frac{[x_1,x_2]x_3 - x_3[x_1,x_2]}{2}$, for all $x_1, x_2, x_3 \in \mathcal{X}$ (see [13]).

In the following, we define generalized 3D Jensen $\rho-$functional equation and 3-Lie homomorphism-derivation between 3-Lie algebras. The mapping J from $\mathcal{X}$ to $\mathcal{X}$ called 3D Jensen $\rho-$functional equation on 3-Lie algebras if satisfies

$$\begin{aligned}\Lambda_J(x_i) &= J\Big(\frac{x_1+x_2}{2}+x_3\Big) + J\Big(\frac{x_1+x_3}{2}+x_2\Big) + J\Big(\frac{x_2+x_3}{2}+x_1\Big) - 2J(x_1) - 2J(x_2) - 2J(x_3) \\ &- \rho\Bigg(3J\Big(\frac{x_1+x_2+x_3}{3}\Big) + J(x_1) + J(x_2) + J(x_3) - 2J\Big(\frac{x_1+x_2}{2}\Big) - 2J\Big(\frac{x_1+x_3}{2}\Big) - 2J\Big(\frac{x_2+x_3}{2}\Big)\Bigg)\end{aligned} \tag{1}$$

where $\rho \neq 0, \pm 1$ and $x_i \in \mathcal{X}$ where is $i = 1, 2, 3$ such that we will solve above equation is a type of additive mapping. In the next part in this section, we have the concepts of 3-Lie derivations and 3-Lie homomorphisms.

A mapping $\zeta : \mathcal{X} \to \mathcal{X}$ is called a 3-Lie homomorphism, if ζ satisfies (1) and

$$\zeta([[x_1, x_2], x_3]) = [[\zeta(x_1), \zeta(x_2)], \zeta(x_3)] \quad \forall\, x_1, x_2, x_3 \in \mathcal{X}.$$

And mapping $D : \mathcal{X} \to \mathcal{X}$ is called a 3-Lie derivation if D satisfies (1) and

$$D([[x_1, x_2], x_3]) = [[D(x_1), x_2], x_3] + [[x_1, D(x_2)], x_3] + [[x_1, x_2], D(x_3)]$$

for all $x_i \in \mathcal{X}$ where $x_i \in \{1, 3\}$.

In the following section, first, we solve that a function satisfying (1) belongs to a category of additive mappings, accompanied by its $\mathbb{C}$-linearity. Subsequently, employing below theorem, we will inquire into the stability of the relation (1) using Găvruta's control function within the framework of Hyers-Ulam stability. Following the confirmation of the stability of relation (1), we will investigate the concurrent Hyers-Ulam stability of 3-Lie derivations and 3-Lie homomorphisms. Lastly, we will engage in an explanation concerning the outcomes of Rassias' theorem for the presented concepts.

Theorem 1.1. *[14] Let a complete generalized metric space $(\mathfrak{S}, d)$, and suppose $G : \mathfrak{S} \to \mathfrak{S}$ is a strictly contractive mapping with a Lipschitz constant $0 < \iota < 1$. Then, for any given element $x_1 \in \mathfrak{S}$, we have either*

$$d(G^j x_1, G^{j+1} x_1) = \infty$$

for all nonnegative integers j, or there exists a positive integer j_0 such that:
(1) $d(G^j x_1, G^{j+1} x_1) < \infty, \qquad \forall j \geq j_0$;
(2) *the sequence $\{G^j x_1\}$ converges to a fixed point x_2^* of G;*
(3) *x_2^* is the unique fixed point of G in the set $\mathfrak{B} = \{x_1 \in \mathfrak{S} \mid d(G^{j_0} x_1, x_2) < \infty\}$;*
(4) *$d(x_2, x_2^*) \leq \frac{1}{1-\iota} d(x_2, Gx_2)$ for all $x_2 \in \mathfrak{B}$.*

2. Main Results

In this section, First, for ease of work, we will introduce some concepts as follows and then we will prove a lemma for the relation (1), so as to show that the relation (1), is an additive mapping.
let $\mathcal{X}$ be a 3-Lie algebras, $\rho \neq 0, \pm 1$, $\gamma \in \mathbb{T}^1 := \{\gamma \in \mathbb{C} \;:\; |\gamma| = 1\}$ and Lipschitz constant $0 < \iota < 1$.

Lemma 2.1. *If A and B be two normed spaces, and the mapping J from A to B satisfies equation (1), in fact J is a additive mapping.*

Proof. Let J satisfies (1), with putting $x_i = 0$ where $i = 1, 2, 3$ in (1), we have $J(0) = 0$. After that, x_i are equal in (1) where $i = 1, 2, 3$, we have

$$J(2x_1) = 2J(x_1). \tag{2}$$

In the continue, by using (2) and putting $x_i = 0$ where $i = 2, 3$ in (1), we get

$$J(3x_1) = 3J(x_1). \tag{3}$$

Again putting $x_1 := -x_2,\; x_3 = 0$ and by using (2) and (3) in (1), we have

$$J(-x_2) = -J(x_2). \tag{4}$$

By attention to relations (2) and (4) and putting $x_3 := x_2$ in (1), we get

$$J\Big(\frac{x_1 - x_2}{2}\Big) + J\Big(\frac{x_1 + x_2}{2}\Big) - J(x_1) = -2\rho\Bigg(J\Big(\frac{x_1 - x_2}{2}\Big) + J\Big(\frac{x_1 + x_2}{2}\Big) - J(x_1)\Bigg) \tag{5}$$

Finally, putting $x := x_1 + x_2$ and $y := x_1 + x_2$ in (5), we get

$$J(x + y) = J(x) + J(y).$$

i.e., J from A to B is additive. □

In the next lemma, by adding $\gamma \in \mathbb{T}^1$ to relation (1), we check that J is a $\mathbb{C}$-linear mapping.

$$\begin{aligned}\Lambda_{J_\gamma}(x_i) =& J\Big(\frac{\gamma x_1 + \gamma x_2}{2} + \gamma x_3\Big) + J\Big(\frac{\gamma x_1 + \gamma x_3}{2} + \gamma x_2\Big) + J\Big(\frac{\gamma x_2 + \gamma x_3}{2} + \gamma x_1\Big) - 2\gamma J(x_1)\\ &- 2\gamma J(x_2) - 2\gamma J(x_3) - \rho\Bigg(3J\Big(\frac{\gamma x_1 + \gamma x_2 + \gamma x_3}{3}\Big) + \gamma J(x_1) + \gamma J(x_2) + \gamma J(x_3)\\ &- 2\gamma J\Big(\frac{x_1 + x_2}{2}\Big) - 2J\Big(\frac{\upsilon x_1 + \gamma x_3}{2}\Big) - 2\gamma J\Big(\frac{x_2 + x_3}{2}\Big)\Bigg)\end{aligned} \tag{6}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$.

Lemma 2.2. *If A and B be two normed spaces, and the mapping J from A to B satisfies equation (6), in fact J is a $\mathbb{C}$-linear mapping.*

Proof. Using lemma 2.1, we conclude that J is a additive mapping. x_i are equal in (6) where $i = 1, 2, 3$, we have

$$J(\gamma x_1) = \gamma J(x_1) \tag{7}$$

for all $x_1 \in \mathcal{X}$. Using the same reasoning as in the proof of [20, Theorem 2.1], we can deduce that J is a $\mathbb{C}-$linear mapping. □

In the next theorem, by using theorem 1.1, we will obtain the Hyers-Ulam stability of relation (6) with the control function of Găvruta. Before stating the main theorem, we check conditions of Găvruta's control function for relation (6), 3-Lie homomorphism and 3-Lie derivation.
Let the mappings θ, π from $\mathcal{X}^3$ to $[0, \infty)$, satisfying

$$\theta\left(\frac{x_1}{2}, \frac{x_2}{2}, \frac{x_3}{2}\right) \leq \frac{\iota}{2}\theta(x_1, x_2, x_3). \tag{8}$$

and

$$\pi\left(\frac{x_1}{2}, \frac{x_2}{2}, \frac{x_3}{2}\right) \leq \frac{\iota}{2^3}\pi(x_1, x_2, x_3). \tag{9}$$

If we put with putting $x_i = 0$ where $i = 1, 2, 3$ in (8) and (9), then we have $\theta(0, 0, 0) = \pi(0, 0, 0) = 0$. After that, in view of (7) and (9), we get

$$\lim_{n\to\infty} 2^n\theta\left(\frac{x_1}{2^n}, \frac{x_2}{2^n}, \frac{x_3}{2^n}\right) = 0 \tag{10}$$

and

$$\lim_{n\to\infty} 2^{3n}\pi\left(\frac{x_1}{2^n}, \frac{x_2}{2^n}, \frac{x_3}{2^n}\right) = 0 \tag{11}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$. Now, by using theorem 1.1, we can investigate the stability of relation (6), 3-Lie homomorphism and 3-Lie derivation on 3-Lie algebras with the control function of Găvruta.

Theorem 2.3. *Let the mappings θ, π from $\mathcal{X}^3$ to $[0, \infty)$, satisfying (8) and (9), respectively. And let J_j from $\mathcal{X}$ to $\mathcal{X}$ are mappings. In fact*
(i) Suppose that J_j satisfies

$$\|\Lambda_{J_j,\gamma}(x_i)\| \leq \theta(x_1, x_2, x_3), \tag{12}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$ and $j = 1, 2$. Then there is a unique $\mathbb{C}-$linear mapping $\mathfrak{T}$ on $\mathcal{X}$ such that

$$\|J_j(x_1) - \mathfrak{T}(x_1)\| \leq \frac{\iota}{1-\iota}\theta(x_1, x_1, x_1)$$

for all $x_1 \in \mathcal{X}$ and $j = 1, 2$.
(ii) Suppose that J_j satisfying (12) and

$$\begin{aligned}&\Big\|J_1([[x_1, x_2], x_3]) - [[J_1(x_1), J_1(x_2)], J_1(x_3)]\Big\| \\ &+ \Big\|J_2([[x_1, x_2], x_3]) - [[J_2(x_1), x_2], x_3] - [[x_1, J_2(x_2)], x_3] - [[x_1, x_2], J_2(x_3)]\Big\| \leq \pi(x_1, x_2, x_3),\end{aligned} \tag{13}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$. Then there are two unique 3-Lie homomorphism ζ and 3-Lie derivation $\mathfrak{D}$ on $\mathcal{X}$, respectively, such that:

$$\|J_1(x_1) - \zeta(x_1)\| \leq \frac{\iota}{1-\iota}\pi(x_1, x_1, x_1),$$

and

$$\|J_2(x_1) - \mathfrak{D}(x_1)\| \leq \frac{\iota}{1-\iota}\pi(x_1, x_1, x_1)$$

for all $x_1 \in \mathcal{X}$.

In the following corollary, by attention to theorem 2.3, we can investigate the Hyers-Ulam-Rassia stability of relation (6), 3-Lie homomorphism and 3-Lie derivation on 3-Lie algebras with the control function of Rassias. Before stating the corollary, we check conditions of the control function of Rassias for relation (6), 3-Lie homomorphism and 3-Lie derivation. For this work, it is enough to change control functions of theorem (2.3), $\theta(x_1, x_2, x_3)$ and $\pi(x_1, x_2, x_3)$, to $\alpha(\|x_1\|^r + \|x_2\|^r + \|x_3\|^r)$, where $r \neq 1$ and α are nonegative real numbers.

Corollary 2.4. *Let J_j from $\mathcal{X}$ to $\mathcal{X}$ are mappings. In fact*
(i) *Suppose that J_j satisfies*

$$\|\Lambda_{J_j,\gamma}(x_i)\| \leq \alpha(\|x_1\|^r + \|x_2\|^r + \|x_3\|^r), \tag{14}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$ and $j = 1, 2$. Then there is a unique $\mathbb{C}-$linear mapping $\mathfrak{T}$ on $\mathcal{X}$ such that

$$\|J_j(x_1) - \mathfrak{T}(x_1)\| \leq \frac{2\alpha}{2^r - 2}\|x_1\|^r, \qquad r > 1$$

and

$$\|J_j(x_1) - \mathfrak{T}(x_1)\| \leq \frac{2\alpha}{2 - 2^r}\|x_1\|^r, \qquad r < 1,$$

for all $x_1 \in \mathcal{X}$ and $j = 1, 2$.
(ii) *Suppose that J_j satisfying (14) and*

$$\begin{aligned} &\Big\|J_1([[x_1, x_2], x_3]) - [[J_1(x_1), J_1(x_2)], J_1(x_3)]\Big\| + \Big\|J_2([[x_1, x_2], x_3]) - [[J_2(x_1), x_2], x_3] \\ &- [[x_1, J_2(x_2)], x_3] - [[x_1, x_2], J_2(x_3)]\Big\| \leq \alpha(\|x_1\|^r + \|x_2\|^r + \|x_3\|^r), \end{aligned} \tag{15}$$

for all $x_i \in \mathcal{X}$ where $i \in \{1, 3\}$. Then there are two unique 3-Lie homomorphism ζ and 3-Lie derivation $\mathfrak{D}$ on $\mathcal{X}$, respectively, such that:

$$\|J_1(x_1) - \zeta(x_1)\| \leq \frac{2\alpha}{2^r - 2}\|x_1\|^r, \qquad r > 1$$

$$\|J_1(x_1) - \zeta(x_1)\| \leq \frac{2\alpha}{2 - 2^r}\|x_1\|^r, \qquad r < 1$$

$$\|J_2(x_1) - \mathfrak{D}(x_1)\| \leq \frac{2\alpha}{2^r - 2}\|x_1\|^r, \qquad r > 1$$

and

$$\|J_2(x_1) - \mathfrak{D}(x_1)\| \leq \frac{2\alpha}{2 - 2^r}\|x_1\|^r, \qquad r < 1,$$

for all $x_1 \in \mathcal{X}$.

3. Conclusions

Due to the importance and applications of 3-Lie algebras and 3-Lie homomorphisms and 3-Lie derivations in mathematical physics and quantum mechanics. By using the concept of Jensen mapping we defined the new concept of generalized 3D Jensen $\rho-$functional equation where $\rho \neq 0, \pm 1$ on normed spaces. After that, we solved it was a class of additively. Also, we solved it was a $\mathbb{C}-$linear mapping between normed spaces. Finally, by using the fixed point theorem, we investigated generalized 3D Jensen $\rho-$functional equation where $\rho \neq 0, \pm 1$ and 3-Lie homomorphisms and 3-Lie derivations can be Hyers-Ulam stable with two control functions of Găvruta and Rassias between 3-Lie algebras.

References

[1] V. Abramov and E. Zappala, 3-Lie algebras, ternary Nambu-Lie algebras and the Yang-Baxter equation, Journal of Geometry and Physics, **183** (2023). https://doi.org/10.1016/j.geomphys.2022.104687

[2] M. Amyari and M.S. Moslehian, Approximately ternary semigroup homomorphisms, Lett. Math. Phys., **77** (2006), 1–9.

[3] D. R. Anderson and M. Onitsuka, Best constant for Hyers–Ulam stability of two step sizes linear difference equations, Journal of Mathematical Analysis and Applications, 496 (2021), https://doi.org/10.1016/j.jmaa.2020.124807

[4] Y. Awad, On the existence and stability of positive solutions of eigenvalue problems for a class of p-Laplacian ψ-caputo fractional integro-differential equations, Journal of Mathematics, **2023**, Article ID 3458858, 26 pages, (2023). https://doi.org/10.1155/2023/3458858

[5] J. Bagger and N. Lambert, Gauge Symmetry and Supersymmetry of Multiple M2-Branes, Phys. Rev., **D77** 065008 (2008). doi.org/10.1103/PhysRevD.77.065008

[6] J. Bagger. N. Lambert, S. Mukhi and C. Papageorgakis, Multiple membranes in M-theory, Physics Reports, **527** (1) (2013), 1-100. https://doi.org/10.1016/j.physrep.2013.01.006

[7] K. Dhanalakshmi and P. Balasubramaniam, Well posedness of second-order non-instantaneous impulsive fractional neutral stochastic differential equations, Bulletin des Sciences Math$\acute{e}$matiques, **189** (2023). https://doi.org/10.1016/j.bulsci.2023.103350

[8] P. G$\breve{a}$vruta, A generalization of the Hyers–Ulam–Rassias stability of approximately additive mappings, J. Math. Anal. Appl., **184** (1994), 431–436.

[9] M. Haji Shaabani and M. Farzi Haromi, Hyers-Ulam stability of some linear operators on a Hilbert space, J. Math. Ext., 16 (2022).

[10] D.H. Hyers, On the stability of the linear functional equation, Proceedings of the National Academy of Sciences of the United States of America, **27** (1941), 222–224.

[11] R. Kasinathan, R. Kasinathan, V. Sandrasekaran and D. Baleanu, Existence and Hyers–Ulam stability of stochastic integrodifferential equations with a random impulse, Journal of Inequalities and Applications, **2023** (116) (2023). https://doi.org/10.1186/s13660-023-03023-y

[12] V. Keshavarz, M. T. Heydari and D. R. Anderson, Hyers–Ulam stabilities for mth differential operators on H^2_β, Chaos, Solitons & Fractals, **179** (2024). https://doi.org/10.1016/j.chaos.2023.114443

[13] V. Keshavarz and S. Jahedi, Hyperstability of orthogonally 3-Lie homomorphism: an orthogonally fixed point approach, Approximation and Computation in Science and Engineering, In: Daras, N.J., Rassias, T.M. (eds) Approximation and Computation in Science and Engineering. Springer Optimization and Its Applications, **180**. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-84122-5_25

[14] B. Margolis and J.B. Diaz, A fixed point theorem of the alternative for contractions on the generalized complete metric space, Bull. Amer. Math. Soc., **126** (1968), 305–309.

[15] J. Nambu, Generalized Hamiltonian Dynamics, Phys. Rev., **D7** (1973), 2405–2412. https://doi.org/10.1103/PhysRevD.7.2405

[16] M. Obloza, Hyers stability of the linear differential equation, Rocznik Nauk. Dydakt. Prace Mat., **13** (1993), 259–270.

[17] M. Obloza, Connections between Hyers and Lyapunov stability of the ordinary differential equations, Rocznik Nauk. Dydakt. Prace Mat., **14** (1997), 141–146.

[18] M. Onitsuka, Ulam stability for second-order linear differential equations with three variable coefficients, Results in Applied Mathematics, **14** (2022). https://doi.org/10.1016/j.rinam.2022.100270

[19] S. Lie and G. Scheffers, Vorlesungen $\ddot{u}$ber kontinuerliche gruppen, Leipzig, 1893.

[20] C. Park, Homomorphisms between Poisson JC^*-algebras, Bull. Braz. Math. Soc., **36** (2005), 79-–97.

[21] Th. M. Rassias, On the stability of the linear mapping in Banach spaces, Proc. Amer.Math. Soc., **72** (1978), 297–300.

[22] S. M. Ulam , Problems in Modern Mathematics, Chapter IV, Science edn. Wiley, New York, 1960.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–227, pp. 129–132

Hyers-Ulam stability of bi-additive s-functional inequality

Zohreh Heydarpour[a,*], Javad Izadi[b]

[a]Department of Mathematics, Payame Noor University (PNU), P. O. Box 19395-4697, Tehran, Iran
[b]Department of Mathematics, Payame Noor University (PNU), P. O. Box 19395-4697, Tehran, Iran

Article Info

Keywords:
Bi-additive s-functional inequality
Hyers-Ulam-Rassias stability
J.M. Rassias stability
Hyers–Ulam stability

2020 MSC:
39B52
46L05
47B47
46L57

Abstract

In this talk, we introduce a concept of the bi-additive s-functional inequality, where s is a fixed nonzero complex number with $|s| < 1$, on algebras. Subsequently, we demonstrate the solution to the above relation, showing that it is a bi-additive mapping. Then, we explore the bi-additive s-functional inequality, which can be bi $\mathbb{C}-$linear between algebras. Finally, we establish the Hyers–Ulam stability of the bi-additive s-functional inequality, where s is a fixed nonzero complex number with $|s| < 1$.

1. Introduction

In 1940, S.M. Ulam [20] proposed what is now known as the stability problem. This problem deals with mappings between groups and metric groups, specifically focusing on the existence of certain conditions for these mappings, he introduced the stability problem in the following question.
Given a small positive value ε, is there a small positive value δ such that if a mapping $h : G_1 \to G_2$ satisfies the inequality $d(h(xy), h(x)h(y)) < \delta$ for all elements $x, y \in G_1$, then there exists a homomorphism $H : G_1 \to G_2$ with the property that $d(h(x), H(x)) < \varepsilon$ for all $x \in G_1$? In 1941, this problem was successfully resolved by D.H. Hyers [3] in the case of Banach spaces, which are complete normed vector spaces. This achievement marked a significant advancement in the understanding of stability in mathematical structures. Building upon D.H. Hyers' work, in 1978, Th.M. Rassias [17] introduced the concept of generalized Hyers-Ulam stability, which extended the scope of stability considerations to a broader class of mathematical structures. Th.M. Rassias' work introduced a novel approach to stability, incorporating a control function characterized by the expression $\varepsilon(\|a\|^r + \|b\|^r)$, where $\varepsilon > 0$ and $r < 1$. After that, J.M. Rassias [16] followed the innovative approach of the Th.M. Rassias theorem in which he replaced the factor $\|a\|^r + \|b\|^r$ by $\|a\|^r.\|b\|^r$ for $r \in \mathbb{R}$ with $r \neq 1$ refers to a significant advancement in the study of stability problems for functional equations. The stability of various functional equations has been extensively investigated by

*Talker
Email addresses: zohreh.heydarpour@pnu.ac.ir (Zohreh Heydarpour), j.izadi@pnu.ac.ir (Javad Izadi)

numerous researchers for different equations on various spaces such as Banach algebras, orthogonally algebras, and various types of Hilbert space (for example, see [5, 7, 8, 19]).

C. Park [12] defined a class of mathematical inequalities as additive ρ-functional inequalities, and demonstrated the Hyers-Ulam stability of additive ρ-functional inequalities within the framework of both Banach spaces and non-Archimedean Banach spaces. This concept brought researchers towards this type of equation, leading to the introduction of various forms of such equations that in recent years, their stability being proven. For more information see [6, 11, 15].

In 1980, G. Maksa [9] introduced the concept of bi-additive mapping, after seven years he introduced the concept of bi-derivations [10]. So that many researchers worked on stability for these concepts, including ternary bi-derivations, ternary bi-homomorphisms, etc. For more information, refer to [2, 4, 13, 18]. In the next, busing concept of additive s-functional inequality and bi-additive mapping, we explain about the bi-additive s-functional inequality on algebras, where s is a fixed nonzero complex number with $|s| < 1$.
Let $\mathfrak{B}$ be an algebra and let $f : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ be a mapping. If f satisfies

$$\begin{aligned}\Big\| f(x+y, z-w) + f(x-y, z+w) - 2f(x,z) + 2f(y,w) \Big\| \\ \leq \Big\| s\Big(f(x-y, z+w) + f(x+y, z-w) - 2f(x,z) + 2f(-y,-w) \Big) \Big\|\end{aligned}$$

then we called f is the bi-additive s-functional inequality where s is a fixed nonzero complex number with $|s| < 1$.
In this paper, we demonstrate the solution to bi-additive s-functional inequality and showing that bi-additive s-functional inequality can be a bi-additive mapping. After that, we investigate the bi-additive s-functional inequality, which can be bi $\mathbb{C}-$linear between algebras. In the end, by using the control function of Th.M. Rassias and J.M. Rassias, we prove the Hyers–Ulam stability of the bi-additive s-functional inequality, where s is a fixed nonzero complex number with $|s| < 1$.

2. Hyers–Ulam stability of the bi-additive s-functional inequality

In this section, Let $\mathfrak{B}$ be algebra and s is a fixed nonzero complex number with $|s| < 1$, and $\kappa_1, \kappa_2 \in T^1_{\frac{1}{n_0}} := \{e^{i\theta}; \quad 0 \leq \theta \leq \frac{2\pi}{n_0}\}$.
In the next lemma, we solve the bi-additive s-functional inequality and show it is a bi-additive mapping.

Lemma 2.1. *If the mapping $f : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ satisfies*

$$\begin{aligned}\Big\| f(x+y, z-w) + f(x-y, z+w) - 2f(x,z) + 2f(y,w)) \Big\| \\ \leq \Big\| s\Big(f(x-y, z+w) + f(x+y, z-w) - 2f(x,z) + 2f(-y,-w) \Big) \Big\|\end{aligned} \tag{1}$$

for all $x, y, z, w \in \mathfrak{B}$, then f is a bi-additive mapping.

Proof. The mapping f satisfies the equation (1), in fact, by putting $x = y = z = w = 0$ in (1), we have $f(0) = 0$. Then, by putting $y = -y$ and $w = -w$ in (1), we have

$$\begin{aligned}\| f(x-y, z+w) + f(x+y, z-w) - 2f(x,z) + 2f(-y,-w) \| \\ \leq \Big\| \Big(f(x+y, z-w) + f(x-y, z+w) - 2f(x,z) - 2f(y,w) \Big) \Big\|\end{aligned} \tag{2}$$

By attention to, (1) and (2), we get

$$\begin{aligned}\| f(x+y, z-w) + f(x-y, z+w) - 2f(x,z) - 2f(y,w) \| \\ \leq \Big\| s^2 \Big(f(x+y, z-w) + f(x-y, z+w) - 2f(x,z) + 2f(y,w) \Big) \Big\|\end{aligned} \tag{3}$$

So, $f(x+y,z-w)+f(x-y,z+w)-2f(x,z)-2f(y,w)=0$, for all $x,y,z,w\in\mathfrak{B}$. By following the same reasoning as in the proof of Theorem 2.1 in [14], f is a bi-additive mapping. □

In the following lemma, if $\kappa_1,\kappa_2\in T^1_{\frac{1}{n_0}}$, we investigate the mapping f is a bi $\mathbb{C}$-linear mapping.

Lemma 2.2. *If the mapping $f:\mathfrak{B}\times\mathfrak{B}\to\mathfrak{B}$ satisfies*

$$\begin{aligned}&\|f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)+f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(y,w))\|\\&\leq\left\|s\Big(f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)+f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(-y,-w)\Big)\right\|\end{aligned}\tag{4}$$

for all $x,y,z,w\in\mathfrak{B}$, then f is a bi $\mathbb{C}$-linear mapping.

Proof. By attention to, lemma 2.1, the mapping f is a bi-additive mapping. Now, by putting $y=w=0$ (4), we get

$$f(\kappa_1x,\kappa_2z)=\kappa_1\kappa_2f(x,z),\qquad\forall\,x,z\in\ \mathfrak{B}.\tag{5}$$

By following the same reasoning as in the proof of Theorem 2.2 in [14] and Theorem 3.2 in [1], f is a $\mathbb{C}-$linear mapping. □

In the following of this section, let f is odd mapping.
In the next, by using the control function of Th. M. Rassias, we investigate the stability of the bi-additive s-functional inequality on algebras.

Theorem 2.3. *Let $r<2$ and ε are nonegative real numbers. If the mapping $f:\mathfrak{B}\times\mathfrak{B}\to\mathfrak{B}$ satisfies*

$$\begin{aligned}&\Big\|f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)+f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(y,w)\Big\|\\&\leq\left\|s\Big(f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)+f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(-y,-w)\Big)\right\|\\&+\varepsilon(\|x\|^r+\|y\|^r+\|z\|^r+\|w\|^r),\qquad\forall\,x,y,z,w\in\ \mathfrak{B}.\end{aligned}\tag{6}$$

Then there is a unique bi-additive mapping $J:\mathfrak{B}\times\mathfrak{B}\to\mathfrak{B}$ such that

$$\|f(x,y)-J(x,y)\|\leq\frac{6\varepsilon(1-s)}{4-2^r}(\|x\|^r+\|y\|^r)+\frac{4}{3}\|f(0,0)\|,$$

where s is a fixed nonzero complex number with $|s|<1$. The mapping J is given by

$$J(x,y):=\lim_{n\to\infty}\frac{1}{4^n}f\left(2^nx,2^ny\right).$$

In the following two theorems, by using the control function of J. M. Rassias, we prove the stability of the bi-additive s-functional inequality on algebras.

Theorem 2.4. *Let r and ε be positive real numbers with $r<\frac{1}{2}$, and let the mapping $f:\mathfrak{B}\times\mathfrak{B}\to\mathfrak{B}$ satisfies*

$$\begin{aligned}&\Big\|f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)+f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(y,w)\Big\|\\&\leq\left\|s\Big(f(\kappa_1x-\kappa_1y,\kappa_2z+\kappa_2w)+f(\kappa_1x+\kappa_1y,\kappa_2z-\kappa_2w)-2\kappa_1\kappa_2f(x,z)+2\kappa_1\kappa_2f(-y,-w)\Big)\right\|\\&+\varepsilon(\|x\|^r\cdot\|y\|^r\cdot\|z\|^r\cdot\|w\|^r),\qquad\forall\,x,y,z,w\in\ \mathfrak{B}.\end{aligned}\tag{7}$$

Then there is a unique bi-additive mapping $J:\mathfrak{B}\times\mathfrak{B}\to\mathfrak{B}$ such that

$$\|f(x,y)-J(x,y)\|\leq\frac{2\varepsilon(1-s)}{4-2^{4r}}\|x\|^{2r}\cdot\|y\|^{2r}+\frac{4}{3}\|f(0,0)\|,\qquad\forall\,x,y\in\ \mathfrak{B}.$$

Theorem 2.5. *Let r and ε be positive real numbers with $r > \frac{3}{2}$, and let $f : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ be a mapping satisfying $f(0,0) = 0$ and (7). Then there is a unique bi-additive mapping $J : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ such that*

$$\|f(x,y) - J(x,y)\| \leq \frac{2\varepsilon(1-s)}{2^{4r}-4}\|x\|^{2r} \cdot \|y\|^{2r}, \qquad \forall\, x, y \in \mathfrak{B}.$$

In the next corollary, we investigate Hyers-Ulam stability of the bi-additive s-functional inequality on algebras.

Corollary 2.6. *If the mapping $f : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ satisfies*

$$\begin{aligned} &\Big\|f(\kappa_1 x + \kappa_1 y, \kappa_2 z - \kappa_2 w) + f(\kappa_1 x - \kappa_1 y, \kappa_2 z + \kappa_2 w) - 2\kappa_1\kappa_2 f(x,z) + 2\kappa_1\kappa_2 f(y,w)\Big\| \\ &\leq \Big\|s\Big(f(\kappa_1 x - \kappa_1 y, \kappa_2 z + \kappa_2 w) + f(\kappa_1 x + \kappa_1 y, \kappa_2 z - \kappa_2 w) - 2\kappa_1\kappa_2 f(x,z) + 2\kappa_1\kappa_2 f(-y,-w)\Big)\Big\| + \varepsilon, \end{aligned} \tag{8}$$

for all $x, y, z, w \in \mathfrak{B}$, Then there is a unique bi-additive mapping $J : \mathfrak{B} \times \mathfrak{B} \to \mathfrak{B}$ such that

$$\|f(x,y) - J(x,y)\| \leq 2(1-s)\varepsilon + \|f(0,0)\|$$

where s is a fixed nonzero complex number with $|s| < 1$ and for all $x, y \in \mathfrak{B}$.

3. Conclusions

In this paper, we introduced a concept of the bi-additive s-functional inequality, where s is a fixed nonzero complex number with $|s| < 1$, on algebras such that we showed the bi-additive s-functional inequality is a bi-additive mapping. After that, we proved the bi-additive s-functional inequality is a bi $\mathbb{C}-$linear between algebras. Finally, by using the control function of Th. M. Rassias, J. M. Rassias and Hyers, we investigated the stability of the bi-additive s-functional inequality, where s is a fixed nonzero complex number with $|s| < 1$.

References

[1] M. Eshaghi Gordji and A. Fazeli, Stability and superstability of homomorphisms on C*-ternary algebras, An. St. Univ. Ovidius Constanta, **20** (2012), 173–188.

[2] M. Eshaghi Gordji, V. Keshavarz, C. Park and J. R. Lee, Approximate ternary Jordan bi-derivations on Banach Lie triple systems, J. Comput. Anal. Appl., **21** (2017), 45-51.

[3] D. H. Hyers, On the stability of the linear functional equation, Proceedings of the National Academy of Sciences of the United States of America, **27** (1941), 222–224.

[4] S. Jahedi, V. Keshavarz, C. Park and S. Yun, Stability of ternary Jordan bi-derivations on C^*-ternary algebras for bi-Jensen functional equation, J. Comput. Anal. Appl., **26** (2019), 140-145.

[5] V. Keshavarz, M. T. Heydari and D. R. Anderson, Hyers–Ulam stabilities for mth differential operators on H^2_β, Chaos, Solitons & Fractals, **179** (2024). https://doi.org/10.1016/j.chaos.2023.114443

[6] V. Keshavarz and S. Jahedi, Hyperstability of orthogonally 3-lie homomorphism: an orthogonally fixed point approach, Approximation and Computation in Science and Engineering, In: Daras, N.J., Rassias, T.M. (eds) Approximation and Computation in Science and Engineering. Springer Optimization and Its Applications, **180**. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-84122-5_25

[7] S. Khan, K.Shah, A. Debbouche, S. Zeb and V. Antonov, Solvability and Ulam-Hyers stability analysis for nonlinear piecewise fractional cancer dynamic systems, Physica Scripta, **99** (2) (2024).

[8] B. Kumar and M. Malik, Existence, controllability and Hyers–Ulam stability of a hybrid neutral switched system with impulsive effects, International Journal of Systems Science **55** (3) (2024), 517-534.

[9] G. Maksa, A remark on symmetric biadditive function having nonnegative diagonalization, Glas. Mat. Ser. III **46** (1980), 279–282.

[10] G. Maksa, On the trace of symmetric bi-derivations, C. R. Math. Rep. Acad. Sci. Canada **9** (1987), 303–307.

[11] Q. Liu, S. Zhuang, and Y. Li, Additive double ρ-functional inequalities in β-homogeneous F-spaces, J. Math. Inequal, **15** (2021), 605-613.

[12] C. Park, Additive ρ-functional inequalities in non-Archimedean normed spaces, J. Math. Inequal., **9** (2015), 397–407.

[13] C. Park, Biderivations and bihomomorphisms in Banach algebras, Filomat, **33** (8) (2019), 2317–2328.

[14] W.-G. Park, J.-H. Bae, Stability of a Bi-additive functional equation in Banach modules over a C^*-algebra, Discrete Dynamics in Nature and Society, **2012**, Article ID 835893, 12 pages, (2012). https://doi.org/10.1155/2012/835893

[15] C. Park, J. R. Lee and X. Zhang, Additive s-functional inequality and hom-derivations in Banach algebras, Journal of Fixed Point Theory and Applications, **21** (2019), 1-14.

[16] J.M. Rassias, On approximation of approximately linear mappings by linear mappings, J. Funct. Anal., **46** (1982), 126–130.

[17] Th. M. Rassias, On the stability of the linear mapping in Banach spaces, Proc. Amer.Math. Soc., **72** (1978), 297–300.

[18] T.L. Shateri, Bi-additive s-functional inequalities and biderivation in modular spaces, The Journal of Analysis, **29** (4) (2021), 1279-1295.

[19] A. P. Selvam and V. Govindaraj, Investigation of controllability and stability of fractional dynamical systems with delay in control, Mathematics and Computers in Simulation, **220** (2024), 89-104.

[20] S. M. Ulam , Problems in Modern Mathematics, Chapter IV, Science edn. Wiley, New York, 1960.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–237, pp. 133–141

On a class of Kirchhoff type problem involving $p(x)$-biharmonic operator

Maryam Mirzapour[a,*]

[a] *Department of Mathematics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran*

Article Info

Keywords:
Kirchhoff type problem
$p(x)$-Biharmonic
Variable exponent

2020 MSC:
35J60
35B30

Abstract

In this paper, we study the following Kirchhoff type problem with Navier boundary conditions

$$M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}dx\Big)\Delta(|\Delta u|^{p(x)-2}\Delta u) = \lambda|u|^{q(x)-2}u + \mu|u|^{\gamma(x)-2}u \quad \text{in } \Omega,$$
$$u = \Delta u = 0 \quad \text{on } \partial\Omega.$$

where Ω is a bounded domain in $\mathbb{R}^N$ with smooth boundary $\partial\Omega$, $N \geq 1$. $M : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous function, $p(x)$, $q(x)$ and $\gamma(x)$ are continuous functions on $\overline{\Omega}$, λ and μ are parameters. Using variational methods, we establish some existence and non-existence results of solutions for this problem.

1. Introduction

In recent years, the study of differential equations and variational problems with $p(x)$-growth conditions was an interesting topic, which arises from nonlinear electrorheological fluids and elastic mechanics. In that context we refer the reader to Ruzicka [24], Zhikov [31] and the reference therein and also see [13, 15, 16, 18].
Fourth order equations appears in many context. Some of theses problems come from different areas of applied mathematics and physics such as Micro Electro-Mechanical systems, surface diffusion on solids, flow in Hele-Shaw cells (see [19]). In addition, this type of equations can describe the static from change of beam or the sport of rigid body.
El Amrouss et al. [12] studied a class of $p(x)$-biharmonic of the form

$$\Delta(|\Delta u|^{p(x)-2}\Delta u) = \lambda|u|^{p(x)-2}u + f(x,u) \quad \text{in } \Omega,$$
$$u = \Delta u = 0 \quad \text{on } \partial\Omega,$$

where Ω is a bounded domains in $\mathbb{R}^N$, with smooth boundary $\partial\Omega$, $N \geq 1$, $\lambda \leq 0$ and some assumptions on the Caratheodory function $f : \Omega \times \mathbb{R} \to \mathbb{R}$. They obtained the existence and multiplicity of solutions.

---

*Talker
Email address: m.mirzapour@cfu.ac.ir (Maryam Mirzapour)

In [1] G.A. Afrouzi et al. have considered problem

$$\begin{cases} M\left(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\right)\Delta(|\Delta u|^{p(x)-2}\Delta u) = f(x,u) & \text{in } \Omega, \\ u = \Delta u = 0 & \text{on } \partial\Omega, \end{cases}$$

in two cases when the Carathéodory function $f(x,u)$ having special structure. Using variational methods, they have established the existence and multiplicity of solutions of the problem. Moreover, we refer the reader to [2, 25, 27], in which, by variational approaches some existence results are given.

The aim of the present paper is to study the existence and multiplicity of weak solutions of the following fourth order elliptic equation with Navier boundary conditions

$$\begin{cases} M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}dx\Big)\Delta(|\Delta u|^{p(x)-2}\Delta u) = \lambda|u|^{q(x)-2}u + \mu|u|^{\gamma(x)-2}u \quad \text{in } \Omega, \\ u = \Delta u = 0 \quad \text{on } \partial\Omega. \end{cases} \tag{1}$$

where Ω is a bounded domain in $\mathbb{R}^N$ with smooth boundary $\partial\Omega$, $N \geq 1$, $M : \mathbb{R}^+ \to \mathbb{R}^+$, $p(x), q(x)$ and $\gamma(x)$ are continuous functions on $\overline{\Omega}$ with $\inf_{x\in\overline{\Omega}} p(x) > 1, \inf_{x\in\overline{\Omega}} q(x) > 1$, $\inf_{x\in\overline{\Omega}} \gamma(x) > 1$ and λ and μ are parameters. Throughout the paper, we assume that $\lambda^2 + \mu^2 \neq 0$.

(1) is called a nonlocal problem because of the presence of the term M, which implies that the equation in (1) is no longer pointwise identities. This provokes some mathematical difficulties which make the study of such a problem particularly interesting. Nonlocal differential equations are also called Kirchhoff-type equations because Kirchhoff [20] has investigated an equation of the form

$$\rho\frac{\partial^2 u}{\partial t^2} - \left(\frac{\rho_0}{h} + \frac{E}{2L}\int_0^L \left|\frac{\partial u}{\partial x}\right|^2 dx\right)\frac{\partial^2 u}{\partial x^2} = 0, \tag{2}$$

which extends the classical D'Alembert's wave equation, by considering the effect of the changing in the length of the string during the vibration. A distinguishing feature of Eq. (2) is that the equation contains a nonlocal coefficient $\frac{\rho_0}{h} + \frac{E}{2L}\int_0^L \left|\frac{\partial u}{\partial x}\right|^2 dx$ which depends on the average $\frac{1}{2L}\int_0^L \left|\frac{\partial u}{\partial x}\right|^2 dx$, and hence the equation is no longer a pointwise identity. The parameters in (2) have the following meanings: L is the length of the string, h is the area of the cross-section, E is the Young modulus of the material, ρ is the mass density and P_0 is the initial tension. Lions [22] has proposed an abstract framework for the Kirchhoff-type equations. After the work of Lions [22], various equations of Kirchhoff-type have been studied extensively, see e.g. [4]-[11].

2. Notations and preliminaries

To study $p(x)$-Laplacian problems, we need some results on the spaces $L^{p(x)}(\Omega)$ and $W^{k,p(x)}(\Omega)$, and properties of $p(x)$-Laplacian, which we use later.

Let Ω be a bounded domain of $\mathbb{R}^N$, denote

$$C_+(\overline{\Omega}) = \{h(x);\ h(x) \in C(\overline{\Omega}),\ h(x) > 1,\ \forall x \in \overline{\Omega}\};$$

For any $h \in C_+(\overline{\Omega})$, we define

$$h^+ = \max\{h(x);\ x \in \overline{\Omega}\}, \quad h^- = \min\{h(x);\ x \in \overline{\Omega}\};$$

For any $p \in C_+(\overline{\Omega})$, we define the *variable exponent Lebesgue space*

$$L^{p(x)}(\Omega) = \Big\{u;\ u \text{ is a measurable real-valued function such that } \int_\Omega |u(x)|^{p(x)}dx < \infty\Big\},$$

endowed with the so-called *Luxemburg norm*

$$|u|_{p(x)} = \inf\Big\{\mu > 0;\ \int_\Omega |\frac{u(x)}{\mu}|^{p(x)}dx \leq 1\Big\},$$

and $(L^{p(x)}(\Omega), |\cdot|_{p(x)})$ becomes a Banach space.

Proposition 2.1 (See Fan and Zhao [17]). *The space* $(L^{p(x)}(\Omega), |\cdot|_{p(x)})$ *is separable, uniformly convex, reflexive and its conjugate space is* $L^{q(x)}(\Omega)$ *where* $q(x)$ *is the conjugate function of* $p(x)$*, i.e.,*

$$\frac{1}{p(x)} + \frac{1}{q(x)} = 1,$$

for all $x \in \Omega$*. For* $u \in L^{p(x)}(\Omega)$ *and* $v \in L^{q(x)}(\Omega)$*, we have*

$$\Big|\int_\Omega uvdx\Big| \leq \Big(\frac{1}{p^-} + \frac{1}{q^-}\Big)|u|_{p(x)}|v|_{q(x)} \leq 2|u|_{p(x)}|v|_{q(x)}.$$

The Sobolev space with variable exponent $W^{k,p(x)}(\Omega)$ is defined as

$$W^{k,p(x)}(\Omega) = \{u \in L^{p(x)}(\Omega) : \ D^\alpha u \in L^{p(x)}(\Omega), \ |\alpha| \leq k\},$$

where $D^\alpha u = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\ldots\partial x_N^{\alpha_N}} u$, with $\alpha = (\alpha_1, \ldots, \alpha_N)$ is a multi-index and $|\alpha| = \sum_{i=1}^N \alpha_i$. The space $W^{k,p(x)}(\Omega)$ equipped with the norm

$$\|u\|_{k,p(x)} = \sum_{|\alpha|\leq k} |D^\alpha u|_{p(x)},$$

also becomes a separable and reflexive Banach space. For more details, we refer the reader to [14, 17, 23, 28]. Denote

$$p_k^*(x) = \begin{cases} \frac{Np(x)}{N-kp(x)} & \text{if} \quad kp(x) < N, \\ +\infty & \text{if} \quad kp(x) \geq N \end{cases}$$

for any $x \in \overline{\Omega}$, $k \geq 1$.

Proposition 2.2 (See Fan and Zhao [17]). *For* $p, r \in C_+(\overline{\Omega})$ *such that* $r(x) \leq p_k^*(x)$ *for all* $x \in \overline{\Omega}$*, there is a continuous embedding*

$$W^{k,p(x)}(\Omega) \hookrightarrow L^{r(x)}(\Omega).$$

If we replace $\leq$ *with* $<$*, the embedding is compact.*

We denote by $W_0^{k,p(x)}(\Omega)$ the closure of $C_0^\infty(\Omega)$ in $W^{k,p(x)}(\Omega)$. Note that the weak solutions of problem (1) are considered in the generalized Sobolev space

$$X = W^{2,p(x)}(\Omega) \cap W_0^{k,p(x)}(\Omega)$$

equipped with the norm

$$\|u\| = \inf\Big\{\mu > 0 : \int_\Omega \Big|\frac{\Delta u(x)}{\mu}\Big|^{p(x)} dx \leq 1\Big\}.$$

Remark 2.3. According to [29], the norm $\|\cdot\|_{2,p(x)}$ is equivalent to the norm $|\Delta\cdot|_{p(x)}$ in the space X. Consequently, the norms $\|\cdot\|_{2,p(x)}$, $\|\cdot\|$ and $|\Delta\cdot|_{p(x)}$ are equivalent.

Proposition 2.4 (See El Amrouss et al. [12]). *If we denote* $\rho(u) = \int_\Omega |\Delta u|^{p(x)}\, dx$*, then for* $u, u_n \in X$*, we have*

$$(1)\|u\| < 1(respectively = 1; > 1) \iff \rho(u) < 1(respectively = 1; > 1);$$
$$(2)\|u\| \leq 1 \Rightarrow \|u\|^{p^+} \leq \rho(u) \leq \|u\|^{p^-};$$
$$(3)\|u\| \geq 1 \Rightarrow \|u\|^{p^-} \leq \rho(u) \leq \|u\|^{p^+};$$
$$(4)\|u\| \to 0\ (respectively \to \infty) \iff \rho(u) \to 0\ (respectively \to \infty).$$

It is clear that the energy functional associated to (1) is defined by

$$I_{\lambda,\mu}(u) = \widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u|^{\gamma(x)}\,dx.$$

where $\widehat{M}(t) = \int_0^t M(\tau)d\tau$. Let us define the functional

$$J(u) = \int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx.$$

It is well known that J is well defined, even and C^1 in X. Moreover, the operator $L = J' : X \to X^*$ defined as

$$\langle L(u), v\rangle = \int_\Omega |\Delta u|^{p(x)-2}\Delta u \Delta v\,dx$$

for all $u, v \in X$ satisfies the following assertions.

Proposition 2.5 (See El Amrouss et al. [12]).

(1) *L is continuous, bounded and strictly monotone.*

(2)*L is a mapping of (S_+) type, namely*

$$u_n \rightharpoonup u \ \textit{and} \ \limsup_{n\to+\infty} L(u_n)(u_n - u) \le 0, \ \textit{implies} \ u_n \to u.$$

(3) *L is a homeomorphism.*

We assume throughout this paper that the Kirchhoff function M satisfies the following hypotheses:

($\mathbf{M_1}$) There exist $m_2 \ge m_1 > 0$ and $\beta \ge \alpha > 1$ such that for all $t \in \mathbb{R}^+$, $m_1 t^{\alpha-1} \le M(t) \le m_2 t^{\beta-1}$.

($\mathbf{M_2}$) For all $t \in \mathbb{R}^+$, $\widehat{M}(t) \ge M(t)t$.

3. Main results and proofs

In this section, we study the existence and non-existence of weak solutions for problem (1). We use the letter c_i in order to denote a positive constant.

Theorem 3.1. *Assume that $q(x), \gamma(x) \in C_+(\overline{\Omega})$, $(p^+)^\alpha < q^- \le q(x) < p_2^*(x)$, $\gamma^+ < \alpha p^-$ and $\beta p^+ < q^-$ for any $x \in \overline{\Omega}$. Then we have*

(**i**) *For every $\lambda > 0$, $\mu \in \mathbb{R}$, (1) has a sequence of weak solutions $(\pm u_k)$ such that $I_{\lambda,\mu}(\pm u_k) \to +\infty$ as $k \to +\infty$.*

(**ii**) *For every $\mu > 0$, $\lambda \in \mathbb{R}$, (1) has a sequence of weak solutions $(\pm v_k)$ such that $I_{\lambda,\mu}(\pm v_k) < 0$ and $I_{\lambda,\mu}(\pm v_k) \to 0$ as $k \to +\infty$.*

(**iii**) *For every $\lambda < 0$, $\mu < 0$, (1) has no nontrivial weak solution.*

We will use the following Fountain theorem to prove (**i**) and the Dual of the Fountain theorem to prove (**ii**).

Lemma 3.2 (See Zhao [30]). *Let X be a reflexive and separable Banach space, then there exist $\{e_j\} \subset X$ and $\{e_j^*\} \subset X^*$ such that*

$$X = \overline{span\,\{e_j \ : j = 1, 2, \dots\}}, \qquad X^* = \overline{span\,\{e_j^* \ : j = 1, 2, \dots\}},$$

and

$$\langle e_i, e_j^*\rangle = \begin{cases} 1 & \textit{if} \quad i = j, \\ 0 & \textit{if} \quad i \ne j, \end{cases}$$

We define

$$X_j = \text{ span } \{e_j\}, \qquad Y_k = \bigoplus_{j=1}^{k} X_j, \qquad Z_k = \overline{\bigoplus_{j=k}^{\infty} X_j}. \tag{3}$$

Then we have the following Lemma.

Lemma 3.3 (See El Amrouss et al. [12]). *If* $q(x), \gamma(x) \in C_+(\overline{\Omega}),\ q(x) < p_2^*(x),$ *and* $\gamma(x) < p_2^*(x)$ *for all* $x \in \overline{\Omega}$, *denote*

$$\beta_k = \sup\{|u|_{q(x)};\ \|u\| = 1,\ u \in Z_k\}$$
$$\theta_k = \sup\{|u|_{\gamma(x)};\ \|u\| = 1,\ u \in Z_k\},$$

then $\lim_{k\to\infty} \beta_k = 0,\ \lim_{k\to\infty} \theta_k = 0.$

Lemma 3.4. *(Fountain Theorem, see Willem [26]). Let*

($\mathbf{A1}$) $I \in C^1(X, \mathbb{R})$ *be an even functional, where* $(X, \|\cdot\|)$ *is a separable and reflexive Banach space, the subspaces* X_k, Y_k *and* Z_k *are defined by* (3.2).
If for each $k \in \mathbb{N}$, *there exist* $\rho_k > r_k > 0$ *such that*
($\mathbf{A2}$) $\inf\{I(u)\ :\ u \in Z_k,\ \|u\| = r_k\} \to +\infty$ *as* $k \to +\infty$.
($\mathbf{A3}$) $\max\{I(u)\ :\ u \in Y_k,\ \|u\| = \rho_k\} \leq 0$.
($\mathbf{A4}$) I *satisfies the (PS) condition for every* $c > 0$.

Then I *has an unbounded sequence of critical points.*

Lemma 3.5. *(Dual Fountain Theorem, see Willem [26]). Assume* ($\mathbf{A1}$) *is satisfied and there is* $k_0 > 0$ *so that, for each* $k \geq k_0$, *there exist* $\rho_k > r_k > 0$ *such that*

($\mathbf{B1}$) $a_k = \inf\{I(u)\ :\ u \in Z_k,\ \|u\| = \rho_k\} \geq 0$.
($\mathbf{B2}$) $b_k = \max\{I(u)\ :\ u \in Y_k,\ \|u\| = r_k\} < 0$.
($\mathbf{B3}$) $d_k = \inf\{I(u)\ :\ u \in Z_k,\ \|u\| \leq \rho_k\} \to 0$ *as* $k \to +\infty$.
($\mathbf{B4}$) I *satisfies the* $(PS)_c^*$ *condition for every* $c \in [d_{k_0}, 0)$.

Then I *has a sequence of negative critical values converging to* 0.

Definition 3.6. We say that $I_{\lambda,\mu}$ satisfies the $(PS)_c^*$ condition (with respect to (Y_n)), if any sequence $\{u_{n_j}\} \subset X$ such that $n_j \to +\infty$, $u_{n_j} \in Y_{n_j}$, $I_{\lambda,\mu}(u_{n_j}) \to c$ and $(I_{\lambda,\mu}|_{Y_{n_j}})'(u_{n_j}) \to 0$, contains a subsequence converging to a critical point of $I_{\lambda,\mu}$.

Proof of Theorem 3.1
(**i**) First we verify $I_{\lambda,\mu}$ satisfies the (PS) condition. Suppose that $(u_n) \subset X$ is (PS) sequence, i.e.,

$$|I_{\lambda,\mu}(u_n)| \leq c_9, \quad I'_{\lambda,\mu}(u_n) \to 0 \quad \text{as } n \to \infty.$$

By Propositions 2.2 and 2.1, we know that if we denote

$$\phi(u) = -\lambda \int_\Omega \frac{1}{q(x)} |u|^{q(x)}\, dx, \qquad \psi(u) = -\mu \int_\Omega \frac{1}{\gamma(x)} |u|^{\gamma(x)}\, dx,$$

then they are both weakly continuous and their derivative operators are compact. By Proposition 2.5, we deduce that $I'_{\lambda,\mu} = L + \phi' + \psi'$ is also of type (S_+). Thus it is sufficient to verify that (u_n) is bounded. Assume $\|u_n\| > 1$ for convenience. For n large enough, we have

$$
\begin{aligned}
c_9 + 1 + \|u_n\| &\geq I_{\lambda,\mu}(u_n) - \frac{1}{q^-}\langle I'_{\lambda,\mu}(u_n), u_n\rangle \\
&= \Big[\widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_n|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u_n|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u_n|^{\gamma(x)}\,dx\Big] \\
&\quad - \frac{1}{q^-}\Big[M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_n|^{p(x)}dx\Big)\int_\Omega \frac{1}{p(x)}|\Delta u_n|^{p(x)}\,dx - \lambda\int_\Omega |u_n|^{q(x)}\,dx \\
&\quad - \mu\int_\Omega |u_n|^{\gamma(x)}\,dx\Big] \\
&\geq \Big(\frac{1}{p^+} - \frac{1}{q^-}\Big)M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_n|^{p(x)}\,dx\Big)\int_\Omega |\Delta u_n|^{p(x)}\,dx - \lambda\int_\Omega |u_n|^{q(x)}\,dx \\
&\quad - \mu\int_\Omega |u_n|^{\gamma(x)}\,dx\Big] \\
&\geq \Big(\frac{1}{p^+} - \frac{1}{q^-}\Big)\frac{m_1}{(p^+)^{\alpha-1}}\Big(\int_\Omega |\Delta u_n|^{p(x)}\,dx\Big)^\alpha - \lambda\int_\Omega |u_n|^{q(x)}\,dx - \mu\int_\Omega |u_n|^{\gamma(x)}\,dx \\
&\geq \Big(\frac{1}{p^+} - \frac{1}{q^-}\Big)\frac{m_1}{(p^+)^{\alpha-1}}\|u_n\|^{\alpha p^-} - c_{10}\|u_n\|^{\gamma^+}.
\end{aligned}
\tag{4}
$$

Since $q^- > p^+$ and $\alpha p^- > \gamma^+$, we know that $\{u_n\}$ is bounded in X. In the following we will prove that if k is large enough, then there exist $\rho_k > r_k > 0$ such that (**A2**) and (**A3**) hold.
(**A2**) For any $u \in Z_k,\ \|u\| = r_k > 1$ (r_k will be specified below), we have

$$
\begin{aligned}
I_{\lambda,\mu}(u) &= \widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_n|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u|^{\gamma(x)}\,dx \\
&\geq \frac{m_1}{\alpha(p^+)^\alpha}\Big(\int_\Omega |\Delta u|^{p(x)}\,dx\Big)^\alpha - \frac{\lambda}{q^-}\int_\Omega |u|^{q(x)}\,dx - \frac{c_{11}|\mu|}{\gamma^-}\|u\|^{\gamma^+} \\
&\geq \frac{m_1}{\alpha(p^+)^\alpha}\|u\|^{\alpha p^-} - \frac{\lambda}{q^-}\int_\Omega |u|^{q(x)}\,dx - \frac{c_{11}|\mu|}{\gamma^-}\|u\|^{\gamma^+}
\end{aligned}
$$

Since $\alpha p^- > \gamma^+$, there exists $r_0 > 0$ large enough such that $\frac{c_{11}|\mu|}{\gamma^-}\|u\|^{\gamma^+} \leq \frac{m_1}{2(p^+)^\alpha}\|u\|^{\alpha p^-}$ as $r = \|u\| \geq r_0$. If $|u|_{q(x)} \leq 1$ then $\int_\Omega |u|^{q(x)}dx \leq |u|_{q(x)}^{q^-} \leq 1$. However, if $|u|_{q(x)} > 1$ then $\int_\Omega |u|^{q(x)}\,dx \leq |u|_{q(x)}^{q^+} \leq (\beta_k\|u\|)^{q^+}$. So, we conclude that

$$
I_{\lambda,\mu}(u) \geq \begin{cases} \frac{m_1}{2(p^+)^\alpha}\|u\|^{\alpha p^-} - \frac{\lambda c_{12}}{q^-} & \text{if } |u|_{q(x)} \leq 1, \\ \frac{m_1}{2(p^+)^\alpha}\|u\|^{\alpha p^-} - \frac{\lambda}{q^-}(\beta_k\|u\|)^{q^+} & \text{if } |u|_{q(x)} > 1. \end{cases}
$$

$$
\geq \frac{m_1}{2(p^+)^\alpha}\|u\|^{\alpha p^-} - \frac{\lambda}{q^-}(\beta_k\|u\|)^{q^+} - c_{13},
$$

choose $r_k = \Big(\frac{2\lambda}{m_1 q^-}q^+\beta_k^{q^+}\Big)^{\frac{1}{\alpha p^- - q^+}}$, we have

$$
I_{\lambda,\mu}(u) = \frac{m_1}{2}\Big(\frac{1}{(p^+)^\alpha} - \frac{1}{q^+}\Big)r_k^{\alpha p^-} - c_{13} \to \infty \quad \text{as} \quad k \to \infty,
$$

because of $(p^+)^\alpha < q^- \leq q^+$ and $\beta_k \to 0$.

(**A3**) Let $u \in Y_k$ such that $\|u\| = \rho_k > r_k > 1$. Then

$$\begin{aligned} I_{\lambda,\mu}(u) &= \widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u|^{\gamma(x)}\,dx \\ &\le \frac{m_2}{\beta}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big)^\beta - \frac{\lambda}{q^+}\int_\Omega |u|^{q(x)}\,dx + \frac{|\mu|}{\gamma^-}\int_\Omega |u|^{\gamma(x)}\,dx \\ &\le \frac{m_2}{\beta(p^-)^\beta}\|u\|^{\beta p^+} - \frac{\lambda}{q^+}\int_\Omega |u|^{q(x)}\,dx + \frac{|\mu|}{\gamma^-}\int_\Omega |u|^{\gamma(x)}\,dx. \end{aligned}$$

Since $\dim Y_k < \infty$, all norms are equivalent in Y_k, we obtain

$$I_{\lambda,\mu}(u) \le \frac{m_2}{\beta(p^-)^\beta}\|u\|^{\beta p^+} - \frac{\lambda}{q^+}\|u\|^{q^-} + \frac{|\mu|}{\gamma^-}\|u\|^{\gamma^+}.$$

We get that: $I_{\lambda,\mu}(u) \to -\infty$ as $\|u\| \to +\infty$ since $q^- > \beta p^+$ and $\gamma^+ < \alpha p^-$. So (**A2**) holds. From the proof of (**A2**) and (**A3**), we can choose $\rho_k > r_k > 0$. Obviously $I_{\lambda,\mu}$ is even and the proof of (**i**) is completed.
(**ii**) We use the Dual Fountain theorem to prove conclusion (**ii**). Now we prove that there exist $\rho_k > r_k > 0$ such that if k is large enough (**B1**), (**B2**) and (**B3**) are satisfied.
(**B1**) For any $u \in Z_k$ we have

$$\begin{aligned} I_{\lambda,\mu}(u) &= \widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u|^{\gamma(x)}\,dx \\ &\ge \frac{m_1}{\alpha(p^+)^\alpha}\Big(\int_\Omega |\Delta u|^{p(x)}\,dx\Big)^\alpha - \frac{|\lambda|}{q^-}\int_\Omega |u|^{q(x)}\,dx - \frac{\mu}{\gamma^-}\int_\Omega |u|^{\gamma(x)}\,dx \\ &\ge \frac{m_1}{\alpha(p^+)^\alpha}\|u\|^{\alpha p^+} - \frac{c_{14}|\lambda|}{q^-}\|u\|^{q^-} - \frac{\mu}{\gamma^-}\int_\Omega |u|^{\gamma(x)}\,dx \end{aligned}$$

Since $q^- > \alpha p^+$, there exists $\rho_0 > 0$ small enough such that $\frac{c_{14}|\lambda|}{q^-}\|u\|^{q^-} \le \frac{m_1}{2\alpha(p^+)^\alpha}\|u\|^{\alpha p^+}$ as $0 < \rho = \|u\| \le \rho_0$. Then from the proof above, we have

$$I_{\lambda,\mu}(u) \ge \begin{cases} \frac{m_1}{\alpha(p^+)^\alpha}\|u\|^{\alpha p^+} - \frac{\mu c_{15}}{\gamma^-} & \text{if } |u|_{\gamma(x)} \le 1, \\ \frac{m_1}{\alpha(p^+)^\alpha}\|u\|^{p^+} - \frac{\mu}{\gamma^-}(\theta_k\|u\|)^{\gamma^+} & \text{if } |u|_{\gamma(x)} > 1. \end{cases} \tag{5}$$

Choose $\rho_k = \left(\frac{2\mu}{m_1\gamma^-}(p^+)^\alpha\theta_k^{\gamma^+}\right)^{\frac{1}{\alpha p^+ - \gamma^+}}$, then

$$I_{\lambda,\mu}(u) = \frac{m_1}{2(p^+)^\alpha}(\rho_k)^{\alpha p^+} - \frac{m_1}{2(p^+)^\alpha}(\rho_k)^{\alpha p^+} = 0.$$

Since $\alpha p^- > \gamma^+$, $\theta_k \to 0$, we know $\rho_k \to 0$ as $k \to \infty$.
(**B2**) For $u \in Y_k$ with $\|u\| \le 1$, we have

$$\begin{aligned} I_{\lambda,\mu}(u)&\widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u|^{\gamma(x)}\,dx \\ &\le \frac{m_2}{\beta(p^-)^\beta}\Big(\frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big)^\beta + \frac{|\lambda|}{q^-}\int_\Omega |u|^{q(x)}dx - \frac{\mu}{\gamma^+}\int_\Omega |u|^{\gamma(x)}\,dx \\ &\le \frac{m_2}{\beta(p^-)^\beta}\|u\|^{\beta p^-} + \frac{|\lambda|}{q^-}\int_\Omega |u|^{q(x)}\,dx - \frac{\mu}{\gamma^+}\int_\Omega |u|^{\gamma(x)}\,dx. \end{aligned}$$

Since $\dim Y_k = k$, conditions $\gamma^+ < \alpha p^- < \beta p^- < \beta(p^-)^\beta$ and $\beta p^+ < q^-$ imply that there exists a $r_k \in (0, \rho_k)$ such that $I_{\lambda,\mu}(u_n) < 0$ when $\|u\| = r_k$. So we obtain

$$\max_{u\in Y_k,\ \|u\|=r_k} I_{\lambda,\mu}(u) < 0,$$

i.e., (**B2**) is satisfied.
(**B3**) Because $Y_k \cap Z_k \neq \emptyset$ and $r_k < \rho_k$, we have

$$d_k = \inf_{u \in Z_k, \|u\| \le \rho_k} I_{\lambda,\mu}(u) \le b_k = \max_{u \in Y_k, \|u\| = r_k} I_{\lambda,\mu}(u) < 0.$$

From (5) , for $u \in Z_k$, $\|u\| \le \rho_k$ small enough we can write

$$\begin{aligned} I_{\lambda,\mu}(u) &\ge \frac{m_1}{2(p^+)^\alpha}\|u\|^{\alpha p^+} - \frac{\mu}{\gamma^-}\theta_k^{\gamma^+}\|u\|^{\gamma^+} \\ &\ge -\frac{\mu}{\gamma^-}\theta_k^{\gamma^+}\|u\|^{\gamma^+}, \end{aligned}$$

Since $\theta_k \to 0$ and $\rho_k \to 0$ as $k \to \infty$, (**B3**) holds. Finally we verify the $(PS)_c^*$ condition. Suppose $\{u_{n_j}\} \subset X$ such that

$$n_j \to +\infty, \quad u_{n_j} \in Y_{n_j}, \quad I_{\lambda,\mu}(u_{n_j}) \to c_{16} \quad \text{and} \quad (I_{\lambda,\mu}|_{Y_{n_j}})'(u_{n_j}) \to 0.$$

If $\lambda \ge 0$, similar to (4), we can get the boundedness of $\|u_{n_j}\|$. Assume $\|u_{n_j}\| \ge 1$ for convenience. If $\lambda < 0$, for $n > 0$ large enough we have

$$\begin{aligned} c_{16} + 1 + \|u_{n_j}\| &\ge I_{\lambda,\mu}(u_{n_j}) - \frac{1}{q^+}\langle I'_{\lambda,\mu}(u_{n_j}), u_{n_j}\rangle \\ &= \Big[\widehat{M}\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_{n_j}|^{p(x)}\,dx\Big) - \lambda\int_\Omega \frac{1}{q(x)}|u_{n_j}|^{q(x)}\,dx - \mu\int_\Omega \frac{1}{\gamma(x)}|u_{n_j}|^{\gamma(x)}\,dx\Big] \\ &\quad - \frac{1}{q^+}\Big[M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u_{n_j}|^{p(x)}dx\Big)\int_\Omega \frac{1}{p(x)}|\Delta u_{n_j}|^{p(x)}\,dx - \lambda\int_\Omega |u_{n_j}|^{q(x)}\,dx \\ &\quad - \mu\int_\Omega |u_{n_j}|^{\gamma(x)}\,dx\Big] \\ &\ge \Big(\frac{1}{p^+} - \frac{1}{q^+}\Big)\frac{m_1}{(p^+)^{\alpha-1}}\Big(\int_\Omega |\Delta u_{n_j}|^{p(x)}\,dx\Big)^\alpha - \lambda\int_\Omega |u_{n_j}|^{q(x)}\,dx \\ &\quad - \mu\int_\Omega |u_{n_j}|^{\gamma(x)}\,dx\Big] \\ &\ge \Big(\frac{1}{p^+} - \frac{1}{q^+}\Big)\frac{m_1}{(p^+)^{\alpha-1}}\|u_{n_j}\|^{\alpha p^-} - c_{17}\|u_{n_j}\|^{\gamma^+}. \end{aligned}$$

Since $\alpha p^- > \gamma^+$ and $p^+ < (p^+)^\alpha < q^-$, we know that $\{u_{n_j}\}$ is bounded in X. Hence there exists $u \in X$ such that $u_{n_j} \to u$ in x. Observe now that $X = \overline{\cup_{n_j} Y_{n_j}}$, then we can find $v_{n_j} \in Y_{n_j}$ such that $v_{n_j} \to u$. We have

$$\langle I'_{\lambda,\mu}(u_{n_j}), u_{n_j} - u\rangle = \langle I'_{\lambda,\mu}(u_{n_j}), u_{n_j} - v_{n_j}\rangle + \langle I'_{\lambda,\mu}(u_{n_j}), v_{n_j} - u\rangle.$$

Having in mind that $(u_{n_j} - v_{n_j}) \in Y_{n_j}$, it yields

$$\langle I'_{\lambda,\mu}(u_{n_j}), u_{n_j} - u\rangle = \langle (I_{\lambda,\mu}|_{Y_{n_j}})'(u_{n_j}), u_{n_j} - v_{n_j}\rangle + \langle I'_{\lambda,\mu}(u_{n_j}), v_{n_j} - u\rangle \to o \quad \text{as} \quad n \to \infty. \tag{6}$$

By Proposition 2.5, the operator $I'_{\lambda,\mu}$ is obviously of (S_+) type. Using this fact with (6), we deduce that $u_{n_j} \to u$ in X, furthermore $I'_{\lambda,\mu}(u_{n_j}) \to I'_{\lambda,\mu}(u)$.
We claim now that u is in fact a critical point of $I_{\lambda,\mu}$. Taking $\omega_k \in Y_k$, notice that when $n_j \ge k$ we have

$$\begin{aligned} \langle I'_{\lambda,\mu}(u), \omega_k\rangle &= \langle I'_{\lambda,\mu}(u) - I'_{\lambda,\mu}(u_{n_j}), \omega_k\rangle + \langle I'_{\lambda,\mu}(u_{n_j}), \omega_k\rangle \\ &= \langle I'_{\lambda,\mu}(u) - I'_{\lambda,\mu}(u_{n_j}), \omega_k\rangle + \Big\langle (I_{\lambda,\mu}|_{Y_{n_j}})'(u_{n_j}), \omega_k\Big\rangle. \end{aligned}$$

Going to the limit on the right side of the above equation reaches

$$\langle I'_{\lambda,\mu}(u), \omega_k\rangle = 0, \qquad \forall \omega_k \in Y_k,$$

so $I'_{\lambda,\mu}(u) = 0$, this show that $I_{\lambda,\mu}$ satisfies the $(PS)^*_c$ condition for every $c \in \mathbb{R}$.
(**iii**) Assume for the sake of contradiction, $u \in X\backslash\{0\}$ is a weak solution of problem (1). Then multiplying the equation in (1) by u, integrating by parts we get

$$M\Big(\int_\Omega \frac{1}{p(x)}|\Delta u|^{p(x)}\,dx\Big)\int_\Omega |\Delta u|^{p(x)}\,dx = \lambda\int_\Omega |u|^{q(x)}\,dx + \mu\int_\Omega |u|^{\gamma(x)}\,dx.$$

This leads to contradiction and the proof is complete.

References

[1] G.A.Afrouzi, M. Mirzapour, N.T. Chung, Existence and multiplicity of solutions for Kirchhoff type problems involving $p(x)$-biharmonic operators, *Z. Anal. Anwend.*, **33** no. 3, (2014), 289-303.
[2] Z.E. Allali, M.K. Hamdani, S. Taarabti, Three solutions to a Neumann boundary value problem driven by p(x)-biharmonic operator, *J. Elliptic Parabol. Equ.*, (2024), https://doi.org/10.1007/s41808-023-00257-1.
[3] A. Ambrosetti, P.H. Rabinowitz, Dual variational methods in critical points theory and applications, *J. Funct. Anal.*, **14** (1973), 349-381.
[4] A. Arosio, S. Panizzi, On the well-posedness of the Kirchhoff string, *Trans. Amer. Math. Soc.*, **348** (1996), 305-330.
[5] M.M. Cavalcanti, V.N. Domingos Cavalcanti, J.A. Soriano, Global existence and uniform decay rates for the Kirchhoff-Carrier equation with nonlinear dissipation, *Adv. Differential Equations*, **6** (2001), 701-730.
[6] F.J.S.A. Corrêa, G.M. Figueiredo, On an elliptic equation of p-Kirchhoff type via variational methods, *Bull. Aust. Math. Soc.*, **74** (2006), 263-277.
[7] F.J.S.A. Corrêa, G.M. Figueiredo, On a p-Kirchhoff equation via Krasnoselskii's genus, *Appl. Math. Letters*, **22** (2009), 819-822.
[8] G. Dai, R. Hao, Existence of solutions for a $p(x)$-Kirchhoff-type equation, *J. Math. Anal. Appl.*, **359** (2009), 275-284.
[9] G. Dai, R. Ma, Solutions for a $p(x)$-Kirchhoff-type equation with Neumann boundary data, *Nonlinear Anal.*, **12** (2011), 2666-2680.
[10] M. Dreher, The Kirchhoff equation for the p-Laplacian, *Rend. Semin. Mat. Univ. Politec. Torino*, **64** (2006), 217-238.
[11] M. Dreher, The wave equation for the p-Laplacian, *Hokkaido Math. J.*, **36** (2007), 21-52.
[12] A. El Amrouss, F. Moradi, M. Moussaoui, Existence of solutions for fourth-order PDEs with variable exponentsns, *Electron. J. Differ. Equ.*, **153** (2009), 1-13.
[13] A. El Hamidi, Existence Results to Elliptic Systems with Nonstandard Growth Conditions *J. Math. Anal. Appl.*, **300** (2004), 30-42.
[14] X.L. Fan, X. Fan, A Knobloch-type result for p(t) Laplacian systems, *J. Math. Anal. Appl.*, **282** (2003), 453-464.
[15] X.L. Fan, X.Y. Han, Existence and multiplicity of solutions for $p(x)$-Laplacian equations in $\mathbb{R}^N$, *Nonlinear Anal. T.M.A.*, **59** (2004), 173-188.
[16] X.L. Fan, Q.H. Zhang, Existence of solutions for $p(x)$-Laplacian Dirichlet problems, *Nonlinear Anal.T.M.A.*, **52** (2003), 1843-1852.
[17] X.L. Fan, D. Zhao, On the spaces $L^{p(x)}$ and $W^{m,p(x)}$, *J. Math. Anal. Appl.*, **263** (2001), 424-446.
[18] X.L. Fan, Solutions for $p(x)$-Laplacian Dirichlet problems with singular coefficients, *J. Math. Anal. Appl.*, **312** (2005), 464-477.
[19] A. Ferrero, G. Warnault, On a solutions of second and fourth order elliptic with power type nonlinearities, *Nonlinear Anal. T.M.A.*, **70** (2009), 2889-2902.
[20] G. Kirchhoff, *Mechanik*, Teubner, Leipzig, 1883.
[21] L. Li, C.L. Tang, Existence and multiplicity of solutions for a class of $p(x)$-Biharmonic equations, *Acta Mathematica Scientia*, **33** (2013), 155-170.
[22] J.L. Lions, On some questions in boundary value problems of mathematical physics, *in: Proceedings of international Symposium on Continuum Mechanics and Partial Differential Equations, Rio de Janeiro 1977, in: de la Penha, Medeiros (Eds.), Math. Stud., North-Holland*, **30** (1978), 284-346.
[23] M. MihĂilescu, Existence and multiplicity of solutions for a Neumann problem involving the p(x)-Laplace operator, *Nonlinear Anal. T.M.A.*, **67** (2007), 1419-1425.
[24] M. Ruzicka, Electrorheological fluids: modeling and mathematical theory, *Lecture Notes in Mathematics 1748, Springer-Verlag, Berlin*, 2000.
[25] S. Taarabti, Z. El Allali, K. B. Hadddouch , On the p(X)-Kirchhoff-Type Equation Involving the p(X)-Biharmonic Operator via the Genus Theory , *Ukr. Math. J.*, **72** (2020), https://doi.org/10.1007/s11253-020-01836-4.
[26] M. Willem, *Minimax Theorems*, Birkhäuser, Boston, 1996.
[27] Y. Wu, S. Taarabti, Z. El Allali, K. Ben Hadddouch, J. Zuo, A Class of Fourth-Order Symmetrical Kirchhoff Type Systems, *Symmetry*, **14** (2022), https://doi.org/10.3390/sym14081630.
[28] J. Yao, Solutions for Neumann boundary value problems involving $p(x)$-Laplace operators, *Nonlinear Anal. T.M.A.*, **68** (2008), 1271-1283.
[29] A. Zang, Y. Fu, Interpolation inequalities for derivatives in variable exponent Lebesgue- Sobolev spaces, *Nonlinear Anal. T.M.A.*, **69** (2008), 3629-3636.
[30] J.F. Zhao, Structure Theory of Banach Spaces, *Wuhan University Press, Wuhan* , (1991) (in Chinese).
[31] V.V. Zhikov, Averaging of functionals of the calculus of variations and elasticity theory, *Math. USSR. Izv.*, **9** (1987), 33-66.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–251, pp. 142–145

Existence of Fixed Point Theorems for F-Contractive Type Fuzzy Mapping

Parastoo Heiatian Naeini[a,*]

[a] *Department of Mmathematics, Payame Noor University , P. O. Box 19395-4697, Tehran, Iran*

Article Info

Keywords:
fixed point
controlled metric space of type (γ, β)
extended b-metric space

2020 MSC:
37C25
47H10

Abstract

In this article, inspired by the concepts of $\mathbb{G}$-metric spaces, we introduce the notion of F-contractive type fuzzy mappings in $\mathbb{G}$-metric spaces. Using this new idea, some fixed point theorems are proved.

1. Introduction

Throughout this article, denoted by $\mathbb{R}$, $\mathbb{R}^+$ and $\mathbb{N}$ are the set of all real numbers, positive real numbers and natural numbers, respectively. Also, $(\mathcal{F}; d)$, ($\mathcal{F}$ for short), represents a metric space with the metric d. Fixed point theory is a renowned and huge field of research in mathematical sciences. This field is known as the combination of analysis which includes topology, geometry and algebra. The first most well known result in fixed point theory with metric space structure is the Banach fixed point theorem (which is also called the contraction map- ping principle). In the literature, there are several extensions of the Banach contraction principle, which states that every self mapping S defined on a complete metric space $(\mathcal{F}; d)$ satisfying for all $\tau, \sigma \in \mathcal{F}$, $d(S\tau, S\sigma) \leq \mathcal{K}d(\tau, \sigma)$; where $\mathcal{K} \in (0, 1)$; has a unique fixed point. Some improvements of the Banach fixed point theorem concern the contrac- tive inequality while others deal with generalizing the space. A particular extension of metric space is the so-called $\mathbb{G}$-metric space initiated by Mustafa and Sims [1] in 2006. In the first paper on $\mathbb{G}$-metric spaces, Sims and Mustafa [1] introduced some propertiesof $\mathbb{G}$-metric spaces and also discussed its topology, compactness, completeness, product and the criteria regarding the convergence and continuity of sequences in $\mathbb{G}$-metric space. Some theorems concerning these properties were also proved. Another famous general- ization of the contraction mapping principle due to Banach was presented by Wardowski [2], the concept of which is called F-contraction. The idea of F-contractions has been extended both for single-valued and set-valued mappings For some comprehensive surveys in this direction, we refer the interested reader to the work of Taskovic or Rhoades. As a natural extension of crisp sets, fuzzy sets was introduced initially by Zadeh.

*Talker
Email address: parastoo.heiatiannaeini@pnu.ac.ir (Parastoo Heiatian Naeini)

After the introduction of this concept, several researches were conducted on various applications and improvements of fuzzy sets in different directions. Along this trend, Heilpern introduced the concept of fuzzy mapping and proved a fixed point theorem for fuzzy contraction mappings which is a generalization of the fixed point theorem for multi-valued mappings of Nadler. Thereafter, other authors have studied the existence of fixed point of fuzzy mappings. The aim of this paper is to establish fixed point theorems, common fixed point theorems for F-contraction type fuzzy mappings in $\mathbb{G}$-metric spaces. Our results generalize and extend a few known results in the comparable literature.
In this following, we recall some basic concepts that are necessary in the establishment of our main results. Most of these preliminaries are recorded from [1, 4].

Definition 1.1. Let $\mathcal{F} \neq \emptyset$; and $\mathbb{G} : \mathcal{F} \times \mathcal{F} \times \mathcal{F} \to \mathbb{R}^+$ be a function such that the following conditions are satisfied:

- $\mathbb{G}1$ $\mathbb{G}(\tau, \sigma, \upsilon) = 0$ if $\tau = \sigma = \upsilon$,
- $\mathbb{G}2$ $\mathbb{G}(\tau, \tau, \sigma) > 0$ for all $\tau, \sigma \in \mathcal{F}$ with $\tau \neq \sigma$,
- $\mathbb{G}3$ $\mathbb{G}(\tau, \tau, \sigma) \leq \mathbb{G}(\tau, \sigma, \upsilon)$ for all $\tau, \sigma, \upsilon \in \mathcal{F}$ with $\upsilon \neq \sigma$,
- $\mathbb{G}4$ $\mathbb{G}(\tau, \sigma, \upsilon) = \mathbb{G}(\tau, \upsilon, \sigma) = \mathbb{G}(\sigma, \upsilon, \tau) = ...$ (symmetric with respect to τ, σ, υ),
- $\mathbb{G}5$ $\mathbb{G}(\tau, \sigma, \upsilon) \leq \mathbb{G}(\tau, a, a) + \mathbb{G}(\sigma, \sigma, \upsilon)$ for all $\tau, \sigma, \upsilon, a \in \mathcal{F}$ (rectangular prop- erty).

Then $\mathbb{G}$ is called a $\mathbb{G} - M$ function and $(\mathcal{F}, \mathbb{G})$ is said to be a $\mathbb{G}$-metric space.

Definition 1.2. Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space. A sequence $\{\tau_e\}$ in $\mathcal{F}$ is $\mathbb{G}$-Convergent sequence if, for any $\delta > 0$; there exists $\tau \in \mathcal{F}$, $O(\delta) \in \mathbb{N}$ such that $\mathbb{G}(\tau, \tau_e, \tau_\rho) < \delta$, for all $e, \rho \geq O(\delta)$. We call the limit of the sequence and write $\tau_e \to \tau$ or $\lim_{e\to\infty} \tau_e = \tau$.

Definition 1.3. Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space. A sequence feg in $\{\tau_e\}$ is called $\mathbb{G}$-Cauchy sequence if, for any $\delta > 0$; there exists $O(\delta) \in \mathbb{N}$ such that $\mathbb{G}(\tau_\zeta, \tau_e, \tau_\rho) < \delta$ for each $e, \rho \geq O(\delta)$, that is, $\mathbb{G}(\tau_\zeta, \tau_e, \tau_\rho) \to o$ as $\rceil, \rho, \zeta \to \infty$.

Definition 1.4. Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space. A sequence $\{\tau_e\}$ in $\mathcal{F}$ is called $\mathbb{G}$-Complete if every $\mathbb{G}$-Cauchy sequence in $(\mathcal{F}; \mathbb{G})$ is convergent in $\mathcal{F}$.

Lemma 1.5. *Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space and $\{\tau_e\}$ be a sequence in $\mathcal{F}$. Then the following statements are equivalent:*

- *(i)* $\{\tau_e\}$ *is $\mathbb{G}$- convergent to τ.*
- *(ii)* $\mathbb{G}(\tau_e, \tau_e, \tau) \to 0$ *as e approaches infinity.*
- *(iii)* $\mathbb{G}(\tau_e, \tau, \tau) \to 0$ *as e approaches infinity.*
- *(iv)* $\mathbb{G}(\tau_e, \tau_\rho, \tau) \to 0$ *as e, ρ approaches infinity.*

Definition 1.6. Kaewcharoen and Kaewkhao introduced the concept of Hausdorff $\mathbb{G}$-distance as follows: Let $\mathcal{F}$ be a $\mathbb{G}$-metric space and $\mathcal{CB}(\mathcal{F})$ be the family of all non empty closed and bounded subsets of $\mathcal{F}$. Then, the Hausdorff $\mathbb{G}$-distance function is defined as follows:

$$\mathcal{H}_{\mathbb{G}}(\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_3) = max\{sup_{\tau\in\mathcal{Z}_1}\mathbb{G}(\tau, \mathcal{Z}_2, \mathcal{Z}_3), sup_{\tau\in\mathcal{Z}_2}\mathbb{G}(\tau, \mathcal{Z}_1, \mathcal{Z}_3), sup_{\tau\in\mathcal{Z}_3}\mathbb{G}(\tau, \mathcal{Z}_1, \mathcal{Z}_2)\},$$

where

$$\mathbb{G}(\tau, \mathcal{Z}_2, \mathcal{Z}_3) = \mu_{\mathbb{G}}(\tau, \mathcal{Z}_2) + \mathbb{G}(\mathcal{Z}_2, \mathcal{Z}_3) + \mu_{\mathbb{G}}(\tau, \mathcal{Z}_3),$$
$$\mathbb{G}(\tau, \mathcal{Z}_2) = inf_{\sigma\in\mathcal{Z}_2}\mu_{\mathbb{G}}(\tau, \sigma),$$
$$\mathbb{G}(\mathcal{Z}_1, \mathcal{Z}_2) = inf_{\tau\in\mathcal{Z}_1,\sigma\in\mathcal{Z}_2}\mu_{\mathbb{G}}(\tau, \sigma).$$

Definition 1.7. [3]. Let $(\mathcal{F}; d$ be a metric space and $\mathcal{Z}_1, \mathcal{Z}_2 \in \mathfrak{Z}(\mathcal{F})$ such that $[\mathcal{Z}_1]_\alpha$ and $[\mathcal{Z}_2]_\alpha$ are compact subsets of $\mathcal{F}$, the following identities are defined as,

$$\begin{aligned} P_\alpha(\mathcal{Z}_1, \mathcal{Z}_2) &= inf_{\tau \in [\mathcal{Z}_1], \sigma \in [\mathcal{Z}_2]_\alpha} \mu(\tau, \sigma), \\ P(\mathcal{Z}_1, \mathcal{Z}_2) &= sup_\alpha P_\alpha(\mathcal{Z}_1, \mathcal{Z}_2), \\ D_\alpha(\mathcal{Z}_1, \mathcal{Z}_2) &= \mu_{\mathcal{H}}([\mathcal{Z}_1]_\alpha, [\mathcal{Z}_2]_\alpha). \end{aligned}$$

Definition 1.8. [3]. Let $(\mathcal{F}; d)$ be a metric space. The distance function $D_\infty : \mathfrak{Z}(\mathcal{F}) \times \mathfrak{Z}(\mathcal{F}) \to \mathbb{R}$ is defined as:

$$\mathcal{D}_\infty(\mathcal{Z}_1, \mathcal{Z}_2) = sup_\alpha P_\alpha(\mathcal{Z}_1, \mathcal{Z}_2).$$

Definition 1.9. Let $(\mathcal{F}; d)$ be a metric space, $\mathcal{T} : \mathcal{F} \to I^{\mathcal{F}}$ and $\mathcal{Q} : \mathcal{F} \to I^{\mathcal{F}}$ be two fuzzy mappings. A point $\nu \in \mathcal{F}$ is called

(i) fuzzy fixed point of $\mathcal{T}$ if $\nu \in [\mathcal{T}_\nu]_\alpha$ for some $\alpha \in [0, 1]$:

(ii) common fuzzy fixed point if $\nu \in [\mathcal{T}_\nu]_\alpha \cap [\mathcal{Q}_\upsilon]_\alpha$.

Definition 1.10. Let $F : \mathbb{R}^+ \to \mathbb{R}$ be a mapping satisfying:

(F1) F is strictly increasing, i.e. for all $\beta, \gamma \in \mathbb{R}_+$ such that $\beta < \gamma$, $F(\beta) < F(\gamma)$.

(F2) For each sequence $\{\beta_n\}_{n\in\mathbb{N}}$ of positive numbers $\lim_{n\to\infty} \beta_n = 0$ if and only if $\lim_{n\to\infty} F(\beta_n) = -\infty$.

(F3) There exists $k \in (0, 1)$ such that $\lim_{\beta\to 0^+} \beta^k F(\beta) = 0$. Subsequently, Altun et al. [4] modified the above definition by adding comprehensive condition (F4) which is stated as:

(F4) $F(infA) = infF(A)$ for all $A \subset (0, \infty)$ with $infA > 0$.

We denote the set of all functions satisfying properties (F1) - (F4) by $\mathcal{X}$.

Definition 1.11. [5]Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space. A mapping $\mathcal{T} : \mathcal{F} \to \mathcal{F}$ is said to be an F-contraction if there exists $\omega > 0$ such that for all $\tau, \sigma, \upsilon \in \mathcal{F}$,

$$\mathbb{G}(\mathcal{T}\tau, \mathcal{T}\sigma, \mathcal{T}\upsilon) > 0 \Rightarrow \omega + F(\mathbb{G}(\mathcal{T}\tau, \mathcal{T}\sigma, \mathcal{T}\upsilon)) \leq F(\mathcal{G}(\tau, \sigma, \upsilon)).$$

Lemma 1.12. *[4] Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space and $\mathcal{Z}_1, \mathcal{Z}_2 \in \mathcal{CB}(\mathcal{F})$, then for each $\tau \in \mathbb{Z}_1$, we have*

$$\mathbb{G}(\tau, \mathcal{Z}_2, \mathcal{Z}_2) \leq \mathcal{H}_{\mathbb{G}}(\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_2).$$

Lemma 1.13. *[4] Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}-$ metric space. If $\mathcal{Z}_1, \mathcal{Z}_2 \in \mathcal{CB}(\mathcal{F})$ and $\tau \in \mathcal{Z}_1$, then for each $\epsilon > 0$ there exists$\sigma \in \mathcal{Z}_2$ s.t.*

$$\mathbb{G}(\tau, \sigma, \sigma) \leq \mathcal{H}_{\mathbb{G}}(\mathcal{Z}_1, \mathcal{Z}_2, \mathcal{Z}_2) + \epsilon.$$

Lemma 1.14. *Let $\mathcal{V}$ be a metric linear space, $\mathcal{T} : \mathcal{F} \to \mathcal{W}(\mathcal{V})$ and $\tau_0 \in \mathcal{V}$. Then there exists $\tau_1 \in \mathcal{V}$ such that $\{\tau_1\} \subset \mathcal{T}(\tau_0)$.*

2. Main Result

We begin this section with some auxiliary concepts as follows.

Definition 2.1. Let $(\mathcal{F}; \mathbb{G})$ be a $\mathbb{G}$-metric space, $\mathcal{F} \in \mathcal{X}$ and $\mathcal{T} : \mathcal{F} \to I^{\mathcal{F}}$ be a fuzzy mapping. Then $\mathcal{T}$ is said to be an $\mathbb{F}$-contractive type fuzzy mapping if there exists $\omega > 0$ such that

$$\omega + \mathcal{H}_{\mathbb{G}}(|\mathcal{T}_{\mathcal{T}}|_\lambda, |\mathcal{T}_\sigma|_\lambda, |\mathcal{T}_\upsilon|_\lambda) \leq F(\mathbb{G}(\tau, \sigma, \upsilon)),$$

for all $\tau, \sigma, \upsilon \in \mathcal{F}$ with $\mathcal{H}_{\mathbb{G}}(|\mathcal{T}_{\mathcal{T}}|_\lambda, |\mathcal{T}_\sigma|_\lambda, |\mathcal{T}_\upsilon|_\lambda) > 0$ and $\lambda \in [0; 1]$.

Theorem 2.2. *Let $(\mathcal{F};\mathbb{G})$ be a $\mathbb{G}$-complete metric space and let $\mathcal{S},\mathcal{T}:\mathcal{F}\to I^{\mathcal{F}}$ IF be fuzzy mapping such that for each $\tau,\sigma\in\mathcal{F}$, there exist $\alpha\in(0;1]$ with $|\mathcal{S}_{\mathcal{T}}|_\lambda, |\mathcal{T}_{\mathcal{T}}|_\lambda\in\mathcal{C}(\mathcal{F})$. Assume there exist some $F\in\mathcal{X}$ and $\tau>0$ such that*

$$\omega+F(\mathcal{H}_{\mathbb{G}}(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{T}_\sigma|_\lambda,|\mathcal{T}_\upsilon|_\lambda))\leq F(\mathbb{G}(\tau,\sigma,\upsilon)),$$

for all $\tau,\sigma,\upsilon\in\mathcal{F}$ with $\mathcal{H}_{\mathbb{G}}(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{T}_\sigma|_\lambda,|\mathcal{T}_\upsilon|_\lambda)>0$ and $\lambda\in[0;1]$. Then S and T have a common fixed point.

Corollary 2.3. *Let $(\mathcal{F};\mathbb{G})$ be a $\mathbb{G}$-complete metric space and let $\mathcal{S}:\mathcal{F}\to I^{\mathcal{F}}$ for each $\tau,\sigma\in\mathcal{F}$, there exist $\lambda\in(0;1]$ such that $[S_\tau]_\lambda\in\mathcal{C}(\mathcal{F})$: Assume there exist some $F\in\mathcal{X}$ and $\tau>0$ such that*

$$\omega+F(\mathcal{H}_{\mathbb{G}}(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{S}_\sigma|_\lambda,|\mathcal{S}_\upsilon|_\lambda))\leq F(\mathbb{G}(\tau,\sigma,\upsilon)),$$

for all $\tau,\sigma,\upsilon\in\mathcal{F}$ with $\mathcal{H}_{\mathbb{G}}(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{S}_\sigma|_\lambda,|\mathcal{S}_\upsilon|_\lambda)>0$. Then S has a fixed point.

Corollary 2.4. *Let $(\mathcal{F};\mathbb{G})$ be a $\mathbb{G}$-complete metric space and let $\mathcal{P}_1,\mathcal{P}_2:\mathcal{F}\to\mathcal{C}(\mathcal{F})$. Suppose that there exist some $F\in\mathcal{X}$ and $\tau>0$ such that*

$$F(\mathcal{H}_{\mathbb{G}}(\mathcal{P}_1\mathcal{T},\mathcal{P}_2\sigma\ ,\mathcal{P}_2\upsilon))\leq F(\mathbb{G}(\tau,\sigma,\upsilon)),$$

for all $\tau,\sigma,\upsilon\in\mathcal{F}$ with $\mathcal{H}_{\mathbb{G}}(\mathcal{P}_1\mathcal{T},\mathcal{P}_2\sigma\ ,\mathcal{P}_2\lambda))>0$. Then $\mathcal{P}_1$ and $\mathcal{P}_2$ have a common fixed point.

Corollary 2.5. *Let $(\mathcal{F};\mathbb{G})$ be a $\mathbb{G}$-complete metric space and let $\mathcal{P}:\mathcal{F}\to\mathcal{C}(\mathcal{F})$. Suppose that there exist some $\mathcal{F}\in\mathcal{X}$ and $\tau>0$ such that*

$$F(\mathcal{H}_{\mathbb{G}}(\mathcal{P}\mathcal{T},\mathcal{P}\sigma\ ,\mathcal{P}\lambda))\leq F(\mathbb{G}(\tau,\sigma,\upsilon)),$$

for all $\tau,\sigma,\upsilon\in\mathcal{F}$ with $\mathcal{H}_{\mathbb{G}}(\mathcal{P}\mathcal{T},\mathcal{P}\sigma\ ,\mathcal{P}\lambda)>0$. Then $\mathcal{P}$ has a common fixed point.

Corollary 2.6. *Let $(\mathcal{F};\mathbb{G})$ be a $\mathbb{G}$-complete metric linear space and $\mathcal{S},\mathcal{T}:\mathcal{F}\to\mathcal{W}(\mathcal{F})$ Suppose that there exist some $\mathcal{F}\in\mathcal{X}$ and $\tau>0$ such that*

$$\omega+F(\mathbb{G}_\infty(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{T}_\sigma|_\lambda,|\mathcal{T}_\upsilon|_\lambda))\leq F(\mathbb{G}(\tau,\sigma,\upsilon)),$$

for all $\tau,\sigma,\upsilon\in\mathcal{F}$ with $\mathbb{G}_\infty(|\mathcal{S}_{\mathcal{T}}|_\lambda,|\mathcal{T}_\sigma|_\lambda,|\mathcal{T}_\upsilon|_\lambda)>0$. There exists $u\in\mathcal{F}$ such that $\{u\}\subset|\mathcal{S}_u|_\lambda$ and $\{u\}\subset|\mathcal{T}_u|_\lambda$.

References

[1] Z. Mustafa, B. Sims, A new approach to generalized metric spaces, J. Nonl. Convex Anal. 7 (2) (2006) 289-297.
[2] D.Wardowski, Fixed points of a new type of contractive mappings in complete metric spaces, Fixed Point Theory Appl. 2012 (1) (2012) 94-100.
[3] A. Azam, Fuzzy fixed points of fuzzy mappings via a rational inequality, Hacettepe J. Math. Stat. 40 (3) (2011) 421-431.
[4] Z. Mustafa, M. Arshad, S.U. Khan, J. Ahmad, M. Jaradat, Common fixed points for multivalued mappings in $\mathcal{G}-$metric spaces with applications. J. Nonlinear Sci. Appl. 10 (2017) 2550-2564
[5] A. Gupta, P.S. Kaurav, Fixed points of a new type of contractive mappings in G- metric spaces, Int. J. Math. And Appl. 7 (4) (2019) 95-100.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–252, pp. 146–149

Existence fixed point theory for two class of generalized nonexpansive mappings

Parastoo Heiatian Naeini[a,*]

[a] *Department of Mmathematics, Payame Noor University , P. O. Box 19395-4697, Tehran, Iran*

Article Info

Keywords:
Nonexpansive mapping
Demiclosedness principle
Fixed point

2020 MSC:
37C25
47H10

Abstract

In this paper we introduce two new classes of generalized nonexpansive mapping and we study both the existence of fixed points and their asymptotic behavior.

1. Introduction

Let C be a nonempty subset of a Banach space X. It is well known that a mapping T: $C \to X$ is said to be nonexpansive whenever $\|Tx - Ty\| \leq \|x - y\|$ for all $x, y \in C$. Among the most important features of nonexpansive mappings are the following facts.

i) If C is closed convex and bounded and $T : C \to C$ is nonexpansive, then there exists a sequence (x_n) in C such that $\|x_n - Tx_n\| \to 0$. Such a sequence is called almost fixed point sequence for T (a.f.p.s. in short).

ii) Even when C is a weakly compact convex subset of X, a nonexpansive self-mapping of C need not have fixed points. Nevertheless, if the norm of X has suitable geometric properties (as for instance uniform convexity, among many others), every nonexpansive self-mapping of every weakly compact convex subset of X has a fixed point. In this case X is said to have the weak fixed point property (WFPP in short).

In a recent paper [1], Suzuki defined a class of generalized nonexpansive mappings as follows.

Definition 1.1. Let C be a nonempty subset of a Banach space X. We say that a mapping $T : C \to X$ satisfy condition (C) on C if for all $x, y \in C$,
$\frac{1}{2}\|x - Tx\| \leq \|x - y\|$ implies $\|Tx - Ty\| \leq \|x - y\|$.

*Talker
Email address: parastoo.heiatiannaeini@pnu.ac.ir (Parastoo Heiatian Naeini)

Of course, every nonexpansive mapping $T : C \to X$ satisfies condition (C) on C, but in [1] some examples of non continuous mappings satisfying condition (C) are given.
In spite that the class of mapping satisfying condition (C) is broader than the class of nonexpansive mappings, when C is a convex bounded subset of X, every mapping $T : C \to C$ which satisfies condition (C) on C has a.f.p. sequences, thatis, it shares (i) with nonexpansive mappings (see [1]), as well as (ii), because for some Banach spaces (see [1]) mappings satisfying (C) leaving invariant weakly compact convex subsets have fixed points. (See also [4])
In this paper we define two kind of generalizations of condition (C). This will lead us to some classes of mappings which are wider than those which satisfy condition (C) but preserving their fixed point properties.

2. Notations and preliminaries

Throughout this note we assume that $(X, \|.\|)$ is a real Banach space whose zero vector is 0_X. As it is usual, we will denote by $B[x, r]$ and $S[x, r]$ the closed ball and the sphere of the Banach space$(X, \|.\|)$ with radius r and center $x \in X$, respectively. In particular we will write $B_X := B[0_X , 1]$and $S_X := S[0_X , 1]$.
We will use $x_n \to x$ to denote that the sequence (x_n) in X is weakly convergent to $x \in X$.
Let C be a nonempty closed and convex subset of X, and let (x_n) be a bounded sequence in X. For $x \in X$ the asymptotic radius of (x_n) at x is the number $r(x, (x_n)) := \lim_{n\to\infty} sup\|x - x_n\|$.
The real number $r(C, (x_n)) := \inf\{r(x, (x_n)) : x \in C\}$ is called the asymptotic radius of (x_n) relative to C and finally the set $A(C, (x_n)) = \{x \in C : r(x, (x_n)) = r(C, (x_n))\}$, is called the asymptotic center of (x_n) relative to C.
It is well known that $A(C, (x_n))$ consists of exactly one point whenever the space X is uniformly convex in every direction (UCED), and that $A(C, (x_n))$ is nonempty and convex when C is weakly compact and convex.

3. A class more general than type (C) mappings

We generalize condition (C) as follows.

Definition 3.1. Let C be a nonempty subset of a Banach space X. For $\mu \geq 1$ we say that a mapping $T : C \to X$ satisfy condition (E_μ) on C if there exists $\mu \geq 1$ such that for all $x, y \in C$,

$$\|x - Ty\| \leq \mu\|x - Tx\| + \|x - y\|$$

We say that T satisfies condition (E) on C whenever T satisfies (E_μ) for some $\mu \geq 1$.

3.1 It is obvious that if $T : C \to X$ is nonexpansive, then it satisfies condition (E_1).

The converse is not true.

3.2 From Lemma 7 in [1] we know that if $T : C \to C$ satisfies condition (C) on C, then is satisfies condition (E_3).

Then, the following result is also obvious.

Proposition 3.2. *Let $T : C \to X$ be a mapping which satisfies condition (E) on C. If T has some fixed point, then T is quasi-nonexpansive. The converse is not true.*

Proposition 3.3. *Let $T : C \to X$ be a mapping which satisfies condition (E) on C. Then the following statements hold.*

a) If $TC \subset C$ then for all $x \in C$, $\|x - T^2x\| \leq (\alpha + 1)\|x - Tx\|$.

b) If $TC \subset C$ then for all $x, y \in C$, $\|Tx - Ty\| \leq \alpha\|Tx - T^2x\| + \|Tx - y\|$.

c) If $r \in (0, 1)$ then the mappings $T_r : C \to X$ defined as $T_r = rT + (1 - r)I$ (where I is the identity mapping), satisfy the condition (E_α) on C.

Proposition 3.4. *(Alternative principle). Let C be a bounded subset of X. Let $T : C \to C$ be an arbitrary mapping. Then one at least of the following statements hold.*

a) There exists an a.f.p.s. for T in C.

b) T satisfies condition (E)on C.

Theorem 3.5. *Let C be a nonempty subset of a Banach space X. Let $T : C \to X$ be a mapping. If*

a) There exists an a.f.p.s. (x_n) for T in C such that $x_n \to z \in C$,

b) T satisfies condition (E) on C, and

c) $(X, \|.\|)$ satisfies the Opial condition.

Then,$Tz = z$.

Corollary 3.6. *Let C be a nonempty weakly compact subset of a Banach space X. Suppose that $(X, \|.\|)$ satisfies the Opial condition. Let $T : C \to X$ be a mapping which satisfies condition (E) on C. Then T has a fixed point in C if and only if T admits an a.f.p.s.*

Theorem 3.7. *Let C be a nonempty compact subset of a Banach space X. Let $T : C \to X$ be a mapping which satisfies condition (E) on C. Then T has a fixed point in C if and only if T admits an a.f.p.s.*

Theorem 3.8. *Let C be a nonempty weakly compact convex subset of a (UCED) Banach space X. Let $T : C \to X$ be a mapping. If*

a) T satisfies condition (E) on C, and

b) $inf\{\|x - Tx\| : x \in C\} = 0$.

Then, T has a fixed point.

4. A direct generalization of condition (C)

Definition 4.1. For $\lambda \in (0, 1)$ we say that a mapping $T : C \to X$ satisfy condition $(C\lambda)$ on C if for all $x, y \in C$ with $\lambda\|x - Tx\| \leq \|x - y\|$. one has that$\|Tx - Ty\| \leq \|x - y\|$.

Of course, if $\lambda = \dfrac{1}{2}$ we recapture the class of mappings satisfying condition (C). Notice that if $0 < \lambda_1 < \lambda_2 < 1$ then the condition (C_{λ_1}) implies condition (C_{λ_2}).

Proposition 4.2. *Let C be a subset of a Banach space X. If $T : C \to X$ satisfies the condition (C_λ) for some $\lambda \in (0, 1)$, then for every $r \in (\lambda, 1)$ the mapping $T_r : C \to X$ defined by $T_r(x) = rTx + (1 - r)x$ satisfy the condition $(C_{\lambda/r})$.*

The class of mappings satisfying condition (C) on a convex bounded subset C of X shares with the class of nonexpansive mappings the existence of almost fixed point sequences.

Theorem 4.3. *Let C be a bounded convex subset of a Banach space X. Assume that $T : C \to C$ satisfies condition (C_λ) on C for some $\lambda \in (0, 1)$. For $r \in (\lambda, 1)$ define a sequence (x_n) in C by tacking $x_1 \in C$ and $x_{n+1} = rT(x_n) + (1-r)x_n$ for $n \geq 1$. Then (x_n) is an a.f.p.s. for T , that is, the mappings T_r are asymptotically regular.*

Lemma 4.4. *Let C be a subset of a Banach space X. Let $T : C \to X$ be a mapping satisfying condition (C_λ) for some $\lambda \in (0, 1)$. Let (x_n) be a bounded a.f.p.s. for T . Then $\lim_{n\to\infty} sup\|x_n - Ty\| \leq \lim_{n\to\infty} sup\|x_n - y\|$ holds for all $y \in C$ with $\liminf_n \|x_n - y\| > 0$.*

Proposition 4.5. *Let $T : C \to X$ be a Lipschitzian mapping with Lipschitz constant $Lip(T)$ satisfying condition $(C\lambda)$ for some $\lambda \in (0, 1)$. Then, T satisfies condition (E_μ) for $\mu = \max\{1, 1 + \lambda(Lip(T) - 1)\}$.*

Definition 4.6. Given a mapping $T : C \to X$, we say that $I - T$ is strongly demiclosed at 0_X if for every sequence (x_n) in C strongly convergent to $z \in C$ and such that $x_n - Tx_n \to 0_X$ e have that $z = Tz$.

This is a weaker version of the well-known demiclosedness principle, in which the weak convergence has been replaced by the strong convergence. Notice that for every continuous mapping (in particular for every Lipschitzian mapping) $T : C \to X$, $I - T$ is strongly demiclosed at 0_X .

Proposition 4.7. *Let* C *be a nonempty subset of a Banach space X. If* $T : C \to X$ *satisfies condition* (E) *on C, then* $I - T$ *is strongly demiclosed at* 0_X .

Theorem 4.8. *Let* C *be a nonempty weakly compact convex subset of a Banach space* X. *Let* $T : C \to C$ *be a mapping. If*

a) T *satisfies condition* (C_λ) *on* C,

) $(X, \|.\|)$ *satisfies the Opial condition, and*

c) $I - T$ *is strongly demiclosed at* 0_X .

Then, $Tz = z = \mu Tx_n + (1 - \mu)x_n$ *for* $n \in \mathbb{N}$, *where* μ *is a real number belonging to* $[\lambda, 1)$. *Then* (x_n) *converges weakly to a fixed point of* T .

Theorem 4.9. *Let* C *be a nonempty weakly compact convex subset of a* $(UCED)$ *Banach space* X. *Let* $T : C \to C$ *be a mapping. If*

a) T *satisfies condition* (C_λ) *on* C, *and*

b) $I{-}T$ *is strongly demiclosed at* 0_X .

Then, T *has a fixed point.*

Theorem 4.10. *Let* C *be a convex subset of a Banach space* X. *Let* $T : C \to C$ *be a mapping satisfying* (E) *and* (C_λ) *for some* $\lambda \in (0, 1)$. *Assume either of the following holds.*

a) C *is weakly compact and* $(X, \|.\|)$ *satisfies the Opial condition.*

b) C *is compact.*

c) C *is weakly compact and* X *is* $(UCED)$.

Theorem 4.11. *Let* T *be a mapping on a locally weakly compact convex subset* C *of a Banach space* X. *Assume that* X *satisfies the Opial condition,* T *satisfies condition* (C_λ) *for some* $\lambda \in (0, 1)$ *and* T *has a fixed point. Define a sequence* (x_n) *in* C *by* $x_1 \in C$ *and* $x_{n+1}ze$ *Then,* T *has a fixed point.*

The following theorem tells that if C is a closed interval of R and T satisfies (C_λ) for some $\lambda \in [0, 3/4]$, then T has a fixed point.

Theorem 4.12. *Let* C *be a closed interval of* $\mathbb{R}$. *Let* T *be a mapping on* C *satisfying condition* $(C_{3/4})$. *Then* T *has a fixed point.*

References

[1] T. Suzuki, *Fixed point theorems and convergence theorems for some generalized nonexpansive mappings*, J. Math. Anal. Appl. 340 (2008) 1088–1095.

[2] K. Goebel, *Concise Course on Fixed Point Theorems*, Yokohama Publ., 2002.

[3] Kazimierz Goebel, W.A. Kirk, *Iteration processes for nonexpansive mappings, in: Topological Methods in Nonlinear Functional Analysis*, Toronto, Ont.,1982, Contemp. Math., vol. 21, Amer. Math. Soc., Providence, RI, 1983, pp. 115–123.

[4] S. Dhompongsa, W. Inthakon, A. Kaewkhao, *Edelstein's method and fixed point theorems for some generalized nonexpansive mappings*, J. Math. Anal. Appl. 350 (2009) 12–17.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–258, pp. 150–153

Approximation of Cubic Numerical Range of Block Operator Matrix

Maryam Fazlollahi*

Department of Mathematics, Payame Noor University, P.O. Box 19395-3697 Tehran, Iran.

Article Info

Keywords:
cubic numerical range
block operator matrix
spectrum

2020 MSC:
47A10
47A12
15A60

Abstract

In this paper we consider bounded 3×3 operator matrix $\mathcal{A}$ and obtain an approximation of the cubic numerical range, that may give a better localization of the spectrum than the numerical range.

1. Introduction

The notion of the numerical range has been defined in various ways in the literatures [1, 2, 7]. The concept of the quadratic numerical range of a 2×2 block operator matrix $\mathcal{A}$ in a Hilbert space $\mathcal{H}$ with respect to a decomposition $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2$ $(W^2(\mathcal{A}))$ has been introduced in [3] and further investigated in[4, 5, 8]. In particular, it has been shown that it is a subset of the numerical range and that its closure still contains the spectrum $\mathcal{A}$. The notion of the cubic numerical range of a 3×3 block operator matrix $\mathcal{A}$ in a Hilbert space $\mathcal{H}$ with respect to a decomposition $\mathcal{H} = \mathcal{H}_1 \times \mathcal{H}_2 \times \mathcal{H}_3$ has been defined in [9] and further investigated in [6]. It is contained in the quadratic numerical range for bounded block operator matrix and, like the numerical range and quadratic numerical range, it contains the spectrum of $\mathcal{A}$ in its closure.

2. The definition and preliminaries

If the complex Hilbert space $\mathcal{H}$ is the product of three Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$, $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$, then every bounded linear operator $\mathcal{A} \in L(\mathcal{H})$ has a block operator matrix representation

$$\mathcal{A} := \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{1}$$

*Talker
Email address: m_fazlolahi@pnu.ac.ir (Maryam Fazlollahi)

with bounded linear operators $A_{ij} \in L(\mathcal{H}_j, \mathcal{H}_i), i, j = 1, 2, 3$. For unit vectors x, y, z in $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$ respectively, define $\mathcal{A}_{x,y,z} \in M_3(\mathbb{C})$ by

$$\mathcal{A}_{x,y,z} := \begin{bmatrix} < A_{11}x, x > & < A_{12}y, x > & < A_{13}z, x > \\ < A_{21}x, y > & < A_{22}y, y > & < A_{23}z, y > \\ < A_{31}x, z > & < A_{32}y, z > & < A_{33}z, z > \end{bmatrix}. \tag{2}$$

Trettre introduced the cubic numerical range of operator $\mathcal{A}$ (with respecting to the block operator representation 1) by

$$W^3(\mathcal{A}) := \{\lambda \in \mathbb{C} : \lambda \in \sigma(\mathcal{A}_{x,y,z}), \| x \| = \| y \| = \| z \| = 1, x \in \mathcal{H}_1, y \in \mathcal{H}_2, z \in \mathcal{H}_3\}.$$

Theorem 2.1. *[9] Let $\mathcal{A}$ be a block operator matrix on $\mathcal{H}$ with respect to the block operator representation 1 .The following properties hold:*
1) $\sigma_p(\mathcal{A}) \subseteq W^3(\mathcal{A})$,
2)$\sigma(\mathcal{A}) \subseteq \overline{W^3(\mathcal{A})}$,
3) $W^3(\mathcal{A}) \subseteq W(\mathcal{A})$,
4) $W^3(\mathcal{A}^*) := \{\lambda : \overline{\lambda} \in W^3(\mathcal{A})\}$.

3. main result

Theorem 3.1. *Let $\mathcal{A}$ be as in 1 and $(U_k)_{k=1}^{\infty}$, $(V_k)_{k=1}^{\infty}$ and $(W_k)_{k=1}^{\infty}$ be nested families of spaces in $\mathcal{H}_1$, $\mathcal{H}_2$ and $\mathcal{H}_3$ respectively, given by $U_k = span\{\varphi_1, ..., \varphi_k\}$, $V_k = span\{\psi_1, ..., \psi_k\}$ and $W_k = span\{\xi_1, ..., \xi_k\}$ where $(\varphi_k)_{k=1}^{\infty}$, $(\psi_k)_{k=1}^{\infty}$ and $(\xi_k)_{k=1}^{\infty}$ are orthonormal. Consider*

$$\mathcal{A}_{i,j,\tau} := \begin{bmatrix} (A_{11})_{i\times i} & (A_{12})_{i\times j} & (A_{13})_{i\times \tau} \\ (A_{21})_{j\times i} & (A_{22})_{j\times j} & (A_{23})_{j\times \tau} \\ (A_{31})_{\tau\times i} & (A_{32})_{\tau\times j} & (A_{33})_{\tau\times \tau} \end{bmatrix} \tag{3}$$

where $((A_{11})_{i\times i})_{pq} :=< A_{11}\varphi_p, \varphi_q >, p, q = 1, \dots, i$, $((A_{12})_{i\times j})_{pq} :=< A_{12}\psi_p, \varphi_q >, p = 1, \dots, i, q = 1, \dots, j$, $((A_{13})_{i\times \tau})_{pq} :=< A_{13}\xi_p, \varphi_q >, p = 1, \dots, i, q = 1, \dots, \tau$, $((A_{21})_{j\times i})_{pq} :=< A_{21}\varphi_q, \psi_p >, p = 1, \dots, j, q = 1, \dots, i$, $((A_{22})_{j\times j})_{pq} =< A_{22}\psi_p, \psi_q >, p, q = 1, \dots, j$, $((A_{23})_{j\times \tau})_{pq} :=< A_{23}\xi_p, \psi_q >, p = 1, \dots, j, q = 1, \dots, \tau$, $((A_{31})_{\tau\times i})_{pq} :=< A_{31}\varphi_q, \xi_p >, p = 1, \dots, \tau, q = 1, \dots, i$, $((A_{32})_{\tau\times j})_{pq} :=< A_{32}\psi_q, \xi_p >, p = 1, \dots, \tau, q = 1, \dots, j$, $((A_{33})_{\tau\times \tau})_{pq} :=< A_{33}\xi_p, \xi_q >, p, q = 1, \dots, \tau$ then $W^3(\mathcal{A}_{i,j,\tau}) \subseteq W^3(\mathcal{A})$.

Proof. First we consider $\lambda \in W^3(\mathcal{A}_{i,j,\tau})$. Then for some $\alpha \in \mathbb{C}^i$, $\beta \in \mathbb{C}^j$ and $\gamma \in \mathbb{C}^\tau$ with $\| \alpha \| = \| \beta \| = \| \gamma \| = 1$, λ is an eigenvalue of

$$(\mathcal{A}_{i,j,\tau})_{\alpha,\beta,\gamma} := \begin{bmatrix} < (A_{11})_{i\times i}\alpha, \alpha > & < (A_{12})_{i\times j}\beta, \alpha > & < (A_{13})_{i\times \tau}\gamma, \alpha > \\ < (A_{21})_{j\times i}\alpha, \beta > & < (A_{22})_{j\times j}\beta, \beta > & < (A_{23})_{j\times \tau}\gamma, \beta > \\ < (A_{31})_{\tau\times i}\alpha, \gamma > & < (A_{32})_{\tau\times j}\beta, \gamma > & < (A_{33})_{\tau\times \tau}\gamma, \gamma > \end{bmatrix} \tag{4}$$

Define isometries $f : U_i \to \mathbb{C}^i$, $g : V_j \to \mathbb{C}^j$ and $h : W_\tau \to \mathbb{C}^\tau$ by $f(\alpha_1\varphi_1 + \cdots + \alpha_i\varphi_i) := (\alpha_1, \cdots, \alpha_i)$, $g(\beta_1\psi_1 + \cdots + \beta_j\psi_j) := (\beta_1, \cdots, \beta_j)$ and $h(\gamma_1\xi_1 + \cdots + \gamma_\tau\xi_\tau) := (\gamma_1, \cdots, \gamma_\tau)$.
Choose $x \in U_i$, $y \in V_j$, $z \in W_\tau$ such that $f(x) = \alpha$, $g(y) = \beta$, $h(z) = \gamma$ and $\| x \| = \| y \| = \| z \| = 1$.
A simple calculation shows that

$$(\mathcal{A}_{i,j,\tau})_{\alpha,\beta,\gamma} = \mathcal{A}_{x,y,z} = \begin{bmatrix} < A_{11})x, x > & < A_{12}y, x > & < A_{13}z, x > \\ < A_{21}x, y > & < A_{22}y, y > & < A_{23}z, y > \\ < A_{31}x, z > & < A_{32}y, z > & < A_{33}z, z > \end{bmatrix} \tag{5}$$

So $\lambda \in W^3(\mathcal{A})$ □

Lemma 3.2. *Let $(U_k)_{k=1}^{\infty}$, $(V_k)_{k=1}^{\infty}$, $(W_k)_{k=1}^{\infty}$ and $\mathcal{A}_{i,j,\tau}$ be as in 3.1. Hence $W^3(\mathcal{A}_{i,j,\tau}) \subseteq W^3(\mathcal{A}_{\mu,\nu,\omega})$ for $\mu \geq i, \nu \geq j, \omega \geq \tau$.*

Proof. This result is an immediate consequence of the fact that $\mathbb{C}^\tau$ is a subspace of $\mathbb{C}^\omega$ for $\tau \leq \omega$. In detail: suppose $\tau \leq \omega$ and suppose $\lambda \in W^3(\mathcal{A}_{i,j,\tau})$. Then there exist $\alpha \in \mathbb{C}^i$, $\beta \in \mathbb{C}^j$ and $\gamma \in \mathbb{C}^\tau$ with $\| \alpha \| = \| \beta \| = \| \gamma \| = 1$ such that in the notation of 4, λ is an eigenvalue of $(\mathcal{A}_{i,j,\tau})_{\alpha,\beta,\gamma}$. Choose $\zeta_1 \in \mathbb{C}^\mu, \zeta_2 \in \mathbb{C}^\nu, \zeta_3 \in \mathbb{C}^\omega$ by setting $\zeta_1 = (\alpha_1, \cdots, \alpha_i, 0, \cdots, 0)^T$, $\zeta_2 = (\beta_1, \cdots, \beta_j, 0, \cdots, 0)^T$ and $\zeta_3 = (\gamma_1, \cdots, \gamma_\tau, 0, \cdots, 0)^T$.
A simple calculation shows that $(\mathcal{A}_{i,j,\tau})_{\alpha,\beta,\gamma} = (\mathcal{A}_{\mu,\nu,\omega})_{\zeta_1,\zeta_2,\zeta_3}$ and so $\lambda \in W^3(\mathcal{A}_{\mu,\nu,\omega})$. □

Theorem 3.3. *Let $\mathcal{A}$ and $\mathcal{A}_{i,j,\tau}$ be as in theorem 3.1 Suppose that A_{31} is $A_{11}-$ bounded, A_{12} is $A_{22}-$ bounded and A_{13} is $A_{33}-$ bounded that $\mathcal{C}_1 := span((\varphi_k)_{k=1}^\infty)$ is a core of A_{11}, $\mathcal{C}_2 := span((\psi_k)_{k=1}^\infty)$ is a core of A_{22} and $\mathcal{C}_3 := span((\xi_k)_{k=1}^\infty)$ is a core of A_{33}. Hence $\overline{W^3\mathcal{A}} = \overline{\cup_{i,j,\tau\in\mathbb{N}} W^3(\mathcal{A}_{i,j,\tau})} = \overline{\cup_{i\in\mathbb{N}} W^3(\mathcal{A}_{i,i,i})}$.*

Proof. First we observe that by lemma 3.2 $W^3(\mathcal{A}_{i,j,\tau}) \subseteq W^3(\mathcal{A}_{l,l})$ where $l = max\{i, j, \tau\}$; this establishes that $\overline{\cup_{i,j,\tau\in\mathbb{N}} W^3(\mathcal{A}_{i,j,\tau})} = \overline{\cup_{i\in\mathbb{N}} W^3(\mathcal{A}_{i,i,i})}$. In view of theorem 3.1 it therefore now suffices to show $W^3(\mathcal{A}) \subseteq \overline{\cup_{i,j,\tau\in\mathbb{N}} W^3(\mathcal{A}_{i,j,\tau})}$.
Suppose $\lambda \in W^3(\mathcal{A})$. Choose $x \in \mathcal{H}_1$, $y \in \mathcal{H}_2$ and $z \in \mathcal{H}_3$ such that λ is an eigenvalue of $\mathcal{A}_{x,y,z}$ as defined in 2. Since $\mathcal{C}_1$ is a core of A_{11} there axists a sequence $(x_k)_{k=1}^\infty$, with each $x_k \in span\{\varphi_1, \cdots, \varphi_{s_k}\}$ for some $s_k > 0$ such that $\| x - x_k \| \to 0$ and $\| A_{11}x - A_{11}x_k \| \to 0$. Because A_{31} is $A_{11}-$ bounded then $\| A_{31}x - A_{31}x_k \| \to 0$. Since $\mathcal{C}_2$ is a core of A_{22} and A_{12} is $A_{22}-$ bounded, in a similar way, a sequence $(y_k)_{k=1}^\infty$ with each $y_k \in span\{\psi_1, \cdots, \psi_{t_k}\}$ for some $t_k > 0$ such that $\| y - y_k \| \to 0$, $\| A_{22}y - A_{22}y_k \| \to 0$ and $\| A_{12}y - A_{12}y_k \| \to 0$. Also since $\mathcal{C}_3$ is a core of A_{33} and A_{13} is $A_{33}-$ bounded. We may also find, in a similar way, a sequence $(z_k)_{k=1}^\infty$ with each $z_k \in span\{\xi_1, \cdots, \xi_{u_k}\}$ for some $u_k > 0$ such that $\| z - z_k \| \to 0$, $\| A_{33}z - A_{33}z_k \| \to 0$ and $\| A_{13}z - A_{13}z_k \| \to 0$. In particular this means that in the notation of 2, $\| \mathcal{A}_{x_k,y_k,z_k} - \mathcal{A}_{x,y,z} \| \to 0$ as $k \to 0$.
Fix $k > 0$. Let $f : span\{\varphi_1, \cdots, \varphi_{s_k}\} \to \mathbb{C}^{s_k}$, $g : span\{\psi_1, \cdots, \psi_{t_k}\} \to \mathbb{C}^{t_k}$ and $h : span\{\xi_1, \cdots, \xi_{u_k}\} \to \mathbb{C}^{u_k}$ are the isometries in the proof of theorem 3.1. Define $\tilde{\alpha_k} \in \mathbb{C}^{s_k}$, $\tilde{\beta_k} \in \mathbb{C}^{t_k}$ and $\tilde{\gamma_k} \in \mathbb{C}^{u_k}$ by $\tilde{\alpha_k} := f(x_k)$, $\tilde{\beta_k} := g(y_k)$ and $\tilde{\gamma_k} := h(z_k)$.
Consider the matrix

$$\mathcal{A}_k := \begin{bmatrix} \frac{< (A_{11})_{s_k\times s_k}\tilde{\alpha_k}, \tilde{\alpha_k} >}{(\| \tilde{\alpha_k} \|)^2} & \frac{< (A_{12})_{t_k\times s_k}\tilde{\beta_k}, \tilde{\alpha_k} >}{\| \tilde{\alpha_k} \|\| \tilde{\beta_k} \|} & \frac{< (A_{13})_{u_k\times s_k}\tilde{\gamma_k}, \tilde{\alpha_k} >}{\| \tilde{\alpha_k} \|\| \tilde{\gamma_k} \|} \\ \frac{< (A_{21})_{s_k\times k_k}\tilde{\alpha_k}, \tilde{\beta_k} >}{\| \tilde{\alpha_k} \|\| \tilde{\beta_k} \|} & \frac{< (A_{22})_{t_k\times t_k}\tilde{\beta_k}, \tilde{\beta_k} >}{(\| \tilde{\beta_k} \|)^2} & \frac{< (A_{23})_{u_k\times t_k}\tilde{\gamma_k}, \tilde{\beta_k} >}{\| \tilde{\beta_k} \|\| \tilde{\gamma_k} \|} \\ \frac{< (A_{31})_{s_k\times u_k}\tilde{\alpha_k}, \tilde{\gamma_k} >}{\| \alpha_k \|\| \tilde{\gamma_k} \|} & \frac{< (A_{32})_{t_k\times u_k}\beta_k, \tilde{\gamma_k} >}{\| \tilde{\beta_k} \|\| \tilde{\gamma_k} \|} & \frac{< (A_{33})_{u_k\times u_k}\tilde{\gamma_k}, \tilde{\gamma_k} >}{(\| \tilde{\gamma_k} \|)^2} \end{bmatrix} \tag{6}$$

A simple calculation shows that $A_k = \mathcal{A}_{x_k,y_k,z_k}$. Since $\| \mathcal{A}_{x_k,y_k,z_k} - \mathcal{A}_{x,y,z} \| \to 0$ as $k \to \infty$. We have $\| A_k - \mathcal{A}_{x,y,z} \| \to 0$. The eigenvalues of A_k are elements of $W^3(\mathcal{A}_{s_k,t_k,u_k})$ by definition of $\mathcal{A}_{s_k,t_k,u_k}$ and 6. Hence there exists $\lambda_k \in W^3(\mathcal{A}_{s_k,t_k,u_k})$ such that $\lambda_k \to \lambda$. In view of lemma 3.2 this immediately gives $\lambda \in \overline{\cup_{i,j,\tau\in\mathbb{N}} W^3(\mathcal{A}_{i,j,\tau})}$. □

Remark 3.4. Let P_k, Q_k and R_k denote orthogonal projection onto U_k, V_k and W_k respectively. If $\mathcal{A}$ is bounded, then the hypotheses that $\mathcal{C}_1$, $\mathcal{C}_2$ and $\mathcal{C}_3$ are cores of $\mathcal{A}_{11}$, $\mathcal{A}_{22}$ and $\mathcal{A}_{33}$ respectively, are equivalent to the statements $P_k \to I$, $Q_k \to I$, $R_k \to I$ strongly as $k \to \infty$. The following example shows that these are necessary.

Example 3.5. Let

$$\mathcal{A} := \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \tag{7}$$

is a block operator matrix in $H = l^2(\mathbb{N}) \times l^2(\mathbb{N}) \times l^2(\mathbb{N})$ where $A_{11} = diag\{\frac{1}{n}\}_{n=1}^\infty$ and $A_{12} = A_{13} = A_{21} = A_{22} = A_{23} = A_{31} = A_{32} = A_{33} = 0$. Let $(U_k)_{k\in\mathbb{N}}$, $(V_k)_{k\in\mathbb{N}}$ are nested families of subspaces in $l^2(\mathbb{N})$ with $U_k = span\{e_2, \cdots, e_{k+1}\}$ where $e_j = j^{th}$ standard basis vector and $V_k = span\{\psi_1, \cdots, \psi_k\}$ where $(\psi_k)_{k=1}^\infty$ is any orthonormal sequence. Hence performing an analysis analogous to theorem 3.3. We see that the top left-hand corner of A_k in 6 is not convergent to $< A_{11}x, x >$ unless x is orthogonal to e_1.

References

[1] K. E. Gustafson, and D. K. M. Rao. Numerical Range. Springer-Verlag, Berlin, Heidelberg. New York, 1997.
[2] R. A. Horn, and C. R. Johnson. Topics in Matrix Analysis. Cambridge University Press, 1991.
[3] H. Langer, and C. Tretter. Spectral decomposition of some nonselfadjoint block operator matrices. J. Operator Theory 39(2), 339-359, 1998.
[4] H. Langer, A. Markus, V. Mastaev, and C. Tretter. A new concept for block operator matrices: The quadratic numerical range. Linear Algebra and Appl., 330 (1-3), 89-112, 2001.
[5] H. Langer, A. Markus, and C. Tretter. Corners of numerical ranges. Operator Theory. Adv. Appl., 385-400, 2001.
[6] M.Sharipova. Usual, quadratic and cubic numerical renges corresponding to a 3×3 operator matrices. Bulletin of National Univ. of Uzbek, no. 5(4), 242-249, 2022.
[7] T. Rasulov, and E. Dilmurodov. Investigations of the numerical range of a operator matrix. J. Samara State Tech. Univ., Ser. Phys. and Math. Sci. 35:2, 50-63, 2014.
[8] T. Rasulov, and E. Dilmurodov. Estimates for quadratic numerical range of a operator matrix. Uzbek Mathematical Journal, no. 1, 64-74, 2015.
[9] C. Tretter. Spectral theory of block operator matrices and applications. Imperial College Press, London, 2008.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–271, pp. 154–156

Maps preserving the pseudo spectrum of operator products

Rohollah Parvinianzadeh[a,*], Morad Ali Peyvand[b]

[a] *Department of Mathematics, University of Yasouj, Yasouj 75918, Iran*
[b] *Department of Mathematics, University of Yasouj, Yasouj 75918, Iran*

Article Info

Keywords:
Local spectrum
Rank-one operators
the single-valued extension property

2020 MSC:
47A11
47A15

Abstract

Let A be a standard operator algebra on complex Hilbert spaces H. Fix $\epsilon > 0$ and $T \in A$, let $\sigma_\epsilon(T)$ denote the ϵ-pseudo spectrum of T. In this paper, we prove if the bijective map φ on A satisfies

$$\sigma_\epsilon(T^*S + ST^*) = \sigma_\epsilon(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*), \quad (T, S \in A),$$

then there exists a unitary operator on H such that $\varphi(T) = \lambda UTU^*$ for every $T \in A$, where $\lambda \in \{-1, 1\}$.

1. Introduction and Preliminaries

Throughout this paper, $B(H)$ stands for the algebra of all bounded linear operators acting on an infinite dimensional complex Hilbert space H and its unit will be denoted by I. For an operator $T \in B(H)$, the adjoint, the spectrum and the spectral radius of T are denoted by T^*, $\sigma(T)$ and $r(T)$, respectively. The peripheral spectrum of an element $T \in B(H)$ is defined by

$$\sigma_\pi(T) = \{\lambda \in \sigma(T) : r(T) = \|T\|\}.$$

Note that $\sigma_\pi(T) \subseteq \sigma(T)$. For $\epsilon > 0$, the ϵ-pseudo spectrum of T, $\sigma_\epsilon(T)$, is defined by $\sigma_\epsilon(T) = \{\lambda \in \mathbb{C} : \|(\lambda I - T)^{-1}\| \geq \epsilon^{-1}$ with the convention that $\|(\lambda I - T)^{-1}\| = \infty$ if $\lambda \in \sigma(T)$. It is a compact subset of $\mathbb{C}$ and contains $\sigma(T)$, the spectrum of T. Unlike the spectrum, which is a purely algebraic concept, the ϵ-pseudo spectrum depends on the norm. Pseudo spectra are a useful tool for analyzing operators, furnishing a lot of information about the algebraic and geometric properties of operators and matrices. They play a very natural role in numerical computations, especially in those involving spectral perturbations. The book [6] gives an extensive account of the pseudo spectra, as well as investigations and applications in numerous fields.
Linear preserver problems, in the most general setting, demands the characterization of maps between algebras that leave a certain property, a particular relation, or even a subset invariant. In all cases that have been studied by now, the maps are either supposed to be linear, or proved to be so. This subject is very old and goes back well over a century to the so-called first linear preserver problem, due to Frobenius [5], who characterized linear maps that preserve the

*Talker
Email addresses: r.parvinian@yu.ac.ir (Rohollah Parvinianzadeh), peyvand@yu.ac.ir (Morad Ali Peyvand)

determinant of matrices. The study of nonlinear pseudo spectrum preserver problems attracted the attention of a number of authors. Cui et al. [2, Theorem 3.3] characterized maps on $M_n(\mathbb{C})$ that preserve the ϵ-pseudo spectrum of the usual product of matrices. They proved that a map φ on $M_n(\mathbb{C})$ satisfies

$$\sigma_\epsilon(\varphi(T)\varphi(S)) = \sigma_\epsilon(TS) \ \ (T, S \in M_n(\mathbb{C}))$$

if and only if there exist a scalar $c = \pm 1$ and a unitary matrix $U \in M_n(\mathbb{C})$ such that $\varphi(T) = cUTU^*$ for all $T \in M_n(\mathbb{C})$. This result was extended to the infinite dimensional case by Cui et al. [4, Theorem 4.1], where the authors showed that a surjective map φ on $B(H)$ preserves the ϵ-pseudo spectrum of the product of operators if and only if it is a unitary similarity transform up to a scalar $c = \pm 1$. The aim of this note is to characterize mappings on $B(H)$ that preserve the ϵ-pseudo spectral of the product "$TS - ST^*$" of operators. For two nonzero vectors x and y in H, let $x \otimes y$ stands for the operator of rank at most one defined by

$$(x \otimes y)z = \langle z, y\rangle\, x, \qquad \forall\, z \in H.$$

Note that every rank one operator in $B(H)$ can be written in this form, and that every finite rank operator $T \in B(H)$ can be written as a finite sum of rank one operators; i.e., $T = \sum_{i=1}^{n} x_i \otimes y_i$ for some $x_i, y_i \in H$ and $i = 1, 2, ..., n$. We denote by $F(H)$ the set of all finite rank operators in $B(H)$ and $F_n(H)$ the set of all operators of rank at most n, n is a positive integer.
In the following proposition, we collects some known properties of the ϵ-pseudo spectrum which are needed in the proof of the main result.
Let $\epsilon > 0$ be arbitrary and $D(0, \epsilon) = \{\mu \in \mathbb{C} : |\mu - a| < \epsilon\}$, where $a \in \mathbb{C}$.

Proposition 1.1. *(See [4, 6].)*
Let $\alpha > 0$ and let $T \in B(H)$.
(1) $\sigma(T) + D(0, \epsilon) \subseteq \sigma_\epsilon(T)$.
(2) If T is normal, then $\sigma_\epsilon(T) = \sigma(T) + D(0, \epsilon)$.
(3) For any $\alpha \in \mathbb{C}, \sigma_\epsilon(T + \alpha I) = \alpha + \sigma_\epsilon(T)$.
(4) For any nonzero $\alpha \in \mathbb{C}, \sigma_\epsilon(\alpha T) = \alpha\sigma_{\frac{\epsilon}{|\alpha|}}(T)$.
(5) For any $\alpha \in \mathbb{C}$, we have $\sigma_\epsilon(T) = D(\alpha, \epsilon)$ if and only if $T = \alpha I$.

Theorem 1.2. *(See [3, Theorem 3.1].) Let A and B be standard operator algebras on complex Hilbert spaces H and K, respectively. Assume that $\varphi : A \to B$ is a map of which range contains all operators with rank at most two. Then φ satisfies that*

$$\sigma_\pi(T^*S + ST^*) = \sigma_\pi(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*),$$

for all $T, S \in A$, if and only if there exist a scalar λ with $|\lambda| = 1$ and a unitary operator of H into K such that either $\varphi(T) = \lambda UTU^*$ for every $T \in A$, or $\varphi(T) = \lambda UT^tU^*$ for every $T \in A$, where T^t is the transpose of T for an arbitrarily but fixed orthonormal basis of H.

2. Main Results

The following lemma is a key tool for the proof of main result and describes the spectrum of the skew Lie product $(x \otimes y)T - T(x \otimes y)^*$ for any nonzero vectors $x, y \in H$ and operator $T \in B(H)$.

Lemma 2.1. *(See [1, Lemma 2.1].) For any nonzero vectors $x, y \in H$ and $T \in B(H)$, set*

$$\Delta_T(x, y) = (\langle Tx, y\rangle + \langle Ty, x\rangle)^2 - 4\|x\|^2 \langle T^2y, y\rangle$$

and

$$\Lambda_T(x, y) = (\langle x, Ty\rangle + \langle Tx, y\rangle)^2 - 4\|x\|^2 \langle Tx, Ty\rangle$$

Then
(1) $\sigma((x \otimes y)T - T(x \otimes y)^*) = \frac{1}{2}\{0, \langle Tx, y\rangle - \langle Ty, x\rangle \pm \sqrt{\Delta_T(x, y)}\,\}$,
(2) $\sigma(T(x \otimes y) - (x \otimes y)T^*) = \frac{1}{2}\{0, \langle Tx, y\rangle - \langle x, Ty\rangle \pm \sqrt{\Lambda_T(x, y)}\,\}$.

Corollary 2.2. *(See [1, Lemma 2.1].) For any $x \in H$ and $T \in B(H)$, we have*

$$\sigma(T(x \otimes x) + (x \otimes x)T) = \{0, \langle Tx, x\rangle \pm \sqrt{\langle T^2x, x\rangle}\ \}.$$

Lemma 2.3. *Let $\varphi : A \to A$ be a bijective map, where A is a standard operator algebra on complex Hilbert spaces H. If*

$$\sigma(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*) = \sigma(T^*S + ST^*)$$

for every $T, S \in A$, then

$$\sigma_\pi(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*) = \sigma_\pi(T^*S + ST^*)$$

for every $T, S \in A$.

Proof. Assume that relation $\sigma(T^*S + ST^*) = \sigma(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$ holds for every $T, S \in A$. Let $\lambda \in \sigma_\pi(T^*S + ST^*)$ then $r(T^*S + ST^*) = |\lambda|$ and $\lambda \in \sigma(T^*S + ST^*)$. Since $\sigma(T^*S + ST^*) = \sigma(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$ then $\lambda \in \sigma(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$. But from $r(T^*S + ST^*) = r(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$ we have $r(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*) = |\lambda|$ and therefore $\lambda \in \sigma_\pi(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$. Thus $\sigma_\pi(T^*S + ST^*) \subseteq \sigma_\pi(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$. Since φ is a bijection, easy computation shows that $\sigma(T^*S + ST^*) = \sigma(\varphi^{-1}(T)^*\varphi^{-1}(S) + \varphi^{-1}(S)\varphi^{-1}(T)^*)$ for every operators $T, S \in A$. By a similar reasoning one can easily shows that $\sigma_\pi(T^*S + ST^*) \supseteq \sigma_\pi(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*)$ for every operators $T, S \in A$. □

The followig theorem is our main result.

Theorem 2.4. *Let A be a standard operator algebras on complex Hilbert spaces H. Assume that a bijective map $\varphi : A \to A$ satisfies*

$$\sigma_\epsilon(T^*S + ST^*) = \sigma_\epsilon(\varphi(T)^*\varphi(S) + \varphi(S)\varphi(T)^*), \ \ (T, S \in A).$$

Then there exists a unitary operator U on H such that $\varphi(T) = \lambda UTU^*$ for every $T \in A$, where $\lambda \in \{-1, 1\}$.

Proof. The proof of it will be completed after checking several claims.

Claim 1. φ preserves self-adjoint and anti-self adjoint operators in both directions.

Claim 2. There exists a unitary or conjugate unitary operator U on H such that $\varphi(T) = UTU^*$ for every projection P .

Claim 3. There exists a unitary operator U on H such that $\varphi(iT) = \lambda U(iT)U^*$ for every self-adjoint operator T, where $\lambda \in \{-1, 1\}$.

Claim 4. The result in the theorem holds. □

References

[1] E. Alzedani and M. Mabrouk, *Maps preserving the spectrum of skew lie oroduct of operators*, Kragujevac Journal of Mathematics. **64** (4) (2022) pp. 525–532.

[2] J. Cui, V. Forstall, C.K. Li and V. Yannello, *Properties and preservers of the pseudospectrum*, Linear Algebra Appl. **436** (2012) pp. 316–325.

[3] J. Cui, C.K. Li, *Maps preserving peripheral spectrum of Jordan products of operators*, Operators and Matrices. **6**1 (2012) pp. 129–146.

[4] J. Cui, C.K. Li and Y.T. Poon, *Pseudospectra of special operators and pseudospectrum preservers*, J. Math. Anal. Appl. **419** (2014) pp. 1261–1273.

[5] G. Frobenius, *Ueber die Darstellung der endlichen Gruppen durch lineare Substitutionen*, Berl Ber. Appl. **203** (1897) pp. 994–1015.

[6] L.N. Trefethen and M. Embree, *Spectra and pseudospectra*, Princeton Univ. Press, Princeton, NJ, 2005.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–282, pp. 157–159

Nonlinear maps preserving the fixed point of difference of operators

Zynab Izadi[a,*]

[a] *Department of Mathematics, Faculty of Sciences, Payame Noor University, P. O. Box 19395-3697, Tehran, Iran*

Article Info

Keywords:
Non-linear preserver problem
Fixed point
Algebraic singularity.

2020 MSC:
47H10
54H25.

Abstract

Let $\mathcal{H}$ be a Hilbert space and $\mathcal{B}(\mathcal{H})$ the algebra of all bounded linear operators on $\mathcal{H}$ .suppose $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ is a surjective map that satisfying the following property:

$$Fix(A - B) = Fix(\Phi(A) - \Phi(B)), \qquad A, B \in \mathcal{B}(\mathcal{H})$$

then we characterized the form of Φ, that $Fix(A)$ is the set of all fixed point of an opereator A.

1. Introduction

Recently non-linear Preserver problems have been investigated by many authors, see for instance [1–3, 6]. Among them, in [4], Oudghiri and Souilah characterized all surjective maps $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ that preserve operator pairs whose their difference is a non-invertible algebraic operator.

They proved that if $\Phi(I) = I + \Phi(0)$, then there exists an invertible either linear or conjugate linear operator $A : \mathcal{H} \to \mathcal{H}$ such that

$$\Phi(T) = ATA^{-1} + \Phi(0)$$

or

$$\Phi(T) = AT^*A^{-1} + \Phi(0), \qquad T \in \mathcal{B}(\mathcal{H}).$$

In this paper, we are interested to determine the general structure of Φ when it restricts to the real Jordan algebra $\mathcal{S}(\mathcal{H})$.

Now we recall some notions and definitions that will be used in the sequel. Through out this paper $\mathcal{H}$ stands for an infinite dimensional separable complex Hilbert space. We denote by $\mathcal{B}(\mathcal{H})$ the algebra of all bounded linear operators

*Talker
Email address: zynab- izadi@pnu.ac.ir (Zynab Izadi)

on $\mathcal{H}$ and its self-adjoint part by $\mathcal{S}(\mathcal{H})$.

The set of all finite rank operators in $\mathcal{S}(\mathcal{H})$ will be denoted by $\mathcal{F}(\mathcal{H})$. For $g, h \in \mathcal{H}$, $< g, h >$ stands for the inner product of g and h. For every $T \in \mathcal{B}(\mathcal{H})$, we use the notations $rank(T),\ ker(T),\ ran(T)$ and $\sigma(T)$ for the rank, kernel, range and the spectrum of T, respectively.

for all $x, y \in \mathcal{H}$. The identity operator on $\mathcal{H}$ will be denoted by I. $x \in \mathcal{H}$ is a fixed point of an operator $A \in \mathcal{B}(\mathcal{H})$, whenever we have $Ax = x$. The set of all fixed point of A is denote by $Fix(A)$. in [5] the form of product preserving maps in banach space has been investigate.

We denote by $\mathcal{A}(\mathcal{H}), \mathcal{NIA}(\mathcal{H})$ and $\mathcal{IA}(\mathcal{H})$, the set of all algebraic, non-invertible algebraic and invertible algebraic operators in $\mathcal{S}(\mathcal{H})$, respectively. A surjective map $\Lambda : \mathcal{S}(\mathcal{H}) \longrightarrow \mathcal{S}(\mathcal{H})$ is said to preserves fixed point of difference of operators if for every $S, T \in \mathcal{S}(\mathcal{H})$

$$Fix(S - T) = Fix(\Lambda(S) - \Lambda(T)).$$

For every nonzero $h \in \mathcal{H}$ and $f \in \mathcal{H}$ symbol $h \otimes f$ stands for the rank-one linear operator on $\mathcal{H}$ defined by $(x \otimes f)y = f(y)x$ for any $y \in \mathcal{H}$.

every rank one operator in $\mathcal{B}(\mathcal{H})$ can be written in this way.the rank one operator $h \otimes f$ is idempotent if and only if $f(x) = 1$ and is nilpotent if and only if $f(x) = 0$. the set of all rank one operator and the set of all rank one idempotent operators in $\mathcal{B}(\mathcal{H})$ denote by $\mathcal{F}_1(\mathcal{H})$ and $\mathcal{P}_1(\mathcal{H})$ respectively.

the linear preserving maps fined characteristics of a map which preserves some properties under some sutable conditions. for example the maps preserving the spectrum, invertibly, numerical rang and minimum moduli. Molnar considered the product of two operators as the theorem below.

Theorem 1.1. *Let X be a Banach space and let $\Phi : \mathcal{B}(X) \to \mathcal{B}(X)$ be a serjective function with the property that*

$$\sigma_p(\phi(A)\phi(B)) = \sigma_p(AB) \quad (A, B \in \mathcal{B}(X))$$

where $\sigma_p(A) = \{\lambda \in \mathcal{C} : (A - \lambda I)X \neq X\}$
if X is infinite dimensional, then there is an invertible linear operator $T \in \mathcal{B}(X)$ such that either

$$\phi(A) = TAT^{-1} \quad (A \in \mathcal{B}(X))$$
$$or$$
$$\phi(A) = TAT^{-1} \quad (A \in \mathcal{B}(X))$$

in this paper we characterize the form of surjective maps on $\mathcal{B}(\mathcal{H})$ such that preserve the fixed point of difference of operators.

2. Main results

Lemma 2.1. *Let $S, T \in \mathcal{S}(\mathcal{H})$. Then $S = T$, under any of the following conditions.*

(i) *For every $N \in \mathcal{S}(\mathcal{H})$, $S - N \in \mathcal{NIA}(\mathcal{H})$ if and only if $T - N \in \mathcal{NIA}(\mathcal{H})$.*

(ii) *For every $N \in \mathcal{IA}(\mathcal{H})$, $S - N \in \mathcal{IA}(\mathcal{H})$ if and only if $T - N \in \mathcal{IA}(\mathcal{H})$.*

(ii) *For every* $N, S \in \mathcal{B}(\mathcal{H})$*, Fix(ST)=Fix(NT), for all* $T \in \mathcal{F}_1(\mathcal{H})$

Lemma 2.2. *let* $A, B \in \mathcal{B}(\mathcal{H})$ *be non scalar operators. if* $Fix(AR) = Fix(BR)$ *for every* $R \in \mathcal{P}_1(\mathcal{H})$*, then* $B = \lambda I + (1-\lambda)A$ *for some* $\lambda \in \mathcal{C} \setminus \{1\}$*.*

Lemma 2.3. *let* $A, B \in \mathcal{B}(\mathcal{H})$*. Then* $A \in \mathcal{C}^* I$ *if and only if* $Fix(AR) = \{0\}$*, for every* $R \in \mathcal{P}_1(\mathcal{H})$

Now we state the main result of this paper. The main idea for proving this theorem is taken from [4] (see Theorem B), however a lot of new phenomena takes place.

Theorem 2.4. *Let* $\mathcal{H}$ *be a complex Banach space with dim* $\mathcal{H} \geq 3$ *suppose* $\Phi : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ *is a surjective map with the property that*

$$Fix(\Phi(S) - \Phi(T)) = Fix(S - T), \qquad (S \in \mathcal{B}(\mathcal{H}), T \in \mathcal{F}_1(\mathcal{H}))$$

then $\Phi(S) = k(S)S$*, where* $K : \mathcal{B}(\mathcal{H}) \to \mathcal{C}$ *is a function such that for any non - scalar operator* $S \in \mathcal{B}(\mathcal{H})$ *we have* $k(A) = 1 or - 1$ *and for any scalar operator* λI *we have* $k(\lambda I) = \gamma(\lambda)I$*, where* $\gamma : \mathcal{C} \to \mathcal{C}$ *is a bijective map.*

Proof. step1: $\Phi(I) = 0, \Phi(0) = 0$, and Φ is injective
step2 Φ preserves idempotent operators.
step3: $\Phi(I - T) = I - \Phi(T)$ for any $T \in \mathcal{H}$.
step4: $\Phi(\frac{1}{2}S) = \frac{1}{2}\Phi(S)$, for any idempotent operator S.
step5: Φ preserves the order og idempotent operator.
step6: Φ preserves rank one idempotents . by step 2 and 5 and [7] we find that Φ preserves rank one idempotent.
step7: $\Phi(S - T) = \Phi(S) - \Phi(T)$ for any $S, T \in \mathcal{H}$.
step8: there exists an invertible bounded linear or conjugate linear operator $T : \mathcal{H} \to \mathcal{H}$ such that
$\Phi(P) = k(P)P$ for any non scalar operator $P \in \mathcal{H}$ and $\Phi(\lambda I) = \lambda I$. □

References

[1] Jafarioan A and Sourour A,Spectrum preserving linear maps, J. Funct Anal.66(1986),255-261.
[2] Bourhim A, Mashreghi J and Stepanyan A. Non-linear maps preserving the minimum and surjectivity moduli, Linear Algebra and its Applications 2014; 463: 171-189.
[3] Costara C, Repovs D. Non-linear mappings preserving at least one eigenvalue, Studia Mathematica 2010; 200: 79-89.
[4] Havlicek H, Semrl P. From geometry to invertibility preserves, Studia Mathematica 2006; 174: 99-109.
[5] Oudghiri M, Souilah K. Non-linear maps preserving singular algebraic operators, Proyecciones Journal of Mathematics 2016; 35: 301-316.
[6] Pearcy C, Topping D. Sums of small numbers of idempotents, Michigan Math. J 1967; 14: 453-465.
[7] Semrl P. Symmetries on bounded observables: a unified approach based on adjacency preserving maps, Integr. Equ. Oper. Theory 2012; 72: 7-66.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–293, pp. 160–166

Fixed point theorems in non-Archimedean G-fuzzy metric spaces

Razieh Farokhzad Rostami[a,*]

[a] *Gonbad Kavous University of Iran*

Article Info

Keywords:
Fixed point
Contractive mapping
Weak Contractive mapping
Fuzzy metric space

2020 MSC:
47H10
54H25

Abstract

In this article, we extend some recent fixed point theorems in the setting of G-fuzzy metric spaces. We introduce some new concepts of contractions called γ-contractions and γ-weak contractions. We prove some fixed point theorems for mappings providing γ-contractions and γ-weak contractions. On the other hand, we consider a more general class of auxiliary functions in the contractivity condition.

1. Introduction

Fixed point theory is a very important concept in mathematics. In 1922, Banach created a famous result called Banach contraction principle in the concept of the fixed point theory which states sufficient conditions for the existence and uniqueness of a fixed point[1].

There are two well-known extensions of the notion of metric space in which imprecise models are considered: fuzzy metric spaces (see [7]) and probabilistic metric spaces [3, 8, 9]. The two concepts are very similar, but they are different in nature. The concept of a fuzzy metric space was introduced in different ways by some authors (see [2, 4]). Gregori and Sapena [4] introduced the notion of fuzzy contractive mappings and gave some fixed point theorems for complete fuzzy metric spaces in the sense of George and Veeramani, and also for Kramosil and Michalek's fuzzy metric spaces which are complete in Grabiec's sense. Mihet [6] developed the class of fuzzy contractive mappings of Gregori and Sapena, considered these mappings in non-Archimedean fuzzy metric spaces in the sense of Kramosil and Michalek, and obtained a fixed point theorem for fuzzy contractive mappings. Lots of different types of fixed point theorems has been presented by many authors by expanding the Banach's result, simultaneously (see [10, 11]).

In this work, using a mapping $\gamma : [0, 1) \to \mathbb{R}$ we introduce some new types of contractions called γ-contractions and γ-weak contractions. Later, we prove some fixed point theorems for mappings providing γ-contractions and γ-weak

*Talker
Email address: first.author@gmail.com (Razieh Farokhzad Rostami)

contractions in non-Archimedean G-fuzzy metric spaces. Also, some examples are supplied in order to support the usability of our results. On the other hand, we consider a more general class of auxiliary functions which generate some contractive conditions, and we show that the function $t \to 1/t - 1$ (which appears in many fixed point theorems in the fuzzy context) can be replaced by more appropriate and general functions.

Before proving our main results, we recall some basic definitions and facts which will be used later in this paper.

Definition 1.1. [8] A binary operation $* : [0,1] \times [0,1] \to [0,1]$ is called a continuous triangular norm (in short, continuous t-norm) if it satisfies the following conditions:

(TN-1) $*$ is commutative and associative,
(TN-2) $*$ is continuous,
(TN-3) $*(a,1) = a$ for every $a \in [0,1]$,
(TN-4) $*(a,b) \leq *(c,d)$ whenever $a \leq c$, $b \leq d$ and $a,b,c,d \in [0,1]$.

Definition 1.2. [11] A G-fuzzy metric space is an ordered triple $(X, G, *)$ such that X is a nonempty set, $*$ is a continuous t-norm, and G is a fuzzy set on $X^3 \times (0,\infty)$, satisfying the following conditions, for all $s,t > 0$:

(GF-1) $G(x,x,y,t) < 1$ for all $x,y \in X$ with $x \neq y$,
(GF-2) $G(x,x,y,t) \leq G(x,y,z,t)$ for all $x,y,z \in X$ with $y \neq z$,
(GF-3) $G(x,y,z,t) = 1$, then $x = y = z$,
(GF-4) $G(x,y,z,t) = G(p(x,y,z),t)$, where p is a permutation function,
(GF-5) $G(x,y,z,t+s) \geq G(x,a,a,s) * G(a,y,z,t)$ for all $x,y,z,a \in X$,
(GF-6) $G(x,y,z,.) : (0,\infty) \to [0,1]$ is continuous.

If, in the above definition, the triangular inequality (GF-5) is replaced by

$$G(x,y,z,max\{s,t\}) \geq G(x,a,a,s) * G(a,y,z,t)$$

for all $x,y,z,a \in X$ and $s,t > 0$, or equivalently,

$$G(x,y,z,t) \geq G(x,a,a,t) * G(a,y,z,t) \tag{1}$$

the triple $(X,G,*)$ is called a non-Archimedean G-fuzzy metric space [5].

Example 1.3. Let X be a nonempty set and let G be a G-metric on X. Denote $*(a,b) = ab$ for all $a,b \in [0,1]$. For each $t > 0$, $G(x,y,z,t) = t/(t + G(x,y,z))$ is a G-fuzzy metric on X.

Definition 1.4. Let $\{x_n\}$ be a sequence in a G-fuzzy metric space $(X,G,*)$. We will say that:

- $\{x_n\}$ converges to x if and only if $\lim_{n\to\infty} G(x_n,x_n,x,t) = 1$; i.e., for all $t > 0$ and all $\lambda \in (0,1)$, there exists $n_0 \in \mathbb{N}$ such that $G(x_n,x_n,x,t) > 1 - \lambda$ for all $n \geq n_0$ (in such a case, we will write $\{x_n\} \to x$);
- $\{x_n\}$ is a Cauchy sequence if and only if for all $t > 0$ and all $\lambda \in (0,1)$, there exists $n_0 \in \mathbb{N}$ such that $G(x_n,x_n,x_m,t) > 1 - \lambda$ for all $n,m \geq n_0$. $\{x_n\}$ is a G-Cauchy sequence if and only if for all $t > 0$ and all $\lambda \in (0,1)$, there exists $n_0 \in \mathbb{N}$ such that $G(x_n,x_n,x_{n+p},t) > 1 - \lambda$ for all $n \geq n_0$ and $p > 0$; in other words, $\lim_{n\to\infty} G(x_n,x_n,x_{n+p},t) = 1$.
- The G-fuzzy metric space $(X,G,*)$ is called complete (G-complete) if every Cauchy (G-Cauchy) sequence is convergent.

Lemma 1.5. *(see [11]) Let $(X,G,*)$ be a G-fuzzy metric space. Then, $G(x,y,z,t)$ is nondecreasing with respect to t for all $x,y,z \in X$.*

Lemma 1.6. *(see [11]) Let $(X, G, *)$ be a G-fuzzy metric space. Then, G is a continuous function on $X^3 \times (0, \infty)$.*

It is easy to prove that a $G(x, y, z, t)$ in a non-Archimedean G-fuzzy metric space $(X, G, *)$ is also nondecreasing with respect to t and a continuous function for all $x, y, z \in X$.

2. New types of contractive mappings

Definition 2.1. Let $\gamma : [0, 1) \to \mathbb{R}$ be a strictly increasing continuous mapping and for each sequence $\{a_n\}_{n\in\mathbb{N}}$ of positive numbers $\lim_{n\to\infty} a_n = 1$ if and only if $\lim_{n\to\infty} \gamma(a_n) = \infty$. Let Γ be the family of all γ functions.

Let $(X, G, *)$ be a non-Archimedean G-fuzzy metric space. A mapping $T : X \to X$ is said to be a γ-contraction if there exists a $\delta > 0$ such that

$$G(Tx, Ty, Tz, t) < 1 \Rightarrow \gamma(G(Tx, Ty, Tz, t)) \geq \gamma(G(x, y, z, t)) + \delta \tag{2}$$

for all $x, y, z \in X,\ t > 0$ and $\gamma \in \Gamma$.

When we consider in (2) the different types of the mapping γ, then we obtain a variety of contractions, some of them are of a type known in the literature. See the following example:

Example 2.2. The different types of the mapping $\gamma \in \Gamma$ are as follows:

$$\gamma_1 = \frac{1}{(1-x)}, \quad \gamma_2 = ln\frac{1}{(1-x)}, \quad \gamma_3 = \frac{1}{(1-x)} + x, \quad \gamma_4 = \frac{1}{(1-x^2)}, \quad \gamma_5 = \frac{1}{\sqrt{1-x}}.$$

If $\gamma = ln\frac{1}{(1-x)}$. Then each mapping $T : X \to X$ satisfying (2) is a γ-contraction such that

$$G(Tx, Ty, Tz, t) \geq k(\delta)G(x, y, z, t),$$

for all $x, y, z \in X,\ t > 0$ and $G(Tx, Ty, Tz, t) < 1$, in which $k(\delta) = \frac{G(x,y,z,t)-1+e^{\delta}}{e^{\delta}G(x,y,z,t)} \geq 1$.

Note that from γ and (2) it is easy to conclude that every γ-contraction T is a contractive mapping, that is,

$$G(Tx, Ty, Tz, t) > G(x, y, z, t) \tag{3}$$

for all $x, y, z \in X$, such that $Tx \neq Ty \neq Tz$. Thus every γ-contraction is a continuous mapping.

Now we state one of the main results of the present manuscript.

Theorem 2.3. *Let $G(X, G, *)$ be a complete non-Archimedean G-fuzzy metric space and let $T : X \to X$ be a γ-contraction. Then T has a unique fixed point in X.*

Proof. Let $x_0 \in X$ be arbitrary and fixed. Define sequence $\{x_n\}$ by

$$Tx_n = x_{n+1}, \quad for\ all\ n \in \mathbb{N}. \tag{4}$$

If $x_n = x_{n+1}$, then x_{n+1} is the fixed point of T; then the proof is finished. Suppose that $x_n \neq x_{n+1}$ for all $n \in \mathbb{N}$. Therefore by (2), we get

$$\gamma(G(Tx_{n-1}, Tx_{n-1}, Tx_n, t)) \geq \gamma(G(x_{n-1}, x_{n-1}, x_n, t)) + \delta. \tag{5}$$

Repeating this process, we have

$$\begin{aligned}\gamma(G(Tx_{n-1}, Tx_{n-1}, Tx_n, t)) &\geq \gamma(G(x_{n-1}, x_{n-1}, x_n, t)) + \delta \\ &= \gamma(G(Tx_{n-2}, Tx_{n-2}, Tx_{n-1}, t)) + \delta \\ &\geq \gamma(G(x_{n-2}, x_{n-2}, x_{n-1}, t)) + 2\delta... \\ &\geq \gamma(G(x_0, x_0, x_1, t)) + n\delta.\end{aligned} \tag{6}$$

Letting $n \to \infty$, from (6) we get

$$\lim_{n\to\infty} \gamma(G(Tx_{n-1}, Tx_{n-1}, Tx_n, t)) = +\infty. \tag{7}$$

Then, we have

$$\lim_{n\to\infty} G(Tx_{n-1}, Tx_{n-1}, Tx_n, t) = 1. \tag{8}$$

With the same process, we have

$$\lim_{n\to\infty} G(Tx_{n-1}, Tx_n, Tx_n, t) = 1.$$

Now, we want to show that $\{x_n\}$ is a Cauchy sequence. Suppose to the contrary, we assume that $\{x_n\}$ is not a Cauchy sequence. Then there are $\lambda \in (0, 1)$ and $t_0 > 0$ such that for all $k \in \mathbb{N}$ there exist $n(k), m(k) \in \mathbb{N}$ with $n(k) > m(k) > k$ and

$$G(x_{n(k)}, x_{n(k)}, x_{m(k)}, t_0) \leq 1 - \lambda. \tag{9}$$

Assume that $m(k)$ is the least integer exceeding $n(k)$ satisfying inequality (9). Then, we have

$$G(x_{n(k)}, x_{n(k)}, x_{m(k)-1}, t_0) > 1 - \lambda, \tag{10}$$

and so, for all $k \in \mathbb{N}$ and from (1), we get

$$\begin{aligned} 1 - \lambda &\geq G(x_{n(k)}, x_{n(k)}, x_{m(k)}, t_0) \\ &= G(x_{m(k)}, x_{n(k)}, x_{n(k)}, t_0) \\ &\geq G(x_{m(k)}, x_{m(k)-1}, x_{m(k)-1}, t_0) * G(x_{m(k)-1}, x_{n(k)}, x_{n(k)}, t_0) \\ &\geq G(x_{m(k)}, x_{m(k)-1}, x_{m(k)-1}, t_0) * (1 - \lambda). \end{aligned} \tag{11}$$

Letting $k \to \infty$ in (11) and using (8), we obtain

$$\lim_{k\to\infty} G(x_{n(k)}, x_{n(k)}, x_{m(k)}, t_0) = 1 - \lambda. \tag{12}$$

From (1), we get

$$\begin{aligned} G(x_{m(k)+1}, x_{n(k)+1}, x_{n(k)+1}, t_0) \geq{}& G(x_{m(k)+1}, x_{m(k)}, x_{m(k)}, t_0) \\ &* G(x_{m(k)}, x_{n(k)}, x_{n(k)}, t_0) * G(x_{n(k)}, x_{n(k)+1}, x_{n(k)+1}, t_0), \end{aligned}$$

so, letting $k \to \infty$ and using (8), we have

$$\lim_{k\to\infty} G(x_{m(k)+1}, x_{n(k)+1}, x_{n(k)+1}, t_0) \geq 1 - \lambda. \tag{13}$$

From (9), we obtain

$$\begin{aligned} 1 - \lambda &\geq G(x_{m(k)}, x_{n(k)}, x_{n(k)}, t_0) \\ &\geq G(x_{m(k)}, x_{m(k)+1}, x_{m(k)+1}, t_0) * G(x_{m(k)+1}, x_{n(k)+1}, x_{n(k)+1}, t_0) * \\ &\quad G(x_{n(k)+1}, x_{n(k)}, x_{n(k)}, t_0), \end{aligned} \tag{14}$$

and so by taking the limit as $k \to \infty$ in (14) and from (8) and (13), we have

$$\lim_{k\to\infty} G(x_{m(k)+1}, x_{n(k)+1}, x_{n(k)+1}, t_0) = 1 - \lambda. \tag{15}$$

By applying inequality (2) with $x = y = x_{n(k)}$ and $z = x_{m(k)}$

$$\gamma(G(x_{n(k)+1}, x_{n(k)+1}, x_{m(k)+1}, t_0)) \geq \gamma(G(x_{n(k)}, x_{n(k)}, x_{m(k)}, t_0)) + \delta. \tag{16}$$

Taking the limit $k \to \infty$ in (16), applying (2), from (12), (15), and the continuity of γ, we obtain

$$\gamma(1 - \lambda) \geq \gamma(1 - \lambda) + \delta,$$

which is a contradiction. Thus $\{x_n\}$ is a Cauchy sequence in X. From the completeness of $(X, G, *)$ there exists $x \in X$ such that

$$\lim_{n\to\infty} x_n = x.$$

Finally, the continuity of T and G yields

$$G(Tx, Tx, x, t) = \lim_{n\to\infty} G(Tx_n, Tx_n, x_n, t) = \lim_{n\to\infty} G(x_{n+1}, x_{n+1}, x_n, t) = 1.$$

Now, we show that T has a unique fixed point. Suppose that x and y are two fixed points of T. Indeed, if for $x, y \in X$, $Tx = x \neq y = Ty$, then we get

$$\gamma(G(x, x, y, t)) \geq \gamma(G(x, x, y, t)) + \delta,$$

which is a contradiction. Thus, T has a unique fixed point. Hence, the proof is completed. □

Example 2.4. Let $X = [0, 1)$, $*(a, b) = min\{a, b\}$, and

$$G(x, y, z, t) = \begin{cases} 1, & \text{if } x = y = z, \\ \frac{1}{1+max\{x,y,z\}}, & \text{otherwise.} \end{cases} \tag{17}$$

for all $t > 0$. Let $\gamma : [0, 1) \to \mathbb{R}$ such that $\gamma(x) = 1/1 - x$ for all $x \in [0, 1)$ and define $T : X \to X$ by $T(x) = 2x^2/5$ for all $x \in X$. Clearly, $(X, G, *)$ is a complete non-Archimedean G-fuzzy metric space.
Case 1. We assume that $x, y, z \in (0, 1)$. Since $x^2 < x,\ y^2 < y$ and $z^2 < z$, then $max\{x, y, z\} > max\{Tx, Ty, Tz\}$. So, there exists a $\delta > 0$ such that

$$\frac{1}{max\{Tx, Ty, Tz\}} + 1 \geq \frac{1}{max\{x, y, z\}} + 1 + \delta.$$

It is easy to see that

$$\gamma(G(Tx, Ty, Tz, t)) \geq \gamma(G(x, y, z)) + \delta.$$

Case 2. Let $x = 0$ and $y, z \in (0, 1)$. Since $x^2 = 0,\ y^2 < y$ and $z^2 < z$, then $max\{x, y, z\} = max\{y, z\} > max\{Tx, Ty, Tz\} = max\{Ty, Tz\}$. Hence, we have

$$G(Tx, Ty, Tz, t) = \frac{1}{1 + max\{Tx, Ty, Tz\}} > \frac{1}{1 + max\{x, y, z\}} = G(x, y, z, t).$$

So, there exists a $\delta > 0$ such that

$$\gamma(G(Tx, Ty, Tz, t)) \geq \gamma(G(x, y, z, t)) + \delta.$$

Case 3. Let $x = y = 0$ and $z \in (0, 1)$, it is easy to see that,

$$\gamma(G(Tx, Ty, Tz, t)) \geq \gamma(G(x, y, z, t)) + \delta.$$

Therefore, T is a γ-contraction. Then all the conditions of Theorem (2.3) hold and T has the unique fixed point $x = 0$.

Definition 2.5. Let $(X, G, *)$ be a non-Archimedean G-fuzzy metric space. A mapping $T : X \to X$ is said to be a γ-weak contraction if there exists a $\delta > 0$ such that

$$\begin{gathered} G(Tx, Ty, Tz, t) < 1 \Rightarrow \\ \gamma(G(Tx, Ty, Tz, t)) \geq \\ \gamma(min\{G(x, y, z, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}) + \delta, \end{gathered} \tag{18}$$

for all $x, y, z \in X$ and $\gamma \in \Gamma$.

Note that every γ-contraction is a γ-weak contraction. But the converse is not true.

Example 2.6. Let $X = A \cup B$, where $A = \{1/10, 1/2, 1, 2, 3\}$, $B = [4, 5]$. $*(a, b) = min\{a, b\}$ and $G(x, y, x, t) = min\{x, y, z\}/max\{x, y, z\}$ for all $t > 0$. Clearly, $(X, G, *)$ is a complete non-Archimedean G-fuzzy metric space. Let $\gamma : [0, 1) \to \mathbb{R}$ such that $\gamma(x) = 1/\sqrt{1-x}$ for all $x \in [0, 1)$ and define $T : X \to X$ by

$$\begin{cases} \frac{1}{10}, & \text{if } x \in A, \\ \frac{1}{2}, & \text{if } x \in B. \end{cases}$$

Since T is not continuous, T is not γ-contraction by (3).

Now, we show that T is a γ-weak contraction for all $x \in X$.
Case 1. Let $x = 1$ and $y, z \in B$,

$$G(Tx, Ty, Tx, t) = \frac{1}{5} > \frac{1}{10} = min\{\frac{1}{max\{y, z\}}, \frac{1}{10}, \frac{1}{2y}, \frac{1}{2z}\} = \\ min\{G(x, y, x, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}.$$

So, there exists a $\delta > 0$ such that

$$\gamma(G(Tx, Ty, Tz, t)) \geq \\ \gamma(min\{G(x, y, z, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}) + \delta.$$

Case 2. Let $x \in \{2, 3\}$ and $y, z \in B$,

$$G(Tx, Ty, Tx, t) = \frac{1}{5} > \frac{1}{10x} = min\{\frac{x}{max\{y, z\}}, \frac{1}{10x}, \frac{1}{2y}, \frac{1}{2z}\} = \\ min\{G(x, y, x, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}.$$

So, there exists a $\delta > 0$ such that

$$\gamma(G(Tx, Ty, Tz, t)) \geq \\ \gamma(min\{G(x, y, z, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}) + \delta.$$

Case 3. Let $x \in \{1/10, 1/2\}$ and $y, z \in B$,

$$G(Tx, Ty, Tx, t) = \frac{1}{5} > \frac{x}{max\{y, z\}} = min\{\frac{x}{max\{y, z\}}, \frac{1}{10}, \frac{1}{2y}, \frac{1}{2z}\} = \\ min\{G(x, y, x, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}.$$

So, there exists a $\delta > 0$ such that

$$\gamma(G(Tx, Ty, Tz, t)) \geq \\ \gamma(min\{G(x, y, z, t), G(x, x, Tx, t), G(y, y, Ty, t), G(z, z, Tz, t)\}) + \delta.$$

By proving the rest of cases, we get T is a γ-weak contraction.

Theorem 2.7. *Let $(X, G, *)$ be a complete non-Archimedean G-fuzzy metric space and let $T : X \to X$ be a γ-weak contraction. Then T has a unique fixed point in X.*

Example 2.8. Let $(X, G, *)$ be the non-Archimedean G-fuzzy metric space and let T be considered in Example (2.6). Let $\gamma : [0, 1) \to \mathbb{R}$ such that $\gamma(x) = 1/(1 - x^2)$ for all $x \in [0, 1)$. So, T is a γ-weak contraction. Therefore, Theorem (2.7) can be applicable to T and the unique fixed point of T is 1/10.

References

[1] S. Banach, *Sur les opérations dans les ensembles abstraits etleur application aux équations intégrales*, Fundamenta Mathematicae, vol. 3, (1922), 133–181.
[2] Z. Deng, *Fuzzy pseudo-metric spaces*, Journal of Mathematical Analysis and Applications, 86 no. 1 (1982), 74-95.
[3] M. Grabiec, YJ. Cho and V. Radu, *On Nonsymmetric Topological and Probabilistic Structures*, Nova Science Publishers, New York, 2006.
[4] V. Gregori and A. Sapena, *On fixed-point theorems in fuzzy metric spaces, Fuzzy Sets and Systems*, 125 no. 2 (2002), 245-252.
[5] V. Istratescu, *An Introduction to Theory of Probabilistic Metric Spaces with Applications*, Politehnica University of Bucharest, Bucharest, Romania, 1974.
[6] D. Mihet, *Fuzzy ψ-contractive mappings in non-Archimedean fuzzy metric spaces*, Fuzzy Sets and Systems, 159 no. 6 (2008), 739-744.
[7] A. Roldán, J. Martinez-Moreno and C. Roldán, *On interrelationships between fuzzy metric structures*, Iran. J. Fuzzy Syst., vol. 10, (2013), 133-150.
[8] B. Schweizer and A. Sklar, *Statistical metric spaces*, Pacific Journal of Mathematics, vol. 10, (1960), 313–334.
[9] B. Schweizer and A. Sklar, *Probabilistic Metric Spaces*, Dover, New York, 2005.
[10] S. Sedghi, N. Shobkolaei, T. Došenović, and S. Radenović, *Suzuki-type of common fixed point theorems in fuzzy metric spaces*, Mathematica Slovaca, 68 no. 2 (2018), 451-462.
[11] G. Sun and K. Yang, *Generalized fuzzy metric spaces with properties*, Research journal of Applied Sciences, Engineering and Technology, vol. 2, no. 7, (2010), 673–678.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–294, pp. 167–172

Modified Homotopy Perturbation method for Solving time-Fractional Fisher equation

M. Eslami[a,*]

[a] *Department of Applied Mathematics, Faculty of Mathematical Sciences, University of Mazandaran, Babolsar, Iran*

Article Info

Keywords:
Efficient method
fractional partial differential equations

2020 MSC:
msc1
msc2

Abstract

In this paper, a novel algorithm based on new modified Homotopy Perturbation Method, called NHPM, for fractional fisher equations was proposed. The solution process is elucidated including how to construct a suitable homotopy equation and how to choose an initial solution. Some examples are given to reveal the effectiveness and convenience of the method.

1. Introduction

In recent years, considerable research and interest in fractional differential equations has been stimulated due to their numerous applications in many areas like physics, and engineering [1]. Many important phenomena in electromagnetics, acoustics, viscoelasticity, electrochemistry, corrosion and material science are well described by differential equations of fractional order [2, 3].
Many new numerical techniques have been widely applied to fractional differential equations. Based on homotopy, which is a basic concept in topology, general analytical method namely the homotopy perturbation method (HPM) was established by He [4–7] in the year 1998, to obtain a series of solutions to the nonlinear differential equations. This simple method has been applied to solve Blasius equation [8], fluid mechanics equations [9], fractional KdV-Burgers equation [10], some boundary value problems, and many other equations in subjects of different disciplines [11–20]. In this study, a new version of the HPM, which efficiently solves fractional differential equations, is being introduced.

2. Basic definitions

In this section some basic definitions and properties of the fractional calculus theory used in this work will be discussed [21, 22].

Definition 2.1. A real function $f(x), x > 0$, in the space $C_\mu, \mu \in R$ if there exists a real number $p > \mu$, such that $f(x) = x^p f_1(x)$ where $f_1(x) \in C[0, \infty]$ and it is said to be in the space C_μ^m if $f^{(m)} \in C_\mu, m \in N$.

*Talker
Email address: mostafa.eslami@umz.ac.ir (M. Eslami)

Definition 2.2. The Riemann–Liouville fractional integral operator of order $\alpha \geq 0$, of a function $f \in C_\mu, \mu \geq -1$, is defined as:

$$J^\alpha f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} f(t)\, dt, \qquad \alpha > 0,\ x > 0,\ J^0 f(x) = f(x).$$

The general and detailed properties of the operator J^α can be found in reference [22]. For this study, where $f \in C_\mu$,$\mu \geq -1$,$\alpha, \beta \geq 0$ and $\gamma > -1$:

$$\begin{aligned}
&(1) \quad J^\alpha J^\beta f(x) = J^{\alpha+\beta} f(x),\\
&(2) \quad J^\alpha J^\beta f(x) = J^\beta J^\alpha f(x),\\
&(3) \quad J^\alpha x^\gamma = \frac{\Gamma(\gamma+1)}{\Gamma(\alpha+\gamma+1)} x^{\alpha+\gamma}.
\end{aligned}$$

It is worth mentioning here that, the Riemann–Liouville derivative method has some disadvantages when used to model real-world phenomena with fractional differential equations. Therefore, a modified fractional differential operator, D_*^α , should be introduced to overcome those weaknesses in the previous models. Such modified fractional differential operators, D_*^α , were first proposed by Caputo [22], in his work on the theory of viscoelasticity.

Definition 2.3. The fractional derivative of $f(x)$ according to Caputo [22], is defined as:

$$D_*^\alpha f(x) = J^{m-\alpha} D^m f(x) = \frac{1}{\Gamma(m-\alpha)} \int_0^x (x-t)^{m-\alpha-1} f^{(m)}(t)\, dt,$$

for: $m-1 < \alpha \leq m, m \in \mathbb{N}, x > 0, f \in C_{-1}^m$.

The following two properties of this operator will be used in what comes next.

Lemma 2.4. *If $m-1 < \alpha \leq m$ and $f \in C_\mu^m, \mu \geq -1$, then $D_*^\alpha J^\alpha f(x) = f(x)$, and* $J^\alpha D_*^\alpha f(x) = f(x) - \Sigma_{k=0}^{m-1} f(0^+) \frac{x^k}{k!}, x > 0$.

Definition 2.5. For m as the smallest integer that exceeds α , the Caputo time-fractional derivative operator of order $\alpha > 0$, is defined as:

$$D_{*t}^\alpha u(x,t) = \frac{\partial^\alpha u(x,t)}{\partial t^\alpha} = \begin{cases} \frac{1}{\Gamma(m-\alpha)} \int_0^t (t-\tau)^{m-\alpha-1} \frac{\partial^m u(x,\tau)}{\partial \tau^m} d\tau, & m-1 < \alpha < m, \\ \frac{\partial^m u(x,t)}{\partial t^m}, & \alpha = m \in \mathbb{N}. \end{cases}$$

For more information on the mathematical properties of fractional derivatives and integrals, one can consult the above mentioned references.

3. Theory of the Method

To illustrate the application and methodology of using the proposed new method, the following fractional differential equation will be considered:

$$A(u(X,t) - f(r)) = 0, r \in \Omega, \tag{1}$$

$$B\left(u(X,t), \frac{\partial u}{\partial n}\right) = 0, r \in \Gamma, \tag{2}$$

where A is a general differential operator, $f(r)$ is a known analytic function, B is a boundary condition, Γ is the boundary of the domain Ω , and $X = (x_1, x_2, \ldots, x_n)$.
In general, the operator A can be divided into two operators, L and N, where L is a linear operator, while N is a non-linear operator.

In this case, equation 1 can be re-written as follows:

$$L(u) + N(u) - f(r) = 0 \tag{3}$$

Using the homotopy technique, a homotopy $U(r,p) : \Omega \times [0,1] \to \mathbb{R}$ could be constructed, which satisfies: In this case, equation 1 can be re-written as follows:

$$H(U,p) = (1-p)[L(U) - L(u_0)] + p[A(U) - f(r)] = 0, \qquad p \in [0,1],\ r \in \Omega, \tag{4}$$

or

$$H(U,p) = L(U) - L(u_0) + pL(u_0) + p[N(U) - f(r)] = 0. \tag{5}$$

Where $p \in [0,1]$, is called homotopy parameter, and u_0 is an initial approximation for the solution of Eq.1, which satisfies the boundary conditions.
Obviously from Eq.4 and Eq.5, Eq.6 and Eq.7 could be derived and written as:

$$H(U,0) = L(U) - L(u_0) = 0, \tag{6}$$

$$H(U,1) = A(U) - f(r) = 0. \tag{7}$$

It is assumed that the solution of Eq.6 or Eq.7 could be expressed as a series in p, as follows:

$$U = U_0 + pU_1 + p^2U_2 + \dots \tag{8}$$

Setting $p = 1$, produces the approximate solution of Eq.1, which could be written in the following form:

$$u = \lim_{p\to 1} U = U_0 + U_1 + U_2 + K$$

Now Eq.5 will be written in the following form:

$$L(U(X,t)) = u_0(X,t) + p[f(r(X,t)) - u_0(X,t) - N(U(X,t))]. \tag{9}$$

By applying the inverse operator, L_1, to both sides of Eq.9, Eq.10 could be derived:

$$U(X,t) = L^{-1}(u_0(X,t)) + p(L^{-1}(f(r)) - L^{-1}(u_0(X,t) - L^{-1}(N(U(X,t))). \tag{10}$$

Suppose that the initial approximation of Eq.1 has the form:

$$u_0(X,t) = \Sigma_{n=0}^{\infty} a_n(X) P_n(t) \tag{11}$$

where $a_1(X)$,$a_2(X)$,$a_3(X)$,... are unknown coefficients, and $P_0(t)$,$P_1(t)$,$P_2(t)$,... are specific functions dependent on the problem. Now by substituting Eq.8 and Eq.11 into Eq.10, we get:

$$\begin{aligned}\Sigma_{n=0}^{\infty} U_n(X,t) = U(X,t) = L^{-1}\left(\Sigma_{n=0}^{\infty} a_n(X) P_n(t)\right) + \\ p\left(L^{-1}(f(r)) - L^{-1}\left(\Sigma_{n=0}^{\infty} a_n(X) P_n(t)\right) - L^{-1}\left(N\Sigma_{n=0}^{\infty} p^n U_n(X,t)\right)\right).\end{aligned} \tag{12}$$

Comparing the coefficients of terms with the identical powers of p, leads to:

$$\begin{aligned}
&p^0 : U_0(X,t) = L^{-1}(\Sigma_{n=0}^{\infty} a_n(X)P_n(t)),\\
&p^1 : U_1(X,t) = L^{-1}(f(r)) - L^{-1}(\Sigma_{n=0}^{\infty} a_n(X)P_n(t)) - L^{-1}N(U_0(X,t)),\\
&p^2 : U_2(X,t) = -L^{-1}N(U_0(X,t), U_1(X,t)),\\
&p^3 : U_3(X,t) = -L^{-1}N(U_0(X,t), U_1(X,t), U_2(X,t)),\\
&M\\
&p^j : U_j(X,t) = -L^{-1}N(U_0(X,t), U_1(X,t), U_2(X,t), K, U_{j-1}(X,t)),\\
&M
\end{aligned} \tag{13}$$

Now, if the above equations are solved in such a way that $U_1(X,t)=0$, then Eq.13 results in:

$$U_1(X,t)=U_2(X,t)=\cdots=0,$$

Therefore, the exact solution may be obtained as follows:

$$u(X,t)=U_0(X,t)=L^{-1}\left(\Sigma_{n=0}^{\infty}a_n(X)P_n(t)\right),$$

To show the capability of this method, it will be applied to some examples in the next section.

4. Illustrative examples

Example

Consider the following fractional Fisher equation:

$$\frac{\partial^\alpha u}{\partial t^\alpha}=\frac{\partial^2 u}{\partial x^2}+u(1-u), \qquad 0<\alpha\leq 1. \tag{14}$$

with a constant initial condition:

$$u(x,0)=\lambda.$$

For solving Eq.**??**, the following homotopy should be constructed:

$$\frac{\partial^\alpha U}{\partial t^\alpha}(x,t)=u_0(x,t)-p\left(u_0(x,t)+\frac{\partial^\alpha U}{\partial t^\alpha}-\frac{\partial^2 U}{\partial x^2}-U(1-U)\right), \tag{15}$$

Applying the inverse operator, J_t^α to both sides of the above equation, results in:

$$U(x,t)=U(x,0)+J_t^\alpha u_0(x,t)-J_t^\alpha\left(u_0(x,t)+\frac{\partial^\alpha U}{\partial t^\alpha}-\frac{\partial^2 U}{\partial x^2}-U(1-U)\right), \tag{16}$$

Suppose the solution of Eq.16 have the form shown in Eq.8, then substituting Eq.8 into Eq.16 and equating the coefficients of p with the same power, leads to:

$$\begin{aligned}
&p^0: U_0(x,t)=U(x,t)+J_t^\alpha u_0(x,t),\\
&p^1: U_1(x,t)=,\\
&p^2: U_2(x,t)=,\\
&M\\
&p^j: U_j(x,t)=,\\
&M
\end{aligned}$$

Assuming $u_0(x,t)=\Sigma_{n=0}^{\infty}a_n(x)t^{\alpha k}$, $U(x,0)=u(x,0)$, and solving the above equation for $U_1(x,t)$, leads to the following result:

$$\begin{aligned}
U_1(x,t)=&\left(-\frac{1}{\Gamma(\alpha+1)}a_0(x)+\frac{\lambda}{\Gamma(\alpha+1)}-\frac{\lambda^2}{\Gamma(\alpha+1)}\right)t^\alpha\\
&+\left(-\frac{\Gamma(\alpha+1)}{\Gamma(2\alpha+1)}a_1(x)+\frac{1}{\Gamma(2\alpha+1)}a_0''(x)+\frac{1}{\Gamma(2\alpha+1)}a_0(x)-2\lambda\frac{1}{\Gamma(2\alpha+1)}a_0(x)\right)t^{2\alpha}\\
&+\begin{pmatrix}-\frac{\Gamma(2\alpha+1)}{\Gamma(3\alpha+1)}a_2(x)+\frac{\Gamma(\alpha+1)}{\Gamma(3\alpha+1)}a_1''(x)+\frac{\Gamma(\alpha+1)}{\Gamma(3\alpha+1)}a_1(x)\\ -2\lambda\frac{\Gamma(\alpha+1)}{\Gamma(3\alpha+1)}a_1(x)-\frac{\Gamma(2\alpha+1)}{\Gamma(\alpha+1)\Gamma(3\alpha+1)\Gamma(\alpha+1)}a_0^2(x)\end{pmatrix}t^{3\alpha}\\
&+\ldots
\end{aligned}$$

Furthermore, if it is assumed that $U_1(x,t)=0$, then:

$$a_0(x)=\lambda(1-\lambda), a_1(x)=\frac{\lambda(1-\lambda)(1-2\lambda)}{\Gamma(\alpha+1)}, a_2(x)=-\frac{\lambda(1-\lambda)(1-2\lambda)^2}{\Gamma(2\alpha+1)}+\frac{(\lambda-2\lambda)^2}{(\Gamma(\alpha+1))^2},\dots$$

Therefore, the solution of the fractional differential equation can be expressed as follows:

$$\begin{aligned} u(x,t)=U_0(x,t) &= \frac{\lambda(1-\lambda)}{\Gamma(\alpha+1)}t^{\alpha}+\frac{\lambda(1-\lambda)(1-2\lambda)}{\Gamma(2\alpha+1)}t^{2\alpha} \\ &+\left(-\frac{\lambda(1-\lambda)(1-2\lambda)^2}{\Gamma(2\alpha+1)}+\frac{(\lambda-2\lambda)^2}{(\Gamma(\alpha+1))^2}\right)\frac{\Gamma(2\alpha+1)}{\Gamma(3\alpha+1)}t^{3\alpha}+\dots, \end{aligned}$$

For the special case $\alpha=1$, the solution will be as follows:

$$\begin{aligned} u(x,t) &= \lambda(1-\lambda)t^{\alpha}+\frac{\lambda(1-\lambda)(1-2\lambda)}{2}t^2 \\ &+\left(-\frac{\lambda(1-\lambda)(1-2\lambda)(1-6\lambda+6\lambda^2)}{6}\right)t^3+\dots \\ &= \frac{\lambda e^t}{1-\lambda+\lambda e^{t'}} \end{aligned}$$

which is an exact solution.

5. Results and discussion

In this manuscript, a novel algorithm for solving fractional differential equations was successfully developed and tested. The proposed method is simple and it finds exact solution to all equations using initial condition only. This method is also very powerful in finding solutions to various types of physical problems in many important practical applications. One of the other main advantages of this method is its fast convergence to the solution.

6. Conclusion

We note that novel algoritm solutions were computed via a simple algorithm and without any need for perturbation techniques, special transformations, linearization, or discretization. Thus, it can be concluded that the new method is an effective numerical tool for solving functional equations. All computations are performed by using Maple 15.

References

[1] B.J. West, M. Bolognab, P. Grigolini, Physics of Fractal Operators, Springer, New York, 2003.
[2] K.S. Miller, B. Ross, An Introduction to the Fractional Calculus and Fractional Differential Equations, Wiley, New York, 1993.
[3] S.G. Samko, A.A. Kilbas, O.I. Marichev, Fractional Integrals and Derivatives: Theory and Applications, Gordon and Breach, Yverdon, 1993.
[4] J.H.He, Homotopy perturbation technique. Comput Meth Appl Mech Eng 1999;178:257–62.
[5] J.H.He,. A coupling method of homotopy technique and perturbation technique for nonlinear problems. Int J Non-linear Mech 2000;35(1):37–43.
[6] J.H.He, New interpretation of homotopy perturbation method. Int J Mod Phys B 2006;20:2561–8.
[7] J.H.He, Recent development of homotopy perturbation method. Topol Meth Nonlinear Anal 2008;31:205–9.
[8] S. Abbasbandy, A numerical solution of Blasius equation by Adomian's decomposition method and comparison with homotopy perturbation method. Chaos Soliton Fract 2007;31:257–60.
[9] A.M. Siddiqui, A. Zeb, Q.K.Ghori, Homotopy perturbation method for thin film flow of a fourth grade fluid down a vertical cylinder. Phys Lett A 2006;352:404–10.
[10] Q.i. Wang Homotopy perturbation method for fractional KdV-Burgers equation. Chaos Soliton Fract 2008;35(5):843–50.
[11] J. Biazar, M. Eslami, and H. Ghazvini, "Homotopy perturbation method for systems of partial differential equations", International Journal of Nonlinear Sciences and Numerical Simulation, 8 (3)(2007) 411-16.D.D.

[12] J. Biazar, M. Eslami, Approximate solutions for Fornberg-Whitham type equations, International Journal of Numerical Methods for Heat & Fluid Flow, 22: (6) 803 – 812.

[13] A. M. A. El-Sayed, A. Elsaid, I. L. El-Kalla, and D. Hammad, "A homotopy perturbation technique for solving partial differential equations of fractional order in finite domains," Applied Mathematics and Computation, vol. 218, no. 17, pp. 8329–8340, 2012.

[14] S. Abbasbandy, Numerical solutions of the integral equations: homotopy perturbation and Adomian's decomposition method. Appl Math Comput 2006;173:493–500.

[15] F. I. Compean, D. Olvera, F. J. Campa, L. N. Lopez de Lacalle, A. Elias-Zuniga, and C. A. Rodriguez, "Characterization and stability analysis of a multivariable milling tool by the enhanced multistage homotopy perturbation method," International Journal of Machine Tools and Manufacture, 57, 27–33, 2012.

[16] U. Filobello-Nino, H. Vazquez-Leal, R. Castaneda-Sheissa et al., "An approximate solution of blasius equation by using hpm method," Asian Journal of Mathematics and Statistics, 5, 50–59, 2012.

[17] Y.-G. Wang, W.-H. Lin, and N. Liu, "A homotopy perturbation-based method for large deflection of a cantilever beam under a terminal follower force," International Journal for Computational Methods in Engineering Science and Mechanics, 13, 197–201, 2012.

[18] L.Cveticanin, Homotopy-perturbation method for pure nonlinear differential equation. Chaos Soliton Fract 2006;30:1221–30.

[19] Turgut Ozis, Ahmet Yıldırım. A note on He's homotopy perturbation method for van der Pol oscillator with very strong nonlinearity. Chaos Soliton Fract 2007;34(3):989–91.

[20] E. Yusufoglu, Homotopy perturbation method for solving a nonlinear system of second order boundary value problems. Int J Nonlinear Sci Numer Simul 2007;8:353–8.

[21] I. Podlubny, Fractional Differential Equations, Academic Press, San Diego, 1999.

[22] M. Caputo, Linear models of dissipation whose Q is almost frequency independent, Part II, J. Roy. Astr. Soc. 13 (1967) 529.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–133, pp. 173–182

B-Splines Operational Matrix Method for Solving Fractional Volterra Integral Equations

Zohreh Abdollahi[a,*], M. Mohseni Moghadam[a], H. Saeedi[b], M. J. Ebadi[c]

[a] Department of Mathematics, Kerman Branch Islamic Azad University, Kerman, Iran
[b] Department of Applied Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran
[c] Department of Mathematics, Chabahar Maritime University, Chabahar, Iran

Article Info

Keywords:
Wavelets
B-splines
Fractional Volterra integral equation

Abstract

In this paper, we propose an efficient matrix method based on B-spline scaling functions (in short B-splines) for the numerical solution of linear fractional Volterra integral equations. B-splines are preferred for function approximation due to their regularity, symmetry, compact support, and the best approximation properties in comparison to wavelets of the same order. By deriving the fractional integration operational matrix and employing a matrix-vector representation, we reduce the linear integral equation to a linear system of algebraic equations. Numerical experiments are included to demonstrate the effectiveness and accuracy of the method.

1. Introduction

Fractional calculus is a branch of mathematics that studies integral and derivatives of non-integer order. Integral and differential equations of fractional order have received lots of attention from mathematicians, physicians and engineers over recent years. The main reason is that these equations appear in a large number of physical and engineering phenomena like electric transmission lines [9], propagation of sound waves [10], viscoelasticity theory [5, 30], fluid mechanics [12, 16, 21] statistical mechanics [19], polymer [20] and water movement in soils [31]. Moreover, some initial and boundary value problems associated with both ODEs and PDEs can be converted into integral equations. Volterra was the first person who worked systematically on the theory of integral equations [1]. The general form of linear Volterra integral equations is as follows:

$$w(x)f(x) = g(x) + \lambda \int_0^x K(x,t)f(t)dt,$$

where $w(x)$ and $g(x)$ are known functions, λ is a constant and $K(x,t)$ is the kernel. If the kernel $K(x,t)$ is infinite at some points of the integration domain, then the integral equation is called singular. The Riemann–Liouville fractional

*Talker
Email addresses: abdolahi_zohre@yahoo.com (Zohreh Abdollahi), mohseni@uk.ac.ir (M. Mohseni Moghadam), saeedi@uk.ac.ir (H. Saeedi), ebadi@cmu.ac.ir (M. J. Ebadi)

integral of order α of a function $f(x)$ is defined as

$$I^{\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} f(t) dt,$$

where $\alpha > 0$ and Γ is the Gamma function. In view of this definition, the following form of integral equations is called fractional Volterra integral equation which is a singular integral equation:

$$f(x) = g(x) + \mu(x) I^{\alpha} \left(k(x,t) f(x)\right) = g(x) + \frac{\mu(x)}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} k(x,t) f(t) dt. \tag{1}$$

As it is very difficult or sometimes impossible to analytically solve fractional differential and integral equations, various numerical methods have been developed in the literature for solving these equations. The authors in [28], considered Legendre multi-wavelet functions as a basis for approximating the solution of a parabolic differential equation with an unknown time-dependent coefficient in an inverse problem. In [38], temperature variations in a two-dimensional field is simulated using three distinct deep learning-based methods including a convolutional neural network (CNN)-based model, a model utilizing convolutional kernels that capture the desired pattern of the phenomenon, and a reduced-dimension model employing the autoencoder technique. In [23], the authors introduce a highly effective numerical approach for solving first-order nonlinear singularly perturbed differential equations. They utilize a hybrid block scheme with four hybrid points on a non-uniform mesh. In addition, among the methods suggested to numerically solve fractional differential and integral equations are a method based on operational matrix of triangular functions [25], B-spline operational matrix method [13], Haar wavelet operational matrix [24], block-pulse operational matrix method [32], linear B-spline scaling function operational matrix method [26], operational matrix of Legendre functions [2], finite element methods of high-order [14], sinc-Legendre collocation method [22], Jacobi-Gauss-Lobatto collection method [6], cubic B-spline wavelet collocation method [18], semi-analytical Adomain decomposition method (ADM) [8, 29], iterative methods [7, 12], sumudu decomposition method [4], Bernstein approximation method [33], Bernoulli wavelet least squares support vector regression [34], Genocchi wavelet neural networks and least squares support vector regression [35], orthonormal Bernoulli wavelets neural network method [36] and an improved composite collocation method based on fractional Chelyshkov wavelets [37]. B-spline scaling functions have explicit formulae and they have the properties of regularity, symmetry and compact support. These special properties encouraged us to examine their application in numerically solving the fractional Volterra integral equation (1). Indeed, we construct a linear B-spline wavelet-based operational method based on the linear B-spline wavelet operational matrix of fractional integration (OMFI). We then derive a matrix technique to convert the original linear problem into a linear system of algebraic equations. Finally, we propose another approach that converts the original fractional integral equation to an optimization problem. This paper is organized as follows: In Section 2, the linear B-spline scaling functions and wavelets, as well as their relationship are presented and an error analysis for function approximation using B-splies wavelets is given. By obtaining the linear B-spline wavelet operational method (LBWOM) in Section 3, we solve a class of fractional Volterra integral equations. In Section 4, the efficiency and accuracy of the proposed method are verified by presenting some examples. Finally, Section 5 is devoted to some conclusions.

2. Preliminaries

This section is devoted to B-splines and wavelets, function approximation and its convergence analysis.

2.1. B-splines and wavelets

Wavelet families arise from expansion as well as transfer of a function ψ, call mother wavelet. As a result of the continuous change of the expansion and transfer parameters, the following continuous wavelet families are raised:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right), \quad a, b \in \mathbb{R}, a \neq 0,$$

in which, a and b are respectively expansion and transfer parameters. If the parameters a and b are limited to the discrete values, i.e. $a_0 > 1, b_0 > 0, a = a_0^{-j}, b = kb_0a_0^{-j}$, where k is a positive integer number, then we can write

$$\psi_{a,b}(t) := \psi_{j,k}(t) = |a_0|^{j/2}\psi\left(a_0^j t - kb_0\right). \tag{2}$$

Note that $\psi_{j,k}(t)$ is a basis wavelet for $L^2(\mathbb{R})$.
Now, according to [11], by constructing semi-orthogonal wavelets from m-order B-splines, the lowest octave level $j = j_0$ is as follows:

$$2^{j_0} \geq 2m - 1, \tag{3}$$

which gives a minimum of a complete wavelet set on the range [0, 1]. In the current research, we utilize a wavelet constructed by linear splines, i.e., cardinal B-spline basis functions of second order. Using (3), the lowest-level B-spline of second order that should be an integer, is calculated for $j_0 = 2$, which limits all octave levels to $j \geq 2$. As it is also true about semi-orthogonal wavelets, the B-splines of second-order can also act as scaling functions. Equations (4)–(6) below give the B-spline/scaling functions of second-order as:

$$\phi_{j,k}(x) = \begin{cases} 2 - (x_j - k), & 0 \leq x_j \leq 1, \ k = -1, \\ 0, & otherwise. \end{cases} \tag{4}$$

$$\phi_{j,k}(x) = \begin{cases} x_j - k, & k \leq x_j \leq k+1, \\ 2 - (x_j - k), & k+1 \leq x_j \leq k+2, \ k = 0, \cdots, 2^j - 2, \\ 0, & otherwise. \end{cases} \tag{5}$$

$$\phi_{j,k}(x) = \begin{cases} x_j - k, & k \leq x_j \leq k+1, \ k = 2^j - 1, \\ 0, & otherwise. \end{cases} \tag{6}$$

Note that, by setting $x_j = 2^j x$ the actual coordinate position x is associated with x_j. Moreover, the following equations give the B-spline wavelets of second-order:

$$\psi_{j,k}(x) = \frac{1}{6}\begin{cases} -6 + 23x_j, & 0 \leq x_j \leq \frac{1}{2}, \\ 14 - 17x_j, & \frac{1}{2} \leq x_j \leq 1, \\ -10 + 7x_j, & 1 \leq x_j \leq \frac{3}{2}, \ k = -1, \\ 2 - x_j, & \frac{3}{2} \leq x_j \leq 2, \\ 0, & otherwise. \end{cases} \tag{7}$$

$$\psi_{j,k}(x) = \begin{cases} x_j - k, & k \leq x_j \leq k + \frac{1}{2}, \\ 4 - 7(x_j - k), & k + \frac{1}{2} \leq x_j \leq k+1, \\ -19 + 16(x_j - k), & k+1 \leq x_j \leq k + \frac{3}{2}, \\ 29 - 16(x_j - k), & k + \frac{3}{2} \leq x_j \leq k+2, \ k = 0, \cdots, 2^j - 3, \\ -17 + 7(x_j - k), & k+2 \leq x_j \leq k + \frac{5}{2}, \\ 3 - (x_j - k), & k + \frac{5}{2} \leq x_j \leq k+3, \\ 0, & otherwise. \end{cases} \tag{8}$$

$$\psi_{j,k}(x) = \begin{cases} 2 - (k + 2 - x_j), & k \leq x_j \leq k + \frac{1}{2}, \\ -10 + 7(k + 2 - x_j), & k + \frac{1}{2} \leq x_j \leq k+1, \\ 14 - 17(k + 2 - x_j), & k+1 \leq x_j \leq k + \frac{3}{2}, \ k = 2^j - 2, \\ -6 + 23(k + 2 - x_j), & k + \frac{3}{2} \leq x_j \leq k+2, \\ 0, & otherwise. \end{cases} \tag{9}$$

2.2. Function approximation

Using B-spline wavelets, a function $f(x)$ may be expanded as [15]

$$f(x) = \sum_{k=-1}^{3} c_k\phi_{2,k} + \sum_{i=2}^{\infty}\sum_{j=-1}^{2^i-2} d_{i,j}\psi_{i,j}, \tag{10}$$

where $\phi_{2,k}$ are scaling functions and $\psi_{i,j}$ are wavelets functions. By truncating the infinite series in (10), $f(x)$ can be approximated as follows:

$$f(x) \approx \sum_{k=-1}^{3} c_k \phi_{2,k} + \sum_{i=2}^{M} \sum_{j=-1}^{2^i-2} d_{i,j}\psi_{i,j} = C^T \Psi, \tag{11}$$

where C and Ψ indicate $(2^{(M+1)} + 1) \times 1$ vectors given by

$$C = \left[c_{-1}, c_0, \cdots, c_3, d_{2,-1}, \cdots, d_{2,2}, d_{3,-1}, \cdots, d_{3,6}, \cdots, d_{M,-1}, \cdots, d_{M,2^M-2}\right]^T, \tag{12}$$

$$\Psi = \left[\phi_{2,-1}, \phi_{2,0}, \cdots, \phi_{2,3}, \psi_{2,-1}, \cdots, \psi_{2,2}, \psi_{3,-1}, \cdots, \psi_{3,6}, \cdots, \psi_{M,-1}, \cdots, \psi_{M,2^M-2}\right]^T, \tag{13}$$

with

$$c_k = \int_0^1 f(x) \overset{\sim}{\phi}_{2,k} (x)dx, \quad k = -1, 0, \cdots, 3, \tag{14}$$

$$d_{i,j} = \int_0^1 f(x) \overset{\sim}{\psi}_{i,j} (x)dx, \quad i = 2, 3, \cdots, M, \quad j = -1, 0, \cdots, 2^M - 2. \tag{15}$$

Also, $\overset{\sim}{\phi}_{2,k} (x)$ and $\overset{\sim}{\psi}_{i,j} (x)$ are dual functions of $\phi_{2,k}(x)$ and $\psi_{i,j}(x)$, respectively, which are defined in the following lemma.

Lemma 2.1 ([15]). *The dual functions $\overset{\sim}{\phi}_{2,k} (x)$ and $\overset{\sim}{\psi}_{i,j} (x)$ can be obtained by linear combinations of $\phi_{2,k},\ k = -1, \cdots, 3$ and $\psi_{i,j},\ i = 2, \cdots, M,\ j = -1, \cdots, 2^M - 2$, as follows*

$$\overset{\sim}{\Phi} (x) = P_1^{-1}\Phi(x), \quad \overset{\sim}{\bar{\Psi}} (x) = P_2^{-1}\bar{\Psi}(x),$$

where $\Phi(x)$ is a 5-vector and $\bar{\Psi}(x)$ denotes an $(2^{(M+1)}+1)\times 1$-vector and P_1 and P_2 are appropriate matrices given below.

Now, suppose that

$$\Phi = \left[\phi_{2,-1}(x), \phi_{2,0}(x), \phi_{2,1}(x), \phi_{2,2}(x), \phi_{2,3}(x)\right]^T, \tag{16}$$

$$\bar{\Psi} = \left[\psi_{2,-1}(x), \psi_{2,0}(x), \cdots, \psi_{M,2^M-2}(x)\right]^T, \tag{17}$$

then for $j = 2$ and using (4)–(9), and (16)–(17) we get

$$\int_0^1 \Phi\Phi^T dx = P_1 = \begin{bmatrix} \frac{1}{12} & \frac{1}{24} & 0 & 0 & 0 \\ \frac{1}{24} & \frac{1}{6} & \frac{1}{24} & 0 & 0 \\ 0 & \frac{1}{24} & \frac{1}{6} & \frac{1}{24} & 0 \\ 0 & 0 & \frac{1}{24} & \frac{1}{6} & \frac{1}{24} \\ 0 & 0 & 0 & \frac{1}{24} & \frac{1}{12} \end{bmatrix}, \tag{18}$$

and

$$\int_0^1 \bar{\Psi}\bar{\Psi}^T dx = P_2 = \begin{bmatrix} N_{4\times4} & & & \\ & \frac{1}{2}N_{8\times8} & & \\ & & \ddots & \\ & & & \frac{1}{2^{M-2}}N_{2^M\times2^M} \end{bmatrix}, \tag{19}$$

where P_1 is a 5×5 matrix and P_2 is a $(2^{M+1} - 4) \times (2^{M+1} - 4)$ matrix, and N denotes a 5-diagonal matrix given by

$$N = \begin{bmatrix} \frac{2}{27} & \frac{1}{96} & -\frac{1}{864} & 0 & 0 & \cdots & 0 \\ \frac{1}{96} & \frac{1}{16} & \frac{5}{432} & -\frac{1}{864} & 0 & \cdots & 0 \\ -\frac{1}{864} & \frac{5}{432} & \frac{1}{16} & \frac{5}{432} & -\frac{1}{864} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -\frac{1}{864} & \frac{5}{432} & \frac{1}{16} & \frac{5}{432} & -\frac{1}{864} \\ 0 & \cdots & 0 & -\frac{1}{864} & \frac{5}{432} & \frac{1}{16} & \frac{1}{96} \\ 0 & \cdots & 0 & 0 & -\frac{1}{864} & \frac{1}{96} & \frac{2}{27} \end{bmatrix}. \tag{20}$$

Next, let $\widetilde{\Phi}$ and $\widetilde{\bar{\Psi}}$ are the dual functions of Φ and $\bar{\Psi}$, respectively, given by

$$\widetilde{\Phi}=\left[\widetilde{\phi}_{2,-1}(x),\widetilde{\phi}_{2,0}(x),\widetilde{\phi}_{2,1}(x),\widetilde{\phi}_{2,2}(x),\widetilde{\phi}_{2,3}(x)\right]^T, \tag{21}$$

$$\widetilde{\bar{\Psi}}=\left[\widetilde{\psi}_{2,-1}(x),\widetilde{\psi}_{2,0}(x),\cdots,\widetilde{\psi}_{M,2^M-2}(x)\right]^T. \tag{22}$$

Using (14)-(17) and (21)–22 we have

$$\int_0^1 \widetilde{\Phi}\,\Phi^T dx = I_1, \qquad \int_0^1 \widetilde{\bar{\Psi}}\,\bar{\Psi}^T dx = I_2, \tag{23}$$

where I_1 and I_2 are 5×5 and $(2^{(M+1)}-4)\times(2^{(M+1)}-4)$ identity matrices, respectively. Then (18),(19) and (23) imply that

$$\widetilde{\Phi}(x) = P_1^{-1}\Phi(x), \qquad \widetilde{\bar{\Psi}}(x) = P_2^{-1}\bar{\Psi}(x). \tag{24}$$

2.3. Convergence analysis

Theorem 2.2. *Suppose that the representation (10) of B-spline wavelets is utilized for $f \in C^2[0,1]$, where Ψ has 2 universality moments. Then we have*

$$|d_{j,k}| \le \alpha\beta\eta^2 \tfrac{2^{-3j}}{2!},$$

where $\alpha = max|f''(t)|_{t\in[0,1]}$, $\beta = \displaystyle\int_{-k}^{2^j-k} \tilde{\psi}(x)dx$ and $\eta \in (-k, 2^j-k)$.

Theorem 2.3. *Let $e_j(x)$ be the error in approximating f. With the assumptions of Theorem 2.2, we have*

$$|e_M(x)| = O(2^{-2M}).$$

3. Methodology

3.1. Problem statement

As stated before, we utilize B-spline wavelets to numerically solve the fractional Volterra integral equation (1) with $\mu(x) = 1$, i.e.,

$$f(x) - \tfrac{1}{\Gamma(\alpha)}\int_0^x (x-t)^{\alpha-1}k(x,t)f(t)dt = g(x), \tag{25}$$

where f denotes an unknown function and $\alpha > 0$. To this end, we first use (11) to approximate the unknown function f and the known functions $g(x)$ and $k(x,t)$ as follows:

$$f(x) = C^T\Psi(x), \quad g(x) = G^T\Psi(x), \quad k(x,t) = \Psi^T(x)K\Psi(t), \tag{26}$$

where K is a $(2^{M+1}+1)\times(2^{M+1}+1)$ matrix as

$$K = \begin{bmatrix} SS & SW \\ WS & WW \end{bmatrix},$$

in which

$$\begin{aligned} SS &= [(ss)_{i,j}]_{5\times 5}, \\ SW &= [(sw)_{i,l}]_{5\times(2^{M+1}+1)}, \\ WS &= [(ws)_{l,i}]_{(2^{M+1}+1)\times 5}, \\ WW &= [(ww)_{l,n}]_{(2^{M+1}+1)\times(2^{M+1}+1)}, \end{aligned}$$

are matrices whose entries are as follows (see (24)):

$$(ss)_{i,j} = \int_0^1 \int_0^1 k(x,t)\, \widetilde{\phi}_{2,i}\,(x)\, \widetilde{\phi}_{2,j}\,(t)dxdt, \quad -1 \le i,j \le 3,$$
$$(sw)_{i,l} = \int_0^1 \int_0^1 k(x,t)\, \widetilde{\phi}_{2,i}\,(x)\, \widetilde{\psi}_l\,(t)dxdt, \quad -1 \le i \le 3, \quad -1 \le l \le 2^{M+1}-6,$$
$$(ws)_{l,i} = \int_0^1 \int_0^1 k(x,t)\, \widetilde{\psi}_l\,(x)\, \widetilde{\phi}_{2,i}\,(t)dxdt, \quad -1 \le i \le 3, \quad -1 \le l \le 2^{M+1}-6,$$
$$(ww)_{l,n} = \int_0^1 \int_0^1 k(x,t)\, \widetilde{\psi}_l\,(x)\, \widetilde{\psi}_n\,(t)dxdt, \quad -1 \le l,n \le 2^{M+1}-6.$$

3.2. Linear B-spline wavelets OMFI

Definition 3.1. Suppose that

$$I^\alpha \Phi(x) \simeq Q_\Phi^\alpha \Phi(x), \quad I^\alpha \Psi(x) \simeq Q_\Psi^\alpha \Psi(x), \tag{27}$$

then $(2^{M+1}+1) \times (2^{M+1}+1)$-matrices Q_Φ^α and Q_Ψ^α are called OMFIs for the linear B-spline scaling functions and wavelets, respectively.

Theorem 3.2. *[15, 27] The OMFI of B-spline scaling functions is shown as follows:*

$$Q_\Phi^\alpha = \begin{bmatrix} 0 & \zeta_1 & \zeta_2 & \zeta_3 & \cdots & \zeta_{2^{M+1}} \\ 0 & \eta_1 & \eta_2 & \eta_3 & \cdots & \eta_{2^{M+1}} \\ 0 & 0 & \eta_1 & \eta_2 & \cdots & \eta_{2^{M+1}-1} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \eta_1 & \eta_2 \\ 0 & \cdots & \cdots & 0 & 0 & \eta_1 \end{bmatrix},$$

where

$$\zeta_i = \frac{1}{2^{(M+1)\alpha}\Gamma(\alpha+2)} \left(i^\alpha (i+\alpha+1) + (i-1)^{\alpha-1} \right), \quad i = 1,2,\cdots,2^{M+1},$$
$$\eta_1 = \frac{1}{2^{(M+1)\alpha}\Gamma(\alpha+2)},$$
$$\eta_i = \frac{1}{2^{(M+1)\alpha}\Gamma(\alpha+2)} \left(i^{\alpha+1} - 2(i-1)^{\alpha+1} + (i-2)^{\alpha+1} \right), \quad i = 2,3,\cdots,2^{M+1}.$$

Moreover, the vector $\Psi(x)$ is expressed by scaling functions as

$$\Psi(x) = H\Phi(x),$$

where H is a $(2^{M+1}+1) \times (2^{M+1}+1)$ matrix as follows [17]:

$$H = \begin{bmatrix} \beta_2 \times \beta_3 \times \cdots \times \beta_M \\ L_2 \times \beta_3 \times \cdots \times \beta_M \\ \vdots \\ L_{M-2} \times \beta_{M-1} \times \beta_M \\ L_{M-1} \times \beta_M \\ L_M \end{bmatrix},$$

in which β_j and L_j, $j = 2, 3, \cdots, M$ are respectively $(2^j + 1) \times (2^{j+1} + 1)$ and $2^j \times (2^{j+1} + 1)$ matrices given by

$$\beta_j = \begin{bmatrix} 1 & \frac{1}{2} & & & & & & & & \\ & \frac{1}{2} & 1 & \frac{1}{2} & & & & & & \\ & & & \frac{1}{2} & 1 & \frac{1}{2} & & & & \\ & & & & \ddots & \ddots & \ddots & & & \\ & & & & & \frac{1}{2} & 1 & \frac{1}{2} & & \\ & & & & & & & \frac{1}{2} & 1 & \frac{1}{2} \\ & & & & & & & & & \frac{1}{2} & 1 \end{bmatrix},$$

$$L_j = \begin{bmatrix} -1 & \frac{11}{12} & \frac{-1}{2} & \frac{1}{12} & & & & & & & & & \\ & \frac{1}{12} & \frac{-1}{2} & \frac{5}{6} & \frac{-1}{2} & \frac{1}{12} & & & & & & & \\ & & & \frac{1}{12} & \frac{-1}{2} & \frac{5}{6} & \frac{-1}{2} & \frac{1}{12} & & & & & \\ & & & & \ddots & \ddots & \ddots & \ddots & \ddots & & & & \\ & & & & & \frac{1}{12} & \frac{-1}{2} & \frac{5}{6} & \frac{-1}{2} & \frac{1}{12} & & & \\ & & & & & & & \frac{1}{12} & \frac{-1}{2} & \frac{5}{6} & \frac{-1}{2} & \frac{1}{12} & \\ & & & & & & & & & \frac{1}{12} & \frac{-1}{2} & \frac{11}{12} & -1 \end{bmatrix}.$$

Theorem 3.3. *The OMFI of B-spline wavelets is shown in the following [15]:*

$$Q_{\Psi}^{\alpha} = H Q_{\Phi}^{\alpha} H^{-1}. \tag{28}$$

3.3. The solution method

Using (27) and (28), the following relation is obtained:

$$I^{\alpha}\Psi(x) = \frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1}\Psi(x) \simeq Q_{\Psi}^{\alpha}\Psi(x) = H Q_{\Phi}^{\alpha} H^{-1}\Psi(x). \tag{29}$$

Furthermore, if we write

$$\Psi(t)\Psi^T(t)C = U^T\Psi(t) \tag{30}$$

then we would have

$$U^T = \int_0^1 \Psi(t)\Psi^T(t)C \stackrel{\sim}{\Psi}^T (t)dt \tag{31}$$

In addition, suppose that we write

$$\frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} k(x,t) f(t) dt = V^T\Psi(x), \tag{32}$$

where

$$V^T = \left[v_{-1}, v_0, \cdots, v_3, w_{2,-1}, \cdots, w_{2,2}, w_{3,-1}, \cdots, w_{3,6}, \cdots, w_{M,-1}, \cdots, w_{M,2^M-2}\right].$$

Then we have

$$v_n = \int_0^1 \left(\frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} k(x,t) f(t) dt \right) \stackrel{\sim}{\phi}_{2,n} (x)dx, \quad n = -1, 0, 1, 2, 3, \tag{33}$$

and

$$\begin{aligned} w_{i,j} &= \int_0^1 \left(\frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} k(x,t) f(t) dt \right) \stackrel{\sim}{\psi}_{i,j} (x)dx, \\ & i = 2, 3, \cdots, M, \quad j = -1, 0, \cdots, 2^M - 2. \end{aligned} \tag{34}$$

Now, using (26), (27) and by (31), can rewrite (33) as follows

$$v_n = \int_0^1 \Psi^T(x) K U^T \left(\frac{1}{\Gamma(\alpha)} \int_0^x (x-t)^{\alpha-1} \Psi(t) dt \right) \widetilde{\phi}_{2,n}(x) dx.$$

Further, by (29), we get

$$v_n = \int_0^1 \Psi^T(x) K U^T Q_\Psi^\alpha \Psi(x) \widetilde{\phi}_{2,n}(x) dx.$$

This implies that

$$v_n = K U^T Q_\Psi^\alpha \int_0^1 \Psi^T(x) \Psi(x) \widetilde{\phi}_{2,n}(x) dx,$$

and

$$w_{i,j} = K U^T Q_\Psi^\alpha \int_0^1 \Psi^T(x) \Psi(x) \widetilde{\psi}_{i,j}(x) dx.$$

In conclusion, we deduce that

$$V = K U^T Q_\Psi^\alpha \int_0^1 \Psi^T(x) \Psi(x) \widetilde{\Psi}(x) dx.$$

Finally, equation (25) is reduced to a linear system of algebraic equations as follows:

$$C^T \Psi - V^T \Psi = G^T \Psi \Longrightarrow C - V = G.$$

Solving this linear algebraic system using an appropriate solver, the vector C of unknowns is obtained. Substituting it into (26) provides an approximation for $f(x)$.

4. Numerical Examples

We examined the proposed method by solving some fractional Volterra integral equations to demonstrate the applicability and accuracy of the method. Let the error be measured by

$$\| e_M(x) \|_2 = \left(\int_0^1 e_M^2(x) dx \right)^{\frac{1}{2}} = \left(\int_0^1 (f(x) - f_M(x))^2 dx \right)^{\frac{1}{2}},$$

where $f(x)$ and $f_M(x)$ are the exact solution and approximate solutions, respectively.

Example 4.1. Consider the following fractional Volterra integral equation:

$$f(x) - \frac{1}{\Gamma(\frac{3}{2})} \int_0^x (x-t)^{\frac{1}{2}} \frac{1}{t} f(t) dt = x^3 - 0.172 x^{\frac{7}{2}}.$$

The exact solution to this problem is $f(x) = x^3$. The L^2-norm of error function of B-spline wavelets approximation of the problem (4.1) for some values of M is shown in Table 1. Moreover, a comparison between the exact and approximate solutions is given in Fig. 1.

Table 1. $\| e_M \|_2$ for some values of M in Example 4.1.

$M=2$	$M=3$	$M=4$	$M=5$	$M=6$	$M=7$
2.39×10^{-2}	1.08×10^{-2}	7.65×10^{-3}	3.75×10^{-3}	1.87×10^{-3}	8.13×10^{-4}

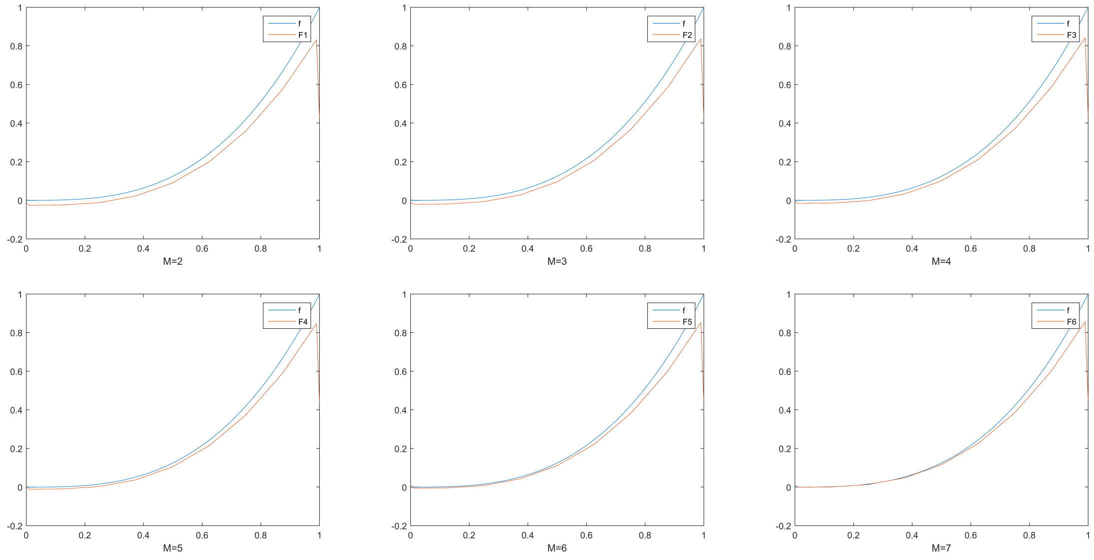


Fig. 1. Comparison of exact solution with approximate solution for problem 4.1.

5. Conclusions

In this research, we successfully obtained a solution for fractional Volterra integral equations using an effective method based on B-splines. By deriving the operational matrix of fractional integration (OMFI), we were able to transform the original equations into a system of linear algebraic equations. Through numerical example, we demonstrated the applicability and effectiveness of the proposed technique. The error analysis of the suggested function approximation revealed its efficiency in solving problems within the realm of applied science and engineering that are modeled as fractional Volterra integral equations (FVIEs). Looking ahead, we are going to extend the proposed method to tackle two-dimensional fractional Fredholm-Volterra integro-differential equations, as well as systems of nonlinear Volterra integro-differential equations. Another direction for further research is to investigate how a fractional integral equation can be converted to a calculus of variation problem and how the proposed B-splines technique is employed to solve it directly.

References

[1] M. Rahman, Integral Equations and their Applications, Wit Press, Southampton, Boston, 2007.

[2] S. Abbasbandy, S. Kazem, M. S. Alhuthali and H. H. Alsulami, Application of the operational matrix of fractional-order Legendre functions for solving the timefractional convection-diffusion equation, Appl. Math. Comput. 266 (2015) 31–40.

[3] G. Ala, M. L. Silvestre, E. Francomano and A. Tortorici, An advanced numerical model in solving thin-wire integral equations by using semi-orthogonal compactly supported spline wavelets, IEEE Trans. Electromagn. Compact. 45(2) (2003) 218–228.

[4] K. Al-khaled, Numerical solution of time-fractional partial differential equations using Sumudu decomposition method, Rom. J. Phys. 60(1–2) (2015) 99–110.

[5] R. Bagley, On the equivalence of the Riemann–Liouville and the Capotu fractional order derivatives in modeling of linear viscoelastic materials, Frac. Calculus Appl. Anal. 10(2) (2007) 123–126.

[6] A. H. Bhrawy, E. H. Doha, M. A. Abdelkawy and R. A. Van Gorder, Jacobi–Gauss– Lobatto collocation method for solving nonlinear reaction-diffusion equations subject to Dirichlet boundary conditions, Appl. Math. Model. in Press (2015).

[7] V. Daftardar-Gejji and S. Bhalekar, An iterative method for solving fractional differential equations, PAMM. Proc. Appl. Math. Mech. 7 (2007) 2050017–2050018.

[8] V. Daftardar-Gejji and H. Jafari, Adomian decomposition: A tool for solving a system of fractional differential equations, J. Math. Anal. Appl. 301 (2005) 508–518.

[9] M. Dalir and M. Bashour, Applications of fractional calculus, Appl. Math. Sci. 4(21) (2010) 1021–1032.

[10] Z. E. A. Fellah, C. Depollier and M. Fellah, Application of fractional calculus to the sound waves propagation in rigid porous materials: Validation via ultrasonic measurement, Acta Acoustica United With Acoustica 88 (2002) 34–39.

[11] J. C. Goswami, A. K. Chan, andC. K. Chui, On solving first-kind integral equations using wavelets on a bounded interval, IEEE Trans. Antennas Propagat. 43 (1995), no. 6, 614–622.

[12] A. A. Hemeda, Solution of fractional partial differential equations in fluid mechanics by extension of some iterative method, Abstract Appl. Anal. 2013 (2013) 1–9.
[13] T. Ismaeelpour, A. Askari Hemmat and H. Saeedi, B-spline operational matrix of fractional integration Optik-Int. J. Light Elect. Opt. 130 (2017) 291–305.
[14] Y. Jiang and J. Ma, High-order finite element methods for time-fractional partial differential equations, J. Comput. Appl. Math. 235 (2011) 3285–3290.
[15] Z. Kargar, H. Saeedi, B-spline wavelet operational for numerical solution of time-space fractional partial differential equations, International Journal of Wavelets, Multiresolution and Information Processing, Vol. 15, No. 4(2017)
[16] V. V. Kulish and J. L. Lage, Application of fractional calculus to fluid mechanics, J. Fluids Eng. 124 (2002) 803–806.
[17] M. Lakestani, M. Razzaghi and M. Dehghan, Semiorthogonal spline wavelets approximation for fredholm integro-differential equations, Math. Prob. Eng. 2006 (2006) 1–12.
[18] X. Li, Numerical solution of fractional differential equations using cubic B-spline wavelet collocation method, Commun. Nonlinear Sci. Numer. Simulat. 17 (2012) 3934–3946.
[19] F. Mainardi, Fractional calculus: Some basic problems in continuum and statistical mechanics, in Fractals and Fractional Calculus in Continuum Mechanics, eds. A. Carpinteri and F. Mainardi (Springer-Verlag, Wien and NewYork, 1997), pp. 291–348.
[20] F. Metzler, W. Schick, H. G. Kilian and T. F. Nonnenmacher, Relaxation in filled polymers: A fractional calculus approach, J. Chem. Phys. 103 (1995) 7180–7186.
[21] K. Moaddy, S. Momani and I. Hashim, The non-standard finite difference scheme for linear fractional PDEs in fluid mechanics, Comput. Math. Appl. 61 (2011) 1209–1216.
[22] H. M. Osama, S. F. Fadhel and M. G. S. AL-Safi, Numerical solution for the timefractional diffusion-wave equations by using Sinc-Legendre collocation method, Math. Theory Model. 5(1) (2015) 49–57.
[23] M. Bisheh Niasar and A. Mahdipour, An adaptive mesh hybrid block method for solving nonlinear singularly perturbed differential equations. Soft Computing Journal, 11(1) (2022). DOI: 10.22052/SCJ.2023.248415.1105
[24] H. Saeedi, Application of Haar wavelets in solving nonlinear fractional Fredholm integro-differential equations, J. Mahani Math. Res. Cent. 2(1) (2013) 15–28.
[25] H. Saeedi and G. N. Chuev, Triangular functions for operational matrix of nonlinear fractional Volterra integral equations, J. Appl. Math. Comput. 49(1) (2015) 213–232.
[26] T. Ismaeelpour, A. Askari Hemmat and H. Saeedi, B-spline operational matrix of fractional integration Optik-Int. J. Light Elect. Opt. 130 (2017) 291–305.
[27] H. Saeedi, On the linear B-spline scaling function operational matrix of fractional integration and its applications in solving fractional order differential equations, Iranian J. Sci. Technol. (2015).
[28] H. R. Tabrizidooz and F. Hajiramezanali, A numerical algorithm for determining time-dependent coefficient in a parabolic inverse problem using Legendre multiwavelet base. Soft Computing Journal, 10(2) (2022) 110-123. DOI: 10.22052/SCJ.2022.243311.1028
[29] A. M. Shukur, Adomian decomposition method for certain space-time fractional partial differential equations, IOSR J. Math. (IOSR-JM) 11 (2015) 55–65.
[30] E. Soczkiewicz, Application of fractional calculus in the theory of viscoelasticity, Mol. Quant. Acoustics 23 (2002) 397–404.
[31] N. Su, Mass-time and space-time fractional partial differential equations of water movement in soils: Theoretical framework and application to infiltration, J. Hydro. 519 (2014) 1792–1803.
[32] M. Yi, J. Huang and J. Wei, Block pulse operational matrix method for solving fractional partial differential equation, Appl. Math. Comput. 221 (2013) 121–131.
[33] F. Usta, Numerical analysis of fractional Volterra integral equations via Bernstein approximation method, J. Comput. Appl. Math. 384 (2021) 113198.
[34] P. Rahimkhani, Y. Ordokhani, S. Sabermahani, Bernoulli wavelet least squares support vector regression: Robust numerical method for systems of fractional differential equations, Math. Methods. Appl. Sci. 46(02) (2023) DOI:10.1002/mma.9522.
[35] P. Rahimkhani, Y. Ordokhani, Performance of Genocchi wavelet neural networks and least squares support vector regression for solving different kinds of differential equations, Comput. Appl. Math. 42(2) (2023) DOI:10.1007/s40314-023-02220-1.
[36] P. Rahimkhani, Y. Ordokhani, Orthonormal Bernoulli wavelets neural network method and its application in astrophysics, Comput. Appl. Math. 40 (78) (2021), https://doi.org/10.1007/s40314-021-01475-w.
[37] P. Rahimkhani, Y. Ordokhani, P.M. Lima, An improved composite collocation method for distributed-order fractional differential equations based on fractional Chelyshkov wavelets, Appl. Numer. Math. 145(6) (2019), DOI:10.1016/j.apnum.2019.05.023.
[38] S. Afzali, M. K. Moayyedi, and F. Fotouhi, Reduced Order Framework for 2D Heat Transfer Simulation Under Variations of Boundary Conditions Based on Deep Learning Algorithm. Soft Computing Journal. (2022). DOI: 10.22052/SCJ.2022.246176.1054.
[39] Podlubny, Igor. Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. Elsevier, 1998.
[40] Miller, Kenneth S., and Bertram Ross. An introduction to the fractional calculus and fractional differential equations. Wiley, 1993.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–167, pp. 183–189

An Operational Haar Wavelet Method for Solving Two Dimensional Fractional Volterra Nonlinear Integral Equations

Zohreh Abdollahi[a,*], M. Mohseni Moghadam[a], H. Saeedi[b], M. J. Ebadi[c]

[a] *Department of Mathematics, Kerman Branch Islamic Azad University, Kerman, Iran*
[b] *Department of Applied Mathematics, Faculty of Mathematics and Computer, Shahid Bahonar University of Kerman, Kerman, Iran*
[c] *Department of Mathematics, Chabahar Maritime University, Chabahar, Iran*

Article Info

Keywords:
Two Dimensional Fractional Volterra Integral Equations
Two Dimensional Haar Wavelets

Abstract

In this paper, we will first introduce the Haar wavelets operational matrix for fractional integration in two dimensional spaces. Haar wavelet approximation is then utilized to reduce a two dimension fractional Volterra nonlinear integral equation, to a system of algebraic equations.

1. Introduction

The conception of the fractional derivatives was introduced for the first time in the middle of the 19th century by Riemann and Liouville. Later on the number of researches and studies about the fractional calculus has rapidly increased, because some physical processes such as anomalous diffusion[2], complex viscoelasticity [10], behavior of biological systems [9], rheology[11] and etc. can't be described by classical models of fractional derivatives. Fractional differential equations have been discussed in many papers and in most of them, they are transformed into fractional Volterra integral equations. Fractional differential equations are solved by different types of wavelets such as Haar wavelets. Haar wavelets are used due to their useful properties such as, orthogonality, compact support and simple applicability. Compact support of the Haar wavelets basis permits straight inclusion of the different types of boundary conditions in the numerical algorithms. Another good feature of these wavelets is the possibility to integrate them analytically. Haar wavelets are used in system analysis by Chen and Hsiao [3], they derived Haar operational matrix of the Haar function vector. In [5] a Haar product matrix and a coefficient matrix proposed for the integrals in linear time delayed systems. Haar wavelets operational matrix of fractional order integration has been applied to solve fractional differential equations in [21]. Sufficient conditions for the existence and uniqueness of solutions for a class of fractional partial differential equations using Haar wavelets operational matrix of fractional order integration are obtained in[8]. In [16] Haar wavelets operational matrix is used to find the solution of fractional Bagley-Torvik equation. A new Haar wavelets method based on operational matrices of fractional order integration are used to solve several types of

*Talker
Email addresses: abdolahi_zohre@yahoo.com (Zohreh Abdollahi), mohseni@uk.ac.ir (M. Mohseni Moghadam), saeedi@uk.ac.ir (H. Saeedi), ebadi@cmu.ac.ir (M. J. Ebadi)

fractional order differential equations numerically in [7]. A Haar wavelets operational matrix is applied for solving fractional Volterra integral equations in [18]. This paper is organized as follows: In section 2 we state the main concepts of Riemann-Liouville fractional integral operator and Haar wavelets in one and two dimensional spaces. Also two dimensional fractional Volterra integral equations, which are going to be solved in this paper, are introduced. In section 3 we will introduce the function approximation via two dimension Haar wavelets, the operational matrix of integration, the product operational matrix and the generalized operational matrix for fractional integration. Next we will utilized these matrices for solving the two dimensional fractional Volterra integral equations.

2. Preliminaries

In this section, we will present the main definitions and properties of fractional integration operator and Haar wavelets in one and two dimensions.

2.1. Riemann-Liouville fractional integration

The Riemann-Liouville fractional integral operator of order $\alpha \geq 0$ of the function $f(x)$ is defined as [12, 14]:

$$I^{\alpha} f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)} \int_0^x (x-s)^{\alpha-1} f(s)ds, & \alpha > 0, \\ f(x), & \alpha = 0. \end{cases} \tag{1}$$

Here, Γ is the so called Gamma function. Similarly, the Riemann-Liouville fractional integral operator of order $\alpha, \beta > 0$ of the function $f(x, y)$ with respect to the variables x and y is given by:

$$I^{\alpha,\beta} f(x,y) = \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \int_0^y (x-s)^{\alpha-1}(y-t)^{\beta-1} f(s,t)dsdt.$$

2.2. Haar Wavelets

The orthogonal set of the Haar wavelets $h_n(x), n = 0, 1, 2, ...$, is a group of square waves defined as follows:

$$h_0(x) = \begin{cases} 1, & 0 \leq \text{x} < 1; \\ 0, & \text{elsewhere.} \end{cases}$$

$$h_1(x) = \begin{cases} 1, & 0 \leq \text{x} < \frac{1}{2}; \\ -1, & \frac{1}{2} \leq \text{x} < 1; \\ 0, & \text{elsewhere.} \end{cases} \tag{2}$$

and for $n = 2, 3, ...$, we choose the numbers $j, k \in \mathbb{N} \cup \{0\}$ and $0 \leq k < 2^j$, such that $n = 2^j + k$ and:

$$h_n(x) = h_1(2^j x - k). \tag{3}$$

On the other hand, we can also define Haar wavelets by using the Heaviside step function as follows [17]:

$$u(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

$$h_0(x) = u(x) - u(x-1),$$

$$h_n(x) = u(x - \tfrac{k}{2^j}) - 2u(x - \tfrac{k+1/2}{2^j}) + u(x - \tfrac{k+1}{2^j}), \tag{4}$$

$$n = 2^j + k,\ j, k \in \mathbb{N} \cup \{0\};\ 0 \leq k < 2^j.$$

In fact

$$h_n(x) = \begin{cases} 1, & \frac{k}{2^j} \leqslant x < \frac{k+0.5}{2^j}, \\ -1, & \frac{k+0.5}{2^j} \leqslant x < \frac{k+1}{2^j}, \\ 0, & o.w. \end{cases} \tag{5}$$

The Haar wavelets satisfy the following orthogonality property:

$$\int_0^1 h_n(x)h_m(x)dx = 2^{-j}\delta_{nm} = 2^{-i}\delta_{nm}, \tag{6}$$

where δ_{nm} is the Kronecker delta and

$$n = 2^j + k,\ j,k \in \mathbb{N} \cup \{0\};\ 0 \leq k < 2^j.$$
$$m = 2^i + l,\ i,l \in \mathbb{N} \cup \{0\};\ 0 \leq l < 2^i.$$

For more details see [2, 6, 19].
Now we define the two dimensional Haar wavelets as follows [6]:

$$h_{n,m}(x,y) := h_n(x)h_m(y) \tag{7}$$

The $\{h_{n,m}(x,y)\}_{n,m=0}^{\infty}$ is an orthogonal basis for the space $L^2[0,1) \times [0,1)$.

2.3. Two dimensional Fractional Volterra nonlinear integral equation

A two dimensional fractional Volterra nonlinear integral equation is given by:

$$f(x,y) - \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \int_0^y (x-s)^{\alpha-1}(y-t)^{\beta-1}k(x,y,s,t)f^q(s,t)dsdt = g(x,y). \tag{8}$$

where f is the unknown function, the kernel $k(x,y,s,t)$, the right-hand side function $g(x,y)$ and $\alpha, \beta > 0$ are given and $q \geq 1$ is a positive integer.

2.4. Existence and uniqueness of the solution

In this subsection according to [15], we introduce a sufficient condition so that Eq.(8) has a unique solution. We denote the exact solution by $f(x,y)$ and approximate solution by $f_{m,n}(x,y)$. We assume $|(x-s)^{\alpha-1}(y-t)^{\beta-1}k(x,y,s,t)| \leq N_1$ and $\mid f^q(x,y) - f^q_{m,n}(x,y) \mid \leq L \mid f(x,y) - f_{m,n}(x,y) \mid$, where L is a constant.

Theorem 2.1. *The solution of Eq.(8) exists and is unique, if* $\frac{N_1 L}{\Gamma(\alpha)\Gamma(\beta)} < 1$.

3. Methodology

3.1. Function approximation

A square integrable function $f(x,y)$ on $[0,1) \times [0,1)$ can be expanded into two dimensional Haar series:

$$f(x,y) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} f_{n,m}h_{n,m}(x,y),\quad x,y \in [0,\ 1) \times [0,\ 1). \tag{9}$$

For computing the Haar coefficients, $f_{n,m}$, of $f(x,y)$,in Eq.(9), we multiply both sides of Eq.(9) by $h_{n,m}(x,y)$. Therefore, by using Eq.(6) and (7)we have:

$$f_{n,m} = 2^{(i+j)} \int_0^1 \int_0^1 f(x,y)h_{n,m}(x,y)dxdy \tag{10}$$

The series (9) is not useful for approximation, so by truncating and rewriting it, we have:

$$f(x,y) \cong \sum_{i=0}^{I} \sum_{l=0}^{2^i-1} \sum_{j=0}^{J} \sum_{k=0}^{2^j-1} f_{2^j+k,2^i+l}h_{2^j+k,2^i+l}(x,y) = h^T(x,y)F = F^T h(x,y), \tag{11}$$

where F and $h(x,y)$ are the vectors given by:
$F^T = \begin{bmatrix} f_{0,0} & \dots & f_{r-1,r-1} \end{bmatrix}, h^T(x,y) = \begin{bmatrix} h_{0,0}(x,y) & \dots & h_{r-1,r-1}(x,y) \end{bmatrix}$, the superscript T denotes the transpose and $r = 2^{I+1} = 2^{J+1}$.
For the positive integer powers of a function we have:

$$f^q(x,y) = h^T(x,y)D = \sum_{i=0}^{I} \sum_{l=0}^{2^i-1} \sum_{j=0}^{J} \sum_{k=0}^{2^j-1} d_{2^j+k,2^i+l} h_{2^j+k,2^i+l}(x,y), \tag{12}$$

where entries of matrix D are:

$$D_{m,n} = 2^{(i+j)} \int_0^1 \int_0^1 (h^T(x,y)F)^q h_{m,n}(x,y) dx dy. \tag{13}$$

Now, considering Eq.(8), let

$$g(x,y) \cong G^T h(x,y) \tag{14}$$

Similarly, the corresponding integral kernel $k(x,y,s,t)$ on the region $[0,1)\times[0,1)\times[0,1)\times[0,1)$ can be approximated as:

$$k(x,y,s,t) \cong h^T(x,y) K h(s,t), \tag{15}$$

where:

$$K = \begin{bmatrix} k_{0,0,0,0} & \dots & k_{0,0,r-1,r-1} \\ \vdots & \ddots & \vdots \\ k_{r-1,r-1,0,0} & \dots & k_{r-1,r-1,r-1,r-1} \end{bmatrix},$$

and the entries of the matrix K are:

$$k_{m,n,p,q} = 2^{(i+j)} 2^{(a+b)} \int_0^1 \int_0^1 \left(\int_0^1 \int_0^1 k(x,y,s,t) h_{m,n}(x,y) dx dy \right) h_{p,q}(s,t) ds dt, \tag{16}$$

with

$$\begin{aligned} m &= 2^i + l,\ i,l \in \mathbb{N} \cup \{0\};\ 0 \le l < 2^i, \\ n &= 2^j + k,\ j,k \in \mathbb{N} \cup \{0\};\ 0 \le k < 2^j, \\ p &= 2^a + c,\ a,l' \in \mathbb{N} \cup \{0\};\ 0 \le c < 2^a, \\ q &= 2^b + d,\ b,d \in \mathbb{N} \cup \{0\};\ 0 \le d < 2^b. \end{aligned}$$

By substituting Eq.(11), Eq.(12), Eq.(14) and Eq.(15) and in Eq.(8), we have:

$$h^T(x,y)F - \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \int_0^y (x-t)^{\alpha-1}(y-s)^{\beta-1} h^T(x,y) K h(s,t) h^T(s,t) D ds dt \cong h^T(x,y)G. \tag{17}$$

If the integral term of Eq.(17) can be written as a linear combination of Haar wavelets, as:

$$\tfrac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \int_0^y (x-t)^{\alpha-1}(y-s)^{\beta-1} h^T(x,y) K h(s,t) h^T(s,t) D ds dt \cong h^T(x,y) V, \tag{18}$$

then, (17) can be converted into:

$$h^T(x,y)F - h^T(x,y)V \cong h^T(x,y)G, \tag{19}$$

and therefore:

$$F - V \cong G. \tag{20}$$

Eq.(20) represents a nonlinear system of algebraic equations for the unknown F. The details of computing the vector components V are given in the next subsection.

3.2. Determining the vector V

First, the following theorem is presented:

Theorem 3.1. *If $D = \begin{bmatrix} d_{0,0} & \dots & d_{r-1,r-1} \end{bmatrix}^T$ is a vector of dimension r^2, then*

$$h(x,y)h^T(x,y)D = Nh(x,y),$$

where

$$\begin{aligned} h(x,y) &= h(x) \otimes h(y) \\ F_{r^2} &= C_r \otimes C_r, \end{aligned} \tag{21}$$

where $\otimes$ is known as Kronecker product.
Also $C_r = [f_0, \cdots f_{r-1}]^T$. By [17], we have

$$h(x)h^T(x)C_r = M_{r\times r}h(x), \tag{22}$$

and $N = (M \otimes M)^q$. Now, consider Eq.(18). If $V^T = \begin{bmatrix} v_{0,0}, & \dots, & v_{m,n}, & \dots, & v_{r-1,r-1} \end{bmatrix}$, then:

$$\begin{aligned} v_{m,n} &= 2^{i+j}\int_0^1\int_0^1 \frac{1}{\Gamma(\alpha)\Gamma(\beta)}\left(\int_0^x\int_0^y (x-s)^{\alpha-1}(y-t)^{\beta-1}h^T(x,y)KNh(s,t)dsdt\right)h_{m,n}(x,y)dxdy \\ &= 2^{i+j}\int_0^1\int_0^1 h^T(x,y)KN\frac{1}{\Gamma(\alpha)\Gamma(\beta)}\left(\int_0^x\int_0^y (x-t)^{\alpha-1}(y-s)^{\beta-1}h(s,t)dsdt\right)h_{m,n}(x,y)dxdy. \end{aligned} \tag{23}$$

For simplifying the integral:

$$\tfrac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^x\int_0^y (x-t)^{\alpha-1}(y-s)^{\beta-1}h(s,t)dsdt, \tag{24}$$

we recall the following Lemmas.

Lemma 3.2. *The expansion of the fractional integral of order α of $h(x)$ into a Haar series including P_m^α as a Haar coefficient matrix is given as below:*

$$\tfrac{1}{\Gamma(\alpha)}\int_0^x (x-t)^{\alpha-1}h(t)dt \cong P_m^\alpha h(x). \tag{25}$$

Proof. See [17]. □

Lemma 3.3. *If P^α and P^β are the operational matrices of fractional integration of orders α and β, respectively, in Lemma 3.2, then*

$$\tfrac{1}{\Gamma(\alpha)\Gamma(\beta)}\int_0^x\int_0^y (x-t)^{\alpha-1}(y-s)^{\beta-1}h(s,t)dsdt \cong (P^\alpha \otimes P^\beta)h(x,y),$$

Now, Eq.(23) can be rewritten as follows using Lemma 3.3:

$$\begin{aligned} v_{m,n} &= 2^{i+j}\int_0^1\int_0^1 h^T(x,y)KNP^\alpha \otimes P^\beta h(x,y)h_{m,n}(x,y)dxdy \\ &= 2^{i+j}\int_0^1\int_0^1 h^T(x,y)Ah(x,y)h_{m,n}(x,y)dxdy, \end{aligned}$$

where $A = KNP^\alpha \otimes P^\beta$. Also we can write

$$v_{m,n} = 2^{i+j}\int_0^1\int_0^1 Bh_{m,n}(x,y)dxdy, \tag{26}$$

where $B = h^T(x,y)Ah(x,y)$. Hence, we can compute $v_{m,n}$ with Eq.(26) and we have the nonlinear system $F - V = G$ for F.

4. Error Approximation

Theorem 4.1. *Suppose that f maps a convex open set $D \subseteq R^2$ into R, f is differentiable in D, and there is a real number L such that*

$$\|f^{'}(t)\| \leq L$$

for every $t \in D$. Then

$$|f(b) - f(a)| \leq L|b - a|$$

for all $a, b \in D$.

Proof. See [15]. □

Let $f(x,y) \in L^2[0,1)^2$ and $F^T h(x,y) = \tilde{f}_{m,n}(x,y)$ are respectively the exact and approximate solutions of Eq. (8). We define $e_{m,n}(x,y) = f(x,y) - \tilde{f}_{m,n}(x,y)$ and call $e_{m,n}$ the error function of 2D HW approximation. Where $f_{m,n}$ is given by Eq.(10), the error $e_{m,n}$ is minimum in the 2-norm sense. By Eq.(11) we have

$$e_{m,n}(x,y) = \sum_{i=I+1}^{\infty} \sum_{l=0}^{2^i-1} \sum_{j=J+1}^{\infty} \sum_{k=0}^{2^j-1} f_{2^j+k,2^i+l} h_{2^j+k,2^i+l}(x,y).$$

Theorem 4.2. *If the function $f(x,y)$ satisfies the following Lipschitz condition on $[0,1) \times [0,1)$:*

$$\exists L > 0; \forall x, y, s, t \in [0,1) \times [0,1): \quad | f(x,y) - f(s,t) | \leq L \mid (x,y) - (s,t) \mid . \tag{27}$$

Then, the HWM will be convergent with the meaning that the error function $e_{m,n}(x,y)$ tends to zero in L^2-norm when m, n go to infinity. In addition, the order of convergence is two, that is

$$\| e_{m,n}(x,y) \|_2 = O(\frac{1}{n}\frac{1}{m}).$$

5. Numerical Examples

In this section, we examine the our proposed method for solving some fractional two-dimensional Volterra integral equations. Notice that

$$\| e_{m,n}(x,y) \|_2 = \left(\int_0^1 \int_0^1 (e_{m,n}(x,y))^2 dxdy \right)^{\frac{1}{2}} = \left(\int_0^1 \int_0^1 (f(x,y) - \tilde{f}_{m,n}(x,y))^2 dxdy \right)^{\frac{1}{2}},$$

where $f(x,y)$ is the exact solution and $\tilde{f}_{m,n}(x,y)$ is the approximate solution obtained by (12).

Example 5.1. Consider the following two dimensional nonlinear fractional Volterra integral equation:

$$\begin{gathered} f(x,y) - \tfrac{1}{\Gamma(2)\Gamma(2)} \int_0^x \int_0^y (x-t)(y-s)xystf^2(s,t)dsdt \\ = xy(x-1)(y-1) - (x^6y^6(10x^2 - 28x + 21)(10y^2 - 28y + 21))/176400, \end{gathered}$$

where $f(x,y) = xy(1-x)(1-y)$ is the exact solution. Table 1 shows the L^2- norm of the two dimensional Haar approximation for some m and n. See Table 1 for example.

6. Conclusions

In this paper, the generalized 2D-FVNIEs were solved by using an OM method Based on 2D-HWs. The 2D-OM of fractional nonlinear integration was obtained to achieve a matrix system of algebraic linear equations. The simplicity and effectiveness of our method were shown in various problems. The error analyses of the proposed method show that this technique is effective to solve the applied science and engineering problems modelled as FVNIEs. For future research, it is suggested to extend our proposed method for solving the fractional Fredholm-Volterra nonlinear integro-differential equations.

Table 1. $\| e_{m,n} \|_2$ for some m,n in Example.

$m=n=2$	$m=n=4$	$m=n=8$	$m=n=16$	$m=n=32$	$m=n=64$
1.84×10^{-2}	1.05×10^{-2}	5.3×10^{-3}	2.7×10^{-3}	1.3×10^{-3}	6.72×10^{-4}

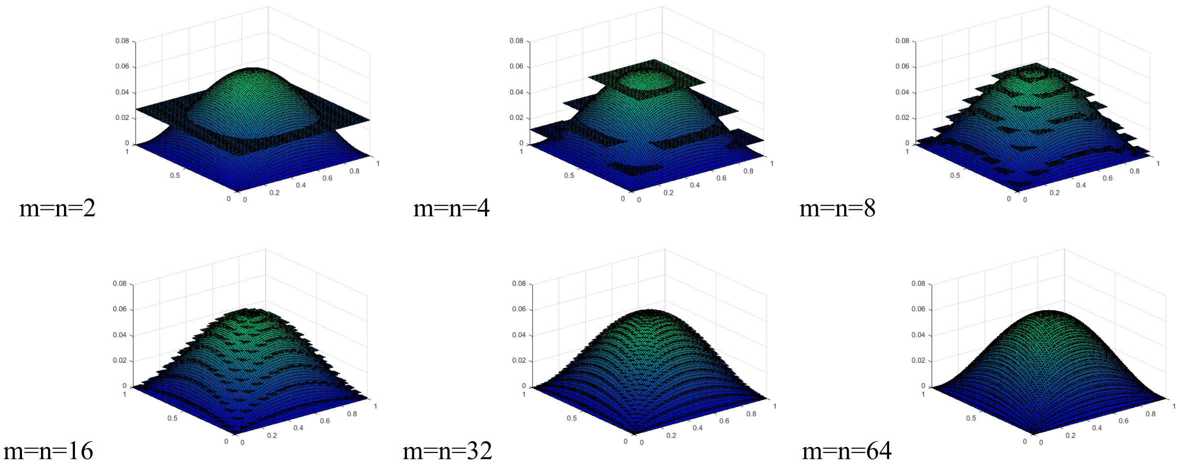


Fig. 1. Exact solution vs. approximate solution for Example 5.1

References

[1] A.A. Kilbas, H.M. Srivastava, J.J. Trujillo, Theory and Applications of Fractional Differential Equations, Elsevier, San Diego,2006.

[2] Akansu A.N., Haddad R.A. (1981). Multiresolution Signal Decomposition, Academic Press Inc., San Diego, CA.

[3] Chen C.F., Hsiao C.H. (1997). Haar wavelet method for solving lumped and distributed parameter systems, IEE Proceedings: Control Theory and Applications 144(1): 87-94.

[4] Chena W., Suna H., Zhang X., Koraosak D. (2010). Anomalous diffusion modeling by fractal and fractional derivatives, Computers & Mathematics with Applications 59(5): 265-274.

[5] C.H. Hsiao, States analysis of linear time delayed systems via Haar wavelets, Math. Comput. Simul. 44 (1997) 457–470.

[6] Derili H.A, Sohrabi S., Arzhang A. ,Two-dimensional wavelets for numerical solution of integral equations

[7] Firdous A. Shah, Abass R., Haar Wavelet Operational Matrix Method for the Numerical Solution of Fractional Order Differential Equations, Nonlinear. Engin, 4(4) (2015) 203–213.

[8] Wang L., Ma Y., Meng Z.(2014). Haar wavelet method for solving fractional partial differential equations numerically, Appl Math. Comput., 227: 66-76.

[9] Magin, R.L. . Fractional calculus in bioengineering. Part 2, Critical Reviews in Bioengineering 32:(2004) 105–193.

[10] Meral F.C., Royston T.J., Magin R., Fractional calculus in viscoelasticity: An experimental study, Commun. Nonlinear Sci. Numer. Simulat. 15 (2010) 939 945.

[11] Metzler R., Nonnenmacher T.F.(2003). Fractional relaxation processes and fractional rheological models for the description of a class of viscoelastic materials, Int. J. Plast. 19: 941-959.

[12] Miller K., Ross B.(1993). An introduction to the fractional calculus and fractional differential equations John Wiley and Sons.

[13] Najafalizadeh S., Ezzati R., Numerical methods for solving two-dimensional nonlinear integral equations of fractional order by using two-dimensional block pulse operational matrix, Applied Mathematics and Computation, 280 (2016) 46-56.

[14] Podlubny I.(1999). Fractional differential equations, New York: Academic Press.

[15] Rudin W. Principles of mathematical analysis. Singapore: McGraw-Hill; 1976.

[16] Ray S.(2012). On Haar wavelet operational matrix of general order and its application for the numerical solution of fractional Bagley-Torvik equation, Appl. Math. Comput., 218: 5239-5248.

[17] Saeedi H., Mollahasani N., Mohseni Moghadam M., Chuev G. N., An operational Haar wavelet method for solving fractional Volterra integral equations, Journal of Applied Mathematics and Computer Science, 21 (3) (2011) 535-547.

[18] Saeedi H., Mollahasani N., Mohseni Moghadam M., Chuev G. N.(2011). An operational Haar wavelet method for solving fractional Volterra integral equations, Journal of Applied Mathematics and Computer Science, 21(3): 535-547.

[19] Strang G.(1989). Wavelets and dilation equations, SIAM 31: 614-627.

[20] Vetterli M., Kovacevic J.(1995). Wavelets and subband coding, Prentice-Hall, Englewood Cliffs, New Jersey.

[21] Y. Li, W. Zhao, Haar wavelet operational matrix of fractional order integration and its applications in solving the fractional order differential equations, Appl. Math. Comput., 216 (2010) 2276-2285.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–176, pp. 190–192

Solving generalized tensor eigenvalue problems

Mehri Pakmanesh[a,], Hamireza Afshin[a], Masoud Hajarian[b]

[a] Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran
[b] Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran

Article Info

Keywords:
modified normalized Newton method
$\mathcal{Z}$-eigenpairs
$\mathcal{H}$-eigenpairs
generalized tensor eigenvalue problem
generalized eigenproblem adaptive power method

Abstract

The problem of generalized tensor eigenvalue is the focus of this paper. To solve the problem, we suggest using the normalized Newton generalized eigenproblem approach (NNGEM). Since the rate of convergence of the spectral gradient projection method (SGP), the generalized eigenproblem adaptive power (GEAP), and other approaches is only linear, they are significantly improved by our proposed method, which is demonstrated to be locally and cubically convergent.

1. Introduction

A numerical algorithm for a generalized tensor eigenproblem was proposed by Kolda and Mayo [1]. They solved the optimization problem in order to apply the adaptive shifted power method (GEAP).

$$\max_{\|y\|=1} \frac{\mathcal{A}y^m}{\mathcal{B}y^m}\|y\|^m.$$

The shifted symmetric higher-order power method (SS-HOPM) for locating $\mathcal{Z}$-eigenpairs is extended by this method, which also includes an adaptive technique for automatically selecting the shift. With this method, the rate of convergence is only linear. An adaptive gradient (AG) approach was presented by Yu et al. in 2016 to address generalized tensor eigenvalue problems. Under certain appropriate conditions, the method establishes both global convergence and linear convergence rates [2]. Moreover, two convergent gradient projection techniques were presented by Zhao, Yang, and Liu to address the weakly symmetric tensors generalized eigenvalue problem [3]. One could consider AGP to be an adaptation of the GEAP technique. The BB method plus the gradient projection method yields SGP, which is better than the GEAP, AG, and AGP methods. Nevertheless, not all potential $\mathcal{Z}$-eigenpairs and generalized eigenvalues of tensors can be found using the GEAP, AG, AGP, and SGP methods.
The computation of a symmetric tensor's $\mathcal{Z}$-, $\mathcal{H}$-, and generalized eigenpairs constitutes the paper's major contribution. We present the normalized Newton method (NNGEM) for solving the generalized tensor eigenproblem, which is based on the relationship between the even degree homogeneous polynomial and even order symmetry. Every NNGEM iteration involves computing a new approximation in two steps; in particular, the second step can be computed quickly by utilizing the qualities that were computed in the first step. The method consistently leads to B-eigenpairs of a

cubical rate symmetric tensor. The NNGEM approach has a higher rate of convergence than the SGP and GEAM approaches. Compared to current methods, ours is a more efficient one, as shown by numerical examples.

2. Preliminaries

A symmetric tensor is a tensor that is invariant under a permutation of its vector arguments:

$$\mathcal{A}(i_1 \dots i_m) = \mathcal{A}(\sigma_1 \dots \sigma_m)$$

for every permutation σ of the symbols $\{1, 2, \dots, r\}$. Denote the set of all mth order, n-dimensional real symmetric tensors by $S^{[m,n]}$. For any $y \in \mathbb{R}^n$, the following frequently used notations are worthy of consideration. If y is a vector, $\mathcal{A}y^{m-1}$ is a vector with

$$(\mathcal{A}y^{m-1})_i = \sum_{i_2,\dots,i_m=1}^{n} a_{ii_2\dots i_m} y_{i_2} y_{i_3} \cdots y_{i_m},$$

for $i = 1, 2, \dots, n$, $\mathcal{A}y^m$ is a scalar as

$$\mathcal{A}y^m = \sum_{i_1,\dots,i_m=1}^{n} a_{i_1\dots i_m} y_{i_1} y_{i_2} \cdots y_{i_m}$$

and $\mathcal{A}y^{m-2} \in \mathbb{R}^{n\times n}$ is a matrix defined by

$$\mathcal{A}(y) := \mathcal{A}y^{m-2} = \left(\sum_{i_3,i_4,\dots,i_m=1}^{n} a_{iji_3\dots i_m} y_{i_1} y_{i_3} \cdots y_{i_m} \right). \tag{1}$$

An mth order, n-dimensional real tensor $\mathcal{A}$ is called positive definite if $\mathcal{A}y^m > 0, \forall y \in \mathbb{R}^n \setminus \{0\}$.

Definition 2.1. Let $\mathcal{A}$ and $\mathcal{B}$ be two mth order, n-dimensional symmetric real tensors. Assume that $\mathcal{A}y^{m-1}$ and $\mathcal{B}y^{m-1}$ are not identical to zero. If $y \in \mathbb{R}^n \setminus \{0\}$ and $\lambda \in \mathbb{R}$ satisfy

$$\mathcal{A}y^{m-1} = \lambda \mathcal{B}y^{m-1}, \tag{2}$$

then λ is called a $\mathcal{B}$-eigenvalue of $\mathcal{A}$, and y is called a $\mathcal{B}$-eigenvector of $\mathcal{A}$. Assume that m is even and $\mathcal{B}$ is positive definite. It is clear that any $\mathcal{B}$-eigenpair in (2) satisfies

$$\lambda = \frac{\mathcal{A}y^m}{\mathcal{B}y^m}. \tag{3}$$

We designate a number $\lambda \in \mathbb{C}$ as an eigenvalue of $\mathcal{A}$ if it satisfies the following homogeneous polynomial equations when combined with a nonzero vector $y \in \mathbb{C}^n$.

$$(\mathcal{A}y^{m-1})_i = \lambda y_i^{m-1}, \qquad \text{for } i = 1, \dots, n. \tag{4}$$

We refer to y in this case as an eigenvector of $\mathcal{A}$ that is connected to the eigenvalue λ. Then, (4) can be simply expressed as

$$\mathcal{A}y^{m-1} = \lambda y^{[m-1]}. \tag{5}$$

If $y^{[m-1]} = (y_1^{m-1}, y_2^{m-1}, \dots, y_n^{m-1})^T$ as a vector in $\mathbb{R}^n$, then x is an $\mathcal{H}$-eigenvector of $\mathcal{A}$ connected to the $\mathcal{H}$-eigenvalue λ . This is equivalent to a generalized tensor eigenpair with $b_{i_1,i_2,\dots,i_m} = \delta_{i_1,i_2,\dots,i_m}$.
If $y \in \mathbb{R}^n \setminus \{0\}$ and $\lambda \in \mathbb{R}$ satisfy

$$\mathcal{A}y^{m-1} = \lambda y \qquad and \qquad y^T y = 1, \tag{6}$$

y is referred to as $\mathcal{A}$'s $\mathcal{Z}$-eigenvector, and λ is referred to as its $\mathcal{Z}$-eigenvalue. In terms of a generalized tensor eigenpair, this is represented by $\mathcal{B} = \mathcal{E} = (e_{i_1,i_2,\ldots,i_m})$, with entries defined by

$$e_{i_1,i_2,\ldots,i_m} = \begin{cases} 1 & i_1 = i_2, i_3 = i_4, \ldots, i_{m-1} = i_m \\ 0 & otherwise. \end{cases}$$

In this case, we have $\mathcal{B}y^{m-1} = x$ and $\mathcal{B}y^m = 1$. Then, (2) is equivalent to

$$\mathcal{A}y^{m-1} = \lambda y. \tag{7}$$

3. Main results

Theorem 3.1. Assume that $\mathcal{A} = (a_{i_1,\ldots,i_m})$, $\mathcal{B} = (b_{i_1,\ldots,i_m}) \in S^{[m,n]}$, that $\mathcal{B}$ is positive definite, and that c is a nonzero real number. The following nonlinear system of equations has (λ_*, y_*) as a nonzero solution if and only if (λ_*, y_*) is a $\mathcal{B}$-eigenpair of $\mathcal{A}$.

$$E(y; b) := \mathcal{A}y^{m-1} - \lambda_F(y)\mathcal{B}y^{m-1} + \frac{b}{2}(y^T y - 1)y = 0, \tag{8}$$

and $\lambda_* = \lambda_F(y_*)$, where $\lambda_F(y) = \frac{\mathcal{A}y^m}{\mathcal{B}y^m}$.

We define the following notations:

$$\lambda_F(y) := \phi(x), \qquad E(y; b) := f(y), \qquad B(y) = (m-1)\mathcal{A}y^{m-2} - (m-1)\mathcal{B}y^{m-2}\lambda_F(y) - \mathcal{B}y^{m-1}(\lambda'_F(y))^T.$$

Specifically, at $\|y\| = 1$, we have

$$\begin{aligned} E'(y; b) &= (m-1)\mathcal{A}y^{m-2} - (m-1)\mathcal{B}y^{m-2}\lambda_F(y) - \mathcal{B}y^{m-1}(\lambda'_F(y))^T + byy^T. \\ &= B(y) + byy^T \end{aligned} \tag{9}$$

Theorem 3.2. *Let $\mathcal{A}, \mathcal{B} \in S^{[m,n]}$ and $\mathcal{B}$ be a positive definite tensor. Assume that (λ_*, y_*) a $\mathcal{B}$-eigenpair of $\mathcal{A}$. Then $H'(y_*; b)$ is nonsingular matrix.*

3.1. Example of $\mathcal{B}$-eigenpair

[1] Let $\mathcal{A}, \mathcal{B} \in S^{[4,3]}$ be defined by

$$\begin{aligned} a_{1111} = 0.4982, a_{1112} = -0.0582, a_{1113} = -1.1719, a_{1122} = 0.2236, \\ a_{1123} = -0.0171, a_{1133} = 0.4597, a_{1222} = 0.4880, a_{1223} = 0.1852, \\ a_{1233} = -0.4087, a_{1333} = 0.7639, a_{2222} = 0.0000, a_{2223} = -0.6162, \\ a_{2233} = 0.1519, a_{2333} = 0.7631, a_{3333} = 2.6311, \end{aligned}$$

and

$$\begin{aligned} b_{1111} = 3.0800, b_{1112} = 0.0614, b_{1113} = 0.2317, b_{1122} = 0.8140, \\ b_{1123} = 0.0130, b_{1133} = 2.3551, b_{1222} = 0.0486, b_{1223} = 0.0616, \\ b_{1233} = 0.0482, b_{1333} = 0.5288, b_{2222} = 1.9321, b_{2223} = 0.0236, \\ b_{2233} = 1.8563, b_{2333} = 0.0681, b_{3333} = 16.0480. \end{aligned}$$

In this case, we want to calculate a symmetric tensor's $\mathcal{B}$-eigenpairs.

References

[1] Kolda TG, Mayo JR. An adaptive shifted power method for computing generalized tensor eigenpairs. SIAM J Matrix Anal Appl 2014; 35(4), 1563-1581.

[2] Yu G, Yu Z, Xu Y. An adaptive gradient method for computing generalized tensor eigenpairs. Comput Optim Appl 2016; 65, 781-797 . https://doi.org/10.1007/s10589-016-9846-9.

[3] Zhao N, Yang Q, Liu Y. Computing the generalized eigenvalues of weakly symmetric tensors. Comput Optim Appl 2017; 66, 285-307. https://doi.org/10.1007/s10589-016-9865-6.

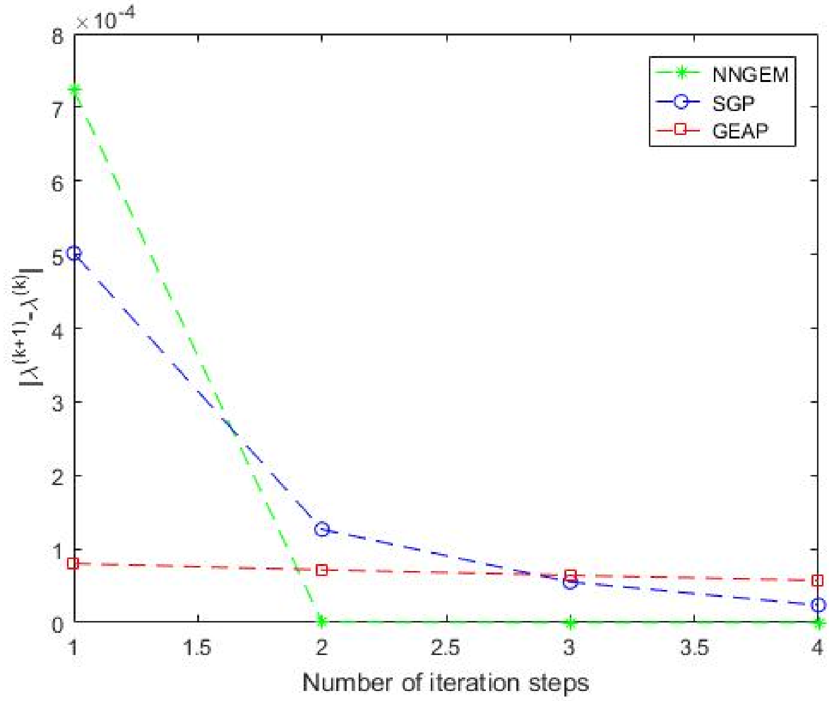


Fig. 1. As the number of iterations increases, the error value changes.

Table 1. Comparison of results for computing $\mathcal{B}$-eigenpairs of $\mathcal{A}$ in Example **??**.

	GEAP			SGP			NNMGET		
λ	**Its**	**Time**	**Error**	**Its**	**Time**	**Error**	**Its**	**Time**	**Error**
0.4359	117	1.241331	0	30	1.045496	2.5403e-08	8	0.293241	0
0.2219	500	4.898877	0	31	1.024953	4.9453e-08	5	0.214433	0
0.5356	206	2.094713	0	16	0.554923	1.6935e-08	8	0.287557	0

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–212, pp. 193–197

Generation of the Structured Light beams by Special Functions

Fazel Saadati-Sharafeh[a,*]

[a]*Department of Mathematics, Payame Noor University(PNU), Tehran, Iran*

Article Info

Keywords:

Bessel beam
Hermite-Gaussian beam
Laguerre-Gaussian beam

Abstract

This paper introduces some of the structured light beams generated by special mathematical functions. For this purpose, the solutions of the Helmholtz equation are obtained in Cartesian, Circular Cylindrical, Elliptical Cylindrical, and Parabolic Cylindrical coordinate systems. We introduce several types of well-known structured light beams.

1. Introduction

In the 16th and 17th centuries, some scientists, such as Newton, considered the physics of light as a particle, while others, such as Huygens, considered it as a wave. But in the 20th century, it was found that light has both wave and particle properties, and at the same time, it is neither of these two. This state of light is called wave-particle duality. Maxwell's equations are set of differential equations with partial derivatives that describe the time dependence of the two electric and magnetic fields of electromagnetic waves. He showed that by using these four equations, all electrical and magnetic phenomena can be described. These equations in free space are as[1]:

$$\nabla \cdot \mathbf{E} = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}, \tag{1}$$

where the vectors $\mathbf{E}$ and $\mathbf{B}$ are the electric field and the magnetic field, respectively, and c is the speed of light in vacuum. Considering that light is electromagnetic phenomenon, Maxwell's equations must also be applicable light. Getting ∇ from Maxwell's equations, we have

$$\nabla^2 \mathbf{E} - \frac{1}{c^2}\frac{\partial^2 \mathbf{E}}{\partial t^2} = 0, \quad \nabla^2 \mathbf{B} - \frac{1}{c^2}\frac{\partial^2 \mathbf{B}}{\partial t^2} = 0, \tag{2}$$

which are known as Almbert's homogeneous wave equation. In practice, due to the measurable nature of the electric field, the electric field is used more than the magnetic field. Now we have to find the solutions of eq.2. A separable monotonic wave function as a solution of the wave equation is in the form $\mathbf{E}(\mathbf{r}, t) = e^{-i\omega t}\mathbf{E}(\mathbf{r})$, where ω is a positive real number, the frequency of the wave and r is the position vector. By putting this function in the wave eq.2, we have

$$\nabla^2 \mathbf{E}(\mathbf{r}) + k^2 \mathbf{E}(\mathbf{r}) = 0, \tag{3}$$

*Talker
Email address: fazelsaadati@pnu.ac.ir (Fazel Saadati-Sharafeh)

where $k = \frac{\omega}{c}$ is the wave number. This equation is known as the Helmholtz equation. Therefore, the propagation of an optical field is described by the three-dimensional Helmholtz equation. Solving this equation in different coordinate systems, it can be seen that this equation is separable in eleven orthogonal coordinate systems, which is symmetrical in only four of them and divided into two longitudinal and transverse parts. These coordinate systems are Cartesian, Circular Cylindrical, Elliptical Cylindrical, and Parabolic Cylindrical Coordinate Systems.[2]
There are a set of special functions such as Gaussian and Hypergeometric functions and a set of orthogonal polynomials or special functions such as Bessel, Laguerre, associated Laguerre, Hermite functions, which are very important in mathematical physics. These functions and their combinations are used in the production of structured light beams that have a particular mathematical and physical structure and a particular total momentum value.[3].

2. Solving the Helmholtz equation in two-dimensional orthogonal coordinate systems

Using the method of separation of variables to solve differential equations with second-order partial derivatives of n variables, turns it into a system of n second-order ordinary differential equations. Now, using this method, we want to obtain the exact solutions of the Helmholtz equation in the four mentioned orthogonal coordinate systems[4].

2.1. Solving the Helmholtz equation in the Cartesian coordinate system

In this orthogonal coordinate system, we are looking for solutions in the form of $E(x, y) = X(x)Y(y)$ for equation 3. As a result, the Helmholtz equation becomes the form:

$$X''Y + XY'' + k^2XY = 0. \tag{4}$$

This equation can be arranged as follows by dividing by XY:

$$\frac{X''}{X} = -\frac{Y''}{Y} - k^2 \tag{5}$$

where the left side of the relationship is only a function of x and the right side is only a function of y and this is only possible when both sides of the equation are equal to a constant number such as $-h^2$ be equal, which is called the separation constant. Therefore, the equation 5 will be equivalent to the following two equations

$$X''(x) + h^2X(x) = 0, \quad Y''(y) + (k^2 - h^2)Y(y) = 0. \tag{6}$$

The general solutions of the first equation with condition $h \neq 0$ are

$$X_1(x) = \exp(ihx), \quad X_2(x) = \exp(-ihx). \tag{7}$$

While with the condition $k^2 - h^2 \neq 0$, the general solutions of the second equation will be

$$Y_1(y) = \exp(i\sqrt{k^2 - h^2}y), \quad Y_2(y) = \exp(-i\sqrt{k^2 - h^2}y). \tag{8}$$

Therefore, the solutions of the eq.3 can be written as a linear combination of solutions type 7 and 8 as follows:

$$E_k(x, y) = \sum_{i,j=1}^{2} A_{i,j}X_i(x)Y_j(y) \tag{9}$$

where the complex coefficients of $A_{i,j}$ are arbitrary. As an example of the answers we have

$$E_h(x, y) = X_1(x)Y_1(y) = \exp\left[i(hx + \sqrt{k^2 - h^2}y)\right]. \tag{10}$$

2.2. Solving the Helmholtz equation in the polar coordinate system

The polar coordinate system has

$$x = r\cos\theta, \quad y = r\sin\theta. \tag{11}$$

With this assumption, the Helmholtz equation transforms into the following form:

$$\left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial \theta^2} + k^2\right)E(r,\theta) = 0. \tag{12}$$

Assuming $\mathbf{E}(r,\theta) = \mathbf{R}(r)\theta(\theta)$, we have

$$\frac{r^2\mathbf{R}'' + r\mathbf{R}' + r^2k^2}{\mathbf{R}} = -\frac{\theta''}{\theta}(= h^2) \tag{13}$$

which leads to the equation $\theta'' + h^2\theta = 0$ with the general solution $\theta(\theta) = \exp(\pm ih\theta)$ and Bessel's differential equation

$$r^2\mathbf{R}'' + r\mathbf{R}' + (r^2k^2 - h^2)\mathbf{R} = 0 \tag{14}$$

With the solution $\mathbf{R}(r) = J_{\pm h}(kr)$. So

$$E_h(r,\theta) = J_h(kr)\exp(ih\theta) \tag{15}$$

is another exact solution of the Helmholtz equation.

2.3. Solving the Helmholtz equation in the elliptic coordinate system

In the elliptic coordinate system (ξ,η), we have

$$x = d\cosh\xi\cos\eta, \quad y = d\sinh\xi\sin\eta \tag{16}$$

where $\xi \geq 0, 0 \leq \eta \leq 2\pi$. With this transformation, the ellipses and hyperbolas obtained for each $\xi \in [0,+\infty]$ and $\eta \in [0,2\pi]$ are the orthogonal lines of this system. With this transformation, the Helmholtz equation will be transformed into the form

$$\left(\frac{\partial^2}{\partial \xi^2} + \frac{\partial^2}{\partial \eta^2} + d^2k^2(\cosh^2\xi - \cos^2\eta)\right)E(\xi,\eta) = 0 \tag{17}$$

which obtained the following two equations by assuming $E(\xi,\eta) = U(\xi)V(\eta)$

$$U'' + (d^2k^2\cosh^2\xi + h^2)U = 0, \quad V'' - (d^2k^2\cos^2\eta + h^2)V = 0. \tag{18}$$

These equations are known as the Mathieu equations and generate the Mathieu functions.

2.4. Solving the Helmholtz equation in the parabolic coordinate system

Using $x = \frac{1}{2}(u^2 - v^2), y = uv$, the Cartesian coordinate system (x,y), transforms into the parabolic coordinate system (u,v). With this transformation, the coordinate lines

$$u = \sqrt{\sqrt{x^2+y^2}+x}, \quad v = \pm\sqrt{\sqrt{x^2+y^2}-x} \tag{19}$$

form two orthogonal families. With this transformation, the Helmholtz equation becomes the form:

$$\left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2} + (u^2+v^2)k^2\right)E(u,v) = 0. \tag{20}$$

Assuming $E(u,v) = U(u)V(v)$, we have

$$U'' + (k^2u^2 - h^2)U = 0, \quad V'' + (k^2v^2 + h^2)V = 0 \tag{21}$$

where k^2 is the separation constant. Parabolic beams are solutions of these equations.

3. Structured light beams generated by special functions

3.1. The Bessel beams

Bessel beams are one of the most important structured light beams whose solution to the Helmholtz equation is in circular cylindrical coordinate system and their amplitude are described by the first kind Bessel function. Bessel-Gaussian beams with limited energy were introduced by Gori and Guattari in[5]. The amplitude of these beams is expressed by the product of the Gaussian function and the first kind Bessel functions of nth order. These beams carry orbital angular momentum and their intensity distribution has circular symmetry. Considering the ability of rotating beams carrying orbital angular momentum in improving the efficiency of communication, Bessel beams are of high importance and are suitable for directing atoms. The complex amplitude of a Bessel beam in cylindrical coordinate system is defined as[6]:

$$E_n(r,\phi,z) = \exp\left[in\phi + iz\sqrt{k^2-\alpha^2}\right] J_n(\alpha r). \tag{22}$$

where k is the wave number, $\alpha = k\sin(\theta_0)$, where θ_0 is the cone wave angle and J_n is the first kind Bessel functions of nth order. Figure 1 depict the bessel beams of order 0 to 5 with $\lambda = 532\mu m$, $\alpha = \frac{1}{10\times\lambda}$ and $-100\mu m \leq x, y \leq 100\mu m$ respectively. obviousely, as the order of the Bessel beam increases, the diagonal of the main central ring of the beam also increases.

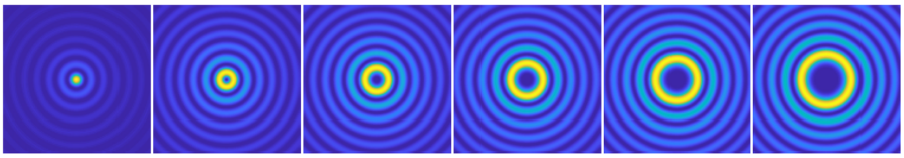


Fig. 1. The Bessel beams of order 0 to 5 with $\lambda = 532\mu m$, $\alpha = \frac{1}{10\times\lambda}$ and $-100\mu m \leq x, y \leq 100\mu m$ respectively.

3.2. The Laguerre-Gaussian beams

Among structured light beams, Laguerre-Gaussian (LG) beams are one of the well-defined optical fields. The shape of the transverse intensity of these fields is uniform during propagation in a homogeneous environment and they have axial symmetry in cylindrical coordinate system. These vortex beams also carry orbital angular momentum [7] and special attention has been paid to these types of beams for the transmission of quantum information [8] and the manipulation of microparticles [9]. Complex amplitude of the LG beams in Cartesian coordinate system is defined as[10]:

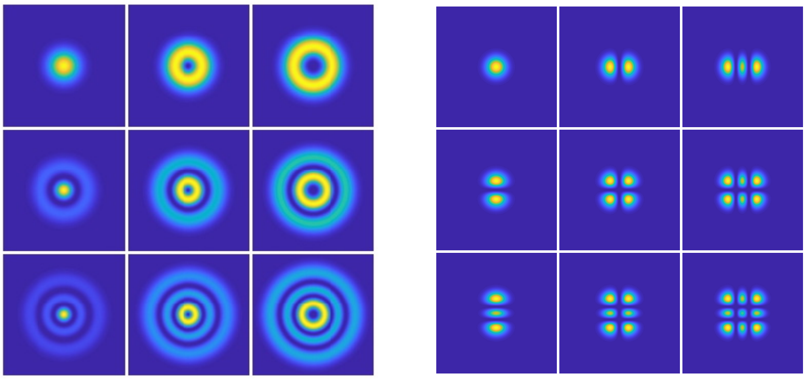
$$E_{nm}(x,y,z) = \frac{w_0}{w(z)}\left(\frac{\sqrt{2}}{w(z)}\right)^m (x\pm iy)^m L_n^m\left[\frac{2(x^2+y^2)}{w^2(z)}\right]\exp\left[-\frac{(x^2+y^2)}{w^2(z)}\right]\exp\left[-i\left((m+2n+1)\zeta(z)-\frac{k(x^2+y^2)}{2R(z)}\right)\right] \tag{23}$$

where w_0 is the waist radius of the Gaussian beam, λ is the wavelength, $w(z) = w_0\sqrt{1+\left(\frac{z}{z_R}\right)^2}$, beam radius, $z_R = \frac{\pi w_0^2}{\lambda}$, Rayleigh range, $\zeta(z) = \arctan\left(\frac{z}{z_R}\right)$ is the Gouy phase, m is the topological charge, and L_n^m is the associated Laguerre polynomial. Figure 2a shows the intensity distribution of the LG beams with n and $m = 0, 1, 2$.

3.3. The Hermite-Gaussian beams

The Hermite-Gaussian (HG) beams as an another kind of structured light beams in cartesian coordinate system have been known in optics by Kogelnik[11]. This kind of beams have no orbital angular momentum (OAM). Complex amplitude of the HG beams in Cartesian coordinates is defined as follows[12]:

$$E_{nm}(x,y,z) = \frac{w_0}{w(z)} H_n\left[\frac{\sqrt{2}}{w(z)}x\right] H_m\left[\frac{\sqrt{2}}{w(z)}y\right]\exp\left[-\frac{(x^2+y^2)}{w^2(z)}\right]\exp\left[-i\left(kz-(1+n+m)\zeta(z)+\frac{k(x^2+y^2)}{2R(z)}\right)\right] \tag{24}$$



(a) The LG beams (b) The HG beams

Fig. 2. Intensity distribution of the LG and HG beams with $n = 0, 1, 2$ in rows and $m = 0, 1, 2$ in columns, respectively.

where w_0 is the waist radius of the Gaussian beam, λ is the wavelength, $w(z) = w_0\sqrt{1 + \left(\frac{z}{z_R}\right)^2}$, beam radius, $z_R = \frac{\pi w_0^2}{\lambda}$, Rayleigh range, $\zeta(z) = \arctan\left(\frac{z}{z_R}\right)$ is the Gouy phase, and H_n is the Herimtian polynomials of order n. The intensity distribution of the HG beams with n and $m = 0, 1, 2$ is shown in Fig 2b.

4. Conclusion

In this study, we introduce some of the well-known structured light beams like Bessel, Hermite-Gaussian and Laguerre-Gaussian beams, generated by special mathematical functions. For this purpose, the solutions of the Helmholtz equation are obtained in Cartesian and Circular Cylindrical coordinate systems.

References

[1] Pedrotti. F. L., L. S. Pedrotti, 1993. Introduction To Optics (Second Edition). Prentice Hall.

[2] Khonina, S.N., A.V. Ustinov, and A.P. Porfirev, 2018. Aberration laser beams with autofocusing properties. *Applied Optics*, 57(6): 1410-1416.

[3] Arefken. G., 1985. Mathematical Methods for Physicists (Third Edition). Academic Press Inc.

[4] Miller. W., 1977. Symmetry and Seperation of Variables. Addison-Wesley, Reading, Massachusetts.

[5] Gori, F., and G. Guattari, 1987. Bessel-Gauss beams. *Optics Communications*, 64: 491-495.

[6] Kovalev, A.A., V.V. Kotlyar, and A.P. Porfirev, 2015. Shifted nondiffractive Bessel beams. *Physical Review A*, 91: 053840.

[7] Allen, L., M.W. Beijersbergen, R.J.C. Spreeuw, and J.P. Woerdman, 1992. Orbital angular momentum of light and the transformation of laguerre-gaussian laser modes. *Physical Review A*, 110(11): 8185-8189.

[8] Cheng, M., L. Guo, J. and Li, Y. Zhang, 2016. Channel capacity of the OAM-based free-space optical communication links with Bessel-Gauss beams in turbulent ocean. *IEEE Photon. J*, 8: 7901411.

[9] Simpson, N.B., K. Dholakia, L. Allen, and M.J. Padgett, 1997. Mechanical equivalence of spin and orbital angular momentum of light: an optical spanner. *Optics Letter*, 22: 5254.

[10] A. A. Kovalev, V. V. Kotlyar, and A. P. Porfirev, 2016. Asymmetric Laguerre-Gaussian beams. *PHYSICAL REVIEW A*, 93: 063858.

[11] H. Kogelnik and T. Li, 1966. □Laser beams and resonators. *Proc. IEEE*, 54: 1312□1329.

[12] V. V. Kotlyar and A. A. Kovalev,2014. Hermite - Gaussian modal laser beams with orbital angular momentum. *J. Opt. Soc. Am. A*, 31(2)274-282.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–213, pp. 198–202

Laguerre-Gaussian beams and their quantified self-healing characteristic

Fazel Saadati-Sharafeh[a,*]

[a]Department of Mathematics, Payame Noor University(PNU), Tehran, Iran

Article Info

Keywords:
Huygens Convolution Method
Laguerre-Gaussian beams
Self-healing property

Abstract

In this paper, the self-healing property of the Laguerre-Gaussian (LG) beams were reported in detail. In all cases, the self-healing value were precisely quantified. For numerical simulation and calculation the similarity function, the Huygens Convolution Method (HCM) were used.

1. Introduction

In recent years, using mathematical special functions in generation and propagation of structured light are noticed. They are easily generated and have a mathematical representation[1]. Also a lot of them have structured stability in propagation during. Among these structural lights, the LG beams consist a class of well-studied optical fields. The transverse intensity shapes of these fields in propagation during in homogeneous (isotropic) medium are uniform and have axial symmetry in cylindrical coordinates. These vortex beams, carrying orbital angular momentum too[2].
There are several works to change various beams, carrying angular momentum by several researchers. For instance, Wada et al.[3], clarified the entire range of transformations that an LG beam with astigmatism can go through in free space. They classified all patterns of beam transformation and map them to initial beam conditions. Mei et al.[4], proposed the concept of vectorial Laguerre-Bessel-Gaussian beams, and based on vectorial Rayleigh-Sommerfeld formulas, they derived the analytical formulas for the non paraxial propagation properties of Laguerre-Bessel-Gaussian beams. The generation of several vortex beams by coaxial superposition of two, three, and four LG beams is proposed by Huang et al.[5]. Plick et al.[6], explained the physical meaning of the radial index of LG beams. Stilgoe et al.[7], investigated the energy, momentum, and propagation of the solutions of the paraxial wave equation like, LG, Hermite-Gaussian and Ince-Gaussian.
Obviously, one of the most interesting and important properties of the structured light beams, is their self-healing characteristic, that describes the ability of the self-construction of the amplitude of beam at the minimum distance beyond the obstruction. Chu et al.[8], analytically studied the self-healing characteristic of the Airy beams in free space. Vaity et al.[9] generated optical ring lattice structures of superposition of two coaxial LG beams with same waist position and waist parameter, experimentally. The self-healing property of the optical Airy beam is investigated by Zhang et al.[10], analytically. Litvin et al.[11] studied theoretically and experimentally, the self-healing property of

*Talker
Email address: fazelsaadati@pnu.ac.ir (Fazel Saadati-Sharafeh)

Bessel-like beams. They found that self-construction of them are similar to Bessel beams but depends on the distance between the initial field and the obstruction. A family of asymmetric LG beams, is introduced by Kovalev et al.[12], via a complex shift of LG beams in Cartesian coordinates, and used asymmetric LG beams to optical trapping and moving of microparticles[13].
The self-healing property of beams is very useful in optical tweezing, Optical manipulation, Optical microscopy, optical trapping, atmospheric sciences, optical communications, especially, in inhomogeneous media. In this paper, We have reported the detailed quantitatively analyzed self-healing behavior of the LG.

2. Theory of the LG beams

Complex amplitude of the LG beams in Cartesian coordinates is described by[12]:

$$E_{nm}(x,y,z) = \frac{w_0}{w(z)}\left[\frac{\sqrt{2}}{w(z)}\right]^m (x+iy)^m L_n^m\left[\frac{2(x^2+y^2)}{w^2(z)}\right]\exp\left[-\frac{(x^2+y^2)}{w^2(z)}+\frac{ik(x^2+y^2)}{2R(z)}-i(m+2n+1)\zeta(z)\right], \quad (1)$$

where w_0, is the Gaussian beam waist radius, λ, is the wave length, $z_R = \frac{\pi w_0^2}{\lambda}$, is the Rayleigh range, $w(z) = w_0\sqrt{1+(\frac{z}{z_R})^2}$, L_m^n, is the associated Laguerre polynomial, $\zeta(z) = \arctan\left(\frac{z}{z_R}\right)$, is the Gouy phase, $R(z) = z[1+(\frac{z_R}{z})^2]$, and m is the vortex topological change.

3. Quantification of the self-healing property of beams

Depicting numerical values of the beam's self-healing, we have define the following similarity function[14], that gives the similarity percent of the propagated beams without mask A and those propagated beams with circular mask B on propagation during.

$$similarity(A,B) = \left(1-\sqrt{\frac{\|A-B\|}{\|A+B\|}}\right)\times 100, \quad (2)$$

where the power of a matrix is the power of all entries and the norm is the Frobenius norm of matrix. To more clarity, we use positive self-healing value and negative self-healing value terminologically. The positive self-healing value means that the similarity percent of constructive beam respect to original beam is more than the similarity percent of masked beam at initial plane and the negative self-healing value means that the similarity percent of constructive beam respect to original beam is less than the similarity percent of masked beam at initial plane.

4. Simulation of propagation

There are several methods to calculate diffraction of the beams in their propagation during. Among them, we can list Fresnel Transform method, Angular Spectrum method and Huygens Convolution method(HCM). We apply HCM to the beams' propagation simulation. The Huygens Convolution integral is described by[15]:

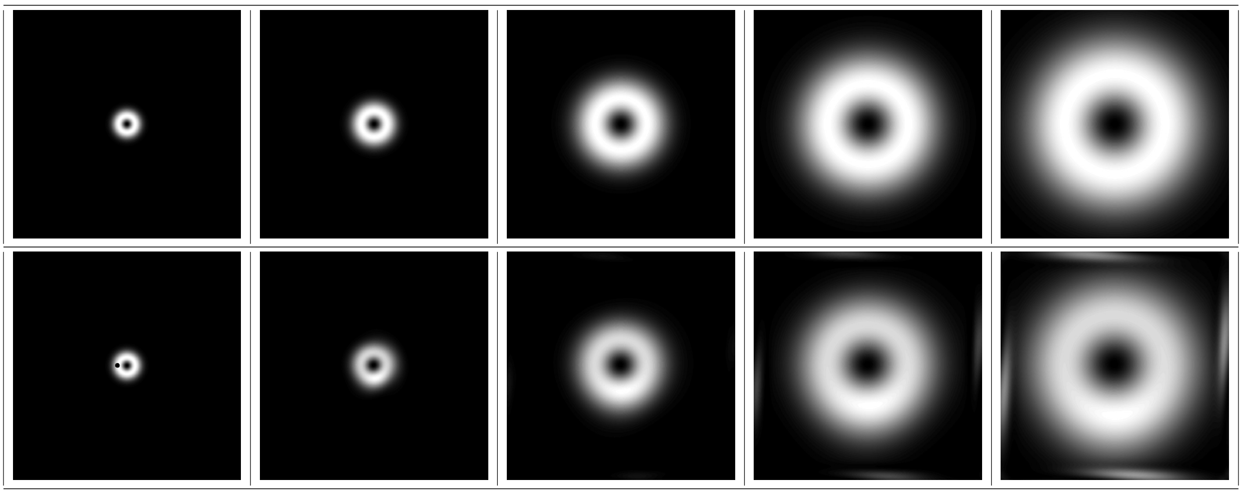
$$E(x,y,z) = \mathfrak{F}^{-1}\{\mathfrak{F}\{E_0(x_0,y_0)\}[k_x,k_y]\mathfrak{F}\{S_H(x_0,y_0)\}[k_x,k_y]\}[x,y], \quad (3)$$

where $\mathfrak{F}$, is fourier transform and the point propagate function of Huygens is

$$S_H(x,y,z) = -\frac{ik}{2\pi z}\exp(ik\sqrt{x^2+y^2+z^2}). \quad (4)$$

5. Detailed Quantification of Self-Healing Property of beams

In this section, we detail observationally and quantitatively, the self-healing property of the LG beams setting two area of circular mask 314 (Mask1) and mask 707 (Mask2) pixels for various topological charges $m = 1, 2, 5$. The self-healing Behavior of beams is investigated in 1000×1000 pixels sampled array with beam waist radius 1000λ. We propagate the beams with $w_0 = 1000\lambda$ in $0 - 10$ meter range.
The LG beams as one of important beams are noticed by optics researchers that have many applications. Table (1) depict the LG beams with topological charge $m = 1$, 1000λ beam waist radius, and Mask1 propagated via HCM (raw 2), in 5 arbitrary distances and similarity values of them respect to propagated beams with no mask (raw 1). For instance, similarity value of the LG beams respect to the LG beam with no mask in initial plane $z = 0$ equals 95.06 percent and maximum similarity value at $z_{opt} = 5$ meter distance of initial plane is 96.46 percent that has positive self-healing value.

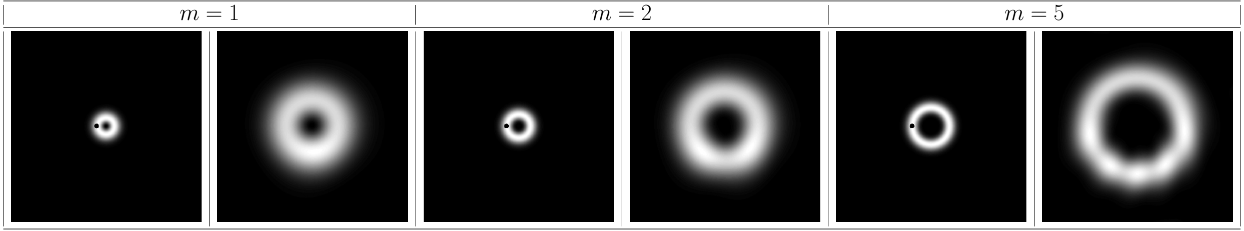


$z = 0\ m$	$z = 2\ m$	$z_{opt} = 5\ m$	$z = 8\ m$	$z = 10\ m$
95.06	95.91	96.46	95.87	94.69

Table 1. propagation of masked LG beams via HCM with topological charge $m = 1$, 1000λ beam waist and their similarity values respect to original propagated beams

In table (2), have been studied quantified self-healing behavior of the LG beams propagated via HCM with 1000λ beam waist and Mask1 for various topological charge $m = 1, 2, 5$. Results show that LG beams have self-healing property observationally, and have positive self-healing quantitatively. Intensity profile of the LG beams are asymmetric. The bright spots are centered on the number of topological charge m on the LG ring. For topological charge $m = 2$ ($m =$

$m = 1$		$m = 2$		$m = 5$	



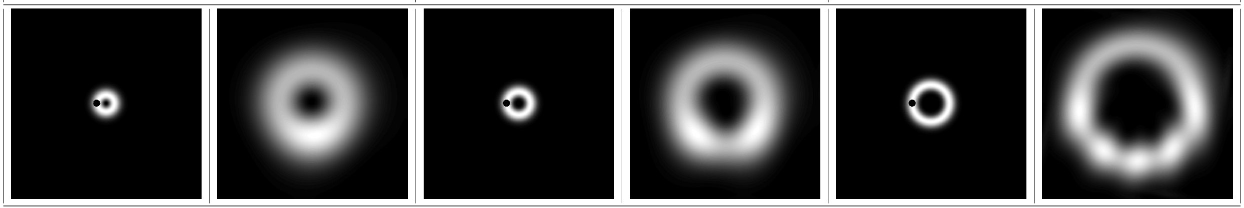
$z = 0\ m$	$z_{opt} = 5\ m$	$z = 0\ m$	$z_{opt} = 4.6\ m$	$z = 0\ m$	$z_{opt} = 4\ m$
95.06	96.46	95.71	96.66	95.56	96.93

Table 2. propagation of the LG beams via HCM with topological charge $m = 1, 2, 5$, 1000λ beam waist radius, circular mask of 314 pixels, and their self-healing values respect to original propagated beam

5), the self-healing value of the LG beam with circular mask after propagation, from 95.71 (96.56) percent at initial

plane $z = 0$, increases to 96.66 (96.3) percent at $z_{opt} = 4.6\ m$ ($z_{opt} = 4\ m$). In this case, both of them have positive self-healing value. Obviously, by increasing of topological charge m, difference of self-healing value of beams at initial and optimum planes with exerting mask decreases consistently. Also because of mask1 in this table The bright spots are centered on the number of topological charge m on the LG ring, too.
Table (3), depicts quantified self-healing behavior of the LG beams propagated via HCM with 1000λ beam waist and Mask2 for various topological charge $m = 1, 2, 5$. Results show that LG beams have self-healing property observationally, while only the LG beam with topological charge $m = 1$ has positive self-healing quantitatively, and others have negative self-healing quantitatively. Like table (2), Intensity profile of the LG beams are asymmetric, and the bright spots are centered on the number of topological charge m on the LG ring.

$m = 1$		$m = 2$		$m = 5$	



$z = 0\ m$	$z_{opt} = 6.4\ m$	$z = 0\ m$	$z_{opt} = 5.8\ m$	$z = 0\ m$	$z_{opt} = 5\ m$
92.75	92.94	93.75	95.54	95.03	94.31

Table 3. propagation of the LG beams via HCM with topological charge $m = 1, 2, 5$, 1000λ beam waist radius, circular mask of 707 pixels, and their self-healing values respect to original propagated beam

According to table (2)-(3) and noticed to the mask sizes, the LG beams with mask1 reach to their maximum self-healing values faster than the LG beams with mask2.

6. Conclusions

In this paper, are investigated the self-healing property of the LG beams with topological charge $m = 1, 2, 5$, quantitatively and observationally. Results show the LG beams have observational self-healing property. In all cases, self-healing values of the LG beams with small mask are positive self-healing value. Mostly, increasing topological charges caused to decreasing self-healing value difference. Also, bright spots are centered on the number of m on the LG ring.

References

[1] D. L. Andrews, ”Structured light and it’s applications: An Introduction to Phase-Structured Beams and Nanoscale Optical Forces” (2008).

[2] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman.”Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes”.Physical review A,. 45, 8185-8189 (1992).

[3] A. Wada, T. Ohtani, Y. Miyamoto, and M. Takeda, ”Propagation analysis of the Laguerre-Gaussian beam with astigmatism” J. Opt. Soc. Am. A/Vol. 22, No. 12/December 2005. Pp2746-2755.

[4] Z. Mei, and D. Zhao,”Nonparaxial analysis of vectorial Laguerre-Bessel-Gaussian beams”,2007, Vol. 15, No. 19,OPTICS EXPRESS.

[5] S. Huang, Z. Miao, C. He, F. Pang, Y. Li, and T. Wang,”Composite vortex beams by coaxial superposition of Laguerre-Gaussian beams” Opt. Lasers Eng. 78, 132 (2016).

[6] W. N. Plick and M. Krenn, ”Physical meaning of the radial index of Laguerre-Gauss beams” Phys. Rev. A 92, 063841 (2015).

[7] A. B. Stilgoe, T. A. Nieminen, and H. Rubinsztein-Dunlop, J.”Energy, momentum and propagation of non-paraxial high-order Gaussian beams in the presence of an aperture” Opt. 17,125601 (2015).

[8] X. Chu, G. Zhou, and R. Chen ”Analytical study of the self-healing property of Airy beams” PHYSICAL REVIEW A 85, 013815 (2012).

[9] P. Vaity and R. P. Singh, "Self-healing property of optical ring lattice" OPTICS LETTERS , Vol. 36, No. 15, 2011.

[10] L. Zhang, F. Ye, M. Cao, D. Wei, P. Zhang, H. Gao, and F. Li, ”Investigating the self-healing property of an optical Airy beam” Vol. 40, No. 21, 2015, Optics Letters.

[11] I. Litvin, L. Burger, and A. Forbes, ”Self-healing of Bessel-like beams with longitudinally dependent cone angles”

[12] A. A. Kovalev, V. V. Kotlyar, and A. P. Porfirev,”Asymmetric Laguerre-Gaussian beams” PHYSICAL REVIEW A 93, 063858 (2016).

[13] A. A. Kovalev, V. V. Kotlyar, and A. P. Porfirev, ”Optical trapping and moving of microparticles by using asymmetrical Laguerre-Gaussian beams” Vol. 41, No. 11, 2016, Optics Letters. pp 2426-2429.

[14] Saadati-Sharafeh. F., A. Borhanifar, A.P. Porfirev, P. Amiri, E.A. Akhlaghi, S.N. Khonina, and Y. Azizian-Kalandaragh, 2020. The superposition of the Bessel and mirrored Bessel beams and investigation of their self-healing characteristic. *International Journal for Light and Electron Optics (Optik)*.

[15] Kim. M. K., 2011. Digital Holographic Microscopy: Principles, Techniques, and Applications. Springer Series in Optical Sciences.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–233, pp. 203–209

Generalized log orthogonal functions for solving a class of cordial Volterra integral equations

S. Tayebinejad[1,*], F. Mohammadi[a]

[a]*Department of Mathematics, University of Hormozgan, Bandarabbas, P. O. Box 3995, Iran.*

Article Info

Keywords:
Cordial Volterra integral equation
Mittage−Leffler function
Generalized log orthogonal functions

2020 MSC:
45D05
42C05

Abstract

This paper deals with numerical solution of a class cordial Volterra integral equations with Mittag−Leffler solution. A numerical approach based the generalized log orthogonal functions is proposed to solve this kind of Volterra integral equations. By using the generalized log orthogonal functions as basis function, the presented numerical method can effectively approximate the solution of problems with singular behavior. The error estimate with respect to L^2-norm is investigated. Finally, the accuracy of the method is illustrated through a numerical example.

1. Introduction

Cordial Volterra integral operators are a special class of Volterra integral operators with weak singular kernels that appear in the study of heat conduction problems with mixed boundary conditions and some Volterra integral operators with certain kernel singularities [1, 2]. Such operators have the form

$$(\mathcal{V}_\varphi u)(t) = \int_0^t t^{-1}\varphi(t^{-1}s)k(t,s)u(s)ds, \qquad t \in I := [0,T], \tag{1}$$

which is inspired by Vainikko's studies [3, 4]. The function $\varphi \in L^1(0,1)$ is the core of the operator, and $k \in C^m(D)$ for some $m \geq 0$ where $D = \{(t,s) : 0 \leq s \leq t \leq T\}$. Cordial Volterra integral operators and the associated Volterra integral equations have been studied by Vainikko [5, 6] and several other authors [7–9]. This work is concerned with numerical solution for the second kind cordial Volterra integral equation (CVIE) of the form

$$u(t) = f(t) + a(\mathcal{V}_\varphi u)(t), \tag{2}$$

whose solution can be expressed in terms of the Mittag−Leffler function defined by $E_d(z) := \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(1+nd)}$, $z \in \mathbb{C}, d > 0$, in which Γ denotes the gamma function. The function $f \in C^m(I)$, a stands for an arbitrary constant, $\varphi(t^{-1}s) = \frac{t^b(1-t^{-1}s)^{b-1}}{\Gamma(b)}$, $0 < b < 1$, and without loss of generality we assume $k(t,s) = 1$.

*Talker
Email addresses: s.tayebi65@yahoo.com (S. Tayebinejad), f.mohammadi62@gmail.com (F. Mohammadi)

It is of interest to know CVIEs have singular behavior at the initial point $t = 0$. In general, facing singular problems, in order to develop accurate spectral methods, there are some strategies such as employing a local adaptive procedure in finite differences/finite elements [10], singular functions method [11], the enriched spectral methods [12, 13], and mapped spectral methods [14, 15]. In [16] authors suggested that mapped spectral methods on the non-uniformly Sobolev weighted spaces are more suitable for equations with singular behaviors. Actually, these methods lead to better convergence results than numerical methods for instance finite-element, finite-difference and spectral methods on usual Sobolev spaces. The choice of a log mapping to generalized Laguerre polynomials seemed to be the best adapted to their theory. Thus, they introduced two new classes of orthogonal functions on the non-uniformly Sobolev weighted spaces, log orthogonal functions (LOFs) and generalized log orthogonal functions (GLOFs). Now, in order to solve numerically (2), we apply the spectral collocation method using the GLOFs as basis functions.
The layout of this paper is as follows: Section 2, presents definitions and some properties of the LOFs and their generalized type, approximation by the GLOFs along with operational matrices, and applies the well-known spectral collocation method for solving (2). The error estimation of the approximate solution will be studied in section 3. In Section 4, a numerical example is given to clarify the effectiveness of the proposed method. Finally, in the last section, we present our conclusion.

2. Generalized log orthogonal functions

In this section, our concepts agree with those given in [16].

Indeed, it is the log mapping

$$x(t) := -(\beta+1)\log(t), \quad \beta > -1, \qquad t \in (0,1),$$

to generalized Laguerre polynomials $\mathbf{L}_n^{(\alpha)}(x)$, $\alpha > -1$, that makes the definition of the LOFs possible as follows.

Definition 2.1. For $\alpha, \beta > -1$ the LOFs are defined by

$$\mathcal{S}_n^{(\alpha,\beta)}(t) := \mathbf{L}_n^{(\alpha)}(x(t)) = \mathbf{L}_n^{(\alpha)}(-(\beta+1)\log(t)), \qquad n = 0, 1, \dots,$$

with satisfying the following properties:

- Three-term recurrence relation

$$\begin{aligned}
&\mathcal{S}_0^{(\alpha,\beta)}(t) = 1,\\
&\mathcal{S}_1^{(\alpha,\beta)}(t) = (\beta+1)\log(t) + \alpha + 1,\\
&\mathcal{S}_{n+1}^{(\alpha,\beta)}(t) = \frac{2n+\alpha+1+(\beta+1)\log(t)}{n+1}\mathcal{S}_n^{(\alpha,\beta)}(t) - \frac{n+\alpha}{n+1}\mathcal{S}_{n-1}^{(\alpha,\beta)}(t), \qquad n = 1, 2, \dots,
\end{aligned}$$

- Orthogonality

$$\int_0^1 \mathcal{S}_n^{(\alpha,\beta)}(t)\mathcal{S}_m^{(\alpha,\beta)}(t)(-\log(t))^{\alpha} t^{\beta} dt = \gamma_n^{(\alpha,\beta)}\delta_{nm}, \qquad \gamma_n^{(\alpha,\beta)} = \frac{\Gamma(n+\alpha+1)}{(\beta+1)^{\alpha+1}\Gamma(n+1)}. \tag{3}$$

Definition 2.2. For $\alpha, \beta > -1$, $\lambda \in \mathbb{R}$, the GLOFs are defined by

$$\mathcal{S}_n^{(\alpha,\beta,\lambda)}(t) := t^{(\beta-\lambda)/2}\mathcal{S}_n^{(\alpha,\beta)}(t), \quad \lambda \in \mathbb{R}, \quad n \geq 0,$$

with satisfying orthogonality condition

$$\int_0^1 \mathcal{S}_n^{(\alpha,\beta,\lambda)}(t)\mathcal{S}_m^{(\alpha,\beta,\lambda)}(t)(-\log(t))^{\alpha} t^{\lambda} dt = \gamma_n^{(\alpha,\beta)}\delta_{mn},$$

in which $\gamma_n^{(\alpha,\beta)}$ is already defined in (3). It is interesting to know that by choosing $\lambda = \beta$, the GLOFs are the same as the LOFs.

2.1. *Approximation by the GLOFs*

To obtain an approximation of any function $f \in L^2[0,1]$ in terms of the GLOFs, one can write

$$f(t) = \sum_{i=0}^{\infty} c_i \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t), \qquad f(t) \simeq f_n(t) = \sum_{i=0}^{n} c_i \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t) := C^T \Phi(t) = \Phi^T(t) C$$

where $c_i = \langle f(t), \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t) \rangle$ in which $\langle .,. \rangle$ denotes the inner product with respect to the weight function $\chi^{\alpha,\lambda}(t) := (-\log(t))^\alpha t^\lambda$, and $(n+1)$-order vectors C, $\Phi(t)$ are given by

$$C = [c_0, c_1, \dots, c_n]^T, \qquad \Phi(t) = [\mathcal{S}_0^{(\alpha,\beta,\lambda)}(t), \mathcal{S}_1^{(\alpha,\beta,\lambda)}(t), \dots, \mathcal{S}_n^{(\alpha,\beta,\lambda)}(t)]^T. \tag{4}$$

Similarly, approximation of a two-variable function $\mathcal{K}(t,s) \in L^2([0,1] \times [0,1])$ is as follows

$$\mathcal{K}(t,s) \simeq \mathcal{K}_n(t,s) = \sum_{i=0}^{n} \sum_{j=0}^{n} \mathcal{K}_{ij} \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t) \mathcal{S}_j^{(\alpha,\beta,\lambda)}(s) = \Phi^T(t) \mathcal{K} \Phi(s), \tag{5}$$

where $\mathcal{K}$ is an $(n+1) \times (n+1)$ matrix with coefficients $\mathcal{K}_{ij}$ are given by

$$\mathcal{K}_{ij} = \langle \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t), \langle \mathcal{K}(t,s), \mathcal{S}_j^{(\alpha,\beta,\lambda)}(s) \rangle_{\chi^{\alpha,\lambda}(s)} \rangle_{\chi^{\alpha,\lambda}(t)}, \qquad i,j = 0,1,\dots,n.$$

2.2. *Operational matrices*

we can approximate the integration of the vector $\Phi(t)$ defined in (4) as follows

$$\int_0^t \Phi(\tau) d\tau \simeq \mathcal{P} \Phi(t), \tag{6}$$

where $\mathcal{P}$ is the GLOFs operational matrix of integration of order $(n+1) \times (n+1)$ with coefficients $\mathcal{P}_{ij}$ are given by

$$\mathcal{P}_{ij} = \frac{\left\langle \int_0^t \mathcal{S}_i^{(\alpha,\beta,\lambda)}(\tau) d\tau, \mathcal{S}_j^{(\alpha,\beta,\lambda)}(t) \right\rangle}{\left\langle \mathcal{S}_j^{(\alpha,\beta,\lambda)}(t), \mathcal{S}_j^{(\alpha,\beta,\lambda)}(t) \right\rangle} \tag{7}$$

Furthermore, we have

$$\Phi(t) \Phi^T(t) C \simeq \tilde{C}^T \Phi(t), \tag{8}$$

where $\tilde{C}$ is the product operation matrix of two GLOFs vector whose entries are related to vector C.

2.3. *Solution of equation*

In order to solve (2) using the collocation method, we approximate the functions $u(t)$, $f(t)$, and $\mathcal{K}(t,s) := at^{-1}\varphi(t^{-1}s)k(t,s)$ by the GLOFs with coefficients determined by collocating (2) at the nodal points $\{t_i\}_{i=0}^n$, which are $(n+1)$ roots of Chebyshev polynomials $T_{n+1}(t)$ of degree $(n+1)$ on $[0,1]$. Assume

$$u(t) \simeq \sum_{i=0}^{n} c_i \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t) = C^T \Phi(t) = \Phi^T(t) C, \quad f(t) \simeq \sum_{i=0}^{n} f_i \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t) = F^T \Phi(t), \quad \mathcal{K}(t,s) \simeq \Phi^T(t) \mathcal{K} \Phi(s), \tag{9}$$

where C, $\mathcal{K}$ are defined in (4), (5), respectively, and $F = [f_0, f_1, \dots, f_n]^T$ is a known vector defined similarly to C. It is obtained by substituting (9) into (2)

$$\begin{aligned} C^T \Phi(t) &\simeq F^T \Phi(t) + \int_0^t \Phi^T(t) \mathcal{K} \Phi(s) \Phi^T(s) C ds = F^T \Phi(t) + \Phi^T(t) \mathcal{K} \tilde{C}^T \int_0^t \Phi(s) ds \\ &= F^T \Phi(t) + \Phi^T(t) \mathcal{K} \tilde{C}^T \mathcal{P} \Phi(t). \end{aligned}$$

It follows that

$$\left(C^T - F^T - \Phi^T(t)\mathcal{K}\tilde{C}^T\mathcal{P}\right)\Phi(t) \simeq 0. \tag{10}$$

Now, if we collocate (10) in $(n+1)$ points $\{t_i\}_{i=0}^n$ and replace $\simeq$ with $=$, we achieve

$$\left(C^T - F^T - \Phi^T(t_i)\mathcal{K}\tilde{C}^T\mathcal{P}\right)\Phi(t_i) = 0, \qquad i = 0, 1, \ldots, n, \tag{11}$$

The equation (11) produces a linear system of $(n+1)$ equations and $(n+1)$ unknowns that can be solved for the unknown vector C. Thus, the approximate solution of (2) will be obtained by $u(t) \simeq C^T\Phi(t)$.

3. Convergence

In this section, we present the approximation error by the GLOFs. The construction is due to Chen and Shen [16].

First of all, suppose that $u_n(t)$ is the approximate solution of (2). The error function will be

$$e_n(t) = u(t) - u_n(t) = \sum_{i=n+1}^{\infty} c_i \mathcal{S}_i^{(\alpha,\beta,\lambda)}(t).$$

Consider a pseudo−derivative with respect to the LOFs as follows

$$\hat{\partial}_t u := t\partial_t u.$$

Assume

$$A^k_{\alpha,\beta}(I) := \left\{\nu \in L^2_{\chi^{\alpha,\beta}}(I) : \hat{\partial}^j_t \nu \in L^2_{\chi^{\alpha+j,\beta}}(I), j = 1, 2, ..., k\right\}, \qquad k \in \mathbb{N},$$

is a non−uniformly weighted Sobolev space equipped with the semi−norm and norm

$$|\,\nu\,|_{A^m_{\alpha,\beta}} :=\|\,\hat{\partial}^m_t \nu\,\|_{\chi^{\alpha+m,\beta}}, \qquad \|\,\nu\,\|_{A^m_{\alpha,\beta}} := \left(\sum_{k=0}^m |\,\nu\,|^2_{A^k_{\alpha,\beta}}\right)^{1/2}.$$

The pseudo−derivative with respect to the GLOFs can be defined as

$$\hat{\partial}_{\gamma,t} u := t^{1+\gamma}\partial_t\left\{t^{-\gamma}u\right\}.$$

Furthermore, to better describe the approximability of $u_n(t)$ by the GLOFs, we need to define a non−uniformly weighted Sobolev space as

$$A^k_{\alpha,\beta,\lambda}(I) := \left\{\nu \in L^2_{\chi^{\alpha,\lambda}}(I) : \hat{\partial}^j_{\frac{\beta-\lambda}{2},t}\nu \in L^2_{\chi^{\alpha+j,\lambda}}(I), j = 1, 2, ..., k\right\}, \qquad k \in \mathbb{N},$$

equipped with semi−norm and norm as

$$|\,\nu\,|_{A^m_{\alpha,\beta,\lambda}} :=\|\,\hat{\partial}^m_{\frac{\beta-\lambda}{2},t}\nu\,\|_{\chi^{\alpha+m,\lambda}}, \qquad \|\,\nu\,\|_{A^m_{\alpha,\beta,\lambda}} := \left(\sum_{k=0}^m |\,\nu\,|^2_{A^k_{\alpha,\beta,\lambda}}\right)^{1/2}.$$

Theorem 3.1. *[16] Given $f(t) = t^r(-\log(t))^k, r \geq 0, k \in \mathbb{N}_0$. Let $\lambda > -1-2r, \alpha, \beta > -1$ and $\beta > \lambda$. Then, we have*

$$f \in L^2_{\chi^{\alpha,\lambda}} \text{ and } \mathbf{R}_{r,\beta,\lambda} = \left|\frac{2r+\lambda-\beta}{2r+2+\lambda+\beta}\right| < 1,$$

and

$$\parallel f - f_n \parallel_{\chi^{\alpha,\lambda}} \leq c(k+1)! n^{\frac{\alpha+1}{2}+k} (\mathbf{R}_{r,\beta,\lambda})^n \ when\ n > -\frac{2k+\alpha+2}{2log(\mathbf{R}_{r,\beta,\lambda})},$$

where

$$c \approx \sqrt{\frac{2^{\alpha+1+k}(\beta+1)^{2\alpha+2-k}}{(\beta+\lambda+2r+2)^{\alpha+1+k}}}.$$

In particular, if $\alpha = \lambda = 0$, then an accurate estimate for the GLOFs to singular functions in L^2-norm is obtained as

$$\parallel f - f_n \parallel \leq \sqrt{2}^k (\beta+1)^{-k} k! n^k \sqrt{2(\beta+1)n} \left| \frac{2r-\beta}{2r+\beta+2} \right|^{n-k}.$$

Also, for $f(t) = t^r$, $r \geq 0$, we have

$$\parallel f - f_n \parallel \leq \sqrt{2(\beta+1)n} \left| \frac{2r-\beta}{2r+\beta+2} \right|^{n}.$$

Theorem 3.2. *[16] Let $m, n, k \in \mathbb{N}, \lambda \in \mathbb{R}$ and $\alpha, \beta > -1$. For any $u \in A^m_{\alpha,\beta,\lambda}(I)$ and $0 \leq k \leq \tilde{m} = min\{m, n+1\}$, we have*

$$\parallel \hat{\partial}^k_{\frac{\beta-\lambda}{2},t}(u - u_n) \parallel_{\chi^{\alpha+k,\lambda}} \leq \sqrt{(\beta+1)^{k-\tilde{m}} \frac{(n-\tilde{m}+1)!}{(n-k+1)!}} \parallel \hat{\partial}^{\tilde{m}}_{\frac{\beta-\lambda}{2},t} u \parallel_{\chi^{\alpha+\tilde{m},\lambda}},$$

In particular, in the case of $\alpha = \beta = \lambda = k = 0$ and $m < n+1$, it holds that

$$\parallel u - u_n \parallel \leq cn^{-m/2} \parallel \hat{\partial}^m_t u \parallel_{\chi^m},$$

where $\chi^m = \chi^{m,0} = (-logt)^m$.

4. Numerical examples

In this section, we implement the collocation method given in Subsection 2.3 numerically for

$$u(t) = 1 - \sqrt{\pi} \int_0^t t^{-1} \varphi(t^{-1}s) u(s) ds, \qquad 0 \leq t \leq 1, \tag{12}$$

which has the exact solution $u(t) = E_{1/2}(\sqrt{\pi t})$. Here $\varphi(t^{-1}s) = \frac{t^{1/2}(1-t^{-1}s)^{-1/2}}{\Gamma(1/2)}$. Table 1 and Figure 1 illustrate the asymptotics of the proposed method numerically for this example. Table 1 exhibits the errors obtained by using the GLOFs with $\alpha = 0, \beta = 1, \lambda = -1$. It can be seen that as the number of the GLOFs increases the accuracy of the solution will reasonably improve. In Figure 1, we have shown the graphic representation of the exact and approximate solution of (12) for $n = 6$ with $\alpha = 0$, $\beta = 1$, $\lambda = -1$.

5. Conclusion

The log orthogonal functions and their generalized type were introduced. The distinctive feature of these functions is that they are very useful in resolving singularities. These functions were used to numerically solve equation (2). An illustrative example is presented to assess the effectiveness of the method.

Table 1. The absolute errors with $\alpha = 0$, $\beta = 1$, $\lambda = -1$ in (12).

t	$n = 2$	$n = 4$	$n = 6$
0	0	0	0
0.125	$9.507E-2$	$1.921E-3$	$6.181E-6$
0.250	$9.409E-2$	$9.815E-4$	$3.570E-6$
0.375	$3.022E-2$	$5.985E-4$	$5.473E-6$
0.500	$3.158E-2$	$9.382E-5$	$2.173E-6$
0.625	$6.673E-2$	$2.710E-4$	$1.425E-6$
0.750	$6.411E-2$	$2.620E-5$	$1.797E-6$
0.875	$1.839E-2$	$1.970E-4$	$1.320E-6$
1.000	$7.298E-2$	$1.648E-4$	$1.292E-6$

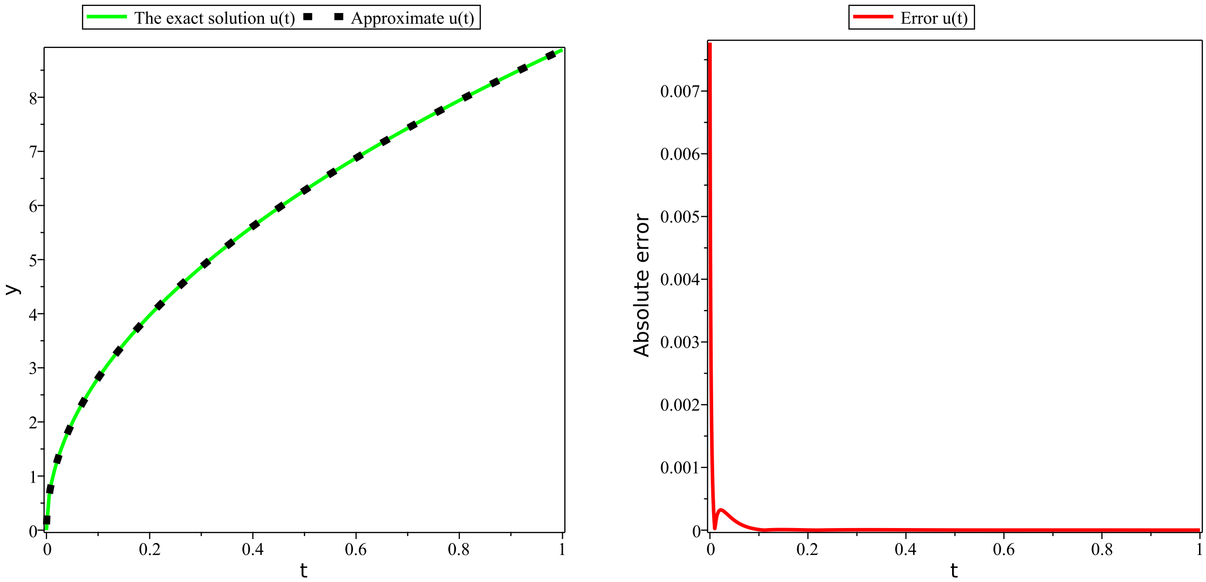


Fig. 1. The approximate and the exact solution (left) and the absolute error (right) for $n = 6$ with $\alpha = 0$, $\beta = 1$, $\lambda = -1$ in (12).

References

[1] Brunner, H. (2017). Volterra integral equations: an introduction to theory and applications (Vol. 30). Cambridge University Press.

[2] Majidian, H. (2017). Efficient quadrature rules for a class of cordial Volterra integral equations: A comparative study. Bulletin of the Iranian Mathematical Society, 43(5), 1245−1258.

[3] Vainikko, G. (2009). Cordial Volterra integral equations 1. Numerical Functional Analysis and Optimization, 30(9−10), 1145−1172.

[4] Vainikko, G. (2010). Cordial Volterra integral equations 2. Numerical functional analysis and optimization, 31(2), 191−219.

[5] Vainikko, G. (2010). Spline collocation for cordial Volterra integral equations. Numerical functional analysis and optimization, 31(3), 313−338.

[6] Vainikko, G. (2010). Spline collocation-interpolation method for linear and nonlinear cordial Volterra integral equations. Numerical functional analysis and optimization, 32(1), 83−109.

[7] Yang, Z. W. (2015). Second-kind linear Volterra integral equations with noncompact operators. Numerical Functional Analysis and Optimization, 36(1), 104−131.

[8] Darbenas, Z., & Oliver, M. (2019). Uniqueness of solutions for weakly degenerate cordial Volterra integral equations. Journal of Integral Equations and Applications, 31(3), 307−327.

[9] Diogo, T., & Vainikko, G. (2013). Applicability of spline collocation to cordial Volterra equations. Mathematical Modelling and Analysis, 18(1), 1−21.

[10] Morin, P., Nochetto, R. H., & Siebert, K. G. (2002). Convergence of adaptive finite element methods. SIAM review, 44(4), 631−658.

[11] Strang, G., & Fix, G. J. (1973). An analysis of the finite element method. Prentice-Hall Series in Automatic Computation, Englewood Cliffs, N. J: Prentice-Hall, Inc.

[12] Chen, S., & Shen, J. (2018). Enriched spectral methods and applications to problems with weakly singular solutions. Journal of Scientific Computing, 77, 1468−1489.

[13] Chen, S., & Shen, J. (2020). An efficient and accurate numerical method for the spectral fractional Laplacian equation. Journal of Scientific Computing, 82(1), 1−25.

[14] Wang, L. L., & Shen, J. (2005). Error analysis for mapped Jacobi spectral methods. Journal of Scientific Computing, 24, 183−218.

[15] Shen, J., & Wang, Y. (2016). Müntz−Galerkin Methods and Applications to Mixed Dirichlet−Neumann Boundary Value Problems. SIAM Journal on Scientific Computing, 38(4), A2357−A2381.

[16] Chen, S., & Shen, J. (2022). Log orthogonal functions: approximation properties and applications. IMA Journal of Numerical Analysis, 42(1), 712−743.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–235, pp. 210–219

Numerical investigation of time varying fractional optimal control problems by Fibonacci polynomials

Mohsen Alipour[a,*]

[a]*Department of Mathematics, Faculty of Basic Science, Babol Noshirvani University of Technology, Shariati Ave., Babol, 47148-71167, Iran*

Article Info

Keywords:
Time varying fractional optimal control problems
Fibonacci polynomials
operational matrix
Caputo derivative

2020 MSC:
65-XX
49-XX

Abstract

In this paper, we present a method for solving time varying fractional optimal control problems by Fibonacci polynomials. Firstly, we derive the Fibonacci polynomials (FPs) operational matrix for the fractional derivative in the Caputo sense, which has not been undertaken before. This method reduces the problems to a system of algebraic equations. The results obtained are in good agreement with the existing ones in open literatures and the solutions approach to classical solutions as the order of the fractional derivatives approach to 1.

1. **Introduction**

Although the optimal control theory is an area in mathematics which has been under development for years but the fractional optimal control theory is a very new area in mathematics. Recent contributions in this field were reported by several authors [1-4].
In this paper, we consider the time varying fractional optimal control problem as follows:

$$Minimize \quad J\left(x(t), u(t)\right) = \frac{1}{2}\int_0^1 x^2(t) + u^2(t)\,dt\,, \; (1)$$

subject to the dynamic constrain

$$D^{\alpha}x(t) = a_1(t)\,x(t) + a_2(t)\,u(t), \quad 0 < t \leq 1\;, \;\; 0 < \alpha \leq 1\,, \; (2)$$

and the initial condition

$$x(0) = x_0\,, \; (3)$$

where $x(t)$ and $u(t)$ are the state function and the control function, respectively. When $\alpha = 1$, the above problem reduces to a standard optimal control problem.

---

*Talker
Email address: m.alipour2323@gmail.com; m.alipour@nit.ac.ir (Mohsen Alipour)

The rest of this paper is as follows. In Section 2, we present some preliminaries in fractional calculus. In Section 3, FBs are introduced and then we approximate functions by using FBs and we show the properties of FBs by several Lemmas and corollaries. We make a new operational matrix for fractional derivative by FBs in Section 4. In Section 5, we apply FBs for solving time varying fractional optimal control problems. In Section 6, numerical examples are simulated to demonstrate the high performance of the proposed method. Finally, Section 7 concludes our work in this paper.

2. Some preliminaries in fractional calculus

In this section, we give some basic definitions and properties of the fractional calculus which are used further in this paper.

Definition 2.1. (See [5]) We define
$C_\mu = \{f(t)\,|\,f(t) > 0 \ for \ t > 0$ and $f(t) = t^p f_1(t)$ where $p > \mu$ and $f_1(t) \in C\,[0\,,\,\infty)$, and $C_\mu^n = \left\{f(t)\,\middle|\,f^{(n)}(t) \in C_\mu\right\}$ where $n \in N\,,\ \mu \in R.$

Definition 2.2. (See [5]) The Riemann-Liouville fractional integral operator of order $\alpha \geq 0$, of a function $f \in C_\mu\,,\ \mu \geq -1$, is defined as

$$\begin{array}{ll} I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-x)^{\alpha-1} f(x)\,dx\,, \quad \alpha > 0,\ t > 0, \\ I^0 f(t) = f(t), \end{array} \quad (4)$$

and for $n-1 < \alpha \leq n\,,\ n \in \mathrm{N}\,,\ t > 0\,,\ f \in C_{-1}^n$, the fractional derivative of $f(t)$ in the Caputo sense is defined as

$$D^\alpha f(t) = I^{n-\alpha} D^n f(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-x)^{n-\alpha-1} f^{(n)}(x)\,dx\,. \,(5)$$

Property 2.3. (See [6-8]) For $f \in C_\mu\,,\ \mu \geq -1\,,\ \alpha\,,\ \beta \geq 0$ we have

$$I^\alpha t^\gamma = \frac{\Gamma(\gamma+1)}{\Gamma(\alpha+\gamma+1)}\, t^{\alpha+\gamma},\ (6)$$

and for $n-1 < \alpha \leq n\,,\ n \in \mathrm{N}$ and $f \in C_\mu^n\,,\ \mu \geq -1$ we see the following properties

$$1.\ D^\alpha I^\alpha f(t) = f(t), \quad (7)$$

$$2.\ I^\alpha D^\alpha f(t) = f(t) - \sum_{k=0}^{n-1} f^{(k)}(0^+)\,\frac{x^k}{k!}\,, \quad t > 0\,,\ (8)$$

$$3.\ D^\beta f(t) = I^{\alpha-\beta} D^\alpha f(t)\,.\ (9)$$

3. Fibonacci polynomials and function approximation

Definition 3.1. For any positive real number k, the k-Fibonacci sequence $\{F_{k,n}\}_{n\in N}$ is defined recurrently by

$$F_{k,n+1} = kF_{k,n} + F_{k,n-1}, \qquad n \geq 1, (10)$$

with initial conditions

$$F_{k,0} = 0, \qquad F_{k,1} = 1.$$

Particular cases of the k-Fibonacci sequence are constructed from the following relations.
If k=1 the classical Fibonacci sequence is obtained.

$$F_0 = 0, \quad F_1 = 1, \quad F_{n+1} = F_n + F_{n-1}, \quad n = 1,$$

If k=2, the Pell sequence appears:

$$P_0 = 0, \quad P_1 = 1, \quad P_{n+1} = 2P_n + P_{n-1}, \quad n = 1,$$

If k=3, the following sequence appears:

$$H_0 = 0, \quad H_1 = 1, \quad H_{n+1} = 3H_n + H_{n-1}, \quad n = 1.$$

If k be a real variable x then $F_{k,n} = F_{x,n}$ and they correspond to the Fibonacci polynomials defined by

$$F_{n+1}(x) = \begin{cases} 1, & n = 0, \\ x, & n = 1, \\ xF_n(x) + F_{n-1}(x), & n > 1, \end{cases} \quad (11)$$

and from these expressions, as for the k-Fibonacci numbers we can write:

$$F_{n+1}(x) = \sum_{i=0}^{\lfloor \frac{n}{2} \rfloor} \begin{pmatrix} n-i \\ i \end{pmatrix} x^{n-2i}, \qquad n = 0,(12)$$

where $\lfloor \frac{n}{2} \rfloor$ denotes the greatest integer in $\frac{n}{2}$.
Note that $F_{2n}(0) = 0$ and $x = 0$ is the only real root, while $F_{2n+1}(0) = 1$ with no real roots. Also for $x = k \in N$ we obtain the elements of the k-Fibonacci sequences.
The Fibonacci polynomials are normalized so that $F_n(1) = F_n$, where the F_n is nth Fibonacci number.
Note first, that the equations for the Fibonacci polynomials may be written in matrix form as

$$F(x) = AT(x), (13)$$

where $F(x) = [F_1(x), F_2(x), F_3(x), \ldots, F_N(x)]^T$, $T(x) = \left[1, x, x^2, x^3, \ldots, x^{N-1}\right]^T$, and A is the lower triangular matrix with entrances the coefficients appearing in the expansion of the Fibonacci polynomials in increasing powers of x. if N is odd,

$$A = \begin{pmatrix} 1 & 0 & & & \cdots & 0 & & \\ 0 & 1 & 0 & & \cdots & & & \\ 1 & 0 & 1 & 0 & \cdots & & & \\ 0 & 2 & 0 & 1 & 0 & \cdots & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & & & & & 0 \\ 0 & \frac{\left(\frac{N+1}{2}\right)!}{\left(\frac{N-1}{2}\right)!} & 0 & \ldots & 0 & N-1 & 0 & 1 \end{pmatrix}_{(N+1)\times(N+1)}$$

If N is even

$$A = \begin{pmatrix} 1 & 0 & & & \cdots & & & 0 \\ 0 & 1 & 0 & & \cdots & & & \\ 1 & 0 & 1 & 0 & \cdots & & & \\ 0 & 2 & 0 & 1 & 0 & \cdots & & \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ & & & & & & & 0 \\ 0 & \frac{\left(\frac{N+2}{2}\right)!}{\left(\frac{N-2}{2}\right)!} & 0 & \ldots & 0 & N-1 & 0 & 1 \end{pmatrix}_{(N+1)\times(N+1)}$$

Note that in matrix A the non-zero entrances build precisely the diagonals of the Pascal triangle and the sum of the elements in the same row gives the classical Fibonacci sequence. In addition, matrix A is invertible and therefore x^n may be written as a linear combination of Fibonacci polynomials that is given in closed form in the following theorem, which is the version of the Zeckendorfs theorem for the Fibonacci polynomials.

Corollary 3.2. For every integer $n = 1$, x^{n-1} may be written in a unique way as linear combination of the n first Fibonacci polynomials as

$$x^{n-1} = \sum_{n=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^i \left[\begin{pmatrix} n \\ i \end{pmatrix} - \begin{pmatrix} n \\ i-1 \end{pmatrix} \right] F_{n-2i}(x), (14)$$

where $\left(\begin{array}{c} n \\ -1 \end{array} \right) = 0$.

Lemma 3.3. (see [9]) Let $L^2\left[0,1\right]$ be a Hilbert space with the inner product $\langle f,g\rangle = \int_0^1 f\left(x\right)g\left(x\right)dx \; and \; y \in L^2\left[0,1\right]$. Then, we can find the unique vector $C = \left[c_1, c_2, \ldots, c_N\right]^T$ such that

$$y\left(x\right) \simeq \sum_{n=1}^{N} C_n\, F_n\left(x\right) = C^T F\left(x\right), (15)$$

where

$$C^T = \left[c_1, c_2, \ldots, c_N\right], F\left(x\right) = \left[F_1\left(x\right), F_2\left(x\right), \ldots, F_N\left(x\right)\right]^T.$$

Corollary 3.4. In lemma 3.3 we have $C^T = \langle y, F\rangle\, Q^{-1}$ such that

$$\langle y, F\rangle = \int_0^1 y\left(x\right)F^T\left(x\right)dx = \left[\langle y, F_1\rangle,\ \langle y, F_2\rangle, \ldots, \langle y, F_N\rangle\right] \; and$$

$$Q_{ij} = \int_0^1 F_i\left(x\right)F_j\left(x\right)dx = \int_0^1 \sum_{p=0}^{\left[\frac{i-1}{2}\right]} \left(\begin{array}{c} i-p-1 \\ p \end{array} \right) x^{i-2p-1} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \left(\begin{array}{c} j-r-1 \\ j \end{array} \right) x^{j-2r-1} dx$$

$$= \sum_{p=0}^{\left[\frac{i-1}{2}\right]} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \left(\begin{array}{c} i-p-1 \\ p \end{array} \right) \left(\begin{array}{c} j-r-1 \\ j \end{array} \right) \int_0^1 x^{i+j-2p-2r-2} dx$$

$$\sum_{p=0}^{\left[\frac{i-1}{2}\right]} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \left(\begin{array}{c} i-p-1 \\ p \end{array} \right) \left(\begin{array}{c} j-r-1 \\ j \end{array} \right) \frac{1}{i+j-2p-2r-1},\ \ i,j = 1, \ldots,\ N.$$

Lemma 3.5. Suppose that $C_{\mathrm{N}\times 1}$ is an arbitrary vector. The operational matrix of product $\hat{C}_{N\times N}$ using FPs can be given as follows:

$$C^T F\left(x\right)\ F(x)^T \cong F(x)^T \hat{C}. (16)$$

Proof. Since $F\left(x\right) = A\,T\left(x\right)$ we have

$$C^T F\left(x\right)F^T\left(x\right) = C^T F\left(x\right)T^T\left(x\right)A^T = \left[C^T F\left(x\right), xC^T F\left(x\right), \ldots, x^{N-1}C^T F\left(x\right)\right]A^T$$

$$= \left[\sum_{i=1}^{N} c_i F_i\left(x\right), \sum_{i=1}^{N} c_i x F_i\left(x\right), \ldots, \sum_{i=1}^{N} c_i x^{N-1} F_i\left(x\right)\right] A^T$$

Now, we approximate all functions $x^k F_i\left(x\right)$ in terms of $F\left(x\right)$ as

$$x^k F_i\left(x\right) \cong e_{k,i}^T\ F\left(x\right), \quad k = 0, \ldots, N-1 \; and \; i = 1, \ldots, N$$

where $e_{k,i} = \left[e_{k,i}^1, e_{k,i}^2, \ldots, e_{k,i}^N\right]^T$ and

$$e_{k,i} = Q^{-1}\int_0^1 x^k F_i\ F\left(x\right)dx$$

$$= Q^{-1}\left[\int_0^1 x^k F_I\left(x\right)\ F_1\left(x\right)dx,\ \int_0^1 x^k F_i\left(x\right)F_2\left(x\right)dx\ , \ldots,\ \int_0^1 x^k\, F_i\left(x\right)F_N\left(x\right)dx\right]^T.$$

So we get

$$\int_0^1 x^k F_i\left(x\right)F_j\left(x\right)dx =$$

$$\int_0^1 x^k \sum_{p=0}^{\left[\frac{j-1}{2}\right]} \begin{pmatrix} i-p-1 \\ p \end{pmatrix} x^{i-2p-1} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \begin{pmatrix} j-r-1 \\ j \end{pmatrix} x^{j-2r-1} dx$$

$$\sum_{p=0}^{\left[\frac{i-1}{2}\right]} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \begin{pmatrix} i-p-1 \\ p \end{pmatrix} \begin{pmatrix} j-r-1 \\ j \end{pmatrix} \int_0^1 x^{k+i+j-2p-2r-2} dx =$$

$$\sum_{p=0}^{\left[\frac{i-1}{2}\right]} \sum_{r=0}^{\left[\frac{j-1}{2}\right]} \begin{pmatrix} i-p-1 \\ p \end{pmatrix} \begin{pmatrix} j-r-1 \\ j \end{pmatrix} * \frac{1}{k+i+j-2p-2r-1}.$$

Then we have

$$\sum_{i=1}^{N} c_i x^k F_i(x) = \sum_{i=1}^{N} c_i \sum_{j=1}^{N} e_{k,i}^j . F_j(x) = \sum_{j=1}^{N} F_j(x) \left(\sum_{j=1}^{N} c_i e_{k,i}^j \right) = F^T(x)\ v_k\ C,$$

where $v_k\ (k=\ 1,\ \ldots,\ N)$ is a $N \times N$ matrix that has vectors $e_{k,i}\ (i=1,\ldots,N)$ for each column. If we define $\overline{C} = [v_1 C, v_2 C, \ldots, v_N C]$, then we obtain

$$C^T F(x)\ F(x)^T \cong F(x)^T \overline{C} A^T,$$

and therefore we get the operational matrix of product $\hat{C} = \overline{C} A^T$.

4. FPs operational matrix for fractional derivative

In this section, we obtain the operational matrix for the fractional derivative. We can write

$$D^\alpha F(t) = \frac{1}{(n-\alpha)} \int_0^t (t-x)^{n-\alpha-1} F^{(n)}(x)$$

$$= \frac{1}{(n-\alpha)} \int_0^t (t-x)^{n-\alpha-1} A\, T^{(n)}(x)\, dx$$

$$= A\left[D^\alpha(1),\ D^\alpha t,\ \ldots,\ D^\alpha t^{N-1}\right] (17)$$

where

$$D^\alpha t^j = \begin{cases} 0 & j=0,\ \ldots,\ \lceil\alpha\rceil - 1, \\ \frac{\Gamma(j+1)}{\Gamma(j+1-\alpha)}\, t^{j-\alpha} & j = \lceil\alpha\rceil,\ \ldots,\ N-1. \end{cases} .(18)$$

Therefore we have

$$D^\alpha T(t) = \tilde{\Sigma}\tilde{T}(t),\ (19)$$

where $\tilde{\Sigma}$ and $\tilde{T}$ are a $N \times N$ diagonal matrix and a $N \times 1$ matrix, respectively as follows:

$$\tilde{\Sigma} = \left(\tilde{\Sigma}_{i,j}\right)_{i,j=1}^{m+1},\ \tilde{\Sigma}_{i+1,j+1} = \begin{cases} \frac{\Gamma(j+1)}{\Gamma(j+1-\alpha)} & i,j = \lceil\alpha\rceil,\ \ldots,\ N-1\ and\, i=j, \\ 0 & otherwise, \end{cases} \quad (20)$$

and

$$\tilde{T} = \left[t^{-\alpha},\ t^{1-\alpha},\ ...,\ t^{N-1-\alpha}\right]^T . (21)$$

Now, we approximate $t^{i-\alpha}\ (i = \lceil\alpha\rceil,\ \ldots,\ N-1)$ with respect to FBs by using (15). Therefore, we can write

$$t^{i-\alpha} \approx P_i^T F(t),\ (22)$$

where $P_i,\ (i = \lceil\alpha\rceil,\ \ldots,\ N-1)$ is a vector $N \times 1$. So, we hav

$$P_i = Q^{-1}\left(\int_0^1 t^{i-\alpha} F(t)\, dt\right) = Q^{-1}\left[\int_0^1 t^{i-\alpha} F_1(t)\, dt\ ,\ \int_0^1 t^{i-\alpha} F_2(t)\, dt\, ,\ \ldots,\ \int_0^1 t^{i-\alpha} F_N(t)\, dt\right]^T = Q^{-1}\bar{P}_i,$$

where

$$\bar{P}_i = [\bar{P}_{i\,,1}\,,\,\bar{P}_{i\,,2}\,,\,\ldots\,,\,\bar{P}_{i\,,N}]^T, \ (23)$$

$$\overline{P}_{i,j} = \int_0^1 t^{i-\alpha} F_{j,m}\,(t)\,dt = \sum_{r=0}^{[\frac{j-1}{2}]} \binom{j-1-r}{r} \int_0^1 t^{i-\alpha+j-2r-1dt} = \sum_{r=0}^{[\frac{j-1}{2}]} \binom{j-1-r}{r} \frac{1}{i-\alpha+j-2r}\,,$$

$$i = \lceil \alpha \rceil\,, \ldots, N-1 \; and \; j = 1, \ldots, N. \quad (24)$$

Now, we suppose P is an $N \times N$ matrix that has vector zero in $\lceil \alpha \rceil$ first column and vector P_i in (i+1)th column's for $i = \lceil \alpha \rceil\ ,\ \ldots\ ,\ N-1$.
Finally, from $(17)-(24)$, we obtain

$$D^{\alpha} F(t) \approx D_{\alpha}\, F(t), \ (25)$$

where

$$D_{\alpha} \approx A\; \tilde{\Sigma}\, P^T\ , \ (26)$$

is called the Fibonacci polynomials operational matrix of fractional derivative.

6. FPs for solving time varying fractional optimal control problems

Using Lemma 3.3, we can approximate the known and unknown functions in (1) and (2) as follows:

$$x(t) \approx c^T F(t), \ (27)$$

$$u(t) \approx b^T F(t), \ (28)$$

$$a_1(t) \approx a_1^T F(t), \ (29)$$

$$a_2(t) \approx a_2^T F(t), \ (30)$$

where $c\,,\; b\; \in R^{N\times 1}$ are unknown vectors and $a_1\,,\; a_2 \in R^{N\times 1}$ are known vectors.
by (25) and (27) we can write

$$D^{\alpha} x(t) \approx c^T D_{\alpha} F(t).(31)$$

Therefore, the problem $(1)-(3)$ reduce to the following problem:

$$Minimize \quad \frac{1}{2}\int_0^1 c^T F(t)F(t)^T c + b^T F(t)F(t)^T b\, dt, (32)$$

subject to the dynamic constraint

$$c^T D_{\alpha} F(t) = a_1^T F(t)F(t)^T c + a_2^T F(t)F(t)^T b\,, (33)$$

and the initial condition

$$c^T F(0) = x_0.(34)$$

Now, using Corollary 3.4 for (32) we can write

$$\begin{aligned} Minimize \quad J\,(c,b) &= \tfrac{1}{2} c^T \left(\int_0^1 F(t)F(t)^T\, dt\right)\, c + \tfrac{1}{2}\, b^T \left(\int_0^1 F(t)F(t)^T dt\right)\, b \\ &= \tfrac{1}{2} c^T Q\, c + \tfrac{1}{2} b^T Q\, b \end{aligned}$$

Also by Lemma 3.5 for (33) we have

$$c^T D_{\alpha} F(t) = F(t)^T \hat{A}_1 c + F(t)^T \hat{A}_2 b\,.(36)$$

Now, by using tau method [10] we can generate algebraic equations from (36) as follows

$$G_j\,(c,b)=\int_0^1 \left(c^T D_\alpha - c^T \hat{A}_1^T - b^T \hat{A}_2^T\right)\,F(t)F_j(t)\,dt=0,\quad j=1,\,\dots\,,\,N\,,(37)$$

and from (34) we set $G_N=c^T F(0)-x_0$.
Finally, the problem $(1)-(3)$ has been reduced to a parameter optimization problem which can be stated as follows: Find $c$and b which

$$Minimize \qquad J\,(c,b)\;=\frac{1}{2}c^T Q\,c+\frac{1}{2}b^T Q\,b,\ (38)$$

subject to the system of algebraic equations

$$G_j\,(c,b)=0,\quad j=1,\dots,N.(39)$$

For solving the above problem we use the Lagrange multipliers method. So, we define Lagrange function for the problem (38) and (39) as follows:

$$L\,(c,b,\lambda)=\frac{1}{2}c^T Q\,c+\frac{1}{2}b^T Q\,b\,+\sum_{j=1}^{N}\lambda_j\,G_j\,(c,b)\ ,(40)$$

where $\lambda=[\lambda_0,\dots,\lambda_m]^T$is the unknown Lagrange multiplier. Now, we consider the necessary conditions for the extremum and obtain the following systems of algebraic equations

$$\frac{\partial L}{\partial c}=0,\qquad (41)$$

$$\frac{\partial L}{\partial b}=0,(42)$$

$$\frac{\partial L}{\partial \lambda}=0.(43)$$

Equations $(41)-(43)$ can be solved for $c,\ b$ and λ by Newton's iterative method. Then, we get the approximate value of the state functions $x(t)$and the control functions $u(t)$ from (27) and (28), respectively.

7. Numerical examples

To demonstrate the applicability and to validate the numerical scheme, we apply the present method for the following examples.

Example 1. Consider the following time invariant problem [1, 2]

$$Minimize \quad J=\frac{1}{2}\int_0^1 x^2(t)+u^2(t)\,dt\,,$$

subject to the system dynamics

$$D^\alpha x(t)=-x(t)+u(t)\,,$$

with initial condition $x(0)=1$. For this problem we have the exact solution in the case of α = 1 as follows

$$x(t)=\cosh(\sqrt{2}\,t)+\beta\,\sinh(\sqrt{2}\,t)\,,\quad u(t)=(1+\sqrt{2}\,\beta)\,\cosh(\sqrt{2}\,t)+(\sqrt{2}+\beta)\,\sinh(\sqrt{2}\,t)\,,$$

where $\beta=-\frac{\cosh(\sqrt{2})+\sqrt{2}\,\sinh(\sqrt{2})}{\sqrt{2}\,\cosh(\sqrt{2})+\sinh(\sqrt{2})}$.
Figures 1 and 2 show the state and the control variables, respectively, as a function of time for $N=5$, for different values of α. These figures show that as α approaches close to 1, the numerical solutions for both the state and the control variables approach to the analytical solutions for $\alpha=1$ as expected. In Figs. 3 and 4, we see the absolute error of obtained results for $N=5$ and $\alpha=1$. Also, in Table 1, the absolute error of $x(t)$ for when $\alpha=1$ and $N=5$ is demonstrated and is compared with [1].

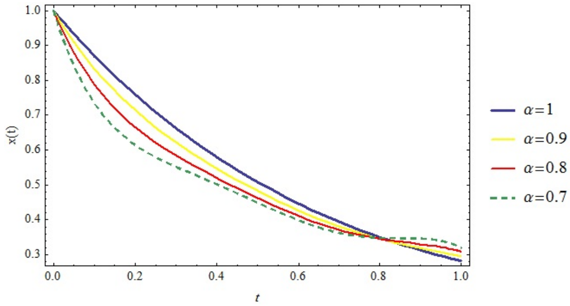


Fig. 1. Approximate solutions of $x(t)$ for $N = 5$ in example 1.

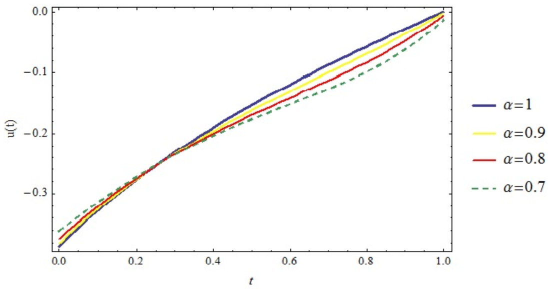


Fig. 2. Approximate solutions of $u(t)$ for $N = 5$ in example 1.

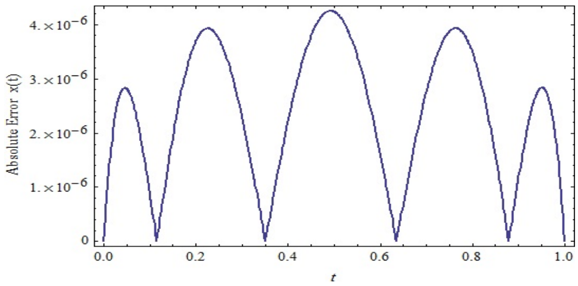


Fig. 3. Plot of absolute error function $x(t)$ for $\alpha = 1$ and $N = 5$ in example 1.

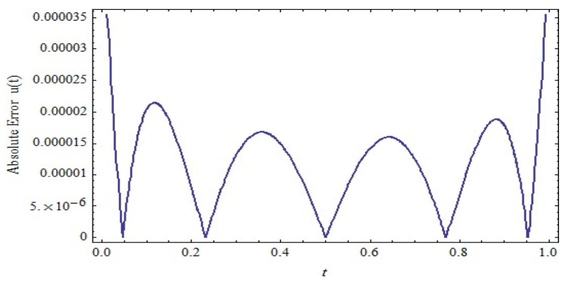


Fig. 4. Plot of absolute error function $u(t)$ for $\alpha = 1$ and $N = 5$ in example 1.

Example 2. This example considers a time varying fractional optimal control problem [1, 2]. Find the control $u(t)$

Table 1. Absolute error $x(t)$ for $\alpha = 1,\ N = 5$ and different values t in example 1.

t	[1] $N = 5$	Present Method $N = 5$
0	0.00000625	0
0.2	0.0000212	3.13415×10^{-6}
0.4	0.0000473	1.53012×10^{-6}
0.6	0.0000749	1.17532×10^{-7}
0.8	0.000107	3.26531×10^{-6}

which minimizes the performance index J given in Example 1 subject to the following dynamical system

$$D^{\alpha}x(t) = t\,x(t) + u(t)\,,$$

with initial condition $x(0) = 1$. Figs. 5 and 6 show the state $x(t)$ and the control $u(t)$ as functions of t for different values of α. These figures show that as α approaches close to 1, the numerical solutions for both the state and the control variables approach to the solutions for $\alpha = 1$ as expected. The numerical solution obtained with the proposed method for fractional orders of α matches those found in the literature.

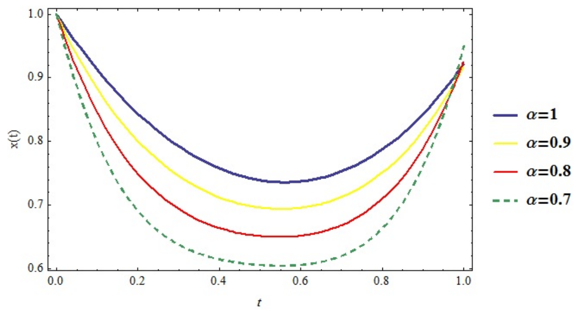


Fig. 5. Approximate solutions of $x(t)$ for $N = 5$ in example 2.

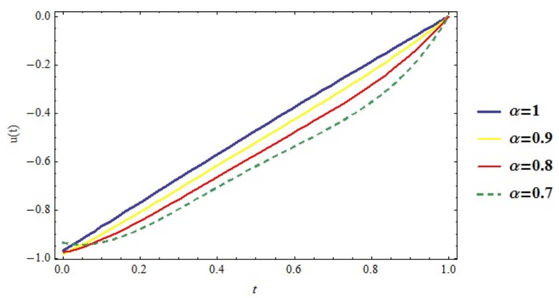


Fig. 6. Approximate solutions of $u(t)$ for $N = 5$ in example 2.

8. Conclusion

In this work, by Fibonacci polynomials we obtained operational matrices of the product and fractional derivative. Then we reduced the time varying fractional optimal control problem to a system of algebraic equations that can be solved easily. We saw that the obtained results in examples were in good agreement with the exact solution and approximate solution of other methods. Also, we observed that the solutions for the fractional optimal control problems approach to the solutions for standard optimal control problems as the order of the fractional derivative approaches to 1.

References

[1] A. Lotfi, M. Dehghan, S.A. Yousefi, A numerical technique for solving fractional optimal control problems, Computers and Mathematics with Applications 62 (2011) 1055-1067.

[2] O.P. Agrawal, A general formulation and solution scheme for fractional optimal control problems, Nonlinear Dynam. 38 (2004) 323-337.

[3] O.P. Agrawal, D. Baleanu, A Hamiltonian formulation and a direct numerical scheme for fractional optimal control problems, J. Vib. Control 13 (2007) 1269-1281.

[4] D. Baleanu, O. Defterli, O.P. Agrawal, A central difference numerical scheme for fractional optimal control problems, J. Vib. Control 15 (**??**) (2009) 583-597.

[5] M. Caputo, Linear models of dissipation whose Q is almost frequency independent II, Geophys. J. Roy. Astronom. Soc. 13 (1967) 529-539.

[6] Y. Luchko and R. Gorneflo, The initial value problem for some fractional differential equations with the Caputo derivative, Preprint series A08-98, Fachbereich Mathematik und Informatik, Freie Universitat Berlin, 1998.

[7] K. S. Miller, B. Ross, An Introduction to the Fractional Calculus and Fractional Differential Equations, John Wiley and Sons, Inc., New York, 1993.

[8] K.B. Oldham and J. Spanier , The Fractional Calculus, Academic Press, New York, 1974.

[9] E. Kreyszig, Introduction Functional Analysis with Applications, John Wiley and Sons Incorporated, 1978.

[10] C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang, Spectral Methods in Fluid Dynamic, Prentice-Hall, Englewood Cliffs, NJ, 1988.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–236, pp. 220–229

A new method based on Bernstein polynomials for numerical solution of two-dimensional fractional heat conduction equation

Mohsen Alipour[a,*]

[a]Department of Mathematics, Faculty of Basic Science, Babol Noshirvani University of Technology, Shariati Ave., Babol, 47148-71167, Iran

Article Info

Keywords:
Bernstein polynomials
fractional two-dimensional heat conduction equation
Operational matrices
Caputo fractional derivative
Riemann–Liouville fractional integral

2020 MSC:
65-XX
26A33
35K05

Abstract

In this paper, we develop a new scheme for numerical solutions of the fractional two dimensional heat conduction equation on a rectangular plane. Our main aim is to generalize the Bernstein operational matrices of derivatives and integrals to the three dimensional case. By the use of these operational matrices, we reduce the corresponding fractional order partial differential equations to a system of easily solvable algebraic equations. The results we obtain are compared with the exact solutions and we find that the error is negligible.

1. Introduction

The diffusion equation is of great importance in many engineering problems such as heat conduction, chemical diffusion, fluid flow, mass transfer, refrigeration and traffic analysis and so on. After the development of fractional derivatives it is found that most of these phenomena can well be explained by fractional order partial differential equations (FPDEs), see for example [1,2] and the references quoted therein. We consider the problem in generalized form as

$$\frac{\partial^\alpha u(t,x,y)}{\partial t^\alpha} = C_1 \frac{\partial^\beta u(t,x,y)}{\partial x^\beta} + C_2 \frac{\partial^\beta u(t,x,y)}{\partial y^\beta} + g\left(x,y,t\right), \quad (1)$$

subject to initial condition

$$u\left(0,\ x,\ y\right) = \ f\ \left(x,\ y\right), \quad (2)$$

where C_1 and C_2 are constants, $0 < \alpha \leq 1$, t, x, y $\in [0, 1]$. Conventionally various methods such as smoothed partial hydrodynamic method [3], meshless method [4,2], homotopy perturbation method [5,6], Tau method [7], method of local radial functions [8], Sinc–Legendre collocation method [9,10] are used for the solutions of such type of problems.

*Talker

Email address: m.alipour2323@gmail.com; m.alipour@nit.ac.ir (Mohsen Alipour)

Recently some approximate solutions for integer order heat conduction equations are obtained by Exp-function method [11], variational iteration method and energy balance method [12, 13]. These methods are very efficient and provide very good approximations to the solutions but due to high computational complexities these methods are not so easy to apply to fractional order partial differential equations in higher dimensions. Therefore, we need an easy and efficient method to solve such type of problems. More recently, the techniques based on operational matrices are extensively used for approximate solutions of a wide class of differential equations as well as partial differential equations, see for example, [14,15] and references quoted therein. The technique based on the operational matrices is simple and provides high accuracy but up to now this technique is used only to solve partial differential equations (PDEs) with only two variables. We generalize the technique to solve PDEs with three variables.
We use Bernstein polynomials and develop new matrices of fractional order differentiations and integrations to solve the corresponding fractional order partial differential equations without actually discretizing the problem. Our method reduces the FPDEs to a system of easily solvable algebraic equations of Sylvester type which can be easily solved by any computational software. Generally, large systems of algebraic equations may lead to greater computational complexity and large storage requirements. However our technique is simple and reduces the computational complexity of the resulting algebraic system. It is worthwhile to mention that, the method based on using the operational matrix of orthogonal functions for solving FPDEs is computer oriented. We use Mathematica software to perform necessary calculations.
The article is organized as follows. We begin by introducing some necessary definitions and mathematical preliminaries of the fractional calculus and Bernstein polynomials which are required for establishing main results. In Section 3, the Bernstein operational matrices of fractional derivatives and fractional integrals are obtained. Section 4 is devoted to the application of the Bernstein operational matrices of fractional derivatives and fractional integrals to solve the transient state two-dimensional fractional heat conduction equation on a rectangular plane. In section 5, the proposed method is applied to an example. Conclusion is made in Section 6.

2. Preliminaries

For convenience, this section summarizes some concepts, definitions and basic results from fractional calculus.

Definition 2.1. ([16–18]). Given an interva$l$ $[a, b] \subset R$ The Riemann–Liouville fractional order integral of a function $\emptyset \in \left(L^1(a,b), R\right)$ of order $\alpha \in R_+$is defined by

$$I^{\alpha}\emptyset(x) = \frac{1}{\Box(\alpha)}\int_a^x (x-s)^{\alpha-1}\emptyset(s)\,ds,$$

provided that the integral on the right hand side exists

Definition 2.2. ([16–18]) For a given function $\emptyset(x) \in c^n[a,b]$, the Caputo fractional order derivative is defined as

$$D^{\alpha}\emptyset(x) = \frac{1}{\Box(n-\alpha)}\int_a^x \frac{\emptyset^{(n)}(s)}{(x-s)^{\alpha+1-n}}ds, \qquad n-1 \le \alpha < n,\ n\epsilon N,$$

where $n = [\alpha] + 1$. Hence, it follows that

$$D^{\alpha}x^k = \frac{\Box(1+k)}{\Box(1+k-\alpha)}x^{k-\alpha}, \quad I^{\alpha}x^k = \frac{\Box(1+k)}{\Box(1+k+\alpha)}x^{k+\alpha}, \ \ D^{\alpha}c = 0 \ \text{ for a constant C.} \quad (3)$$

2.1. Properties of Bernstein polynomials

The well-known Bernstein polynomials of the nth degree are defined on the interval $[0, 1]$ as [19, 20]

$$b_i^n(x) = \binom{n}{i} x^i (1-x)^{n-i}, i = 0, \ldots, n \quad (4)$$

These Bernstein polynomials form a complete basis on over the interval $[0, 1]$. A recursive definition also can be used to generate these polynomials

$$b_i^n(x) = (1-x)\, b_i^{n-1}(x) + x b_{i-1}^{n-1}, i = 0, \ldots, n,$$

where $b_{-1}^{n-1}(x) = 0$ and $b_{n}^{n-1}(x) = 0$. Since the power basis $\{1, x, x^2, \ldots, x^n\}$, forms a basis for the space of polynomials of degree less than or equal to n, any Bernstein polynomial of degree n can be written in terms of the power basis. This can be directly calculated using the binomial expansion of $(1-x)^{n-i}$, one can show that

$$b_i^n(x) = \sum_{j=i}^{n} (-1)^{j-i} \binom{n}{i} \binom{n-i}{j-i} x^j, i = 0, \ldots, , n \quad (5)$$

On the other hand, the fact that they are not orthogonal turns out to be their disadvantage when used in the least-squares approximation. As said in [21] one approach to direct least-squares approximation by polynomials in Bernstein form relies on construction of the basis $\{d_0^n(x), d_1^n(x), \ldots, d_n^n(x)\}$ that is "dual" to the Bernstein basis of degree n on $x \in [0,1]$ .This dual basis is characterized by the property

$$\int_0^1 b_i^n(x)\, d_j^n(x)\, dx = \begin{cases} 1 \ for & i = j, \\ 0 \ for & i \neq j, \end{cases}$$

for $i, j = 0, 1, \ldots, n$. A function $f(x)$, square integrable in $[0,1]$ may be expressed in terms of the Bernstein basis in practice, only the first $(n+1)$ term Bernstein polynomials are considered. Hence if we write

$$f(x) \cong \sum_{i=0}^{n} c_i b_i^n(x) = C^T B(x), \quad (6)$$

where the Bernstein coefficient vector C and the Bernstein vector $B(x)$ are given by

$$C^T = [c_0, \ldots, c_n], \ B(x) = [b_0^n(x), b_1^n(x), \ldots, b_n^n(x)]^T, \quad (7)$$

then

$$c_i = \int_0^1 f(x)\, d_i^n(x), \quad i = 0, 1, \ldots, n.$$

Author of [22] has derived explicit representations

$$d_j^n(x) = \sum_{k=0}^{n} \lambda_{jk} b_k^n(x), \quad j = 0, 1, \ldots, n,$$

for the dual basis functions, defined by the coefficients

$$\lambda_{jk} = \frac{(-1)^{j+k}}{\binom{n}{j}\binom{n}{k}} \sum_{i=0}^{\min(j,k)} (2i+1) \binom{n+i+1}{n-j} \binom{n-i}{n-j} \binom{n+i+1}{n-k} \binom{n-i}{n-k} \quad (8)$$

for $j, k = 0, 1, \ldots, n$.

It is worth to mention here that, orthogonal bases such as the Legendre polynomials present very nice stability properties and are very useful in approximation. Bernstein polynomials and Legendre polynomials both span the same spaces and the transformation between Legendre and Bernstein polynomials is comparatively well-conditioned [21]. The Bernstein polynomials is advantageous for practical computations, on account of its intrinsic numerical stability [23]. One of the useful property of Bernstein basis polynomials is that they all vanish at end points of the interval, except the first and the last one, which are equal to one at $x = 0$ and $x = 1$, respectively. This provides greater flexibility in which to impose boundary conditions at the end points of the interval [24]. Also, Bernstein polynomials have two main properties: their sum equals 1 and every $b_i^n(x)$ is positive for all real x belonging to the interval $x \in (0,1)$. Moreover, as pointed by [25], the Bernstein basis polynomials have the following properties:

1) $b_i^n(x)$ has a root with multiplicity i at point $x = 0$ (note if i is 0 there is no root at 0).
2) $b_i^n(x)$ has a root with multiplicity $n-i$ at point $x = 1$ (note if $n = i$ there is no root at 1).

While for the Legendre polynomials, no explicit formula of the roots is known.

2.1. Approximation by Bernstein polynomials

By (6) we see that a function $\mathrm{f}(\mathrm{x})$ can be approximated by Bernstein polynomials as follows:

$$\mathrm{f}(\mathrm{x}) \approx \sum_{a=0}^{m} c_a p_a(x) \quad (9)$$

where

$$c_a = \int_0^1 f(x)\, d_a(x)\, dx, \quad a = 0, 1, \ldots, n, \quad (10)$$

In vector notation, we write

$$f(x) \approx K_M^T \ \hat{P}_M(x), \quad (11)$$

where $M = m + 1$, K is the coefficient vector and is $\hat{P}_M(x)$, M terms function vector. The notion was extended to the two-dimensional space and the two-dimensional Bernstein polynomials of order M are defined as a product function of two Bernstein polynomials

$$P_n(x,y) = P_a(x) P_b(y) \quad , n = Ma + b + 1, \quad a = 0, 1, 2, \ldots, m, \quad b = 0, 1, 2, \ldots, m. \quad (12)$$

The orthogonality condition of $P_n(x,\ y)$ is

$$\int_0^1 \int_0^1 P_a(x) P_b(y) d_i(x) d_j(y)\, dxdy = \begin{cases} 1 & i = a, j = b, \\ 0 & o.w. \end{cases} \quad (13)$$

Any $f\ (x,\ y) \in C([0,\ 1] \times [0,\ 1])$ can be approximated by the polynomials $Pn(x,\ y)$ as follows:

$$f(x,y) \approx \sum_{a=0}^{m} \sum_{b=0}^{m} c_{ab} P_a(x) P_b(y) \quad (14)$$

where

$$\mathrm{c_{ab}} = \int_0^1 \int_0^1 f(x) f(y) d_a^n(x)\, d_b^n(y)\, dxdy \quad (15)$$

For simplicity, we use the notation $C_n = C_{ab}$ where $n = Ma + b + 1$, and rewrite (14) as follows:

$$f(x,y) \approx \sum_{n=1}^{M^2} c_n P_n(x,y) = K_{M^2} \widehat{\psi}(x,y) \quad (16)$$

in vector notation, where K_{M^2} is the $1 \times M^2$ coefficient row vector and $\widehat{\psi}(x,y)$ is the$M^2 \times 1$ column vector of functions defined by

$$\widehat{\psi}(x,y) = (\psi_{11}(x,y), \ldots, \psi_{1M}(x,y), \psi_{21}(x,y), \ldots \psi_{2M}(x,y), \ldots \psi_{MM}(x,y))^T \quad (17)$$

where

$$\psi_{i+1,j+1}(x,\ y) = P_i(x) P_j(y),\ i,\ j = 0,\ 1,\ 2,\ \ldots, m, \quad (18)$$

where

$$c_{Mi+j+1} = \int_0^1 \int_0^1 f(x,y) d_i(x) d_j(y)\, dxdy = \sum_{n=1}^{M^2} c_n \int_0^1 \int_0^1 P_a(x) P_b(y) d_i(x) d_j(y)\, dxdy.$$

3. The operational matrices

3.1. Operational matrices of Riemann–Liouville fractional integral

Lemma 3.1. Let $\psi(x)$ be the function vector, then the integration of order α of $\psi(x)$ is generalized as

$$I^{\alpha}(\psi(x)) \cong P^{\alpha}\psi(x) \quad (19)$$

where P^{α} is the operational matrix of integration of order α and is defined by

$$P^{\alpha} = \begin{pmatrix} \sum_{l=0}^{0}\theta_{0,0,l} & \sum_{l=0}^{0}\theta_{0,1,l} & \cdots & \sum_{l=0}^{0}\theta_{0,j,l} & \cdots & \sum_{l=0}^{0}\theta_{0,m,l} \\ \sum_{l=0}^{1}\theta_{1,0,l} & \sum_{l=0}^{1}\theta_{1,1,l} & \cdots & \sum_{l=0}^{1}\theta_{1,j,l} & \cdots & \sum_{l=0}^{1}\theta_{1,m,l} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{l=0}^{i}\theta_{i,0,l} & \sum_{l=0}^{i}\theta_{i,1,l} & \cdots & \sum_{l=0}^{i}\theta_{i,j,l} & \cdots & \sum_{l=0}^{i}\theta_{i,m,l} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \sum_{l=0}^{m}\theta_{m,0,l} & \sum_{l=0}^{m}\theta_{m,1,l} & \cdots & \sum_{l=0}^{m}\theta_{m,j,l} & \cdots & \sum_{l=0}^{m}\theta_{m,m,l} \end{pmatrix}, \quad (20)$$

where

$$\theta_{i,j,l} = (-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i}\frac{\Gamma(j+1)}{\Gamma(j+1+\alpha)}u_{l,j}. \quad (21)$$

Proof. Using (3) and (5) we have

$$I^{\alpha}b_i^n(x) = \sum_{j=i}^{n}(-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i}I^{\alpha}(x)^j = \sum_{j=i}^{n}(-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i}\frac{\Gamma(j+1)}{\Gamma(j+1+\alpha)}x^{j+\alpha} \quad (22)$$

Now, approximate $x^{j+\alpha}$ by by $m + 1$ termsBernstein polynomials, we have

$$x^{j+\alpha} \cong \sum_{l=0}^{n}u_{l,j}b_l^n(x),$$

$$u_{l,j} = \int_0^1 x^{j+\alpha}d_l^n(x)\,dx = \sum_{k=0}^{n}\lambda_{l,k}\int_0^1 x^{j+\alpha}b_k^n(x)\,dx = \sum_{k=0}^{n}\lambda_{l,k}\sum_{s=k}^{n}(-1)^{s-k}\binom{n}{k}\binom{n-k}{s-k}\int_0^1 x^{j+\alpha+s}dx$$

$$= \sum_{k=0}^{n}\lambda_{l,k}\sum_{s=k}^{n}(-1)^{s-k}\binom{n}{k}\binom{n-k}{s-k}\frac{1}{j+\alpha+s+1} = \sum_{k=0}^{n}\lambda_{l,k}\,{\mu'}_{k,j}.$$

Consequently, (22) takes the form

$$I^{\alpha}b_i^n(x) \approx \sum_{j=i}^{n}\sum_{l=0}^{n}(-1)^{j-i}\binom{n}{i}\binom{n-i}{j-i}\frac{\Gamma(j+1)}{\Gamma(j+1+\alpha)}u_{l,j}b_l^n(x) = \sum_{l=0}^{n}\sum_{j=i}^{n}\theta_{i,j,l}b_l^n(x)$$

$$= \left[\sum_{j=i}^{n}\theta_{i,j,0}, \sum_{j=i}^{n}\theta_{i,j,1}, \ldots, \sum_{j=i}^{n}\theta_{i,j,n}\right]B(x).$$

So we have

$$I^{\alpha}(\psi(x)) \cong P^{\alpha}\psi(x). \quad (23)$$

Lemma 3.2. Let $\widehat{\psi}(x,y)$ be the function vector as defined in (17), then the fractional derivative of order of $\psi(x,y)$ with respect to y is given by

$$D_y^{\alpha}\left(\widehat{\psi}(x,y)\right) \cong H^{\alpha,y}_{M^2\times M^2}\widehat{\psi}(x,y), \quad (24)$$

where $H^{\alpha}_{M^2\times M^2}$ is the operational matrix of differentiation of order α and is defined as

$$H^{\alpha,y}_{M^2\times M^2} = \begin{pmatrix} \Delta_{1,1,k} & \Delta_{1,2,k} & \cdots & & \Delta_{1,r,k} & \cdots & \Delta_{1,M^2,k} \\ \Delta_{2,1,k} & \Delta_{2,1,k} & \cdots & & \Delta_{2,r,k} & \cdots & \Delta_{2,M^2,k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \Delta_{q,1,k} & \Delta_{q,2,k} & \cdots & & \Delta_{q,r,k} & \cdots & \Delta_{q,M^2,k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ \Delta_{M^2,1,k} & \Delta_{M^2,2,k} & \cdots & & \Delta_{M^2,r,k} & \cdots & \Delta_{M^2,M^2,k} \end{pmatrix}, \quad (25)$$

where $q = Mi + j + 1, r = Ma + b + 1, \quad i,j,a,b = 0,1,2,\ldots,m$ and

$$\Delta_{q,r,k} = C_{i,j,a,b} = \sum_{k=\max(b,\lceil\alpha\rceil)}^{m} \delta_{i,a} \sum_{k_1=0}^{m} \lambda_{jk_1}(-1)^{k-b}\binom{m}{b}\binom{m-b}{k-b}\frac{m!\Box(1-\alpha+k+k_1)\Box(k+1)}{k_1!\Box(2-\alpha+k+m)\Box(k+1-\alpha)}. \ (26)$$

Proof. Taking the element $P_n(x,\ y)$ defined by (12), the fractional order partial derivative of $P_n(x,\ y)$ with respect to y is given by relation

$$D_y^{\alpha}(b_m(x,y)) = D_y^{\alpha} b_a^m(x)\, b_b^m(y) = b_a^m(x)\sum_{k=b}^{m}(-1)^{k-b}\binom{m}{b}\binom{m-b}{k-b}D_y^{\alpha}\, y^k$$

$$= \sum_{k=\max(b,\lceil\alpha\rceil)}^{m}(-1)^{k-b}\binom{m}{b}\binom{m-b}{k-b}b_a(x)\frac{\Box(k+1)}{\Box(k+1-\alpha)}y^{k-\alpha}$$

$$= \sum_{k=\max(b,\lceil\alpha\rceil)}^{m}(-1)^{k-b}\binom{m}{b}\binom{m-b}{k-b}\frac{\Box(k+1)}{\Box(k+1-\alpha)}b_a(x)\, y^{k-\alpha}. \quad (27)$$

Approximating $b_a(x)\, y^{k-\alpha}$ by Bernstein polynomials in two variables, yields

$$b_a(x)\, y^{k-\alpha} \approx \sum_{i=0}^{m}\sum_{j=0}^{m} c_{ij}^{k,a} P_i(x)\, P_j(y), \quad (28)$$

where

$$c_{ij}^{k,a} = \int_0^1\int_0^1 b_a(x)\, y^{k-\alpha} d_i(x)\, d_j(y)\, dxdy = \int_0^1 b_a(x)\, d_i(x)\, dx\int_0^1 y^{k-\alpha} d_j(y)\, dy \ = \delta_{ia}\int_0^1 y^{k-\alpha} d_j(y)\, dy,$$

where

$$\delta_{ia} = \begin{cases} 1 & i = a \\ 0 & i \neq a \end{cases}, d_j^n(y) = \sum_{k_1=0}^{n}\lambda_{jk_1} b_{k_1}^n(y), \quad j = 0,1,\ldots,n,$$

$$c_{ij}^{k,a} = \delta_{ia}\sum_{k_1=0}^{m}\lambda_{jk_1}\int_0^1 b_{k_1}^n(y)\, y^{k-\alpha}dy = \delta_{ia}\sum_{k_1=0}^{m}\lambda_{jk_1}\frac{m!\Box(1-\alpha+k+k_1)}{k_1!\Box(2-\alpha+k+m)} \quad (29)$$

hence it follows that

$$D_y^{\alpha} b_a^m(x)\, b_b^m(y) \approx \sum_{k=\max(b,\lceil\alpha\rceil)}^{m} (-1)^{k-b} \binom{m}{b}\binom{m-b}{k-b} \frac{\square\,(k+1)}{\square\,(k+1-\alpha)} \sum_{i=0}^{m}\sum_{j=0}^{m} c_{ij}^{k,a} P_i(x)\, P_j(y)$$

$$= \sum_{i=0}^{m}\sum_{j=0}^{m}\sum_{k=\max(b,\lceil\alpha\rceil)}^{m} (-1)^{k-b} \binom{m}{b}\binom{m-b}{k-b} \frac{\square\,(k+1)}{\square\,(k+1-\alpha)} c_{ij}^{k,a} P_i(x)\, P_j(y)$$

$$\cong \sum_{i=0}^{m}\sum_{j=0}^{m} c_{i,j,a,b} P_i(x)\, P_j(y), \qquad (30)$$

where

$$c_{i,j,a,b} = \sum_{k=\max(b,\lceil\alpha\rceil)}^{m} \delta_{i,a} \sum_{k_1=0}^{m} \lambda_{jk_1} (-1)^{k-b} \binom{m}{b}\binom{m-b}{k-b} \frac{m!\square(1-\alpha+k+k_1)\square(k+1)}{k_1!\square(2-\alpha+k+m)\square(k+1-\alpha)}. \quad (31)$$

Using the notations, $q = Mi + j + 1,\ r = Ma + b + 1$ and $\square q, r, k = C_{i,j,b,a,k}\ for\ i,\ j,\ a,\ b = 0,\ 1,\ 2,\ 3,\ \ldots, m$, we get the desired result .

Lemma 3.3. Let $\square(x,\ y)$ be as defined in $(1, thenthederivativeoforder\beta$ of $\square(x,\ y)$ with respect to x is given by

$$D_x^{\beta}(\psi(x,y)) \cong H_{M^2\times M^2}^{\beta,x} \psi(x,y), \qquad (32)$$

where $H_{M^2\times M^2}^{\beta,x}$ is the operational matrix of derivative of order β and is defined as

$$H_{M^2\times M^2}^{\beta,x} = \begin{pmatrix} \theta_{1,1,k} & \theta_{1,2,k} & \cdots & \theta_{1,r,k} & \ldots & \theta_{1,M^2,k} \\ \theta_{2,1,k} & \theta_{2,1,k} & \ldots & \theta_{2,r,k} & \ldots & \theta_{2,M^2,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \theta_{q,1,k} & \theta_{q,2,k} & \ldots & \theta_{q,r,k} & \ldots & \theta_{q,M^2,k} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \theta_{M^2,1,k} & \theta_{M^2,2,k} & \ldots & \theta_{M^2,r,k} & \ldots & \theta_{M^2,M^2,k} \end{pmatrix}, \qquad (33)$$

where $q = Mi + j + 1, r = Ma + b + 1, i, j, a, b = 0, 1, 2, \ldots, m$ and

$$\theta_{q,r,k} = C_{i,j,a,b} = \sum_{k=\max(a,\lceil\alpha\rceil)}^{m} \delta_{j,b} \sum_{k_1=0}^{m} \lambda_{ik_1} (-1)^{k-a} \binom{m}{a}\binom{m-a}{k-a} \frac{m!\square(1-\alpha+k+k_1)\square(k+1)}{k_1!\square(2-\alpha+k+m)\square(k+1-\alpha)}. \quad (34)$$

Proof. The proof of this lemma 3.3 is similar as the above lemma 3.2.

4. Application of the Bernstein operational matrices for two-dimensional fractional heat conduction problem

Now we apply the Bernstein operational matrices to solve the two-dimensional fractional heat equation (1) and (2) as follows:

Approximate

$$\frac{\partial^{\alpha} u(t,x,y)}{\partial t^{\alpha}} \approx \psi^T(t)\, K_{M\times M^2} \widehat{\psi}(x,y). \qquad (35)$$

Applying fractional integration of order with respect to t on Eq. (35), we have

$$I^{\alpha} \frac{\partial u(t,x,y)}{\partial t^{\alpha}} \approx \psi^T(t)\, P_{M\times M}^{\alpha^T} K_{M\times M^2} \widehat{\psi}(x,y),$$

that implies

$$u\left(t,x,y\right)-c_1\approx\psi^T\left(t\right)P_{M\times M}^{\alpha^T}K_{M\times M^2}\widehat{\psi}\left(x,y\right)\qquad(36)$$

The initial condition $u(0,\ x,\ y)\ =\ f\ (x,\ y)$ yields $c_1\ =\ f\ (x,\ y)$ and the above equation can be written as

$$u\left(t,x,y\right)\approx\psi^T\left(t\right)P_{M\times M}^{\alpha^T}K_{M\times M^2}\widehat{\psi}\left(x,y\right)+f\left(x,y\right),\quad(37)$$

which implies that

$$f\left(x,y\right)\approx\psi^T\left(t\right)F_{M\times M^2},\quad(38)$$

$$u\left(t,x,y\right)\approx\psi^T\left(t\right)\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)\widehat{\psi}\left(x,y\right).\quad(39)$$

Taking β order derivative of $u(t,\ x,\ y)$, we get

$$\frac{\partial^\beta u\left(t,x,y\right)}{\partial x^\beta}\approx\psi^T\left(t\right)\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,x)}\widehat{\psi}\left(x,y\right).\quad(40)$$

and

$$\frac{\partial^\beta u(t,x,y)}{\partial y^\beta}\approx\psi^T\left(t\right)\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,y)}\widehat{\psi}\left(x,y\right).\quad(41)$$

$$g\left(x,y,t\right)\approx\psi^T\left(t\right)G_{M\times M^2}\ \widehat{\psi}\left(x,y\right),\quad(42)$$

Using (35), $(40)-(42)$ in (1) we get

$$\psi^T\left(t\right)K_{M\times M^2}\widehat{\psi}\left(x,y\right)=C_1\psi^T\left(t\right)\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,x)}\widehat{\psi}\left(x,y\right)$$

$$+C_2\psi^T\left(t\right)\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}\quad+F_{M\times M^2}\right)H^{(B,y)}\widehat{\psi}\left(x,y\right)+\psi^T\left(t\right)G_{M\times M^2}\ \widehat{\psi}\left(x,y\right),\quad(43)$$

which can be rewritten as

$$\psi^T\left(t\right)\left(K_{M\times M^2}-C_1(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,x)}$$

$$-C_2\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,y)}-G_{M\times M^2}\ )\widehat{\psi}\left(x,y\right)=0.$$

Hence it follows that

$$K_{M\times M^2}-c_1\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,x)}-c_2\left(P_{M\times M}^{\alpha^T}K_{M\times M^2}+F_{M\times M^2}\right)H^{(B,y)}-G_{M\times M^2}=0\quad(44)$$

Using the value of $K_{M\times M^2}$ in (39) we can get the approximate solution of the problem (1) and (2).

5. Illustrative example

Example . Consider the following two-dimensional fractional heat conduction equation

$$\frac{\partial^\alpha u(t,x,y)}{\partial t^\alpha}=\frac{\partial^\beta u(t,x,y)}{\partial x^\beta}+\frac{\partial^\beta u(t,x,y)}{\partial y^\beta}+2t-6x-6y,$$

subject to the initial condition $u\left(0,x,y\right)=x^3+y^3$ where 0 < ≤ 1 and t, x, y ∈ [0, 1]. The exact solution for = 1 and = 2 is known and is given by $u\left(t,\ x,\ y\right)=t^2+x^3+y^3$. We calculate the absolute error at $t\ =\ 0.5$ found that the error is much more less than 10^{-15}as evident from Fig. 1 show this phenomenon. Also we use the method to approximate solution for the fractional value of $\alpha\&\beta$ and it is found that as $\alpha\ \rightarrow\ 1$ and $\beta\ \rightarrow\ 2$ the approximate solution becomes equal to the exact solution. This phenomenon is shown in Fig. 2 where we observe the approximate solution at t = 0.5.

Fig. 1. The absolute error at time t = 0.5 and M = 4.

Fig. 2. The approximate solution at time t = 0.5 for M = 4 and different values of α and β.

6. Conclusion

The presented method is a simple method for the numerical solution of the fractional order heat conduction equation on a rectangular region. The method easily reduces the problem to a system of easily solvable algebraic equations. We saw that the obtained results in example were in good agreement with the exact solution. Also, we observed that the solutions approach to the solutions for standard problems $(\alpha = 1, \;\; \beta = 2)$ as $\alpha \to \; 1 \, and \, \beta \; \to \; 2$.

References

[1] Akbar Mohebbi, Mostafa Abbaszadeh, Mehdi Dehghan, A high-order and unconditionally stable scheme for the modified anomalous fractional subdiffusion equation with a nonlinear source term, J. Comput. Phys. 240 (2013) 36–48.
[2] Akbar Mohebbi, Mostafa Abbaszadeh, Mehdi Dehghan, The use of a meshless technique based on collocation and radial basis functions for solving the time fractional nonlinear Schrodinger equation arising in quantum mechanics, Eng. Anal. Bound. Elem. 37 (2013) 475–485.
[3] J.H. Jeong, M.S. Jhon, J.S. Halow, V. Osdol, Smoothed particle hydrodynamics: applications to heat conduction, Comput. Phys. Comm. 153 (2003) 71–84.
[4] Y. Liu, X. Zhang, M.-W. Lu, A meshless method based on least squares approach for steady and unsteady state heat conduction problems, Numer. Heat Transfer 47 (2005) 257–275.
[5] M. Akbarzade, J. Langari, Application of Homotopy perturbation method and variational iteration method to three dimensional diffusion problem, Int. J. Math. Anal. 5 (2011) 871–880.
[6] Mehdi Dehghan, Jalil Manafian, Abbas Saadatmandi, The solution of the linear fractional partial differential equations using the homotopy analysis method, Z. Naturforsch. 65a (2010) 935–949.
[7] Abbas Saadatmandi, Mehdi Dehghan, A tau approach for solution of the space fractional diffusion equation, Comput. Math. Appl. 62 (2011) 1135–1142.
[8] S. Soleimani, D. Domiri Ganji, E. Ghasemi, M. Jalaal, Bararnia, Meshless local RBF-DQ for 2-D heat conduction: a comparative study, Therm. Sci. 15 (**??**)
(2011) 117–121.
[9] Abbas Saadatmandi, Mehdi Dehghan, Mohammad-Reza Azizi, The Sinc–Legendre collocation method for a class of fractional convection–diffusion equations with variable coefficients, Commun. Nonlinear Sci. Numer. Simul. 17 (2012) 4125–4136.
[10] Abbas Saadatmandi, Mehdi Dehghan, A Legendre collocation method for fractional integro-differential equations, J. Vib. Control 17 (2013) 2050–2058.
[11] J.H. He, Xu-Hong Wu, Exp-function method for nonlinear wave equations, Chaos Solitons Fractals 30 (**??**) (2006) 700–708.
[12] J.H. He, Non-perturbative methods for strongly nonlinear problems, Dissertation, De-Verlag im Internet GmbH, Berlin, 2006.
[13] Mehdi Deghan, Y.A. Yousefi, A. Lotfi, The use of He's variational iteration method for solving the telegraph and fractional telegraph equations, Comm. Numer. Methods Engrg. 27 (2011) 219–231.
[14] Hammad Khalil, Rahmat Ali Khan, New operational matrix of inegration and coupled system of Fredholm integral equations. Transformation 10.1: 2.
[15] Hammad Khalil, Rahmat Ali Khan, A new method based on legender polynomials for solution of system of fractional order partial differential equation, Int. J. Comput. Math. (2014).
[16] B. Ross (Ed.), Fractional Calculus and its Applications, in: Lecture Notes in Mathematiucs, vol. 457, 1975.
[17] K.B. Oldhalm, J. Spainer, The Fractional Calculus, Academic Press, New York, 1974.

[18] I. Podlubny, Fractional Differential Equations, Academic Press, New York, 1999.
[19] S.A. Yousefi, M. Behroozifar, Operational matrices of Bernstein polynomials and their applications, Int. J. Syst. Sci. 41 (2010) 709–716.
[20] S.A. Yousefi, M. Behroozifar, M. Dehghan, The operational matrices of Bernstein polynomials for solving the parabolic equation subject to specification of the mass, J. Comput. Appl. Math. 235 (2011) 5272–5283.
[21] R.T. Farouki, Legendre-Bernstein basis transformations, J. Comput. Appl. Math. 119 (2000) 145–160.
[22] B. Jüttler, The dual basis functions for the Bernstein polynomials, Adv. Comput. Math. 8 (1998) 345–352.
[23] R.T. Farouki, T.N.T. Goodman, On the optimal stability of the Bernstein basis, Math. Comput. 64 (1996) 1553–1566.
[24] M. Idrees Bhatti, P. Bracken, Solutions of differential equations in a Bernstein polynomial basis, J. Comput. Appl. Math. 205 (2007) 272–280.
[25] S.A. Yousefi, Z. Barikbin, M. Dehghan, Bernstein Ritz-Galerkin method for solving an initial-boundary value problem that combines Neumann and integral condition for the wave equation, Numer. Methods Partial Diff. Eq. 26 (2010) 1236–1246.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–287, pp. 230–235

Preconditioned Global CGS Method to Solve Saddle Point Problems with Multiple Right-Hand Sides

Mohammad Mahdi Izadkhah[a,*]

[a]*Department of Computer Science, Faculty of Computer & Industrial Engineering, Birjand University of Technology, Birjand, Iran*

Article Info

Keywords:
Saddle point problems
Global CGS method
Indefinite preconditioning

2020 MSC:
65F10
65F45

Abstract

The conjugate gradient squared method is a desirable approach for solving nonsymmetric linear system of equations due to its low computational cost, compared to the other iterative solvers. In this paper, we present a preconditioned global conjugate gradient squared method to solve the saddle point problems with multiple right-hand sides arising from Stocks equation. Finally, some numerical experiments are provided to illustrate the feasibility and validity of the algorithm proposed.

1. Introduction

Recently, systems with multiple right-hand sides have been occurred in many research and engineering application areas, such as mixed or mixed-hybrid finite element disceritization of partial differential equations in computational fluid dynamics [1], constrained optimization [2] and so on.
In this paper we deal with the following saddle point problem with multiple right-hand sides

$$\begin{bmatrix} A & B \\ \epsilon B^T & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix, $B \in \mathbb{R}^{n\times m}$ has a full column rank, $X \in \mathbb{R}^{n\times s}, Y \in \mathbb{R}^{m\times s}$, $F_1 \in \mathbb{R}^{n\times s}$, $F_2 \in \mathbb{R}^{m\times s}$ and ϵ is a given scalar. To convenience, we denote system (1) as follow:

$$\mathcal{A}\mathcal{X} = \mathcal{B}, \tag{2}$$

where $\mathcal{X} = [\mathcal{X}^{(1)}, \dots, \mathcal{X}^{(s)}]$ and $\mathcal{B} = [\mathcal{B}^{(1)}, \dots, \mathcal{B}^{(s)}]$. Under the aforementioned assumptions, the coefficient matrix $\mathcal{A}$ is nonsingular.
In the last decade, several efficient iterative methods have been developed to solve the saddle point problems, such as the Uzawa method, the HSS method and related variants [3] and Krylov subspace methods [1]. For the purpose of the improvement to the efficiency of standard iterative solvers, many preconditioners have been proposed in the literature.

*Talker
Email address: izadkhah@birjandut.ac.ir (Mohammad Mahdi Izadkhah)

For example, the block diagonal preconditioners, the block triangular preconditioners, and the parameterized block triangular preconditioners [2]. In this paper, we investigate the preconditioned global BiCG (PGl-BiCG) method [5] and preconditioned global conjugate gradient squared (PGl-CGS) method for solving linear system with multiple right-hand sides (1). We concentrate on the use of the indefinite preconditioner

$$P = \begin{bmatrix} I & B \\ \epsilon B^T & 0 \end{bmatrix}, \tag{3}$$

where I is the identity matrix of order n. This choice has been shown to be particulary effective on problems associated with constrained nonlinear programming. Throughout this paper, we use the following notations. Inner product for two $n \times s$ matrices X and Y is defined as $\langle X, Y\rangle_F = \mathrm{tr}(X^T Y)$, where $\mathrm{tr}(Z)$ denotes the trace of the square matrix Z. The associated norm is the Frobenius norm denoted by $\|\cdot\|_F$. We will use the notation $\langle\cdot,\cdot\rangle_2$ for the usual inner product in $\mathbb{R}^n$, and the related norm will be denoted by $\|\cdot\|_2$. For a matrix $V \in \mathbb{R}^{n\times s}$, the matrix Krylov subspace $\mathcal{K}_m(A, V)$ is defined by

$$\mathcal{K}_m(A, V) = \mathrm{span}\{V, AV, \ldots, A^{m-1}V\}.$$

Moreover, $Z \in \mathcal{K}_j(A, V)$ means that $Z = \sum_{i=0}^{j-1} \xi_i A^i V$, where $\xi_i \in \mathbb{R}$, for $i = 0, 1, \ldots, j-1$. Finally, $0_s, I_s$, and $0_{l\times s}$ will denote the zero, the identity, and zero matrices in $\mathbb{R}^{s\times s}, \mathbb{R}^{s\times s}$, and $\mathbb{R}^{l\times s}$, respectively. For brevity, we use the MATLAB-like notation $[v; w]$ to represent the vector $[v^T w^T]^T$.

2. Properties of the indefinite preconditioner

In this work, we utilize the global version of the CGS (Gl-CGS) [4] for the solution of the nonsymmetric saddle point problems with multiple right-hand sides (2). The convergence behavior of this method without a good preconditioner is very slow, when applied to the saddle point problems with multiple right-hand sides. In order to accelerate the convergence, we use the indefinite matrix P defined in (3) as a right preconditioner for the Gl-CGS algorithm applied to the problem (1) as

$$\mathcal{A}P^{-1} \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix} = \mathcal{B}, \quad \begin{bmatrix} X \\ Y \end{bmatrix} = P^{-1} \begin{bmatrix} \tilde{X} \\ \tilde{Y} \end{bmatrix}, \tag{4}$$

where

$$P^{-1} = \begin{bmatrix} I - \Pi & \frac{1}{\epsilon} B(B^T B)^{-1} \\ (B^T B)^{-1} B^T & -\frac{1}{\epsilon}(B^T B)^{-1} \end{bmatrix}, \quad \mathcal{A}P^{-1} = \begin{bmatrix} G & S \\ 0 & I \end{bmatrix},$$

with $\Pi = B(B^T B)^{-1} B^T$, $G = A(I - \Pi) + \Pi$, and $S = \frac{1}{\epsilon}(A - I)B(B^T B)^{-1}$. Once an approximate solution $[\tilde{X}_k; \tilde{Y}_k]$ is determined, an approximate solution to the unpreconditioned problem is recovered as $[X_k; Y_k] = P^{-1}[\tilde{X}_k; \tilde{Y}_k]$. Choosing the vector $[\tilde{X}_0, \tilde{Y}_0] = [0; F_2]$ as the initial approximate solution, the starting residual is given by

$$R_0 = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} - \mathcal{A}P^{-1} \begin{bmatrix} \tilde{X}_0 \\ \tilde{Y}_0 \end{bmatrix} = \begin{bmatrix} F_1 - SF_2 \\ 0 \end{bmatrix} = \begin{bmatrix} R_0^{(1)} \\ 0 \end{bmatrix},$$

so that the second block component of R_0 is identically zero. Problem (4) can thus be reformulated as determining an approximation $[\bar{X}_k; \bar{Y}_k]$ to the solution $[\bar{X}; \bar{Y}]$ of the system

$$\mathcal{A}P^{-1} \begin{bmatrix} \bar{X} \\ \bar{Y} \end{bmatrix} = R_0, \tag{5}$$

so that $[\tilde{X}_k; \tilde{Y}_k] = [\tilde{X}_0; \tilde{Y}_0] + [\bar{X}_k; \bar{Y}_k]$. In addition, for every $[U; 0] \in \mathbb{R}^{(m+n)\times s}$, we have

$$\mathcal{A}P^{-1} \begin{bmatrix} U \\ 0 \end{bmatrix} = \begin{bmatrix} GU \\ 0 \end{bmatrix}, \text{ and } P^{-1} \begin{bmatrix} U \\ 0 \end{bmatrix} = \begin{bmatrix} (I - \Pi)U \\ (B^T B)^{-1} B^T U \end{bmatrix}. \tag{6}$$

3. The preconditioned global CGS method

At the first, we employ the PGl-BiCG method [5] to solve system (4). Let $X_0 \in \mathbb{R}^{n\times s}$ be an initial guess with the residual $R_0 = \mathcal{B} - \mathcal{A}X_0$ and let $\tilde{R}_0$ be an arbitrary $n \times s$ matrix. The residual R_k generated by the step k of the PGl-BiCG method is such that, $R_k - R_0$ lies in the right matrix Krylov subspace $\mathcal{K}_k(\mathcal{A}, R_0)$ and R_k is F-orthogonal to the left matrix Krylov subspace $\mathcal{K}_k(\mathcal{A}^T, \tilde{R}_0)$. We say that two set of matrices $\{V_i\}$ and $\{W_i\}$ are F-orthogonal if and only if $\langle V_i, W_j\rangle_F = 0$ for $i \neq j$ and $\langle V_i, W_i\rangle_F = 1$ for $i = 1, 2, \dots, m$.
For solving (1) by preconditioned global biconjogate gradient (PGl-BiCG) algorithm by using the preconditioner P defined in (3) and equations (4) and (5), we choose $[\tilde{X}_0; \tilde{Y}_0] = [0; F_2]$ as the initial guess. So, $R_0 = [R_0^{(1)}; 0]$. We set $\tilde{R}_0 = P^{-1}R_0 = [(I-\Pi)R_0^{(1)}; (B^TB)^{-1}B^TR_0^{(1)}], P_0 = R_0, \tilde{P}_0 = \tilde{R}_0$, and we obtain $\tilde{R}_k = P^{-1}R_k = [(I-\Pi)R_k^{(1)}; (B^TB)^{-1}B^TR_k^{(1)}]$ and $\tilde{P}_k = P^{-1}P_k = [(I-\Pi)P_k^{(1)}; (B^TB)^{-1}B^TP_k^{(1)}]$. Therefore, the iterates $\tilde{R}_k$ and $\tilde{P}_k$ can be computed explicitly from R_k and P_k, and the auxiliary "tilde" recurrence can be omitted. Now, by using the relations (6) and ignoring from the last m rows of the matrices that are zero, we can summarize the PGl-BiCG algorithm for solving (2) as the following Algorithm 1.
Algorithm 1: The right PGl-BiCG method

1 Set $\begin{bmatrix} \tilde{X}_0 \\ \tilde{Y}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ F_2 \end{bmatrix}$ and $\begin{bmatrix} \bar{X}_0 \\ \bar{Y}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

2. Compute $R_0^{(1)} = F_1 - SF_2$ and set $R_0 = \begin{bmatrix} R_0^{(1)} \\ 0 \end{bmatrix}$ and $P_0 = \begin{bmatrix} P_0^{(1)} \\ P_0^{(2)} \end{bmatrix} = \begin{bmatrix} R_0^{(1)} \\ 0 \end{bmatrix}$

3. **for** $k = 0, 1, 2, \dots$ **until convergence**

4. $\alpha_k = \frac{\langle R_k^{(1)}, (I-\Pi)R_k^{(1)}\rangle_F}{\langle GP_k^{(1)}, (I-\Pi)P_k^{(1)}\rangle_F}$

5. $\bar{X}_{k+1} = \bar{X}_k + \alpha_k P_k^{(1)}, \qquad \bar{Y}_{k+1} = 0_{m\times s}$

6. $R_{k+1}^{(1)} = R_k^{(1)} - \alpha_k GP_k^{(1)}, \qquad R_{k+1}^{(2)} = 0$

7. $\beta_k \frac{\langle R_{k+1}^{(1)}, (I-\Pi)R_{k+1}^{(1)}\rangle_F}{\langle R_k^{(1)}, (I-\Pi)R_k^{(1)}\rangle_F}$

8. $P_{k+1}^{(1)} = R_{k+1}^{(1)} + \beta_k P_k^{(1)}, \qquad P_{k+1}^{(2)} = 0$

9. **end**

10. $\tilde{X}_{k+1} = \tilde{X}_0 + \bar{X}_{k+1}, \qquad \tilde{Y}_{k+1} = \tilde{Y}_0 + \bar{Y}_{k+1}$

11. $X_{k+1} = (I-\Pi)\tilde{X}_{k+1} + \frac{1}{\epsilon}B(B^TB)^{-1}\tilde{Y}_{k+1}, \qquad Y_{k+1} = (B^TB)^{-1}B^T\tilde{X}_{k+1} - \frac{1}{\epsilon}(B^TB)^{-1}\tilde{Y}_{k+1}$

In practical implementation of **Algorithm 1**, we can factorize B as $B = Q\begin{bmatrix} R \\ 0 \end{bmatrix}$ and use the relation $(B^TB)^{-1} = R^{-1}R^{-T}$. From **Algorithm 1**, the first block of residuals and the first block of matrix directions can be expressed as follows:

$$R_k^{(1)PGl-BiCG} = \mathcal{R}_k(G)R_0^{(1)}, \quad P_k^{(1)PGl-BiCG} = \mathcal{P}_k(G)R_0^{(1)}, \tag{7}$$

where $\mathcal{R}_k(t)$ and $\mathcal{P}_k(t)$ are the polynomials of degree k with scalar coefficients satisfying $\mathcal{R}_k(0) = 1$ and $\mathcal{P}_k(0) = 1$, respectively. The polynomials $\mathcal{R}_k(t)$ and $\mathcal{P}_k(t)$ are related together with the recurrence formulas as follows:

$$\mathcal{R}_{k+1}(t) = \mathcal{R}_k(t) - \alpha_k t\mathcal{P}_k(t), \tag{8}$$
$$\mathcal{P}_{k+1}(t) = \mathcal{R}_{k+1}(t) + \beta_k\mathcal{P}_k(t). \tag{9}$$

In the PGl-CGS method, the matrix residual and the matrix direction satisfies

$$R_k = \mathcal{R}_k(\mathcal{A}P^{-1})^2 R_0, \qquad P_k = \mathcal{P}_k(\mathcal{A}P^{-1})^2 R_0, \tag{10}$$

where $\mathcal{R}_k$ and $\mathcal{P}_k$ are defined as before. We introduce the matrix Q_k as

$$Q_k = \mathcal{R}_{k+1}(\mathcal{A}P^{-1})\mathcal{P}_k(\mathcal{A}P^{-1})R_0.$$

If the PGl-CGS method is written for system (5) with $R_0 = [R_0^{(1)}; 0]$, then the relation (10) can be written as

$$R_k = \begin{bmatrix} \mathcal{R}_k(G)^2 R_0^{(1)} \\ 0 \end{bmatrix}, \tag{11}$$

and for solving the problem (2), the PGl-CGS algorithm can be summarized as **Algorithm 2**.

Algorithm 2: The right PGl-CGS method

1 Set $\begin{bmatrix} \tilde{X}_0 \\ \tilde{Y}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ F_2 \end{bmatrix}$ and $\begin{bmatrix} \bar{X}_0 \\ \bar{Y}_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

2. Compute $R_0^{(1)} = F_1 - SF_2$ and set $R_0 = \begin{bmatrix} R_0^{(1)} \\ 0 \end{bmatrix}$ and $P_0 = \begin{bmatrix} P_0^{(1)} \\ P_0^{(2)} \end{bmatrix} = \begin{bmatrix} R_0^{(1)} \\ 0 \end{bmatrix}$

3. Set $U_0 = P_0$

3. **for** $k = 0, 1, 2, \ldots$ **until convergence**

4. $\alpha_k = \frac{\langle R_k^{(1)}, (I-\Pi)R_0^{(1)}\rangle_F}{\langle GP_k^{(1)}, (I-\Pi)R_0^{(1)}\rangle_F}$

5. $Q_k = U_k - \alpha_k G P_k$

6. $\bar{X}_{k+1} = \bar{X}_k + \alpha_k(U_k + Q_k), \qquad \bar{Y}_{k+1} = 0$

7. $R_{k+1}^{(1)} = R_k^{(1)} - \alpha_k G(U_k + Q_k), \qquad R_{k+1}^{(2)} = 0$

8. $\beta_k = \frac{\langle R_{k+1}^{(1)}, (I-\Pi)R_0^{(1)}\rangle_F}{\langle R_k^{(1)}, (I-\Pi)R_0^{(1)}\rangle_F}$

9. $U_{k+1} = R_{k+1}^{(1)} + \beta_k Q_k$

10. $P_{k+1} = U_{k+1} + \beta_k(Q_k + \beta_k P_k),$

11. **end**

12. $\tilde{X}_{k+1} = \tilde{X}_0 + \bar{X}_{k+1}, \qquad \tilde{Y}_{k+1} = \tilde{Y}_0 + \bar{Y}_{k+1}$

13. $X_{k+1} = (I - \Pi)\tilde{X}_{k+1} + \frac{1}{\epsilon}B(B^TB)^{-1}\tilde{Y}_{k+1}, \qquad Y_{k+1} = (B^TB)^{-1}B^T\tilde{X}_{k+1} - \frac{1}{\epsilon}(B^TB)^{-1}\tilde{Y}_{k+1}$

4. Numerical experiments

In this section, we consider the stokes equation as

$$\begin{aligned} -\nu\Delta\mathbf{u} + \bigtriangledown p &= \tilde{f}, \quad \text{in } \Omega \\ \bigtriangledown \cdot \mathbf{u} &= \tilde{g}, \quad \text{in } \Omega \\ \mathbf{u} &= 0, \quad \text{on } \partial\Omega \\ \int_\Omega p(x)\,dx &= 0, \end{aligned} \tag{12}$$

where $\Omega = (0,1) \times (0,1) \subseteq \mathbb{R}^2$, $\partial\Omega$ is the boundary of Ω, ν is the viscosity scalar, and $\mathbf{u}$ and p denote the velocity and the pressure, respectively. By discretizing (12), we obtain te system of linear equations as

$$\begin{bmatrix} A & B \\ -B^T & 0 \end{bmatrix} \begin{bmatrix} U \\ P \end{bmatrix} = \begin{bmatrix} F_1 \\ -F_2 \end{bmatrix} \tag{13}$$

in which

$$A = \begin{bmatrix} I\otimes T + T\otimes I & 0 \\ 0 & I\otimes T + T\otimes I \end{bmatrix} \in \mathbb{R}^{2q^2\times 2q^2} \quad B = \begin{bmatrix} I\otimes F \\ F\otimes I \end{bmatrix} \in \mathbb{R}^{2q^2\times q^2}$$

where T and F are tridiagonal matrices given by

$$T = \frac{\nu}{h^2}\text{tridiag}(-1,2,-1) \in \mathbb{R}^{q\times q}, \qquad F = \frac{1}{h}\text{tridiag}(-1,1,0) \in \mathbb{R}^{q\times q}$$

and $\otimes$ denotes the Kronecker product. Also $h = \frac{1}{q+1}$ is the discretization mesh size. We set $n = 2q^2$ and $m = q^2$. Hence the total number of variables is $n + m = 3q^2$. We choose the right-hand side such that the exact solution of the saddle point problem (13) is a matrix of ones. In this example, we use zero tensor as the initial guess and

$$ERR \equiv \frac{\|R_k\|}{\|R_0\|} \leq 10^{-5},$$

as the stopping criterion and the maximum number of iterations Max-Iter $= 500$. The computations was performed in double-precision floating-point arithmetic in MATLAB codes.
For this example, we obtain numerical results of the number of iteration (Iter) and the elapsed time (CPU time) in Table 1. In Figure 1, we display the convergence history of the PGl-CGS and PGl-BiCG algorithms for Stokes problem (12) with $s = 5$, $\nu = 0.01$ and $q = 16$. The results in Figure 1 show that the PGl-CGS method is more effective and smoother than the PGl-BiCG method.

Table 1. Numerical results for $\nu = 0.01$

Grid	q=8	q=12	q=16
PGl-BiCG			
Iter	21	31	39
CPU time	0.0372	0.3994	2.1306
PGl-CGS			
Iter	10	14	19
CPU time	0.0185	0.1595	1.0264

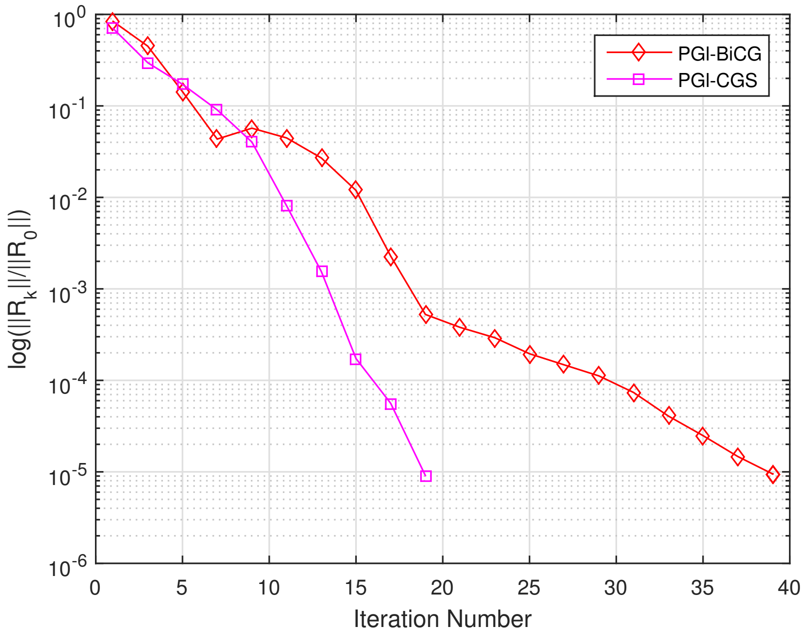


Fig. 1. Comparison of convergence histories for the case $s = 5$, $\nu = 0.01$ and $q = 16$

References

[1] Badahmane, A., Bentbib, A.H., and Sadok, H. *Preconditioned global Krylov subspace methods for solving saddle point problems with multiple right-hand sides*, Electron. Trans. Numer. Anal. 51 (2019), 495-511.

[2] Badahmane, A., Bentbib, A.H. and Sadok, H. *Preconditioned Krylov subspace and GMRHSS iteration methods for solving the nonsymmetric saddle point problems*, Numer. Algorithms, 84 (2020), 1295-1312.

[3] Bai Z.-Z. and Benzi, M. *Regularized HSS iteration methods for saddlepoint linear systems*, BIT Numerical Mathematics, 57 (2017), 287-311.

[4] P. Sonneveld, *CGS, A Fast Lanczos-Type Solver for Nonsymmetric Linear systems*, SIAM Journal on Scientific and Statistical Computing, 13 (1989) 36-52.

[5] Zhang, J. and Dai, H. *Global GPBiCG method for complex non-Hermitian linear systems with multiple right-hand sides*, Comput. Appl. Math. 35 (2016), 171-185.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–101, pp. 236–239

Introducing new nonparametric coefficient of variation (NCV) for detection of the stable cultivars in multi-environment trials

Naser Sabaghnia[a,*], Mohsen Janmohammadi[a]

[a]*Department of Plant Genetics and Production, University of Maragheh, Maragheh, Iran.*

Article Info

Keywords:
GE
interaction
percentiles

Abstract

Performance of new released cultivars in multi-environment trials are analyzed by various parametric and nonparametric methods for exploring stabile cultivars via biometricians. For a nonparametric estimation of stability in test environments, a new nonparametric statistic as NCV (nonparametric coefficient of variation) has been introduced which is based upon the ranks of the cultivars in each test environment. This new statistic use median as a nonparametric central tendency, and nonparametric index of statistical dispersion as inter-percentile range. The NCV nonparametric stability statistic which presented here is similar to the concept of environmental coefficient of variation which was previously proposed for detection of the stable cultivars in multi-environment trials. Our research showed that the most stable cultivars based on the lowest values of this nonparametric statistic, had the highest mean yield among studied genotypes. Plotting of mean yield versus NCV verified the above results and showed that the highest mean yielding cultivar is identified as the most stable one. This nonparametric statistic would be useful for simultaneous selection for yield and stability, so this model can provide some flexible tool for biometricians.

1. Introduction

Multi-environment yield trials are carried out over several years in experimental stations and are central to plant breeding programs to evaluate and improve various crops. These trials are the most common and important experiments in agricultural research, and various statistical methods have been extensively developed and discussed to effectively analyze yield trials. In most cases, only simple statistical methods are needed when cultivars show similar results between test environments. Evaluation in terms of relative performance and utility is complicated by the failure of two or more cultivars to respond similarly to a test environment, known as cultivar by environment interactions. such interactions had important effects on improving cultivar buffering because they impede the extrapolation of agronomic assessments from one environment to another, so more knowledge about the magnitude of cultivar by environment interactions and different sources of variation in cultivar by environment interactions is needed [1]. Ignoring cultivar by environment interactions is problematic when it is larger than the main effect of cultivar, which is a common

*Talker

Email addresses: sabaghnia@maragheh.ac.ir; sabaghnia@yahoo.com (Naser Sabaghnia), mohsen_janmohammadi@yahoo.com (Mohsen Janmohammadi)

problem in multiple environment yield experiments. Additionally, cultivar by environment interactions complicate cultivar recommendations, as cultivars must target specific test sites. In most cases, analysis of variance estimates the existence, importance, and magnitude of cultivar by environment interactions, but cannot explain their significance. Therefore, several statistical strategies were developed to analyze cultivar by environment interaction patterns. Whether this statistical strategy is sufficient to explain cultivar by environment interactions remains a matter of debate. The stabilization methods for both strategies mentioned above are all parametric. In contrast, the third strategy is a nonparametric stability statistic, which is largely independent of the data distribution. These stability methods are rank-based, and a given cultivar is considered stable if its rank is constant across environments. Several nonparametric stability statistics have been developed to account for cultivar by environment interactions. The nonparametric stability statistics separate cultivars based on their similarity of response to a range of test environments. The nonparametric strategy is based on ranks of cultivars and provides an important alternative to the parametric strategies including univariate and multivariate statistics. According to Sabaghnia et al. [2], the nonparametric strategy has some advantages over the parametric strategies such as: reduction of the bias caused by outliers, no assumptions are required about the data distribution, easy to use and to interpret, additions or deletions of few cultivars or environments do not cause much variation of estimates, and for many applications such as selection in plant breeding programs and cultivar testing trials, the rank order of cultivars is the most essential information. The good ability of the nonparametric stability statistics for detecting the most stable cultivars as well as cultivar by environment interaction investigation have been demonstrated in different crops. The objective of this study was estimation of stability performance of cultivars in different test environments using a new nonparametric statistic as NCV (nonparametric coefficient of variation) which is based upon the ranks of the cultivars in each environment.

2. Materials and Methods

If x_{ij} is denoted as observed mean value of the ith cultivar in the jth environment $(i = 1, 2, \ldots, M; j = 1, 2, \ldots, N)$. Then, r_{ij} is considered as the rank of cultivar i in environment j which the lowest value is rank 1 and the highest value is rank of K. The concept of stability is practicable; a cultivar is the most stable over test environments if its ranks are similar over environments, and so maximum stability = equal ranks over all test environments. The new nonparametric stability statistics as NCV (nonparametric coefficient of variation) which is proposed in this paper is:

$$NCV = \frac{(P_{95} - P_5)}{M_{d_i}}.$$

In the above nonparametric statistic, $P_{95} - P_5$ is the inter-percentile range, is a nonparametric index of statistical dispersion, being equal to the difference between the upper and lower percentiles. Mdi is the median of the cultivars' ranks in the test environments. The NCV nonparametric stability statistic which presented here is similar to the nature and concept of ECV (environmental coefficient of variation) [3] as parametric CV (coefficient of variation), and two nonparametric stability statistics, $NS_i^{(1)}$ and $NS_i^{(2)}$, as nonparametric CV [4]. The important central tendency of ranks

Table 1. Two-way layout cited from the reference [5].

Cultivar	E1	E2	E3	E4	E5	E6	Mean	Median	P_5	P_{95}	NCV
Man.	162	247	185	219	165	155	189	2.5	1	3	0.80
Sva.	188	258	182	183	139	144	182	1	1	3	2.00
Vel.	200	263	195	220	166	146	198	4	2	5	0.75
Tre.	197	339	271	266	151	194	236	5	2	5	0.60
Pea.	183	254	219	201	184	190	205	3	2	5	1.00

is the median and its related measures of dispersion can be inter-percentile range. It would be interesting that compare this nonparametric statistic with the environmental coefficient of variation. The ECV was designed to exploration in investigation on the physiological basis for stability [3], and was found more practical to characterize cultivars on a group basis rather than individually. However, this procedure and its related concept could be used in the breeding because it represents a simple and descriptive tool for investigation of cultivars' stability. Considering these benefits of

ECV concept, using new nonparametric stability statistic (NCV) could be useful in cultivar by environment interaction interpreting and identification of the most stable cultivars especially in nonparametric strategy.

3. Findings

A barley multi-environmental trails [5] dataset is used in this research and its two-way layout of yield for five barley cultivars at six environments. The cultivars mean, the ranks of cultivars in environments and the median of these ranks are given in Table 1. According to the obtained results, cultivar #4 (Tre.) was the most stable cultivar based on the lowest values of NCV statistic. This cultivar had the high mean yield performance among studied cultivars (Table 1). Plotting of mean yield versus NCV (Fig. 1) verified the above results and indicated that the high mean yielding cultivar is identified as the most stable cultivar. In other word, this nonparametric statistic would be useful for simultaneous selection for mean yield and stability. Simultaneous selection for mean yield and stability of performance is an important issue in breeding programs. Several stability models for simultaneous selection for yield and stability of performance have been investigated. According to literature, most nonparametric stability statistics of could not be useful for simultaneous selection of mean yield and stability [2], while this nonparametric stability statistic (NCV), showed good detection power for both yield and stability.

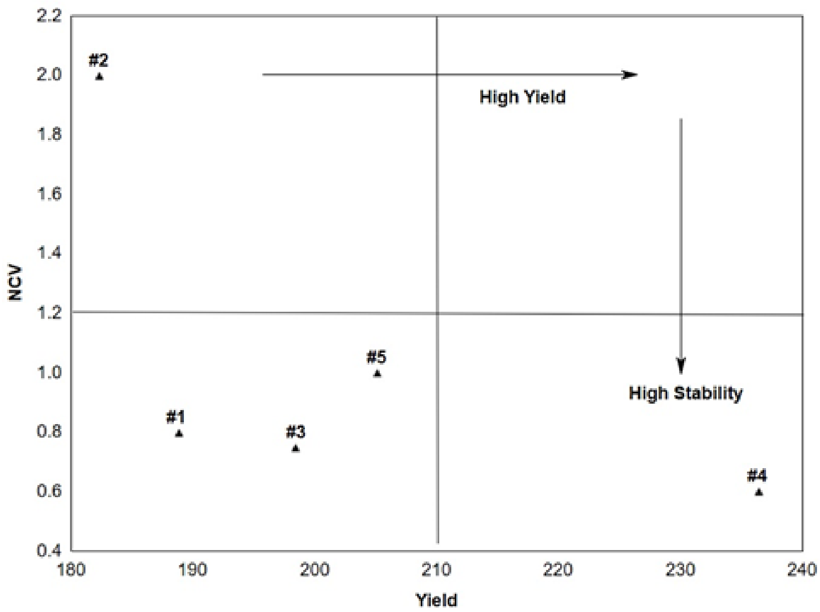


Fig. 1. Plot of mean yield versus the NCV nonparametric statistic. Cultivars are: 1: Man., 2: Sva., 3: Vel., 4: Tre., 5: Pea.

4. Conclusion

It seems that the new nonparametric stability statistic (NCV) has similar nature of ECV and so benefits from Type I of stability. In contrast, detection of high yielding cultivar as the stable one benefits from dynamic concept of stability. However, for simultaneous selection of mean yield and stability, it is necessary to use mean yield in the formula of each stability statistic. This method thus provides some flexibility in the hands of plant biometricians for simultaneous selection for yield and stability. Many parametric and nonparametric statistics of stability have been presented in the literature [1], while for making recommendation, it is essential to investigate the association among these statistics and compare their statistical powers through biometricians.

References

[1] F. Flores M. T. Moreno, and J.I. Cubero. A comparison of univariate and multivariate methods to analyze $G \times E$ interaction. Field Crops Research, 56: 271–286, 1998.

[2] N. Sabaghnia, H. Dehghani, S.H. Sabaghpour. Nonparametric methods for interpreting genotype × environment interaction of lentil genotypes. Crop Science, 46: 1100–1106, 2006.

[3] T.R. Francis, L.W. Kannenberg. Yield stability studies in short-season maize: I. A descriptive method for grouping genotypes. Canadian Journal of Plant Science, 58: 1029–1034, 1978.

[4] N. Sabaghnia. Identification of the most stable genotypes in multi-environment trials by using nonparametric methods. Acta Agriculturae Slovenica, 105(1): 103-110, 2015.

[5] F. Yates, W.G. Cochran. The analysis of groups of experiments. Journal of Agricultural Science, 28: 556–580, 1938.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–102, pp. 240–249

Nonparametric Bayesian optimal designs for Unit Exponential regression model with Respect to Prior Processes(with Polya Urn Scheme as the base measure)

Anita Abdollahi Nanvapisheh[a,*], habib Jafari[a], Soleiman Khazaei[a]

[a] *Department of Statistics, Razi University, Kermanshah, Iran*

Article Info

Keywords:
D-optimal design
Nonparametric Bayesian optimal design
Unit Exponential model (UE)

2020 MSC:
60E05
60Exx

Abstract

Nonlinear regression models find extensive applications across various scientific disciplines. It is vital to accurately fit the optimal nonlinear model while considering the biases of the Bayesian optimal design. By utilizing the Dirichlet process as a prior, we present a Bayesian optimal design. The Dirichlet process serves as a fundamental tool in the exploration of Nonparametric Bayesian inference, offering multiple representations that are well-suited for application. This research paper introduces a novel one-parameter model, referred to as the ”Unit-Exponential distribution”, specifically designed for the unit interval. Additionally, we employ a representation to approximate the D-optimality criterion considering the Dirichlet process as a functional tool. Through this approach, we aim to identify a Nonparametric Bayesian optimal design.

1. Introduction

Within the realm of experimental design, the concept of optimal design refers to a specific category of designs that are classified based on certain statistical criteria. It is widely acknowledged that a well-designed experiment can significantly enhance the accuracy of statistical analyses. Consequently, numerous researchers have dedicated their efforts to address the challenge of constructing optimal designs for nonlinear regression models. Experimental design plays a pivotal role in scientific research domains, including but not limited to biomedicine and pharmacokinetics. Its application in these fields enables researchers to conduct rigorous investigations and yield valuable insights.
Optimal designs are sought using optimality criteria, typically based on the information matrix. Until 1959, research primarily focused on linear models, where the models were linear with respect to the parameters. However, in nonlinear models, the presence of unknown parameters introduced complexities in the design problem, as the optimality criteria depended on these unknown parameters [3, 5, 26]. To address this challenge, researchers proposed various solutions, including local optimal designs [1, 6, 7, 11, 18, 29], sequential optimal designs, minimax optimal designs, Bayesian optimal designs [27, 20-24], and pseudo-Bayesian designs [25]. Chernoff (1953) introduced the concept of local optimality, which involves specifying fixed values for the unknown parameters and optimizing a function of the

*Talker
Email addresses: anita.abdollahi@yahoo.com (Anita Abdollahi Nanvapisheh), h.jafari@razi.ac.ir (habib Jafari), s.khazaei@razi.ac.ir (Soleiman Khazaei)

information matrix to determine the design for these specified parameter values. This approach aimed to overcome the difficulties associated with the dependence of the design problem on unknown parameters in nonlinear models.
The selection of unknown parameters in local designs is typically obtained from previous studies or experiments specifically conducted for this purpose. The effectiveness of local designs heavily relies on the appropriate selection of these parameters. However, a significant challenge arises when the investigated problem lacks robustness in relation to weak parameter estimation. To address this, an alternative approach for local optimal designs involves utilizing a prior distribution for the unknown parameters instead of relying solely on initial guess. In the Bayesian method, the first step is to represent the available information in the form of a probability distribution for the model parameter, known as the prior distribution. A Bayesian optimal design aims to maximize the relevant optimality criterion over this prior distribution. Nevertheless, it is crucial to acknowledge that the selection of the prior distribution within the Bayesian framework can be problematic and may potentially lead to erroneous results. The choice of the prior distribution is subjective, relying on the researcher's beliefs, and it significantly influences the final outcome. Unfortunately, the Bayesian approach lacks a definitive method for selecting the prior distribution. Numerous researchers have investigated the effect of the prior distribution on determining design points in various types of optimal designs. For instance, Chaloner and Lorentz [10], Chaloner and Duncan [8], Burghaus and Dette [4], Chaloner and Vardinelli [9], Pronzato and Walter [28], Mukhopadhyay and Haines [25], Dette and Ngobauer [12, 13], Fedorov [14, 15], and Firth and Hinde [17] have contributed extensively to this field. Chapter 18 of Atkinson et al.'s book [2] provides further reading on this topic. Moreover, in situations where there is insufficient evidence from previous studies on the topic of interest, specifying an appropriate prior distribution becomes challenging. In such cases, subjective or noninformative prior distributions are used, incorporating all available information regarding the uncertainty of the parameter values.For more information, refer to Burghaus and Dette [4].
This research paper presents the introduction of a novel one-parameter model, referred to as the UE distribution, specifically designed for the unit interval in section 2. In Section3, the optimal design for nonlinear models is derived. In the fourth section, the nonparametric Bayesian D-optimal design, Dirichlet process and Polya Urn Scheme are introduced in this section. Finally, Section 5 concludes the paper with some closing remarks.

2. The Unit-Exponential distribution

The exponential distribution is continuous distribution in statistics and probability theory. If $Y \sim Exp(\theta)$, then using the transformation $X = \dfrac{Y}{1+Y}$ we have a new distribution with support on the unit-interval that the CDF and the PDF of the resulting distribution are respectively:

$$F(x \mid \theta) = 1 - Exp(\frac{-\theta x}{1-x});\ 0 \leq x < 1,\ \theta > 0, \tag{1}$$

$$f(x \mid \theta) = \frac{\theta}{(1-x)^2} Exp(\frac{-\theta x}{1-x});\ 0 \leq x < 1,\ \theta > 0. \tag{2}$$

The Hazard Rate Function (HRF) of this distribution is as follows:

$$h(x \mid \theta) = \frac{f(x \mid \theta)}{1 - F(x \mid \theta)} = \frac{\theta}{(1-x)^2};\ 0 \leq x < 1,\ \theta > 0. \tag{3}$$

In the following figure, the PDF and the HRF of this distribution are plotted for different values of the parameter θ. Acording to this figurs, it can be seen that the HRF is increasing in $0 \leq x < 1$,.

3. Optimal Design for Nonlinear Models

In the context of nonlinear experimental design, a common issue arises where the relationship between the response variable y and the independent variable x is given by the equation $y = \eta(x, \boldsymbol{\theta}) + \epsilon$ where $x \in \chi \subseteq$ R and y is a response variable and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is the unknown parameter vector and ϵ is a normally distributed residual value with mean 0 and known variance $\sigma^2 > 0$. For simplicity, we assume $\sigma^2 = 1$ in this problem. If $\eta(x, \boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$ then, the information matrix at a given point x can be represented as follows:

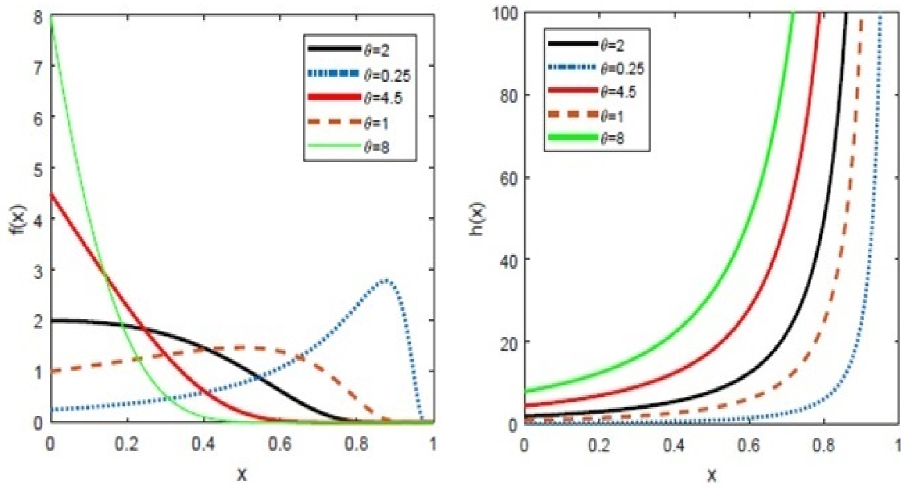


Fig. 1. Plot of density function (left) and hrf (right)

$$I(\xi,\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}}\eta(x,\boldsymbol{\theta})\frac{\partial}{\partial\boldsymbol{\theta}^T}\eta(x,\boldsymbol{\theta}). \tag{4}$$

There exist several optimality criteria used to obtain the optimal design, including D-optimality and A-optimality. These criteria are functions of the information matrix and can be expressed as follows:

$$\Psi_D(\xi,\boldsymbol{\theta}) = -\log(det(M(\xi,\boldsymbol{\theta})))\ ,\ \Psi_A(\xi,\boldsymbol{\theta}) = tr(M^{-1}(\xi;\boldsymbol{\theta})),$$

where ξ denotes a design with two components; the first component represents specific values from the design space χ and the second component corresponds to the weights assigned to these values, so that design ξ can be defined as follows:

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \dots & x_\ell \\ w_1 & w_2 & \dots & w_\ell \end{array} \right\} \in \boldsymbol{\Xi}, \tag{5}$$

where

$\boldsymbol{\Xi}$=$\{\xi \mid 0 \leq w_j \leq 1\,; \sum_{j=1}^{\ell} w_j = 1\,,\, x \in \boldsymbol{\chi}\}$, [25].

When considering a discrete probability measure ξ with finite support, the information function of ξ can be expressed as follows [3]:

$$M(\xi,\boldsymbol{\theta}) = \sum_{j=1}^{\ell} w_j I(x_j,\boldsymbol{\theta}). \tag{6}$$

Because of the dependence of the information matrix $M(\xi,\boldsymbol{\theta})$ to the unknown parameter $\boldsymbol{\theta}$, one approach to address this issue is to employ the Bayesian method and incorporate a prior distribution for the parameter vector. The Bayesian D-optimality criterion can be formulated as follows:

$$\Psi_\Pi(\xi) = E(\psi(\xi;\boldsymbol{\theta})) = \int_\Theta \psi(\xi;\boldsymbol{\theta})d\Pi(\boldsymbol{\theta}) = \int_\Theta -\log(det(M(\xi,\boldsymbol{\theta})))d\Pi(\boldsymbol{\theta}), \tag{7}$$

where Π represents the prior distribution for $\boldsymbol{\theta}$ and the Bayesian D-optimal design is attained by minimizing (7). According to Dette and Neugebauer [12], in the general case of optimal designs which can include designs with two and more points, if the support of the prior distribution has n points, then the maximum number of Bayesian optimal

design points is given by $n \quad \frac{p(p+1)}{2}$. Hence, in the specific scenario of nonlinear models with one parameter ($p = 1$), this implies that the support of the Bayesian optimal design does not contain more points than the support of the prior distribution.
In certain situations, specifying a prior distribution on the parameter space $\boldsymbol{\Theta}$ can be challenging for the experimenter. In such cases, an alternative approach is to consider an unknown prior distribution Π for the parameter $\boldsymbol{\theta}$. In this condition, Π is treated as a parameter itself. Consequently, equation (7) becomes a random functional, and it becomes necessary to determine its distribution or approximation. From a Bayesian perspective, we construct a prior distribution on the space of all distribution functions to address this issue. Ferguson (1973) introduced the concept of the Dirichlet process in this context, that in the section 4.1.1 an overview of the Dirichlet process will be provided.

4. Nonparametric Bayesian D-optimal design

Now suppose we have the following regression model:

$$E(y|x) = \eta(x, \boldsymbol{\theta}) = \frac{\theta}{(1-x)^2} \exp(\frac{-\theta x}{1-x}), 0 \leq x < 1, \theta > 0. \tag{8}$$

In this section, we introduce the nonparametric Bayesian optimal designl. In the nonparametric Bayesian framework, it is assumed that $\theta \mid P \sim P$, where P is a random probability distribution and $P \sim \Pi$. General method of construction a random measure is to start with the stochastic processes. Ferguson (1973) formulated the requirements which must be imposed on a prior distribution and proposed a class of prior distributions, named Dirichlet processes. One of the main argument in using the Dirichlet distribution in practical applications is based on the fact that this distribution is a good approximation of many parametric probability distributions. Below we give the definition of the Dirichlet process.

4.1. Dirichlet Process (DP)

To have a random distribution G distributed according to a Dirichlet process (DP), its marginal distributions must follow a Dirichlet distribution. Specically, let H be a distribution over $\boldsymbol{\Theta}$ and α be a positive real number. For any finite measurable partition $A_1, A_2, ..., A_r$ of $\boldsymbol{\Theta}$ the vector $(G(A_1), G(A_2), ..., G(A_r))$ is random since G is random. We say G is Dirichlet process distributed with base distribution H and concentration parameter α, written $G \sim$ DP(α,H), if the following conditions hold:

$$(G(A_1), G(A_2), ..., G(A_r)) \sim Dir(\alpha H(A_1), ..., \alpha H(A_r)), \tag{9}$$

for every finite measurable partition $A_1, A_2, ..., A_r$ of $\boldsymbol{\Theta}$. The parameters of H and α play intuitive roles in the definition of the DP. The base distribution H represents the mean of the Dirichlet process, such that for any measurable set $A \subset \boldsymbol{\Theta}$ we have $E[\text{G(A)}] = H(A)$. On the other hand, the concentration parameter α can be viewed as an inverse variance: $V[G(A)]$=$H(A)(1 - H(A))$/(α +1). The larger α is, the smaller the variance, and the DP will concentrate more of its mass around the mean. The concentration parameter is also referred the strength parameter, referring to the strength of the prior when using the DP as a nonparametric prior in Bayesian nonparametric modelsl, It can be interpreted as the amount of mass or sample size associated with the observations. It is worth noting that α and H only appear as their product in the definition of the Dirichlet process (equation 4.2). Consequently, some authors treat $\tilde{\text{H}}$=α H, as the same as the single (positive measure) parameter of the DP, writing DP($\tilde{\text{H}}$) instead of DP(α,H). This parametrization can be notationally convenient, but loses the distinct roles α and H play in describing the DP.
As the concentration parameter α increases, the mass of the DP becomes more concentrated around its mean. Consequently, when α approaches infinity ($\alpha \to \infty$), $G(A)$ approaches $H(A)$ for any measurable set A, indicating weak or pointwise convergence of G to H. However, it's important to note that this does not imply a direct convergence of G to H as a whole. In fact, as we will explore later, samples drawn from a DP will typically be discrete distributions with probability one, even if the base distribution H is smooth. Therefore, G and H may not be absolutely continuous with respect to each other. Despite this, some authors still utilize the DP as a nonparametric extension of a parametric model represented by H. However, if the desire is to maintain smoothness, it is possible to extend the DP by convolving G with kernels, resulting in a random distribution with a density function.

An alternative definition of the Dirichlet process is proposed by Ferguson [16] that defined a random probability measure which is a Dirichlet process on (**Θ**, B(**Θ**)), as:

$$P(.) = \sum_{i=1}^{\infty} p_i \delta_{\theta_i}(.), \tag{10}$$

where θ_i $(i > 1)$ be a sequence of $i.i.d.$ random variables with common distribution Q, δ_{θ_i} represents a probability measure that is degenerate at θ where δ_{θ_i}=1 if $\theta_i \in A$ and 0 otherwise, and $p_i^{,}$ s are the random weights satisfying p_i>0 and $\sum_{i=1}^{\infty} p_i$=1. The random distribution P is discrete with probability one. Several authors have proposed alternative series representations of the Dirichlet process. Bondesson [6], Sethuraman [30], and Zarepour and Al Labadi [31] are among those who have contributed to this area. A method of producing samples from the Dirichlet process is to use the Polya urn process that in the upcoming section, we will discuss about it. Then the nonparametric Bayesian D-optimal design for the UE model is discussed.

4.2. Polya Urn Scheme

Polya Urn Scheme was used by Blackwell and McQueen (1973) to demonstrate the existence of the Dirichlet process. The method of producing a sample of the Dirichlet process is to use a Polya Urn Scheme [19]. Consider a Polya urn with $a(\chi)$ balls of which $a(i)$ are of color i ; $i = 1, 2, ..., k$.[For the moment assume that $a(i)$ are whole numbers or 0]. Draw balls at random from the urn, replacing each ball drawn by two balls of the same color. Let $X_i = j$ if the i th ball is of color j. Then:

$$P(X_1 = j) = \frac{a(j)}{a(\chi)}, \tag{11}$$

$$P(X_2 = j \mid X_1) = \frac{a(j) + \delta_{X_1}(j)}{a(\chi) + 1}, \tag{12}$$

and in general

$$P(X_{n+1} = j \mid X_1, X_2, ..., X_n) = \frac{a(j) + \sum_1^n \delta_{X_i}(j)}{a(\chi) + n}, \tag{13}$$

That n is the number of extracted balls and $\delta_{X_i}(j)$ is equal to one if $X_i = j$, otherwise it is equal to zero.

4.3. Nonparametric Bayesian D-optimal design for UE model

Now let's consider the regression model (8), Therefore, the Bayesian D-optimality criterion, denoted as $\Psi_\Pi(\xi)$ can be expressed as follows:

$$\Psi_\Pi(\xi) = E(\psi(\xi;\boldsymbol{\theta})) = \int_\Theta \psi(\xi;\theta) d\Pi(\boldsymbol{\theta}) = \int_\Theta -\log(\sum_{j=1}^{\ell} w_j [\exp(\frac{-\theta x_j}{1-x_j})(\frac{1}{(1-x_j)^2} - \frac{\theta x_j}{(1-x_j)^3})]^2) d\Pi(\theta) \tag{14}$$

where Π is the prior distribution for θ. The Bayesian D-optimal design is attained by minimizing equation (14). In the nonparametric Bayesian framework, we consider $P \sim$ DP(α,P_0) and its collective representation as $P(.) = \sum_{i=1}^{\infty} p_i\ \delta_{\theta_i}(.)$. In this context, the optimality criterion can be expressed as follows:

$$\Psi_\Pi(\xi) = \sum_{i=1}^{\infty} p_i (-\log(\sum_{j=1}^{\ell} w_j [\exp(\frac{-\theta_i x_j}{1-x_j})(\frac{1}{(1-x_j)^2} - \frac{\theta_i x_j}{(1-x_j)^3})]^2)). \tag{15}$$

Chernoff [7] demonstrated that when searching for a local optimal design, there exists an optimal design where all the mass is concentrated at a single point within the design's support. Caratheodory's theorem also confirms the existence of a one-point optimal design. However, when employing the Bayesian optimality criterion, a more complex situation arises. Brice and Dette showed that with a uniform prior distribution, as the support of the prior distribution

increases, the number of optimal design points for the single-parameter model also increases. Challoner suggested that if the researcher aims to obtain a one-point optimal design, it is advisable to consider a small support for the uniform prior distribution. The same principle applies to nonparametric Bayesian designs. In this case, assuming a uniform distribution over the interval [1, B] as the basic distribution, the one-point optimal design can be achieved.
Equation (7) is a stochastic function of the Dirichlet process. According to Ferguson's definition of the Dirichlet process, the calculation (8) is not easily possible, so to solve this problem in obtaining the optimal nonparametric Bayesian criterion, methods such as the stick breaking process are used to approximate this criterion. Another method has been presented by Zarepour and Ellabadi [31] whose simulation speed and accuracy is much higher than the stick breaking process.
Since the weights produced by the stick breaking process don't follow a decreasing trend, therefore, the Dirichlet process can be simulated in a way where the weights are produced in a decreasing manner. The reason for this is that the speed of reaching the cutting point increases. Zarepour and Ellabadi presented a finite collective representation of the Dirichlet process in order to generate data from the Dirichlet process, which almost certainly converges to the Ferguson collective representation that we present below, and the weights produced from this method are uniformly descending, while the weights produced by the stick breaking method are randomly descending. Ferguson showed that the Dirichlet process with parameters (α, P_0) can be presented using the following series representation:

$$P_n^{Ferg.}(.)=\sum_{i=1}^{\infty}\frac{N^{-1}(\Gamma_i)}{\sum_{i=1}^{\infty}N^{-1}(\Gamma_i)}\delta_{\theta_i}(.),$$

where

$$N(x)=\alpha\int_x^{\infty}\frac{\exp(-t)}{t}dt, x>0. \tag{16}$$

is the Levy measure of a Gamma(α, 1) random variable and $\delta_{\theta_i}(.)$ denotes the Dirac measure. Now, in this section, we present the finite sum representation of the Dirichlet process presented by Zarepour and Ellabadi [31]. Let X_n be a random variable with distribution Gamma($\frac{\alpha}{n}$,1) and with survival and quantile function, respectively as follows:

$$G_n(x)=P(X_n>x)=\int_x^{\infty}\frac{1}{\Gamma(\frac{\alpha}{n})}\exp(-t)\,t^{\frac{\alpha}{n}-1}dt,$$
$$G_n^{-1}(y)=inf\{x:G_n(x)\leq y\}\text{ , }0<y<1.$$

According to the dominated convergence theorem, n $\to\infty$, we have:

$$nG_n(x)\to N(x).$$

Notice that the left hand side of the above quantitative is a sequence of monotone functions converging to a monotone function. We have:

$$G_n^{-1}(x/n)\to N^{-1}(y).$$

Zarepour and Ellabadi showed that for each $E_i{\sim}Exp(1), i=1,2,...,n$ and for each $\theta_i\sim P_0$, that $\Gamma_i{=}E_1{+}E_2{+}...{+}E_i$ the obtained approximation almost certainly converges to Ferguson's collective representation; that's mean:

$$P_n^{New.}(.)=\sum_{i=1}^{n}\frac{G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}{\sum_{i=1}^{n}G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}\delta_{\theta_i}(.)\to P^{Ferg.}(.)=\sum_{i=1}^{\infty}\frac{N^{-1}(\Gamma_i)}{\sum_{i=1}^{\infty}N^{-1}(\Gamma_i)}\delta_{\theta_i}(.), \tag{17}$$

where n is as follows:

$$n = inf\{m : \frac{G_m^{-1}(\frac{\Gamma_m}{\Gamma_{m+1}})}{\sum_{i=1}^{m} G_m^{-1}(\frac{\Gamma_i}{\Gamma_{m+1}})} < \epsilon\}. \tag{18}$$

It is important to emphasize that unlike in the previously discussed truncation approximations, the weights:

$$p_i = \frac{G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}{\sum_{i=1}^{n} G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})}, \tag{19}$$

decrease monotonically for any fixed positive integer n, that leads to the fact that the speed of simulation and its accuracy is much higher than the stick breaking process. In the following, a nonparametric Bayesian optimal design is obtained for different selections of Dirichlet process parameters. For this purpose, at first we generate p_i from (19). We obtain n from (18) and generate $E_i \sim Exp(1)$, $i = 1, 2, ..., n$, and let Γ_i=$E_1 + E_2 + ... + E_i$. We calculate $G_n^{-1}(\frac{\Gamma_i}{\Gamma_{n+1}})$, $i = 1, 2, ...,n$ from the equation $G_n^{-1}(y)$= $inf\{x : G_n(x) < y\}$, and generate θ_i from Base measure P_0. Finally, we evaluate the functional:

$$\Psi_P(\xi) = \sum_{i=1}^{\infty} p_i\ [-\log(det(M(\xi, \theta_i)))],$$

and obtain ξ^* from the following equation:

$$\xi^*= \arg min\ \Psi_P(\xi).$$

Now, in this section we consider Polya Urn Scheme as the base measure in DP. We get the results by using a nonlinear optimization programing with R package $Rsolnp$. For better understanding of the effect of the α parameter, we tabulate the results for four different values of α=1, 5, 10, 50, in Tables 1-4. We also fixed ϵ=10^{-10}. Without loss of generality, we consider a bounded design space χ=[0, 1].

Tables 1-4 represent the results when the concentration parameter and uncertainty in the base measure increase. According to the results, when the value of α increases, the support points in two points design do not significantly change. The weight of minimum point increases rapidly and the smallest point will have the most weight that this weight almost increases or remains fixed by increasing the concentration parameter. Also for three points design, minimum support point has the greatest weight. In addition, in the range under investigation, the results show that we don't have a three - point design for $\mu = 5, \sigma = 2$, and in fact, it converts to the design by less points. This observation is more clear for larger concentration parameter. But, by increasing the parameter space, optimal two and three - point designs are obtained.

Table 1. Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=1. First row: support points; second row: weights.

$Parameters$	$Design$	$Two\ points$		$Three\ points$		
$\mu = 5, \sigma = 2$	x	0.00039	0.23748	$-$	$-$	$-$
	w	0.99738	0.00262	$-$	$-$	$-$
$\mu = 50, \sigma = 30$	x	0.03358	0.18865	0.038036	0.18630	0.29563
	w	0.97085	0.02915	0.949231	0.050768	0.0000001
$\mu = 150, \sigma = 90$	x	0.01520	0.19838	0.01595	0.19625	0.29908
	w	0.99393	0.00607	0.98983	0.00813	0.00204
$\mu = 1000, \sigma = 500$	x	0.002302	0.19991	0.00275	0.20004	0.29995
	w	0.999998	0.000002	0.999999	0.0000006	0.0000005

Now, if we assume the mean of the base distribution to be constant and increase the variance, it can be seen that in the two-point designs, the smallest point has the most weight. The results related to this case has been presented in the table 5.

Table 2. Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=5. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00068	0.22791	−	−	−
	w	0.99734	0.00266	−	−	−
$\mu = 50, \sigma = 30$	x	0.03373	0.18516	0.03688	0.17411	0.28843
	w	0.96606	0.03394	0.91814	0.07302	0.00884
$\mu = 150, \sigma = 90$	x	0.01457	0.19748	0.14979	0.19533	0.29895
	w	0.99192	0.00808	0.9979605	0.0020304	0.0000001
$\mu = 1000, \sigma = 500$	x	0.002169	0.19994	0.00245	0.19965	0.29999
	w	0.999999	0.000001	0.9999908	0.0000001	0.0000001

Table 3. Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=10. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00031	0.23268	−	−	−
	w	0.99735	0.00265	−	−	−
$\mu = 50, \sigma = 30$	x	0.03269	0.17711	0.03922	0.17145	0.29027
	w	0.94385	0.05615	0.92004	0.07552	0.00444
$\mu = 150, \sigma = 90$	x	0.01361	0.19819	0.01601	0.19434	0.29891
	w	0.99798	0.00202	0.9879202	0.0120707	0.0000001
$\mu = 1000, \sigma = 500$	x	0.00226	0.20000	0.00251	0.19948	0.29997
	w	0.999992	0.0000008	0.9979901	0.0020004	0.0000005

Table 4. Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=50. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 5, \sigma = 2$	x	0.00119	0.21958	−	−	−
	w	0.98904	0.01096	−	−	−
$\mu = 50, \sigma = 30$	x	0.03164	0.18099	0.03504	0.16948	0.29022
	w	0.96095	0.03905	0.91685	0.07641	0.00674
$\mu = 150, \sigma = 90$	x	0.01236	0.19683	0.14969	0.19641	0.29876
	w	0.99596	0.00404	0.9979801	0.0020107	0.0000002
$\mu = 1000, \sigma = 500$	x	0.00221	0.19988	0.002462	0.19979	0.29999
	w	0.9999999	0.00000001	0.9999908	0.0000006	0.0000006

Table 5. Nonparametric Bayesian D-optimal designs with truncated normal base distribution and concentration parameter when α=1. First row: support points; second row: weights.

Parameters	*Design*	*Two points*		*Three points*		
$\mu = 50, \sigma = 30$	x	0.03342	0.18889	0.03342	0.19076	0.29333
	w	0.97926	0.02074	0.96545	0.03227	0.00228
$\mu = 50, \sigma = 90$	x	0.02209	0.19534	0.02491	0.19477	0.29463
	w	0.98720	0.01280	0.98060	0.1939609	0.0000001
$\mu = 50, \sigma = 500$	x	0.00567	0.20004	0.008246	0.19944	0.29863
	w	0.99797	0.00203	0.9979821	0.0020104	0.0000005

5. Concluding Remarks And Future Works

Nonlinear regression models are widely used in various scientific fields, and the Bayesian method is commonly employed to obtain optimal designs in such models. However, one of the challenges in the Bayesian framework is the subjective selection of the prior distribution, which can potentially lead to incorrect results. The choice of the prior distribution is often based on the researcher's beliefs, and it strongly influences the final outcome. Unfortunately, the Bayesian approach lacks a systematic method for selecting the prior distribution. To overcome these limitations and reduce reliance on restrictive parametric assumptions, nonparametric Bayesian methods are pursued. In this study, we consider the prior distribution as an unknown parameter and utilize the Dirichlet process to derive nonparametric

Bayesian D-optimal designs. Specifically, we focus on a nonlinear model with one parameter, namely the Unit-Exponential distribution. We investigate the Bayesian D-optimal design for the unit exponential regression model (equation 8) using a truncated normal prior distribution, examining various parameter values. By adopting a non-parametric Bayesian approach and utilizing the Dirichlet process, we aim to address the challenges associated with selecting the prior distribution in Bayesian optimal design construction. This allows us to account for uncertainty and mitigate the impact of restrictive parametric assumptions, providing more flexible and robust designs for nonlinear regression models.

In this study, we focus on utilizing the Polya Urn Scheme as the base distribution in the Dirichlet process. To better understand the influence of the concentration parameter α, we present the results in tables for four different values of α=1, 5, 10, 50. These tables provide valuable insights into the nonparametric Bayesian optimal designs, showcasing the distribution of weights and support points. By analyzing the results for different values of α, we can better understand the impact of this parameter on the design outcomes. This approach allows us to explore and evaluate the performance of the nonparametric Bayesian optimal designs under varying levels of concentration parameter α.

In the investigated range, the results reveal interesting findings. For small parameter values, there are no three-point designs observed. However, By increasing uncertainty in the base measure, another optimal point is obtained with a very small weight, resulting in a design where the smallest point has the highest weight.

Moreover, as the uncertainty in the base measure and the concentration parameter in the Dirichlet process increase, the support points in the two-point designs do not undergo significant changes. The weight of the smallest point increases rapidly, and it becomes the point with the highest weight. This weight tends to either increase or remain relatively stable with an increase in the concentration parameter.

It is important to note that this approach can be applied to other optimality criteria and various models with two or more parameters. For example, nonparametric Bayesian optimal designs using the A- or E-optimality criterion for the nonlinear model discussed in this paper, along with a Dirichlet process prior, hold potential for further research. We hope to report new results in this area in the near future.

References

[1] Aminnejad, M., and H. Jafari. *Bayesian A and D-optimal designs for gamma regression model with inverse link function,*. Communications in Statistics-Simulation and Computation 46 (10):8166–89. doi:10.1080/03610918.2016.1271888, 2017.

[2] Atkinson, A. C., A. N. Donev, and R. D. Tobias. *Optimum experimental design, with SAS,*. Oxford, UK: Oxford University Press, 2007.

[3] Blackwell, D. *Discreteness of Ferguson selections,*. The Annals of Statistics, 1(2), 356-358, 1973.

[4] Burghaus, I., and H. Dette. *Optimal designs for nonlinear regression models with respect to non-informative priors,*. Journal of Statistical Planning and Inference 154:12–25. doi:10.1016/ j.jspi.2014.05.009, 2014.

[5] Burkner PC, Schwabe R, Holling H. *Optimal designs for the generalized partial credit model,*. British Journal of Mathematical and Statistical Psychology. doi:10.1111/bmsp. 12148, 2019.

[6] Bondesson, Lennart. *On simulation from in nitely divisible distributions,*. Advances in Applied Probability 855,869, 1982.

[7] Chernoff, H. *Locally optimal designs for estimating parameters,*. The Annals of Mathematical Statistics 24 (4):586–602. doi:10.1214/aoms/1177728915, 1953.

[8] Chaloner, K. M., and Duncan, G. T. *Assessment of a beta prior distribution:PM elicitation,*. The Statistician, 174-180, 1983.

[9] Chaloner, K. and Verdinelli, I. *Bayesian experimental design: a review,*. Statistical Science, 10, 273-304, 1995.

[10] Chaloner, K. and Larntz, K. Optimal Bayesian design applied to logistic regression experiments. Journal of Statistical Planning and Inference, 21, 191-208, 1989.

[11] Dette, H., Melas, V. B., Wong, W.K. *Locally D-optimal designs for exponential regression Statistica Sinica,*. 16, 789–803, 2006.

[12] Dette, H. and Neugebauer, H. M. *Bayesian optimal one point designs for one parameter nonlinear models,*. Journal of Statistical Planning and Inference, 52(1), 17-31, 1996.

[13] Dette, H. and Neugebauer, H. M. *Bayesian D-optimal designs for exponential regression models,*. Journal of Statistical Planning and Inference, 60(2), 331-349, 1997.

[14] Fedorov V, Hackl P. *Model-Oriented Design of Experiments,*. volume 125. Springer Science and Business Media. doi:10.1007/978-1-4612-0703-0, 2012.

[15] Fedorov VV, Leonov SL. *Optimal design for nonlinear response models,*. CRC Press. bability Letters 82 (5):916–24. doi:10.1016/j.spl.2012.01.020, 2013.

[16] Ferguson, T. S. *A bayesian analysis of some nonparametric problems,*. The Annals of Statistics 1 (2):209–30. doi:10.1214/aos/1176342360, 1973.

[17] Firth D, Hinde J. *On Bayesian D-optimum Design Criteria and the Equivalence Theorem in Non-linear Models,*. Journal of the Royal Statistical Society B, 59(4), 793–797. doi:10.1111/1467-9868.00096, 1997.

[18] Ford, I., Torsney, B. and Wu, C. F. J. *The use of a canonical form in theconstruction of locally optimal designs for non-linear problems,*.J. Roy. Statist. Soc. (Ser.) B, 54, 569-583, 1992.

[19] Ghosh, J. and Ramamoorthi, R. *Bayesian Nonparametrics,*. New York: Springer-Verlag, 2003.

[20] Goudarzi, M., H. Jafari,and S. Khazaei. *Nonparametric Bayesian optimal designs for exponential regression model,*. Communications in Statistics-Simulation and Computation.doi:10.1080/03610918.2019.1593454, 2019.

[21] Grashoff U, Holling H, Schwabe R *Optimal designs for the Rasch model,*. Psychometrika, 77(4), 710–723. doi:10.1007/s11336-012-9276-2, 2012.

[22] Kiefer, J. *Optimum experimental designs,*. J.R. Statist. Soc. B. 21, 272-319, 1959.

[23] Kiefer, J., and Wolfowitz, J. *Optimum designs in regression problems,*. The Annals of Mathematical Statistics, 30(2), 271-294, 1959.

[24] Kiefer, J. *General equivalence theory for optimum designs (approximate theory),*. The Annals of Statistics 2 (5):849–79. doi:10.1214/aos/1176342810, 1974.

[25] Mukhopadhyay, S., and Haines, L. M. *Bayesian D-optimal designs for the exponential growth model,*. Journal of Statistical Planning and Inference, 44(3), 385-397, 1995.

[26] Niaparast, M. *Optimal designs for mixed effects poisson regression models*. PhD thesis, Faculty of Mathematics, Magdeborg, 2010.

[27] Parsamaram, P., and H. Jafari. *Bayesian D-optimal Design for the logistic regression model with exponential distribution for random intercept,*. Journal of Statistical Computation and Simulation 86:1856–68. doi:10.1080/00949655.2015.1087525, 2016.

[28] Pronzato, L, E. walter. *Robust experiment design via stochastic approximation,*. mathematical BIOSCIENCES. doi.org/10.1016/0025-5564(85)90068-9, 1985.

[29] Rodriguez-Torreblanca C. and Rodríguez-Díaz J. M. *Locally D- and c-optimal designs for Poisson and negative binomial regression models,*. Metrika, vol. 66, pp 161–172, 2007.

[30] Sethuraman, J. *A Constructive De nition of Dirichlet Priors,*. Statis-tica Sinica 4, 639,650, 1994.

[31] Zarepour, Mahmoud and Al Labadi, Luai. *On a rapid simulation of the Dirichlet process,*. Statistics and Probability Letters 82.5, 916,924, 2012.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–103, pp. 250–252

Assessments of the general and specific adaptive abelites under different conditions for stability analysis

Naser Sabaghnia[a,*], Mohsen Janmohammadi[b]

[a]*Department of Plant Genetics and Production, University of Maragheh, Maragheh, Iran.*
[b]*Department of Plant Genetics and Production, University of Maragheh, Maragheh, Iran.*

Article Info

Keywords:
stability
yield
variance

2020 MSC:

Abstract

A statistical model for statistical analysis has been proposed for biometricians based on testing the n varieties in m environments which lets to explore the general and specific adaptive ability of varieties. The stability characterizes of statistics, are compared across environments for their ability to differentiate the varieties. It was shown that the both general and specific adaptive ability have their variances give more precise information about the stability and differentiative ability of environments. A method for estimation of selection value of a genotype as relative stability (RS) based on the general and specific adaptive ability of varieties is proposed in selection of highly productive and stable ones. Selection for specific adaptive ability taking into stability can be recommended, because the first direction of selection can transform the structure of crop. Predictably given environmental conditions, it is advisable to carry out selection for specific adaptive ability. However, in unpredictable cases, selection of general adaptive ability cause to achieve stability and aid biometricians.

1. Introduction

Multi-environment Attention of biometricians is involved in the development of models for assessing sustainable stability, general adaptive ability as well specific adaptive ability of varieties in order to optimize these important properties of crops. There are various approaches to assessing stability of cultivars and some overviews of these models are presented in the work of Cheshkova et al. [1]. The terms general and specific adaptation have become firmly established in the literature while term stability reflecting the general reaction of the variety throughout set of locations and years. The study of varieties in different environments can also provide information about the environments as backgrounds for selection. Kilchevskiy and Hotyleva [2] proposed a method of genetic analysis based on the testing varieties in different environments and allowing biometricians to identify a general and the specific adaptive ability of varieties, their stability and also compare performance according to their ability to differentiate varieties. In this case, it is possible to evaluate the breeding value of the variety and carry out selection according to adaptive ability depending on the selected selection tasks. The method also makes it possible to obtain phenotypic variances of the

*Talker
Email addresses: sabaghnia@maragheh.ac.ir; sabaghnia@yahoo.com (Naser Sabaghnia), mohsen_janmohammadi@yahoo.com (Mohsen Janmohammadi)

studied population on the basis of general and specific adaptive capacity for the purpose of comparing populations and choosing methods for genetic improvement.

2. Methods

The adaptive ability modifies the ability of variety under keeping its characteristic phenotypic expression of a trait under certain environmental conditions. General adaptive capacity of the variety as defined GAA (general adaptive ability) characterizes the average value of a trait under different conditions environment, while the specific adaptive ability (SAA) is deviation from GAA in a specific environment. The developed method for assessing GAA and SAA is based on a test population of i varieties in j environments, and the number of repetitions is equal to k. Then,

$$x_{ijk} = \mu + \nu_i + e_j + \nu e_{ij} + \epsilon_{ijk}, \tag{1}$$

where x_{ijk} is the phenotypic value of the variety i grown in the j-th environment in the k-th repetition, and μ is the general mean among the entire set of phenotypic values, ν_i is the effect of variety i; e_i, is the effect of the j-th environment; νe_{ij} is the effect of the interaction of variety i with the environment j; ϵ_{ijk}, residual of formula, caught by random reasons and assigned to error. To establish the significance of the contributions of varieties, environments and interaction actions between them in the phenotypic variability of the population, two-factor analysis of variance is performed. Interpretation of the results of variance analysis depends on the reprovisions regarding the material being studied. If varieties and environments are non-random samples, we consider their effects to be fixed data and can obtain information about specific varieties and mean. In this case, the first option is used, in which all effects except independently varying environments are considered constant values, and normally distributed with zero mean and variance equal to one is supposed. To establish the significance of the effects of varieties, environments and interactions, the p-criterion is used. At the same time, the corresponding averages the squares are compared with the mean square of random deviations error. Let us accept the assumptions of the first option, then, according to definition, the effect the general adaptive capacity of the i-th variety of GAA is equal to vi. Deviation from the sum $\nu + e$, will show the effect of a specific adaptive ability of variety in A-that environment - SAS. This effect consists of the linear (the effect of the j-th environment) and the nonlinear part (the effect of interaction effect νe_{ij}). To determine the effects of GAA and SAA, we calculate the average values of the phenotypic in terms of repetition. Then the results of testing i varieties in j environment, can be written in the form of a two-way table variety $\times$ environment $\times$ repetition. The effects of GAA and SAA are calculated using the following formulas:

$$x_{ij} = \mu + \nu_i + e_j + \nu e_{ij}, \tag{2}$$
$$x_{ij} = \mu + GAA_i + SAA_{ij}. \tag{3}$$

To obtain the standard error of the differences it is necessary to take the square root of the corresponding variation. A number of varieties in terms of general adaptive ability can be determined lead by comparing the GAA. As a measure of the stability of each variety, it is proposed to use this formula:

$$\sigma^2_{SAA} = \frac{1}{m-1}\sum_k (\nu_i + \nu e_{ij})^2 - \frac{m-1}{m}\sigma^2. \tag{4}$$

The concept of relative stability (RS) of the variety can be calculated as:

$$RS = \frac{\sigma_{SAA_i}}{\mu + GAA_i} \times 100. \tag{5}$$

This indicator will allow you to compare the results of experiments, carried out with a different set of crops, varieties, environments and studied traits. Essentially, relative stability is similar to the coefficient of variation in the case in a number of environments.

3. Findings

There are big disagreements in the selection interpretation of the evaluation of stability of varieties. Thus, the optimal variety based on regression analysis, is one with a high average value of the trait, to effectively regression rate close to unity and the smallest deviations from regression line. However, Olivoto et al. [3] showed that as a result of selection it is possible to achieve any combination of the average value of the characteristic and the environment index of sensitivity. optimal variety that, having a high general adaptive capacity, gives the greatest harvest in favorable environments and provides maximum stability. The final assessment of the breeding material will depend on the adaptive selection. The following main directions are possible adaptive selection. Under selection of varieties for SAA in a certain environment, selection is desirable only in controlled environmental conditions (closed ground. irrigated lands, etc.). At the same time, there is a danger of depletion of the basis of general fitness and decreased stability variety. Under selection for GAA of varieties for a number of environments, the average phenotype in all environments and, there is a maximum increase in the trait compared to selection in the favorable or unfavorable environments and average environmental sensitivity activity, the latter is not controlled during the selection process.

4. Conclusion

Finally, selection for SAA taking into account stability can be recommended, because the first direction of adaptive selection which transforms the structure and functions of organisms, the second leads to stabilization of morphogenesis. Predictably given environmental conditions, it is advisable to carry out selection for specific adaptive ability, in unpredictable cases, to select for GAA taking into account volume of stability. Experimental verification of comparative effectiveness is necessary because methods of adaptive selection on various objects with depended to genetic analysis of the obtained material.

References

[1] A. F. Cheshkova, P. I. Stepochkin, A. F. Aleynikov, I. G. Grebennikova, V. I. Ponomarenko. A comparison of statistical methods for assessing winter wheat grain yield stability. Vavilov Journal of Genetics and Breeding, 24(3): 267-274, 2020.

[2] A.V. Kilchevskiy, L.V. Hotyleva. Methods of assessments of the adaptive abilities and stability of genotypes differentiating the abilities of environment. Messages 1. Justification of the method. Genetics, 21:1481 –1490, 1985.

[3] T. Olivoto, A. D. Lúcio, J. A. da Silva, V. S. Marchioro, V. Q. de Souza, E. Jost. Mean performance and stability in multi-environment trials I: combining features of AMMI and BLUP techniques. Agronomy Journal, 111(6), 2949-2960, 2019.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–105, pp. 253–258

The optimal method selection for considering the dependence degree determining of students examinations

Hakim Bekrizadeh[a,*]

[a]*Department of Statistics, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
Canonical correlation analysis
Neural network
Kernel method
Gaussian-kernel function
Hyperbolic tangent function

Abstract

Canonical correlation analysis is a technique applied for extracting common properties of multivariate data pair to find their linear transformation in a way that coefficient correlation becomes maximum. There are also other techniques including non-linear neural network and kernel canonical correlation analysis method that are in fact generalization of canonical correlation analysis. The whole purpose of these techniques is determination of the correlation or relation among several variable sets. In this paper, canonical correlation analysis methods were considered and an optimal way for determining the correlation degree of student's examinations was selected.

1. Introduction

Since regression analysis is not reliable for very great dimensions, it is necessary to decrease the spatial dimension in some ways to solve this problem. For this reason the canonical correlation analysis is introduced as a technique to extract common features from a pairs of multivariate data. Therefore canonical correlation analysis is used for decreasing spatial dimension and finds the linear transformation of a pairs of multivariates so that correlation coefficient is maximized (Hotelling, 1936; Bartlett, 1941; Foucat, 1999). The initial objective is to find the maximal correlation between a chosen linear combination of the first set of variables and a chosen linear combination of the second set of variables. Maximizing this method tries to focus on a relation with high dimension which is illustrated between two sets of variables to a number of canonical variable pairs. However, if the purpose is regression, the large values of correlation coefficient are crucially necessary. The reason that correlation coefficients are small can be considered in the following cases:

1. X and Y does not have almost any relation.
2. There is strong nonlinear relation between X and Y.

It is impassible to improve the first case. However, in the second case, we can obtain the relation by some method. One of those methods is to allow the nonlinear transformation.
If the relationship between variables pairs is stated in nonlinear form it is proposed to use the neural network model which approximately optimizes the nonlinear canonical correlation analysis (Lai and Fyfe, 2000). However, this

*Talker
Email address: Bekrizadeh@pnu.ac.ir (Hakim Bekrizadeh)

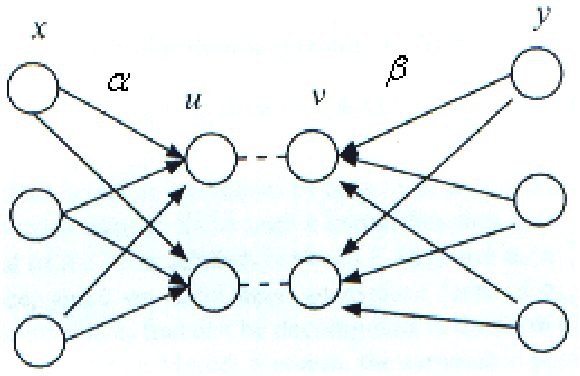


Fig. 1. CCA

model requires a lot of computation time and it also has a lot of local optima. To solve this problem a kernel canonical correlation analysis is used which enable us to decrease voluminous computations and pretend unwanted local optima. A kernel model analyses the data to its basic elements and keeps them in the memory (Akaho, 2001). In this way it learns those data. The whole purpose of these techniques is to determine the correlation or relation among several variable sets.
In this paper, canonical correlation analysis methods were considered and an optimal way for determining the correlation degree of student's examinations was selected.

2. Analysis methods

2.1. Canonical Correlation Analysis (CCA)

CCA has been proposed by Hotelling(1936). Suppose there is a pair's multi-variate and CCA finds a pairs of linear transformation such that the correlation coefficient between extracted features is maximized (Fig.1).
Suppose two linear combinations of and . Correlation between and is as following:

$$\rho\left(v,u\right)=\frac{E\left(v_u\right)}{\sqrt{var(u)var(v)}}=\frac{\alpha'\Sigma_{xy}\beta}{\sqrt{\left(\alpha'\Sigma_{xx}\alpha\right)\left(\beta'\Sigma_{yy}\beta\right)}} \tag{1}$$

where, $\sum_{xx}=E\left(x-\mu\right)\left(x-\mu\right)'$, $\sum_{\text{yy}}=E\left(y-\eta\right)\left(y-\eta\right)'$, and $\sum_{xy}=E\left(x-\mu\right)\left(y-\eta\right)'$. We have to further assume

$$\begin{cases} var(u)=\alpha'\Sigma_{xx}\alpha=1 \\ var(v)=\beta'\Sigma_{yy}\beta=1 \end{cases}. \tag{2}$$

To reduce the freedom of scaling of u and v. Using the Lagrangean,

$$L\left(\lambda,\alpha,\beta\right)=\alpha'\Sigma_{xy}\beta-\frac{\lambda_x}{2}\left(\alpha'\Sigma_{xx}\alpha-1\right)-\frac{\lambda_y}{2}\left(\beta'\Sigma_{yy}\beta-1\right) \tag{3}$$

where λ_x and λ_y are langrangean coefficients. Deriving the equation (3) for and parameters gives their scores.

2.2. Nonlinear neural network

Simulation techniques with neural network have had successful results e.g. in functions values approximation, realizing patterns and nonlinear processes estimation (Scholkopf, et. all, 1999; Vert and Kanehisa, 2003; Asoh and Takechi, 1994; Hsieh, 2000). The main advantage of the neural network model is learning. The operational function for all variables is,

$$f\left(x_1,x_2\right)=arctanh\left(\frac{\|x_1\|^2}{\|x_2\|^2}\right)$$

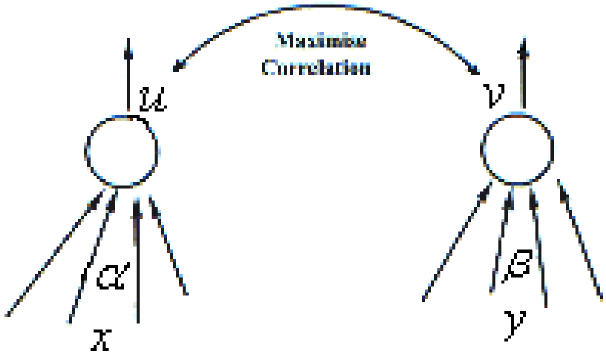


Fig. 2. Nonlinear neural network

If we choose appropriate initial coefficient, the above will be able to relate between income and outcome. In this method the main purpose is to maximize the correlation and (Fig.2).That is;

$$\begin{cases} u = \sum_j \alpha_j f(x_j) = \alpha f_1 \\ v = \sum_j \beta_j f(y_j) = \beta f_2 \end{cases} . \tag{4}$$

To calculate the and coefficient used Lagrangean,

$$J = E\left\{(uv) + \frac{\lambda_1}{2}\left(1-u^2\right) + \frac{\lambda_2}{2}\left(1-v^2\right)\right\} \tag{5}$$

Deriving the equation (5) for and parameters gives their scores. e.g.,

$$\begin{cases} \frac{\partial J}{\partial \alpha} = f_1\left(v - \lambda_1 u\right) \\ \frac{\partial J}{\partial \beta} = \alpha\left(1 - f_1^2\right) x\left(v - \lambda_1 u\right) \end{cases} . \tag{6}$$

. This model which approximately optimizes the nonlinear canonical correlation analysis. However, this model requires a lot of computation time and it also has a lot of local optima. To solve this problem a kernel canonical correlation analysis is used which enable as to decrease voluminous computations and pretend unwanted local optima.

2.3. Kernel canonical correlation analysis (KCCA)

KCCA is a method which generalizes classical CCA and which we now recall. The kernel canonical correlation analysis is used which enables as to decrease voluminous computations and pretend unwanted local optima (Bach and Jordan, 2002; Melzer, et. all, 2001). Its goal is to detect correlation between two data-sets and . To this end, the objects and are mapped to some Hilbert space and by a mapping and . Classical CCA can then be performed between the images and as follows. The goal is to find two directions and such that the features

$$\begin{cases} u = (\alpha, \Phi_X(x)) \\ v = (\beta, \Phi_Y(y)) \end{cases} \tag{7}$$

be maximally correlated (Fig.3). They can therefore be expressed as:

$$\begin{cases} \alpha = \sum_i \alpha_i \Phi_X(x_i) \\ \beta = \sum_i \beta_i \Phi_Y(y_i) \end{cases} \tag{8}$$

In that case the corresponding and can be rewritten as

$$\begin{cases} u = \sum_i \alpha_i \left(\Phi_X(x_i), \Phi_X(x)\right) \\ v = \sum_i \beta_i \left(\Phi_Y(y_i), \Phi_Y(y)\right) \end{cases} \tag{9}$$

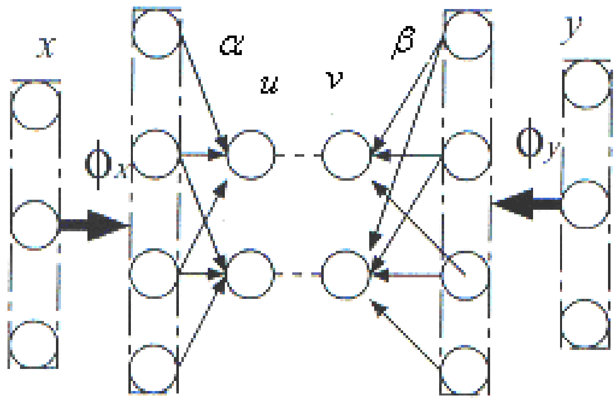


Fig. 3. KCCA

The and can now be found by solving the Lagrangean

$$L_0 = E\left(\left(u - E\left(u\right)\right)\left(v - E\left(v\right)\right)\right) - \frac{\rho_x}{2} E\left(u - E\left(u\right)\right)^2 - \frac{\rho_y}{2} E\left(v - E\left(v\right)\right)^2 \tag{10}$$

Where and are lagrange multipliers. Any Gaussian kernel function $K\left(x_1, x_2\right) = \exp\left(\frac{\|x_1 - x_2\|^2}{2\sigma^2}\right) K\left(.,.\right)on\chi^2$ defines a Hilbert space and a mapping$\Phi_\chi(.)$. Such that

$$K_X(x_1, x_2) = \left(\Phi_X(x_1), \Phi_X(x_2)\right), \forall (x_1, x_2) \in \chi^2 \tag{11}$$

Now let $(K_x)_{ij} = K_x\left(x_i, x_j\right)$and $(K_y)_{ij} = K_y\left(y_i, y_j\right)$be two matrices, assumed to be centered. Then L can be rewritten as

$$L = \alpha' K_x K_y \beta - \frac{\rho_x}{2} \alpha' \left(K_x + \lambda_x I\right)^2 \alpha - \frac{\rho_y}{2} \beta' \left(K_y + \lambda_y I\right)^2 \beta \tag{12}$$

This shows that regularization parametersλ_x and λ_y control the trade-off between maximizing the correlation and penalizing the complexity of α and β. Maximizing this lagrangean can be done by solving the following generalized eigenvalue problem. After finding, kernel canonical correlation scores (KCC scores) can be recovered by $u = K_x\alpha$ and $v = K_x\beta$.

3. Data Analysis

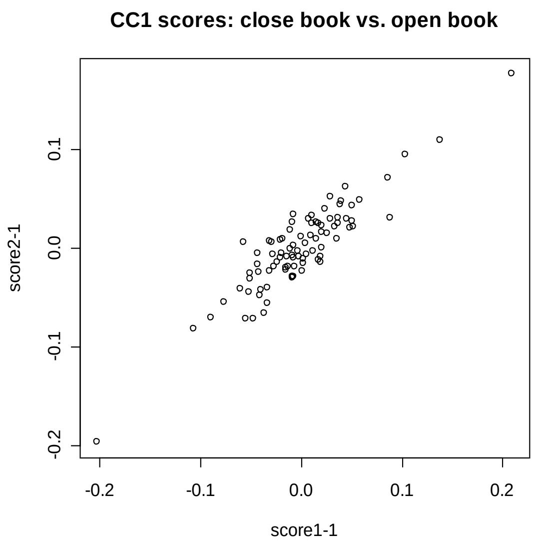
Knowing how to find scores (dependent variable) will be affected by pervious exam's or other variables such as age or sex (explanatory variable), maybe useful for canceling or other objectives. Suppose we have some final scores from different field, how can we combine and mix these scores or obtain the mean? The direction is to the mean. But this way always is not the appropriate way. For example if some scores from certain examination, will change more consistently than others we may give them different value and this lead to seeking linear combination which is, in one sense, optimal. Some times some exams will be classified in to more than one group. For example we may have "open book exams" or" closed book" ones. In such cases we may tend to use linear combination in each group separately. To study this, consider, data related to open and closed books, so that we have five variables, two closed books (Mechanic and vectors) and three open books (statistics, analysis and algebra). In this case we want to know how much a student has aptitude in open books. You should know that (u_1, v_1) and (u_2, v_2) are new artificial pair variables based on canonical and kernel methods. Table 1 shows that the first and second correlation obtained by kernel are higher than canonical methods. Table 2 shows that the first correlation (the most) obtained by neural network and kernel methods are higher than those obtained by canonical method. Of course correlation obtained by kernel method is much higher than those obtained by other statistical analysis. Diagram 1 and 2 show the dispersion of the data obtained by normal & kernel method. As clear from diagram 1 it is the best interpretation for is that, this variable shows the high frequency for closed book exams. The above results obtained by R1.9.0.

Table 1. shows that the first and second correlation

	CCA		*KCCA*	
	v_1	v_2	v_1	v_2
u_1	0.663	0	0.736	0
u_1	0	0.031	0	0.419

Table 2. Statistical methods

1. Standard statistics maximum correlation	0.663
2. Canonical maximum correlation	0.663
3. Nonlinear neural network maximum correlation	0.674
4. Kernel maximum correlation	0.736



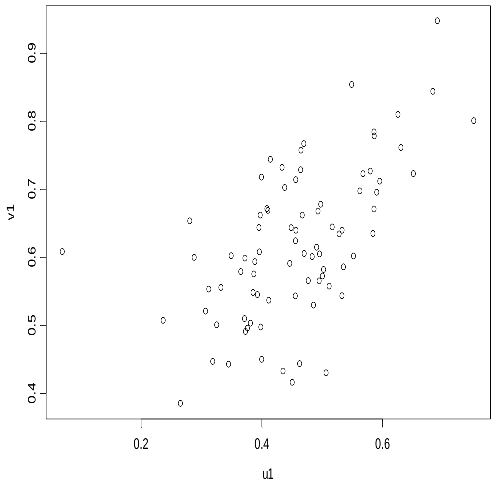


Fig. 4. Diagram 1 and 2 show the dispersion of the data obtained by normal & kernel method.

4. Conclusion

The kernel method compared to canonical method is preferred in space with high domination one .If all canonical correlation are significant, all correlation will have an approximate one. We can apply kernel method for discriminant analysis, principal Components analysis and regression. Generally, for nonlinear data we can use kernel method to optimize correlation in canonical method is both linear & significant, we can show this relationship by kernel method much better. The optimal method for student's exams results is kernel method. For optimizing exams values or coefficient, the first step is correlation or relationships between the data. Kernel method shows that student's aptitude in closed book depends highly on his or her aptitude in open book. Other words we can use open book results for predicting close book ones or visa versa. Therefore one appropriate method for exams result analysis in determining correlation or relationship is kernel analysis. These results, of course, should be studied carefully since firstly, in kernel method it is necessary that quantity of each sample data be similar. Secondly, obtaining correlation or relationship in each set of data is possible only on the condition that we have a generalized kernel regression.

References

[1] Akaho, S. (2001). A kernel method for canonical correlation analysis. International Meeting of Psychometric Society (IMPS2001).
[2] Asoh, H. and Takechi, O. (1994). An approximation of nonlinear canonical correlation analysis by multilayer perceptron's. Proc. of int. conf. artificial neural network, pp. 713-716.
[3] Bach, F. R. and Jordan, M. I. (2002). Kernel independent component analysis. Journal of Machine Learning Research, 3, 1-48.
[4] Bafield, J. D. and Raftery, A. E. (1993). Model based Gaussian and non-Gaussian Clustering. Biometrics, 49, 803-821.
[5] Bartlett, M. S. (1941).The statistical significance of canonical correlation. Biometrika, 32, 29-38.
[6] Foucat, T. (1999). Multiple linear regression on canonical correlation variables. Biometrical, 5, 559-572.
[7] Hotelling, H. (1936). Relation between sets of variables. Biometrika, 28, 312-377.
[8] Hsieh, W. W. (2000). Nonlinear canonical correlation analysis by neural networks. Neural Networks, 13, 1095-1105.
[9] Lai, P. L. and Fyfe, C. (2000). Kernel and nonlinear canonical correlation. International Journal of neural systems, 10(5), 365-377.
[10] Melzer, T. & Reiter, M. and Bischof, H. (2001). Kernel CCA: a nonlinear extension of canonical correlation analysis. Submitted to IEEE trans. neural network-works.
[11] Scholkopf, B. & Mika, S. & Burges, C. & Knirch, P. & Muller, K. R. & Ratsch, C. and Sloma, J.(1999). Input space vs feature space in kernel-based methods. IEEE transformation on neural networks, 10, 1000-1017.
[12] Vert, J. P. and Kanehisa, M. (2003). Graph driven features extraction from microarray data using diffusion kernels and kernel CCA. To appear in advances in neural information processing systems 15. MIT press, Cambridge, MA.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–107, pp. 259–264

Assessing Outlier Sensitivity in Nonparametric and Model-Based Classification Techniques

Mostafa Tamandi[a,*]

[a] *Department of Statistics, Vali-e-Asr University of Rafsanjan*

Article Info

Keywords:
Outlier
Classification
Random forests
Decision tree
Finite mixture models.

2020 MSC:
62H30
62R07

Abstract

Dealing with outliers poses a significant challenge in the realm of data classification. These outliers can inadvertently inflate the number of predicted categories and result in classification errors. Classification methods can be broadly divided into two categories: parametric and non-parametric. One example of a parametric approach is model-based classification, which relies on utilizing a finite mixture of probability density functions. Notable among the non-parametric methods are decision trees and random forests. In this research, we explore the impact of outliers on classification algorithms through Monte Carlo simulations and conduct a comparative analysis. The findings indicate that parametric methods exhibit superior accuracy compared to the decision trees and random forests algorithms.

1. Introduction

Classification methods are a fundamental aspect of data analysis and machine learning. They are utilized to categorize and predict the class or category to which a particular data sample belongs. In this article, we will delve into the world of classification methods, specifically focusing on the comparison between parametric and nonparametric approaches and their sensitivity to outliers.

There are two general categories for these methods: parametric and nonparametric. Parametric methods involve making assumptions about the underlying distribution of the data. By making these assumptions, the model can estimate the parameters that define the distribution and then use those estimated parameters to make predictions. Finite mixture (FM) models, specifically, incorporate a combination of probability density functions, which can accommodate multi-modal distributions. Each density function within the model has weight coefficients that determine the probability of assigning each observation to a specific group or class. Since the sum of these coefficients equals one, this method can be used to classify data into predefined groups. The main condition for using this method is that the data distribution is known in advance.

On the other hand, nonparametric methods do not make any restrictive assumptions about the underlying data distribution. Instead, they rely on the data itself to determine the decision boundaries and make predictions. Nonparametric

*Talker
Email address: Tamandi@vru.ac.ir (Mostafa Tamandi)

methods include techniques such as k-nearest neighbors (KNN), support vector machines (SVM), decision trees, and random forests.
One crucial aspect to consider when working with classification methods is the presence of outliers in the data. Outliers are data points that deviate significantly from the overall pattern or distribution of the data. These outliers can have a substantial impact on the classification results, potentially skewing the decision boundaries and misclassifying data points.
In this article, we compare the behavior of some finite mixture models and nonparametric tree-based models for classification in the presence of outliers. In the following section, we will examine a few classification models. Section 3 will include simulated studies to assess the classification performance of each model when dealing with data containing outliers.

2. Classification schemes

2.1. Finite mixture models

Consider n independent random variables $Y_1, Y_2, ..., Y_n$, which are taken from a mixture of some distributions. The pdf of a g-component mixture model is given by

$$f(y_i|\boldsymbol{\Theta}) = \sum_{j=1}^{g} \pi_j \; f_Z(y_i; \boldsymbol{\theta}_j), \; i = 1, \cdots, n, \tag{1}$$

where $\pi_j \geq 0$'s, $\sum_{j=1}^{g} \pi_j = 1$, are mixing proportions and $f_Z(y_i; \boldsymbol{\theta}_j)$ denotes the pdf of the f_Z and $\boldsymbol{\Theta} = (\pi_1, ..., \pi_{g-1}, \theta_1, ..., \theta_g)$ represents all unknown parameters.
McLachlan and Peel (2004) is a good reference for two well-known FM-Normal and FM-t models. In these models, $f_Z(y_i; \boldsymbol{\theta}_j)$ is the pdf of Normal and t-student distributions, respectively. Nonetheless, the stability provided by the FM-Normal and FM-t models might not be adequate to handle significant extreme values in the observed data's tails. To mitigate this problem, Lin et al. introduced finite mixture models of the skew-normal (SN) and skew-t (ST) distributions in separate studies published in 2007.
The SN distribution (Azzalini, 1985) is conventionally denoted by $Y \sim SN\left(\xi, \sigma^2, \lambda\right)$ with pdf given by

$$f_{SN}(y; \xi, \sigma^2, \lambda) = \frac{2}{\sigma}\phi(u)\Phi(\lambda u), \tag{2}$$

where $u = (y - \xi)/\sigma$, and $\phi(\cdot)$ and $\Phi(\cdot)$, respectively, stand for the pdf and cdf of the standard normal distribution, namely $N(0, 1)$. Moreover, the skew-t (Azzalini and Capitanio, 2003) distribution has the pdf as

$$f_{ST}(y; \xi, \sigma^2, \lambda, \nu) = \frac{2}{\sigma}t(u; \nu)T\left(\lambda u\sqrt{\frac{\nu+1}{\nu+u^2}}; \nu+1\right), \quad y \in \mathbb{R},$$

where $t(\cdot; \kappa)$ and $T(\cdot; \kappa)$ denote the pdf and cdf of the Student-t distribution with κ degrees of freedom (df), respectively. The FM-SN and FM-t distributions are defined by getting $f_Z(y_i; \boldsymbol{\theta}_j)$ in (1) as the pdf of SN and ST distributions, respectively.

2.2. Tree-based classification models

In this section, we give a summary of each nonparametric method that is used to classify data. The focus is on providing an overview rather than diving into the mathematical formulas or technical algorithms of each approach. For in-depth understanding, readers are encouraged to explore the referenced sources associated with each model. In addition, there are various machine learning algorithms available for this purpose. However, for the purpose of comparison, we will only focus on two tree-based models that are developed using distance measurements between observations and their similarities and dissimilarities.

- **DT:** The decision tree is a type of supervised learning model that involves dividing the predictor space into multiple simple regions. In this model, each observation is assigned to the most common class of training observations within its specific region. This is achieved by constructing a tree with different nodes and leaves. Each observation can be classified into a terminal node of the decision tree based on input variables, with a conditional probability. To create a classification tree, the misclassification error rate or Gini index can be used as criteria to determine when the algorithm should stop. Decision trees are easily displayed visually, making them simple to explain to non-experts. However, compared to other parametric classification approaches, decision trees may not always provide the same level of predictive accuracy, as noted by some authors (James et al., 2013).

- **RF:** The random forest algorithm is a method for improving the predictive accuracy of decision trees by combining several of them together. It works by constructing multiple decision trees using different training samples, selecting a random set of input variables for each tree, and stopping the algorithm when a certain accuracy level is reached. Because these trees are uncorrelated and chosen through bootstrap sampling, the random forest provides a more accurate classification compared to individual decision trees. (Wager and Athey, 2018)

In the next section, we'll employ these algorithms on simulated data and compare the performance of them in presence of outliers data.

3. Simulation scheme

Outliers can have a large impact on the estimation of model parameters and the prediction of future outcomes. Therefore, it is desirable to have models that are not overly influenced by outliers and can handle them appropriately. In this section, we will compare the models that we have introduced in the previous section in terms of their sensitivity to outliers. We will use some synthetic and real data sets to illustrate the effect of outliers on each model and discuss the advantages and disadvantages of each model in dealing with outliers.
The faithful data set is a classic example of how to apply classification techniques to real-world problems. This data set contains 272 observations of the Old Faithful geyser in Yellowstone National Park, USA. The data set records the duration of each eruption (eruption) and the time elapsed since the previous eruption (waiting). By plotting these two variables, we can see that there are two distinct clusters of points, corresponding to two types of eruptions: short and long. The goal of classification is to assign each observation to one of these two classes, based on the values of the variables. This can help us understand the patterns and dynamics of the geyser's behavior, and also predict when the next eruption will occur.
We wanted to see how outliers affect the classification performance of different models on the data set we used. We also wanted to compare the models under different scenarios of outlier contamination. So, we created three simulation schemes, where we added outliers to the original data set at different rates: 10%, 20% and 30%. We generated the outliers from a uniform distribution by adding 0.2 to maximum values of each variable.
Figure 1 shows the faithful data and added outliers in three different schems.

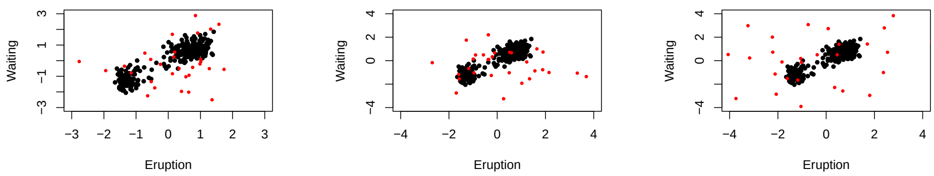


Fig. 1. Faithful data (black cyrcles) and added outliers (gray) in three different schems. Left: 10 percent, middle: 20 percent, Right: 30 percent.

Then, we apply the classification models that we discussed in the previous chapter to the three data sets with different levels of outlier contamination. These models are based on either distributional assumptions or tree-based structures.

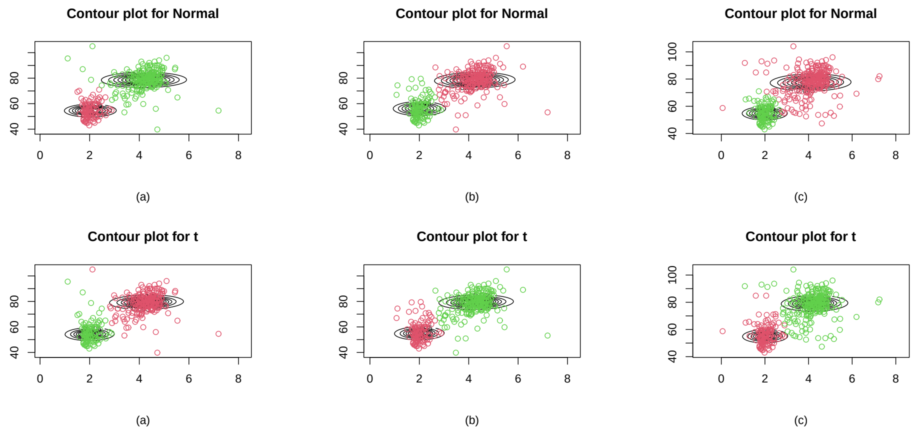


Fig. 2. The contour plots of fitted Normal and t distributions to faithful data with differnet outlier percents: (a) 10 percent (b) 20 percent (c) 30 percent.

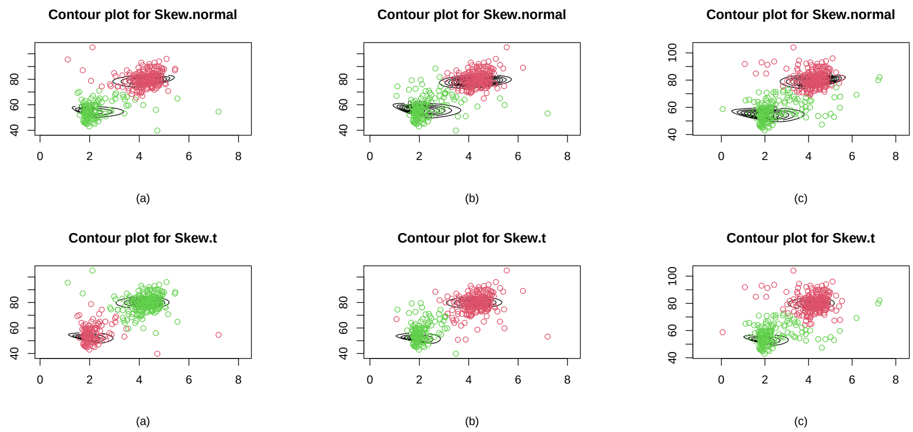


Fig. 3. The contour plots of fitted skew-normal and skew-t distributions to faithful data with differnet outlier percents: (a) 10 percent (b) 20 percent (c) 30 percent.

We plot the contour lines of the fitted distributions for each data set in Figures 2 and 3. We can see that the skew-normal and skew-t distributions are more robust to outliers than the normal and t distributions.
We also compare the classification performance of the different models on the original data and the data with outliers. We use two metrics to measure the performance: classification accuracy and sensitivity. The accuracy index measures the proportion of correct predictions out of the total number of predictions made by the model. It is calculated by dividing the number of true positives and true negatives by the total number of examples. A high accuracy index means that the model can correctly classify most of the examples.
The sensitivity index measures the proportion of positive examples that are correctly classified by the model. It is also known as the true positive rate or the recall. It is calculated by dividing the number of true positives by the sum of true positives and false negatives. A high sensitivity index means that the model can correctly identify most of the positive examples. The confusion matrix for a classification model is shown as:
The accuracy and sensitivity indices for this model are:

$$\textsc{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)}, \quad \textsc{Sensitivity} = \frac{TP}{(TP+FN)}$$

Figures 4 and 5 show the accuracy and sensitivity of compared models in 250 replicated of simulation schemes. Figure

Table 1. The confusion matrix of a classification model

	Predicted positive	Predicted negative
Actual positive	TP	FN
Actual negative	FP	TN

4 shows that in all three cases, the finite mixture models have higher accuracy than the nonparametric models. Among the parametric models, the skew-normal model is better than the other models as the number of outliers increases. According to Figure 5, the sensitivity of tree-based models decreases as the number of outliers increases. While the accuracy of the finite mixture models increases. Also, in general, in all three cases, the sensitivity of the finite mixture models is higher than the nonparametric models.

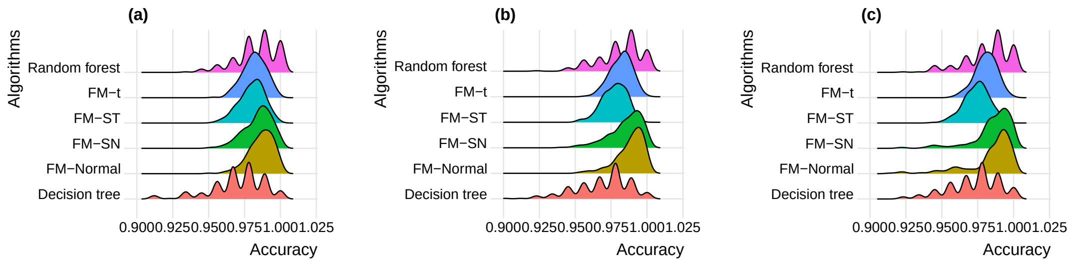


Fig. 4. The accuracy of finite mixture and tree-based models in three situations: (a) 10 percent outlier (b) 20 percent outlier (c) 30 percent outlier.

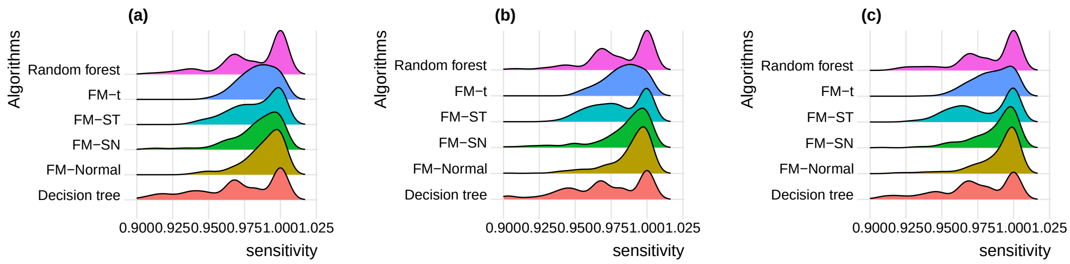


Fig. 5. The sensitivity of finite mixture and tree-based models in three situations: (a) 10 percent outlier (b) 20 percent outlier (c) 30 percent outlier.

4. Conclusions

In this paper, we have compared the parametric and nonparametric classification models in the presence of outliers. We observed that the parametric models have higher accuracy and sensitivity in classifying the data and as the number of outliers increases, the accuracy and sensitivity of the parametric models also increases. In this paper, we only used tree-based models, which are one of the most popular classification models and classify the observations based on their distance. There are other nonparametric models that can be compared, which researchers can also do this comparison on them.

References

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones. Scandinavian Journal of Statistics 12:171–178.

[2] Azzalini, A., and Capitaino, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65:367–389.

[3] James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). An introduction to statistical learning. New York: springer.
[4] Lin, T. I., Lee, J. C., and Hsieh, W. J. (2007a). Robust mixture modeling using the skew t distribution. Statistics and Computing 17:81–92
[5] Lin, T. I., Lee, J. C., and Yen, S. Y. (2007b). Finite mixture modelling using the skew normal distribution. Statistica Sinica 17:909–927.
[6] McLachlan, G., and Peel, D. (2004). Finite mixture models. John Wiley & Sons.
[7] Wager, S., and Athey, S. (2018). Estimation and inference of heterogeneous treatment effects using random forests. Journal of the American Statistical Association, 113(523), 1228-1242.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–108, pp. 265–268

Modeling stock return indices using a finite mixture of skewed distributions

Mostafa Tamandi[a,*]

[a]*Department of Statistics, Vali-e-Asr University of Rafsanjan*

Article Info

Keywords:
Skewed distribution
Finite mixtures
Stock returns
Kurtosis.

2020 MSC:
62E15

Abstract

In this paper, we introduce some novel finite mixture models that are based on scale and shape mixtures of skew-normal and skew-t distributions. These models are useful for analyzing the returns of financial assets, which often exhibit skewness and heavy tails. We compare the performance of these models with existing ones using real data and demonstrate that our models provide better fit and prediction accuracy.

1. Introduction

In the field of financial marketing, the analysis of stock returns serves as a versatile tool for predicting future market trends. Additionally, modeling stocks can provide insights into the behavior of other indices such as hedge funds, private equity, and money markets. Various authors have proposed probability distributions to effectively model financial and actuarial indices, with commonly utilized models including the Generalized Hyperbolic, inverse Gaussian, Logistic, and the well-established normal distribution. More recently, literature has discussed skewed distributions that are related to the normal model. For instance, Eling et al. (2014) applied skew-normal (SN) and skew-t (ST) distributions to asset returns. It is noteworthy that these models share the common property of being unimodal; however, in some cases, the histogram of the data suggests that the fitted models may be multimodal or a mixture with more than one component.
Finite mixture models, which involve expressing results as a linear combination of component probability density functions mixed in varying proportions, have been widely utilized as an analytical method in various scientific domains such as density estimation, supervised classification, unsupervised clustering, data mining, image analysis, pattern recognition, and machine learning (McLachlan and Peel, 2004). In the literature, Lin et al. (2007a) and Lin et al. (2007b) presented finite mixture models of the ST and SN distributions, respectively. Basso et al. (2010) conducted an investigation into finite mixture modeling in the univariate scale mixtures of SN (SMSN) distributions. Additionally, Tamandi and Jamaliadeh (2019) introduced a new family of finite mixtures based on skew scale-shape mixtures of the normal distribution. The objective of this study is to apply finite mixture models to stock returns datasets and demonstrate the performance of these models.

*Talker
Email address: Tamandi@vru.ac.ir (Mostafa Tamandi)

2. Main results

The family of skew scale-shape mixtures of normal(SKSSMN) distributions, introduced by Tamandi and Jamalizadeh (2019) is a big family includes many distributions. We briefly discuss on some of them and fitted these models to stock markets data.
Let the PDF of the random variable Y is given by

$$f_{SKSSMN}(y;\xi,\sigma^2,\lambda,\alpha,\boldsymbol{\nu}) = 2g(u;\boldsymbol{\nu})\Phi(\frac{\lambda u}{\sqrt{1+\alpha u^2}}),\ \ y\in\mathbb{R}, \tag{1}$$

where $u = (y-\xi)/\sigma$, $g(u;\boldsymbol{\nu}) = \int_0^\infty \tau^{1/2}\phi\left(\tau^{1/2}u\right)h(\tau;\boldsymbol{\nu})d\tau$, and τ is a positive random variable with PDF $h(\tau;\boldsymbol{\nu})$. Moreover, $\phi(\cdot)$ and $\Phi(\cdot)$ are the PDF and CDF of a standard normal distribution. Then generally the distribution of Y is on the family of skew scale-shape mixtures of normal(SKSSMN) distributions and is denoted by $Y\sim SKSSMN(\xi,\sigma^2,\lambda,\alpha,\boldsymbol{\nu})$.
Some special cases of SKSSMN family of distributions are:

(a) Normal distribution: If $\tau = 1$ and $\lambda = 0$, then the pdf in (1) is equal to a standard normal distribution.

(b) Skew-normal distribution: If $\tau = 1$ and $\alpha = 0$, then the pdf in (1) is equal to skew-normal, introduced by Azzalini(1983).

(c) Skew-t-normal (STN): Let $\tau\sim\Gamma(\frac{\nu}{2},\frac{\nu}{2})$ and $\alpha = 0$, where $\Gamma(\alpha,\beta)$ denotes the gamma distribution with mean α/β, we have STN distribution which is introduced by Gomez (2007). The range of kurtosis of STN distribution is wider than ST and hence this model is extensively used in many applications as an alternative to the ST distribution.

(d) Skew-contaminated-normal (SCN): The skew-contaminated-normal distribution is studied by Ferreira et al. (2011) as a particular member of the skew scale mixture of normal (SSMN) family. Now, let τ be a discrete random variable taking one of two states with pdf $h(\tau;\nu_1,\nu_2) = \nu_1 I_{(\tau=\nu_2)} + (1-\nu_1)I_{(\tau=1)}$. Further, let $\alpha = 0$ in (1), so the SCN distribution is achieved.

(e) Skew-slash (SSL): Another member of the SSMN family of distributions, that is discussed by Ferreira et al. (2011) is the skew-slash distribution. This model is obtained by assuming $\tau\sim\beta(\nu,1)$ and $\alpha = 0$ in (1).

(f) shape mixture of STN (SMSTN): Tamandi et al. (2018) generalized the STN distribution by introducing a shape mixture of the STN distribution. They showed that although the SMSTN distribution provides the same skewness range as the STN case, but allows for kurtosis ranging over a slightly wider interval than the STN distribution. To obtain this model from (1), it is sufficient to assume $\alpha > 0$ in case (c).

(g) and (h) If in the cases (d) and (e), we assume $\alpha > 0$, shape mixtures of SCN (SMSCN) and SSL (SMSSL) are obtained, which are discussed by Tamandi and Jamalizadeh (2019).

The well-known t and skew-t(ST; Azzalini and Capitanio(2003)) distribution are not in this family. But we discuss on them in our application for the sake of comparison.

2.1. Finite mixtures

Consider $Y_1, Y_2, ..., Y_n$, is a random sample from a mixture of the SKSSMN distributions. The pdf of a g-component FM-SKSSMN model is given by

$$f(y_i|\boldsymbol{\Theta}) = \sum_{j=1}^{g}\pi_j\ f_{SKSSMN}(y_i;\xi_j,\sigma_j,\lambda_j,\alpha_j,\boldsymbol{\nu}_j),\ i=1,\cdots,n, \tag{2}$$

where $\pi_j\geq 0$'s, $\sum_{j=1}^{g}\pi_j = 1$, are mixing proportions and $f_{SKSSMN}(y_i;\xi_j,\sigma_j,\lambda_j,\alpha_j,\boldsymbol{\nu}_j)$ denotes the pdf of the SKSSMN defined in (1) and $\boldsymbol{\Theta} = (\pi_1,...,\pi_{g-1},\theta_1,...,\theta_g)$ represents all unknown parameters. In this case, the component vector θ_j consists of $(\xi_j,\sigma_j,\lambda_j,\alpha_j,\boldsymbol{\nu}_j)$.

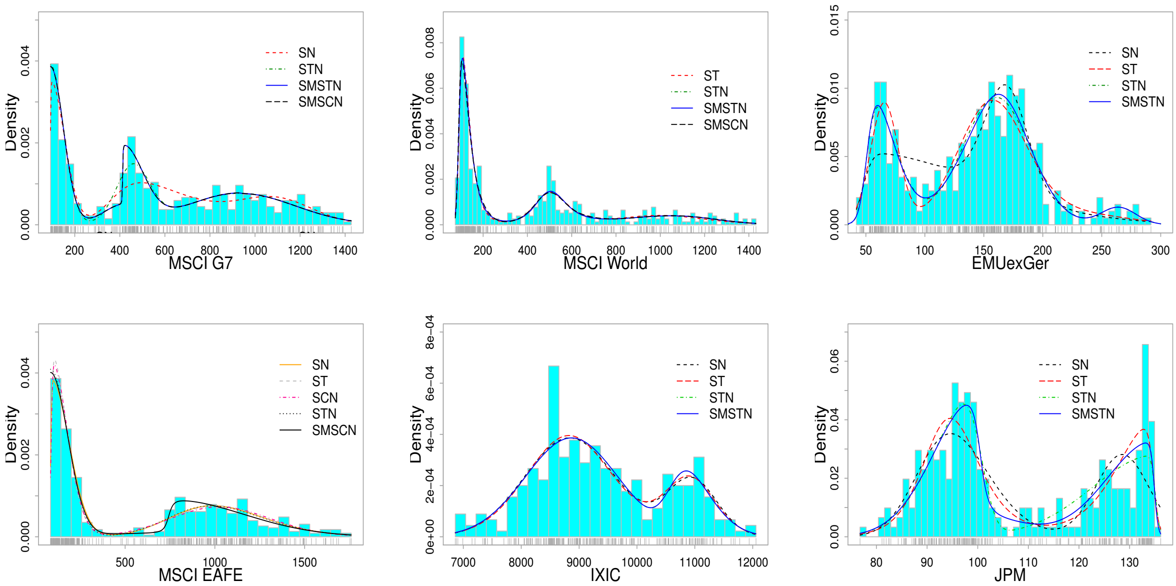


Fig. 1. Stock markets: Histograms with superimposed fitted densities for some concentrations.

Usually, mixture models have a widespread parameter's space and the PDF of these models have a complicated form. So, estimation of the parameters in a finite mixture model is a challenging aspect. Tamandi and Jamalizadeh (2019) introduced an expectation-conditional-maximization (ECM) algorithm to obtain the maximum likelihood estimates of the parameters of FM-SKSSMN model. For more details on FM-SKSSMN model and its ECM algorithm, see Tamandi and Jamalizadeh (2019)

3. Application

To model stock returns using finite mixture distributions, we employ six indices: JP Morgan Chase & Co. (JPM) and NAZDAQ (IXIC) Adj.close prices from October 2019 to October 2020 (the data are given from Yahoo finance), MSCI WORLD, MSCI G7, EMU ex Germany (EMUexGer) and MSCI EAFE from MSCI Inc.
Fig. 1 shows the histogram of data sets superimposed with the fitted density of the competitors. In all cases the well-known Gaussian mixture and t-mixture models are not adequate because of the big kurtosis in the data. Also the goodness of fit test based on Kolmogrov-Smirnov(KS) test statistics is significant in all cases and so the null hypothesis is rejected in all sudied stock indices.
Table 1 gives the number of components and some information criteria for finite mixture models of the studied distributions in section 2. The FM models with a big criteria are not reported in this table. Furthermore, we perform the Kolmogorov-Smirnov (KS) goodness of fit test to compare the quality of fit among the studied finite mixture distributions. The results are also shown in Table 1. To determine the optimal number of components of a mixture model, we fit the studied models to data sets with $g = 2 - 5$ components. Recently, McNicholas and Murphy (2008) have demonstrated the effectiveness of BIC in selecting the number of components for Gaussian mixture models.

4. Conclusions

In conclusion, this study has delved into the application of finite mixture models, particularly focusing on the comparison of shape mixture of STN and SCN distributions with older classes such as SN and ST distributions in analyzing

Table 1. Information criteria and goodness of fit test for some stock market indices.

Stock Index	g	Model	BIC	EDC	ICL	KS test	p-value
JPM	2	SN	4204.954	4190.909	4375.044	0.3845	0
		ST	4192.464	4178.418	4239.381	**0.0296**	0.847
		STN	4193.899	4177.847	**4230.132**	0.0316	0.829
		SMSTN	4199.920	4181.862	4235.387	0.0325	0.775
IXIC	2	SN	4252.007	4235.509	**4273.485**	0.0379	0.837
		ST	4250.76	4234.262	4274.957	**0.0337**	0.92
		STN	4255.994	4237.139	4274.764	0.0349	0.905
		SMSTN	4261.519	4240.308	4280.280	0.0348	0.914
MSCI WORLD	3	ST	5254.266	5232.195	**5295.167**	0.0404	0.533
		STN	5254.452	5230.374	5306.572	0.0336	0.742
		SMSTN	5260.456	5234.372	5312.955	0.0335	0.759
		SMSCN	5266.507	5238.417	5314.687	**0.0332**	0.775
MSCI G7	3	SN	5556.816	5534.745	5632.99	0.0522	0.212
		STN	5540.685	5516.607	5581.855	**0.0251**	0.955
		SMSTN	5533.938	5507.854	**5576.301**	0.0255	0.956
		SMSCN	5539.862	5511.771	5582.026	0.0254	0.95
MSCI EMUexGER	3	SN	4229.013	4206.942	4510.956	0.0669	0.053
		ST	4211.657	4189.585	4291.01	0.0380	0.64
		STN	4195.669	4171.592	**4219.099**	0.0337	0.767
		SMSTN	4200.463	4174.379	4224.239	**0.0308**	0.838
MSCI EAFE	2	SN	5459.532	5445.486	5462.143	0.0347	0.713
		ST	5462.069	5448.024	5467.158	0.0346	0.717
		STN	5454.647	5438.595	**5457.455**	0.0264	0.939
		SCN	5462.108	5448.063	5467.637	0.0294	0.868
		SMSCN	5459.153	5439.088	5469.956	**0.0250**	0.95

stock returns. Through rigorous empirical analysis, we have demonstrated that the utilization of the newer distributions, SMSTN and SMSCN, offer significant advantages over traditional distributions in capturing the complexities and nuances inherent in stock return data.

References

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* 12:171–178.

[2] Azzalini, A., Capitaino, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65:367–389.

[3] Basso, R. M., Lachos, V. H., Cabral, C. R. B., Ghosh, P. (2010). Robust mixture modeling based on scale mixtures of skew-normal distributions. *Computational Statistics and Data Analysis* 54:2926–2941.

[4] Eling M (2014) Fitting asset returns to skewed distributions: Are the skew-normal and skew-student good models? Insur Math Econ 59:45–56

[5] Ferreira, C., Bolfarine, H., Lachos, V. H. (2011). Skew scale mixtures of normal distributions: properties and estimation. *Statistical Methodology* 8:154–171.

[6] Gómez, H. W., Venegas, O., Bolfarine, H. (2007). Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics* 18:395–407.

[7] Lin, T. I., Lee, J. C., Hsieh, W. J. (2007a). Robust mixture modeling using the skew t distribution. *Statistics and Computing* 17:81–92.

[8] Lin, T. I., Lee, J. C., Yen, S. Y. (2007b). Finite mixture modelling using the skew normal distribution. *Statistica Sinica* 17:909–927.

[9] McLachlan, G., and Peel, D. (2004). Finite mixture models. John Wiley & Sons.

[10] McNicholas, P. D., Murphy, T. B., McDaid, A. F., Frost, D. (2010). Serial and parallel implementations of model-based clustering via parsimonious Gaussian mixture models. *Computational Statistics and Data Analysis* 54:711–723.

[11] Tamandi, M., Jamalizadeh, A., Lin, T. I. (2019). Shape mixtures of skew-t-normal distributions: characterizations and estimation, *Computational Statistics* 34: 323-347

[12] Tamandi, M., Jamalizadeh, A. (2019). Finite mixture modeling using shape mixtures of the skew scale mixtures of normal distributions. *Communications in Statistics-Simulation and Computation* 1-22.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–140, pp. 269–272

On testing exponentiality based on Gini index

Masoumeh Akbari

[a]*Department of Statistics, University of Mazandaran, Babolsar, Iran.*

Article Info

Keywords:
characterization
exponential distribution
Gini index
testing

2020 MSC:
62G10
62E10

Abstract

This paper first proves a characterization of exponential distribution in terms of the Gini index. After that three test statistics for exponentiality are introduced. Then, Monte Carlo simulations are carried out to assess their performances. Finally a real data set is used to illustrate introduced methods.

1. Introduction

The exponential distribution is one of the continuous distributions that is used frequently in different statistical fields like reliability theory and stochastic process. So, it is important to verify whether a random sample is taken from exponential distribution or not. There are various results for testing exponentiality in the literature, see [1] as the recent work. To the best of our knowledge, there is no test for exponentiality based on the Gini index. Gini [2] introduced the Gini coefficient which is a measure of statistical dispersion to represent income inequality, wealth inequality the consumption inequality. The rest of the paper is organized as follows. In Section 2, we present a new characterization of exponential distribution and propose three new test statistics for exponentiality. In Section 3, by Monte Carlo simulation, the critical values of proposed tests are computed. Also, their empirical power are derived for different alternatives which have different hazard functions. Finally, in Section 4, usefulness of these methods are illustrated with a real data set.

2. Characterization and test statistics

Let X and Y be independent and identically distributed (iid) non-negative continuous random variables (rvs). The notation $X \overset{d}{=} Y$ means that X and Y have the same distribution. Puri and Rubin[3] proved that

$$|X - Y| \overset{d}{=} X, \tag{1}$$

Email address: m.akbari@umz.ac.ir (Masoumeh Akbari)

if and only if they have exponential distribution. Suppose X_1 and X_2 are iid non-negative rvs with finite mean μ. Then, the Gini index is given by

$$G = \frac{E|X_1 - X_2|}{2\mu}. \tag{2}$$

Theorem 2.1. *Let X and Y be iid non-negative continuous rvs. Then $G = \frac{1}{2}$ if and only if X belongs to family of exponential distributions.*

Proof. The result can be easily obtained by (1) and (2). ☐

There exist several estimations of G based on the sample $X_1, \ldots, X_n$ as follows.

$$\begin{aligned}
G_1 &= \frac{2}{n\sum_{i=1}^{n} X_i} \sum_{i=1}^{n} X_{(i)} \left(i - \frac{1}{2}\right) - 1, \\
G_2 &= \frac{2\sum_{i=1}^{n-1} iX_{(i+1)}}{(n-1)\sum_{i=1}^{n} X_i} - 1, \\
G_3 &= \frac{2}{n\sum_{i=1}^{n} X_i} \sum_{i=1}^{n} iX_{(i)} - 1,
\end{aligned}$$

where $X_{(i)}$ is the ith order statistics of the sample. For more details about G_1 and G_2 see [4] and [5], respectively. G_3 is a new proposed estimator for G. Therefore three following statistics for exponentiality are proposed as follows.

$$T_1 = |G_1 - 0.5|, \quad T_2 = |G_2 - 0.5|, \quad T_3 = |G_3 - 0.5|.$$

3. Power comparison

In this section, using Monte Carlo simulations, we get the critical values of T_1, T_2 and T_3 for sample sizes $n = 10, 20, 30$. The results are shown in Table 1. We also obtain the empirical power of the proposed tests for different alternative distributions Weibull $(W(\theta))$, gamma $(G(\theta))$, half-normal (HN), lognormal $(LN(\theta))$ and standard uniform $(U(0,1))$ with following densities, respectively

$$\begin{aligned}
f(x) &= \theta x^{\theta-1} \exp\{-x^{\theta}\}; \;\; x > 0, \\
f(x) &= \frac{1}{\Gamma(\theta)} x^{\theta-1} \exp\{-x\}; \;\; x > 0, \\
f(x) &= \sqrt{\frac{2}{\pi}} \exp\{-\frac{x^2}{2}\}; \;\; x > 0, \\
f(x) &= \frac{1}{\theta x\sqrt{2\pi}} \exp\{\frac{-\log^2 x}{2\theta^2}\}; \;\; x > 0, \\
f(x) &= 1; \;\; 0 < x < 1.
\end{aligned}$$

Table 1. Critical values at significance level $\alpha = 0.05$.

n	T_1	T_2	T_3
10	0.191	0.187	0.195
20	0.132	0.129	0.132
30	0.082	0.810	0.081

From the Tables 2-5, it is obvious that test T_1 is the most powerful for increasing failure rate (IFR) distributions, and T_3 is the best for decreasing failure rate (DFR) distributions. Also T_2 is the most powerful for unimodal increasing-decreasing failure rate (UFR) distributions.

Table 2. Empirical power of proposed tests for $\alpha = 0.05$ and $n = 10$.

	IFR				
	$G(1)$	$W(1.4)$	$G(2)$	HN	$U(0,1)$
T_1	0.056	0.275	0.360	0.194	0.505
T_2	0.050	0.169	0.228	0.114	0.363
T_3	0.046	0.019	0.029	0.014	0.079

Table 3. Empirical power of proposed tests for $\alpha = 0.05$ and $n = 10$.

	UFR	UFR	DFR	DFR
	$LN(0.8)$	$LN(1.5)$	$W(0.8)$	$G(0.4)$
T_1	0.092	0.094	0.041	0.210
T_2	0.093	0.096	0.149	0.488
T_3	0.079	0.078	0.206	0.579

Table 4. Empirical power of proposed tests for $\alpha = 0.05$ and $n = 50$.

	IFR				
	$G(1)$	$W(1.4)$	$G(2)$	HN	$U(0,1)$
T_1	0.048	0.845	0.936	0.619	0.995
T_2	0.049	0.788	0.905	0.539	0.990
T_3	0.049	0.679	0.832	0.402	0.978

Table 5. Empirical power of proposed tests for $\alpha = 0.05$ and $n = 50$.

	UFR	UFR	DFR	DFR
	$LN(0.8)$	$LN(1.5)$	$W(0.8)$	$G(0.4)$
T_1	0.129	0.131	0.371	0.971
T_2	0.150	0.149	0.476	0.984
T_3	0.160	0.163	0.548	0.990

4. Application

In this section, we explain the procedures of proposed tests to detect the accuracy of exponentiality. It is used one following data set. In every test, the corresponding p-value is obtained based on 10000 replications. Also, assessing the suitability of exponential distribution is done at 5 percent nominal level.

Data1 : $74; 57; 48; 29; 502; 12; 70; 21; 29; 386; 59; 27; 153; 26; 326$

This is from Prochan [6] that shows the time between consecutive failures of the air conditioner in a Boeing 720 air plain. So far several authors verified that exponential distribution is a good fit for this data. See for example, Shanker et al.[7], Jose and Sathar[8].
The p-values of our proposed tests T_1, T_2 and T_3 are $0.349, 0.133$ and 0.073, respectively that confirm exponential distribution can be a fit for its distribution with mean 121.267.

References

[1] P. Xiong, W. Zhuang and G. Qiu. Testing exponentiality based on the extropy of record values. *Journal of Applied Statistics*, 49: 782–802, 2022.
[2] C. Gini.On the measure of concentration with special reference to income and statistics. *Colorado College Publication, General Series*, 208: 73–79, 1963.

[3] P.S. Puri, H. Rubin, A characterization based on absolute difference of two I.I.D. random variables, *Annals of Mathematical Statistics* , 41: 2113–2122, 1970.
[4] R. Davidson,Reliable inference for the Gini index, *Journal of Econometrics*, 150: 30–40, 2009.
[5] K.K. Sudheesh and I. Dewan. A simple non-parametric estimator of the Gini index, *Preprint series, Indian Statistical Institute at Delhi*,1–15, 2013.
[6] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*,5: 375–383, 1963.
[7] R. Shanker, F. Hagos and S. Sujatha. On the modelling of lifetimes data using exponential and Lindley distributions. *Biometrics* & *Biostatistics International Journal* ,2: 1-9, 2015.
[8] Jose, J., and Sathar, E.I.A.. Symmetry being tested through si- multaneous application of upper and lower k-records in extropy, *Journal of Statistical Computation and Simulation*,92: 830-846, 2022.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–180, pp. 273–277

Prediction of air quality index using machine learning algorithms: A case study of Tehran

Arshia Azizi[a], Sajjad Rahmany[a,*]

[a]*School of Mathematics and Computer Sciences, Damghan University, Damghan, Iran*

Article Info

Keywords:
air pollution
air quality index
machine learning
algorithms
mean absolute error
root mean square error

Abstract

Air quality index (AQI) forecasting is a useful tool for increasing the general public 's awareness of the state of the air in the next days. This is one of the most significant problems facing any country. In this study, machine learning algorithms are used to predict the AQI in Tehran. The six important regression models are applied to forecast AQI on a daily basis. Models were compared and evaluated using statistical measures such as Mean Absolute Error (MAE), coefficient of determination, and root mean square error (RMSE). Based on these evaluations, the best model was selected. ExtraTreesRegressor is thought to be the best model for forecasting AQI in all seasons based on its outcomes. The results demonstrate that the ExtraTreesRegressor 's determination coefficient is nearly 1, and that the values of MAE and RMSE are respectively 0.002 and 0.004.

1. Introduction

Tehran is one of several cities throughout the world that are affected by air pollution, which is a serious environmental issue. Recent years have seen a considerable decline in air quality due to the city's rapid population increase, heavy traffic, and industrial activity. In order to assist authorities in making sound decisions and executing the necessary steps to enhance the quality of the air, there is an urgent need for efficient monitoring and prediction systems[3][6]. Recently, many researchers have focused on forecasting air pollution using machine learning[10], neural networks[3][12], and deep learning[6]. Machine learning techniques are widely used in environmental sciences, including weather forecasting, soil erosion, waste management, dust storms, and air pollution [1]. Conventional air pollution prediction techniques can be divided into statistical methods , artificial intelligence , and numerical forecasting [2]. Sharma et al.[11] time-series analysis of data from 2009 to 2017 was used to predict the air quality in New Delhi. To create a forecasting model based on deep learning, Kaya and Oguducu [7] used PM10 hourly data from Istanbul (Turkey) between 2014 and 2018. Gocheva-llieva et al. [5] developed a model for daily prediction that had 90% accuracy using the classification and regression tree technique.
The rest of the paper follows the materials and methods in Section 2, the results including the data preparation and refinement and air pollution prediction are presented in Section 3, and conclusions is presented in Section 4.

*Talker
Email addresses: arshiyaazizi1999@gmail.com (Arshia Azizi), s_rahmani@du.ac.ir (Sajjad Rahmany)

2. Materials and Methods

The data used in this paper consist of daily air quality data from Tehran Air Quality Control Company (AQCC) from several monitoring stations across Tehran from March 21, 2020, and March 21, 2023.
After researching and reading several articles about Tehran's air pollution, the following 14 features have been selected in Table2. [4][14][9][8][13]

Symbol	Feature
PM	Particulate Matter
NO2	Nitrogen Oxides
SO2	Sulfur Dioxide
CO	Carbon Monoxide
O3	Ozone
Temperature	A physical quantity known as temperature expresses quantitatively how hot or cold something is
Humidity	The concentration of water vapor in the air is known as humidity
Precipitation	Precipitation is any byproduct of atmospheric water vapor condensation that falls from clouds as a result of gravitational pull
Wind Gust	A wind gust is a momentary increase in wind speed
Wind Speed	The most important component of the atmosphere is wind speed, which is the rate at which air shifts from high to low pressure
Sea Level Pressure	Pressure within the atmosphere of Earth
Visibility	The measurement of the distance at which a light or item can be seen clearly
Solar Radiation	Solar irradiance is the surface power density of electromagnetic radiation that is received from the Sun in the wavelength range of the measuring device
UV Index	The ultraviolet index, or UV index, is a globally recognized indicator of the amount of UV light that can cause sunburns at a specific location and time

All of the data used in this work was normalized as scaled to range (0,1) in order to ensure that all numerical values were on the same scale and that large values did not dominate smaller ones. Figure 1 shows the general structure of the suggested model.

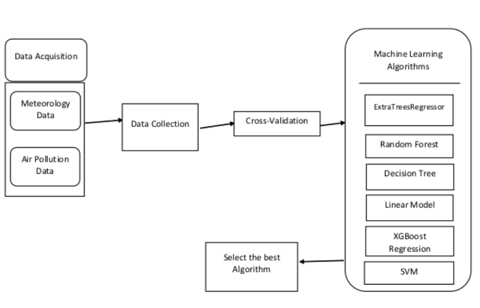


Fig. 1. The proposed best air pollution prediction model.

The performance of our approach was confirmed using the 10-fold cross-validation method. Applying machine learning algorithms like ExtraTreesRegressor, Random Forest, Decision Tree, Linear Model and XGBoost Regression will be done in the coming steps. In order to evaluate the prediction performance of the proposed model, we used the following measure:
• The mean absolute error is the average absolute difference between the predicted value and actual value, and is

calculated as follows:

$$MAE = \frac{\sum_{i=1}^{N} |y_i - x_i|}{N}$$

• The root mean square error is the square root of the distance between the predicted and actual value:

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}(Predicted_i - Actual_i)^2}{N}}$$

• The coefficient of determination, sometimes called coefficient, is the fraction of the variation in the dependent variable that is predicted from the independent variable(s), denoted R^2:

$$R^2 = 1 - \frac{RSS}{TSS}$$

Our results show that the ExtraTreesRegressor model performed well in predicting AQI. The most important variables for predicting AQI concentration were found to be temperature, humidity, wind speed, and traffic volume.
The ExtraTreesRegressor algorithm is a variant of the popular Random Forest method, which uses an ensemble of decision trees to make predictions. The ExtraTreesRegressor algorithm adds an additional level of randomness to the decision tree construction process, resulting in improved performance and faster training times.
The most important variables for predicting AQI concentration were found to be PM2.5, PM10, O3, Visibility, NO2 and traffic volume. Figure 2 shows the relative importance of each variable in the model.

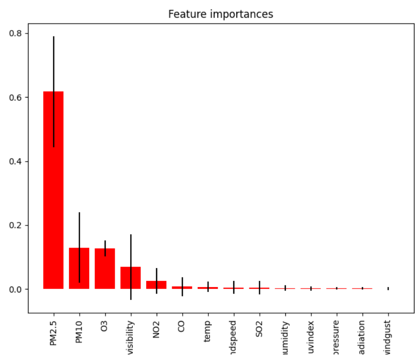


Fig. 2. Relative importance of each variable in predicting AQI concentration using the ExtraTreesRegressor algorithm.

3. Results

We evaluated the performance of the model using several metrics, including mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE) and coefficient of determination (R^2). The mean absolute error (MAE) measures the difference in errors between paired observations describing the same occurrence. The results of the algorithms applied to the data are shown in Table 3 and Figure 3 shows a comparison of the predicted and actual AQI

Algorithm	MAE	MSE	RMSE	R^2
ExtraTreesRegressor	0.002	1.94	0.004	0.996
Random Forest	0.002	0.0001	0.010	0.97
Decision Tree	0.002	4.84	0.006	0.99
Linear Model	0.015	0.0005	0.02	0.91
XGBoost Regression	0.003	3.43	0.005	0.99
SVM	0.04	0.002	0.046	0.62

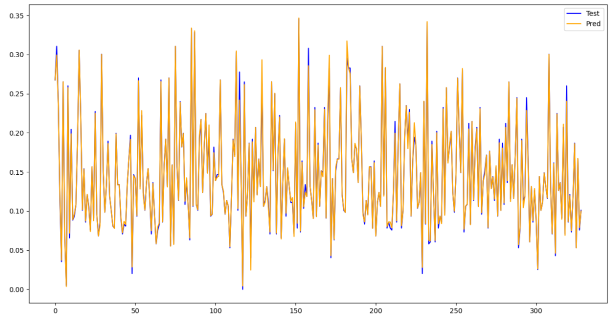


Fig. 3. Comparison of predicted and actual AQI concentrations for the testing set using the ExtraTreesRegressor algorithm.

concentrations for the testing set.

Overall, our results suggest that the ExtraTreesRegressor algorithm can be a useful tool for predicting air pollution in Tehran. The default setting for the ExtraTreesRegressor algorithm used in this study is 100 trees. The ExtraTreesRegressor algorithm produces the results shown below by altering the number of trees:

Number of Trees	MAE	MSE	RMSE	R^2
100	0.002	2.23	0.004	0.993
500	0.002	2.35	0.004	0.995
1000	0.002	1.94	0.004	0.996

4. conclusions

The focus of the research conducted in this article was on predicting air quality in Tehran by utilizing machine learning algorithms. Specifically, the ExtraTreesRegressor algorithm was employed to predict the concentrations of air pollutants in the city, based on various environmental and meteorological factors. Our study's findings indicated that the ExtraTreesRegressor algorithm was successful in predicting air pollutant concentrations, with an overall accuracy rate of 99% for predicting pollutant AQI.

One of the major advantages of our study was the use of a comprehensive and novel dataset that encompassed various meteorological and environmental factors. This allowed us to identify the crucial factors responsible for air pollutant concentrations and develop an accurate model to predict pollutant concentrations. Our research also highlights the potential of machine learning algorithms in predicting air pollutant concentrations, which can be leveraged to inform public health policies and decrease the adverse effects of air pollution on public health.

However, it's worth noting that our study's scope was limited to air pollutant concentrations in Tehran, which may not be generalizable to other regions or cities. Additionally, our research did not consider the impact of human behavior and activity patterns on air pollutant concentrations, which could be an important factor to consider in future studies.

In conclusion, our study provides crucial insights into the potential of machine learning algorithms for predicting air pollutant concentrations, emphasizing the necessity for further research in this domain. By enhancing and refining these algorithms, we can gain a better understanding of the factors contributing to air pollution and develop more effective approaches to mitigate its negative impacts on public health.

References

[1] S. Agarwal, S. Sharma, R. Suresh, and et al. Air quality forecasting using artificial neural networks with real time dynamic error correction in highly polluted regions. *Science of The Total Environment*, 735:139454, 2020. doi: 10.1016/j.scitotenv.2020.139454.

[2] L. Bai, J. Wang, X. Ma, and H. Lu. Air pollution forecasts: An overview. *Int. J. Environ. Res. Public Health*, 15:1–44, 2018. doi: 10.3390/ijerph15040780.

[3] M. Delavar, A. Gholami, G. Shiran, Y. Rashidi, G. Nakhaeizadeh, K. Fedra, and S. Hatefi Afshar. A novel method for improving air pollution prediction based on machine learning approaches: A case study applied to the capital city of tehran. *ISPRS Int. J. Geo-Inf.*, 8:99, 2019. doi: 10.3390/ijgi8020099.

[4] Z. Ghaemi, A. Alimohammadi, and M. Farnaghi. Lasvm-based big data learning system for dynamic prediction of air pollution in tehran. *Environ Monit Assess*, 190:300, 2018. doi: 10.1007/s10661-018-6659-6.

[5] S.G. Gocheva-Ilieva, D.S. Voynikova, M.P. Stoimenova, and et al. Regression trees modeling of time series for air pollution analysis and forecasting. *Neural Computing and Applications*, 31:9023–9039, 2019. doi: 10.1007/s00521-019-04432-1.

[6] A. Heydari, M. Majidi Nezhad, D. Astiaso Garcia, and et al. Air pollution forecasting application based on deep learning model and optimization algorithm. *Clean Techn Environ Policy*, 24:607–621, 2021. doi: 10.1007/s10098-021-02080-5.

[7] K. Kaya and S. Gunduz Oguducu. Deep flexible sequential (dfs) model for air pollution forecasting. *Sci. Rep*, 10:3346, 2020. doi: 10.1038/s41598-020-60102-6.

[8] H. Luo, Q. Guan, J. Lin, Q. Wang, and et al. Air pollution characteristics and human health risks in key cities of northwest china. *Science of The Total Environment*, 269:110791, 2020. doi: 10.1016/j.jenvman.2020.110791.

[9] R. Munsif, M. Zubair, A Aziz, and M. N. Zafar. Industrial air emission pollution: Potential sources and sustainable mitigation. *Environ Emissions*, pages 1–9, 2021. doi: 10.5772/intechopen.93104.

[10] A.K. Rad, R.R. Shamshiri, A. Naghipour, S.-O. Razmi, M. Shariati, F. Golkar, and S.K. Balasundram. Machine learning for determining interactions between air pollutants and environmental parameters in three cities of iran. *Sustainability*, 14:2–6, 2022. doi: 10.3390/su14138027.

[11] N. Sharma, S. Taneja, and V. Sagar. Forecasting air pollution load in delhi using data analysis tools. *Procedia Comput. Sci*, 132:1077–1085, 2018. doi: 10.1016/j.procs.2018.05.023.

[12] O. Wen, S. Liu, X. Yao, L. Peng, X. Li, and et al. A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Science of The Total Environment*, 654:1091–1099, 2019. doi: 10.1016/j.scitotenv.2018.11.086.

[13] L. Wu, N. Li, and Y. Yang. Prediction of air quality indicators for the beijing-tianjin-hebei region. *Cleaner Production*, 196:682–687, 2018. doi: 10.1016/j.jclepro.2018.06.068.

[14] M. Zamani Joharestani, C. Cao, X. Ni, B. Bashir, and S. Talebiesfandarani. Pm2.5 prediction based on random forest, xgboost, and deep learning using multisource remote sensing data. *Atmosphere*, 10:373, 2019. doi: 10.3390/atmos10070373.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–185, pp. 281–284

A Case Study about Identification of Probability Distributions

Leader Navaei[a,*]

[a]*Department of Statistics, Payame Noor University, (PNU), Tehran, Iran.*

Article Info

Keywords:
Identification
Asymptotically optimal
Independent Objects

2020 MSC:
62P30

Abstract

In this paper the problem of logarithmically asymptotically optimal (LAO) identification of probability distributions of two independent objects, is studied.

1. Introduction

The problem of identification of distribution for one object was considered in [1] and for two objects in [2]. We revealed certain in formulation and proof of the Theorem about indentification in [2]. Similar inaccuracy is remarded also in paper [4]. It is convinient to apply the definitions and notations of the paper [2]. The problem of identification of distributions for one and two independent Markov chains to the subject reliobility Critrion was studied in [5] and [6]. Also we add some assertion in the formulation of the Theorem. Let X_1 and X_2 be independent random variables (RV) taking values in the same finite set χ with one of M probability (PDs). They are characteristics of corræponding independent objects. The random vector (X_1, X_2) assumes values $(x^1, x^2) \in \chi \times \chi$.
Let $(x^1, x^2) = ((x_1^1, x_1^2), \ldots, (x_n^1, x_n^2), \ldots, (x_N^1, x_N^2))$, $x_n^i \in \chi$, $i = \overline{1, 2}$, $n = \overline{1, N}$, be two dimentional vectors of results of N independent observations of the pair (X_1, X_2). The statistician must define unknown PDs of the objects on the base of observed data. The selection for each object must be made from the same known set of hypotheses: $H_m : G = G_m, m = \overline{1, M}$. We call the procedure of making decision on the base of N pairs of observations the test for two objects and denote it by Φ_N. Because of the objects independence test Φ_N may be considered as the pair of the tests ϕ_N^1 and ϕ_N^2 for the respective separate objects. We shall denote the infinite sequence of compound tests by $\Phi_N = (\phi^1, \phi^2)$.
Let $\alpha_{l_1,l_2|m_1,m_2}(\Phi_N)$ be the probability of the erroneous acceptance by test Φ_N of the hypotheses pair (H_{l_1}, H_{l_2}) provided that the pair (H_{m_1}, H_{m_2}), is true, where $(m_1, m_2) \neq (l_1, l_2)$, $m_i, l_i = \overline{1, M}$, $i = 1, 2$. The probability to reject a true pair of hypothess (H_{m_1}, H_{m_2}) is following:

$$\alpha_{m_1,m_2|m_1,m_2}(\Phi_N) \triangleq \sum_{(l_1,l_2)\neq(m_1,m_2)} \alpha_{l_1,l_2|m_1,m_2}(\Phi_N) \tag{1}$$

*Talker
Email address: l.navaei@pnu.ac.ir (Leader Navaei)

Corresponding limits $E_{l_1,l_2|m_1,m_2}(\Phi_N)$ of the error probability exponents of the sequence of tests Φ, are called reliabilities:

$$E_{l_1,l_2|m_1,m_2}(\Phi) \triangleq \overline{\lim_{N\to\infty}} -\frac{1}{N}\log\alpha_{l_1,l_2|m_1,m_2}(\Phi_N), \quad m_i,l_i=\overline{1,M}, \quad i=1,2. \tag{2}$$

It is clear that

$$E_{m_1,m_2|m_1,m_2}(\Phi)=\min_{(l_1,l_2)\neq(m_1,m_2)} E_{l_1,l_2|m_1,m_2}(\Phi). \tag{3}$$

Here we call the test sequence Φ^* logarithmically asymptotically optimal (LAO) for the model with two objects if for given positive values of certain $2(M-1)$ elements of the reliability matrix the procedure provides maximal values for all other elements of it.

2. Problem Statament and formoulation of results.

For identification the statistician have to answer to the question whether the pair of distributions $(r_1,r_2), r_l,r_2\in[1,M]$ occurred or not. Let us consider two kinds of error probabilities for each pair (r_1,r_2). We denote by $\alpha^N_{(l_1,l_2)\neq(r_1,r_2)|(m_1,m_2)=(r_1,r_2)}$ the probability.that pair (r_1,r_2) is true, but it is rejected, that is accepted pair (l_1,l_2) do not coinsides with (r_1,r_2). Note that this probability is equal to probability $\alpha_{r_1,r_2|r_1,r_2}(\Phi_N)$ in testing. Let $\alpha^N_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)}$ be the probability that the pair (r_1,r_2) is accepted, when it is not correct. The corresponding reliabilities are $E_{(l_1,l_2)\neq(r_1,r_2)|(m_1,m_2=(r_1,r_2)}=E_{r_1,r_2|r_1,r_2}$ and $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}$. Our aim is to determine the depende of $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}$ on given $E_{r_1,r_2|r_1,r_2}$ during optimal. that is LAO, identification.
As in [2] we assume that hypotheses $G_1,G_2,\dots,G_M$ have a priori positive probabilities $Pr(r)$, $r=\overline{1,M}$, and consider the following probability:

$$\begin{aligned}\alpha^N_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)} &= \frac{Pr^N((m_1,m_2)\neq(r_1,r_2),(l_1,l_2)=(r_1,r_2))}{Pr((m_1,m_2)\neq(r_1,r_2))}\\ &= \frac{\sum_{(m_1,m_2):(m_1,m_2)\neq(r_1,r_2)}\alpha_{r_1,r_2|m_1,m_2}Pr(m_1,m_2)}{\sum_{(m_1,m_2)\neq(r_1,r_2)}Pr(m_1,m_2)}.\end{aligned}$$

Using this expression, we can derive that

$$E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}=\min_{(m_1,m_2):(m_1,m_2)\neq(r_1,r_2)} E_{r_1,r_2|m_1,m_2}. \tag{4}$$

For every test $Phi=(\phi_1,\phi_2)$, such that $E_{r_i|m_i}(\phi_i)>0$,$i=\overline{1,2}$, from (4) and Lemma [2] we obtain that

$$E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}=\min[\min_{m_1\neq r_1} E^I_{r_1|m_1}, \min_{m_2\neq r_2} E^{II}_{r_2|m_2}], \tag{5}$$

where E^I and E^{II} are elements of reliability matrices of corresponding objects. Using that the minimal elements of rowes of reliability matrices are the diagonal ones, we find that

$$E_{r_1,r_2|r_1,r_2}=\min_{m_1\neq r_1,m_2\neq r_2}(\min(E^I_{m_1|r_1},E^{II}_{m_2|r_2}))=\min(E^I_{r_1|r_1},E^{II}_{r_2|r_2}). \tag{6}$$

Let us denote for brevity

$$A(r)=\min_{l\neq r} D(G_l||G_r).$$

Let $\mathcal{P}=\{\Phi=(\phi_1,\phi_2):E_{r_1,r_2|r_1,r_2}(\Phi)=E_{r_1,r_2|r_1,r_2}\}$ be the set of tests the reliability matrices of which have diagonal elements equal to some preliminary given number $E_{r_1,r_2|r_1,r_2}$. For each test $\phi\in\mathcal{P}$ we can obtain value of corresponding reliability of indentification $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}$. We must choose such a test, for which the reliability $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}$ is the greatest. For every test $\Phi=(\phi_1,\phi_2)$ we find the reliability

$E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2\neq(r_1,r_2)}$ by equality (5), for which we must find the greater values of reliabilities $E^{I}_{l_1|r_1}(\phi_1)$ and$E^{II}_{l_2|r_2}(\phi_2)$. But only matrices corresponding to LAO tests can have such properties. Hence. the selection must be implemented from the set $\Phi^* = (\phi_1^*, \phi_2^*)$ of LAO tests, such that $E_{r_1,r_2|r_1,r_2}(\Phi^*) = E_{r_1,r_2|r_1,r_2}$.
From Theorem 1 of paper [3] we see that the order of hypotheses is important in formulation of conditions imposed on diagonal elements of the reliability matrix. This conditions depend on elements which are defined by preceding diagonal elements. But if we Consider the element $E_{r|r}$ as the element $E_{1|1}$, it will be possible consider the conditions formulated only by distribution G_m, $m = \overline{1,M}$. Passing to the problem of identification by (5) we can see that it will be usefully to change numeration of hypotheses obtaining formulation of corresponding conditions by distributions G_m, $m = \overline{1,M}$ only. Assume that $E_{r_1,r_2|r_1,r_2} = E^{I}_{r_1|r_1}$ and $E_{r_1,r_2|r_1,r_2} = E^{I}_{r_1|r_1} \leq E^{II}_{r_2|r_2}$. According to the above mentioned argumentation the number $E_{r_1,r_2|r_1,r_2} = E^{I}_{r_1|r_1}$ must satisfy the condition for being LAO test, i.e. $E_{r_1,r_2|r_1,r_2} \in (0, \mathcal{A}(r_1))$. Then we get the best test for the first object. from which we will obtain the least value in the column r_1 of the reliability matrix. It remains only to define the test for the second object applying the condition that its diagonal element $E_{r_2|r_2}$ is not less then $E_{r_1,r_2|r_1,r_2} = E^{I}_{r_1|r_1}$. From the tests with such property we will select only one requiring that in the column r_2 are greater than corresponding elements of the other tests.
Consider a set of numbers $\mathcal{K} = \{E : E_{r_1,r_2|r_1,r_2} \leq E_{r_2|r_2} < A(r_2)\}$. We introduce this numerical set with the goal to include into consideration all the LAO tests corresponding to the second object with a diagonal element $E_{r_2|r_2} \geq E^{I}_{r_1|r_1}$. Taking into consideration the obtained condition we determine the following condition for the preliminary given number

$$E_{r_1,r_2|r_1,r_2} \leq \min[A(r_1), A(r_2)]. \tag{7}$$

For each $E_{r_2|r_2} \in \mathcal{K}$ there exists a LAO test such that elements in column r_2 of its reliablity matrix are greater than corresponding elements of other tests. To examine all such tests we require that the second expression in (5) be the following:

$$\max_{E_{r_2|r_2}\in\mathcal{K}} \min_{m_2\neq r_2} E_{r_2|m_2}(E_{r_2|r_2}). \tag{8}$$

Since the lower bound can only decrease when the set increases, we get

$$\max_{E_{r_2|r_2}\in\mathcal{K}} \min_{m_2\neq r_2} E_{r_2|m_2}(E_{r_2|r_2}) = \min_{m_2\neq r_2} E_{r_2|m_2}(E_{r_1,r_2|r_1,r_2}) :$$

The expression (8) takes its greatest value at the point $E_{r_1,r_2|r_1,r_2}$. We derive the following estimate

$$E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)} = \min[\min_{m_1\neq r_1} E_{r_1|m_1}(E_{r_1,r_2|r_1,r_2}), \min_{m_2\neq r_2} E_{r_2|m_2}(E_{r_1,r_2|r_1,r_2})], \tag{9}$$

where

$$E_{r|m}(E_{r|r}) = \inf_{Q:D(Q||G_r)\leq E_{r|r}} D(Q||G_m).$$

If we assume that $E_{r_1,r_2|r_1,r_2} = E_{r_2|r_2}$, we will again come to formula (9) for calculation $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)}$, where the preliminary by given $E_{(r_1,r_2)|(r_1,r_2)}$ elements must meet condition (7).
If (7) is violated then the reliability which we investigate is equal to zero. The main result is can be formulated now in the following

Theorem 2.1. *If the distributions G_m, $m = \overline{1,M}$, are different and the given strictly positive number $E_{r_1,r_2|r_1,r_2}$ satisfy condition (7), then the reliability $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)}$ is defined in (9).*

If condition (7) is violated, then the reliability $E_{(l_1,l_2)=(r_1,r_2)|(m_1,m_2)\neq(r_1,r_2)}$ is equal to zero.

References

[1] R. F. Ahlswede and E. A. Haroutunian, ”On logarithmically asymptotically testing of hypot,heses and identification”. Lecture Notes in Computer Science 4123, ” General Theory of Information Transfer and Combinatorics” Springer. pp 478, 2006.

[2] E. A. Haroutunian and P. M. Hakobyan, ”On Identification of Distributions ci Independent Objects”. Mathematical Problems of Computer Science. vol. 28. pp. 119. 2007.
[3] E. A. Haroutunian. ”Logarithmically asymptotically optimal testing of multipli - tistical hypotheses”. Problerns of Control and Information Theory, vol. 19. pp. 413-421. 1990.
[4] L. Navaei . ”On reliable identification of two independent Markov chain” . Mathematical Problems of Computer Science vol. 32, pp. 75-79, 2009.
[5] L. Navaei. ”Identification of optimality of multiple Hypotheses Testing”. Lamdert Academic Publishiy, Germany. 2010.
[6] L. Navaei and R.Akbari. ”On optimal identification of distribution for two independent Markov Chains to the subject reliability criterion”. Control and Optimization in Applied Mathematics (COAM), Vol. 6, No. 1, pp. 31-41, 2021.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–186, pp. 285–289

On Many Hypotheses Optimal Testing and Identification of Probability Distributions

Leader Navaei[a,*]

[a]*Department of Statistics, Payame Noor University, (PNU), Tehran, Iran.*

Article Info

Keywords:
Hypotheses Testing
Dependent Objects
Method of Types

2020 MSC:
62P30

Abstract

The problem of logarithmically asymptotically optimal hypotheses (LAO) testing and identification for a model consisting of two stochastically related objects is studied. It is supposed that L_1 possible probability distributions are known for the first object and the second object is distributed according to one of $L_l \times L_2$ given conditional distributions depending on the distribution index and the current observed state of the first object.

1. Introduction

As a development of the results on multiple hypotheses testing concerning probability distributions of one object. Haroutunian and Hakobyan considered in [3] the problem of many hypotheses testing and in [4] the problem of the indentification of distributions for two independent objects. Solutions of analogical problems concerning Markov distributions are obtained in works of Haroutunian and Grigoryan [8], Haroutunian and Navaei [10]. In [6], [7] and [9] Haroutunian and Yessayan solved the problem of many hypotheses testing for two objects under different kinds of dependence. We study characteristics of procedures of logarithmically asymptotically optimal testing and identification of probability distributions of two stochastically dependent objects.
Let X_l and X_2 be random variables (RVs) taking values in the same finite set χ and $\mathcal{P}(\chi)$ be the space of all possible distributions on X. If X_l and X_2 take values in different sets χ_l and χ_2 only the notations became more complicated. so we omit this "generalization". There are given L_l probability distributions (PDs) $G_{l_1} = G_{l_1}(x^1), x^1 \in \chi, l_1 = \overline{1, L_1}$, from $\mathcal{P}(\chi)$. The first object is characterized by $RV\ X_l$ which has one of these L_1 PDs and the second object is dependent on the first and is characterized by $RV\ X_2$ which can have one of $L_l \times L_2$ conditional PDs $G_{l_2/l_1} = \{G_{l_2/l_1}(x^2|x^1), x^1, x^2 \in \chi\}$, $l_1 = \overline{1, L_1}$, $l_2 = \overline{1, L_2}$. Let $(X_1, X_2) = ((x_1^1, x_1^2), (x_2^1, x_2^2), \ldots, (x_N^1, x_N^2))$ be a sequence of results of N independent observations of pair of objects. Joint (PDs) $G_{l_1,l_2}(X^1, X^2)$, $l_1 = \overline{1, L_1}$, $l_2 = \overline{1, L_2}$ where $G_{l_1,l_2}(x^1, x^2) = G_{l_1}(x^1)G_{l_2/l_1}(x^2|x^1)$. The probability $G^N_{l_1,l_2}(X_1, X_2)$ of vector (X_1, X_2) is following product:

$$G^N_{l_1,l_2}(X_1, X_2) = G^N_{l_1}(x_1)G^N_{l_2/l_1}(X_2|X_1) = \prod_{n=1}^{N} G_{l_1}(x_n^1)G_{l_2/l_1}(x_n^2|x_n^1). \tag{1}$$

*Talker
Email address: l.navaei@pnu.ac.ir (Leader Navaei)

with $G^N_{l_1}(x_1) = \prod_{n=1}^N G_{l_1}(x^1_n)$ and $G^N_{l_2/l_1}(X_2|X_1) = \prod_{n=1}^N G_{l_2/l_1}(x^2_n|x^1_n)$. For the object characterized by X_l the non-randomized test $\phi^N_1(X_1)$ can be determined partition of the sample space χ^N on L_1 disjoint subsets $\mathcal{A}^N_{l_1} = \{X_1 : \phi^N_1(X_1) = l_1\}$, $l_1 = \overline{1, L_1}$, i.e. the set $\mathcal{A}^N_{l_1}$ consists of vectors X_1 for which the PDG_{l_1} is adopted. The probability $\alpha^N_{l_1|m_1}(\phi^N_1)$ of the erroneous acceptance of PDG_{l_1} provided that G_{m_1} is true, $l_1, m_1 = \overline{1, L_1}$, $m_1 \neq l_1$, is defined by the set $\mathcal{A}^N_{l_1}$

$$\alpha^N_{l_1|m_1}(\phi^N_1) \triangleq G^N_{m_1}(\mathcal{A}^N_{l_1}).$$

We define the probability to reject G_{m_1}, when it is true, as follows

$$\alpha^N_{m_1|m_1}(\phi^N_1) \triangleq \sum_{l_1 \neq m_1} \alpha^N_{l_1|m_1}(\phi^N_1) = G^N_{m_1}(\overline{\mathcal{A}^N_{m_1}}). \tag{2}$$

Denote by ϕ_1, ϕ_2 and Φ the infinite sequences of tests. Corresponding error probability exponents $E_{l_1|m_1}(\phi_1)$ for testϕ_1 are defined as

$$E_{l_1|m_1}(\phi_1) \triangleq \overline{\lim_{N\to\infty}} - \frac{1}{N} \log \alpha^N_{l_1|m_1}(\phi^N_1), \quad l_1, m_1 = \overline{1, L_1}. \tag{3}$$

For brevity we call them reliabilities. It follows from (2) and (3) that

$$E_{m_1|m_1}(\phi_1) = \min_{l_1 \neq m_1} E_{l_1|m_1}(\phi_1), \quad l_1, m_1 = \overline{1, L_1}, \quad l_1 \neq m_1. \tag{4}$$

We shall reformulate now the Theorem from [1] for the case of one object with L_l hypotheses. This requires some notions and notations. For some PD $Q = \{Q(x^1), x^1 \in \chi\}$ the entropy $H_Q(X_1)$ and the informational divergence $D(Q||G_{l_1}, l_1 = \overline{1, L_1}$,

$$H_Q(X_1) \triangleq -\sum_{x^1 \in \chi} Q(x^1) \log Q(x^1),$$

$$D(Q||G_{l_1}) \triangleq -\sum_{x^1 \in \chi} Q(x^1) \log \frac{Q(x^1)}{G_{l_1}(x^1)}.$$

For given positive numbers $E_{1|1}, \ldots, E_{L-1|L-1}$ let us consider the following sets of PDs $Q = \{Q(x^1),\ x^1 \in \chi\}$:

$$\mathcal{R}_{l_1} \triangleq \{Q :\ D(Q||G_{l_1}) \leq E_{l_1|l_1}\}, \quad l_1 = \overline{1, L_1 - 1} \tag{5}$$

$$\mathcal{R}_{L_1} \triangleq \{Q :\ D(Q||G_{l_1}) > E_{l_1|l_1}\}, \quad l_1 = \overline{1, L_1 - 1}. \tag{6}$$

and the elements of the reliability matrix E^* of the LAO test:

$$E^*_{l_1|l_1} = E^*_{l_1|l_1}(E_{l_1|l_1}) \triangleq E_{l_1|l_1}, \quad l_1 = \overline{1, L_1 - 1}, \tag{7}$$

$$E^*_{l_1|m_1} = E^*_{l_1|m_1}(E_{l_1|l_1}) \triangleq \inf_{Q \in R_{l_1}} D(Q||G_{m_1}),\ m_1 = \overline{1, L_1},\ m_1 \neq l_1,\ l_1 = \overline{1, L_1 - 1}, \tag{8}$$

$$E^*_{l_1|m_1} = E^*_{L_1|m_1}(E_{1|1}, E_{2|2}, \ldots, E_{L_1-1|L_1-1}) \triangleq \inf_{Q \in R_{L_1}} D(Q||G_{m_1}),\ m_1 = \overline{1, L_1 - 1},\ m_1 \neq l_1,\ l_1 = \overline{1, L_1 - 1}, \tag{9}$$

$$E^*_{L_1|L_1} = E^*_{L_1|L_1}(E_{1|1}, E_{2|2}, \ldots, E_{L_1-1|L_1-1}) \triangleq \min_{l_1 = \overline{1, L_1 - 1}} E^*_{l_1|L_1}. \tag{10}$$

Theorem 1.1. *If all distributions $G_{l_1}, l_1 = \overline{1, L_1}$, are different in the sense that $D(G_{l_1}||G_{m_1}) > 0$, $L_1 \neq m_1$ and the positive numbers $(E_{1|1}, E_{2|2}, \dots, E_{L_1-1|L_1-1}$ are such that the inequalities hold*

$$\begin{aligned} &E_{1|1} < \min_{l_1=\overline{2,L_1}} D(G_{l_1}||G_1) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &E_{m_1|m_1} < \min(\min_{l_1=\overline{m_1+1,L_1}} D(G_{l_1}||G_{m_1}), \min_{l_1=\overline{1,m_1-1}} E^*_{l_1|m_1}(E_{l_1|l_1})),\ m_1 = \overline{2, L_1-1} \end{aligned} \tag{11}$$

exists a LAO sequence of tests ϕ_1^*, the reliability matric of which $E^* = \{E_{l_1|m_1}(phi_1^*)\}$ is defined in (7)–(10) and all elements of it are positive.

Inequalities (11) are necessary for existence of tests sequence unth reliability matrix E having in diagonal given elements $E^*_{l_1|l_1}$, $l_1 = \overline{1, L_1-1}$, and all other elements positive.

2. LAO Testing and Identification of the Probability Distributions for Two Stochastically Coupled Objects.

Firest it is necessary to formulate the concept of LAO approach to the identification problem for one object, which was introduced in [1] and [2], see also [9]. We have one object, and there are known $L_1 \geq 2$ possible PDs. Identification Is the answer to the question: whether $r_1 - th$ distribution is correct, or not. As in the testing problem, the answer must be given on the base of a sample x with the help of an appropriate test.
There are two error probabilities for each $r_1 \in [1, L_1]$: the probability $\alpha_{l_1\neq r_1|m_1=r_1}(\phi_N)$ to accept $l-th$ PD different from r_1, when PD r_1 is in reality, and the probability $\alpha_{l_1=r_1|m_1\neq r_1}(\phi_N)$ that r_1 is accepted, when it is not correct.
The probability $\alpha_{l_1\neq r_1|m_1=r_1}(\phi_N)$ coincides with the probability $\alpha_{r_1|r_1}(\phi_N)$ which is equal to $\sum_{l_1:l_1\neq r_1} \alpha_{l_1|r_1}(\phi^N)$.
The corresponding reliability $E_{l_1\neq r_1|m_1=r_1}(\phi)$ is equal to $E_{r_1|r_1}(\phi)$ which satisfies the equality (4).
Reliability approach to identification means to determine the optimal dependence of $E^*_{l_1=r_1|m_1\neq r_1}$ upon given $E^*_{l_1\neq r_1|m_1=r_1} = E^*_{r_1|r_1}$, which can be an assigned value satisfying conditions (11). Solution of this problems uses knowledge of some a priori PD of the hypotheses. The result from paper [1] is valid for the first object.

Theorem 2.1. *In the case of distinct PDs $G_1, G_2, \dots, G_{L_1}$, under condition that the probabilities of all L_1 hypotheses are positive the reliability of $E_{l_1=r_1|m_1\neq r_1}$ for given $E_{l_1\neq r_1|m_1=r_1} = E_{r_1|r_1}$ is the following:*

$$E_{l_1=r_1|m_1\neq r_1}(E_{r_1|r_1}) = \min_{m_1:m_1\neq r_1} \inf_{Q:D(Q||G_{r_1})\leq E_{r_1|r_1}} D(Q||G_{m_1}),\ r_1 \in [1, L_1].$$

The test, which we denote by Φ^N, is a procedure of making decision about unknown indices of PDs on the base of results of N observations (X_l, x_2). For the objects characterize by X_1, X_2 the non-randomized test $\Phi^N(X_l, x_2)$ can be determined by partition of the sample space $(\chi \times \chi)^N$ on $L_1 \times L_2$ disjoint subsets $\mathcal{A}^N_{l_1,l_2} = \{X_1, X_2 : \Phi^N(X_l, x_2) = l_1, l_2\}$, $l_1 = \overline{1, L_1}$, $l_2 = \overline{1, L_2}$ i.e. the set $\mathcal{A}^N_{l_1,l_2}$ consists of vectors X_1, X_2 for which the PD G_{l_1,l_2} is adopted. The true, $l_1, m_1 = \overline{1, L_1}$, $l_2, m_2 = \overline{1, L_2}$ $(m_1, m_2) \neq (l_1, l_2)$ is defined by the set $\mathcal{A}^N_{l_1,l_2}$

$$\alpha^N_{l_1,l_2|m_1,m_2}(\Phi^N) \triangleq G^N_{m_1,m_2}(\mathcal{A}^N_{l_1,l_2}). \tag{12}$$

We define the probability to reject G_{m_1,m_2}, when it is true, as follows

$$\alpha^N_{m_1,m_2|m_1,m_2}(\Phi^N) \triangleq \sum_{(l_1,l_2)\neq(m_1,m_2)} \alpha^N_{l_1,l_2|m_1,m_2}(\Phi^N) = G^N_{m_1,m_2}(\overline{\mathcal{A}^N_{m_1,m_2}}). \tag{13}$$

We study the reliabilities of the sequence of tests Φ

$$E_{l_1,l_2|m_1,m_2}(\Phi) \triangleq \overline{\lim_{N\to\infty}} -\frac{1}{N} \log \alpha^N_{l_1,l_2|m_1,m_2}(\Phi^N), \quad l_1, m_1 = \overline{1, L_1},\ l_2, m_2 = \overline{1, L_2} \tag{14}$$

From (13) and (14) we have

$$E_{m_1,m_2|m_1,m_2}(\Phi) = \min_{(l_1,l_2)\neq(m_1,m_2)} E_{l_1,l_2|m_1,m_2}(\Phi), \quad l_1, m_1 = \overline{1, L_1},\ l_2, m_2 = \overline{1, L_2} \tag{15}$$

We call the matrix $E(\Phi) = \{E_{l_1,l_2|m_1,m_2}(\Phi), l_1, m_1 = \overline{1, L_1},\ l_2, m_2 = \overline{1, L_2}\}$ the reliability matrix of the sequence of tests Φ. Our aim is to investigate the reliability matrix of optimal tests, and the conditions ensuring positivity of all its elements.
We use some notions and estimates from [11], [12]. For given positive numbers $E_{1,1|1,1}, \dots, E_{L_1,L_2-1|L_1,L_2-1}$ let us consider the following sets of $PDs\ QoV \triangleq \{Q(x^1)V(x^2|x^1), x^1, x^2 \in \chi\}$:

$$\mathcal{R}_{l_1,l_2} \triangleq \{QoV : D(QoV||G_{l_1,l_2}) \leq E_{l-1,l_2|l_1,l_2}\},\ l_1 = \overline{1, L_1},\ l_2 = \overline{1, L_2 - 1} \tag{16}$$

$$\mathcal{R}_{L_1,L_2} \triangleq \{QoV : D(QoV||G_{l_1,l_2}) > E_{l-1,l_2|l_1,l_2}\},\ l_1 = \overline{1, L_1},\ l_2 = \overline{1, L_2 - 1} \tag{17}$$

and the elements of the reliability matrix E^* of the LAO test:

$$E^*_{l_1,l_2|l_1,l_2} = E^*_{l_1,l_2|l_1,l_2}(l_1, l_2|l_1, l_2) \triangleq E_{l_1,l_2|l_1,l_2}, \quad l_1 = \overline{1, L_1},\ l_2 = \overline{1, L_2 - 1}, \tag{18}$$

$$\begin{aligned} &E^*_{l_1,l_2|m_1,m_2} = E^*_{l_1,l_2|m_1,m_2}(E_{l_1,l_2|l_1,l_2}) \triangleq \inf_{QoV \in R_{l_1,l_2}} D(QoV||G_{m_1,m_2}), \\ &m_1 = \overline{1, L_1},\ m_2 = \overline{1, L_2}\ (m_1, m_2) \neq (l_1, l_2),\ l_1 = \overline{1, L_1},\ l_2 = \overline{1, L_2 - 1}, \end{aligned} \tag{19}$$

$$\begin{aligned} E^*_{L_1,L_2|m_1,m_2} = E^*_{L_1,L_2|m_1,m_2}(E_{1,1|1,1}, E_{1,2|1,2}, \dots, E_{L_1,L_2-1|L_1,L_2-1}) \triangleq \\ \inf_{QoV \in R_{L_1,L_2}} D(QoV||G_{m_1,m_2}),\ m_1 = \overline{1, L_1},\ m_2 = \overline{1, L_2} \end{aligned} \tag{20}$$

$$E^*_{L_1,L_2|L_1,L_2} = E^*_{L_1,L_2|L_1,L_2}(E_{1,1|1,1}, E_{1,2|1,2}, \dots, E_{L_1,L_2-1|L_1,L_2-1}) \triangleq \min_{l_1=\overline{1,L_1}} \min_{l_2=\overline{1,L_2-1}} E^*_{l_1,l_2|L_1,L_2}. \tag{21}$$

For simplicity we can take $\overline{(X_l, X_2)} = Y$, $\chi \times \chi = \mathcal{Y}$ and $y = (y_1, y_2, \dots, y_N) \in \mathcal{Y}$, where $y_n = (x^1_n, x^2_n)$, $n = \overline{1, N}$, then we will have $L_1 \times L_2 = L$ new hypotheses for one object $G_{1,1}(X_1, X_2) = F_1(y)$, $G_{1,2}(X_1, X_2) = F_2(y)$, $G_{1,3}(X_1, X_2) = F_3(y)$,. . .,$G_{1,L_2}(X_1, X_2) = F_{L_2}(y)$, $G_{2,1}(X_1, X_2) = F_{L_1+1}(y)$,. . .,$G_{l_1,l_2}(X_1, X_2) = F_{(l_1-1)L_1+L_2}(y)$, $l_1 = \overline{1, L_1}$, $l_1 = \overline{1, L_2}$, $\alpha_{l_1,l_2|m_1,m_2} = \alpha^{'}_{l|m}$, $l_1 = \overline{1, L_1}$, $l_1 = \overline{1, L_2}$, $E_{l_1,l_2|m_1,m_2} = E^{'}_{l|m}$, $l_1 = \overline{1, L_1}$, $l_1 = \overline{1, L_2}$, and thus we have brought the original problem to the case of one object with L hypotheses. Now we can reformulate Theorem 1.1 and Theorem 2.1 for these notations:

Theorem 2.2. *If all distributions F_l, $l = \overline{1, L}$, are different in the sense that $D(F_l||F_m) > 0$, $l \neq m$, and the positive numbers $E^{'}_{1|1}$,$E^{'}_{2|2}$,. . ., $E^{'}_{L-1|L-1}$ are such that the following inequalities hold*

$$\begin{aligned} &E^{'}_{1|1} < \min_{l=\overline{2,L}} D(F_l||F_l) \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ &E^{'}_{m|m} < \min(\min_{l=\overline{m+1,L}} D(F_1||F_{m_1}), \min_{l=\overline{1,m-1}} E^{*'}_{l|m}(E^{'}_{l|l})),\ m_1 = \overline{2, L} \end{aligned} \tag{22}$$

then there exists a LAO sequence of tests Φ^*, the reliability matric of which $E^{*'} = \{E^{'}_{l|m}(\Phi^*)\}$ is defined in (18)-(21) and all elements of it are positive.
Inequalities (22) are necessary for ecistence of tests sequence with reliability matric E' hving in diagonal given elements ${E^*}'_{l|l_1}$, $l = \overline{1, L-1}$, and all other elements positive.

Theorem 2.3. *In the case of distinct PDs $F_1, F_2, \dots, F_L$, under condition that the probabilities of all L hypotheses are positive the reliability of $E^{'}_{l=r_1|m\neq r_1}$ for given $E^{'}_{l\neq r_1|m=r_1} = E_{r_1|r_1}$ is the following:*

$$E^{'}_{l=r_1|m\neq r_1}(E^{'}_{r|r}) = \min_{m:m\neq r_1} \inf_{QoV:D(QoV||F_r)\leq E^{'}_{r|r}} D(QoV||F_m),\ r \in [1, L].$$

References

[1] R. F. Ahlswede and E. A. Haroutunian, ”On logarithmically asymptotically optimal testing of hypotheses and identification”. Lecture Notes in Computer Science, vol. 4123 ”General Theory of Information Transfer and Combinatorics” , Springer, pp. 462–478. 2006.

[2] E. A. Haroutunian. ”Reliability in multiple hypotheses testing and identification”. proceedings of the NATO ASI. Yerevan 2003, NATO Science Series. III: Computer and Systems Sciences. vol. 198, IOS Press, pp. 189-201. 2005.

[3] E. A. Haroutunian and P. M. Hakobyan, ” On LAO testing of multiple hypotheses for pair of objects:• . Mathematical Problems of Computer Science. vol. XX V, pp. 92–100. 2006.

[4] E. A. Haroutunian and P. M. Hakobyan, ”On identification of distributions of two independent objects”. Mathematical Problems of Computer Science. vol. XXVIII. pp.114–119. 2007.

[5] E. A. Haroutunian and P. M. Hakobyan . O’Multiple hypotheses LAO testing for many independent object”. International journal ”Scholarly Research Ecchange”. 15pp. 2009.

[6] E. A. Haroutunian and A. O. Yessayan, On hypotheses testing for two differentlly distributed objects” . Mathematical Problems of Computer Science, vol. XXVI. pp. 91–96. 2006.

[7] E. A. Haroutunian and A. O. Yessayan, ” On logarithmically asymptotically optimal hypothesis testing for pair of statistically dependent objects” Mathematical of Computer Science, vol. XXIX, pp. 97—103, 2007.

[8] E. A. Haroutunian and N. A. Grigoryan , ” On reliability approach for testing of many distributions for pair of Markov chains” , Mathematical Problems of Computer Science vol. XXIX. pp. 89- 96, 2007.

[9] A. O. Yessayan. On reliability approach to identification of probability distributtions of two statistically dependent objects” , Mathematical Problems of Computer Science vol. XXXII. pp. 65– 69. 2009.

[10] E. A. Haroutunian and L. Navaei. ”On optimal identification of Markov chain distributtions subject to the reliability criteria” , Mathematical Problems of Computer Science vol.XXXII. pp. 56–64, 2009.

[11] I. Csiszar and J. Korner. Information theory: Coding Theorems for Discrete Memoryless Systems. New York: Academic Press, 1981 (Russian translation, Mir, Moscow, 1985).

[12] E. A. Haroutunian, M. E. Haroutunian, and A. N. Harutyunyan, ” Reliability criteria in information theory and in statistical hypotheses testing” , Foundations and Trends in Communications and Information Theory, vol. 4, no. 2-3, 2008.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–264, pp. 290–296

Parameter estimation for the Weibull-half-logistic distribution based on lower record values and inter-record times

Mohammad Vali Ahmadi[a,*], Gholam Reza Ahmadi[b]

[a] *Department of Statistics, University of Bojnord, Bojnord, Iran*
[b] *Department of Mathematics, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
Record data
Weibull-half-logistic distribution
Maximum likelihood estimation
Bayesian estimation
Lindley's approximation

2020 MSC:
62Fxx
62N01
62N02
62N05

Abstract

Here, the classical and Bayesian inferences for estimating the parameters of the Weibull-half-logistic distribution have been studied, when the lower record values along with the number of observations following the record values (inter-record times) have been observed. In classical inference, the maximum likelihood estimation and the asymptotic confidence intervals for the parameters are obtained. In Bayesian inference, two approximation Bayes estimates and highest posterior density intervals for the parameters are discussed. The Bayes estimates of the parameters have been provided by using Lindley's approximation and Markov Chain Monte Carlo method. A data set is also analyzed for illustration aims.

1. Introduction

The Weibull distribution is one of the most widely used distributions in the reliability and survival studies. The exponential and Rayleigh distributions are two special cases of this model. For more detail on applications of the Weibull distribution, see, for example, Murthy et al. [10] and some references therein. The cumulative distribution function (CDF) of the Weibull distribution is $F_{\alpha,\beta}(x) = 1 - e^{-\alpha x^{\beta}}$, $x > 0$ with parameters $\alpha > 0$ and $\beta > 0$. Let $G_\theta(x)$ be a continuous baseline CDF, where θ is a parameter vector. By replacing x with $G_\theta(x)/(1 - G_\theta(x))$ in Weibull CDF, the CDF of Weibull-G distribution is defined by

$$F_{\alpha,\beta,\theta}(x) = 1 - \exp\left\{-\alpha\left(\frac{G_\theta(x)}{1 - G_\theta(x)}\right)^{\beta}\right\}, \ x \in D \subset \mathbb{R},\ \alpha > 0,\ \beta > 0,\ \theta \in \Theta. \tag{1}$$

* Talker
Email addresses: mv.ahmadi@ub.ac.ir (Mohammad Vali Ahmadi), ghrahmadi@pnu.ac.ir (Gholam Reza Ahmadi)

For half-logistic distribution, we suppose $G(x) = (e^x - 1)/(e^x + 1)$ and then $G(x)/(1 - G(x)) = (e^x - 1)/2$. So, from (1), the CDF of the Weibull-half-logistic (WHL) distribution is defined by

$$F_{\alpha,\beta}(x) = 1 - e^{-\alpha 2^{-\beta}(e^x - 1)^\beta}, \; x > 0. \tag{2}$$

Hence, the probability density function (PDF) of the WHL distribution becomes

$$f_{\alpha,\beta}(x) = \alpha\beta 2^{-\beta} \left(e^x - 1\right)^{\beta-1} e^{x - \alpha 2^{-\beta}(e^x - 1)^\beta}, \; x > 0. \tag{3}$$

Let $X_1, X_2, \ldots$ be a sequence of identical and independent continuous random variables. The random variable X_i is a lower record value if its value is smaller than all preceding values $X_1, X_2, \ldots, X_{i-1}$. Considering this fact, the sequence of lower record values can be rewritten as $(\boldsymbol{W}, \boldsymbol{K}) \equiv (W_1, K_1, W_2, K_2 \ldots, W_r, K_r)$ where W_i is the ith record value or new minimum and K_i is the number of trials following the observation of W_i that are needed to obtain a new record value W_{i+1}. Also, we set $K_r = 1$. The values $K_1, \ldots, K_r$ are called inter-record times. An analogous definition can be provided for upper record values.

Statistical inferences based on record data have been discussed extensively in the literature, see, for example, Arnold et al. [1]. The Bayesian and non-Bayesian estimates for the parameters of the two-parameter exponential distribution based on record values and their corresponding inter-record times was considered by Doostparast [3]. Optimal confidence intervals of the parameters as well as uniformly most powerful tests for one-sided alternatives and generalized likelihood ratio and uniformly unbiased and invariant tests for two-sided alternatives were derived by Doostparast and Balakrishnan [4]. The optimal statistical procedures including point and interval estimation as well as most powerful tests based on record data from a two-parameter Pareto model were studied by Doostparast and Balakrishnan [5]. The Bayesian and non-Bayesian point estimates as well as asymptotic confidence intervals for the parameters of the lognormal distribution were obtained by Doostparast et al. [6]. When the underlying distribution is the generalized exponential distribution, Bayesian and non-Bayesian point estimates as well as asymptotic confidence intervals and highest probability density credible intervals for the parameters were obtained by Kizilaslan and Nadar [8]. Also, Pak and Dey [11] developed inference procedures for the estimation of the parameters and prediction of future record values for the power Lindley distribution using record data and inter-record times.

In this paper, based on a sequence of record data from the WHL distribution, the unknown parameters of the baseline distribution are estimated. The rest of the paper is organized as follows. In Section 2, we describe the construction of the likelihood function for record data and corresponding inter-record times and then obtain the Maximum likelihood (ML) estimates of the parameters of distribution. In Section 3, we provide the asymptotic confidence intervals for the unknown parameters. The Bayesian estimates are another estimates which we will obtain. Assuming statistically independent gamma priors for the parameters and SE loss function, the Bayes estimates of the parameters are discussed by applying the Lindley's approximation and the Markov chain Monte Carlo (MCMC) method in Section 4. Moreover, the highest probability density (HPD) credible intervals for the parameters are derived on the basis of the MCMC method in Section 4. Section 5 includes an illustrative example.

2. Maximum likelihood estimation

Suppose that $(\boldsymbol{W}, \boldsymbol{K})$ is be a sequence of record data from the WHL distribution with CDF and PDF (2) and (3), respectively. Then, following Samaniego and Whitaker [12], the likelihood function associated with the sequence $(\boldsymbol{W}, \boldsymbol{K})$ is given by

$$\begin{aligned} L \equiv L(\alpha, \beta | \boldsymbol{w}, \boldsymbol{k}) &= \prod_{i=1}^{r} f_{\alpha,\beta}(w_i) \left(1 - F_{\alpha,\beta}(w_i)\right)^{k_i - 1} \\ &= \left(\alpha\beta 2^{-\beta}\right)^r e^{-\alpha 2^{-\beta} \sum_{i=1}^{r} k_i (e^{w_i} - 1)^\beta} \prod_{i=1}^{r} e^{w_i} (e^{w_i} - 1)^{\beta - 1}, \end{aligned} \tag{4}$$

and the corresponding log-likelihood function is

$$\log L = r\left(\log\alpha + \log\beta - \beta\log 2\right) - \alpha 2^{-\beta}\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\beta} + (\beta-1)\sum_{i=1}^{r}\log(e^{w_i}-1) + \sum_{i=1}^{r} w_i. \tag{5}$$

Taking the first partial derivatives of log-likelihood (5) with respect to α and β and equating each to zero, we obtain

$$\frac{\partial \log L}{\partial \alpha} = \frac{r}{\alpha} - 2^{-\beta}\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\beta} = 0, \tag{6}$$

and

$$\frac{\partial \log L}{\partial \beta} = r\left(\frac{1}{\beta} - \log 2\right) - \alpha 2^{-\beta}\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\beta}\log\frac{e^{w_i}-1}{2} + \sum_{i=1}^{r}\log(e^{w_i}-1) = 0. \tag{7}$$

From (6), we have

$$\hat{\alpha} = r 2^{\hat{\beta}}\left[\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\hat{\beta}}\right]^{-1}, \tag{8}$$

where $\hat{\beta}$ is the solution of the nonlinear equation in (7). Therefore, $\hat{\beta}$ can be obtained by using the fixed point method or Newton–Raphson method or other numerical methods.

3. Asymptotic Confidence Interval

In this section, we obtain the asymptotic confidence intervals for α and β by the asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$. First, we need to compute the Fisher information matrix given by

$$I(\alpha,\beta) = \begin{bmatrix} -\mathrm{E}\left(\frac{\partial^2 \log L}{\partial \alpha^2}\right) & -\mathrm{E}\left(\frac{\partial^2 \log L}{\partial \alpha\partial\beta}\right) \\ -\mathrm{E}\left(\frac{\partial^2 \log L}{\partial \beta\partial\alpha}\right) & -\mathrm{E}\left(\frac{\partial^2 \log L}{\partial \beta^2}\right)\end{bmatrix} = \begin{bmatrix} I_{11} & I_{12} \\ I_{21} & I_{22}\end{bmatrix}. \tag{9}$$

But, it is difficult to obtain exact mathematical expectations I_{11}, I_{12}, I_{21} and I_{22} in (9). Therefore, we take the observed Fisher information matrix as

$$\hat{I}(\alpha,\beta) = \begin{bmatrix} -\frac{\partial^2 \log L}{\partial \alpha^2} & -\frac{\partial^2 \log L}{\partial \alpha\partial\beta} \\ -\frac{\partial^2 \log L}{\partial \beta\partial\alpha} & -\frac{\partial^2 \log L}{\partial \beta^2}\end{bmatrix} = \begin{bmatrix} \hat{I}_{11} & \hat{I}_{12} \\ \hat{I}_{21} & \hat{I}_{22}\end{bmatrix}. \tag{10}$$

From (6) and (7), we have

$$\hat{I}_{11} = \frac{r}{\alpha^2}, \quad \hat{I}_{12} = \hat{I}_{21} = 2^{-\beta}\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\beta}\log\frac{e^{w_i}-1}{2}, \tag{11}$$

$$\hat{I}_{22} = \frac{r}{\beta^2} + \alpha 2^{-\beta}\sum_{i=1}^{r} k_i (e^{w_i}-1)^{\beta}\left(\log\frac{e^{w_i}-1}{2}\right)^2. \tag{12}$$

Theorem 3.1. Let $\hat{\alpha}$ and $\hat{\beta}$ be the ML estimators for α and β. So

$$\begin{bmatrix}\hat{\alpha}-\alpha \\ \hat{\beta}-\beta\end{bmatrix} \xrightarrow{D} \mathrm{N}_2\left(0, \hat{I}^{-1}(\alpha,\beta)\right), \tag{13}$$

where $\xrightarrow{D}$ denotes convergence in distribution and $\hat{I}^{-1}(\alpha, \beta)$ is the inverse of the matrix $\hat{I}(\alpha, \beta)$ with

$$\hat{I}^{-1}(\alpha, \beta) = \frac{1}{\hat{I}_{11}\hat{I}_{22} - \hat{I}_{12}\hat{I}_{21}} \begin{bmatrix} \hat{I}_{22} & -\hat{I}_{12} \\ -\hat{I}_{21} & \hat{I}_{11} \end{bmatrix} = \begin{bmatrix} \mathrm{Var}(\hat{\alpha}) & \mathrm{Cov}(\hat{\alpha}, \hat{\beta}) \\ \mathrm{Cov}(\hat{\alpha}, \hat{\beta}) & \mathrm{Var}(\hat{\beta}) \end{bmatrix}. \tag{14}$$

Proof. From the asymptotic normality of the ML estimators, the theorem resulted. □

On the basis of Theorem 3.1, the asymptotic $100(1-\delta)\%$ confidence intervals for α and β are

$$\hat{\alpha} \pm z_{1-\frac{\delta}{2}} \sqrt{\widehat{\mathrm{Var}(\hat{\alpha})}}, \tag{15}$$

and

$$\hat{\beta} \pm z_{1-\frac{\delta}{2}} \sqrt{\widehat{\mathrm{Var}(\hat{\beta})}}, \tag{16}$$

where z_γ is the lower γ-th quantile of the standard normal distribution. respectively.
The confidence intervals (15) and (16) for α and β, respectively, may lead to negative lower bounds. For this reason, we apply the logarithmic transformation and use the delta method to obtain the asymptotic normality distribution of $\log \hat{\alpha}$ and $\log \hat{\beta}$, respectively, as

$$\log \hat{\alpha} - \log \alpha \xrightarrow{D} \mathrm{N}\left(0, \frac{\mathrm{Var}(\hat{\alpha})}{\alpha^2}\right), \tag{17}$$

and

$$\log \hat{\beta} - \log \beta \xrightarrow{D} \mathrm{N}\left(0, \frac{\mathrm{Var}(\hat{\beta})}{\beta^2}\right). \tag{18}$$

Now, the asymptotic $100(1-\delta)\%$ confidence intervals for $\log \alpha$ and $\log \beta$ are

$$\log \hat{\alpha} \pm z_{1-\frac{\delta}{2}} \frac{\sqrt{\widehat{\mathrm{Var}(\hat{\alpha})}}}{\hat{\alpha}} \equiv (L_1, U_1),$$

and

$$\log \hat{\beta} \pm z_{1-\frac{\delta}{2}} \frac{\sqrt{\widehat{\mathrm{Var}(\hat{\beta})}}}{\hat{\beta}} \equiv (L_2, U_2),$$

respectively. Finally, using the inverse logarithmic transformation, the asymptotic $100(1-\delta)\%$ confidence intervals for α and β are derived, respectively, as

$$(e^{L_1}, e^{U_1}), \tag{19}$$

and

$$(e^{L_2}, e^{U_2}). \tag{20}$$

4. Bayes Estimation

Bayesian inference is a useful method for analyzing record-breaking data. Since records are often scarce, incorporating prior information becomes invaluable. In the Bayesian inference, the most commonly used loss function is the squared error (SE) loss, $L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$, where $\hat{\theta}$ is an estimator for θ. The Bayes estimate of θ under SE loss function is the posterior mean of θ.
In this section, we derive the Bayes estimators for the parameters α and β by using the SE loss function. We assume that α and β are random variables that follow the gamma prior distributions with PDFs as

$$\pi_1(\alpha) = \frac{b_1^{a_1}}{\Gamma(a_1)} \alpha^{a_1 - 1} e^{-b_1 \alpha}, \tag{21}$$

and

$$\pi_2(\beta) = \frac{b_2^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}e^{-b_2\beta}, \tag{22}$$

respectively, where the positive hyperparameters a_i and b_i for $i = 1, 2$ are chosen to reflect the prior knowledge about α and β.

Using (4), (21) and (22), the joint posterior PDF of α and β can be written as

$$\pi(\alpha, \beta|\boldsymbol{w}, \boldsymbol{k}) = \frac{L(\alpha, \beta|\boldsymbol{w}, \boldsymbol{k})\pi_1(\alpha)\pi_2(\beta)}{\int_0^\infty \int_0^\infty L(\alpha, \beta|\boldsymbol{w}, \boldsymbol{k})\pi_1(\alpha)\pi_2(\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta}. \tag{23}$$

As shown in the joint posterior PDF (23), it is not possible to obtain the Bayes estimates of parameters in closed form. Therefore, we employ Lindley's approximation to obtain them. Following Lindley [9], this method can be demonstrated as follows.

Suppose that $g(\lambda)$ is a desired function of parameter λ. The Bayes estimate of $g(\lambda)$, under the SE loss function, is

$$\mathrm{E}(g(\lambda)|\text{data}) = \frac{\int g(\lambda)e^{Q(\lambda)}\mathrm{d}\lambda}{\int e^{Q(\lambda)}\mathrm{d}\lambda}, \tag{24}$$

where $Q(\lambda) = l(\lambda) + \rho(\lambda)$, $l \equiv l(\lambda)$ and $\rho \equiv \rho(\lambda)$ are the logarithm of the likelihood function and the logarithm of the prior density of λ, respectively. Lindley approximated $\mathrm{E}(g(\lambda)|\text{data})$ in (24) as

$$\mathrm{E}(g(\lambda)|\text{data}) \cong g + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(g_{ij} + 2g_i\rho_j)\sigma_{ij} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{h=1}^{m} l_{ijk}\sigma_{ij}\sigma_{kh}g_h\Big|_{\lambda=\hat{\lambda}},$$

where $\lambda = (\lambda_1, \ldots, \lambda_m)$, $g \equiv g(\lambda)$, $g_i = \partial g/\partial\lambda_i$, $g_{ij} = \partial^2 g/\partial\lambda_i\partial\lambda_j$, $l_{ijk} = \partial^3 l/\partial\lambda_i\partial\lambda_j\partial\lambda_k$, $\rho_j = \partial\rho/\partial\lambda_j$, σ_{ij} is (i, j)-th element in the inverse of the matrix $[-l_{ij}]$ with $l_{ij} = \partial^2 l/\partial\lambda_i\partial\lambda_j$. Also, $\hat{\lambda}$ is the ML estimate of λ.

In the case of two parameters $\lambda = (\lambda_1, \lambda_2)$, Lindley's approximation leads to

$$\begin{aligned}\mathrm{E}(g(\lambda)|\text{data}) \cong\ & g + g_1(\rho_1\sigma_{11} + \rho_2\sigma_{12}) + g_2(\rho_1\sigma_{21} + \rho_2\sigma_{22}) + g_{12}\sigma_{12} + \frac{1}{2}(g_{11}\sigma_{11} + g_{22}\sigma_{22}) \\ & + \frac{1}{2}\Big[(l_{111}\sigma_{11} + 2l_{121}\sigma_{12} + l_{221}\sigma_{22})(g_1\sigma_{11} + g_2\sigma_{12}) \\ & + (l_{112}\sigma_{11} + 2l_{122}\sigma_{12} + l_{222}\sigma_{22})(g_1\sigma_{21} + g_2\sigma_{22})\Big],\end{aligned}$$

calculated at $\hat{\lambda} = (\hat{\lambda}_1, \hat{\lambda}_2)$. In our case $(\lambda_1, \lambda_2) = (\alpha, \beta)$. From (11) and (12), we have

$$\begin{aligned} l_{111} &= \frac{2r}{\alpha^3}, \quad l_{112} = l_{121} = l_{211} = 0, \\ l_{221} &= l_{212} = l_{122} = -2^{-\beta}\sum_{i=1}^{r} k_i\,(e^{w_i} - 1)^\beta \left(\log\frac{e^{w_i} - 1}{2}\right)^2, \\ l_{222} &= \frac{2r}{\beta^3} - \alpha 2^{-\beta}\sum_{i=1}^{r} k_i\,(e^{w_i} - 1)^\beta \left(\log\frac{e^{w_i} - 1}{2}\right)^3. \end{aligned}$$

Using (21) and (22), it is easy to see that $\rho_1 = (a_1 - 1)/\alpha - b_1$ and $\rho_2 = (a_2 - 1)/\beta - b_2$. When $g(\alpha, \beta) = \alpha$, we obtain $g_1 = 1$, $g_2 = 0$ and $g_{ij} = 0$ for $i, j = 1, 2$. Also, when $g(\alpha, \beta) = \beta$, we have $g_1 = 0$, $g_2 = 1$ and $g_{ij} = 0$ for $i, j = 1, 2$. Hence, the Bayes estimators for the parameters α and β under the SE loss function are obtained as

$$\hat{\alpha}_{\text{Lind}} \cong \alpha + \rho_1\sigma_{11} + \rho_2\sigma_{12} + \frac{1}{2}\Big[\sigma_{11}(l_{111}\sigma_{11} + l_{221}\sigma_{22}) + \sigma_{21}(2l_{122}\sigma_{12} + l_{222}\sigma_{22})\Big], \tag{25}$$

$$\hat{\beta}_{\text{Lind}} \cong \beta + \rho_1\sigma_{21} + \rho_2\sigma_{22} + \frac{1}{2}\Big[\sigma_{12}(l_{111}\sigma_{11} + l_{221}\sigma_{22}) + \sigma_{22}(2l_{122}\sigma_{12} + l_{222}\sigma_{22})\Big], \tag{26}$$

respectively. All parameters (α, β) in (25) and (26) are replaced by the ML estimators, i.e., $(\hat{\alpha}, \hat{\beta})$.
Clearly, constructing the HPD credible interval is not possible, by applying the Lindley's approximation. So, we approximate the Bayes estimates of the parameters and construct the corresponding HPD credible intervals using the MCMC method.
Referring to (23), it is easy to show that the conditional posterior PDFs of α and β are as follows:

$$\alpha|\beta, \text{data} \sim \Gamma\left(r + a_1, 2^{-\beta}\sum_{i=1}^{r} k_i(e^{w_i} - 1)^{\beta} + b_1\right),$$

$$\pi\left(\beta|\alpha, \text{data}\right) \propto \beta^{r+a_2-1}2^{-r\beta}e^{-\alpha 2^{-\beta}\sum_{i=1}^{r} k_i(e^{w_i}-1)^{\beta} - b_2\beta}\prod_{i=1}^{r}\left(e^{w_i} - 1\right)^{\beta-1}. \tag{27}$$

Note that the conditional posterior PDF in (27) is not a well-known distribution and, therefore, it is not possible to generate a sample from this distribution, directly. When a posterior PDF is roughly symmetric and unimodal, it can be approximated by using the normal distribution, see, for example, Gelman et al. [7]. Consequently, we employ the Metropolis-Hastings method with the normal proposal distribution to generate random samples from the distribution with PDF (27). So, the hybrid Metropolis-Hastings and Gibbs sampling algorithm is as follows:

1. Choose an initial value (α_0, β_0) for the parameters (α, β) and set $t = 1$.
2. Using the Metropolis-Hastings algorithm under the proposal distribution $N(\beta, 1)$ with PDF denoted by $q(x|\beta)$, generate β_t from $\pi\left(\beta|\alpha_{t-1}, \text{data}\right)$ as follows:
   - Generate $\beta^\star$ from the proposal distribution $q(x|\beta_{t-1})$.
   - Compute the quantity
   $$p = \min\left\{1, \frac{\pi(\beta^\star|\alpha_{t-1}, \text{data})\, q(\beta_{t-1}|\beta^\star)}{\pi(\beta_{t-1}|\alpha_{t-1}, \text{data})\, q(\beta^\star|\beta_{t-1})}\right\}. \tag{28}$$
   If $p = 1$, then set $\beta_t = \beta^\star$. If $p < 1$, then set $\beta_t = \beta^\star$ with probability p and $\beta_t = \beta_{t-1}$ with probability $1 - p$. To do this, generate u from the standard uniform distribution. If $u < p$, accept $\beta^\star$, otherwise accept β_{t-1}. Note that the property of symmetry of the normal distribution causes $q(\beta_{t-1}|\beta^\star) = q(\beta^\star|\beta_{t-1})$ and so the quantity p in (28) can be simplified to $p = \min\{1, \pi(\beta^\star|\alpha_{t-1}, \text{data})/\pi(\beta_{t-1}|\alpha_{t-1}, \text{data})\}$.
3. Generate α_t from $\Gamma\left(r + a_1, 2^{-\beta_t}\sum_{i=1}^{r} k_i(e^{w_i} - 1)^{\beta_t} + b_1\right)$.
4. Set $t = t + 1$.
5. Reiterate steps 2–4, T times and obtain (α_t, β_t) for $t = 1, \ldots, T$.

By using the random sample generated from the above Gibbs sampling algorithm, the Bayes estimators for the parameters α and β under SE loss function are approximated by

$$\hat{\alpha}_{\text{MC}} \cong \frac{1}{T}\sum_{t=1}^{T}\alpha_t, \tag{29}$$

and

$$\hat{\beta}_{\text{MC}} \cong \frac{1}{T}\sum_{t=1}^{T}\beta_t, \tag{30}$$

respectively. Also, using the method of Chen and Shao [2], a $100(1-\delta)\%$ HPD credible interval for the parameters α and β is provided as follows. Let $\alpha_1^\star < \cdots < \alpha_T^\star$ be the ordered values of $\alpha_1, \ldots, \alpha_T$. The HPD credible interval for α is the shortest length interval through the following $100(1-\delta)\%$ intervals

$$\left(\alpha_t^\star, \alpha_{t+[T(1-\delta)]}^\star\right) \quad \text{for} \quad t = 1\ldots, T - [T(1-\delta)],$$

where $[\gamma]$ denotes the largest integer less than or equal to γ. The $100(1-\delta)\%$ HPD credible interval for β is derived in a similar manner.

5. Illustrative eample

In order to explain how the methods proposed in previous sections can be applied in practice, we consider a simulated data set for $(\boldsymbol{W}, \boldsymbol{K})$ generated form the WHL distribution with parameters $\alpha = 2$ and $\beta = 3$. This data set is presented in Table 1.

Table 1. A sequence of record data $(\boldsymbol{W}, \boldsymbol{K})$ of size $r = 5$ simulated from WHL distribution with $\alpha = 2$ and $\beta = 3$.

i	1	2	3	4	5
W_i	0.72227	0.25851	0.10432	0.09783	0.05564
K_i	4	155	541	2725	1

Based on the data presented in Table 1 and using Equations (7) and (8), the ML estimates of the parameters are derived $\hat{\alpha} = 1.76468$ and $\hat{\beta} = 2.67931$. On the basis of Equations (19) and (20), the asymptotic 95% confidence intervals for the parameters α and β are obtained, respectively, as $(0.24485, 12.71835)$ and $(1.93238, 3.71495)$. Under the Bayesian framework, we assumed that the hyperparameters in prior densities (21) and (22) are $(a_1, b_1) = (1, 0.1)$ and $(a_2, b_2) = (2, 1)$. From (25) and (26), the Bayes estimates of α and β under Lindley's approximation are $\hat{\alpha}_{\text{Lind}} = 3.27719$ and $\hat{\beta}_{\text{Lind}} = 2.86272$, respectively. Also, from (29) and (30), the Bayes estimates of these parameters using the MCMC method are $\hat{\alpha}_{\text{MC}} = 3.64587$ and $\hat{\beta}_{\text{MC}} = 2.18444$, respectively. Finally, the 95% HPD credible intervals for α and β are, respectively, as $(0.18784, 7.1039)$ and $(2.18444, 3.42092)$.

References

[1] B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja. *Records*. John Wiley, New York, 1998.
[2] M. H. Chen and Q. M. Shao. Monte carlo estimation of bayesian credible and hpd intervals. *Journal of computational and Graphical Statistics*, 8:69–92, 1999.
[3] M. Doostparast. A note on estimation based on record data. *Metrika*, 69:69–80, 2009.
[4] M. Doostparast and N. Balakrishnan. Optimal record-based statistical procedures for the two-parameter exponential distribution. *Journal of Statistical Computation and Simulation*, 81:2003–2019, 2011.
[5] M. Doostparast and N. Balakrishnan. Pareto analysis based on records. *Statistics*, 47:1075–1089, 2013.
[6] M. Doostparast, S. Deepak, and A. Zangoie. Estimation with the lognormal distribution on the basis of records. *Journal of Statistical Computation and Simulation*, 83:2339–2351, 2013.
[7] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, New York, 1998.
[8] F. Kizilaslan and M. Nadar. Estimation with the generalized exponential distribution based on record values and inter-record times. *Journal of Statistical Computation and Simulation*, 85:978–999, 2015.
[9] D. V. Lindley. Approximate bayesian methods. *Trabajos de Estadistica*, 31:223–245, 1980.
[10] D. N. P. Murthy, M. Xie, and R. Jiang. *Weibull Models*. John Wiley, New York, 2004.
[11] A. Pak and S. Dey. Statistical inference for the power lindley model based on record values and inter-record times. *Journal of Computational and Applied Mathematics*, 347:156–172, 2019.
[12] F. J. Samaniego and L. R. Whitaker. On estimating popular characteristics from record breaking observations. i. parametric results. *Naval Research Logistics*, 33:531–543, 1986.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–265, pp. 297–304

Bayesian estimation and prediction based on upper record values from a Weibull-half-logistic distribution

Mohammad Vali Ahmadi[a,*], Gholam Reza Ahmadi[b]

[a]*Department of Statistics, University of Bojnord, Bojnord, Iran*
[b]*Department of Mathematics, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
Upper record values
Weibull-half-logistic distribution
Bayesian inference
Lindley's approximation
Prediction

2020 MSC:
62Fxx
62N01
62N02
62N05

Abstract

The problem of point and interval estimation of the two parameters of the Weibull-half-logistic distribution through a Bayesian approach is discussed, when the upper record values have been observed. The Bayes estimates are derived on the basis of a bivariate prior distribution for the parameters. To do this, the Bayes estimates are studied by using Lindley's approximation and Markov Chain Monte Carlo method under symmetric and asymmetric loss functions. Also, point and interval predictions for future upper records are provided from a Bayesian viewpoint. A data set is analyzed for illustrating the findings.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed continuous random variables. If the observation X_j is upper (lower) than all previous observations, then it is called a upper (lower) record value. That is, X_j is an upper (lower) record value if $X_j > X_i$ ($X_j < X_i$) for every $i < j$. Generally, let us define $T_1 = 1$ and for $n \geq 2$,

$$T_n = \min\{j : j > T_{n-1}, X_j > X_{T_{n-1}}\},$$

So, the sequence $T_1, T_2, \ldots$ is called the upper record times. Then, the sequence of upper record values becomes

$$R_n = X_{T_n},\ n = 1, 2, \ldots,$$

where $R_1 = X_1$. An analogous definition deals with lower record times and lower record values. Since these record values arise in many practical situations, such as industrial stress testing, meteorology, seismology and athletic events,

*Talker
Email addresses: mv.ahmadi@ub.ac.ir (Mohammad Vali Ahmadi), ghrahmadi@pnu.ac.ir (Gholam Reza Ahmadi)

they are extensively employed in statistical applications and modeling. Properties of record values have been widely discussed in the literature. Chandler [7] studied record values and documented many of basic properties of them. We refer the intersted readers to Glick [10], Arnold et al. [4] and Nevzorov [14] for studing in this area.
Many studies have been done concerning the statistical inference on the basis of record values. See, for example, Ahmadi and Arghami [1], Ahmadi et al. [2], Baklizi [6], Ahsanullah and Shakil [3], Seo and Kim [17] and Zhao et al. [20]. Recently, Asgharzadeh [5] presented the exact confidence intervals and joint confidence regions for the parameters of Gumbel and inverse Weibull distributions based on record data. Also, Piriaei et al. [16] discussed the E-Bayesian estimations for the unknown parameters, when the observed data set is a sequence of record values coming from the exponential distribution.
The Weibull distribution is one of the most widely used distributions in the reliability and survival studies. The exponential and Rayleigh distributions are two special cases of this model. For more detail on applications of the Weibull distribution, see, for example, Murthy et al. [12] and some references therein. The cumulative distribution function (CDF) of the Weibull distribution is $F_{\alpha,\beta}(x) = 1 - e^{-\alpha x^{\beta}}$, $x > 0$ with parameters $\alpha > 0$ and $\beta > 0$. Let $G_\theta(x)$ be a continuous baseline CDF, where θ is a parameter vector. By replacing x with $G_\theta(x)/(1 - G_\theta(x))$ in Weibull CDF, the CDF of Weibull-G distribution is defined by

$$F_{\alpha,\beta,\theta}(x) = 1 - \exp\left\{-\alpha\left(\frac{G_\theta(x)}{1 - G_\theta(x)}\right)^{\beta}\right\},\ x \in D \subset \mathbb{R},\ \alpha > 0,\ \beta > 0,\ \theta \in \Theta. \tag{1}$$

For half-logistic distribution, we suppose $G(x) = (e^x - 1)/(e^x + 1)$ and then $G(x)/(1 - G(x)) = (e^x - 1)/2$. So, from (1), the CDF of the Weibull-half-logistic (WHL) distribution is defined by

$$F_{\alpha,\beta}(x) = 1 - e^{-\alpha 2^{-\beta}(e^x - 1)^{\beta}},\ x > 0. \tag{2}$$

Hence, the probability density function (PDF) of the WHL distribution becomes

$$f_{\alpha,\beta}(x) = \alpha\beta 2^{-\beta}\left(e^x - 1\right)^{\beta - 1} e^{x - \alpha 2^{-\beta}(e^x - 1)^{\beta}},\ x > 0. \tag{3}$$

In this paper, in Section 2, based on upper record data from the WHL distribution, Bayes estimators for the unknown parameters are discussed. In order to obtain the Bayes estimates under symmetric and asymmetric loss functions, Lindley's approximation and Markov Chain Monte Carlo (MCMC) method are used in Subsections 2.1 and 2.2, respectively. The highest probability density (HPD) credible intervals for the unknown parameters are also constructed on the basis of the MCMC sample. In Section 3, the Bayesian approach is employed to develop the point and interval predictors for the future records. Finally, a simulated data set analysis is provided in Section 4.

2. Bayesian estimation

Let $\boldsymbol{R} = (R_1, \ldots, R_m)$ be the first m upper record values from a distribution with PDF $F(\cdot;\theta)$ and CDF $f(\cdot;\theta)$. Then, the likelihood function associated with the observed first m upper record values $\boldsymbol{r} = (r_1, \ldots, r_m)$ (for more details, see Arnold et al. [4]) is given by

$$L(\theta|\boldsymbol{r}) = f(r_m;\theta)\prod_{i=1}^{m-1} h(r_i;\theta), \tag{4}$$

where $-\infty < r_1 < \cdots < r_m < \infty$ and

$$h(r_i;\theta) = \frac{f(r_i;\theta)}{\bar{F}(r_i;\theta)},$$

with $\bar{F}(\cdot;\theta) = 1 - F(\cdot;\theta)$. Under the WHL distribution with CDF (2) and PDF (3), the the likelihood function in (4) is reduced to

$$L(\alpha,\beta|\boldsymbol{r}) = \left(\alpha\beta 2^{-\beta}\right)^m e^{-\alpha 2^{-\beta}(e^{r_m} - 1)^{\beta} + \sum_{i=1}^m r_i}\prod_{i=1}^{m}(e^{r_i} - 1)^{\beta - 1},\ \ \alpha > 0,\ \beta > 0. \tag{5}$$

Following Nadar et al. [13], we consider a bivariate prior distribution for the parameters α and β with the joint prior PDF as

$$\pi(\alpha,\beta) = \pi_1(\beta|\alpha)\pi_2(\alpha), \tag{6}$$

where α has a gamma prior distribution with shape and scale parameters a_2 and b_2, respectively, and the conditional prior distribution of β given α is gamma with shape and scale parameters a_1 and α/b_1, respectively. In other words, we have

$$\pi_1(\beta|\alpha) = \frac{\alpha^{a_1}}{\Gamma(a_1)b_1^{a_1}}\beta^{a_1-1}e^{-\frac{\alpha\beta}{b_1}},\ \beta>0, \tag{7}$$

and

$$\pi_2(\alpha) = \frac{b_2^{a_2}}{\Gamma(a_2)}\alpha^{a_2-1}e^{-b_2\alpha},\ \alpha>0, \tag{8}$$

where (a_1,b_1) and (a_2,b_2) are positive hyperparameters. Using Equations (5) and (6), the joint posterior PDF of α and β can be rewritten as

$$\begin{aligned}\pi(\alpha,\beta|\boldsymbol{r}) &= \frac{L(\alpha,\beta|\boldsymbol{r})\pi(\alpha,\beta)}{\int_0^\infty\int_0^\infty L(\alpha,\beta|\boldsymbol{r})\pi(\alpha,\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta}\\ &= C(\boldsymbol{r})\alpha^{m+a_1+a_2-1}\beta^{m+a_1-1}2^{-m\beta}e^{-\alpha\left[2^{-\beta}(e^{r_m}-1)^\beta+b_2+\frac{\beta}{b_1}\right]}\prod_{i=1}^{m}(e^{r_i}-1)^{\beta-1},\ \ \alpha>0,\ \beta>0,\end{aligned} \tag{9}$$

where

$$[C(\boldsymbol{r})]^{-1} = \Gamma(m+a_1+a_2)\int_0^\infty \frac{\beta^{m+a_1-1}2^{-m\beta}\prod_{i=1}^m(e^{r_i}-1)^{\beta-1}}{\left(2^{-\beta}(e^{r_m}-1)^\beta+b_2+\frac{\beta}{b_1}\right)^{m+a_1+a_2}}\,\mathrm{d}\beta.$$

The choice of loss function is crucial in the Bayesian approach to point estimation. The most commonly employed loss function is the squared error (SE) loss functoin defined as

$$L(\delta,\theta) = (\delta-\theta)^2, \tag{10}$$

where δ is an estimate of θ. This symmetric loss function assigns equal importance to both overestimation and underestimation. However, in many situations, overestimation can have more severe consequences than underestimation, or vice versa. Therefore, the use of symmetric loss functions may not be suitable in such cases. This is particularly true in the estimation of reliability and failure rate functions, where an overestimate is often more consequential than an underestimate. This realization necessitates the adoption of an asymmetric loss function. A valuable alternative to the SE loss function is a convex but asymmetric loss function, called the LINEX (linear-exponential) loss function, which was proposed by Varian [19]. It is defined as

$$L(\delta,\theta) = e^{\gamma(\delta-\theta)} - \gamma(\delta-\theta) - 1, \tag{11}$$

where $\gamma \neq 0$ is the shape parameter. The sign and magnitude of γ determine the direction and degree of asymmetry, respectively. When γ tends to zero, the LINEX loss function converges to the SE loss functon. For more details, see Parsian and Kirmani [15]. The Bayes estimates of δ that minimize $\mathrm{E}\left(L(\delta,\theta)|\text{data}\right)$ can be shown to be $\mathrm{E}(\theta|\text{data})$ and $-\log\mathrm{E}(e^{-\gamma\theta}|\text{data})/\gamma$ under SE and LINEX loss functions in (10) and (11), respectively.
Using (5) and (6), the Bayes estimate of a given measurable function of α and β, say $g(\alpha,\beta)$, under the SE loss function is given by

$$\mathrm{E}(g(\alpha,\beta)|\boldsymbol{r}) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\beta)L(\alpha,\beta|\boldsymbol{r})\pi(\alpha,\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta}{\int_0^\infty\int_0^\infty L(\alpha,\beta|\boldsymbol{r})\pi(\alpha,\beta)\,\mathrm{d}\alpha\,\mathrm{d}\beta}. \tag{12}$$

As we recognize, Equation (12) leads us to some complexities and it is not possible to compute it analytically and then the Bayes estimates cannot be obtained in the closed form. So, we use the Lindley's approximation and MCMC methods to obtain Equation (12).

2.1. Lindley's approximation

Lindley [11] suggested a method to approximate the ratio of integrals such as Equation (12). This procedure has been employed by many authors to derive a approximation for the Bayes estimates. For two parameter case (α, β), Lindley's approximation can be represented as

$$\mathrm{E}(g(\alpha,\beta)|\boldsymbol{r}) \cong g(\tilde{\alpha},\tilde{\beta}) + \frac{1}{2}\Big[B + Q_{30}B_{12} + Q_{21}C_{12} + Q_{12}C_{21} + Q_{03}B_{21}\Big], \tag{13}$$

where $B = \sum_{i=1}^{2}\sum_{j=1}^{2} g_{ij}\tau_{ij}$, $Q_{ij} = \partial^{i+j}Q/\partial\alpha^i\partial\beta^j$ for $i, j = 0, 1, 2, 3$ with $i+j = 3$, $g_1 = \partial g/\partial\alpha$, $g_2 = \partial g/\partial\beta$, $g_{ij} = \partial^2 g/\partial\alpha^i\partial\beta^j$ for $i, j = 0, 1, 2$, $B_{ij} = (g_i\tau_{ii} + g_j\tau_{ij})\tau_{ii}$, $C_{ij} = 3g_i\tau_{ii}\tau_{ij} + g_j(\tau_{ii}\tau_{ij} + 2\tau_{ij}^2)$ for $i \neq j$, τ_{ij} is the (i, j)-th element in the inverse of the matrix $Q^\star = (-Q^\star_{ij})$ with $Q^\star_{ij} = \partial^2 Q/\partial\theta_i\partial\theta_j$ for $i, j = 1, 2$, where $(\theta_i, \theta_j) = (\alpha, \beta)$ and Q is the logarithm of the posterior PDF in (9). Also, $(\tilde{\alpha}, \tilde{\beta})$ is the mode of Q. Notice that the parameters α and β in (13) are replaced by $\tilde{\alpha}$ and $\tilde{\beta}$, respectively.

From (9), we have

$$\begin{aligned}Q =& \log C(\boldsymbol{r}) + (m + a_1 + a_2 - 1)\log\alpha + (m + a_1 - 1)\log\beta - m\beta\log 2 - \alpha\left[\left(\frac{e^{r_m}-1}{2}\right)^{\beta} + b_2 + \frac{\beta}{b_1}\right]\\ &+ (\beta - 1)\sum_{i=1}^{m}\log(e^{r_i} - 1).\end{aligned}$$

The mode of Q is derived from the equations $\partial Q/\partial\alpha = 0$ and $\partial Q/\partial\beta = 0$. Then, we obtain

$$\tilde{\alpha} = \frac{m + a_1 + a_2 - 1}{2^{-\tilde{\beta}}\left(e^{r_m} - 1\right)^{\tilde{\beta}} + b_2 + \frac{\tilde{\beta}}{b_1}},$$

where $\tilde{\beta}$ is the solution of the following nonlinear equation

$$\frac{m + a_1 - 1}{\tilde{\beta}} - m\log 2 - \frac{m + a_1 + a_2 - 1}{2^{-\tilde{\beta}}\left(e^{r_m} - 1\right)^{\tilde{\beta}} + b_2 + \frac{\tilde{\beta}}{b_1}}\left[\left(\frac{e^{r_m}-1}{2}\right)^{\tilde{\beta}}\left(\log\frac{e^{r_m}-1}{2}\right) + \frac{1}{b_1}\right] + \sum_{i=1}^{m}\log(e^{r_i} - 1) = 0.$$

To obtain $\tilde{\beta}$, we can apply the Newton-Raphson method or fixed point method or other numerical methods. Also, the elements of $Q^\star$ are given by

$$\begin{aligned}Q^\star_{11} &= \frac{\partial^2 Q}{\partial\alpha^2} = -\frac{m + a_1 + a_2 - 1}{\alpha^2},\\ Q^\star_{12} = Q^\star_{21} &= \frac{\partial^2 Q}{\partial\alpha\,\partial\beta} = -\left(\frac{e^{r_m}-1}{2}\right)^{\beta}\left(\log\frac{e^{r_m}-1}{2}\right) - \frac{1}{b_1},\\ Q^\star_{22} &= \frac{\partial^2 Q}{\partial\beta^2} = -\frac{m + a_1 - 1}{\beta^2} - \alpha\left(\frac{e^{r_m}-1}{2}\right)^{\beta}\left(\log\frac{e^{r_m}-1}{2}\right)^2.\end{aligned}$$

So, we have

$$\begin{bmatrix}\tau_{11} & \tau_{12}\\ \tau_{21} & \tau_{22}\end{bmatrix} = \frac{1}{Q^\star_{11}Q^\star_{22} - Q^\star_{12}Q^\star_{21}}\begin{bmatrix}-Q^\star_{22} & Q^\star_{12}\\ Q^\star_{21} & -Q^\star_{11}\end{bmatrix}.$$

Moreover, it is easy to show that

$$\begin{aligned}Q_{30} &= \frac{\partial^3 Q}{\partial\alpha^3} = \frac{2(m + a_1 + a_2 - 1)}{\alpha^3}, \qquad Q_{21} = \frac{\partial^3 Q}{\partial\alpha^2\partial\beta} = 0,\\ Q_{12} &= \frac{\partial^3 Q}{\partial\alpha\partial\beta^2} = -\left(\frac{e^{r_m}-1}{2}\right)^{\beta}\left(\log\frac{e^{r_m}-1}{2}\right)^2,\\ Q_{03} &= \frac{\partial^3 Q}{\partial\beta^3} = \frac{2(m + a_1 - 1)}{\beta^3} - \alpha\left(\frac{e^{r_m}-1}{2}\right)^{\beta}\left(\log\frac{e^{r_m}-1}{2}\right)^3.\end{aligned}$$

Therefore, based on the Lindley's approximation in (13), the Bayes estimates of α and β under SE and LINEX loss functions are, respectively, given by

$$\hat{\alpha}_{\mathrm{BS}} \cong \tilde{\alpha} + \frac{1}{2}\Big[Q_{30}\tau_{11}^2 + 3Q_{21}\tau_{11}\tau_{12} + Q_{12}\tau_{21}\left(\tau_{22} + 2\tau_{21}\right) + Q_{03}\tau_{21}\tau_{22}\Big], \tag{14}$$

$$\hat{\beta}_{\mathrm{BS}} \cong \tilde{\beta} + \frac{1}{2}\Big[Q_{30}\tau_{11}\tau_{12} + Q_{21}\tau_{12}(\tau_{11} + 2\tau_{12}) + 3Q_{12}\tau_{21}\tau_{22} + Q_{03}\tau_{22}^2\Big], \tag{15}$$

$$\hat{\alpha}_{\mathrm{BL}} \cong \tilde{\alpha} - \frac{1}{\gamma}\log\Big\{1 + \frac{\gamma}{2}\Big(\gamma\tau_{11} - Q_{30}\tau_{11}^2 - 3Q_{21}\tau_{11}\tau_{12} - Q_{12}\tau_{21}(\tau_{22} + 2\tau_{21}) - Q_{03}\tau_{21}\tau_{22}\Big)\Big\}, \tag{16}$$

$$\hat{\beta}_{\mathrm{BL}} \cong \tilde{\beta} - \frac{1}{\gamma}\log\Big\{1 + \frac{\gamma}{2}\Big(\gamma\tau_{22} - Q_{30}\tau_{11}\tau_{12} - Q_{21}\tau_{12}(\tau_{11} + 2\tau_{12}) - 3Q_{12}\tau_{21}\tau_{22} - Q_{03}\tau_{22}^2\Big)\Big\}. \tag{17}$$

It is noteworthy that the parameters α and β are evaluated at $\tilde{\alpha}$ and $\tilde{\beta}$, respectively.

2.2. MCMC method

In the preceding subsection, we derived the Bayes estimates of the parameters α and β using Lindley's approximation under the SE and LINEX loss functions. Due to the unavailability of exact probability distributions for these estimates, evaluating Bayesian credible intervals for the parameters becomes challenging. Nevertheless, by employing the MCMC method, we can obtain estimates of α and β and subsequently construct HPD credible intervals. To do this, we refer to the joint posterior PDF of α and β in (9). From (9), we have

$$(\alpha|\beta, \boldsymbol{r}) \sim \text{Gamma}\left(m + a_1 + a_2, \left(\frac{e^{r_m} - 1}{2}\right)^{\beta} + b_2 + \frac{\beta}{b_1}\right),$$

and

$$\pi(\beta|\alpha, \boldsymbol{r}) \propto \beta^{m+a_1-1}2^{-m\beta}\prod_{i=1}^{m}(e^{r_i} - 1)^{\beta-1}e^{-\alpha\left[2^{-\beta}(e^{r_m}-1)^{\beta}+b_2+\frac{\beta}{b_1}\right]}.$$

Generation the random sample from the posterior distribution of α given β is straightforward. But, the posterior distribution of β given α cannot be analytically reduced to a well-known distribution, making it impossible to sample directly using standard methods. Following the approach by Gelman et al. [9], in cases where the posterior PDF is unimodal and roughly symmetric, it is often convenient to approximate it with a normal distribution. Since, the posterior PDF of β given α is unimodal and roughly symmetric (as determined through experimentation), we employ the Metropolis-Hasting algorithm with a normal proposal distribution to generate the sample from the posterior distribution of β given α. Therefore, for generating samples from the joint posterior distribution of α and β, the following hybrid Metropolis-Hastings and Gibbs sampling algorithm, proposed by Tierney [18], will be used.

1. Let initial values of the parameters to be (α_0, β_0) and set $t = 1$.
2. Using the Metropolis-Hastings algorithm under the proposal distribution $N(\beta, V_\beta)$, where $V_\beta = (-\partial^2\log\pi(\beta|\alpha_{t-1}, \boldsymbol{r})/\partial\beta^2)^{-1}$ with the corresponding PDF $q(x|\beta)$, generate β_t from $\pi\left(\beta|\alpha_{t-1}, \text{data}\right)$ as follows:
   - Generate $\beta^\star$ from the proposal distribution $q(x|\beta_{t-1})$.
   - Compute the quantity
   $$p = \min\left\{1, \frac{\pi(\beta^\star|\alpha_{t-1}, \boldsymbol{r})\,q(\beta_{t-1}|\beta^\star)}{\pi(\beta_{t-1}|\alpha_{t-1}, \boldsymbol{r})\,q(\beta^\star|\beta_{t-1})}\right\}. \tag{18}$$
   If $p = 1$, then set $\beta_t = \beta^\star$. If $p < 1$, then set $\beta_t = \beta^\star$ with probability p or $\beta_t = \beta_{t-1}$ with probability $1 - p$. To do this, generate u from the standard uniform distribution. If $u < p$, accept $\beta^\star$, otherwise accept β_{t-1}.
3. Generate α_t from Gamma $\Big(m + a_1 + a_2, 2^{-\beta_t}\left(e^{r_m} - 1\right)^{\beta_t} + b_2 + \beta_t/b_1\Big)$.
4. Set $t = t + 1$.
5. Repeat steps 2–4, N' times, where N' is a large number, and obtain (α_t, β_t) for $t = 1, \ldots, N'$.

Also, we employ a thinning approach, whereby only every k-th generated sample is retained, to diminish the autocorrelation of the chain. Then, the generated samples from $\pi(\alpha,\beta|\boldsymbol{r})$ are

$$(\alpha_{M+1},\beta_{M+1}),(\alpha_{M+k+1},\beta_{M+k+1}),\ldots,(\alpha_{M+k(N-1)+1},\beta_{M+k(N-1)+1}),$$

where M is the burn-in period, k is the thinning parameter and N is the desired size of the generated sample. On the basis of the random samples generated from the above algorithm, the Bayes estimators for the parameters α and β under SE and LINEX loss functions are, respectively, given by

$$\hat{\hat{\alpha}}_{\text{BS}} \cong \frac{1}{N}\sum_{t=1}^{N}\alpha_t \quad , \quad \hat{\hat{\beta}}_{\text{BS}} \cong \frac{1}{N}\sum_{t=1}^{N}\beta_t, \tag{19}$$

and

$$\hat{\hat{\alpha}}_{\text{BL}} \cong -\frac{1}{\gamma}\log\left(\frac{1}{N}\sum_{t=1}^{N}e^{-\gamma\alpha_t}\right) \quad , \quad \hat{\hat{\beta}}_{\text{BL}} \cong -\frac{1}{\gamma}\log\left(\frac{1}{N}\sum_{t=1}^{N}e^{-\gamma\beta_t}\right). \tag{20}$$

Also, using the method of Chen and Shao [8], a $100(1-\delta)\%$ HPD credible interval for the parameters α and β is provided as follows. Let $\alpha_1^\star < \cdots < \alpha_N^\star$ be the ordered values of $\alpha_1,\ldots,\alpha_T$. The HPD credible interval for α is the shortest length interval through the following $100(1-\delta)\%$ intervals

$$\left(\alpha_t^\star,\alpha_{t+[N(1-\delta)]}^\star\right) \quad \text{for} \quad t=1\ldots,N-[N(1-\delta)],$$

where $[\kappa]$ denotes the largest integer less than or equal to κ. The $100(1-\delta)\%$ HPD credible interval for β is derived in a similar manner.

3. Bayesian prediction

This section deals with the prediction of the future records based on the past records using a Bayesian approach. Suppose that we have the first m upper records $\boldsymbol{R}=(R_1,\ldots,R_m)$ with observed values $\boldsymbol{r}=(r_1,\ldots,r_m)$ from the WHL distribution with parameters α and β. Also, let $Y=R_s$ with $s>m$ be the s-th upper record. It is proved that the sequence $\{R_1,R_2,\ldots\}$ is a Markov chain; that is, the conditional PDF of $Y=R_s$ given $\boldsymbol{R}=\boldsymbol{r}$ is just the conditional pdf of $Y=R_s$ given $R_m=r_m$. For more details, see Arnold et al. [4]. It follows that the conditional PDF of $Y=R_s$ given $\boldsymbol{R}=\boldsymbol{r}$ is given by

$$f(y|\boldsymbol{r},\theta)=\frac{f(y;\theta)\big(H(y;\theta)-H(r_m;\theta)\big)^{s-m-1}}{\Gamma(s-m)\bar{F}(r_m;\theta)}, \quad r_m<y<\infty, \tag{21}$$

where $H(\cdot;\theta)=-\log\bar{F}(\cdot;\theta)$. We can rewrite the conditional PDF of $Y=R_s$ given $\boldsymbol{R}=\boldsymbol{r}$ in (21) as follows

$$f(y|\boldsymbol{r},\theta)=\frac{f(y;\theta)}{\Gamma(s-m)\bar{F}(r_m;\theta)}\sum_{j=0}^{s-m-1}\binom{s-m-1}{j}(-1)^j\left(\log\bar{F}(y;\theta)\right)^j\left(\log\bar{F}(r_m;\theta)\right)^{s-m-1-j}, \; r_m<y<\infty.$$

Hence, the conditional CDF of $Y=R_s$ given $\boldsymbol{R}=\boldsymbol{r}$ is

$$\begin{aligned}F(y|\boldsymbol{r},\theta)&=\int_{r_m}^{y}f(t|\boldsymbol{r},\theta)\,\mathrm{d}t\\&=\frac{1}{\Gamma(s-m)\bar{F}(r_m;\theta)}\sum_{j=0}^{s-m-1}\binom{s-m-1}{j}\left(\log\bar{F}(r_m;\theta)\right)^{s-m-1-j}\\&\quad\times\Big[\Gamma\big(j+1,-\log\bar{F}(y;\theta)\big)-\Gamma\big(j+1,-\log\bar{F}(r_m;\theta)\big)\Big], \quad r_m<y<\infty,\end{aligned}$$

where $\Gamma(c,x) = \int_0^x u^{c-1}e^{-u}\,\mathrm{d}u$ is the incomplete gamma function. For a special case, when $s = m+1$, the conditional PDF and CDF of $Y = R_s$ given $R_m = r_m$ are reduced to $f(y|\boldsymbol{r},\theta) = f(y;\theta)/\bar{F}(r_m;\theta)$ and $F(y|\boldsymbol{r},\theta) = 1-\bar{F}(y;\theta)/\bar{F}(r_m;\theta)$, respectively.
Replacing Equations (2) and (3) into (21), the conditional PDF of $Y = R_s$ given $\boldsymbol{R} = \boldsymbol{r}$, under the WHL distribution with parameters α and β, is obtained as

$$f(y|\boldsymbol{r},\alpha,\beta) = \frac{\alpha\beta 2^{-\beta}\left(e^y-1\right)^{\beta-1} e^{y-\alpha 2^{-\beta}(e^y-1)^\beta}\left(\alpha 2^{-\beta}\left(e^y-1\right)^\beta - \alpha 2^{-\beta}\left(e^{r_m}-1\right)^\beta\right)^{s-m-1}}{\Gamma(s-m)\,e^{-\alpha 2^{-\beta}(e^{r_m}-1)^\beta}}, \quad r_m < y < \infty. \tag{22}$$

Then, the Bayes predictive PDF of $Y = R_s$ given $\boldsymbol{R} = \boldsymbol{r}$ is

$$h(y|\boldsymbol{r}) = \int_0^\infty\int_0^\infty f(y|\boldsymbol{r},\alpha,\beta)\pi(\alpha,\beta|\boldsymbol{r})\,\mathrm{d}\alpha\,\mathrm{d}\beta.$$

Based on Equations (9) and (22), it is easy to check that $h(y|\boldsymbol{r})$ is not computed analytically and then the Bayes predictive PDF of Y given $\boldsymbol{R} = \boldsymbol{r}$ has not a closed form. However, we can derive a consistent estimator for $h(y|\boldsymbol{r})$ using the hybrid Metropolis-Hastings and Gibbs sampling method explained in Subsection 2.2. Let $\{(\alpha_t,\beta_t);\ t = 1,\ldots,N\}$ be the MCMC sample obtained from the posterior PDF $\pi(\alpha,\beta|\boldsymbol{r})$ on the basis of the hybrid Metropolis-Hastings and Gibbs sampling method. So, the consistent estimator for $h(y|\boldsymbol{r})$ is given by

$$\hat{h}(y|\boldsymbol{r}) = \frac{1}{N}\sum_{t=1}^{N} f(y|\boldsymbol{r},\alpha_t,\beta_t).$$

Also, a consistent estimator for the Bayes predictive CDF of $Y = R_s$ given $\boldsymbol{R} = \boldsymbol{r}$, based on the MCMC sample, is as follows

$$\hat{H}(y|\boldsymbol{r}) = \frac{1}{N}\sum_{t=1}^{N} F(y|\boldsymbol{r},\alpha_t,\beta_t). \tag{23}$$

Moreover, the point predictors for $Y = R_s$ given $\boldsymbol{R} = \boldsymbol{r}$, under SE and LINEX loss functions, are, respectivley, given by

$$\hat{Y}_{BS} \cong \int_{r_m}^\infty y\,\hat{h}(y|\boldsymbol{r})\,\mathrm{d}y = \frac{1}{N}\sum_{t=1}^{N}\int_{r_m}^\infty y f(y|\boldsymbol{r},\alpha_t,\beta_t)\,\mathrm{d}y, \tag{24}$$

and

$$\hat{Y}_{BL} \cong -\frac{1}{\gamma}\log\left\{\int_{r_m}^\infty e^{-\gamma y}\,\hat{h}(y|\boldsymbol{r})\,\mathrm{d}y\right\} = -\frac{1}{\gamma}\log\left\{\frac{1}{N}\sum_{t=1}^{N}\int_{r_m}^\infty e^{-\gamma y} f(y|\boldsymbol{r},\alpha_t,\beta_t)\,\mathrm{d}y\right\}. \tag{25}$$

Finally, for constructing the equi-tailed two-sided $100(1-\delta)\%$ Bayesian prediction interval (L,U) for $Y = R_s$, we have

$$\frac{\delta}{2} = P(Y < L|\boldsymbol{r}) = H(L|\boldsymbol{r}) \quad \text{and} \quad 1-\frac{\delta}{2} = P(Y < U|\boldsymbol{r}) = H(U|\boldsymbol{r}).$$

From (23), we can obtain the lower bound L and upper bound U by solving the nonlinear equations $\delta/2 = \hat{H}(L|\boldsymbol{r})$ and $1-\delta/2 = \hat{H}(U|\boldsymbol{r})$ using the Newton-Raphson method.

4. Illustrative example

In this section, we illustrate how the methods presented in two previous sections can be employed using an numerical example. For this purpose, we simulated a set of upper record data with $m = 5$ form the WHL distribution with parameters $\alpha = 2$ and $\beta = 3$. This simulated data set is

$$0.6946668 \quad 0.7850423 \quad 1.1602795 \quad 1.1807104 \quad 1.3466174.$$

Let the hyperparameters of the gamma prior distributions in (7) and (8) be $(a_1, b_1) = (1, 1)$ and $(a_2, b_2) = (2, 2)$, respectively. On the basis of the above record data, from Equations (14) to (17), the Bayes estimates of α and β under SE and LINEX loss functions and using the Lindley's approximation are obtained as

$$\hat{\alpha}_{\text{BS}} = 1.38612, \quad \hat{\beta}_{\text{BS}} = 2.06941, \quad \hat{\alpha}_{\text{BL}} = 1.53841, \quad \hat{\beta}_{\text{BL}} = 1.78531.$$

However, using the MCMC method and based on Equations (19) and (20), these Bayes estimates are derived as

$$\hat{\hat{\alpha}}_{\text{BS}} = 1.47188, \quad \hat{\hat{\beta}}_{\text{BS}} = 1.80343, \quad \hat{\hat{\alpha}}_{\text{BL}} = 1.29781, \quad \hat{\hat{\beta}}_{\text{BL}} = 1.45382.$$

Also, the 95% HPD credible intervals for α and β are computed as $(0.37147, 2.78904)$ and $(0.17588, 3.64702)$, respectively.

Moreover, from Equations (24) and (25), the point Bayesian predictors for $Y = R_{m+1}$ under SE and LINEX loss functions are obtained as $\hat{Y}_{BS} = 1.54365$ and $\hat{Y}_{BL} = 1.51395$, respectively. Finally, the 95% Bayesian prediction interval for $Y = R_{m+1}$ is derived as $(1.35035, 2.24837)$.

References

[1] J. Ahmadi and N.R. Arghami. Comparing the fisher information in record values and iid observations. *Statistics*, 37:435–441, 2003.

[2] J. Ahmadi, M. Doostparast, and A. Parsian. Estimation and prediction in a two parameter exponential distribution based on k-record values under linex loss function. *Communications in Statistics-Theory and Methods*, 34:795–805, 2003.

[3] M. Ahsanullah and M. Shakil. Characterizations of rayleigh distribution based on order statistics and record values. *Bulletin of the Malaysian Mathematical Sciences Society*, 36:625–635, 2013.

[4] B.C. Arnold, N. Balakrishnan, and H.N. Nagaraja. *Records*. John Wiley, New York, 1998.

[5] A. Asgharzadeh, M. Abdi, and S. Nadarajah. Interval estimation for gumbel distribution using climate records. *Bulletin of the Malaysian Mathematical Sciences Society*, 39:257–270, 2016.

[6] A. Baklizi. Likelihood and bayesian estimation of Pr($X < Y$) using lower record values from the generalized exponential distribution. *Computational Statistics and Data Analysis*, 52:3468–3473, 2008.

[7] K.N. Chandler. The distribution and frequency of record values. *Journal of the Royal Statistical Society, Series B*, 14:220–228, 1952.

[8] M.H. Chen and Q.M. Shao. Monte carlo estimation of bayesian credible and hpd intervals. *Journal of computational and Graphical Statistics*, 8:69–92, 1999.

[9] A. Gelman, J.B. Carlin, H.S. Stern, and D.B. Rubin. *Bayesian Data Analysis*. Chapman and Hall/CRC, New York, 1998.

[10] N. Glick. Breaking records and breaking boards. *The American Mathematical Monthly*, 85:2–26, 1978.

[11] D.V. Lindley. Approximate bayesian methods. *Trabajos de Estadistica*, 31:223–245, 1980.

[12] D.N.P. Murthy, M. Xie, and R. Jiang. *Weibull Models*. John Wiley, New York, 2004.

[13] M. Nadar, A. Papadopoulos, and F. Kizilaslan. Statistical analysis for kumaraswamy's distribution based on record data. *Statistical Papers*, 54:355–369, 2013.

[14] V. Nevzorov. Records: mathematical theory. *Translations of mathematical monographs*, 2000.

[15] A. Parsian and S.N.U.A. Kirmani. Estimation under linex loss function. In *Handbook of applied econometrics and statistical inference*, pages 75–98. CRC Press, 2002.

[16] H. Piriaei, G.H. Yari, and R. Farnoosh. E-bayesian estimations for the cumulative hazard rate and mean residual life based on exponential distribution and record data. *Journal of Statistical Computation and Simulation*, 90:271–290, 2020.

[17] J.I. Seo and Y. Kim. Statistical inference on gumbel distribution using record values. *Journal of the Korean Statistical Society*, 45:342–357, 2016.

[18] L. Tierney. Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728, 1994.

[19] H.R. Varian. A bayesian approach to real estate assessment. *Studies in Bayesian econometrics and statistics in honor of Leonard J. Savage*, 1975.

[20] J. Zhao, W. Cheng, H. Chen, and M. Wu. Comparisons of several pareto distributions based on record values. *Communications in Statistics-Theory and Methods*, 47:2456–2468, 2018.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–267, pp. 305–308

Application of unified skewed distribution in financial model

Asieh Abtahi[a,*]

[a]*Department of Mathematics and Statistics, Islamic Azad University, Shiraz, Iran*

Article Info

Keywords:
Skewed disribution
Stein Lemma
Portfolio selection

2020 MSC:
msc1
msc2

Abstract

Returns on financial assets are not normally distributed. They exhibit both skewness and kurtosis. There are so many proposals in construction of skewed distributions and it is worth to find an overall class which covers all of these proposals. In this paper the multivariate unified skew-symmetric distributions is introduced. This unified multivariate representation of skewed distributions includes all of the multivariate skewed distributions in the literature. The purpose of this paper is to use one version of this unified form which is an attractive model for applications in finance and is suitable for portfolio selection in the presence of skewness and offers a number of different insights into the sources of expected return and risk in a portfolio. This article shows that, investors who are expected utility maximizers will be located on a single mean-variance-skewness efficient surface, regardless of their choice of utility function.

1. Introduction

Markowitz (1952) was the first to propose a quantitative approach to optimal portfolio selection. This method is equivalent to assuming that investors maximize the expected utility when the utility function used is quadratic in portfolio return. The use of quadratic utility functions in finance, however, is criticized, see for example Pratt(1964), on the grounds that there must be circumstances in which an investor appears to prefer less wealth to more wealth. This criticism, coupled with the natural desire to achieve higher portfolio returns or lower portfolio volatility or both, has lead to the search for what might be called better utility functions. When returns on financial assets have a multivariate normal distribution, the consequence of Stein's Lemma is that all well behaved utility functions will lead to a point on Markowitz' mean-variance efficient frontier.
When two random variable have a bivariate normal distribution, Stein's Lemma provides an expression for the covariance of the first variable with a function of the second. Stein's Lemma has many application in statistics and probability and it plays an important role in modern finance. Stein's Lemma states that, if $\mathbf{X}$ is a random vector which has a multivariate normal distribution with mean vector μ and covariance matrix Σ and $h(\mathbf{x})$ is a scaler valued function which satisfies certain regularity conditions, then $Cov\{\mathbf{X}, h(\mathbf{X})\} = \Sigma E\{\nabla h(\mathbf{X})\}$, where $\nabla h(\mathbf{X})$ is the vector of first derivatives of $h(.)$ with respect to the elements of $\mathbf{X}$.
However, it is well known that returns on financial assets are not normally distributed. They exhibit both skewness and kurtosis. The construction of skewed multivariate distributions have received some considerable attention in the past

*Talker
Email address: as.abtahi@iau.ac.ir (Asieh Abtahi)

few decades. There are so many proposals in construction of skewed distributions and it is worth to find an overall class which covers all of these proposals. Unification of families of skewed distributions under a new general formulation has been one of the important subject considered in recent years. Abtahi and Towhidi (2013) introduced, the unified representation of multivariate skewed distributions. They showed that, this new unified multivariate representation of skewed distributions includes all of the multivariate skewed distributions in the literature and introduced in the next section.

2. Unified multivariate skewed distribution

Definition 2.1. A random vector $\mathbf{X}_{f,p}$ is said to have a unified multivariate skewed distribution with functional parameters f and p, if its *pdf* be of the form

$$s(\mathbf{x}|f,p) = f(\mathbf{x})\, p[F(x_1), F(x_2|x_1), ..., F(x_k|x_1, ..., x_{k-1})]; \quad \mathbf{x} \in \Re^k, \tag{1}$$

where f is a symmetric multivariate pdf corresponding to a random vector $\mathbf{U}$ on $\Re^k$ and p is a multivariate pdf on $[0,1]^k$.
Note that for $i = 2, ...k$

$$F(x_i|x_1, ..., x_{i-1}) = F_{U_i}(x_i|U_1 = x_1, U_2 = x_2, ..., U_{i-1} = x_{i-1}) \tag{2}$$

is the conditional cumulative distribution of $U_i|U_1 = x_1, U_2 = x_2, ..., U_{i-1} = x_{i-1}$.

$X_{f,p}$ has the unified multivariate skewed distribution, is denoted by $X_{f,p} \sim UMSD(f,p)$, if its pdf is given by (1). In the above definition the density p specifies a skewing mechanism. Different skewing mechanisms lead to different skewed symmetric distributions.
Unified definition of multivariate skewed distributions introduced here, separated the skewing mechanism from the symmetric distribution. The following theorem states that any multivariate continuous skewed distribution can be interpreted as a skewed version of a symmetric distribution F, generated by a skewing mechanism p, in the form (2).

Theorem 2.2. *Every continuous multivariate skewed density function $s(.)$ can be expressed in the form of $USS_k(f,p)$.*

The following example is the direct result of the above theorem.

Example 2.3. The probability distribution of $\mathbf{X}$ is multivariate skew-normal with parameters μ, Σ, λ and τ, denoted as $X \sim MSN(\mu, \Sigma, \lambda, \tau)$. The probability density function of this distribution is

$$s(\mathbf{x}) = \phi_n(\mathbf{x}, \mu + \lambda\tau, \square + \lambda\lambda^T)\Phi(\frac{\tau + \lambda^T \square^{-1}(x-\mu)}{\sqrt{1 + \lambda^T \square^{-1}\lambda}})\frac{1}{\Phi(\tau)} \tag{3}$$

This density function is introduced by Azzalini and Dalla Valle's (1996) with a change of notation and generalization to accommodate a non zero value of τ. By defining $f(\mathbf{x}) = f_U(x) = \phi_n(x)$and $\Phi(x_i|x_1, ..., x_{i-1}) = F_{U_i(x_i|U_1=x_1,...,U_{i-1}=x_{i-1})}$,the MSN can be stated as (1) with the skewing mechanism

$$p(v) = \Phi[\acute{\alpha}(g_1(v_1), g_2(v_1, v_2), ..., g_k(v_1, v_2, ..., v_k))' - (\mu_1, ..., \mu_k)]\frac{1}{\Phi(\tau)} \tag{4}$$

where $\alpha' = \frac{\tau + \lambda'\Sigma^{-1}}{\sqrt{1+\lambda'\Sigma^{-1}\lambda}}$ and $g_i(v_1, ..., v_i)$ defined as follows

$$\begin{cases} x_1 = F^{-1}(v_1) = g_1(v_1) \\ x_2 = F^{-1}(v_2|g_1(v_1)) = g_2(v_1,\ v_2) \\ . \\ . \\ . \\ x_k = F^{-1}(v_k|g_1(v_1), ..., g_{k-1}(v_1), ..., v_{k-1})) = g_k(v_1, ..., v_k) \end{cases}$$

and $F^{-1}(x_i|x_1, ..., x_{i-1})$ is the inverse of the cumulative distribution function of $(U_i|U_1 = x_1, ..., U_{i-1} = x_{i-1})$.

This distribution that is introduced in the above example is an attractive model for application in finance. This is because it offers a parsimonious model for the skewness which is often observed in the return distributions of risky financial assets. The setup and method of proof follow Adcock(2007).

Theorem 2.4. *Let* $\mathbf{X}$ *be an* n *vector that has the distribution* $MSN(\mu, \Sigma, \Lambda, \tau)$. *For any scalar valued function* $h(\mathbf{x})$ *such that* $\frac{\partial h(\mathbf{x})}{\partial x_i}$ *is continuous almost everywhere and* $E\{\frac{\partial h(\mathbf{x})}{\partial x_i}\} < \infty$, $i = 1, ..., n$,

$$Cov(\mathbf{X}, h(\mathbf{X})) = (\Sigma + \lambda\lambda^T)E\{\nabla h(\mathbf{X})\} + \lambda[E_N h(\mathbf{X}) - E\{h(\mathbf{X})\}]\xi_1(\tau), \tag{5}$$

where $E\{.\}$ *denotes expectation taken over the* MSN *distribution defined above and* $E_N\{.\}$ *denotes expectation taken over a multivariate normal distribution with mean vector* μ *and covariance matrix* Σ. *The vector* $\nabla h(\mathbf{X})$ *is the vector of first derivatives of* $h(.)$ *with respect to the elements of* $\mathbf{X}$.

Proof. First write the vector $\mathbf{X}$ as $\mathbf{X} = \square^{\frac{1}{2}}\mathbf{Z} + \gamma$, where $\square$ and γ are as defined as follow:

$$\gamma = \mu + \lambda\tau, \Omega = \Sigma + \lambda\lambda^T, \beta = \{1 + \lambda^T\Sigma^{-1}\lambda\}^{\frac{1}{2}}\lambda = \alpha^{-1}\lambda$$

and define $g(\mathbf{Z}) = h(\Omega^{\frac{1}{2}}\mathbf{Z} + \gamma)$. Noting that the expected value of $\mathbf{X}$ is δ which is defined as, $\delta = E(\mathbf{X}) = \mu + \lambda\{\tau + \xi_1(\tau)\}$. It follows that $cov\{\mathbf{X}, h(\mathbf{X})\} = \square^{\frac{1}{2}}cov\{\mathbf{Z}, g(\mathbf{Z})\} - \lambda\xi_1(\tau)E\{g(\mathbf{Z})\}$.
Element i of the vector of covariance of $\mathbf{Z}$ with $g(\mathbf{Z})$ is $\int ... \int z_i g(\mathbf{z}) f_Z(\mathbf{z}) d\mathbf{z}$. The distribution of Z has a probability density function given by

$$f_Z(\mathbf{z}) = \phi_n(\mathbf{x}; \mathbf{0}, \mathbf{I}_n)\frac{\Phi(\tau\{1+\eta^T\eta\}^{\frac{1}{2}} + \eta^T\mathbf{z})}{\Phi(\tau)}, \eta = \square^{\frac{-1}{2}}\beta$$

The integral with respect to z_i only is $\int z_i\phi(z_i)g(\mathbf{z})\Phi(\tau\{1 + \eta^T\eta\}^{\frac{1}{2}} + \eta^T\mathbf{z})/\Phi(\tau)dz_i$, which, using integration by parts, can be expressed as

$$\int \phi(z_i)\frac{\partial g(\mathbf{z})}{\partial z_i}\frac{\Phi(\tau\{1+\eta^T\eta\}^{\frac{1}{2}} + \eta^T\mathbf{z})}{\Phi(\tau)}dz_i + \int \phi(z_i)g(\mathbf{z})\eta_i\frac{\Phi(\tau\{1+\eta^T\eta\}^{\frac{1}{2}} + \eta^T\mathbf{z})}{\Phi(\tau)}dz_i$$

Integration of the first term with respect to all elements of $\mathbf{z}$ gives $E\{\partial g(Z)/\partial Z_i\}$. After re-arrangement of the arguments of the two $\phi(.)$ functions, the second integral with respect to $\mathbf{z}$ is $\xi_1(\tau)\eta_i\int ... \int g(\mathbf{z})\phi_n(\mathbf{z}, -\eta\tau, \mathbf{I}_n + \eta\eta^T)d\mathbf{z}$. This may be expressed in terms of $\mathbf{X}$, the original vector of variables, as $\alpha\xi_1(\tau)\eta_i\int ... \int h(\mathbf{x})\phi_n(\mathbf{x}, \mu, \Sigma)d\mathbf{x}$. To complete the proof, note that $\alpha\Omega^{\frac{1}{2}}\eta = \lambda$ and that the vector of derivatives $\partial g(Z)/\partial Z_i, i = 1, ..., n$ is given by $\Omega^{\frac{1}{2}}\nabla h(.)$. $\square$

3. Application To Portfolio Selection

A portfolio is a set of investment weights or proportions $\{w_i\}, i = 1, ..., n$, defined such that an investor invests $100w_i$ of wealth in asset i. It is conventionally assumed that the weights sum to one. If the return on asset i is denoted by the random variable $R_i, i = 1, ...n$, then the return on the portfolio with weights $\{w_i\}$ is $R_P = \sum_{i=1}^n w_i R_i = \mathbf{w}^T\mathbf{R}$, where **w** and **R** are vectors of length n containing the investment weights and asset returns, respectively. For portfolio selection, the investor conventionally chooses the vector of portfolio weights **w** to maximize the expected utility of portfolio return.
For a general utility function, $U(R_p)$, the expected utility is, $\int ... \int U(r_P)f(\mathbf{r})d\mathbf{r} = \Xi(\mathbf{w})$ where $f(\mathbf{r})$ denote the density function of the multivariate probability distribution of the vector of return **R**.
Following Kallberg and Ziemba(1983), the investor who is an expected utility maximizer solves, $max_{(w)}\Xi(\mathbf{w}) - \eta(\mathbf{u}^T\mathbf{w} - \mathbf{1})$, where, **u** is a vector of length n containing ones and η is the Lagrange multiplier of the budget constraint. Ignoring this constraint for simplicity, the first order conditions for the weight for asset i are, $\frac{\partial \Xi}{\partial w_i} = \int ... \int r_i U'(r_P)f(\mathbf{r})d\mathbf{r}$. The right hand side of this last expression may be written as

$$E(R_i)E\{U'(R_P)\} + \int ... \int r_i E[R_i]U'(r_P)f(\mathbf{r})d\mathbf{r} = Cov(R_i, U'(R_P)) + E(R_i)E\{U'(R_P)\}$$

When returns follow the multivariate skew normal distribution in example 2.3, application of theorem 3.4 gives the vector of first order conditions for all assets.

$$\delta E(U') + (\Box + \lambda\lambda^T)E(\nabla U') + \lambda\{E_N(U') - E(U')\}\xi_1(\tau)$$

where the vector $\nabla U'$ is given by $\nabla U' = \frac{\partial U'}{\partial \mathbf{R}} = \mathbf{w}U''$. The first order conditions may re-expressed in terms of Θ, the covarianve of the returns as

$$\delta E(U') + \Theta\mathbf{w}E(U'') + \lambda[(E_N(U') - E(U'))\xi_1(\tau) - \lambda^T\mathbf{w}E(U'')\xi_2(\tau)]$$

This equation in **w**is the same for all investors, except for the three scalar quantities, which are functions of certain expected values of $U'(.)$ and $U''(.)$. When all elements of the vector λ are equal to zero, asset returns have a multivariate normal distribution and Kallberg and Ziemba's(1983) result is obtained. That is, the portfolios of all investors who are expected utility maximizers are located on Markowitz's mean-variance efficient frontier. Under the multivariate skew normal distribution, investors' portfolios are located on the mean variance-skewness-efficient surface.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–291, pp. 309–314

Using the copula density function in the estimation of the conditional density function

Esmaeil Shirazi[a,*]

[a] *Faculty of Science, Gonbad Kavous university, Gonbad Kavous, Iran*

Article Info

Keywords:
1. Conditional density
2. Copula density
3. Wavelet-based method

2020 MSC:
msc1
msc2

Abstract

In this article, a new wavelet-based method for estimating the conditional density function using the wavelet method is investigated. Based on this method, we will explain how to obtain an estimate with the optimal convergence rate for the conditional density function based on the information in the quantiles and using the copula density function. We also discuss the convergence rate of the new estimator.

1. Introduction

One of the important statistical subjects with wide application in many practical problems, especially problems related to forecasting, is the estimation of the conditional probability function. For example, this function plays a central role in financial econometrics (see [2] and [7]). Now, suppose that $(X_i, Y_i)_{i\geq 1}$ be mutually independent random vectors of a pair (X, Y). We can show the structure of function $f(y \mid x)$ as below

$$f(y \mid x) = \frac{f_{XY}(x, y)}{f_X(x)} \tag{1}$$

where $f_{XY}(x, y)$ and $f_X(x)$ denote the joint density of (X, Y) and X, respectively. The kernel estimators of these functions are given by

$$\hat{f}_{n,XY}(x, y) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x)K_h(Y_i - y)$$
$$\hat{f}_{n,X}(x) = \frac{1}{n}\sum_{i=1}^{n} K'_{h'}(X_i - x)$$

*Talker
Email address: shirazi.esmaeel@gmail.com (Esmaeil Shirazi)

respectively. Here, $K_h(.) = 1/hK(./h)$ and $K'_{h'}(.) = 1/h'K(./h')$ are (rescaled) kernels whose associated sequences $h = h_n$ and $h' = h'_n$ of bandwidth vanish as $n \to \infty$. Accordingly, an estimator of $f(y \mid x)$ is given by

$$\hat{f}(y \mid x) = \frac{\hat{f}_{n,XY}(x,y)}{\hat{f}_{n,X}(x)}$$

Kernel estimators have been studied in various fields of application due to their ease of working with them. But these estimators create limitations, especially in the estimation of bounds in discontinuous functions. For this reason, it is more preferred in this situation to suggest alternative estimation methods such as wavelets, which have higher efficiency. The complete background on wavelets can be found in [10].

Following the idea of [3], we propose two new wavelet estimators for $f(y \mid x)$: a linear estimator and a non-linear hard thresholding estimator. The later estimator is entirely adaptive. We evaluate their performance by taking the mean integrated error (MISE) over a wide class of functions. We show that the introduced estimator obtains the near optimal convergence rate $(\ln n/n)^{2s/(2s+1)}$, where s represents the smoothness parameter which we will discuss about in section 2.

The following is how the paper is structured. In Section 2, we give some details about the wavelets and Besov balls and then we introduce the considered block threshold estimator for the conditional density function. The main assumptions and the main theoretical result, with discussions are presented in Section 3. Some potential applications are listed in Section 4. Finally, the proofs are collected in Section 5.

2. Wavelets and Estimator

The mathematical context of the multiresolution analysis as well as the considered estimator are presented in this section.

2.1. Wavelets

We consider an orthonormal wavelet basis generated by dilation and translation of a compactly supported "father" wavelet $\phi(.)$ and a compactly supported "mother" wavelet $\psi(.)$. For the purposes of this paper, we use the periodized wavelet bases on the unit interval. For any $x \in [0,1]$, any integer i and any $j \in \{0, ..., 2^i-1\}$, let

$$\phi_{i_0 j}(x) = 2^{i_0/2}\phi(2^{i_0}x - j),\ \psi_{ij}(x) = 2^{i/2}\psi(2^i x - j)$$

be the elements of the wavelet basis and

$$\phi^{per}_{i,j}(x) = \sum_{k \in Z} \phi_{i,j}(x-k),\ \psi^{per}_{i,j}(x) = \sum_{k \in Z} \psi_{i,j}(x-k)$$

there periodized version. There exists an integer τ such that the collection $\Omega =$ defined by $\Omega = \{\phi^{per}_{\tau,j},\ j = 0, ..., 2^\tau-1;\ \psi^{per}_{i,j},\ i = \tau, ..., \infty,\ j = 0, ..., 2^i-1\}$, constitutes an orthonormal basis of $L^2([0,1])$. In what follows, the superscript "per" will be suppressed from the notations for convenience. For any integer $k \geq \tau$, a function $f \in L^2([0,1])$ can be expanded into a wavelet series as

$$f(x) = \sum_{j \in \{0,...,2^{i_0}-1\}} \alpha_{i_0,j}\phi_{i_0,j}(x) + \sum_{i \geq i_0} \sum_{j \in \{0,...,2^i-1\}} \beta_{i,j}\psi_{i,j}(x),$$

where the scaling coefficient $\alpha_{i_0,j}$ and the wavelet coefficient $\beta_{i,j}$ are given by

$$\alpha_{i_0,j} = \int f(x)\phi_{i_0,j}(x)dx, \quad and \quad \beta_{i,j} = \int f(x)\psi_{i,j}(x)dx.$$

All the details about these wavelet bases, including the expansion into wavelet series as described above, can be found in, for example, [10] and [6].

Let $M > 0,\ s > 0,\ p \geq 1$, and $q \geq 1$. We say that a function $f(.) \in L_2(R)$ pertains to the Besov balls $B^s_{p,q}(M)$ if and only if the associated wavelet coefficients (2) satisfy

$$\left(\sum_{k=0}^{2^\tau-1} \mid \alpha_{\tau,k} \mid^p\right)^{1/p} + \left(\sum_{j=\tau}^{\infty}\left(2^{j\sigma}\left(\sum_{k=0}^{2^j-1} \mid \beta_{jk} \mid^p\right)^{1/p}\right)^q\right)^{1/q} \leq M \tag{2}$$

where $\sigma = s + 1/2 - 1/p$. and with the usual modifications for $p = \infty$ or $q = \infty$.
These sets contain function classes of significant spatial inhomogeneity, including Sobolev balls and Holder balls. Details about Besov balls can be found in, for example [10].

2.2. *A product-shaped estimator*

The kernel-based approach described above suffers from several drawbacks. From a practical point of view, the Nadaraya-Watson estimator (and its local polynomial counterpart) may be numerically unstable when the denominator is close to zero. The large-sample behavior of the estimators is also difficult to track down, due to the quotient form. This problem is usually addressed by linearizing the inverse after centering the numerator and the denominator individually; see, e.g., [4] or [1] for details. At a conceptual level, one could also argue that implementing regression estimation techniques in this setting is somewhat artificial: estimating a density, albeit a conditional one, should resort to density estimation techniques only.
To remedy these problems, we propose an estimator which builds on the idea of using pseudo-observations, i.e., a transformation of the original data. To be specific, a quantile transform of the data will be seen to lead, through a copula representation, to a product-form estimator

$$\hat{f}_n(y \mid x) = \hat{f}_Y(y)\hat{c}_n\{\hat{F}_n(x), \hat{G}_n(y)\} \tag{3}$$

where $\hat{f}_Y, \hat{c}_n, \hat{F}_n(x), \hat{G}_n(y)$ are estimators of the density f_Y of Y , the copula density c , the c.d.f. F of X and G of Y , respectively.
As will be shown, the properties of $\hat{f}_n(y \mid x)$ are easily deduced from existing results in nonparametric density estimation.

2.3. *The quantile transform*

Data transformations are common. They are often used to improve the range of applicability and performance of classical estimation techniques to deal with skewed data, heavy tails, or restrictions on the support, among others; see, e.g., Chapter 14 of [5]. In order to make inference on Y from X , an appropriate choice of transformation must be made. We will see below that, in our context, a natural candidate is the quantile transform, i.e., the mapping $X \longmapsto U = F(X)$ which turns a continuous random variable X with c.d.f. F into a uniform random variable U on the interval $[0, 1]$.

2.4. *The copula representation*

A copula is a cumulative distribution function whose margins are uniform on the interval $[0, 1]$. [9] proved the following fundamental result:

Theorem 2.1. *For any bivariate c.d.f.* $F_{X,Y}$ *on* $\mathbb{R}^2$*, with marginal c.d.f.* F *of* X *and* G *of* Y *, there exists some function* $C : [0,1]^2 \longrightarrow [0,1]$*, called the dependence or copula function, such as*

$$F_{X,Y}(x, y) = C\{F(x), G(y)\}, \quad -\infty < x, y < \infty \tag{4}$$

If F and G are continuous, this representation is unique. The copula C is itself a c.d.f. on $[0,1]^2$ *with uniform margins.*

This theorem gives a representation of the bivariate c.d.f. as a function of each univariate c.d.f. In other words, the copula captures the dependence structure in the pair (X, Y), irrespectively of the marginal distribution F and G. Simply put, it allows one to deal with the randomness of the dependence structure and the randomness of the margins separately.
Copulas are naturally linked with the quantile transform as formula (4) entails that

$$C(u, v) = F_{X,Y}\{F^{-1}(u), G^{-1}(v)\}$$

For more details regarding copulas and their properties, see, e.g., the book of Joe [12].
From now on, we assume that the copula function $C(u, v)$ has a density

$$c(u, v) = \frac{\partial^2}{\partial u \partial v} C(u, v)$$

with respect to the Lebesgue measure on $[0, 1]^2$ and that F and G are strictly increasing and differentiable with densities f and g. $C(u, v)$ and $c(u, v)$ are then the c.d.f. and density of the transformed variables $(U, V) = (F(X), G(Y))$. Upon differentiating both sides of (4), we get the joint density, viz.

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) = f(x)g(y)c\{F(x), G(y)\}.$$

This leads to the following explicit formula of the conditional density:

$$f(y \mid x) = \frac{f_{X,Y}(x, y)}{f(x)} = g(y)c\{F(x), G(y)\}. \tag{5}$$

2.5. Main Estimator

Starting from the previously stated product-type formula (5), a natural plug-in approach to build an estimator of the conditional density is using:
1- a wavelet based nonparametric estimator of the density g of Y , viz

$$\hat{g}(y) = \sum_{j \in Z} \hat{\theta}_{i_0,j} \phi_{i_0,j}(y) + \sum_{i=i_0}^{R} \sum_{j \in Z} \hat{\xi}_{ij} \psi_{i,j}(y),$$

where $\hat{\xi}_{ij} = \hat{\omega}_{i,j} I_{\left(|\hat{\omega}_{i,j}| > c\sqrt{n^{-1} \ln n}\right)}$ and $\hat{\theta}_{i_0,j},\ \hat{\omega}_{i,j}$ are defined as follow

$$\hat{\theta}_{i_0,j} = \frac{1}{n} \sum_{k=1}^{n} \phi_{i_0,j}(Y_k), \qquad \hat{\omega}_{i,j} = \frac{1}{n} \sum_{k=1}^{n} \psi_{i_0,j}(Y_k)$$

2- Given that $c(u, v)$ is the joint density of the transformed variables $(U, V) = (F(X), G(Y))$, it could be estimated in principle by the BlockShrink estimator,

$$\hat{c}_{k'}(u, v) = \sum_{k} \hat{\alpha}_{i_0,j} \phi_{i_0 j}(u, v) + \sum_{i=i_1}^{i_2} \sum_{j \in Z} \hat{\beta}^{\epsilon}_{i,j} 1_{\{|\hat{\beta}^{\epsilon}_{i,j}| \geq d\lambda_n\}} \psi^{\epsilon}_{i,j}(u, v), \quad u, v \in [0, 1].$$

Where $\lambda_n = \sqrt{\frac{\log(n-k')}{n-k'}}$, and the natural estimator of $\alpha_{j_0 k}$ and $\beta^{\epsilon}_{j,k}$ are given by

$$\hat{\alpha}_{j_0,k} = \frac{1}{n} \sum_{i=1}^{n} \phi_{j_0,k}(F(X_i), G(Y_i)), \quad \hat{\beta}^{\epsilon}_{j,k} = \frac{1}{n} \sum_{i=1}^{n} \psi^{\epsilon}_{j,k}(F(X_i), G(Y_i)).$$

These coefficients cannot be evaluated since the distributions functions associated to the marginal distributions F and G are unknown. We propose to replace these unknown distributions functions by their corresponding empirical distributions functions F_n and G_n. The modified empirical coefficients are

$$\tilde{\alpha}_{j_0,k} = \frac{1}{n}\sum_{i=1}^{n} \phi_{j_0,k}(F_n(X_i), G_n(Y_i)), \quad \tilde{\beta}^{\epsilon}_{j,k} = \frac{1}{n}\sum_{i=1}^{n} \psi^{\epsilon}_{j,k}(F_n(X_i), G_n(Y_i)). \tag{6}$$

where the empirical distribution functions are given by

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} I(X_i \leq x), \qquad G_n(y) = \frac{1}{n}\sum_{k=1}^{n} I(Y_i \leq y)$$

So the wavelet-based estimator for copula density is as

$$\tilde{c}_{k'}(u,v) = \sum_{k} \tilde{\alpha}_{j_0,k}\phi_{j_0 k}(u,v) + \sum_{i=i_1}^{i_2}\sum_{j\in Z} \tilde{\beta}^{\epsilon}_{i,j} 1_{\{|\tilde{\beta}^{\epsilon}_{i,j}| \geq d\lambda_n\}} \psi^{\epsilon}_{i,j}(u,v), \quad u,v \in [0,1]. \tag{7}$$

where where the smoothing parameters i_1 and i_2 satisfying $2^{i_1} \simeq \log_2^n$ and $2^{i_2} \simeq n(\log_2^n)^{-2}$.

3. Main Results

At the beginning of this section, we formulate some basic assumptions that are needed to prove the main results.

3.1. Assumptions

- **A1** There is some compact subset Ω_x such that for any $c > 0$, we have

$$\sup_{y\in\Omega_x} f(x,y) \leq c$$

- **A2** There exists a known constant $c' > 0$ such that $f(x) \geq c' > 0$

These assumptions are only technical and necessary to prove our main theorem.

3.2. Main Results

In the following two theorems, an upper bound for the convergence rate of the bivariate density function and conditional density function estimators is investigated. We also have two important propositions, which are used to prove these theorems.

Theorem 3.1. *Suppose that* **A1** *hold and let $c \in B^s_{p,q}(M)$ for all $M > 0$, $s > 2/p$ and $p,q \geq 1$. Consider the estimator $\tilde{c}_{k'}(u,v)$ characterized by (7) with i_1 and i_2 satisfying $2^{i_1} \simeq \log_2^n$ and $2^{i_2} \simeq n(\log_2^n)^{-2}$. Then there exists a constant $C > 0$ such that*

$$E\left(\int_{[0,1]} \mid \tilde{c}_{k'}(u,v) - c(u,v) \mid^2 dy\right) \leq C\left(\frac{\ln n}{n}\right)^{-\frac{2s}{1+2s}}$$

Theorem 3.2. *Suppose that* **A1** *and* **A2** *hold and let $c \in B^s_{p,q}(M)$ for all $M > 0$, $s > 2/p$ and $p,q \geq 1$. Consider the estimation of conditional density function $f(y \mid x)$ with estimates (6) and (7). Then for a large enough d, there exists a value of $C > 0$ such that we have*

$$E\left(\int_{[0,1]} \mid \hat{f}(y \mid x) - f(y \mid x) \mid^2 dy\right) \leq Cn^{-\frac{2s}{1+2s}}$$

The proof of Theorem 3.1 is based on a suitable decomposition of the MISE and the statistical properties of (6) presented in Lemma 2 and Lemma 3 which are discussed in [8]. Also to prove Theorem 3.2, we follow the lines of [3], Proof of Theorem 3.1.

Theorems 3.1 and 3.2 provide theoretical guarantees about the optimal convergence rate for estimators (7) and (2.2) under mild boundedness assumptions. In addition, it should be noted that in Theorem 3.2, the value of $n^{-\frac{2s}{1+2s}}$ is the standard near optimal convergence rate for the nonlinear hard threshold wavelet estimator in the standard one-dimensional density estimation problem (see [2]).

References

[1] D. Bosq. *Nonparametric Statistics for Stochastic Processes*. Springer, New York, second edition, 1998.

[2] C. Chesneau and H. Doosti. A note on the adaptive estimation of a conditional continuous-discrete multivariate density by wavelet methods. *Chinese Journal of Mathematics*, 2016:1–8, 2016.

[3] O. P. Faugeras. A quantile-copula approach to conditional density estimation. *Journal of Multivariate Analysis*, 100:2083–2099, 2009.

[4] Q. Yao J. Fan. *Nonlinear Time Series*. Springer, New York, 2005.

[5] G. Lugosi L. Devroye. *Combinatorial Methods in Density Estimation*. Springer-Verlag, New York, 1987.

[6] S. Mallat. A wavelet tour of signal processing, the sparse way, third ed. Academic Press, 2009.

[7] E. Shirazi and O. P. Faugeras. A new wavelet-based estimation of conditional density via block threshold method. *Communications in Statistics - Theory and Methods*, 2023.

[8] E. Shirazi and B. Ghanbari. Nonlinear wavelet estimation of a bivariate density function under nsd assumption. *Communications in Statistics Part B: Simulation and Computation, Under review*, 2023.

[9] A. Sklar. Fonctions de repartition a n dimensions et leurs marges. *Publ. Inst. Statist, Univ, Paris*, 8:229–231, 1959.

[10] D. Picard A. Tsybakov W. Hărdle, G. Kerkyacharian. Wavelets, approximations, and statistical applications. In *Lecture Notes in Statistics*. Springer-Verlag, Berlin, 1998.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, ST–301, pp. 315–318

A note on characterization of exponential family of distribution

Abdolsaeed Toomaj[a,*]

[a]*Department of Mathematics and Statistics, Gonbad Kavous University, Gonbad Kavous, Iran*

Article Info

Keywords:
Characterizations
Memoryless property
Exponential distribution
Exponential family of distribution

2020 MSC:
msc1
msc2

Abstract

Recently, there has been significant attention given to the characterization of probability distribution functions for random variables. This paper presents an approach to characterize the distributions of real-valued random variables based on a real, continuous, and strictly increasing function $\psi(t)$. By examining specific scenarios, we demonstrate the versatility and applicability of the developed techniques.

1. Introduction

Modern probability theory emerged in the late 19th and early 20th centuries, reaching its pinnacle with the introduction of Kolmogorov's axioms in 1933. The development of mathematical statistics came even later, with its fundamental concepts solidifying in the early 1940s. Following World War II, there was a surge in the publication of periodicals and monographs in this field, and this trend continues to this day, with an ever-increasing pace. Probability theory and mathematical statistics rely on several specific concepts that become clearer when translated into pure mathematical language. One such concept is the distribution function. Characterization theorems, situated at the intersection of probability theory and mathematical statistics, make use of various classical tools from mathematical analysis, including advanced topics such as complex variable functions, different types of differential equations, series theory, and the theory of functional equations. One significant distribution is the exponential distribution where its cumulative distribution function (CDF) is expressed as $F(x) = e^{-\lambda x}$, where x and λ are positive values greater than zero. The lack of memory property of a random variable or its distribution can be conveniently illustrated using the analogy of the lifespan of an industrial device. Let's consider a random variable X representing the lifespan of an item. It is evident that X takes values greater than zero. We say that X or its distribution function $F(x)$ exhibits the lack of memory property if, for all x and t greater than zero, such that $P(X > z) > 0$, it satisfies the following condition:

$$P(X - t > x|X > t) = P(X > x),$$

*Talker
Email address: ab.toomaj@gmail.com (Abdolsaeed Toomaj)

Later advancements in this field have superseded the aforementioned characterization. Cox [1], Cundy [2], and Shanbhag [10] have introduced a new characterization that utilizes the concept of constant expected residual life. Additional relevant results can be found in Dallas [3], Galambos and Hagwood [5], Rao and Shanbhag [7], and other related references. For instance, if a non-negative and non-degenerate random variable X satisfies the following condition:

$$E(G(X-t)|X>t)=c,\ t>0,$$

If a non-negative and non-degenerate random variable X satisfies certain conditions, where the function G fulfills mild requirements, then X follows an exponential distribution. Additionally, Zoroa *et al.* [11] established necessary and sufficient conditions for a real function $h(y)$ to be the conditional expectation $E(g(X)|X \geq t)$ of a continuous random variable Z, considering a given real, continuous, and strictly monotone function g. Ruiz *et al.* [8] extended this result to continuous multivariate random variables. For discrete cases, Ruiz and Navarro [9] and Marin *et al.* [6] provided related characterizations. Franco and Ruiz [4] also contributed corresponding characterizations for order statistics and record values, respectively. The aim of this paper is to provide some characterization results based on the exponential family of distribution.

2. Main results

Hereafter, we aim to characterize a broad range of distribution functions, dividing it into three subsections. The first subsection focuses on the characterization of the exponential family of distributions. Consider a real, continuous, and strictly increasing function $\psi(t)$. The literature pays particular attention to the exponential family (EF) of distributions based on the weight cumulative function $\psi(t)$, where we have an absolutely continuous nonnegative random variable defined by its survival function, which can be expressed as follows:

$$\overline{G}(t)=e^{-\theta\psi(t)},\ \alpha<t<\beta, \tag{1}$$

where $\theta>0$, is a positive parameter such $\psi(\alpha)=0$ and β may take the value of infinity. Specifically, if we set $\psi(t)=F(t)$, where $F(t)$ represents the cumulative distribution function, we obtain a wide range of distributions, including the following:

$$\overline{G}(t)=e^{-\theta F(t)},\ \alpha<t<\beta.$$

In the specific case where $\psi(t)=t$, the distribution reduces to the exponential distribution when $t>0$. It is widely recognized that the exponential distribution exhibits the following property:

$$\overline{G}(x+t)=\overline{G}(x)\overline{G}(t),\ x,t>0.$$

This property, known as the memoryless property, has been extensively utilized in various fields of applied probability, such as queueing theory. Now, we present the following theorem.

Theorem 2.1. *Let X be an absolutely continuous random variable with survival function $\overline{F}(x)$, PDF $f(x)$ and hazard rate function of $\lambda(x)=f(x)/\overline{F}(x)$ with the support $D=(\alpha,\beta)$. Moreover, let $\psi(x)$ be a given real, continuous and strictly increasing function. Then*

$$P(\psi(X)-\psi(t)>y|X>t)=P(\psi(X)>y),\ y>0,\ t\in D, \tag{2}$$

holds if and only if the underlying distribution is EF distribution.

Proof. Let us define $\overline{F}_{\psi}(y)=P(\psi(X)>y)$. The equation (2) can be equivalently expressed as follows:

$$\overline{F}_{\psi}(y+\psi(t))=\overline{F}_{\psi}(y)\overline{F}(t),\ y>0,\ t\in D. \tag{3}$$

By taking the derivative of (3) with respect to t and y and dividing the resulting expressions, we obtain the following relationship:

$$\phi(t)=\frac{\lambda(t)}{\lambda_{\psi}(y)},\ y>0,\ t\in D.$$

Taking into account

$$\lambda_\psi(y) = \frac{f(\psi^{-1}(y))}{\phi(\psi^{-1}(y))\overline{F}(\psi^{-1}(y))} = \frac{\lambda(\psi^{-1}(y))}{\phi(\psi^{-1}(y))},\ y > 0, \tag{4}$$

and setting $x = \psi^{-1}(y)$, then we have

$$\psi^{-1}(0) < \psi^{-1}(y) < \psi^{-1}(\infty) \iff \alpha < x < \beta.$$

So, we obtain

$$\lambda(x) = \theta\phi(x),\ \alpha < x < \beta,$$

where $\theta = \frac{\lambda(\alpha)}{\phi(\alpha)}$ is a constant. This represents the hazard rate function of the exponential family of distributions, and it concludes the proof.

□

An immediate consequence of the preceding theorem is presented in the following corollary.

Corollary 2.2. *Under the conditions of Theorem 2.1, let us define $\psi(x) = -\log \overline{F}(x)$. Then, we have the following:*

$$P(-\log \overline{F}(X) + \log \overline{F}(t) > x | X > t) = P(-\log \overline{F}(X) > t),\ x > 0,\ t \in D, \tag{5}$$

holds if and only if the underlying distribution is the proportional hazard rate model with the survival function $\overline{G}(x) = [\overline{F}(x)]^\theta$.

Let's now write Equation (5) in detail for several specific distributions that are commonly used in various applications. For each of the distributions, we will rewrite only one property of the exponential distribution that is representative and informative for the transformed distribution. Here are the characterizing properties for each distribution.

(i) The power distribution: if $\psi(x) = -\theta\log(1-x)$, from Corollary 2.2, we can conclude that

$$P(-\theta\log(1-X) + \theta\log(1-t) > x | X > t) = P(-\theta\log(1-X) > t),\ 0 < x < 1, 0 < t < 1,$$

holds if and only if the underlying distribution is the power distribution with the survival function $\overline{F}(x) = (1-x)^\theta,\ \theta > 0$, for $0 \le x \le 1$.

(ii) The Weibull distribution: if $\psi(X) = X^\theta$, from Corollary 2.2, we can conclude that

$$P(X^\theta - t^\theta > x | X > t) = P(X^\theta > t),\ x, t > 0,$$

holds if and only if the underlying distribution is the Weibull distribution with the survival function $\overline{F}(x) = e^{-x^\theta},\ \theta > 0$, for $x > 0$.

(iii) The Pareto distribution: if $\psi(X) = \theta\log X$, from Corollary 2.2, we can conclude that

$$P(\theta\log X - \theta\log t > x | X > t) = P(\theta\log X > t),\ x > 1,\ t > 1,$$

holds if and only if the underlying distribution is the Pareto distribution with the survival function $\overline{F}(x) = x^{-\theta},\ \theta > 0$, for $x \ge 1$.

The Table 1 displays some continuous distribution functions that can be characterized under a suitable choice of $\psi(x)$. Because of the obvious flexibility of Eq. (2), it is very useful in several characterization results,

3. Conclusion

The characterization of probability distribution functions for random variables has received considerable attention in recent times. This paper introduced a novel approach to characterize the distributions of real-valued random variables using a real, continuous, and strictly increasing function $\psi(t)$. Through the examination of specific scenarios, we have showcased the versatility and practical applicability of the developed techniques. By applying these methods to concrete examples, we have demonstrated their effectiveness in capturing the essential characteristics of various distributions. The proposed approach offers a valuable tool for understanding and analyzing the probability distributions of real-valued random variables, with potential applications in a wide range of domains.

Table 1. Examples of Exponential Family of Distributions.

Distribution	Survival function	Parameter	$\psi(x)$
Exponential	$\overline{F}(x) = e^{-\theta x},\ x > 0,$	θ	x
Weibull	$\overline{F}(x) = e^{-(\lambda x)^{\alpha}},\ x > 0,$	λ^{α}	x^{α}
Pareto Type I	$\overline{F}(x) = \left[\frac{x}{\mu}\right]^{-\alpha},\ x > \mu,$	α	$\log(\frac{x}{\mu})$
Pareto Type II	$\overline{F}(x) = \left[1 + \frac{x-\mu}{\sigma}\right]^{-\alpha},\ x > \mu,$	α	$\log(1 + \frac{x-\mu}{\sigma})$
Lomax	$\overline{F}(x) = \left[1 + \frac{x}{\sigma}\right]^{-\alpha},\ x > 0,$	α	$\log(1 + \frac{x}{\sigma})$
Gompertz	$\overline{F}(x) = e^{-\lambda(e^{bx}-1)},\ x > 0,$	λ	$e^{bx} - 1$
Beta of the first kind	$\overline{F}(x) = (1-x)^{\alpha},\ 0 < x < 1,$	α	$-\log(1-x)$
Beta of the second kind	$\overline{F}(x) = (1+x)^{-1},\ x > 0,$	1	$\log(1+x)$
Log logistic	$\overline{F}(x) = (1+ax^b)^{-1},\ x > 0,$	1	$\log(1+ax^b)$
Burr type XII	$\overline{F}(x) = (1+ax^b)^{-\theta},\ x > 0,$	θ	$\log(1+ax^b)$

References

[1] DR Cox. Renewal theory, methuen & co. ltd, 1962.

[2] H Martyn Cundy. 138. birds and atoms. *The Mathematical Gazette*, 50(373):294–295, 1966.

[3] AC Dallas. Characterizing the pareto and power distributions. *Annals of the Institute of Statistical Mathematics*, 28:491–497, 1976.

[4] M Franco and JM Ruiz. On characterization of continuous distributions with adjacent order statistics. *Statistics: A Journal of Theoretical and Applied Statistics*, 26(4):375–385, 1995.

[5] Janos Galambos and Charles agwood. The characterization of a distribution function by the second moment of the residual life. *Communications in Statistics-Theory and Methods*, 21(5):1463–1468, 1992.

[6] J Marín, JM Ruiz, and P Zoroa. Characterization of discrete random vectors by conditional expectations. *journal of multivariate analysis*, 58 (1):82–95, 1996.

[7] CR Rao and DN Shanbhag. Choquet-deny type functional equations with applications to stochastic models, 1994.

[8] JM Ruiz, J Marin, and P Zoroa. A characterization of continuous multivariate distributions by conditional expectations. *Journal of statistical planning and inference*, 37(1):13–21, 1993.

[9] Jose M Ruiz and Jorge Navarro. Characterization of distributions by relationships between failure rate and mean residual life. *IEEE Transactions on Reliability*, 43(4):640–644, 1994.

[10] DN Shanbhag. The characterizations for exponential and geometric distributions. *Journal of the American Statistical Association*, 65(331): 1256–1259, 1970.

[11] P Zoroa, JM Ruiz, and J Marin. A characterization based on conditional expectations. *Communications in Statistics-Theory and Methods*, 19 (8):3127–3135, 1990.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–125, pp. 319–321

On inner uniform acts

Mohammad Roueentan[a,*], Roghaieh Khosravi[b], Elham Rahimi[c]

[a] *College of Engineering, Lamerd Higher Education Center, Lamerd, Iran*
[b] *Department of Mathematics, Faculty of Sciences, Fasa University, Fasa, Iran*
[c] *Department of Mathematics, Shiraz Branch, Islamic Azad University, Shiraz, Iran*

Article Info

Keywords:
act
inner uniform
monoid

2020 MSC:
20M30

Abstract

In the present work the concept of inner uniform acts is introduced and investigated. We bring out some general properties of these classes of acts and also their relations with some concepts for example duo, coregular and quasi-injective acts are studied.

1. Introduction

Throughout this paper S will denote a monoid and an S-act A_S (or A) is a right S-act. Recall that a subact B of a right S-act A is called $large$ (or $essential$) in A if any S-homomorphism $g : A \longrightarrow C$ such that $g|_B$ is a monomorphism is itself a monomorphism (see [1]). Also from [4] an S-act A is said to be $uniform$ if any nonzero subact of A is essential. As a special case of the concept of uniform acts in this paper we introduce and study the notion of inner uniform acts. We investigate conditions that are relevant to these classes of acts and under some special conditions we characterize monoids which over them any cyclic S-act is inner uniform. It is shown that for an injective act A the concepts uniformity and inner uniformity are equivalent. For an S-act A a nonzero subact B of A is called $inner\ essential$ and denoted by $B \subseteq^{ie} A$ if any endomorphism $g : A \longrightarrow A$ such that $g|_B$ is a monomorphism is itself a monomorphism and A is called $inner\ uniform$ if any nonzero subact of A is inner essential. Here it is necessary to remember some concepts.

A right S-act A is called *injective* if for any S-act B, any subact C of B and any homomorphism $f : C \longrightarrow A$, there exists a homomorphism $\bar{f} : B \longrightarrow A$ such that $\bar{f} |_C = f$ (see [1]). Also from [5] the right S-act A is called *quasi-injective* if it is injective relative to all inclusions from its subacts. For the sake of simplicity, we denote "quasi-injective" by "Q-injective". Recall from [6], the right S-act A is called *C-injective* if it is injective relative to every inclusion from cyclic acts. For an S-act A_S, by $E(A)$, we mean the injective envelope of A. We refer the reader to [1] for all concepts and basic properties of S-acts not defined here.

---

*Talker

Email addresses: m.rooeintan@yahoo.com (Mohammad Roueentan), khosravi@fasau.ac.ir (Roghaieh Khosravi), rahimie@shirazu.ac.ir (Elham Rahimi)

2. Main Results

Definition 2.1. Let A be an S-act over a monoid S. A nonzero subact B of A is called $inner\ essential$ and denoted by $B \subseteq^{ie} A$ if any endomorphism $g : A \longrightarrow A$ such that $g|_B$ is a monomorphism is itself a monomorphism and A is called $inner\ uniform$ if any nonzero subact of A is inner essential. Also a monoid S is called $right\ inner\ uniform$ if the right S-act S_S is inner uniform.

Regarding the above definition it is clear that any uniform S-act is inner uniform and also clearly $\Theta \sqcup \Theta$ is inner uniform.

Proposition 2.2. *Suppose A is an S-act over a monoid S. Then the following conditions are equivalent:*

(i) *Any nonzero subact of A is inner essential.*

(ii) *Any nonzero finitely generated subact of A is inner essential.*

(iii) *Any nonzero cyclic subact of A is inner essential.*

(iv) *Any nonzero indecomposable subact of A is inner essential.*

Proposition 2.3. *Over a monoid S an injective S-act A is uniform if and only if it is inner uniform.*

Proposition 2.4. *Suppose a monoid S is right inner uniform. Then for any elements $x, y \in S$ if xy is left cancellable then x is also left cancellable.*

From [2] a subact B of an S-act A is called $fully\ invariant$ if for any endomorphism $f : A \longrightarrow A, f(B) \subseteq B$. Also A is said to be duo if any subact of A is fully invariant.

Proposition 2.5. *Suppose $B \subseteq C \subseteq A$ are S-acts over a monoid S. Then the following hold:*

(i) *If $B \subseteq^{ie} A$, then $C \subseteq^{ie} A$.*

(ii) *If A is Q-injective and $B \subseteq^{ie} A$ then $B \subseteq^{ie} C$.*

(iii) *If C is a fully invariant subact of A and $B \subseteq^{ie} C \subseteq^{ie} A$, then $B \subseteq^{ie} A$.*

Corollary 2.6. *Suppose $B \subseteq C \subseteq A$ are S-acts over a monoid S and A is a Q-injective duo act. Then $B \subseteq^{ie} A$ if and only if $B \subseteq^{ie} C \subseteq^{ie} A$.*

Corollary 2.7. *The following hold over a monoid S:*

(i) *Any nonzero subact of any Q-injective inner uniform S-act is inner uniform.*

(ii) *Any essential duo extension of any inner uniform S-act is also inner uniform.*

Proposition 2.8. *Suppose for an S-act $A, E(A)$ is a duo S-act. Then the following conditions are equivalent:*

(i) *A is inner uniform.*

(ii) *$E(A)$ is inner uniform.*

(iii) *$E(A)$ is uniform.*

(iv) *A is uniform.*

From [3] an S-act A is called $coregular$ if any cyclic subact of A is injective.

Corollary 2.9. *Suppose an S-act A is coregular. Then the following conditions are equivalent:*

(i) *A is inner uniform.*

(ii) *A is uniform.*

(iii) *A is a zero simple S-act.*

Corollary 2.10. *Suppose over a monoid S any S-act is C-injective. Then the following conditions are equivalent:*

(i) *Any cyclic S-act is inner uniform $(uniform)$.*

(ii) *$S = G \sqcup \Theta$ where G is a group and Θ is the one element S-act.*

References

[1] Kilp, M., Knauer, U., Mikhalev, A. V.: Monoids, Acts and Categories, With Application to Wreath Product, Berlin; New York(2000)
[2] Roueentan, M., Ershad, M., Strongly duo and duo right S-acts, Italian Journal of Pure and Applied mathematics, 32, 143-154 (2014).
[3] Roueentan, M., Khosravi, R., On coregular S-acts, the 52^{th} Annual Iranian Mathematics Conference, (2021.
[4] Roueentan, M., Sedaghatjoo, M., On uniform acts over semigroups, Semigroup Forum, 97, 229-243 (2018)
[5] Satyanarayana, M., Quasi and weakly injective S-systems. Math. Nachr, 71(1), 183–190 (1973).
[6] Zhang, X., Knauer, U., Chen, Y., Classification of monoids by injectivities I. C-injectivity, Semigroup Forum, 76, 169-176 (2008).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–127, pp. 322–324

Stanley's Conjecture and Stanley Cohen-Macaulay Ideals

Seyed Mohammad Ajdani[a,*], Kamal Ahmadi[a], Asghar Madadi[a]

[a] *Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*

Article Info

Keywords:
Stanley depth
Stanley Cohen-Macaulay
Prime filtration

2020 MSC:
13A30
13C12, 13F55

Abstract

We introduce the notion of Stanley Cohen-Macaulay modules and show that if $u \in S$ is a regular element on S/I, where $I \subset S = K[x_1, \ldots, x_n]$ is a monomial ideal, then I is a Stanley Cohen-Macaulay ideal if and only if (I, u) is a Stanley Cohen-Macaulay ideal.

1. Introduction

Let $S = K[x_1, \ldots, x_n]$ be a polynomial ring in n variables over a field K and M be a finitely generated $\mathbb{Z}^n$-graded S-module. Let $m \in M$ be a homogeneous element in M and $Z \subseteq \{x_1, \ldots, x_n\}$. We denote by $mK[Z]$ the K-subspace of M generated by all elements mf where f is a monomial in $K[Z]$. The $\mathbb{Z}^n$-graded K-subspace $mK[Z] \subset M$ is called a Stanley space of dimension $|Z|$, if $mK[Z]$ is a free $K[Z]$-module. A Stanley decomposition of M is a presentation of the K-vector space M as a finite direct sum of Stanley spaces

$$\mathcal{D} : M = \bigoplus_{i=1}^{r} m_i K[Z_i].$$

Set $(\mathcal{D}) = \min\{|Z_i| : i = 1, \ldots, r\}$. The number

$$(M) = \max\{(\mathcal{D}) : \mathcal{D} \text{ is a Stanley decomposition of M}\}$$

is called Stanley depth of M. R. P. Stanley [2] conjectured that $(M) \leq (M)$ for all finitely generated $\mathbb{Z}^n$-graded S-modules M.

A chain of $\mathbb{Z}^n$-graded submodules $\mathcal{F} : 0 = M_0 \subset M_1 \subset \ldots \subset M_r = M$ is called a prime filtration of M if $M_i/M_{i-1} \cong S/P_i(-a_i)$ Where $a_i \in \mathbb{Z}^n$ and each P_i is a monomial prime ideal. We call the set $\{P_1, \ldots, P_r\}$ the

---

*Talker

Email addresses: majdani2@yahoo.com (Seyed Mohammad Ajdani), kamal.ahmadi.math@gmail.com (Kamal Ahmadi), amadadi2009@gmail.com (Asghar Madadi)

support of $\mathcal{F}$ and denote it $(\mathcal{F})$.
Herzog, Vladoiu, Zheng proved in [[1], Proposition 1.3] that if $\mathcal{F}$ is a prime filtration of M, then
$\min\{\dim(S/P) : P \in \mathcal{F}\} \leq (M), (M) \leq \min\{\dim(S/P) : P \in Ass(M)\}$.
We say that a finitely generated $\mathbb{Z}^n$-graded S-module M is Stanley Cohen-Macaulay module if $(M) = (M)$, where $(M) = \min\{\dim(S/P) : P \in (M)\}$. This paper is organized as follows. In Section 1 we recall some notation and definitions which will be needed later. A monomial ideal I is called Stanley Cohen-Macaulay ideal if S/I is a Stanley Cohen-Macaulay S-module. As the main result of Section 2 we prove that if I is a monomial ideal and $u \in S$ is a regular element on S/I, then I is a Stanley Cohen-Macaulay ideal if and only if (I, u) is a Stanley Cohen-Macaulay ideal, see Theorem 3.4.
Example 3.2 shows that there exists Stanley Cohen-Macaulay modules which are not Cohen-Macaulay.

2. Preliminaries

In this section we fix some notation and recall some definitions.

Definition 2.1. Let M be a finitely generated $\mathbb{Z}^n$-graded S-module. Then the Stanley dimension of M is given by

$$(M) = \min\{\dim(S/P) : P \in (M)\}$$

Definition 2.2. Let M be a finitely generated $\mathbb{Z}^n$-graded S-module. M is Stanley Cohen-Macaulay module if $(M) = (M)$.

We also say that $x \in S$ is an M-regular element if $xz = 0$ for $z \in M$ implies $z = 0$, in other words, if x is not a zero-divisor on M.

Definition 2.3. A monomial ideal I is called Stanley Cohen-Macaulay ideal if S/I is a Stanley Cohen-Macaulay S-module.

3. Stanley Cohen-Macaulay ideals

As the main result of this paper we prove that if I is a monomial ideal and $u \in S$ is a regular element on S/I, then I is a Stanley Cohen-Macaulay ideal if and only if (I, u) is a Stanley Cohen-Macaulay ideal.

Remark 3.1. Let M be a Cohen-Macaulay finitely generated $\mathbb{Z}^n$-graded S-module and Stanley's Conjecture holds for M. Then M is a Stanley Cohen-Macaulay module.

The following example shows that converse of the above remark is not true in general.

Example 3.2. Let $S = K[x_1, x_2]$ and $M = S/(x_1^2, x_1x_2)$. Then we have $(M) = (M) = 0$. So M is a Stanley Cohen-Macaulay module. But $(M) = 0$, $\dim(M) = 1$. Thus M is not Cohen-Macaulay.

Proposition 3.3. *Let $I = (u_1, \ldots, u_r)$ be a monomial ideal, and $u = \prod_{k=1}^{t} x_{j_k}^{a_k} \in S$ be a monomial regular on S/I. Then*

(i) $x_{j_k}^{a_k} \nmid u_i$ *for all* $k = 1, \ldots, t$ *and all* $i = 1, \ldots, r$,

(ii) $I = \bigcap_{i=1}^{l} Q_i$ *is the minimal primary decomposition of* I *if and only if* $(I, u) = \bigcap_{i=1}^{l}\bigcap_{k=1}^{t}(Q_i, x_{j_k}^{a_k})$ *is the minimal primary decomposition of* (I, u).

Proof. (i) Suppose on the contrary that there exists $d \in [r]$ such that $x_{j_k}^{a_k} \mid u_d$ for all $k = 1, \ldots, t$. So $u_d = x_{j_k}^{a_k} f_d$ for some $f_d \in S$ and this implies that $f_d \notin I$ and $x_{j_k}^{a_k}(f_d + I) = I$ which is a contradiction.
(ii) Let $I = \bigcap_{i=1}^{l} Q_i$ is the minimal primary decomposition of I. We claim that $\bigcap_{i=1}^{l}\bigcap_{k=1}^{t}(Q_i, x_{j_k}^{a_k})$ is the minimal primary decomposition of (I, u). We first prove that $(I, u) = \bigcap_{i=1}^{l}\bigcap_{k=1}^{t}(Q_i, x_{j_k}^{a_k})$. Let $w \in (I, u)$. Then we have to consider two cases: If $u \nmid w$, then one has $w \in I$ and $w \in Q_i$ for all $i = 1, \ldots, l$. This implies that $w \in (Q_i, x_{j_k}^{a_k})$ for all $k = 1, \ldots, t$ and all $i = 1, \ldots, l$. Hence $w \in \bigcap_{i=1}^{l}\bigcap_{k=1}^{t}(Q_i, x_{j_k}^{a_k})$. If $u \mid w$, then $w \in (Q_i, x_{j_k}^{a_k})$ for all

$k = 1, \dots, t$ and all $i = 1, \dots, l$. So $w \in \bigcap_{i=1}^{l} \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$. Assume $w' \in \bigcap_{i=1}^{l} \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$. Then we have $w' \in (Q_i, x_{j_k}^{a_k})$ for all $i = 1, \dots, l$ and all $k = 1, \dots, t$. If $x_{j_k}^{a_k} \mid w'$ for all $k = 1, \dots, t$, then $w' \in (I, u)$. Otherwise, we have $w' \in Q_i$ for all $i = 1, \dots, l$ thus $w' \in I$ and $w \in (I, u)$. Now We show that $\bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$ is a primary ideal for all $i = 1, \dots, l$ and $k = 1, \dots, t$. Let $fg \in \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$, where $f, g \in S$ and $g \notin \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$. If $fg \in Q_i$, then there exists $n \in \mathbb{N}$ such that $f^n \in Q_i$ and $f^n \in \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$.
If $fg \in (x_{j_k}^{a_k})$ and $g \notin (x_{j_k}^{a_k})$, then there exists $m \in \mathbb{N}$ such that $f^m \in (x_{j_k}^{a_k})$ and $f^m \in \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$. It suffices to show that decomposition is minimal. Suppose on the contrary that there exists $d \in [l]$ such that $\bigcap_{i=1, i \neq d}^{l} \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k}) \subset \bigcap_{k=1}^{t} (Q_d, x_{j_k}^{a_k})$. So $\bigcap_{i=1, i \neq d}^{l} Q_i \subset Q_d$ which is a contradiction. Let $(I, u) = \bigcap_{i=1}^{l} \bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k})$ is the minimal primary decomposition of (I, u). We know that $\bigcap_{k=1}^{t} (Q_i, x_{j_k}^{a_k}) = (Q_i, u)$ and $(I, u) = \bigcap_{i=1}^{l} (Q_i, u)$. Therefore $(I, u) \setminus (u) = \bigcap_{i=1}^{l} ((Q_i, u) \setminus (u))$ and $I = \bigcap_{i=1}^{l} Q_i$ is the minimal primary decomposition of I. $\square$

Theorem 3.4. *Let $I \subset S$ be a monomial ideal of $S = K[x_1, \dots, x_n]$ and $u \in S$ be a monomial regular on S/I. Then $(S/(I, u)) = (S/I) - 1$. In particular, I is a Stanley Cohen-Macaulay ideal if and only if (I, u) is a Stanley Cohen-Macaulay ideal.*

Proof. By part (ii) Proposition 3.3 and the definition of Stanley dimension we have $(S/(I, u)) = (S/I) - 1$. Let I be a Stanley Cohen-Macaulay ideal. Then $(S/I) = (S/I)$. Also $(S/(I, u)) = (S/I) - 1$.
On the other hand, A. Rauf [[3],Theorem 2.4.1] proved that

$$(S/(I, u)) = (S/I) - 1.$$

Therefore $(S/(I, u)) = (S/(I, u))$. Now let (I, u) is a Stanley Cohen-Macaulay ideal. Then $(S/(I, u)) = (S/I) - 1 = (S/(I, u)) = (S/I) - 1$. So $(S/I) = (S/I)$. $\square$

References

[1] J. Herzog, M. Vladoiu, X. Zheng, How to compute the stanley depth of a monomial ideal, Journal of Algebra **322(9)**, (2009), 3151-3169.
[2] R. P. Stanley, Linear Diophantine Equations and Local Cohomology, Invent. Math. **68**(1982), 175-193.
[3] A. Rauf, Stanley Decomposition of Multigraded Modules and Reductions Modulo Regular Elements, ph.d Thesis.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

The 4th National Congress on Mathematics and Statistics, AL–129, pp. 325–329

On non-commuting sets in some finite groups

Maryam Ghorbani[a,*]

[a] *Department of Mathematics, Faculty of Science, University of Science and Technology of Mazandaran, P. O. Box 48518-78195, Behshahr, Iran.*

Article Info

Keywords:
pairwise non-commuting elements,
$AC-$group,
finite $p-$group,
centralizer

2020 MSC:
20C15

Abstract

Let G be a finite group. A subset X of a group G is a set of pairwise of noncommuting elements if $xy \neq yx$ for all $x \neq y \in X$. If $|X| \geq |Y|$ for any other subset Y of pairwise noncommuting elements, then X is called a maximal subset of pairwise noncommuting elements and the size of such a set is denoted by $\omega(G)$. Here, we determine the cardinality of a maximal subset of pairwise non-commuting elements for some p-groups and generalise some results that was computed for certain $p-$groups.

1. Introduction

Let G be a non-abelian group and let X be a maximal subset of pairwise non-commuting elements of G. The cardinality of such a subset is denoted by $\omega(G)$. Also, $\omega(G)$ is the maximal clique size in the non-commuting graph of a group G. Let $Z(G)$ be the center of G. The non-commuting graph of a group G is a graph with $G \setminus Z(G)$ as the vertices and join two distinct vertices x and y, whenever $xy \neq yx$. By a famous result of Neumann [11], answering a question of Erd$\ddot{o}$s, the finiteness of $\omega(G)$ in G is equivalent to the finiteness of the factor group $\frac{G}{Z(G)}$. Pyber [13] has shown that there is a constant c such that $|G : Z(G)| \leq c^{\omega(G)}$. Moreover, figure out $\omega(G)$ for various families of groups are attracted by the several authors, for example, $[1, 2, 3, 4, 5, 7, 8, 9, 12]$. A finite $p-$group G is called extraspecial if the center, the Frattini subgroup and the derived subgroup of G all coincide and are cyclic of order p. Chin [7] obtained upper and lower bounds for $\omega(G)$ for an extra-special $p-$group G, where p is an odd prime number.
For $p = 2$, Issacs $[6\ p.40]$ showed that $\omega(G) = 2n + 1$ for any extra-special group G of order 2^{2n+1}. Mason [10] gave a bound for $\omega(G)$ by covering the group G by $(\dfrac{1}{2}|G| + 1)$ abelian subgroups. The cardinalities of maximal subsets of pairwise non-commuting elements of extraspecial $p-$groups are important as they provide combinatorial information which can be used to calculate their cohomology lengths. (The cohomology length of a nonelementary abelian $p-$group is a cohomology invariant derived from a theorem of Serre [15]). Azad [5] proved that $\omega(G) = p+1$ for any finite $p-$group G with central quotient of order p^2, where p is a prime number. Moreover, they also determined $\omega(G)$ for any nonabelian group of order p^4. Orfi [12] determined $\omega(G)$ for $p-$groups of order p^5. Fouladi and Orfi

*Talker
Email address: m_ghorbani@mazust.ac.ir (Maryam Ghorbani)

[9] proved that $\omega(G) = |G^{'}|(p+1)/p$, where G is a finite nonabelian metacyclic $p-$group with p^2. Further, Fouladi and Orfi [8] determined $\omega(G)$ for some $p-$groups G of maximal class. In this paper, we generalise the results of [5], namely Lemmas $[3.1\ ,\ 3.2]$.

2. Notation and preliminary results

Throughout of the paper, G denotes a finite group. Then $Z(G), C_G(x)$ denotes respectively the center and the centralizer of an element $x \in G$. If $x, y \in G$ then $[x, y] = x^{-1}y^{-1}xy$. A group G is called $AC-$group if the centralizer of every non-central element of G is abelian. In this section we give some basic results that are needed for the rest of the paper. We start the following lemma, while is an easy exercise.

Lemma 2.1. *Let G be a finite group. Then*
1) For any subgroup H of G, $\omega(G) \leq \omega(G)$.
2) For any normal subgroup N of G, $\omega(\frac{G}{N}) \leq \omega(G)$.

Lemma 2.2. $[14,\ Lemma\ 3.2]$ *The following statements are equivalent:*
1) G is an $AC-$*group.*
2) If $[x, y] = 1$, *then* $C_G(x) = C_G(y)$, *where* $x, y \in G \setminus Z(G)$.
3) If $[x, y] = [x.z] = 1$, *then* $[y, z] = 1$, *where* $x \in G \setminus Z(G)$.
4) If A and B are subgroups of G and $Z(G) < C_G(A) \leq C_G(B) < G$, *then* $C_G(A) = C_G(B)$.

Remark 2.3. If G is an $AC-$group, then $\{C_G(x) | x \in G \setminus Z(G)\}$ is the set of maximal abelian subgroups.

Lemma 2.4. $[8,\ Lemma\ 2.2]$. *Let G be an* $AC-$*group. Then*
1) If $x, y \in Z(G)$ *with distinct centralisers, then* $C_G(x) \bigcap C_G(G) = Z(G)$.
2) If $G = \bigcup_{i=1}^{k} C_G(x_i)$, *where* $C_G(x_i)$ *and* $C_G(y_i)$ *are distinct for* $1 \leq i < j \leq k$, *then* $\{x_1, x_2, ..., x_k\}$ *is a maximal set of pairwise non-commuting elements of G.*

Lemma 2.5. $[3,\ Lemma\ 2.3]$ *Let G be a finite* $AC-$*group. Then* $G = \bigcup_{i=1}^{k} C_G(x_i)$, *where* $C_G(x_i)$ *are distinct for* $i \neq j$ *and* $\{x_1, ..., x_k\}$ *is a maximal set of pairwise non-commuting elements of G.*

Lemma 2.6. $[12,\ Lemma\ 3.4]$. *Let* $G = H \times K$, *where H and K are nonabelian subgroups of G. Then,* $\omega(G) \geq \omega(H)\omega(K)$.

Lemma 2.7. *Let H and K be groups. Then*
1) If K is an $AC-$*group and* $H^{'} = 1$ *then* $H \times K$ *is also an* $AC-$*group.*
2) If H, K *and* $H \times K$ *all are* $AC-$*groups, then* $\omega(H \times K) = \omega(H)\omega(K)$.
3) If H is a nilpotent $AC-$*group, then H is a metabelian.*

Lemma 2.8. $[14,\ Proposition\ 3.10]$. *Let G be a* $p-$*group. Then*
1) If G has an abelian subgroup of index p, then G is an $AC-$ *group.*
2) If G has an abelian subgroup A of index p^2, *but no abelian subgroup of index p, then G is an* $AC-$*group if and only if* $C_G(x) \bigcap A = Z(G)$ *for every* $x \in G \setminus A$.

Lemma 2.9. *Let G be a $p-$group of maximal class and order p^n $(n \geq 4)$ with positive degree of commutativity which possesses an abelian maximal subgroup. Then :*
1) *G is an $AC-$group,*
2) $\omega(G) = p^{p-2} + 1$.

Lemma 2.10. *If G is a $2-$group of maximal class and order 2^n, then $\omega(G) = 2^{n-2} + 1$*

Corollary 2.11. *Let G be a $3-$group of maximal class and order 3^n.*
1) *If G possesses an abelian maximal subgroup, then $\omega(G) = 3^{n-2} + 1$.*
2) *If G possesses no abelian maximal subgroup, then $\omega(G) = 3^{n-2} + 4$.*

Corollary 2.12. *Let G be a $p-$group of maximal class and order p^5.*
1) *If G possesses an abelian maximal subgroup, then $\omega(G) = p^3 + 1$.*
2) *If G possesses no abelian maximal subgroup, then $\omega(G) = p^3 + p + 1$.*

3. Main results

In this section, we generalize the results of [5], namely below Lemmas:

Lemma 3.1. *Let G be a group of order p^n with the central quotient of order p^2, where p is a prime number. Then $\omega(G) = p + 1$.*

Lemma 3.2. *Let G be a group of order p^n with the central quotient of order p^3, where p is a prime number.*
1) G is an $AC-$group.
2) If G possesses an abelian maximal subgroup, then there exists an element x in $G \setminus Z(G)$ such that $C_G(x)$ is of order p^{n-1} and $C_G(x)$ is uniquely determined.

Lemma 3.3. *Let G be a group of order p^n with the central quotient of order p^3, where p is a prime number.*
1) If G possesses no abelian maximal subgroup, then $\omega(G) = p^2 + p + 1$.
2) If G possesses an abelian maximal subgroup, then $\omega(G) = p^2 + 1$.

Lemma 3.4. *Let G be a non-abelian group of order p^4.*
1) If G is of maximal class, then $\omega(G) = 1 + p^2$.
2) If G is of class two, then $\omega(G) = 1 + p$.

Theorem 3.5. *Let G be a non abelian $p-$group of order p^n. Suppose $|Z(G)| = p^r$ with $n - r \geq 3$ and G has an abelian maximal subgroup. Then there exist an element $x \in G \setminus Z(G)$ such that $|C_G(x)| = p^{n-1}$ and $C_G(x)$ is uniquely determined. Moreover, $\omega(G) = p^{n-r-1} + 1$.*

Proof. Since G has an abelian maximal subgroup, G is an $AC-$ group by Lemma 2.8(1). By Remark 2.3, there exist a non-central element x such that $|C_G(x)| = p^{n-1}$. Suppose for $y \neq x$, we have $|C_G(y)| = p^{n-1}$ such that $C_G(y) \neq C_G(x)$. Then

$$p^{n-r} = |\frac{G}{Z(G)}| = |\frac{G}{C_G(x) \bigcap C_G(y)}| \leq |\frac{G}{C_G(x)}||C_G(y)| = p^2$$

,
which is impossible.Hence $|C_G(x)|$ is uniquely determined. Next, we determine the $\omega(G)$ in the following two cases.......

□

Corollary 3.6. *Let G be a non abelian $p-$group of order p^n. Suppose $|Z(G)| = p^r$ with $n - r \geq 3$ and G has an abelian maximal subgroup. Then*
1) *If* $|\frac{G}{Z(G)}| = p^3$, *then* $\omega(G) = p^{n-2} + 1$.
2) *If* $|\frac{G}{Z(G)}| = p^4$, *then* $\omega(G) = p^{n-3} + 1$.
3)*If* $|\frac{G}{Z(G)}| = p^5$, *then* $\omega(G) = p^{n-4} + 1$.

Theorem 3.7. *Let G be a $p-$group of order p^6. Suppose G is not $AC-$group and $|\frac{G}{Z(G)}| = p^4$. Then we have the following:*
1) *There exist a non-central element x such that*

$$|C_G(x)| = p^5 \quad and \quad \frac{C_G(x)}{Z(C_G(x))} = C_p \times C_p.$$

2) *If there is an element x such that $|C_G(x)| = p^3 \quad or \quad p^4, \quad then \ C_G(x)$ is abelian.*

Theorem 3.8. *Let G be an an $AC-$ group with $|\frac{G}{Z(G)} = p^5|$ where p is odd. Suppose G has no abelian maximal subgroup and let $X = \{x_1, x_2, ...x_k\}$ is a maximal set of non-commutative elements in G. Then we have the following:*

1) *If G has a non-central element x such that $|C_G(x)| = p^{n-2}$, then $C_G(x)$ is uniquely determined. Further, suppose $|C_G(x_i)| = p^{n-3}$ with $2 \leq i \leq r+1$ and $|C_G(x_j)| = p^{n-4}$ for $r+2 \leq j \leq k$. Then $\omega(G) = p^4 + p^3 - rp + 1$ and $r \neq k - r - 1$.*

2) *If G has no non-central element x such that $|G(x)| = p^{n-2}$, then $\omega(G) = p(p+1)(p^2+1) + 1 - r(p+1) + r$, where $|C_G(x_i)| = p^{n-3}$ for $1 \leq i \leq r$, $|C_G(x_j)| = p^{n-4}$ for $r+1 \leq j \leq k$ and $rneqk$.*

Proof. By the hypothesis, G has no abelian maximal subgroup and hence cardinality of centralizer of any non-central element is either $p^{n-4} \quad or \quad p^{n-3} \quad or \quad p^{n-2}$. Write $G = \bigcup_{i=1}^{k} C_G(x_i)$........

□

References

[1] A. Abdollahi, A. Akbari and H. R. Maimani, Non-commuting graph of a group, J. Algbera, 298(2006), 468-492.
[2] C.P. Anil Kumar, S. K. prajapati, Maximal Non-commutative Sets in Unipotent Upper-triangular Linear Groups $UU_4(\mathbb{F}_{\shortparallel})$, submitted for publication.
[3] A. Azad and Cheryl E. Praeger, Maximal subsets of pairwise non-commuting elements of three-dimensional general linear groups, Bull. Aust. Math. Soc., 80(2009), 91-104.
[4] A. Azad and M. A. Iranmanesh, Cheryl E. Praeger and P. Spiga, Abelian coverings of finite general linear groups and an application to their non-commuting graphs, J. Algebr. Comb., 34(2011), 683-710.
[5] A. Azad, S. Fouladi and R. Orfi, Maximal subsets of pairwise non-commuting elements of some finite $p-$groups, Bull. Iran. Math. Soc., 39(1) (2013), 187-192.
[6] E. A. Bertram, Some applications of graph theory to finite groups, Discrete Math., 44(1983), 31-43.
[7] A. Y. M. Chin, On non-commuting sets in an extraspecial $p-$group, J. Group Theory, 8 (2) (2005), 189-194.
[8] S. Fouladi and R. Orfi, Maximal Subgroups of pairwise non-commuting elements of some $p-$groups of maximal class, Bull. Aust. Math. Soc., 84 (2011), 447-4510
[9] S. Fouladi and R. Orfi, Maximal size of subsets of pairwise non-commuting elements in finite metacyclic $p-$groups, Bull. Aust. Math. Soc., 87 (2013), 18-23.
[10] D. R. Mason, On coverings of a finite group by abelian subgroups, Math. Proc. Combridge Philos. Soc., 83(2) (1978), 205-209.

[11] B. H. Neumann, A problem of Paul Erdos on groups, J. Austral. Math. Soc., 21(Series A) (1976), 467-472.
[12] R. Orfi, Maximal Subsets of pairwise non-commuting elements of $p-$groups of order less than p^6, Int. J. Group Theory, 3 (1) (2014) 65-72.
[13] L. Pyber, The number of pairwise non-commuting elements and the index of the center in a finite group, J. London Math. Soc., 35(2), 1987.
[14] D. M. Rocke, $p-$groups with abelian centralizers, Proc. London Math. Soc., 30 (3) (1975) 55-75.
[15] J. P. Serre, Sur la dimension cohomologique des groups profinis, Topology, 3 (1965), 413-420.

”

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–130, pp. 330–334

Quasi-permutation representations of special orthogonal groups

Maryam Ghorbani[a,*]

[a] *Department of Mathematics, Faculty of Science, University of Science and Technology of Mazandaran, P. O. Box 48518-78195, Behshahr, Iran.*

Article Info

Keywords:
character table
quasi-permutation representation
orthogonal group

2020 MSC:
20C15

Abstract

Let G be a finite group of degree n. We shall say that G is a quasi-permutation group if the trace of every element of G is a non-negative rational integer. By a quasi-permutation matrix we mean a square matrix over the complex field C with non-negative integral trace. For a given finite group G, let $q(G)$ denote the minimal degree of a faithful representation of G by quasi-permutation matrices over the rational field Q, and let $c(G)$ be the minimal degree of a faithful representation of G by complex quasi-permutation matrices. Finally $r(G)$ denotes the minimal degree of a faithful rational valued complex character of G. The purpose of this paper is to calculate above quantities for finite special orthogonal groups.

1. Introduction

In [13], [14] Wong defined a quasi-permutation group of degree n to be a finite group G of automorphisms of an n-dimensional complex vector space such that every element of G has non-negative integral trace. Also Wong studied the extent to which some facts about permutation groups generalize to the quasi-permutation group situation. Then in 1994 Hartley with their colleague investigated further the analogy between permutation groups and quasi-permutation groups by studying the relation between the minimal degree of a faithful permutation representation of a given finite group G and the minimal degree of a faithful quasi-permutation representation. They also worked over the rational field and found some interesting results (See [4]).

If F is a subfield of the complex numbers C, then a square matrix over F with non-negative integral trace is called a quasi-permutation matrix over F. Thus every permutation matrix over C is a quasi-permutation matrix. For a given finite group G, let $p(G)$ denote the minimal degree of a faithful permutation representation of G or of a faithful representation of G by permutation matrices, let $q(G)$ denote the minimal degree of a faithful representation of G by quasi-permutation matrices over the rational field Q and let $c(G)$ be the minimal degree of a faithful representation of G by complex quasi-permutation matrices.

By a rational valued character we mean a character χ corresponding to a complex representation of G such that $\chi(g) \in Q$ for all $g \in G$. As the values of the characters of a complex representation are algebraic numbers, a rational valued character is in fact integer valued. A quasi-permutation representation of G is then simply a complex representation of G whose character values are rational and non-negative. The module of such a representation will be called a quasi-permutation module. We will call a homomorphism from G to $GL(n, Q)$ a rational representation

*Talker
Email address: m_ghorbani@mazust.ac.ir (Maryam Ghorbani)

of G and its corresponding character will be called a rational character of G. Let $r(G)$ denote the minimal degree of a faithful rational valued character of G. It is easy to see that for a finite group G the following inequalities hold

$$r(G) \leq c(G) \leq q(G) \leq p(G).$$

It is easy to see that if G is a symmetric group of degree 6, then $c(G) = q(G) = p(G) = 6$. If G is a cyclic group of order 6, then $c(G) = q(G) = 4$ and $p(G) = 5$, while on the other hand, if G is a quaternion group of order 8, then $c(G) = 4$ and $q(G) = p(G) = 8$. Thus, both inequalities can be strict. It is not too hard to see that for the group $SL(2,5)$, both inequalities are strict. Our principal aim in this paper is to investigate these quantities and inequalities further.

Finding the above quantities have been carried out in some papers, for example in [5], [6], [7] and [9] we found these for the groups $GL(2,q)$, $SU(3,q^2)$, $PSU(3,q^2)$, $SL(3,q)$, $PSL(3,q)$ and $G_2(2^n)$ respectively. In [3] we found the rational character table and the values of $r(G)$, $c(G)$, $q(G)$ and $p(G)$ for the group $PGL(2,q)$. In this paper we will apply the algorithms in [1] to the special orthogonal groups.

2. Notation and preliminary results

Assume that E is a splitting field for G and that F is a subfield of E. If $\chi, \psi \in \mathrm{Irr}_E(G)$ we say that χ and ψ are Galois conjugate over F if $F(\chi) = F(\psi)$ and there exists $\sigma \in \mathrm{Gal}(F(\chi)/F)$ such that $\chi^\sigma = \psi$, where $F(\chi)$ denotes the field obtained by adding the values $\chi(g)$, for all $g \in G$, to F. It is clear that this defines an equivalence relation on $\mathrm{Irr}_E(G)$. (see [11])

Let η_i for $0 \leq i \leq r$ be Galois conjugacy classes of irreducible complex characters of G. For $0 \leq i \leq r$ let φ_i be a representative of the class η_i, with $\varphi_o = 1_G$. Write $\Psi_i = \sum_{\chi_i \in \eta_i} \chi_i$ and $K_i = ker\varphi_i$. We know that $K_i = ker\Psi_i$. For $I \subseteq \{0,1,2,\cdots,r\}$, put $K_I = \bigcap_{i \in I} K_i$.

By definition of $r(G)$, $c(G)$ and using above notations we have:

$r(G) = \min\{\xi(1) : \xi = \sum_{i=1}^{r} n_i\Psi_i, n_i \geq 0, K_I = 1 \; for \; I = \{i, i \neq 0, n_i > 0\}\}$

$c(G) = \min\{\xi(1) : \xi = \sum_{i=0}^{r} n_i\Psi_i, n_i \geq 0, K_I = 1 \text{ for } I = \{i, i \neq 0, n_i > 0\}\}$

where $n_0 = -\min\{\xi(g) | g \in G\}$.

In [1] we defined $d(\chi), m(\chi)$ and $c(\chi)$ [See Definition 3.4]. Here we can redefine it as follows:

Definition 2.1. Let χ be a complex charater of G, such that $\ker \chi = 1$ and $\chi = \chi_1 + \cdots + \chi_n$ for some $\chi_i \in Irr(G)$. Then define

(1) $d(\chi) = \sum_{i=1}^{n} |\Gamma_i(\chi_i)|\chi_i(1)$,

(2) $m(\chi) = \begin{cases} 0 & \text{if } \chi = 1_G, \\ |\min\{\sum_{i=1}^{n} \sum_{\alpha \in \Gamma_i(\chi_i)} \chi_i^\alpha(g) : g \in G\}| & \text{otherwise,} \end{cases}$

(3) $c(\chi) = \sum_{i=1}^{n} \sum_{\alpha \in \Gamma_i(\chi_i)} \chi_i^\alpha + m(\chi)1_G$.

So

$$r(G) = \min\{d(\chi) : \ker \chi = 1\},$$

and

$$c(G) = \min\{c(\chi)(1) : \ker \chi = 1\}.$$

We can see all the following statements in [1], [2].

Lemma 2.2. *Let χ be a character of G. Then $Ker\chi = Ker\sum_{\alpha\in\Gamma(\chi)}\chi^{\alpha}$. Moreover χ is faithful if and only if $\sum_{\alpha\in\Gamma(\chi)}\chi^{\alpha}$ is faithful.*

Lemma 2.3. *Let $\chi \in Irr(G)$,then $\sum_{\alpha\in\Gamma(\chi)}\chi^{\alpha}$ is a rational valued character of G . Moreover $c(\chi)$ is a non-negative rational valued character of G and $c(\chi) = d(\chi) + m(\chi)$.*

Lemma 2.4. *Let $\chi \in Irr(G), \chi \neq 1_G$. Then $c(\chi)(1) \geq d(\chi) + 1 \geq \chi(1) + 1$.*

Corollary 2.5. *Let $\chi \in Irr(G)$. Then*
(**1**) $c(\chi)(1) \geq d(\chi) \geq \chi(1)$ *;*
(**2**) $c(\chi)(1) \leq 2d(\chi)$ *. Equality occurs if and only if $Z(\chi)/ker\chi$ is of even order.*

3. Quasi-permutation representations

The orthogonal group of degree n over a field F (write as $O(n, F)$) is the group of n-by-n orthogonal matrices with entries from F, with the group operation that of matrix multiplication. This is a subgroup of the general linear group $GL(n, F)$ given by

$$O(n, F) = \{Q \in GL(n, F) \mid Q^T Q = QQ^T = I\}.$$

where Q^T is the transpose of Q. The classical orthogonal group over the real numbers is usually just written $O(n)$.More generally the orthogonal group of a non-singular quadratic form over F is the group of matrices preserving the form. The Cartan-Dieudonne theorem describes the structure of the orthogonal group. Every orthogonal matrix has determinant either 1 or -1. The orthogonal n-by-n matrices with determinant 1 form a normal subgroup of $O(n, F)$ known as the special orthogonal group $SO(n, F)$. If the characteristic of F is 2, then $1 = -1$, hence $O(n, F)$ and $SO(n, F)$ coincide ; otherwise the index of $SO(n, F)$ in $O(n, F)$ is 2.
For the convenience in the discussion of the characterization of the conjugacy classes of $SO(3, q)$ and $O(3, q)$, we let

$$\begin{aligned}
\Theta &= \{\theta \in GF(q) | \theta \neq 0, 4,\ \theta \ \ is\ a\ square\ and\ \theta - 4 \ \ is\ a\ square\},\\
\Gamma &= \{\gamma \in GF(q) | \gamma \neq 0, 4,\ \gamma \ \ is\ a\ nonsquare\ and\ \gamma - 4 \ \ is\ a\ nonsquare\},\\
\Pi &= \{\pi \in GF(q) \mid \pi \neq 0, 4\ ,\ \pi \ \ is\ a\ square\ and\ \pi - 4 \ \ is\ a\ nonsquare\},\\
\Xi &= \{\xi \in GF(q) \mid \xi \neq 0\ ,\ \ \xi \ \ is\ a\ nonsquare\ \ and\ \ \xi - 4 \ \ is\ \ a\ square\}.
\end{aligned}$$

Then $|\Theta| = \frac{q-5}{4}$, and $|\Gamma| = |\Pi| = |\Xi| = \frac{q-1}{4}$. We shall denote the elements of Θ, Γ, Π, Ξ by $\theta_i, \gamma_j, \pi_k$ and ξ_l respectively , where $1 \leq i \leq \frac{q-5}{4}$ and $1 \leq j, k, l \leq \frac{q-1}{4}$.

Now let $L = \{\begin{pmatrix} ad+bc & ac & bd \\ 2ab & a^2 & b^2 \\ 2cd & c^2 & d^2 \end{pmatrix} \mid ad - bc = 1\ and\ a, b, c, d \in GF(q)\}$.
Dickson [8, Theorem 178] shows that L is a normal subgroup of index 2 of $SO(3, q)$ and is isomorphic to $PSL(2, q) = SL(2, q)/\{\pm I\}$ through the isomorphism

$$\pm\begin{pmatrix} a & b \\ c & d \end{pmatrix} \longmapsto \begin{pmatrix} ad+bc & ac & bd \\ 2ab & a^2 & b^2 \\ 2cd & c^2 & d^2 \end{pmatrix}, \qquad (**)$$

for each $\pm\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL(2, q)$.

Theorem 3.1. *Let $G = O(3, q)$, then*

a) $r(G) = q - 1$

b) $c(G) = q(G) = 2(q - 1)$.

c)$Lim_{q \longrightarrow \infty} \frac{c(G)}{r(G)} = 2$.

Proof. By [12] we know that the characters $\psi_1 \times \rho_1$, $\psi_1 \times \rho_q$, $\psi_1 \times \rho_1^{'}$, $\psi_1 \times \rho_q^{'}$, $\psi_1 \times \rho_{q+1}^{(n)}$; $1 \leq n \leq \frac{q-3}{2}$, $\psi_1 \times \rho_{q-1}^{(m)}$; $1 \leq m \leq \frac{q-1}{2}$ are not faithful and (x, D) belongs to kernel of all characters , so $\bigcap_\chi ker\chi \neq 0$. But the characters $\psi_2 \times \rho_1$, $\psi_2 \times \rho_q$,$\psi_2 \times \rho_1^{'}$, $\psi_2 \times \rho_q^{'}$, $\psi_2 \times \rho_{q+1}^{(n)}$; $1 \leq n \leq \frac{q-3}{2}$, $\psi_2 \times \rho_{q-1}^{(m)}$; $1 \leq m \leq \frac{q-1}{2}$ are faithful and by Corollary 4.8 we know that $\psi_2 \times \rho_{q-1}^{(\frac{q+1}{4})}$ and $\psi_2 \times \rho_{q+1}^{(\frac{q-1}{4})}$ are rational valued characters.
Now by definition of $d(\chi)$ and $c(\chi)$ and Table (3) we have
$d(\psi_2 \times \rho_{q-1}^{(m)}) = |\Gamma|(q-1) \geq q-1$ where $\Gamma = \Gamma(Q(\psi_2 \times \rho_{q-1}^{(m)}) : Q)$ and $m(\psi_2 \times \rho_{q-1}^{(m)}) \geq q-1$ and so $c(\psi_2 \times \rho_{q-1}^{(m)})(1) \geq 2(q-1)$ and equality holds if $m = \frac{q+1}{4}$.
And $d(\psi_2 \times \rho_{q+1}^{(n)}) \geq q+1$ and $m(\psi_2 \times \rho_{q+1}^{(n)}) \geq q+1$ and so $c(\psi_2 \times \rho_{q+1}^{(n)})(1) \geq 2(q+1)$ and equality hold if $m = \frac{q-1}{4}$.
For other characters we have $d(\psi_2 \times \rho_q) = d(\psi_2 \times \rho_q^{'}) = q$ and $m(\psi_2 \times \rho_q) = m(\psi_2 \times \rho_q^{'}) = q$ and so $c(\psi_2 \times \rho_q)(1) = c(\psi_2 \times \rho_q^{'})(1) = 2q$.
The values are set out in the following table :

Table (1)

χ	$d(\chi)$	$c(\chi)(1)$
$\psi_2 \times \rho_q$	q	$2q$
$\psi_2 \times \rho_q^{'}$	q	$2q$
$\psi_2 \times \rho_{q+1}^{(n)}$	$\geq q+1$	$\geq 2(q+1)$
$\psi_2 \times \rho_{q-1}^{(m)}$	$\geq q-1$	$\geq 2(q-1)$

Now by Definition 2.1 and Table (1) we have

min $\{d(\chi) :\ ker\chi = 1\} = q-1$ and min $\{c(\chi)(1) :\ ker\chi = 1\} = 2(q-1)$.
Now by Lemmas $2.3, 2.5$ and [10] Schur index of each irreducible characters of the group $O(3, q)$ is 1, and so $c(G) = q(G)$.

□

References

[1] H. Behravesh, *Quasi-Permutation representations of p-groups of class 2*, J. London Math. Soc. (2) 55 (1997) 251-260.
[2] H. Behravesh, *The rational character table of special linear groups* J. Sci. I.R.I. vol. 9, no. 2 (1998), 173-180.
[3] H. Behravesh,A. Daneshkhah, M.R. Darafsheh and M. Ghorbany, *The rational character table and quasi-permutation representations of the group $PGL(2, q)$*, Italian Journal of Pure and Applied Mathematics N.11, (2001) 9-18.
[4] J.M. Burns, B. Goldsmith, B. Hartley and R. Sandling, *on quasi-permutation representations of finite groups*, Glasgow Math, J. 36, 301-308 (1994).
[5] M.R. Darafsheh, M. Ghorbany, A. Daneshkhah and H. Behravesh, *Quasi-permutation representation of the group $GL(2, q)$*, Journal of Algebra (2001)243, 142-167.
[6] M.R. Darafsheh and M.Ghorbany, *Quasi-permutation representations of the groups $SU(3, q^2)$ and $PSU(3, q^2)$*, Southest Asian Bulletin of Mathemetics (2002)26, 395-406.
[7] M.R. Darafsheh and M.Ghorbany, *Quasi-permutation representations of the groups $SL(3, q)$ and $PSL(3, q)$*, Iranian Journal of Science and Technology, Tra.A, Vol. 26, No. A1 (2002)145-154.

[8] L.E. Dickson, *Linear groups with an exposition of the Galois Field Theory* New York:Dover (1958).
[9] M.Ghorbany, *Special representations of the group $G_2(2^n)$ with minimal degrees*, Southest Asian Bulletin of Mathemetics (2006)30: 663-670.
[10] R. Gow, *Schur indices of some groups of Lie type* J. Algebra 42 (1976), 102-120.
[11] I.M. Isaacs, *Character theory of finite groups* , Academic Press, New York (1976).
[12] R. Steinberg, *The representation of $GL(3,q)$, $GL(4,q)$, $PGL(3,q)$ and $PGL(4,q)$*, Can.J.Math. (1951)3, 225-235.
[13] W.J. Wong, *Linear groups analogous to permutation groups* , J. Austral. Math. Soc (Sec. A) 3 (1963), 180-184.
[14] W.J. Wong , *On linear $p-$groups*, J. Austral. Math. Soc. Ser. A 4 (1964), 174-178.

”

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–136, pp. 335–339

Permutation and quasi-permutation representations of Steinberg groups

Maryam Ghorbani[a,*]

[a]Department of Mathematics, Faculty of Science, University of Science and Technology of Mazandaran, P. O. Box 48518-78195, Behshahr, Iran.

Article Info

Keywords:
character table
quasi-permutation
representation
steinberg group

2020 MSC:
20C15

Abstract

A square matrix over the complex field with non-negative integral trace is called a quasi-permutation matrix. Thus every permutation matrix over C is a quasi-permutation matrix. For a given finite group G, let $p(G)$ denote the minimal degree of a faithful permutation representation of G. The minimal degree of a faithful representation of G by quasi-permutation matrices over the rational and the complex numbers are denoted by $q(G)$ and $c(G)$ respectively. Finally $r(G)$ denotes the minimal degree of a faithful rational valued complex character of G. In this paper we will calculate $p(G)$, $q(G)$, $c(G)$ and $r(G)$ for the Steinberg groups.

1. Introduction

In [14], [15] Wong defined a quasi-permutation group of degree n to be a finite group G of automorphisms of an n-dimensional complex vector space such that every element of G has non-negative integral trace. Also Wong studied the extent to which some facts about permutation groups generalize to the quasi-permutation group situation. Then in 1994 Hartley with their colleague investigated further the analogy between permutation groups and quasi-permutation groups by studying the relation between the minimal degree of a faithful permutation representation of a given finite group G and the minimal degree of a faithful quasi-permutation representation. They also worked over the rational field and found some interesting results (See [4]).

If F is a subfield of the complex numbers C, then a square matrix over F with non-negative integral trace is called a quasi-permutation matrix over F. Thus every permutation matrix over C is a quasi-permutation matrix. For a given finite group G, let $p(G)$ denote the minimal degree of a faithful permutation representation of G or of a faithful representation of G by permutation matrices, let $q(G)$ denote the minimal degree of a faithful representation of G by quasi-permutation matrices over the rational field Q and let $c(G)$ be the minimal degree of a faithful representation of G by complex quasi-permutation matrices.

By a rational valued character we mean a character χ corresponding to a complex representation of G such that $\chi(g) \in Q$ for all $g \in G$. As the values of the characters of a complex representation are algebraic numbers, a rational valued character is in fact integer valued. A quasi-permutation representation of G is then simply a complex representation of G whose character values are rational and non-negative. The module of such a representation will

*Talker
Email address: m_ghorbani@mazust.ac.ir (Maryam Ghorbani)

be called a quasi-permutation module. We will call a homomorphism from G to $GL(n, Q)$ a rational representation of G and its corresponding character will be called a rational character of G. Let $r(G)$ denote the minimal degree of a faithful rational valued character of G. It is easy to see that for a finite group G the following inequalities hold

$$r(G) \leq c(G) \leq q(G) \leq p(G).$$

It is easy to see that if G is a symmetric group of degree 6, then $c(G) = q(G) = p(G) = 6$. If G is a cyclic group of order 6, then $c(G) = q(G) = 4$ and $p(G) = 5$, while on the other hand, if G is a quaternion group of order 8, then $c(G) = 4$ and $q(G) = p(G) = 8$. Thus, both inequalities can be strict. It is not too hard to see that for the group $SL(2, 5)$, both inequalities are strict. Our principal aim in this paper is to investigate these quantities and inequalities further.

Finding the above quantities have been carried out in some papers, for example in [5], [6], [7] and [8] we found these for the groups $GL(2, q)$, $SU(3, q^2)$, $PSU(3, q^2)$, $SL(3, q)$, $PSL(3, q)$ and $G_2(2^n)$ respectively. In [3] we found the rational character table and the values of $r(G)$, $c(G)$, $q(G)$ and $p(G)$ for the group $PGL(2, q)$. Let G be a finite group and χ be an irreducible complex character of G. Let $m_Q(\chi)$ denote the Schur index of χ over Q and $\Gamma(\chi)$ be the Galois group $Q(\chi)$ over Q. It is known that

$$\sum_{\alpha \in \Gamma(\chi)} m_Q(\chi)\chi^{\alpha} \tag{1}$$

is a character of an irreducible QG-module [[12], Corollary 10.2(b)]. So by knowing the character table of a group and Suchr indices of each of the irreducible characters of G, we can find the irreducible rational characters.

In this paper we will apply the algorithms in [1] to the Steinberg groups.

2. Notation and preliminary results

Assume that E is a splitting field for G and that F is a subfield of E. If $\chi, \psi \in \mathrm{Irr}_E(G)$ we say that χ and ψ are Galois conjugate over F if $F(\chi) = F(\psi)$ and there exists $\sigma \in \mathrm{Gal}(F(\chi)/F)$ such that $\chi^{\sigma} = \psi$, where $F(\chi)$ denotes the field obtained by adding the values $\chi(g)$, for all $g \in G$, to F. It is clear that this defines an equivalence relation on $\mathrm{Irr}_E(G)$. (see [11])

Let η_i for $0 \leq i \leq r$ be Galois conjugacy classes of irreducible complex characters of G. For $0 \leq i \leq r$ let φ_i be a representative of the class η_i, with $\varphi_o = 1_G$. Write $\Psi_i = \sum_{\chi_i \in \eta_i} \chi_i$ and $K_i = ker\varphi_i$. We know that $K_i = ker\Psi_i$. For $I \subseteq \{0, 1, 2, \cdots, r\}$, put $K_I = \bigcap_{i \in I} K_i$.

By definition of $r(G)$, $c(G)$ and using above notations we have:

$r(G) = \min\{\xi(1) : \xi = \sum_{i=1}^{r} n_i\Psi_i, n_i \geq 0, K_I = 1 \; for \; I = \{i, i \neq 0, n_i > 0\}\}$

$c(G) = \min\{\xi(1) : \xi = \sum_{i=0}^{r} n_i\Psi_i, n_i \geq 0, K_I = 1 \text{ for } I = \{i, i \neq 0, n_i > 0\}\}$

where $n_0 = -\min\{\xi(g) | g \in G\}$.

In [1] we defined $d(\chi), m(\chi)$ and $c(\chi)$ [See Definition 3.4]. Here we can redefine it as follows:

Definition 2.1. Let χ be a complex charater of G, such that $\ker\chi = 1$ and $\chi = \chi_1 + \cdots + \chi_n$ for some $\chi_i \in Irr(G)$. Then define

(1) $d(\chi) = \sum_{i=1}^{n} |\Gamma_i(\chi_i)|\chi_i(1)$,

(2) $m(\chi) = \begin{cases} 0 & \text{if } \chi = 1_G, \\ |\min\{\sum_{i=1}^{n} \sum_{\alpha \in \Gamma_i(\chi_i)} \chi_i^{\alpha}(g) : g \in G\}| & \text{otherwise,} \end{cases}$

(3) $c(\chi) = \sum_{i=1}^{n} \sum_{\alpha\in\Gamma_i(\chi_i)} \chi_i^\alpha + m(\chi)1_G.$

We can see all the following statements in [1], [2].

Lemma 2.2. *Let χ be a character of G. Then $Ker\chi = Ker\sum_{\alpha\in\Gamma(\chi)}\chi^\alpha$. Moreover χ is faithful if and only if $\sum_{\alpha\in\Gamma(\chi)}\chi^\alpha$ is faithful.*

Lemma 2.3. *Let $\chi \in Irr(G)$, then $\sum_{\alpha\in\Gamma(\chi)}\chi^\alpha$ is a rational valued character of G. Moreover $c(\chi)$ is a non-negative rational valued character of G and $c(\chi) = d(\chi) + m(\chi)$.*

Lemma 2.4. *Let $\chi \in Irr(G), \chi \neq 1_G$. Then $c(\chi)(1) \geq d(\chi) + 1 \geq \chi(1) + 1$.*

Corollary 2.5. *Let $\chi \in Irr(G)$. Then*
(**1**) $c(\chi)(1) \geq d(\chi) \geq \chi(1)$;
(**2**) $c(\chi)(1) \leq 2d(\chi)$. *Equality occurs if and only if $Z(\chi)/ker\chi$ is of even order.*

Now according to Corollary 3.11 of [1] and above statements the following result is useful for calculation of $r(G)$, $c(G)$ and $q(G)$.

Lemma 2.6. *Let G be a finite group with a unique minimal normal subgroup. Then*

1) $r(G) = \min\{d(\chi) : \chi$ *is a faithful irreducible complex character of G*$\}$

2) $c(G) = \min\{c(\chi)(1) : \chi$ *is a faithful irreducible complex character of G*$\}$

3) $q(G) = \min\{m_Q(\chi)c(\chi)(1) : \chi$ *is a faithful irreducible complex character of G*$\}$.

If the Schur index of each non-principal irreducible character of G over Q is equal to m, then from [1] Corollary 3.15 we have $q(G) = mc(G)$.

3. Permutation and Quasi-permutation representations

Using definition of $p(G)$ it is proved in [1] that

$$p(G) = \min\{\sum_{i=1}^{n}[G : H_i] : H_i \leq G, \bigcap_{i=1}^{n}\bigcap_{x\in G} H_i^x = 1\}$$

By [1] Corollary 2.4 we know that if G is a finite group with a minimal normal subgroup then $p(G)$ is the smallest index of a subgroup with trivial core. Therefore for the simple group ${}^3D_4(q)$ we introduce the maximal subgroups of ${}^3D_4(q)$ and then calculate $p(G)$ for this group.

We can see all the following statements in [10], [11], [13].

Now we introduce the maximal subgroups of the group ${}^3D_4(q)$.
Throughout this paper, H_0 denotes the finite simple Steinberg triality group ${}^3D_4(q)$ of order $q^{12}(q^6-1)^2(q^4-q^2+1)$, where $q = p^n$ and p is prime. We define H_1 =Aut(H_0) and we let H be any group with socle H_0. Thus

$${}^3D_4(q) \cong H_0 \leq H \leq H_1 \cong Aut({}^3D_4(q)). \quad (1)$$

Theorem 3.1. *1 Let H be as in (1) and assume that M is a maximal subgroup of H not containing H_0. Then $M_0 = M \cap H_0$ is H_0-conjugate to one of the following groups:*

Proof. **Table (I)**

Groups	Structure	Remarks
P_a	$[q^9] : (SL_2(q^3) \circ (Z_{q-1})).d$	$Parabolic, d = (2, q-1)$
P_b	$[q^{11}] : ((Z_{q^3-1}) \circ SL_2(q)).d$	$Parabolic, d = (2, q-1)$
$C_{H_0}(g_1)$	$G_2(q)$	
$C_{H_0}(g_2)$	$PGL_3^{\varepsilon}(q)$	$2 < q \equiv \varepsilon 1 (mod\, 3), \varepsilon = \pm$
$C_{H_0}(\phi_\alpha)$	${}^3D_4(q_0)$	$q = q_0^\alpha, \alpha\, prime, \alpha \neq 3$
$N_{H_0}(F)$	$L_2(q^3) \times L_2(q)$	$p = 2, F \cong L_2(q)$ a fundamental subgroup
$C_{H_0}(S_2)$	$(SL_2(q^3) \circ SL_2(q)).2$	p odd, involution centralizer
$N_{H_0}(< S_4 >)$	$((Z_{q^2+q+1}) \circ SL_3(q)).f_+.2$	$f_+ = (3, q^2+q+1)$
$N_{H_0}(< S_9 >)$	$((Z_{q^2-q+1}) \circ SU_3(q)).f_-.2$	$f_- = (3, q^2-q+1)$
$N_{H_0}(T_3)$	$(Z_{q^2+q+1})^2.SL_2(3)$	
$N_{H_0}(T_4)$	$(Z_{q^2-q+1})^2.SL_2(3)$	
$N_{H_0}(T_5)$	$(Z_{q^4+q^2+1}).4$	

Conversely, if $L \leq H_0$ is conjugate to one of these groups, then $N_H(L)$ is maximal in H.

□

Theorem 3.2. **1)** *Let $G =^3 D_4(q)$, where q is odd, then*

$$p(G) = (q^2+q+1)(q^4-q^2+1)(q^3+1)/2.$$

2)*Let $G =^3 D_4(q)$, where q is even, then $p(G) = (q^2+q+1)(q^4-q^2+1)(q^3+1)$.*

Theorem 3.3. *Let $G =^3 D_4(q)$, where $q = p^n, p$ is a prime number and n is an integr. then*

1) $r(G) = q(q^4 - q^2 + 1)$

2) $c(G) = q(G) = q^5$

3)$Lim_{q \longrightarrow \infty} \frac{c(G)}{r(G)} = 1.$

References

[1] H. Behravesh, *Quasi-Permutation representations of p-groups of class 2*, J. London Math. Soc. (2) 55 (1997) 251-260.

[2] H. Behravesh, *The rational character table of special linear groups* J. Sci. I.R.I. vol. 9, no. 2 (1998), 173-180.

[3] H. Behravesh,A. Daneshkhah, M.R. Darafsheh and M. Ghorbany, *The rational character table and quasi-permutation representations of the group $PGL(2,q)$*, Italian Journal of Pure and Applied Mathematics N.11, (2001) 9-18.

[4] J.M. Burns, B. Goldsmith, B. Hartley and R. Sandling, *on quasi-permutation representations of finite groups*, Glasgow Math, J. 36, 301-308 (1994).

[5] M.R. Darafsheh, M. Ghorbany, A. Daneshkhah and H. Behravesh, *Quasi-permutation representation of the group $GL(2,q)$*, Journal of Algebra (2001)243,142-167.

[6] M.R. Darafsheh and M.Ghorbany, *Quasi-permutation representations of the groups $SU(3,q^2)$ and $PSU(3,q^2)$*, Southest Asian Bulletin of Mathemetics (2002)26, 395-406.

[7] M.R. Darafsheh and M.Ghorbany, *Quasi-permutation representations of the groups $SL(3,q)$ and $PSL(3,q)$*, Iranian Journal of Science and Technology, Tra.A, Vol. 26, No. A1 (2002)145-154.

[8] M.Ghorbany,*Special representations of the group $G_2(2^n)$ with minimal degrees*, Southest Asian Bulletin of Mathemetics (2006)30: 663-670.

[9] R. Gow, *Schur indices of some groups of Lie type* J. Algebra 42 (1976), 102-120.

[10] F.Himstedt, *Character tables of parabolic subgroups of Steinberg's triality groups ${}^3D_4(2^n)$*, Journal of Algebra, (2007)doi:10.1016/j.jaljebra.2007.01.010, in press.

[11] F.Himstedt, *Character tables of parabolic subgroups of Steinberg's triality groups*, Journal of Algebra, 281(2004) 774-822.
[12] I.M. Isaacs, *Character theory of finite groups*, Academic Press, New York (1976).
[13] B.Peter Kleidman,*The maximal subgroups of the Steinberg Triality groups* $^3D_4(q)$ *and their automorphism groups*, Journal of Algebra 115, 182-199(1988).
[14] W.J. Wong, *Linear groups analogous to permutation groups* , J. Austral. Math. Soc (Sec. A) 3 (1963), 180-184.
[15] W.J. Wong, *On linear* $p-$*groups*, J. Austral. Math. Soc. Ser. A 4 (1964),174-178.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–139, pp. 340–342

Some results on short ρ-exact sequences of S-acts

Roghaieh Khosravi[a,∗], Mohammad Roueentan[b], Elham Rahimi[c]

[a] Department of Mathematics, Faculty of Sciences, Fasa University, Fasa, Iran
[b] College of Engineering, Lamerd Higher Education Center, Shiraz University of Technology, Lamerd, Iran
[c] Department of Mathematics, Shiraz Branch, Islamic Azad University, Shiraz, Iran

Article Info

Keywords:
S-act
Exact Sequence
Congruence

2020 MSC:
20M30

Abstract

In this paper we generalize the notion of (short) Rees exact sequence for S-acts to the notion of (short) exact sequence for S-acts with respect to a congruence. Then, some results concerning commutative diagrams are discussed. Specially, We focus our scope to the short five lemma.

1. Preliminaries and Introduction

Throughout the paper, S and A_S are used to denote a monoid and a right S-act, respectively. A non-empty set A is called a *right S-act*, usually denoted A_S, if S acts on A unitarily from the right; that is, there exists a mapping $A \times S \to A,\ (a, s) \mapsto as$, satisfying the conditions $(as)t = a(st)$ and $a1 = a$, for all $a \in A$ and all $s, t \in S$. A non-empty subset B of a right S-act A_S is called a *subact* of A if $bs \in B$ for all $s \in S$ and $b \in B$. Recall that $\Theta = \{\theta\}$ with the action $\theta s = \theta$ for all $s \in S$ is called *the one element S-act*. An S-act is said to be *simple* if it has no subacts other than itself, and it is called *θ-simple* if it has no subacts other than itself and the one-element subact Θ. Also, products and coproducts of non-empty families of S-acts are the cartesian product and disjoint unions, respectively.
Unlike the case of groups, rings and modules, in the category of S-acts their congruences are not defined by special subacts, and so we have to use congruence for the desired characterizations. Recall from [4, Definition 2.4.18.] that an equivalence relation ρ on an S-act A_S is said to be a *congruence* on A_S if $a\,\rho\,a'$ implies $as\,\rho\,a's$ for any $a, a' \in A_S$ and $s \in S$. The set of all congruences on A_S is denoted by $Con(A_S)$. Any subact B_S of A_S defines the *Rees congruence* ρ_B on A, by setting $a\,\rho_B\,a'$ if $a, a' \in B$ or $a = a'$. For more information on S-acts, we refer the reader to [4].
Let $f : A \longrightarrow B$ be an S-morphism, the kernel of f is defined by

$$\mathcal{K}_f = \ker f := \{(a, a') \in A \times A : f(a) = f(a')\},$$

and $f(A) = \{f(a) :\ a \in A\}$. Note that $f(A)$ is a subact of B_S. Let $\mathcal{I}_f = (f(A) \times f(A)) \cup \Delta_B$, where $\Delta_B = \{(b, b) :\ b \in B\}$. Recall from [3], suppose that A, B, C are S-acts, and $f : A \longrightarrow B$, $g : B \longrightarrow C$ are S-morphisms.

∗Talker
Email addresses: khosravi@fasau.ac.ir (Roghaieh Khosravi), m.rooeintan@yahoo.com (Mohammad Roueentan), rahimie@shirazu.ac.ir (Elham Rahimi)

Then the sequence

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is *Rees exact* at B if $\mathcal{I}_f = \mathcal{K}_g$, and it is called a *Rees short exact sequence* if g is surjective, f is a monomorphism, and $\mathcal{I}_f = \mathcal{K}_g$.

Studying the properties of short exact sequence is a key tool in homological algebra. Preliminary works on Rees short exact sequence of S-acts were done in [2, 3]. Then, in [5], the authors investigated conditions under which a spacial property can be transferred under Rees exact sequences. Also, in [1], the authors introduced the concept of quasi exact sequences as a generalization of Rees exact sequences in the category of S-acts. In this paper we continue the investigation of quasi Rees short exact sequences of S-acts.

2. Main Results

In [1], the authors substitute a congruence ρ on C_S for the trivial congruence Δ_{C_S} in the definition of Rees exact sequence to obtain the notion of ρ-exact sequence in the category of S-acts for a congruence ρ on C, as follows. We recall that $\mathcal{K}_g(\rho) = \{(b, b') \in B \times B : g(b)\, \rho\, g(b')\}$. In fact, $\mathcal{K}_g = \ker g = \mathcal{K}_g(\Delta_{C_S})$.

Definition 2.1. ([1]) Suppose that A, B, C are S-acts, ρ a congruence on C_S, and $f : A \longrightarrow B$, $g : B \longrightarrow C$ are S-morphisms. Then the sequence

$$A \xrightarrow{f} B \xrightarrow{g} C$$

is *ρ-exact (quasi exact)* at B if $\mathcal{I}_f = \mathcal{K}_g(\rho)$, and it is called a *short ρ-exact sequence* if g is surjective, f is a monomorphism, and $\mathcal{I}_f = \mathcal{K}_g(\rho)$.

First, we start with the examples which illustrate how ρ-exact sequences accrue naturally.

Example 2.2. Let A_i, B_i, C_i be S-acts, for $i \in I$, and ρ_i a congruence on C_i. It is easily checked that

$$\rho = \prod_{i \in I} \rho_i = \{(\{a_i\}_{i\in I}, \{b_i\}_{i\in I}) |\ (a_i, b_i) \in \rho_i, \text{ for all } i \in I\}$$

is a congruence on $\prod_{i\in I} C_i$. Suppose that the sequences

$$A_i \xrightarrow{f_i} B_i \xrightarrow{g_i} C_i$$

are ρ_i-exact at B_i for all $i \in I$. Define $\prod_{i\in I} f_i : \prod_{i\in I} A_i \longrightarrow \prod_{i\in I} B_i$ by $\prod_{i\in I} f_i(\{a_i\}_{i\in I}) = \{f_i(a_i)\}_{i\in I}$. Then

$$\prod_{i\in I} A_i \xrightarrow{\prod_{i\in I} f_i} \prod_{i\in I} B_i \xrightarrow{\prod_{i\in I} g_i} \prod_{i\in I} C_i$$

is ρ-exact at $\prod_{i\in I} B_i$.

If ρ is a congruence on C_S, and $h : C \longrightarrow D$ an S-morphism, then it is easily checked that

$$h(\rho) = \{(h(a), h(b)) |\ (a, b) \in \rho\}$$

is a congruence on D_S.

Example 2.3. Let ρ and σ be congruences on S-acts C_S and E_S, respectively. Suppose $A \xrightarrow{f} B \xrightarrow{g} C$ and $C \xrightarrow{h} D \xrightarrow{k} E$ are ρ-exact and σ-exact at B and D, respectively. Then

$$A \xrightarrow{f} B \xrightarrow{hg} D \xrightarrow{k} E$$

is $h(\rho)$-exact and σ-exact at B and D, respectively.

The following result be easily obtained.

Proposition 2.4. *Let $A \xrightarrow{f} B \xrightarrow{g} C$ be a ρ-exact sequence at B_S, and g is an epimorphism. If A_S is generated by a set A' and C_S is generated by a set C', then B_S is generated by $f(A') \cup g^{-1}(C')$.*

Corollary 2.5. *Let $A \xrightarrow{f} B \xrightarrow{g} C$ be a ρ-exact sequence at B_S, and g is an epimorphism. If A_S and C_S are finitely generated, then so is B_S.*

Proposition 2.6. *Let $A \xrightarrow{f} B \xrightarrow{g} C$ be a ρ-exact sequence at B_S and B' be a subact of B such that $B'' = B' \cap f(A) \neq \emptyset$. Take $C' = g(B'),\ A' = f^{-1}(B''),\ f' = f|_{A'},\ g' = g|_{B'}$ and $\sigma = \rho \cap \mathcal{I}_{g'}$. Then $A' \xrightarrow{f'} B' \xrightarrow{g'} C'$ is σ-exact sequence at B'.*

Similar to the proof of [5, Corollary], we could generalize short five lemma for short ρ-exact sequences.

Lemma 2.7. *Let the following diagram of S-morphisms of acts be commutative rows.*

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} \\ L & \xrightarrow{h} & M & \xrightarrow{r} & N \end{array}$$

Suppose that the upper row is short ρ-exact at B_S and the lower row is short σ-exact at M_S, where ρ, σ are congruence of C_S and N_S, respectively. Then the followings hold:

(i) *If α and γ are monomorphisms, then β is a monomorphism.*
(ii) *If α and γ are epimorphisms, then β a is epimorphism.*
(iii) *If α and γ are isomorphisms, then β is a isomorphism.*

Next, we generalize [2, Lemma 3.2] for ρ-exact sequences.

Lemma 2.8. *Suppose that $A_S, B_S, C_S, L_S, M_S, N_S$ are S-acts, and f, g, h, r are S-morphisms. Let the following diagram be commutative:*

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} \\ L & \xrightarrow{h} & M & \xrightarrow{r} & N \end{array}$$

where α, β, γ are isomorphisms. Then the followings hold:

(i) *The upper row is ρ-exact at B_S if and only if the lower row is $\gamma(\rho)$-exact at M_S.*
(ii) *The upper row is a short ρ-exact sequence if and only if the lower row is a short $\gamma(\rho)$-exact sequence.*

Suppose that A, B are S-acts, and $f : A \longrightarrow B$ is an S-morphism. Clearly, $\mathcal{K}_f$ is a subact of $A \prod A$. Now, we generalize Snake Lemma for ρ-exact sequences.

Lemma 2.9. *Suppose that A, B, C, L, M, N are S-acts, and $f, g, h, r,\ \alpha, \beta, \gamma$ are S-morphisms. Let the following diagram be commutative:*

$$\begin{array}{ccccc} A & \xrightarrow{f} & B & \xrightarrow{g} & C \\ \downarrow{\scriptstyle\alpha} & & \downarrow{\scriptstyle\beta} & & \downarrow{\scriptstyle\gamma} \\ L & \xrightarrow{h} & M & \xrightarrow{r} & N \end{array}$$

where the upper and lower rows are ρ-exact at B and $\gamma(\rho)$-exact at M, respectively. Then there exists a sequence $\mathcal{K}_\alpha \longrightarrow \mathcal{K}_\beta \longrightarrow \mathcal{K}_\gamma$ which is δ-exact at $\mathcal{K}_\beta$, where

$$\delta = \rho \times \rho = \{((a_1, b_1), (a_2, b_2))|\ (a_1, b_1), (a_2, b_2) \in \rho\}.$$

References

[1] R. Aminizadeh, H. Rasouli, and A.Tehranian, Quasi-exact Sequences of S-Acts. *Bull. Malays. Math. Sci. Soc.* **42** (2019), 2225–2235.
[2] Y. Chen, Projective S-acts and Exact Functors, *Algebra Colloquium* **7(1)** (2000), 113–120.
[3] Y. Chen and K. P. Shum, Rees short exact sequences of S-systems, *Semigroup Forum* **65** (2002), 141–148.
[4] M. Kilp, U. Knauer, A. Mikhalev, *Monoids, Acts and Categories,* W. de gruyter. Berlin, (2000). https://doi.org/10.4134/CKMS.c210114.
[5] M. Jafari, A. Golchin, and H. Mohammadzadeh Saany, Rees short exact sequence and flatness properties. *Semigroup Forum* **99** (2019), 32–46.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–148, pp. 343–346

The automorphisms of autocentral and autocommutator series terms

Sara Barin[a,*], Mohammad Mehdi Nasrabadi[a]

[a] *Department of Mathematics, University of Birjand, Birjand, Iran*

Article Info

Keywords:
Automorphism
autocentral automorphism
upper autocentral series
lower autocentral series
autosoluble series

2020 MSC:
20D45
20D15

Abstract

In this paper, we define new automorphisms on the autocentral series and the autocommutator series and we identify the relationships of these automorphisms with Aut(G), $Aut_L(G)$, Inn(G), and each other. Also, we present some results that generalize two important theorem.

1. Introduction and preliminaries

The upper and lower autocentral series and the autocommutator series are necessary for important definitions such as autonilpotency and autosolubility of groups. On the other hand, All kinds of automorphisms also have interesting properties. Hence, the automorphisms have been the idea of many researchers articles.
Let G be a group and j be any positive integer. Let us denote by $Z_j(G)$, G', Aut(G) and Inn(G), respectively the j-th term of the upper central series, the commutator subgroup, the full automorphism group and the inner automorphisms. Bachmuth [1] in 1965 defined an IA-automorphism of a group G as

$$IA(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) = [g, \alpha] \in G', \ \forall \, g \in G\}.$$

For any group G, $Inn(G) \leqslant IA(G) \trianglelefteq Aut(G)$.
The investigation of the IA-group has been of interest in different contexts. P. Hall, for example, has shown that the IA-group of a nilpotent group of class c is nilpotent of class c-1 [3]; M. Zyman has remarked that if G is finitely generated nilpotent, so too is IA(G) [7]; and Bonanome et al. have studied the IA-group of a group G for which the upper central series stalls at some point [2]. Bonanome et al. defined the group of j-central automorphism of G, denoted by $Aut_{c_j}(G)$, as the kernel of the natural homomorphism from Aut(G) to $Aut\big(G/Z_j(G)\big)$, i.e.

$$Aut_{c_j}(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in Z_j(G), \ \forall \, g \in G\}.$$

*Talker
Email addresses: s.barin@birjand.ac.ir (Sara Barin), mnasrabadi@birjand.ac.ir (Mohammad Mehdi Nasrabadi)

Hegarty [4] in 1994 introduced the absolute center as

$$L(G) = \{g \in G \mid g^{-1}\alpha(g) = 1, \ \forall\, \alpha \in Aut(G)\}.$$

Also, he defined the autocommutator subgroup as

$$K(G) = \langle [g, \alpha] \mid g \in G, \ \alpha \in Aut(G) \rangle.$$

An automorphism α of G is called an absolute central automorphism if $x^{-1}\alpha(x) \in L(G)$ for each $x \in G$. The set of all absolute central automorphisms of G is denoted by $Aut_L(G)$.
The concept of autonilpotent and autosoluble groups were introduced by Parvaneh and Moghaddam [5] in 2010. They defined the upper autocentral series of G in the following way:

$$\langle 1 \rangle = L_0(G) \subseteq L_1(G) = L(G) \subseteq L_2(G) \subseteq \cdots \subseteq L_n(G) \subseteq \cdots$$

where

$$L_n(G) = \{g \in G \mid [g, \alpha_1, \alpha_2, \ldots, \alpha_n] = 1, \ \forall\, \alpha_1, \alpha_2, \ldots, \alpha_n \in Aut(G)\}$$

and $L_n(G)$ is nth-absolute centre of G. Also, They defined the autocommutator subgroup of weight n+1 in the following way:

$$\begin{aligned} K_n(G) &= [K_{n-1}(G), Aut(G)] \\ &= \langle [g, \alpha_1, \alpha_2, \ldots, \alpha_n] \mid g \in G, \ \alpha_1, \alpha_2, \ldots, \alpha_n \in Aut(G) \rangle, \end{aligned}$$

for all $n \geq 1$, and obtained a the lower autocentral series of G as follows:

$$\cdots \subseteq K_n(G) \subseteq \cdots \subseteq K_2(G) \subseteq K_1(G) = K(G) \subseteq K_0 = G.$$

Also, they called a group to be autonilpotent of class at most c if $K_c(G) = 1$, or $L_c(G) = G$ for some positive integer c. They defined the autosoluble series as a descending series

$$\cdots \subseteq K^{(n)}(G) \subseteq \cdots \subseteq K^{(2)}(G) \subseteq K^{(1)}(G) = K(G) \subseteq K^{(0)}(G) = G.$$

of subgroups of G inductively as follows

$$\begin{aligned} K^{(n)}(G) &= [K^{(n-1)}(G), Aut(G)] \\ &= \langle [g, \alpha] \mid g \in K^{(n-1)}(G), \ \alpha \in Aut\big(K^{(n-1)}(G)\big) \rangle, \end{aligned}$$

for all $n \geq 1$ which is called the nth-autocommutator subgroup of G. Also, they called a group to be autosoluble of length l if $K^{(l)}(G) = 1$ and $K^{(l-1)}(G) \neq 1$, for some positive integer l.
In this paper, we study the automorphisms on the upper and lower autocentral series, the autocommutator series and the relationships of them.

2. Main results

In this section and in each subsection, after some new definitions, we give our main results about the automorphisms on series.

2.1. The automorphisms of autosoluble series

Definition 2.1. The kernel of the natural homomorphism from Aut(G) to $Aut\big(G/K^{(j)}(G)\big)$ is called the group of K^j-automorphism and denoted by $Aut_{K^j}(G)$.
According to the above definition, a K^j-automorphism group acts as the identity on G modulo $K^{(j)}(G)$, Thus:

$$Aut_{K^j}(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in K^{(j)}(G), \ \forall\, g \in G\} \trianglelefteq Aut(G).$$

Also, we have $Aut_{K^1}(G) = Aut(G)$ and $IA(G) \leqslant Aut_{K^j}(G)$, for every j.

Notation 2.2. We use the notation $Aut_{a^j}(G) = Aut_L(G) \cap Aut_{K^j}(G)$ and we refer to $Aut_{a^j}(G)$ as the group of a^j-automorphism. Another definition of $Aut_{a^j}(G)$ is given by

$$Aut_{a^j}(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in L(G) \cap K^{(j)}(G),\ \forall\, g \in G\} \trianglelefteq Aut(G).$$

Proposition 2.3. *For any group G,*

a) $\varphi \in Aut_{K^j}(G)$ *if and only if* $[\alpha, \varphi] \in Aut_{K^j}(G)$*, for all* $\alpha \in Aut(G)$.

b) $Aut_{a^j}(G)$ *is a normal subgroup of* $Aut_L(G)$ *and we have* $\dfrac{Aut_L(G)}{Aut_{a^j}(G)} \cong \dfrac{Aut_L(G)Aut_{K^j}(G)}{Aut_{K^j}(G)}$.

Proof. a) It is obvious by the normality of $Aut_{K^j}(G)$.
b) Let $\psi \in Aut(G)$ and $[g, \alpha] \in K^{(j)}(G)$. Because $K^{(j)}(G)$ is a characteristic subgroup of G, so $\psi([g, \alpha]) \in K^{(j)}(G)$. Now, let $\sigma \in Aut_L(G)$ and $\beta \in Aut_{a^j}(G)$. We will show that $\sigma^{-1}\beta\sigma \in Aut_{a^j}(G)$. For every $g \in G$, we have

$$g^{-1}\big(\sigma^{-1}\beta\sigma\big)(g) = g^{-1}\sigma^{-1}\Big(\sigma(g)\big(\sigma(g)\big)^{-1}\beta\big(\sigma(g)\big)\Big) = g^{-1}g\sigma^{-1}\Big(\underbrace{\big(\sigma(g)\big)^{-1}\beta\big(\sigma(g)\big)}_{\in\, L(G)\cap K^{(j)}(G)}\Big).$$

Because the intersection of two characteristic subgroups is a characteristic subgroup, the first part is proved.
For the second part, the result follows by the definition of $Aut_{a^j}(G)$ and the third isomorphism theorem. □

Corollary 2.4. *For any group G,* $[Aut(G), Aut_{K^j}(G)] \leqslant Aut_{K^j}(G)$.

Theorem 2.5. *Let G be a group. If* $Aut_L\big(G/K^{(j)}(G)\big) = Inn\big(G/K^{(j)}(G)\big)$*, then*

$$Aut_L(G) \leqslant Inn(G)Aut_{K^j}(G).$$

Proof. Let $\alpha \in Aut_L(G)$. By hypothesis, there exists $g \in G$ such that for all $x \in G$, $\alpha(x)K^{(j)}(G) = x^gK^{(j)}(G)$. Hence,

$$x^{-g}\alpha(x) = \Big(x^{-1}\big(\alpha(x)\big)^{g^{-1}}\Big)^g \in K^{(j)}(G).$$

Thus,

$$x^{-1}\big(\alpha(x)\big)^{g^{-1}} = x^{-1}g\big(\alpha(x)\big)g^{-1} = x^{-1}\varphi_g^{-1}\alpha(x) \in K^{(j)}(G)$$

where $\varphi_g \in Inn(G)$.
Consequently, $\varphi_g^{-1}\alpha \in Aut_{K^j}(G)$, i.e., $\alpha = \varphi_g\varphi_g^{-1}\alpha \in Inn(G)Aut_{K^j}(G)$. □

2.2. The automorphisms of lower autocentral series

Definition 2.6. The kernel of the natural homomorphism from Aut(G) to $Aut\big(G/K_j(G)\big)$ is called the group of K_j-automorphism and denoted by $Aut_{K_j}(G)$. In other words,

$$Aut_{K_j}(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in K_j(G),\ \forall\, g \in G\} \trianglelefteq Aut(G).$$

Also, $Aut_{K_1}(G) = Aut(G)$ and $IA(G) \leqslant Aut_{K_j}(G) \leqslant Aut_{K^j}(G)$, for every j.

Notation 2.7. We use the notation

$$\begin{aligned} Aut_{a_j}(G) &= Aut_L(G) \cap Aut_{K_j}(G) \\ &= \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in L(G) \cap K_j(G),\ \forall\, g \in G\} \trianglelefteq Aut(G) \end{aligned}$$

and we refer to $Aut_{a_j}(G)$ as the group of a_j-automorphism.

In the following, the results are similar to the results of the subsection 2.1, so we do not prove them.

Proposition 2.8. *For any group G,*

a) $\varphi \in Aut_{K_j}(G)$ *if and only if* $[\alpha, \varphi] \in Aut_{K_j}(G)$*, for all* $\alpha \in Aut(G)$.

b) $Aut_{a_j}(G)$ *is a normal subgroup of* $Aut_L(G)$ *and* $\dfrac{Aut_L(G)}{Aut_{a_j}(G)} \cong \dfrac{Aut_L(G)Aut_{K_j}(G)}{Aut_{K_j}(G)}$.

Corollary 2.9. *For any group G,* $[Aut(G), Aut_{K_j}(G)] \leqslant Aut_{K_j}(G)$.

Theorem 2.10. *Let G be a group. If* $Aut_L\big(G/K_j(G)\big) = Inn\big(G/K_j(G)\big)$*, then*

$$Aut_L(G) \leqslant Inn(G)Aut_{K_j}(G).$$

2.3. The automorphisms of upper autocentral series

Definition 2.11. The kernel of the natural homomorphism from Aut(G) to $Aut\big(G/L_j(G)\big)$ is called the group of L_j-automorphism and denoted by

$$Aut_{L_j}(G) = \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in L_j(G),\ \forall\, g \in G\} \trianglelefteq Aut(G).$$

Also, we have $Aut_{L_1}(G) = Aut_L(G)$ and $Aut_{L_j}(G) \leqslant Aut_{c_j}(G)$, for every j.

Notation 2.12. We use the notation

a) $Aut_{\underline{a}_j}(G) = Aut_{L_j}(G) \cap Aut_{K_j}(G)$
$= \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in L_j(G) \cap K_j(G),\ \forall\, g \in G\} \trianglelefteq Aut(G).$

b) $Aut_{\overline{a}_j}(G) = Aut_{L_j}(G) \cap Aut_{K^j}(G)$
$= \{\alpha \in Aut(G) \mid g^{-1}\alpha(g) \in L_j(G) \cap K^{(j)}(G),\ \forall\, g \in G\} \trianglelefteq Aut(G).$

Proposition 2.13. *For any group G,*

a) $\varphi \in Aut_{L_j}(G)$ *if and only if* $[\alpha, \varphi] \in Aut_{L_j}(G)$*, for all* $\alpha \in Aut(G)$.

b) $\dfrac{Aut_{K_j}(G)}{Aut_{\underline{a}_j}(G)} \cong \dfrac{Aut_{K_j}(G)Aut_{L_j}(G)}{Aut_{L_j}(G)}.$

c) $\dfrac{Aut_{K^j}(G)}{Aut_{\overline{a}_j}(G)} \cong \dfrac{Aut_{K^j}(G)Aut_{L_j}(G)}{Aut_{L_j}(G)}.$

Corollary 2.14. *For any group G,* $[Aut(G), Aut_{L_j}(G)] \leqslant Aut_{L_j}(G)$.

Theorem 2.15. *Let G be a group.*
a) If $Aut_{K_j}\big(G/L_j(G)\big) = Inn\big(G/L_j(G)\big)$*, then* $Aut_{K_j}(G) \leqslant Inn(G)Aut_{L_j}(G)$.
b) If $Aut_{K^j}\big(G/L_j(G)\big) = Inn\big(G/L_j(G)\big)$*, then* $Aut_{K^j}(G) \leqslant Inn(G)Aut_{L_j}(G)$.

2.4. Some generalization

In introduction, we stated two theorem from Hall and Zyman. In this subsection, we generalize them as follows.

Theorem 2.16 (Generalization of Hall's theorem)**.** *The IA-group of a autonilpotent(or autosoluble) group of class c is solvable and its derived length is at most* $[log_2\ c] + 1$.

Proof. This idea is motivated by the fact that nilpotent groups are solvable. Now the result follow by [6, 5.1.2]. □

Similary, we have

Theorem 2.17 (Generalization of Zyman's theorem)**.** *If G is finitely generated autonilpotent or autosoluble, then IA(G) is finitely generated solvable.*

References

[1] S. Bachmuth, *Automorphisms of free metabelian groups*, Trans. Amer. Math. Soc, 118 (1965), 93-104.
[2] M. Bonanome, M. H. Dean and M. Zyman, *IA-automorphisms of groups with almost constant upper central series*, Contemporary Mathematica, 582 (2011).
[3] P. Hall, *The Edmonton notes on nilpotent groups*, Queen Mary College Mathematics Notes (1969).
[4] P. V. Hegarty, *The absolute centre of a group*, J. Algebra, 169 (1994), 929-935.
[5] F. Parvaneh and M. R. R. Moghaddam, *Some properties of autosoluble groups*, J. Math. Ext, 5(1) (2010), 13-19.
[6] D. J. Robinson, *A Course in the Theory of Groups*, Springer-verlag, New York/Berlin (1982).
[7] M. Zyman, *IA-automorphisms and localization of nilpotent groups*, Ph.D. Thesis-City University of New York (2007).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–149, pp. 347–348

On hypercyclic monoids

Mohammad Roueentan[a,*], Roghaieh Khosravi[b], Elham Rahimi[c]

[a] *College of Engineering, Lamerd Higher Education Center, Lamerd, Iran*
[b] *Department of Mathematics, Faculty of Sciences, Fasa University, Fasa, Iran*
[c] *Department of Mathematics, Shiraz Branch, Islamic Azad University, Shiraz, Iran*

Article Info

Keywords:
act
hypercyclic
monoid

2020 MSC:
20M30

Abstract

The main purpose of this paper is to introduce and study the notion of hypercyclic monoids as monoids which for them any cyclic right act has cyclic injective envelope. We investigate some properties of some classes of acts over such monoids.

1. Introduction

In [7] the authores have investigated the monoids for which any cyclic right act is injective and also in [8] in more limited conditions these classes of monoids have been studied. In this article we introduce and study the concept of hypercyclic monoids (monoids for which any cyclic right act has cyclic injective envelope). It is shown that over a commutative hypercyclic monoids any cyclic act is quasi-injective and contains a zero element. Also it is proved that for a hypercyclic monoid S any non-zero right ideal I of S is essential in S if and only if $S_S = E(I)$.
Throughout this article S will denote a monoid and an S-act A_S (or A) is a right S-act. A subact B of an act A is called $large$ (or $essential$) in A denoted by $B \subseteq' A$, if any S-homomorphism $g : A \longrightarrow C$ such that $g|_B$ is a monomorphism is itself a monomorphism (see [1]). From [5] an S-act A is said to be $uniform$, provided that every non-zero subact of A is essential. For an S-act A, by $E(A)$, we mean the injective envelope of A. Recall that an S-act A is injective if for any right S-act N, any subact M of N, and any homomorphism $f \in Hom(M, A)$, there exists a homomorphism $g \in Hom(N, A)$ which extends f, i.e., $g|_M = f$. Also regarding the following diagram an S-act A is called:
(i) *C-injective* if M is a cyclic subact of N ([7]).
(ii) *CC-injective* if M and N are cyclic acts ([8]).
(iii) *quasi-injective* if $N = A$ ([6]).

*Talker
Email addresses: m.rooeintan@yahoo.com (Mohammad Roueentan), khosravi@fasau.ac.ir (Roghaieh Khosravi), rahimie@shirazu.ac.ir (Elham Rahimi)

$$\begin{array}{ccc} M & \subseteq & N \\ f\downarrow & \swarrow \exists\, g & \\ A. & & \end{array}$$

We encourage the reader to see [1] for basic results and definitions relating to acts over monoids.

2. Main Results

Definition 2.1. A monoid S is said to be $hypercyclic$ if any cyclic right S-act has cyclic injective envelope.

Let A, B be right S-acts. From [1] the *trace* of B in A is defined by $tr(B, A) = \bigcup_{f \in Hom(B,A)} f(B)$. Recall that an S-act A is *strongly duo* if for every subact B of A, $tr(B, A) = B$ ([3]). Also from [4] an S-act A is siad to be $coregular$ if any cyclic subact of A is injectice.

Proposition 2.2. *Suppose S is a commutative hypercyclic monoid. Then*

(i) *Every cyclic S-act is quasi-injective $(strongly\ duo)$ and contains a zero element.*

(ii) *Eny strongly faithful S-act is coregular.*

(iii) *The right S-act S_S is injective.*

Theorem 2.3. *Let S be a commutative hypercyclic monoid. Then S_S is uniform if and only if for any non-trivial right ideal I of S, $E(I) = S_S$.*

Proposition 2.4. *Suppose S is a hypercyclic monoid. Then the following conditions are equivalent:*

(i) *Every cyclic S-act is injective.*

(ii) *Eny finite product of cyclic S-acts are quasi-injective.*

Proposition 2.5. *Let S be a hypercyclic monoid and A be a uniform S-act which contains a non-zero element. Then $E(A)$ is a cyclic S-act.*

Theorem 2.6. *A monoid S is hypercyclic and every S-act is CC-injective if and only if all S-acts are C-injective.*

Corollary 2.7. *Suppose S is a monoid with central idempotents. Then S is a regular hypercylic monoid if and only if any S-act is C-injective.*

From [2] a monoid S is called $Rees\ artinian$ if S satisfies the descending chain condition on right ideals.

Proposition 2.8. *Let S be a hypercylc Rees artinian monoid. Then the right S-act S_S is injective and for any right ideal I of S, $E(I) = eS$ for some idempotent $e \in S$.*

Proof. Suppose $E(S_S) = aS$ is a cyclic S-act. Thus by projectivity of S there exists a homomorphism $f : S_S \longrightarrow S_S$ such that $\lambda f = i$ where $\lambda : S_S \longrightarrow aS$ is defined by $\lambda(s) = as$ for any $s \in S$ and i is the inlusion map $i : S_S \longrightarrow E(S_S)$. Hence f is a monomorphism. Now since S satisfies the descending chain condition on right ideals we can see that f is an epimorphism and so i is an isomorphism which implies the result. For the second part note that by injectivity of S_S, for any right ideal I of S, $E(I)$ is a retract of S and consequently is indecomposable and projective. Thus we imply the result. □

As a direct concequence of the previouse proposition if S is a finite hypercylci monoid, the S_S is injective.

References

[1] Kilp, M., Knauer, U., Mikhalev, A. V.: Monoids, Acts and Categories, With Application to Wreath Product, Berlin; New York(2000).

[2] Khosravi, R., Roueentan, M. Chain conditions on (Rees) congruences of S-acts, Journal of Algebra and Its Applications, DoI: 10.1142/S0219498824500853.

[3] Roueentan, M., Ershad, M., Strongly duo and duo right S-acts, Italian Journal of Pure and Applied mathematics, 32, 143-154 (2014).

[4] Roueentan, M., Khosravi, R., On coregular S-acts, the 52^{th} Annual Iranian Mathematics Conference, (2021.

[5] Roueentan, M., Sedaghatjoo, M., On uniform acts over semigroups, Semigroup Forum, 97, 229-243 (2018).

[6] Satyanarayana, M., Quasi and weakly injective S-systems. Math. Nachr, 71(1), 183–190 (1973).

[7] Zhang, X., Knauer, U., Chen, Y., Classification of monoids by injectivities I. C-injectivity, Semigroup Forum, 76, 169-176 (2008).

[8] Zhang, X., Knauer, U., Chen, Y., Classification of monoids by injectivities II. CC-injectivity, Semigroup Forum, 76 ,177-184 (2008).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–150, pp. 349–352

Some results on the vanishing and coassociated prime ideals of the top generalized local cohomology module

Ali Fathi*

Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.

Article Info

Keywords:
Generalized local cohomology
Lichtenbaum-Hartshorne vanishing theorem
Coassociated prime ideal

2020 MSC:
13D45
13E05
13E10

Abstract

Let R be a commutative Noetherian ring, and let $\mathfrak{a}$ be a proper ideal of R. Let M be a non-zero finitely generated R-module with the finite projective dimension p, and let N be a non-zero finitely generated R-module with $N \neq \mathfrak{a}N$. Assume that c is the greatest non-negative integer with the property that $\mathrm{H}^i_{\mathfrak{a}}(N)$, the i-th local cohomology module of N with respect to $\mathfrak{a}$, is non-zero. It is known that $\mathrm{H}^i_{\mathfrak{a}}(M,N)$, the i-th generalized local cohomology module of M and N with respect to $\mathfrak{a}$, is zero for all $i > p+c$. In this paper, we obtain the coassociated prime ideals of $\mathrm{H}^{p+c}_{\mathfrak{a}}(M,N)$. Using this, in the case when R is a local ring and c is equal to the dimension of N, we give a necessary and sufficient condition for the vanishing of $\mathrm{H}^{p+c}_{\mathfrak{a}}(M,N)$ which extends the Lichtenbaum-Hartshorne vanishing theorem for generalized local cohomology modules.

1. Introduction

Throughout this paper, let R be a commutative Noetherian ring with non-zero identity. Let $\mathfrak{a}$ be an ideal of R and N be an R-module. The i-th local cohomology module of N with respect to $\mathfrak{a}$ was defined by Grothendieck as follows:

$$\mathrm{H}^i_{\mathfrak{a}}(N) := \varinjlim_{n\in\mathbb{N}} \mathrm{Ext}^i_R(R/\mathfrak{a}^n, N);$$

see [2] for more details. For a pair of R-modules (M,N), the i-th generalized local cohomology module of (M,N) with respect to $\mathfrak{a}$ was introduced by Herzog as follows:

$$\mathrm{H}^i_{\mathfrak{a}}(M,N) := \varinjlim_{n\in\mathbb{N}} \mathrm{Ext}^i_R(M/\mathfrak{a}^n M, N);$$

see [1] for more details. It is clear that $\mathrm{H}^i_{\mathfrak{a}}(R,N) = \mathrm{H}^i_{\mathfrak{a}}(N)$. The cohomological dimension of N with respect to $\mathfrak{a}$ and the cohomological dimension of (M,N) with respect to $\mathfrak{a}$ are defined, respectively, as follow:

$$\mathrm{cd}_{\mathfrak{a}}(N) := \sup\{i \in \mathbb{N}_0 : \mathrm{H}^i_{\mathfrak{a}}(N) \neq 0\}$$

*Talker
Email address: alif1387@gmail.com (Ali Fathi)

and
$$\operatorname{cd}_{\mathfrak{a}}(M,N) := \sup\{i \in \mathbb{N}_0 : \mathrm{H}^i_{\mathfrak{a}}(M,N) \neq 0\}.$$
Assume that N is finitely generated with finite dimension d. By [2, Theorem 6.1.2 and Exercise 7.1.7], $\operatorname{cd}_{\mathfrak{a}}(N) \leq d$ and $\mathrm{H}^d_{\mathfrak{a}}(N)$ is Artinian. When R is local, Dibaei and Yassemi proved in [3, Theorem A] that

$$\operatorname{Att}_R(\mathrm{H}^d_{\mathfrak{a}}(N)) = \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \operatorname{cd}_{\mathfrak{a}}(R/\mathfrak{p}) = d\}.$$

This equality also holds without the hypothesis that R is local (see [4, Theorem 2.5]). If M is finitely generated with finite projective dimension p, then $\operatorname{cd}_{\mathfrak{a}}(M,N) \leq p+d$ and $\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)$ is Artinian (see [1, Lemma 5.1] and [11, Theorem 2.9]). When R is local, as a generalization of the theorem of Dibaei and Yassemi, Gu and Chu show in [8, Theorem 2.3] that
$$\operatorname{Att}_R(\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)) = \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \operatorname{cd}_{\mathfrak{a}}(M,R/\mathfrak{p}) = p+d\}.$$
In [7, Theorem 5.3], the author of the present paper, Tehranian and Zakeri proved this equality in the case when R is not necessarily local. Also, it is shown in [7, Theorem 5.6] that

$$\operatorname{Att}_R(\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)) = \operatorname{Supp}_R(\operatorname{Ext}^p_R(M,R)) \cap \operatorname{Att}_R(\mathrm{H}^d_{\mathfrak{a}}(N)) \tag{†}$$

whenever $R/\operatorname{Ann}_R(\mathrm{H}^d_{\mathfrak{a}}(N))$ is a complete semilocal ring. This equality allows us to compute the set of attached prime ideals of the top generalized local cohomology module $\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)$ from the set of attached prime ideals of the top local cohomology module $\mathrm{H}^d_{\mathfrak{a}}(N)$.
Now we assume that $N \neq \mathfrak{a}N$ and set $c := \operatorname{cd}_{\mathfrak{a}}(N)$. For all $i > p+c$, $\mathrm{H}^i_{\mathfrak{a}}(M,N) = 0$; see [9, Proposition 2.8]. Since $c \leq d$, $p+c$ yields a sharper upper bound for $\operatorname{cd}_{\mathfrak{a}}(M,N)$. Note that $\mathrm{H}^c_{\mathfrak{a}}(N)$ and $\mathrm{H}^{p+c}_{\mathfrak{a}}(M,N)$ are not necessarily Artinian. In Theorem 3.1, using the set of coassociated prime ideals of $\mathrm{H}^c_{\mathfrak{a}}(N)$, we compute the set of coassociated prime ideals of $\mathrm{H}^{p+c}_{\mathfrak{a}}(M,N)$. More precisely, we show that

$$\operatorname{Coass}_R\left(\mathrm{H}^{p+c}_{\mathfrak{a}}(M,N)\right) = \{\mathfrak{p} \in \operatorname{Supp}_R(M) \cap \operatorname{Coass}_R\left(\mathrm{H}^c_{\mathfrak{a}}(N)\right) : \operatorname{proj\,dim}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = p\}.$$

As a consequence of this equality, we prove in Corollary 3.2 that the equality (†) holds even if $R/\operatorname{Ann}_R(\mathrm{H}^d_{\mathfrak{a}}(N))$ is not a complete semilocal ring, and we show that

$$\operatorname{Att}_R\left(\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)\right) = \{\mathfrak{p} \in \operatorname{Supp}_R(M) \cap \operatorname{Ass}_R(N) : \operatorname{proj\,dim}_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = p, \operatorname{cd}_{\mathfrak{a}}(R/\mathfrak{p}) = d\}. \tag{‡}$$

In particular, if R is local and M is Cohen-Macaulay, then it is shown in Corollary 3.3 that

$$\operatorname{Att}_R(\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)) = \{\mathfrak{p} \in \operatorname{Supp}_R(M) \cap \operatorname{Ass}_R(N) : \operatorname{cd}_{\mathfrak{a}}(R/\mathfrak{p}) = d\}.$$

Finally, we will extend the Lichtenbaum-Hartshorne vanishing theorem for generalized local cohomology modules. More precisely, when R is a local ring, we prove in Theorem 3.4 that $\mathrm{H}^{p+d}_{\mathfrak{a}}(M,N)$ is zero if and only if for all $\mathfrak{P} \in \operatorname{Supp}_{\widehat{R}}(\widehat{M}) \cap \operatorname{Ass}_{\widehat{R}}(\widehat{N})$ with $\dim_{\widehat{R}}(\widehat{R}/\mathfrak{P}) = d$ and $\operatorname{proj\,dim}_{\widehat{R}_{\mathfrak{P}}}(\widehat{M}_{\mathfrak{P}}) = p$, $\dim_{\widehat{R}}(\widehat{R}/(\mathfrak{a}\widehat{R}+\mathfrak{P})) > 0$.

2. Preliminaries

Let M be an R-module. We denote the localization of M at $\mathfrak{p}$ by $M_{\mathfrak{p}}$, and the set of all prime ideals $\mathfrak{p}$ of R such that $M_{\mathfrak{p}}$ is non-zero is called the support of M and denoted by $\operatorname{Supp}_R(M)$. The annihilator of M in R, denoted by $\operatorname{Ann}_R(M)$, is defined to be the set $\{r \in R : rx = 0 \text{ for all } x \in M\}$. If $\mathfrak{p} := \operatorname{Ann}_R(Rx)$ is a prime ideal of R for some $x \in M$, then $\mathfrak{p}$ is called an associated prime ideal of M, and we denote the set of all associated prime ideals of M by $\operatorname{Ass}_R(M)$. We will denote the set of all positive integers (respectively, non-negative integers) by $\mathbb{N}$ (respectively, $\mathbb{N}_0$).
The concepts of attached prime ideal and secondary representation as the duals of the concepts of associated prime ideal and primary decomposition were introduced by Macdonald in [10]. An R-module M is said to be secondary if $M \neq 0$ and, for each $r \in R$, the endomorphism $\mu_r : M \to M$ defined by $\mu_r(x) = rx$ (for $x \in M$) is either surjective or nilpotent. If M is secondary, then $\mathfrak{p} := \sqrt{\operatorname{Ann}_R(M)}$ is a prime ideal and we say that M is $\mathfrak{p}$-secondary. A prime ideal $\mathfrak{p}$ is called an attached prime ideal of M if M has a $\mathfrak{p}$-secondary quotient. We denote the set of all attached prime

ideals of M by $\mathrm{Att}_R(M)$. If M can be written as a finite sum of its secondary submodules, then we say that M has a secondary representation. Such a secondary representation

$$M = M_1 + \cdots + M_t \quad \text{with } \mathfrak{p}_i := \sqrt{\mathrm{Ann}_R(M_i)} \text{ for } i = 1, \ldots, t$$

of M is said to be minimal when none of the modules M_i $(1 \leq i \leq t)$ is redundant (that is, $M_i \not\subseteq M_1 + \cdots + M_{i-1} + M_{i+1} + \cdots + M_t$) and the prime ideals $\mathfrak{p}_1, \ldots, \mathfrak{p}_t$ are distinct. Since the sum of two $\mathfrak{p}$-secondary submodules of M is again $\mathfrak{p}$-secondary, so if M has a secondary representation, then it has a minimal one. When the above secondary representation is minimal, then $\mathrm{Att}_R(M) = \{\mathfrak{p}_1, \ldots, \mathfrak{p}_t\}$, and hence t and the set $\{\mathfrak{p}_1, \ldots, \mathfrak{p}_t\}$ are independent of the choice of minimal secondary representation of M. Artinian modules have secondary representation. We refer the readers to [10] for more details.
Yassemi [12] has introduced the coassociated prime ideal as a dual of associated prime ideal. In Yassemi's definition, we do not need to assume that the module has a secondary representation, and note that if a module has a secondary representation, then its sets of coassociated prime ideals and attached prime ideals are same (see [12, Theorem 1.14]).

Definition 2.1. We say that an R-module M is cocyclic when M is a submodule of $E(R/\mathfrak{m})$ for some maximal ideal $\mathfrak{m}$ of R, where $E(R/\mathfrak{m})$ denotes the injective envelope of $R/\mathfrak{m}$.

Definition 2.2. We say that a prime ideal $\mathfrak{p}$ of R is a coassociated prime ideal of an R-module M when there exists a cocyclic homomorphic image L of M such that $\mathfrak{p} = \mathrm{Ann}_R(L)$. We denote by $\mathrm{Coass}_R(M)$ the set of all coassociated prime ideals of M.

3. Main results

In the following theorem, using the set of coassociated prime ideals of the top local cohomology module, we compute the set of coassociated prime ideals of the top generalized local cohomology module.

Theorem 3.1 ([6])**.** *Let $\mathfrak{a}$ be an ideal of R and M be a non-zero finitely generated R-module with finite projective dimension p. Let N be an R-module such that $N \neq \mathfrak{a}N$ and $c := \mathrm{cd}_{\mathfrak{a}}(N)$. Then*

$$\mathrm{Coass}_R\left(\mathrm{H}_{\mathfrak{a}}^{p+c}(M,N)\right) = \{\mathfrak{p} \in \mathrm{Supp}_R(M) \cap \mathrm{Coass}_R\left(\mathrm{H}_{\mathfrak{a}}^{c}(N)\right) : \operatorname{proj}\dim_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = p\}.$$

Now, let the notations and assumptions be as in Corollary 3.2. The author of the present paper, Tehranian and Zakeri, in [7, Theorem 5.6], proved that

$$\mathrm{Att}_R(\mathrm{H}_{\mathfrak{a}}^{p+d}(M,N)) = \mathrm{Supp}_R(\mathrm{Ext}_R^p(M,R)) \cap \mathrm{Att}_R(\mathrm{H}_{\mathfrak{a}}^{d}(N))$$

whenever $B := R/\operatorname{Ann}_R(\mathrm{H}_{\mathfrak{a}}^{d}(N))$ is a complete semilocal ring. In the following corollary it is shown that the above equality holds without the hypothesis that B is a complete semilocal ring.

Corollary 3.2 ([6])**.** *Let $\mathfrak{a}$ be an ideal of R, and M, N be two non-zero finitely generated R-modules such that $p := \operatorname{proj}\dim_R(M) < \infty$ and $d := \dim_R(N) < \infty$. Then $\mathrm{H}_{\mathfrak{a}}^{p+d}(M,N)$ is Artinian and*

$$\mathrm{Att}_R(\mathrm{H}_{\mathfrak{a}}^{p+d}(M,N)) = \{\mathfrak{p} \in \mathrm{Supp}_R(M) \cap \mathrm{Ass}_R(N) : \operatorname{proj}\dim_{R_{\mathfrak{p}}}(M_{\mathfrak{p}}) = p, \mathrm{cd}_{\mathfrak{a}}(R/\mathfrak{p}) = d\}.$$

Corollary 3.3 ([6])**.** *Let R be a local ring and $\mathfrak{a}$ be an ideal of R. Let M and N be two non-zero finitely generated R-modules such that M is Cohen-Macaulay, $p := \operatorname{proj}\dim_R(M) < \infty$ and $d := \dim_R(N)$. Then we have*

$$\mathrm{Att}_R\,\mathrm{H}_{\mathfrak{a}}^{p+d}(M,N) = \{\mathfrak{p} \in \mathrm{Supp}_R(M) \cap \mathrm{Ass}_R(N) : \mathrm{cd}_{\mathfrak{a}}(R/\mathfrak{p}) = d\}.$$

Theorem 3.4 (The Lichtenbaum-Hartshorne vanishing theorem for generalized local cohomology modules [6])**.** *Let $(R,\mathfrak{m})$ be a local ring and $\mathfrak{a}$ be a proper ideal of R. Let M and N be two non-zero finitely generated R-modules such that $p := \operatorname{proj}\dim_R(M) < \infty$ and $d := \dim_R(N)$. Then the following statements are equivalent:*

(i) $\mathrm{H}_{\mathfrak{a}}^{p+d}(M,N) = 0$;

(ii) *for each* $\mathfrak{P} \in \operatorname{Supp}_{\widehat{R}}(\widehat{M}) \cap \operatorname{Ass}_{\widehat{R}}(\widehat{N})$ *satisfying* $\operatorname{proj\,dim}_{\widehat{R}_{\mathfrak{P}}}(\widehat{M}_{\mathfrak{P}}) = p$ *and* $\dim_{\widehat{R}}(\widehat{R}/\mathfrak{P}) = d$, *we have* $\dim_{\widehat{R}}(\widehat{R}/(\mathfrak{a}\widehat{R} + \mathfrak{P})) > 0$.

Remark 3.5. Let $(R, \mathfrak{m})$ be a local ring. Let M and N be two non-zero finitely generated R-modules such that $p := \operatorname{proj\,dim}_R(M) < \infty$ and $d := \dim_R(N)$. By Grothendieck's vanishing and non-vanishing theorems [2, Theorems 6.1.2 and 6.1.4], we have $\operatorname{cd}_{\mathfrak{m}}(N) = \dim_R(N)$. The exact value of $\operatorname{cd}_{\mathfrak{m}}(M, N)$ is unknown under the above assumptions. However, if in addition R is Cohen-Macaulay, then Divaani-Aazar and Hajikarimi in [5, Theorem 3.5] proved that

$$\operatorname{cd}_{\mathfrak{m}}(M, N) = \dim_R(R) - \operatorname{grade}_R(\operatorname{Ann}_R(N), M).$$

We know that $p + d$ is an upper bound for $\operatorname{cd}_{\mathfrak{m}}(M, N)$. If we set $\mathfrak{a} := \mathfrak{m}$ in Theorem 3.4, then it is not true to say that since $\dim_{\widehat{R}}(\widehat{R}/\mathfrak{m}\widehat{R} + \mathfrak{P}) = 0$ for all prime ideals $\mathfrak{P}$ of $\widehat{R}$, $\mathrm{H}^{p+d}_{\mathfrak{m}}(M, N)$ is non-zero and so $\operatorname{cd}_{\mathfrak{m}}(M, N) = p + d$. The following example shows that $p + d$ can be a strict upper bound for $\operatorname{cd}_{\mathfrak{m}}(M, N)$. In fact, if there does not exist a prime ideal $\mathfrak{P}$ in $\operatorname{Supp}_{\widehat{R}}(\widehat{M}) \cap \operatorname{Ass}_{\widehat{R}}(\widehat{N})$ satisfying $\dim_{\widehat{R}}(\widehat{R}/\mathfrak{P}) = d$ and $\operatorname{proj\,dim}_{\widehat{R}_{\mathfrak{P}}}(\widehat{M}_{\mathfrak{P}}) = p$, then the statement (ii) in Theorem 3.4 is true and hence $\mathrm{H}^{p+d}_{\mathfrak{m}}(M, N) = 0$.

Example 3.6. Let K be a field and $R := K[[x, y]]$ be the ring of formal power series over K in indeterminates x, y. Then R is a complete regular local ring of dimension 2 with maximal ideal $\mathfrak{m} := (x, y)$. We set $M := R/(x^2, xy)$. It follows from $\operatorname{Ass}_R(M) = \{(x), (x, y)\}$ that $\operatorname{depth}_R(M) = 0$ and $\dim_R(M) = \dim_R(R/(x)) = 1$. Since R is regular, M has finite projective dimension and so the Auslander-Buchsbaum formula implies that $\operatorname{proj\,dim}_R(M) = 2$. Therefore $\operatorname{proj\,dim}_R(M) + \dim_R(R) = 4$. Now since $\operatorname{Ass}_R(R) = \{0\}$, $\operatorname{Supp}_R(M) \cap \operatorname{Ass}_R(R) = \emptyset$ and so, by Theorem 3.4 or Corollary 3.2, we obtain $\mathrm{H}^4_{\mathfrak{m}}(M, R) = 0$. Hence

$$\operatorname{cd}_{\mathfrak{m}}(M, R) < \operatorname{proj\,dim}_R(M) + \dim_R(R).$$

Furthermore, since R is Cohen-Macaulay and M has finite projective dimension, the Divaani-Azar–Hajikarimi formula implies that

$$\operatorname{cd}_{\mathfrak{m}}(M, R) = \dim_R(R) - \operatorname{grade}_R(\operatorname{Ann}_R(R), M) = 2 - 0 = 2.$$

References

[1] M. H. Bijan-Zadeh. A common generalization of local cohomology theories. *Glasgow Math. J.*, 21(2):173–181, 1980.

[2] M. P. Brodmann and R. Y. Sharp. *Local Cohomology: An Algebraic Introduction with Geometric Applications*, volume 60. Cambridge University Press, Cambridge, 1998.

[3] M. T. Dibaei and S. Yassemi. Attached primes of the top local cohomology modules with respect to an ideal. *Arch. Math. (Basel)*, 84(4): 292–297, 2005.

[4] K. Divaani-Aazar. Vanishing of the top local cohomology modules over noetherian rings. *Proc. Indian Acad. Sci. Math. Sci.*, 119(1):23–35, 2009.

[5] K. Divaani-Aazar and A. Hajikarimi. Generalized local cohomology modules and homological gorenstein dimensions. *Comm. Algebra*, 39 (6):2051–2067, 2011.

[6] A. Fathi. Lichtenbaum–hartshorne vanishing theorem for generalized local cohomology modules. *J. Adv. Math. Model.*, 13(2):250–258, 2023.

[7] A. Fathi, A. Tehranian, and H. Zakeri. Filter regular sequences and generalized local cohomology modules. *Bull. Malays. Math. Sci. Soc.*, 38 (2):467–482, 2015.

[8] Y. Gu and L. Chu. Attached primes of the top generalized local cohomology modules. *Bull. Aust. Math. Soc.*, 79(1):59–67, 2009.

[9] S. H. Hassanzadeh and A. Vahidi. On vanishing and cofiniteness of generalized local cohomology modules. *Comm. Algebra*, 37(7):2290–2299, 2009.

[10] I. G. Macdonald. Secondary representation of modules over a commutative ring. *Symp. Math.*, 11:23–43, 1973.

[11] A. Mafi. On the associated primes of generalized local cohomology modules. *Comm. Algebra*, 34(7):2489–2494, 2006.

[12] S. Yassemi. Coassociated primes. *Comm. Algebra*, 23(4):1473–1498, 1995.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–155, pp. 353–356

On the annihilator of first non-zero local cohomology module

Ali Fathi*

Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.

Article Info

Keywords:
Local cohomology
Annihilator
First non-zero local cohomology

2020 MSC:
13D45
13E05

Abstract

Let R be a commutative Noetherian ring, $\mathfrak{a}$ a proper ideal of R and N a non-zero finitely generated R-module with $N \neq \mathfrak{a}N$. Let d be the least non-negative integer i for which the local cohomology $\mathrm{H}^i_{\mathfrak{a}}(N)$ is non-zero. In this paper, we provide a sharp bound under inclusion for the annihilator of the local cohomology module $\mathrm{H}^d_{\mathfrak{a}}(N)$ and we show that the analogue version of Lynch's conjecture does not hold for $\mathrm{H}^d_{\mathfrak{a}}(N)$. Also, when R is a local ring with maximal ideal $\mathfrak{n}$ and t is an arbitrary non-negative integer, a sharp bound for the annihilator of $\mathrm{H}^t_{\mathfrak{n}}(N)$ is given.

1. Introduction

Throughout this note, we assume that R is a commutative Noetherian ring with non-zero identity. Let $\mathfrak{a}$ be an ideal of R and N be an R-module. For an integer i, the *i-th local cohomology* of N with respect to $\mathfrak{a}$ was defined by Grothendieck as follows

$$\mathrm{H}^i_{\mathfrak{a}}(N) := \varinjlim_{n\in\mathbb{N}} \mathrm{Ext}^i_R(R/\mathfrak{a}^n, N);$$

see [1] for more details. We denote the ideal $\{r \in R : rx = 0 \text{ for all } x \in N\}$ of R by $\mathrm{Ann}_R(N)$. Let N be finitely generated. We recall that the cohomological dimension (respectively, the depth) of N with respect to $\mathfrak{a}$, denoted by $\mathrm{cd}_R(\mathfrak{a}, N)$ (respectively, $\mathrm{depth}_R(\mathfrak{a}, N)$), is defined as the supremum (respectively, infimum) of the non-negative integers i such that $\mathrm{H}^i_{\mathfrak{a}}(N)$ is non-zero. The N-height of $\mathfrak{a}$ is defined as

$$\mathrm{ht}_N(\mathfrak{a}) := \inf\{\dim_{R_{\mathfrak{p}}}(N_{\mathfrak{p}}) : \mathfrak{p} \in \mathrm{Supp}_R(N) \cap \mathrm{V}(\mathfrak{a})\},$$

where $\mathrm{V}(\mathfrak{a})$ denotes the set of all prime ideals of R containing $\mathfrak{a}$. We denote the set of all minimal elements of $\mathrm{Ass}_R(N)$ by $\mathrm{MinAss}_R(N)$. For a submodule L of N and $\mathfrak{p} \in \mathrm{Supp}_R(N)$, we denote the contraction of $L_{\mathfrak{p}}$ under the canonical map $N \to N_{\mathfrak{p}}$ by $C^N_{\mathfrak{p}}(L)$. Also, we denote the set of positive integers (respectively, non-negative integers) by $\mathbb{N}$ (respectively, $\mathbb{N}_0$). For any unexplained notation or terminology, we refer the reader to [1] and [5].

*Talker
Email address: alif1387@gmail.com (Ali Fathi)

Assume that $N \neq \mathfrak{a}N$ and $d := \operatorname{depth}_R(\mathfrak{a}, N)$. We show, in Theorem 2.4, that there is the following bound under inclusion for the annihilator of $\mathrm{H}^d_{\mathfrak{a}}(N)$:

$$\operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Delta} C^N_{\mathfrak{p}}(0)\right) \subseteq \operatorname{Ann}_R\left(\mathrm{H}^d_{\mathfrak{a}}(N)\right) \subseteq \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Sigma} C^N_{\mathfrak{p}}(0)\right) \cap \left(\bigcap_{\mathfrak{p}\in\Sigma'} \mathfrak{p}\right),$$

where $\Delta := \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \mathfrak{a}+\mathfrak{p} \neq R\}$, $\Sigma := \{\mathfrak{p} \in \operatorname{MinAss}_R(N) : \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p}) = d\}$ and $\Sigma' := \{\mathfrak{p} \in \operatorname{Ass}_R(N) \setminus \operatorname{MinAss}_R(N) : \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p}) = d\}$. If in addition N is Cohen-Macaulay, then we show, in Corollary 2.5, that the last inclusion is equality.
Also, when R is a local ring with maximal ideal $\mathfrak{n}$ and N is a non-zero finitely generated R-module, then it is shown in Theorem 2.1 that, for an arbitrary non-negative integer t, there is the following bound for the annihilator of $\mathrm{H}^t_{\mathfrak{n}}(N)$:

$$\operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Delta(t)} C^N_{\mathfrak{p}}(0)\right) \subseteq \operatorname{Ann}_R\left(\mathrm{H}^t_{\mathfrak{n}}(N)\right) \subseteq \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Sigma(t)} C^N_{\mathfrak{p}}(0)\right) \cap \left(\bigcap_{\mathfrak{p}\in\Sigma'(t)} \mathfrak{p}\right),$$

where $\Delta(t) := \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \dim_R(R/\mathfrak{p}) \geq t\}$, $\Sigma(t) := \{\mathfrak{p} \in \operatorname{MinAss}_R(N) : \dim_R(R/\mathfrak{p}) = t\}$ and $\Sigma'(t) := \{\mathfrak{p} \in \operatorname{Ass}_R(N) \setminus \operatorname{MinAss}_R(N) : \dim_R(R/\mathfrak{p}) = t\}$. Finally, in the case when R is a homomorphic image of a Cohen-Macaulay local ring, the dimension of $R/\operatorname{Ann}_R(\mathrm{H}^t_{\mathfrak{n}}(N))$ is considered (see Proposition 2.7 and Corollary 2.9) and we show that the analogue version of Lynch's conjecture is not true for $\mathrm{H}^{\operatorname{depth}_R R}_{\mathfrak{n}}(R)$ even if R is a homomorphic image of a complete regular local ring (see Example 2.8).

2. Main results

In the following theorem, when R is a local ring, we provide a sharp bound for the annihilator of local cohomology module of a finitely generated module with respect to the maximal ideal.

Theorem 2.1 (see [2, Theorem 3.2] and [3, Lemma 5.1]). *Let $(R, \mathfrak{n})$ be a local ring and let N be a non-zero finitely generated R-module. For each $t \in \mathbb{N}_0$, set $\Delta(t) := \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \dim_R(R/\mathfrak{p}) \geq t\}$, $\Sigma(t) := \{\mathfrak{p} \in \operatorname{MinAss}_R(N) : \dim_R(R/\mathfrak{p}) = t\}$ and $\Sigma'(t) := \{\mathfrak{p} \in \operatorname{Ass}_R(N) \setminus \operatorname{MinAss}_R(N) : \dim_R(R/\mathfrak{p}) = t\}$. Then there is the following bound under inclusion for the annihilator of $\mathrm{H}^t_{\mathfrak{n}}(N)$:*

$$\operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Delta(t)} C^N_{\mathfrak{p}}(0)\right) \subseteq \operatorname{Ann}_R\left(\mathrm{H}^t_{\mathfrak{n}}(N)\right) \subseteq \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Sigma(t)} C^N_{\mathfrak{p}}(0)\right) \cap \left(\bigcap_{\mathfrak{p}\in\Sigma'(t)} \mathfrak{p}\right).$$

Let the situation be as in the above theorem and $\mathfrak{p} \in \operatorname{Supp}_R(N)$. Then it is clear that $\operatorname{Ann}_R(N/C^N_{\mathfrak{p}}(0)) \subseteq \mathfrak{p}$. Example 2.6 shows that to improve the upper bound for the annihilator of $\mathrm{H}^t_{\mathfrak{n}}(N)$, we cannot replace $\operatorname{MinAss}_R(N)$ by $\operatorname{Ass}_R(N)$ in the definition of $\Sigma(t)$.

Proposition 2.2 ([3, Lemma 5.2]). *Let N be a finitely generated R-module and let $\mathfrak{a}$ an ideal of R such that $N \neq \mathfrak{a}N$. Let $\mathfrak{p} \in \operatorname{Ass}_R(N)$ with $\mathfrak{a}+\mathfrak{p} \neq R$ and $n := \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p})$. Then*

$$\operatorname{Ann}_R\left(\mathrm{H}^n_{\mathfrak{a}+\mathfrak{p}}(N)\right) \subseteq \mathfrak{p}.$$

If, in addition, $\mathfrak{p} \in \operatorname{MinAss}_R(N)$, then

$$\operatorname{Ann}_R\left(\mathrm{H}^n_{\mathfrak{a}+\mathfrak{p}}(N)\right) \subseteq \operatorname{Ann}_R\left(N/C^N_{\mathfrak{p}}(0)\right).$$

The following corollary is an immediate consequence of Proposition 2.2.

Corollary 2.3. *Let N be a finitely generated R-module such that $|\operatorname{Ass}_R(N)| = 1$, and let $\mathfrak{a}$ be an ideal of R with $N \neq \mathfrak{a}N$. Then*

$$\operatorname{Ann}_R\left(\mathrm{H}^{\operatorname{ht}_N(\mathfrak{a})}_{\mathfrak{a}}(N)\right) = \operatorname{Ann}_R(N).$$

The next theorem provides a sharp bound for the annihilator of the first non-zero local cohomology module of a finitely generated module with respect to an arbitrary ideal in the case when R is not necessarily local.

Theorem 2.4 ([3, Theorem 5.4]). *Let N be a finitely generated R-module, $\mathfrak{a}$ be an ideal of R such that $N \neq \mathfrak{a}N$ and $d := \operatorname{depth}_R(\mathfrak{a}, N)$. There is the following bound for the annihilator of $\mathrm{H}^d_{\mathfrak{a}}(N)$:*

$$\operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Delta} C^N_{\mathfrak{p}}(0)\right) \subseteq \operatorname{Ann}_R\left(\mathrm{H}^d_{\mathfrak{a}}(N)\right) \subseteq \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Sigma} C^N_{\mathfrak{p}}(0)\right) \cap \left(\bigcap_{\mathfrak{p}\in\Sigma'} \mathfrak{p}\right),$$

where $\Delta := \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \mathfrak{a} + \mathfrak{p} \neq R\}$, $\Sigma := \{\mathfrak{p} \in \operatorname{MinAss}_R(N) : \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p}) = d\}$ and $\Sigma' := \{\mathfrak{p} \in \operatorname{Ass}_R(N) \setminus \operatorname{MinAss}_R(N) : \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p}) = d\}$.

Corollary 2.5 ([3, Corollary 5.5]). *Let N be a finitely generated Cohen-Macaulay R-module and let $\mathfrak{a}$ be an ideal of R such that $N \neq \mathfrak{a}N$. Let $d := \operatorname{depth}_R(\mathfrak{a}, N)$ and $\Sigma := \{\mathfrak{p} \in \operatorname{Ass}_R(N) : \operatorname{ht}_{R/\mathfrak{p}}((\mathfrak{a}+\mathfrak{p})/\mathfrak{p}) = d\}$. Then*

$$\operatorname{Ann}_R\left(\mathrm{H}^d_{\mathfrak{a}}(N)\right) = \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{p}\in\Sigma} C^N_{\mathfrak{p}}(0)\right).$$

In particular,

$$\operatorname{ht}_N\left(\operatorname{Ann}_R\left(\mathrm{H}^d_{\mathfrak{a}}(N)\right)\right) = 0$$

and

$$\dim_R\left(R/\operatorname{Ann}_R\left(\mathrm{H}^d_{\mathfrak{a}}(N)\right)\right) = \dim_R(N).$$

Example 2.6 ([3, Example 5.6]). Let K be a field and $R := K[[x,y]]$ be the ring of formal power series over K in indeterminates x, y. Set $N := R/(Rx^2 + Rxy)$, $\mathfrak{p} := Rx$ and $\mathfrak{n} := Rx + Ry$. Then we have $\Gamma_{\mathfrak{n}}(N) \cong R/\mathfrak{n}$, $\mathrm{H}^1_{\mathfrak{n}}(N) \cong \operatorname{Hom}_R(R/\mathfrak{p}, E_R(R/\mathfrak{n}))$ and $\mathrm{H}^i_{\mathfrak{n}}(N) = 0$ for all $i \geq 2$. Thus $\operatorname{depth}_R(N) = 0$, $\operatorname{cd}_R(\mathfrak{n}, N) = \dim_R(N) = 1$ and

$$\operatorname{Ann}_R\left(\mathrm{H}^{\operatorname{depth}_R(N)}_{\mathfrak{n}}(N)\right) = Rx + Ry,$$

$$\operatorname{Ann}_R\left(\mathrm{H}^{\dim_R(N)}_{\mathfrak{n}}(N)\right) = Rx.$$

On the other hand, $\operatorname{Supp}_R(N) = \{\mathfrak{p}, \mathfrak{n}\}$ and if Δ is a subset of $\operatorname{Supp}_R(N)$, then

$$\operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{q}\in\Delta} C^N_{\mathfrak{q}}(0)\right) = \begin{cases} R & \text{if } \Delta = \emptyset, \\ Rx & \text{if } \Delta = \{\mathfrak{p}\}, \\ Rx^2 + Rxy & \text{otherwise.} \end{cases}$$

Therefore the following statements hold.
(i) $\operatorname{Ann}_R\left(\mathrm{H}^{\operatorname{depth}_R(N)}_{\mathfrak{n}}(N)\right) \neq \operatorname{Ann}_R\left(N/\bigcap_{\mathfrak{q}\in\Delta} C^N_{\mathfrak{q}}(0)\right)$ for all subsets Δ of $\operatorname{Supp}_R(N)$,

(ii) By setting $\mathfrak{a} := \mathfrak{n}$, this example shows that to improve the upper bound for the annihilator of $\mathrm{H}^{\operatorname{depth}_R(\mathfrak{a},N)}_{\mathfrak{a}}(N)$ in Theorem 2.4, we can not replace $\operatorname{MinAss}_R(N)$ by $\operatorname{Ass}_R(N)$ in the index set Σ.

Proposition 2.7 ([3, Proposition 5.8]). *Let $(R, \mathfrak{n})$ be a homomorphic image of a Cohen-Macaulay local ring and let N be a non-zero finitely generated R-module. Assume that $t \in \mathbb{N}_0$ is such that $\mathrm{H}^t_{\mathfrak{n}}(N) \neq 0$. Then*

$$\dim_R(R/\operatorname{Ann}_R(\mathrm{H}^t_{\mathfrak{n}}(N)) \leq t$$

and equality holds when there exists $\mathfrak{p} \in \operatorname{Ass}_R(N)$ with $\dim_R(R/\mathfrak{p}) = t$.

The following example shows that there is a local ring $(A, \mathfrak{n})$ which is a homomorphic image of a complete regular local ring such that

$$\dim_A(A/\operatorname{Ann}_A(\mathrm{H}^{\operatorname{depth}_A(A)}_{\mathfrak{n}}(A))) < \operatorname{depth}_A(A) = \operatorname{depth}_A(A/\Gamma_{\mathfrak{n}}(A)).$$

Therefore the inequality in Proposition 2.7 may be strict. Also, this shows that the analogue version of Lynch's conjecture [4, Conjecture 1.2] is not true for $\mathrm{H}^{\operatorname{depth}_A(A)}_{\mathfrak{n}}(A)$.

Example 2.8 ([3, Example 5.9])**.** Let K be a field and let $R := K[[x, y, z, w]]$ be the ring of formal power series over K in indeterminates x, y, z, w. Set $\mathfrak{m} := (x, y, z, w)$ and $I := (x, y) \cap (z, w)$. Then $A := R/I$ is a local ring with maximal ideal $\mathfrak{n} := \mathfrak{m}/I$. By [2, Example 2.8] and the Independence Theorem, we have $\Gamma_{\mathfrak{n}}(A) \cong \Gamma_{\mathfrak{m}}(R/I) = 0$ and $\mathrm{H}^1_{\mathfrak{n}}(A) \cong \mathrm{H}^1_{\mathfrak{m}}(R/I) \cong R/\mathfrak{m} \cong A/\mathfrak{n}$. Therefore

$$\operatorname{depth}_A(A) = \operatorname{depth}_A(A/\Gamma_{\mathfrak{n}}(A)) = 1,$$

$$\dim_A(A/\operatorname{Ann}_A(\mathrm{H}^{\operatorname{depth}_A(A)}_{\mathfrak{n}}(A))) = 0.$$

Thus $\dim_A(A/\operatorname{Ann}_A(\mathrm{H}^{\operatorname{depth}_A(A)}_{\mathfrak{n}}(A)))$ is not equal to $\operatorname{depth}_A(A)$ or $\operatorname{depth}_A(A/\Gamma_{\mathfrak{n}}(A))$ and the inequality in Proposition 2.7 may be strict.

Let $(R, \mathfrak{n})$ be a local ring and let N be a non-zero finitely generated R-module. Then, for each $\mathfrak{p} \in \operatorname{Ass}_R(N)$, $\operatorname{depth}_R(N) \leq \dim_R(R/\mathfrak{p})$. We say that N has *maximal depth* if $\operatorname{depth}_R(N) = \dim_R(R/\mathfrak{p})$ for some $\mathfrak{p} \in \operatorname{Ass}_R(N)$. Cohen-Macaulay modules and sequentially Cohen-Macaulay modules have maximal depth; see [6] for more details. Now assume in addition that R is a homomorphic image of a Cohen-Macaulay local ring. Proposition 2.7 shows that $\dim_R(R/\operatorname{Ann}_R(\mathrm{H}^{\operatorname{depth}_R(N)}_{\mathfrak{n}}(N))) \leq \operatorname{depth}_R(N)$ and Example 2.8 shows that this inequality may be strict. In the following corollary we see that the equality holds if N has maximal depth.

Corollary 2.9. *Let $(R, \mathfrak{n})$ be a homomorphic image of a Cohen-Macaulay local ring and let N be a non-zero finitely generated R-module which has maximal depth. Then*

$$\dim_R\left(R/\operatorname{Ann}_R\left(\mathrm{H}^{\operatorname{depth}_R(N)}_{\mathfrak{n}}(N)\right)\right) = \operatorname{depth}_R(N).$$

Proof. It is an immediate consequence of Proposition 2.7. □

References

[1] M. P. Brodmann and R. Y. Sharp. *Local Cohomology: An Algebraic Introduction with Geometric Applications*, volume 60. Cambridge University Press, Cambridge, 1998.
[2] A. Fathi. Some bounds for the annihilators of local cohomology and ext modules. *Czechoslovak Math. J.*, 72(1):265–284, 2022.
[3] A. Fathi. Annihilator of top local cohomology and lynch's conjecture. *J. Algebra Appl.*, ?(?):2450200, 2024. doi: 10.1142/S0219498824502001.
[4] L. R. Lynch. Annihilators of top local cohomology. *Comm. Algebra*, 40(2):542–551, 2012.
[5] H. Matsumura. *Commutative ring theory*, volume 8. Cambridge University Press, Cambridge, 1986.
[6] A. Rahimi. Maximal depth property of finitely generated modules. *J. Algebra Appl.*, 17(11):1850202, 2018.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–160, pp. 357–359

Some Studies on the Multiplicative Commutators of a Division Ring

Asghar Madadi[a,*], Seyed Mohammad Ajdani[b], Kamal Ahmadi[c]

[a]*Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*
[b]*Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*
[c]*Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*

Article Info

Keywords:
Multiplicative Commutator
Trace Function
Division Ring

2020 MSC:
17A35
17C60

Abstract

Let D be a division ring with center F. An element of the form $xyx^{-1}y^{-1} \in D$ is called a multiplicative commutator. Let $T(D)$ be the vector space over F generated by all multiplicative commutators in D. In this paper it is shown that if D is algebraic over F and $Char(D) = 0$, then $D = T(D)$. Among other results it is shown that in characteristic zero if $T(D)$ is algebraic over F, then D is algebraic over F.

1. Introduction and Preliminaries

Throughout this paper D is a division ring with center F. An element of the form $xyx^{-1}y^{-1} \in D$ is called a *multiplicative commutator*, and D' and $[D, D]$ denote the multiplicative and additive commutator subgroup of D, respectively. Also we denote by $T(D)$ the vector space generated by the set of all multiplicative commutators of D over F. An element $a \in D$ is said to be *algebraic* over F if a satisfies a non-zero polynomial in $F[x]$. A set $S \subseteq D$ is called *algebraic* if each of its elements is algebraic over F. When K is a finite dimensional extension of F, then we denote by $Tr_{K/F}$, the regular trace of K over F. If $a \in D$, then $F(a)$ denotes the subfield of D generated by F and $\{a\}$. An element $a \in D$ is said to be *radical* over F if there exists an integer $n = n(a)$ such that $a^n \in F$. A set $S \subseteq D$ is said to be radical over F, if each element of S is radical over F. Also we denote by $Char(D)$ the characteristic of D.

The division ring generated by additive commutators or multiplicative commutators of D is the whole D [7, pp. 205, 211]. In the algebraic and zero characteristic case, it was proved that D is generated as a vector space over the center by the union of its additive commutators and the unity, see [1]. As a generalization we provide another generating structure for a division ring as a vector space generated by all multiplicative commutators over its center. We prove this in the algebraic case when the characteristic is zero. Besides, consider a special property P in a ring (for example commutativity, algebricity or some finiteness conditions), there are a lot of studies in literature to specify a set or a substructure S, such that the property P for S implies the property P for the whole ring, for instance see [1–6] and

*Talker
Email addresses: amadadi2009@gmail.com (Asghar Madadi), majdani2@yahoo.com (Seyed Mohammad Ajdani), kamal.ahmadi.math@gmail.com (Kamal Ahmadi)

[8–10]. Most of these studies have focused on the set of multiplicative and additive commutators and their generating subgroups D' and $[D, D]$ in division rings. We show that the subspace $T(D)$ reflects some properties to the whole division ring.

2. Results

We start this section with the following lemma.

Lemma 2.1. *Let D be a division ring with center F. Then for each algebraic element $a \in D$, there exists an element $d \in T(D) \cap F(a)$ such that $Tr_{F(a)/F}(a) = ad$.*

The following lemma is an immediate consequence of the previous lemma.

Lemma 2.2. *Let D be a division ring with center F and $a \in D$ be an algebraic element over F such that $Tr_{F(a)/F}(a) \neq 0$. Then $a^{-1} \in T(D)$.*

Now, we have the following theorem.

Theorem 2.3. *Let D be a division ring with center F. Then $T(D)$ contains all separable elements of D.*

The following is a simple corollary of the previous theorem.

Corollary 2.4. *Let D be an algebraic division ring over its center F. If $Char(D) = 0$, then $T(D) = D$.*

3. Some other aspects of $T(D)$

A theorem due to Kaplansky [7, p. 246] states that if D is radical over F, then $D = F$. There are various kinds of generalizations of this theorem. The next theorem is another one.

Theorem 3.1. *Let D be a division ring with center F. If $T(D)$ is radical over F, then $D = F$.*

Mahdavi-Hezavehi together with his colleagues in [10] conjectured that one could conclude the algebracity of a division ring over its center from the algebracity of all its multiplicative commutators. They were able to deduce the conjecture is true in the case D' is algebraic over the center, but in general the problem is still open. Also it is proved that when $[D, D]$ is algebraic over F, then D is algebraic over F, see [1]. In what follows we give an affirmative answer to the conjecture when the center of division ring is uncountable.

Theorem 3.2. *Let D be a division ring with uncountable center F such that all of its multiplicative commutators are algebraic over F. Then D is algebraic over F. In particular, if D is a division ring with uncountable center F and $T(D)$ is algebraic over F, then D is algebraic over F.*

Now, a question is naturally proposed: Whether a division ring is finite dimensional when $T(D)$ is a finite dimensional vector space over F? This is the content of the following theorem. This theorem is along of some analogous results which state that if each element of a specific set has a minimal polynomial of bounded degree, then D is finite dimensional over F [10]. Note that finite dimensionality of $T(D)$ as a vector space does not imply that all elements of $T(D)$ or multiplicative commutators are algebraic over F.

Theorem 3.3. *Let D be a division ring with center F. If $dim_F T(D) = n < \infty$, then $dim_F D < \infty$.*

Now, we would like to prove that if D is a division ring with center F and $T(D)$ is algebraic over F, then D is algebraic over F. Before stating the proof we need a lemma.

Lemma 3.4. *Let D be a division ring with center F, $T(D)$ is algebraic over F and $Char(D) = 0$. Then for any two algebraic elements $a, b \in D$, the set $S = \{a + b, aba, a^2 b\}$ is algebraic over F.*

To prove the main theorem we need the following well known result.

Theorem 3.5. *(Cartan-Brauer-Hua [7, p. 211]) Let D be a division ring with center F and K be its subdivision ring such that for any non-zero element $x \in D$, $xKx^{-1} \subseteq D$ and $K \neq D$. Then $K \subseteq F$.*

Theorem 3.6. *Let D be a division ring with center F and $Char(D) = 0$. Then $T(D)$ is algebraic over F if and only if D is algebraic over F.*

Proof. Let $T(D)$ be algebraic over F. If a and b are algebraic elements of D over F, then by the previous lemma $(a+1)^2b$ is algebraic over F and so $ab = 1/2((a+1)^2b - b - a^2b)$ is algebraic over F. Therefore the set A containing all algebraic elements of D over F forms an algebraic subdivision ring containing $T(D)$ which is also invariant under conjugation. Now, by Theorem 3.5 , the proof is complete. $\square$

References

[1] S. Akbari, M. Ariannejad, M. L. Mehrabadi, On additive commutator groups in division rings, *Results Math.* **33** (1998) 9-21.
[2] M. Ariannejad, Some studies on additive commutator groups in division rings, *Bull. Iranian Math. Soc.* **26**(1) (2000) 1-6.
[3] M. A. Chebotar, Y. Fong, P. H. Lee, On division rings with algebraic commutators of bounded degree, *Manuscripta Math.* **113** (2004) 153-164.
[4] K. Chiba, Skew fields with a non-trivial generalized power central rational identity, *Bull. Austr. Math. Soc.* **49** (1994) 85-90.
[5] I. N. Herstein, Multiplicative commutators in division rings, *Israel J. Math.* **131**(2) (1978) 180-188.
[6] I. N. Herstein, Multiplicative commutators in division rings II, *Rend. Circ. Mat. Palermo II* **29** (1980) 485-489.
[7] T. Y. Lam, *A First Course in Noncommutative Rings*, 2nd edn., Graduate Texts in Mathematics, Vol. 131 (Springer, New York, 2001).
[8] M. Mahdavi-Hezavehi, On derived groups of division rings, *Commun. Algebra* **23**(3) (1995) 913-926.
[9] M. Mahdavi-Hezavehi, S. Akbari-Feyzaabaadi, On a question of Herstein concerning commutators in division rings, *Rev. Roumaine Math. Pures Appl.* **41**(3-4) (1996) 201-204.
[10] M. Mahdavi-Hezavehi, S. Akbari-Feyzaabaadi, M. Mehraabaadi, H. Haji-Abolhassan, On derived groups of division rings II, *Commun. Algebra* **23**(8) (1995) 2881-2887.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–161, pp. 360–362

Finite Groups With Some ν-permutable Subgroups

Hassan Jafarian Dehkordy[a,*]

[a]*Department of Mathematics, Payme Noor University, P.O.Box 19395-4697, Tehran, Iran.*

Article Info

Keywords:
S-permutable
S-semipermutable
Weakly S-semipermutable
SS-quasi-normal
p-nilpotent.

2020 MSC:
20D15
20D20
20F19

Abstract

Let G be a finite group. If $H \leq G$, recall that H is Weakly S-semipermutable in G provided there exists a normal subgroup A of G such that HA is S-permutable in G, and $H \cap A$ is S-semipermutable in G. The purpose of this survey note is to show Weakly S-semipermutability or SS-quasi-normality of especial types of subgroups in a group G can help us to determine the properties of G. This relationship Weakly S-semipermutability or SS-quasi-normality of special subgroups of group G and, p-supersolubility group G. Moreover, we investigate, if some subgroups of G be Weakly S-semipermutable or SS-quasi-normal in G, then G is a p-nilpotent group.

1. Introduction

All groups mentioned in this paper are considered to be finite. Let $\pi(G)$ stand for the set of all prime divisors of the order of a group G. Recall that two subgroups H and K of a group G are said to be permutable if $HK = KH$. The subgroup H is said to be S-permutable (S-quasinormal or π-quasinormal) in G if H permutes with every Sylow subgroup of G, i.e., $HG_p = G_pH$ for any Sylow subgroup G_p of G. This concept was introduced by O. H. Kegel [3]. A subgroup H of G is said to be S-semipermutable in G, if $HG_p = G_pH$ for every Sylow p-subgroup G_p of G with $(|H|, p) = 1$. This concept was introduced by Chen(1998). Clearly, every S-permutable subgroup of G is S-semipermutable in G but the converse does not hold: for example, a Sylow 2-subgroup of S_3 (the symmetric group of degree 3) is semipermutable in S_3, so is S-semipermutable in S_3 but does not S-permutable in S_3. Kegel proved that, for every S-permutable subgroup H of G, H is subnormal subgroup of G.
H.Zhangjia (2010) proved that, let p be an odd prime dividing the order of a group G and G_p be a Sylow p-subgroup of G. Suppose that $N_G(G_p)$ be a p-nilpotent group and there exists a subgroup D of G_p with $1 < |D| < |G_p|$ such that every subgroup H of G_p with order $|D|$ is S-semipermutable in G. Then G is a p-nilpotent.
Let G be a group and $H \leq G$. H is called SS-quasi-normal in G (Supplement Sylow-quasi-normal subgroup) if there exists a Supplement K of H in G, such that $HB_p = B_pH$ for all Sylow p-subgroups B_p of K, Li [4].
S.E. Mirdamadi and G.R. Rezaeezadeh [5] introduced SS-semipermutability's concept, which is a general of SS-quasi-normality and semipermutability.

*Talker
Email address: h.jafarian.d@pnu.ac.ir (Hassan Jafarian Dehkordy)

2. Preliminaries

Below are some of the Lemmas and Theorems needed to prove the main results.

Lemma 2.1. *[1, Lemma 2.5] Let G be a group and $p \in \pi(G)$ a prime with $(|G|, p-1) = 1$. Then the subsequent statements stand:*

1. *If $N \trianglelefteq G$ and $|N| = p$, then $N \leq Z(G)$.*
2. *If G has a Sylow p-subgroup such that it is cyclic, then G is p-nilpotent.*
3. *If $M \leq G$ such that $|G : M| = p$, then $M \trianglelefteq G$.*

Lemma 2.2. *[1, Lemma 2.4] Let G be a minimal non-nilpotent group. Then the following hold:*

1. *For some $p \in \pi(G)$, there exists $G_p \in Syl_p(G)$ such that $G_p \trianglelefteq G$ and $G = G_pQ$, where Q is a cyclic non-normal Sylow q-subgroup of G for some prime $q \neq p$.*
2. *If $p > 2$, then G_p has exponent p. If $p = 2$, then G_p has exponent 2 or 4.*
3. *If G_p is abelian, then G_p is elementary abelian.*
4. *$\Phi(G_p) \leq Z(G)$.*
5. *$G_p/\Phi(G_p)$ is a chief factor of G.*

Lemma 2.3. *[2, Lemma 2.1] Suppose that a subgroup H of a group G is S-permutable in G and N is a normal subgroup of G. Then the following hold:*

1. *If $H \leq K \leq G$ then H is S-permutable in K.*
2. *HN and $H \cap N$ are S-permutable in G, HN/N is S-permutable in G/N.*
3. *$H \cap K$ is S-permutable in K.*
4. *If H is a p-subgroup of G, then $H \subseteq O_p(G)$ and $O^p(G) \leq N_G(H)$.*
5. *H/H_G is nilpotent.*

Lemma 2.4. *[1, Lemma 2.2] Let G be a group, $p \in \pi(G)$ be a prime and A be a p-subgroup of G. Then A is S-permutable in G if and only if $O^p(G)$ normalizes A.*

Lemma 2.5. *[1, Lemma 2.3] Let G be a group and $p \in \pi(G)$ be a prime and $A \leq O_p(G)$. If A is S-semipermutable in G, then A is S-permutable in G.*

Lemma 2.6. *[2, Lemma 2.2] Suppose that a subgroup H of a group G is S-semipermutable in G and N is a normal subgroup of G, then the following hold:*

1. *If $H \leq K \leq G$, then H is S-semipermutable in K.*
2. *If H is a p-subgroup for some prime $p \in \pi(G)$, then HN/N is S-semipermutable in G/N.*
3. *If $(|H|, |N|) = 1$, then HN/N is S-semipermutable in G/N.*
4. *If $H \leq O_p(G)$, then H is S-permutable in G.*

Lemma 2.7. *[4, Lemma 2.1, 2.2] Let $H \leq G$ be the SS-quasi-normal in G, $H \leq L \leq G$, and $N \trianglelefteq G$.*

1. *H is SS-quasi-normal in L.*
2. *HN/N is SS-quasi-normal in G/N.*
3. *If for some prime $p \in \pi(G)$, H is a p-subgroup of G, then $HG_q = G_qH$ for every $G_q \in Syl_q(G)$ with $q \neq p$ ($q \in \pi(G)$).*
4. *If $H \leq F(G)$, then H is S-quasi-normal in G.*

G.R. Rezaeezadeh and H. Jafarian Dehkordy[2] define the Weakly S-semipermutable (ν-permutable) subgroups.

Definition 2.8. Let G be a group and $H \leq G$. Then H is said to be ***Weakly S-semipermutable*** (ν-permutable) in G provided there exists $A \trianglelefteq G$ such that HA is S-permutable in G and $H \cap A$ is S-semipermutable in G .

It is clear that if $K \leq G$ is S-semipermutable in G, then K is ν-permutable in G. However, the converse is not true. For instance, let K denote the subgroup $\langle(12)\rangle$ of S_4 (symmetric group of degree 4). Then K is easily seen to be ν-permutable in S_4, but K is not S-semipermutable in S_4.
The purpose of this survey paper is to show ν-permutability of some subgroups of a group G can help us to determine the p-supersolubility or p-nilpotency of G.
G.R.Rezaeezadeh and H.Jafarian Dehkordy [2],[1] prove the following results.

Lemma 2.9. *[2, Lemma 2.7] Suppose that a subgroup H of a group G is ν-permutable in G and N is a normal subgroup of G, then the following statements hold:*

1. *If $H \leq K \leq G$, then H is ν-permutable in K.*
2. *If $(|H|, |N|) = 1$, then HN/N is ν-permutable in G/N.*
3. *If $H \leq K \trianglelefteq G$, then G has a normal subgroup L contained in K such that HL is S-permutable in G and $H \cap L$ is S-semipermutable in G.*
4. *If H is a p-subgroup of G, then HN/N is ν-permutable in G/N.*
5. *If $N \leq U$ and U/N is ν-permutable in G/N, then U is ν-permutable in G.*

Lemma 2.10. *[2, Lemma 2.8] Let $N \trianglelefteq G$ be a minimal normal and elementary abelian subgroup. Then N has no nontrivial proper subgroup K such that any subgroup of N with order $|K|$ is ν-permutable in G.*

Theorem 2.11. *[2, Theorem 3.2] Let G be a group, $p \in \pi(G)$ with $(|G|, p-1) = 1$ and G_p is a Sylow p-subgroup of G. If every maximal subgroup of G_p is ν-permutable in G, then G is p-nilpotent.*

Theorem 2.12. *[2, Theorem 3.3] Let G be a group, $p \in \pi(G)$ with $(|G|, p-1) = 1$ and G_p be a Sylow p-subgroup of G. If every cyclic subgroup of G_p with order p or 4 (if G_p is a nonabelian 2-group) has a p-nilpotent supplement in G or is ν-permutable in G, then G is p-nilpotent.*

Theorem 2.13. *[1, Theorem 3.1] Let G be a p-soluble group and $G_p \in Syl_p(G)$ where $p \in \pi(G)$. If each of the maximal subgroups of G_p is ν-permutable in G, then G is p-supersoluble.*

3. Main results

Theorem 3.1. *Let G be a group, p be the smallest prime dividing the order of G, and $G_p \in Syl_p(G)$. If each of the maximal subgroups of G_p is either ν-permutable in G or SS-quasi-normal in G, then G is p-nilpotent.*

Theorem 3.2. *Let G be a group, p be the smallest prime dividing the order of G, and $G_p \in Syl_p(G)$. If there exists a subgroup L of G_p with $1 < |L| < |G_p|$ such that all subgroups K of G_p with $|K| = |L|$ or $|K| = 2|L|$ (G_p is a non-abelian 2-group) is either ν-permutable or SS-quasi-normal in G, then G is p-nilpotent.*

Theorem 3.3. *Let G be a group and $p \in \pi(G)$ with $(|G|, p-1) = 1$. Let G_p be a Sylow p-subgroup of G. Suppose that any maximal subgroup of G_p, that does not have a p-nilpotent supplement in G, is ν-permutable in G. Then G is p-nilpotent.*

References

[1] H Jafarian Dehkordy, GR Rezaeezadeh, and M Bibak. Structure of finite groups with some weakly s-semipermutable subgroups. *Journal of Mahani Mathematical Research Center*, 13(1):459–468, 2024.

[2] Hassan Jafarian Dehkordy and Gholamreza Rezaeezadeh. Weakly s-semipermutable subgroups and p-nilpotency of groups. *Rendiconti del Seminario Matematico della Università di Padova*, (150):137–148, 2023.

[3] Otto H Kegel. Sylow-gruppen und subnormalteiler endlicher gruppen. *Mathematische Zeitschrift*, 78(1):205–221, 1962.

[4] Shirong Li, Zhencai Shen, Jianjun Liu, and Xiaochun Liu. The influence of ss-quasinormality of some subgroups on the structure of finite groups. *Journal of Algebra*, 319(10):4275–4287, 2008.

[5] SE Mirdamadi and Gh R Rezaeezadeh. On finite groups in which ss-semipermutability is a transitive relation. *Journal of Algebraic Systems*, 4(1):29–36, 2016.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–162, pp. 363–366

On $\delta - \oplus$-supplemented modules

Behnam Talaee[a,*], Mehran Motiee[b]

[a] *Department of Mathematics, Faculty of Basic Sciences, Babol Noshirvani University of Technology, Shariati Ave., Babol, Iran Post Code:47148-71167*

[b] *Department of Mathematics, Faculty of Basic Sciences, Babol Noshirvani University of Technology, Shariati Ave., Babol, Iran Post Code:47148-71167*

Article Info

Keywords:
Weak $\delta-$coclosed submodules
$\oplus-$supplemented modules
$\delta - \oplus-$supplemented modules

2020 MSC:
16D80
16D90

Abstract

$\delta - \oplus-$supplemented module is a generalization of $\oplus-$supplemented modules. We introduce (weak) $\delta-$coclosed submodules and give the nature of $\delta - \oplus-$supplemented modules under these submodules. We show that if M is a $\delta - \oplus-$supplemented module and N is a $\delta-$coclosed submodule of M such that the intersection of N with any direct summand of M is a direct summand of N, then N is $\oplus-$supplemented and especially $\delta - \oplus-$supplemented. It will be shown any $\delta - \oplus-$supplemented module M has a decomposition $M = M_1 \oplus M_2$, where $\delta(M_1) \ll_\delta M_1$ and $\delta(M_2) = M_2$. The relationship of this type of modules with some other modules and also the nature of $\delta - \oplus-$supplemented modules under direct sum and direct summand will be investigated.

1. Introduction

A submodule L of M is called *small* in M (denoted by $L \ll M$) if, for every proper submodule K of M, $L+K \neq M$. A submodule N of M is called *essential* in M (denoted by $N \trianglelefteq_e M$) if $N \cap K \neq 0$ for every nonzero submodule K of M. The *singular* submodule of a module M (denoted by $Z(M)$) is $Z(M) = \{x \in M \mid Ix = 0$ for some ideal $I \trianglelefteq_e R\}$. A module M is called *singular* (*nonsingular*) if $Z(M) = M$ (resply. $Z(M) = 0$).
A submodule N of M is called *fully invariant* in M if $f(N) \leq N$ for every $f \in End(M)$.
A module M is called *lifting* if, for every submodule N of M, M has a decomposition $M = A \oplus B$ such that $A \leq N$ and $N \cap B \ll B$. M is said to *satisfy condition* D_3 if, whenever A and B are direct summands of M with $M = A+B$, then $A \cap B$ is also a direct summand of M.
For two submodules N and K of the module M, N is called a *supplement* of K in M if N is minimal with respect to the property $M = K + N$, equivalently $M = K + N$ and $N \cap K \ll N$. The module M is called *supplemented* if every submodule of M has a supplement in M. The module M is called $\oplus-$supplemented if every submodule of M has a supplement that is a direct summand of M. M is called completely $\oplus-$supplemented if every direct summand of M is $\oplus-$supplemented (see [1]).

*Talker
Email addresses: behnamtalaee@nit.ac.ir (Behnam Talaee), motiee@nit.ac.ir (Mehran Motiee)

Let M be a module and $B \leq A \leq M$. If $A/B \ll M/B$, then B is called a *cosmall* submodule of A in M. The submodule A of M is called *coclosed* if A has no proper $cosmall$ submodule. Also B is called a *coclosure* of A in M if B is a cosmall submodule of A and B is coclosed in M.

2. $\delta-$upplemented and $\delta-\oplus-$upplemented modules

$\delta-$*small* submodules were defined as a generalization of small submodules by Zhou in [5]. After that, many authors have done researches related to this type of modules (for example see [2–4] . Let M be a module and $L \leq M$. Then L is called $\delta-$*small* in M (denoted by $L \ll_\delta M$) if, for any submodule N of M with M/N singular, $M = N + L$ implies that $M = N$. The sum of all $\delta-$small submodules of M is denoted by $\delta(M)$.
Let K, N be submodules of module M. Then N is called a $\delta-$*supplement* of K in M if $M = N+K$ and $N \cap K \ll_\delta N$. The module M is called $\delta-$*supplemented* if every submodule of M has a $\delta-$supplement in M.
The module M is called $\delta-$*lifting* if, for any submodule N of M, there is a decomposition $M = M_1 \oplus M_2$ such that $M_1 \leq N$ and $N \cap M_2 \ll_\delta M_2$.
In this paper we investigate generalizations of $\oplus-$supplemented and completely $\oplus-$supplemented modules, namely $\delta-\oplus-$supplemented modules and completely $\delta-\oplus-$supplemented modules. A module M is called a $\delta-\oplus-$supplemented module if every submodule of M has a $\delta-$supplement that is a direct summand of M. Also M is called a completely $\delta-\oplus-$supplemented module if every direct summand of M is $\delta-\oplus-$supplemented.
Let M be any module and $B \leq A$ be submodules of M. Then B is called a $\delta-$*cosmall* submodule of A in M if $A/B \ll_\delta M/B$. A submodule A of M is called $\delta-$*coclosed* in M if N has no proper $\delta-$cosmall submodule in M, that is, if $B \leq A$ such that $A/B \ll_\delta M/B$, then $A = B$. A submodule A of M is *weak* $\delta-$*coclosed* in M if, given $B \leq A$ such that A/B is singular and $A/B \ll_\delta M/B$, then $A = B$. For a submodule N of M, $A \leq N$ is called a $\delta-$*coclosure* of N in M if A is $\delta-$coclosed in M and $N/A \ll_\delta M/A$ and A is called a *weak* $\delta-$*coclosure* of N in M if A is weak $\delta-$coclosed in M and $N/A \ll_\delta M/A$.
It is easy to see that every small submodule of a module M is $\delta-$small in M, so $Rad(M) \subseteq \delta(M)$ and, if M is singular, all $\delta-$small submodules of M are small and so $Rad(M) = \delta(M)$. Also any non-singular semisimple submodule of M is $\delta-$small in M.

Example 2.1. Let R be a semisimple ring and M a left $R-$module. Since R has no essential ideal, all submodules of M are non-singular. Hence all submodules of are $\delta-$small in M, while non of its nonzero submodules are small. Especially all submodules of the $\mathbb{Z}_6-$modules, $\mathbb{Z}_6$ are $\delta-$small but it has no nonzero small submodules.

Lemma 2.2. *Let M and N be modules. Then*

(1) $\delta(M) = \sum\{L \leq M \mid L \ll_\delta M\} = \bigcap\{K \leq M \mid M/K \text{ is singular simple }\}$.

(2) *If $f : M \to N$ is an $R-$homomorphism, then $f(\delta(M)) \subseteq \delta(N)$. Therefore $\delta(M)$ is a fully invariant submodule of M. In particular, if $K \leq M$, then $\delta(K) \subseteq \delta(M)$.*

(3) *If $M = \oplus_{i\in I} M_i$, then $\delta(M) = \oplus_{i\in I}\delta(M_i)$.*

(4) *If every proper submodule of M is contained in a maximal submodule of M, then $\delta(M)$ is the unique largest $\delta-$small submodule of M. In particular if M is finitely generated, then $\delta(M)$ is $\delta-$-small in M.*

Lemma 2.3. *Let M be a module and $\delta(M) \leq K \leq M$. Then the following hold:*

(1) *If $\delta(M)$ is $\delta-$small in M and $\delta(M)$ is a $\delta-$cosmall submodule of K in M, then K is $\delta-$small in M.*

(2) *If $\delta(M)$ is $\delta-$small in M, then $\delta(M/\delta(M)) = 0$.*

3. Main Results

Lemma 3.1. *Let M be a module and $N \subseteq^{\oplus} M$. Then N is weak $\delta-$coclosed in M.*

Lemma 3.2. *Let M be a module and $A \leq N \leq M$ such that N is weak $\delta-$coclosed in M. Then $A \ll_\delta M$ implies that $A \ll_\delta N$.*

Lemma 3.3. *Let M be a module and $A \leq N \leq M$ be such that N is $\delta-$coclosed in M. Then $A \ll_\delta M$ implies that $A \ll N$.*

Theorem 3.4. *Let M be a $\delta-\oplus-$supplemented module and N be a weak $\delta-$coclosed submodule of M. If the intersection of N with any direct summand of M is a direct summand of N, then N is $\delta-\oplus-$supplemented. In particular, if M has the summand intersection property, then every direct summand of M is $\delta-\oplus-$supplemented.*

Corollary 3.5. *Let M be a $\delta-\oplus-$supplemented module and N be a $\delta-$coclosed submodule of M. If the intersection of N with any direct summand of M is a direct summand of N, then N is $\oplus-$supplemented.*

Corollary 3.6. *Let M be a $\delta-\oplus-$supplemented module and N be a weak $\delta-$coclosed ($\delta-$coclosed) submodule of M. If $eN \leq N$ for all $e = e^2 \in End(M)$, then N is $\delta-\oplus-$supplemented ($\oplus-$supplemented). In particular any fully invariant $\delta-$coclosed submodule of M is $\oplus-$supplemented.*

Lemma 3.7. (1) *Let M be a module and N, L be submodules of M. Then $N + L \ll_\delta M$ if and only if $N \ll_\delta M$ and $L \ll_\delta M$.* (2) *Let M, N be modules with $K \leq M$. If $f : M \to N$ is a homomorphism, then $K \ll_\delta M$ implies that $f(K) \ll_\delta N$. In particular $K \ll_\delta M \leq N$ implies $K \ll_\delta N$.*
(3) *Let $M = M_1 \oplus M_2$, $K_1 \leq M_1 \leq M$ and $K_2 \leq M_2 \leq M$. Then $K_1 \oplus K_2 \ll_\delta M_1 \oplus M_2$ if and only if $K_1 \ll_\delta M_1$ and $K_2 \ll_\delta M_2$. In particular if $X \ll_\delta N$ and $Y \ll_\delta M$, then $X + Y \ll_\delta M + N$.*

Theorem 3.8. *Any finite direct sum of $\delta-\oplus-$supplemented modules is a $\delta-\oplus-$supplemented module.*

Lemma 3.9. *The following are equivalent :*
(1) M is $\delta-D_1$.
(2) Every submodule A of M can be written as $A = N \oplus S$ with $N \subseteq^{\oplus} M$ and $S \ll_\delta M$.

A module M is called *distributive* if its lattice of submodules is distributive, equivalently for all submodules L, K, N of M, $N + (K \cap L) = (N + K) \cap (N + L)$ or $N \cap (K + L) = (N \cap K) + (N \cap L)$.

Corollary 3.10. *Any distributive $\delta-D_1$ module is a $\delta-\oplus-$supplemented module.*

Corollary 3.11. *Let $M_1, M_2, \ldots, M_n$ be modules where each M_i is distributive and $\delta-$lifting. Then $\oplus_{i=1}^{n} M_i$ is a $\delta-\oplus-$supplemented module.*

Recall that a module M is called $\delta-$hollow if every proper submodule of M is $\delta-$small in M. It is clear that any $\delta-$hollow module is $\delta-$lifting and so any finite direct sum of distributive $\delta-$hollow modules is a $\delta-\oplus-$supplemented module.

Theorem 3.12. *Any $\delta-\oplus-$supplemented module M has a decomposition $M = M_1 \oplus M_2$, where $\delta(M_1) \ll_\delta M_1$ and $\delta(M_2) = M_2$.*

The submodule $Z^*(M)$ of M is defined by $Z^*(M) = \{m \in M : mR$ is small in $E(mR)\}$ where $E(mR)$ is the injective hull of mR (see for example [4]). We define the submodule $\delta^*(M)$ of M by $\delta^*(M) = \{m \in M : mR$ is $\delta-$small in $E(mR)\}$. Note that if $M = M_1 \oplus M_2$ we have $Z^*(M) = Z^*(M_1) \oplus Z^*(M_2)$ and $\delta^*(M) = \delta^*(M_1) \oplus \delta^*(M_2)$.

Theorem 3.13. *Let M be a $\delta-\oplus-$supplemented module. Then there exists a decomposition $M = M_1 \oplus M_2$, where $\delta^*(M_1) \ll_\delta M_1$ and $\delta^*(M_2) = M_2$.*

Lemma 3.14. *Let M be a nonzero module and N a fully invariant submodule of M. If $M = M_1 \oplus M_2$, then $N = (N \cap M_1) \oplus (N \cap M_2)$.*

Theorem 3.15. *Let N be a fully invariant submodule of the module M. If M is $\delta-\oplus-$supplemented, then M/N is $\delta-\oplus-$supplemented. Moreover if N is a direct summand of M, then N is also $\delta-\oplus-$supplemented.*

A module N is called *radical* if $Rad(N) = N$. The sum of all radical submodules N of M is denoted by $P(M)$. Here we say that a module N is $\delta-$*radical* if $\delta(N) = N$ and denote the sum of all $\delta-$radical submodules N of M by $P_\delta(M)$.

Corollary 3.16. *Let M be a module. If M is $\delta-\oplus-$supplemented, then $M/P_\delta(M)$ is $\delta-\oplus-$supplemented. Moreover if $P_\delta(M) \subseteq^{\oplus} M$ then $P_\delta(M)$ is $\delta-\oplus-$supplemented.*

Example 3.17. (1) Let M denote the $\mathbb{Z}-$module $\mathbb{Z}/6\mathbb{Z}$. The proper nonzero submodules of M are $2\mathbb{Z}/6\mathbb{Z}$ and $3\mathbb{Z}/6\mathbb{Z}$ and we have $\mathbb{Z}/6\mathbb{Z} = 2\mathbb{Z}/6\mathbb{Z} \oplus 3\mathbb{Z}/6\mathbb{Z}$. Hence M is a $\delta-\oplus-$supplemented module.
(2) The $\mathbb{Z}-$module $\mathbb{Z}$ is not a $\delta-\oplus-$supplemented module.

References

[1] HARMANCIA., KESKIN D., AND SMITH P. F., *On $\oplus-$supplemented modules*, Acta Math., (1999), 83, 161-169.
[2] KOSAN M. T., *$\delta-$lifting and $\delta-$supplemented Modules*, Alg. Coll., (2007), 14 (1), 53-60.
[3] B. Talaee, *$G-\square\delta-\square M$ Modules and Torsion Theory Cogenerated by Such Modules*, Iranian Journal of Science and Technology, Transactions A: Science, 42: 141 - 146, (2018).
[4] B. Talaee, *A Generalization of M-Small Modules*, Journal of Sciences, Islamic Republic of Iran, 26(2): 179 - 185, (2015).
[5] ZHOU Y., *Generalizations of perfect, semiperfect, and semiregular rings*, Alg. Coll. 7 (3) (2000), 305-318.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–178, pp. 367–369

Realizing the singularity category as a Spanier-Whitehead triangulated category

Abdolnaser Bahlekeh[a,*]

[a] *Gonbad-Kavous University, Gonbad-Kavous, Iran*

Article Info

Keywords:
singularity category
Spanier-Whitehead category
maximal Cohen-Macaulay modules

2020 MSC:
13C60
13D05

Abstract

Let $(R, \mathfrak{m})$ be a commutative noetherian Cohen-Macaulay ring with a canonical module ω. In this paper, we describe the singularity $\mathsf{D}_{\mathsf{sg}}(R)$ of R in terms of the Spanier-Whitehead category of the category of maximal Cohen–Macaulay R-modules.

1. Introduction

Let $(R, \mathfrak{m})$ be a commutative noetherian local ring. The *singularity category* of R is the Verdier quotient

$$\mathsf{D}_{\mathsf{sg}}(R) = \mathsf{D}^{\mathsf{b}}(\mathsf{mod}R)/\mathsf{P}(R)$$

of the bounded derived category of finitely generated R-modules $\mathsf{D}^{\mathsf{b}}(\mathsf{mod}R)$ by the full subcategory $\mathsf{P}(R)$ of bounded complexes of finitely generated projective R-modules. This category measures the homological singularity of R in the sense that R has finite global dimension if and only if its singularity category is trivial. The singularity category was introduced by Buchweitz [3] in the 1980s, and studied actively ever since the relation with mirror symmetry was found by Orlov [6]. Assume that $\mathsf{Gp}(R)$ is the category of all finitely generated Gorenstein projective R-modules. A nice feature of this category is that its stable version $\underline{\mathsf{Gp}}(R)$ modulo projectives is a triangulated category. It is known that the natural triangulated functor $F : \underline{\mathsf{Gp}}(R) \longrightarrow \mathsf{D}_{\mathsf{sg}}(R)$ sending each object X to the complex concentrated in degree zero is fully faithful. A fundamental result of Buchweitz and Happel [3, 4] states that F is an equivalence, provided that R is Gorenstein. So a natural question is arisen: Is there a description of the singularity category of a Cohen-Macaulay ring? The aim of this paper is to give an affirmative answer to this question. In this direction, as already mentioned above, the singularity category of a Gorenstein ring can be described via the stable category Gorenstein projective modules and Since over gorenstein rings, the category of Gorenstein projective modules is the

*Talker
Email address: first.author@gmail.com (Abdolnaser Bahlekeh)

same as the category of maximal Cohen-Macaulay modules, an appropriate candidate will be the category of maximal Cohen-Macaulay modules. But the problem is that the stable category of maximal Cohen-Macaulay modules is not a triangulated category. To overcome this hole, we use the notion of the Spanier–Whitehead category, which has been defined by Heller [5]. Indeed, inspired by a well-known construction in algebraic topology, he defined the Spanier–Whitehead category for each left triangulated category by formally inverting the suspension and proved that it is always a triangulated category. Assume that R is a Cohen-Macaulay ring with a canonical module and $\mathsf{CM}(R)$ is the category of maximal Cohen-Macaulay R-modules. Since $\mathsf{CM}(R)$ is a resolving subcategory of the category of finitely generated R-modules, its stable category modulo projectives forms a left triangulated category with the syzygy functor being the suspension functor. So by Heller's result, the Spanier–Whitehead category of the stable category of maximal Cohen-Macaulay modules, $SW(\underline{\mathsf{CM}})$, is a triangulated category. The main result of this paper is to show that the singularity category $\mathsf{D_{sg}}(R)$ of R can be described within $SW(\underline{\mathsf{CM}})$.

2. Results

Let us begin this section wby recalling the definition of the left triangulated category.

Definition 2.1. Let $\mathcal{C}$ be an additive cateory and $\Omega : \mathcal{C} \to \mathcal{C}$ an additive covariant functor. Assume that Δ is a collection of sequences of the form $\Omega X \xrightarrow{h} Z \xrightarrow{g} Y \xrightarrow{f} X$ in $\mathcal{C}$, called *left triangles*. The triple $(\mathcal{C}, \Omega, \Delta)$ is said to be a *left triangulated category*, if the following conditions are satisfied:
(LT0): Any sequence which is isomorphic to a left triangle is a left triangle. Moreover, for any object $X \in \mathcal{C}$, the left triangle $0 \longrightarrow X \xrightarrow{id} X \longrightarrow 0$ belongs to Δ.
(LT1): For any morphism $f : Y \longrightarrow X$ in $\mathcal{C}$, there is a left triangle $\Omega X \xrightarrow{h} Z \xrightarrow{g} Y \xrightarrow{f} X$ lies in Δ.
(LT2): For a given left triangle $\Omega X \xrightarrow{h} Z \xrightarrow{g} Y \xrightarrow{f} X$ in Δ, the left triangle $\Omega F \xrightarrow{-\Omega f} \Omega X \xrightarrow{h} Z \xrightarrow{g} Y$ also belongs to Δ.
(LT3): For any commutative diagram of the form

$$\begin{array}{ccccccc} \Omega X & \xrightarrow{h} & Z & \xrightarrow{g} & Y & \xrightarrow{f} & X \\ {\scriptstyle\Omega\gamma}\downarrow & & & & {\scriptstyle\beta}\downarrow & & {\scriptstyle\gamma}\downarrow \\ \Omega X' & \xrightarrow{h'} & Z' & \xrightarrow{g'} & Y' & \xrightarrow{f'} & X', \end{array}$$

where the rows are left triangles, there is a morphism $\alpha : Z \longrightarrow Z'$ in $\mathcal{C}$, making the completed diagram commutative.
(LT4): For any two left triangles $\Omega X \xrightarrow{m} Z' \xrightarrow{l} Y \xrightarrow{k} X$ and $\Omega Y \xrightarrow{h} X' \xrightarrow{g} Z \xrightarrow{f} Y$ in Δ, there exist a third left triangle $\Omega X \xrightarrow{j} Y' \xrightarrow{i} Z \xrightarrow{kof} X$ in Δ and two morphisms $\alpha : X' \longrightarrow Y'$ and $\beta : Y' \longrightarrow Z'$ in $\mathcal{C}$, such that the following diagram is commutative, where the second column from the left is a left triangle in Δ

$$\begin{tikzcd}
 & \Omega Z' \arrow[d, "ho\Omega l"] & & \\
\Omega Y \arrow[r, "h"] \arrow[d, "\Omega k"] & X' \arrow[r, "g"] \arrow[d, "\alpha"] & Z \arrow[r, "f"] \arrow[d, equal] & Y \arrow[d, "k"] \\
\Omega X \arrow[r, "j"] \arrow[d, equal] & Y' \arrow[r, "i"] \arrow[d, "\beta"] & Z \arrow[r, "kof"] \arrow[d, "f"] & X \arrow[d, equal] \\
\Omega X \arrow[r, "m"] & Z' \arrow[r, "l"] & Y \arrow[r, "k"] & X
\end{tikzcd}$$

Remark 2.2. SPANIER-WHITEHEAD CATEGORY. Assume that $(\mathcal{C}, \Omega, \Delta)$ is a left triangulated category. The Spanier-Whitehead category of $\mathcal{C}$, $SW(\mathcal{C}, \Omega)$ is defined as follows: the objects have the form $X[n]$ where X is an object of $\mathcal{C}$ and $n \in \mathbb{Z}$. Moreover, for any two objects $X[n], Y[m]$ in $SW(\mathcal{C}, \Omega)$, their Hom-set is defined by

$$\mathsf{Hom}_{SW}(X[n], Y[m]) = \varinjlim_{i \geq n,m} \mathsf{Hom}_{\mathcal{C}}(\Omega^{i-n} X, \Omega^{i-m} Y).$$

For the basic properties of Spanier-Whitehead categories, the reader is refereed to [5]. In the remainder, we write $SW(\mathcal{C})$ for simplicity instead of $SW(\mathcal{C}, \Omega)$.

Remark 2.3. Let $\mathcal{X}$ be a resolving subcategory and let $\underline{\mathcal{X}}$ denote the stable category of $\mathcal{X}$ modulo projective modules. Recall that for an R-module X, its first syzygy ΩX is the kernel of its projective (pre)cover $\pi_X : P_X \longrightarrow X$. Now fix the short exact sequence $0 \longrightarrow \Omega X \xrightarrow{\lambda_X} P_X \xrightarrow{\pi_X} X \longrightarrow 0$. This gives rise to the syzygy functor $\Omega : \underline{\mathcal{X}} \longrightarrow \underline{\mathcal{X}}$. Take an R-homomorphism $f \in \mathsf{Hom}_{\mathcal{X}}(X, Y)$ and consider the following pullback diagram;

$$\begin{array}{ccccccccc}
0 & \longrightarrow & \Omega Y & \xrightarrow{\alpha(f)} & C_f & \xrightarrow{\beta(f)} & X & \longrightarrow & 0 \\
 & & {\scriptstyle id}\downarrow & & {\scriptstyle \theta_f}\downarrow & & {\scriptstyle f}\downarrow & & \\
0 & \longrightarrow & \Omega Y & \xrightarrow{\lambda_Y} & P_Y & \xrightarrow{\pi_Y} & Y & \longrightarrow & 0.
\end{array}$$

In particular, we have the sequence $\Omega Y \xrightarrow{\bar{\alpha(f)}} C_f \xrightarrow{\bar{\beta(f)}} X \xrightarrow{\bar{f}} Y$ in $\underline{\mathcal{X}}$ which is called left triangle. Put all sequences which are equivalent to one of this from in Δ. In view of [2, Theorem 3.1], $(\underline{\mathcal{X}}, \Omega, \Delta)$ is a left triangulated category. Therefore, as we have mentioned above, $SW(\underline{\mathcal{X}})$ is a triangulated category.

Assume that $(R, \mathfrak{m})$ is a (commutative) Cohen–Macaulay local ring. Denote the category of all maximal Cohen–Macaulay by $\mathsf{CM}(R)$. Since $\mathsf{CM}(R)$ is a resolving subcategory of $\mathsf{mod}R$, $(\underline{\mathsf{CM}}(R), \Omega)$ is a left triangulated category and, in particular, $SW(\underline{\mathsf{CM}}(R))$ is a triangulated category.
Now we are ready to state our results.

Proposition 2.4. *Let R be a d-dimensional commutative Cohen–Macaulay local ring. Then there is a triangle equivalence $SW(\underline{\mathsf{CM}}(R)) \cong SW(\underline{\mathsf{mod}}R)$.*

A result due to Beligiannis [1, Theorem 3.8] asserts that there is a triangle equivalence $SW(\mathsf{mod}R) \cong \mathsf{D}_{\mathsf{sg}}(R)$. This fact, in conjunction with Proposition 2.4 yields the main result of this paper.

Theorem 2.5. *Let R be a commutative Cohen–Macaulay local ring. Then there is a triangle equivalence $SW(\underline{\mathsf{CM}}(R)) \cong \mathsf{D}_{\mathsf{sg}}(R)$.*

It is known that the category of finitely generated Gorenstein projective R-modules, $\mathsf{Gp}(R)$, is a is a Frobenius category , and so, its stable category is triangulated. Thus one may apply [1, Corollary 3.9.(3)] and get a triangle equivalence $\underline{\mathsf{Gp}}(R) \cong SW(\underline{\mathsf{Gp}}(R))$. This leads us to recover a fundamental result of Buchweitz and Happel, as follows.

Corollary 2.6. *Let R be a commutative Gorenstein local ring. Then there is a triangle equivalence $\underline{\mathsf{Gp}}(R) \cong \mathsf{D}_{\mathsf{sg}}(R)$.*

References

[1] A. Beligiannis. The homological theory of contravariantly finite subcategories: Auslander-buchweitz contexts, gorenstein categories and (co)stabilization. *Comm. Algebra*, 10:4547–4596, 2000.
[2] A. Beligiannis and N. Marmaridis. Left triangulated categories arising from contravariantly finite subcategories. *Communication is algebra*, 12:5021–5036, 1994.
[3] R.-O. Buchweitz. Maximal Cohen-Macaulay modules and Tate-cohomology over Gorenstein rings. unpublished paper, 1987, available from http://hdl.handle.net/1807/16682.
[4] D. Happel. On gorenstein algebras, progress in math. *Birkh ä user, Verlag, Basel*, (95):389–404, 1991.
[5] A. Heller. Stable homotopy categories. *Bulletin of A. M. S.*, (74):28–63, 1968.
[6] D. Orlov. Triangulated categories of singularities and D-branes in Landau-Ginzburg models. *(Russian) Tr. Mat. Inst. Steklova*, (246):28–63, 2004.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–179, pp. 370–371

Pure semisimplity of the Spanier-Whitehead category of modules over Cohen-Macaulay rings

Abdolnaser Bahlekeh[a,*]

[a] *Gonbad-Kavous University, Gonbad-Kavous, Iran*

Article Info

Keywords:
Spanier-Whitehead category
Cohen-Macaulay rings
pure-semisimple

2020 MSC:
13C60
13D05

Abstract

Let $(R, \mathfrak{m})$ be a commutative noetherian Cohen-Macaulay ring of finite Cohen-Macaulay type. The aim of this paper is to show that the Spanier-Whitehead of R-modules is pure-semisimple.That is, every object of this category is a direct sum of finitely generated ones.

1. Introduction

The pure-semisimple conjecture asserts that every left pure-semisimple ring (a ring over which every left module is a direct sum of finitely generated ones) is of finite representation type. A result due to Chase [4, Theorem 4.4] yields that left pure-semisimple rings are left artinian. A famous result of Auslander [1, 2] implies the validity of the pure-semisimple conjecture for artin algebras. Indeed, he has proved that an artin algebra Λ is of finite representation type if and only if every left Λ-module is a direct sum of finitely generated modules. Recall that an associative ring is said to be of finite representation type, if the set of isomorphism classes of indecomposable finitely generated modules is finite. Motivaited by the Auslander's result, Beligiannis investigated decomposability of Gorebstein projective modules. Precisely, in a successful attempt, he showed that a virtually Gorenstein algebra Λ is of finite Cohen-Macaulay type, in the sense that there are only finitely many isomorphism classes of indecomposable finitely generated Gorenstein projective Λ-modules if and only if any left Gorenstein projective Λ-module is a direct sum of finitely generated ones, see [3, Theorem 4.10]. This result has been proved by Chen [5] for Gorenstein artin algebras.In this paper, we make progress twords examining the pure-semisimplicity of modules over Cohen-Macaulay rings. To be more precise, assume that $(R, \mathfrak{m})$ is a commutative noetherian Cohen-Macaulay local ring with a cononical module. As a main result, we show that if R is of finite Cohen-Macaulay type, then Spanier-Whitehead category of R-modules, $SW(\mathsf{Mod}R)$, is pure-semisimple.

---

*Talker
Email address: first.author@gmail.com (Abdolnaser Bahlekeh)

2. Results

Let us begin this section with our convention.

Remark 2.1. Throughout the paper, $(R, \mathfrak{m})$ is a commutative Cohen-Macaulay local ring with a canonical module ω. The category of all R-modules will be denoted by $\mathsf{Mod}R$. By $\mathcal{X}_\omega$ we mean a subcategory of $\mathsf{Mod}R$ consisting of all modules M such that there is an exact sequence of R-modules;

$$0 \longrightarrow M \longrightarrow w_0 \xrightarrow{d_0} w_1 \xrightarrow{d_1} \cdots \xrightarrow{d_{i-1}} w_i \xrightarrow{d_i} \cdots ,$$

with $w_i \in \mathsf{Add}\omega$. Here $\mathsf{Add}\omega$) stands for the full subcategory of $\mathsf{Mod}R$ consisting of all modules isomorphic to direct summands of direct sums (resp. finite direct sums) of copies of ω.

Remark 2.2. It is evident that maximal Cohen–Macaulay modules are exactly those finitely generated modules in $\mathcal{X}_\omega$. Moreover, it is known that over a d-dimensional Cohen–Macaulay ring, dth syzygy of any finitely generated module is a maximal Cohen–Macaulay module. The result below proves the validity of this result for not necessarily finitely generated modules.

Proposition 2.3. *Let $(R, \mathfrak{m})$ be a d-dimensional commutative Cohen–Macaulay ring with a canonical module ω. Then for any R-module M, $\square^d M \in \mathcal{X}_\omega$.*

Recall that a given R-module M is said to be *fully decomposable*, provided that it is a direct sum of finitely generated modules.

Proposition 2.4. *Let $(R, \mathfrak{m})$ be a Cohen–Macaulay local ring which is of finite* CM*-type. Then for a given object $M \in \mathcal{X}_\omega$, $\square M$ is fully decomposable.*

The result below is a direct consequence of the two previous results.

Corollary 2.5. *Let $(R, \mathfrak{m})$ be a d-dimensional Cohen–Macaulay local ring of finite Cohen-Macaulay type. Then $\square^{d+1}\mathsf{Mod}R$ is pure-semisimple.*

Remark 2.6. Spanier-Whitehead category. Assume that $(\mathcal{C}, \Omega, \Delta)$ is a left triangulated category. The Spanier-Whitehead category of $\mathcal{C}$, $SW(\mathcal{C})$ is defined as follows: the objects have the form $X[n]$ where X is an object of $\mathcal{C}$ and $n \in \mathbb{Z}$. Moreover, for any two objects $X[n], Y[m]$ in $SW(\mathcal{C}, \Omega)$, their Hom-set is defined by

$$\mathrm{Hom}_{SW}(X[n], Y[m]) = \varinjlim_{i \geq n,m} \mathrm{Hom}_{\mathcal{C}}(\Omega^{i-n}X, \Omega^{i-m}Y).$$

A given element $f \in \mathrm{Hom}_{SW}(X[n], Y[m])$ is said to have an i-th representative $f_i : \Omega^{i-n}X \longrightarrow \Omega^{i-m}Y$, provided that the canonical image f_i by $\lambda_i = \lambda_i(X[n], Y[m]) : \mathrm{Hom}_{\mathcal{C}}(\Omega^{i-n}X, \Omega^{i-m}Y) \longrightarrow \mathrm{Hom}_{SW}(X[n], Y[m])$ equals f.

Theorem 2.7. *Let $(R, \mathfrak{m})$ be a Cohen–Macaulay local ring of finite* CM*-type. Then $SW(\mathsf{Mod}R)$ is pure-semisimple.*

References

[1] M. Auslander. Representation theory of artin algebras. ii. *Comm. Algebra*, 1:269–310, 1974.
[2] M. Auslander. A functorial approach to representation theory, in representations of algebra. *Workshop Notes of the Third Inter. Confer., Lecture Notes Math.*, 944:105–179, 1982. Springer-Verlag.
[3] A. Beligiannis. On algebras of finite cohen-macaulay type. *Adv. Math.*, 226(2):1973–2019, 2011.
[4] S. U. Chase. Direct product of modules. *Trans. Amer. Math. Soc.*, 97:457–473, 1960.
[5] X. W. Chen. An auslander-type result for gorenstein projective modules. *Adv. Math.*, 218:2043–2050, 2008.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–182, pp. 372–374

On Polynilpotent Covering Lie Algebra of a Polynilpotent Lie Algebra

M. Alizadeh Sanati[a,*]

[a]*Department of Mathematics, Faculty of Sciences, Golestan University, Gorgan 79142-49361, Iran.*

Article Info

Keywords:
Polynilpotent multiplier
polynilpotent Lie algebra
$(c_1, \ldots, c_t)$-cover

2020 MSC:
17B05
17B99

Abstract

Let $\mathcal{N}_{c_1,\ldots,c_t}$ be the variety of polynilpotent Lie algebras of class row $(c_1, \ldots, c_t)$. In this paper, we show that a polynilpotent Lie algebra L of class row $(c_1, \ldots, c_t)$ has no any $\mathcal{N}_{c_1,\ldots,c_t,c_{t+1}}$-covering lie algebra if its Baer-invariant with respect to the variety $\mathcal{N}_{c_1,\ldots,c_t,c_{t+1}}$ is nontrivial. As an immediate consequence, we can conclude that a solvable Lie algebra L of length c with nontrivial solvable multiplier, $\mathcal{S}_n M(L)$, has no $\mathcal{S}_n$-covering Lie algebra for all $n > c$, where $\mathcal{S}_n$ is the variety of solvable Lie algebras of length at most n.

1. Introduction and preliminaries

Definition 1.1. Let F be a field. A Lie algebra over F is an F-vector space L, together with a bilinear map, the Lie bracket $L \times L \to L, (x, y) \mapsto [x, y]$, for all $x, y, z \in L$ satisfying the following properties:
$(L1)$ $[x, x] = 0$,
$(L2)$ *(Jacobi identity)* $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$.

Suppose that I and J are ideals of a Lie algebra L. Then it is defined $[I, J]$ as the ideal generated by $\{[x, y] : x \in I, y \in J\}$. For each positive integers c, it is defined $(I, J)_c$ inductively as $(I, J)_0 = I$ and $(I, J)_{c+1} = [(I, J)_c, J]$. The ideal $L' := [L, L]$ is called *derived algebra* of L. The *lower central series* of L is the series $L \supseteq L_1 \supseteq L_2 \supseteq \cdots$ with terms $L_1 = L'$ and $L_{c+1} = [L, L_c]$ for $c \geq 2$. Also, The *upperer central series* of L is the series $\{0\} \subseteq Z_1(L) := Z(L) \subseteq Z_2(L) \subseteq \cdots$ in which $Z(L) = \{x \in L | [x, y] = 0,\ \forall y \in L\}$ and $Z_{c+1}(L)/Z_c(L) = Z\left(L/Z_c(L)\right)$ for $c \geq 2$.
Let $t \geq 2$ and $c_1, \ldots, c_t$ be arbitrary positive integers. By notation of [1], we denote recursively $L_{c_1,\ldots,c_t} = \left(L_{c_1,\ldots,c_{t-1}}\right)_{c_t}$, and $Z_{c_1,\ldots,c_t}(L)$ as numrator of the c_1-th center of $L/Z_{c_2,\ldots,c_t}(L)$. Also, we set

$$(I, J)_{c_1,\ldots,c_t} = \left[(I, J)_{c_1,\ldots,c_{t-1}}, {}_{c_t} J_{c_1,\ldots,c_{t-1}}\right].$$

It is obvious that $L_{c_1,\ldots,c_t} = (L, L)_{c_1,\ldots,c_t}$ and also,

*Talker
Email address: m.alizadeh@gu.ac.ir (M. Alizadeh Sanati)

$$(I, L)_{c_1,\dots,c_t} = 0 \text{ if and only if } I \subseteq Z_{c_1,\dots,c_t}(L).$$

A Lie algebra is called *polynilpotent Lie algebra of class row* $(c_1, \dots, c_t)$, even $L \in \mathcal{N}_{c_1,\dots,c_t}$ or equivalentely $L_{c_1,\dots,c_t} = 0$.
The following key-lemma in [1] is needed for proving our main result.

Lemma 1.2 ([1]). *Let I and J be ideals of a Lie algebra L. Then*
(i) For all $c > 1$, $[I, J_c] = (I, J)_{c+1}$.
(ii) For all $t > 1$, $[I, J_{c_1,\dots,c_t}] \subseteq (I, J)_{c_1,\dots,c_t}$.

Definition 1.3. Let $0 \to R \to F \to L \to 0$ be a free presentation for L. Then the *polynilpotent multiplier* of L with respect to the variety $\mathcal{N}_{c_1,\dots,c_t}$, denoted by $\mathcal{M}^{c_1,\dots,c_t}(L)$, is defined to be

$$\mathcal{M}^{c_1,\dots,c_t}(L) = \frac{R \cap F_{c_1,\dots,c_t}}{(R, F)_{c_1,\dots,c_t}}.$$

It is always abelian and independent of the choice of the free presentation of L, [6]. If $t = 1$, the Lie algebra $\mathcal{M}^{c_1}(L) = (R \cap F_{c_1})/(R, F)_{c_1}$ is c_1-nilpotent multiplier of L and $\mathcal{M}^{(1)}(L) = \mathcal{M}(L)$ is the well-known Schur multiplier of L.

Definition 1.4. An $\mathcal{N}_{c_1,\dots,c_t}$*-covering Lie algebra* of L is a Lie algebra L^* with an ideal M such that
(i) $L \simeq L^*/M$,
(ii) $M \simeq \mathcal{M}^{c_1,\dots,c_t}(L^*)$ and
(iii) $M \subseteq Z_{c_1,\dots,c_t}(L^*) \cap L^*_{c_1,\dots,c_t}$.
We also call L^* a $(c_1, \dots, c_t)$-polynilpotent covering Lie algebra of L.

In 1994, Moneyhun proved that all covers of finite dimensional Lie algebras are isomorphic. Also, it is a well-known fact that every finite dimention Lie algebra has at least a covering Lie algebra (see [2], [3] and [5]). But it is an open problem whether the cover for an arbitrary Lie algebra exists. As a result of [4], every nilpotent Lie algebras of class at most n with non-trivial c-nilpotent multiplier does not have any c-covering whenever $c > n$.
Now, in this paper, we concentrate on nonexistence of polynilpotent covering Lie algebra of some polynilpotent Lie algebra.

2. Nonexistence of polynilpotent covering

Let L be a Lie algebra and $\mathcal{V}$ a variety of Lie algebras. It is clear, by definition, that if $\mathcal{V}M(L) = 1$, then L is the only $\mathcal{V}$-covering Lie algebra of itself. So it is natural to put the condition $\mathcal{V}M(L) \neq 1$ for nonexistence of $\mathcal{V}$-covering Lie algebra L.

Theorem 2.1. *If L is a polynilpotent Lie algebra of class row $(c_1, \dots, c_t)$ such that $\mathcal{M}^{c_1,\dots,c_t,c_{t+1}}(L) \neq 0$, then L has no $(c_1, \dots, c_t, c_{t+1})$-polynilpotent covering Lie algebra of L.*

Proof. Suppose on the contrary that L^* is a $(c_1, \dots, c_{t+1})$-polynilpotent covering Lie algebra of L. There exists ideal M of L^* such that

$$L \simeq L^*/M,\ M \simeq \mathcal{M}^{c_1,\dots,c_{t+1}}(L^*) \text{ and } M \subseteq Z_{c_1,\dots,c_{t+1}}(L^*) \cap L^*_{c_1,\dots,c_{t+1}}.$$

Since L is polynilpotent of class row $(c_1, \dots, c_t)$, $L_{c_1,\dots,c_t} = 1$. By property (i) of cover, we conclude $(L^*/M)_{c_1,\dots,c_t} = 1$ and so $L^*_{c_1,\dots,c_t} \subseteq M$. On the other hand by property (iii), $M \subseteq L^*_{c_1,\dots,c_t,c_{t+1}}$. Hence $L^*_{c_1,\dots,c_t} \subseteq L^*_{c_1,\dots,c_t,c_{t+1}}$. Clearly $L^*_{c_1,\dots,c_t,c_{t+1}} = \left(L^*_{c_1,\dots,c_t}\right)_{c_{t+1}} \subseteq L^*_{c_1,\dots,c_t}$, so $L^*_{c_1,\dots,c_t,c_{t+1}} = L^*_{c_1,\dots,c_t}$. It is concluded that

$$L^*_{c_1,\dots,c_t,c_{t+1}} = L^*_{c_1,\dots,c_t,c_{t+1}-1} = \cdots = L^*_{c_1,\dots,c_t,1} = L^*_{c_1,\dots,c_t}. \qquad (*)$$

On the other hand $M \subseteq Z_{c_1,\dots,c_{t+1}}(L^*)$ imples $L^*_{c_1,\dots,c_t} \subseteq Z_{c_1,\dots,c_{t+1}}(L^*)$ and thus $(L^*_{c_1,\dots,c_t}, L^*)_{c_1,\dots,c_t,c_{t+1}} = 0$. Now, by Lemma 1.2 (ii) we can write $[L^*_{c_1,\dots,c_t}, L^*_{c_1,\dots,c_t,c_{t+1}}] = 0$. The relation $(*)$ implies $[L^*_{c_1,\dots,c_t}, L^*_{c_1,\dots,c_t}]$ is vanished. It means $\left(L^*_{c_1,\dots,c_t}\right)' = 0$ or $L^*_{c_1,\dots,c_t,1} = 0$ and this by $(*)$ implies $L^*_{c_1,\dots,c_t,c_{t+1}} = 0$ and so $M = 0$. This is contradiction. $\square$

Note that, $\mathcal{S}_n$, the variety of solvable Lie algebra of derived length at most n is in fact the variety of polynilpotent Lie algebra of class row $\underbrace{(1, \ldots, 1)}_{n-times}$. Hence the following result is a consequence of Theorem 2.1.

Corollary 2.2. *Let L be a solvable Lie algebra with derived length at most c. If the l-solvable multiplier of L, $\mathcal{S}_n M(L)$, is nontrivial, then L has no any $\mathcal{S}_n$-covering Lie algebra, for all $n > c$.*

References

[1] A. Aslizadeh, A. R. Salemkar, Z. Riyahi *Polynilpotent capability of Lie algebras* , Commun. Algebra, 74(19) (2019), 1390–1400.
[2] P. Batten and E. Stitzinger, *On covers of Lie algebras*, Commun. Algebra 24 (1996), 4301–4317.
[3] K. Moneyhun, *Isoclinisms in Lie algebras*, Algebras Groups Geom. 11 (1994), 9–22.
[4] H. Safa and H. Arabyani, *On c-nilpotent multiplier and c-covers of a pair of Lie algebras*, Commun. Algebra 45(10) (2017), 4429–434.
[5] A. R. Salemkar, V. Alamian and H. Mohammadzadeh, *Some properties of the Schur multiplier and covers of Lie algebras*, Commun. Algebra 36 (2008), 697–707.
[6] A. R. Salemkar and Z. Riyahi, *The polynilpotent multiplier of Lie algebras*, J. Algebra Appl. 12(02): (2013) 1250154.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–189, pp. 375–378

Remark's on Applications of Nakayama Lemma in the Categroy **Act**-S

Kamal Ahmadi[a,*], Seyed Mohammad Ajdani[b], Asghar Madadi[c]

[a] *Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*
[b] *Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*
[c] *Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*

Article Info

Keywords:
monoids
S-act
Nakayama's lemma

2020 MSC:
20M30, 20M50
08A60, 08B25

Abstract

Let S be a monoid. In this talk we study Nakayama's Lemma on **Act**-S. We also include some consequences which are useful in other settings. As an application of Nakayama's Lemma we prove Krull intersection theorem for S-acts.

1. Introduction and Preliminaries

Let S be a monoid with identity 1. Recall that a *(right) S-act* is a non-empty set A equipped with a map $\mu : A \times S \to A$ called its action, such that, denoting $\mu(a, s)$ by as, we have $a1 = a$ and $a(st) = (as)t$, for all $a \in A$, and $s, t \in S$. An element $\theta \in A$ is called a *zero element* of A if $\theta s = \theta$ for every $s \in S$. Let A be an S-act and $B \subseteq A$ a non-empty subset. Then B is called a *subact* of A if $bs \in B$ for all $s \in S$ and $b \in B$. In particular, if I is a (proper) ideal of S, then

$$AI := \{as \mid a \in A, s \in I\} \tag{1}$$

is a subact of A. An equivalence relation ρ on an S-act A is called a *congruence* on A if $a\rho a'$ implies $(as)\rho(a's)$ for $a, a' \in A$ and $s \in S$. Any subact $B \subseteq A$ defines the *Rees congruence* ρ_B on A, by setting $a\rho_B a'$ if $a, a' \in B$ or $a = a'$. We denote the resulting factor act by A/B and call it the *Rees factor* act of A by the subact B. Clearly, A/B has a zero which is the class consisting of B, all other classes are one-element sets. Moreover, any subact $B \subseteq A$ gives rise to a kernel congruence $\ker \pi$ where $\pi : A \to A/B$ is the canonical epimorphism. The category of all S-acts, with action-preserving (S-act) maps ($f : A \to B$ with $f(as) = f(a)s$, for $s \in S$, $a \in A$), is denoted by **Act**-S. Clearly S itself is an S-act with its operation as the action.

*Talker
Email addresses: kamal.ahmadi.math@gmail.com (Kamal Ahmadi), majdani2@yahoo.com, (Seyed Mohammad Ajdani), amadadi2009@gmail.com (Asghar Madadi)

Throughout this paper, S is a monoid with at least a right non-invertible element, all S-acts will be right S-acts and all ideals of S are right ideals, zero element of an S-act, if it exists, is unique. If S-act A has a unique zero element θ, then $\theta \in B$ for any subact B of A. The set of all idempotents of S is denoted by $E(S)$. It is known that the set

$$\{s \mid s \text{ is a right non-invertible element of } S\} \tag{2}$$

is the only maximal right ideal of S. In this note, we reserve $\mathfrak{M}$ to denote, always, this unique maximal right ideal of S. For more information on S-acts we refer the reader to [9]. Nakayama's Lemma was first discovered in the special case of ideals in a commutative ring by W. Krull and then in general by G. Azumaya [5]. The lemma is named after the Japanese mathematician T. Nakayama and introduced in its present form in [10]. D. D. Anderson and E. W. Johnson have proved some versions of Nakayama's Lemma for lattices and commutative monoids with a unique zero element in [2] and [1], respectively. Some generalizations of Nakayama's Lemma have been given and studied, in the literatures. For example, A. Azizi [4] introduced Nakayama property for modules over a commutative ring with identity. He says that an R-module M has Nakayama property if $IM = M$, where I is an ideal of R, implies that there exists $a \in R$ such that $aM = 0$ and $a - 1 \in I$. Then Nakayama's Lemma states that every finitely generated R-module has Nakayama property. He has proved that R is a perfect ring if and only if every R-module has Nakayama property. Besides, we remark that there are generalizations in other contexts, we refer the reader to [6] and [11]. It is a significant tool in algebraic geometry, because it allows local data on algebraic varieties, in the form of modules over local rings, to be studied pointwise as vector spaces over the residue field of the ring. Nakayama's Lemma for R-modules governs the interaction between the Jacobson radical of a ring and its finitely generated modules. There are several equivalent forms of Nakayama's Lemma in algebra. We express one here. Let R be a ring with identity 1, and A a finitely generated right R-module. If I is a right ideal of R contained in the Jacobson radical of R, $J(R)$, and $AI = A$ then $A = 0$.

$$\mathfrak{M} = \{s \in S \mid st \neq 1 \text{ for all } t \in S\} \tag{3}$$

is the only maximal right ideal of S and then for any proper ideal I of S we have $I \subseteq \mathfrak{M}$. We can not talk about Jacobson radical of a monoid, because $\mathfrak{M}$ is the only maximal right ideal of it. We therefore consider Nakayama's Lemma in **Act**-S where S is a monoid with a unique two-sided maximal ideal $\mathfrak{M}$. In [9, Example 3.18.10], the authors present a monoid S in which $\mathfrak{M}$ is not a two-sided ideal. Moreover, there are other examples of monoids S that are not commutative, but their maximal ideals are two-sided. For example, given $S = (M_n(\mathbb{R}), \cdot)$, the monoid of all $n \times n$ matrices with real number entries under usual multiplication of matrices. Since $ab = 1$ implies $ba = 1$ for $a, b \in S$, the unique maximal ideal of S is two-sided (see Lemma 2.1 below). Besides, there are many examples of finitely generated S-acts A with a zero element θ in which for a proper ideal I of S, $AI = A$, but $A \neq \{\theta\}$; take any monoid S and an arbitrary finite set A with $|A| > 1$. Then A becomes a right S-act by trivial action, i.e., $as = a$ for all $a \in A, s \in S$. Therefore, $AI = A$ for every proper ideal I, although, $A \neq \{\theta\}$. Finally, as an application of Nakayama's Lemma we prove Krull intersection theorem for S-acts.

2. Some Forms of Nakayama's Lemma

In this section first we present a lemma which determines monoids in which their unique maximal ideals are two-sided.

Lemma 2.1. *Let S be a monoid. Then the following statements are equivalent:*
1) $\mathfrak{M}$ *is a two-sided ideal of* S.
2) $st = 1$ *implies* $ts = 1$, *for all* $s, t \in S$.

The next lemma guarantees that every finitely generated S-act with a unique zero element has a maximal subact.

Lemma 2.2. *Let S be a monoid and let A be a finitely generated S-act with a unique zero element θ and $A \neq \{\theta\}$. Then every proper subact of A is contained in a maximal subact. In particular, A has a maximal subact.*

Now we are ready to state the first version of Nakayama's lemma.

Theorem 2.3. *Let S be a monoid in which its unique maximal right ideal $\mathfrak{M}$ is two-sided. Moreover, let A be an S-act and B a maximal subact of A in which there exists $a \in A \setminus B$ such that $\mathfrak{M} = \{s \in S \mid as \in B\}$. Then for every proper ideal I of S we have $AI \neq A$.*

The next example illustrates Theorem 2.3.

Example 2.4. Let $S = (\mathbb{N}, \cdot)$ be the monoid of natural numbers with the usual multiplication. Then $A = \{2, 3, \cdots\}$ is a subact of $\mathbb{N}$. Let U be the set of all prime numbers. Then U is a set of generating elements of A. Note that U is the least generating set of A, i.e. A is not finitely generated. The set $B = \{3, 4, \cdots\}$ is a maximal subact of A in which $\mathfrak{M} = \{n \in \mathbb{N} \mid 2n \in B\}$. All assumptions of Theorem 2.3 hold for A. Therefore for any proper ideal I of S, $AI \neq A$.

In the next lemma we will see that the second condition of Theorem 2.3 is equivalence to the implication that '$as = a$ implies $s \notin \mathfrak{M}$'. More precisely, we have:

Lemma 2.5. *Let A be an S-act and B a maximal subact of A. Then for every $a \in A \setminus B$ the following statements are equivalent:*
1) $\mathfrak{M} = \{s \in S \mid as \in B\}$.
2) $as = a$ *implies* $s \notin \mathfrak{M}$.

Definition 2.6. Let A be an S-act. A nonzero element $a \in A$ is called *quasi-strongly faithful*, if for $s \in S$ the equality $as = a$ implies that $s \notin \mathfrak{M}$. One calls A a *quasi-strongly faithful S-act* if all of its nonzero elements are quasi-strongly faithful.

Proposition 2.7. *Let S be a monoid in which its unique maximal right ideal $\mathfrak{M}$ is two-sided.*
(i) *Let A be an S-act and B a maximal subact of A such that there exists a quasi-strongly faithful element in $A \backslash B$. Then:*
a) *$AI = A$ if and only if $I = S$;*
b) *For every ideal I of S with $I^2 = I$, $AI \cong A$ if and only if $I = S$.*
(ii) *Let A be a quasi-strongly faithful S-act and B a maximal subact of A. Then $AI = A$ if and only if $I = S$.*

Now we state a second version of Nakayama's Lemma which is quite similar to module theory. Also, in the next corollary we give another similar consequence.

Theorem 2.8. *Let S be a monoid in which its unique maximal right ideal $\mathfrak{M}$ is two-sided. Let A be a finitely generated quasi-strongly faithful S-act with a unique zero element θ. If $AI = A$ for some proper ideal I of S, then $A = \{\theta\}$.*

Corollary 2.9. Let S be a monoid in which its unique maximal right ideal $\mathfrak{M}$ is two-sided. Let A be a finitely generated quasi-strongly faithful S-act. If $B \cup AI = A$ for some proper ideal I of S and some subact B of A, then $A = B$.

3. Krull's Intersection Theorem

Let I be an ideal of a commutative Noetherian ring R such that $I \subseteq J(R)$. Then $\bigcap_{n \in \mathbb{N}} I^n = 0$. This is known as Krull's Intersection Theorem in the theory of modules over commutative rings. (see [12, Corollary 8.25])
In the following we prove a counterpart of this result for S-acts. First we need the next theorem. Recall [13] that an S-act A is *Noetherian* if A satisfies the ascending chain condition on its subacts, that is, every ascending chain of subacts of A is finite.

Lemma 3.1. *Let S be a monoid and A, B be two S-acts.*
1) *$A \dot{\cup} B$ is Noetherian if and only if A and B themselves are;*
2) *if B is a subact of A, then A is Noetherian if and only if B and A/B themselves are;*
3) *if ρ is a congruence on Noetherian S-act A, then A/ρ is Noetherian;*
4) *if S is a Noetherian monoid and A is a finitely generated S-act then A is Noetherian.*

Theorem 3.2. *Let S be a commutative Noetherian monoid, I an ideal of S and A a finitely generated S-act. If $B = \bigcap\limits_{n \in \mathbb{N}} AI^n$, then $BI = B$.*

Corollary 3.3. (**Krull Intersection Theorem**) *Suppose that S is a commutative Noetherian monoid. Let A be a finitely generated quasi-strongly faithful S-act with a unique zero element θ. Then $\bigcap\limits_{n \in \mathbb{N}} AI^n = \{\theta\}$ for every proper ideal I of S.*

Corollary 3.4. *Let S be a commutative Noetherian monoid with a unique zero element θ. Then:*
1) *S as an S-act is quasi-strongly faithful if and only if $\bigcap_{n\in\mathbb{N}} \mathfrak{M}^n = \{\theta\}$;*
2) *if S as an S-act is quasi-strongly faithful, then $\bigcap_{n\in\mathbb{N}} I^n = \{\theta\}$ for every proper ideal I of S.*

In the following proposition we will see the converse of Corollary 3.4, of course with some additional conditions.

Proposition 3.5. *Let S be a commutative monoid in which as an S-act is quasi-strongly faithful and has a unique zero element θ. Then S is Noetherian if and only if $\bigcap_{n\in\mathbb{N}} \mathfrak{M}^n = \{\theta\}$ and $\mathfrak{M}$ is a finitely generated ideal.*

References

[1] D. D. Anderson, E. W. Johnson, Ideal theory in commutative semigroups, Semigroup Forum, 30: 127-158 (1984).
[2] D. D. Anderson, E. W. Johnson, Join-principally generated multiplicative lattices, Algebra Univers., 19: 74-82 (1984).
[3] F. W. Anderson, K. R. Fuller, Rings and Categories of Modules, Second edition, Gradugate Texts in Mathematics, Springer-Verlag, New York, (1992).
[4] A. Azizi, On Generalization of Nakayama's lemma, Glasg. Math. J., 52: 605-617 (2010).
[5] G. Azumaya, On maximally central algebras, Nagoya Math. J., 2: 119-150 (1951).
[6] P. Balister, S. Howson. Note on Nakayama's lemma for compact Λ-Modules, Asian J. Math., 1: 224-229 (1997).
[7] P. M. Cohen, Further Algebra and Applications, Springer-Verlag, (2003).
[8] D. Eisenbud, Commutative Algebra with a view toward Algebraic Geometry, Graduate Texts in Mathemtaics, Springer-Verlag, (1994).
[9] M. Kilp, U. Knauer, A. V. Mikhalev, Monoids, Act and Categories, Walter de Grupter, Berlin (2000).
[10] T. Nakayama, A remark on finitely generated modules, Nagoya Math. J., 3: 139-140 (1951).
[11] A. Ogus, G. Bergman, Nakayama's lemma for half-exact functors, Proc. Amer. Math. Soc, 31: 67-74 (1972).
[12] R. Y. Sharp, Steps in Commutative Algebra, Second Edition, Cambridge University Press (2000).
[13] X. Zhang, Y. Chen, Y. Wang, Chain conditions on essential subacts, Northeast Math. J., 22(3): 357-369 (2006).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–196, pp. 379–380

On noncyclic division algebras

Mehran Motiee[a,*], Behnam Talaee[a]

[a]*Faculty of Basic Sciences, Babol Noshirvani University of Technology, Babol, Iran*

Article Info

Keywords:
Division algebra
Cyclic algebra
Valuation

2020 MSC:
11R52
16W60

Abstract

Let F be a field and D be a central F-division algebra. D is called cyclic if it contains a maximal subfield K such that K/F is a cyclic Galois extension. We also say that D is cyclically split if D has a cyclic splitting filed. In this note we give some conditions under which D is not cyclically split.

1. Introduction

Let F be a field and D be a central F-division algebra, i.e., D is a division algebra whose center is F. Recall that the dimension of D over F, as a vector space, is always a square [3, p. 31, Cor. 5]. In the other words $\dim_F D = n^2$ for some $n \geq 1$. Such n is called the degree of D and is denoted by $\deg(D)$. Recall that a field extension L/F is called a splitting field of D if $D \otimes_F L \cong M_n(L)$ where $n = \deg(D)$. D is said to be cyclic if there is a maximal subfield K of F such that K/F is a cyclic extension. In a more general context, we say that D is cyclically split if there exists a cyclic extension L/F which splits D. It is well known that all division algebras of degrees two, three and six are cyclic (see [6, §15]). The existence of a non-cyclic division algebra remained as an open problem till 1932 when Albert presented the first example of a non-cyclic division algebra in [1]. After that, other examples of noncyclic division algebras were given by different authors. As an example, in [2] Amitsur proved that for a given n there exists a division algebra D of degree n such that every maximal subfield of D is Galois over the center and its Galois group is a direct product of cyclic groups of prime order. So in the special case that n is divisible by a square of a prime, D is noncyclic. However, in general, it is very difficult to recognize that a given division algebra is cyclic or cyclically split. In fact, there are very limited methods to detect this properties. One of the most important results in relation to this problem is that if D is a tame and totally ramified valued F-central division algebra such that its relative value group has rank at least 3, then D is not cyclically split [8, Th. 4.7]. To see more results concerning noncyclic alegbars, we refer the reader to [4, 5, 7]. The aim of this note is to give a generalization of [8, Th. 4.7] in the level of inertially split division algebras.

*Talker
Email addresses: motiee@nit.ac.ir (Mehran Motiee), behnamtalaee@nit.ac.ir (Behnam Talaee)

2. Results

To state our results we need to recall some preliminaries from the theory of valued division algebras.
Let Γ be an additive totally ordered abelian group. By a valuation on D with values in Γ, we mean a map $v : D^* \to \Gamma$ satisfying, for all $a, b \in D^*$, $v(ab) = v(a) + v(b)$ and $v(a + b) \geq \min\{v(a), v(b)\}$. We write Γ_D for the value group of v, i.e., $\Gamma_D = v(D^*)$. We denote the valuation ring by $V_D = \{d \in D^* \mid v(d) \geq 0\} \cup \{0\}$. It can be seen that this ring has a unique maximal ideal denoted by $M_D = \{d \in D^* \mid v(d) > 0\} \cup \{0\}$. We denote the residue division ring of v by $\overline{D} = V_D/M_D$. When we restrict v to F^*, we obtain a valuation w on the field F. The objects for w corresponding to those for v are denoted by Γ_F, V_F, M_F and $\overline{F}$. Since $V_F \cap M_D = M_F$, one can consider the residue field $\overline{F}$ as a subalgebra of $\overline{D}$. F is called *Henselian* if its valuation has a unique extension to any algebraic extension of F. In this setting, we say D is *tame* if $Z(\overline{D})/\overline{F}$ is separable and char $\overline{F} \nmid n$. D is called *totally ramified* if $\overline{D} = \overline{F}$. We also say that D is *inertially split* if it has a valued splitting field L such that $\Gamma_L = \Gamma_F$. Finally, we recall that if A is an abelian group then its rank, denoted by rank(A), is the smallest cardinality of a generating set for A.
Now, we are ready for the following results.

Theorem 2.1. *Let D be a tame and inertially split F-central division algebra. If* $\text{rank}(\Gamma_D/\Gamma_F) \geq 3$*, then D is not cyclically split.*

Theorem 2.2. *Let F be a Henselian field. Let D be a tame and inertially split central F-division algebra. If* $\text{rank}(\Gamma_D/\Gamma_F) = 2$ *and F does not contain any n-th root of unity then D is not cyclically split.*

References

[1] A. A. Albert. A construction of non-cyclic normal division algebras. *Bull. Amer. Math. Soc.*, 38(6):449–456, 1932.
[2] S. A. Amitsur. On central division algebras. *Israel J. Math.*, 12:408–420, 1972.
[3] P. K. Draxl. *Skew fields*, volume 81 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1983.
[4] W. Kuyk. Generic construction of non-cyclic division algebras. *J. Pure Appl. Algebra*, 2(2):121–130, 1972.
[5] E. Matzri, L. H. Rowen, and U. Vishne. Non-cyclic algebras with n-central elements. *Proc. Amer. Math. Soc.*, 140(2):513–518, 2012.
[6] R. S. Pierce. *Associative algebras*, volume 88 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1982. Studies in the History of Modern Science, 9.
[7] L. Risman. Non-cyclic division algebras. *J. Pure Appl. Algebra*, 11(1-3):199–215, 1977/78.
[8] J.-P. Tignol and A. R. Wadsworth. Totally ramified valuations on finite-dimensional division algebras. *Trans. Amer. Math. Soc.*, 302(1): 223–250, 1987.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–199, pp. 381–382

Relation between stable range one and uniquely clean in ring theory

Asghar Taheri[a,*]

[a] *Department of Mathematics, Payme Noor University, P. O. Box 19395-4697 Tehran, IRAN.*

Article Info

Keywords:
clean range
strongly clean range
uniquely clean range
stable range one.

2020 MSC:
16U99
16D50

Abstract

It is well known that every unit regular ring is clean and has stable range one. We construct an example of a ring that has stable range one and uniquely clean but isn't unit regular.

1. Introduction

An element a in a ring R is called a clean element if a is a sum of a unit and an idempotent in R. On the other hand, We have $a = e + u$, with $e \in idem(R)$ and $u \in U(R)$, a clean decomposition for a. A ring is clean if every element has a clean decomposition. Clean rings were first studied by Nicholson [2] in connection with exchange rings and lifting of idempotents. Recall that an element in a ring R is called unit-regular if it can be expressed as a product of a unit and an idempotent. The ring R is called a unit regular ring if every element is unit-regular [1]. In [4], Camillo and Khurana proved that every unit-regular ring is clean. An element $a \in R$ is strongly clean if it has a clean decomposition $a = e + u$ in which $eu = ue$. An element is uniquely clean if it has exactly one clean decomposition. Strongly clean rings and uniquely clean rings are defined similarly. A ring R is said to have stable range one if for any $a, b \in R$ satisfying $aR + bR = R$, there exists $y \in R$ such that $a + by$ is a unit [1].

2. Stable range one with uniquely clean do not implies unit regularity

Lemma 2.1. *If $aR + bR = R$ implies a or b is unit then R has stable range one.*

Proof. If $aR + bR = R$ and a is a unit then $a + (b.0)$ is a unit so R has stable range one. □

*Talker
Email address: a.taheri@pnu.ac.ir (Asghar Taheri)

Example 2.2. Let $R = \{m/n \in Q : n\ is\ odd\}$,then R has stable range one also uniquely clean and do not unit regular.
If $(m/n)R+(p/q)R = R$, we conclude that there exist (x/y) and (r/t) in R such that $(m/n)(x/y)+(p/q)(r/t) = 1$. So, we have $qtmx + nypr = nyqt$. Then both of m and p can't even since the $nyqt$ is odd. Let m is odd then (m/n) is a unit with inverse (n/m), by applying lemma 2.1, R has stable range one. If $(m/n) \in R$ and m is odd then (m/n) is a unit and all of idempotent of R is 0 and 1 so $(m/n) = (m/n) + 0$ is unique presentations in clean form and if m is even then $m - n$ is odd so $(\dfrac{m-n}{n})$ is a unit and $(m/n) = (\dfrac{m-n}{n}) + 1$ is unique presentations in clean form so R is uniquely clean. It is obvious that if m is even. So, we can't present (m/n) of product a unit and idempotent. Hence R isn't unit-regular.

3. Unit regularity do not implies uniquely cleanity

Lemma 3.1. *Let $e \in R$ is an idempotent. If $r \in R$, then $e - (er - ere)$ is an idempotent and $1 - (er - ere)$ is a unit.*

Proof. $1 - (er - ere)$ is a unit with inverse $1 + (er - ere)$.
Now, we pay attention if R is uniquely clean then by lemma 3.1 and the fact that $[e-(er-ere)]+1 = e+[1-(er-ere)]$ implies that $e - (er - ere) = e$ because R is uniquely clean. Then, we obtain that $er = ere$, and similarly $re = ere$. Hence, idempotents in uniquely clean ring are central. □

Theorem 3.2. *Every unit-regular ring is not uniquely clean.*

Proof. In view of [4, Theorem1], if the ring R is unit-regular and $a \in R$ then, there exist $u \in U(R)$ and idempotent $e \in R$ such that $a = e + u$ and e and u constructed in the proof do not commute But by Lemma 3.1 if R is a uniquely clean ring any idempotent is central, so R is not uniquely clean. □

Let $_RM_a$ leftt R-module. $_RM$ is said to have the (full) exchange property if for every module $_RA$ and any two decompositions of $_RA$,

$$A = M' \oplus N = \oplus_{i \in I} A_i,$$

With M' $\approx M$ there exist submodules $A'_i \subseteq A_i$ such that

$$A = M' \oplus (\oplus_{i \in I} A'_i).$$

The module $_RM$ has the finite exchange property if the above condition is satisfied whenever the index set I is finite. Warfield [5] introduced the class of exchange rings. He called a ring R an exchange ring if $_RR$ has the exchange property above and proved that this definition is left-right symmetric. Independently, Goodearl and Nicholson [2] obtained the very useful characterization that R is an exchange ring if and only if for any $a \in R$ there exist $e^2 = e \in R$ such that $e \in aR$ and $1 - e \in (1 - a)R$

Theorem 3.3. *Let R be an exchange ring. If for any regular $a \in R$, there exist an idempotent $e \in R$ and $u \in U(R)$ such that $a = e + u$ and $aR \bigcap eR = 0$ then R has stable range one.*

Proof. Given that any regular $a \in R$. So, there exist an idempotent $e \in R$ and $u \in U(R)$ such that $a = e + u$ and $aR \bigcap eR = 0$. As a result, $(au^{-1} - 1)a = eu^{-1}a \in aR \bigcap eR = 0$; hence, $a = au^{-1}a$. According to [3, Theorem 3], we complete the proof of Theorem 3.3. □

Author in [2, Proposition 1.8] was obtained that every clean ring is an exchange ring. So, by using Theorem 3.3, we derive the following corollary.

Corollary 3.4. *If R is a clean ring such that every regular element of R is a unit regular then R has stable range one.*

References

[1] K.R. Goodearl. *Von Neumann Regular Rings*. Pitman, 1979.
[2] W. K. Nicholson. Lifting idempotents and exchange rings. *Trans. Amer. Math. Soc.*, 229:269–278, 1977.
[3] H. P. Yu V. P. Camillo. Stable range one for rings with many idempotents. *Trans. Amer. Math. Soc.*, 347:3141–3147, 1995.
[4] Dinesh Khurana Victor P. Camillo. A characterization of unit regular rings. *Comm. Algebra*, 29:2293–2295, 2001.
[5] R.B. Warfield. Warfield, exchange rings and decompositions of modules. *Math. Ann.*, 199:31–36, 1972.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–200, pp. 383–387

Bass numbers of generalized local cohomology modules

Alireza Vahidi

Department of Mathematics, Payame Noor University, Tehran, Iran

Article Info

Keywords:
Bass numbers
Generalized local cohomology modules

2020 MSC:
13D45

Abstract

Let R be a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ an ideal of R, M a finitely generated R-module, X an arbitrary R-module, and t a non-negative integer. In this paper, we compare the Bass numbers of generalized local cohomology module $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ with the Bass numbers of some other generalized local cohomology modules $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ where $i \neq t$.

1. Introduction

Throughout R is a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ is an ideal of R, M is a finite (i.e., finitely generated) R-module, and X is an arbitrary R-module which is not necessarily finite. For a prime ideal $\mathfrak{p}$ of R, the number $\mu^i(\mathfrak{p}, X) = \dim_{\kappa(\mathfrak{p})}(\mathrm{Ext}^i_{R_{\mathfrak{p}}}(\kappa(\mathfrak{p}), X_{\mathfrak{p}}))$ is known as the ith Bass number of X with respect to $\mathfrak{p}$, where $\kappa(\mathfrak{p}) = R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$. When R is local with maximal ideal $\mathfrak{m}$, we write $\mu^i(X) = \mu^i(\mathfrak{m}, X)$ and $\kappa = R/\mathfrak{m}$. The ith generalized local cohomology module

$$\mathrm{H}^i_{\mathfrak{a}}(M, X) \cong \varinjlim_{n \in \mathbb{N}} \mathrm{Ext}^i_R(M/\mathfrak{a}^n M, X)$$

was introduced by Herzog in [6]. It is clear that $\mathrm{H}^i_{\mathfrak{a}}(R, X)$ is just the ordinary local cohomology module $\mathrm{H}^i_{\mathfrak{a}}(X)$ of X with respect to $\mathfrak{a}$. For basic results, notations, and terminology not given in this paper, the reader is referred to [1, 2, 10].
An important problem in commutative algebra is to determine when the Bass numbers of the local cohomology module $\mathrm{H}^i_{\mathfrak{a}}(X)$ are finite. The following conjecture, which was made by Huneke, has been studied by several authors (see [7, Conjecture 4.4]).

Conjecture 1.1. Let R be a regular local ring. Then for any prime ideal $\mathfrak{p}$ of R, the Bass numbers $\mu^i(\mathfrak{p}, \mathrm{H}^j_{\mathfrak{a}}(R))$ is finite for all i and j.

Email address: vahidi.ar@pnu.ac.ir (Alireza Vahidi)

Even though there is a negative answer to this conjecture (over a non-regular ring) (see [5]), there is evidence that this conjecture is true (see [8] and [9]). Since the conjecture does not hold over a non-regular ring, several attempts made to find some conditions for the ideal $\mathfrak{a}$ to have finiteness for the Bass numbers of the local cohomology modules with respect to the ideal $\mathfrak{a}$.
In [3] and for the local case, Dibaei and Yassemi compared the Bass numbers of local cohomology module $\mathrm{H}^t_{\mathfrak{a}}(X)$ with the Bass numbers of X and those of some other local cohomology modules $\mathrm{H}^i_{\mathfrak{a}}(X)$ where $i \neq t$. They proved, in [3, Theorem 2.6], that the inequality

$$\mu^s(\mathrm{H}^t_{\mathfrak{a}}(X)) \leq \mu^{s+t}(X) + \sum_{i=0}^{t-1} \mu^{s+t+1-i}(\mathrm{H}^i_{\mathfrak{a}}(X)) + \sum_{i=t+1}^{s+t-1} \mu^{s+t-1-i}(\mathrm{H}^i_{\mathfrak{a}}(X))$$

holds for all $s \geq 0$. In this paper, we generalize this result by showing that if M is a finite R-module, then the inequality

$$\mu^s(\mathrm{H}^t_{\mathfrak{a}}(M,X)) \leq \sum_{i=0}^{s+t} \dim_\kappa(\mathrm{Ext}^{s+t-i}_R(\mathrm{Tor}^R_i(\kappa,M),X)) + \sum_{i=0}^{t-1} \mu^{s+t+1-i}(\mathrm{H}^i_{\mathfrak{a}}(M,X)) + \sum_{i=t+1}^{s+t-1} \mu^{s+t-1-i}(\mathrm{H}^i_{\mathfrak{a}}(M,X))$$

is true for all $s \geq 0$.

2. Main Results

The following theorem is the main result of this paper which generalizes [3, Theorem 2.6]. In the proof, we use the isomorphism

$$\mathrm{H}^i_{\mathfrak{a}}(M,X) \cong \mathrm{H}^i(\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))),$$

for all i, where $\mathrm{E}^{\bullet X}$ is a deleted injective resolution of X (see [4, Lemma 2.1(i)]).

Theorem 2.1. *Let R be a local ring, M a finite R-module, and X an R-module. Then the inequality*

$$\mu^s(\mathrm{H}^t_{\mathfrak{a}}(M,X)) \leq \sum_{i=0}^{s+t} \dim_\kappa(\mathrm{Ext}^{s+t-i}_R(\mathrm{Tor}^R_i(\kappa,M),X)) + \sum_{i=0}^{t-1} \mu^{s+t+1-i}(\mathrm{H}^i_{\mathfrak{a}}(M,X)) + \sum_{i=t+1}^{s+t-1} \mu^{s+t-1-i}(\mathrm{H}^i_{\mathfrak{a}}(M,X))$$

holds for all non-negative integers s and t.

Proof. Suppose that the right-hand side of the inequality is a finite number. Let

$$\mathrm{E}^\bullet = 0 \longrightarrow X \longrightarrow E^0 \longrightarrow \cdots \longrightarrow E^i \longrightarrow \cdots$$

be an injective resolution of X and apply $\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(-))$ to its deletion $\mathrm{E}^{\bullet X}$ to get the complex

$$\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X})) = 0 \longrightarrow \mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(E^0)) \longrightarrow \cdots \longrightarrow \mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(E^i)) \longrightarrow \cdots .$$

Let

$$0 \longrightarrow \mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X})) \longrightarrow T^{\bullet,0} \longrightarrow T^{\bullet,1} \longrightarrow \cdots \longrightarrow T^{\bullet,i} \longrightarrow \cdots$$

be a Cartan-Eilenberg injective resolution of $\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))$, which exists by [10, Theorem 10.45], and consider the third quadrant bicomplex $\mathcal{T} = \{\mathrm{Hom}_R(\kappa,T^{p,q})\}$. We denote the total complex of $\mathcal{T}$ by $\mathrm{Tot}(\mathcal{T})$.
The first filtration has ${}^I\mathrm{E}_2$ term the iterated homology $\mathrm{H}'^p\,\mathrm{H}''^{p,q}(\mathcal{T})$. We have

$$\begin{aligned}\mathrm{H}''^{p,q}(\mathcal{T}) &\cong \mathrm{H}^q(\mathrm{Hom}_R(\kappa,T^{p,\bullet}))\\ &\cong \mathrm{Ext}^q_R(\kappa,\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^p)))\\ &\cong \mathrm{Hom}_R(\mathrm{Tor}^R_q(\kappa,M),\Gamma_{\mathfrak{a}}(\mathrm{E}^p))\end{aligned}$$

from [1, Proposition 2.1.4] and [10, Corollary 10.63]. Therefore

$$\begin{aligned}{}^I\mathrm{E}^{p,q}_2 &\cong \mathrm{H}'^p\,\mathrm{H}''^{p,q}(\mathcal{T})\\ &\cong \mathrm{H}^p(\mathrm{Hom}_R(\mathrm{Tor}^R_q(\kappa,M),\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X})))\\ &\cong \mathrm{H}^p_{\mathfrak{a}}(\mathrm{Tor}^R_q(\kappa,M),X)\end{aligned}$$

which yields, by [12, Lemma 2.5(c)], the third quadrant spectral sequence

$$^{I}\,\mathrm{E}_2^{p,q} := \mathrm{Ext}_R^p(\mathrm{Tor}_q^R(\kappa, M), X) \underset{p}{\Longrightarrow} \mathrm{H}^{p+q}(\mathrm{Tot}(\mathcal{T})).$$

For all $i \leq s+t$, we have $^{I}E_\infty^{s+t-i,i} = {}^{I}E_{s+t+2}^{s+t-i,i}$ because $^{I}E_j^{s+t-i-j,i+j-1} = 0 = {}^{I}E_j^{s+t-i+j,i+1-j}$ for all $j \geq s+t+2$; so that $\dim_\kappa({}^{I}E_\infty^{s+t-i,i}) \leq \dim_\kappa({}^{I}E_2^{s+t-i,i})$ from the fact that $^{I}E_{s+t+2}^{s+t-i,i}$ is a subquotient of $^{I}E_2^{s+t-i,i}$. There exists a finite filtration

$$0 = \phi^{s+t+1}H^{s+t} \subseteq \phi^{s+t}H^{s+t} \subseteq \cdots \subseteq \phi^1 H^{s+t} \subseteq \phi^0 H^{s+t} = H^{s+t}$$

such that $^{I}E_\infty^{s+t-i,i} \cong \phi^{s+t-i}H^{s+t}/\phi^{s+t-i+1}H^{s+t}$ for all $i \leq s+t$. Now the exact sequences

$$0 \longrightarrow \phi^{s+t-i+1}H^{s+t} \longrightarrow \phi^{s+t-i}H^{s+t} \longrightarrow {}^{I}E_\infty^{s+t-i,i} \longrightarrow 0,$$

for all $i \leq s+t$, show that

$$\begin{aligned} \dim_\kappa(H^{s+t}) &\leq \sum_{i=0}^{s+t} \dim_\kappa({}^{I}E_\infty^{s+t-i,i}) \\ &\leq \sum_{i=0}^{s+t} \dim_\kappa({}^{I}E_2^{s+t-i,i}) \\ &= \sum_{i=0}^{s+t} \dim_\kappa(\mathrm{Ext}_R^{s+t-i}(\mathrm{Tor}_i^R(\kappa, M), X)). \end{aligned}$$

On the other hand, the second filtration has $^{II}\,\mathrm{E}_2$ term the iterated homology $\mathrm{H}''^p\,\mathrm{H}'^{q,p}(\mathcal{T})$. Note that every short exact sequence of injective modules splits and so it remains split after applying the functor $\mathrm{Hom}_R(\kappa, -)$. By using this fact and the fact that $T^{\bullet,\bullet}$ is a Cartan-Eilenberg injective resolution of $\mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))$, we get

$$\begin{aligned} \mathrm{H}'^{q,p}(\mathcal{T}) &\cong \mathrm{H}^q(\mathrm{Hom}_R(\kappa, T^{\bullet,p})) \\ &\cong \mathrm{Hom}_R(\kappa, \mathrm{H}^q(T^{\bullet,p})) \\ &\cong \mathrm{Hom}_R(\kappa, \mathrm{H}^{q,p}). \end{aligned}$$

Therefore

$$\begin{aligned} {}^{II}\,\mathrm{E}_2^{p,q} &\cong \mathrm{H}''^p\,\mathrm{H}'^{q,p}(\mathcal{T}) \\ &\cong \mathrm{H}^p(\mathrm{Hom}_R(\kappa, \mathrm{H}^{q,\bullet})) \\ &\cong \mathrm{Ext}_R^p(\kappa, \mathrm{H}_{\mathfrak{a}}^q(M, X)) \end{aligned}$$

which gives the third quadrant spectral sequence

$$^{II}\,\mathrm{E}_2^{p,q} := \mathrm{Ext}_R^p(\kappa, \mathrm{H}_{\mathfrak{a}}^q(M, X)) \underset{p}{\Longrightarrow} \mathrm{H}^{p+q}(\mathrm{Tot}(\mathcal{T})).$$

Thus there exists a finite filtration

$$0 = \psi^{s+t+1}H^{s+t} \subseteq \psi^{s+t}H^{s+t} \subseteq \cdots \subseteq \psi^1 H^{s+t} \subseteq \psi^0 H^{s+t} = H^{s+t}$$

such that $^{II}E_\infty^{s+t-i,i} \cong \psi^{s+t-i}H^{s+t}/\psi^{s+t-i+1}H^{s+t}$ for all $i \leq s+t$. Hence $\dim_\kappa({}^{II}E_\infty^{s,t}) \leq \dim_\kappa(\psi^s H^{s+t}) \leq \dim_\kappa(H^{s+t})$. Therefore

$$\dim_\kappa({}^{II}E_{s+t+2}^{s,t}) \leq \dim_\kappa(H^{s+t})$$

because $^{II}E_{s+t+2}^{s,t} = {}^{II}E_\infty^{s,t}$ from the fact that $^{II}E_j^{s-j,t+j-1} = 0 = {}^{II}E_j^{s+j,t+1-j}$ for all $j \geq s+t+2$. For all $r \geq 2$, let $^{II}Z_r^{s,t} = \mathrm{Ker}({}^{II}E_r^{s,t} \longrightarrow {}^{II}E_r^{s+r,t+1-r})$ and $^{II}B_r^{s,t} = \mathrm{Im}({}^{II}E_r^{s-r,t+r-1} \longrightarrow {}^{II}E_r^{s,t})$. We have the exact sequences

$$0 \longrightarrow {}^{II}Z_r^{s,t} \longrightarrow {}^{II}E_r^{s,t} \longrightarrow {}^{II}E_r^{s,t}/{}^{II}Z_r^{s,t} \longrightarrow 0$$

and

$$0 \longrightarrow {}^{II}B_r^{s,t} \longrightarrow {}^{II}Z_r^{s,t} \longrightarrow {}^{II}E_{r+1}^{s,t} \longrightarrow 0.$$

Thus we get

$$\begin{aligned}\dim_\kappa({}^{II}E_r^{s,t}) &\le \dim_\kappa({}^{II}E_{r+1}^{s,t}) + \dim_\kappa({}^{II}E_r^{s,t}/{}^{II}Z_r^{s,t}) + \dim_\kappa({}^{II}B_r^{s,t})\\ &\le \dim_\kappa({}^{II}E_{r+1}^{s,t}) + \dim_\kappa({}^{II}E_r^{s+r,t+1-r}) + \dim_\kappa({}^{II}E_r^{s-r,t+r-1})\\ &\le \dim_\kappa({}^{II}E_{r+1}^{s,t}) + \dim_\kappa({}^{II}E_2^{s+r,t+1-r}) + \dim_\kappa({}^{II}E_2^{s-r,t+r-1}).\end{aligned}$$

Therefore we have

$$\begin{aligned}&\dim_\kappa({}^{II}E_2^{s,t})\\ &\le \dim_\kappa({}^{II}E_3^{s,t}) + \dim_\kappa({}^{II}E_2^{s+2,t-1}) + \dim_\kappa({}^{II}E_2^{s-2,t+1})\\ &\le \dim_\kappa({}^{II}E_4^{s,t}) + (\dim_\kappa({}^{II}E_2^{s+3,t-2}) + \dim_\kappa({}^{II}E_2^{s+2,t-1})) + (\dim_\kappa({}^{II}E_2^{s-2,t+1}) + \dim_\kappa({}^{II}E_2^{s-3,t+2}))\\ &\le \cdots\\ &\le \dim_\kappa({}^{II}E_{s+t+2}^{s,t}) + (\sum_{i=0}^{t-1} \dim_\kappa({}^{II}E_2^{s+t+1-i,i})) + (\sum_{i=t+1}^{s+t-1} \dim_\kappa({}^{II}E_2^{s+t-1-i,i}))\\ &\le \dim_\kappa({}^{II}E_{s+t+2}^{s,t}) + (\sum_{i=0}^{t-1} \dim_\kappa(\mathrm{Ext}_R^{s+t+1-i}(\kappa, \mathrm{H}_\mathfrak{a}^i(M,X))) + (\sum_{i=t+1}^{s+t-1} \dim_\kappa(\mathrm{Ext}_R^{s+t-1-i}(\kappa, \mathrm{H}_\mathfrak{a}^i(M,X))))\end{aligned}$$

which completes the proof. □

As applications of the above theorem, we have the following corollaries. For a finite R-module M with finite projective dimension and for an arbitrary R-module X, we denote the largest integer i in which $\mathrm{H}_\mathfrak{a}^i(M,X)$ is not zero by $\mathrm{cd}_\mathfrak{a}(M,X)$. We also write $\mathrm{cd}_\mathfrak{a}(X) = \mathrm{cd}_\mathfrak{a}(R,X)$.

Corollary 2.2. *Let R be a local ring, M a finite R-module with finite projective dimension, and X a finite R-module such that* $\mathrm{depth}_R(\mathfrak{a} + \mathrm{Ann}_R(M), X) = \mathrm{cd}_\mathfrak{a}(M,X)$. *Then the inequality*

$$\mu^s(\mathrm{H}_\mathfrak{a}^{\mathrm{cd}_\mathfrak{a}(M,X)}(M,X)) \le \sum_{i=0}^{s+\mathrm{cd}_\mathfrak{a}(M,X)} \dim_\kappa(\mathrm{Ext}_R^{s+\mathrm{cd}_\mathfrak{a}(M,X)-i}(\mathrm{Tor}_i^R(\kappa,M),X))$$

is true for all $s \ge 0$.

Proof. By [12, Corollary 2.14], $\mathrm{H}_\mathfrak{a}^i(M,X) = 0$ for all $i \ne \mathrm{cd}_\mathfrak{a}(M,X)$. Thus the assertion follows from Theorem 2.1. □

Corollary 2.3. *Let R be a local Cohen-Macaulay ring with maximal ideal $\mathfrak{m}$ and let M be a finite R-module with finite projective dimension. Then*

$$\mu^s(\mathrm{H}_\mathfrak{m}^{\dim(R)}(M,R)) \le \sum_{i=0}^{s} \dim_\kappa(\mathrm{Ext}_R^{s+\dim(R)-i}(\mathrm{Tor}_i^R(\kappa,M),R))$$

for all $s \ge 0$.

Proof. By [12, Corollary 2.9] and [11, Corollary 3.2], $\mathrm{H}_\mathfrak{m}^i(M,R) = 0$ for all $i \ne \dim(R)$. Thus the assertion follows from Theorem 2.1. Note that, $\dim_\kappa(\mathrm{Ext}_R^{s+\dim(R)-i}(\mathrm{Tor}_i^R(\kappa,M),R)) = 0$ for all $i > s$ because R is a local Cohen-Macaulay ring. □

The next results follow by putting $M = R$ in Theorem 2.1, Corollary 2.2, and Corollary 2.3.

Corollary 2.4. (see [3, Theorem 2.6]) *Let R be a local ring and X an R-module. Then the inequality*

$$\mu^s(\mathrm{H}_\mathfrak{a}^t(X)) \le \mu^{s+t}(X) + \sum_{i=0}^{t-1} \mu^{s+t+1-i}(\mathrm{H}_\mathfrak{a}^i(X)) + \sum_{i=t+1}^{s+t-1} \mu^{s+t-1-i}(\mathrm{H}_\mathfrak{a}^i(X))$$

holds for all non-negative integers s and t.

Corollary 2.5. *Let R be a local ring and X a finite R-module such that* $\operatorname{depth}_R(\mathfrak{a}, X) = \operatorname{cd}_{\mathfrak{a}}(X)$. *Then the inequality*

$$\mu^s(\mathrm{H}_{\mathfrak{a}}^{\operatorname{cd}_{\mathfrak{a}}(X)}(X)) \leq \mu^{s+\operatorname{cd}_{\mathfrak{a}}(X)}(X)$$

is true for all $s \geq 0$.

Corollary 2.6. *Let R be a local Cohen-Macaulay ring with maximal ideal $\mathfrak{m}$. Then*

$$\mu^s(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(R)) \leq \mu^{s+\dim(R)}(R)$$

for all $s \geq 0$.

References

[1] M. P. Brodmann and R. Y. Sharp, Local cohomology: an algebraic introduction with geometric applications, Cambridge Stud. Adv. Math., 60, *Cambridge University Press*, *Cambridge*, 1998. MR1613627.

[2] W. Bruns and J. Herzog, Cohen-Macaulay rings, Cambridge Stud. Adv. Math., 39, *Cambridge University Press*, *Cambridge*, 1993. MR1251956.

[3] M. T. Dibaei and S. Yassemi, Bass numbers of local cohomology modules with respect to an ideal, *Algebr. Represent. Theory* **11** (2008), no. 3, 299–306. MR2403295.

[4] K. Divaani-Aazar, R. Sazeedeh, and M. Tousi, On vanishing of generalized local cohomology modules, *Algebra Colloq.* **12** (2005), no. 2, 213–218. MR2127246.

[5] R. Hartshorne, Affine duality and cofiniteness, *Invent. Math.* **9** (1970), 145–164. MR0257096.

[6] J. Herzog, Komplexe, auflösungen und dualität in der lokalen Algebra, Habilitationsschrift, *Universitat Regensburg*, *Regensburg*, 1970.

[7] C. Huneke, Problems on local cohomology, *Free resolutions in commutative algebra and algebraic geometry (Sundance, UT, 1990)*, 93–108, Res. Notes Math., 2, *Jones and Bartlett*, *Boston*, *MA*, 1992. MR1165320.

[8] C. Huneke and R. Y. Sharp, Bass numbers of local cohomology modules, *Trans. Amer. Math. Soc.* **339** (1993), no. 2, 765–779. MR1124167.

[9] G. Lyubeznik, Finiteness properties of local cohomology modules (an application of D-modules to commutative algebra), *Invent. Math.* **113** (1993), no. 1, 41–55. MR1223223.

[10] J. J. Rotman, An introduction to homological algebra, Second edition, Universitext, *Springer*, *New York*, 2009. MR2455920.

[11] N. Tu Cuong and N. Van Hoang, On the vanishing and the finiteness of supports of generalized local cohomology modules, *Manuscripta Math.* **126** (2008), no. 1, 59–72. MR2395249.

[12] A. Vahidi and M. Aghapournahr, Some results on generalized local cohomology modules, *Comm. Algebra* **43** (2015), no. 5, 2214–2230. MR3316846.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–201, pp. 388–391

Betti numbers of generalized local cohomology modules

Alireza Vahidi

Department of Mathematics, Payame Noor University, Tehran, Iran

Article Info

Keywords:
Betti numbers
Generalized local cohomology modules

2020 MSC:
13D45

Abstract

Let R be a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ an ideal of R, M an R-module with a finite free resolution, X an arbitrary R-module, and t a non-negative integer. In this paper, we compare the Betti numbers of generalized local cohomology module $\mathrm{H}^t_{\mathfrak{a}}(M,X)$ with the Betti numbers of some other generalized local cohomology modules $\mathrm{H}^i_{\mathfrak{a}}(M,X)$ where $i \neq t$.

1. Introduction

Throughout R is a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ is an ideal of R, M is a finite (i.e., finitely generated) R-module, and X is an arbitrary R-module which is not necessarily finite. For basic results, notations, and terminology not given in this paper, the reader is referred to [1, 2, 5].
The ith generalized local cohomology module

$$\mathrm{H}^i_{\mathfrak{a}}(M,X) \cong \varinjlim_{n\in\mathbb{N}} \mathrm{Ext}^i_R(M/\mathfrak{a}^n M, X)$$

was introduced by Herzog in [3]. It is clear that $\mathrm{H}^i_{\mathfrak{a}}(R,X)$ is just the ordinary local cohomology module $\mathrm{H}^i_{\mathfrak{a}}(X)$ of X with respect to $\mathfrak{a}$. For a finite R-module M with finite projective dimension and for an arbitrary R-module X, we denote the largest integer i in which $\mathrm{H}^i_{\mathfrak{a}}(M,X)$ is not zero by $\mathrm{cd}_{\mathfrak{a}}(M,X)$. We also write $\mathrm{cd}_{\mathfrak{a}}(X) = \mathrm{cd}_{\mathfrak{a}}(R,X)$.
For a prime ideal $\mathfrak{p}$ of R, the number $\beta_i(\mathfrak{p},X) = \dim_{\kappa(\mathfrak{p})}(\mathrm{Tor}^{R_{\mathfrak{p}}}_i(\kappa(\mathfrak{p}),X_{\mathfrak{p}}))$ is known as the ith Betti number of X with respect to $\mathfrak{p}$, where $\kappa(\mathfrak{p}) = R_{\mathfrak{p}}/\mathfrak{p}R_{\mathfrak{p}}$. When R is local with maximal ideal $\mathfrak{m}$, we write $\beta_i(X) = \beta_i(\mathfrak{m},X)$ and $\kappa = R/\mathfrak{m}$.
In [7] and for the local case, the author compared the Betti numbers of local cohomology module $\mathrm{H}^t_{\mathfrak{a}}(X)$ with the Betti numbers of X and those of some other local cohomology modules $\mathrm{H}^i_{\mathfrak{a}}(X)$ where $i \neq t$. He proved, in [7, Theorem 3.2], that the inequality

$$\beta_s(\mathrm{H}^t_{\mathfrak{a}}(X)) \leq \beta_{s-t}(X) + \sum_{i=0}^{t-1} \beta_{s-t+i-1}(\mathrm{H}^i_{\mathfrak{a}}(X)) + \sum_{i=t+1}^{\mathrm{cd}_{\mathfrak{a}}(X)} \beta_{s-t+i+1}(\mathrm{H}^i_{\mathfrak{a}}(X))$$

Email address: vahidi.ar@pnu.ac.ir (Alireza Vahidi)

holds for all $s \geq 0$. In this paper, we generalize this result by showing that if M is an R-module with a finite free resolution of length l, then the inequality

$$\beta_s(\mathrm{H}^t_{\mathfrak{a}}(M,X)) \leq$$
$$\sum_{i=s-t}^{\mathrm{ara}(\mathfrak{a})+l+s-t} \dim_\kappa(\mathrm{Ext}^{t-s+i}_R(M, \mathrm{Tor}^R_i(\kappa, X))) + \sum_{i=0}^{t-1} \beta_{s-t+i-1}(\mathrm{H}^i_{\mathfrak{a}}(M,X)) + \sum_{i=t+1}^{\mathrm{cd}_{\mathfrak{a}}(M,X)} \beta_{s-t+i+1}(\mathrm{H}^i_{\mathfrak{a}}(M,X))$$

is true for all $s \geq 0$. Here, $\mathrm{ara}(\mathfrak{a})$ is the arithmetic rank of $\mathfrak{a}$.

2. Main Results

The following theorem is the main result of this paper which generalizes [7, Theorem 3.2].

Theorem 2.1. *Let R be a local ring, M an R-module with a finite free resolution of length l, and X an R-module. Then the inequality*

$$\beta_s(\mathrm{H}^t_{\mathfrak{a}}(M,X)) \leq$$
$$\sum_{i=s-t}^{\mathrm{ara}(\mathfrak{a})+l+s-t} \dim_\kappa(\mathrm{Ext}^{t-s+i}_R(M, \mathrm{Tor}^R_i(\kappa, X))) + \sum_{i=0}^{t-1} \beta_{s-t+i-1}(\mathrm{H}^i_{\mathfrak{a}}(M,X)) + \sum_{i=t+1}^{\mathrm{cd}_{\mathfrak{a}}(M,X)} \beta_{s-t+i+1}(\mathrm{H}^i_{\mathfrak{a}}(M,X))$$

holds for all non-negative integers s and t.

Proof. Suppose that the right-hand side of the inequality is a finite number. Let $c = \mathrm{ara}(\mathfrak{a})$, $u = c + l - t$, and $n = s + u$. There exist elements $x_1, \ldots, x_c$ of R such that $\sqrt{\mathfrak{a}} = \sqrt{(x_1, \ldots, x_c)}$. Let

$$F_\bullet = \cdots \longrightarrow F_{i+1} \longrightarrow F_i \longrightarrow F_{i-1} \longrightarrow \cdots \longrightarrow F_1 \longrightarrow F_0 \longrightarrow 0$$

and

$$F'_\bullet = 0 \longrightarrow F'_l \longrightarrow F'_{l-1} \longrightarrow \cdots \longrightarrow F'_1 \longrightarrow F'_0 \longrightarrow 0$$

be, respectively, free resolutions of κ and M such that F'_i is finite for all $0 \leq i \leq l$. Consider the first quadrant bicomplex $\mathcal{T} = \{F_p \otimes_R (\bigoplus_{i+j=c+l-q} C^i \otimes_R \mathrm{Hom}_R(F'_j, X))\}$, where $C^\bullet$ is the Čech complex of R with respect to $x_1, \ldots, x_c$. We denote the total complex of $\mathcal{T}$ by $\mathrm{Tot}(\mathcal{T})$.
The second filtration has ${}^{II}E^2$ term the iterated homology $H''_p H'_{q,p}(\mathcal{T})$. We have

$$\begin{aligned} H'_{q,p}(\mathcal{T}) &\cong H_q(F_\bullet \otimes_R (\textstyle\bigoplus_{i+j=c+l-q} C^i \otimes_R \mathrm{Hom}_R(F'_j, X))) \\ &\cong \textstyle\bigoplus_{i+j=c+l-q} C^i \otimes_R \mathrm{Hom}_R(F'_j, H_q(F_\bullet \otimes_R X)) \\ &\cong \textstyle\bigoplus_{i+j=c+l-q} C^i \otimes_R \mathrm{Hom}_R(F'_j, \mathrm{Tor}^R_q(\kappa, X)). \end{aligned}$$

Thus, by [4, Theorem 2.8] and [8, Lemma 2.5(c)],

$$\begin{aligned} {}^{II}E^2_{p,q} &\cong H^{c+l-p}(\textstyle\bigoplus_{i+j=\bullet} C^i \otimes_R \mathrm{Hom}_R(F'_j, \mathrm{Tor}^R_q(\kappa, X))) \\ &\cong \mathrm{H}^{c+l-p}_{\mathfrak{a}}(M, \mathrm{Tor}^R_q(\kappa, X)) \\ &\cong \mathrm{Ext}^{c+l-p}_R(M, \mathrm{Tor}^R_q(\kappa, X)) \end{aligned}$$

which gives the first quadrant spectral sequence

$${}^{II}E^2_{p,q} := \mathrm{Ext}^{c+l-p}_R(M, \mathrm{Tor}^R_q(\kappa, X)) \underset{p}{\Longrightarrow} H_{p+q}(\mathrm{Tot}(\mathcal{T})).$$

For all $i \leq n$, we have ${}^{II}E^\infty_{n-i,i} = {}^{II}E^{n+2}_{n-i,i}$ because ${}^{II}E^j_{n-i+j,i+1-j} = 0 = {}^{II}E^j_{n-i-j,i-1+j}$ for all $j \geq n+2$. So that $\dim_\kappa({}^{II}E^\infty_{n-i,i}) \leq \dim_\kappa({}^{II}E^2_{n-i,i})$ as ${}^{II}E^{n+2}_{n-i,i}$ is a subquotient of ${}^{II}E^2_{n-i,i}$. There exists a finite filtration

$$0 = \phi^{-1}H_n \subseteq \phi^0 H_n \subseteq \cdots \subseteq \phi^{n-1}H_n \subseteq \phi^n H_n = H_n$$

such that ${}^{II}E^{\infty}_{n-i,i} = \phi^{n-i}H_n/\phi^{n-i-1}H_n$ for all $i \leq n$. Now the exact sequences

$$0 \longrightarrow \phi^{n-i-1}H_n \longrightarrow \phi^{n-i}H_n \longrightarrow {}^{II}E^{\infty}_{n-i,i} \longrightarrow 0,$$

for all $i \leq n$, show that

$$\begin{aligned}
\dim_\kappa(H_n) &= \sum_{i=0}^{n} \dim_\kappa({}^{II}E^{\infty}_{n-i,i}) \\
&\leq \sum_{i=0}^{n} \dim_\kappa({}^{II}E^{2}_{n-i,i}) \\
&= \sum_{i=0}^{\operatorname{ara}(\mathfrak{a})+l+s-t} \dim_\kappa(\operatorname{Ext}^{t-s+i}_R(M, \operatorname{Tor}^R_i(\kappa, X))) \\
&= \sum_{i=s-t}^{\operatorname{ara}(\mathfrak{a})+l+s-t} \dim_\kappa(\operatorname{Ext}^{t-s+i}_R(M, \operatorname{Tor}^R_i(\kappa, X))).
\end{aligned}$$

On the other hand, the first filtration has ${}^{I}E^2$ term the iterated homology $H'_pH''_{p,q}(\mathcal{T})$. Again by [4, Theorem 2.8], we have

$$\begin{aligned}
H''_{p,q}(\mathcal{T}) &= H^{c+l-q}(F_p \otimes_R (\textstyle\bigoplus_{i+j=\bullet} C^i \otimes_R \operatorname{Hom}_R(F'_j, X))) \\
&= F_p \otimes_R H^{c+l-q}(\textstyle\bigoplus_{i+j=\bullet} C^i \otimes_R \operatorname{Hom}_R(F'_j, X)) \\
&= F_p \otimes_R \operatorname{H}^{c+l-q}_{\mathfrak{a}}(M, X).
\end{aligned}$$

Hence

$$\begin{aligned}
{}^{I}E^2_{p,q} &= H_p(F_\bullet \otimes_R \operatorname{H}^{c+l-q}_{\mathfrak{a}}(M, X)) \\
&= \operatorname{Tor}^R_p(\kappa, \operatorname{H}^{c+l-q}_{\mathfrak{a}}(M, X))
\end{aligned}$$

which yields the first quadrant spectral sequence

$${}^{I}E^2_{p,q} := \operatorname{Tor}^R_p(\kappa, \operatorname{H}^{c+l-q}_{\mathfrak{a}}(M, X)) \underset{p}{\Longrightarrow} H_{p+q}(\operatorname{Tot}(\mathcal{T})).$$

Thus there exists a finite filtration

$$0 = \psi^{-1}H_n \subseteq \psi^0 H_n \subseteq \cdots \subseteq \psi^{n-1}H_n \subseteq \psi^n H_n = H_n$$

such that ${}^{I}E^{\infty}_{n-i,i} = \psi^{n-i}H_n/\psi^{n-i-1}H_n$ for all $i \leq n$. Hence $\dim_\kappa({}^{I}E^{\infty}_{s,u}) \leq \dim_\kappa(\psi^s H_n) \leq \dim_\kappa(H_n)$. Therefore

$$\dim_\kappa({}^{I}E^{n+2}_{s,u}) \leq \dim_\kappa(H_n)$$

because we have ${}^{I}E^{n+2}_{s,u} = {}^{I}E^{\infty}_{s,u}$ as ${}^{I}E^{j}_{s+j,u+1-j} = 0 = {}^{I}E^{j}_{s-j,u-1+j}$ for all $j \geq n+2$. For all $r \geq 2$, let ${}^{I}Z^r_{s,u} = \operatorname{Ker}({}^{I}E^r_{s,u} \longrightarrow {}^{I}E^r_{s-r,u-1+r})$ and ${}^{I}B^r_{s,u} = \operatorname{Im}({}^{I}E^r_{s+r,u+1-r} \longrightarrow {}^{I}E^r_{s,u})$. We have the exact sequences

$$0 \longrightarrow {}^{I}Z^r_{s,u} \longrightarrow {}^{I}E^r_{s,u} \longrightarrow {}^{I}E^r_{s,u}/{}^{I}Z^r_{s,u} \longrightarrow 0$$

and

$$0 \longrightarrow {}^{I}B^r_{s,u} \longrightarrow {}^{I}Z^r_{s,u} \longrightarrow {}^{I}E^{r+1}_{s,u} \longrightarrow 0$$

which show that

$$\begin{aligned}
\dim_\kappa({}^{I}E^r_{s,u}) &= \dim_\kappa({}^{I}E^{r+1}_{s,u}) + \dim_\kappa({}^{I}E^r_{s,u}/{}^{I}Z^r_{s,u}) + \dim_\kappa({}^{I}B^r_{s,u}) \\
&\leq \dim_\kappa({}^{I}E^{r+1}_{s,u}) + \dim_\kappa({}^{I}E^r_{s-r,u-1+r}) + \dim_\kappa({}^{I}E^r_{s+r,u+1-r}) \\
&\leq \dim_\kappa({}^{I}E^{r+1}_{s,u}) + \dim_\kappa({}^{I}E^2_{s-r,u-1+r}) + \dim_\kappa({}^{I}E^2_{s+r,u+1-r}).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\dim_\kappa({}^{I}E^2_{s,u}) \\
&\leq \dim_\kappa({}^{I}E^3_{s,u}) + \dim_\kappa({}^{I}E^2_{s-2,u+1}) + \dim_\kappa({}^{I}E^2_{s+2,u-1}) \\
&\leq \dim_\kappa({}^{I}E^4_{s,u}) + (\dim_\kappa({}^{I}E^2_{s-2,u+1}) + \dim_\kappa({}^{I}E^2_{s-3,u+2})) + (\dim_\kappa({}^{I}E^2_{s+3,u-2}) + \dim_\kappa({}^{I}E^2_{s+2,u-1})) \\
&\leq \cdots \\
&\leq \dim_\kappa({}^{I}E^{n+2}_{s,u}) + (\sum_{i=0}^{t-1} \dim_\kappa(\operatorname{Tor}^R_{s-t+i-1}(\kappa, \operatorname{H}^i_{\mathfrak{a}}(M, X)))) + (\sum_{i=t+1}^{\operatorname{cd}_{\mathfrak{a}}(M,X)} \dim_\kappa(\operatorname{Tor}^R_{s-t+i+1}(\kappa, \operatorname{H}^i_{\mathfrak{a}}(M, X))))
\end{aligned}$$

which completes the proof. □

The following corollaries are immediate applications of the above theorem.

Corollary 2.2. *Let R be a local ring, M an R-module with a finite free resolution of length l, and X a finite R-module such that* $\operatorname{depth}_R(\mathfrak{a} + \operatorname{Ann}_R(M), X) = \operatorname{cd}_{\mathfrak{a}}(M, X)$*. Then the inequality*

$$\beta_s(\mathrm{H}_{\mathfrak{a}}^{\operatorname{cd}_{\mathfrak{a}}(M,X)}(M, X)) \leq \sum_{i=s-\operatorname{cd}_{\mathfrak{a}}(M,X)}^{\operatorname{ara}(\mathfrak{a})+l+s-\operatorname{cd}_{\mathfrak{a}}(M,X)} \dim_\kappa(\operatorname{Ext}_R^{\operatorname{cd}_{\mathfrak{a}}(M,X)-s+i}(M, \operatorname{Tor}_i^R(\kappa, X)))$$

is true for all $s \geq 0$. In particular, for all $s \geq 0$,

$$\beta_s(\mathrm{H}_{\mathfrak{a}}^{\operatorname{cd}_{\mathfrak{a}}(M,R)}(M, R)) \leq \dim_\kappa(\operatorname{Ext}_R^{\operatorname{cd}_{\mathfrak{a}}(M,R)-s}(M, \kappa))$$

when $\operatorname{depth}_R(\mathfrak{a} + \operatorname{Ann}_R(M), R) = \operatorname{cd}_{\mathfrak{a}}(M, R)$.

Proof. By [8, Corollary 2.14], $\mathrm{H}_{\mathfrak{a}}^i(M, X) = 0$ for all $i \neq \operatorname{cd}_{\mathfrak{a}}(M, X)$. Thus the assertion follows from Theorem 2.1. □

Corollary 2.3. *Let R be a local Cohen-Macaulay ring with maximal ideal $\mathfrak{m}$ and M an R-module with a finite free resolution. Then*

$$\beta_s(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(M, R)) \leq \dim_\kappa(\operatorname{Ext}_R^{\dim(R)-s}(M, \kappa))$$

for all $s \geq 0$. In particular, $\beta_s(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(M, R)) = 0$ for all $s > \dim(R)$.

Proof. By [8, Corollary 2.9] and [6, Corollary 3.2], $\mathrm{H}_{\mathfrak{m}}^i(M, R) = 0$ for all $i \neq \dim(R)$. Thus the assertion follows from Theorem 2.1. □

The next results follow by putting $M = R$ in Theorem 2.1, Corollary 2.2, and Corollary 2.3.

Corollary 2.4. (see [7, Theorem 3.2]) *Let R be a local ring and X an R-module. Then the inequality*

$$\beta_s(\mathrm{H}_{\mathfrak{a}}^t(X)) \leq \beta_{s-t}(X) + \sum_{i=0}^{t-1} \beta_{s-t+i-1}(\mathrm{H}_{\mathfrak{a}}^i(X)) + \sum_{i=t+1}^{\operatorname{cd}_{\mathfrak{a}}(X)} \beta_{s-t+i+1}(\mathrm{H}_{\mathfrak{a}}^i(X))$$

holds for all non-negative integers s and t.

Corollary 2.5. *Let R be a local ring and X a finite R-module such that* $\operatorname{depth}_R(\mathfrak{a}, X) = \operatorname{cd}_{\mathfrak{a}}(X)$*. Then the inequality*

$$\beta_s(\mathrm{H}_{\mathfrak{a}}^{\operatorname{cd}_{\mathfrak{a}}(X)}(X)) \leq \beta_{s-\operatorname{cd}_{\mathfrak{a}}(X)}(X)$$

is true for all $s \geq 0$. In particular, $\beta_s(\mathrm{H}_{\mathfrak{a}}^{\operatorname{cd}_{\mathfrak{a}}(X)}(X)) = 0$ for all $s < \operatorname{cd}_{\mathfrak{a}}(X)$.

Corollary 2.6. *Let R be a local Cohen-Macaulay ring with maximal ideal $\mathfrak{m}$. Then*

$$\beta_s(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(R)) \leq \beta_{s-\dim(R)}(R)$$

for all $s \geq 0$. That means $\beta_{\dim(R)}(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(R)) \leq 1$ and $\beta_s(\mathrm{H}_{\mathfrak{m}}^{\dim(R)}(R)) = 0$ for all $s \neq \dim(R)$.

References

[1] M. P. Brodmann and R. Y. Sharp, Local cohomology: an algebraic introduction with geometric applications, Cambridge Stud. Adv. Math., 60, *Cambridge University Press*, *Cambridge*, 1998. MR1613627.

[2] W. Bruns and J. Herzog, Cohen-Macaulay rings, Cambridge Stud. Adv. Math., 39, *Cambridge University Press*, *Cambridge*, 1993. MR1251956.

[3] J. Herzog, Komplexe, auflösungen und dualität in der lokalen Algebra, Habilitationsschrift, *Universitat Regensburg*, *Regensburg*, 1970.

[4] T. T. Nam, N. M. Tri, and N. V. Dong, Some properties of generalized local cohomology modules with respect to a pair of ideals, *Internat. J. Algebra Comput.* **24** (2014), no. 7, 1043–1054. MR3286150.

[5] J. J. Rotman, An introduction to homological algebra, Second edition, Universitext, *Springer*, *New York*, 2009. MR2455920.

[6] N. Tu Cuong and N. Van Hoang, On the vanishing and the finiteness of supports of generalized local cohomology modules, *Manuscripta Math.* **126** (2008), no. 1, 59–72. MR2395249.

[7] A. Vahidi, Betti numbers and flat dimensions of local cohomology modules, *Canad. Math. Bull.* **58** (2015), no. 3, 664–672. MR3372880.

[8] A. Vahidi and M. Aghapournahr, Some results on generalized local cohomology modules, *Comm. Algebra* **43** (2015), no. 5, 2214–2230. MR3316846.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–202, pp. 392–396

Cofiniteness of generalized local cohomology modules with respect to Serre categories of modules

Alireza Vahidi

Department of Mathematics, Payame Noor University, Tehran, Iran

Article Info

Keywords:
Cofinite modules
Generalized local cohomology modules

2020 MSC:
13D45

Abstract

Let R be a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ an ideal of R, $\mathcal{S}$ a Serre subcategory of the category of R-modules, M a finitely generated R-module, X an arbitrary R-module, and t a non-negative integer. In this paper, we show that if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X) \in \mathcal{S}$ for all i and $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all $i \neq t$, then $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module. In particular, if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X)$ is a finitely generated R-module for all $i \leq \mathrm{ara}(\mathfrak{a})$ and $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ is an $\mathfrak{a}$-cofinite R-module for all $i \neq t$, then $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ is an $\mathfrak{a}$-cofinite R-module.

1. Introduction

Throughout, let R denote a commutative Noetherian ring with non-zero identity, $\mathfrak{a}$ an ideal of R, M and N finite (i.e., finitely generated) R-modules, X an arbitrary R-module which is not necessarily finite, and t a non-negative integer. For basic results, notations, and terminology not given in this paper, readers are referred to [3, 4, 11].
It is a known fact that $\mathrm{H}^i_{\mathfrak{m}}(N)$ is an Artinian R-module and hence $\mathrm{Hom}_R(R/\mathfrak{m}, \mathrm{H}^i_{\mathfrak{m}}(N))$ is a finite R-module for all i whenever R is a local ring with maximal ideal $\mathfrak{m}$. This led Grothendieck to conjecture that $\mathrm{Hom}_R(R/\mathfrak{a}, \mathrm{H}^i_{\mathfrak{a}}(N))$ is a finite R-module for all i [6, Expose XIII, Conjecture 1.1]. Hartshorne showed in [7, Section 3] that this conjecture is not true in general. However, he defined an $\mathfrak{a}$-torsion R-module X to be $\mathfrak{a}$-*cofinite* if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X)$ is a finite R-module for all i and asked the following question:
(see [7, First Question]) Under what hypotheses, is $\mathrm{H}^i_{\mathfrak{a}}(N)$ an $\mathfrak{a}$-cofinite R-module for all i?
Recall that a subcategory of the category of R-modules is said to be *Serre* if it is closed under taking submodules, quotients, and extensions. The class of zero R-modules, the class of finite R-modules, and the class of Artinian R-modules are some examples of Serre subcategories of the category of R-modules. In this paper, $\mathcal{S}$ stands for a Serre subcategory of the category of R-modules. We say that X is an $(\mathcal{S}, \mathfrak{a})$-*cofinite* R-*module* if X is an $\mathfrak{a}$-torsion R-module and $\mathrm{Ext}^i_R(R/\mathfrak{a}, X) \in \mathcal{S}$ for all i [1, Definition 4.1]. Note that, if $\mathcal{S}$ is the class of finite R-modules, then X is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module if and only if X is an $\mathfrak{a}$-cofinite R-module. Therefore, as a generalization of Question 1, we have the following question.
Under what hypotheses, is $\mathrm{H}^i_{\mathfrak{a}}(N)$ an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all i?

Email address: vahidi.ar@pnu.ac.ir (Alireza Vahidi)

In [9, Proposition 2.5], Marley and Vassilev proved that $\mathrm{H}^t_{\mathfrak{a}}(N)$ is an $\mathfrak{a}$-cofinite R-module whenever $\mathrm{H}^i_{\mathfrak{a}}(N)$ is an $\mathfrak{a}$-cofinite R-module for all $i \neq t$. As an improvement of [9, Proposition 2.5], by [10, Proposition 3.11], $\mathrm{H}^t_{\mathfrak{a}}(X)$ is an $\mathfrak{a}$-cofinite R-module if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X)$ is a finite R-module for all i and $\mathrm{H}^i_{\mathfrak{a}}(X)$ is an $\mathfrak{a}$-cofinite R-module for all $i \neq t$. As a generalization of [10, Proposition 3.11], the authors in [1, Theorem 4.2] showed that $\mathrm{H}^t_{\mathfrak{a}}(X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module when $\mathrm{Ext}^i_R(R/\mathfrak{a}, X) \in \mathcal{S}$ for all i and $\mathrm{H}^i_{\mathfrak{a}}(X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all $i \neq t$.
The ith generalized local cohomology module

$$\mathrm{H}^i_{\mathfrak{a}}(M, X) \cong \varinjlim_{n\in\mathbb{N}} \mathrm{Ext}^i_R(M/\mathfrak{a}^n M, X)$$

was introduced by Herzog in [8]. It is clear that $\mathrm{H}^i_{\mathfrak{a}}(R, X)$ is just the ordinary local cohomology module $\mathrm{H}^i_{\mathfrak{a}}(X)$ of X with respect to $\mathfrak{a}$. As generalizations of Questions 1 and 1, we have the following questions.
Under what hypotheses, is $\mathrm{H}^i_{\mathfrak{a}}(M, N)$ an $\mathfrak{a}$-cofinite R-module for all i?
Under what hypotheses, is $\mathrm{H}^i_{\mathfrak{a}}(M, N)$ an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all i?
In this paper, we study the above questions. We generalize [1, Theorem 4.2] by showing that if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X) \in \mathcal{S}$ for all i and $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all $i \neq t$, then $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module. In particular, if $\mathrm{Ext}^i_R(R/\mathfrak{a}, X)$ is a finite R-module for all $i \leq \mathrm{ara}(\mathfrak{a})$ and $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ is an $\mathfrak{a}$-cofinite R-module for all $i \neq t$, then $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ is an $\mathfrak{a}$-cofinite R-module. Here, $\mathrm{ara}(\mathfrak{a})$ is the arithmetic rank of $\mathfrak{a}$.

2. Main Results

The following lemma is needed in this paper

Lemma 2.1. *Let N be a finite R-module, X an arbitrary R-module, and t a non-negative integer such that* $\mathrm{Ext}^i_R(N, X) \in \mathcal{S}$ *for all $i \leq t$. Then* $\mathrm{Ext}^i_R(L, X) \in \mathcal{S}$ *for each finite R-module L with* $\mathrm{Supp}_R(L) \subseteq \mathrm{Supp}_R(N)$ *and for all $i \leq t$.*

Proof. Let L be a finite R-module such that $\mathrm{Supp}_R(L) \subseteq \mathrm{Supp}_R(N)$. By Gruson's theorem [13, Theorem 4.1], there is a finite filtration

$$0 = L_0 \subset L_1 \subset \cdots \subset L_{n-1} \subset L_n = L$$

of submodules of L such that for all $1 \leq j \leq n$, there exists a short exact sequence

$$0 \longrightarrow L'_j \longrightarrow N^{\alpha_j} \longrightarrow L_j/L_{j-1} \longrightarrow 0,$$

where L'_j is a finite R-module with $\mathrm{Supp}_R(L'_j) \subseteq \mathrm{Supp}_R(N)$ and α_j is an integer. Thus we get the long exact sequence

$$\begin{array}{lllllll} 0 & \longrightarrow & \mathrm{Hom}_R(L_j/L_{j-1}, X) & \longrightarrow & \mathrm{Hom}_R(N^{\alpha_j}, X) & \longrightarrow & \mathrm{Hom}_R(L'_j, X) \\ & \longrightarrow & \cdots & & & & \\ & \longrightarrow & \mathrm{Ext}^{i-1}_R(L_j/L_{j-1}, X) & \longrightarrow & \mathrm{Ext}^{i-1}_R(N^{\alpha_j}, X) & \longrightarrow & \mathrm{Ext}^{i-1}_R(L'_j, X) \\ & \longrightarrow & \mathrm{Ext}^{i}_R(L_j/L_{j-1}, X) & \longrightarrow & \mathrm{Ext}^{i}_R(N^{\alpha_j}, X) & \longrightarrow & \mathrm{Ext}^{i}_R(L'_j, X) \\ & \longrightarrow & \cdots & & & & \end{array}$$

for all $1 \leq j \leq n$. Also, by the short exact sequences

$$0 \longrightarrow L_{j-1} \longrightarrow L_j \longrightarrow L_j/L_{j-1} \longrightarrow 0,$$

we get the long exact sequences

$$\begin{array}{lllllll} 0 & \longrightarrow & \mathrm{Hom}_R(L_j/L_{j-1}, X) & \longrightarrow & \mathrm{Hom}_R(L_j, X) & \longrightarrow & \mathrm{Hom}_R(L_{j-1}, X) \\ & \longrightarrow & \cdots & & & & \\ & \longrightarrow & \mathrm{Ext}^{i}_R(L_j/L_{j-1}, X) & \longrightarrow & \mathrm{Ext}^{i}_R(L_j, X) & \longrightarrow & \mathrm{Ext}^{i}_R(L_{j-1}, X) \\ & \longrightarrow & \cdots & & & & \end{array}$$

for all $1 \leq j \leq n$.

We use induction on t. Let $t = 0$ and $1 \leq j \leq n$. Since $\mathrm{Hom}_R(N, X) \in \mathcal{S}$, $\mathrm{Hom}_R(N^{\alpha_j}, X) \in \mathcal{S}$. Thus $\mathrm{Hom}_R(L_j/L_{j-1}, X) \in \mathcal{S}$. Therefore $\mathrm{Hom}_R(L_j, X) \in \mathcal{S}$ whenever $\mathrm{Hom}_R(L_{j-1}, X) \in \mathcal{S}$. Hence $\mathrm{Hom}_R(L, X) \in \mathcal{S}$. Suppose that $t > 0$ and that the case $t-1$ is settled. Let i and j be integers such that $i \leq t$ and $1 \leq j \leq n$. Since $\mathrm{Ext}^i_R(N, X) \in \mathcal{S}$, $\mathrm{Ext}^i_R(N^{\alpha_j}, X) \in \mathcal{S}$. Therefore $\mathrm{Ext}^i_R(L_j/L_{j-1}, X) \in \mathcal{S}$ because $\mathrm{Ext}^{i-1}_R(L'_j, X) \in \mathcal{S}$ from the induction hypothesis. Hence $\mathrm{Ext}^i_R(L_j, X) \in \mathcal{S}$ whenever $\mathrm{Ext}^i_R(L_{j-1}, X) \in \mathcal{S}$. Thus $\mathrm{Ext}^t_R(L, X) \in \mathcal{S}$. □

The following theorem is the main result of this paper which generalizes [1, Theorem 4.2]. In the proof, we use the isomorphism

$$\mathrm{H}^i_{\mathfrak{a}}(M, X) \cong \mathrm{H}^i(\mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))),$$

for all i, where $\mathrm{E}^{\bullet X}$ is a deleted injective resolution of X (see [5, Lemma 2.1(i)]).

Theorem 2.2. *Let M be a finite R-module, X an arbitrary R-module such that $\mathrm{Ext}^i_R(R/\mathfrak{a}, X) \in \mathcal{S}$ for all i, and t a non-negative integer such that $\mathrm{H}^i_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module for all $i \neq t$. Then $\mathrm{H}^t_{\mathfrak{a}}(M, X)$ is an $(\mathcal{S}, \mathfrak{a})$-cofinite R-module.*

Proof. Let s be a non-negative integer. We show that $\mathrm{Ext}^s_R(R/\mathfrak{a}, \mathrm{H}^t_{\mathfrak{a}}(M, X)) \in \mathcal{S}$. Assume that

$$\mathrm{E}^{\bullet X} : 0 \longrightarrow E^0 \longrightarrow \cdots \longrightarrow E^i \longrightarrow \cdots$$

is a deleted injective resolution of X. By applying $\mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(-))$ to $\mathrm{E}^{\bullet X}$, we get the complex

$$\mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X})) : 0 \longrightarrow \mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(E^0)) \longrightarrow \cdots \longrightarrow \mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(E^i)) \longrightarrow \cdots.$$

Assume that

$$0 \longrightarrow \mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X})) \longrightarrow T^{\bullet,0} \longrightarrow \cdots \longrightarrow T^{\bullet,j} \longrightarrow \cdots$$

is a Cartan-Eilenberg injective resolution of $\mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))$ which exists from [11, Theorem 10.45]. Now, consider the third quadrant bicomplex $\mathcal{T} = \{\mathrm{Hom}_R(R/\mathfrak{a}, T^{p,q})\}$. We denote the total complex of $\mathcal{T}$ by $\mathrm{Tot}(\mathcal{T})$. The first filtration has ${}^{I}\mathrm{E}_2$ term the iterated homology $\mathrm{H}'^{p}\,\mathrm{H}''^{p,q}(\mathcal{T})$. We have

$$\begin{aligned} \mathrm{H}''^{p,q}(\mathcal{T}) &\cong \mathrm{H}^q(\mathrm{Hom}_R(R/\mathfrak{a}, T^{p,\bullet})) \\ &\cong \mathrm{Ext}^q_R(R/\mathfrak{a}, \mathrm{Hom}_R(M, \Gamma_{\mathfrak{a}}(\mathrm{E}^p))) \\ &\cong \mathrm{Hom}_R(\mathrm{Tor}^R_q(R/\mathfrak{a}, M), \Gamma_{\mathfrak{a}}(\mathrm{E}^p)) \end{aligned}$$

from [3, Proposition 2.1.4] and [11, Corollary 10.63]. Therefore, by [5, Lemma 2.1(i)],

$$\begin{aligned} {}^{I}\mathrm{E}_2^{p,q} &\cong \mathrm{H}'^{p}\,\mathrm{H}''^{p,q}(\mathcal{T}) \\ &\cong \mathrm{H}^p(\mathrm{Hom}_R(\mathrm{Tor}^R_q(R/\mathfrak{a}, M), \Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))) \\ &\cong \mathrm{H}^p_{\mathfrak{a}}(\mathrm{Tor}^R_q(R/\mathfrak{a}, M), X) \end{aligned}$$

which yields, by [12, Lemma 2.5(c)], the third quadrant spectral sequence

$${}^{I}\mathrm{E}_2^{p,q} := \mathrm{Ext}^p_R(\mathrm{Tor}^R_q(R/\mathfrak{a}, M), X) \underset{p}{\Longrightarrow} \mathrm{H}^{p+q}(\mathrm{Tot}(\mathcal{T})).$$

For all $i \leq s+t$, we have ${}^{I}E_{\infty}^{s+t-i,i} = {}^{I}E_{s+t+2}^{s+t-i,i}$ because ${}^{I}E_j^{s+t-i-j,i-1+j} = 0 = {}^{I}E_j^{s+t-i+j,i+1-j}$ for all $j \geq s+t+2$; so that ${}^{I}E_{\infty}^{s+t-i,i}$ is in $\mathcal{S}$ from the fact that ${}^{I}E_{s+t+2}^{s+t-i,i}$ is a subquotient of ${}^{I}E_2^{s+t-i,i}$ which is in $\mathcal{S}$ by assumption and Lemma 2.1. There exists a finite filtration

$$0 = \phi^{s+t+1}H^{s+t} \subseteq \phi^{s+t}H^{s+t} \subseteq \cdots \subseteq \phi^1 H^{s+t} \subseteq \phi^0 H^{s+t} = H^{s+t}$$

such that ${}^{I}E_{\infty}^{s+t-i,i} \cong \phi^{s+t-i}H^{s+t}/\phi^{s+t-i+1}H^{s+t}$ for all $i \leq s+t$. Now the exact sequences

$$0 \longrightarrow \phi^{s+t-i+1}H^{s+t} \longrightarrow \phi^{s+t-i}H^{s+t} \longrightarrow {}^{I}E_{\infty}^{s+t-i,i} \longrightarrow 0,$$

for all $i \leq s+t$, show that H^{s+t} is in $\mathcal{S}$.

On the other hand, the second filtration has ${}^{II}\,\mathrm{E}_2$ term the iterated homology $\mathrm{H}''^{p}\,\mathrm{H}'^{q,p}(\mathcal{T})$. Note that every short exact sequence of injective modules splits and so it remains split after applying the functor $\mathrm{Hom}_R(R/\mathfrak{a},-)$. By using this fact and the fact that $T^{\bullet,\bullet}$ is a Cartan-Eilenberg injective resolution of $\mathrm{Hom}_R(M,\Gamma_{\mathfrak{a}}(\mathrm{E}^{\bullet X}))$, we get

$$\begin{aligned}\mathrm{H}'^{q,p}(\mathcal{T}) &\cong \mathrm{H}^q(\mathrm{Hom}_R(R/\mathfrak{a},T^{\bullet,p}))\\ &\cong \mathrm{Hom}_R(R/\mathfrak{a},\mathrm{H}^q(T^{\bullet,p}))\\ &\cong \mathrm{Hom}_R(R/\mathfrak{a},\mathrm{H}^{q,p}).\end{aligned}$$

Therefore, by [5, Lemma 2.1(i)],

$$\begin{aligned}{}^{II}\,\mathrm{E}_2^{p,q} &\cong \mathrm{H}''^{p}\,\mathrm{H}'^{q,p}(\mathcal{T})\\ &\cong \mathrm{H}^p(\mathrm{Hom}_R(R/\mathfrak{a},\mathrm{H}^{q,\bullet}))\\ &\cong \mathrm{Ext}_R^p(R/\mathfrak{a},\mathrm{H}_{\mathfrak{a}}^q(M,X))\end{aligned}$$

which gives the third quadrant spectral sequence

$${}^{II}\,\mathrm{E}_2^{p,q} := \mathrm{Ext}_R^p(R/\mathfrak{a},\mathrm{H}_{\mathfrak{a}}^q(M,X)) \underset{p}{\Longrightarrow} \mathrm{H}^{p+q}(\mathrm{Tot}(\mathcal{T})).$$

Thus there exists a finite filtration

$$0=\psi^{s+t+1}H^{s+t}\subseteq\psi^{s+t}H^{s+t}\subseteq\cdots\subseteq\psi^1H^{s+t}\subseteq\psi^0H^{s+t}=H^{s+t}$$

such that ${}^{II}E_\infty^{s+t-i,i}\cong\psi^{s+t-i}H^{s+t}/\psi^{s+t-i+1}H^{s+t}$ for all $i\leq s+t$. Since H^{s+t} is in $\mathcal{S}$, ψ^sH^{s+t} is in $\mathcal{S}$. Hence ${}^{II}E_\infty^{s,t}\cong\psi^sH^{s+t}/\psi^{s+1}H^{s+t}$ is in $\mathcal{S}$. Therefore ${}^{II}E_{s+t+2}^{s,t}$ is in $\mathcal{S}$ because ${}^{II}E_{s+t+2}^{s,t}={}^{II}E_\infty^{s,t}$ from the fact that ${}^{II}E_j^{s-j,t-1+j}=0={}^{II}E_j^{s+j,t+1-j}$ for all $j\geq s+t+2$. For all $r\geq2$, let ${}^{II}Z_r^{s,t}=\mathrm{Ker}({}^{II}E_r^{s,t}\longrightarrow{}^{II}E_r^{s+r,t+1-r})$ and ${}^{II}B_r^{s,t}=\mathrm{Im}({}^{II}E_r^{s-r,t-1+r}\longrightarrow{}^{II}E_r^{s,t})$. We have the exact sequences

$$0\longrightarrow{}^{II}Z_r^{s,t}\longrightarrow{}^{II}E_r^{s,t}\longrightarrow{}^{II}E_r^{s,t}/{}^{II}Z_r^{s,t}\longrightarrow0$$

and

$$0\longrightarrow{}^{II}B_r^{s,t}\longrightarrow{}^{II}Z_r^{s,t}\longrightarrow{}^{II}E_{r+1}^{s,t}\longrightarrow0.$$

Since ${}^{II}E_2^{s+r,t+1-r}$ and ${}^{II}E_2^{s-r,t-1+r}$ are in $\mathcal{S}$ by assumptions, ${}^{II}E_r^{s+r,t+1-r}$ and ${}^{II}E_r^{s-r,t-1+r}$ are also in $\mathcal{S}$, and so ${}^{II}E_r^{s,t}/{}^{II}Z_r^{s,t}$ and ${}^{II}B_r^{s,t}$ are in $\mathcal{S}$. It shows that ${}^{II}E_r^{s,t}$ is in $\mathcal{S}$ whenever ${}^{II}E_{r+1}^{s,t}$ is in $\mathcal{S}$. Therefore ${}^{II}E_2^{s,t}=\mathrm{Ext}_R^s(R/\mathfrak{a},\mathrm{H}_{\mathfrak{a}}^t(M,X))\in\mathcal{S}$ which completes the proof. □

As applications of the above theorem, we have the following corollaries.

Corollary 2.3. *Let M be a finite R-module, X an arbitrary R-module such that $\mathrm{Ext}_R^i(R/\mathfrak{a},X)$ is a finite R-module for all $i\leq\mathrm{ara}(\mathfrak{a})$, and t a non-negative integer such that $\mathrm{H}_{\mathfrak{a}}^i(M,X)$ is an $\mathfrak{a}$-cofinite R-module for all $i\neq t$. Then $\mathrm{H}_{\mathfrak{a}}^t(M,X)$ is an $\mathfrak{a}$-cofinite R-module.*

Proof. Follows from Theorem 2.2 and [2, Theorem 3.3]. □

For a finite R-module M with finite projective dimension and for an arbitrary R-module X, we denote the largest integer i in which $\mathrm{H}_{\mathfrak{a}}^i(M,X)$ is not zero by $\mathrm{cd}_{\mathfrak{a}}(M,X)$. We also write $\mathrm{cd}_{\mathfrak{a}}(X)=\mathrm{cd}_{\mathfrak{a}}(R,X)$.

Corollary 2.4. *Let M be a finite R-module with finite projective dimension and X a finite R-module such that $\mathrm{depth}_R(\mathfrak{a}+\mathrm{Ann}_R(M),X)=\mathrm{cd}_{\mathfrak{a}}(M,X)$. Then $\mathrm{H}_{\mathfrak{a}}^i(M,X)$ is an $\mathfrak{a}$-cofinite R-module for all i.*

Proof. By [12, Corollary 2.14], $\mathrm{H}_{\mathfrak{a}}^i(M,X)=0$ for all $i\neq\mathrm{cd}_{\mathfrak{a}}(M,X)$. Thus the assertion follows from Corollary 2.3. □

The next results follow by putting $M=R$ in Theorem 2.2, Corollary 2.3, and Corollary 2.4.

Corollary 2.5. (see [1, Theorem 4.2]) *Let X be an arbitrary R-module such that $\mathrm{Ext}_R^i(R/\mathfrak{a},X)\in\mathcal{S}$ for all i and t a non-negative integer such that $\mathrm{H}_{\mathfrak{a}}^i(X)$ is an $(\mathcal{S},\mathfrak{a})$-cofinite R-module for all $i\neq t$. Then $\mathrm{H}_{\mathfrak{a}}^t(X)$ is an $(\mathcal{S},\mathfrak{a})$-cofinite R-module.*

Corollary 2.6. (see [10, Proposition 3.11]) *Let X be an arbitrary R-module such that $\mathrm{Ext}^i_R(R/\mathfrak{a}, X)$ is a finite R-module for all $i \leq \mathrm{ara}(\mathfrak{a})$ and t a non-negative integer such that $\mathrm{H}^i_{\mathfrak{a}}(X)$ is an $\mathfrak{a}$-cofinite R-module for all $i \neq t$. Then $\mathrm{H}^t_{\mathfrak{a}}(X)$ is an $\mathfrak{a}$-cofinite R-module.*

Corollary 2.7. *Let X be a finite R-module such that $\mathrm{depth}_R(\mathfrak{a}, X) = \mathrm{cd}_{\mathfrak{a}}(X)$. Then $\mathrm{H}^i_{\mathfrak{a}}(X)$ is an $\mathfrak{a}$-cofinite R-module for all i.*

References

[1] M. Aghapournahr, A. J. Taherizadeh, and A. Vahidi, Extension functors of local cohomology modules, *Bull. Iranian Math. Soc.* **37** (2011), no. 3, 117–134. MR2901589.
[2] K. Bahmanpour and M. Aghapournahr, A note on cofinite modules, *Comm. Algebra* **44** (2016), no. 9, 3683–3691. MR3503378.
[3] M. P. Brodmann and R. Y. Sharp, Local cohomology: an algebraic introduction with geometric applications, Cambridge Stud. Adv. Math., 60, *Cambridge University Press*, *Cambridge*, 1998. MR1613627.
[4] W. Bruns and J. Herzog, Cohen-Macaulay rings, Cambridge Stud. Adv. Math., 39, *Cambridge University Press*, *Cambridge*, 1993. MR1251956.
[5] K. Divaani-Aazar, R. Sazeedeh, and M. Tousi, On vanishing of generalized local cohomology modules, *Algebra Colloq.* **12** (2005), no. 2, 213–218. MR2127246.
[6] A. Grothendieck, Cohomologie locale des faisceaux cohérents et théorèmes de Lefschetz locaux et globaux (SGA 2), (French) Augmenté d'un exposé par Michèle Raynaud, Séminaire de Géométrie Algébrique du Bois-Marie, 1962, Adv. Stud. Pure Math., Vol. 2, *North-Holland Publishing Co.*, *AmsterdamMasson* & *Cie, Éditeur, Paris*, 1968. MR0476737.
[7] R. Hartshorne, Affine duality and cofiniteness, *Invent. Math.* **9** (1970), 145–164. MR0257096.
[8] J. Herzog, Komplexe, auflösungen und dualität in der lokalen Algebra, Habilitationsschrift, *Universitat Regensburg*, *Regensburg*, 1970.
[9] T. Marley and J. C. Vassilev, Cofiniteness and associated primes of local cohomology modules, *J. Algebra* **256** (2002), no. 1, 180–193. MR1936885.
[10] L. Melkersson, Modules cofinite with respect to an ideal, *J. Algebra* **285** (2005), no. 2, 649–668. MR2125457.
[11] J. J. Rotman, An introduction to homological algebra, Second edition, Universitext, *Springer*, *New York*, 2009. MR2455920.
[12] A. Vahidi and M. Aghapournahr, Some results on generalized local cohomology modules, *Comm. Algebra* **43** (2015), no. 5, 2214–2230. MR3316846.
[13] W. V. Vasconcelos, Divisor theory in module categories, North-Holland Math. Stud., 14, *North-Holland Publishing Co.*, *Amsterdam-OxfordAmerican Elsevier Publishing Co.*, *Inc.*, *New York*, 1974. MR498530.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–207, pp. 397–399

A characterization on locally-primitive 1-transitive graphs of valency $3p$

Fateme Abedi[a,*]

[a] *Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran*

Article Info

Keywords:
Arc-transitive graph
s-Transitive graph
Locally-primitive graph

2020 MSC:
05C25
20B25

Abstract

Let X be a connected locally-primitive $(G, 1)$-transitive graph for some $G \leq Aut(X)$. In this paper, we determine the structure of the vertex-stabilizer G_v when X has valency $3p$ where $p \geq 5$ is a prime.

1. Introduction

Throughout this paper, we consider finite, undirected graphs without loops or multiple edges. For a graph X, let $V(X)$, $E(X)$ and $Aut(X)$ denote the vertex set, the edge set and the full automorphism group, respectively. For $u, v \in V(X)$, $\{u, v\}$ is the edge incident to u and v in X and $X_1(v)$ is the set of vertices adjacent to v in X. Let $G \leq Aut(X)$. We denote the vertex-stabilizer of $v \in V(X)$ in G by G_v. Denote by $G_v^{X_1(v)}$ the constituent of G_v acting on $X_1(v)$ and by G_v^* the kernel of G_v acting on $X_1(v)$. Then, $G_v^{X_1(v)} \cong G_v/G_v^*$. For an edge $\{u, v\} \in E(X)$, we consider $G_{uv} = G_u \cap G_v$ and $G_{uv}^* = G_u^* \cap G_v^*$.

For each integer $s \geq 0$, an *s-arc* in graph X is a sequence $(v_0, v_1, ..., v_{s-1}, v_s)$ of vertices such that v_{i-1} is adjacent to v_i and $v_{i-1} \neq v_{i+1}$ for all admissible i. For a subgroup $G \leq Aut(X)$, X is said to be *(G, s)-arc-transitive* if G is transitive on the set of s-arcs in X. A (G, s)-arc-transitive graph which is not $(G, s+1)$-arc-transitive is called *(G, s)-transitive*. A graph X is called *s-arc-transitive* or *s-transitive* if it is $(Aut(X), s)$-arc-transitive or $(Aut(X), s)$-transitive, respectively. In particular, X is called *symmetric* if it is $(Aut(X), 1)$-arc-transitive. A G-arc-transitive graph X is called *locally-primitive* if G_v acts on $X_1(v)$ primitively, that is, the induced permutation group $G_v^{X_1(v)}$ is primitive.

As we all know, a graph X is G-arc-transitive if and only if G is transitive on $V(X)$ and G_v is transitive on $X_1(v)$. So the structure of G_v plays an important role in the study of symmetric graphs and to investigate such graphs, we need the information of the vertex-stabilizer G_v. Let X be a connected (G, s)-transitive graph for some $s \geq 1$. It is a

*Talker
Email address: std_f.abedi@khu.ac.ir (Fateme Abedi)

well-known result that $s \leq 7$ and $s \neq 6$ [6]. Note that the only connected graphs of valency two are cycles which are s-arc-transitive for any positive integer s. So the valency of a s-transitive graph is greater than 2. Up to now, we know the structure of G_v when X has prime or twice a prime valency [4, 5]. In this paper, we characterize the structure of G_v when X is a locally-primitive 1-transitive graph of valency $3p$.

Let p a prime and n a positive integer. We denote by n the cyclic group of order n, by A_n and S_n the alternating group and the symmetric group of degree n. For two groups M and N, we denote by $N.M$ an extension of N by M and $N : M$ stands for a semidirect product of N by M. For any group G, denote the largest normal p-subgroup of G by $\mathbf{O}_p(G)$ All the notation and terminology used throughout this paper are standard. For group and graph theoretic concepts not defined here, we refer the reader to [1, 3].

2. Preliminaries

In this section, we collect some results which have an important role in characterizing the structure of G_v in symmetric graphs. The first proposition is about sufficient and necessary conditions for such graphs. Its proof is straightforward.

Proposition 2.1. *Let X be a graph and $G \leq Aut(X)$. Then we have;*

(i) *X is G-arc-transitive if and only if G is transitive on $V(X)$ and G_v is transitive on $X_1(v)$ for each $v \in V(X)$.*

(ii) *X is $(G, 2)$-arc-transitive if and only if G is transitive on $V(X)$ and G_v is 2-transitive on $X_1(v)$ for each $v \in V(X)$.*

Lemma 2.2. *Let X be a (G, s)-arc-transitive graph for some $G \leq Aut(X)$ and $s \geq 1$. Let $\{u, v\} \in E(X)$. Then we have;*

(i) $G_v \cong G_v^*.G_v^{X_1(v)} \cong (G_{uv}^*.(G_v^*)^{X_1(u)}).G_v^{X_1(v)}$.

(ii) $G_v^{*X_1(u)} \trianglelefteq G_{uv}^{X_1(u)} \cong G_{uv}^{X_1(v)}$.

Proof. (i) Since $G_v^{X_1(v)} \cong G_v/G_v^*$ and $(G_v^*)^{X_1(u)} \cong G_v^*/G_{uv}^*$, by group extensions we obtain the result.

(ii) Since G is transitive on the set of arcs, then $G_{uv}^{X_1(u)} \cong G_{uv}^{X_1(v)}$. By isomorphism we have, $(G_v^*)^{X_1(u)} \cong G_v^*/G_{uv}^* \cong G_v^*/G_u^* \cap G_v^* \cong G_u^*G_v^*/G_u^* \trianglelefteq G_{uv}/G_u^* \cong G_{uv}^{X_1(u)}$.

□

Combining a result from ([2], Corollary 2.3) and ([5], Theorem 4.3), we have the following lemma.

Lemma 2.3. *Let X be a connected graph, $\{u, v\} \in E(X)$ and $G \leq Aut(X)$. Suppose that X is a G-locally-primitive arc-transitive graph. Then G_{uv}^* is a r-group for some prime r. Moreover, either $G_{uv}^* = 1$ or $\mathbf{O}_r((G_v^*)^{X_1(u)}) \neq 1$ and $\mathbf{O}_r(({G_v}^{X_1(v)})_u) \neq 1$.*

In view of ([1], Section7.7), we have the following observation.

Lemma 2.4. *Let H be a primitive but not 2-transitive permutation group of degree $3p$ where $p \geq 5$ is a prime and let $\alpha \in \Omega$. Up to isomorphism H is one of the groups in Table 1.*

Table 1.

Degree	H	H_v
15	A_6	S_4
15	S_6	$S_4 \times 2$
21	$PGL_2(7)$	$PSL_2(7)$
21	A_7	S_5
21	S_7	$S_5 \times 2$
57	$PSL_2(19)$	A_5

3. Main results

Let X be a connected locally-primitive $(G, 1)$-transitive graph of valency $3p$ for some $G \leq Aut(X)$ and $p \geq 5$ a prime. According to Proposition 2.1, we get that $G_v^{X_1(v)} \cong G_v/G_v^* \leq S_{3p}$ is a primitive but not 2-transitive group of degree $3p$. So by Lemma 2.4, we know the possibilities for p and $G_v^{X_1(v)}$. Hence the main result of the paper is as follows.

Theorem 3.1. *Let X be a finite connected locally-primitive $(G, 1)$-transitive graph of valency $3p$ where $G \leq Aut(X)$ and $p \geq 5$ is a prime. Let $v \in V(X)$. Then one of the following holds.*

(i) $p = 5$, $G_v \cong A_6$, S_6, $\mathbf{O}_2(G_v).N.A_6$ *or* $\mathbf{O}_2(G_v).N.S_6$ *where* $N \trianglelefteq S_3$.
(ii) $p = 7$, $G_v \cong PGL_2(7)$ *or* $PSL_2(7).PGL_2(7)$.
(iii) $p = 7$, $G_v \cong A_7$, S_7, $A_5.A_7$, $S_5.A_7$, $A_5.S_7$, $S_5.S_7$ *or* $\mathbf{O}_2(G_v).N.S_7$ *where* $N \trianglelefteq S_5$.
(iv) $p = 19$, $G_v \cong PSL_2(19)$ *or* $A_5.PSL_2(19)$.

References

[1] J. Dixon and B. Mortimer, *Permutation Groups*, Springer-Verlag, New York, Berlin, 1996.
[2] A. Gardiner, *Arc transitivity in graphs*, Q. J. Math. **24** (1973), 399-407.
[3] C. Godsil and G.F. Royle, *Algebraic Graph Theory*, Springer, New York, 2001.
[4] J.J. Li, B. Ling and G.D. Liu, *A characterization on arc-transitive graphs of prime valency*, Applied Mathematics and Computation, 325 (2018), 227-233.
[5] G. Li, Z. Lu and X, Zhang, *Locally-primitive arc-transitive 10-valent graphs of square-free order*, Algebra Colloq. **25** (02) (2018), 243-264.
[6] R.M. Weiss, *The nonexistence of 8-transitive graphs*, Combinatorica 1 (1981) 309-311.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–224, pp. 400–403

On real closed domains

Amir Veisi[a,*], Zahra Keshtkar[b]

[a] *Faculty of Petroleum and Gas, Yasouj University, Gachsaran, Iran*
[b] *Faculty of Petroleum and Gas, Yasouj University, Gachsaran, Iran*

Article Info

Keywords:
totally ordered domain
divisibility property
real closed rings

2020 MSC:
13G05
06F25

Abstract

In this article, some results and corollaries on real closed domains are derived. We show that the field of fractions F_D of an ordered domain D is real closed when D is real closed. It is observed that an ordered domain has divisibility property if and only if it is a convex subring of F_D. It is also shown that an ordered domain is real closed if and only if it is a convex subring of a real closed field.

1. Introduction

Let $(D, \leq)$ (in brief, D) be a commutative totally ordered integral domain with unity. Briefly, D is called an ordered domain. Define a relation $\sim$ on $D \times D$ by $(a, b) \sim (c, d) \Leftrightarrow ad = bc$. It is observed that $\sim$ is an equivalence relation on $D \times D$. If we let $[(a, b)]$ be the equivalence class of (a, b) with respect to $\sim$, then we have $[(a, b)] = \{(c, d) : (c, d) \sim (a, b)\}$. For simplicity, $D \times D$ is denoted by F_D and $[(a, b)]$ is also denoted by $\frac{a}{b}$. In the recent fraction, b can be taken as a positive element of D. Hence, we have $F_D = \{\frac{a}{b} : a, b \in D, \text{ and } b > 0\}$. It is easy to check that $(F_D, +, \times)$ has a field structure, where $+$ and $\times$ are defined as follows:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}, \text{ and } \frac{a}{b} \times \frac{c}{d} = \frac{ac}{bd}.$$

Note that F_D is the smallest field containing D, i.e., if F is a field containing D, then it contains F_D. Moreover, define $\leq'$ on F_D by $\frac{a}{b} \leq' \frac{c}{d} \Leftrightarrow ad \leq bc$. By the fact that D is an ordered domain, one can show that $(F_D, \leq')$ (in brief, F_D) is also an ordered domain. As usual, $\mathbb{Z}$, $\mathbb{Q}$, and $\mathbb{R}$ are the sets of integers, rational numbers, and real numbers respectively. In Section 2, we show that F_D is real closed if D is real closed. We observed that D has divisibility property (in brief, DP) if and only if D is a convex subring of F_D. It is also proved that D is real closed if and only if D is a convex subring of a real closed field. Moreover, it is equivalent to: every nonnegative element of D has a square root in D, and every monic polynomial of odd degree in $D[x]$ has a root in D, if D satisfies DP. Real closed fields are introduced and investigated in [5, Chapter 13], and real closed domains are also studied in [1], [2], [3], and [4]. Throughout this article, D is an ordered domain.

*Talker
Email addresses: aveisi@yu.ac.ir (Amir Veisi), zahra.keshtkar@gmail.com (Zahra Keshtkar)

2. Topics and Results

Definition 2.1. ([2, Definition 1]) A domain D is called real closed if it satisfies the intermediate value property for polynomials in $D[x]$, i.e., for any polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0 \in D[x]$ and any $a, b \in D$; $a < b$ and $p(a)p(b) < 0$ gives $p(c) = 0$ for some $c \in D$ such that $a < c < b$.

Example 2.2. The field $\mathbb{R}$ is real closed while $\mathbb{Z}$ and $\mathbb{Q}$ are not real closed. To see this, let $p(x) = 2x - 1 \in \mathbb{Z}[x]$ on $[0,\ 1]$. Then $p(0)p(1) < 0$ but the root of p does not belong to $\mathbb{Z}$. Also, for $q(x) = x^3 - 2 \in \mathbb{Q}[x]$ on $[0,\ 2]$ we have $q(0)q(2) < 0$, but q does not have a root in $\mathbb{Q}$.

Definition 2.3. ([2, Definition 20]) A domain D is called satisfying *Divisibility Property* (in brief, DP) whenever $0 < a < b$ gives $b \mid a$, i.e., $b = ac$ for some $c \in D$. In other words, $a \in (b)$ where (b) is the generated ideal by b.

For instance, every ordered division ring D satisfies DP, since if $a, b \in D$ and $0 < a < b$, then $a = (ab^{-1})b$. So $b \mid a$. Similarly, $a \mid b$. In fact, each of the two nonzero elements of a division ring divides the other.

Remark 2.4. (i) Every ordered field satisfies DP.
(ii) For any domain D; F_D is the smallest field containing D.
(iii) A subring of a ring with divisibility property (DP) need not necessarily satisfy the same property, for example, consider $\mathbb{Z}$ and $\mathbb{Q}$.

Proposition 2.5. *If D satisfies DP, then for every $c \in F_D$; $c \in D$ or $c^{-1} \in D$.*

Proof. Let $0 < c = \frac{m}{n} \in F$. If $0 < m < n$, then by hypothesis $n \mid m$. So $m = nk$ for some $k \in D$ and hence $c = \frac{m}{n} = k \in D$. If $0 < n < m$, then $m \mid n$. So $c^{-1} = \frac{n}{m} \in D$. □

Proposition 2.6. *Let D be a real closed domain. Then the following hold.*

(i) D satisfies DP.
(ii) Any nonnegative element $a \in D$ has a square root in D, i.e., $b^2 = a$ for some $b \in D$.
(iii) Any monic polynomial of odd degree of $D[x]$ has a root in D.
(iv) The field of fractions F_D of D is real closed.

Proof. (i). Let $a, b \in D$ and $0 < a < b$. Consider polynomial $p(x) = bx - a$ on $[0,\ 1]$. Then $p(0) = -a < 0$ and $p(1) = b - a > 0$. Since D is real closed, there exists $0 < c < 1$ such that $p(c) = bc - a = 0$. Hence, $bc = a$, i.e., $b \mid a$.
(ii). Let $a > 0$. Take $p(x) = x^2 - a \in D[x]$. If $0 < a < 1$, then values of p at the endpoints of the interval $[0,\ 1]$ are $p(0) = -a < 0$ and $p(1) = 1 - a > 0$, and if $a > 1$, then consider p on $[1\ \ a]$. So $p(1) < 0$ and $p(a) = a^2 - a > 0$. By intermediate value property, there exists $0 < c < 1$ or $1 < c < a$ such that $p(c) = c^2 - a = 0$, i.e. $c^2 = a$.
(iii). Let $p(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1 x + a_0 \in D[x]$ and n is an odd integer. Let

$$M = 1 + |a_{n-1}| + |a_{n-2}| + \cdots + |a_1| + |a_0|.$$

It is sufficient to get values of p at the endpoints of $[-M,\ M]$.
(iv). Let

$$p(x) = \tfrac{c_n}{d_n}x^n + \tfrac{c_{n-1}}{d_{n-1}}x^{n-1} + \tfrac{c_{n-2}}{d_{n-2}}x^{n-2} + \tfrac{c_{n-3}}{d_{n-3}}x^{n-3} + \cdots + \tfrac{c_1}{d_1}x + \tfrac{c_0}{d_0} \in F[x].$$

Take $\frac{m}{n}, \frac{m'}{n'} \in F_D$ such that $\frac{m}{n} < \frac{m'}{n'}$. Suppose $p(\frac{m}{n}) < 0$, and $p(\frac{m'}{n'}) > 0$. We are going to find an element $c \in F_D$ lying between $\frac{m}{n}$ and $\frac{m'}{n'}$ such that $p(c) = 0$. Put

$$d = d_0 d_1 d_2 \cdots d_{n-1} d_n \quad \text{and} \quad p_1(x) = d^n p(x).$$

Since $mn' < m'n$, we have $dmn' < dm'n$. Furthermore, $p_1(x) \in D[x]$; $p_1(\frac{m}{n}) < 0$ and $p_1(\frac{m'}{n'}) > 0$. Now, consider polynomial q as below:

$$q(x) = p_1(\frac{x}{dnn'}) = d^n p(\frac{x}{dnn'}).$$

Clearly, $q \in D[x]$. It is easy to see that for some $t \in D$; $q(t) = 0$. So $p(\frac{t}{dnn'}) = 0$, and we are done. □

A subset S of a partially ordered ring R is called *convex, if* $a \leq c \leq b$ *and* $a, b \in S$*, then* $c \in S$.

Proposition 2.7. *An ordered domain satisfies DP if and only if it is a convex subring of its field of fractions.*

Proof. $(\Rightarrow)$: Let D be an ordered domain satisfying DP and $0 < a = \frac{m}{n} < b$, where $b \in D$. Since $m < bn$, by hypothesis $bn \mid m$, i.e., $m = bnd$ for some $d \in D$. Hence, $a = \frac{m}{n} = bd \in D$.
$(\Leftarrow)$: Let $a, b \in D$ and $0 < a < b$. Then $0 < \frac{a}{b} < 1$ (note that $a < b$ if and only if $\frac{a}{b} < 1$). Since D is convex in F_D, we obtain $\frac{a}{b} \in D$. So $a = bd$ for some $d \in D$, i.e. D satisfies DP. □

Proposition 2.8. *If D satisfies property (i) or (ii) below*

(i) every nonnegative element of D has a square root in D,
(ii) every monic polynomial of odd degree in $D[X]$ has a root in D,

then F_D satisfies the same property.

Proof. Suppose first that D satisfies (i) and $0 < \frac{m}{n} \in F_D$. So $0 < mn \in D$ and hence there exists $c \in D$ such that $c^2 = mn$. Therefore, $(\frac{c}{n})^2 = \frac{c^2}{n^2} = \frac{mn}{n^2} = \frac{m}{n}$, i.e., $\frac{c}{n}$ is a square root of $\frac{m}{n}$, in other words, F_D satisfies (i). Next, suppose that D satisfies (ii), and

$$p(x) = x^n + \tfrac{c_{n-1}}{d_{n-1}} x^{n-1} + \tfrac{c_{n-2}}{d_{n-2}} x^{n-2} + \tfrac{c_{n-3}}{d_{n-3}} x^{n-3} + \cdots + \tfrac{c_1}{d_1} x + \tfrac{c_0}{d_0} \in F_D[x]$$

where n is an odd integer. Put $d = d_0 d_1 d_2 \ldots d_{n-1}$, and $q(x) = d^n p(\frac{x}{d})$.
Since q is a monic polynomial of $D[x]$, by hypothesis it has a root in D, say λ. So $q(\lambda) = d^n p(\frac{\lambda}{d}) = 0$. On the other hand, since D is domain and $d \neq 0$, we get $p(\frac{\lambda}{d}) = 0$, as desired. □

The converse of Proposition 2.8(i) is always true. But the converse of (ii) is not true in general.

Corollary 2.9. *If D satisfies properties (i) and (ii) in the above proposition, then F_D is real closed.*

According to Example 2.2, a subring of a real closed field need not necessarily be real closed. The following result gives an equivalent condition for the real closeness of an ordered domain.

Theorem 2.10. *A domain D is real closed if and only if it is a convex subring of a real closed field F.*

Proof. $(\Rightarrow)$: Since D is real closed, parts (i) and (iv) of Proposition 2.6 respectively imply that D satisfies DP, and F_D is real closed. On the other hand, by Proposition 2.7, D is a convex subring of F_D. So the result holds.
$(\Leftarrow)$: Let F be a real closed field in which D is convex. Let $p(x) \in D[x] \leq F[x]$ and $a, b \in D$ such that $a < b$ and $p(a)p(b) < 0$. Since F is real closed, there exists $c \in F$ such that $a < c < b$ and $p(c) = 0$. By convexity of D in F, we obtain $c \in D$, i.e., D is real closed. □

Corollary 2.11. *Suppose D satisfies DP. Then D is real closed if and only if F_D is real closed.*

Proof. Necessity: It follows from Proposition 2.6 (iv).
Sufficiency: Since D satisfies DP, D is convex in F_D, by Proposition 2.7. The result is now obtained by Theorem 2.10. □

Theorem 2.12. *Suppose D satisfies DP. Then the following are equivalent.*

(i) D is real closed.
(ii) Every nonnegative element of D has a square root in D, and every monic polynomial of odd degree in $D[X]$ has a root in D.
(iii) Every nonnegative element of F_D has a square root in F_D, and every monic polynomial of odd degree in $F_D[X]$ has a root in F_D.
(iv) F_D is real closed.

Proof. $(i) \Rightarrow (ii)$ follows from Proposition 2.6.
$(ii) \Rightarrow (iii)$ follows from Proposition 2.8.
$(iii) \Rightarrow (iv)$ If every nonnegative element of F_D has a square root in F_D, then D clearly satisfies the same property. Furthermore, suppose that every monic polynomial of odd degree in $F_D[X]$ has a root in F_D. Since D satisfies DP, it inherits the same property. Now, Corollary 2.9 gives F_D is real closed.
$(iv) \Rightarrow (i)$ follows from Corollary 2.11. $\square$

Note that for an ordered field F; $F = F_D$ and it clearly satisfies DP. So the following result is immediate.

Corollary 2.13. *An ordered field F is real closed if and only if every nonnegative element of F has a square root in F, and every monic polynomial of odd degree in $F[X]$ has a root in F.*

Lemma 2.14. *Let $(D_1, \leq_1)$, $(D_2, \leq_2)$ be ordered domains and $\varphi : D_1 \to D_2$ be a surjective homomorphism that preserves order. If D_1 satisfies DP, then D_2 satisfies DP.*

Proof. Let $b_1, b_2 \in D_2$ such that $0 < b_1 < b_2$. Then there exist $a_1, a_2 \in D_1$ such that $a_1 \in \varphi^{-1}(b_1)$ and $a_2 \in \varphi^{-1}(b_2)$. It is claimed that $a_1 < a_2$. Otherwise, $a_1 \geq a_2$. Since φ preserves order, we obtain $b_1 = \varphi(a_1) \geq b_2 = \varphi(a_2)$ which is a contradiction. So $0 < a_1 < a_2$. By hypothesis, $a_2 \mid a_1$. So $a_1 = a_2\, k$ for some $k \in D_1$. Therefore, $b_1 = b_2\, \varphi(k)$, i.e., $b_2 \mid b_1$. This means D_2 satisfies DP. $\square$

According to [5, Chapter 5], we have the following:

(a) An ideal I in a partially ordered ring A is called *convex* whenever $a \leq b \leq c$ and $a, c \in I$, then $b \in I$. Equivalently, if $0 \leq a \leq b$ and $b \in I$, then $a \in I$.
(b) Let $(A, \leq)$ be a partially ordered ring and I be an ideal of A. Then $\frac{A}{I}$ is a partially ordered ring via the following definition:

$$I(a) \geq 0 \text{ if and only if there exists } 0 \leq x \in A \text{ such that } a \equiv x \pmod{I}, i.e., a - x \in I.$$

Theorem 2.15. *([5, Theorem 5.2]) Let $(A, \leq)$ be a partially ordered ring and I be an ideal of A. Then $\frac{A}{I}$ is a partially ordered ring (via the above definition) if and only if I is convex.*

Proposition 2.16. *Let D be a domain and P be a prime convex ideal in D. If D satisfies DP, then $\frac{D}{P}$ satisfies DP.*

Proof. By Theorem 2.15, $\frac{D}{P}$ is a partially ordered domain. Let $P(a_1),\ P(a_2) \in \frac{D}{P}$. If $a_1 < a_2$, then $a_2 - a_1 > 0$ and therefore $P(a_2) - P(a_1) = P(a_2 - a_1) \geq 0$. So $P(a_1) \leq P(a_2)$ and hence $\frac{D}{P}$ is totally ordered. On the other hand, the natural mapping $\pi : D \to \frac{D}{P}$ with $\pi(a) = P(a)$ is a surjective homomorphism and preserves order. The result is now obtained by Lemma 2.14. $\square$

References

[1] G. Cherlin, *Model-theoretic algebra*, J. Symbolic Logic, 41 (1976), 537-545.
[2] G. Cherlin and M.A. Dickmann, *Real closed rings I*, Fund. Math, 126 (1986), 147-183.
[3] G. Cherlin and M.A. Dickmann, *Real closed rings II*, Model theory, Ann. Pure and Applied Logic, 25 (1983), 213-231.
[4] M.A. Dickmann, *On polynomials over real closed rings*, Model Theory of Algebra and Arithmetic, Lecture Notes in Math. 834 (1980), 117-135.
[5] L. Gillman and M. Jerison, *Rings of Continuous Functions*, Springer-Verlag, 1976.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–266, pp. 404–409

A Cohomological Invariant via Linkage

Maryam Jahangiri[a,*], Khadijeh Sayyari[b,]

[a] *Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran Iran.*
[b] *Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran Iran.*

Article Info

Keywords:
Linkage of ideals
local cohomology modules.

2020 MSC:
13C40
13D45.

Abstract

Let R be a commutative Noetherian ring and M be a finitely generated R-module. In this talk we concider the cohomological dimension of M with respect to the linked ideals over it. We show that for every ideal $\mathfrak{b}$ which is geometrically linked with $\mathfrak{a}$, cd $(\mathfrak{a}, H_{\mathfrak{b}}^{\text{grade}\ _R\mathfrak{b}}(R))$ is constant and doese not depend on $\mathfrak{b}$.

1. Introduction

Linkage theory is an important topic in commutative algebra and algebraic geometry. It refers to Halphen (1870) and M. Noether (1882) who worked to classify space curves. In 1974 the significant work of Peskine and Szpiro [7] brought breakthrough to this theory and stated it in the modern algebraic language; two proper ideals $\mathfrak{a}$ and $\mathfrak{b}$ in a Cohen-Macaulay local ring R is said to be linked if there is a regular sequence $\underline{x}$ in their intersection such that $\mathfrak{a} = (\underline{x}) :_R \mathfrak{b}$ and $\mathfrak{b} = (\underline{x}) :_R \mathfrak{a}$.
A new progress in the linkage theory is the work of Martsinkovsky and Strooker [5] which established the concept of linkage of modules. Also, in [3], the authors introduced the concept of linkage of ideals over a module and studied some of its basic properties.
Let R be a commutative Noetherian ring, $\mathfrak{a}$ and $\mathfrak{b}$ be two non-zero ideals of R and M denotes a non-zero finitely generated R-module. Assume that $\mathfrak{a}M \neq M \neq \mathfrak{b}M$ and let $I \subseteq \mathfrak{a} \cap \mathfrak{b}$ be an ideal generating by an M-regular sequence. Then the ideals $\mathfrak{a}$ and $\mathfrak{b}$ are said to be linked by I over M, denoted by $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$, if $\mathfrak{b}M = IM :_M \mathfrak{a}$ and $\mathfrak{a}M = IM :_M \mathfrak{b}$. Also, we say that $\mathfrak{a}$ is linked over M if there exist ideals $\mathfrak{b}$ and I of R such that $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$. This concept is the classical concept of linkage of ideals in [7], where $M := R$. Note that these two concepts do not coincide [3, 2.6] although, in some cases they do (e.g. Example [3, 2.4]). We can also characterize linked ideals over R, see [4, 2.7].
As an application of this generalization, one may characterize Cohen-Macaulay modules in terms of the type of linked ideals over it, see [4, 3.5].

---

*Talker
Email addresses: jahangiri@khu.ac.ir (Maryam Jahangiri), sayarikh@gmail.com (Khadijeh Sayyari)

One of the most important invariants in local cohomology theory is the cohomological dimension of an R-module X with respect to the ideal $\mathfrak{a}$, i.e.

$$\operatorname{cd}(\mathfrak{a}, X) := \operatorname{Sup}\{i \in \mathbb{N}_0 | H^i_{\mathfrak{a}}(X) \neq 0\}.$$

In this talk, we consider the above generalization of linkage of ideals over a module and study the cohomological dimension of an R-module M with respect to the ideals which are linked over M. In particular, in Theorem 2.6 we show that if $\mathfrak{a}$ is an ideal of R which is linked by I over M, then

$$\operatorname{cd}(\mathfrak{a}, M) \in \{\operatorname{grade}_M \mathfrak{a}, \operatorname{cd}(\mathfrak{a}, H^{\operatorname{grade}_M \mathfrak{a}}_{\mathfrak{c}}(M)) + \operatorname{grade}_M \mathfrak{a}\},$$

where $\mathfrak{c} := \bigcap_{\mathfrak{p} \in \operatorname{Ass} \frac{M}{IM} - V(\mathfrak{a})} \mathfrak{p}$.

And in Corollary 2.10 it is shown that for every ideal $\mathfrak{b}$ which is geometrically linked with $\mathfrak{a}$ over M, $\operatorname{cd}(\mathfrak{a}, H^{\operatorname{grade}_M \mathfrak{b}}_{\mathfrak{b}}(M))$ is constant.

Also, we show that if $\operatorname{cd}(\mathfrak{b}, R) < \dim(R)$ for any linked ideal $\mathfrak{b}$ over R, then $\operatorname{cd}(\mathfrak{a}, R) < \dim(R)$ for any ideal $\mathfrak{a}$ (Corollary 2.15).

Throughout the paper, R denotes a commutative Noetherian ring with $1 \neq 0$, $\mathfrak{a}$ and $\mathfrak{b}$ are two non-zero proper ideals of R and M denotes a non-zero finitely generated R-module.

2. Cohomological dimension

The cohomological dimension of an R-module X with respect to $\mathfrak{a}$ is defined by

$$\operatorname{cd}(\mathfrak{a}, X) := \operatorname{Sup}\{i \in \mathbb{N}_0 | H^i_{\mathfrak{a}}(X) \neq 0\}.$$

It is a significant invariant in local cohomology theory and attracts lots of interest. In this section, we study this invariant via "linkage". We begin by the definition of our main tool.

Definition 2.1. *Assume that $\mathfrak{a}M \neq M \neq \mathfrak{b}M$ and let $I \subseteq \mathfrak{a} \cap \mathfrak{b}$ be an ideal generated by an M-regular sequence. Then we say that the ideals $\mathfrak{a}$ and $\mathfrak{b}$ are linked by I over M, denoted $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$, if $\mathfrak{b}M = IM :_M \mathfrak{a}$ and $\mathfrak{a}M = IM :_M \mathfrak{b}$. The ideals $\mathfrak{a}$ and $\mathfrak{b}$ are said to be geometrically linked by I over M if $\mathfrak{a}M \cap \mathfrak{b}M = IM$. Also, we say that the ideal $\mathfrak{a}$ is linked over M if there exist ideals $\mathfrak{b}$ and I of R such that $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$. $\mathfrak{a}$ is M-selflinked by I if $\mathfrak{a} \sim_{(I;M)} \mathfrak{a}$. Note that in the case where $M = R$, this concept is the classical concept of linkage of ideals in [7].*

The following lemma, which will be used in the next proposition, finds some relations between local cohomology modules of M with respect to ideals which are linked over M.

Lemma 2.2. *Assume that I is an ideal of R such that $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$. Then*

(i) $\sqrt{I + \operatorname{Ann} M} = \sqrt{(\mathfrak{a} \cap \mathfrak{b}) + \operatorname{Ann} M}$. *In particular, $H^i_{\mathfrak{a}\cap\mathfrak{b}}(M) \cong H^i_I(M)$, for all i.*

(ii) *Let $I = 0$. Then, $\sqrt{0 :_{\frac{R}{\operatorname{Ann} M}} \mathfrak{a}} = \sqrt{\frac{\operatorname{Ann} \mathfrak{a}M}{\operatorname{Ann} M}} = \sqrt{\frac{\mathfrak{b} + \operatorname{Ann}_R M}{\operatorname{Ann}_R M}}$. Therefore, $H^i_{\operatorname{Ann}_R M :_R \mathfrak{a}}(M) \cong H^i_{\operatorname{Ann} \mathfrak{a}M}(M) \cong H^i_{\mathfrak{b}}(M)$. In other words, if M is faithful, then $H^i_{\mathfrak{b}}(M) \cong H^i_{0:_R \mathfrak{a}}(M)$.*

Proposition 2.3. *Let I be an ideal of R such that $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$ and set $t := \operatorname{grade}_M I$. Then $\operatorname{cd}(\mathfrak{a} + \mathfrak{b}, M) \leq \operatorname{Max}\{\operatorname{cd}(\mathfrak{a}, M), \operatorname{cd}(\mathfrak{b}, M), t + 1\}$. Moreover, if $\operatorname{cd}(\mathfrak{a} + \mathfrak{b}, M) \geq t + 1$, e.g. $\mathfrak{a}$ and $\mathfrak{b}$ are geometrically linked over M, then the equality holds.*

The following corollary, which is immediate by the above proposition, shows that, in spite of [2, 21.22], parts of an R-regular sequence can not be linked over R.

Corollary 2.4. *Let $(R, \mathfrak{m})$ be local and $x_1, ..., x_n \in \mathfrak{m}$ be an R-regular sequence, where $n \geq 4$. Then $(x_{i_1}, ..., x_{i_j}) \nsim (x_{i_{j+1}}, ..., x_{i_{2j}})$, for all $1 < j \leq [\frac{n}{2}]$ and any permutation $(i_1, ..., i_{2j})$ of $\{1, ..., 2j\}$.*

Let $M \neq \mathfrak{a}M$. It is known that, e.g. by [1, 1.3.9], grade $_M\mathfrak{a} \leq \operatorname{cd}(\mathfrak{a}, M)$. Then M is said to be relative Cohen-Macaulay with respect to $\mathfrak{a}$ if

$$\operatorname{cd}(\mathfrak{a}, M) = \operatorname{grade}\ _M\mathfrak{a}.$$

In the following proposition we compute the cohomological dimension of an R-module M with respect to $\mathfrak{a}$ in two cases.

Proposition 2.5. *Let I be an ideal of R generating by an M-regular sequence and $\mathfrak{a} \sim_{(I;M)} \mathfrak{b}$.*

(i) If M is relative Cohen-Macaulay with respect to $\mathfrak{a}+\mathfrak{b}$, then $H^i_{\mathfrak{a}}(M) = 0$ for all $i \notin \{\operatorname{grade}\ _M\mathfrak{a}, \operatorname{grade}\ _M\mathfrak{a}+\mathfrak{b}\}$.

(ii) If $I = 0$, then $\operatorname{cd}(\mathfrak{a}, M) = \operatorname{cd}(\mathfrak{a}, \frac{M}{\mathfrak{b}M})$.

The next theorem, which is our main result, provides a formula for $\operatorname{cd}(\mathfrak{a}, M)$ in the case where $\mathfrak{a}$ is linked over M.

Theorem 2.6. *Let I be an ideal of R generating by an M-regular sequence such that $\operatorname{Ass}\frac{M}{IM} = \operatorname{Min} Ass\frac{M}{IM}$ and $\mathfrak{a}$ is linked by I over M. Then*

$$\operatorname{cd}(\mathfrak{a}, M) \in \{\operatorname{grade}\ _M\mathfrak{a}, \operatorname{cd}(\mathfrak{a}, H^{\operatorname{grade}\ _M\mathfrak{a}}_{\mathfrak{c}}(M)) + \operatorname{grade}\ _M\mathfrak{a}\},$$

where $\mathfrak{c} := \bigcap_{\mathfrak{p}\in \operatorname{Ass}\frac{M}{IM} - V(\mathfrak{a})} \mathfrak{p}$.

Proof. Note that, by [3, 2.7], $\operatorname{Ass}\frac{M}{\mathfrak{a}M} \subseteq \operatorname{Ass}\frac{M}{IM}$. Set $t := \operatorname{grade}\ _M\mathfrak{a}$. Without loss of generality, we may assume that $\operatorname{cd}(\mathfrak{a}, M) \neq t$. Hence, there exists $\mathfrak{p} \in \operatorname{Ass}\frac{M}{IM} - V(\mathfrak{a})$, else, $\sqrt{I + \operatorname{Ann} M} = \sqrt{\mathfrak{a} + \operatorname{Ann} M}$ which implies that $\operatorname{cd}(\mathfrak{a}, M) = t$. We claim that

$$\operatorname{grade}\ _M(\mathfrak{a} + \mathfrak{c}) > t. \tag{1}$$

Suppose the contrary. So, there exist $\mathfrak{p} \in \operatorname{Ass}\frac{M}{IM}$ and $\mathfrak{q} \in \operatorname{Ass}\frac{R}{\mathfrak{c}}$ such that $\mathfrak{a} + \mathfrak{q} \subseteq \mathfrak{p}$. By the assumption, $\mathfrak{p} = \mathfrak{q}$ which is a contradiction to the structure of $\mathfrak{c}$.
Let $A := \{\mathfrak{p}|\mathfrak{p} \in \operatorname{Ass}\frac{M}{IM} \cap V(\mathfrak{a})\}$. Then, in view of [3, 2.7],

$$\sqrt{\mathfrak{a} + \operatorname{Ann} M} = \bigcap_{\mathfrak{p}\in \operatorname{Min} Ass\frac{M}{\mathfrak{a}M}} \mathfrak{p} \supseteq \bigcap_{\mathfrak{p}\in A} \mathfrak{p}.$$

On the other hand, let $\mathfrak{p} \in \operatorname{Min} A$. Then, there exists $\mathfrak{q} \in \operatorname{Min} Ass\frac{M}{\mathfrak{a}M}$ such that $\mathfrak{q} \subseteq \mathfrak{p}$. Hence, again by [3, 2.7], $\mathfrak{q} \in A$ and, by the structure of $\mathfrak{p}$, $\mathfrak{q} = \mathfrak{p}$. Therefore,

$$\sqrt{\mathfrak{a} + \operatorname{Ann} M} = \bigcap_{\mathfrak{p}\in A} \mathfrak{p}. \tag{2}$$

Whence, using (2), it follows that

$$\sqrt{I + \operatorname{Ann} M} = \bigcap_{\mathfrak{p}\in \operatorname{Ass}\frac{M}{IM}} \mathfrak{p} = \bigcap_{\mathfrak{p}\in \operatorname{Ass}\frac{M}{\mathfrak{a}M}} \mathfrak{p} \cap \mathfrak{c} = \sqrt{\mathfrak{a} \cap \mathfrak{c} + \operatorname{Ann} M}.$$

Now, in view of (1), we have the following Mayer-Vietoris sequence

$$0 \longrightarrow H^t_{\mathfrak{a}}(M) \oplus H^t_{\mathfrak{c}}(M) \longrightarrow H^t_I(M) \longrightarrow N \longrightarrow 0 \tag{3}$$

for some $\mathfrak{a}$-torsion R-module N. Applying $\Gamma_{\mathfrak{a}}(-)$ on (3), we get the exact sequence

$$0 \to H^t_{\mathfrak{a}}(M) \oplus \Gamma_{\mathfrak{a}}(H^t_{\mathfrak{c}}(M)) \to \Gamma_{\mathfrak{a}}(H^t_I(M)) \to N \xrightarrow{f} H^1_{\mathfrak{a}}(H^t_{\mathfrak{c}}(M)) \to H^1_{\mathfrak{a}}(H^t_I(M)) \to 0$$

and the isomorphism

$$H^i_{\mathfrak{a}}(H^t_{\mathfrak{c}}(M)) \cong H^i_{\mathfrak{a}}(H^t_I(M)), \text{for all } i > 1.$$

Also, using [6, 3.4], we have $H_{\mathfrak{a}}^{i+t}(M) \cong H_{\mathfrak{a}}^{i}(H_I^t(M))$, for all $i \in \mathbb{N}_0$. This implies that

$$H_{\mathfrak{a}}^{i}(M) \begin{cases} \cong H_{\mathfrak{a}}^{i-t}(H_{\mathfrak{c}}^{t}(M) & i > t+1, \\ \cong \frac{H_{\mathfrak{a}}^{1}(H_{\mathfrak{c}}^{t}(M)}{im(f)} & i = t+1, \\ \neq 0 & i = t, \\ 0 & \text{otherwise.} \end{cases}$$

Now, the result follows from the above isomorphisms. □

The following corollary, which follows from the above theorem, provides a precise formula for $\operatorname{cd}(\mathfrak{a}, M)$ in the case where $\mathfrak{a}$ is geometrically linked over M and shows how far $\operatorname{cd}(\mathfrak{a}, M)$ is from $\operatorname{grade}_M \mathfrak{a}$. Note that by [1, 1.3.9], $\operatorname{grade}_M \mathfrak{a} \leq \operatorname{cd}(\mathfrak{a}, M)$.

Corollary 2.7. *Let I be an ideal of R generating by an M-regular sequence and $\mathfrak{a}$ and $\mathfrak{b}$ be geometrically linked by I over M. Also, assume that M is not relative Cohen-Macaulay with respect to $\mathfrak{a}$. Then*

$$\operatorname{cd}(\mathfrak{a}, M) = \operatorname{cd}(\mathfrak{a}, H_{\mathfrak{b}}^{\operatorname{grade}_M \mathfrak{a}}(M)) + \operatorname{grade}_M \mathfrak{a}.$$

Proposition 2.8. *Let I be an ideal of R generating by an M-regular sequence and $\mathfrak{a}$ and $\mathfrak{b}$ be geometrically linked by I over M. Then* $\operatorname{grade}_M \mathfrak{a} + \mathfrak{b} = \operatorname{grade}_M I + 1$.

Remark 2.9. A linked ideal can be linked with more than one ideal. Let R be local and x, y, z be an M-regular sequence. Then Rx and Ry are geometrically linked over M. Also Rx and Rz are geometrically linked over M. The following corollary shows that for all ideals $\mathfrak{b}$ which are geometrically linked with $\mathfrak{a}$ over M, $\operatorname{cd}(\mathfrak{a}, H_{\mathfrak{b}}^{\operatorname{grade}_M \mathfrak{b}}(M))$ is constant.

Corollary 2.10. *Let $\mathfrak{a}$ be linked over M. Then, for every ideal $\mathfrak{b}$ which is geometrically linked with $\mathfrak{a}$ over M,* $\operatorname{cd}(\mathfrak{a}, H_{\mathfrak{b}}^{\operatorname{grade}_M \mathfrak{b}}(M))$ *is constant. In particular,*

$$\operatorname{cd}(\mathfrak{a}, H_{\mathfrak{b}}^{\operatorname{grade}_M \mathfrak{b}}(M)) = \begin{cases} 1, & M \text{ is relative Cohen-Macaulay with respect to } \mathfrak{a}, \\ \operatorname{cd}(\mathfrak{a}, M) - \operatorname{grade}_M \mathfrak{a}, & \text{otherwise.} \end{cases}$$

Proof. Assume that M is relative Cohen-Macaulay with respect to $\mathfrak{a}$ and $\mathfrak{a}$ and $\mathfrak{b}$ are geometrically linked by some M-regular sequence I of length t over M. Then, by [3, 2.8], we have the following Mayer-Vietoris sequence

$$0 \longrightarrow H_{\mathfrak{a}}^{t}(M) \oplus H_{\mathfrak{b}}^{t}(M) \longrightarrow H_I^t(M) \longrightarrow N \longrightarrow 0 \tag{4}$$

for some $\mathfrak{a}$-torsion R-module N. Applying $\Gamma_{\mathfrak{a}}(-)$ on (4), we get the exact sequence

$$H_{\mathfrak{a}}^{i-1}(N) \to H_{\mathfrak{a}}^{i}(H_{\mathfrak{b}}^{t}(M)) \to H_{\mathfrak{a}}^{i}(H_I^t(M)),$$

for $i > 1$. Now, by [6, 3.4] and the assumption, we get $H_{\mathfrak{a}}^{i}(H_{\mathfrak{b}}^{t}(M)) = 0$, for $i > 1$. On the other hand, again by [3, 2.8] , $\Gamma_{\mathfrak{a}}(H_{\mathfrak{b}}^{t}(M)) = 0$.
Therefore, using the convergence of spectral sequence

$$H_{\mathfrak{a}}^{i}(H_{\mathfrak{b}}^{j}(M)) \Rightarrow_i H_{\mathfrak{a}+\mathfrak{b}}^{i+j}(M)$$

and the assumption, we get $H_{\mathfrak{a}}^{1}(H_{\mathfrak{b}}^{t}(M)) \cong H_{\mathfrak{a}+\mathfrak{b}}^{t+1}(M)$. Now, by 2.8, $H_{\mathfrak{a}}^{1}(H_{\mathfrak{b}}^{t}(M)) \neq 0$ and $\operatorname{cd}(\mathfrak{a}, H_{\mathfrak{b}}^{\operatorname{grade}_M \mathfrak{b}}(M)) = 1$.
In the case where M is not relative Cohen-Macaulay with respect to $\mathfrak{a}$, the result follows from 2.7. □

Convetion 2.11. Assume that I is an ideal of R which is generated by an M-regular sequence. We define the set

$$S_{(I;M)} := \{\mathfrak{a} \lhd R | I \subsetneqq \mathfrak{a}, \mathfrak{a} = IM :_R IM :_M \mathfrak{a}\}.$$

$S_{(I;R)}$ actually contains of all linked ideals by I.

The following proposition, which is needed in the next two items, shows that any ideal $\mathfrak{a}$ with $\mathfrak{a}M \neq M$ can be embedded in a redical ideal $\mathfrak{a}'$ of $S_{(I;M)}$ for some I.

Proposition 2.12. *Assume that* $\mathfrak{a}M \neq M$. *Then,*

(i) There exists an ideal I*, generating by an* M*-regular sequence, such that* $\mathfrak{a}$ *can be embedded in a radical element* $\mathfrak{a}'$ *of* $S_{(I;M)}$ *with* $\operatorname{grade}_M \mathfrak{a}' = \operatorname{grade}_M \mathfrak{a} =: t$. *Also,* $\mathfrak{a}'$ *can be chosen to be the smallest radical ideal with this property.*

(ii) Let $\mathfrak{a}'$ *be as in (i). Then* $\operatorname{Ass} H^t_{\mathfrak{a}}(M) = \operatorname{Ass} \frac{R}{\mathfrak{a}'}$. *In particular,* $\mathfrak{a}' = \bigcap_{\mathfrak{p} \in \operatorname{Ass} H^t_{\mathfrak{a}}(M)} \mathfrak{p}$ *and it is independent of the choice of the ideal* I.

(iii) Let $\mathfrak{a}$ *be a linked ideal over* M. *Then,* $\sqrt{\mathfrak{a} + \operatorname{Ann} M} = \bigcap_{\mathfrak{p} \in \operatorname{Ass} H^t_{\mathfrak{a}}(M)} \mathfrak{p}$.

The following theorem provides some conditions in order to have $cd(\mathfrak{a}, M) < \dim M$.

Theorem 2.13. *Let* $(R, \mathfrak{m})$ *be local and* $\underline{\mathfrak{x}} = x_1, ..., x_t$ *be an* M*-regular sequence of length* t. *Assume that* $H^{\dim M}_{\mathfrak{p}}(M) = 0$ *for all* $\mathfrak{p} \in \operatorname{Ass}_R \frac{M}{(\underline{\mathfrak{x}})M}$. *Then,* $H^{\dim M}_{\mathfrak{a}}(M) = 0$ *for any ideal* $\mathfrak{a} \supseteq (\underline{\mathfrak{x}})$ *with* $\operatorname{grade}_M \mathfrak{a} = t$.

Remark 2.14. Let the situations be as in the above theorem and assume, in addition, that $(R, \mathfrak{m})$ is complete. Let $\mathfrak{a} \supseteq (\underline{\mathfrak{x}})$ be an ideal with $\operatorname{grade}_M \mathfrak{a} = t$. Then, $\mathfrak{a}$ can not be coprimary with a member of Assh M,i.e. there is no $\mathfrak{p} \in$ Assh M with $\sqrt{\mathfrak{a} + \mathfrak{p}} = \mathfrak{m}$.

Corollary 2.15. *Let* $H^{\dim R}_{\mathfrak{b}}(R) = 0$ *for any linked ideal* $\mathfrak{b}$ *over* R. *Then* $H^{\dim R}_{\mathfrak{a}}(R) = 0$ *for any ideal* $\mathfrak{a}$.

References

[1] Brodmann, M. P. and Sharp, R. Y. (2013), *Local cohomology: an algebraic introduction with geometric applications*. Cambridge Studies in Advanced Mathematics. 136. Cambridge University Press, Cambridge.

[2] Eisenbud, D. (1995), *Commutative algebra with a view toward algebraic geometry*, Graduate Texts in Math., vol. 150, Springer-Verlag, Berlin and New York.

[3] Jahangiri, M. and Sayyari, Kh. (2021), *Linkage of ideals over a module*, Journal of Algebraic Systems Vol. 8, No. 2, pp 267-279.

[4] Jahangiri, M. and Sayyari, Kh. (2020), *Charaterization of some special rings via linkage*, Journal of Algebra and Related Topics Vol. 8, No 1, pp 67-81.

[5] Martsinkovsky, A. and Strooker, J. R. (2004), *Linkage of modules*, Jornal Algebra **271** 587-626.

[6] Nagel, U. and Schenzel, P. (1994), *Annihilator and Castelnovo-Mumford regularity*, in: Commutative Algebra: Syzygies, Multiplicities and Birational Algebra, in: Contemp. Math., vol. 159, Amer. Math. Soc, Providence, RI, pp. 307-328.

[7] Peskine, C. and Szpiro, L. (1974), *Liasion des variétés algébriques*, I, Inv. math. **26** 271-302.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–268, pp. 410–415

Rate, change of rings and tensor product

Maryam Jahangiri[a,*]

[a] *Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran Iran.*

Article Info

Keywords:
Regularity
Rate
Minimal free resolutions.

2020 MSC:
13A02
13D07
13D02.

Abstract

The rate of a graded module is a measure of the growth of shifts in its minimal graded free resolution. In this paper, we consider this invariant and study its behaviour under some special change of rings and also tensor product.

1. Introduction

Let R be a standard graded algebra over a field K, i.e. R is generated as an K-algebra by finitely many elements of degree 1. Also, let $\mathfrak{m} := \bigoplus_{i\in\mathbb{N}} R_i$ be the homogeneous maximal ideal of R and M be a finitely generated graded R-module . One of the important invariants attached to M is the Castelnuovo-Mumford regularity of M, defined by

$$\mathrm{reg}_R(M) := \sup\{t_i^R(M) - i : i \geq 0\},$$

where for all $i \geq 0$, $t_i^R(M)$ denotes the maximum degree of minimal generators of the ith syzygy module of M, i.e. $t_i(M) := \sup\{j \in \mathbb{Z} : Tor_R^i(M,K)_j \neq 0\}$. This invariant plays an important role in the study of homological properties of M. Regularity of a module can be infinite. Avramov and Peeva in [6] showed that $\mathrm{reg}_R(K)$ is zero or infinite. The K-algebra R is called Koszul if $\mathrm{reg}_R(K) = 0$. From certain point of views, Koszul algebras behave homologically as polynomial rings. Avramov and Eisenbud in [5] showed that if R is Koszul, then the regularity of every finitely generated graded R-module is finite.
Another important invariant of M is *rate*. The notion of rate for algebras introduced by Backelin ([7]) and it is generalized in [3] for graded modules. The rate of a finitely generated graded module M over R is defined by

$$\mathrm{rate}_R(M) := \sup\{t_i^R(M)/i : i \geq 1\},$$

This invariant is always finite (see [3]). The Backelin rate of the algebra R is denoted by $\mathrm{Rate}(R)$ and is equal to $\mathrm{rate}_R(\mathfrak{m}(1))$, the rate of the unique homogenous maximal ideal of R which is shifted by 1. By the definition, one can

*Talker
Email address: jahangiri@khu.ac.ir (Maryam Jahangiri)

see that $\mathrm{Rate}(R) \geq 1$ and the equality holds if and only if R is Koszul, so that the $\mathrm{Rate}(R)$ can be taken as a measure of how much R deviates from being Koszul.
The goal of this talk is to study the rate of modules under some change rings and also tensor product. Let $\varphi : R \to S$ be a surjective homomorphism of standard graded K-algebras and M be a finitely generated graded S-module. Aramova et al. in [3, Proposition 1.2] studied the behavior of rate via change of rings. They showed that

$$\mathrm{rate}_S(M) \leq \max\{\mathrm{rate}_R(M), \mathrm{rate}_R(S)\}.$$

But, one can see that their proof works for the case where M is non-negatively graded over S. In this talk we consider this problem. More precisely, with the above assumptions, we show that

$$\mathrm{rate}_R(M) \leq \max\{\mathrm{rate}_S(M), \mathrm{rate}_R(S)\} + \max\{0, t_0^S(M)\}.$$

By a result of Backelin and Fröberg ([8]) the tensor product of two Koszul algebras are Koszul. Here we extend this result for the rate of modules. More precisely, we show that if R and S are two standard graded K-algebras and M and N are finitely generated graded modules over R and S respectively, then $\mathrm{rate}_T(M \otimes_K N) \leq \max\{\mathrm{rate}_R(M), \mathrm{rate}_S(N)\}$, where $T = R \otimes_K S$ (Proposition 4.2).
Throughout, K is a field and $R = \bigoplus_{i \in \mathbb{N}_0} R_i$ is a standard graded K-algebra. Also, $M = \bigoplus_{i \in \mathbb{Z}} M_i$ denotes a finitely generated graded R-module.

2. Preliminaries

First of all, we prepare some notations and preliminaries which will be used in the paper.

Remark 2.1.

1. For each $d \in \mathbb{Z}$ we denote by $M(d)$ the graded R-module with $M(d)_p = M_{d+p}$, for all $p \in \mathbb{Z}$.
   Denote by $\mathfrak{m}$ the maximal homogeneous ideal of R, that is $\mathfrak{m} = \oplus_{i \in \mathbb{N}} R_i$. Then we may consider K as a graded R-module via the identification $K = R/\mathfrak{m}$.
2. A minimal graded free resolution of M as an R-module is a complex of free R-modules
   $$\mathbf{F} = \cdots \to F_i \xrightarrow{\partial_i} F_{i-1} \to \cdots \to F_1 \xrightarrow{\partial_1} F_0 \to 0$$
   such that $\mathrm{H}_i(\mathbf{F})$, the i-th homology module of $\mathbf{F}$, is zero for $i > 0$, $\mathrm{H}_0(\mathbf{F}) = M$ and $\partial_i(F_i) \subseteq \mathfrak{m}F_{i-1}$, for all $i \in \mathbb{N}_0$. Each F_i is isomorphic to a direct sum of copies of $R(-j)$, for $j \in \mathbb{Z}$. Such a resolution exists and any two minimal graded free resolutions are isomorphic as complexes of graded R-modules. So, for all $j \in \mathbb{Z}$ and $i \in \mathbb{N}_0$ the number of direct summands of F_i isomorphic to $R(-j)$ is an invariant of M, called the ij-th graded Betti number of M and denoted by $\beta_{ij}^R(M)$.
   Also, by definition, the i-th Betti number of M as an R-module, denoted by $\beta_i^R(M)$, is the rank of F_i. By construction, one has $\beta_i^R(M) = \dim_k(\mathrm{Tor}_i^R(M, K))$ and $\beta_{ij}^R(M) = \dim_k(\mathrm{Tor}_i^R(M, K)_j)$.
3. For every integer i we set
   $$t_i^R(M) := \max\{j : \beta_{ij}^R(M) \neq 0\}.$$
   If $\beta_i^R(M) = 0$ we set $t_i^R(M) = 0$.
4. The Castelnuovo-Mumford regularity of M is defined by
   $$\mathrm{reg}_R(M) := \sup\{t_i^R(M) - i : i \in \mathbb{N}_0\}.$$
   This invariant is, after Krull dimension and multiplicity, perhaps the most important invariant of a finitely generated graded R-module.

Definition and Remark 2.2. Following [13], M is called Koszul if the associated graded module $M^g := \oplus_{i \geq 0} \mathfrak{m}^i M/\mathfrak{m}^{i+1} M$ has 0-linear resolution. The ring R is Koszul if the residue field K, as an R-module, is Koszul.
The Castelnuovo-Mumford regularity plays an important role in the study of homological properties of M and it is clear that $\mathrm{reg}_R(M)$ can be infinite. Avramov and Peeva in [6] proved that $\mathrm{reg}_R(K)$ is zero or infinite. Also, Avramov and Eisenbud in [5] showed that if R is Koszul, then the regularity of every finitely generated graded R-module is finite.

3. The rate of modules

Definition and Remark 3.1.

1. The Backelin rate of R is defined as

$$\text{Rate}(R) := \sup\{(t_i^R(K) - 1)/i - 1 : i \geq 2\},$$

and generalization of this for modules is defined by

$$\text{rate}_R(M) := \sup\{t_i^R(M)/i : i \geq 1\}.$$

2. A comparison with Bakelin's rate shows that $\text{Rate}(R) = \text{rate}_R(\mathfrak{m}(1))$. Also, it turns out that the rate of any module is finite (see [3, Corollary 1.3]).
3. One can see that $\text{rate}_R(M) \leq \text{reg}_R(M) + 1$.

Remark 3.2. Consider a minimal presentation of R as a quotient of a polynomial ring, i.e.

$$R \cong S/I,$$

where $S = k[X_1, \cdots, X_n]$ is a polynomial ring and I is an ideal generated by homogeneous elements of degree ≥ 2. I is called a defining ideal of R. Let $m(I)$ denotes the maximum of the degree of a minimal homogeneous generator of I. Then, by (the graded version of) [9, 2.3.2], $t_2^R(K) = m(I)$. Therefore, one has

$$\text{Rate}(S/I) \geq m(I) - 1. \tag{1}$$

From the above inequality one can see that $Rate(R) \geq 1$ and the equality holds if and only if R is Koszul. So that $Rate(R)$ can be taken as a measure of how much R deviates from being Koszul. Also, for a module M with $\text{indeg}(M) = t_0^R(M) = 0$ we have $\text{rate}_R(M) \geq 1$ and the equality holds if and only if M is Koszul.

Lemma 3.3. *Let*

$$\cdots \to L_n \to L_{n-1} \to \cdots \to L_1 \to L_0 \to L \to 0$$

be an exact sequence of graded R-modules and homogeneous homomorphisms. Then

$$t_n^R(L) \leq \max\{t_{n-i}^R(L_i) : 0 \leq i \leq n\}.$$

Proof. We prove the claim by induction on n.
In the case $n = 0$, the result follows using the surjection

$$\text{Tor}_0^R(L_0, K)_j \to \text{Tor}_0^R(L, K)_j$$

and Remark 2.1(3). Now, let $n > 0$ and suppose that the result has been proved for smaller values of n. Let K_1 be the kernel of the homomorphism $L_0 \to L$. Then, using the exact sequence

$$\cdots \to L_i \to \cdots \to L_1 \to K_1 \to 0$$

and the inductive hypothesis, we have

$$t_{n-1}^R(K_1) \leq \max\{t_{n-1-i}^R(L_{i+1}) | 0 \leq i \leq n-1\}. \tag{2}$$

Also, the exact sequence

$$0 \to K_1 \to L_0 \to L \to 0$$

and Remark 2.1(4) implies that

$$t_n^R(L) \leq \max\{t_n^R(L_0), t_{n-1}^R(K_1)\}.$$

Now, the result follows from 2 and the above inequality. $\square$

In the following theorem we study the rate of modules via change of rings.

Theorem 3.4. *Let $\varphi : R \to S$ be a surjective homogeneous homomorphism of standard graded K-algebras. Assume that M is a finitely generated graded S-module. Then*

$$\mathrm{rate}_R(M) \leq \max\{\mathrm{rate}_S(M), \mathrm{rate}_R(S)\} + \max\{0, t_0^S(M)\}.$$

Proof. Let

$$\mathbf{F} : \cdots \to F_n \to F_{n-1} \to \cdots \to F_0 \to 0$$

be the minimal graded free resolution of M as an S-module. Applying Lemma 3.3 to $\mathbf{F}$, one has

$$t_n^R(M) \leq \max\{t_0^R(F_n), t_1^R(F_{n-1}), \cdots, t_n^R(F_0)\}.$$

Note that

$$t_i^R(F_j) = t_i^R(\oplus_{r\in\mathbb{Z}} S(-r)^{\beta_{jr}^S(M)}) = t_i^R(S) + t_j^S(M).$$

In particular, since $t_0^R(S) = 0$, we have $t_0^R(F_j) = t_j^S(M)$. Therefore, for any integer $n \geq 1$, we get

$$\frac{t_n^R(M)}{n} \leq \max\{\frac{t_n^S(M)}{n}, \frac{t_1^R(S) + t_{n-1}^S(M)}{n}, \cdots, \frac{t_n^R(S) + t_0^S(M)}{n}\}. \tag{3}$$

Let $a = \max\{\mathrm{rate}_R(S), \mathrm{rate}_S(M)\}$ then for all $j > 0$, we have $t_j^S(M) \leq ja$ and $t_j^R(S) \leq ja$. Now, by (3), one has

$$\frac{t_n^R(M)}{n} \leq \max\{a, a + t_0^S(M)\}$$

and this implies our desired inequality. □

Remark 3.5. Let the situation be as in the above theorem.

1. Aramova et al. in [3, Proposition 1.2] studied the behaviour of rate via change of rings. They showed that

$$\mathrm{rate}_S(M) \leq \max\{\mathrm{rate}_R(M), \mathrm{rate}_R(S)\}.$$

   But, one can see that their proof works for the case where M is non-negatively graded over S.
2. We remark that in particular case in the above theorem, if $t_0^S(M) \leq 0$ and $\mathrm{rate}_R(S) = 1$, then

$$\mathrm{rate}_R(M) \leq \mathrm{rate}_S(M).$$

3. In view of (1) and (2), if M is generated in degree zero and $\mathrm{rate}_R(S) = 1$, then

$$\mathrm{rate}_R(M) = \mathrm{rate}_S(M).$$

4. Bounding Rate and regularity of tensor product

In this section we find upper bounds for the rate and regularity of tensor product of modules over standard graded K-algebras in terms of rate and regularity of those modules.
The following lemma will be used in the next theorem.

Lemma 4.1. *Let R and S be two standard graded K-algebras and M and N be two finitely generated graded R and S-modules, respectively. Set $T := R \otimes_K S$. Then*

$$t_n^T(M \otimes_K N) = \max\{t_i^R(M) + t_j^S(N) :\ i, j \geq 0 \text{ and } i + j = n\}.$$

Proof. Let

$$\mathbf{F} : \cdots \to F_i \to F_{i-1} \to \cdots \to F_0 \to M \to 0$$

and

$$\mathbf{G} : \cdots \to G_j \to G_{j-1} \to \cdots \to G_0 \to N \to 0$$

be the minimal graded free resolutions of M and N as an R and S-module, respectively. We show that $\mathbf{F} \otimes_K \mathbf{G}$ is the minimal graded free resolution of $M \otimes_K N$ as T-module:
Note that $T_+ = R_+ \otimes_K S \oplus R \otimes_K S_+$ is the homogeneous maximal ideal of T. Let $a \in F_i$ and $b \in G_j$ with $i+j = n$. Then

$$\begin{aligned}
\partial^{\mathbf{F}\otimes_K \mathbf{G}}(a \otimes b) &= \partial^{\mathbf{F}}(a) \otimes b \pm a \otimes \partial^{\mathbf{G}}(b) \\
&\in R_+ F_{i-1} \otimes_K G_j \oplus F_i \otimes_K S_+ G_{j-1} \\
&\subseteq T_+(F_{i-1} \otimes_K G_j \oplus F_i \otimes_K G_{j-1}).
\end{aligned}$$

Now consider the Künneth map

$$\begin{aligned}
K^{\mathbf{FG}} \quad : \quad & H(\mathbf{F}) \otimes_K H(\mathbf{G}) \longrightarrow H(\mathbf{F} \otimes_K \mathbf{G}) \\
& K^{\mathbf{FG}}(cls(a) \otimes cls(b)) = cls(a \otimes b)
\end{aligned}$$

Since $H(\mathbf{F})$ is a free K-module, $K^{\mathbf{FG}}$ is an isomorphism (see [4, Proposition 1.3.4]). Therefore, $\mathbf{F} \otimes_K \mathbf{G}$ is acyclic and augmented to $M \otimes_K N$.
Now, let $r, s \in \mathbb{N}_0$ with $r + s = n$, $F_r = \bigoplus_{i\in\mathbb{Z}} R(-i)^{\beta^R_{ri}(M)}$ and $G_s = \bigoplus_{j\in\mathbb{Z}} S(-j)^{\beta^S_{sj}(N)}$. Then we have

$$\begin{aligned}
t^T_n(M \otimes_K N) &= \max\{i + j : \beta^R_{ri}(M)\beta^S_{sj}(N) \neq 0\} \\
&= \max\{t^R_r(M) + t^S_s(N) : r + s = n\},
\end{aligned}$$

as desired.

□

Backelin and Fröberg in [8] showed that the tensor product $R \otimes_K S$ of two standard graded K-algebras is Koszul if and only if R and S are both Koszul. In the following theorem we consider this problem for modules. More precisely, we find an upper bound for the rate and regularity of tensor product of two modules in terms of the rate and regularity of each of them. As a result, it gives a sufficient condition for the Koszulness of tensor product of two modules.

Theorem 4.2. *Let the situations be as in Lemma 4.1. Then*

1. $\operatorname{rate}_T(M \otimes_K N) \leq \max\{\operatorname{rate}_R(M), \operatorname{rate}_S(N)\}$. *In particular, if R and S are Koszul algebras, then so is $R \otimes_K S$.*
2. $\operatorname{reg}_T(M \otimes_K N) \leq \operatorname{reg}_R(M) + \operatorname{reg}_S(N)$.

Proof. 1) Let $\operatorname{rate}_R(M) = a$ and $\operatorname{rate}_S(N) = b$. Then, by definition, for all $i, j \geq 1$, we have

$$t^R_i(M) \leq ai \quad \textit{and} \quad t^S_j(N) \leq bj.$$

So, if $i + j = n$ then

$$t^R_i(M) + t^S_j(N) \leq ai + bj \leq \max\{a, b\}n.$$

Now, by the above lemma,

$$t^T_n(M \otimes_K N) \leq \max\{a, b\}n.$$

Therefore,

$$\operatorname{rate}_T(M \otimes_K N) \leq \max\{\operatorname{rate}_R(M), \operatorname{rate}_S(N)\},$$

as desired.
2) The proof of (2) is similar and follows from Lemma 4.1 . □

References

[1] Abhyankar, S. (1967), *Local rings of high embedding dimension*, Amer. J. Math., 89 1073–-1077.
[2] Ahangari Maleki, R. (2014), *On the regularity and Koszulness of modules over local rings*, Comm. Alg., 42, 3438-3452
[3] Aramova, A. Bărcănescu, S. and Herzog, J. (1995), *On the rate of relative Veronese submodules*, Rev. Roumaine Math. Pures Appl., 40(3-4) 243–-251.
[4] Avramov, L. L. (1996), *Infinite free resolutions*, in: Six Lectures on Commutative Algebra, Bellaterra, in: Progr. Math., vol.166, Birkhäuser, Basel, 1998, 1–-118.
[5] Avramov, L. L. and Eisenbud, D. (1992),*Regularity of modules over a Koszul algebra*, J.Algebra, 153, 85-90.
[6] Avramov, L. L. and Peeva, I. (2001),*Finite regularity and Koszul algebras*, Amer. J. Math., 123, 275-–281.
[7] Backelin, J. (1986),*On the rates of growth of the homologies of Veronese subrings*, Algebra, Algebraic Topology, and their interactions, ed. J.-E. Roos, Springer Lect. Notes in Math. 1183, 79-100.
[8] Backelin, J. and Fröberg, R. (1985), *Veronese subrings, Koszul algebras and rings with linear resolutions*, Rev. Roum. Pures Appl., 30, 85– 97.
[9] Bruns, W. and Herzog, J. (1993), *Cohen-Macaulay rings*, Cambridge Studies in Advanced Mathematics, 39 Cambridge University Press, Cambridge.
[10] Conca, A., De Negri, E. and Rossi, M. E. (2001), *On the rate of points in projective spaces*, Israel J. Math. 124, 253–-265.
[11] Conca, A., Rossi M. E. and Valla, G. (2001), *Gröbner flags and Gorenstein algebras*, Compositio Math., 129(1) 95-–121.
[12] Conca, A. , Trung, N. V. and valla, G. (2001), *Koszul property for points in projective spaces*, Math. Scand. 89(2) 201–-216.
[13] Herzog, J. and Iyengar, S. (2005), *Koszul modules*, J. Pure Appl. Algebra 201(1-3) 154-–188.
[14] Huneke, C. and Swanson, I. (2006), *Integral closure of ideals, rings, and modules*, London Mathematical Lecture Note Series, 336 Cambridge University Press.
[15] Macaulay, F. S. (1927), *Some Properties of Enumeration in the Theory of Modular Systems* , Proc. London. Math. Soc., 26, 531–555.
[16] Rossi, M. E. and Valla, G. (1988), *Multiplicity and t-isomultiple ideals*, Nagoya Math. J., 110, 81–-111.
[17] Sally, J. (1979), *Stretched Gorenstein rings*, J. London Math. Soc. 20, 19–-26.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–276, pp. 416–418

G(D) and division algebras

R. Fallah-Moghaddam[a,*]

[a]*Department of Mathematics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran*

Article Info

Keywords:
Division ring
Maximal subgroup
G(D)

2020 MSC:
16K20
20H25

Abstract

Assume that D be an division algebra with centre F of index n. Let D' the commutator subgroup of the multiplicative group D^*. For a given subgroup G of D^*, G is maximal in D^* if for any subgroup H of D^* such that $G \subseteq H$, we have $H = D^*$. We define $G(D) := D^*/RN(D^*)D'$, where $RN(D^*)$ is the image of D^* under the reduced norm of D to F. In this paper, we investigate the structure of $G(D)$. Let $A = M_n(D)$ be an F-central simple algebra of finite dimension its center. Assume that $Br(F) = \oplus\mathbb{Z}_2$ and $F^{*n} \neq F^*$. Then either $G(A) := A*/F^*A' \neq 1$ or $F^{*2} = RN_{D/F}(D^*)$.

1. Introduction

Assume that D be an division algebra with centre F of index n. Let D' the commutator subgroup of the multiplicative group D^*. For a given subgroup G of D^*, G is maximal in D^* if for any subgroup H of D^* such that $G \subseteq H$, we have $H = D^*$. We define $G(D) := D^*/RN(D^*)D'$, where $RN(D^*)$ is the image of D^* under the reduced norm of D to F, is an abelian periodic group of a bounded exponent dividing the index of D over F. We know that this group is not trivial in general. For example, when D is the algebra of real quaternions, we have $G(D)$ is trivial whereas for rational quaternions $G(D)$ is isomorphic to a direct product of copies of $\mathbb{Z}_2$. Consider that G(D) is not trivial, so by Prufer- Baer Theorem, we obtain that G(D) is isomorphic to a direct product of $\mathbb{Z}_{r_i}$, when r_i divides the index of D over F. In addition, we conclude that the existence of normal maximal subgroups of finite index in D^*. Thus, when $G(D)$ is not trivial, then D^* contains maximal subgroups.

The problem of whether the multiplicative group of D contains non-cyclic free subgroups seems to be posed first by Lichtman in [8]. In [5] and [6] stronger versions of this problem have been investigated which deal with the existence of non-cyclic free subgroups in normal or subnormal subgroups of $GL_1(D)$. Also, the question on the existence of non-cyclic free subgroups in linear groups over a field was studied by Tits in [14] which asserts that in the characteristic 0, every subgroup of the general linear group over a field F either contains a non-cyclic free subgroup or is soluble-by-finite, and every finitely generated subgroup either contains a non-cyclic free subgroup or is soluble-by-finite in the case of prime characteristic. This result of Tits is now referred as the Tits Alternative. Lichtman in [8] showed that there exists a finitely generated group which is not soluble-by-finite and does not contain a non-cyclic free subgroup.

For more information on these concepts, please refer to [2], [3], [7], [11],[12] and [13].

*Talker
Email address: r.fallahmoghaddam@cfu.ac.ir (R. Fallah-Moghaddam)

2. Main Result

In references [1], [3]and [7] various studies have been performed on the maximal subgroups of multiplicative subgroups of division algebras as well as the G (D) structure. Proven for example:

Theorem A. Let D be an F-central division algebra of index p^e such that F contains a primitive p-th root of unity and $G(D) = 1$. Then D is a quaternion algebra.

Theorem B. Given an F-central division algebra D of index n, the following conditions are equivalent:
(1) $G(D) = 1$;
(2) $SK_1(D) = 1$ and $F^{*2} = F^{*2n}$;
(3) $G_0(D) = 1$ and $F^{*2} = F^{*2n}$;
(4) $D^*$is F_π-perfect where π is the set of all primes dividing $ind(D)$.

Also, examples show that $G(D)$ is not stable under the extension over formal Laurent series. We then have an analogue of the Lipnickii Theorem by showing that there exists a field F and F-central division algebra D of odd index such that $G(D)$ (or $G(D)$) can be any finite cyclic group.

Assume that a field F is a field, set F^* the multiplicative group of F. For any prime number p, set $\mathbb{F}_p$ be its prime subfield. An absolutely algebraic field, denoted by aaf, is an algebraic extension of $\mathbb{F}_p$. It is easily checked that for any aaf F we obtain that $F = \bigcup_{n\in S} \mathbb{F}_{p^n}$, where S is a nonempty subset of the positive integers such that for any $n, m \in S$ we conclude that $\mathbb{F}_{p^{lcm(n,m)}} \subseteq F$. Also, if $n \in S$ and $x|n$, then $\mathbb{F}_{p^x} \subseteq F$. These conditions are necessary and sufficient conditions for when F is an absolutely algebraic field (aaf). It is also clear that any aaf is perfect. For this field it is proved that $Br(F) = 0$. Let $A = M_n(D)$ be an F-central simple algebra of finite dimension its center. Let $G(A) := A^*/F^*A' \neq 1$. So for any aaf F, $G(A) := F^*/F^{*n}$.

In this manner, we prove the following result:

Main Result. Let $A = M_n(D)$ be an F-central simple algebra of finite dimension its center. Assume that $Br(F) = \oplus\mathbb{Z}_2$ and $F^{*n} \neq F^*$. Then either $G(A) := A^*/F^*A' \neq 1$ or $F^{*2} = RN_{D/F}(D^*)$.

Proof. On the contrary, consider that $G(A) = 1$. We know that $A = M_n(D)$, when D is a division ring of finite dimensional over center F. By Theorem 5.7 of [10], we obtain that $G(A) = A^*/F^*A' = D^*/F^{*n}D'$. If $F = D$, then $F^{*n} = F^*$, which is a contradiction. Then, $F \neq D$ with $F^{*n}D' = D^*$, and so $G(D) = 1$. Since $Br(F) = \oplus\mathbb{Z}_2$ we conclude that D is a quaternion division algebra. Therefore, there exist $\alpha, \beta \in F^*$such that $D = (\frac{\alpha,\beta}{F})$ and $D = a + bi + cj + dk|a, b, c, d \in F^*$ with $i^2 = \alpha, j^2 = \beta, ij = -ji = k$, and $RN_{D/F}(a + bi + cj + dk) = a^2 - \alpha b^2 - \beta c^2 + \alpha\beta d^2$. Thus $F^*D' = D^*$, for any $x \in D^*$ we have $x = ac$ with $a \in F^*$ and $c \in D'$. This implies that $RN_{D/F}(x) = a^2 \in F^{*2}$, and hence $F^{*2} = RN_{D/F}(D^*)$, which is a contradiction.

3. ACKNOWLEDGMENT

The author thanks the Research Council of Farhangian University for support.

References

[1] H. Dorbidi, R. Fallah-Moghaddam, M. Mahdavi-Hezavehi, Soluble maximal subgroups of $GL_n(D)$, J. Algebra Appl., **9**(6):921-932, 2010.
[2] P. K. Draxl, Skew fields, Cambridge Univ. Press, Cambridge, 1983.
[3] R. Hazrat, SK_1-like Functors for Division Algebras, J. of. Algebra 239(2001), 573-588.
[4] R. Fallah-Moghaddam, M. Mahdavi-Hezavehi, *Locally finite conditions on maximal subgroups of $GL_n(D)$*, Algebra Colloq., **19**(1):73-86, 2012.
[5] J. Z. Goncalves, A. Mandel, *Are there free groups in division rings*, Israel J. Math., **51**(1):69-80, 1986.

[6] J. Z. Goncalves, *Free groups in subnormal subgroups and the residual nilpotence of the group of units of group rings*, Canad. Math. Bull., **27**:365-370, 1982.
[7] R. Hazrat, M. Mahdavi-Hezavehi, B. Mirzaii, Reduced K-Theory and the Group $G(D) = D^*/F^*D'$, Algebraic K-theory and its applications (Trieste, 1997), 403-409.
[8] A. E. Lichtman, *On subgroups of the multiplicative group of skew fields*, Proc. Amer. Math. Soc., **63**:15-16, 1977.
[9] M. Mahdavi-Hezavehi, *Free subgroups in maximal subgroups of* $GL_1(D)$, J. Algebra, **241**:720-730, 2001.
[10] M. Mahdavi-Hezavehi, Commutators in division rings revisted, Bull. Iranian Math. Soc. 26(2) (2000) 7–88.
[11] W. R. Scott, Group theory, Dover, New York, 1987.
[12] M. Shirvani, B. A. F. Wehrfritz, Skew linear groups, London Math. Soc. Lecture Note Ser., Vol. 118, 1986.
[13] D. A. Suprunenko, Matrix groups, Amer. Math. Soc. Providence, RI, 1976.
[14] J. Tits, *Free subgroups in linear groups*, J. Algebra, **20**:250-270, 1972.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–277, pp. 419–421

Frattini subgroup and skew linear groups

R. Fallah-Moghaddam[a,*]

[a]*Department of Mathematics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran*

Article Info

Keywords:
Division ring
Maximal subgroup
Frattini subgroup

2020 MSC:
16K20
20H25

Abstract

The Frattini subgroup of a group G is defined as the intersection of all maximal subgroups of G. If we have no maximal subgroups in G, we set $\phi(G) = G$. For finite groups G the Frattini subgroup, being a nilpotent subgroup. For infinite groups, we have a result of Wehrfritz in [10], which says that the Frattini subgroup of a finitely generated linear group is nilpotent. Here we investigate the Frattini subgroup of the general skew linear group $A^* = GL_n(D)$ for various division rings D. Let $A = M_n(D)$ be an F-central simple algebra of finite dimension its center. Assume that F is an absolutely algebraic field such that contains no q-th root of unity for any prime number q, then $F^* \bigcap \phi(A^*) = F^* Z(A')$.

1. Intrduction

Assume that D is a division algebra, when $F = Z(D)$. Set $A := M_n(D)$ the full $n \times n$ matrix ring over D and also let $A^* := GL_n(D)$ the general skew linear group over D. Consider that $A' := SL_n(D)$ the derived group of A^*. We say that a subgroup M of A^* is maximal in A^*if for any subgroup L of A^* with $M \subseteq L$, we obtain that $L = A^*$. When the dimension $[A : F]$ of A/F is finite, it is known that the group $G(A) := A^*/F^*A'$ is periodic of a bounded exponent dividing the index of A. Consequently, by Prufer–Baer Theorem , we obtain that normal maximal subgroups in A^*.

The Frattini subgroup of a group G is defined as the intersection of all maximal subgroups of G. If we have no maximal subgroups in G, we set $\phi(G) = G$. For finite groups G the Frattini subgroup, being a nilpotent subgroup. It is known that $\Phi(G)$ is equal to the set of all non-generators or non-generating elements of G.Notice that non-generating elements of a greoup G are elements such that can always be removed from a generating set. Which means an element a of G such that whenever X is a generating set of G containing a. Also, $\Phi(G)$ is always a characteristic subgroup of G; and it is a normal subgroup of G. When G is finite, we know that $\Phi(G)$ is nilpotent. Where G is a finite p-group, thus $\Phi(G) = G^p[G, G]$. In addition, the Frattini subgroup is the smallest normal subgroup N such that the quotient group G/N is an elementary abelian group and isomorphic to a direct sum of cyclic groups of order p. Also,, if the quotient group $G/\Phi(G)$ has order p^k, then k is the smallest number of generators for G. Also, a finite p-group is cyclic if and only if its Frattini quotient is cyclic. We say that a finite p-group is elementary abelian if and only if its

*Talker
Email address: r.fallahmoghaddam@cfu.ac.ir (R. Fallah-Moghaddam)

Frattini subgroup is the trivial group.

It seems that the most difficult case in the investigation of the structure of Frattini subgroup of $GL_n(D)$ is the case $n = 1$, since we do not know in general if we have maximal subgroups in $GL_1(D)$ for a noncommutative division ring D (cf. [5]). However, for the real quaternion division algebra H, we will be able to determine the Frattini subgroup of $GL_n(H)$ for all $n \geq 1$. In fact, it is shown that $Frat(GL_n(H)) = Z(GL_n(H))$. It is proved that, if F is an algebraic number field, then the group of roots of unity, denoted by $\mu(F)$, contains $2t$ elements. In this case it is proved that

$$Frat(F^*) = Frat(\mu(F)) \cong Frat(\mathbb{Z}_{2t}).$$

Also, if A is an F-central simple algebra of finite dimension over an algebraic number field, then

$$Frat(\mu(F))Z(A') \subseteq F^* \cap Frat(A^*) \subseteq \cap_p(\mu(F)^p)Z(A'),$$

where p is a prime with $p|(2t/(2t, [A : F]))$. If $(2t, [A : F]) = 1$, then $F^* \cap Frat(A^*) = Frat(\mu(F))$. If F is a local field with residue class field of p elements, set $U^{(n)} := 1 + p^n O$, where O is the valuation ring. It is proved that for any prime number $q \neq p$, $(U^{(1)})^q = U^{(1)}$ and $Frat(F^*) = (Frat(\mu_{p-1})) \times (U^{(1)})^p$. Furthermore, if A is of index $m := \sqrt{[A : F]}$ and $G := F^* \cap Frat(A^*)$, it is shown that $G = (\cap_q F^{*q})Z(A')$, where q is a prime with $(q, m) = 1$.

For more results see [1], [2],[3], [4], [8], [9], [10] and [11].

2. Main Result

For infinite groups, we have a result of Wehrfritz in [10], which says that the Frattini subgroup of a finitely generated linear group is nilpotent. Not much is known about the structure of the Frattini subgroup for skew linear groups. Here, we investigate the Frattini subgroup of the general skew linear group $A^* = GL_n(D)$ for various division rings D. If $n > 1$, Lemma 1 shows that $\phi(A^*)$ is central. For example, it is proved that:

Theorem A. Given a global field F, let A be a finite dimensional F-central simple algebra. Assume that the group $\mu(F)$ of roots of unity of F contains r elements. Then, we have

$$\phi(\mu(F))Z(A') \subset F^* \bigcap \phi(A^*) \subset \bigcap_p (\mu(F)^p)Z(A')$$

where p is a prime with

$$p|(r/(r, [A : F])).$$

If $(r, [A : F]) = 1$, then $F^* \bigcap \phi(A^*) = \phi(\mu(F))$.

In this manner, we prove the following result:

Main Result. Let $A = M_n(D)$ be an F-central simple algebra of finite dimension its center. Assume that F is an absolutely algebraic field such that contains no q-th root of unity for any prime number q, then $F^* \bigcap \phi(A^*) = F^*Z(A')$.

Proof. For any prime p, set $\mathbb{F}_p$ be its prime subfield, where $Char(F) = p$. An absolutely algebraic field, or an aaf, is an algebraic extension of $\mathbb{F}_p$. One may easily check that for any aaf F we have $F = \bigcup_{n \in S} \mathbb{F}_{p^n}$, where S is a nonempty subset of the positive integers such that for any $n, m \in S$ we have $\mathbb{F}_{p^{lcm(n,m)}} \subseteq F$. And, if $n \in S$ and $x|n$, then $\mathbb{F}_{p^x} \subseteq F$. The above conditions are necessary and sufficient conditions for when F is an absolutely algebraic field or aaf. It is also clear that any aaf is perfect.

We have $F^* \bigcap \phi(A^*) \subseteq F^* \bigcap_{i \in I} M_i$, where M_i are normal maximal subgroups of A^* of index p with $(p, [A : F]) = 1$. When $F^* = F^{*p}Z(A')$, then F^* does not any maximal subgroup of index p containing $Z(A')$. Consequently, A^* has not any normal maximal subgroup M of index p such that $F^* \nsubseteq M$ and thus $F^* \bigcap_{i \in I} M_i \subseteq F^{*p}Z(A') = F^*$.

Consider that $F^* \neq F^{*p}Z(A')$. Using Prufer-Baer Theorem, we conclude that $F^*/F^{*p}Z(A')$ is isomorphic to a direct product $\mathbb{Z}_p$.
On the other hand, we know that for an commutative group G, $\phi(G) = \bigcap G^p$, when p is an arbitrary prime number. it is clear that $F^{*p} = F^*$. But, since F contains no q-th root of unity for any prime number q, so $F^{*q} = F^*$. Now, by simple calculation we conclude that $F^* \bigcap \phi(A^*) = F^*Z(A')$.

3. ACKNOWLEDGMENT

The author thanks the Research Council of Farhangian University for support.

References

[1] H. Dorbidi, R. Fallah-Moghaddam, M. Mahdavi-Hezavehi, Soluble maximal subgroups of $GL_n(D)$, J. Algebra Appl., **9**(6):921-932, 2010.
[2] P. K. Draxl, Skew fields, Cambridge Univ. Press, Cambridge, 1983.
[3] R. Hazrat and A. U. Vishne, Triviality of the functor Coker($K_1(F) \to K_1(D)$) for division algebras, Commun. Algebra 33 (2005) 1427–1435.
[4] G. Karpilosky, Unit Groups of Classical Rings (Clarendon Press, Oxford, 1988). $G(D) = D^*/F^*D'$, Algebraic K-theory and its applications (Trieste, 1997), 403-409.
[5] T. Keshavarzipour, M. Mahdavi-Hezavehi, On the non-triviality of $G(D)$ and the existence of maximal subgroups of $GL_1(D)$, J. Algebra 285(2005)213-221.
[6] T. Y. Lam, A First Course in Noncommutative Rings, in: Grad.Texts in Math., Vol. 131, Springer-Verlag, Berlin, 2001.
[7] M. Mahdavi-Hezavehi, Commutators in division rings revisted, Bull. Iranian Math. soc.26(200), no. 2, 7-88.
[8] M. Mahdavi-Hezavehi, Commutators in division rings revisted, Bull. Iranian Math. Soc. 26(2) (2000) 7–88.
[9] D. J. S. Robinson, A Course in the Theory of Groups, in Graduate Text in Math., Vol. 80 (Springer-Verlag, 1982).
[10] M. Shirvani, B. A. F. Wehrfritz, Skew linear groups, London Math. Soc. Lecture Note Ser., Vol. 118, 1986.
[11] D. A. Suprunenko, Matrix groups, Amer. Math. Soc. Providence, RI, 1976.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–278, pp. 422–424

Crossed product condition and skew linear groups

R. Fallah-Moghaddam[a,*]

[a] *Department of Mathematics Education, Farhangian University, P.O. Box 14665-889, Tehran, Iran*

Article Info

Keywords:
Division ring
Maximal subgroup
Crossed Product

2020 MSC:
16K20
20H25

Abstract

Let D be a noncommutative F-central division algebra and N a subnormal subgroup of $GL_n(D)$. If M is a non-abelian maximal subgroup of N, such that M' is a locally finite group, then, $n = 1$ and D is cyclic of prime degree p with a maximal cyclic subfield K/F such that the groups $Gal(K/F)$ and $M/(K^* \cap N)$ are isomorphic. Furthermore, for any $x \in M \setminus K^*$, we have $x^p \in F^*$ and $D = F[M] = \bigoplus_{i=1}^{p} Kx^i$. So, $M_n(D)$ is a cyclic crossed product.

1. Introduction

Let R be a ring and assume that S and K are subsets of ring R, the subring generated by K and S is showed by $K[S]$. The multiplicative group of R is written as R^*. When G is a group and subset $S \subset G$, $Z(G)$ and $C_G(S)$ are the center and the centralizer of S in G, also the same notations are applied for R. $N_G(S)$ is used for the normalizer of S in G and G' for the derived group.
Assume that R be a ring and S a sub-ring of R and G be a subgroup of the multiplicative group R normalizing S such that $R = S[G]$. Suppose that $N = S \cap G$ is a normal subgroup of G and $R = \bigoplus_{t \in T} tS$, where T is some transversal of N to G. Set $E = G/N$. Then, we say that R is a crossed product of S by E and we denote it by (R, S, G, E).

A locally finite group is a group for when every finitely generated subgroup is finite. The cyclic subgroups of a locally finite group are finitely generated hence finite, every element has finite order, and so the group is torsion. In group theory, a locally finite group is a type of group that can be studied in ways analogous to a finite group. Sylow subgroups, Carter subgroups, and abelian subgroups of locally finite groups have been studied. This topic is credited to work in the 1930s by Russian mathematician Sergei Chernikov. A field of positive characteristic is called locally finite if every finite subset of the field is contained in a finite subfield.

Assume that a division ring D be a division ring with center F and also consider that G a subgroup of $GL_n(D)$, the space of column n-vectors $V = D^n$ over D is a G-D bi-module. G is called irreducible (resp. completely reducible, reducible) if V is irreducible (resp. completely reducible, reducible) as G-D bimodule. Furthermore, G is absolutely

*Talker
Email address: r.fallahmoghaddam@cfu.ac.ir (R. Fallah-Moghaddam)

irreducible if $F[G] = M_n(D)$.

For more information on these concepts, please refer to [2], [3], [4], [5], [8], [9] and [10].

D is said to be supersoluble crossed product if $Gal(K/F)$ is supersoluble. We also recall that a subgroup G of D^* is irreducible if $F[G] = D$. When $n = p$, a prime, it is shown that D is cyclic if and only 1 if D^* contains a nonabelian soluble subgroup. this result is generalized to a division algebra of arbitrary degree n. To be more precise, it is proved that D is supersoluble crossed product if and only if D^* contains an irreducible abelianby- supersoluble subgroup. it is presented a criterion for D to be nilpotent (abelian or cyclic) crossed product. In fact, it is shown that a noncommutative finite dimensional F-central division algebra D is nilpotent (abelian or cyclic) crossed product if and only if there exist an irreducible subgroup G of D^* and an abelian normal subgroup A of G such that G/A is nilpotent (abelian or cyclic). We recall that soluble subgroups of the multiplicative group of a division ring were first studied by Suprunenko.

2. Main Results

In references [1], [6], [4], [5] and [8], various conditions have been examined to determine which group properties cause the condition of the crossed product to occur. For example. it is proved that:

Theorem A. Let D be a noncommutative F-central division algebra and N a subnormal subgroup of $GL_n(D)$. If M is a non-abelian soluble maximal subgroup of N, then, $n = 1$ and D is cyclic of prime degree p with a maximal cyclic subfield K/F such that the groups $Gal(K/F)$ and $M/(K^* \cap N)$ are isomorphic. Furthermore, for any $x \in M \setminus K^*$, we have $x^p \in F^*$ and $D = F[M] = \bigoplus_{i=1}^{p} Kx^i$. So, $M_n(D)$ is a cyclic crossed product.

In this manner, we prove the following result:

Main Result. Let D be a noncommutative F-central division algebra and N a subnormal subgroup of $GL_n(D)$. If M is a non-abelian maximal subgroup of N, such that M' is a locally finite group, then, $n = 1$ and D is cyclic of prime degree p with a maximal cyclic subfield K/F such that the groups $Gal(K/F)$ and $M/(K^* \cap N)$ are isomorphic. Furthermore, for any $x \in M \setminus K^*$, we have $x^p \in F^*$ and $D = F[M] = \bigoplus_{i=1}^{p} Kx^i$. So, $M_n(D)$ is a cyclic crossed product.

Proof. In order not to prolong the discussion, we present a brief proof. Since $M^{'}$ is a locally finite normal subgroup of M, then by using Corollary 5.4.6 of [8], $M/C_M(M')$ is locally finite and therefore, it has a metabelian normal subgroup of finite index. Consider that G is a normal subgroup of M such that G/F^* is a metabelian normal subgroup of M/F^* and $[M/F^* : G/F^*] < \infty$. Thus, we conclude that $[M : G] < \infty$ and $G^{''} \subseteq F$. We may conclude that either $F[G] = M_n(D)$ or $G \subseteq Z(M) = F^*$. Consider that $G \subseteq F^*$. Thus $[M : F^*] < \infty$. We obtain at a contradiction. So, G is non-central and $F[G] = M_n(D)$.

Now, consider that $G^{'} \nsubseteq F$. We know that $G^{'} \subseteq M^{'}$ and $M^{'}$ is a locally finite group, we conclude that $G^{'}$ is a locally finite group. Using Theorem 1.1.14 of [8], we conclude that $F[G^{'}]$ is a simple Artinian ring. We obtain that $M \subseteq N_N(F[G^{'}]^*)$, by maximality of M, two cases may occur, i.e., either $M = N_N(F[G^{'}]^*)$ or $N = N_N(F[G^{'}]^*)$.

If $M = N_{GL_n(D)}(F[G^{'}]^*)$, then $F[G^{'}]^* \cap N \subseteq M$. We obtain that $G^{'}$ is non-central abelian normal subgroup of M, which reduces to the previous case.

Consider $N = N_{GL_n(D)}(F[G^{'}]^*)$. Using Theorem 14.3.8 of [7] and Corollary 1 of [**?**], we have either $G^{'} \subseteq F$ or $F[G^{'}] = M_n(D)$. The first case cannot happen. Now, let $F[G^{'}] = M_n(D)$. So, $G^{''} \subseteq F$ we have $G^{''} \subseteq Z(G^{'})$. This conclude that $G^{'}$ is nilpotent. Thus, by 2.5.2 of [8], $G^{'}$ is abelian-by-finite. The group ring $FG^{'}$ satisfies a polynomial identity. This implies that $F[G^{'}]$ satisfies a polynomial identity, and hence D satisfies a polynomial identity. Now, we have $[D : F] < \infty$. Since M is an absolutely irreducible skew linear group, we conclude that M is an irreducible linear group (cf. [9, p. 100]). So, by Theorem 6 of [9, p. 135], M contains an abelian normal subgroup H of finite index. If $H \subseteq F^*$, thus M/F^* is finite. We arrive at a contradiction. So, H is non-central, which reduces to the previous case.

Now, assume that $G^{'} \subseteq F$. Then, G is nilpotent. With a similar argument as before, we coclude that either $M' \cap G \subseteq F^*$ or $F[M' \cap G] = M_n(D)$. $M' \cap G$ is a locally finite nilpotent group. Therefore, the second case cannot happen. Now, assume $M' \cap G \subseteq F^*$. Since $F[M'] = M_n(D)$ and $[M' : M' \cap G] < \infty$, we conclude that $[D : F] < \infty$. Finally, we conclude that M is soluble and conclusion is obtained.

3. ACKNOWLEDGMENT

The author thanks the Research Council of Farhangian University for support.

References

[1] H. Dorbidi, R. Fallah-Moghaddam, M. Mahdavi-Hezavehi, Soluble maximal subgroups of $GL_n(D)$, J. Algebra Appl., **9**(6):921-932, 2010.
[2] P. K. Draxl, Skew fields, Cambridge Univ. Press, Cambridge, 1983.
[3] J. Z. Goncalves, A. Mandel, Are there free groups in division rings, Israel J. Math., **51**(1):69-80, 1986.
[4] J. Z. Goncalves, Free groups in subnormal subgroups and the residual nilpotence of the group of units of group rings, Canad. Math. Bull., **27**:365-370, 1982.
[5] D. S. Passman, Group Ring Satisfying a Polynomial Identity, J. Algebra **20**:103-117, 1972.
[6] M. Ramezan-Nassab, D. Kiani, (Locally soluble)-by-(locally finite) maximal subgroups of $GL_n(D)$, J. Algebra, **376**:1-9, 2013.
[7] W. R. Scott, Group theory, Dover, New York, 1987.
[8] M. Shirvani, B. A. F. Wehrfritz, Skew linear groups, London Math. Soc. Lecture Note Ser., Vol. 118, 1986.
[9] D. A. Suprunenko, Matrix groups, Amer. Math. Soc. Providence, RI, 1976.
[10] B. A. Wehrfritz, Goldie subrings of Artinian rings generated by groups, Quart. J. Math., **40**:501-512, 1989.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–281, pp. 425–431

On left reversible weakly flat acts

Leila Nouri[a], Basira Monfared[b,*]

[a] *Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran*
[b] *Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran*

Article Info

Keywords:
(weakly) right reversible
LR-weakly flat
LR-right reversible

2020 MSC:
20M30
20M50

Abstract

In this paper, a generalization of weak flatness of acts, called left reversible weakly flat (LR-weakly flat) is defined. Also we introduce condition LR-right reversible as a generalization of right reversible, and we show that weakly right reversible does not imply LR-right reversible. Finally, some basic results of property LR-weakly flat is given.

1. Introduction

Throughout this paper, unless otherwise stated, S always will stand for a monoid and LR-weakly flat will stand for a left reversible weakly flat. For S, a nonempty set A is called a right S-act, usually denoted A_S, if S acts on A unitarily from the right, that is, there exists a mapping $A \times S \to A$, $(a, s) \mapsto as$, satisfying the conditions $(as)t = a(st)$ and $a1 = a$, for all $a \in A$ and all $s, t \in S$. Also, we denote by $\Theta_S = \{\theta\}$ a one-element right S-act. A right S-act A_S is called *weakly flat* if the functor $A_S \otimes_S -$ preserves all monomorphisms from left ideals of S into S. This is equivalent to say that, $as = a't$, for $a, a' \in A_S, s, t \in S$ implies $a \otimes s = a' \otimes t$ in $A_S \otimes_S (Ss \cup St)$ (see [4]).
We recall from [3, 4] that a submonoid P of S is called *weakly right reversible* if for every $s, t \in P$, $z \in S$, $sz = tz$ implies the existence of $u, v \in P$ such that $us = vt$, and S is called *right (left) reversible*, if for any $s, t \in S$, there exists $u, v \in S$ such that $us = vt$ $(su = tv)$.

An element s of S is called *right e-cancellable* for an idempotent $e \in S$, if $s = es$ and $ker\rho_s \leq ker\rho_e$. The monoid S is called *left PP* if every principal left ideal of S is projective, as a left S-act. This is equivalent to saying that every element $s \in S$ is right e-cancellable for some idempotent $e \in S$ (see [2]). The monoid S is called *left PSF* if every principal left ideal of S is strongly flat, as a left S-act. This is equivalent to saying that S is right semi-cancellative, that is, whenever $su = s'u$, for $s, s', u \in S$, there exists $r \in S$ such that $u = ru$ and $sr = s'r$ (see [6]). It is obvious that every left PP monoid is left PSF.

*Talker
Email addresses: Leila_Nouri@math.usb.ac.ir (Leila Nouri), Basira.monfared@pgs.usb.ac.ir (Basira Monfared)

2. Basic Results

In this section, we introduce a generalization of weak flatness property and give some basic results. Also we introduce condition LR-right reversible as a generalization of right reversible, and we show that weakly right reversible does not imply LR-right reversible.

Definition 2.1. Let A_S be a right S-act. We say that A_S is *left reversible weakly flat* or *LR-weakly flat* if $as = a't$ and $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$ imply $a \otimes s = a' \otimes t$ in $A_S \otimes {}_S(Ss \cup St)$.

Theorem 2.2. *The following statements are hold:*

(1) *If $\{A_i \mid i \in I\}$ is a chain of subacts of an act A_S and every A_i, $i \in I$ is LR-weakly flat, then $\bigcup_{i\in I} A_i$ is LR-weakly flat.*

(2) *$\coprod_{i\in I} A_i$, where A_i, $i \in I$, are right S-acts, is LR-weakly flat if and only if A_i is LR-weakly flat, for every $i \in I$.*

(3) *The right S-act S_S is LR-weakly flat.*

Proof. Proofs are obvious. □

Lemma 2.3. [4, Lemma 2.5.5] *Let A_S be a right S-act and ${}_SM$ be a left S-act. Then $a\otimes m = a'\otimes m'$, for $a, a' \in A_S$ and $m, m' \in {}_SM$, if and only if there exist $s_1, ..., s_k, t_1, ...t_k \in S$, $b_1, ..., b_{k-1} \in A_S$ and $n_1, ..., n_k \in {}_SM$ such that*

$$\begin{array}{rr} & s_1n_1 = m \\ as_1 = b_1t_1 & s_2n_2 = t_1n_1 \\ b_1s_2 = b_2t_2 & s_3n_3 = t_2n_2 \\ \dots & \dots \\ b_{k-1}s_k = a't_k & m' = t_kn_k. \end{array}$$

Theorem 2.4. *Any retract of any LR-weakly flat right S-act is LR-weakly flat.*

Proof. Let A_S be a retract of B_S, which B_S is LR-weakly flat. Let $as = a't$, $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$. Since A_S is a retract of B_S, thus there exist homomorphisms $\varphi : A_S \to B_S$ and $\psi : B_S \longrightarrow A_S$, such that $\psi\varphi = 1_A$. Then we have $\varphi(as) = \varphi(a't)$ or $\varphi(a)s = \varphi(a')t$. Since $\varphi(a), \varphi(a') \in B_S$, by assumption we have $\varphi(a) \otimes s = \varphi(a') \otimes t$ in $B_S \otimes {}_S(Ss \cup St)$. Hence there exist $t_1, ..., t_k$, $u_1, ..., u_k \in S$, $r_1, ..., r_k \in Ss \cup St$ and $b_1, ..., b_{k-1} \in B_S$ such that

$$\begin{array}{rr} & t_1r_1 = s \\ \varphi(a)t_1 = b_1u_1 & t_2r_2 = u_1r_1 \\ b_1t_2 = b_2u_2 & t_3r_3 = u_2r_2 \\ \dots & \dots \\ b_{k-1}t_k = \varphi(a')u_k & t = u_kr_k. \end{array}$$

Then $\psi(\varphi(a)t_1) = \psi(b_1u_1)$, and so $at_1 = \psi(b_1)u_1$. Similarly, $\psi(b_{i-1})t_i = \psi(b_i)u_i$, for $2 \leq i \leq k-1$, and $\psi(b_{k-1})t_k = a'u_k$. Let $\psi(b_i) = a_i$, for $i \in \{1, ..., k-1\}$. Now if we substitute $\psi(b_i)$ by a_i, for $i \in \{1, ..., k-1\}$, then we obtain $a \otimes s = a' \otimes t$ in $A_S \otimes {}_S(Ss \cup St)$. □

Definition 2.5. A submonoid P of S is called *LR-right reversible* if for every $s, s' \in P$, $z_1, z_2 \in S$, $sz_1 = s'z_2$ implies the existence of $u, v \in P$ such that $us = vs'$.

Indeed, right reversible $\Rightarrow$ LR-right reversible $\Rightarrow$ weakly right reversible. The following example shows that weakly right reversible does not imply LR-right reversible.

Example 2.6. Let $S = \left\{ \begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix} \mid a, b \in \mathbb{Z}, a \neq 0 \right\}$. Then S is a right cancellative monoid, and so it is weakly right reversible, but S is not LR-right reversible, since

$$\begin{pmatrix} 3 & 0 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix},$$

but for every $a, b, c, d \in \mathbb{Z}$, with $ac \neq 0$,

$$\begin{pmatrix} a & 0 \\ b & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 3 & 1 \end{pmatrix} \neq \begin{pmatrix} c & 0 \\ d & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix}.$$

Proposition 2.7. *Θ_S is LR-weakly flat if and only if S is LR-right reversible.*

Proof. Necessity. Let $sz_1 = s'z_2$, for $s, s', z_1, z_2 \in S$. Since $\theta s = \theta s'$, thus $\theta \otimes s = \theta \otimes s'$ in $\Theta_S \otimes {}_S(Ss \cup Ss')$, and so by Lemma 2.3 there exist $s_1, ..., s_n, t_1, ..., t_n, r_1, ..., r_n \in S$, which $r_i \in \{s, s'\}$ for $1 \leq i \leq n$, such that

$$\begin{aligned} t_1 r_1 &= s \\ t_2 r_2 &= s_1 r_1 \\ t_3 r_3 &= s_2 r_2 \\ &\dots \\ s' &= s_n r_n. \end{aligned}$$

Now let $s_0 = 1$ and $r_0 = s$. If $r_i = s$, for $1 \leq i \leq n$, then $s' = s_n s$ implies that S is LR-right reversible, otherwise let j is a first index such that $r_j = s'$. Then $t_j s' = s_{j-1} s$ implies that S is LR-right reversible.

Sufficiency. Let $\theta s = \theta t$ and $sz_1 = tz_2$, for $s, t, z_1, z_2 \in S$. By assumption there exist $u, v \in S$, such that $us = vt$. Hence $\theta \otimes s = \theta u \otimes s = \theta \otimes us = \theta \otimes vt = \theta v \otimes t = \theta \otimes t$ in $\Theta_S \otimes {}_S(Ss \cup St)$, and so Θ_S is LR-weakly flat, as required. ☐

Theorem 2.8. *Let S be left reversible. Then every LR-weakly flat right S-act is weakly flat.*

Proof. Suppose A_S be a LR-weakly flat right S-act. Let $as = a't$, for $a, a' \in A_S$ and $s, t \in S$. Since S is left reversible, there exist $z_1, z_2 \in S$, such that $sz_1 = tz_2$, and so by assumption $a \otimes s = a' \otimes t$ in $A_S \otimes {}_S(Ss \cup St)$. Thus A_S is weakly flat, as required. ☐

Theorem 2.9. *Every LR-weakly flat right S-act is principally weakly flat.*

Proof. Proof is obvious. ☐

Definition 2.10. A right S-act A_S satisfies *Condition LR-W*, if $as = a't$ and $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$, then there exist $a'' \in A_S$ and $w \in Ss \cap St$, such that $as = a't = a''w$.

Theorem 2.11. *A right S-act A_S is LR-weakly flat, if and only if it is principally weakly flat and satisfies Condition LR-W.*

Proof. Necessity. A_S is principally weakly flat, by Theorem 2.9. Now let $as = a't$ and $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$. By assumption, $a \otimes s = a' \otimes t$ in $A_S \otimes {}_S(Ss \cup St)$. Thus there exist $s_1, ..., s_k, t_1, ..., t_k \in S$, $u_1, ..., u_k \in Ss \cup St$, $b_1, ..., b_{k-1} \in A_S$ such that

$$\begin{aligned} & & s_1 u_1 &= s \\ a s_1 &= b_1 t_1 & s_2 u_2 &= t_1 u_1 \\ b_1 s_2 &= b_2 t_2 & s_3 u_3 &= t_2 u_2 \\ &\dots & &\dots \\ b_{k-1} s_k &= a' t_k & t &= t_k u_k. \end{aligned}$$

If $u_j \in Ss$, for $1 \leq j \leq k$, then $t = t_k u_k$ implies that $t = us$, for some $u \in S$. Thus we can take $w = t$ and $a'' = a'$. Now let j be the first index such that $u_j \in St$. If $j = 1$ then $s = s_1 u_1 \in St$, and so $s = vt$, for some $v \in S$. Thus we can take $w = s$ and $a'' = a$. Suppose $j > 1$. Since $u_{j-1} \in Ss$, $s_j u_j = t_{j-1} u_{j-1}$ implies that $w = s_j u_j \in Ss \cap St$ and so $as = as_1 u_1 = b_1 t_1 u_1 = \cdots = b_{j-1} s_j u_j = b_{j-1} w$. Thus we can take $a'' = b_{j-1}$, and so Condition LR-W is satisfied, as required.

Sufficiency. Let $as = a't$, $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$. Since A_S satisfies Condition LR-W, there exist $a'' \in A_S$ and $w \in Ss \cap St$, such that $as = a't = a''w$. Also $w = us = vt$, for some $u, v \in S$. Since A_S is principally weakly flat, thus $as = a''us$ implies that $a \otimes s = a''u \otimes s$ in $A_S \otimes {}_S Ss$, and $a't = a''vt$ implies that $a' \otimes t = a''v \otimes t$ in $A_S \otimes {}_S St$. Hence,

$$a \otimes s = a''u \otimes s = a'' \otimes us = a'' \otimes vt = a''v \otimes t = a' \otimes t$$

in $A_S \otimes {}_S(Ss \cup St)$, and so A_S is LR-weakly flat, as required. □

By using Theorem 2.11, the proof of the following theorem is similar to that of [7, Theorem 2.18].

Theorem 2.12. *For any family $\{A_i\}_{i \in I}$, $\prod_{i \in I} A_i$ is LR-weakly flat if and only if it is principally weakly flat and for any $s, t, z_1, z_2 \in S$, if $sz_1 = tz_2$ and $A_i s \cap A_i t \neq \emptyset$, for every $i \in I$, then for every $(w_i)_I \in \prod_{i \in I}(A_i s \cap A_i t)$, there exist $(a_i'')_I \in \prod_{i \in I} A_i$ and $u \in Ss \cap St$ such that $(w_i)_I = (a_i'')_I u$.*

By a similar argument as in the proof of [7, Theorem 2.21], for a commutative monoid we can show LR-weakly flat can be transferred from products of acts to their components.

Theorem 2.13. *For a commutative monoid S, if $\prod_{i \in I} A_i$ is LR-weakly flat, then A_i is LR-weakly flat, for every $i \in I$.*

Notice that in Theorem 2.13 commutativity of S is a sufficient condition. See the following example.

Example 2.14. If S is not LR-right reversible, then S is not right reversible, and so is not commutative. Since $(\Theta \times S)_S \cong S_S$ and S_S is LR-weakly flat, $(\Theta \times S)_S$ is also LR-weakly flat, but Θ_S is not LR-weakly flat, by Proposition 2.7.

Theorem 2.15. *The following statements are equivalent:*

(1) *If $A_S = \prod_{i \in I} A_i$ is LR-weakly flat, then A_i is LR-weakly flat, for every $i \in I$.*

(2) *If $A_S = \prod_{i \in I} A_i$ is LR-weakly flat, then A_i satisfies Condition LR-W, for every $i \in I$.*

(3) *Θ_S is LR-weakly flat.*

(4) *Θ_S satisfies Condition LR-W.*

(5) *S is LR-right reversible.*

(6) *There exists a LR-weakly flat right S-act containing a zero.*

(7) *There exists a right S-act which containing a zero and satisfies Condition LR-W.*

Proof. Implications $(1) \Rightarrow (2)$, $(3) \Rightarrow (4)$ and $(6) \Rightarrow (7)$ are obvious, by Theorem 2.11. Also $(3) \Leftrightarrow (6)$ and $(4) \Leftrightarrow (7)$, since Θ_S is a retract of any act containing zero. (4) and (5) are equivalent, by Proposition 2.7.

$(2) \Rightarrow (3)$. Since $S_S \cong S_S \times \Theta_S$, it is obvious, by Theorem 2.2.

$(2) \Rightarrow (1)$. It follows from [7, Lemma 2.12] and Theorem 2.11.

$(5) \Rightarrow (2)$. Suppose $A_S = \prod_{i \in I} A_i$ is LR-weakly flat, and let $a_i s = a'_i t$, $sz_1 = tz_2$, for $a_i, a'_i \in A_i$, $s, t, z_1, z_2 \in S$. Since S is LR-right reversible, there exist $u_1, v_1 \in S$ such that $u_1 s = v_1 t$. For every $j \in I \setminus \{i\}$, let a_j be a fix element of A_j, and define

$$c_j = \begin{cases} a_j u_1 & j \neq i \\ a_i & j = i \end{cases}, \quad d_j = \begin{cases} a_j v_1 & j \neq i \\ a'_i & j = i \end{cases}.$$

Thus $(c_j)_I s = (d_j)_I t$, and so by assumption, there exist $(a''_j)_I \in \prod_{i \in I} A_i$ and $w \in Ss \cap St$ such that $(c_j)_I s = (d_j)_I t$ $= (a''_j)_I w$. Hence $a_i s = a'_i t = a''_i w$, and so A_i satisfies Condition LR-W, as required. □

Theorem 2.16. *Let ρ be a right congruence on S. Then the right S-act S/ρ is LR-weakly flat if and only if for every $s, t, z_1, z_2 \in S$, with $s\rho t$ and $sz_1 = tz_2$, there exist $u, v \in S$, such that $1(\rho \vee ker\rho_s)u$, $1(\rho \vee ker\rho_t)v$ and $us = vt$.*

Proof. Necessity. Let $s\rho t$ and $sz_1 = tz_2$, for $s, t, z_1, z_2 \in S$. Then $[1]_\rho s = [1]_\rho t$. Since S/ρ satisfies Condition LR-W, there exist $x, y, w \in S$, such that $xs = yt$ and $[1]_\rho s = [w]_\rho xs = [w]_\rho yt$. If $u = wx$ and $v = wy$, then $us = vt$, $s\rho(us)$ and $t\rho(vt)$. Since S/ρ is principally weakly flat, by [4, Proposition 3.10.7], we get from $s\rho(us)$ that $1(\rho \vee ker\rho_s)u$ and from $t\rho(vt)$ that $1(\rho \vee ker\rho_t)v$.

Sufficiency. Let $[x]_\rho s = [y]_\rho t$ and $sz_1 = tz_2$, for $x, y, s, t, z_1, z_2 \in S$. Then $(xs)\rho(yt)$, and so by assumption there exist $u, v \in S$, such that $1(\rho \vee ker\rho_{xs})u$, $1(\rho \vee ker\rho_{yt})v$ and $uxs = vyt$. By [4, Lemma 3.10.6], we have $[1]_\rho \otimes xs = [u]_\rho \otimes xs$ in $S/\rho \otimes Sxs$ and $[1]_\rho \otimes yt = [v]_\rho \otimes yt$ in $S/\rho \otimes Syt$. Hence

$$\begin{gathered}[x]_\rho \otimes s = [1]_\rho \otimes xs = [u]_\rho \otimes xs = [1]_\rho \otimes uxs = [1]_\rho \otimes vyt = \\ [v]_\rho \otimes yt = [1]_\rho \otimes yt = [y]_\rho \otimes t,\end{gathered}$$

in $S/\rho \otimes_S (Ss \cup St)$. Therefore S/ρ is LR-weakly flat, as required. □

Corollary 2.17. *Let $z \in S$. Then the principal right ideal zS is LR-weakly flat, if and only if for all $s, t, z_1, z_2 \in S$, $zs = zt$ and $sz_1 = tz_2$ imply that there exist $u, v \in S$, such that $us = vt$, $1(ker\lambda_z \vee ker\rho_s)u$ and $1(ker\lambda_z \vee ker\rho_t)v$.*

Proof. Since $zS \cong S/ker\lambda_z$, for every $z \in S$, it is sufficient to apply Theorem 2.16, for $\rho = ker\lambda_z$. □

We know that every weakly flat right S-act is LR-weakly flat. Now from Theorem 2.9 and [4, Theorem 3.12.16] we have the following theorem.

Theorem 2.18. *Let $w, t \in S$, where $wt \neq t$. Then the following statements are equivalent:*

(1) *$S/\rho(wt, t)$ is flat.*

(2) *$S/\rho(wt, t)$ is weakly flat.*

(3) *$S/\rho(wt, t)$ is LR-weakly flat.*

(4) *$S/\rho(wt, t)$ is principally weakly flat.*

(5) *t is a regular element in S.*

Recall from [5] that a right ideal K of S is called *left stabilizing*, if for every $k \in K$ there exists $l \in K$, such that $lk = k$.

Theorem 2.19. *Let K_S be a right ideal of S. Then S/K_S is LR-weakly flat if and only if S is LR-right reversible and K_S is left stabilizing.*

Proof. Necessity. Suppose S/K_S is LR-weakly flat. Then there are two cases that can arise:

Case 1. $K_S = S$. Then $S/K_S \cong \Theta_S$ is LR-weakly flat, and so S is LR-right reversible, by Propotion 2.7.

Case 2. $K_S \neq S$. Then S/K_S is principally weakly flat, by Theorem 2.11, and so K_S is left stabilizing, by [4, Theorem 3.10.11]. To show that S is LR-right reversible, let $sz_1 = tz_2$, for $s, t, z_1, z_2 \in S$. Suppose $k \in K_S$. Then $[k]_{\rho_K} s = [k]_{\rho_K} t$, and since by Theorem 2.11, S/K_S satisfies Condition LR-W, there exist $u, v \in S$ such that $us = vt$.

Sufficiency. Suppose S is LR-right reversible and K_S is left stabilizing. Then there are two cases that can arise:

Case 1. $K_S = S$. Since S is LR-right reversible, $S/K_S \cong \Theta_S$ is LR-weakly flat, by Propotion 2.7.

Case 2. $K_S \neq S$. Let $s\rho_K t$ and $sz_1 = tz_2$. Then here are two cases that can arise:

2.1. $s = t$. If $u = 1 = v$, then S/K_S is LR-weakly flat, by Theorem 2.16.

2.2. $s \neq t$. Then $s, t \in K_S$. Since K_S is left stabilizing, there exist $l_1, l_2 \in K_S$, such that $l_1 s = s$ and $l_2 t = t$, that is $l_1 \ker \rho_s 1$ and $l_2 \ker \rho_t 1$. Since $sz_1 = tz_2$ and S is LR-right reversible, there exist $u', v' \in S$, such that $u's = v't$. Let $u = l_1 u'$ and $v = l_1 v'$. Then $1 \ \ker \rho_s \ l_1 \ \rho_K u$ and so $1(\rho_K \vee \ker \rho_s)u$. Similarly, $1(\rho_K \vee \ker \rho_t)v$. Since $us = l_1 u's = l_1 v't = vt$, S/K_S is LR-weakly flat, by Theorem 2.16.

From above cases S/K_S is LR-weakly flat, as required. □

Now from Theorem 2.9 and [4, Proposition 3.12.19] we have the following theorem.

Theorem 2.20. *Let K_S be a proper right ideal of S. Then the following statements are equivalent:*

(1) *The right S-act $S \coprod^K S$ is flat.*

(2) *The right S-act $S \coprod^K S$ is weakly flat.*

(3) *The right S-act $S \coprod^K S$ is LR-weakly flat.*

(4) *The right S-act $S \coprod^K S$ is principally weakly flat.*

(5) *K_S is left stabilizing.*

It is shown in [1] that for a left PP monoid S, A_S is weakly flat if and only if $as = a't$, for $a, a' \in A_S$, $s, t \in S$ implies that there exist $a'' \in A_S$, $u, v \in S$ and $e, f \in E(S)$ such that $es = s$, $ft = t$, $ae = a''ue$, $a'f = a''vf$ and $us = vt$. In a similar way we can show the following theorem for LR-weakly flat.

Theorem 2.21. *Let S be a left PP monoid. An act A_S is LR-weakly flat if and only if, for every $a, a' \in A_S$ and $s, t, z_1, z_2 \in S$, $as = a't$ and $sz_1 = tz_2$ imply that there exist $a'' \in A_S$, $u, v \in S$, and $e, f \in E(S)$ such that $es = s$, $ft = t$, $ae = a''ue$, $a'f = a''vf$ and $us = vt$.*

Now we give a similar theorem for PSF monoid.

Theorem 2.22. *Let S be a left PSF monoid. An act A_S is LR-weakly flat if and only if, for every $a, a' \in A_S$ and $s, t, z_1, z_2 \in S$, $as = a't$ and $sz_1 = tz_2$ imply that there exist $a'' \in A_S$ and $u, v, r, r' \in S$, such that $rs = s$, $r't = t$, $ar = a''ur$, $a'r' = a''vr'$ and $us = vt$.*

Proof. Necessity. Let $as = a't$, $sz_1 = tz_2$, for $a, a' \in A_S$, $s, t, z_1, z_2 \in S$. Since, by Theorem 2.11, A_S satisfies Condition LR-W, there exist $a'' \in A_S$ and $u, v \in S$ such that $as = a''us$, $a't = a''vt$ and $us = vt$. Again by Theorem 2.11, A_S is principally weakly flat, and so $a \otimes s = a''u \otimes s$ and $a \otimes t = a''v \otimes t$ in $A_S \otimes_S Ss$ and $A_S \otimes_S St$, respectively. Then by Lemma 2.3, there exist $s_1, ..., s_k, t_1, ...t_k \in S$ and $b_1, ..., b_{k-1} \in A_S$ such that

$$\begin{array}{rr} & s_1 s = s \\ as_1 = b_1 t_1 & s_2 s = t_1 s \\ b_1 s_2 = b_2 t_2 & s_3 s = t_2 s \\ \ldots & \ldots \\ b_{k-1} s_k = a'' u t_k & s = t_k s. \end{array}$$

Let $t_0 = 1$ and $s_{k+1} = 1$. Since S is right semi-cancellative, $s_1 s = t_0 s$ implies the existence of $r_1 \in S$, such that $r_1 s = s$ and $s_1 r_1 = t_0 r_1$. Then $s_2 r_1 s = t_1 r_1 s$ implies the existence of $r_2 \in S$ such that $r_2 s = s$ and $s_2 r_1 r_2 = t_1 r_1 r_2$. If $z = r_1 r_2$, then

$$zs = r_1 r_2 s = s, s_1 z = s_1 r_1 r_2 = t_0 r_1 r_2 = t_0 z, s_2 z = t_1 z.$$

Continuing this procedure, there exists $r \in S$, such that $rs = s$ and $s_i r = t_{i-1} r$, for $1 \leq i \leq k+1$. Thus we have

$$\begin{aligned} ar &= a(t_0 r) = a(s_1 r) = (as_1)r = (b_1 t_1)r = b_1(t_1 r) = b_1(s_2 r) = (b_1 s_2)r \\ &= \cdots = (b_{k-1} s_k)r = (a'' u t_k)r = a'' u(t_k r) = a'' u(s_{k+1} r) = a'' ur. \end{aligned}$$

A similar argument shows that there exists $r' \in S$ such that $r't = t$ and $a'r' = a''vr'$.

Sufficiency. Suppose $as = a't$ and $sz_1 = tz_2$, for $a, a' \in A_S$ and $s, t, z_1, z_2 \in S$. By assumption, there exist $a'' \in A_S$ and $u, v, r, r' \in S$, such that $rs = s$, $r't = t$, $ar = a''ur$, $a'r' = a''vr'$ and $us = vt$. Thus

$$\begin{aligned} a \otimes s &= a \otimes rs = ar \otimes s = a''ur \otimes s = a''u \otimes rs = a''u \otimes s = a'' \otimes us \\ &= a'' \otimes vt = a''v \otimes t = a''v \otimes r't = a''vr' \otimes t = a'r' \otimes t \\ &= a' \otimes r't = a' \otimes t, \end{aligned}$$

in $A_S \otimes_S (Ss \cup St)$, and so A_S is LR-weakly flat, as required. □

References

[1] S. Bulman-Fleming, K. McDowell, Monoids over which all weakly flat acts are flat, *Proc. Edinburgh Math. Soc.* **33** (1990), 287-298.
[2] J. Fountain, Right PP monoids with central idempotents, *Semigroup Forum* **13** (1977), 229–237.
[3] A. Golchin and H. Mohammadzadeh, On Condition (P'), *Semigroup Forum* **86** (2013), 413–430.
[4] M. Kilp, U. Knauer and A. Mikhalev, Monoids, Acts and Categories, Walter de Gruyter, Berlin, 2000.
[5] V. Laan, Pullbacks and flatness properties of acts I, *Comm. Algebra* **29(2)** (2001), 829–850.
[6] Z.K. Liu, Y.B. Yang, Monoids Over Which Every Flat Right Act Satisfies Condition (P), *Comm. Algebra* **22**(8) (1994), 2861–2875.
[7] L. Nouri, A. Golchin and H. Mohammadzadeh, On properties of product acts over monoids, *Comm. Algebra* **43** (2015), 1854–1876.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–284, pp. 432–434

On Classification of 4-dimensional nilpotent complex Leibniz algebras

Seyed Jalal Langari

Department of Mathematics, Farhangian University, Tehran, Iran.

Article Info

Keywords:
Leibniz algebra
4-dimensional
Classification

2020 MSC:
17A32
17A36

Abstract

Leibniz algebras introduced by J. L. Loday (1993) are non-antisymmetric generalizations of Lie algebras. The classification problem of complex nilpotent Leibniz algebras were first studied by Loday himself. Ismail Demir, Kailash C. Misra and Ernie Stitzinger obtained the classification of 4-dimensional complex nilpotent Leibniz algebras. Through reviewing their classification, we concluded that some algebras should be omitted from the list.

1. Introduction and preliminaries

Leibniz algebras were first introduced by Loday in [4] as a non-antisymmetric version of Lie algebras. The classification problem of complex nilpotent Leibniz algebras were first studied by Loday himself. He obtained the complete classification of complex nilpotent Leibniz algebras of dimension $n \leq 2$. Later Ayupov and Omirov classified $3-$dimensional complex nilpotent Leibniz algebras in [2]. Albeverio, Omirov and Rakhimov, also I. Demir, K. C. Misra, and E. Stitzinger obtained the classification of 4-dimensional complex nilpotent Leibniz algebras in [1] and [6]. One of the techniques to classify nilpotent Lie algebras were introduced by Skjelbred and Sund in [8]. Rakhimov and Langari were the first researchers who used Skjelbred-Sund method in Leibniz algebras [3]. They also applied in [7] and [3] this technique to obtain the classification of complex nilpotent Leibniz algebras of dimension $n \leq 4$. By comparing our classification with classification in [1] we realized that the Skjelbred-Sund method works also very well. In this part we give the basic definitions and properties for Leibniz algebras.

Definition 1.1. A Leibniz algebra L is a vector space over a field F equipped with a bilinear map

$$[\cdot,\cdot] : L \times L \longrightarrow L$$

satisfying the Leibniz identity

$$[x,[y,z]] = [[x,y],z] - [[x,z],y], \quad \text{for all } \; x,y,z \in L.$$

Email address: jalal_langari@yahoo.com (Seyed Jalal Langari)

Obviously, a Lie algebra is a Leibniz algebra. A Leibniz algebra is a Lie algebra if and only if

$$[x,x]=0, \text{ for all } x\in L.$$

Let n be the dimension of Leibniz algebra L. Let $\{e_1,e_2,...,e_n\}$ be a basis in L. The structural constants of L are the numbers C_{ij}^k given by

$$[e_i,e_j]=\sum_{k=1}^{n}C_{ij}^k e_k \quad (i,j=1,...,n).$$

We can identify the Leibniz identity with its structural constants. These constants satisfy:

$$\sum_{l=1}^{n}(C_{jk}^l C_{il}^s - C_{ij}^l C_{lk}^s + C_{ik}^l C_{lj}^s)=0 \quad (i,j,k,s=1,...,n).$$

Definition 1.2. Let L is a Leibniz algebra. We define

$$L^1=L, L^k=\left[L^{k-1},L\right] \quad (k>1).$$

The series

$$L^1\supseteq L^2\supseteq L^3\supseteq ...$$

is called the descending central series of L. If the series terminates for some positive integer s, then the Leibniz algebra L is said to be nilpotent.

Definition 1.3. A Leibniz algebra L is said to be split if it can be written as a direct sum of two nontrivial ideals. Otherwise, L is called non-split.

Theorem 1.4. *[6] Let A be a non-split non-Lie nilpotent Leibniz algebra of $dim(A)=4$. Then A is isomorphic to a Leibniz algebra spanned by $\{x_1,x_2,x_3,x_4\}$ with the nonzero products given by the following:*

$$\begin{cases}
A_1\colon & [x_1,x_3]=x_4,[x_3,x_2]=x_4;\\
A_2\colon & [x_1,x_3]=x_4,[x_2,x_2]=x_4,[x_2,x_3]=x_4,[x_3,x_2]=-x_4,[x_3,x_1]=x_4;\\
A_3\colon & [x_1,x_2]=x_4,[x_2,x_1]=-x_4,[x_3,x_3]=x_4;\\
A_4\colon & [x_1,x_2]=x_4,[x_2,x_1]=-x_4,[x_2,x_2]=x_4,[x_3,x_3]=x_4;\\
A_5(\alpha)\colon & [x_1,x_2]=x_4,[x_2,x_1]=\alpha x_4,[x_3,x_3]=x_4, \ \alpha\in C\setminus\{-1,1\};\\
A_6\colon & [x_1,x_1]=x_4,[x_2,x_2]=x_4,[x_3,x_3]=x_4;\\
A_7\colon & [x_1,x_1]=x_2,[x_1,x_2]=x_3,[x_1,x_3]=x_4;\\
A_8\colon & [x_1,x_1]=x_4,[x_1,x_2]=x_3,[x_2,x_1]=-x_3;\\
A_9\colon & [x_1,x_1]=x_4,[x_1,x_2]=x_3,[x_2,x_1]=-x_3,[x_2,x_2]=x_4;\\
A_{10}\colon & [x_1,x_1]=x_4,[x_1,x_2]=x_3,[x_2,x_1]=-x_3,[x_1,x_3]=x_4,[x_3,x_1]=-x_4;\\
A_{11}\colon & [x_1,x_2]=x_3,[x_2,x_1]=-x_3,[x_2,x_2]=x_4,[x_1,x_3]=x_4,[x_3,x_1]=-x_4;\\
A_{12}\colon & [x_1,x_1]=x_4,[x_1,x_2]=x_3,[x_2,x_1]=-x_3+x_4,[x_1,x_3]=x_4,[x_3,x_1]=-x_4;\\
A_{13}\colon & [x_1,x_1]=x_3,[x_2,x_1]=-x_3+x_4,[x_2,x_2]=x_4,[x_1,x_3]=x_4,[x_3,x_1]=-x_4;\\
A_{14}\colon & [x_1,x_1]=x_3,[x_1,x_2]=x_4;\\
A_{15}\colon & [x_1,x_1]=x_3,[x_2,x_1]=x_4;\\
A_{16}\colon & [x_1,x_2]=x_4,[x_2,x_1]=x_3,[x_2,x_2]=-x_3;\\
A_{17}(\alpha)\colon & [x_1,x_1]=x_3,[x_1,x_2]=x_4,[x_2,x_1]=\alpha x_4, \ \alpha\in C\setminus\{-1,0\};\\
A_{18}(\alpha)\colon & [x_1,x_1]=x_3,[x_2,x_1]=x_4,[x_1,x_2]=\alpha x_3,[x_2,x_2]=-x_4, \ \alpha\in C\setminus\{-1\};\\
A_{19}\colon & [x_1,x_1]=x_3,[x_1,x_2]=x_3,[x_2,x_1]=x_3+x_4,[x_2,x_3]=x_4;\\
A_{20}\colon & [x_1,x_2]=x_3,[x_1,x_3]=x_4;\\
A_{21}\colon & [x_1,x_2]=x_3,[x_2,x_2]=x_4,[x_1,x_3]=x_4;\\
A_{22}\colon & [x_1,x_2]=x_3,[x_2,x_1]=x_4,[x_1,x_3]=x_4;\\
A_{23}\colon & [x_1,x_2]=x_3,[x_2,x_1]=x_4,[x_2,x_2]=x_4,[x_1,x_3]=x_4;\\
A_{24}\colon & [x_1,x_1]=x_3,[x_2,x_1]=x_4,[x_1,x_3]=x_4;\\
A_{25}\colon & [x_1,x_1]=x_3,[x_2,x_2]=x_4,[x_1,x_3]=x_4;
\end{cases}$$

2. Main results

In this section we prove that $A_3 \cong A_5(\alpha = -1)$, $A_6 \cong A_5(\alpha = 1)$, $A_8 \cong A_{17}(\alpha = -1)$, $A_{14} \cong A_{17}(\alpha = 0)$ and $A_{15} \cong A_{18}(\alpha = -1)$. Therefore, the Leibniz algebra A_3, A_6, A_8, A_{14}, and A_{15} should be omitted from the list 1.4. For example here we put:

A$_8$: $[x_1, x_1] = x_4, [x_1, x_2] = x_3, [x_2, x_1] = -x_3;$

and

A$_{17}(\alpha = -1)$: $[x_1', x_1'] = x_3', [x_1', x_2'] = x_4', [x_2', x_1'] = -x_4'.$

The matrix A representing this linear transformation with respect to these bases is

$$A = \begin{bmatrix} c_{11} & 0 & 0 & 0 \\ c_{21} & c_{22} & 0 & 0 \\ c_{31} & c_{32} & 0 & c_{11}^2 \\ c_{41} & c_{42} & c_{22} & 0 \end{bmatrix};$$

where $\det(A) = -c_{11}^4 c_{22}^2$. We can easily get the following constrains for matrix A such that $\det(A) \neq 0$:

$$c_{21} = c_{31} = c_{32} = c_{41} = c_{42} = 0,\ c_{11} = c_{22} = 1.$$

Thus we can get:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Writing the elements of basis $A_{17}(\alpha = -1)$ in terms of the basis A_8 we have

$$\begin{cases} x_1 = x_1'; \\ x_2 = x_2'; \\ x_3 = x_4'; \\ x_4 = x_3'. \end{cases}$$

It shows that $A_{17}(\alpha = -1) \cong A_8$.

References

[1] S. Albeverio, B. A. Omirov and I.S Rakhimov, *Classification of 4-dimensional nilpotent complex Leibniz algebras*, Extracta mathematicae, 21(3), (2006), 197-210.

[2] Sh. A. Ayupov and B. A. Omirov, *On 3-dimensional Leibniz algebras*, Uzbek. Mat. Zh., (1), (1999), 9-14.

[3] S. J. Langari, *Classification of Low Dimensional Nilpotent Leibniz Algebras Using Central Extensions*, PhD thesis, (2010), University Putra Malaysia.

[4] J.-L. Loday, *Cyclic homology*, Grundl. Math. Wiss. Bd. 301, (1992), Springer-Verlag, Berlin.

[5] I. S. Rakhimov and S. J. Langari, *A cohomological approach for classifying nilpotent Leibniz algebras*, Int. J. Algebra, 4(1-4), (2010), 153-163.

[6] I. Demir, K. C. Misra and E. Stitzinger, On classification of four-dimensional nilpotent Leibniz Algebras, Comm. Algebra 45(3), 1012-1018, 2017

[7] I. S. Rakhimov, S. J. Langari, and M. B. Langari, *On central extensions of nulfiliform Leibniz algebras*, Int. J. Algebra, 3(5-8), (2009), 271-280.

[8] T. Skjelbred, T. Sund, *Technical report: On the classification of nilpotent Lie algebras*, Mathematisk Institutt, Universiteteti Oslo, (1977) .

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–286, pp. 435–437

Feebly Neatness Property of Some Classes of Rings

Neda Pouyan*

Faculty of Engineering, Shohadaye Hoveizeh Campus of Technology, Shahid Chamran University of Ahvaz, Dashte Azadegan, Iran.

Article Info

Keywords:
feebly clean ring
feebly neat ring
FGC ring

2020 MSC:
16U60
Secondary: 16D10, 16S50, 16S34.

Abstract

An element x of ring R has feebly cleanness property if x can be written as $x = f_1 - f_2 + v$ where v is a unit and f_1, f_2 are orthogonal idempotents of R. A ring R is feebly clean if every element in R has feebly cleanness property. We define R is feebly neat if every homomorphic image of R is feebly clean. In this paper, we begin with elementary properties of feebly neat rings. In the following, we investigate feebly cleanness for some classes of FGC rings.

1. Introduction

Throughout this paper all rings are considered associative with identity element. An element in a ring is called clean if it can be written as the sum of an idempotent and a unit. A ring is clean if every it's element is clean. This concept was introduced by Nicholson in 1977 (see[5]). Arora and Kundu [1] introduced the family of feebly clean rings which has the property that every element x can be written as $x = f_1 - f_2 + v$ where v is a unit and f_1, f_2 are orthogonal idempotents. A ring R is feebly clean if every element in R is feebly clean. One of the fundamental properties of feebly clean rings is that every homomorphic image of a feebly clean ring is feebly clean. We define a feebly neat ring to be one for which every proper homomorphic image is feebly clean. In this paper, we begin with elementary properties of feebly neat rings. In the following, we investigate feebly neatness property of some classes of FGC rings.

1.1. Main Results

The aim of this section is to examine the basic properties of feebly neat rings, which will be utilized in the next section. We begin with

Proposition 1.1. *The following are equivalent for a ring R.*

1. *R is feebly neat.*
2. *R/xR is feebly clean for every nonzero $x \in R$.*

*Talker
Email address: n.pouyan@scu.ac.ir & neda.pouyan@gmail.com (Neda Pouyan)

3. *R/xR is feebly neat for every $x \in R$.*
4. *R/J is feebly clean for every nonzero semiprime ideal J.*

Moreover, a homomorphic image of a feebly neat ring is feebly neat.

Corollary 1.2. *If R is a feebly neat ring which is not feebly clean, then R is semiprime.*

Proposition 1.3. *Let R be a decomposable ring. Then R is feebly neat if and only if it is feebly clean.*

Theorem 1.4. *Every zero-dimensional ring is feebly clean.*

Corollary 1.5. *If R is a domain of (Krull) dimension equal to 1, then R is feebly neat.*

Example 1.6. If F is a field and $R = F[X, Y]$, then R is not feebly neat as $R/YR \cong F[X]$ is not feebly clean. $F[X]$ is feebly neat by the previous theorem. Moreover, if $R[X]$ is feebly neat, then R is field.

*1.2. **FGC Rings***

Recall that, an FGC ring is known as a R ring when every finitely generated module is isomorphic to a direct sum of cyclics. This class of rings dates back to Kaplansky [4] who was interested in classifying rings which satisfied the generalization of the Fundamental Theorem of Finitely Generated Abelian Groups. FGC rings are classified in [3]. The classification states that an FGC ring is finite direct product of three types of rings. Before we classify feebly neat FGC rings we recall a few definitions.

Definition 1.7. Let R be a ring and M an R-module. We say M is a linearly compact R-module if every collection of cosets with the finite intersection property has nonempty intersection. It is known that a homomorphic image of a linearly compact R-module is linearly compact (see [3]). If R_R is a linearly compact, then we say R is maximal ring.

Example 1.8. Artinian rings are examples of maximal rings.

Definition 1.9. A ring R is said to be almost maximal if R/I is a linearly compact R-module for every nonzero ideal I of R.

Theorem 1.10. *If the ring A is maximal, then it is a finite direct product of local rings.*

Corollary 1.11. *Every maximal ring is feebly clean. Moreover, an almost maximal ring is feebly neat.*

A ring R is called h-local if it is of finite character and every proper homomorphic image is a pm-ring. To be of finite character means that every element is contained in a finite number of maximal ideals. Recall that a ring is a Bézout ring if every finitely generated ideal is principal.

Definition 1.12. A ring R is called a torch ring if it satisfies the following conditions:

1. R is not local.
2. R has a unique minimal prime ideal P which is nonzero and whose R-submodule form a chain.
3. R/P is an h-local domain.
4. R is a locally almost maximal Bézout ring.

Theorem 1.13. *(Brandal [[3], Theorem 9.1]). A ring is an FGC-ring if and only if it is a finite direct product of the following types of rings:*

1. *Maximal valuation rings.*
2. *Almost maximal Bézout domains.*
3. *Torch rings.*

Lemma 1.14. *A torch ring is never feebly neat.*

Theorem 1.15. *Suppose R is an FGC ring. R is feebly clean if and only if R is a finite direct product of local rings. In this case, it is a finite direct product of almost maximal valuation rings.*

Theorem 1.16. *Suppose R is an FGC ring. R is feebly neat if and only if R is either a feebly clean ring or it is an almost maximal Bézout domain which is not local.*

Corollary 1.17. *An FGC-domain is feebly neat.*

Almost maximal rings are feebly neat. Almost maximal domains were classified by Brandal:

Proposition 1.18. *(Brandal [2]). A ring is an almost maximal domain if and only if it is h-local and locally almost maximal.*

Lemma 1.19. *(Brandal [[3], Lemma 2.4]). Let I be an ideal of R which is contained in a finite number of maximal ideals. Then R/I is a direct sum of indecomposable modules of the form R/J , where $I \leq J$.*

Proposition 1.20. *(Brandal [3], Proposition 2.5]). Let I be an ideal of R such that $W(I)$ is finite. Then R/I is indecomposable if and only if for all nontrivial partitions W_1, W_2 of $W(I)$ there are $N_1 \in W_1$, $N_2 \in W_2$ and a prime ideal P of R such that $I \subseteq P \subseteq N_1 \cap N_2$.*

Proposition 1.21. *Suppose $W(I)$is finite and R/I is a pm-ring. Then R/I is a finite direct product of local rings.*

We are now in position to state our desired theorem whose proof is a consequence of the previous proposition.

Theorem 1.22. *An h-local domain is feebly neat.*

Acknowledgement

The author would like to thank Dr. Feroz Siddique (Associate Professor at University of Wisconsin-Eau Claire)for his valuable comments and insightful suggestions that led to the improvement of the paper.

References

[1] N. Arora and S. Kundu, Commutative feebly clean rings, J. Algebra and its Applications 5(2017), 1750128, 14p.
[2] W. Brandal, Almost maximal integral domains and finitely generated modules, Trans. Amer. Math. Soc. 183 (1973) 203–222.
[3] W. Brandal, Commutative Rings Whose Finitely Generated Modules Decompose, Lecture Notes in Mathematics, vol. 723, Springer, New York, 1979.
[4] I. Kaplansky, Modules over Dedekind rings and valuation rings, Trans. Amer. Math. Soc. 72 (1952) 327–340.
[5] W.K. Nicholson, Extensions of clean rings,Comm. Algebra 29(6), 2589-2595 (2001).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–172, pp. 438–447

A Generalization of The Pascal Functional Matrix Using the Generalized Fibonacci Polynomials Sequence

Morteza Bayat[a,*]

[a] *Department of Mathematics Zanjan Branch Islamic Azad University, Zanjan, Iran*

Article Info

Keywords:
The Pascal functional matrix
The Generalized Fibonacci Polynomials sequence
Characteristic polynomial

2020 MSC:
15A15
11B37
15A23
05A10

Abstract

In this paper, we first introduce the Right Justified Pascal functional matrix with three variables $P_n[x, y, z]$. Then, we obtain a decomposition and inverse of these new matrices using Pascal functional matrices. We also introduce the generalized Fibonacci sequence of three variables $\{F_n(x, y, z)\}_{n\geq 0}$ as a tool for obtaining eigenvalues, eigenvectors and characteristic polynomial of the matrix $P_n[x, y, z]$. Finally, we present some applications of this new kind of matrix including some combinational identities related to the sequence $\{F_n(x, y, z)\}_{n\geq 0}$ and a diagonal decomposition of $P_n[x, y, z]$ based on Vandermond matrix.

1. Introduction

In [1] and [2] Peele and Stăniča studied the matrices with the $(i, j)-$entries the binomial coefficient $\binom{i-1}{j-1}$ and $\binom{i-1}{n-j}$ respectively. Then have also obtained some interesting results above the powers of these matrices. In [3], they obtained eigenvalues, eigenvectors and characteristic polynomial of the matrix $\mathcal{P}_n = \left[\binom{i-1}{n-j}\right]_{0\leq i,j\leq n}$ using the standard Fibonacci sequence. In this paper, as a generalization of previous aforementioned research, we first introduce the Right Justified Pascal functional matrix of three variables $\mathcal{P}_n[x, y, z]$. Next, by generalizing the ordinary Fibonacci sequence to the generalized three variables $\{F_n(x, y, z)\}_{n\geq 0}$, we use the similar strategy of [3] to obtain the eigenvalues, eigenvectors and characteristic polynomial of $\mathcal{P}_n[x, y, z]$. Finally, we present some applications of these new kind of matrices including some combinational identities related to the sequence $\{F_n(x, y, z)\}_{n\geq 0}$ and a diagonal decomposition of $\mathcal{P}_n[x, y, z]$ based on Vandermond matrices.

Definition 1.1. For three variables x, y, z, the generalized Fibonacci matrix of three variables of order $(n+1)\times(n+1)$ is defined by

$$\mathcal{P}_n[x, y, z] = \left[x^{i+j-n} y^{n-j} z^{n-i} \binom{i}{n-j}\right]_{0\leq i,j\leq n}.$$

*Talker
Email address: baayyaatt@gmail.com (Morteza Bayat)

Example 1.2. the right justified Pascal functional matrix of three variables of order 4×4 is as follows

$$\mathcal{P}_4[x,y,z] = \begin{pmatrix} 0 & 0 & 0 & z^3 \\ 0 & 0 & yz^2 & xz^2 \\ 0 & y^2z & 2xyz & x^2z \\ y^3 & 3y^2x & 3yx^2 & x^3 \end{pmatrix}.$$

In the following lemma we present a decomposition and inverse for the right justified Pascal functional matrix of three variables.

Lemma 1.3.

$$\mathcal{P}_n[x,y,z] = \text{diag}(z^n, \cdots, z, 1)\mathcal{P}_n[x]\text{diag}(y^n, \cdots, y, 1), \tag{1}$$

$$\mathcal{P}_n^{-1}[x,y,z] = \left[(-x)^{n-i-j}y^{i-n}z^{j-n}\binom{n-i}{j}\right]_{0\le i,j\le n}. \tag{2}$$

Proof. Clearly we have (1). For finding $\mathcal{P}_n^{-1}[x,y,z]$, it is enough to find $\mathcal{P}_n^{-1}[x]$. Now, consider the matrix $\tilde{I} = [\delta_{i,n-j}]_{0\le i,j\le n}$, where $\delta_{i,n-j}$ is the Kronecker delta. It is easy to see that $\mathcal{P}_n[x] = P_n[x]\tilde{I}$, where $P_n[x] = \left[\binom{i}{j}x^{i-j}\right]_{0\le i,j\le n}$ is the Pascal functional matrix with one variable, has the following properties (see [4–6])

1. $P_n[x]P_n[y] = P_n[x+y]$
2. $P_n[x]P_n[-x] = P_n[0] = I_n$. In Particular, $P_n^{-1}[x] = P_n[-x]$.

Therefore

$$\mathcal{P}_n^{-1}[x] = P_n[x]\tilde{I}P_n[-x] = \left[\binom{n-i}{j}(-x)^{n-i-j}\right]_{0\le i,j\le n}.$$

Therefore, we present the inverse of the right justified Pascal functional matrix of three variables as follows $\mathcal{P}_n^{-1}[x,y,z] = \left[(-x)^{n-i-j}y^{i-n}z^{j-n}\binom{n-i}{j}\right]_{0\le i,j\le n}$. □

Example 1.4.

$$\mathcal{P}_4^{-1}[x,y,z] = \begin{pmatrix} -\frac{x^3}{y^3z^3} & \frac{3x^2}{y^3z^2} & -\frac{3x}{y^3z} & \frac{1}{y^3} \\ \frac{x^2}{y^2z^3} & -\frac{2x}{y^2z^2} & \frac{1}{y^2z} & 0 \\ -\frac{x}{yz^3} & \frac{1}{yz^2} & 0 & 0 \\ \frac{1}{z^3} & 0 & 0 & 0 \end{pmatrix}.$$

2. The Generalized Fibonacci Sequence And $\mathcal{P}_n[x,y,z]$

Definition 2.1. The generalized Fibonacci sequence of three variables x, y, z over the field $\mathbb{F}$ is defined by

$$zF_{n+1}(x,y,z) = yF_{n-1}(x,y,z) + xF_n(x,y,z) \tag{3}$$

$$F_0(x,y,z) = 0, \quad F_1(x,y,z) = 1. \tag{4}$$

By the linear recursive relation (3), the characteristic equation is $z\lambda^2 - x\lambda - y = 0$. If α and β are the roots of the characteristic equation, by the Binet formula, we have

$$F_n(x,y,z) = \frac{\alpha^n - \beta^n}{\alpha - \beta}. \tag{5}$$

Remark 2.2. It is possible that the characteristic equation has no root in the field $\mathbb{F}$. So, we consider the extended field of $\mathbb{F}$.

Example 2.3.

$$F_0(x,y,z)=0,\quad F_1(x,y,z)=1,\quad F_2(x,y,z)=\frac{y}{x},\quad F_3(x,y,z)=\frac{x}{z}+\frac{y^2}{z^2},\cdots.$$

The following theorem is the main result of this paper which gives the relationship between the characteristic polynomial of the right justified Pascal functional matrix of three variables $\mathcal{P}_n[x,y,z]$ and the generalized Fibonacci sequence of three variables $\{F_n(x,y,z)\}_{n\geq 0}$.

Theorem 2.4. *If* $\Big(F_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\Big)_{0\leq i\leq n}$ *is a column vector of* $(n+1)$*-dimension, then*

$$\mathcal{P}_n[x,y,z]\Big(F_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\Big)_{0\leq i\leq n}=z^n\Big(F_{l+1}^{n-i}(x,y,z)F_{l+2}^i(x,y,z)\Big)_{0\leq i\leq n}. \tag{6}$$

In other words the recursive sequence $\Big\{F_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\Big\}_{0\leq i\leq n}$ is generated by the characteristic polynomial of matrix $z^{-n}\mathcal{P}_n[x,y,z]$ for all $0\leq i\leq n$.

Proof. Let $\mathcal{P}_n[x,y,z]\Big(F_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\Big)_{0\leq i\leq n}=[a_{ij}]$, we have

$$\begin{aligned} a_{ij} &= \sum_{k=0}^{n}\binom{i}{n-k}x^{i+k-n}y^{n-k}z^{n-i}F_l^{n-k}(x,y,z)F_{l+1}^k(x,y,z)\\ &= \Big(zF_{l+1}(x,y,z)\Big)^{n-i}\sum_{k=n-i}^{n}\binom{i}{n-k}\Big(yF_l(x,y,z)\Big)^{n-k}\Big(xF_{l+1}(x,y,z)\Big)^{i+k-n},\end{aligned}$$

which by substituting $t=k-n+i+1$, we obtain

$$\begin{aligned} a_{ij} &= \Big(zF_{l+1}(x,y,z)\Big)^{n-i}\sum_{t=0}^{i}\binom{i}{t}\Big(yF_l(x,y,z)\Big)^{i-t}\Big(xF_{l+1}(x,y,z)\Big)^{t}\\ &= \Big(zF_{l+1}(x,y,z)\Big)^{n-i}\Big(yF_l(x,y,z)+xF_{l+1}(x,y,z)\Big)^{i}\\ &= z^n\Big(F_{l+1}^{n-i}(x,y,z)F_{l+2}^i(x,y,z)\Big).\end{aligned}$$

This completes the proof of Theorem 2.4. ☐

Corollary 2.5.

$$\begin{aligned}&F_{l+1}^{n-i}(x,y,z)F_{l+2}^i(x,y,z)=\\ &\sum_{i_1,\cdots,i_l}\binom{i}{n-i_1}\binom{i_1}{n-i_2}\cdots\binom{i_{l+1}}{n-i_l}x^{i-nl+2\sum_{t=1}^{l-1}i_t}y^{nl-\sum_{t=1}^{l}i_t}z^{-\sum_{t=1}^{l-1}i_t}.\end{aligned}$$

Proof. By induction on l and using (6), we have

$$\mathcal{P}_n^l[x,y,z](x^i)_{0\leq i\leq n}=z^n\Big(F_{l+1}^{n-i}(x,y,z)F_{l+2}^i(x,y,z)\Big)_{0\leq i\leq n}.$$

Now, if we consider the $i-$th rows, we get

$$\begin{aligned}&z^nF_{l+1}^{n-i}(x,y,z)F_{l+2}^i(x,y,z)=\Big(\mathcal{P}_n^l[x,y,z](x^i)_{i=0}^n\Big)_{i,0}\\ =&\sum_{i_1,\cdots,i_l}a_{i,i_1}\cdots a_{i_{l-1},i_l}x^{i_l}\\ =&\sum_{i_1,\cdots,i_l}\binom{i}{n-i_1}\binom{i_1}{n-i_2}\cdots\binom{i_{l-1}}{n-i_l}x^{i-nl+2\sum_{t=1}^{l-1}+i_t}y^{nl-\sum_{t=1}^{l}i_t}z^{-\sum_{t=1}^{l-1}i_t},\end{aligned}$$

which completes the proof. $\square$

Example 2.6.

$$
\begin{aligned}
&\mathcal{P}_3[x,y,z](F_l^{3-i}(x,y,z)F_{l+1}^i(x,y,z))_{0\le i\le 3} = \\
= &\begin{pmatrix} 0 & 0 & 0 & z^3 \\ 0 & 0 & yz^2 & xz^2 \\ 0 & y^2z & 2xyz & x^2z \\ y^3 & 3y^2x & 3yx^2 & x^3 \end{pmatrix} \begin{pmatrix} F_l^3(x,y,z) \\ F_l^2(x,y,z)F_{l+1}(x,y,z) \\ F_l(x,y,z)F_{l+1}^2(x,y,z) \\ F_{l+1}^3(x,y,z) \end{pmatrix} \\
= &\; z^3 \begin{pmatrix} F_{l+1}^3(x,y,z) \\ F_{l+1}^2(x,y,z)F_{l+2}(x,y,z) \\ F_{l+1}(x,y,z)F_{l+2}^2(x,y,z) \\ F_{l+1}^3(x,y,z) \end{pmatrix}.
\end{aligned}
$$

Theorem 2.7. *The matrix $[F_i^{n-i}(x,y,z)F_{j+1}^i(x,y,z)]$ is invertible, and hence, the sequence $\left\{F_i^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\right\}$ is nondegenerate of degree $n+1$.*

Proof. If we divide the $j-$th column by $F_{j+1}^i(x,y,z)$, we obtain the Vandermond matrix $\left[\left(\frac{F_j(x,y,z)}{F_{i+1}(x,y,z)}\right)^{n-i}\right]$ which has nonzero determinant. Hence

$$[F_i^{n-i}(x,y,z)F_{j+1}^i(x,y,z)]$$

is invertible and the sequence $\{F_i^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\}$ is nondegenerate of degree $n+1$. $\square$

3. The Characteristic Polynomial of $\mathcal{P}_n[x,y,z]$

Theorem 3.1. *The eigenvalues of $\mathcal{P}_n[x,y,z]$ are*

$$z^n\alpha^n, z^n\alpha^{n-1}\beta, \cdots, z^n\alpha\beta^{n-1}, z^n\beta^n,$$

and the characteristic polynomial of $\mathcal{P}_n[x,y,z]$ is

$$\chi_n(t) = \prod_{i=0}^{n}\Big(t - z^n\alpha^i\beta^{n-i}\Big).$$

Proof. By (5) we have

$$
\begin{aligned}
F_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z) &= \left(\frac{\alpha^l-\beta^l}{\alpha-\beta}\right)^{n-i}\left(\frac{\alpha^{l+1}-\beta^{l+1}}{\alpha-\beta}\right)^i. \\
&= \sum_{\omega\in S_n} A_\omega\omega^{l-1},
\end{aligned}
$$

where $S_n = \{\alpha^n, \alpha^{n-1}\beta, \cdots, \alpha\beta^{n-1}, \beta^n\}$. So the polynomial of $\chi_n(t) = \prod_{\omega\in S_n}^n (t - z^n\omega)$ generates the sequence $\{z^nF_l^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\}$. Again

$$\left\{F_i^{n-i}(x,y,z)F_{l+1}^i(x,y,z)\right\}$$

is non-degenerate of degree $n+1$, and the proof is complete. $\square$

Corollary 3.2.

$$tr(\mathcal{P}_n^k[x,y,z]) = \frac{F_{k(n+1)}(x,y,z)}{F_k(x,y,z)}.$$

Proof. By Theorem 3.1, we have

$$\begin{aligned} tr(\mathcal{P}_n^k[x,y,z]) &= \alpha^{nk} + \alpha^{(n-1)k}\beta^k + \cdots + \beta^{nk} \\ &= \frac{\alpha^{k(n+1)} - \beta^{k(n+1)}}{\alpha^k - \beta^k} \\ &= \frac{F_{k(n+1)}(x,y,z)}{F_k(x,y,z)}. \end{aligned}$$

□

Theorem 3.3.

$$\chi_n(t) = \sum_{i=0}^{n+1}(-1)^{\frac{i(i+1)}{2}}\left(\frac{y}{x}\right)^{\frac{i(i-1)}{2}} z^{ni}\begin{bmatrix} n+1 \\ i \end{bmatrix}_{F_n(x,y,z)} t^{n+1-i},$$

where $\begin{bmatrix} n+1 \\ i \end{bmatrix}_{F_n(x,y,z)}$ *is defined by*

$$\begin{bmatrix} n+1 \\ i \end{bmatrix}_{F_n(x,y,z)} = \begin{cases} 1, & j = 0, m; \\ \frac{F_m(x,y,z)\cdots F_{m-j+1}(x,y,z)}{F_j(x,y,z)\cdots F_1(x,y,z)}, & 0 < j < m. \end{cases}$$

Proof. We use the following identity (see [7])

$$\prod_{j=0}^{n}(1-q^j t) = \sum_{l=0}^{n+1}(-1)^l q^{\frac{l(l-1)}{2}-nl}\begin{bmatrix} n+1 \\ l \end{bmatrix}_q z^l,$$

where $\begin{bmatrix} n+1 \\ l \end{bmatrix}_q$ is the Gauss binomial coefficient which is defined by

$$\begin{bmatrix} n+1 \\ l \end{bmatrix}_q = \frac{(1-q^n)\cdots(1-q^{n-l+1})}{(1-q^l)\cdots(1-q)}.$$

Replacing q in the above equation by $\frac{\beta}{\alpha}$ and using the Binet formula, we have

$$\begin{bmatrix} n+1 \\ i \end{bmatrix}_q = \alpha^{i^2-(n+1)i}\begin{bmatrix} n+1 \\ i \end{bmatrix}_{F_n(x,y,z)}.$$

Therefore

$$\prod_{j=0}^{n}(1-\alpha^{-j}\beta^j t) = \sum_{l=0}^{n+1}(-1)^l \beta^{\frac{l(l-1)}{2}}\alpha^{\frac{l(l+1)}{2}-(n+1)l}\begin{bmatrix} n+1 \\ k \end{bmatrix}_{F_n(x,y,z)} t^l.$$

Substituting t by $z^n\alpha^n t^{-1}$ and using $\alpha\beta = \frac{-y}{z}$, we get

$$\prod_{j=0}^{n}(t - z^n\alpha^{n-j}\beta^j) = \sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{-y}{x}\right)^{\frac{k(k-1)}{2}} z^{nk}\begin{bmatrix} n+1 \\ k \end{bmatrix}_{F_n(x,y,z)} t^{n+1-l},$$

which is the desired result. □

Example 3.4. The characteristic polynomials of $\chi_n(t)$ for $n = 0, 1, 2$ are

$$\begin{aligned} \chi_0(t) &= t - 1 \\ \chi_1(t) &= t^2 - xt - zy \\ \chi_2(t) &= t^3 - (zy + x^2)t^2 - (zx^2y + y^2z^2)t + y^3z^3. \end{aligned}$$

4. Diagonalization of $\mathcal{P}_n[x,y,z]$

Let $n \geq 1$ and $C_n[x,y,z]$ be the component matrix for $\chi_n(t)$, where

$$C_n[x,y,z] = (c_{i,j}(x,y,z)), \quad i,j = 0,1,\cdots,n.$$

$$\begin{cases} c_{0,1} = 1, & \\ c_{i,i+1}(x,y,z) = z^n, & i = 1,\cdots,n-1; \\ c_{n,n-j}(x,y,z) = -(-1)^{\frac{(j+1)(j+2)}{2}} z^n \left(\frac{y}{x}\right)^{\frac{j(j+1)}{2}} \begin{bmatrix} n+1 \\ j+1 \end{bmatrix}_{F_n(x,y,z)}, & j = 0,1,\cdots,n-1; \\ c_{i,j}(x,y,z) = 0, & \text{otherwise}. \end{cases}$$

and

$$R_n[x,y,z] = \Big(r_{i,j}(x,y,z)\Big) \text{ and } M_n[x,y,z] = \Big(m_{i,j}(x,y,z)\Big), \quad i,j = 0,1,\cdots,n,$$

$$\begin{cases} r_{0,j}(x,y,z) = r_{1,j} = z^n \delta_{n,j}, & j = 0,1,\cdots,n, \\ r_{i,j}(x,y,z) = \binom{n}{j} F_{i-1}^{n-j}(x,y,z) F_i^j(x,y,z) \left(\frac{y}{x}\right)^{n-j} z^n, & \end{cases}$$

$$\begin{cases} m_{0,j}(x,y,z) = z^n \delta_{n,j}, & j = 0,1,\cdots,n, \\ m_{i,j}(x,y,z) = \binom{n}{j} F_i^{n-j}(x,y,z) F_{i+1}^j(x,y,z) \left(\frac{y}{x}\right)^{n-j} z^{2n}. & \end{cases}$$

Lemma 4.1. *For every positive integer k, we have*

$$\Big(\mathcal{P}_n^k[x,y,z]\Big)_{nj} = z^{nk} \binom{n}{j} \left(\frac{y}{x} F_k(x,y,z)\right)^{n-j} \Big(F_{k+1}(x,y,z)\Big)^j.$$

Proof. Let n be a fixed natural number. We will prove the assertion by induction on k. The above equality is valid for $k=0$. Now assume the results is valid for $k \geq 0$. Then, since $\mathcal{P}_n^{k+1}[x,y,z] = \mathcal{P}_n^k[x,y,z]\mathcal{P}_n[x,y,z]$, we have

$$\begin{aligned}
&\Big(\mathcal{P}_n^{k+1}[x,y,z]\Big)_{nj} = \sum_{i=0}^{n} \Big(\mathcal{P}_n^k[x,y,z]\Big)_{ni} \Big(\mathcal{P}_n[x,y,z]\Big)_{ij} \\
&= \sum_{i=0}^{n} z^{nk} \binom{n}{i} \left(\frac{y}{x} F_k(x,y,z)\right)^{n-i} \Big(F_{k+1}(x,y,z)\Big)^i \binom{i}{n-j} x^{i+j-n} y^{n-j} z^{n-i} \\
&= z^{n(k+1)} \left(\frac{y}{z} F_{k+1}(x,y,z)\right)^{n-j} \sum_{i=0}^{n} \binom{n}{n-j} \binom{j}{i+j-n} \\
&\times \left(\frac{x}{z} F_{k+1}(x,y,z)\right)^{i-n+j} \left(\frac{y}{z} F_k(x,y,z)\right)^{n-i} \\
&= \binom{n}{j} z^{n(k+1)} \left(\frac{y}{z} F_{k+1}(x,y,z)\right)^{n-j} \sum_{i=0}^{n} \binom{j}{i+j-n} \\
&\times \left(\frac{x}{z} F_{k+1}(x,y,z)\right)^{i+j-n} \left(\frac{y}{z} F_k(x,y,z)\right)^{n-i} \\
&= \binom{n}{j} z^{n(k+1)} \left(\frac{y}{z} F_{k+1}(x,y,z)\right)^{n-j} \sum_{m=0}^{j} \binom{j}{m} \left(\frac{x}{z} F_{k+1}(x,y,z)\right)^m \left(\frac{y}{z} F_k(x,y,z)\right)^{j-m} \\
&= \binom{n}{j} z^{n(k+1)} \left(\frac{y}{z} F_{k+1}(x,y,z)\right)^{n-j} \left(\frac{x}{z} F_{k+1}(x,y,z) + \frac{y}{z} F_k(x,y,z)\right)^j \\
&= \binom{n}{j} z^{n(k+1)} \left(\frac{y}{z} F_{k+1}(x,y,z)\right)^{n-j} (F_{k+2}(x,y,z))^j.
\end{aligned}$$

$\square$

Theorem 4.2.

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{z}\right)^{\frac{k(k-1)}{2}}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}\Big(F_{n-k+1}(x,y,z)\Big)^{n-j}\Big(F_{n-k+2}(x,y,z)\Big)^{j}=0.$$

Proof. The characteristic polynomials of $\mathcal{P}_n^k[x,y,z]$ is

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{z}\right)^{\frac{k(k-1)}{2}}z^{nk}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}t^{n+k-1}=0.$$

Now by the Cayley-Hamilton's Theorem, we get

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{x}\right)^{\frac{k(k-1)}{2}}z^{nk}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}(\mathcal{P}_n^k[x,y,z])^{n-k+l}=O, \tag{7}$$

where O denotes the $(k+1)\times(k+1)$ zero matrix. So by Lemma 4.1 and substituting this result into (7), we obtain

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{z}\right)^{\frac{k(k-1)}{2}}z^{nk}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}\Big(\mathcal{P}_n^{n-k+1}[x,y,z]\Big)_{nj}=0,$$

or

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{z}\right)^{\frac{k(k-1)}{2}}z^{nk}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}z^{n(n-k+1)}\binom{n}{j}$$
$$\times\left(\frac{y}{z}F_{n-k+1}(x,y,z)\right)^{n-j}(F_{n-k+2}(x,y,z))^{j}=0.$$

Therefore

$$\sum_{k=0}^{n+1}(-1)^{\frac{k(k+1)}{2}}\left(\frac{y}{z}\right)^{\frac{k(k-1)}{2}}z^{nk}\begin{bmatrix}n+1\\k\end{bmatrix}_{F_n(x,y,z)}\Big(F_{n-k+1}(x,y,z)\Big)^{n-j}\Big(F_{n-k+2}(x,y,z)\Big)^{j}=0.$$

□

Theorem 4.3. *For all n, we have*

$$M_n[x,y,z]=C_n[x,y,z]R_n[x,y,z]=R_n[x,y,z]\mathcal{P}_n[x,y,z].$$

Furthermore,

$$\mathcal{P}_n[x,y,z]=R_n^{-1}[x,y,z]C_n[x,y,z]R_n[x,y,z].$$

Proof. We first prove $M_n[x,y,z]=C_n[x,y,z]R_n[x,y,z]$. In fact, multiplying the first n rows of $C_n[x,y,z]$ by

$R_n[x,y,z]$, clearly we get the first n rows of $M_n[x,y,z]$. For the last row, for each $0 \le j \le n$, we have

$$\begin{aligned}
&\Big(C_n[x,y,z]R_n[x,y,z]\Big)_{nj} = \sum_{k=0}^{n}\Big(C_n[x,y,z]\Big)_{n,n-k}\Big(R_n[x,y,z]\Big)_{n-k,j}\\
&= \sum_{k=0}^{n} -(-1)^{\frac{(k+1)(k+2)}{2}} z^n \left(\frac{y}{z}\right)^{\frac{k(k+1)}{2}} \begin{bmatrix} n+1 \\ k+1 \end{bmatrix}_{F_n(x,y,z)}\\
&\times \binom{n}{j} F_{n-k-1}^{n-j}(x,y,z) F_{n-k}^{j}(x,y,z) \left(\frac{y}{z}\right)^{n-j} z^n\\
&= z^{2n}\binom{n}{j}\left(\frac{y}{z}\right)^{n-j}\sum_{t=1}^{n+1} -(-1)^{\frac{t(t+1)}{2}}\left(\frac{y}{z}\right)^{\frac{t(t-1)}{2}}\\
&\begin{bmatrix} n+1 \\ t \end{bmatrix}_{F_n(x,y,z)} F_{n-t}^{n-j}(x,y,z)F_{n-t+1}^{j}(x,y,z)\\
&= z^{2n}\binom{n}{j}\left(\frac{y}{z}\right)^{n-j}\Big[F_n^{n-j}(x,y,z)F_{n+1}^{j}(x,y,z)\\
&+\sum_{t=0}^{n+1} -(-1)^{\frac{t(t+1)}{2}}\left(\frac{y}{z}\right)^{\frac{t(t-1)}{2}}\begin{bmatrix} n+1 \\ t \end{bmatrix}_{F_n(x,y,z)} F_{n-t}^{n-j}(x,y,z)F_{n-t+1}^{j}(x,y,z)\Big]\\
&= z^{2n}\binom{n}{j}\Big(F_n(x,y,z)\Big)^{n-j}F_{n+1}^{j}(x,y,z)\left(\frac{y}{z}\right)^{n-j},
\end{aligned}$$

which is true by Theorem 4.2. This proves,

$$M_n[x,y,z] = C_n[x,y,z]R_n[x,y,z].$$

Since for each i, j with $0 \le i \le j \le n$, we have

$$\begin{aligned}
&\Big(R_n[x,y,z]\mathcal{P}_n[x,y,z]\Big)_{ij} = \sum_{k=0}^{n}\Big(R_n[x,y,z]\Big)_{ik}\Big(\mathcal{P}_n[x,y,z]\Big)_{kj}\\
&= \sum_{k=0}^{n}\binom{n}{k}\Big(F_{i-1}(x,y,z)\Big)^{n-k}\Big(F_i(x,y,z)\Big)^{k}\left(\frac{y}{z}\right)^{n-k} z^n x^{k+j-n}y^{n-j}z^{n-k}\binom{k}{n-j}\\
&= \binom{n}{j} z^{2n}\sum_{k=0}^{n}\binom{j}{n-k}x^{k+j-n}y^{2n-j-k}z^{-n}\Big(F_{i-1}(x,y,z)\Big)^{n-k}\Big(F_i(x,y,z)\Big)^{k}\\
&= \binom{n}{j} z^{2n}x^{j-n}y^{n-j}\sum_{t=0}^{j}\binom{j}{t}\left(\frac{y}{z}F_{i-1}(x,y,z)\right)^{t}\left(\frac{x}{z}F_i(x,y,z)\right)^{n-t}
\end{aligned}$$

$$\begin{aligned}
&= \binom{n}{j} z^{2n}x^{j-n}y^{n-j}\left(\frac{y}{z}F_{i-1}(x,y,z)+\frac{x}{z}F_i(x,y,z)\right)^{j}\\
&= \binom{n}{j} z^{2n}x^{j-n}y^{n-j}F_{i+1}^{j}(x,y,z)\left(\frac{x}{z}F_i(x,y,z)\right)^{n-j}\\
&= \binom{n}{j}\left(\frac{y}{z}\right)^{n-j}F_{i+1}^{j}(x,y,z)F_i^{n-j}(x,y,z)z^{2n}\\
&= \Big(M_n[x,y,z]\Big)_{ij},
\end{aligned}$$

we get $M_n[x,y,z] = R_n[x,y,z]\mathcal{P}_n[x,y,z]$. □

Example 4.4.

$$M_3[x,y,z]=\begin{pmatrix} 0 & 0 & 0 & z^3 \\ z^3y^3 & 3y^3z^3x & 3z^3yx^2 & z^3x^3 \\ x^3y^3 & 3y^2x^2(yz+x^2) & 3yx(yz+x^2)^2 & (yz+x^2)^3 \\ \frac{y^3}{z^4}(yz+x^2)^3 & \frac{3y^2x}{z^3}(yz+x^2)^2(2yz+x^2) & \frac{3yx^2}{z^3}(yz+x^2)(2yz+x^2)^2 & \frac{x^3}{z^3}(2yz+x^2)^3 \end{pmatrix}$$

$$\begin{aligned} C_3[x,y,z] &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & z^3 & 0 \\ 0 & 0 & 0 & z^3 \\ -y^6 & \frac{-y^3x}{z^3}(2yz+x^2) & \frac{y}{z^2}(yz+x^2)(2yz+x^2) & (2yz+x^2)x \end{pmatrix}, \\ R_3[x,y,z] &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & z^3 \\ y^3 & 3y^2x & 3yx^2 & x^3 \\ \frac{x^3y^3}{z^3} & \frac{3y^2x^2}{z^3}(yz+x^2) & \frac{3yx}{z^3}(yz+x^2)^2 & \frac{1}{z^3}(yz+x^2)^3 \end{pmatrix}, \end{aligned}$$

and so

$$M_3[x,y,z]=C_3[x,y,z]R_3[x,y,z].$$

Also,

$$\begin{aligned} R_3[x,y,z] &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & z^3 \\ y^3 & 3y^2x & 3yx^2 & x^3 \\ \frac{x^3y^3}{z^3} & \frac{3y^2x^2}{z^3}(yz+x^2) & \frac{3yx}{z^3}(yz+x^2)^2 & \frac{1}{z^3}(yz+x^2)^3 \end{pmatrix}, \\ \mathcal{P}_3[x,y,z] &= \begin{pmatrix} 0 & 0 & 0 & z^3 \\ 0 & 0 & z^2y & z^3 \\ 0 & y^2z & 2yxz & zx^2 \\ y^3 & 3y^2x & 3yx^2 & x^3 \end{pmatrix}, \end{aligned}$$

and therefore $M_3[x,y,z]=R_3[x,y,z]\mathcal{P}_3[x,y,z]$.

Let V_n be the Vandermonde matrix which is defined by

$$\begin{aligned} V_n &= \Delta[z^n\alpha^n, z^n\alpha^{n-1}\beta, \cdots, z^n\alpha\beta^{n-1}, z^n\beta^n] \\ &= \begin{pmatrix} 1 & 1 & \cdots & 1 & 1 \\ z^n\alpha^n & z^n\alpha^{n-1}\beta & \cdots & z^n\alpha\beta^{n-1} & z^n\beta^n \\ (z^n\alpha^n)^2 & (z^n\alpha^{n-1}\beta)^2 & \cdots & (z^n\alpha\beta^{n-1})^2 & (z^n\beta^n)^2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ (z^n\alpha^n)^n & (z^n\alpha^{n-1}\beta)^n & \cdots & (z^n\alpha\beta^{n-1})^n & (z^n\beta^n)^n \end{pmatrix}. \end{aligned}$$

By the relation between the component matrix and the Vandermonde matrix, we can obtain Theorems 4.6 and 4.7.

Lemma 4.5. *If A be the following matrix*

$$A=\begin{pmatrix} 0 & a_1 & 0 & \cdots & 0 \\ 0 & 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{n-1} \\ p_1 & p_2 & p_3 & \cdots & p_n \end{pmatrix},$$

that $a_1, a_2, \cdots, a_{n-1}$ are non-zero, then its eigenvalues are the roots of $p_1+p_2\lambda+\cdots+p_n\lambda^{n-1}=\lambda^n$ and

$$v_1=(\alpha,\alpha\lambda,,\alpha\lambda^2,\cdots,\alpha\lambda^{n-1})^T$$

is an eigenvector coresponding to the root λ.

Proof. Let $X = (x_1, \cdots, x_n)^T$. We can rewrite the equation $AX = \lambda X$ as follows

$$a_1x_2 = \lambda x_1, a_2x_3 = \lambda x_2, \cdots, a_{n-1}x_n = \lambda x_{n-1} \quad (a_1, \cdots, a_{n-1} \neq 0),$$

$$p_1x_1 + p_2x_2 + \cdots + p_nx_n = \lambda x_n.$$

If we put $x_1 = \alpha$, then the eigenvectors of A are

$$v_1 = (\alpha, \alpha\lambda, \alpha\lambda^2, \cdots, \alpha\lambda^{n-1})^T$$

where

$$a_1a_2 \cdots a_{n-1}p_1 + a_2 \cdots a_{n-1}p_2\lambda + \cdots + p_n\lambda^{n-1} = \lambda^n.$$

So,

$$p_1 + p_2\lambda + \cdots + p_n\lambda^{n-1} = \lambda^n,$$

where $v_1 = (\alpha, \alpha\lambda, , \alpha\lambda^2, \cdots, \alpha\lambda^{n-1})^T$. □

Theorem 4.6. *Eigenvectors of* $C_n[x, y, z]$ *are* V_n*, and also eigenvectors of* $\mathcal{P}_n[x, y, z]$ *are* $E_n[x, y, z] = R_n^{-1}[x, y, z]V_n$.

Proof. According to Lemma 4.5, columns of V_n are eigenvectors of $C_n[x, y, z]$. □

Theorem 4.7.

$$\Big(R_n^{-1}[x, y, z]V_n\Big)^{-1}\mathcal{P}_n[x, y, z]\Big(R_n^{-1}[x, y, z]V_n\Big) = \mathrm{diag}(z^n\alpha^n, z^n\alpha^{n-1}\beta, \cdots, z^n\alpha\beta^{n-1}, z^n\beta^n).$$

References

[1] R. Peele and P. Stănică, *Matrix powers of column justified Pascal triangles and Fibonacci sequences*, Fibonacci Qurts. **40**(2002), 146-152.

[2] P. Stănică, *Netted matrices*, J. Math. Sci. **39**(2003), 2507-2518.

[3] P. Stănică and R. Peele, *Spectral Properties of a Binomial Matrix*, unpublished manuscript, http://aexive.org/abs/math/0011111, 2000.

[4] G.S. Call and D.J. Velleman, *Pascal's matrices*, Amer. Math. Monthly **100**(1993), 372-376.

[5] M. Bayat and H. Teimoori, The linear algebra of the generalized Pascal functional matrix, *Linear Algebra and its Application*, **295**(1999), 81-89.

[6] M. Bayat, Generalized Pascal k-eliminted functional matrix with $2n$ variables, *Electronic Journal of Linear Algebra*, **22**(2010), 419-429.

[7] A. Mericier, *Identities containing Gauss binomial coefficients*, Discrate Math. **76**(1989), 67-73.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–195, pp. 448–454

Atom-bond connectivity and Geometric-arithmetic indices of generalized Sierpiński graph

Mahsa Khatibi[a], Ali Behtoei[a]

[a] *Department of Mathematics, Faculty of Science,, Imam Khomeini International University, Qazvin, Iran*

Article Info

Keywords:
Atom-bond connectivity index
geometric-arithmetic index
Sierpiński

2020 MSC:
05C05
05C07

Abstract

Sierpiński graph are a family of fractal features with many applications in topology, Tower of Hanoi mathematics, and computer science.Serpinski graphs are constructed by generalize (Duplication) the original graph and creating self-similar graphs.Some graph variables that are introduced as topological indices are used to determine some graph properties such as physical-chemical properties, thermodynamic properties, biological activity of chemical graphs. In this article, we investigate and calculate atom-bond connectivity index and geometric-arithmetic index for Sierpiński graphs like C_n ,P_n and Petersen.

1. Introduction

All graphs considered in this paper are simple, connected and finite. Let $G = (V(G), E(G))$ be a connected graph of order $n = |V(G)|$ and of size $m = |E(G)|$. The degree of a vertex $v \in V(G)$ is the number of vertices adjacent to v and is denoted by $\deg_G(v)$. The distance between two vertices $u, v \in V(G)$, denoted by $d_G(u, v)$, is the length of a shortest path between u and v in G. A topological index is a real number assigned to a graph which that is not depended on the labeling or pictorial representation of the graph.

Let k be an integer and G be a finite undirected graph on the vertex set $\{1, ..., k\}$. In the following, vertices of graphs will be identified with words on integers. We denote by $\{1, ..., k\}^t$ the set of words of size n on alphabet $\{1, ..., k\}$. The letters of a word u in $\{1, ..., k\}^t$ are denoted by $u = u_1u_2 \cdots u_t$. The Sierpiński graph $S(K_n, t)$ is a graph whose vertex set is $\{1, ..., k\}^t$ and uv is an edge in it if and only if there exists $i \in \{1, ..., t\}$ such that:[4]

(i) $u_j = v_j \quad if \quad j < i$;

(ii) $u_i \neq v_i$;

(iii) $u_j = v_i \quad and \quad v_j = u_i \;\; if \;\; j > i.$

For Sierpiński graphs $S(n, t)$, if $t = 1$, then $S(n, 1)$ is isomorphic to the complete graph K_n; if $n = 1$, then $S(1, t)$ is isomorphic to the complete graph K_1 consisting of one vertex and no edges; if $n = 2$, $S(2, t)$ is isomorphic to the path of length $2^t - 1$ and to the state graph of Chinese Rings[1].

Figure 1 shows the Sierpiński graph $S(K_4, 3)$. The generalized Sierpiński graph, $S(G, t)$, as the graph with vertex $\{1, 2, \ldots, n\}^t$ and edge set defined as follow, $\{u, v\}$ is an edge if and only if there exists $i \in \{1, , ..., t\}$ such that:

---

*Talker

Email addresses: Mahsa72kh@gmail.com (Mahsa Khatibi), a.behtoei@sci.ikiu.ac.ir (Ali Behtoei)

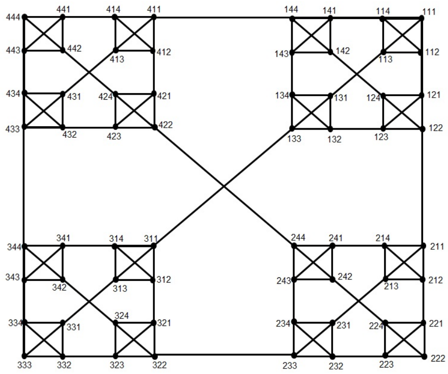


Fig. 1. generalized Sierpiński graph: $S(K_4, 3)$.

(i) $u_j = v_j \quad if \quad j < i$;
(ii) $u_i \neq v_i \quad and \quad \{u_i, v_i\} \in E$;
(iii) $u_j = v_i \quad and \quad v_j = u_i \ if \ j > i$.
Figure 2 shows the generalized Sierpiński graph $S(G, 2)$ for arbitrary graph G. note that $S(G, 1)$ is isomorphic to the

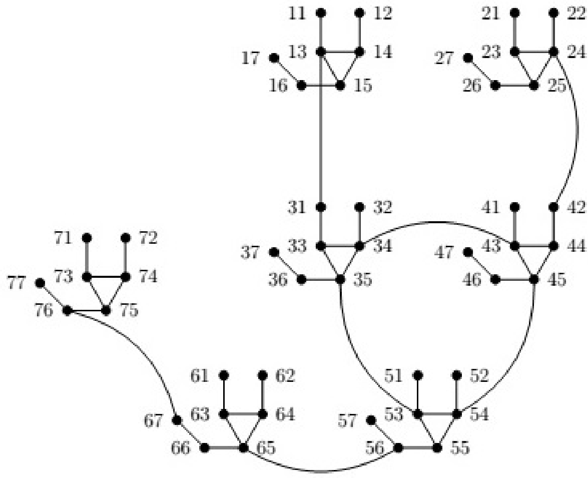


Fig. 2. generalized Sierpiński graph: $S(G, 2)$.

graph G and $S(G, 2)$ can be constructed by copying n times $S(G, 1)$ and adding an edge between the ith vertex of the jth copy and the jth vertex of the ith copy of $S(G, 1)$ whenever $\{i, j\}$ is an edge in G. In fact $S(G, t)$ is a fractal-like graph that uses G as a building block.
In mathematical chemistry and chemical graph theory, two important topological indices introduced about forty years ago by Ivan Gutman and Trinajstic [3] are the first Zagreb index $M_1(G)$ and second Zagreb index $M_2(G)$ which are defined as:

$$M_1(G) = \sum_{v \in V(G)} (\deg_G(v))^2 \qquad , \qquad M_2(G) = \sum_{uv \in E(G)} \deg(u)\deg(v)$$

Estrada et al. [2] introduced a well known topological index, called the atom-bond connectivity (ABC) index that is defined as:

$$ABC(G) = \sum_{uv \in E(G)} \sqrt{\frac{\deg(u) + \deg(v) - 2}{\deg(u)\deg(v)}} \tag{1}$$

The degree based topological index geometric-arithmetic (GA) index was introduced by vukicevic et al. in [7] and is defined as:

$$GA(G) = \sum_{uv \in E(G)} \frac{2\sqrt{\deg(u)\deg(v)}}{\deg(u) + \deg(v)} \tag{2}$$

2. Main results

Let $E_{i,j}$ be the set of edges in a graph whose end points are of degree i and j.

Remark 2.1. The generalized Sierpiński graph $S(C_n, t)$ has three kind of edges, i.e $E(S(C_n, t)) = E_{2,2} \cup E_{2,3} \cup E_{3,3}$ which their sizes are listed below.(see Figure 3 for instance)

$E_{i,j}$	$\mid E_{i,j} \mid$
$E_{2,2}$	$n^{t-1}(n-4)$
$E_{2,3}$	$\frac{4n}{n-1}(n^{t-1} - 2n^{t-2} + 1)$
$E_{3,3}$	$\frac{n}{n-1}(n^{t-1} + 4n^{t-2} - 5)$

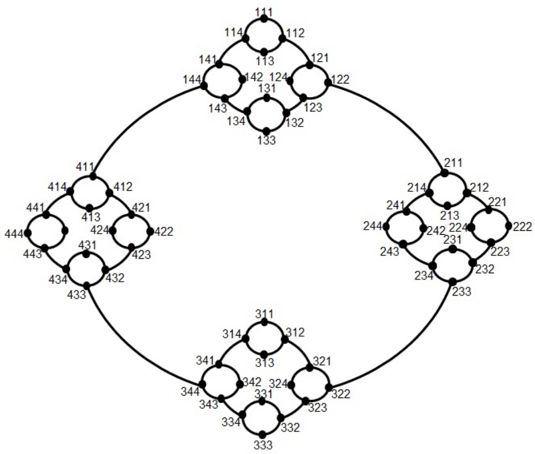


Fig. 3. Sierpiński graph: $S(C_4, 3)$.

Theorem 2.2. *Let C_n be a cycle with n vertices and $t \geqslant 4$. Then*

$$ABC(S(C_n, t)) = \frac{2n}{3(n-1)}(n^{t-1} + 4n^{t-2} - 5) + \frac{4\sqrt{2}n}{2(n-1)}(n^{t-1} - 2n^{t-2} + 1) + \frac{\sqrt{2}n^{t-1}}{2}(n-4).$$

Proof. By Using formula 1 and remark 2.1 we have

$$\begin{aligned}
ABC(S(C_n,t)) &= \sum_{uv\in E(S(C_n,t)} \sqrt{\frac{\deg(u)+\deg(v)-2}{\deg(u)\deg(v)}} \\
&= (\frac{n}{n-1}(n^{t-1}+4n^{t-2}-5))(\sqrt{\frac{3+3-2}{3\times 3}}) \\
&+ (\frac{4n}{n-1}(n^{t-1}-2n^{t-2}+1))(\sqrt{\frac{2+3-2}{2\times 3}}) \\
&+ n^{t-1}(n-4)(\sqrt{\frac{2+2-2}{2\times 2}}) \\
&= \frac{2n}{3(n-1)}(n^{t-1}+4n^{t-2}-5)+\frac{4\sqrt{2}n}{2(n-1)}(n^{t-1}-2n^{t-2}+1)+\frac{\sqrt{2}n^{t-1}}{2}(n-4)
\end{aligned}$$

□

Theorem 2.3. *geometric-arithmetic index of C_n for $t \geqslant 4$ is given by*

$$GA(S(C_n,t)) = \frac{n}{n-1}(n^{t-1}+4n^{t-2}-5)+\frac{8\sqrt{6}}{5(n-1)}(n^{t-1}-2n^{t-2}+1)+n^{t-1}(n-4).$$

Proof. from formula 2 and remark 2.1 we obtain.

$$\begin{aligned}
GA(S(C_n,t)) &= \sum_{uv\in E(S(C_n,t)} \frac{2\sqrt{\deg(u)\deg(v)}}{\deg(u)+\deg(v)} \\
&= (\frac{n}{n-1}(n^{t-1}+4n^{t-2}-5))(\frac{2\sqrt{3\times 3}}{3+3}) \\
&+ (\frac{4n}{n-1}(n^{t-1}-2n^{t-2}+1))(\frac{2\sqrt{2\times 3}}{2+3}) \\
&+ n^{t-1}(n-4)(\frac{2\sqrt{2\times 2}}{2+2}) \\
&= \frac{n}{n-1}(n^{t-1}+4n^{t-2}-5)+\frac{8\sqrt{6}}{5(n-1)}(n^{t-1}-2n^{t-2}+1)+n^{t-1}(n-4)
\end{aligned}$$

□

Remark 2.4. Let P be the petersen graph as shown in Figure 4. The generalized Sierpiński graph $S(P,t)$ has three kind of edges, i.e $E(S(P,t)) = E_{3,3} \cup E_{3,4} \cup E_{4,4}$ which their sizes are listed below.

$E_{i,j}$	$\mid E_{i,j} \mid$
$E_{3,3}$	$6\times 10^{t-1}$
$E_{3,4}$	$10^t - 2\times 10^{t-1}+10$
$E_{4,4}$	$\frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2}$

Theorem 2.5. *For Petersen graph and $t \geq 3$ we have*

$$ABC(S(Peterson,t)) = \sqrt{\frac{3}{8}}(\frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2})+\sqrt{\frac{5}{12}}(10^t-2\times 10^{t-1}+10)+4\times 10^{t-1}$$

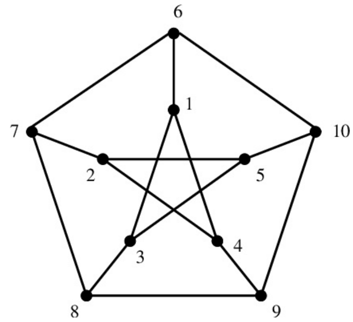


Fig. 4. Petersen graph.

Proof. By Using formula 1 and remark 2.4 we have:

$$\begin{aligned}
ABC(S(Peterson,t)) &= \sum_{uv\in E(S(C_n,t)}\sqrt{\frac{\deg(u)+\deg(v)-2}{\deg(u)\deg(v)}}\\
&= (\frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2})\sqrt{\frac{4+4-2}{4\times 4}}\\
&+ (10^{t}-2\times 10^{t-1}+10)\sqrt{\frac{4+3-2}{4\times 3}}\\
&+ (6\times 10^{t-1})\sqrt{\frac{3+3-2}{3\times 3}}\\
&= \sqrt{\frac{3}{8}}(\frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2})+\sqrt{\frac{5}{12}}(10^{t}-2\times 10^{t-1}+10)+4\times 10^{t-1}
\end{aligned}$$

□

Theorem 2.6. *geometric-arithmetric index of P for $t \geqslant 3$ is given by*

$$AG(S(Peterson,t)) = \frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2}+\frac{2\sqrt{12}}{7}(10^{t}-2\times 10^{t-1}+10)+6\times 10^{t-1}$$

Proof. From formula 2 and remark 2.4 we obtain

$$\begin{aligned}
AG(S(Peterson,t)) &= \sum_{uv\in E(S(C_n,t)}\frac{2\sqrt{\deg(u)\deg(v)}}{\deg(u)+\deg(v)}\\
&= (\frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2})\frac{2\sqrt{4\times 4}}{4+4}\\
&+ (10^{t}-2\times 10^{t-1}+10)\frac{2\sqrt{4\times 3}}{4+3}\\
&+ (6\times 10^{t-1})\frac{2\sqrt{3\times 3}}{3+3}\\
&= \frac{35}{3}(10^{t-2}-1)+15\times 10^{t-2}+\frac{2\sqrt{12}}{7}(10^{t}-2\times 10^{t-1}+10)+6\times 10^{t-1}
\end{aligned}$$

□

Remark 2.7. The generalized Sierpiński graph $S(P_n,t)$ has five kind of edges, i.e $E(S(P_n,t)) = E_{2,2}\cup E_{1,3}\cup E_{1,2}\cup E_{2,3}\cup E_{3,3}$ which their sizes are listed below.

$E_{i,j}$	$\mid E_{i,j} \mid$
$E_{2,2}$	$(n-4)(n-3)+2$
$E_{1,3}$	4
$E_{1,2}$	$2n-6$
$E_{2,3}$	$4n-10$
$E_{3,3}$	$n-3$

Theorem 2.8. *If P_n be a path with n vertices and $t=2$, Then*

$$ABC(S(P_n,2)) = \frac{\sqrt{2}}{2}(n^2-n-4)+\frac{2}{3}(n-3)+\frac{4\sqrt{6}}{3}+\sqrt{2}$$

Proof. by Using formula 1 and remark 2.7 we will have:

$$\begin{aligned} ABC(S(P_n,2)) &= \sum_{uv\in E(S(C_n,t))}\sqrt{\frac{\deg(u)+\deg(v)-2}{\deg(u)\deg(v)}} \\ &= ((n-4)(n-3)+2)\sqrt{\frac{2+2-2}{2\times 2}}+4\times\sqrt{\frac{1+3-2}{1\times 3}} \\ &+ (2n-6)\sqrt{\frac{1+2-2}{1\times 2}}+(4n-10)\sqrt{\frac{2+3-2}{2\times 3}}+(n-3)\sqrt{\frac{3+3-2}{3\times 3}} \\ &= \frac{\sqrt{2}}{2}(n^2-n-4)+\frac{2}{3}(n-3)+\frac{4\sqrt{6}}{3}+\sqrt{2} \end{aligned}$$

□

Theorem 2.9. *If P_n is a path with n vertices and $t=2$, Then*

$$GA(S(P_n,2)) = (n-3)^2+\frac{2\sqrt{2}}{3}(2n-6)+\frac{2\sqrt{6}}{5}(4n-10)+2(1+\sqrt{3})$$

Proof. From formula 2 and remark 2.7 we obtain:

$$\begin{aligned} GA(S(P_n,2)) &= \sum_{uv\in E(S(C_n,t))}\frac{2\sqrt{\deg(u)\deg(v)}}{\deg(u)+\deg(v)} \\ &= ((n-4)(n-3)+2)\frac{2\sqrt{2\times 2}}{2+2}+4\times\frac{2\sqrt{1\times 3}}{1+3} \\ &+ (2n-6)\frac{2\sqrt{1\times 2}}{1+2}+(4n-10)\frac{2\sqrt{2\times 3}}{2+3}+(n-3)\frac{2\sqrt{3\times 3}}{3+3} \\ &= (n-3)^2+\frac{2\sqrt{2}}{3}(2n-6)+\frac{2\sqrt{6}}{5}(4n-10)+2(1+\sqrt{3}) \end{aligned}$$

□

References

[1] S. Afriat, The Ring of Linked Rings, Duckworth, London, 1982.
[2] E. Estrada, L. Torres, L. Rodriguez, I. Gutman. An atom-bond connectivity index: modelling the enthalpy of formation of alkanes. Indian J Chem 1998;37A:849-55.
[3] I. Gutman, N. Trinajstic, *Graph theory and molecular orbitals. Total π-electron energy of alternant hydrocarbons*, Chem. Phys. Lett., 17 (1972) 535-538.

[4] S. Klavˇzar and U. Milutinovi´c, Graphs S(n, k) and a variant of the Tower of Hanoi problem, Czechoslovak Math. J. 47 (122) (1997) 95–104.
[5] S. Klavˇzar and U. Milutinovi´c and C. petr, 1-perfect codes in sierpinski graphs, Bull. Austral. Math. Soc., 66 (2002) 369-473.
[6] J. Liu, B. Ali, M. AslamMalik, H. Afzal Siddiqui, and M. Imran, *Reformulated Zagreb Indices of Some Derived Graphs*, MDPI, (2019).
[7] D. vukicevic, B. furtula. Topological index based on the ratios of geometrical and arithmetical means of end-vertex degrees of edges. J Math Chem 2009;46:1369-76.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–219, pp. 455–464

Remarks on the some parameters of domination in the middle neighborhood graph of a graph

Mohammad Hadi Akhbari[a,*], Fateme Movahedi[b]

[a] *Department of Mathematics, Estahban Branch, Islamic Azad University, Estahban, Iran*
[b] *Department of Mathematics, Faculty of Sciences, Golestan University, Gorgan, Iran*

Article Info

Keywords:
Middle neighborhood graph
Domination number
Independent domination
k-connected component domination

2020 MSC:
05C69
05C72

Abstract

For a simple graph $G = (V, E)$, the middle neighborhood graph $M_{nd}(G)$ of G is the graph with the vertex set $V \cup S$ where S is the set of all open neighborhood sets of G in which two vertices u and v are adjacent if $u, v \in S$ and $u \cap v \neq \emptyset$ or $u \in V$ and v is an open neighborhood set of G containing u. In this paper, we obtain the domination number, the independent domination number and the k-connected component domination number in the middle neighborhood graph.

1. Introduction

Let G be a graph with the vertex set $V(G)$ of size n and the edge set of size m. A fundamental concept in graph theory is domination which has been studied extensively [1]. The studies of domination set are important in the control of engineering systems [2]. Many studies are done on the parameters of domination of graphs [3–7].

A dominating set is a set D of vertices such that every vertex outside D is dominated by some vertex of D. The domination number of G, denoted by $\gamma(G)$, is the minimum size of the dominating set of G. A dominating set D is called an independent dominating set if D is an independent set. The independent domination number of G denoted by $\gamma_i(G)$ is the minimum size of an independent dominating set of G [1]. The subset D of the set of vertices $V(G)$ is a connected dominating set in G if D is a dominating set and the subgraph induced by D is connected. The minimum cardinality of any connected dominating set in G is called the connected domination number of G and it is denoted by $\gamma_c(G)$ [8].

Assume the graph G consists of k connected components. A subset D of $V(G)$ is a k-connected component dominating set of G if and only if $D = \bigcup_{i=1}^{k} D_i$ where D_i for $1 \leq i \leq k$ is a connected dominating set for every component of G and D has exactly k components. The k-connected component domination number of G is denoted by $\gamma_c^k(G)$ that is the minimum cardinality of a k-connected component dominating set of G. If G is connected, then it is clear that

*Talker
Email addresses: hadi.akhbari@iau.ac.ir (Mohammad Hadi Akhbari), f.movahedi@gu.ac.ir (Fateme Movahedi)

$\gamma_c^1(G) = \gamma_c(G)$. In the case that there is not a k-connected component dominating set of G, we can define $\gamma_c^k(G) = 0$. In this paper, we use $u \sim v$ and $G_1 \simeq G_2$ to denote two vertices u and v of G are adjacent and two graphs G_1 and G_2 are isomorphic, respectively.
Kulli in [9] introduced a neighborhood graph $N(G)$ of graph G and studied some properties of this graph. The neighborhood graph $N(G)$ of a graph G is the graph with the vertex set $V \cup S$ where S is the set of all open neighborhood sets of G and two vertices u and v in $N(G)$ are adjacent if $u \in V$ and v is an open neighborhood set containing u. Akhbari et. al in [10] determined the domination number, the total domination number and the independent domination number in the neighborhood graph. They also investigated these parameters of domination on the join and the corona of two neighborhood graphs. In [11], some parameters of domination number and the 2-domination polynomial of the edge neighborhood graph for certain graphs are determined.
Kulli introduced a new graph called by $M_{nd}(G)$ and obtained some results about it. The middle neighborhood graph $M_{nd}(G)$ of G is the graph with the vertex set $V \cup S$ where S is the set of all open neighborhood sets of G. Two vertices u and v of $M_{nd}(G)$ are adjacent if $u, v \in S$ and $u \cap v \neq \emptyset$ or $u \in V$ and v is an open neighborhood set of G containing u. In Figure 1, a graph G and its middle neighborhood graph $M_{nd}(G)$ are shown. The open neighborhood sets in the graph G are $N(1) = \{2, 3, 4\}$, $N(2) = \{1, 3\}$, $N(3) = \{1, 2\}$ and $N(4) = \{1\}$ [12].

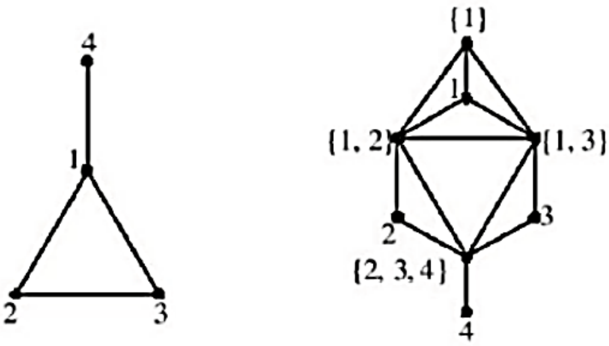


Fig. 1. The graph G and the middle neighborhood graph of G.

Kulli in [12] showed that for a graph G, the neighborhood graph $N(G)$ of G is a spanning subgraph of $M_{nd}(G)$. In this paper, we determine some properties of the middle neighborhood graph of a graph. We also obtain the domination number, independent domination number and k-connected component domination number for the middle neighborhood graph of a graph G.

2. General results for the middle neighborhood graph of some certain graphs

In this section, we recall some results that in our work. Also, we obtain some properties of the middle neighborhood graph on some certain graphs.

Lemma 2.1. [12] *For any graph G, the neighborhood graph $N(G)$ of G is a spanning subgraph of $M_{nd}(G)$.*

Lemma 2.2. [12] *For $n \geq 1$, $M_{nd}(G) = 2nK_2$ if and only if $G = nK_2$.*

Lemma 2.3. [12] *For $n \geq 2$, $M_{nd}(G) = G \cup K_n$ if and only if $G = K_{1,n}$.*

Lemma 2.4. [12] *If v is a vertex of a graph G, then the degree of the corresponding vertex of v in $M_{nd}(G)$ is the same with the degree v in G.*

Lemma 2.5. [13] *For any graph G of order n with no isolated vertex and the minimum degree δ,*

$$\gamma(G) \leq \frac{1}{2}(n + 2 - \delta).$$

Lemma 2.6. [14] *If G is an isolated-free graph of order n, then*

$$\gamma_i(G) \le n + 2 - 2\sqrt{n}.$$

Theorem 2.7. *For $n \ge 2$,*

(i) If n is even, then $M_{nd}(P_n) = 2G_P$ in which G_P is shown in Figure 2(a).

(ii) If n is odd, then $M_{nd}(P_n) = G_{1P} \cup G_{2P}$ where G_{iP}, for $i = 1, 2$, is shown in Figure 2(b).

Proof. Let P_n be a path of order n that the vertices are labeled by i, $1 \le i \le n$. Then by the definition of the middle neighborhood graph of P_n and its structure, we can consider the following cases.

(i) If n even, then $M_{nd}(P_n) = G_1 \cup G_2$ in which G_i is connected component of $M_{nd}(P_n)$ with n vertices. The vertex set and the edge set of G_i for $i = 1, 2$ is as follows,

$$V(G_1) = \{2i-1, N(2i) \,:\, 1 \le i \le \frac{n}{2}\},$$

$$\begin{aligned} E(G_1) = \{1 \sim N(2)\} &\cup \big\{2i-1 \sim N(2(i-1)),\, 2i-1 \sim N(2i) \,:\, 2 \le i \le \frac{n}{2}\big\} \\ &\cup \big\{N(2i) \sim N(2(i+1)) \,:\, 1 \le i \le \frac{n}{2} - 1\big\}, \end{aligned}$$

and

$$V(G_2) = \{N(2i-1), 2i \,:\, 1 \le i \le \frac{n}{2}\},$$

$$\begin{aligned} E(G_2) = \{n \sim N(n-1)\} &\cup \big\{2i \sim N(2i-1),\, 2i \sim N(2i+1) \,:\, 1 \le i \le \frac{n}{2} - 1\big\} \\ &\cup \big\{N(2i-1) \sim N(2i+1) \,:\, 1 \le i \le \frac{n}{2} - 1\big\}. \end{aligned}$$

Each of G_i's, for $i = 1, 2$ is a graph of order n and consists $\frac{n-2}{2}$ triangles and one end-vertex. So, it is clear to see that $G_i \simeq G_P$ for $i = 1, 2$ that is shown in Figure 2(a). Therefore, $M_{nd}(P_n) = 2G$ that $G \simeq G_P$.

(ii) If n is odd, then $M_{nd}(P_n)$ consists of two connected components G_i, $i = 1, 2$ of order n where their vertex sets and their edge set are as follows,

$$V(G_1) = \{2i-1, N(2i) \,:\, 1 \le i \le \frac{n}{2}\} \cup \{n\},$$

$$\begin{aligned} E(G_1) = \{1 \sim N(2),\, n \sim N(n-1)\} &\cup \big\{2i-1 \sim N(2(i-1)),\, 2i-1 \sim N(2i) \,:\, 2 \le i \le \frac{n-1}{2}\big\} \\ &\cup \big\{N(2i) \sim N(2(i+1)) \,:\, 1 \le i \le \frac{n-1}{2} - 1\big\}, \end{aligned}$$

and

$$V(G_2) = \{N(2i-1), 2i \,:\, 1 \le i \le \frac{n}{2}\} \cup \{N(n)\},$$

$$\begin{aligned} E(G_2) = &\big\{2i \sim N(2i-1),\, 2i \sim N(2i+1) \,:\, 1 \le i \le \frac{n}{2} - 1\big\} \\ &\cup \big\{N(2i-1) \sim N(2i+1) \,:\, 1 \le i \le \frac{n}{2} - 1\big\}. \end{aligned}$$

We can easily see that G_1 consists of $\frac{n-3}{2}$ triangles and two end-vertices and G_2 consists $\frac{n-1}{2}$ triangles. So, according to Figure 2(b) $G_1 \simeq G_{1P}$ and $G_2 \simeq G_{2P}$. Therefore, the proof is complete.

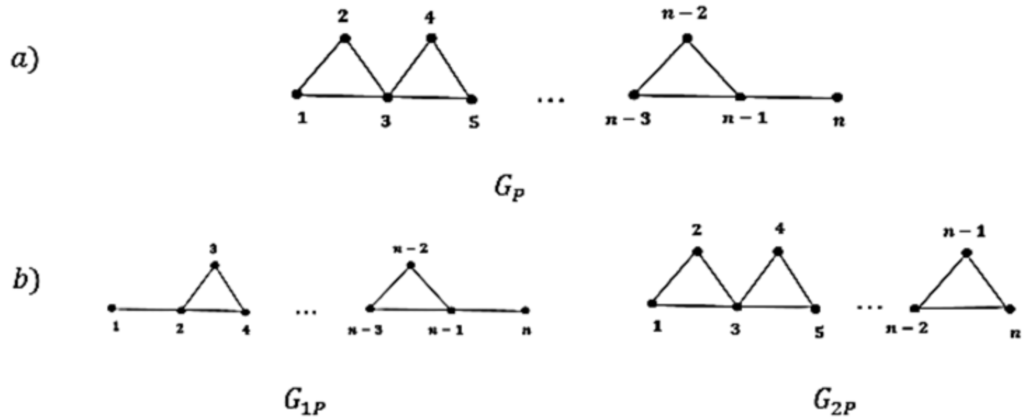


Fig. 2. The middle neighborhood graph of P_n. *a)* If n is even, then $M_{nd}(P_n) = 2G_P$. b) If n is odd, then $M_{nd}(P_n) \simeq G_{1P} \cup G_{2P}$.

□

Theorem 2.8. *For $n \geq 3$,*

(i) If n is even, then $M_{nd}(C_n) = 2G_C$ in which G_C is shown in Figure 3 for $k = n$.

(ii) If n is odd, then $M_{nd}(C_n) = G_C$ where G_C is shown in Figure 3 for $k = 2n$.

Proof. Let C_n be a cycle of order n that the vertices of C_n are labeled by i, $1 \leq i \leq n$. Using the structure of the middle neighborhood graph of C_n and its structure, we can consider the following cases.

(i) If n even, then $M_{nd}(C_n) = G_1 \cup G_2$ in which G_i is one connected component of $M_{nd}(C_n)$ of order n. The vertex set and the edge set of G_i for $i = 1, 2$ are as follows,

$$V(G_1) = \{1 + 2i \,:\, 0 \leq i \leq \frac{n}{2} - 1\} \cup \{N(2i) \,:\, 1 \leq i \leq \frac{n}{2}\},$$

$$E(G_1) = E(C_n) \cup \{N(2i) \sim N(2(i+1)) \,:\, 1 \leq i \leq \frac{n}{2}\} \cup \{N(2) \sim N(n)\},$$

also, we have

$$V(G_2) = \{N(2i+1), 2(i+1) \,:\, 0 \leq i \leq \frac{n}{2} - 1\},$$

$$E(G_2) = E(C_n) \cup \{N(2i-1) \sim N(2i+1) \,:\, 1 \leq i \leq \frac{n-1}{2}\} \cup \{N(1) \sim N(n)\}.$$

It is clear to see that $G_i \simeq G_C$ for $i = 1, 2$ that is shown in Figure 3 with $k = n$. Therefore, $M_{nd}(C_n) = 2G \simeq 2G_C$.

(ii) If n is odd, then $M_{nd}(C_n) = G'$ in which G' a graph of order $2n$. Let the vertices of G' are labeled by u_i for $1 \leq i \leq 2n$. The vertex set and the edge set of graph G' are as follows,

$$V(G') = \{u_i = 2i - 1\,,\, u_{2i} = N(2i) \,:\, 1 \leq i \leq \frac{n}{2}\} \cup \{u_n = n\}$$

$$\cup\{u_{2i} = N(2(i - \frac{n-1}{2}) - 1)\,,\, u_{2i+1} = 2(i - \frac{n-1}{2})\ :\ \frac{n+1}{2} \le i \le n\},$$
$$E(G') = E(C_{2n}) \cup \{u_{2i} \sim u_{2(i+1)}\ :\ 1 \le i \le n\} \cup \{u_2 \sim u_{2n}\}.$$

It is easy to see that $G' \simeq G_C$ in which G_C is shown in Figure 3 with $k = 2n$ vertices. Therefore, the proof is complete.

□

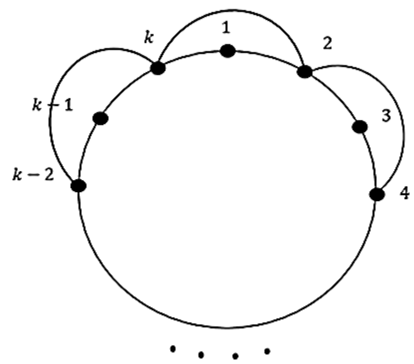


Fig. 3. The middle neighborhood graph of C_n. If n is even, then $M_{nd}(C_n) \simeq 2G_C$ where $k = n$. If n is odd, then $M_{nd}(C_n) \simeq G_C$ for $k = 2n$.

3. The results for the domination parameters of $M_{nd}(G)$

In this section, we propose the obtain results on some parameters of domination. We obtain the domination number and independent domination number of $M_{nd}(G)$ for a graph G. The results for k-connected component domination number of $M_{nd}(G)$ is proposed.

Theorem 3.1. *For $n \ge 1$,*
$$\gamma(M_{nd}(K_n)) = \gamma_i(M_{nd}(K_n)) = 2.$$

Proof. Using the definition of the middle neighborhood graph of K_n, we have $V(M_{nd}(K_n)) = V(K_n) \cup S$ where S is the open neighborhood sets of K_n. If the vertices of K_n are labeled by i, $1 \le i \le n$, then $M_{nd}(K_n)$ is a graph with $2n$ vertices with label u_i for every $1 \le i \le 2n$ that

$$V(M_{nd}(K_n)) = \{u_i = i\ :\ 1 \le i \le n\} \cup \{u_i = N(i - n)\ :\ n+1 \le i \le 2n\},$$

$$E(M_{nd}(K_n)) = \{u_i \sim u_j\ :\ 1 \le i \le n\,,\, n+1 \le j \le 2n\ and\ i \ne j - n\}$$
$$\cup\,\{u_i \sim u_j\ :\ n+1 \le i \le 2n\ and\ j = i+1, i+2, \cdots, 2n\}. \quad (1)$$

Using the structure of $M_{nd}(K_n)$, the vertex $u_1 = 1$ is adjacent to $n-1$ vertices u_j for $n+2 \le j \le 2n$. Also, $u_{n+1} = N(1)$ is adjacent to $n-1$ vertices u_i for $2 \le i \le n$. It is shown that the set $D = \{u_1, u_{n+1}\}$ dominates all of vertices of $M_{nd}(K_n)$. So, $\gamma(M_{nd}(K_n)) = 2$.
On the other hand, by (1), u_1 and u_{n+1} are not adjacent in $M_{nd}(K_n)$. So, D is an independent set in this graph. Therefore, $\gamma_i(M_{nd}(K_n)) = 2$. □

Theorem 3.2. *For $1 \le m \le n$,*
$$\gamma(M_{nd}(K_{m,n})) = \gamma_i(M_{nd}(K_{m,n})) = 2.$$

Proof. Let $V(K_{m,n}) = A \cup B$ where $|A| = m$ and $|B| = n$. Using the definition of $M_{nd}(K_{m,n})$, we can obtain $M_{nd}(K_{m,n}) = G_1 \cup G_2$. If $A = \{u_i \ : \ 1 \leq i \leq m\}$ and $B = \{v_j \ : \ 1 \leq j \leq n\}$, then the vertices of $M_{nd}(K_{m.n})$ is the set $A \cup B \cup \{N(u_i) \ : \ 1 \leq i \leq m\} \cup \{N(v_j) \ : \ 1 \leq j \leq n\}$. Each vertex $N(u_i)$ is adjacent to the vertices v_j for $1 \leq j \leq n$ and each of vertices $N(v_j)$ is adjacent to the vertices u_i for $1 \leq i \leq m$. Also, $N(u_i)$ and $N(u_j)$ are adjacent for $1 \leq i \leq m$ and $j = i+1, i+2, \cdots, 2m$. Similarly, for every $1 \leq i \leq n$ and $j = i+1, i+2, \cdots, 2n$, two vertices $N(v_i)$ and $N(v_j)$ are adjacent. Thus, $M_{nd}(K_{m,n}) = G_1 \cup G_2$ such that $V(G_1) = A \cup N(B)$ and $V(G_2) = B \cup N(A)$ and $K_{m,n}$ is a spanning subgraph of G_1 and G_2.
We consider $D_1 = \{N(u_1)\}$ and $D_2 = \{N(v_1)\}$ as the dominating sets for graph G_1 and graph G_2, respectively. Therefore, $D = D_1 \cup D_2$ is a dominating set of $M_{nd}(K_{m,n})$. Since D is an independent set,

$$\gamma(M_{nd}(K_{m.n})) = \gamma_i(M_{nd}(K_{m.n})) = 2.$$

□

Theorem 3.3. *for* $n \geq 1$,

(i) $\gamma(M_{nd}(K_{1,n})) = \gamma_i(M_{nd}(K_{1,n})) = 2$.

(i) $\gamma(M_{nd}(nK_2)) = \gamma_i(M_{nd}(nK_2)) = 2n$.

Proof. (i) Using Lemma 2.3 and Theorem 3.2, the result holds.

(ii) Using Lemma 2.2, $M_{nd}(nK_2) = 2nK_2$. Therefore,

$$\gamma(M_{nd}(nK_2)) = 2n\gamma(K_2).$$

Since $\gamma(K_2) = 1$, the result completes. Also, the dominating set consists of $2n$ vertices of $M_{nd}(nK_2)$ where these are not adjacent together. This completes that $\gamma_i(M_{nd}(nK_n)) = 2n$.

□

Theorem 3.4. *For any graph* P_n *with* $n \geq 4$,

$$\gamma(M_{nd}(P_n)) = \gamma_i(M_{nd}(P_n)) = \begin{cases} \frac{n}{2} & if\, n \equiv 0 \ (mod\, 4), \\ \frac{n}{2}+1 & if\, n \equiv 2 \ (mod\, 4), \\ \frac{n+1}{2} & if\, n \equiv 1,3 \ (mod\, 4). \end{cases}$$

Proof. Using Theorem 2.7, we consider the following cases.
Case 1: In n is even, then $M_{nd}(P_n) = 2G$ where G is isomorphism with the graph G_P in Figure 2(a). So, it is sufficient to obtain a dominating set for graph G_P.
Subcase 1: If $n \equiv 0$ (mod 4), then the set $S_1 = \bigcup_{i=0}^{\frac{n}{4}-1}\{3+4i\}$ is a dominating set for graph G_P. Thus, $\gamma(G_P) \leq |S| = \frac{n}{4}$.
Let D be a dominating set of G_P. We must show that $|D| \geq \frac{n}{4}$. Otherwise, assume $|D| \leq \frac{n}{4} - 1$. Since vertex n is an end-vertex, then $\{n-1\} \subseteq D$. The vertex $n-1$ is adjacent to $\{n-3, n-2, n\}$. So, $n-4$ remained vertices of G_P are dominated by $D \setminus \{n-1\}$. If each vertex of $D \setminus \{n-1\}$ dominates at most four vertices, then $D \setminus \{n-1\}$ dominates at most $4(\frac{n}{4}-2)$ vertices of G_P. It is a contradiction. Because there are at least four vertices of G_P any vertices of D can not dominate them. So, $\gamma(G_P) \geq \frac{n}{4}$.
Since $\gamma(M_{nd}(P_n)) = 2\gamma(G_P)$ we get, $\gamma(M_{nd}(P_n)) = \frac{n}{2}$. Also, the set S_1 is the independent dominating set of $M_{nd}(P_n)$. So, using $\gamma(G) \leq \gamma_i(G)$ for every graph G, we have

$$\frac{n}{4} \leq \gamma_i(G_P) \leq |S_1| = \frac{n}{4}.$$

Therefore, $\gamma_i(M_{nd}(P_n)) = 2\gamma_i(G_P) = \frac{n}{2}$.
Subcase 2: If $n \equiv 2$ (mod 4), we consider the set $S_2 = \bigcup_{i=0}^{\frac{n+2}{4}}\{1+4i\}$ as a dominating set of graph G_P. So,

$\gamma(G_P) \leq |S| = \frac{n+2}{4}$.
Similarly, we can obtain $\gamma(G_P) = \frac{n+2}{4}$. On the other hand, the set S_2 is the independent set in G_P. So,

$$\gamma(M_{nd}(P_n)) = \gamma_i(M_{nd}(P_n)) = 2\gamma_i(G_P) = 2\left(\frac{n+2}{4}\right) = \frac{n+2}{2},$$

and therefore, $\gamma_i(M_{nd}(P_n)) = \frac{n}{2} + 1$.
Case 2: In n is odd, using Theorem 2.7(ii) $M_{nd}(P_n) \simeq G_{1P} \cup G_{2P}$ such that G_{iP} is shown in Figure 2(b) for $i = 1, 2$. So, we obtain the dominating set of G_{iP} for $i = 1, 2$.
Subcase 1: If $n \equiv 1$ (mod 4), $S_1' = \bigcup_{i=0}^{\frac{n-5}{4}} \{2+4i\} \cup \{n-1\}$ is a dominating set of G_{1P}. So, $\gamma(G_{1P}) \leq |S_1'| = \frac{n+3}{4}$. Let D is a dominating set of G_{1P} and $|D| \leq \frac{n-1}{4}$. Since vertices 1 and n are the end-vertices, $\{1, n\} \subseteq D$. Thus, $D \setminus \{1, n\}$ must dominate $n-6$ remained vertices of G_{1P}. If every vertex of $D \setminus \{1, n\}$ dominates at most four vertices of G_{1P}, then there are at least three vertices of G_{1P} that D can not dominate it. It is a contradiction. So, $\gamma(G_{1P}) = |D| = \frac{n+3}{4}$.
For graph G_{2P}, we consider $S_2' = \bigcup_{i=0}^{\frac{n-5}{4}} \{3+4i\}$ as a dominating set. We obtain $\gamma(G_{2P}) = \frac{n-1}{4}$. Therefore,

$$\gamma(M_{nd}(P_n)) = \gamma(G_{1P}) + \gamma(G_{2P}) = \frac{n+1}{2}.$$

It is clear to see that S_1' and S_2' are the independent sets in graphs G_{1P} and G_{2P}, respectively. So, using $\gamma(G) \leq \gamma_i(G)$, the result holds.
Subcase 2: If $n \equiv 3$ (mod 4), then for graph G_{1P} we consider the set $S_1'' = \bigcup_{i=0}^{\frac{n-3}{4}} \{2+4i\}$ as the dominating set. So, $\gamma(G_{1P}) \leq \frac{n+1}{4}$.
Let D be a dominating set of G_{1P} that $|D| \leq \frac{n-3}{4}$. According to Theorem 2.7(ii) and Figure 2(b), $\{1, n\} \subseteq D$. So, $D \setminus \{1, n\}$ must dominate $n-6$ remained vertices of G_{1P}. If any vertex of $D \setminus \{1, n\}$ dominates at most four vertices of G_{1P}, then there are at least five vertices that D can not dominate them. It is a contradiction and so, $\gamma(G_{1P}) = \frac{n+1}{4}$.
Similarly, for graph G_{2P}, $S_2'' = \bigcup_{i=0}^{\frac{n-3}{4}} \{3+4i\}$ is the dominating set. Therefore, $\gamma(G_{2P}) = \frac{n+1}{4}$. Since S_1'' in the graph G_{1P} and S_2'' in the graph G_{2P} are the independent sets, then

$$\gamma_i(M_{nd}(P_n)) = \gamma_i(G_{1P}) + \gamma_i(G_{2P}) = \frac{n+1}{2}.$$

□

Theorem 3.5. *For any $n \geq 3$,*

$$\gamma(M_{nd}(C_n)) = \gamma_i(M_{nd}(C_n)) = \begin{cases} \frac{n}{2} & \text{if } n \equiv 0 \ (mod\, 4), \\ \frac{n}{2}+1 & \text{if } n \equiv 2 \ (mod\, 4), \\ \frac{n+1}{2} & \text{if } n \equiv 1, 3 \ (mod\, 4). \end{cases}$$

Proof. Using Theorem 2.8 we have
Case 1: If n is odd, then $M_{nd}(C_n) \simeq G_C$ where G_C is a graph of order $2n$ (see Figure 3). Any vertex of G_C with labeling $2i$ for $1 \leq i \leq n$ is adjacent to four vertices $\{2i-2, 2i-1, 2i+1, 2i+2\}$. The set $S = \bigcup_{i=0}^{\frac{n-3}{2}} \{2+4i\} \cup \{2n-1\}$ is a dominating set of G_C. So,

$$\gamma(M_{nd}(C_n)) = \gamma(G_C) \leq |S| = \frac{n+1}{2}.$$

Let D be the dominating set of G_C with $|D| \leq \frac{n-1}{2}$. Without loss of generality suppose $\{2\} \subseteq D$. So, the vertex 2, dominates the vertices $\{2n, 1, 3, 4\}$. So, $D \setminus \{2\}$ must dominate $2n-5$ remained vertices of G_C. If any vertex of $D \setminus \{2\}$ dominate at most four vertices of G_C, then $2(n-3)$ vertices of the graph are dominated by $D \setminus \{2\}$. It is a contradiction. Because at least one vertex of G_C is not dominated by D. So, $\gamma(M_{nd}(C_n)) = \gamma(G_C) \geq \frac{n+1}{2}$. On the other hand, S is the independent set of G_C. Thus,

$$\frac{n+1}{2} = \gamma(M_{nd}(C_n)) \leq \gamma_i(M_{nd}(C_n)) \leq \frac{n+1}{2}.$$

Case 2: Let n is even. According to Theorem 2.8 and Figure 3, $M_{nd}(C_n) \simeq 2G_C$ for $k = n$. So, we can consider the following cases.

Subcase 1: If $n \equiv 0 \pmod 4$, then $S' = \bigcup_{i=0}^{\frac{n}{4}-1}\{2+4i\}$ is a dominating set of G_C. So, $\gamma(G_C) \leq |S| = \frac{n}{4}$.
Let D be a dominating set of G_C with $|D| \leq \frac{n}{4} - 1$. Without loss of generality suppose that $\{2\} \subseteq D$. So, $n - 5$ remained vertices are dominated by $D \setminus \{2\}$. But, $D \setminus \{2\}$ dominates at most $n - 8$ vertices of G_C. It is a contradiction. So, $\gamma(G_C) = \frac{n}{4}$. Therefore, $\gamma(N_{nd}(C_n)) = 2\gamma(G_C) = \frac{n}{2}$. Since S' is an independent set, in this case, $\gamma_i(M_{nd}(C_n)) = \frac{n}{2}$.

Subcase 2: Assume $n \equiv 2 \pmod 4$, then $S'' = \bigcup_{i=0}^{\frac{n-6}{4}}\{2+4i\} \cup \{2n-1\}$. It is clear to see that S'' is a dominating set of G_C. So, $\gamma(G_C) \leq |S| = \frac{n+2}{4}$.
Similarly, in this case, S'' is an independent set. So, the result holds. □

Theorem 3.6. *Let G be a graph of order n without isolated vertices and with the minimum degree δ. Then*

$$\gamma(M_{nd}(G)) \leq n + 1 - \frac{\delta}{2}.$$

Proof. For any graph G of order n, $M_{nd}(G)$ is a graph with $2n$ vertices. Using Lemma 2.1 and Lemma 2.4, we can obtain that the minimum degree of $M_{nd}(G)$ is equal with the minimum degree of G. So, $\delta_{M_{nd}(G)} = \delta_G = \delta$.
According to Lemma 2.5, for graph $M_{nd}(G)$ of order $2n$,

$$\gamma(M_{nd}(G)) \leq \frac{1}{2}(2n + 2 - \delta) = n + 1 - \frac{\delta}{2}.$$

□

Theorem 3.7. *Let G be a graph without isolated vertices. Then*

$$\gamma_i(M_{nd}(G)) \leq 2(n + 1 - \sqrt{2n}).$$

Proof. Using Lemma 2.4, since G doesn't have any isolated vertex, $M_{nd}(G)$ of G is a graph without isolated vertex. Therefore, using Lemma 2.6, for graph $M_{nd}(G)$ with $2n$ vertices, we have

$$\gamma_i(M_{nd}(G)) \leq 2n + 2 - 2\sqrt{2n}.$$

□

Theorem 3.8. *For $n \geq 3$,*

$$\gamma_c^2(M_{nd}(P_n)) = n - 2.$$

Proof. **Case 1**: If n is even, using Theorem 2.7 and Figure 2(a), $M_{nd}(P_n) \simeq 2G_P$. So, graph $M_{nd}(P_n)$ consists of two connected components G_P. It is sufficient to obtain a connected dominating set of G_P. We easily see that $S_1 = \bigcup_{i=1}^{\frac{n}{2}-1}\{1+2i\}$ is a connected dominating set of G_P. So, $\gamma_c(G_P) \leq |S_1| = \frac{n}{2} - 1$.
Let D be a connected dominating set of G_P and $|D| \leq \frac{n}{2} - 2$. According to Figure 2(a), $\{n-1\} \subseteq D$. So, $D \setminus \{n-1\}$ must dominate $n - 4$ remained vertices of G_P. Since D is a connected set, every vertex of D dominates at most two vertices of G_P. So, $n - 6$ vertices are dominated by $D \setminus \{n - 1\}$. It is a contradiction. Because there are at least two vertices of G_P that can not dominate by D. So, $\gamma_c(G_P) \geq \frac{n}{2} - 1$. Since $M_{nd}(P_n)$ has two connected dominating sets,

$$\gamma_c^2(M_{nd}(P_n)) = 2\gamma_c(P_n) = n - 2.$$

Case 2: If n is odd, by Theorem 2.7 and Figure 2(b), $M_{nd}(P_n) \simeq G_{1P} \cup G_{2P}$. Similarly, we can consider $S_1 = \bigcup_{i=1}^{\frac{n-1}{2}}\{2i\}$ as a connected dominating set of G_{1P} and $S_2 = \bigcup_{i=1}^{\frac{n-3}{2}}\{1+2i\}$ as a dominating set of G_{2P}. So, $\gamma_c(G_{1P}) = \frac{n-1}{2}$ and $\gamma_c(G_{2P}) = \frac{n-3}{2}$. Therefore,

$$\gamma_c^2(M_{nd}(P_n)) = \gamma_c(G_{1P}) + \gamma_c(G_{2P}) = n - 2.$$

□

Theorem 3.9. *For $n \geq 3$,*

(i) If n is odd, $\gamma_c(M_{nd}(C_n)) = n$.

(ii) If n is even, $\gamma_c^2(M_{nd}(C_n)) = n$.

Proof. (i) If n is odd, then according to Theorem 2.8 $M_{nd}(C_n) \simeq G_C$ where G_C is the graph in Figure 3 for $k = 2n$. So, $S = \bigcup_{i=1}^{n}\{2i\}$ is a connected dominating set of G_C. Thus, $\gamma_c(G_C) \leq n$. Let D be a connected dominating set of G_C and $|D| \leq n-1$. Every vertex of D dominates at most two vertices of G_C. So, $2n-2$ vertices of G_C are dominated by D. It is a contradiction. Because there are at least two vertices of G that are not dominated by D. So, $\gamma_c(G_C) \geq n$. Therefore,

$$\gamma_c(M_{nd}(C_n)) = \gamma_c(G_C) = n.$$

(ii) If n is odd, then using Theorem 2.8 $M_{nd}(C_n) \simeq 2G_C$ where G_C is a graph with n vertices as is shown in Figure 3. It is easy to show that the set $S = \bigcup_{i=1}^{\frac{n}{2}}\{2i\}$ is a connected dominating set of G_C. In a similar way of (i), we can obtain $\gamma_c(G_C) = \frac{n}{2}$. Therefore,

$$\gamma_c^2(M_{nd}(C_n)) = 2\gamma_c(M_{nd}(G_C)) = n.$$

□

Theorem 3.10. *For $n \geq 1$,*

$$\gamma_c(M_{nd}(K_n)) = 3.$$

Proof. Let D be a connected dominating set of $M_{nd}(K_n)$. Since $M_{nd}(K_n)$ is a connected graph, D is one component and we have $3 \leq |D| = \gamma_c(M_{nd}(K_n))$. According to label the vertices of Theorem 3.1, we consider $S = \{1, N(1), N(2)\}$. It is clear that S is a connected dominating set of $M_{nd}(K_n)$. Therefore, the proof is complete. □

Theorem 3.11. *For $1 \leq m \leq n$,*

(i) $\gamma_c^2(M_{nd}(K_{m,n})) = 2$.

(ii) $\gamma_c^2(M_{nd}(K_{1,n})) = n$.

(iii) $\gamma_c^{2n}(M_{nd}(nK_2)) = 2n$.

Proof. (i) Using Theorem 3.2 and since $M_{nd}(K_{m,n}) = G_1 \cup G_2$, $M_{nd}(K_{m,n})$ consists of two connected components. Thus, $M_{nd}(K_{m,n})$ has 2-connected component dominating set. So, every G_i has a connected dominating set. Using Theorem 3.2(i), the proof is complete.

(ii) This is the result of the segment (i).

(iii) Using Lemma 2.2, the graph $M_{nd}(nK_2)$ consists of $2n$ connected components. So, we consider a connected dominating set for every component. Thus, $M_{nd}(nK_2)$ has $2n$-connected component dominating set. Using Theorem 3.3(ii) the result holds.

□

References

[1] T. W. Haynes, S. T. Hedetniemi and P. J. Slater, Fundamentals of Domination in Graphs, Marcel Dekker Inc., New York, (1998).
[2] I. Stojmenovic, M. Seddigh, and J. Zunic, Dominating sets and neighbour elimination based broadcasting algorithms in wireless networks, IEEE Transactions on Parallel and Distributed Systems, 13 (2002), 14-25.
[3] K. Kalyan, M. James, Domination Number of Sierpinski Cycle Graph of Order n, Journal of Computer and Mathematical Sciences, 10(4) (2019), 796-818.
[4] H. Tan, L. Liu, TOTAL k-domination in special graphs, Mathematical Foundations of Computing, 1(3) (2018), 255-263.
[5] S. M. Hosamani, Correlation of domination parameters with physicochemical properties of octane isomers, Applied Mathematics and Non-linear Sciences 1(2) (2016), 345-352.
[6] Y. Zhu, C. Zhao, Connected Domination Partition of Regular Graphs, Applied Mathematical Sciences, 13(6) (2019), 275-281.
[7] R. E. Leonida, R. C. Allosa, Independent 2-Domination in Graphs, Applied Mathematical Sciences, 12(12) (2018), 581-585.
[8] E. Sampathkumar, H. Walikar, The connected domination number of a graph, Journal of Mathematical and Physical Sciences, 13 (1979), 607-613.
[9] V.R. Kulli, The neighborhood graph of a graph, International Journal of Fuzzy Mathematical Archive, 8 (2015), 93-99.
[10] M. H. Akhbari, F. Movahedi, V. Kulli, Some Parameters of Domination on the Neighborhood Graph, International Journal of Mathematical Combinatorics, 4 (2017), 138-150.
[11] F. Movahedi, M. H. Akhbari, A note on some parameters of domination on the edge neighborhood graph of a graph, Caspian Journal of Mathematical Sciences, 11(1) (2022), 40-51.
[12] V.R. Kulli, On Middle Neighborhood Graphs, International Journal of Mathematics And its Applications, 3 (2015), 79-83.
[13] A. Poghosyan, The probabilistic method for upper bounds in dominion theory, thesis of Ph. D, Faculty of Environment and Technology University of the West of England, Bristol, (2010).
[14] A. Gagarin, A. Poghosyan and V. Zverovich, Upper bounds for the α-domination parameters. Graphs and Combinatorics, 25 (2009), 513-520.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–220, pp. 465–470

General sum-connectivity, general Randić and arithmetic-geometric indices of generalized Sierpiński gasket graph

Fatemeh Attarzadeh[a,*], Ali Behtoei[b], Ahmad Abbasi[a]

[a] *Department of Pure Mathematics, Faculty of Mathematical Sciences, University of Guilan, Rasht, Iran,*
[b] *Department of Mathematics, Faculty of Science, Imam Khomeini International University, Qazvin, Iran.*

Article Info

Keywords:
Generalized Sierpiński gasket graph
general sum-connectivity index
general Randić index
arithmetic-geometric index

2020 MSC:
05C05
05C07

Abstract

Sierpiński gasket graphs have many applications in diverse areas like topology, chemistry and dynamical systems. In this paper, we determine some topological indices for the generalized Sierpiński gasket of cycles and complete graphs.

1. Introduction

All graphs considered in this paper are assumed to be simple connected. Throughout $G = (V, E)$ is assumed to be a non-empty graph of order n with the vertex set $V = \{1, 2, \ldots, n\}$ and the edge set E. The degree of a vertex v of G is denoted by $\deg_G(v)$ which is the size of the set of its neighbourhood $N_G(v)$. In mathematical chemistry and chemical graph theory, topological indices are numerical parameter of a graph that are invariant under graph isomorphism. One of the first degree-based indices is Randić index defined in [7] as follows

$$R(G) = \sum_{uv\in E(G)} [\deg(u)\deg(v)]^{-\frac{1}{2}}.$$

The sum-connectivity index of a graph G defined in [11] as follows

$$\chi(G) = \sum_{uv\in E(G)} [\deg(u) + \deg(v)]^{-\frac{1}{2}}.$$

*Talker

Email addresses: prs.attarzadeh@gmail.com (Fatemeh Attarzadeh), a.behtoei@sci.ikiu.ac.ir (Ali Behtoei), aabbasi@guilan.ac.ir (Ahmad Abbasi)

It is shown in [3] that the Randić and sum-connectivity indices are highly interdependent quantities. The Randić index has been extended and generalized as [1]

$$R_\alpha(G) = \sum_{uv \in E(G)} [\deg(u)\deg(v)]^\alpha,$$

in which α is real a number. Similarly, the general sum-connectivity index is defined in [10] as follows

$$\chi_\alpha(G) = \sum_{uv \in E(G)} [\deg(u) + \deg(v)]^\alpha.$$

Also, in [9] Arithmetic-Geometric index is defined as follows

$$AG_1(G) = \sum_{uv \in E(G)} \frac{\deg(u) + \deg(v)}{2\sqrt{\deg(u) \times \deg(v)}}.$$

2. Some topological indices of generalized Sierpiński gasket graph

Decomposition into special substructures that inherit remarkable features is an important method used for the investigation of some mathematical structures, specifically when the regarded structures have self-similarity features. Klavžar et al. for the first time, introduced the Sierpiński graph $S(K_n, t)$, see [5] and [6]. One of the most important families of these self similar graphs is the family of Sierpiński gasket graphs, see [8]. The Sierpiński gasket graph S_t is made from the Sierpiński graph $S(3, t)$ by contracting all the edges that are in no triangle, see [4] for more details. In this paper, we generalized this structure to each graph G, and call it the generalized Sierpiński gasket graph and denote it by $S[G, t]$.

Definition 2.1. [2] Let $G = (V, E)$ be a graph of order $n \geqslant 2$, t be a positive integer and denote the set of words of length t on the alphabet V by V^t. The letters of a word $u \in V(G)$ (of length t) are denoted by $u_1u_2 \dots u_t$. The generalized Sierpiński graph of G of dimension t, denoted by $S(G, t)$, is the graph with vertex set V^t and $\{u, v\}$ is an edge if and only if there exists $i \in \{1, , ..., t\}$ such that:
(i) $u_j = v_j \quad if \quad j < i,$
(ii) $u_i \neq v_i \quad and \quad \{u_i, v_i\} \in E,$
(iii) $u_j = v_i \quad and \quad v_j = u_i \;\; if \;\; j > i.$

Definition 2.2. Let G be a simple graph of order $n \geq 2$, with the vertex set $V = \{1, 2, \dots n\}$ and t be a positive integer. If i is adjacent to j in G, then by contracting the new edge between two copies i and j (the linking edge) in the generalized Sierpiński graph, the generalized Sierpiński gasket graph is obtained. In other words, if i is adjacent to j in G, then the vertex $\mathbf{u} = v_1v_2 \dots v_r ji \dots i$ is adjacent to $\mathbf{v} = v_1v_2 \dots v_r ij \dots j$, in $S(G, t)$, $0 \leq r \leq t-2$, the edge $\mathbf{uv}$ will be contracted in $S[G, t]$, and this new vertex will be denoted by $v_1v_2 \dots v_r\{i, j\}_{t-r}$ or shortly by $v_{(r)}\{i, j\}_{t-r}$, see Figure 1 and Figure 2.

Similar to the structure of the generalized Sierpiński graph $S(G, t)$, $S[G, t]$ is constructed inductively by inserting a copy of $S[G, t-1]$ instead of each vertex of G ($S_i[G, t]$ for $i \in V(G)$) and then by contracting the new $|E(G)|$ linking edges (of $S(G, t)$). More precisely, when i is adjacent to j in the graph G, then the linking edge between $ijj \dots j$ and $jii \dots i$ is contracted and the new vertex is shown by $\{i, j\}_t$ in $S[G, t]$. Note that the vertex $\{i, j\}_t$ is the unique common shared vertex between two copies $S_i[G, t]$ and $S_j[G, t]$.
In what follows, we determine some of the topological indices of $S[G, t]$ for special graphs G.

Theorem 2.3. *The general sum-connectivity index of generalized Sierpiński gasket graph of K_n in step t is given by*

$$\chi_\alpha(S[K_n, t]) = (n^2 - n)(3n - 3)^\alpha + (mn^{t-1} - n^2 + n)(4n - 4)^\alpha.$$

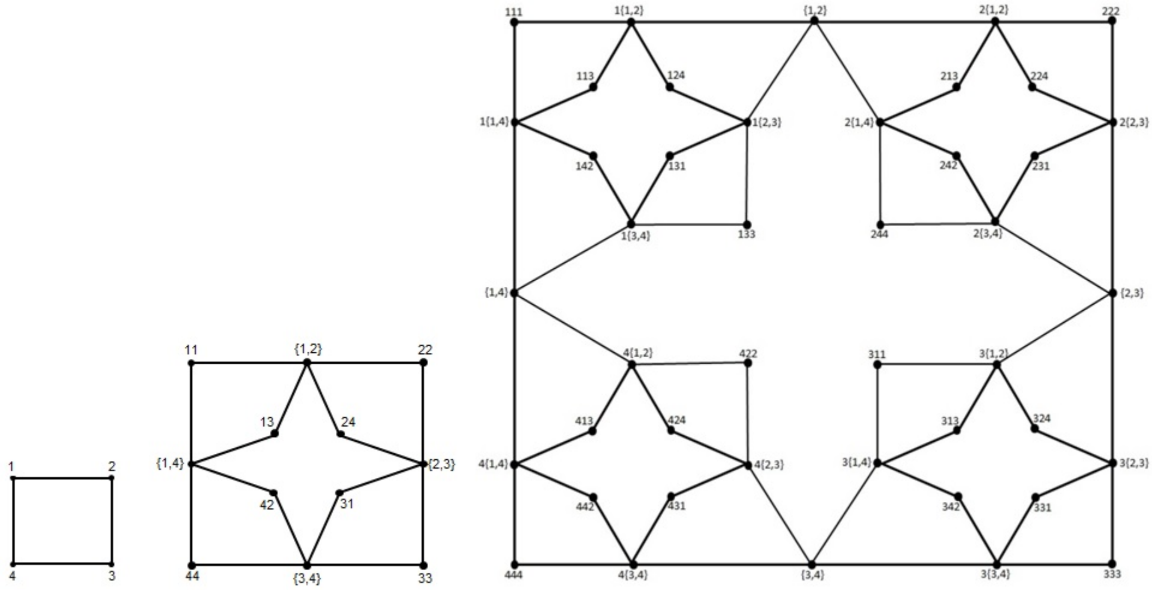


Fig. 1. C_4, $S[C_4, 2]$ and $S[C_4, 3]$.

Proof. Since $S[K_n, t]$ has $n(n-1)$ edges with endpoint degrees $(n-1)$ and $2(n-1)$, and the other edges have endpoints of degree $2(n-1)$, we get

$$\begin{aligned}
\chi_\alpha(S[K_n,t]) &= \sum_{uv\in E(S[K_n,t])} [\deg(u)+\deg(v)]^\alpha \\
&= \sum_{uv\in E(S[K_n,t])} ((n-1)+(2n-2))^\alpha + \sum_{uv\in E(S[K_n,t])} ((2n-2)+(2n-2))^\alpha \\
&= (n^2-n)(3n-3)^\alpha + (mn^{t-1}-n^2+n)(4n-4)^\alpha.
\end{aligned}$$

□

Corollary 2.4. *If in Theorem 2.3, we let $\alpha = -\frac{1}{2}$, then sum-connectivity index of $S[K_n, t]$ is equal to*

$$\chi(S[K_n,t]) = \frac{(n^2-n)}{\sqrt{3n-3}} + \frac{(mn^{t-1}-n^2+n)}{\sqrt{4n-4}}.$$

Proof. Note that $S[K_n, t]$ has n vertices of degree $n-1$ and $(n^t - m\frac{n^{t-1}-1}{n-1}) - n$ vertices of degree $2(n-1)$. Hence, one has

$$\begin{aligned}
\chi(S[K_n,t]) &= \sum_{uv\in E(S[K_n,t])} [\deg(u)+\deg(v)]^{-\frac{1}{2}} \\
&= \sum_{uv\in E(S[K_n,t])} ((n-1)+(2n-2))^{-\frac{1}{2}} + \sum_{uv\in E(S[K_n,t])} ((2n-2)+(2n-2))^{-\frac{1}{2}} \\
&= \frac{(n^2-n)}{\sqrt{3n-3}} + \frac{(mn^{t-1}-n^2+n)}{\sqrt{4n-4}}.
\end{aligned}$$

□

Theorem 2.5. *The general Randić index of generalized Sierpiński gasket graph of K_n in step t is given by*

$$R_\alpha(S[K_n,t]) = (n^2-n)(2n^2-4n+2)^\alpha + (mn^{t-1}-n^2+n)(2n-2)^{2\alpha}.$$

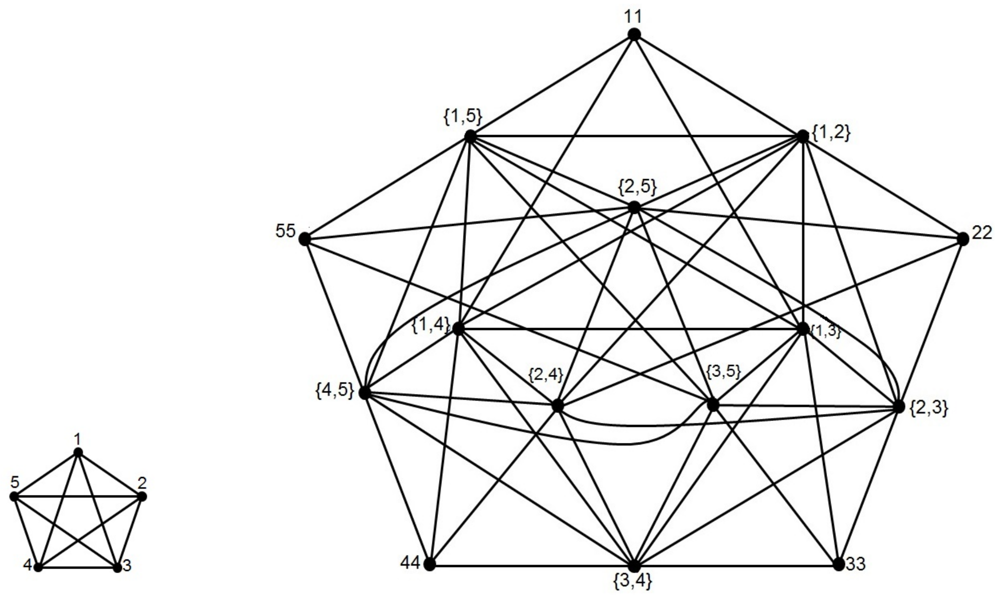


Fig. 2. K_5 and $S[K_5, 2]$.

Proof. By considering structure of generalized sierpiński gasket graph, in $S[K_n, t]$, we have n extreme vertices in form $ii \dots i$, $1 \leqslant i \leqslant n$, and other vertices are in contracted form. Since each extreme vertex has degree $n-1$ and its neighbours have degree $2(n-1)$, we have

$$\begin{aligned} R_{\alpha}(S[K_n,t]) &= \sum_{uv \in E(S[K_n,t])} [\deg(u) \times \deg(v)]^{\alpha} \\ &= \sum_{uv \in E(S[K_n,t])} ((n-1)(2n-2))^{\alpha} + \sum_{uv \in E(S[K_n,t])} ((2n-2)(2n-2))^{\alpha} \\ &= (n^2-n)((n-1)(2n-2))^{\alpha} + (mn^{t-1} - n^2 + n)((2n-2))^{2\alpha} \\ &= (n^2-n)(2n^2-4n+2)^{\alpha} + (mn^{t-1} - n^2 + n)(2n-2)^{2\alpha}. \end{aligned}$$

□

Corollary 2.6. *If in Theorem 2.5, $\alpha = -\frac{1}{2}$, then, the Randić index of $S[K_n, t]$ is equal to*

$$R(S[K_n,t]) = \frac{(n^2-n)}{\sqrt{2n^2-4n+2}} + \frac{(mn^{t-1}-n^2+n)}{2n-2}.$$

Proof. The Randić index of $S[K_n, t]$ is as follows:

$$\begin{aligned} R(S[K_n,t]) &= \sum_{uv \in E(S[K_n,t])} [\deg(u) \times \deg(v)]^{-\frac{1}{2}} \\ &= \sum_{uv \in E(S[K_n,t])} ((n-1)(2n-2))^{-\frac{1}{2}} + \sum_{uv \in E(S[K_n,t])} ((2n-2)(2n-2))^{-\frac{1}{2}} \\ &= \frac{(n^2-n)}{\sqrt{2n^2-4n+2}} + \frac{(mn^{t-1}-n^2+n)}{2n-2}. \end{aligned}$$

□

Theorem 2.7. *The Arithmetic-Geometric index of generalized Sierpiński gasket graph of K_n in step t is given by*

$$AG_1(S[K_n,t]) = \frac{3n^3-6n^2+3n}{2\sqrt{2n^2-4n+2}} + (mn^{t-1}-n^2+n).$$

Proof. The Arithmetic-Geometric index of $S[K_n,t]$ is as follows:

$$\begin{aligned} AG_1(S[K_n,t]) &= \sum_{uv\in E(S[K_n,t])} \frac{\deg(u)+\deg(v)}{2\sqrt{\deg(u)\times\deg(v)}} \\ &= \sum_{uv\in E(S[K_n,t])} \frac{(n-1)+(2n-2)}{2\sqrt{(n-1)(2n-2)}} + \sum_{uv\in E(S[K_n,t])} \frac{(2n-2)+(2n-2)}{2\sqrt{(2n-2)(2n-2)}} \\ &= \frac{3n^3-6n^2+3n}{2\sqrt{2n^2-4n+2}} + (mn^{t-1}-n^2+n). \end{aligned}$$

□

Theorem 2.8. *The general sum-connectivity index of generalized Sierpiński gasket graph of C_n in step t is given by*

$$\chi_\alpha(S[C_n,t]) = n^{t-1}(n-4)(4)^\alpha + \frac{4n(n^{t-2}-1)}{n-1}(8)^\alpha + \frac{4n+n^{t-1}(4n-8)}{n-1}(6)^\alpha.$$

Proof. we have three types of edges, there are $n^{t-1}(n-4)$ edges whose two endpoints have degree 2, and $\frac{4n(n^{t-2}-1)}{n-1}$ edges whose endpoints have degree 4, and $\frac{4n+n^{t-1}(4n-8)}{n-1}$ edges whose endpoints have degree 2 and 4. Thus,

$$\begin{aligned} \chi_\alpha(S[C_n,t]) &= \sum_{uv\in E(S[C_n,t])} [\deg(u)+\deg(v)]^\alpha \\ &= \sum_{uv\in E(S[C_n,t])} (2+2)^\alpha + \sum_{uv\in E(S[C_n,t])} (4+4)^\alpha + \sum_{uv\in E(S[C_n,t])} (2+4)^\alpha \\ &= n^{t-1}(n-4)(4)^\alpha + \frac{4n(n^{t-2}-1)}{n-1}(8)^\alpha + \frac{4n+n^{t-1}(4n-8)}{n-1}(6)^\alpha. \end{aligned}$$

□

Corollary 2.9. *If in Theorem 2.8, $\alpha=-\frac{1}{2}$, then sum-connectivity index of $S[C_n,t]$ is equal to*

$$\chi(S[C_n,t]) = \frac{1}{2}n^{t-1}(n-4) + \frac{1}{2\sqrt{2}}\frac{4n(n^{t-2}-1)}{n-1} + \frac{1}{\sqrt{6}}\frac{4n+n^{t-1}(4n-8)}{n-1}.$$

Theorem 2.10. *The general Randić index of generalized Sierpiński gasket graph of C_n is determined by*

$$R_\alpha(S[C_n,t]) = n^{t-1}(n-4)(4)^\alpha + \frac{4n(n^{t-2}-1)}{n-1}(16)^\alpha + \frac{4n+n^{t-1}(4n-8)}{n-1}(8)^\alpha.$$

Corollary 2.11. *The Randić index of $S[C_n,t]$ is obtained*

$$R(S[C_n,t]) = \frac{1}{2}n^{t-1}(n-4) + \frac{n(n^{t-2}-1)}{n-1} + \frac{\sqrt{2}(n+n^{t-1}(n-2))}{n-1}.$$

Theorem 2.12. *The Arithmetic-Geometric index of generalized Sierpiński gasket graph of C_n in step t is given by*

$$AG_1(S[C_n,t]) = n^{t-1}(n-4) + \frac{4n(n^{t-1}-1)}{n-1} + 3\sqrt{2}\,(\frac{n+n^{t-1}(n-2)}{n-1}).$$

References

[1] B. Bollabás, P. Erdős, Graphs of extremal weights, Ars Combin. 50 (1998) 225-233.

[2] S. Gravier, M. Kovše, and A. Parreau, Generalized Sierpiński graphs, in: Posters at EuroComb'11, Budapest, 2011, http://www.renyi.hu/conferences/ec11/posters/parreau.pdf

[3] O. Ivanciuc, T. Ivanciuc, D. Cabrol-Bass, A.T. Balaban, Evaluation in quantitative structure-property relationship models of structural descriptors derived from information-theory operators, J. Chem. Inf. Comput. Sci. 40 (2000) 631-643.

[4] Klavžar, S. Coloring Sierpiński graphs and Sierpiński gasket graphs Taiwanese Journal of Mathematics., 12 (2008), 513-522.

[5] S. Klavžar and U. Milutinovic, Graphs $S(n, k)$ and a variant of the Tower of Hanoi problem, Czechoslov. Math. J. 47 (1997), 95-104.

[6] S. Klavžar, U. Milutinovic and C. Petr, 1-perfect codes in Sierpiński graphs, Bull. Aust. Math. Soc. 66 (3) (2002), 369-384.

[7] M. Randić, On characterization of molecular branching, J. Am. Chem. Soc. 97 (1975) 6609-6615.

[8] R.S. Scorer, P.M. Grundy and C.A.B. Smith, Some binary games, Math. Gaz., 28 (1944), 96-103.

[9] V. S. Shigehalli and Rachanna Kanabur, Computation of New Degree-Based topological indices of graphene, J. Math., 6 (2016), 1-7

[10] B. Zhou, N. Trinajstić, On general sum-connectivity index, J. Math. Chem. 47 (2010) 210-218.

[11] B. Zhou, N. Trinajstić, On a novel connectivity index, J. Math. Chem. 46 (2009) 1252-1270.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–272, pp. 471–474

Some results on Nil clean graphs

Moloukkhatoon Bozorgzadeh[a,*], Hosein Fazaeli Moghimi[a], Mahdi Samiei[b]

[a]*Department of Mathematics, University of Birjand, Birjand, Iran*
[b]*Department of Mathematics, University of Neyshabur, P.O.Box 91136-899, Neyshabur, Iran*

Article Info

Keywords:
Nil clean ring
Nil clean graph
orthogonal set
chromatic number

2020 MSC:
msc1
msc2

Abstract

Let R be a commutative ring with nonzero identity. The Nil clean graph $N.G(R)$ is a simple graph with vertex set the Nil clean element of R; and two element x and y are adjoint if and only if xy or $x - y$ in $Nil(R)$. In this paper some of the basic graph theoric properties like the diameter, maximum and mimimum degree, orthogonal set , choromatic number of nil clean graph are studied.

1. Introduction

When one assigns a graph with an algebraic structure, numerous interesting algebraic problems arise from the translation of some graph-theoretic parameters such as clique number, chromatic number, diameter, radius and so on. There are many papers about the attribution of a graph to a ring. Due to the fact that the idempotents, nilpotents and unit elements in a ring are key instruments for the acknowledgment of the structure of a ring, different definitions for the graphs related with rings were given utilizing these concepts, each of which characterized the ring structure in a few way. In these structures, the condition of the connecting vertices is regularly that the sum, difference or product of the vertices could be a zero divisor, a unit or a nilpotent element, for example, zero divisor graphs [2], unitary Cayley graph [1], a kind of graph structure [4]
Presenting the concept of nil clean elements in [8], The elements are written as a sum of an idempotent element and a nilpotent element, has made a wide field of exciting investigate topic, leading to numerous captivating comes about, for occasion see, [3, 5, 6, 9].

Let R be a commutative ring with nonzero unitary. An element $e \in R$ is called idempotent if $e^2 = e$, and $n \in R$ is called nilpotent, if there exists a poitive integer k such that $n^k = 0$ and denoted by Id(R) and $Nil(R)$ respectively. An element a of R is nil clean when a can be written as a sum of idempotent and nilpotent element. The subset N.C(R)

*Talker
Email addresses: M.bozorgzadeh@birjand.ac.ir, M.bozorgzadeh83@gmail.com (Moloukkhatoon Bozorgzadeh), hfazaeli@birjand.ac.ir (Hosein Fazaeli Moghimi), m.samiei@neyshabur.ac.ir, m.samiei91@yahoo.com (Mahdi Samiei)

of a ring R consists all the nonzero nil clean elements of R. A pair of idempotents e and f are said to be orthogonal if $ef = 0$. Let e be an idempotent element of R. The set of all nonzero orthogonal idempotents of e is denoted by O_e. In fact, $O_e = \{0 \neq f \in \mathrm{Id}(R) : \ ef = 0\}$.
In this paper we focouse on nil clean graph. The main object of this paper is to study the interplay between the ring-theoretic properties of R and the graph-theoretic properties of $N.G(R)$. This study helps illuminate the structure of NC(R). In [11] is defined following nil clean graph and show that basic properties of this graph and obtained some result about graph from ring and conversly.

Definition 1.1. The nil clean graph of a ring R, denoted by $N.G(R)$ is defined by vertices $V(N.G(R)) = \mathrm{N.C}(R)$, and for distinct $x, y \in \mathrm{N.C}(R)$, the vertices x and y are adjacent if and only if $xy \in \mathrm{Nil}(R)$ or $x - y \in \mathrm{Nil}(R)$.

In [11][2.3] is expressed an equivalent condition for adjacency of two vertices, which plays a key role in identifying Nil clean graph.

Theorem 1.2. *For a commutative ring R, the vertices $x = e + n$ and $y = f + m$ are adjacent, where $e, f \in \mathrm{Id}(R)$ and $n, m \in \mathrm{Nil}(R)$, if and only if $e = f$ or $ef = 0$.*

2. Main results

The focus of this section will be on elucidating the concept of the nil clean graph of a commutative ring and emphasizing its notable characteristics.
We introduce two subgraphs as follows:

$$N.C_1(R) = \{(0, n) : \ n \in Nil(R)\}$$

and

$$N.C_2(R) = \{(e, n) : \ 0 \neq e \in Id(R) \text{ and } n \in Nil(R)\}.$$

It is clear that

$$N.G(R) = N.G_1(R) \vee N.G_2(R)$$

where $N.G_1(R)$ is a complete subgraph of $N.G(R)$ with vertices $N.C_1(R)$ and $N.G_2(R)$ is a subgraph of $N.G(R)$ with vertices $N.C_2(R)$, thus to explor main properties of $N.G_2(R)$, we need to study $N.G_2(R)$.
Let R be a reduced ring. Then $V(N.G_2(R)) = \mathrm{Id}(R)\backslash\{0\}$. So, in this case, $\mathrm{N.C}(R)$ is a subset of nonzero zero-divisor elements of R and the next Theorem shows that two vertices x and y are adjacent if and only if $xy = 0$.

Theorem 2.1. *Let R be a reduced ring. Then $N.G_2(R)$ is a subgraph of the zero-divisor graph of R.*

Corollary 2.2. *For a ring R, if $N.G_2(R)$ is connected, then R is non-reduced. The convere is not true in general.*

The ring of integers modulo 12 is a non-reduced ring in which $N.G_2(\mathbb{Z}_{12})$ is a not a connected graph.

Theorem 2.3. *Let R be a ring with a non trivial idempotent then $N.G_2(R)$ is not connected graph.*

In [11] is showed for a commutative ring R with a non trivial idempotent, $\mathrm{diam}(N.G_2(R)) \leq 3$. Moreover, If $N.G_2(R)$ contains a cycle, then $g(N.G_2(R)) \leq 7$.
The aim of this paper is to introduce and study some of the basic properties of the nil clean graph $N.G_2(R)$, some of its subgraphs $N.G_{21} = \{(1, n) : n \in Nil(R)\}$ and $N.G_{22} = \{(e, n) : 1, 0 \neq e \in Id(R), n \in Nil(R)\}$.

$$N.G_2(R) = N.G_{21} \vee N.G_{22}$$

by definition nil clean graph, we see if R be a non indecomposible ring then any vertex of $N.G_{21}$ is not conjected with vertex of $N.G_{22}$ and $N.G_2(R)$ is a nonconnected graph that composed of two subgraph $N.G_{21}$ and $N.G_{22}$.

Corollary 2.4. *for a ring R*

(1) $N.G_{21}$ *is isomorphic to* $K_{|Nil(R)|}$

(2) For a ring R, $N.G_{22}$ *is a connected graph.*

Corollary 2.5. *If* R *has a nontrivial nilpotent and nontrivial idempotent then* $N.G_{22}(R)$ *has At least a triangle.*

Proof. suppose that e is nontrivial idempotent and n is nontrivial nilpotent of R. then we have the cycle $(e, n) \sim (1-e, n) \sim (1-e, 0) \sim (e, n)$, and we are done. □

Corollary 2.6. *If* $|Nil(R)| \geq 3$ *then* $N.G_2(R)$ *has At least a triangle.*

Proof. suppose that n_1, n_2 are nontrivial nilpotent of R then we have the cycle $(e, 0) \sim (e, n_1) \sim (e, n_2) \sim (e, 0)$, for every $0 \neq e \in Id(R)$. and we are done. □

Theorem 2.7. *Let* R *be a ring then* $gr(N.G_2(R)) \in \{3, \infty\}$

Proof. If R be indecomposible and reduced then $(1, 0)$ is isolated vertices and $N.G_2(R)$ has no cycle so $gr(N.G_2(R)) = \infty$. If R not indecomposible and reduced by corollary 2.5 $N.G_2(R)$ has a cycle with lenght 3. □

The degree of $v \in V(G)$ denoted by $deg(v)$, is the number of edges of G incident with v. We denoted the minimum and the maximum degrees of G, by $\delta(G)$ and $\Delta(G)$, respectively.

Theorem 2.8. *If* R *a finite ring then* $\delta(N.G_2(R))=N-1$ *and* $\Delta(N.G_2(R))=|Nil(R)|(|O_{ei}|+1)-1$ *where* e_i *is idempotent with maximum orthogonal idempotents.*

Lemma 2.9. *For the nil clean graph* $N.G_2(R)$ *we have the following:*

(1) If R *is reduced, then* $\deg(1, 0) = 0$ *and for every* $0, 1 \neq e \in Id(R), deg(e, 0) = |O_e|$.

(2) If $|\mathrm{Nil}(R)| = 2$*, then*

(2.1) $\deg(1, 0) = \deg(1, n) = 1$ *where* n *is nonzero.*

(2.2) $\deg(e, n) = 2|O_e| + 1$ *where* e *is a non-trivial idempotent of* R *and* n *is nonzero.*

(3) If $|\mathrm{Nil}(R)| = k \geq 3$*, then* $\deg(e, n) = k(1 + |O_e|) - 1$ *where* e *is nonzero.*

The next result [10] provides us a plenty source of examples.

Corollary 2.10. *Let* $R_1, R_2, \cdots, R_n$ *be indecomposable reduced rings. If* $R = R_1 \times R_2 \times \cdots R_n$*, then* $N.G_2(R)$ *is finite and moreover,*

(1) $V(N.G_2(R)) = \{(e_1, \cdots, e_n) : e_i = 0$ *or* 1 *in* $R_i\} - \{(0, 0, \cdots, 0)\}$.

(2) $|V(N.G_2(R))| = 2^n - 1$.

(3) If $x = (e_1, \cdots, e_n)$ *where* $k < n$ *number of the* e_i*'s are equal to zero, then* $deg(x) = 2^k - 1$.

Weddernburn-Artin theorem state that any semisimple commutative ring is a finite direct product of fields. Thus, by this fact the following example follows.

Example 2.11. For a semisimple ring $R \cong F_1 \times F_2 \times \cdots \times F_n$, $N.G_2(R)$ is finite. In addition, the statements of corollary 2.10 are valid.

Proposition 2.12. *If* $x = (x_1, \cdots .x_n)$ *and* $y = (y_1, \cdots, y_n)$ *are adjacent vertices in the ring* $R_1 \times R_2 \times \cdots R_n$*, then* x_i *and* y_i *are adjacent vertices in* R_i *for all* $1 \leq i \leq n$*. The converse is not true in general.*

Now we would like to determine the clique number and the chromatic number of the idempotent graphs and the relationship between them. First we define M_O the maximal set of orthogonal idempotent element. for every idempotent element e_i we defined $M_{O_{ei}} = \{f \in O_{ei} : fg = 0, g \in O_{ei}\}$, $M_{O_j} = Max\{M_{O_{ei}} : e_i \in Id(R)\}$. may be this set is not unique but they have the same number of members. Now we start with the following theorem.

Theorem 2.13. *Let R be a ring and $x = (e, n)$ be a vertex of degree at least $2|Nil(R)|$ then $\omega(N.G_2(R)) \geq 3|Nil(R)|$*

Proof. x is a vertex of degree at least $2|Nil(R)|$ so there exists a vertex $y = (f, m)$ where $f \neq 1 - e$ such that e and f are orthogonal. now all of vertices $x = (e, n = n_1) \sim (e, n_2) \sim .... \sim (e, n_{|Nil(R)|}) \sim (1 - e - f, n_1) \sim .... \sim (1 - e - f, n_{|Nil(R)|}) \sim (f, n_1) \sim .... \sim (f, n_{|Nil(R)|} = m)$ are adjacent, so this a clique in $N.G_2(R)$ and $\omega(N.G_2(R)) \geq 3|Nil(R)|$. □

Theorem 2.14. *»et R be a ring*

(1)]f R be indecomposible ring then $\omega(N.G_2(R))$=$|Nil(R)|$

(2)]f R is not a indecomposible ring then $\omega(N.G_2(R)) = |M_O||Nil(R)|$

now we want to obtain chromatic number of $N.G_2(R)$.

Theorem 2.15. *]f R be a ring and $N.G_2(R)$ be Nil clean graph then we have $\omega(N.G_2(R)) = \chi(N.G_2(R))$*

References

[1] R. Akhtar, M. Boggess, T. Jackson-Henderson, I. Jim´enez, R. Karpman, A. Kinzel, D. Pritikin, *On the unitary Cayley graph of a finite ring*, Electron. J. Combin., 16 (2009), https://doi.org/10.37236/206.
[2] D. F. Anderson and P. S. Livingston, *The zero-divisor graph of a commutative ring*, J. Algebra 217 (1999) 434-447.
[3] D. K. Basnet and J. Bhattacharyya, *Nil Clean Graphs of Rings*, Algebra Colloq. 24(03) (2017), 481-492.
[4] P. W. Chen, *A kind of graph structure of rings*, Algebra Colloq., 10(2) (2003), 229-238.
[5] P. V. Danchev and W. W. McGovern, *Commutative weakly nil clean unital rings*. J. Algebra, 425 (2015), 410-422.
[6] P. Danchev and M. Samiei, *The probability when a finite commutative ring is nilclean*, Trans. A. Razmadze Math. Inst. 176 (2022).
[7] R. Diestel, Graph Theory, Springer-Verlag, New York, 1997.
[8] A.J. Diesl, *Nil clean rings*, J. Algebra 383 (2013) 197-211.
[9] T. Koşan, Z. Wang and Y. Zhou, *Nil-clean and strongly nil-clean rings*, J. Pure Appl. Algebra, 220(2) (2016), 633-646.
[10] M. Samiei, *Commutative rings whose proper homomorphic images are nil clean*. Novi Sad J. Math. 50, 1 (2020), 37-44.
[11] M.Bozorgzadeh, H.F. Moghimi and M.Samiei, *Nil clean graphs associted with commutative rings*, J.of Algebra and its applications, 2023

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GC–288, pp. 475–478

The Game Chromatic Number and The Game Chromatic Index of Cartesian Product of Path and Cycle

Elham Sharifi Yazdi[a,*]

[a]*Department of Computer Engineering, Imam Javad University College, Yazd, Iran*

Article Info

Keywords:
game chromatic number
game chromatic index
Cartesian product

2020 MSC:
05C15
05C57

Abstract

In this paper, we determine the exact value of game chromatic number of Cartesian product $P_3 \square C_n$ and game chromatic index of Cartesian product $P_2 \square C_n$ where C_n is a cycle on n vertices.

1. Introduction

Let G be a simple graph and $X = \{1, 2, ..., k\}$ be a set of colors. Consider two players. Alice and Bob take turns in playing the game. They alternately color a vertex of G with a color from X. Alice always moving first. In this coloring no two adjacent vertices recieve the same color. In the end if all the vertices of G are colored properly, Alice wins and Bob wins if at any stage of the game before the G completely colored, one of the players has no legal move. The game chromatic number of G, denoted by $\chi_g(G)$, is the least number k for which Alice has a winning strategy. The game coloring introduced by Bodlaender [7]. There are some results for game chromatic number of graph, see [3, 5, 6, 8, 10, 11, 15, 16]. It is clear that the game chromatic number $\chi_g(G)$ satisfies

$$\chi(G) \le \chi_g(G) \le \Delta(G) + 1. \tag{1}$$

where $\chi(G)$ and $\Delta(G)$ are chromatic number and the largest vertex degree of the graph G, respectively.
It is defined similarly for edge coloring. That means Alice and Bob will be coloring the edges of a graph G instead of its vertices with Alice start playing first. They alternately color an edge of G with a color from X such that two adjacent edges have different colors. Alice wins the game if it is possible to color all the edges of G with colors in X . Otherwise Bob wins. The game chromatic index $\chi'_g(G)$ is the least number of colors such that Alice has winning strategy. The edge game coloring problem of graphs was introduced by Lam et al. [13] and it was studied by many

*Talker
Email address: sharifiyazdielham@yahoo.com (Elham Sharifi Yazdi)

people, see [1, 2, 4, 9, 12, 14]. The trivial bounds are:

$$\Delta(G) \leq \chi'_g(G) \leq 2\Delta(G) - 1. \tag{2}$$

where $\Delta(G)$ denotes the maximum degree of G.
The Cartesian product of graphs G and H, denoted by $G\Box H$, where two vertices (u, v) and (u', v') are adjacent if and only if $u = u'$ and $vv' \in E(H)$ or $v = v'$ and $uu' \in E(G)$. In this paper we determine the exact value of $\chi_g(P_3\Box C_n)$ and $\chi'_g(P_2\Box C_n)$, where C_n is a cycle with n vertices.

2. Game Chromatic Number of $P_3\Box C_n$

We denote vertices of copies C_n of graph $P_3\Box C_n$ by $V_1 = \{v_1, ..., v_n\}$, $V_2 = \{v'_1, ..., v'_n\}$ and $V_3 = \{v''_1, ..., v''_n\}$.

Theorem 2.1. *For every integer $n \geq 3$, $\chi_g(P_3\Box C_n) \geq 4$.*

Proof. At first, we show that Bob has a winning strategy using 3 colors. For Alice first move consider following the cases:
Case 1: Alice colors a vertex of degree 3.
Without loss of generality, suppose Alice colors vertex v_1 with color 1. Then Bob replies with vertex v'_2 with color 2. In the second move, if Alice colors a vertex of v_j $(2 \leq j \leq n)$ or v'_i or v''_i $(3 \leq i \leq n)$ then Bob replies with vertex v''_1 with color 3 and Bob wins. Suppose That in the second move Alice colors one of the vertices of $\{v_1, v''_1, v''_2\}$ then Bob colors vertex v_3 with color 3 and he wins.
Case 2: Alice colors a vertex of degree 4.
Without loss of generality, suppose Alice colors vertex v'_1 with color 1. Then Bob colors vertex v''_2 with color 2. In the next move, if Alice colors a vertex of v_i $(1 \leq i \leq n)$ or v'_j $(1 \leq j \leq n-1)$, then Bob replies with vertex v''_n with color 3 and Bob wins. Also if Alice colors a vertex of v''_i $(1 \leq i \leq n, i \neq 2)$ or v'_n, then Bob replies with vertex v_2 with color 3 and Bob wins □

Theorem 2.2. *For every integer $n \geq 3$, $\chi_g(P_3\Box C_n) \leq 4$.*

Proof. Let $X = \{1, ..., 4\}$ be a color set. Now we give a winning strategy for Alice with color set X. At first, Alice colors vertex v_2 with color 1. In the following moves, if Bob colors a vertex $v_i(v''_i)$ from $V_1(V_3)$ with color j then Alice colors vertex $v''_i(v_i)$ from $V_3(V_1)$ with the same color. Now if Bob colors a vertex v'_i from V_2 with color j then Alice colors a vertex among v'_{i+2} or v'_{i-2} $(i \leq n-2)$ with color j. If it is not possible, she colors any vertex with an available color. In this strategy any uncolored vertex is adjacent to colored neighbors with at most 3 distinct colors and Alice wins. □

Corollary 2.3. *For every integer $n \geq 3$, we have $\chi_g(P_3\Box C_n) = 4$.*

3. Game Chromatic Index of $P_2\Box C_n$

We denote edges of copies C_n of graph $P_2\Box C_n$ by $E_1 = \{e_1, ..., e_n\}$, $E_2 = \{e''_1, ..., e''_n\}$ and edges of copies P_2 by $E_3 = \{e'_1, ..., e'_n\}$ such that edge e'_i is neighbor with e_i and e''_i $(1 \leq i \leq n)$.

Theorem 3.1. *For every integer $n \geq 3$, $\chi'_g(P_2\Box C_n) \geq 4$.*

Proof. We show that Bob has a winning strategy using three colors. For Alice first move consider following the cases:
Case 1: Alice colors an edge from E_3.
Without loss of generality, suppose Alice colors vertex e'_1 with color 1 then Bob colors e_2 with color 2. For the Alice next move consider the following cases:
Case 1.1: Alice colors an edge from E_2.

Then Bob colors e_n with color 3 and Bob wins.
Case 1.2: Alice colors an edge from E_1.
If Alice colors e_1 then Bob replies with edge e''_2 with color 1 and Bob wins. Now suppose Alice colors an edge e_i $(3 \leq i \leq n)$ in her second move then Bob colors e'_2 with color 3 and Bob wins.
Case 1.3: Alice colors an edge from E_3.
If Alice colors e'_2 then Bob replies with edge e_n with color 3 and Bob wins. Now suppose Alice colors an edge e'_i $(3 \leq i \leq n)$ then Bob colors e''_2 with color 3 and Bob wins.
Case 2: Alice colors an edge from E_1 or E_2.
Without loss of generality, suppose Alice colors e_1 with color 1 then Bob colors e''_2 with color 2. For the Alice second move consider the following cases:
Case 2.1: Alice colors an edge from E_2.
Then Bob colors e_2 with color 3 and Bob wins.
Case 2.2: Alice colors an edge from E_1.
Then Bob colors e''_1 with color 3 and Bob wins.
Case 2.3: Alice colors an edge from E_3.
If Alice colors e'_1 then Bob replies with edge e_2 with color 3 and Bob wins. Consider Alice colors e'_2 then Bob replies with edge e''_n with color 1 and Bob wins. Now suppose Alice colors an edge e'_i $(3 \leq i \leq n)$ then Bob colors e''_1 with color 3 and again Bob wins. □

Theorem 3.2. *For every integer $n \geq 3$, $\chi'_g(P_2 \square C_n) \leq 4$.*

Proof. We show that Alice has a strategy to win the game using four colors. Alice, in the second move, has this strategy. In the first priority, she colors an edge such that it has three colored neighbors with three colors. In second priority, if Bob colors an edge $e_i(e''_i)$ from $E_1(E_2)$ with color j, if it is possible, Alice colors edge $e''_i(e_i)$ from $E_2(E_1)$ with the same color j, otherwise she colors an edge of distance 2 from $e_i(e''_i)$. Similarly, if Bob colors edge e'_i then Alice colors edge e'_{i+1} or e'_{i-1} with color j and if it is not possible she colors an edge of distance 2 from e'_i .By using this strategy, at any stage, any uncolored edge has two neighbors colored with the same color or if it has three colored neighbors with three colors then Alice colors it. □

Corollary 3.3. *For every integer $n \geq 3$, we have $\chi'_g(P_2 \square C_n) = 4$.*

4. Conclusion

In this paper we have determined the exact value of game chromatic number of graph $P_3 \square C_n$ and game chromatic index of graph $P_2 \square C_n$ where C_n is a cycle on $n \geq 3$ vertices and the results are obtained $\chi_g(P_3 \square C_n) = \chi'_g(P_2 \square C_n) = 4$.

References

[1] R. Alagammai, and V. Vijayalakshmi. Game chromatic number and game chromatic index of the Mycielski graphs of some families of graphs. *Appl. Math. Inf. Sci.*,13(S1):261–265, 2019.
[2] S.D. Andres. The game chromatic index of forests of maximum degree $\Delta \geq 5$. *Discrete Appl. Math.*, 154:1317–1323, 2006.
[3] S.D. Andres, and W. Hochsättler. The game colouring number of powers of forests. *Discrete Math. Theor. Comput. Sci.*, 18:#2, 2015.
[4] S.D. Andres, W. Hochsättler, and C. Schalluck. The game chromatic index of wheels. *Discrete Appl. Math.*, 159:1660–1665, 2011.
[5] S.D. Andres, and A. Theuser. Note on the game colouring number of powers of graphs. *Discuss. Math. Graph Theory*, 36:31–42, 2016.
[6] T. Bartnicki, B. Breˇsar, J. Grytczuk, M. Kovˇse, Z. Miechowicz, and I. Peterin. Game chromatic number Cartesian product graphs. *Electron. J. Combin.*, 15:R72, 2008.
[7] H.L. Bodlaender. On the complexity of some colouring games. *Int. J. Found. Comput. Sci.*, 2(2):133–147, 1989.
[8] C.J. Destacamento, A.D. Rodriguez, and L. Aquino-Ruiviva. The game chromatic number of some classes of graphs. *DLSU Research Conggress*, De La Salle University, Manila, Philippines, March 6-8, 2014.
[9] P.L. Erd¨os, U. Faigle, W. Hochs¨attler, and W. Kern. Note on the game chromatic index of trees, *Theoret. Comput. Sci.*, 313:371–376, 2004.
[10] A.L.C. Furtado, M.A.D.R. Palma, S.Dantas, and C.M.H. de Figueiredo. On the degree of trees with game chromatic number 4, *RAIRO-Oper. Res.*, 57(5):2757– 2767, 2023.
[11] M.Imran1,S.A.U.H.Bokhary,M.S.Akhtar, N.Matsumoto. Game chromatic number of honeycomb related networks, *Communications in Combinatorics and Optimization*, 1-11, 2023.
[12] R. Keusch. A new upper bound on the game chromatic index of graphs. *Electron. J. Combin. 25(2)*, P2.33, 2018.

[13] P.C.B. Lam, W.C. Shiu, and B. Xu. Edge game coloring of graphs. *Graph Theory Notes N. Y.*, 37:17–19, 1999.
[14] E. Sharifi Yazdi. The game chromatic index of special graphs. *52nd Annual Iranian Mathematics Conference*, Shahid Bahonar University of Kerman, IEEE, 30 August- 02 September, 2021.
[15] E. Sharifi Yazdi. The game chromatic number of power two of Cartesian product of two paths. *The 3rd National Congress on Mathematics and Statistics*, Gonbad Kavous University, 444–446, 24 Feb, 2022.
[16] C. Sia. The game chromatic number of some families of Cartesian product graphs. *AKCE Int. J. Graphs and Combin.*, 6(2):315–327, 2009.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–190, pp. 479–482

Non-Existence of $\alpha-$ biharmonic Maps

Seyed Mehdi Kszrmi Torbaghan[a,*], Amir Shahnavaz[b]

[a] *Department of mathematics, Faculty of Basic Sciences, University of Bojnord*
[b] *Department of Mathematics, University of Sistan and Baluchestan, Zahedan, Iran*

Article Info

Keywords:
Conformal vector field
$\alpha-$ biharmonic map
Riemannian manifold

2020 MSC:
53C43
58E20.

Abstract

In this paper, it is shown that any $\alpha-$ biharmonic map from an arbitrary Riemannian manifold to a Riemannian manifold admitting a conformal vector field with potential function $\mu > 0$, is constant.

1. introduction

The theory of harmonic maps was first introduced in mathematics in 1954, almost six years earlier than a particular example of it, the $O(4) \sim SU(2) \times SU(2)$ non-linear sigma model, was independently introduced in physics [2] to explain the low energy interaction of π mesons, the name coming from the sigma resonance which had by then been observed in this system. Generalized sigma models (which are typically harmonic maps) were introduced in physics later(in 1978), [5]. At present one often hears it said colloquially that ' mathematicians call harmonic maps what physicists call sigma model'. Therefore, harmonic maps has become one of the effective tools for studying in many fields of mathematical physics such as theory of relativity, theory of electromagnetism, gravitational theories and etc. , [1].
Let $\phi : (M, g) \longrightarrow (N, h)$ be a smooth map between Riemannian manifolds. The bienergy functional of ϕ is denoted by $e(\phi)$ and defined as follows

$$E_2(\phi) = \frac{1}{2} \int_M \mid \tau(\phi) \mid^2 \upsilon_g, \tag{1}$$

where $\tau(\phi) := Tr_g \nabla d\phi$ is the tension field of ϕ and υ_g is the volume element of (M, g) . A smooth map ϕ is called a biharmonic map if ϕ is a critical point of the bienergy functional E_2. This is equivalent to saying that ϕ satisfies the Euler-Lagrange equation corresponding to E_2 given by

$$\tau_2(\phi) := Tr_g \left[\nabla d\phi + R^N(d\phi, \tau(\phi))d\phi\right] = 0, \tag{2}$$

---

*Talker

Email addresses: m.kazemi@ub.ac.ir (Seyed Mehdi Kszrmi Torbaghan), amir.shahnavaz@pgs.usb.ac.ir (Amir Shahnavaz)

here ∇ is the induced connection on the pull-back bundle $\phi^{-1}TN$. The section $\tau_2(\phi) \in \Gamma(\phi^{-1}TN)$ described in (2), is known as the *bitension field* of ϕ.
Biharmonic maps from a closed compact Riemannian surface and their variants are useful tools in both physics and mathematics for studying the geometry of a Riemannian manifold. In mathematical physics, they are borderline examples for the Palais- Smale condition and so can not be reached directly by using normal techniques. For solving, Sacks and Uhlenbeck in their prominent paper [3] in 1981, introduced perturbed bienergy functional that satisfied the Palais- Smale condition, and thus obtained well-known α-harmonic maps as critical points of perturbed functional to approximate biharmonic maps. They also proved that if M is a compact Riemannian manifold, then any non-trivial class in $\pi_2(M)$ can be represented by a sum of smooth biharmonic mappings, $\phi_j : S^2 \longrightarrow M,\ j = 1, \cdots, n$, for some positive integer n. In 2019, K. Uhlenbeck, as the first woman, won prestigious Abel prize for her prominent works on $\alpha-$ harmonic maps from minimal surfaces.

2. $\alpha-$biharmonic maps

In this section we recall the notion of $\alpha-$harmonic maps. In this regard, we firstly study the concepts of $\alpha-$bienergy functional, $\alpha-$bitension field and $\alpha-$biharmonic maps. Then we calculate the first variation formula for $\alpha-$bienergy functional via a new technical method.
Let $\phi : (M^m, g) \longrightarrow (N^n, h)$ be a smooth map between Riemannian manifolds. From now on, it is supposed that (M, g) is an $m-$dimensional oriented compact Riemannian manifold unless otherwise stated. Furthermore, the Riemannian connections on Riemannian manifolds M and N are denoted by ∇^M and ∇^N, respectively. Moreover, denote the induced connection on the pullback bundle $\phi^{-1}TN$ by ∇ and defined as follows

$$\nabla_Z W = \nabla^N_{d\phi(Y)} V,$$

for any smooth vector field $Z \in \chi(M)$ and section $W \in \Gamma(\phi^{-1}TN)$
For $\alpha > 1$ and any smooth map $\phi : (M, g) \longrightarrow (N, h)$ between Riemannian manifolds, the $\alpha-$ energy functional of the map ϕ is denoted by $E_\alpha(\phi)$ and defined as follows:

$$E_{2,\alpha}(\phi) := \int_M (1+ \mid \tau(\phi) \mid^2)^\alpha \upsilon_g, \tag{3}$$

This function could be considered as a perturbation of the bienergy functional, E_2, that is, in contrast to E_2, satisfying the Palais-Smale condition. Furthermore, $E_{2,\alpha}$ assuage the Ljusternik- Schnirelman theory and the Morse theory, [4]. By Green's theorem, the first variation formula of $\alpha-$ harmonic maps can be obtained as follows

Theorem 2.1. *(The first variation formula) Let $\{\phi_t : M \longrightarrow N\}_{-\varepsilon<t<\varepsilon}$ be a 1-parameter smooth variation of $\phi : (M, g) \longrightarrow (N, h)$ such that $\phi_0 = \phi$. Setting $\vartheta := \frac{d\phi_t}{dt} \mid_{t=0}$. Then we have*

$$\frac{d}{dt} E_{2,\alpha}(\phi_t) \mid_{t=0} = -\int_M \langle \tau_{2\alpha}(\psi), \vartheta \rangle d\upsilon_g \tag{4}$$

where

$$\begin{aligned} \tau_{2,\alpha}(\psi) &:= Tr_g \nabla(2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi) \\ &= 2\alpha(1+ \mid d\phi \mid^2)^{\alpha-1} \tau(\phi) \\ &+ d\psi(grad(2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1}), \end{aligned} \tag{5}$$

here $\tau(\phi)$ is the tension field of ϕ.

The section $\tau_{2,\alpha}(\phi) \in \Gamma(\phi^{-1}TN)$ is said to be the $\alpha-$ bitension field of ϕ and the critical points of $\alpha-$bienergy functional is called $\alpha-$biharmonic maps.

Theorem 2.2. *$\alpha-$harmonicity of $\phi : (M, g) \longrightarrow (N, h)$ is equivalent by $\tau_{2,\alpha}(\phi) = 0$.*

By (4) and (5) toghther with Theorem 2.2, we get the following result.

Corollary 2.3. *$\alpha-$ biharmonicity of $\phi : (M^m, g) \longrightarrow (N^n, h)$ implies that*

$$\int_M \langle 2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi, \nabla\vartheta\rangle dV_g = 0, \tag{6}$$

for every compactly supported vector field ϑ along ϕ.

Proof. Let p_0 be a point on M and fix it, and let $\{\xi_k\}$ be an orthonormal frame around p_0 with $\nabla\xi_k = 0$ at p_0 for $k = 1, \cdots, m$. Define a 1-form δ as follows

$$\delta(Y) := \langle 2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi(Y), \vartheta\rangle, \tag{7}$$

for any vector field Y on M. By (4), (7) and the first equality of (5), it is obtained that

$$\begin{aligned}
&\frac{d}{dt} E_{2,\alpha}(\phi_t) \mid_{t=0} \\
&= -\int_M \langle Tr_g \nabla(2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi), \vartheta\rangle \\
&= -\int_M \sum_i \langle \nabla_{\xi_i}(2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi(\xi_i)), \vartheta\rangle \upsilon_g \\
&= -\int_M \sum_k \{\xi_k(\langle 2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi(e_k), \vartheta\rangle) \\
&\quad + \langle 2\alpha(1+ \mid d\phi \mid^2)^{\alpha-1} d\phi(\vartheta_k), \nabla_{\xi_i}\vartheta\rangle\}\upsilon_g \\
&= -\int_M div\,\delta + \langle 2\alpha(1+ \mid \tau(\phi) \mid^2)^{\alpha-1} d\phi, \nabla\vartheta\rangle \upsilon_g.
\end{aligned} \tag{8}$$

By (8) and Green's theorem, Corollary 2.3, follows. □

3. Existence of $\alpha-$harmonic maps

In this section, we investigate the existence of $\alpha-$biharmonic maps. First, it is studied the concepts of conformal and killing vector fields and their physical applications. Then the existence of $\alpha-$biharmonic maps between Riemannian manifolds whose target admitting a conformal vector field with potential vector field $\mu > 0$, is studied.

Definition 3.1. A smooth vector field Y on a Riemannian manifold (M, g) is called conformal with the potential function μ If $\mathcal{L}_Y g = 2\mu g$, where $\mathcal{L}_Y g$ is the Lie derivative of the Riemannian metric g on M with respect to Y. Moreover, A conformal vector field Y with the potential μ is said to be a killing vector field If $\mu = 0$ or, equivalently, the flow of Y consists of isometries of (M, g).

Theorem 3.2. *Any conformal vector field Y with potential function μ on M satisfying the following equation*

$$g(\nabla^M_X Y, Z) + g(X, \nabla^M_Z Y) = \mu g(X, Z), \tag{9}$$

for any $X, Z \in \chi(M)$.

Proof. It can be found the proof of this Theorem in [5, page 3] □

Remark 3.3. Conformal vector fields are key tools in physics for constructing various solutions that are then employed in the analysis of physical parameters in modified gravity theories (MGTs). For example, by using conformal vector fields of $pp-$wave space-times in the $f(R)$ theory of gravity, it is possible to discover that the plane fronted gravitational waves (GWs), as a very specific class of $pp-$waves, has a solution in the $f(R)$ theory of gravity, [3]. Killing vector fields are also important in many domains of mathematical physics, including gravitation, quantum and teleparallel theory, and so on. For instance, Sharif and Amir in [4] launched an approach to exploring symmetries in teleparallel theory by introducing the teleparallel version of the Lie derivative for Killing vector fields.

According to the above notations, we investigate the non-existence of non-constant $\alpha-$biharmonic maps as follows:

Theorem 3.4. *Any $\alpha-$biharmonic map $\phi : (M, g) \longrightarrow (N, h)$ between Riemannian manifolds whose target admitting a conformal vector field Y with the potential function $\mu > 0$, is constant.*

Proof. Choose an arbitrary point $p \in M$ and fix it. Let $\{\xi_k\}$ be a normal orthonormal frame at p. Setting

$$\delta(X) := h(Y \circ \phi, (1+ \mid \tau(\phi) \mid^2)^\alpha d\phi(X)), \tag{10}$$

for any smooth vector field X on M. By (9) and (10), the divergence of δ can be calculated as follows

$$\begin{aligned} div\, \delta &= \xi_k(h(Y \circ \phi, (1+ \mid \tau(\phi) \mid^2)^\alpha d\phi(\xi_k))) \\ &= (1+ \mid \tau(\phi) \mid^2)^\alpha h(\nabla_{\xi_k} Y \circ \phi, d\phi(\xi_k)) \\ &= \mu \circ \phi(1+ \mid \tau(\phi) \mid^2)^\alpha h(d\phi(\xi_k), d\phi(\xi_k)) \\ &= \mu \circ \phi \mid \tau(\phi) \mid^2 (1+ \mid d\phi \mid^2)^\alpha, \end{aligned} \tag{11}$$

where we use the $\alpha-$biharmonicity of ϕ for the second equality. From (11), we get

$$\begin{aligned} 0 &= \int_M divX \, dV_g \\ &= \int_M \mu \circ \phi \mid \tau(\phi) \mid^2 (1+ \mid \tau(\phi) \mid^2)^\alpha dV_g. \end{aligned} \tag{12}$$

This implies that $\mid \tau(\phi) \mid= 0$. Then, ϕ is constant and hence completes the proof. $\square$

References

[1] Bajnok, Z. , et al. : Instanton effects vs resurgence in the $O(3)$ sigma model. Physics Letters B. 829, 137073 (2022).
[2] Gell-Mann, M. , Lévy, M. : The axial vector current in beta decay. Il Nuovo Cimento (1955-1965), 16(4), 705-726,(1960).
[3] Sacks, J. , Uhlenbeck, K. : The existence of minimal immersions of 2-spheres, Annals of mathematics.1-24 (1981).
[4] Li, J. , Zhu, X. : Energy identity and recklessness for a sequence of Sacks–Uhlenbeck maps to a sphere. Annales de l'Institut Henri Poincaré C, Analyse non linéaire. 36(1), 103-118. Elsevier Masson (2019).
[5] Zakharov, V. E. , Mikhailov, A. V. : Relativistically invariant two-dimensional models of field theory which are integrable by means of the inverse scattering problem method, Sov.Phys.-JETP. 47, 1017-1027 (1978).

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–211, pp. 483–487

Analysis of the symmetry group of the Kuper-Schmidt equation and classification of its subalgebras

Mehdi Jafari[a], Adele Alemahdi[a,*]

[a]*Department of Matematics , Payam Noor University, PO BOX, 19395-4697*

Article Info

Keywords:
Lie symmetry group
Kuper-Schmidt equation
Optimal system

2020 MSC:
msc1
msc2

Abstract

In this paper, we solve the Kuper-Schmidt equation with the symmetry method. This PDE equation is one of the most important and widely used equations in physics and chemistry. This 5th order nonlinear equation appears as a wave phenomenon in mechanical engineering. In this article, by using the symmetry method by classifying the adjoint representation circuits of the symmetry group on its Lie algebra, we obtain the optimal system of one-dimensional subalgebras of the Kuper-Schmidt equation.

1. Introduction

The Kuper-Schmidt equation is widely used in the diagnosis of various diseases and medical experiments inoptical fibers. It can be used to identify internal body defects, in dentistry and to measure blood and fluids [6].
The Kuper-Schmidt equation plays an important role in the scattering of nonlinear waves.
These waves maintain a stable form. Due to the dynamic balance and non-linearity of this equation, an approximate solution has been presented in many articles [1, 3]. In this paper, we examine the symmetry group of the Kuper-Schmidt equation.
This equation is in the following form:

$$u_t = u_{xxxxx} + 10\, u_{xxx}\, u + 25\, u_{xx}\, u_x + 20\, u^2\, u_x \tag{1}$$

Lie's symmetry method was first proposed by Sophus Lie in the middle of the 19th century and then attracted the attention of mathematicians and researchers [2].
In fact, the most important and practical method to solve the differential equation is the symmetric or classical method. In this article, we want to obtain new solutions to the Kuper-Schmidt equation using the lie symmetry method. As it is known, the symmetry method plays an important role in the analysis of differential equations.
Lie's method is a method to find the solution of the differential equation with the help of its symmetry group. These solutions are called group invariant solutions and obtained by solving the reduced system of differential equation

*Talker
Email addresses: m.jafarii@pnu.ac.ir (Mehdi Jafari), Almhdyadlh@gmail.com (Adele Alemahdi)

having fewer independent variables than the original system.In this paper we use this method to compute the invariant solutions of Kuper-Schmidt equation and classify them.
In section 2, with the help of Lie's method, we express the symmetries of the Kuper-Schmidt equation and obtain its symmetry groups. In section 3, we make an optimal system of one-dimensional sub-algebras of the Kuper-Schmidt equation.

2. Lie symmetries of Kuper-Schmidt equation

In this part, we get to know classical symmetries of nonlinear differential equation with partial derivatives and its calculation method.

Definition 2.1. Suppose $\Delta\ = 0$ be a system of differential equation with p as the independent Variabl $x = (x^1\,, \ldots, x^p)$ and q as the dependent variable $u = (u^1\,, \ldots, u^q)$.
A solution for this system is a function of the form $u = f(x)$ where;

$$u^\alpha = \ f^\alpha\ \left(x^1\,, \ldots, x^p\right), \qquad \alpha = 1, \ldots, q \tag{2}$$

Are smooth function.

Definition 2.2. A symmetry group for system differential equation $\Delta\ = 0$, is a locally group of transformations such as G acting on an open subset of E such as O in such a way that it transforms any solution of the system into another solution.
In other words, if $u = f(x)$ be an answer to system, to every $g \in G$ where f and g can be defined, $g.f(x)$ be solution too.

Theorem 2.3 (Invariance theorem of differential equations)**.** *suppose $\Delta\ = 0$ be a system of differential equations of maximal rank defined on the open subset O of E, and suppose G be locally group of transformations where action on O and v be its infinitesimal generator. In this case $\Delta\ = 0$, accepts the group G as a symmetry group, if $v^n\ (\Delta) = 0$ every that $\square = 0$ [5].*

meaning of v^n is the prolong of the nth order the vector field V is on the jet space and defined as follows:

Definition 2.4. suppose :

$$v = \sum_{i=1}^{p} \xi^i(\ x\ ,\ u)\ \frac{\partial}{\partial x^i} + \sum_{\alpha=1}^{q} \emptyset^\alpha\ (x\ ,\ u)\ \frac{\partial}{\partial u^\alpha} \tag{3}$$

be a vector field on open subset O of the set E. The nth order prolong of V is :

$$v^{(n)} = v + \sum_{\alpha=1}^{q}\sum_{J} \emptyset_J^\alpha\ \left(x,\ u^{(n)}\right) \frac{\partial}{\partial u_J^\alpha} \tag{4}$$

Where the coefficients are calculated as follows:

$$\boldsymbol{\emptyset_J^\alpha(\,\mathbf{x}\,,\,u^n) = D_J\ Q_\alpha + \sum_{i=1}^{p} \xi^i u_{J,i}^\alpha} \tag{5}$$

Where J is multiple index with condition $1 \leq j_k \leq p\,, 1 \leq k \leq n$ [5].

In this section , we calculate the symmetries of the Kuper-Schmidt equation using theabove theorem. We remind you that equation (1) is a differential equation with partialderivatives that has a dependent variable u and two independent variables x and t. So the general form of the symmetry algebra generators of this equation is as follows:

$$\mathbf{v} = \boldsymbol{\tau}(\mathbf{x}, \mathbf{t}, \mathbf{U})\boldsymbol{\partial_t} + \boldsymbol{\xi}(\mathbf{x}, \mathbf{t}, \mathbf{U})\boldsymbol{\partial_x} + \boldsymbol{\emptyset}(\mathbf{x}, \mathbf{t}, \mathbf{U})\ \boldsymbol{\partial_U} \tag{6}$$

Since the order of the equation is (5), we must apply the prolong of the fifth order of the above field to the equation and set it equal to zero to obtain the coefficients τ, $\boldsymbol{\xi}$, $\emptyset$.

We have to do this with the help of the maple :

$$\begin{aligned}
\emptyset &= c_1\, u \\
\tau &= -\frac{5\, c_1\, t}{2} + c_2 \\
\xi &= -\frac{c_1\, x}{2} + c_3
\end{aligned}$$

Where c_1 , c_2 and c_3 are constant.
The infinitesimal generators of any Lie group are the symmetry parameters. In Kuper-schmidt equation the linear combination of vector fields is as follows:

$$v_1 = -\frac{x\partial_x}{2} - \frac{5t\partial_t}{2} + \mathbf{u}\partial_u \;\;,\;\; v_2 = \partial_t \;\;,\;\; v_3 = \partial_u \tag{7}$$

The commutation relations between vector fields is given by the following table, the entry in row i and column j $[v_i\ ,\ v_j]$:

Table 1. Commutator table.

$[,]$	v_1	v_2	v_3
v_1	0	$\frac{5v_2}{2}$	$\frac{v_3}{2}$
v_2	$-\frac{5v_2}{2}$	0	0
v_3	$-\frac{v_3}{2}$	0	0

By calculating the flow corresponding to the vector fields above, the one-parameter group $g.k(s)$ generated by b is as follows:

$$\begin{aligned}
g_1\,(s) = (x,t,u) &\longrightarrow (\mathbf{x}e^{-\frac{\varepsilon}{2}}\,,\, t\, e^{-\frac{5\varepsilon}{2}}\,,\, u\, e^{\varepsilon}) \\
g_2\,(s) = (x,t,u) &\longrightarrow (\mathbf{x}\,,\, \varepsilon + t\,,\, u) \\
g_3\,(s) = (x,t,u) &\longrightarrow (\varepsilon + x\,,\, t\,,\, u)
\end{aligned} \tag{8}$$

Therefore, according to the definition of the symmetry group and the above results, the following theorem can be concluded;

Theorem 2.5. *If $f = f(x,t)$ be a solution of equation (1), then The following functions will also be the solutions of this equation,*

$$\begin{aligned}
g_1\,(s)\cdot \mathbf{f} &= \mathbf{f}(\frac{x}{e^{-\frac{\varepsilon}{2}}}\,,\, \frac{t}{e^{-\frac{5\varepsilon}{2}}}\,)\, e^{\varepsilon} \\
g_2\,(s)\cdot \mathbf{f} &= \mathbf{f}(\mathbf{x}\,,\, \mathbf{t} -\ \varepsilon) \\
g_3\,(s)\cdot \mathbf{f} &= \mathbf{f}(\mathbf{x} -\ \varepsilon\,,\, \mathbf{t})
\end{aligned} \tag{9}$$

3. One-dimensional optimal system of subalgebras for the Kuper-Schmidt equation

In this section, we compute the one-dimensional optimal system of Kuper-Schmidt equation by using symmetry group. We know that the linear combination of infinitesimal symmetries is infinitesimal symmetry. Therefore, every group has like G countless one-dimensional subgroups.Therefore,it is very important to know which subgroups result in different solutions. So we look for solutions that do not transform into each other by symmetry transformations. For this, we need the concept of subalgebra optimal system [4]. The optimal system under algebras is determined by taking a representative from each class.

Adjoint representation of each X_i, $i = 1, \ldots, 3$ is defined as follow:

$$\mathbf{Ad}\big(\exp(\varepsilon \mathbf{v})\big)\boldsymbol{W_0} = \boldsymbol{W_0} - \varepsilon\,[\,\boldsymbol{V}\,,\boldsymbol{W_0}\,] + \frac{\varepsilon^2}{2}[\mathbf{V}\,,[\,\mathbf{V}\,,\boldsymbol{W_0}]] - \ldots\,. \tag{10}$$

Where s is a parameter and $[V\ ,\ W_0]$ is the commutator of the Lie algebra for $i\ ,\ j = 1, \ldots, 3$ [4].
Using the commutator table, we obtain all the e adjoint representations related to the lie group of the Kuper-Schmidt equation. This table is displayed in section 1 and table 2.
The lie algebra spanned by:

$$\boldsymbol{v_1} = -\ \frac{\boldsymbol{x\partial_x}}{\mathbf{2}} - \frac{\mathbf{5}\boldsymbol{t\partial_t}}{\mathbf{2}} + \mathbf{u}\boldsymbol{\partial_u} \quad , \quad \boldsymbol{v_2} = \boldsymbol{\partial_t} \quad , \quad \boldsymbol{v_3} = \boldsymbol{\partial_u}$$

Generate the symmetry group of the Kuper-Schmidt equation to compute the adjoint representation. we use the lie series the following:

$$\begin{aligned}\mathbf{Ad}\big(\exp(\varepsilon \mathbf{v})\big)\boldsymbol{W_0} &= \sum_{n=0}^{\infty}\frac{\varepsilon^n}{n!}(\boldsymbol{ad\,V})^n(\boldsymbol{W_0})\\ &= \boldsymbol{W_0} - \varepsilon\,[\,\boldsymbol{V}\,,\boldsymbol{W_0}\,] + \frac{\varepsilon^2}{2}[\mathbf{V},[\mathbf{V}\,,\boldsymbol{W_0}\,]] - \ldots\,.\end{aligned} \tag{11}$$

In conjunction with the commutator table (in table 1), in this manner, we construct the table 2:

Table 2. The adjoint representation table of the infinitesimal generators X_i.

Ad	$\boldsymbol{v_1}$	$\boldsymbol{v_2}$	$\boldsymbol{v_3}$
$\boldsymbol{v_1}$	$\boldsymbol{V_1}$	$e^{-\frac{5}{2}\varepsilon}\,\boldsymbol{V_2}$	$\mathbf{e}^{-\frac{\varepsilon}{2}}\,\mathbf{V_3}$
$\boldsymbol{v_2}$	$e^{\frac{5}{2}\varepsilon}\,\boldsymbol{V_1}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$
$\boldsymbol{v_3}$	$\mathbf{e}^{\frac{\varepsilon}{2}}\,\mathbf{V_3}$	$\boldsymbol{V_2}$	$\boldsymbol{V_3}$

With the (i, j)-th entry indicating $Ad\big(\exp(\varepsilon\ V_i)\big)V_j$.

$$\boldsymbol{T_1}\,,\boldsymbol{T_2}\,,\boldsymbol{T_3} = \begin{bmatrix}1 & 0 & 0\\ 0 & e^{-\frac{5\varepsilon}{2}} & 0\\ 0 & 0 & e^{-\frac{\varepsilon}{2}}\end{bmatrix} \quad , \quad \begin{bmatrix}1 & 0 & 0\\ \frac{5\varepsilon}{2} & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \quad , \quad \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ \frac{\varepsilon}{2} & 0 & 1\end{bmatrix}$$

Now we can state the following theorem:

Theorem 3.1. *A one-dimensional optimal system for Lie algebra of Kuper-Schmidt equation is given by:*

$$\mathbf{1)}\ \mathbf{V_1} \quad , \quad \mathbf{2)}\ \mathbf{V_2} \quad , \quad \mathbf{3)}\ \mathbf{V_3} \quad , \quad \mathbf{4)}\ \mathbf{V_1} \pm \mathbf{V_2} \tag{12}$$

4. Conclusion

In this paper we analyzed the symmetries of one of the important equations called Kuper-schmidt in engineering sciences and medicine. It was proved that the algebra of symmetries of this equation is produced by three generators. In addition, by using the adjoint representation of the symmetry group in algebra, we built an optimal system of one-dimensional subalgebras.

References

[1] Xianjuan Li and Tao Tang. Convergence analysis of jacobi spectral collocation methods for abel-volterra integral equations of second kind. *Frontiers of Mathematics in China*, 7:69–84, 2012.

[2] Sophus Lie. On integration of a class of linear partial differential equations by means of definite integrals. *Arch. Math*, 6(3):328–368, 1881.

[3] Micheline Musette and Caroline Verhoeven. Nonlinear superposition formula for the kaup–kupershmidt partial differential equation. *Physica D: Nonlinear Phenomena*, 144(1-2):211–220, 2000.

[4] Peter J Olver. *Applications of Lie groups to differential equations*, volume 107. Springer Science & Business Media, 1993.

[5] Peter J Olver. *Equivalence, invariants and symmetry*. Cambridge University Press, 1995.

[6] Abdul-Majid Wazwaz. Generalized boussinesq type of equations with compactons, solitons and periodic solutions. *Applied mathematics and computation*, 167(2):1162–1178, 2005.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–221, pp. 488–492

HA-contractibility of simple closed HA-curves

Tayyebe Nasri[a,*]

[a]*Department of Pure Mathematics, Faculty of Basic Sciences, University of Bojnord, Bojnord, Iran.*

Article Info

Keywords:
Marcus–Wyse topology
HA-countractible
HA-homotopy

2020 MSC:
54A10
55P99
68U10

Abstract

In 2017, the H-topology was developed in $\mathbb{Z}^n$, $n \in \mathbb{N}$, as an Alexandroff topological space and a generalization of the Marcus–Wyse topology. Based on the H-topology, several concepts were established such as an H-adjacency derived from an H-topological space, HA-spaces, an HA-map, and an HA-isomorphism between two HA-spaces. Then the author established an HA-homotopy on an HA-space, which can contributed to the classification of digital images in $\mathbb{Z}^n$. In the present paper, we define the concept of an HA-contactible space in the category of HAC and investigate HA-contractiblity of simple closed HA-curves $SC_{HA}^{n,l}$.

1. Introduction

It is well known that the study of 2D digital spaces plays an important role in digital geometry related to the fields of mathematical morphology, computer graphics, image analysis, image processing and so forth. In [2] established two maps such as an MA-map and an MA-isomorphism based on Marcus–Wyse topology and studied its properties. Also he defined the conception of MA-homotopy between two MA-maps and MA-contractibility of a space X in MAC. Han [3] also developed a new topology on $\mathbb{Z}^n$, $n \in \mathbb{N}$, that is called the H-topology and investigated its properties. Also, he introduced the notions of an HA-continuous map and an HA-homeomorphism. The Author [5] introduced the HA-homotopy between two HA-maps. In this paper, we define the concept of an HA-contactible space in the category of HAC and investigate HA-contractiblity of simple closed HA-curves with l elements, $SC_H^{n,l}$. We prove that an HA-connected proper subset of $SC_{HA}^{n,l}$ is HA-contractible, for all $l \geq 4$. Also, every $SC_{HA}^{n,2l}$ is not HA-contractible if $l \geq 3$.

2. Preliminaries

Recall from [2] some of the main concepts from the Marcus–Wyse topology (M-topology, for brevity) on $\mathbb{Z}^2$. To make the paper self-contained, we need to mention some essential notions from M-topology as follows:
The M-topology on $\mathbb{Z}^2$, denoted by $(\mathbb{Z}^2, \gamma^2)$ in the present paper, is induced by the set of all U in (2.1) as a base. For

---

*Talker
Email address: t.nasri@ub.ac.ir (Tayyebe Nasri)

each point $p = (x, y) \in \mathbb{Z}^2$, we define the set $\{U(p) \,|\, p \in \mathbb{Z}^2\}$ in (2.1) below as a base [2], where for each point $p = (x, y) \in \mathbb{Z}^2$

$$U(p) := \begin{Bmatrix} N_4(p) \cup \{p\} \text{ if } x + y \text{ is even, and} \\ \{p\} : \text{ else.} \end{Bmatrix} \tag{2.1}$$

where $N_4(p) = \{(x \pm 1, y), (x, y \pm 1)\} = \{q \,|\, q$ is 4-adjacent to $p\}$. In relation to the further statement of a point in $\mathbb{Z}^2$, a point $p = (x_1, x_2)$ is called to be *doubly even* if each x_i is even, $i \in \{1, 2\}$; *even* if each x_i is odd, $i \in \{1, 2\}$; and *odd* if $x_1 + x_2$ is an odd number. For a set $X \subseteq \mathbb{Z}^2$, one can take the subspace induced by $(\mathbb{Z}^2, \gamma^2)$, denoted by (X, γ^2_X), which has been often studied in the context of digital images.
Hereinafter, given a space (X, γ^2_X), the set X with an M-adjacency is said to be an MA-space. In other words, we may recognize the space (X, γ^2_X) with an M-adjacency as an MA-space. For two MA-spaces $X := (X, \gamma^2_X)$ and $Y := (Y, \gamma^2_Y)$, a function $f : X \longrightarrow Y$ is said to be an MA-*map* at a point $x \in X$ if $f(MN(x)) \subseteq MN(f(x))$. Furthermore, a map $f : X \longrightarrow Y$ is said to be an MA-map if it is an MA-map at every point $x \in X$.
Using MA-maps, the paper [2] established the MA-topological category, denoted by MAC, consisting of two sets.
(1) The set of objects (X, γ^2_X) with an M-adjacency,
(2) For every ordered pair of objects $X := (X, \gamma^2_X)$ and $Y := (Y, \gamma^2_Y)$, the set of all MA-maps $f : X \longrightarrow Y$ as morphisms.
For two MA-spaces X and Y, a map $f : X \longrightarrow Y$ is called an MA-*isomorphism* if f is a bijective MA-map and furthermore, $f^{-1} : Y \longrightarrow X$ is an MA-map. The paper [2, Theorem 6.3] showed that $SC^{l_1}_{MA}$ is MA-isomorphic to $SC^{l_2}_{MA}$ if and only if $l_1 = l_2$.

Let us now recall some basic concepts of the generalized Marcus–Wyse topology on $\mathbb{Z}^n$, $n(\geq 3) \in \mathbb{N}$ from [3]:
On $\mathbb{Z}^n$, $n \geq 3$, the H-topology, denoted by $(\mathbb{Z}^n, \gamma^n)$, was defined as the product topology γ^n induced by the topologies $(\mathbb{Z}^{n-1}, \gamma^{n-1})$, $n \geq 3$, and $(\mathbb{Z}, \gamma)$, where $(\mathbb{Z}, \gamma)$ is the subspace $(\mathbb{Z}, \gamma^2_{\mathbb{Z}})$. Namely, we obtain

$$(\mathbb{Z}^n, \gamma^n) = (\mathbb{Z}^{n-1}, \gamma^{n-1}) \times (\mathbb{Z}, \gamma),\ n \geq 3. \tag{3.1}$$

For instance, we have $(\mathbb{Z}^3, \gamma^3) = (\mathbb{Z}^2, \gamma^2) \times (\mathbb{Z}, \gamma)$, $(\mathbb{Z}^4, \gamma^4) = (\mathbb{Z}^3, \gamma^3) \times (\mathbb{Z}, \gamma)$, and so on. Hereinafter, $(\mathbb{Z}^n, \gamma^n)$, $n \in \mathbb{N}$, is called an H-topological space in this paper. Thus the space $(\mathbb{Z}^n, \gamma^n)$ is a proper subspace of $(\mathbb{Z}^{n+1}, \gamma^{n+1})$ with the relative topology on $\mathbb{Z}^n$ induced by $(\mathbb{Z}^{n+1}, \gamma^{n+1})$, $n \in \mathbb{N}$. In the present paper, in every subspace of $(\mathbb{Z}^n, \gamma^n)$, a black jumbo dot means a point whose all coordinates are odd, the symbols ■ means a point in $\mathbb{Z}^n$ whose all coordinates are even, and • means a mixed point.
For $X \subset \mathbb{Z}^n$, assuming a subspace $(X, \gamma^n_X) := (X, (\gamma^n)_X)$ induced from $(\mathbb{Z}^n, \gamma^n)$, we define an H-adjacency on X [3] as follows:

$$\text{Two distinct points } x \text{ and } y \text{ in } X \text{ are } H\text{-adjacent if } y \in O_H(x) \text{ or } x \in O_H(y), \tag{1}$$

where $O_H(q)$ is the smallest open set of the point $q \in (X, \gamma^n_X)$, $q \in \{x, y\}$. Given a subspace (X, γ^n_X), we consider X with an H-adjacency. Then we call it an HA-space and denote it by just X instead of "X with an H-adjacency". Besides, motivated by the adjacency neighborhood in $(\mathbb{Z}^2, \gamma^2)$ [2], Han [3] established an H-adjacency neighborhood of a given point $x \in (X, \gamma^n)$ as follows:

$$HN(x) := HA(x) \cup \{x\}, \tag{2}$$

where $HA(x) = \{q \,|\, q$ is H-adjacent to $x\}$.

Definition 2.1. [3] We say that a simple closed HA-curve with l elements, denoted by $SC^{n,l}_{HA}$, $l \geq 4$, is a set $(x_i)_{i \in [0, l-1]_{\mathbb{Z}}}$ such that x_i is H-adjacent to x_j if and only if either $j = i + 1(mod\ l)$ or $i = j + 1(mod\ l)$ (or $|\, i - j\,| = \pm 1(mod\ l)$).

for two HA-spaces $(X, \gamma^{n_1}_X) = X$ and $(Y, \gamma^{n_2}_Y) = Y$, a function $f : X \longrightarrow Y$ is said to be an HA-*map* at a point $x \in X$ if

$$f(HN(x)) \subseteq HN(f(x)). \tag{3}$$

Furthermore, a map $f : X \longrightarrow Y$ is called an HA-map if it is an HA-map at every point $x \in X$.
Based on these notions, an HA-space X is called HA-connected [3] if for any two points x and y in X, there is an HA-map $f : [0, m]_{\mathbb{Z}} \to X$ such that $f(0) = x$ and $f(m) = y$.

Using HA-maps, in [3], it was established the HA-category, denoted by HAC, consisting of two sets, as follows:
(1) The set of HA-spaces as objects for $X \subseteq \mathbb{Z}^n, n \in \mathbb{N}$.
(2) For every ordered pair of objects X and Y, the set of all HA-maps $f : X \longrightarrow Y$ as morphisms of HAC.
For two HA-spaces $X \subset \mathbb{Z}^{n_1}$ and $Y \subset \mathbb{Z}^{n_2}$, a map $f : X \longrightarrow Y$ is called an HA-*isomorphism* if f is a bijective HA-map and $f^{-1} : Y \longrightarrow X$ is an HA-map.
It was proved that SC^{n,l_1}_{HA} is HA-isomorphic to SC^{n,l_2}_{HA} if and only if $l_1 = l_2$ [3, Theorem 5.1]. Furthermore, owing to (3), for $SC^{n_i,l_i}_{HA}, i \in \{1, 2\}$, even if $n_1 \neq n_2$, we obtain the following remark.

Remark 2.2. SC^{n_1,l_1}_{HA} is HA-isomorphic to SC^{n_2,l_2}_{HA} if and only if $l_1 = l_2$.

3. Main Results

Let $X \in Ob(HAC)$ and let B be a subset of X. Then (X, B) is called a space pair in HAC. Furthermore, if B is a singleton set $\{x_0\}$, then (X, x_0) is called a pointed space in HAC. We establish the notions of an HA-homotopy relative to a subset $B \subseteq X$ and a pointed HA-contractible, which will be used for studying spaces in HAC.

Definition 3.1. Let (X, B) and Y be a space pair and a space in $Ob(HAC)$, respectively. Let $f, g : X \longrightarrow Y$ be HA-maps. Suppose that there exist $m \in \mathbb{N}$ and

$$\text{a function } F : X \times [0, m]_{\mathbb{Z}} \longrightarrow Y \tag{4.1}$$

such that
(i) for all $x \in X$, $F(x, 0) = f(x)$ and $F(x, m) = g(x)$;
(ii) for all $x \in X$, the induced function $F_x : [0, m]_{\mathbb{Z}} \longrightarrow Y$ given by $F_x(t) = F(x, t)$ for all $t \in [0, m]_{\mathbb{Z}}$ is an HA-map;
(iii) for all $t \in [0, m]_{\mathbb{Z}}$, the induced function $F_t : X \longrightarrow Y$ given by $F_t(x) = F(x, t)$ for all $x \in X$ is an HA-map.
Then we say that F is an HA-homotopy between f and g.
Furthermore, for all $t \in [0, m]_{\mathbb{Z}}$, if $F_t(x) = f(x) = g(x)$ for all $x \in B$, then we call F an HA-homotopy relative to B between f and g, and we say that f and g are HA-homotopic relative to B in Y, denoted by $f \simeq_{HArel.B} g$.

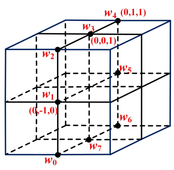


Fig. 1. Configuration of $SC^{3,8}_{HA}$.

Example 3.2. Consider the space $X := SC^{3,8}_{HA} := (w_i)_{i\in[0,7]_{\mathbb{Z}}}$ in Figure **??**. Then one click transformation on X is a kind of an HA-homotopy of X. Consider the map $F : X \times [0, 1]_{\mathbb{Z}} \longrightarrow X$ given by $F(w_i, 0) = w_i$ and $F(w_i, 1) = w_{i+1(mod\ 8)}$. Since $HN(w_i) = \{w_{i-1(mod\ 8)}, w_i, w_{i+1(mod\ 8)}\}$, we have $F_t(HN(w_i)) = HN(F_t(w_i))$ and $F_{w_i}(HN(t)) \subseteq HN(F_{w_i}(t))$, for all $t \in \{0, 1\}$ and $i \in [0, 7]_{\mathbb{Z}}$. Therefore, F satisfies the conditions of Definition 3.1.

Definition 3.3. Let $X \in Ob(HAC)$ and let $x_0 \in X$. If 1_X is HA-homotopic to the constant map in the space $\{x_0\}$ relative to $\{x_0\}$, then we say that (X, x_0) is a pointed HA-contractible. (for brevity HA-contractible if there is no danger of ambiguity).

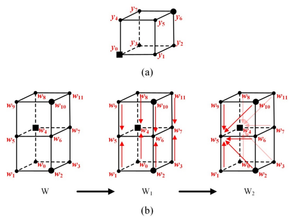


Fig. 2. (a) Configuration of Y used in Lemma 3.4(1); (b) HA-contractibility of W.

Now we exhibit some examples for an HA-contractible space.

Example 3.4. (1) The space $Y := ([0,1]_{\mathbb{Z}})^3$ in Figure 2(a) is an HA-contractible space.
(2) The space $W := X \times [-1,1]_{\mathbb{Z}}$, where $X = \{(0,0),(1,0),(0,1),(1,1)\}$ in Figure 2(b) is HA-contractible.

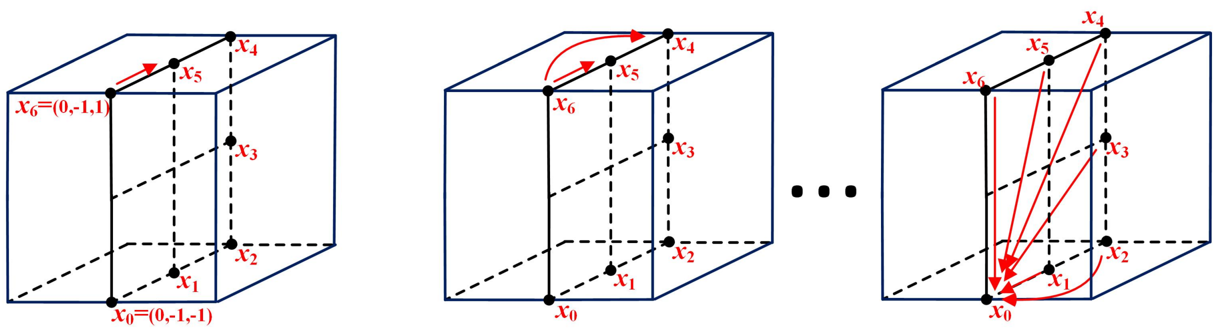


Fig. 3. (a) Configuration of Y used in Lemma 3.4(1); (b) HA-contractibility of W.

Example 3.5. Consider the space $X = (x_i)_{i\in[0,6]_{\mathbb{Z}}} \subseteq SC^{3,8}_{HA}$ defined in Figure 3 starting at $x_0 = (0,-1,-1)$ and finishing at $x_6 = (0,-1,1)$. To guarantee the assertion, let us now consider the following H-adjacency neighborhood:

$$\begin{aligned} HN(x_0) &= \{x_0, x_1\}, & HN(x_1) &= \{x_0, x_1, x_2\},\\ HN(x_2) &= \{x_1, x_2, x_3\}, & HN(x_3) &= \{x_2, x_3, x_4\},\\ HN(x_4) &= \{x_3, x_4, x_5\}, & HN(x_5) &= \{x_4, x_5, x_6\},\\ HN(x_6) &= \{x_5, x_6\}. \end{aligned}$$

We can establish an HA-homotopy F on X relative to the set $\{x_0\}$ between 1_X and the constant map at $\{x_0\}$. Consider the map $F : X \times [0,6]_{\mathbb{Z}} \longrightarrow X$ by (see Figure 3)

$$\begin{cases} F(x_i,0) = x_i & \text{if } i \in [0,6]_{\mathbb{Z}},\\ F(x_i,1) = x_i & \text{if } i \in [0,5]_{\mathbb{Z}} \text{ and } F(x_6,1) = x_5,\\ F(x_i,2) = x_i & \text{if } i \in [0,4]_{\mathbb{Z}} \text{ and } F(x_i,2) = x_4 \text{ if } i \in \{5,6\},\\ \vdots & \\ F(x_i,6) = x_0 & \text{if } i \in [0,6]_{\mathbb{Z}}. \end{cases}$$

Therefore, X is an HA-contractible space.

Using the process of the above example, we obtain the following result.

Theorem 3.6. *An HA-connected proper subset of $SC^{n,l}_{HA}$ is HA-contractible, $l \geq 4$.*

Proof. Put $X_{HA}^{l'} := (x_i)_{i\in[0,l'-1]_{\mathbb{Z}}}$ as an HA-connected proper subset of $SC_{HA}^{n,l}$, where $l' \leq l-1$. Indeed, the set $X_{HA}^{l'}$ is a kind of simple HA-path with l' elements in $SC_{HA}^{n,l}$. Consider the map $F : X_{HA}^{l'} \times [0, l'-1]_{\mathbb{Z}} \longrightarrow X_{H}^{l'}$ by

$$\begin{cases} F(x_i, 0) = x_i & \text{if } \ i \in [0, l'-1]_{\mathbb{Z}}, \\ F(x_i, 1) = x_i & \text{if } \ i \in [0, l'-2]_{\mathbb{Z}} \ \text{ and } \ F(x_i, 1) = x_{l'-2} \ \text{ if } \ i = l'-1, \\ F(x_i, 2) = x_i & \text{if } \ i \in [0, l'-3]_{\mathbb{Z}} \ \text{ and } \ F(x_i, 2) = x_{l'-3} \ \text{ if } \ i \in \{l'-1, l'-2\}, \\ \vdots & \\ F(x_i, l'-1) = x_0 & \text{if } \ i \in [0, l'-1]_{\mathbb{Z}}, \end{cases}$$

which implies that the given map F is an HA-homotopy on $X_{HA}^{l'}$ relative to the set $\{x_0\}$ between the identity map $1_{X_{HA}^{l'}}$ and the constant map at $\{x_0\}$. □

Corollary 3.7. *Every $SC_{HA}^{n,2l}$ is not HA-contractible if $l \geq 3$.*

References

[1] U. Eckhardt and L.J. Latecki, Topologies for the digital spaces $\mathbb{Z}^2$ and $\mathbb{Z}^3$, *Computer Vison and Image Understanding*, 90 (2003), 295–312.

[2] S.E. Han, Generalizations of continuity of maps and homeomorphisms for studying 2D digital topological spaces and their applications, *Topology Appl.* 196 (2015), 468–482.

[3] S.E. Han, Topological Graphs Based on a new Topology on $\mathbb{Z}^n$ and its Applications, *Filomat* 31:20 (2017), 6313–6328.

[4] S.E. Han and W. Yao, Homotopy based on Marcus–Wyse topology and its applications, *Topology Appl.* 201 (2016), 358–371.

[5] T. Nasri, Homotopy based on generalized Marcus–Wyse topology, In Proceedings of the 32rd seminar on mathematical analysis and its applications, 2018.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–225, pp. 493–499

Symmetry Reduction and One-Dimensional Optimal System of the Hunter-Saxton Equation

Mehdi Jafari[a], Sajiyeh Safaei nejad[b,*]

[a] *Corresponding author. Department of mathematics, payame Noor university, PO BOX 19395-3697, Tehran, Iran.*
[b] *Department of mathematics, payame Noor university, PO BOX 19395-3697, Tehran, Iran.*

Article Info

Keywords:
Hunter-Saxton equation
Optimal System
Group-invariant solutions

Abstract

In this paper, we investigate the symmetry group of the Hunter-Saxton equation applying the classical Lie symmetry methods. Also, by utilizing the classification of one-dimensional subalgebras of the symmetry algebra of this equation, we compute the optimal system of one-parameter subalgebras, then by using this optimal system and differential invariants, we reduce the equation and obtain the group-invariant solutions.

1. Introduction

Differential equations play an important role in modeling physical problems, particularly differential equations with partial derivatives have a special role, for example, one of the most important equations used in physics and mathematics is the Hunter- Saxton (H-S) equation. This equation is used to modeling of nematic liquid crystals and it is a second order equation which is defined as follows.

$$(u_t + uu_x)_x = \frac{1}{2}u^2{}_x,$$

The H-S equation that is a well known nonlinear hyperbolic in mathematical model which is described by the partial differential equation (PDE) [9]. Which has two independent variables $x = \left(x^1, x^2\right) = (t, x)$ and one dependent variable $u = u\,(t, x)$. This equation has been first suggested by Hunter and Saxton for the theoretical modeling of nematic liquid crystals [3]. Yao et al [12] tackled the periodic H-S equation, introducing a variable coefficient by employing the classical approach to finding invariant solutions. Johnpillai and Khaliquo [8] also used Lie symmetry analysis to find exact solutions for yet another generalized version of the H-S equation. Usually, two linearly independent vector fields are needed for the complete description of nematic liquid crystals [1]. One for characterizing the fluid flow and one for describing the orientation of the molecules which is the so-called director field [3]. Also order of the sections of this article is as follows. In section 2, we have referred to some definitions and theorem that are used in the later sections. In section 3, we calculate the lie algebra of infinitesimal symmetries, the one-parameter groups generated and G-invariant solutions of H-S equation. In section 4, we obtain the optimal system of one-dimensional sub-algebras

*Talker
Email addresses: m.jafarii@pnu.ac.ir (Mehdi Jafari), hivacaren@gmail.com (Sajiyeh Safaei nejad)

of H-S equation using the classification of orbits of adjoint representation of the symmetry group. In section 5, we find the reduced equation and group invariant solutions for each element of the optimal system. In section 6, we draw conclusion.

2. Definitions and Theorem of infinitesimal criterion

We use definitions 1 and 2 in section 3 and definitions 3 and 4 and theorem in section 4.

Definition 2.1. A system s of n-th order differential equations in p independent and q dependent varables is given as a system of equations,

$$\Delta_v\left(x, u^{(n)}\right) = 0, \qquad v = 1, \cdots, l$$

Involving $x = \left(x^1, \cdots, x^p\right), u = \left(u^1, \cdots, u^q\right)$ and the derivatives of with respect to x up to order n. The functions $\Delta\left(x, u^{(n)}\right) = \left(\Delta_1\left(x, u^{(n)}\right), \cdots, \Delta_l\left(x, u^{(n)}\right)\right)$ will be assumed to be smooth in their arguments, so Δ can be viewed as a smooth map from the jet space $X \times U^{(n)}$ to some l-dimensional Euclidean space

$$\Delta : X \times U^{(n)} \to \mathbb{R}^l$$

The solutions of this differential equation are where the given mapping Δ vanishes on $X \times U^{(n)}$, and thus determine a subvariety [10]

$$\mathrm{s}_\Delta = \left\{\left(x, u^{(n)}\right) : \Delta\left(x, u^{(n)}\right) = 0\right\} \subset X \times U^{(n)}$$

Definition 2.2. Let $M \subset X \times U$ be open and suppose v is a vector field on M with corresponding (local) one-parameter group $\exp(\varepsilon v)$ the n-th prolongation of v denoted $pr^{(n)}v$ will be avector field on the n-jet space $M^{(n)}$ and is defined to be the infinitesimal generator of the corresponding prolonged one-parameter group $pr^{(n)}[\exp(\varepsilon v)]$ in other words, we have

$$pr^{(n)}v\Big|_{\left(x,u^{(n)}\right)} = \frac{d}{d_\varepsilon}|_{\varepsilon=0}\, pr^{(n)}\left[\exp(\varepsilon v)\right]\left(x, u^{(n)}\right),$$

For any $\left(x, u^{(n)}\right) \in M^{(n)}$ [3].

Definition 2.3. Let G be a lie group. For each $g \in G$, group conjugation $K_g(h) \equiv ghg^{-1}, h \in G$, determines a diffeomorphism on G. Moreover, $K_g O K_{g'} = K_{gg'}, K_e = 1_G$, so K_g determines a global group action of G on itself, with each conjugacy map K_g being a group homomorphism: $K_g(hh') = K_g(h) K_g(h')$, etc. The differential $dK_g : TG|_h \to TG\big|_{K_g(h)}$ is readily seen to preserve the right-invariance of vector fields, and hence determines a linear map on the Lie algebra of G, called the adjoint representation [10]

$$Adg(v) \equiv dK_g(v), \qquad v \in \mathfrak{g}$$

Definition 2.4. Let G be a lie group. An optimal system of s-parameter subgroups is a list of conjugacy inequivalent s-parameter subgroups with the property that any other subgroup is conjugate to precisely one subgroup in the list. Similarly, the optimal system for subalgebras is also defined in the same way [10]

Theorem 2.5. *Suppose*

$$\Delta_v\left(x, u^{(n)}\right) = 0, \qquad v = 1, \cdots, l,$$

Is a system of differential equations of maximal rank defined over $M \subset X \times U$ if G is a local group of transformations acting on M , and

$$pr^{(n)}v\left[\Delta_v\left(x, u^{(n)}\right)\right] = 0, \quad v = 1, \cdots, l \quad whenever \quad \Delta\left(x, u^{(n)}\right) = 0,$$

For every infinitesimal generator v of G, then G is a symmetry group of the system [11]. (Infinitesimal criterion)

3. Lie symmetry group for the H-S equation

We consider the second order PDE,

$$\Delta\left(x,t,u^{(2)}\right) = u_{tx} + \frac{u^2{}_x}{2} + uu_{xx},$$

And let,

$$v = \eta(x,t,u)\frac{\partial}{\partial_x} + \tau(x,t,u)\frac{\partial}{\partial_t} + \phi(x,t,u)\frac{\partial}{\partial_u},$$

Be a vector field on $X \times U$.
We want to determine all possible coefficient functions η, τ, ϕ so that the corresponding one-parameter group exp (εv) is a symmetry group of the H-S equation [6, 7]. According to theorem, we need to know the second prolongation of v

$$pr^{(2)}v = v + \phi^x\frac{\partial}{\partial u_x} + \phi^t\frac{\partial}{\partial u_t} + \phi^{xx}\frac{\partial}{\partial u_{xx}} + \phi^{tt}\frac{\partial}{\partial u_{tt}} + \phi^{xt}\frac{\partial}{\partial u_{xt}}.$$

Applying $pr^{(2)}v$, we find the infinitesimal criterion must be satisfied whenever.

$$u_{tx} = -\frac{u^2{}_x}{2} - uu_{xx},$$

Replacing u_{tx} by $-\frac{u^2{}_x}{2} - uu_{xx}$. Equating the coefficient of the various monomials in the first and second order partial derivatives of u, we find the determining equations for the symmetry group of the H-S equation [2]. The solution of the determining equations follows by

$$\begin{aligned}\eta(x,t,u) &= f_2(t)x + f_3(t),\\ \tau(x,t,u) &= f_1(t),\\ \phi(x,t,u) &= f_2(t)u\end{aligned}$$

For linear function f_1, f_2, f_3 depend on t. Let

$$\begin{aligned}f_1(t) &= c_1t + c_2,\\ f_2(t) &= c_3,\\ f_3(t) &= c_4t + c_5,\end{aligned}$$

Where $c_1, \cdots, c_5$ are arbitrary constants [5]. We conclude that the most general infinitesimal symmetry of the H-S equation has cofficient functions as follows.

$$\begin{aligned}\eta(x,t,u) &= c_3x + c_4t + c_5,\\ \tau(x,t,u) &= c_1t + c_2,\\ \phi(x,t,u) &= c_3u,\end{aligned}$$

Thus, the lie algebra of infinitesimal symmetries of H-S equation are spanned by the five vector fields.

$$\begin{array}{ll} v_1 = t\partial_t & Dilatation\\ v_2 = \partial_t & Time\ translation\\ v_3 = x\partial_x + u\partial_u & Dilatation\\ v_4 = t\partial_x & \\ v_5 = \partial_x & Space\ translation\end{array}$$

The commutation relations between these vector fields are given by the following table, the entry in row i and column j representing $[v_i, v_j]$
The one-parameter groups f_i generated by the v_i are given in the pervious table.

	v_1	v_2	v_3	v_4	v_5
v_1	0	$-v_2$	0	v_4	0
v_2	v_2	0	0	v_5	0
v_3	0	0	0	$-v_4$	$-v_5$
v_4	$-v_4$	$-v_5$	v_4	0	0
v_5	0	0	v_5	0	0

Theorem 3.1. *The entries give the transformed point* $\exp(\varepsilon v - i)(x,t,u) = \left(\tilde{x},\tilde{t},\tilde{u}\right)$

$$\begin{aligned} f_1(t) &= (x, te^{\varepsilon}, u), \\ f_2(t) &= (x, \varepsilon + t, u), \\ f_3(t) &= (xe^{\varepsilon}, t, ue^{\varepsilon}), \\ f_4(t) &= (\varepsilon t + x, t, u), \\ f_5(t) &= (\varepsilon + x, t, u). \end{aligned}$$

Corollary 3.2. *Since each group f_i is a symmetry group, implies that if $u = f(x,t)$ is a solution of H-S equation, the following functions are also solutions.*

$$\begin{aligned} u_1 &= f\left(t, \frac{x}{e^{\varepsilon}}\right), \\ u_2 &= f(t, x - \varepsilon), \\ u_3 &= f\left(\frac{t}{e^{\varepsilon}}, x\right) e^{\varepsilon}, \\ u_4 &= f(-\varepsilon x + t, x), \\ u_5 &= f(t - \varepsilon, x), \end{aligned}$$

Where ε is any real number. Therefore, we obtained the G-invariant solutions of the H-S equation.

4. One-Dimensional Optimal System of Subalgebras for the H-S equation

In this section, we obtain the one-dimensional optimal system of H-S equation by using symmetry group. Since every linear combination of infinitesimal symmetries is an infinitesimal symmetry, there is an infinite number of one-dimensional subgroups for G. Therefore, it is important to determine which subgroups give different types of solutions. For this, we must find invariant solutions which cannot be transformed to each other by symmetry transformations in the full symmetry group [5, 7]. Adjoint representation of each $v_i,\ i = 1,\cdots,5$ is defined as follow:

$$Ad(\exp(\varepsilon v_i))\, v_j = v_j - \varepsilon [v_i, v_j] + \frac{\varepsilon^2}{2!}[v_i,[v_i,v_j]] - \cdots$$

Where ε is a parameter and $[v_i, v_j]$ is the commutator of the lie algebra for $i,j = 1,\cdots,5$. Considering the table of commutator, we can compute all the adjoint representations corresponding to the Lie group of the H-S equation. Now we state the following theorem.

Theorem 4.1. *A one-dimensional optimal system for Lie algebra of the H-S equation is given by.*

$$\begin{aligned} &(1) \quad 2v_1 + v_3, \\ &(2) \quad v_3, \\ &(3) \quad v_1, \end{aligned}$$

Proof. Let $T_i^{\varepsilon} : g \to g$ be the adjoint transformation defined by $X \to Ad\left(\exp\left(\varepsilon V_i\right)\right)V$ for $i = 1, \cdots, 5$ The matrix of $T_i^{\varepsilon},\ i = 1, \cdots, 5$, with respect to basis $[T_1, \cdots, T_5]$ is.

$$T^{\varepsilon}{}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & e^{\varepsilon} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & e^{-\varepsilon} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},\ T^{\varepsilon}{}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -\varepsilon & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -\varepsilon & 1 \end{bmatrix},\ T^{\varepsilon}{}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & e^{\varepsilon} & 0 \\ 0 & 0 & 0 & 0 & e^{\varepsilon} \end{bmatrix}$$

$$T^{\varepsilon}{}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \varepsilon & 0 & -\varepsilon & 1 & 0 \\ 0 & \varepsilon & 0 & 0 & 1 \end{bmatrix},T^{\varepsilon}{}_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -\varepsilon & 0 & 1 \end{bmatrix},$$

Now, we try to vanish the coefficients of v that given a nonzero vector:

$$v = a_1v_1 + a_2v_2 + a_3v_3 + a_4v_4 + a_5v_5,$$

By acting the adjoint representations $T^{\varepsilon}{}_i$ on v by choosing suitable parameters ε in each step. Therefor we can simplify v as follows:
Case (1): If $a_3 \neq 0$ we can assume that $a_3 = 1, \varepsilon = \dfrac{a_5}{a_3}$ with $T^{\varepsilon}{}_2$ then we can make the coefficient of v_2 vanish.
With $T^{\varepsilon}{}_5,\ \varepsilon = a_5$ then we can make the coefficient of v_5 vanish.
Case (2): If $a_1 \neq 0$ we can assume that $a_1 = 2, \varepsilon = -\dfrac{a_4}{a_1{}^{-1}}$ with $T^{\varepsilon}{}_4$ then we can make the coefficient of v_4 vanish. so v is reduced:

$$2v_1 + v_3,$$

If $a_1 = 0$ we can assume that $\varepsilon = a_4$ with $T^{\varepsilon}{}_4$ then we can make the coefficient of v_4 vanish. So v is reduced:

$$v_3,$$

If $a_3 = 0$ we can assume that $a_1 = 1, \varepsilon = \dfrac{a_5}{a_1}$ with $T^{\varepsilon}{}_2$ then we can make the coefficient of v_2 vanish.
Case (3): If $a_4 \neq 0$ we can assume that $\varepsilon = a_4$ with $T^{\varepsilon}{}_4$ then we can make the coefficient of v_4 vanish. So v is reduced:

$$v_1,$$

If $a_4 = 0$ with $T^{\varepsilon}{}_5$ we can make the coefficient of v_5 vanish. so v is reduced:

$$v_1,$$

We have found an optimal system of one-dimensional subalgebras to be those spanned by,

$$\begin{aligned} v_1 &= t\partial_t, \\ v_3 &= x\partial_x + u\partial_u, \\ 2v_1 + v_3 &= 2t\partial_t + x\partial_x + u\partial_u, \end{aligned}$$

And the proof is complete. □

5. Similarity Reduction and group-invariant solutions of H-S equation

In this section, the one-dimensional flow equation will be reduced by expressing it in the new coordinates. The H-S equation is expressed in the coordinates (x, t, u),we must search for this equations form in the suitable coordinates for reducing it. These new coordinates will be obtained by looking for independent invariants (z, w, f) corresponding to the generators of the symmetry group. For solving this equation, the following associated characteristic ODE must be solved:

$$\frac{d_x}{x} = \frac{d_t}{2t} = \frac{d_u}{2u}$$

Hence, three functionally independent invariant $z = tx^{-2}, w = tu^{-1}$ and $f = ux^{-2}$ are obtained. The reduced equations are obtained as follows:

$$\begin{aligned} ff_{zz} + \frac{{f_z}^2}{z} &= 0, \\ f_w + \frac{f^2}{2} &= 0, \\ \frac{f^2}{2} &= 0, \end{aligned}$$

The solutions to the above equations in terms of (z, w, f) variables are as follows:

$$\begin{aligned} &f(z) = \left(\frac{3}{2}(c_1 z + c_2)\right)^{\frac{2}{3}}, \\ &f(w) = \frac{2}{2c_1 + w}, \\ &f(z) = 0. \end{aligned}$$

And finally, we found group-invariant solutions of H-S equation.

$$\begin{aligned} &f\left(tx^{-2}\right) = \left(\frac{3}{2}\left(c_1 tx^{-2} + c_2\right)\right)^{\frac{2}{3}}, \\ &f\left(tu^{-1}\right) = \frac{2}{2c_1 + tu^{-1}}, \\ &f\left(tx^{-2}\right) = 0. \end{aligned}$$

6. conclusion

In this paper, by using the theory of Infinitesimal criterion and the influence of prolonged on the vector field v we compute the Lie symmetry group of H-S equation. Also, by determining equation of functions η, τ, ϕ one-parameter groups generator and optimal system of symmetry are obtained for H-S equation, five flows are acquired which two of them are dilatation, another one is time translation and the last one is space translation. Finally, we achieved to the G-invariant solutions of H-S equation under symmetry group G. Also, by using the adjoint representation of the symmetry group on its lie algebra, we have constructed an optimal system of one-dimensional subalgebras. Moreover, we have obtained the similarity reduced equations for each element of optimal system as well as its group invariant solutions of H-S equation.

References

[1] E.G. Virga, *Variational Theories for Liquid Crystals*, Chapman Hall, London (1994).
[2] G. Baumann, *Symmetry Analysis of Differential Equations with Mathematica*; Telos/ Springer: New York, NY, USA,2000.
[3] J.K. Hunter, R.A. Saxton, *Dynamics of director fields*, SIAM J. Appl.Math.1991,51,1498-1521.
[4] Johnpillai, A. G.; Khalique, C. M. Symmetry Reductions, *Exact Solutions, and Conservation Laws of a Modified Hunter-Saxton Equation*, Abstr. Appl. Anal.2013,2013, 204746.
[5] L. V. Ovsiannikov, *Group Analysis of Differential Equations*, Academic Press, New York, (1982).
[6] M. C. Kakuli, W. Sinkala, P. Masemola, *Conservation laws and Symmetry Reductions of the Hunter-Saxton Equation via the Double Reduction Method*, Math.Comput.Appl. 2023, 28-92.

[7] M. Jafari, A. Zaeim, A. Tanhaeivash, *Symmetry group analysis and conservation laws of the potential modified KdV equation using the scaling method*, International Journal of Geometric Methods in Modern Physics 19(07), 2250098.

[8] M. Jafari, S. S. Mahdion, *Investigating the New Conservation Laws of Hunter-Saxton Equation via Lie Symmetries*, International Journal of Mathematical Modelling & Computations 13 (2), (2023), 24-37.

[9] M. Nadjafikhah, S. Shaban, *Computation of symmetry and conservation law for the Camassa-Holm and Hunter-Saxton equation*, AUT Journal. Math. Com,4(2) (2023) 113-122.

[10] P.J. Olver, *Applications of Lie Groups to Differential Equations*, 2nd ed.; Springer: New York, NY, USA,1993.

[11] P.J. Olver, P. Rosenau, *Group-invariant solutions of Differential Equations*, SIAM J.Appl. Math.,47, N 2, (1987),263-278.

[12] Yao, S. W.; Gulsen, S.; Hashemi, M. S.; Inc, M.; Bicer, H. *Periodic Hunter-Saxton equation parametrized by the speed of the Galilean frame: Its new solutions, Nucci's reduction, first integrals and Lie symmetry reduction*. Results Phys.2023,47, 106370.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–248, pp. 500–505

The Koch curve as manifolds and the Chaos Game

Hamed Soroush[a,*]

[a] *Department of Mathematics, Payame Noor University (PNU), Tehran, Iran*

Article Info

Keywords:
Fractal
Chaos
homeomorphisms
Koch curve

2020 MSC:
90C34
90C40
49J52

Abstract

In this paper, a summary of concepts from the Chaos and Fractals theory and some basic theory to understand how the Chaos and Fractal theory fundamentals work, for more to explain how is related the chaos theory with mathematics application.We find homeomorphisms between The Koch and manifolds. We can use this homeomorphisms to endow such fractals with the structure of smooth manifolds.

1. The Chaos theory

Tien Yien Li and James Yorke's 1975 paper, "Period three implies chaos", contained the first instance of this usage of chaos in the scientific literature, and led ultimately to the creation of a new science, chaos theory. [8]
The chaos theory is for some concepts that they offer an alternative that describes and explanations how is the behavior of some nonlinear systems (which are basically almost all naturally occurring physical, chemical, biological or social structures or systems). The name "Chaos" comes from the fact that nonlinear systems seem to behave chaotically or randomly from a traditional linear point of view. There are many natural systems whose behavior that can't be described and explained by simply dividing the whole into its parts and study them separately from the rest of the system. For example, studying the behavior of an individual bee may not provide any insight into a beehive as a system because the bees colony's behavior is driven by the cooperation and pheromone interaction between flowers. In a different example, of course the movement of water molecules in process of boiling the water might seem chaotic and random, but there are patterns of movements that change over time and tend to form similar structure.
Most natural systems change over time and this change does not happen in proportional and regular manner. A concept of proportional change is an idealization because real life phenomena change differently sometimes smoothly, sometimes not smoothly. The Chaos theory provides a theoretical framework and a set of tools for conceptualizing change and the changing system may have appeared to be chaotic from traditional (linear) perspective while it exhibits coherence, structure and patterns of motion from the global and nonlinear perspective.

*Talker
Email address: h_soroush2011@pnu.ac.ir (Hamed Soroush)

Chaos is a fundamental property that possesses nonlinearity and it is very sensitive on initial conditions. Because of the nonlinearity in a chaotic system it becomes very difficult to make an exact or accurate predictions about the system over a given time interval. Weather forecasting is an example of how chaos theory effects the accuracy of predictions over a given time interval, but using the similar structured a meteorologist can predict how is going to move a hurricane. Through analyzing a weather pattern over time and different structures, meteorologists have been able to make better predictions of future weather based on this theory. [2]
The dictionary definition of chaos is turmoil, unpredicted, turbulence, primordial abyss, and undesired randomness, but scientists will tell you that chaos is something extremely sensitive to initial conditions. Chaos also refers to the question of whether or not it is possible to make good long-term predictions about how a system will act. A chaotic system can actually develop in a way that appears very smooth and ordered. Determinism is the belief that every action is the result of preceding actions. It began as a philosophical belief in Ancient Greece thousands of years ago and was introduced into science around 1500 A.D. with the idea that cause and effect rules. Newton was closely associated with the establishment of determinism in modern science. His laws were able to predict systems very accurately. They were deterministic at their core because they implied that everything that would occur would be based entirely on what happened right before. Henry Adams has described like this ”Chaos often breeds life, when order breeds habit”. Henri Poincaré was really the ”Father of Chaos [Theory],”. Chaos theory describes complex motion and the dynamics of sensitive systems. Chaotic systems are mathematically deterministic but is hard or impossible to predict. Chaos is more evident in long-term systems than in short-term systems. Behavior in chaotic systems is not periodic, meaning that no variable describing the state of the system undergoes a regular repetition of values. A chaotic system can actually develop gradually in a way that appears to be smooth and ordered, however. Chaos refers to the issue of whether or not it is possible to make accurate long-term predictions of any system if the initial conditions are known to an accurate degree. Chaos occurs when a system is very sensitive to initial conditions. Initial conditions are the values of measurements at a given starting time. The phenomenon of chaotic motion was considered a mathematical oddity at the time of its discovery, but now physicists know that it is very widespread and may even be the norm in the universe. The weather is an example of a chaotic system. In order to make long-term weather forecasts it would be necessary to take an infinite number of measurements, which would be impossible to do. Also, because the atmosphere is chaotic, tiny uncertainties would eventually overwhelm any calculations and defeat the accuracy of the forecast. The presence of chaotic systems in nature seems to place a limit on our ability to apply deterministic physical laws to predict motions with any degree of certainty. [2]

2. Notations and Preliminaries

Helge von Koch was a Swedish mathematician who, in 1904, introduced what is now called the Koch curve ([8]). Fitting together three suitably rotated copies of the Koch curve produces a figure, which for obvious reasons is called the snowflake curve or the Koch island. The outline of the Koch snowflake (also called Koch island) is composed of three congruent parts, each of which is a Koch curve as shown in figure. The Koch snowflake obviously has some

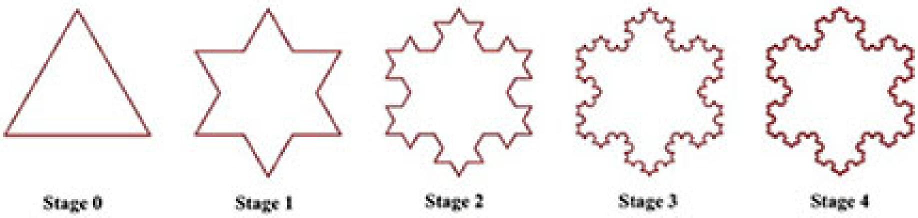


Fig. 1. The Koch snowflake.

similarities with real flakes, some of which are pictured here.
He defined the curve as the limit of an infinite sequence of increasingly wrinkly curves. The finished curve is infinitely long, despite being contained in a finite area. It has no tangent or smoothness anywhere. Slicing the curve at certain angles reveals an infinity of Cantor sets lurking within.
In the paper [The Koch curve as a smooth manifold, Chaos, Soliton and fractals, 38(2008)334-338] the authors try

to find a homeomorphism between the Koch curve and the closed interval $[0, 1]$ to endow the Koch curve with the structure of a smooth manifold with boundary. The method which they used is complicated and strongly depended on the structure of the Koch curve. We give a simple ordinary proof which can be used in similar cases. We first define the family of "Exact limit fractals", which involves many kinds of fractals, such as the Koch curve. Then we find homeomorphisms between exact limit fractals and manifolds.

2. Exact limit fractals

In the following, the topology of the subsets of R^n are induced topology from R^n, and the usual distance of two points x, y in R^n, is denoted by $d(x, y)$.

Definition 2.1. *Let F be a fractal subset of R^n and $\{A_1, A_2, ...\}$ be a sequence of subsets of R^n. Then, F is called an "Exact limit fractal", with the starting set A_1, and the function set $\{f_i\}$, if the following assertions are true:*
(1) For each $i \geq$ there exists a homeomorphism $f_i : A_i \to A_{i+1}$.
(2) For each $x \in A_1$ the following sequence, which we call exact sequence, converges to a point in F

$$x_1 = x, x_2 = f_1(x), ..., x_{n+1} = f_n(x_n), ...$$

(3) For each point $y \in F$ there is a unique exact sequence $\{x_n\}$, such that $y = lim_{n\to\infty} x_n$.
(4) For each $\epsilon > 0$, there is a number $n_\epsilon \in N$ such that for each $x_1 \in A_1$ we have:

$$m \geq n_\epsilon \Rightarrow d(x_m, lim_{n\to\infty} x_n) < \epsilon$$

Theorem 2.2. *Let F be an exact limit fractal in R^n, with the starting set A_1. If A_1 is compact then F is homeomorphic to A_1.*

Proof: Consider $\{A_1, A_2, ...\}$ and $\{f_1, f_2, ...\}$ as definition 2.1, and for each $m \geq 1$ let $f^{(m)} = f_m \circ f_{m-1} \circ ... \circ f_1$. Consider an exact sequence $x_{m+1} = f_m(x_m)$,$m \geq 1$. Define the function $\psi : A_1 \to F$ by

$$\psi(x_1) = lim_{n\to\infty} x_n$$

We show that ψ is a homeomorphism. Let $\epsilon > 0$ be given. By definition 2.1, there is $n_\epsilon \in N$ such that

$$m \geq n_\epsilon \Rightarrow d(x_m, lim_{n\to\infty} x_n) < \frac{\epsilon}{3}$$

Fix a number $m \geq n_\epsilon$. Since $f^{(m)} : A_1 \to A_m$ is a homeomorphism, there exists a $\delta > 0$(related to ϵ), such that

$$d(y_1, x_1) < \delta \Rightarrow d(f^{(m)}(y_1), f^{(m)}(x_1)) < \frac{\epsilon}{3} \Rightarrow d(y_m, x_m) < \frac{\epsilon}{3}$$

Thus

$$d(\psi(y_1), \psi(x_1)) = d(lim_{n\to\infty} y_n, lim_{n\to\infty} x_n)$$

$$\leq d(lim_{n\to\infty} y_n, y_m) + d(y_m, x_m) + d(x_m, lim_{n\to\infty} x_n) \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

That is

$$d(y_1, x_1) < \delta \Rightarrow d(\psi(y_1), \psi(x_1)) < \epsilon$$

Therefore, ψ is continuous. A_1 is compact and by definition 2.1(3), ψ is injective and onto, so it is a homeomorphism (because any continuous bijection from a compact space onto a Hausdorff space is a homeomorphism).

Remark 2.3. *In the definition 2.1, let the conditions (1)-(3) are valid and there exists constant numbers $0 < \alpha < 1$ and $\beta > 0$, such that for each exact sequence $\{x_n\}$, we have $d(x_n, x_{n+1}) \leq \beta(\alpha)^n$. Then the condition (4) in the definition 2.1 is valid.*

Proof : Consider an exact sequence $\{x_n\}$. Let $x = lim_{n\to\infty} x_n \in F$. We have

$$d(x_n, x) = lim_{m\to\infty} d(x_n, x_m) \leq lim_{m\to\infty}(d(x_n, x_{n+1}) + ... + d(x_{m-1}, x_m)) \leq lim_{m\to\infty} \beta(\alpha^n + ... + \alpha^{m-1})$$

$$= \beta \frac{\alpha^n}{1 - \alpha}$$

For given $\epsilon > 0$, if we choose an N_ϵ sufficiently large, then we would have

$$n \geq N_\epsilon \Rightarrow \frac{\beta\alpha^n}{1-\alpha} < \epsilon \Rightarrow d(x_n, x) \leq \epsilon$$

Example 2.4. *The Koch curve with the induced topology from* R^2*, is homeomorphic to* $[0, 1]$.
Proof: Denote the Koch curve by K. Let L be a line segment of the length a. Replace the middle third segment of L by two sides of an equilateral triangle of the side length $\frac{a}{3}$, to get the set $\widehat{L}$. For each x in L, let L_x be the line perpendicular to L at the point x. Define the map g_L as follows

$$g_L : L \to \widehat{L}; g_L(x) = L_x \bigcap \widehat{L}$$

Clearly g_L is a homeomorphism and we have

$$d(g_L(x), x) \leq \frac{\sqrt{3}}{2}a$$

Now consider the usual method for constructing of the Koch curve, starting by the closed interval $A_1 = [(0, 0), (0, 1)]$. Then replacing the middle third by two sides of an equilateral triangle of side length $\frac{1}{3}$ to get A_2. Do the similar thing to the line segments of A_2 to get A_3, and so on. In fact for each n, A_n is a union of 4^{n-1} line segments L_{n_i},$1 \leq i \leq 4^{n-1}$, of the length $(\frac{1}{3})^{n-1}$, and A_{n+1} is the union of the sets $\widehat{L_{n_i}}$. Consider the function $f_n : A_n \to A_{n+1}$, defined by

$$f_n(x) = g_{L_{n_i}}(x), x \in L_{n_i}$$

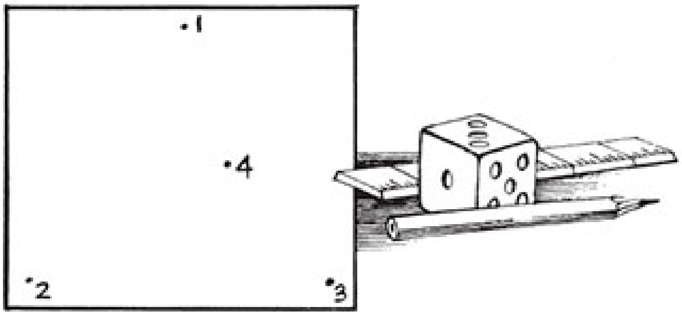
Clearly f_n is a homeomorphism, and we have

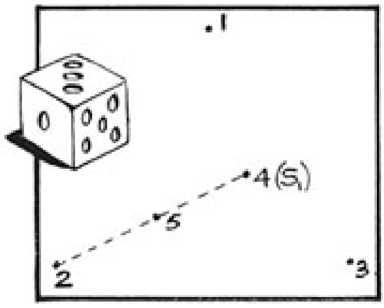
$$d(f_n(x), x) \leq \frac{\sqrt{3}}{2}(\frac{1}{3})^{n-1} = \frac{3\sqrt{3}}{2}(\frac{1}{3})^n$$

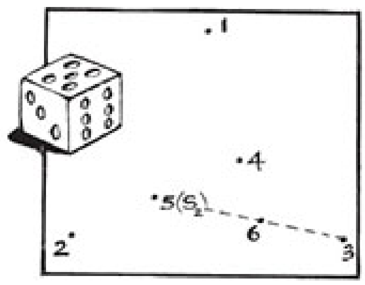
By using Definition 2.1 and Remark 2.3, we get that K is an exact limit fractal, with the starting set A_1. Thus by Theorem 2.2, K is homeomorphic to A_1. This completes the proof.
In the example 2.4, if we start by four sides of a square, instead of closed interval, we can show in the similar way, that the Koch snowflake is homeomorphic to the boundary of a square, so it is homeomorphic to a circle. Thus we have the following corollary.
Corolarry 2.5 *The Koch snowflake is homeomorphic to the circle* S^1.
There are many other fractals, which can be characterized from topological view point, by using the Theorem 2.2. For instance, we can show that, the modified Koch curve ([4] page 121) and self affine curve ([4] page 151) are homeomorhpic to $[0, 1]$.

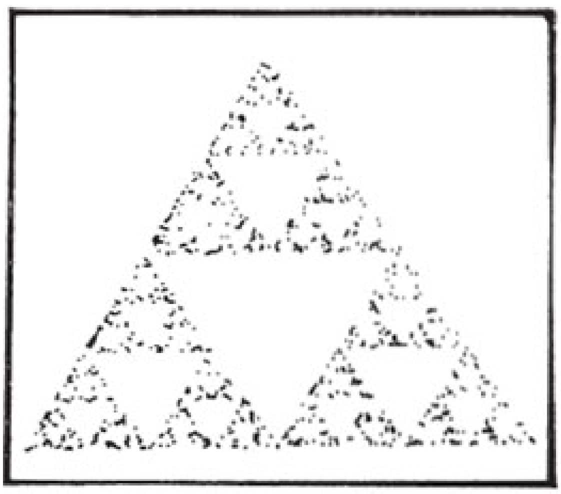
3. The Chaos Game

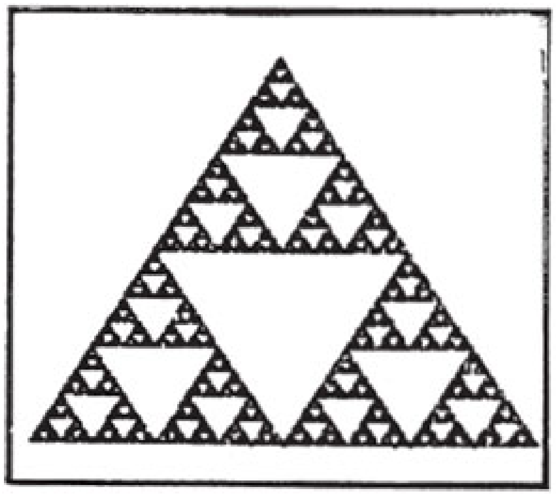
In the 1980s Michael Barnsley [3] discovered another way of generating a fractal. It's a bit like dot-to-dot drawings, only you don't join the dots with lines. You just plot point after point according to some simple rules.
For example: draw three dots in a triangle, and a fourth dot at random, anywhere in the triangle, to be the stating point. Now proceed as follows.
Step 1: Roll a die.
Step 2: If the top face is 1 or 2, draw another dot half-way from the starting point to the first point. If the top face is 3 or 4, draw another dot half-way from the starting point to the second dot. If the top face is 5 or 6, draw another dot half-way from the starting point to the third dot. This new dot then becomes the new stating point, and the whole process is repeated, from Step 1. After a while, a pattern begins to emerge, a familiar pattern: the Sierpinski triangle.
It's not much of a game. With only one player having just one move, the game does not allow for a great deal of choice. After you have chosen your initial point, the future of the game is decided. As Michael Barnsley discovered, if we had chosen different points, we could have generated a fern instead, or any other fractal – or any shape at all, come to that. Every picture can be encoded as a fractal formula like this.

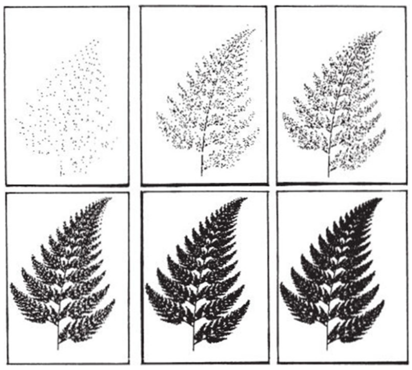

1
4
2
3


1
4(S1)
5
2
3


1
4
5(S2)
6
2
3







References

[1] M. F. Barnsley and S. Demko, *Iterated function systems and the global construction of fractals*, Proc. Roy. Soc. London Ser. A. **399** (1985), 243-275.
[2] M. F. Barnsley, *Fractals Everywhere*, Academic Press, New York, 1998.
[3] M. Espin and J. Sniatycki *The Koch curve as a smooth manifold, Chaos, Soliton and fractals*. **38** (2008), 334-338.
[4] K. Falconer, *Fractal Geometry, Mathematical Foundations and Applications*, J. John Wiley and sons, New york, 2014.
[5] H. Peitgen, H. Jürgens and D. Saupe, *Chaos and Fractals, New Frontiers of Science Second Edition*, Springer, 2004.
[6] B. B. Mandelbrot, *The Fractal Geometry of Nature*, New York, 1977.
[7] J. R. Munkres, *Topology a first course*, Appleton Century Crofts, 2000.
[8] N. Lesmoir-Gordon, W. Rood, *Introducing Fractals: A Graphic Guide*, London, 2013.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, GT–253, pp. 506–509

Clean rings on $C_c(X)$

Zahra Keshtkar[a,*], Amir Veisi[b]

[a]*Faculty of Petroleum and Gas, Yasouj University, Gachsaran, Iran*
[b]*Faculty of Petroleum and Gas, Yasouj University, Gachsaran, Iran*

Article Info

Keywords:
clean ring
F_c-space
$C_c(X)$

2020 MSC:
54C30
13B99

Abstract

In this article, clean elements of $C(X, \mathbb{N})$ and $C(X, \mathbb{Z})$ are determined. It is shown that for every space X, $C_c(X)(x)$, $Q_c(X)$, and $q_c(X)$ are clean rings. We prove that if X is an F_c-space, and S is a multiplicative closed subset of $\mathcal{R}(C_c(X))$, then max $\big(S^{-1}C_c(X)\big)$ and $\beta_0 X$ are homeomorphic, where $\mathcal{R}(C_c(X))$ is the set of non-zerodivisor elements of $C_c(X)$. Furthermore, $S^{-1}C_c(X)$ is a clean ring. It is also observed that $C_c(X)[x]$ is not clean although $C_c(X)[[x]]$ is clean.

1. Introduction

As usual, all topological spaces in this article are infinite Hausdorff completely regular (i.e., infinite Tychonoff spaces). We recall that a *zero-dimensional space* is a Hausdorff space with a base consisting of clopen sets. We denote by $C(X)$ the ring of all real-valued, continuous functions on a space X. The subring of $C(X)$ consisting of those functions with countable image, which is denoted by $C_c(X)$ is an $\mathbb{R}$-subalgebra of $C(X)$. The ring $C_c(X)$ is introduced and studied in [2], [4], [5], and [8]. The *Banaschewski compactification* of a zero-dimensional topological space X is denoted by $\beta_0 X$. It is observed in [2] that $\beta_0 X$ is the maximal ideal space of $C_c(X)$. We say that a space X is an F_c-space whenever the prime ideals of $C_c(X)$ contained in a given maximal ideal form a chain. An element a in a commutative ring R is called a *clean* element if a is a sum of a unit and an idempotent in R, and R is called *clean* if every element in R is clean. R is said to be *zero-dimensional* if every prime ideal in R is maximal. Also, R is called *regular* if, for every $a \in R$, there exists $b \in R$ such that $a = a^2 b$. We remind that a topological space X is called a *CP-space* if $C_c(X)$ is a regular ring. In the following, $\mathbb{Z}$ represents the set of integers, and $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

2. Topics and results

Recall that $e \in C(X)$ is idempotent if and only if for a clopen subset A of X; $e(x) = \begin{cases} 1 & x \in A \\ 0 & x \in X \setminus A \end{cases}$. It is easily seen that the idempotent elements of $C(X)$, $C(X, \mathbb{N})$, and $C(X, \mathbb{Z})$ are the same.

---
*Talker
Email addresses: zahra.keshtkar@gmail.com (Zahra Keshtkar), aveisi@yu.ac.ir (Amir Veisi)

Proposition 2.1. *$u \in C(X, \mathbb{N})$ is a unit if and only if $u = 1$. Moreover, $f \in C(X, \mathbb{N})$ is clean if and only if $f = 1+e$, where e is an idempotent element of $C(X)$.*

Proof. Since u is a unit, there exists $0 \neq v \in C(X, \mathbb{N})$ such that $u(x)\, v(x) = 1$ for every $x \in X$. Then $u(x) = \frac{1}{v(x)} \in \mathbb{N}$. So $v(x) = 1$ and hence $u(x) = 1$ for every $x \in X$, i.e., $u = 1$. The second part is clear. □

Proposition 2.2. *$u \in C(X, \mathbb{Z})$ is a unit if and only if $u(X) \subseteq \{-1, 1\}$. Also, $f \in C(X, \mathbb{Z})$ is clean if and only if for an idempotent $e \in C(X)$ and a clopen subset A of X, we have $f = \begin{cases} e+1 & on\ A \\ e-1 & on\ X \setminus A. \end{cases}$*

Proof. If u is a unit, then there exists $0 \neq v \in C(X, \mathbb{Z})$ such that $u(x)\, v(x) = 1$ for every $x \in X$. Then $u(x) = \frac{1}{v(x)} \in \mathbb{Z}$. So $v(x) = 1$ or $v(x) = -1$. Hence, $u(x) = 1$ or $u(x) = -1$, for every $x \in X$, i.e., $u(X) \subseteq \{-1, 1\}$. The converse is clear.
Let f be clean. Then $f = u + e$, where u is a unit and e is an idempotent element of $C(X, \mathbb{Z})$. So $u(X) \subseteq \{-1, 1\}$. If we put $A = \{x \in X : u(x) = 1\}$, then A is a clopen set. Therefore,
$f(x) = \begin{cases} e+1 & x \in A \\ e-1 & x \in X \setminus A \end{cases}$, and we are done. Now, suppose that for a clopen set $A \subseteq X$ and an idempotent $e \in C(X)$, we have $f = \begin{cases} e+1 & on\ A \\ e-1 & on\ X \setminus A \end{cases}$. Let $u(x) = \begin{cases} 1 & x \in A \\ -1 & x \in X \setminus A \end{cases}$, then u is a unit element of $C(X, \mathbb{Z})$, and $f = u + e$ is clean, as desired. □

The next result determines the general form of clean elements of $C(X, \mathbb{N})$ and $C(X, \mathbb{Z})$.

Corollary 2.3. *An element $f \in C(X, \mathbb{N})$ is clean if and only if there exists a clopen set A of X such that*

$$f(x) = \begin{cases} 2 & x \in A \\ 1 & x \in X \setminus A \end{cases}.$$

An element $f \in C(X, \mathbb{Z})$ is clean if and only if there exist clopen sets A, B of X such that

$$f(x) = \begin{cases} 2 & x \in A \cap B \\ 1 & x \in A \setminus B \\ 0 & x \in B \setminus A \\ -1 & x \in X \setminus (A \cup B). \end{cases}$$

The content ideal of a polynomial $p(x) = a_n x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 \in C_c(X)[x]$ is the ideal of $C_c(X)$ generated by coefficients $a_n, a_{n-1}, ..., a_1, a_0$. We denoted the content ideal $p(x)$ by $c(p)$. Let $S = \big\{p(x) \in C_c(X)[x] : c(p) = C_c(X)\big\}$. Clearly, S is closed under multiplication. The localization of $C_c(X)[x]$ with respect to S is denoted by $C_c(X)(x)$.

Theorem 2.4. *For every space X, $C_c(X)(x)$ is a clean ring.*

Proof. Let $a \in C_c(X)(x)$. Then there exist $p(x),\, q(x) \in C_c(X)[x]$ such that $a = \frac{p(x)}{q(x)}$ and $c(q) = C_c(X)$. So we can take

$$f_0,\, f_1,\, \dots,\, f_n,\, g_0,\, g_1, \dots,\, g_n,\, \alpha_1,\, \alpha_2,\, \dots\, \alpha_n \in C_c(X)$$

such that

$$p = \sum_{i=0}^{n} f_i\, x^i, \quad q = \sum_{i=0}^{n} g_i\, x^i, \text{ and } \quad \sum_{i=0}^{n} \alpha_i g_i = 1.$$

Let $f = 1 - \sum_{i=0}^{n} \alpha_i\, f_i$. Then $f \in C_c(X)$ and hence $f(X) \neq (0,\ 1)$. So there exists $r \in (0,\ 1)$ such that $r \notin f(X)$. Set $U = f^{-1}((r,\ +\infty)) = f^{-1}([r,\ +\infty))$ and define $e,\, h : X \to \mathbb{R}$ as follows:

$$e(x) = \begin{cases} 1 & x \in U \\ 0 & x \in X \setminus U, \end{cases} \quad \text{and } \ h(x) = \begin{cases} \frac{-1}{f} & x \in U \\ \frac{1}{1-f} & x \in X \setminus U. \end{cases}$$

Since U (and hence $X \setminus U$) is a clopen set, e and h are continuous, and so they belong to $C_c(X)$. Furthermore, e is idempotent. Hence, for every $x \in X$ we obtain $\sum_{i=0}^{n} h(x)\,\alpha_i(x)\,\Big(f_i(x) - e(x)g_i(x)\Big) = 1$. Therefore, the ideal of $C_c(X)$ which is generated by the set $\{f_i - eg_i : 0 \leq i \leq n\}$ is the $C_c(X)$. Now, let

$$k = p - eq = \textstyle\sum_{i=0}^{n} (f_i - eg_i)x^i.$$

Then $c(k) = C_c(X)$. Hence, $\frac{k}{q} \in C_c(X)(x)$. Since $c(q) = C_c(X)$, we get $\frac{k}{q}$ is a unit. On the other hand, idempotent elements of $C_c(X)(x)$ and $C_c(X)$ are the same. So e is an idempotent in $C_c(X)(x)$. Therefore,

$$\frac{k}{q} + e = a.$$

This yields a is a clean element in $C_c(X)(x)$, and we are done. □

The ring $C_c(X) < x >$ is the localization of the polynomial ring $C_c(X)[x]$ concerning the set of monic polynomials.

Proposition 2.5. *For any topological space X, the following are equivalent.*

(1) X *is a* CP*-space.*
(2) $C_c(X)$ *is a regular ring.*
(3) $C_c(X)$ *is zero-dimensional.*
(4) $C_c(X) < x >$ *is zero-dimensional.*
(5) $C_c(X) < x >$ *is a clean ring.*
(6) $C_c(X) < x >= C_c(X)(x)$.

Proof. By [4, Theorem 5.8], $(1),\ (2), and\ (3)$ are equivalent. By [7, Theorem 8], $(3),\ (4),\ (5)$, and (6) are equivalent. So the result holds.

□

Theorem 2.6. *Let X be an F_c-space, $\mathcal{R}(C_c(X))$ be the set of non-zerodivisors elements of $C_c(X)$, and let S be a multiplicative closed set of $\mathcal{R}(C_c(X))$. Then the following hold.*

(i) $\max\big(S^{-1}C_c(X)\big)$ *and* $\beta_0 X$ *are homeomorphic.*
(ii) $S^{-1}C_c(X)$ *is a clean ring.*

Proof. (i). The proof is similar to the proof of [3, Proposition 3.2], and so we eliminate it.
(ii). First, we note that $C_c(X)$ is a pm-ring (i.e., every prime ideal of $C_c(X)$ is contained in a unique maximal ideal). Hence, $S^{-1}C_c(X)$ is a is pm-ring. By (i), $\max\big(S^{-1}C_c(X)\big) \simeq \beta_0 X$, and the fact that $\beta_0 X$ is zero-dimensional, we obtain $\max\big(S^{-1}C_c(X)\big)$ is zero-dimensional. Now, [6, Theorem 1.7] gives $S^{-1}\big(C_c(X)\big)$ is a clean ring, and we are done. □

The maximal ring of quotients and the classical ring of quotients of $C_c(X)$ are denoted by $Q_c(X)$ and $q_c(X)$ respectively. These rings have been identified in [8, Theorem 2.12] as follows:

$$Q_c(X) = \varinjlim\{C_c(V) : V \text{ is dense open in } X\}, \text{ and } \ q_c(X) = \varinjlim\{C_c(coz\, f) :\ f \in C_c(X) \text{ and } \ \overline{coz\, f} = X\}.$$

Proposition 2.7. *Let X be zero-dimensional. Then $Q_c(X)$ and $q_c(X)$ are clean rings.*

Proof. According to [8, Corollary 2.17], $Q_c(X)$ is regular and therefore it is clean since a regular ring is a clean ring. On the other hand, $q_c(X)$ is a direct limit of the rings $C_c(V)$ where V ranges over dense cozerosets of X. By [2, Corollary 2.8], $C_c(V)$ is clean. Now, [3, Proposition 2.4] gives $q_c(X)$ is also clean, and we are done. □

Note that $R[x]$ is never clean for any ring R ([1, Proposition 12]). Therefore, $C_c(X)[x]$ is not clean. Furthermore, R is clean if and only if $R[[x]]$ is clean, by [1, Proposition 12]. Now, since $C_c(X)$ is clean, $C_c(X)[[x]]$ is also clean. This fact is emphasized in the following statement.

Proposition 2.8. *For any topological space X, $C_c(X)[[x]]$ is a clean ring.*

References

[1] D.D. Anderson and V.P. Camillo, *Commutative rings whose elements are a sum of a unit and an idempotent*, Comm. Algebra, 30 (2002), 3327–3336.

[2] F. Azarpanah, O.A.S. Karamzadeh, Z. Keshtkar, and A. Olfati, *On maximal ideals of* $C_c(X)$ *and uniformity its localizations*, Rocky Mountain Journal of Mathematics, 48(2) (2018), 345-384.

[3] W.D. Burgess and R. Raphael, *Clean classical rings of quotients of commutative rings, with applications to* $C(X)$, Journal of Algebra and Its Applications, 7(2) (2008), 195-209.

[4] M. Ghadermazi, O. A. S. Karamzadeh, and M. Namdari, *On the functionally countable subalgebra of* $C(X)$, Rend. Sem. Mat. Univ. Padova, 129 (2013), 47-69.

[5] O.A.S. Karamzadeh and Z. Keshtkar, *On c-realcompact spaces*, Quaestiones Mathematicae, 41 (2018), 1135-1167.

[6] W. Wm. McGovern, *Neat rings*, J. Pure Appl. Algebra, 105 (2006), 243–265.

[7] W. Wm. McGovern, F. Richman, *When* $R(X)$ *and* $R < X >$ *are clean a constructive treatment*, Communications in algebra, 43 (2015), 3389-3394.

[8] M. Namdari, A. Veisi, *Rings of quotients of the subalgebra of* $C(X)$ *consisting of functions with countable range*, Int. Math. Forum, 7(2) (2012), 561-571.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–146, pp. 510–517

A survey on results to fractional boundary value problem

Mohsen Khaleghi Moghadam[a,*], Yasser Khalili[b]

[a]*Department of Basic Sciences, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.*
[b]*Department of Basic Sciences, Sari Agricultural Sciences and Natural Resources University, Sari, Iran.*

Article Info

Keywords:
Discrete fractional calculus
Discrete nonlinear boundary value problem
Non trivial solution
Variational methods
Critical point theory

2020 MSC:
26A33
39A10
39A27

Abstract

In this paper, we deal with the existence of a non-trivial solution for the following fractional discrete boundary-value problem for any $k \in [1, T]_{\mathbb{N}_0}$

$$\begin{cases} {}_{T+1}\nabla_k^{\alpha}\left({}_k\nabla_0^{\alpha}(u(k))\right) + {}_k\nabla_0^{\alpha}\left({}_{T+1}\nabla_k^{\alpha}(u(k))\right) = \lambda h(k) g(u(k)) - u(k), \quad k \in [1, T]_{\mathbb{N}_0}, \\ u(0) = u(T+1) = 0, \end{cases}$$

where $0 < \alpha < 1$ and ${}_k\nabla_0^{\alpha}$ is the left nabla discrete fractional difference and ${}_{T+1}\nabla_k^{\alpha}$ is the right nabla discrete fractional difference $f : [1, T]_{\mathbb{N}_0} \times \mathbb{R} \to \mathbb{R}$ is a continuous function, $\lambda > 0$ is a parameter. The technical method is variational approach for differentiable functionals. An example is included to illustrate the main results.

1. Introduction

The first concepts of fractional nabla differences traces back to the works of Gray and Zhang [14]. Discrete fractional calculus with the nabla operator studied in [3]. In [4] authors studied two-point boundary value problems for finite fractional difference equations. This kind of problems play a fundamental role in different fields of research, for example in biological, Atici and Şengül introduced and solved Gompertz fractional difference equation for tumor growth models [5].
We refer the reader to the recent monograph on the introduction to fractional nabla calculus [10].
Also we refer the reader to the new monograph [18] that works for differential and integral equations and systems and for many theoretical and applied problems in mathematics, mathematical physics, probability and statistics, applied computer science and numerical methods.
It is well known that variational methods is an important tool to deal with the problems for differential and difference equations. Fractional difference equations with boundary value conditions have appeared in [9, 16] by using variational methods. More, recently, in [8, 11, 12], the existence and multiplicity of solutions for nonlinear discrete boundary value problems have been investigated by adopting variational methods.
In last decades, some researchers investigated q-fractional difference equations. Later, q-fractional boundary value

*Talker
Email addresses: m.khaleghi@sanru.ac.ir (Mohsen Khaleghi Moghadam), y.khalili@sanru.ac.ir (Yasser Khalili)

problems considered by many researchers; see for instance, [17] and references therein. The other important tool in the study of nonlinear difference equations is fixed point methods. Morse theory is also other tool in the study of nonlinear fractional differential equations [13].
The aim of this paper is to establish the existence of non-trivial solution for the following discrete boundary-value problem

$$\begin{cases} {}_{T+1}\nabla_k^\alpha \left({}_k\nabla_0^\alpha(u(k))\right) + {}_k\nabla_0^\alpha \left({}_{T+1}\nabla_k^\alpha(u(k))\right) = \lambda f(k, u(k)) - u(k), \;\; k \in [1,T]_{\mathbb{N}_0}, \\ u(0) = u(T+1) = 0, \end{cases} \tag{1}$$

where $0 < \alpha < 1$ and ${}_k\nabla_0^\alpha$ is the left nabla discrete fractional difference and ${}_{T+1}\nabla_k^\alpha$ is the right nabla discrete fractional difference and $\nabla u(k) = u(k) - u(k-1)$ is the backward difference operator $f : [1,T]_{\mathbb{N}_0} \times \mathbb{R} \to \mathbb{R}$ is a continuous function, $\lambda > 0$ is a parameter and $T \geq 2$ is fixed positive integer and $\mathbb{N}_1 = \{1,2,3,\cdots\}$ and ${}_T\mathbb{N} = \{\cdots T-2, T-1, T\}$ and $[1,T]_{\mathbb{N}_0}$ is the discrete set $\{1,2,\cdots,T-1,T\} = \mathbb{N}_1 \bigcap {}_T\mathbb{N}$.

In this paper, based on a local minimum theorem (Theorem 2.4) due to Bonanno [6], we ensure an exact interval of the parameter λ, in which the problem (1) admits at least a non-trivial solution. As an example, here, we point out the following special case of our main results.

Theorem 1.1. *Let $h : [1,T] \to \mathbb{R}$ be a positive and essentially bounded function and $g : \mathbb{R} \to \mathbb{R}$ be a nonnegative continuous function and*

$$\lim_{d \to 0^+} \frac{g(d)}{d} = +\infty, \quad \lim_{c \to +\infty} \frac{g(c)}{c} = 0.$$

Then for any

$$\lambda \in \Big]0, +\infty\Big[,$$

the problem

$$\begin{cases} {}_{T+1}\nabla_k^\alpha \left({}_k\nabla_0^\alpha(u(k))\right) + {}_k\nabla_0^\alpha \left({}_{T+1}\nabla_k^\alpha(u(k))\right) = \lambda h(k) g(u(k)) - u(k), \quad k \in [1,T]_{\mathbb{N}_0}, \\ u(0) = u(T+1) = 0, \end{cases} \tag{2}$$

has at least one non-trivial solution in the space $\{u : [0, T+1] \to \mathbb{R} : u(0) = u(T+1) = 0\}$.

2. Preliminaries

The following definitions will be helpful to our discuss.

Definition 2.1. [2] **(i)** Let m be a natural number, then the m rising factorial of t is written as

$$t^{\overline{m}} = \prod_{k=0}^{m-1} (t+k), \quad t^{\overline{0}} = 1. \tag{3}$$

(ii) For any real number, the α rising function is increasing on $\mathbb{N}_0$ and

$$t^{\overline{\alpha}} = \frac{\Gamma(t+\alpha)}{\Gamma(t)}, \quad \text{such that} \;\; t \in \mathbb{R} \backslash \{\cdots, -2, -1, 0\}, \;\; 0^{\overline{\alpha}} = 0. \tag{4}$$

Definition 2.2. Let f be defined on $\mathbb{N}_{a-1} \bigcap {}_{b+1}\mathbb{N}$, $a < b$, $\alpha \in (0,1)$, then the nabla discrete new (left Gerasimov-Caputo) fractional difference is defined by

$$\left({}_k^C\nabla_{a-1}^\alpha f\right)(k) = \frac{1}{\Gamma(1-\alpha)} \sum_{s=a}^{k} \nabla_s f(s) (k - \rho(s))^{\overline{-\alpha}}, \quad k \in \mathbb{N}_a, \tag{5}$$

and the right Gerasimov-Caputo one by

$$\left({}_{b+1}^{C}\nabla_k^{\alpha} f\right)(k) = \frac{1}{\Gamma(1-\alpha)} \sum_{s=k}^{b} (-\Delta_s f)(s)(s-\rho(k))^{\overline{-\alpha}}, \quad k \in {}_b\mathbb{N}, \tag{6}$$

and in the left Riemann-Liouville sense by

$$\begin{aligned} \left({}_{k}^{R}\nabla_{a-1}^{\alpha} f\right)(k) &= \frac{1}{\Gamma(1-\alpha)} \nabla_k \sum_{s=a}^{k} f(s)(k-\rho(s))^{\overline{-\alpha}}, \quad k \in \mathbb{N}_a, & (7)\\ &= \frac{1}{\Gamma(-\alpha)} \sum_{s=a}^{k} f(s)(k-\rho(s))^{\overline{-\alpha-1}}, \quad k \in \mathbb{N}_a, & (8) \end{aligned}$$

and the right Riemann-Liouville one by

$$\begin{aligned} \left({}_{b+1}^{R}\nabla_k^{\alpha} f\right)(k) &= \frac{1}{\Gamma(1-\alpha)} (-\Delta_k) \sum_{s=k}^{b} f(s)(s-\rho(k))^{\overline{-\alpha}}, \quad k \in {}_b\mathbb{N}, & (9)\\ &= \frac{1}{\Gamma(-\alpha)} \sum_{s=k}^{b} f(s)(s-\rho(k))^{\overline{-\alpha-1}}, \quad k \in {}_b\mathbb{N}, & (10) \end{aligned}$$

where $\rho(k) = k-1$ be the backward jump operator.

So, for convenience, from now on we will use the symbol ∇^{α} instead of ${}^{R}\nabla^{\alpha}$ or ${}^{C}\nabla^{\alpha}$.
Now we present summation by parts formula in new discrete fractional calculus.

Theorem 2.3. *([1, Theorem 4.4] Integration by parts for fractional difference) For functions f and g defined on $\mathbb{N}_a \bigcap {}_b\mathbb{N}$, $a \equiv b$ (mod 1), and $0 < \alpha < 1$, one has*

$$\sum_{k=a}^{b} f(k) \left({}_{k}\nabla_{a-1}^{\alpha} g\right)(k) = \sum_{k=a}^{b} g(k) \left({}_{b+1}\nabla_k^{\alpha} f\right)(k). \tag{11}$$

Our main tool is a local minimum theorem due to Bonanno (see [6, Theorem 5.1]), which is recalled below (see also [6, Proposition 2.1]). Such a result is more general than [15, Theorem 2.5] since the critical point, surely, is not zero. First, for given $\Phi,\ \Psi : X \to \mathbb{R}$, we defined the following functions

$$\beta(r_1, r_2) = \inf_{v \in \Phi^{-1}(]r_1, r_2[)} \frac{\sup_{u \in \Phi^{-1}(]r_1, r_2[)} \Psi(u) - \Psi(v)}{r_2 - \Phi(v)}, \tag{12}$$

and

$$\rho(r_1, r_2) = \sup_{v \in \Phi^{-1}(]r_1, r_2[)} \frac{\Psi(v) - \sup_{u \in \Phi^{-1}(]-\infty, r_1[)} \Psi(u)}{\Phi(v) - r_1}, \tag{13}$$

for all $r_1, r_2 \in \mathbb{R}$, with $r_1 < r_2$.

Theorem 2.4. *([6, Theorem 5.1]) Let X be a reflexive real Banach space, $\Phi : X \to \mathbb{R}$ a sequentially weakly semicontinuous coercive and continuously Gâteaux differentiable functional whose Gâteaux derivative admits a continuous inverse on X^* and $\Psi : X \to \mathbb{R}$ a continuously Gâteaux differentiable functional whose Gâteaux derivative is compact. Put $I_\lambda = \Phi - \lambda\Psi$ and assume that there are $r_1, r_2 \in \mathbb{R}$, with $r_1 < r_2$, such that*

$$\beta(r_1, r_2) < \rho(r_1, r_2),$$

where β and ρ are given by (12) and (13). Then, for each

$$\lambda \in \Lambda = \left] \frac{1}{\rho(r_1, r_2)}, \frac{1}{\beta(r_1, r_2)} \right[,$$

there is $u_{0,\lambda} \in \Phi^{-1}(]r_1, r_2[)$ such that $I_\lambda(u_{0,\lambda}) \le I_\lambda(u)$ for all $u \in \Phi^{-1}(]r_1, r_2[)$ and $I'_\lambda(u_{0,\lambda}) = 0$.

In order to give the variational formulation of the problem (1), let us define the finite $T-$dimensional Banach space

$$W := \{u : [0, T+1]_{\mathbb{N}_0} \to \mathbb{R} : u(0) = u(T+1) = 0\},$$

which is equipped with the norm

$$\|u\| := \left(\sum_{k=1}^{T} |u(k)|^2\right)^{\frac{1}{2}}.$$

Let $\Phi : W \to \mathbb{R}$ be the functional

$$\Phi(u) := \frac{1}{2}\sum_{k=1}^{T} |\left({}_k\nabla_0^{\alpha} u\right)(k)|^2 + |\left({}_{T+1}\nabla_k^{\alpha} u\right)(k)|^2 + \frac{1}{2}\sum_{k=1}^{T} |u(k)|^2. \tag{14}$$

An easy computation ensures that Φ turns out to be of class C^1 on W and Gateaux differentiable with

$$\begin{aligned}\Phi'(u)(v) &= \sum_{k=1}^{T} \left({}_k\nabla_0^{\alpha}(u(k))\right)\left({}_k\nabla_0^{\alpha} v(k)\right) + \left({}_{T+1}\nabla_k^{\alpha}(u(k))\right)\left({}_{T+1}\nabla_k^{\alpha} v(k)\right)\\ &+ \sum_{k=1}^{T} |u(k)|^{p-2}u(k)v(k),\end{aligned}$$

for all $u, v \in W$. To study the problem (1), for every $\lambda > 0$, we consider the functional $I_\lambda : W \to \mathbb{R}$ defined by

$$I_\lambda(u) := \Phi(u) - \lambda\Psi(u), \quad \Psi(u) := \sum_{k=1}^{T} F(k, u(k)), \tag{15}$$

where $F(k,u) = \int_0^u f(k,t)dt$.

Lemma 2.5. *The function u be a critical point of I_λ in W, iff u be a solution of the problem (1).*

3. Auxiliary inequalities

Now we provide some inequalities used throughout the paper, which hold on the space W. In the sequel, we will use the following inequality.

Lemma 3.1. *For every $u \in W$, we have*

$$\|u\|_\infty := \max_{k\in[1,T]} |u(k)| \le \|u\|. \tag{16}$$

Lemma 3.2. *For every $u \in W$, we have*

$$\frac{1}{2}\|u\|^2 \le \Phi(u) \le 2T(T+1)\|u\|^2 + \frac{1}{2}\|u\|^2. \tag{17}$$

4. Main Results

First, let us introduce a function for convenience. For given two non-negative constants c and d, put

$$a_d(c) := \frac{\sum_{k=1}^{T} \max_{|\xi|\le c} F(k,\xi) - \sum_{k=1}^{T} F(k,d)}{\frac{(c)^2}{2} - \frac{d^2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2 - \frac{Td^2}{2}}.$$

We state our main result as follows.

Theorem 4.1. *Assume that there exist a non-negative constant c_1 and two positive constants c_2 and d with*

(A0) $\frac{(c_1)^2}{2} < \frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}} \right)^2 + \frac{Td^2}{2} < \frac{(c_2)^2}{2}$,

such that

(A1) $\quad a_d(c_2) < a_d(c_1)$.

Then for any $\lambda \in]\frac{1}{a_d(c_1)}, \frac{1}{a_d(c_2)}[$ *the problem* (1) *has at least one non-trivial solution* $u_0 \in W$.

Proof. Our aim is to apply Theorem 2.4 to our problem. To this end, take $X = W$, and put Φ, Ψ and I_λ as in (14) and (15). We know Φ is a nonnegative continuously Gâteaux differentiable and sequentially weakly lower semicontinuous functional whose Gâteaux derivative admits a continuous inverse on X^*, and Ψ is a continuously Gâteaux differentiable functional whose Gâteaux derivative is compact. By similar arguing in [7], put

$$\bar{v}(k) = \begin{cases} d & k \in [1, T]_{\mathbb{N}_0}, \\ 0 & k = 0, T+1, \end{cases}$$

$r_1 = \frac{(c_1)^2}{2}$ and $r_2 = \frac{(c_2)^2}{2}$. Clearly $\bar{v} \in W$. Since $\bar{v}$ vanishes at the end points that is $\bar{v}(0) = 0 = \bar{v}(T+1)$, thus its nabla Riemann-Liouville and Gerasimov-Caputo fractional differences coincide, hence for any $k \in \mathbb{N}_1 \bigcap {}_T\mathbb{N}$

$$\left({}_{T+1}\nabla_k^\alpha \bar{v}\right)(k) = \left({}_{T+1}^{R}\nabla_k^\alpha \bar{v}\right)(k) = \left({}_{T+1}^{C}\nabla_k^\alpha \bar{v}\right)(k) = \frac{d(T+1-k)^{\overline{-\alpha}}}{\Gamma(1-\alpha)},$$

$$\left({}_k\nabla_0^\alpha \bar{v}\right)(k) = \left({}_k^{R}\nabla_0^\alpha \bar{v}\right)(k) = \left({}_k^{C}\nabla_0^\alpha \bar{v}\right)(k) = \frac{d(k)^{\overline{-\alpha}}}{\Gamma(1-\alpha)}.$$

So, we have

$$\begin{aligned} \Phi(\bar{v}) &= \frac{1}{2}\sum_{k=1}^{T} |\left({}_k\nabla_0^\alpha \bar{v}\right)(k)|^2 + |\left({}_{T+1}\nabla_k^\alpha \bar{v}\right)(k)|^2 + \frac{1}{2}\sum_{k=1}^{T} |\bar{v}(k)|^2 \\ &= \frac{1}{2}\sum_{k=1}^{T} |\frac{d(k)^{\overline{-\alpha}}}{\Gamma(1-\alpha)}|^2 + |\frac{d(T+1-k)^{\overline{-\alpha}}}{\Gamma(1-\alpha)}|^2 + \frac{Td^2}{2} \\ &= \frac{d^2}{2\left(\Gamma(1-\alpha)\right)^2}\sum_{k=1}^{T} |(k)^{\overline{-\alpha}}|^2 + |(T+1-k)^{\overline{-\alpha}}|^2 + \frac{Td^2}{2} \\ &= \frac{d^2}{\left(\Gamma(1-\alpha)\right)^2}\sum_{k=1}^{T} |(k)^{\overline{-\alpha}}|^2 + \frac{Td^2}{2} \\ &= \frac{d^2}{\left(\Gamma(1-\alpha)\right)^2}\sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 + \frac{Td^2}{2}, \end{aligned}$$

and

$$\Psi(\bar{v}) = \sum_{k=1}^{T} F(k, \bar{v}(k)) = \sum_{k=1}^{T} F(k, d).$$

Moreover, for all $u \in W$ such that $\Phi(u) < r_i$, $i = 1, 2$, taking (3.1) and (17) into account, one has $\max_{k\in[1,T]} |u(k)| \leq c_i$, $i = 1, 2$. Therefore,

$$\sup_{u\in\Phi^{-1}(-\infty, r_i)} \Psi(u) = \sup_{\Phi(u)<r_i} \sum_{k=1}^{T} F(k, u(k)) \leq \sum_{k=1}^{T} \max_{|\xi|\leq c_i} F(k, u(k)), \quad i = 1, 2.$$

By (A0), $\bar{v} \in \Phi^{-1}(r_1, r_2)$, hence,

$$\begin{aligned} 0 \leq \beta(r_1, r_2) &\leq \frac{\sup_{u \in \Phi^{-1}(r_1, r_2)} \Psi(u) - \Psi(\bar{v})}{r_2 - \Phi(\bar{v})} \\ &\leq \frac{\sup_{u \in \Phi^{-1}(-\infty, r_2)} \Psi(u) - \Psi(\bar{v})}{r_2 - \Phi(\bar{v})} \\ &\leq \frac{\sum_{k=1}^{T} \max_{|\xi| \leq c_2} F(k,\xi) - \sum_{k=1}^{T} F(k,d)}{\frac{(c_2)^p}{p(T+1)^{\frac{p(p-2)}{4}}} - \frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 - \frac{Td^p}{2}} \\ &= a_d(c_2). \end{aligned}$$

On the other hand, one has

$$\begin{aligned} \rho(r_1, r_2) &\geq \frac{\Psi(\bar{v}) - \sup_{u \in \Phi^{-1}(-\infty, r_1)} \Psi(u)}{\Phi(\bar{v}) - r_1} \\ &\geq \frac{\sum_{k=1}^{T} F(k,d) - \sum_{k=1}^{T} \max_{|\xi| \leq c_1} F(k,\xi)}{\frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 + \frac{Td^p}{2} - \frac{(c_1)^p}{p}} \\ &\geq \frac{\sum_{k=1}^{T} \max_{|\xi| \leq c_1} F(k,\xi) - \sum_{k=1}^{T} F(k,d)}{\frac{(c_1)^p}{2} - \frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 - \frac{Td^p}{2}} = a_d(c_1). \end{aligned}$$

Hence, from Assumption (A1), we get $\beta(r_1, r_2) < \rho(r_1, r_2)$.
Therefore, owing to Theorem 2.4, for each $\lambda \in]\frac{1}{a_d(c_1)}, \frac{1}{a_d(c_2)}[$, the functional I_λ admits one critical point $u_0 \in W$ such that $r_1 < \Phi(u_0) < r_2$. Hence, the proof is completed. □

Here we point out an another immediate consequence of Theorem 4.1 as follows.

Theorem 4.2. *Let $f : [1,T]_{\mathbb{N}_0} \times \mathbb{R} \to \mathbb{R}$ be a nonnegative continuous function and assume that there exist two positive constants c and d with*

(A0) $\frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 + \frac{Td^p}{2} < \frac{(c)^p}{p}$,

such that

(A1) $\frac{\sum_{k=1}^{T} F(k,d)}{\frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 + \frac{Td^p}{2}} > p(T+1)^{\frac{p(p-2)}{4}} \frac{\sum_{k=1}^{T} \max_{|\xi| \leq c} F(k,\xi)}{c^p}$.

Then for any

$$\lambda \in \left] \frac{\frac{d^2}{(\Gamma(1-\alpha))^2} \sum_{k=1}^{T} \left((k)^{\overline{-\alpha}}\right)^2 + \frac{Td^p}{2}}{\sum_{k=1}^{T} F(k,d)}, \frac{c^p}{p(T+1)^{\frac{p(p-2)}{4}} \sum_{k=1}^{T} \max_{|\xi| \leq c} F(k,\xi)} \right[,$$

the problem (1) has at least one non-trivial solution in W.

Proof. Applying Theorem 4.1, we have the conclusion, by picking $c_1 = 0$ and $c_2 = c$. Indeed, owing to our assumptions, one has

$$\begin{aligned} a_d(0) &= \frac{\sum_{k=1}^{T} F(k,d)}{\frac{d^2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+\frac{Td^p}{2}} \\ &> \frac{\sum_{k=1}^{T}\max_{|\xi|\le c} F(k,\xi)-\sum_{k=1}^{T} F(k,d)}{\frac{(c)^p}{p}-\frac{d^2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2-\frac{Td^p}{2}} \\ &= a_d(c). \end{aligned}$$

Hence, the proof is completed. □

Theorem 4.3. *Let $h:[1,T]\to\mathbb{R}$ be a positive and essentially bounded function and $g:\mathbb{R}\to\mathbb{R}$ be a nonnegative continuous function and assume that there exist two positive constants c and d with $d^2\left(\frac{2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+T\right)<c^2$,*

such that

$$\frac{\int_0^d g(t)dt}{d^2} > \left(\frac{2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+T\right)\frac{\int_0^c g(t)dt}{c^2}.$$

Then for any

$$\lambda \in \left]\frac{\frac{2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+T}{2\sum_{k=1}^{T}h(k)}\frac{d^2}{\int_0^d g(t)dt}, \frac{1}{2\sum_{k=1}^{T}h(k)}\frac{c^2}{\int_0^c g(t)dt}\right[,$$

the problem (2) has at least one non-trivial solution in the space $\{u:[0,T+1]\to\mathbb{R}: u(0)=u(T+1)=0\}$.

Remark 4.4. We point out Theorem 4.3 is an immediate consequence of Theorem 4.2, by selecting $p=2$ and $f(k,t)=h(k)g(t)$ for all $(k,t)\in[1,T]_{\mathbb{N}_0}\times\mathbb{R}$ be separable variable which satisfies (A0) and (A1).

Proof of Theorem 1.1:
For fixed $\lambda>0$ as in the conclusion, the condition $\lim_{d\to0^+}\frac{g(d)}{d}=+\infty$ implies $\lim_{d\to0^+}\frac{\int_0^d g(t)dt}{d^2}=+\infty$, therefor there exists positive fixed constant d such that

$$\frac{\frac{2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+T}{2\sum_{k=1}^{T}h(k)}\frac{d^2}{\int_0^d g(t)dt}<\lambda.$$

On the other hand for fixed $\lambda<+\infty$ as in the conclusion, the condition $\lim_{c\to+\infty}\frac{g(c)}{c}=0$ implies $\lim_{c\to+\infty}\frac{\int_0^c g(t)dt}{c^2}=0$, so for fixed d a positive constant c satisfying

$$d^2\left(\frac{2}{(\Gamma(1-\alpha))^2}\sum_{k=1}^{T}\left((k)^{\overline{-\alpha}}\right)^2+T\right)<c^2,$$

can be chosen such that

$$\lambda<\frac{1}{2\sum_{k=1}^{T}h(k)}\frac{c^2}{\int_0^c g(t)dt}.$$

Hence, the conclusion follows from Theorem 4.3. Finally we present an example of Theorem 1.1.

Example 4.5. The following discrete boundary-value problem

$$\begin{cases} {}_{T+1}\nabla_k^\alpha\left({}_k\nabla_0^\alpha(u(k))\right)+{}_k\nabla_0^\alpha\left({}_{T+1}\nabla_k^\alpha(u(k))\right)=\lambda(2-\tanh^2 u(k))-u(k), & k\in[1,T]_{\mathbb{N}_0}, \\ u(0)=u(T+1)=0, \end{cases} \tag{18}$$

for any $\lambda \in \left]0, +\infty\right[$, has at least one non-trivial solution u_0, since $\lim_{c\to+\infty} \frac{2-\tanh^2 c}{c} = 0$ and $\lim_{d\to 0^+} \frac{2-\tanh^2 d}{d} = \infty$.

Funding
This work was supported by Sari agricultural sciences and natural resources university [grant number 03-1401-18].

References

[1] T. Abdeljawad and F. Atici, *On the Definitions of Nabla Fractional Operators,* Abstract and applied Analysis, 2012 (2012), Article ID 406757, 13 pages, doi:10.1155/2012/406757.
[2] FM. Atici and PW. Eloe, *Discrete fractional calculus with the nabla operator,* Electron. J. Qual. Theory Differ. Equ. Spec. Ed. I 2009, 3 (2009).
[3] F.M. Atici and P.W. Eloe, *Discrete fractional calculus with the nabla operator,* Electron. J. Qual. Theory Differ. Equ. Spec. Ed. I (3) (2009) 1–12.
[4] F.M. Atici and P.W. Eloe, *Two-point boundary value problems for finite fractional difference equations,* J. Difference Equ. Appl. 17(04) (2011) 445–456.
[5] F.M. Atici and S. Şengül, *Modeling with fractional difference equations,* J. Math. Anal. Appl. 369 (2010) 1–9.
[6] G. Bonanno, *A Critical point theorem via the Ekeland variational principle,* Nonl. Anal. TMA 75 (2012) 2992–3007.
[7] G. Bonanno, B. Di Bella and D. O'Regan, *Non-trivial solutions for nonlinear fourth-order elastic beam equations,* Compu. Math. Appl. 62 (2011) 1862–1869.
[8] S. Dhar and L. Kong, *A critical point approach to multiplicity results for a fractional boundary value problem,* Bulletin of the Malaysian Mathematical Sciences Society, (2020) 1–17.
[9] W. Dong, J. Xu and D. O□Regan, *Solutions for a fractional difference boundary value problem,* Adv. Difference Equ. 2013, 2013:319
[10] C. Goodrich and A. C. Peterson, *Discrete Fractional Calculus,* Springer international publishing Switzerland, New York, 2015.
[11] M. Khaleghi Moghadam and M. Avci, *Existence results to a nonlinear $p(k)$-Laplacian difference equation,* J. Difference Equ. Appl. 23(10) (2017), 1055–1068.
[12] M. Khaleghi Moghadam and J. Henderson, *Triple solutions for a dirichlet boundary value problem involving a perturbed discrete $p(k)$-laplacian operator,* Open Math. J., 15 (2017), 1075–1089.
[13] P. Li, L. Xu, H. Wang and Y. Wang, *The existence of solutions for perturbed fractional differential equations with impulses via Morse theory,* Boundary Value Problems, 2020(1),(2020) 1–13.
[14] H. L. Gray and N. F. Zhang, *On a new definition of the fractional difference,* Math. Comp. 50 , no. 182, (1988) 513–529.
[15] B. Ricceri, *A general variational principle and some of its applications,* J. Comput. Appl. Math. 113(1-2) (2000) 401–415.
[16] ZS. Xie, YF. Jin and CM. Hou, *Multiple solutions for a fractional difference boundary value problem via variational approach.* Abstr. Appl. Anal. 2012, 143914 (2012)
[17] M.E. Samei, G.K. Ranjbar and V. Hedayati,*Existence of solutions for equations and inclusions of multi-term fractional q-integro-differential with non-separated and initial boundary conditions,* J. Inequal. Appl. 2019, 273 (2019).
[18] E. Shishkina and S. Sitnik, *Transmutations, singular and fractional differential equations with applications to mathematical physics,* Academic Press, 2020.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–152, pp. 518–525

A Model of the Heroin Epidemic

Zahra Ebadi[a,*]

[a] *Faculty of Mathematics, Statistics and Computer Science, Semnan University, Semnan, Iran.*

Article Info

Keywords:
Epidemic model
Backward bifurcation
Global Stability

2020 MSC:
37C75; 37N25; 65C20.
92D30; 34D23; 34C23.

Abstract

In this article, we propose drug modeling based on the principles of mathematical epidemiology. We are investigating under what conditions this epidemic can be eradicated and under what conditions it continues. Also, the backward bifurcation and the global stability of the proposed model have been investigated.

1. Introduction

Epidemiological modeling has made great progress in recent years, but unfortunately, not much has been done in the field of drugs, as we know that drug use harms the physical, mental, or social health of an individual, family, or the whole. Therefore, since the treatment of this disease is very expensive and imposes a psychological burden on the society, it is better to provide solutions for prevention before the disease occurs. The White-Comiskey model in [6] , on heroin epidemics, was one of the first mathematical modeling of drugs, and this model was reviewed by Malone and Straughan in [9].
The mathematical epidemiology of infectious diseases is well developed and can be found in the works of Bailey [2], Anderson and May [1], Murray [5] and Brauer and Castillo-Chavez [3].

In this paper, we have upgraded White-Comiskey's model to four compartments. Then we examined the threshold quantity, backward bifurcation and overall stability of the new model.

2. Model Formulation and Basic Properties

[>=stealth,scale=0.9]
[thick, ->] (0,-1)–(2,-1); (1,-0.5) node Λ;
[thick, ->] (4,-1)–(6,-1); (5,-0.5) node βIS;
[thick, ->] (8,-1)–(10,-1); (9,-0.5) node αE;

*Talker
Email address: ebadi.zahra@semnan.ac.ir (Zahra Ebadi)

[thick, ->] (12,-1)–(14,-1); (13,-0.5) node γI;
[thick] (2,0)–(4,0)–(4,-2)–(2,-2)–(2,0); (3,-1) node S;
[thick, ->] (3,0)–(3,1.5); (3,0.75) node[right] μ;
[thick] (6,0)–(8,0)–(8,-2)–(6,-2)–(6,0); (7,-1) node E;
[thick, ->] (7,0)–(7,1.5); (7,0.75) node[right] μ;
[thick] (10,0)–(12,0)–(12,-2)–(10,-2)–(10,0); (11,-1) node I;
[thick, ->] (11,0)–(11,1.5); (11,0.75) node[right] μ;
[thick] (14,0)–(16,0)–(16,-2)–(14,-2)–(14,0); (15,-1) node T;
[thick, ->] (15,0)–(15,1.5); (15,0.75) node[right] μ;

In this model, society is divided into four compartments: S susceptible individuals at risk of using drugs, E the exposed stage, the stage that the patient himself denies, I drug users, and T drug users under treatment. $S(t)$, $E(t)$, $I(t)$, and $T(t)$ show the number of compartments respectively.

Table 1: Model parameters

Symbol	Description
Λ	Recruitment of susceptible population.
μ	Natural mortality rate.
β	Probability of drug user per unit time.
α	Progression rate for addiction.
γ	The rate of treatment of drug users.

According to the model flow diagram in the figure on the previous page, The ODE system (1) obtained:

$$\begin{cases} \dfrac{dS}{dt} = \Lambda - \beta SI - \mu S \\ \dfrac{dE}{dt} = \beta SI - (\mu+\alpha)E \\ \dfrac{dI}{dt} = \alpha E - (\mu+\gamma)I \\ \dfrac{dT}{dt} = \gamma I - \mu T \end{cases} \tag{1}$$

We first prove the positivity of solution (1)

Theorem 2.1. *If initial data $S(0) > 0$, $E(0) > 0$, $I(0) > 0$ and $T(0) > 0$, then the solution $(S(t), E(t), I(t), T(t))$ of (1) is positive for all $t \geq 0$*

Proof. Let $(S(t), E(t), I(t), T(t))$ be the solution of the system (1) with initial data $S(0) > 0$, $E(0) > 0$, $I(0) > 0$ and $T(0) > 0$. Suppose that the conclusion is not true, then there is a $t^* > 0$ such that,

$$min\{S(t^*), E(t^*), I(t^*), T(t^*)\} = 0$$

and

$$min\{S(t), E(t), I(t), T(t)\} > 0$$

for all $t \in [0, t^*)$. If $min\{S(t^*), E(t^*), I(t^*), T(t^*)\} = S(t^*)$, then we have, $\dfrac{dS}{dt} \geq -\beta SI - \mu S$, for all $t \in [0, t^*)$. Hence, $0 = S(t^*) \geq S(0)exp(-\int_0^{t^*} (\beta I(t) + \mu)\, dt) > 0$, which leads to a contradiction. Similarly, we can obtain contradictions when, $min\{S(t^*), E(t^*), I(t^*), T(t^*)\}$, is equal to other variables of the system. This completes the proof. ☐

This system has a unique DFE $P_0 = (S^*,\ E^*,\ I^*,\ T^*) = (\dfrac{\Lambda}{\mu},\ 0\ , 0,\ 0)$ and has a Jacobian matrix of P_0 as follows:

$$J(P_0) = \begin{pmatrix} -\mu & 0 & -\beta\frac{\Lambda}{\mu} & 0 \\ 0 & -(\mu+\alpha) & \beta\frac{\Lambda}{\mu} & 0 \\ 0 & \alpha & -(\mu+\gamma) & 0 \\ 0 & 0 & \gamma & -\mu \end{pmatrix}$$

$-\mu$, $-\mu$ and $(\mu+\alpha)(\mu+\gamma) - \alpha\beta\frac{\Lambda}{\mu}$ are the eigenvalues of the Jacobin matrix. We define the basic reproduction number as $R_0 = \dfrac{\alpha\beta\Lambda}{\mu(\mu+\alpha)(\mu+\gamma)}$.

Theorem 2.2. *DFE P_0 is unstable when $R_0 > 1$ and DFE P_0 is asymptotically stable when $R_0 < 1$.*

3. Endemic Equilibrium and Backward bifurcation

The endemic equilibrium points of (1) satisfy the following system,

$$\begin{cases} \Lambda - \beta S^* I^* - \mu S^* = 0 \\ \beta S^* I^* - (\mu + \alpha) E^* = 0 \\ \alpha E^* - (\mu + \gamma) I^* = 0 \\ \gamma I^* - \mu T^* = 0 \end{cases} \tag{2}$$

We have (2) relations

$$E^* = \frac{\mu + \gamma}{\alpha} I^*, \quad T^* = \frac{\gamma}{\mu} I^*, \quad S^* = \frac{(\mu + \alpha)(\mu + \gamma)}{\alpha \beta} = \frac{\Lambda}{\mu R_0}, \quad I^* = \frac{\Lambda}{\beta S^*} - \frac{\mu}{\beta} = \frac{\mu}{\beta}(R_0 - 1)$$

In $R_0 > 1$, there is exactly one endemic equilibrium point.

Which yields that, I^* is the positive root of

$$F(I^*) = A(I^*)^2 + BI^* + C = 0 \tag{3}$$

Where

$$\begin{aligned} A &= -\beta^2 \Big(\frac{1}{R_0}(\mu + \alpha) + \beta \Lambda (\mu + \gamma) \Big) \\ B &= \beta^2 \Lambda^2 - \mu \Big(\frac{1}{R_0}(\mu + \alpha) + \beta \Lambda (\mu + \gamma) \Big) \\ C &= 0 \end{aligned}$$

The endemic steady state exists when roots of (3) are positive real numbers. Now since $A < 0$, we must have $B \geq 0$, $\Delta > 0$. Consider the discriminant, $\Delta = B^2 - 4AC = B^2 > 0$, solving $\Delta = 0, (B = 0)$ in terms of R_0, we obtain:

$$R_0^c = \frac{\mu(\mu + \alpha)}{\beta \Lambda \Big(\beta \Lambda - \mu(\mu + \gamma) \Big)}$$

Theorem 3.1. *If $R_0 > 1$, system (1) has a unique endemic equilibrium point, and when $R_0^c < R_0 < 1$ it has two endemic equilibrium points.*

Typically in most epidemics, if $R_0 < 1$ and the initial values of all model compartments are in the area of attraction of the DFE P_0, the disease can be eliminated. Also, in some epidemiological models in the range $R_0 < 1$, there are also endemic balance points that indicate that $R_0 < 1$ is not enough to eliminate the disease. In such problems, backward bifurcation occurs. You can refer to [8]. Now, we use the Castillo-Chavez and Song theorem, see [4], to determine the conditions for the occurrence of backward bifurcation in (1).
Let $s = x_1$, $I = x_2$, $E = x_3$ and $T = x_4$. System (1) transforms to the following system:

$$\begin{cases} \dfrac{dx_1}{dt} = \Lambda - \beta x_1 x_2 - \mu x_1 = f_1 \\ \dfrac{dx_2}{dt} = \alpha x_3 - (\mu + \gamma) x_2 = f_2 \\ \dfrac{dx_3}{dt} = \beta x_1 x_2 - (\mu + \alpha) x_3 = f_3 \\ \dfrac{dx_4}{dt} = \gamma x_2 - \mu x_4 = f_4 \end{cases} \tag{4}$$

Now we apply Castillo-Chavez and Song theorem to show that in (4), backward bifurcation occurs when $R_0 = 1$. The relation $R_0 = 1$ can be interpreted in term of β, as $\beta = \beta^* = \dfrac{\mu(\mu+\alpha)(\mu+\gamma)}{\alpha\Lambda}$. The eigen values of the Jacobian matrix,

$$J(P_0, \beta^*) = \begin{pmatrix} -\mu & -\beta^* \frac{\Lambda}{\mu} & 0 & 0 \\ 0 & \beta^* \frac{\Lambda}{\mu} & -(\mu+\alpha) & 0 \\ 0 & -(\mu+\gamma) & \alpha & 0 \\ 0 & \gamma & 0 & -\mu \end{pmatrix}$$

Let$v = (v_1, v_2, v_3, v_4)$, be the left eigenvector of $\mathcal{A}$ associated with zero eigenvalue is founded by, $v\mathcal{A} = 0$, and turns out to be

$$v = (0, \alpha, \alpha + \mu, 0)$$

On the other hand, $w = (w_1, w_2, w_3, w_4)^T$, be the right eigenvector of $\mathcal{A}$ associated with eigenvalue $\lambda_4 = 0$, founded by, $\mathcal{A}w = 0$. Computation of the solution of this linear system yields: $w = (-\dfrac{(\mu+\alpha)(\mu+\gamma)}{\alpha\mu}, \alpha, \mu+\gamma, \gamma)^T$.

Now we compute the quantities **a** and **b** of Castillo-Chavez and Song theorem , that is,

$$\begin{aligned} \mathbf{a} &= \sum_{k,i,j=1}^{n} v_k w_i w_j \frac{\partial^2 f_k}{\partial x_i \partial x_j}(P_0, \beta^*), \\ &= \sum_{i,j=1}^{4} \alpha w_i w_j \frac{\partial^2 f_2}{\partial x_i \partial x_j}(P_0, \beta^*) + \sum_{i,j=1}^{4} (\mu+\alpha) w_i w_j \frac{\partial^2 f_3}{\partial x_i \partial x_j}(P_0, \beta^*) \\ &= 2(\mu+\alpha) w_1 w_2 = -2\frac{(\mu+\alpha)^2(\mu+\gamma)}{\mu} < 0 \end{aligned}$$

and

$$\begin{aligned} \mathbf{b} &= \sum_{k,i=1}^{n} v_k w_i \frac{\partial^2 f_k}{\partial x_i \partial \phi}(P_0, \beta^*) = \sum_{i=1}^{4} w_i \frac{\partial^2 f_2}{\partial x_i \partial \beta}(P_0, \beta^*) + \sum_{i=1}^{4} w_i \frac{\partial^2 f_3}{\partial x_i \partial \beta}(P_0, \beta^*) \\ &= w_1 x_2 + w_2 x_1 = w_2 \frac{\Lambda}{\mu} = \frac{\alpha\Lambda}{\mu} > 0 \end{aligned}$$

We observe that **b** is positive, and **a** is negative, Using part (4) in the Castillo-Chavez and Song theorem, the back bifurcation occurs.

4. Global stability of equilibrium points

Theorem 4.1. *If $R_0 < 1$, then the system (1) has a unique disease-free equilibrium point, which is globally asymptotically stable. If $R_0 > 1$, then the endemic equilibrium is globally asymptotically stable.*

Proof. We define the following Lyapunov function for $R_0 < 1$.

$$V = \kappa\Big(S - S^* - S^* \ln \frac{S}{S^*}\Big) + \frac{1}{\mu+\alpha} E + \frac{1}{\alpha} I \tag{5}$$

Where $\kappa > 0$.

$$\begin{aligned} \frac{dV}{dt} &= \frac{\partial V}{\partial S}.\frac{dS}{dt} + \frac{\partial V}{\partial E}.\frac{dE}{dt} + \frac{\partial V}{\partial I}.\frac{dI}{dt} = \kappa\Big(1 - \frac{S^*}{S}\Big)S' + \frac{1}{\mu+\alpha}E' + \frac{1}{\alpha}I' \\ &= \kappa\Big(1 - \frac{S^*}{S}\Big)(\Lambda - \beta SI - \mu S) + \Big(\frac{1}{\mu+\alpha}\Big)(\beta SI - (\mu+\alpha)E) + \frac{1}{\alpha}(\alpha E - (\mu+\gamma)I) \\ &= 2\kappa\Lambda - \beta\kappa SI - \kappa\mu S - \frac{\Lambda^2\kappa}{\mu S} + \frac{\Lambda\beta\kappa}{\mu}I + \frac{\beta}{\mu+\alpha}SI - \frac{\mu+\gamma}{\alpha}I \end{aligned} \tag{6}$$

By choosing $\kappa = \dfrac{1}{\mu+\alpha}$, we have

$$\frac{dV}{dt} = -\kappa\Lambda\Big(\frac{\Lambda}{\mu S} + \frac{\mu S}{\Lambda} - 2\Big) + \frac{\mu+\gamma}{\alpha}(R_0 - 1)I$$

Assume $A = \dfrac{\Lambda}{\mu S} > 0$, We have $A + \dfrac{1}{A} - 2 = \dfrac{A^2 - 2A + 1}{A} = \dfrac{(A-1)^2}{A} > 0$. Hence we have $V' < 0$ for all $(S, E, I) \neq (S^*, 0, 0)$. So, by Lyapunov's theorem, the disease-free equilibrium is globally asymptoticlly stable.

We define the following Lyapunov function for $R_0 > 1$.

$$V = \kappa_1\Big(S - S^* - S^* \ln\frac{S}{S^*}\Big) + \kappa_2\Big(E - E^* - E^* \ln\frac{E}{E^*}\Big) + \kappa_3\Big(I - I^* - I^* \ln\frac{I}{I^*}\Big) \tag{7}$$

Where $\kappa_1,\ \kappa_2,\ \kappa_3 > 0$.

$$\begin{aligned} \frac{dV}{dt} &= \frac{\partial V}{\partial S}.\frac{dS}{dt} + \frac{\partial V}{\partial E}.\frac{dE}{dt} + \frac{\partial V}{\partial I}.\frac{dI}{dt} \\ &= \kappa_1\Big(1 - \frac{S^*}{S}\Big)S' + \kappa_2\Big(1 - \frac{E^*}{E}\Big)E' + \kappa_3\Big(1 - \frac{I^*}{I}\Big)I' \\ &= \kappa_1\Big(1 - \frac{S^*}{S}\Big)(\Lambda - \beta SI - \mu S) + \kappa_2\Big(1 - \frac{E^*}{E}\Big)(\beta SI - (\mu+\alpha)E) \\ &+ \kappa_3\Big(1 - \frac{I^*}{I}\Big)(\alpha E - (\mu+\gamma)I) \end{aligned} \tag{8}$$

By placement $\Lambda = \beta S^* I^* + \mu S^*$ and $\kappa_1 = \kappa_2$; $\kappa_3 = \kappa_2 \frac{\mu+\alpha}{\alpha}$, We have

$$\begin{aligned} \frac{dV}{dt} &= -\kappa_1\frac{(S-S^*)^2}{S} + \kappa_1\beta S^* I^*\Big(3 - \frac{S^*}{S} - \frac{E^* SI}{ES^* I^*} - \frac{I^* E}{IE^*}\Big) \\ &+ (\kappa_1\beta S^* - \kappa_3(\mu+\gamma))I + (\kappa_3\alpha - \kappa_2(\mu+\alpha))E \end{aligned} \tag{9}$$

by choose $\kappa_1 = \kappa_2 = 1$, So $\kappa_3 = \dfrac{\mu+\alpha}{\alpha}$

$$\frac{dV}{dt} = -\frac{(S-S^*)^2}{S} + \beta S^* I^*\Big(3 - \frac{S^*}{S} - \frac{E^* SI}{ES^* I^*} - \frac{I^* E}{IE^*}\Big) \tag{10}$$

suppose $x_1 = \frac{S^*}{S}$, $x_2 = \frac{E^* SI}{ES^* I^*}$, $x_3 = \frac{I^* E}{IE^*}$, We have $\frac{x_1+x_2+x_3}{3} \geq \sqrt[3]{x_1\, x_2\, x_3} = 1$ in the result $3 - \frac{S^*}{S} - \frac{E^* SI}{ES^* I^*} - \frac{I^* E}{IE^*} \leq 0$, So $\frac{dV}{dt} \leq 0$

Now we check when $\frac{dV}{dt} = 0$ is equal to zero, it has two states $S^* \neq S$,It is obvious. And when $S^* = S$ in the result

$$2 - \frac{E^* I}{EI^*} - \frac{I^* E}{IE^*} = 0 \Rightarrow 2 - A - \frac{1}{A} = 0 \Rightarrow A = 1 \Rightarrow I = I^*, E = E^*.$$

□

Suppose that in the system (1), there exists a unique endemic equilibrium point. In system (1), the Jacobian matrix at point $(S,\, E,\, I,\, T)$ is as follows:

$$J = \frac{\partial f}{\partial x} = \begin{bmatrix} -\beta I - \mu & 0 & -\beta S & 0 \\ \beta I & -(\mu+\alpha) & \beta S & 0 \\ 0 & \alpha & -(\mu+\gamma) & 0 \\ 0 & 0 & \gamma & -\mu \end{bmatrix} \tag{11}$$

The second order complex matrix of $J = \frac{\partial f}{\partial x}$, is in the form of $M = J^{[2]} = \Big[B_{ij}\Big]_{6\times 6}$. Its arrays are as follows:
$M_{11} = -\beta I - 2\mu - \alpha, \quad M_{12} = \beta S, \quad M_{14} = -\beta S, \quad M_{13} = M_{15} = M_{16} = 0,$
$M_{21} = \alpha, \quad M_{22} = -\beta I - 2\mu - \gamma, \quad M_{23} = M_{24} = M_{25} = M_{26} = 0,$
$M_{32} = \gamma, \quad M_{33} = -\beta I - 2\mu, \quad M_{36} = -\beta S, \quad M_{31} = M_{33} = M_{35} = 0$
$M_{42} = \beta_1 I, \quad M_{44} = -2\mu - \alpha - \gamma, \quad M_{41} = M_{43} = M_{45} = M_{46} = 0,$
$M_{53} = \beta I, \quad M_{54} = \gamma, \quad M_{55} = -2\mu - \alpha, \quad M_{56} = \beta S, \quad M_{51} = M_{52} = 0$
$M_{65} = \alpha, \quad M_{66} = -2\mu - \gamma, \quad M_{61} = M_{62} = M_{63} = M_{64} = 0.$
Consider the matrix function $P = [p_{ij}]$, where $p_{11} = p_{22} = p_{44} = p_{55} = \frac{1}{I},\ p_{33} = p_{66} = 1$ and the rest of the arrays are zero. Therefore this matrix is obtained $P_f P^{-1} = -diag(\frac{I'}{I}, \frac{I'}{I}, 0, \frac{I'}{I}, \frac{I'}{I}, 0)$, thus, $Q = P_f P^{-1} + PMP^{-1} = \Big[A_{ij}\Big]_{6\times 6}$, in which
$A_{11} = -\beta I - \mu - \alpha - \alpha\frac{E}{I} + \gamma, \quad A_{12} = \beta S, \quad A_{14} = -\beta S, \quad A_{13} = A_{15} = A_{16} = 0,$
$A_{21} = \alpha, \quad A_{22} = -\beta I - \mu - \alpha\frac{E}{I}, \quad A_{23} = A_{24} = A_{25} = A_{26} = 0,$
$A_{32} = \gamma, \quad A_{33} = -\beta I - 2\mu, \quad A_{36} = -\beta S, \quad A_{31} = A_{33} = A_{35} = 0$
$A_{42} = \beta_1 I, \quad A_{44} = -\mu - \alpha - \alpha\frac{E}{I}, \quad A_{41} = A_{43} = A_{45} = A_{46} = 0,$
$A_{53} = \beta I, \quad A_{54} = \gamma, \quad A_{55} = -\mu - \alpha - \alpha\frac{E}{I} + \gamma, \quad A_{56} = \beta S, \quad A_{51} = A_{52} = 0$
$A_{65} = \alpha, \quad A_{66} = -2\mu - \gamma, \quad A_{61} = A_{62} = A_{63} = A_{64} = 0.$
Now, using the norm introduced in [7],

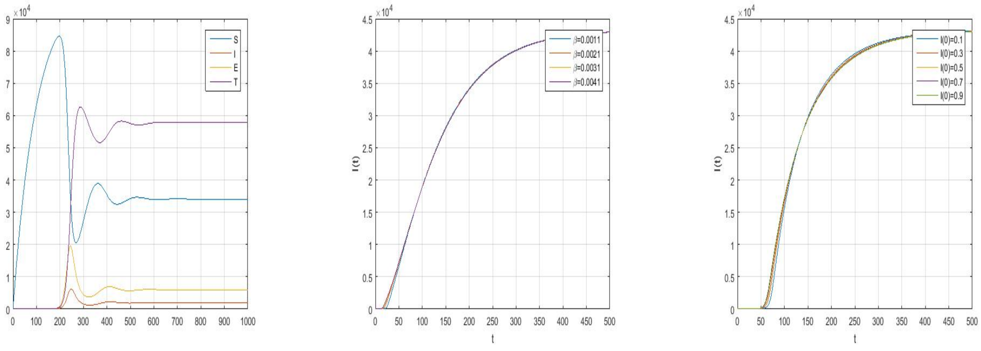
Lemma 4.2. *There is a constant $\chi > 0$, for which $D_+ \parallel z \parallel \leq -\chi \parallel z \parallel$ for all $z \in \mathbb{R}^6$ and all $S,\, E,\, I,\, T$, where z is the solution of $\frac{dz}{dt} = Q(\phi_t(l))z$, provided that $0 < \beta - \gamma < -(\mu+\alpha)$*

Theorem 4.3. *Suppose the inequalities in Lemma 4.2 holds, then:*
(1) when the only equilibrium point is the drug-free equilibrium P_0, then all solutions tend to P_0;
(2) when $R_0 > 1$, then all solutions of (1) tends to the unique endemic equilibrium point;
(3) when there are two endemic equilibrium points, which occurs when $R_0^c < R_0 < 1$, solutions of the system either tend to the drug-free equilibrium P_0 or tend to the upper equilibrium point.

5. Numerical Simulation

In this section, we will simulate the system using MATLAB software, so that the obtained analytical results can be seen numerically. We present two cases.
Case 1. $\frac{\mu(\mu+\alpha)}{\beta\Lambda\Big(\beta\Lambda - \mu(\mu+\gamma)\Big)} < R_0 < 1.$

We choose $\Lambda = 1000,\ \mu = 10^{-1},\ \alpha = 10^{-1}, \beta = 10^{-5},\ \gamma = 3\times 10^{-1}$. In this case, $R_0 \simeq 0.125$ and $\frac{\mu(\mu+\alpha)}{\beta\Lambda\Big(\beta\Lambda - \mu(\mu+\gamma)\Big)} \simeq -2.08$. See figure 1. Case 2. $R_0 > 1, 0 < \beta - \gamma < -(\mu+\alpha)$. See figure 2.



(a) (b) (c)

Fig. 1. (a) shows the plot of the solution of system. (b) shows the sensitivity of $I(t)$ with respect to β respectively. (c) shows the convergence of the $I(t)$, (infectious compartment) of five solution curve of the system to the DFE.

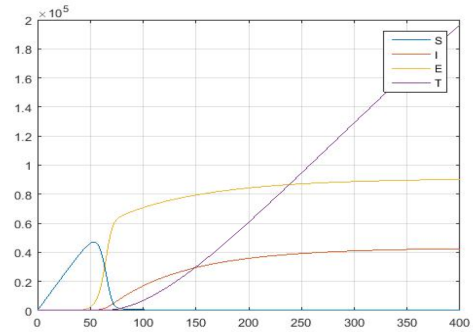


Fig. 2. shows the plot of the solution of system.

6. Conclusion

In this article, we have added two compartments to the model of White and Comiskey heroin epidemics for hard drug users and light drug users. We scrutinized the existence and local stability, and global stability of the model. We have shown that DFE is stable locally and globally in suitable conditions. We obtained sufficient conditions for the local stability and global stability of endemic equilibrium points using compound matrices. We also proved the occurrence of backward bifurcation. Backward bifurcation indicates that the presence of $R_0 < 1$ is not sufficient to control the epidemic.

References

[1] ANDERSON, R. M., MAY, R. M. (1991). Infectious Diseases of Humans, Dynamics and Control. *Oxford University Press*.
[2] BAILEY, N. (1975). The Mathematical Theory of Infectious Diseases. Charles Griffin, London.
[3] BRAUER, F., CASTILLO-CHAVEZ, C. (2000). Mathematical Models in Population Biology and Epidemiology. *Springer, New York*.
[4] CASTILLO-CHAVEZ, C. (2004). Song, B., Dynamical models of tuberculosis and their applications. *Math. Biosci. Eng*, (2), 361-404.
[5] MURRAY, J. D. (2004). Mathematical Biology I and II. *Springer, New York*.s
[6] COMISKEY, C., WHITE , E. (2007). Heroin epidemics treatment and ODE modelling. Math. Biosci., (208), 312-324.
[7] GUMEL, A.B., MCCLUSKEY, C.C., WATMOUGH, J. (2006). An SVEIR model for assessing potential impact of an imperfect anti-SARS vaccine. *Math. Biosci. Eng.*, (3), 485-512.
[8] MARTCHEVA, M. (2015). An Introduction to Mathematical Epidemiology. *Springer, New York*.
[9] MULONE, G., STRAUGHAN, B. (2009). A note on heroin epidemics.*Math. Biosci.*, (218), 138-141.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–165, pp. 526–528

Inverse spectral problems for a conformable fractional Sturm-Liouville equation

Yasser Khalili[a,*], Mohsen Khaleghi Moghadam[a]

[a] *Department of Basic Sciences, Sari Agricultural Sciences and Natural Resources University, 578 Sari, Iran*

Article Info

Keywords:
Inverse problem
Conformable
Sturm-Liouville equation
Spectrum

2020 MSC:
34A55
34A08
34B24

Abstract

In this work, we consider a conformable fractional Sturm-Liouville equation with spectral boundary conditions which includes conformable fractional derivatives. We prove the uniqueness theorem for the inverse problem of the conformable fractional Sturm-Liouville equation from the Weyl function.

1. Introduction

Inverse problems for Sturm-Liouville equations are widely used in various models of quantum and classical mechanics to express physical phenomena and natural sciences [5]. The conformable derivative has been employed in various fields such as the control theory of dynamical systems and other areas, and for this reason, the research has increased in this field in the recent years (see [2]).
The inverse problem for fractional Sturm-Liouville equations is a topic that we can work on it. In this paper, we investigate the inverse spectral problem for the fractional Sturm-Liouville equation which includes the conformable derivative of order $\alpha \in (0, 1]$. By the spectral mappings method, we study the inverse spectral problem for the conformable fractional Sturm-Liouville equation.
We remark that some definitions and properties of the conformable fractional calculus can be found in [6, 7].

2. Preliminaries and main results

Let us consider the conformable fractional Sturm-Liouville equation

$$-D_x^\alpha D_x^\alpha y + q(x)y = \lambda y, \quad x \in (0, \pi), \tag{1}$$

* Talker
Email address: y.khalili@sanru.ac.ir (Yasser Khalili)

with the spectral boundary conditions

$$U(y) := D_x^\alpha y(0) - (h_1\rho + h_0)y(0) = 0, \tag{2}$$
$$V(y) := D_x^\alpha y(\pi) = 0. \tag{3}$$

Here D_x^α is the conformable fractional derivative of order $\alpha \in (0,1]$. The coefficients h_0, h_1 are complex numbers and $\lambda = \rho^2$ is a spectral parameter. The complex-valued function $q(x)$ is continuous on $[0,\pi]$. The boundary value problem (1)-(3) is denoted by $L_\alpha = L_\alpha(q(x), h_0, h_1)$.
Let $\varphi(x,\rho)$ be the solution of Eq. (1) under the initial conditions

$$\varphi(0,\rho) = 1, \ \ D_x^\alpha \varphi(0,\rho) = h_1\rho + h_0.$$

For each fixed x, this function and its derivative with respect to x are entire in ρ. From [1], we have the following asymptotic formulae for sufficiently large $|\rho|$ and $x \in [0,\pi]$,

$$\varphi(x,\rho) = \sqrt{1+h_1^2}\cos\left(\frac{\rho}{\alpha}x^\alpha - \sigma_1\right) + O\left(\frac{1}{\rho}\exp\left(\frac{|\Im\rho|}{\alpha}x^\alpha\right)\right), \tag{4}$$
$$D_x^\alpha\varphi(x,\rho) = -\rho\sqrt{1+h_1^2}\sin\left(\frac{\rho}{\alpha}x^\alpha - \sigma_1\right) + O\left(\exp\left(\frac{|\Im\rho|}{\alpha}x^\alpha\right)\right), \tag{5}$$

where $\sigma_1 = \frac{1}{2i} ln \frac{i-h_1}{i+h_1}$.
Consider the characteristic function of L_α as

$$\Delta_\alpha(\rho) = D_x^\alpha \varphi(\pi),$$

which is the entire function in ρ. Therefore by using (4) and (5), we can give for sufficiently large $|\rho|$,

$$\Delta_\alpha(\rho) = \rho\sqrt{1+h_1^2}\sin\left(\frac{\rho}{\alpha}\pi^\alpha - \sigma_1\right) + O\left(\exp\left(\frac{|\Im\rho|}{\alpha}\pi^\alpha\right)\right). \tag{6}$$

Assume that $\delta > 0$ be fixed and $C_\delta > 0$ be a constant. Put $G_\delta := \{\rho; \mid \rho - \rho_n \mid \geq \delta, \ \forall n\}$. Taking (6) and the known technique [3], one gets

$$|\Delta_\alpha(\rho)| \geq C_\delta |\rho| \exp\left(\frac{|\Im\rho|}{\alpha}\pi^\alpha\right), \quad \rho \in G_\delta. \tag{7}$$

By the Rouche's theorem [4] and the known technique [3], one can give that the roots of the characteristic function $\Delta_\alpha(\rho)$ have the asymptotics

$$\rho_n = \frac{\alpha}{\pi^\alpha}(n\pi + \sigma_1) + o(1),$$

for large enough n.
Let $\psi(x,\rho)$ and $S(x,\rho)$ be the solution of Eq. (1) under the initial conditions

$$\psi(\pi,\rho) = 1, \ \ D_x^\alpha\psi(\pi,\rho) = 0,$$
$$S(0,\rho) = 0, \ \ D_x^\alpha S(0,\rho) = 1.$$

Define the meromorphic function

$$\phi(x,\rho) = -\frac{\psi(x,\rho)}{\Delta_\alpha(\rho)}, \tag{8}$$

which is called the Weyl solution of the boundary value problem L_α. Also considering the initial conditions at $x = 0$, we can give

$$\phi(x,\rho) = S(x,\rho) + M(\rho)\varphi(x,\rho), \tag{9}$$

in which

$$M(\rho) := \phi(0,\rho),$$

and is called the Weyl function of the boundary value problem L_α.
Problem Given the Weyl function $M(\rho)$, construct the coefficients of the boundary value problem L_α.
To show the uniqueness theorem in this section, alongside $L_\alpha = L_\alpha(q, h_1, h_0)$, a boundary value problem $\widetilde{L}_\alpha = L_\alpha(\widetilde{q}, , \widetilde{h}_1, \widetilde{h}_0)$ of the similar form (1)-(3) is considered. We suppose that if α signifies an object relevant to L, then $\widetilde{\alpha}$ will signify the similar object relevant to $\widetilde{L}$.

Theorem 2.1. *Let $M(\rho) = \widetilde{M}(\rho)$. Then $q(x) = \widetilde{q}(x)$, a.e. on $(0, \pi)$ and $h_0 = \widetilde{h}_0$, $h_1 = \widetilde{h}_1$.*

Proof. Let us consider the matrix $P(x, \rho) = (P_{j,k}(x, \rho))_{j,k=1,2}$ and then

$$\varphi(x, \rho) = P_{11}(x, \rho)\widetilde{\varphi}(x, \rho) + P_{12}(x, \rho)D_x^\alpha \widetilde{\varphi}(x, \rho), \tag{10}$$

$$\phi(x, \rho) = P_{11}(x, \rho)\widetilde{\phi}(x, \rho) + P_{12}(x, \rho)D_x^\alpha \widetilde{\phi}(x, \rho). \tag{11}$$

Since the Wronskian of the functions $\varphi(x, \rho)$ and $\phi(x, \rho)$ is one, we have

$$P_{11}(x, \rho) = \varphi(x, \rho)D_x^\alpha \widetilde{\phi}(x, \rho) - \phi(x, \rho)D_x^\alpha \widetilde{\varphi}(x, \rho), \tag{12}$$

$$P_{12}(x, \rho) = \phi(x, \rho)\widetilde{\varphi}(x, \rho) - \varphi(x, \rho)\widetilde{\phi}(x, \rho). \tag{13}$$

Taking (9) and the hypothesis of the theorem, we result that $P_{1k}(x, \rho),\ k = 1, 2$ are entire in ρ for each fixed x. Also by using (12) and the property of the Wronskian for $\varphi(x, \rho)$ and $\phi(x, \rho)$, we will have

$$P_{11}(x, \rho) - 1 = (\varphi(x, \rho) - \widetilde{\varphi}(x, \rho))\, D_x^\alpha \widetilde{\phi}(x, \rho) - (\phi(x, \rho) - \widetilde{\phi}(x, \rho))D_x^\alpha \widetilde{\varphi}(x, \rho). \tag{14}$$

Denote $G_\delta^0 = G_\delta \cap \widetilde{G}_\delta$. By taking (7), (8), (13) and (14), we get as sufficiently large $\rho \in G_\delta^0$,

$$|P_{11}(x, \rho) - 1| \le C_\delta |\rho|^{-1}, \quad |P_{12}(x, \rho)| \le C_\delta |\rho|^{-1}.$$

Therefore $P_{11}(x, \rho) = 1$ and $P_{12}(x, \rho) = 0$. Now together with (10) and (11), this yields $\varphi(x, \rho) = \widetilde{\varphi}(x, \rho)$ and $\phi(x, \rho) = \widetilde{\phi}(x, \rho)$ for all $x,\ \rho$. Thus $q(x) = \widetilde{q}(x)$ a.e. on $(0, \pi)$ and $h_0 = \widetilde{h}_0$, $h_1 = \widetilde{h}_1$. The proof is completed. □

References

[1] Ozkan AS Adalar I. Inverse problems for a conformable fractional sturm-liouville operator. *J Inverse and Ill-posed Problems*, 28(6):775–782, 2020.
[2] Yurko VA Buterin SA. Fractional newton mechanics with conformable fractional derivative. *J Inverse Ill-Posed Probl*, 20:855–881, 2015.
[3] Yurko VA Freiling G. *Inverse Sturm-Liouville Problems and their Applications*. NOVA Scince, New York, 2001.
[4] Conway JB. *Functions of One Complex Variable*. Springer-Verlag, New York, 1995.
[5] Anderssen RS. The effect of discontinuities in density and shear velocity on the asymptotic overtone structure of tortional eigenfrequencies of the earth. *Geophys J R Astr Soc*, 50:303–309, 1997.
[6] Abdeljawad T. On conformable fractional calculus. *J Comput Appl Math*, 279:57–66, 2015.
[7] Chung WS. On conformable fractional calculus. *J Comput Appl Math*, 290:150–158, 2015.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–174, pp. 529–533

Fractional calculations and their applications

Mohammad Molaei[a,], Mohammad Jafari[a]

[a] *Department of Science, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
Fractional Laplace transform; conformable fractional derivative

Abstract

In this paper, we use the fractional Laplace transform to solving a class of ordinary differential equations (ODEs) as well as conformable fractional differential equations (CFDEs). We applying the fractional Laplace transform to convert some of (ODEs) of second-order, as well as conformable fractional differential equations into linear differential equation of first-order. This is done by using of the fractional Laplace transform of $(\alpha + \beta)$ order.

1. Introduction

Fractional calculus has garnered significant attention from researchers in the past and present centuries. However, in recent decades, there has been substantial progress in both the theory and application of fractional calculus and fractional differential equations, thanks to their potential for powerful applications. One particular area that requires further investigation is the conformable fractional derivative. In our study, we explored the fractional Laplace transform, which is compatible with this type of fractional derivative. Unlike traditional fractional calculus, the conformable fractional derivative exhibits favorable behavior in the product rule and chain rule, with less complicated formulas. We present new findings that are valuable in the theory of conformable fractional differential equations. Additionally, we have developed the fractional Laplace transform method as a robust approach for obtaining exact solutions to differential equations, specifically conformable time fractional differential equations. By utilizing the fractional Laplace transform, we can convert certain ordinary differential equations and conformable fractional differential equations into first-order ordinary differential equations.

1.1. Basic definitions of fractional derivative and fractional Laplace transform

While fractional derivatives have gained popularity in recent years, their origins can be traced back to the late 17th century. Over time, various definitions for fractional derivatives have been proposed. One notable definition is the conformable fractional derivative, which was initially introduced by Khalil et al. [2]. Subsequently, Abdeljawad [1] extended the concept by presenting fractional versions of the chain rule, exponential functions, Gronwall's inequality, Taylor power series expansions, and the fractional Laplace transform specifically designed for the conformable derivative. In their work [2], Khalil et al. introduced a novel type of fractional derivatives, outlined as follows:

Definition 1.1. The left conformable fractional derivative of order $0 < \alpha \leq 1$ starting from $a \in \mathbb{R}$ of function $u : [a, +\infty) \to \mathbb{R}$, is defined by

$$({}_tT^a_\alpha u)(t) = \lim_{\epsilon \to 0} \frac{u\left(t + \epsilon(t-a)^{1-\alpha}\right) - u(t)}{\epsilon}. \tag{1}$$

When $a=0$, we have:

$$({}_tT_\alpha^0 u)(t) = {}_tT_\alpha u(t) = \lim_{\epsilon\to 0} \frac{u\left(t+\epsilon t^{1-\alpha}\right)-u(t)}{\epsilon}.$$

If $(T_\alpha^a u)(t)$ exists on $(a,+\infty)$, then $(T_\alpha^a u)(a) = \lim_{t\to a^+}(T_\alpha^a u)(t)$. If $(T_\alpha^a u)(t_0)$ exists and is finite, then we say that u is left $\alpha-$differentiable at t_0.
The right conformable fractional derivative of order $0<\alpha\le 1$ terminating at $b\in\mathbb{R}$ of function $u:(-\infty,b]\to\mathbb{R}$, is defined by

$$({}^bT_\alpha u)(t) = -\lim_{\epsilon\to 0}\frac{u\left(t+\epsilon(b-t)^{1-\alpha}\right)-u(t)}{\epsilon}. \tag{2}$$

If $({}^bT_\alpha u)(t)$ exists on $(-\infty,b)$, then $({}^bT_\alpha u)(a) = \lim_{t\to b^-}({}^bT_\alpha u)(t)$. If $({}^bT_\alpha u)(t_0)$ exists and is finite, then we say that u is right $\alpha-$differentiable at t_0.

Definition 1.2. The conformable fractional Laplace transform (CFLT) of function $u:[0,\infty)\to\mathbb{R}$ for $t>0$, of order $0<\alpha\le 1$, starting from a of u is defined by

$$L_\alpha^a\{u(t)\} = \int_a^\infty e^{-s\frac{(t-a)^\alpha}{\alpha}} u(t)(t-a)^{\alpha-1}dt = U_\alpha^a(s). \tag{3}$$

If a=0, we have

$$L_\alpha^0\{u(t)\} = \int_0^\infty e^{-s\frac{t^\alpha}{\alpha}} u(t)t^{\alpha-1}dt = U_\alpha^0(s) = U_\alpha(s). \tag{4}$$

In particular, if $\alpha=1$, then Eq. (4) is reduced to the definition of the Laplace transform

$$L\{u(t)\} = \int_0^\infty e^{-st}u(t)dt = U(s). \tag{5}$$

Theorem 1.3. *Let $0<\alpha\le 1$, and f, g be left(right) $\alpha-$differentiable functions. Then,* $\left(T_\alpha^a f\right)(t) = (t-a)^{1-\alpha}f'(t)$, $\left({}^bT_\alpha f\right)(t) = -(b-t)^{1-\alpha}f'(t)$, *where* $f'(t) = \lim_{\epsilon\to 0}\left(\frac{f(t+\epsilon)-f(t)}{\epsilon}\right)$.

Theorem 1.4. *Let $u:[a,\infty)\to\mathbb{R}$ be differentiable real valued function and $0<\alpha\le 1$. Then*

$$L_\alpha^a\left\{{}_tT_\alpha^a(u)(t)\right\} = sU_\alpha^a(s) - u(a). \tag{6}$$

Proof. See [1]. ☐

Theorem 1.5. *Let u is piecewise continuous on $[0,\infty)$ and $L_\alpha^a\{u(t)\} = U_\alpha^a(s)$, then*

$$L_\alpha^0\left\{t^{n\alpha}u(t)\right\} = (-1)^n\alpha^n\frac{d^n}{ds^n}\left[U_\alpha^0(s)\right], \quad n\in\mathbb{N}. \tag{7}$$

Proof. See [4]. ☐

Theorem 1.6. *[3]. Let $u:[a,\infty)\to\mathbb{R}$ be twice differentiable on (a,∞), $\alpha,\ \beta>0$ and $\alpha+\beta\le 1$, then*

$${}_tT_\beta[{}_tT_\alpha u(t)] = (1-\alpha)t^{1-(\alpha+\beta)}u'(t) + t^{2-(\alpha+\beta)}u''(t). \tag{8}$$

Theorem 1.7. *[3]. Let $u:[a,\infty)\to\mathbb{R}$ be twice differentiable on (a,∞), $\alpha,\ \beta>0$ and $\alpha+\beta\le 1$, then*

1) $\displaystyle\int_0^\infty e^{-s\frac{t^{\alpha+\beta}}{\alpha+\beta}}\Big(tu''(t)\Big)dt = u(0) + s\int_0^\infty t^{\alpha+\beta}u'(t)e^{-s\frac{t^{\alpha+\beta}}{\alpha+\beta}}dt - sU_{(\alpha+\beta)}(s).$

2) $\displaystyle\int_0^\infty \Big(1-\alpha+st^{\alpha+\beta}\Big)u'(t)e^{-s\frac{t^{\alpha+\beta}}{\alpha+\beta}}dt = (\alpha-1)u(0)-(\alpha+\beta)sU_{(\alpha+\beta)}(s)+(1-\alpha)sU_{(\alpha+\beta)}(s)-(\alpha+\beta)s^2U'_{(\alpha+\beta)}(s).$

Theorem 1.8. *[3]. Let $u : [a, \infty) \to \mathbb{R}$ be twice differentiable on (a, ∞), $\alpha,\ \beta > 0$ and $\alpha + \beta \leq 1$, then*

$$1)\ L^0_{(\alpha+\beta)}\left\{{}_tT_\beta\Big({}_tT_\alpha u(t)\Big)\right\} = \alpha u(0) - (2\alpha+\beta)sU_{(\alpha+\beta)}(s) - (\alpha+\beta)s^2U'_{(\alpha+\beta)}(s).$$

$$2)\ L^0_{(\alpha+\beta)}\left\{{}_tT_\beta\Big({}_tT_\alpha u(t)\Big) + {}_tT_\alpha\Big({}_tT_\beta u(t)\Big)\right\} = (\alpha+\beta)u(0) - (3\alpha+3\beta)sU_{(\alpha+\beta)}(s) - (2\alpha+2\beta)s^2U'_{(\alpha+\beta)}(s).$$

Theorem 1.9. *[3]. Let $s,\ \alpha,\ \beta > 0$ be and $\alpha + \beta \leq 1$, then we have*

$$1)\ L^0_{\alpha+\beta}\left\{t^{-\beta}\Big({}_tT_\alpha u(t)\Big)\right\} = -u(0) + sU_{(\alpha+\beta)}(s).$$

$$2)\ L^0_{\alpha+\beta}\left\{t^{\alpha}\Big({}_tT_\alpha u(t)\Big)\right\} = L^0_{\alpha+\beta}\left\{tu'(t)\right\} = L^0_{\alpha+\beta}\left\{t^{\alpha+\beta}\Big({}_tT_{\alpha+\beta} u(t)\Big)\right\} = -(\alpha+\beta)U_{(\alpha+\beta)}(s) - (\alpha+\beta)sU'_{(\alpha+\beta)}(s).$$

Theorem 1.10. *Let $s,\ \alpha > 0,$ be and $\alpha \leq 1$ such that $L^0_\alpha\{g(t)\} = G_\alpha(s)$, then*

$$L^0_\alpha\left\{{}_tT_\alpha g(t) + t^{2-\alpha}g''(t)\right\} = -\alpha sG_\alpha(s) - \alpha s^2G'_\alpha(s). \tag{9}$$

Proof. From Eq. (7), we have

$$L^0_\alpha\left\{f(t) + \frac{t}{\alpha}f'(t)\right\} = L^0_\alpha\left\{f(t)\right\} + \frac{1}{\alpha}L^0_\alpha\left\{t^\alpha\Big({}_tT_\alpha f(t)\Big)\right\} = F_\alpha(s) + \frac{1}{\alpha}(-\alpha)\Big(sF_\alpha(s) - f(0)\Big)' = -sF'_\alpha(s).$$

Now, by substituting $f(t) = {}_tT_\alpha g(t)$ and using Theorem 1.4, we can write

$$L^0_\alpha\left\{{}_tT_\alpha g(t) + \frac{t}{\alpha}\Big({}_tT_\alpha g(t)\Big)'\right\} = -s\Big(sG_\alpha(s) - g(0)\Big)' = -sG_\alpha(s) - s^2G'_\alpha(s).$$

$Since,\ {}_tT_\alpha g(t) + \frac{t}{\alpha}\Big({}_tT_\alpha g(t)\Big)' = \frac{1}{\alpha}t^{1-\alpha}g'(t) + \frac{1}{\alpha}t^{2-\alpha}g''(t)$, we obtain

$$L^0_\alpha\left\{{}_tT_\alpha g(t) + t^{2-\alpha}g''(t)\right\} = L^0_\alpha\left\{t^{1-\alpha}g'(t) + t^{2-\alpha}g''(t)\right\} = -\alpha sG_\alpha(s) - \alpha s^2G'_\alpha(s).$$

□

Therefore, if $\alpha, \beta > 0$ be and $\alpha + \beta \leq 1$, we have

$$L^0_{\alpha+\beta}\left\{\Big({}_tT_{(\alpha+\beta)}u(t)\Big) + t^{2-(\alpha+\beta)}u''(t)\right\} = L^0_{\alpha+\beta}\left\{t^{1-(\alpha+\beta)}u'(t) + t^{2-(\alpha+\beta)}u''(t)\right\}$$
$$= -(\alpha+\beta)sU_{(\alpha+\beta)}(s) - (\alpha+\beta)s^2U'_{(\alpha+\beta)}(s).$$

2. applications

Proposition 2.1. *Suppose that $u(t)$ be twice differentiable on $(0, \infty)$ and $\alpha, \beta > 0$ be such that $\alpha + \beta \leq 1$. Then, the following CFDE*

$${}_tT_\beta\Big({}_tT_\alpha u(t)\Big) + \Big({}_tT_{\alpha+\beta}u(t)\Big) + t^{2-(\alpha+\beta)}u''(t) = q(t), \tag{10}$$

is given by

$$u(t) = \Big(L^0_{\alpha+\beta}\Big)^{-1}\left\{\Big(\frac{1}{s^{\frac{3\alpha+2\beta}{2\alpha+2\beta}}}\Big)\Big(\int\frac{\alpha u(0) - Q_{\alpha+\beta}(s)}{2(\alpha+\beta)s^{\frac{\alpha+2\beta}{2\alpha+2\beta}}}ds + C\Big)\right\}.$$

Proof. Applying $L^0_{\alpha+\beta}$ to the both sides of equation (10) and using Theorems 1.8 and 1.10, we have

$$\alpha u(0)-(2\alpha+\beta)sU_{(\alpha+\beta)}(s)-s^2(\alpha+\beta)U'_{(\alpha+\beta)}(s)-(\alpha+\beta)sU_{(\alpha+\beta)}(s)-(\alpha+\beta)s^2U'_{(\alpha+\beta)}(s)=Q_{(\alpha+\beta)}(s).$$

Above differential equation can be written as

$$2(\alpha+\beta)s^2U'_{(\alpha+\beta)}(s)+(3\alpha+2\beta)sU_{(\alpha+\beta)}(s)=\alpha u(0)-Q_{(\alpha+\beta)}(s).$$

So, we can write

$$U'_{(\alpha+\beta)}(s)+\Big(\frac{\alpha}{(2\alpha+2\beta)s}+\frac{1}{s}\Big)U_{(\alpha+\beta)}(s)=\frac{\alpha u(0)-Q_{(\alpha+\beta)}(s)}{(2\alpha+2\beta)s^2}.$$

By solving the above ODE of first order, we obtain

$$\begin{aligned}U_{\alpha+\beta}(s)&=\Big(e^{-\int(\frac{\alpha}{(2\alpha+2\beta)s}+\frac{1}{s})ds}\Big)\Big(\int\frac{\alpha u(0)-Q_{\alpha+\beta}(s)}{(2\alpha+2\beta)s^2}e^{\int(\frac{\alpha}{(2\alpha+2\beta)s}+\frac{1}{s})ds}ds+C\Big)\\&=\frac{1}{s^{\frac{3\alpha+2\beta}{2\alpha+2\beta}}}\Big(\int\frac{\alpha u(0)-Q_{\alpha+\beta}(s)}{2(\alpha+\beta)s^{\frac{\alpha+2\beta}{2\alpha+2\beta}}}ds+C\Big).\end{aligned}$$

Thus, solution $u(t)$ results from the CF inverse transform. □

In particular, when $\alpha=\beta=\frac{1}{2}$, and u(0)=1, the solution to the following CFDE

$${}_tT_{\frac{1}{2}}\Big({}_tT_{\frac{1}{2}}u(t)\Big)+{}_tT_{\frac{1}{2}+\frac{1}{2}}u(t)+t^{2-(\frac{1}{2}+\frac{1}{2})}u''(t)=(\frac{3}{2}+2t)e^t,$$

is given by

$$u(t)=L^{-1}\left\{\frac{1}{s^{1+\frac{1}{4}}}\Big(\frac{s^{\frac{5}{4}}}{s-1}+C\Big)\right\}=L^{-1}\left\{\frac{1}{s-1}\right\}+CL^{-1}\left\{\frac{1}{s^{1+\frac{1}{4}}}\right\}=e^t+C\Big(\frac{2\sqrt{2}\Gamma(\frac{3}{4})t^{\frac{1}{4}}}{\pi}\Big).$$

Proposition 2.2. *If $\alpha,\ \beta>0,$ be such that $\alpha+\beta\leq1$, a, b, m, n$\in\mathbb{R}$ and $a\neq0$, then the following ODE of second order,*

$$at^{2-(\alpha+\beta)}u''(t)+bt^{1-(\alpha+\beta)}u'(t)+mu(t)=nq(t),\tag{11}$$

is given by

$$u(t)=\Big(L^0_{\alpha+\beta}\Big)^{-1}\left\{\Big(\frac{e^{\frac{-m}{a(\alpha+\beta)s}}}{s^{1-\frac{b-a}{a(\alpha+\beta)}}}\Big)\Big(\int\frac{(a-b)u(0)-nQ_{\alpha+\beta}(s)}{a(\alpha+\beta)s^{1+\frac{b-a}{a(\alpha+\beta)}}}e^{\frac{m}{a(\alpha+\beta)s}}ds+C\Big)\right\}.$$

Proof. Applying the CF Laplace transform ($L^0_{\alpha+\beta}$) to the both sides of equation (11) and using Theorem **??**, we have

$$a\Big\{\Big(-1-(\alpha+\beta)\Big)sU_{\alpha+\beta}(s)+u(0)-(\alpha+\beta)s^2U'_{\alpha+\beta}(s)\Big\}+b\Big\{-u(0)+sU_{\alpha+\beta}(s)\Big\}+mU_{\alpha+\beta}(s)=nQ_{\alpha+\beta}(s).$$

Therefore we can write

$$U'_{\alpha+\beta}(s)+\Big(\frac{a(\alpha+\beta)s-bs+as-m}{a(\alpha+\beta)s^2}\Big)U_{\alpha+\beta}(s)=\frac{(a-b)u(0)-nQ_{\alpha+\beta}(s)}{a(\alpha+\beta)s^2}.$$

By solving the above ODE of first order, we obtain

$$U_{\alpha+\beta}(s)=\Big(\frac{e^{\frac{-m}{a(\alpha+\beta)s}}}{s^{1-\frac{b-a}{a(\alpha+\beta)}}}\Big)\Big(\int\frac{(a-b)u(0)-nQ_{\alpha+\beta}(s)}{a(\alpha+\beta)s^{1+\frac{b-a}{a(\alpha+\beta)}}}e^{\frac{m}{a(\alpha+\beta)s}}ds+C\Big).$$

Thus, solution $u(t)$ results from the CF inverse transform. □

In particular by substituting $\alpha = \beta = \frac{1}{2}$ in equation (11), the following ODE of second order,

$$atu''(t) + bu'(t) + mu(t) = nq(t),$$

is given by

$$u(t) = L^{-1}\left\{\Big(\frac{e^{\frac{-m}{a(\frac{1}{2}+\frac{1}{2})s}}}{s^{1-\frac{b-a}{a(\frac{1}{2}+\frac{1}{2})}}}\Big)\Big(\int \frac{(a-b)u(0) - nQ_{\frac{1}{2}+\frac{1}{2}}(s)}{a(\frac{1}{2}+\frac{1}{2})s^{1+\frac{b-a}{a(\frac{1}{2}+\frac{1}{2})}}} e^{\frac{m}{a(\frac{1}{2}+\frac{1}{2})s}} ds + C\Big)\right\}$$

$$= L^{-1}\left\{\frac{e^{\frac{-m}{as}}}{s^{2-\frac{b}{a}}}\Big(\int \frac{(a-b)u(0) - nQ(s)}{as^{\frac{b}{a}}} e^{\frac{m}{as}} ds + C\Big)\right\}.$$

For example, if a=b=m=n=1, then the following differential equation

$$tu''(t) + u'(t) + u(t) = te^{-t},$$

is given by

$$u(t) = L^{-1}\Big\{\frac{e^{\frac{-1}{s}}}{s}\Big(\int -\frac{1}{(s+1)^2 s} e^{\frac{1}{s}} ds + C\Big)\Big\} = L^{-1}\Big\{\frac{e^{\frac{-1}{s}}}{s}\Big(\frac{se^{\frac{1}{s}}}{s+1} + C\Big)\Big\} = L^{-1}\Big\{\frac{1}{s+1}\Big\} + L^{-1}\Big\{C\frac{e^{\frac{-1}{s}}}{s}\Big\}$$

$$= e^{-t} + CBesselJ\big(0, 2\sqrt{t}\big).$$

References

[1] T. Abdeljawad, On conformable fractional calculus. Journal of Computational and Applied Mathematics 2015; 279: 57-66. doi: 10.1016/j.cam.2014.10.016

[2] R. Khalil, M. Al Horani, Y. Abdelrahman, S. Mohammad, A new definition of fractional derivative. Journal of Computational and Applied Mathematics 2014; 264: 65-70. doi: 10.1016/j.cam.2014.01.002

[3] M. Molaei, F. Dastmalchi Saei, M. Javidi, Y. Mahmoudi. New analytical methods for solving a class of conformable fractional differential equations by fractional Laplace transform. Computational Methods for Differential Equations. 2021; doi: 10.22034/cmde.2021.40834.1775

[4] MS. Hashemi, Invariant subspaces admitted by fractional differential equations with conformable derivatives. Journal of Chaos, Solitons & Fractals. 2018; 107: 161-169. doi: 10.1016/j.chaos.2018.01.002

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–175, pp. 534–538

Some Properties Of Generalized Conformable Fractional Derivatives(GCFD)

Mohammad Jafari[a], Mohammad Molaei[b]

[a]Department of Science, Payame Noor University, Tehran, Iran
[b]Department of Science, Payame Noor University, Tehran, Iran

Article Info

Keywords:
Conformable derivative
generalized fractional derivative
Laplace transform

2020 MSC:
26A33
34A08

Abstract

In this paper, we discuss on the new generalized fractional operator, and we will discuss some of its properties. The introduced method is the generalization of conformable fractional derivatives which was previously introduced by Kalil et al. and Katugampola. This operator similarly conformable derivative satisfies properties such as the sum, product/quotient, and chain rule

1. Introduction and Preliminaries

The creation of the concept of fractional calculus was formed by L'Hopital in a letter to Leibniz in 1695. The interest in fractional differential equations grew rapidly, and various types of definitions were introduced. One thing that all these have in common, is that they are consist of integral with different singular kernels. The most popular of them, we can mention to Grünwald-Letnikov, Riemann-Liouville, Caputo and Riesz, etc. Although fractional operators are linear, Unfortunately, this class of fractional derivatives is unsatisfying in some properties, such as product rule, quotient rule, chain rule and compositions rule. In 2014, Kalil et al.[1] and Also Katugampola, [2] by modifying the limit definition on the classic derivative, introduced a simple and local type of fractional derivative called conformable fractional derivative. These new definitions are complies with the computational relationship of the first derivative. Recently Mingarelli et al. [3] introduced the so-called "*Generalized Conformable Fractional Derivative*"(GCFD) as a unifying framework for conformable fractional methods. M.Jafari and et al. [4] have discussed GCFD for the 4-order fractional Sturm Liouville problem.
Here, we recall some definitions, notations and properties of fractional calculus theory used in this work.

Definition 1.1. (Ref [1]) The conformable fractional derivative of a function $f : [0, \infty] \to \mathbb{R}$ is defined as:

$$D^{\alpha} f(t) = \lim_{h\to 0} \frac{f(t + ht^{1-\alpha}) - f(t)}{h}, \qquad t > 0, \tag{1}$$

where $0 < \alpha \leq 1$. Also Katugampola present the following definition [2].

Definition 1.2. Let $f : [0, \infty] \to \mathbb{R}$. Then the fractional derivative of f of order $0 < \alpha \leq 1$ is defined by,

$$D^\alpha f(t) = \lim_{h \to 0} \frac{f(te^{ht^{-\alpha}}) - f(t)}{h}, \qquad t > 0. \tag{2}$$

In both definitions, if $(D^\alpha f)(t)$ exists on $(0, \infty)$, then $D^\alpha f(0) = \lim_{t \to 0^+} D^\alpha f(t)$.

2. Generalized Fractional Derivative

Mingarelli [3] extended the mentioned definition of compatible derivative and proposed it in a form .

Definition 2.1. For a given function $f : I \subseteq R \to R$, defined on the range of an appropriate real valued function $p : U_\delta \to R$ where $U_\delta = \{t, h) : t \in I = (a, b), |h| < \delta\}$, by means of the limit

$$D_p^\alpha f(t) = \lim_{h \to 0} \frac{f(p(t, h, \alpha)) - f(t)}{h}, \tag{3}$$

whenever the limit exists and is finite, be called the $(\alpha - p)$**-derivative** of f at t or f is $(\alpha - p)$**-differentiable** at t.

The function p must be satisfies in the following conditions.
$H1^+$ for $t \in I$ and for all sufficiently small $\varepsilon > 0$, the equation $p(t, h) = t + \varepsilon$ has a solution $h = h(t, \varepsilon)$. In addition, $h \to 0$ as $\varepsilon \to 0$.
$H1^-$ for $t \in I$ and for all sufficiently small $\varepsilon > 0$, the equation $p(t, h) = t - \varepsilon$ has a solution $h = h(t, \varepsilon)$. In addition, $h \to 0$ as $\varepsilon \to 0$.
$H_2 \frac{1}{p_h(.,0)} \in L(I) =$ {f: f is Lebesgue integrable function}.
Under the above conditions, if f, g are $(\alpha - p)$ - differential at $t \in I$, then
(i) (The Sum Rule)

$$D_p^\alpha \big(f + g\big)(t) = D_p^\alpha f(t) + D_p^\alpha g(t).$$

(ii) (The Product Rule)

$$D_p^\alpha \big(f.g\big)(t) = g(t).D_p^\alpha f(t) + f(t).D_p^\alpha g(t).$$

(iii) (The Quotient Rule)

$$D_p^\alpha \big(\frac{f}{g}\big)(t) = \frac{g(t).D_p^\alpha f(t) - f(t).D_p^\alpha g(t)}{g^2(t)}.$$

(iv) (The Chain Rule)

$$D_p^\alpha \big(f \circ g\big)(t) = g'\big(f(t)\big) D_p^\alpha f(t).$$

(v)(The relationship between differentiability and $(\alpha - p)$-differentiability) For $\frac{\partial p(t,0,\alpha)}{\partial h} = p_h(t, 0, \alpha) \neq 0$ we have

$$D_p^\alpha f(t) = p_h(t, 0, \alpha) f'(t), \tag{4}$$

that $0 < \alpha < 1$. For more details, the reader can refer to [3].

Example 2.2.

$$a) \qquad D_p^\alpha \Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big) = 1,$$

$$b) \qquad D_p^\alpha \Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big)^n = n\Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big)^{n-1},$$

$$c) \qquad D_p^\alpha \big(e^{c \int_a^t \frac{ds}{p_h(s,0,\alpha)}}\big) = c e^{c \int_a^t \frac{ds}{p_h(s,0,\alpha)}},$$

$$d) \qquad D_p^\alpha \sin\Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big) = a \cos\Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big),$$

$$e) \qquad D_p^\alpha \cos\Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big) = -a \sin\Big(\int_a^t \frac{ds}{p_h(s, 0, \alpha)}\Big).$$

The Wroneskian and Abel's Formula

Here we discuss about the fractional Wroneskian of two functions.

For two functions y_1 and y_2, satisfying $D_p^\alpha\left(D_p^\alpha y\right)+p(x)D_p^\alpha y+q(x)y=0$ we set

$$W_{\mathrm{p}}^\alpha(y_1,y_2)=\begin{vmatrix} y_1 & y_2 \\ D_{\mathrm{p}}^\alpha y_1 & D_{\mathrm{p}}^\alpha y_2 \end{vmatrix}$$

thus

$$\begin{aligned} W_{\mathrm{P}}^\alpha(y_1,y_2) &= y_1 D_{\mathrm{p}}^\alpha y_2 - y_2 D_{\mathrm{p}}^\alpha y_1 = y_1 y_2 p_h(t,0,\alpha) - y_2 y_1 p_h(t,0,\alpha) \\ &= p_h(t,0,\alpha)(y_1 y_2' - y_2 y_1') = p_h W(y_1,y_2). \end{aligned}$$

Generalized Fractional Laplace and Integrals

Definition 2.3. In this section we start with the definition of (α,p) - fractional integral of f in the following sense.

$$I_p^\alpha(f(t))=\int_a^t f(s)d_\alpha(s)=\int_a^t \frac{f(s)}{p_h(s,0,\alpha)}ds.$$

In the sequel we state some theorems that their proofs are routine and we leave for the reader.
If f is a continuous function in the domain of I_p^α, then we have
(1)

$$D_{\mathrm{p}}^\alpha\left(I_p^\alpha(f(t))\right)=f(t),$$

(2)

$$I_{\mathrm{p}}^\alpha D_{\mathrm{p}}^\alpha(f(t))=f(t)-f(a).$$

Also we have

$$\int_a^b f(t)D_p^\alpha(g(t))d_\alpha(t)= f(t)g(t)|_a^b-\int_a^b g(t)D_{\mathrm{p}}^\alpha(f(t))d_\alpha(t).$$

The following definition gives us the adapted Laplace transform to the $(\alpha,p)-$ fractional derivative.

Definition 2.4. The fractional Laplace transform is defined by

$$\mathcal{L}_{\mathrm{p}}^\alpha(f(t))=\int_0^\infty \frac{e^{-s\int_0^t \frac{d}{p_h(s,0,\alpha)}}}{p_h(t,0,\alpha)}f(t)dt=F_{\mathrm{p}}^\alpha(s).$$

Theorem 2.5. *If $f(t)$ is (α,p)-differentiable and $\int_0^t \frac{ds}{p_s(s,0,\alpha)}>0$ then we have*

$$\begin{aligned} \mathcal{L}_{\mathrm{p}}^\alpha\left(D_{\mathrm{p}}^\alpha(f(t))\right) &= sF_p^\alpha(s)-f(0). \\ \mathcal{L}_{\mathrm{p}}^\alpha\left(D_{\mathrm{p}}^\alpha\left(D_{\mathrm{p}}^\alpha f(t)\right)\right) &= s^2F_{\mathrm{p}}^\alpha(s)-sf(0)-p_h(0,0,\alpha)f'(0). \end{aligned}$$

Proof. We prove the first part by the definition (2.4) and item (v), proof of the second part is similar,so we have

$$\mathfrak{L}_{\mathrm{p}}^\alpha\left(D_{\mathrm{p}}^\alpha(f(t))\right)=\int_0^\infty \frac{e^{-\Omega\int_0^1 \frac{d}{p_h(t,0,\alpha)}}}{p_h(t,0,\alpha)}p_h(t,0,\alpha)f'(t)dt,$$

now by means of integration by parts

$$=sF_{\mathrm{p}}^\alpha(s)-f(0).$$

□

.
If $f(t)$ is (α, p)-differentiable, the following relations are immediately concluded.

$$
\begin{aligned}
&1) \qquad \mathfrak{L}_p^\alpha\left(I_{\mathrm{p}}^\alpha(f(t))\right)=\frac{F_{\mathrm{p}}^\alpha(s)}{s},\\
&2) \qquad \mathcal{L}_p^\alpha\left(e^{a\int_0^\varepsilon \frac{d}{p_k(t,0,a)}} f(t)\right)=F_p^\alpha(s-a),\\
&3) \qquad \mathfrak{L}_p^\alpha\left(\int_0^{\mathrm{t}} \frac{ds}{p_h(s,0,\alpha)} f(t)\right)=-\frac{d}{ds}F_p^\alpha(s).
\end{aligned}
$$

Example 2.6. We have the following formulae (By using $u=\int_0^t \frac{ds}{pt(s,0,\alpha)}$ for the proof) then
(a)

$$\mathcal{L}_p^\alpha(1)=\int_0^\infty \frac{e^{-s\int_0^z \frac{du}{p_b(\pi,v,\alpha)}}}{P_h(t,0,\alpha)}dt=\int_0^\infty e^{-su}du=\frac{1}{s}.$$

(b)

$$\mathscr{L}_p^\alpha\left(e^{a\int_0^\varepsilon \frac{dA}{D_B(x,0)}}\right)=\int_0^\infty e^{-(s-a)u}du=\frac{1}{8-a}.$$

(c)

$$\mathcal{L}_p^\alpha\left(\sin a\int_0^t \frac{ds}{p_n(s,0,\alpha)}\right)=\int_0^\infty e^{-su}\sin audu=\frac{a}{s^2+a^2}.$$

(d)

$$\mathfrak{L}_{\mathrm{p}}^\alpha\left(\cos a\int_0^e \frac{ds}{p_h(s,0,\alpha)}\right)=\int_0^\infty e^{-su}\cos audu=\frac{s}{s^2+a^2}.$$

(e)

$$\mathcal{L}_{\mathrm{p}}^\alpha\left(e^{\mathrm{a}\int_0^a \frac{d,}{p_h(s,u,\alpha)}}\cos b\int_0^t \frac{ds}{p_h(8,0,\alpha)}\right)=\int_0^\infty e^{-(s-a)u}\cos budu=\frac{8-a}{(8-a)^2+b^2}.$$

Example 2.7. Consider the fractional initial value problem

$$D_{\mathrm{p}}^\alpha y(t)=\lambda y(t), \quad y(0)=C.$$

Apply the Laplace transform and item (b) to the above equation to obtain

$$Y_p^\alpha(s)=\frac{C}{s-\lambda},$$

and hence, $y(t)=Ce^{\lambda\int_0^t \frac{dt}{p_k(t,0,a)}}$. We also can solve the equation by another method, indeed by above property (iv) we have

$$D_p^\alpha y(t)=p_h(t,0,\alpha)y'(t)=\lambda y(t).$$

Then

$$\frac{y'(t)}{y(t)}=\frac{\lambda}{p_h(t,0,\alpha)},$$

by integration both sides

$$y=Ce^{\lambda\int_0^t \frac{dt}{p_h(t,0,\alpha)}}.$$

Remark 2.8. It is important to note that if we take p such that $p_h(t,0,\alpha)=t^{1-\alpha}, \quad 0<\alpha<1$ then t results of all above examples are agreement with $[1,2]$.

Example 2.9. Consider the fractional initial value problem

$$-D_p^\alpha\left(D_p^\alpha y(t)\right)=\lambda y(t),\quad y(0)=c,y'(0)=d.$$

Apply the Laplace transform to the above equation to obtain

$$-s^2Y_{\mathrm{P}}^\alpha(s)+sy(0)+p_h(0,0,\alpha)y'(0)=\lambda Y_{\mathrm{P}}^\alpha(s),$$

and hence,

$$Y_{\mathrm{p}}^a(s)=\frac{As+B}{s^2+\lambda}.$$

Now by item (c) and (d) we have

$$y(t)=A\sin\left(\sqrt{\lambda}\int_a^t\frac{1}{p_h(s,0)}ds\right)+B\cos\left(\sqrt{\lambda}\int_a^t\frac{1}{p_h(s,0)}ds\right).$$

where A,B are constants and we fix the determination of the square root if $\lambda<0$.

Remark 2.10. Indeed, in [3] under H2 the general solution of $-D_p^\alpha\left(D_p^\alpha y(t)\right)=\lambda y(t)$ for $\lambda\neq 0$ is obtained by Carathéodory solutions.

Example 2.11. Consider the fractional initial value problem

$$D_p^\sigma\left(D_p^\alpha y(t)\right)=0,\quad y(0)=c,y'(0)=d.$$

Apply the I_p^β to the above equation, we obtain

$$I_{\mathrm{p}}^\beta\left(D_{\mathrm{p}}^\beta\left(D_p^\alpha y(t)\right)\right)=0,$$

and hence,

$$I_{\mathrm{p}}^\alpha\left(D_{\mathrm{p}}^\alpha y(t)\right)=I_{\mathrm{p}}^\alpha(A).$$

Now

$$y(t)-y(0)=\int_0^t\frac{A}{p_h(t,0,\alpha)}dt$$

then

$$y(t)=A\int_0^t\frac{1}{p_h(t,0,\alpha)}dt+B$$

Remark 2.12. The reader notice that if we take p such that $p_n(t,0,\alpha)=t^{1-\alpha},0<\alpha<1$ then the solution of the equation

$$D_p^\sigma\left(D_{\mathrm{p}}^a y(t)\right)=0,\quad y(0)=c,y'(0)=d,$$

is $y(t)=A\frac{t^\alpha}{\alpha}+B$.

Example 2.13. Consider the fractional initial value problem

$$D_{\mathrm{p}}^\beta\left(D_{\mathrm{p}}^\alpha y(t)\right)=\lambda y(t),\quad y(0)=c,y'(0)=d.$$

Applying the property (v) to the above equation we obtain

$$p_h(t,0,\alpha)\left(p_h(t,0,\beta)y'(t)\right)'=\lambda y(t).$$

and hence,

$$\left(p_n(t,0,\beta)y'(t)\right)'=\frac{\lambda y(t)}{p_h(t,0,\alpha)}.$$

References

[1] R. Khalil, M. Al Horani, A. Yousef, and M. Sababhehb, *A new definition of fractional derivative*, J. Comp. Applied Math., 264 (2014), 65–70.

[2] U. N. Katugampola, *A new fractional derivative with classical properties*, e-print arXiv:1410.6535 (2014).

[3] A.B. Mingarelli, *On generalized and fractional derivatives and their applications to classical mechanics*, Journal of Physics A: Mathematical and Theoretical. DOI:10.1088/1751-8121/aad345.

[4] M. Jafari, F. Dastmalch, M. Jahangirirad *The generalized conformable derivative for 4α-order Sturm-Liouville problems*, Computational Methods for Differential Equations, Vol. 10, No. 3, 2022, pp. 816-825.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–183, pp. 539–543

Optimization Functions for Neural Network-Based Approximation of Burger's–Fisher Equation: A Comparative Analysis

Haniye Hajinezhad[a], Abulfazl Yavari[b,*]

[a]Department of Mathematics, Payame Noor University, Tehran, Iran.
[b]Department of Computer Engineering and IT, Payame Noor University, Tehran, Iran

Article Info

Keywords:
Burger's–Fisher equation
Neural networks
Optimization functions

2020 MSC:
92B20
65K05
35Q92

Abstract

This study explores the effectiveness of different optimization functions for approximating the solution of Burger's–Fisher equation with initial and boundary conditions based on neural networks. It compares and analyzes the performance of nine common optimization functions, emphasizing computational accuracy. Extensive experiments show that the choice of appropriate optimization function significantly influences the performance of neural network-based solvers for approximating the solution of Burger's–Fisher equation with initial and boundary conditions. The findings provide valuable insights and practical recommendations for researchers applying neural networks to solve Berger's equation in fields such as fluid dynamics and heat transfer.

1. Introduction

Neural networks are a type of machine learning algorithm that are inspired by the structure and function of the human brain. They are used to learn patterns in data and make predictions based on those patterns. In addition to numerical methods, neural networks can be used to approximate the solutions of equations, they differ in their approach. Numerical methods rely on mathematical models to approximate solutions, while neural networks learn the solution directly from data.

Neural networks have been used for solving differential equations for recent decades. One of the key papers in this area is the 1998 work of Lagaris et al. [7], in which they presented a technique for solving differential equations with neural networks.

Nonlinear partial differential equations are significant in the realms of physical science and engineering. One of the key advantages of using neural networks for solving differential equations is their ability to handle complex and non-linear systems, which are often challenging for traditional numerical methods. Additionally, neural networks can learn from data and adapt to new situations, making them particularly well-suited for solving differential equations in real-world applications.

The Burger equation is a highly nonlinear equation that combines convection, diffusion, and reaction mechanisms. Here we review several notable research efforts focused on employing neural networks to solve the Burger equation.

*Talker
Email addresses: H.Hajinezhad@pnu.ac.ir (Haniye Hajinezhad), A.Yavari@pnu.ac.ir (Abulfazl Yavari)

Wen and Chaolu [16] introduced a neural network approach based on Lie series within Lie groups of differential equations for solving Burgers–Huxley equations, incorporating initial or boundary value terms in the loss functions. Cheng and Zhang [2] proposed a method that integrates the Physical Informed Neural Network (PINN) with Resnet blocks to address the Navier-Stokes and Burger's equations, incorporating these equations into the loss function of the deep neural network to guide the model. Additionally, the loss function accounts for the inclusion of initial and boundary conditions. Ye et al. [17] presented the modified PINN method for time-fractional Burger's-type equations, introducing locally adaptive activation functions and two effective weighting strategies to enhance solution accuracy. Lu et al. [10] introduced a residual-based adaptive refinement method to improve the training efficiency of PINNs for approximating solutions of differential equations, including the Burger's equation. This paper emphasizes the significance of optimizer function selection in improving the approximation of Burger's–Fisher equation with initial and boundary conditions.

Adaptive gradient methods play a pivotal role in deep learning. Despite stochastic gradient descent (SGD) being a long-standing favorite, it grapples with challenges related to ill-conditioning and the time needed to train deep neural networks with large datasets. Consequently, more sophisticated algorithms have emerged to address the limitations of SGD. Currently, optimization algorithms for deep learning dynamically adapt their learning rates during training. Adaptive gradient methods essentially fine-tune the learning rate for each parameter, reducing it when gradients of specific parameters are large, and vice versa. Several adaptive methods have been introduced recently, emerging as predominant alternatives to SGD.

In this paper, the authors compare various optimization functions, including SGD [14], Rprop, RMSprop [13], RAdam [8], AdaGrad, Nadam [3], Adam [6], AdamW [9], and Adamax [6], for training neural networks to approximate the solution of the Burger's–Fisher equation with initial and boundary conditions. They evaluate the performance of these optimizers in terms of accuracy.

The paper's structure is as follows: Section 2 introduces the Burger's-Fisher equation with initial and boundary conditions. Section 3 describes the neural network used to approximate the solution of the Burger's-Fisher equation. Section 4 is dedicated to comparing the performance of different optimizer functions. Finally, the conclusion is presented.

2. Description of Burger's–Fisher equation

The generalized Burger's–Fisher equation has been applied in various fields including gas dynamics, number theory, heat conduction, and elasticity [18]. The generalized Burger's–Fisher equation is as follows.

$$u_t = \alpha u^\delta u_x - u_{xx} = \beta u(1-u^\delta), \qquad a \leqslant x \leqslant b, \quad t \geq 0 \tag{1}$$

By assuming appropriate initial and boundary conditions, the exact solution is

$$u(x,t) = (\frac{1}{2} + \frac{1}{2} tanh(\theta_1(x-\theta_2)))^\delta \tag{2}$$

where,

$$\theta_1 = \frac{-\alpha\delta}{2(1+\delta)}, \quad \theta_2 = \frac{\alpha}{1+\delta} + \frac{\beta(1+\delta)}{\alpha}$$

To test the accuracy of the approximate solution obtained by the neural network, the initial and boundary conditions are determined by the exact solution. Numerous researchers have investigated this equation to approximate its solution with numerical methods [12], [4], [15], [11], [5].

3. Description of used Neural Network

The initial step in solving a differential equation using neural networks (NNs) involves reframing the problem as an optimization task, based on the definition of the loss function. Subsequently, the architecture of the neural network needs to be designed. This entails determining the number of hidden layers, the number of neurons in each layer, the volume of training data, the selection of activation function, the number of epochs, and the selection of optimization functions and other relevant parameters. Once the architecture is established, the neural network must be trained using

an appropriate dataset. During the training process, the loss function is used to measure the disparity between the network's predictions and the actual values. The network's parameters (weights and biases) are iteratively adjusted using optimization algorithms such as gradient descent to minimize the loss function. Upon completion of training, the neural network can approximate the solution to the given differential equation for unseen inputs.
In this paper, the idea of Chen and his colleague [1] is used to formulate the loss function. Our neuron network comprises an input layer with two variables x and t, an output layer with variable u(x,t), and an intermediate structure of five hidden layers, each containing 32 neurons. We employ the GELU activation function. Also, 32 points as x and 16 points as t are used for training. The maximum number of epoches is considered 5000, in other cases it is mentioned. A suitable optimization function should be determined to update the network parameters based on the gradients calculated through the backpropagation method.

4. Comparing the accuracy of different optimizer functions

In this section, our objective is to identify the ideal optimization function for training the proposed neural network. For this purpose, the infinity norm and one norm of approximation error is defined, respectively as follows.

$$\|E\|_\infty = Max_{1\le i\le N}|u_{ana_i} - u_{net_i}|,$$
$$\|E\|_1 = \frac{\sum_{i=1}^{N} |u_{ana_i} - u_{net_i}|}{N},$$

where u_{ana_i} and u_{net_i} represent the exact solution and neural network's approximate solution at (x_i, t_i) respectively, and the N is the total number of points used for error computation, set to $N = 10000$.
We conducted a series of experiments using the neural network described in the preceding section to explore the efficacy of nine optimization algorithms, namely SGD [14], Rprop, RMSprop [13], RAdam [8], AdaGrad, Nadam [3], Adam [6], AdamW [9], and Adamax [6]. Each algorithm represents a unique case used to approximate the solution of equation 1 with specific initial and boundary conditions.
In the first test, to ensure statistical reliability, each of these 9 cases was run 30 times, and the average of $\|E\|_\infty$ and $\|E\|_1$ of the resulting 30 runs were computed. Table 1 presents the average of $\|E\|_\infty$ and $\|E\|_1$ for each optimizer function, by assuming $\alpha = 1$, $\beta = 1$, and $\delta = 2$, with $x \in [-10, 20]$ and $t \in [0, 1]$. The columns representing the infinity norm of the error and the one norm of the error are arranged in ascending order. The table indicates that Adamax, AdamW, and Radam provide more accurate approximations, while RMSprop yields less accurate results.

Table 1. Average of $\|E\|_\infty$ and $\|E\|_1$ for Each Optimizer Function

Optimizer	$\|E\|_\infty$	Optimizer	$\|E\|_1$
Adamax	$1.64708244050343 \times 10^{-6}$	Adamax	$7.85898779760152 \times 10^{-6}$
AdamW	$2.52910249324498 \times 10^{-6}$	Radam	$1.46709472871244 \times 10^{-5}$
Radam	$2.82038610701581 \times 10^{-6}$	AdamW	$1.48394629333827 \times 10^{-5}$
Adagrad	$3.00686953460146 \times 10^{-6}$	Adagrad	$1.57488386856474 \times 10^{-5}$
SGD	$5.90508594510665 \times 10^{-6}$	SGD	$2.58074128027252 \times 10^{-5}$
Adam	$6.33155090038516 \times 10^{-6}$	Adam	$3.55779421131220 \times 10^{-5}$
Rprop	$8.64033769476345 \times 10^{-6}$	Rprop	$4.46275454700967 \times 10^{-5}$
Nadam	$4.44236498029124 \times 10^{-5}$	Nadam	$2.87230686660080 \times 10^{-4}$
RMSprop	$3.19363565259401 \times 10^{-2}$	RMSprop	$2.27806677748135 \times 10^{-1}$

For the second test, Tables 2 displays $\|E\|_1$ of each optimizer function by assuming $\alpha = 1$, $\beta = 1$, and $\delta = 2$, with $x \in [-10, 20]$ and $t \in [0, 1]$ and under different conditions, including a maximum of 10000 and 20000 epochs. The results consistently show that Adamax, AdamW, and Radam lead to more accurate approximations, while RMSprop produces less accurate results.
The results of another test are presented in Table 3, showcasing $\|E\|_1$ of each optimizer function by assuming $\alpha = 2$, $\beta = 4$, and $\delta = 2$, with $x \in [-10, 20]$ and $t \in [0, 1]$. Once again, this table shows that Adamax, AdamW, and Radam lead to more accurate approximations, while RMSprop produces less accurate results.

Table 2. $\|E\|_1$ for Each Optimizer Function with 10000 and 20000 Epochs by assuming $\alpha = 1$, $\beta = 1$, and $\delta = 2$

Optimizer	$\|E\|_1$ (Max of epoche=10000)	Optimizer	$\|E\|_1$ (Max of epoche=10000)
Radam	$5.35030649640497 \times 10^{-6}$	Adamax	$2.35291704558418 \times 10^{-6}$
AdamW	$5.35423608295110 \times 10^{-6}$	AdamW	$2.41199352109394 \times 10^{-6}$
Adamax	$5.46882189079660 \times 10^{-6}$	Radam	$3.04817965468302 \times 10^{-6}$
Adagrad	$8.93723642925836 \times 10^{-6}$	Adagrad	$5.15871498623183 \times 10^{-6}$
Rprop	$9.93681627194604 \times 10^{-6}$	Adam	$5.83348527166089 \times 10^{-6}$
SGD	$9.96818089713675 \times 10^{-6}$	SGD	$2.64993640289539 \times 10^{-5}$
Adam	$1.34757426205124 \times 10^{-5}$	Rprop	$2.65192343763965 \times 10^{-5}$
Nadam	$1.36896577896813 \times 10^{-4}$	Nadam	$3.59266351531136 \times 10^{-5}$
RMSprop	$2.48884376692317 \times 10^{-4}$	RMSprop	$1.72365296041790 \times 10^{-4}$

Table 3. $\|E\|_1$ for Each Optimizer Function by assuming $\alpha = 2$, $\beta = 4$, and $\delta = 2$

Optimizer	$\|E\|_1$
Adamax	$7.57195157904586 \times 10^{-5}$
Radam	$3.40890863500638 \times 10^{-4}$
AdamW	$3.89393208257223 \times 10^{-4}$
Adam	$4.12877173177112 \times 10^{-4}$
Nadam	$4.35333281243569 \times 10^{-4}$
Adagrad	$4.80405060820108 \times 10^{-4}$
SGD	$1.64000379193429 \times 10^{-3}$
Rprop	$1.65260399241475 \times 10^{-3}$
RMSprop	$4.66357137269634 \times 10^{-3}$

The final test results are presented in Table 4, showcasing the $\|E\|_1$ of each optimizer function by assuming $\alpha = 3$, $\beta = 2$, and $\delta = 2$, with $x \in [-10, 20]$ and $t \in [0, 1]$. Once again, findings of table 4 show that Adamax, AdamW, and Radam lead to more accurate approximations, while RMSprop produces less accurate results.

Table 4. $\|E\|_1$ for Each Optimizer Function by assuming $\alpha = 3$, $\beta = 2$, and $\delta = 2$

Optimizer	$\|E\|_1$
Adamax	$2.09975415513011 \times 10^{-5}$
AdamW	$2.96310587707187 \times 10^{-5}$
Radam	$3.36089030061656 \times 10^{-5}$
SGD	$3.96262605974220 \times 10^{-5}$
Rprop	$4.02997607315718 \times 10^{-5}$
Adagrad	$5.41389936737727 \times 10^{-5}$
Nadam	$7.69479337684435 \times 10^{-5}$
Adam	$1.22248478864545 \times 10^{-4}$
RMSprop	$2.29861196186085 \times 10^{-3}$

In summary, the results from Tables 1-4 confirm that Adamax, AdamW, and Radam are the most suitable optimizers for approximating the Berger-Fisher equation, while RMSprop consistently yields less accurate approximations.

5. Conclusions

In conclusion, this study presented a comprehensive analysis of various optimization functions for approximating the solution of the Burger's-Fisher equation with initial and boundary conditions using neural networks. Through extensive experiments and comparisons, it was demonstrated that the choice of optimization function significantly impacts the performance of neural network-based solvers for this complex equation. The findings highlight the effectiveness of Adamax, AdamW, and Radam as the most suitable optimizers for accurately approximating the Burger's-Fisher equation with initial and boundary conditions, while RMSProp consistently yielded less accurate results. These insights provide valuable guidance for researchers applying neural networks to solve challenging equations in fluid dynamics and heat transfer fields. The study contributes to advancing neural network applications in solving differential equations.

References

[1] F. Chen et al. Neurodiffeq: A python package for solving differential equations with neural networks. *Journal of Open Source Software*, 5 (46):1931, 2020.

[2] C. Cheng and G.-T. Zhang. Deep learning method based on physics informed neural network with resnet block for solving fluid flow problems. *Water*, 13(4):423, 2021.

[3] T. Dozat. Incorporating nesterov momentum into adam. *International conference on learning representations*, 2016.

[4] B. Gürbüz and M. Sezer. A modified laguerre matrix approach for burgers–fisher type nonlinear equations, numerical solutions of realistic nonlinear phenomena. pages 107–123. Springer, 2020.

[5] H. Hajinezhad. A numerical approximation for the one-dimensional burger–fisher equation. *Asian Research Journal of Mathematics*, 18(5): 22–30, 2022.

[6] D.P. Kingma and J. Ba. Adam: a method for stochastic optimization. 2014.

[7] I. E. Lagaris, A. Likas, and D. I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(5):987–1000, 1998.

[8] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han. On the variance of the adaptive learning rate and beyond. *arXiv preprint arXiv:1908.03265*, 2019.

[9] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. *arXiv preprint arXiv:1711.05101*, 2017.

[10] L. Lu et al. Deepxde: A deep learning library for solving differential equations. *SIAM review*, 63(1):208–228, 2021.

[11] R.K. Mohanty and S. Sharma. A high-resolution method based on off-step non-polynomial spline approximations for the solution of burgers-fisher and coupled nonlinear burgers' equations. *Engineering Computations*, 2020.

[12] M.A. Pirdawood and Y.A. Sabawi. High-order solution of generalized burgers–fisher equation using compact finite difference and dirk methods. *Journal of Physics: Conference Series*, 2021.

[13] M. Riedmiller and H. Braun. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In *Proceedings of the IEEE International Conference on Neural Networks*, page 586–591, 1993.

[14] H. Robbins and S. Monro. A stochastic approximation method. *The Annals of Mathematical Statistics*, 25:400–407, 1951.

[15] A. Singh, S. Dahiya, and S. Singh. A fourth-order b-spline collocation method for nonlinear burgers–fisher equation. *Mathematical Sciences*, 14(1):75–85, 2020.

[16] Y. Wen and T. Chaolu. Study of burgers–huxley equation using neural network method. *Axioms*, 12(5):429, 2023.

[17] Y. Ye et al. Deep neural network method for solving the fractional burgers-type equations with conformable derivative. *Physica Scripta*, 98 (6):065214, 2023.

[18] C.G. Zhu and W.S. Kang. Numerical solution of burgers–fisher equation by cubic b-spline quasi-interpolation. *Appl. Math. Comput.*, 216: 2679–2686, 2010.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–198, pp. 544–553

Exploring the Phenomenon of Multiple-Rogue Waves in a $(3+1)-$Dimensional Extension of the Hirota Bilinear Equation

Majid Madadi[a,*], Esmaeel Asadi[a,b]

[a] *Department of Mathematics, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137–66731, Iran*
[b] *School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O. Box 19395-5746, Tehran, Iran*

Article Info

Keywords:
Rogue wave;
Soliton;
Symbolic computation;
Hirota Bilinear;

2020 MSC:
35C08;
35C11;
45G10;
33F10;

Abstract

This paper investigates rational solutions of an extended Hirota bilinear equation, a foundational mathematical model for studying nonlinear phenomena in shallow water and fluid mechanics. By leveraging the Hirota bilinear equation and employing symbolic computation method, we derive higher-order rational and generalized rational solutions and scrutinize their dynamic characteristics. The derived solutions reveal two distinct waveforms: multi-rogue and multi-soliton. Generalized multi-rogue waves encompass three, six, eight, and ten rogue waves, while multi-soliton waves comprise one to five soliton waves. These waveforms are visually depicted in the paper. Furthermore, we analyze the roots of the rational solutions, uncovering a compelling relationship between these complex roots and solutions of well-known Painlevé II and IV equations.

1. Introduction

Rogue waves, also called freak or monster waves, are notable for their sudden appearance and disappearance, featuring dangerously steep crests that endanger ships at sea [1]. The study of rogue waves spans various fields like nonlinear science, optical communications, and fluid dynamics[2, 3]. Several methods, including the Hirota bilinear, Darboux transformation and inverse scattering methods, have been used to model rogue waves. While progress has been made in understanding single rogue waves, higher-order ones, with multiple rogue wave interactions, remain challenging, influencing phase shifts and energy distribution [4–6].
Symbolic computation method plays a pivotal role in the study of rogue wave phenomena in nonlinear differential equations with constant coefficients. This method, utilizing computer algebra systems, is particularly effective for intricate equations lacking analytical solutions. Notable advancements include rational solutions for equations like the nonlinear Schrödinger (NLS) equation and exploring multiple rogue waves in equations like the Boussinesq equation. On the way, Clarkson et al. [7] utilized the symbolic computation approach to obtain solutions for the Boussinesq equation, which is known to be solvable using the inverse scattering method. Expanding on earlier research, they derived rational solutions for the Boussinesq equation. These solutions exhibit algebraic decay and are characterized

*Talker
Email addresses: mmadadi@iasbs.ac.ir. (Majid Madadi), easadi@iasbs.ac.ir. (Esmaeel Asadi)

by two arbitrary parameters. The expressions for these solutions involve special polynomials derived from a bilinear equation. Notably, these rational solutions display resemblances to the rogue wave solutions found in the NLS equation.

The primary objective of this article is to investigate a novel $(3+1)$-dimensional Hirota bilinear (HB) equation [8] in fluid dynamics, expressed as

$$U_{yt} + c_1 \left(U_{xxxy} + 3(2U_x U_y + U_{xy} U) + 3U_{xx} \int^x U_y dx' \right) + c_2 U_{yy} + c_3 U_{zz} = 0, \tag{1}$$

where $U = U(x, y, z, t)$ and c_1, c_2, and c_3 are arbitrary constants. It is worth noting that when considering $c_3 = 0$, the above $(3+1)$-dimensional HB equation reduces to the $(2+1)$-dimensional HB equation [9]. Equation (1) serves as a valuable mathematical and physical model for studying nonlinear phenomena in the fields of shallow water and fluid mechanics. By incorporating higher-dimensionality, Eq. (1) provides a more precise description of complex motion laws and offers a more comprehensive explanation for various physical problems.

The findings of our study indicate that the model is not Painlevé integrable. However, upon further investigation, we discovered that the model can achieve integrability when setting $c_2 = c_3 = 0$. This integrable reduction, leads us to a previously discovered equation known as the $(2+1)$-dimensional Boiti-Leon-Manna-Pempinelli (BLMP) equation, as referenced in [10]. It is worth mentioning that by the following variable transformations: $y, z \to x$, $t \to -t$, $c_2 = -c_3$ and $u_x \to u$, Eq. (1) can be further reduced to the $(1+1)-$dimensional pioneering Korteweg–de Vries (KdV) equation. Additionally, the model (1) exhibits multiple-soliton solutions.

Numerous solutions have been obtained for Eq (1) through various approaches. For instance, Ref. [8] explored the Bäcklund transformation, Hirota bilinear form, and lump-type solution. Ref. [11] discussed the lump-N-soliton solutions, lump-periodic solutions, and multilump solutions. Additionally, Ref. [12] employed the long-wave limit method and selected specific parameters to derive solutions such as M-lump solutions, rogue waves, breathers, and interaction solutions.

The main research focus of this study revolves around obtaining exact multi-rogue wave solutions for equation (1) using the Hirota bilinear technique and symbolic computation method. The study also explores generalized multi-rogue waves, which are controllable rogue waves with adjustable characteristics.

This work is structured as follows: In Section 2, we delve into the investigation of multiple-rogue waves. We explore the characteristics and behavior of these waves. Moreover, we examine the roots of the rational solutions, revealing an intriguing correlation between these intricate roots and the solutions found in renowned Painlevé II and IV equations. Moving on to Section 3, we shift our focus to the study of generalized controllable rogue waves. Finally, in Section 4, we present the conclusion of our study.

2. Multiple rogue waves

In this section, we outline the approach utilized to verify the proposed rogue wave solution, known as the symbolic computation technique. We start by considering a nonlinear partial differential equation of the form

$$N(U, U_x, U_y, U_t, U_z, U_{tx}, U_{xx}, \dots) = 0, \quad U = U(x, y, t, z). \tag{2}$$

To initiate the analysis, we conduct the Painlevé analysis and introduce the transformation

$$U = T(f), \tag{3}$$

where $f = f(x, y, t, z)$ is a dependent variable function. Substituting Eq. (3) into Eq. (2), we obtain the resulting equation in Hirota bilinear form

$$\mathcal{B}(D_\zeta, D_z; f) = 0, \tag{4}$$

where $\zeta = x + \gamma y - \nu t$, and γ and ν are real parameters. Here, D_ζ and D_z are expressed as the Hirota bilinear operator [13].

Let

$$f=\sum_{m=0}^{n(n+1)/2}\sum_{j=0}^{m}a_{j,m}\zeta^{2j}z^{2(m-j)}. \tag{5}$$

By inserting Eq. (5) into Eq. (4) and setting all the coefficients of different powers of $\zeta^m z^n$, $(m,n=0,1,2,\dots)$ to zero, we obtain a system of algebraic equations. Solving this system will yield the values of the coefficients a_i. Finally, by substituting the obtained values of the coefficients a_i into the test function (5), and subsequently inserting this updated function into the logarithmic transformation (4), one can achieve an $n-$rogue wave solution to Eq. (2).
To obtain multiple-rogue waves for Eq. (1), we set $X=x+\delta y-\eta t$ and $Y=z$ in Eq. (1). This transformation yields

$$(\delta^2c_2-\eta\delta)U_{XX}+c_1\delta(U_{XXXX}+6U_X^2+6UU_{XX})+c_3U_{YY}=0, \tag{6}$$

where δ and η are two real parameters. By applying the logarithmic transformation

$$U_n=2\ln(F_n)_{XX},\quad n\geq 1. \tag{7}$$

Eq. (1) can be equivalently expressed as

$$[(\delta^2c_2-\eta\delta)D_X^2+c_1\delta D_X^4+c_3D_Y^2]F_n.F_n=0. \tag{8}$$

Here, $F_n=F_n(X,Y)\neq 0$ is mentioned in expression (5), and it is a polynomial of degree $\frac{1}{2}n(n+1)$ in X^2 and Y^2, with a total degree of $\frac{1}{2}n(n+1)$. The constants $a_{j,m}$ are unknown constants that need to be determined.
Utilizing upon procedure, one can obtain

$$F_1(X,Y)=-\frac{X^2a_{0,1}L_3}{\delta\left(-\delta L_2+\eta\right)}+Y^2a_{0,1}-\frac{3a_{0,1}L_3L_1}{\delta\left(-\delta L_2+\eta\right)^2}. \tag{9}$$

Consequently, the one-rogue wave solution for the reduced Eq. (6) is constructed as

$$U_1=-\frac{4L_3\left(\delta L_2-\eta\right)\left(\left(\delta L_2L_3-\eta L_3\right)X^2+\left(-\delta^3L_2^2+2\delta^2L_2\eta-\delta\,\eta^2\right)Y^2+3L_3L_1\right)}{\left(\left(\delta L_2L_3-\eta L_3\right)X^2+\left(\delta^3L_2^2-2\delta^2L_2\eta+\delta\,\eta^2\right)Y^2-3L_3L_1\right)^2}. \tag{10}$$

It is worth mentioning that this rogue wave exhibits the following characteristics:

$$\lim_{X\to\pm\infty}U_1(X,Y)=0,\qquad \lim_{Y\to\pm\infty}U_1(X,Y)=0. \tag{11}$$

Moreover, the function $U_1(X,Y)$ possesses the following three critical points.

$$(X,Y)=(0,0),\quad (X,Y)=\left(\frac{3\sqrt{-\left(\delta L_2-\eta\right)L_1}}{\delta L_2-\eta},0\right),\quad (X,Y)=\left(-\frac{3\sqrt{-\left(\delta L_2-\eta\right)L_1}}{\delta L_2-\eta},0\right), \tag{12}$$

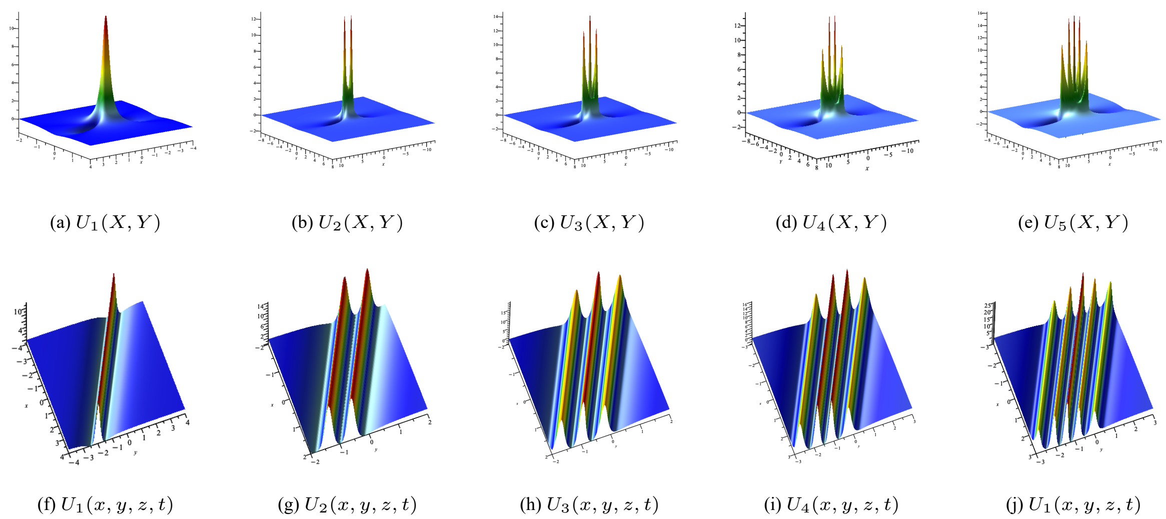
From the expression provided above, it is evident that the amplitude of the peak of one-rogue wave is given by $A_{peak}=\frac{-4\delta L_2+4\eta}{3L_1}$, and the amplitudes of the two troughs can be determined as $A_{trough}=\pm\frac{\delta L_2-\eta}{6L_1}$. The dynamic characteristics of the one-rogue wave solutions are illustrated in Figure 1 (a).
By following the earlier process, for $n=2,3,4,5$, we derive the following expressions for two, three, four, and five-rogue waves:

$$\begin{aligned}F_2(X,Y)&=\left(\frac{3Y^2L_3^2}{\delta^2\left(-\delta L_2+\eta\right)^2}-\frac{25L_3^3L_1}{\delta^3\left(-\delta L_2+\eta\right)^4}\right)X^4+\left(-\frac{3Y^4L_3}{\delta\left(-\delta L_2+\eta\right)}+\frac{90Y^2L_3^2L_1}{\delta^2\left(-\delta L_2+\eta\right)^3}+\frac{125L_1^2L_3^3}{\delta^3\left(-\delta L_2+\eta\right)^5}\right)X^2\\&-\frac{X^6L_3^3}{\delta^3\left(-\delta L_2+\eta\right)^3}+Y^6-\frac{17Y^4L_3L_1}{\delta\left(-\delta L_2+\eta\right)^2}+\frac{475Y^2L_3^2L_1^2}{\delta^2\left(-\delta L_2+\eta\right)^4}-\frac{1875L_3^3L_1^3}{\delta^3\left(-\delta L_2+\eta\right)^6},\end{aligned} \tag{13}$$

$$
\begin{aligned}
F_3(X,Y) &= \left(\frac{98L_1L_3^6}{\delta^6\left(-\delta L_2+\eta\right)^7}-\frac{6L_3^5Y^2}{\delta^5\left(-\delta L_2+\eta\right)^5}\right)X^{10}+\left(-\frac{690L_1L_3^5Y^2}{\delta^5\left(-\delta L_2+\eta\right)^6}+\frac{735L_3^6L_1^2}{\delta^6\left(-\delta L_2+\eta\right)^8}+\frac{15L_3^4Y^4}{\delta^4\left(-\delta L_2+\eta\right)^4}\right)X^8\\
&+\frac{L_3^6X^{12}}{\left(-\delta^7L_2+\delta^6\eta\right)^6}+\left(-\frac{18620L_1^2L_3^5Y^2}{\delta^5\left(-\delta L_2+\eta\right)^7}+\frac{1540L_1L_3^4Y^4}{\delta^4\left(-\delta L_2+\eta\right)^5}+\frac{75460L_3^6L_1^3}{3\delta^6\left(-\delta L_2+\eta\right)^9}-\frac{20L_3^3Y^6}{\delta^3\left(-\delta L_2+\eta\right)^3}\right)X^6\\
&+\left(\frac{37450L_1^2L_3^4Y^4}{\delta^4\left(-\delta L_2+\eta\right)^6}-\frac{220500Y^2L_3^5L_1^3}{\delta^5\left(-\delta L_2+\eta\right)^8}-\frac{1460L_1L_3^3Y^6}{\delta^3\left(-\delta L_2+\eta\right)^4}-\frac{5187875L_3^6L_1^4}{3\delta^6\left(-\delta L_2+\eta\right)^{10}}+\frac{15L_3^2Y^8}{\delta^2\left(-\delta L_2+\eta\right)^2}\right)X^4+Y^{12}\\
&+\left(-\frac{35420L_1^2L_3^3Y^6}{\delta^3\left(-\delta L_2+\eta\right)^5}-\frac{14700Y^4L_3^4L_1^3}{\delta^4\left(-\delta L_2+\eta\right)^7}-\frac{565950Y^2L_3^5L_1^4}{\delta^5\left(-\delta L_2+\eta\right)^9}+\frac{570L_3^2L_1Y^8}{\delta^2\left(-\delta L_2+\eta\right)^3}+\frac{159786550L_3^6L_1^5}{3\delta^6\left(-\delta L_2+\eta\right)^{11}}-\frac{6L_3Y^{10}}{\delta\left(-\delta L_2+\eta\right)}\right)X^2\\
&+\frac{878826025L_1^6L_3^6}{9\delta^6\left(-\delta L_2+\eta\right)^{12}}+\frac{16391725Y^4L_3^4L_1^4}{3\delta^4\left(-\delta L_2+\eta\right)^8}-\frac{58L_3L_1Y^{10}}{\delta\left(-\delta L_2+\eta\right)^2}+\frac{4335L_3^2L_1^2Y^8}{\delta^2\left(-\delta L_2+\eta\right)^4}-\frac{798980Y^6L_1^3L_3^3}{3\delta^3\left(-\delta L_2+\eta\right)^6}-\frac{300896750Y^2L_3^5L_1^5}{3\delta^5\left(-\delta L_2+\eta\right)^{10}}.
\end{aligned}
\tag{14}
$$

For the specific selection of values of parameters $\delta = 2, \eta = 3, L_1 = 1, L_2 = -3, L_3 = -1$, the expressions for F_4 and F_5 are presented in the Appendix.

Based on the expression of F_n, $(n = 1, \ldots, 5)$ available, according to transformation (7), we can obtain various types of waves by choosing different coordinates. For instance, when using (X, Y) coordinates, we can derive multi-rogue waves, while (x, y) coordinates are suitable for multi-soliton waves solitons. Figure 1 (a,...,e) illustrates the $n-$rogue wave solutions, displaying two, three, four and five-rogue respectively. The two-rogue wave solution $U_2(X, Y)$ exhibits two separated peaks, while $U_3(X, Y)$ has one sharper peak in the middle surrounded by two shorter peaks. In the case of $U_4(X, Y)$, there are four divided peaks, with two in the middle and two on the periphery. finally, $U_5(X, Y)$ has one sharper peak in the middle surrounded by four peaks. It is evident that the $n-$rogue solution $U_n(X, Y)$ consists of n separated peaks, with the maximum value located at $Y = 0$ and n maximum values of $U_n(X, Y)$. On the other hand, Figures 1 (f,...,j) showcase multi-soliton wave. From the evolutionary plots, we observe that these waves originate from a plane wave background and reach their maximum amplitude at $y = 0$ before eventually dissipating with time.



(a) $U_1(X, Y)$ (b) $U_2(X, Y)$ (c) $U_3(X, Y)$ (d) $U_4(X, Y)$ (e) $U_5(X, Y)$

(f) $U_1(x, y, z, t)$ (g) $U_2(x, y, z, t)$ (h) $U_3(x, y, z, t)$ (i) $U_4(x, y, z, t)$ (j) $U_1(x, y, z, t)$

Fig. 1. **Panels (a, b, c, d, e)**: the $n-$rogue wave $U_n(X, Y)$ with $\delta = 2, \eta = 3, L_1 = 1, L_2 = -3, L_3 = -1$. **Panels (f, g, h, i , j)**: the $n-$soliton wave $U_n(x, y, z, t)$ with the same parameters as well as $z = 0$ and $t = 0$.

In the context of the connection between the polynomial F_n and the rational solution U_n, and the role of complex roots in wave formation, it is essential to first understand the solutions of Painlevé Equations II and IV. These equations are typically represented in the following forms, respectively.

$$PII := w'' = 2w^3 + zw + \theta, \tag{15}$$

$$PIV := w'' = \frac{(w')^2}{2w} + \frac{3}{2}w^3 + 4zw^2 + 2(z^2 - \theta)w + \frac{\delta}{w}, \tag{16}$$

where $w = w(z)$ is a analytic function, θ and δ are real constants.
The rational solutions of PII (15) were originally expressed in terms of polynomials by Vorob'ev and Yablonskii [14]. These polynomials, now known as the Yablonskii–Vorob'ev polynomials, have been further studied by Clarkson and Mansfield [15]. Similarly, Okamoto [16] derived related polynomials, known as the Okamoto polynomials, which are associated with certain rational solutions of PIV (16). The results of Okamoto were subsequently generalized by Noumi and Yamada [17], who demonstrated that all rational solutions of PIV can be represented using logarithmic derivatives of two sets of special polynomials: the generalized Hermite polynomials and the generalized Okamoto polynomials.
The Yablonskii–Vorob'ev polynomials possess several distinctive characteristics. Firstly, they are monic polynomials with integer coefficients, having a degree of $\frac{1}{2}n(n+1)$. Secondly, all roots of these polynomials are simple, resulting in $N(N+1)/2$ distinct roots. The arrangement of these roots forms a consistently structured triangle, and they are located on circles with the origin as the center. On the other hand, the Okamoto polynomials exhibit different characteristics. They are polynomials of degree $n(n-1)$ and are monic polynomials in $\sqrt{2}z$ with integer coefficients. Additionally, these polynomials are even, specifically monic polynomials in $2z^2$ with a degree of $\frac{1}{2}n(n-1)$. The roots of the Okamoto polynomials are arranged in symmetrically reflected triangles along the real axis.

Remark 2.1. The rational solutions of Eq. (1) can be obtained using the Yablonskii–Vorob'ev and generalized Okamoto polynomials.

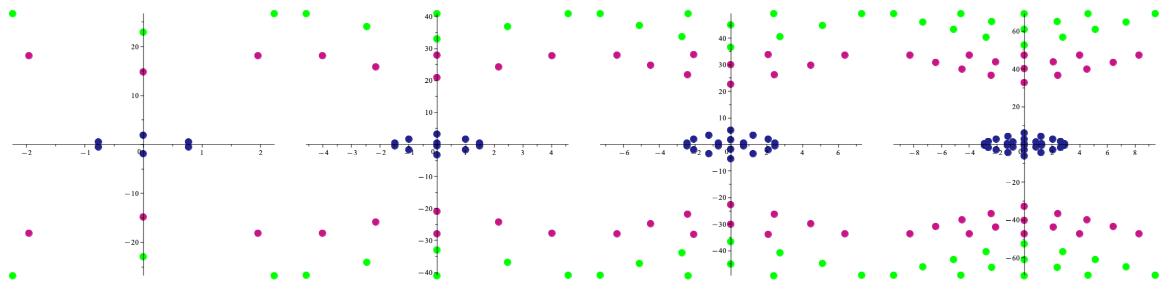
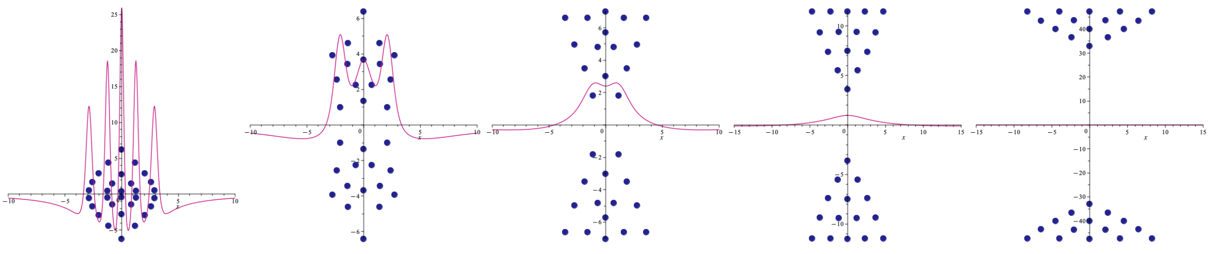


Fig. 2. Complex roots of F_n for $n = 2; 3; 4; 5$ with $Y = 0$ (black) and $Y = 2n$ (red) and $Y = 3n$ (green).

The graphical representations in Figure 2 reveal intriguing patterns in the complex roots of F_n for $Y = 0$, $Y = 2n$ and $Y = 3n$. Notably, the roots form roughly equilateral triangles reminiscent of Pascal triangles, hinting at underlying geometric relationships. Surprisingly, these roots deviate from straight lines, instead aligning along curves. This property arises because these roots are obtained from polynomials with non-integer coefficients. Moreover, the plots exhibit symmetrical properties, remaining unchanged under rotations through $\frac{2}{3}\pi$ and reflections in the real axis. As the parameter Y increases, the roots gradually shift away from the real axis, indicating a dynamic behavior. Particularly striking is the arrangement of roots along concentric circles encircling the imaginary axis, suggesting a deeper structural organization within the polynomial solutions. These observations offer valuable insights into the intricate behavior of F_n polynomials under varying conditions, enriching our understanding of their complex dynamics.
In Figure 3, plots depict the trajectories of the complex roots of F_5, with the solution u_6 superimposed. These plots demonstrate that as the roots move away from the real axis, the solution decays to zero.

3. Generalized rogue wave solutions (Controllable rogue waves)

In the preceding section, we examined the rational solutions of Eq (1) and investigated the inherent connection between the bilinear equation and the rational solution. Now, we will proceed to construct the generalized rational solutions



(a) Y=0 (b) Y=0.5 (c) Y=1 (d) Y=2 (e) Y=10

Fig. 3. Plots of the loci of the complex roots of $F_5(X,Y)$ (black) with the solution $U_5(X,Y)$ (red) superimposed , for different values of Y.

of Eq (1) by introducing two additional parameters, α and β. These parameters possess more intricate characteristics compared to their counterparts. As a result, the generalized rational solution exhibits multi-rogue solutions with two control parameters α and β. By utilizing the bilinear form (8), we can derive the generalized rational solution of the equation using the following theorem.

Theorem 3.1. *Equation (1) has generalized rational solution*

$$\tilde{U}_n(X,Y;\alpha,\beta) = 2\ln\left(\tilde{F}_n(X,Y;\alpha,\beta)\right)_{XX}, \tag{17}$$

with

$$\tilde{F}_{n+1}(X,Y;\alpha,\beta) = F_{n+1}(X,Y) + 2\alpha t P_n(X,Y) + 2\beta x Q_n(X,Y) + (\alpha^2+\beta^2)F_{n-1}(X,Y), \tag{18}$$

where $F_n(X,Y) \neq 0$ is given by (5) and $F_0 = 1$, $F_{-1} = 0$ and

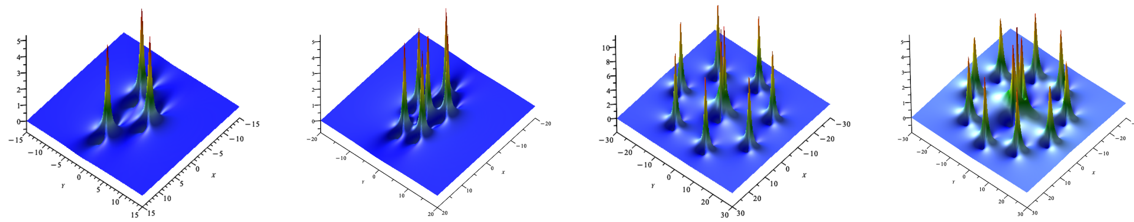
$$P_n(X,Y) = \sum_{m=0}^{n(n+1)/2}\sum_{i=0}^{m} b_{j,m}X^{2j}Y^{2(m-j)}, \quad Q_n(X,Y) = \sum_{m=0}^{n(n+1)/2}\sum_{i=0}^{m} c_{j,m}X^{2j}Y^{2(m-j)}, \tag{19}$$

where $b_{j,m}$ and $c_{j,m}$ are real parameters, and $P_0 = Q_0 = 0$.

By substituting the expression $\tilde{F}_n$ into Eq. (17) and equating the coefficients of all powers of X and Y to zero, we can obtain the expressions for Q_n and P_n. These expressions are derived by solving the resulting system of algebraic equations. With the expressions for F_n obtained in Section 2, we can then obtain the general rational solution $\tilde{U}_n$ by substituting the solutions for F_n, Q_n and P_n into the appropriate equation.
In the subsequent equations, we express Q_n and P_n for $n = 2, ..., 5$ as shown below:

$$\begin{aligned}
p_1 &= \frac{\sqrt{-\beta^2\delta^2 L_2 L_3 + \beta^2\delta\eta L_3 + 9\beta^2 L_3^2 + 9\alpha^2 L_3^2}\,X^2}{(\delta L_2 - \eta)\,\alpha\delta} - \frac{\sqrt{L_3\left(\left(9\beta^2+9\alpha^2\right)L_3 + \beta^2\delta\left(-\delta L_2+\eta\right)\right)}\,Y^2}{3\alpha L_3} \\
&\quad - \frac{5L_1\sqrt{L_3\left(\left(9\beta^2+9\alpha^2\right)L_3 + \beta^2\delta\left(-\delta L_2+\eta\right)\right)}}{3\left(-\delta L_2+\eta\right)^2\alpha\delta}, \\
p_2 &= \left(\frac{9L_3Y^4}{\delta\left(-\delta L_2+\eta\right)} - \frac{190L_3^2L_1Y^2}{\delta^2\left(-\delta L_2+\eta\right)^3} + \frac{665L_1^2L_3^3}{\delta^3\left(-\delta L_2+\eta\right)^5}\right)X^2 + Y^6 + \frac{7L_1L_3Y^4}{\delta\left(-\delta L_2+\eta\right)^2} - \frac{245L_1^2L_3^2Y^2}{\delta^2\left(-\delta L_2+\eta\right)^4}, - \frac{18865L_1^3L_3^3}{3\delta^3\left(-\delta L_2+\eta\right)^6} \\
&\quad - \frac{5L_3^3X^6}{\delta^3\left(-\delta L_2+\eta\right)^3} + \left(-\frac{5L_3^2Y^2}{\delta^2\left(-\delta L_2+\eta\right)^2} - \frac{105L_1L_3^3}{\delta^3\left(-\delta L_2+\eta\right)^4}\right)X^4, \\
q_1 &= \frac{L_3X^2}{3\delta\left(-\delta L_2+\eta\right)} + Y^2 - \frac{L_1L_3}{3\delta\left(-\delta L_2+\eta\right)^2}, \\
q_2 &= -\frac{\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_3^3X^6}{\delta^4\left(-\delta L_2+\eta\right)^4\beta} - \frac{9\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_3^2X^4Y^2}{\delta^3\left(-\delta L_2+\eta\right)^3\beta} + \frac{5\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_3X^2Y^4}{\delta^2\left(-\delta L_2+\eta\right)^2\beta} - \frac{5\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,Y^6}{\delta\left(\delta L_2-\eta\right)\beta} \\
&\quad - \frac{13\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_1L_3^3X^4}{\delta^4\left(-\delta L_2+\eta\right)^5\beta} - \frac{230L_3^2L_1\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,X^2Y^2}{\delta^3\left(-\delta L_2+\eta\right)^4\beta} - \frac{45L_1L_3\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,Y^4}{\delta^2\left(-\delta L_2+\eta\right)^3\beta} \\
&\quad + \frac{245\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_1^2L_3^3X^2}{\delta^4\left(-\delta L_2+\eta\right)^6\beta} + \frac{535\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_1^2L_3^2Y^2}{\delta^3\left(-\delta L_2+\eta\right)^5\beta} - \frac{12005\sqrt{\delta\left(\delta L_2-\eta\right)L_3\beta^2}\,L_1^3L_3^3}{3\delta^4\left(-\delta L_2+\eta\right)^7\beta}.
\end{aligned} \tag{20}$$

For the specific selection of values of parameters $\delta = 2, \eta = 3, L_1 = 1, L_2 = -3, L_3 = -1$, the expressions for Q_n and P_n, $(n = 3, 4)$ are presented in the Appendix. Plots of the solutions $\tilde{U}_n(X, Y; \alpha, \beta)$, $n = 1, \ldots, 4$ of Eq. (6) for specific values of the parameters α and β are given in Figure 4.



(a) $\tilde{U}_1(X, Y; 30, 30)$ (b) $\tilde{U}_2(X, Y; 1000, 1000)$ (c) $\tilde{U}_3(X, Y; 10^7, 10^7)$ (d) $\tilde{U}_3(X, Y; 10^9, 10^9)$

Fig. 4. General rogue wave solutions $\tilde{U}_n(X, Y, \alpha, \beta)$, with $\delta = 1, \eta = 1, L_1 = 1, L_2 = -3, L_3 = -1$.

As observed in Figure 5, for $|\alpha|, |\beta| > 0$, the peaks of each solution are arranged along a specific circle. The circles $C_n(X, Y, r)$ corresponding to the choices of parameters α and β are as $C_1(-0.03238, 0.01011, 5.0719)$, $C_2(-0.0022, -0.0040, 6.294$ $C_3(-0.0022, -0.0040, 19.5882)$ and $C_4(2.35 \times 10^{-6}, -2.98 \times 10^{-12}, 19.7346)$. As the values of α and β increase, the spacing between the peaks becomes larger, and the peaks themselves move along distinct paths in a straight line. These paths correspond to three, five, seven, and so on, different trajectories. For $\tilde{U}_n$ these trajectories are:

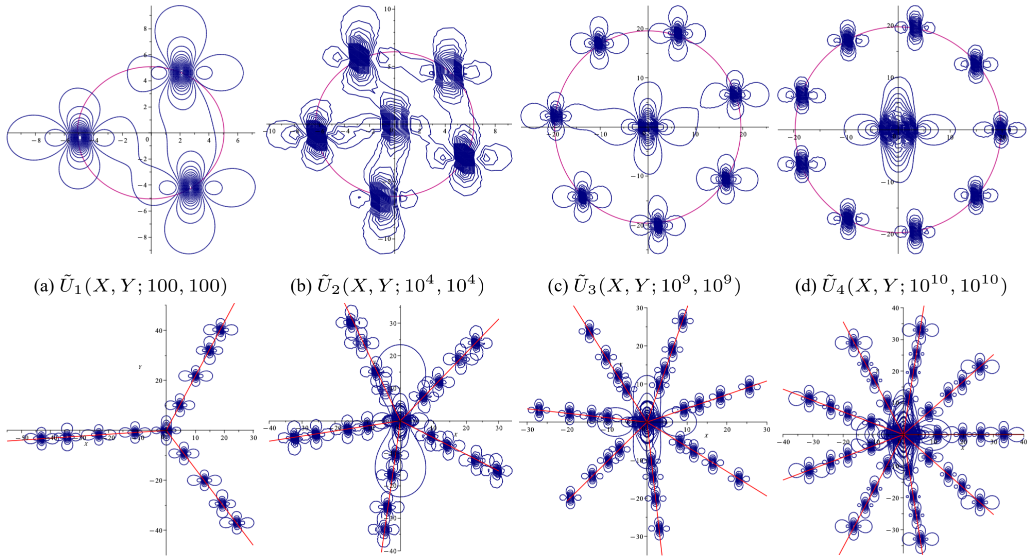
$$
\begin{aligned}
&\tilde{U}_1 : [Y_1 = 2.159X_1 + 0.080, Y_2 = 0.077X_2 + 0.0126, Y_3 = -1.529X_3 - 0.039],\\
&\tilde{U}_2 : [Y_1 = 1.044X_1 - 0.0016, Y_2 = -2.1242X_2 - 0.008, Y_3 = 0.1566X_3 - 0.003, Y_4 = 7.0X_4 + 0.011,\\
&\qquad Y_5 = -0.514X_5 - 0.005],\\
&\tilde{U}_3 : [Y_1 = 0.6476X_1 - 0.0054, Y_2 = 2.9925X_2 + 0.002, Y_3 = 0.3594X_3 - 0.0032, Y_4 = -1.659X_4 - 0.007,\\
&\qquad Y_5 = -0.1156X_5 - 0.0042, Y_6 = 1.0366X_6 - 0.001, Y_7 = -9.3081X_7 - 0.0247],\\
&\tilde{U}_4 : [Y_1 = -5.951 \times 10^{-9}X_1 - 2.975 \times 10^{-12}, Y_2 = 0.8420X_2 - 1.99 \times 10^{-6}, Y_3 = 5.926X_3 - 0.00001,\\
&\qquad Y_4 = -1.7782X_4 + 4.20 \times 10^{-6}, Y_5 = -0.3584X_5 + 8.45 \times 10^{-7}, Y_6 = 0.3584X_6 - 8.46 \times 10^{-7},\\
&\qquad Y_7 = 1.778X_7 - 4.19 \times 10^{-6}, Y_8 = -5.92656X_8 + 0.00001, Y_9 = -0.8420X_9 + 1.99 \times 10^{-6}].
\end{aligned} \tag{21}
$$

4. Conclusion

This paper investigates multiple-rogue wave solutions arising from a novel extension of the Hirota bilinear equation, a crucial model for understanding nonlinear phenomena in shallow water and fluid mechanics. Utilizing the bilinear equation and symbolic computation methods, we have derived rational and generalized rational solutions. Plotting these solutions in various planes reveals two distinct waveforms: multi-rogue and multi-soliton waves, as demonstrated in Figures 1 and 4, which showcases occurrences of two, three, four, and five-rogue waves. Furthermore, Figures 1 illustrate the multi-soliton wave, highlighting the presence of multiple-order line rogue waves. Our numerical simulations unveil that these waves emerge from a background of plain waves. Additionally, we analyze the trajectory of each peak in the first to fifth-order controllable rogue waves.

Moreover, we delve into the intrinsic relationship between F_n and U_n by scrutinizing the connection between the complex roots of F_n and wave formation. Interestingly, our exploration reveals that these roots align with the Yablonskii–Vorob'ev and generalized Okamoto polynomials, further emphasizing the importance of these polynomials in understanding wave dynamics.

This article presents clear and easily digestible ideas, distinguishing itself from methods such as Darboux transformation and KP-reduction. Our approach facilitates the creation of arbitrary n-order rogue wave solutions, enticing exploration into higher-dimensional nonlinear equations. Despite the complexity of the calculations, the conceptual clarity of our method endures. These findings promise a deeper understanding of nonlinear phenomena in high-dimensional systems, propelling further advancements in the field.



(a) $\tilde{U}_1(X,Y;100,100)$ (b) $\tilde{U}_2(X,Y;10^4,10^4)$ (c) $\tilde{U}_3(X,Y;10^9,10^9)$ (d) $\tilde{U}_4(X,Y;10^{10},10^{10})$

(e) $\tilde{U}_1(X,Y;\alpha,\beta)$ (f) $\tilde{U}_2(X,Y;\alpha,\beta)$ (g) $\tilde{U}_3(X,Y;\alpha,\beta)$ (h) $\tilde{U}_4(X,Y;\alpha,\beta)$

Fig. 5. **Panels (a, b, c, d)**: The peaks of each solution $\tilde{U}_n(X,Y,\alpha,\beta)$ are arranged along a specific circle. **Panels (e, f, g, h)**: The Trajectories of each peak (red lines) in the solutions $\tilde{U}_n(X,Y,\alpha,\beta)$. with $\delta=1,\eta=1,L_1=1,L_2=-3,L_3=-1$.

References

[1] Akhmediev, Nail, Adrian Ankiewicz, and Majid Taki. Waves that appear from nowhere and disappear without a trace. Physics Letters A, 373 (2009) 675-678.

[2] Dubard, Ph and Gaillard, P and Klein, C and Matveev, VB, On multi-rogue wave solutions of the NLS equation and positon solutions of the KdV equation, The European Physical Journal Special Topics, 185 (2010) 247-258.

[3] Gaillard, Pierre, Families of quasi-rational solutions of the NLS equation and multi-rogue waves, Journal of Physics A: Mathematical and Theoretical, 44 (2011) 421-440.

[4] Ohta, Yasuhiro and Feng, Bao-Feng, General rogue wave solution to the discrete nonlinear Schrödinger equation, Physica D: Nonlinear Phenomena, 439 (2022) 133-148.

[5] Yang, Bo and Yang, Jianke, General rogue waves in the Boussinesq equation, Journal of the Physical Society of Japan, 89 (2020) 24003-24020.

[6] Yang, Bo and Yang, Jianke, Rogue wave patterns in the nonlinear Schrödinger equation, Physica D: Nonlinear Phenomena, 419 (2021) 132850-132872.

[7] Clarkson, Peter A and Dowie, Ellen, Rational solutions of the Boussinesq equation and applications to rogue waves, Transactions of Mathematics and its Applications, 1 (2017) 879-893.

[8] Hosseini, Kamyar and Samavat, Majid and Mirzazadeh, Mohammad and Ma, Wen-Xiu and Hammouch, Zakia, A new (3+ 1)-dimensional Hirota bilinear equation: its Bäcklund transformation and rational-type solutions, Regular and Chaotic Dynamics, 25 (2020) 383-391.

[9] Lü, Xing and Hua, Yan-Fei and Chen, Si-Jia and Tang, Xian-Feng, Integrability characteristics of a novel (2+ 1)-dimensional nonlinear model: Painlevé analysis, soliton solutions, Bäcklund transformation, Lax pair and infinitely many conservation laws, Communications in Nonlinear Science and Numerical Simulation, 95 (2021) 105-119.

[10] Gilson, CR and Nimmo, JJC and Willox, R, A (2+ 1)-dimensional generalization of the AKNS shallow water wave equation, Physics Letters A, 180 (1993) 337-345.

[11] Tian, Yu and Liu, Jian-Guo, Study on dynamical behavior of multiple lump solutions and interaction between solitons and lump wave, Nonlinear Dynamics, 104 (2021) 1507-1517.

[12] Wang, Binji and Ma, Zhimin and Xiong, Sihan, M-lump, rogue waves, breather waves, and interaction solutions among four nonlinear waves to new (3+ 1)-dimensional Hirota bilinear equation, Nonlinear Dynamics, 111 (2023) 9477-9494.

[13] Hirota, Ryogo, The direct method in soliton theory, Cambridge University Press, 34 (2004) 5423-5440.

[14] YABLONSKII, AL, On rational solutions of the second Painlevé equation, Vesti AN BSSR, ser. fiz-tekh, Nauk, 3 (1959) 30-35.

[15] Clarkson, Peter A and Mansfield, Elizabeth L, The second Painlevé equation, its hierarchy and associated special polynomials, Nonlinearity, 16 (2003) 324-331.

[16] Okamoto, Kazuo, Studies on the Painlevé equations: III. Second and fourth Painlevé equations, P II and P IV, Mathematische Annalen, 275 (1986) 221-255.

[17] Noumi, Masatoshi and Yamada, Yasuhiko, Symmetries in the fourth Painlevé equation and Okamoto polynomials, Nagoya mathematical journal, 153 (1999) 53-86.

Appendix

$$
\begin{aligned}
F_4(X,Y) = {} & 0.497618547178Y^6 - 9.99200722141 \times 10^{-16}Y^7 - 1.35525271557 \times 10^{-19}Y + 0.0000724218232120516 + 0.0409848309576036Y^4 \\
& + 2.39808173313906 \times 10^{-14}Y^{13} + 1.59872115542604 \times 10^{-14}Y^{11} + \left(4.840025824 \times 10^{-10}Y^2 + 1.333476503 \times 10^{-10}\right)X^{18} \\
& + \left(4.307854803 \times 10^{-7} + 0.00046855476Y^{10} + 0.000996077Y^8 + 0.000640579Y^6 + 0.000111788Y^4 + 2.167724802 \times 10^{-6}Y^2\right)X^{10} \\
& + [-5.233391262 \times 10^{-6} + 0.005466472Y^{12} + 0.013498839Y^{10} + 0.01223467Y^8 + 0.003741649724Y^6 + 0.00030252582Y^4 \\
& + 0.00001402735639Y^2]X^8 + [0.00002924767164 + 0.04373177842Y^{14} + 0.1155768430Y^{12} + 0.1366458703Y^{10} + 0.05855958861Y^8 \\
& + 0.01511191583Y^6 + 0.00002565608791Y^4 - 0.00021459555Y^2]X^6 + 2.41322781598Y^{10} + 1.35563829483Y^8 + 3.91069060907Y^{12} \\
& - 5.20417042781913 \times 10^{-18}Y^3 + \left(3.049216270 \times 10^{-8}Y^4 + 2.09503974996372 \times 10^{-8}Y^2 + 1.17155435611833 \times 10^{-9}\right)X^{16} \\
& + \left(1.138374074 \times 10^{-6}Y^6 + 1.33584712782492 \times 10^{-6}Y^4 + 2.72029206611445 \times 10^{-7}Y^2 + 3.53779480222407 \times 10^{-9}\right)X^{14} \\
& + [-2.71820567410810 \times 10^{-8} + 0.00002789016481Y^8 + 0.0000466733370279255Y^6 + 0.0000189051408693252Y^4 \\
& + 1.49882758910389 \times 10^{-6}Y^2]X^{12} + [0.22959183Y^{16} + 0.00011429038857 - 5.09796286764 \times 10^{-17}Y^{15} + 0.593502707139Y^{14} \\
& + 4.58816658088026 \times 10^{-16}Y^{13} + 0.860132682708485Y^{12} + 5.60775915440921 \times 10^{-16}Y^{11} + 0.512883055217146Y^{10} \\
& - 4.07837029411579 \times 10^{-16}Y^9 + 0.147551575299234Y^8 - 1.68870019990732 \times 10^{-16}Y^7 + 0.0255335147481418Y^6 \\
& + 5.97417523552117 \times 10^{-17}Y^5 - 0.00233752715965681Y^4 - 1.31431855181466 \times 10^{-17}Y^3 + 0.00195627417600323Y^2 \\
& - 6.59648515588796 \times 10^{-19}Y]X^4 + [1.596209912Y^{16} + 0.71428571Y^{18} + 2.637293Y^{14} + 2.705186614Y^{12} + 0.52002066Y^{10} \\
& + 0.2487391327Y^8 + 0.1598207859Y^6 + 0.07479018500Y^4 + 0.007307250727Y^2 - 0.00005964150150]X^2 + 0.00416734871516429Y^2 \\
& + 1.59872115542604 \times 10^{-14}Y^{15} + 8.88178419681132 \times 10^{-16}Y^{17} + Y^{20} + 3.50555465719951Y^{14} + 2.40368596374425Y^{16} \\
& + 3.457161303 \times 10^{-12}X^{20} + 1.53061224503500Y^{18} + 1.33226762952170 \times 10^{-15}Y^9 - 1.66533453690212 \times 10^{-16}Y^5,
\end{aligned}
$$

$$
\begin{aligned}
F_5(X,Y) = {} & .053013138Y^6 - 0.00015395657X^4 + 0.2594667676Y^4 + 0.00003792567X^6 - 0.0009109286X^{10}Y^4 + 756.7987265X^2Y^{14} \\
& + 1.583512581X^6Y^{10} + 0.00004752X^{12}Y^4 + 9.235383 \times 10^{-7}X^{14}Y^2 + 1130.156099X^2Y^{16} + X^{10}Y^{20} + 1.841055 \times 10^{-12}X^{26}Y^2 \\
& + 524.3556X^2Y^{28} + 203.91608X^4Y^{26} + 49.09090909X^6Y^{24} + 3943.009791X^2Y^{20} + 1155.39104X^4Y^{18} + 213.48721X^6Y^{16} \\
& + 25.50362054X^8Y^{14} + 1.949113867X^{10}Y^{12} + 0.09605519834X^{12}Y^{10} + 0.003043499719X^{14}Y^8 + 0.00005941625044X^{16}Y^6 \\
& + 6.397024894 \times 10^{-7}X^{18}Y^4 + 3.058057092 \times 10^{-9}X^{20}Y^2 + 4025.053958X^2Y^{22} + 1605.510470X^4Y^{20} + 349.2707372X^6Y^{18} \\
& + 48.05312735X^8Y^{16} + 4.450906556X^{10}Y^{14} + 0.2846558918X^{12}Y^{12} + 0.01260742596X^{14}Y^{10} + 0.0003805218461X^{16}Y^8 \\
& + 7.5249813 \times 10^{-6}X^{18}Y^6 + 9.0207734 \times 10^{-8}X^{20}Y^4 + 5.5648027 \times 10^{-10}X^{22}Y^2 + 2721.44522X^2Y^{24} + 1251.7223X^4Y^{22} \\
& + 0.064771477X^8Y^8 + 0.00339912X^{10}Y^6 + 26.066045X^6Y^{12} - 5.076284 \times 10^{-8}X^{16}Y^2 + 1858.211X^2Y^{18} + 95.646466X^6Y^{14} \\
& + 8.582936871X^8Y^{12} + 0.0002701626970X^{14}Y^6 + 2.438624078 \times 10^{-6}X^{16}Y^4 + 1.007260812 \times 10^{-8}X^{18}Y^2 + 0.4870153464X^6Y^6 \\
& + 25.0761583X^4Y^{10} - 0.000013215X^{12}Y^2 + 73.19658X^4Y^{12} + 4.858407 \times 10^{-6}X^8 + 138.0376439Y^{10} + 0.0001240554X^{10} \\
& + 27.75892269Y^8 + 704.2945294Y^{12} - 3.085426102 \times 10^{-8}X^{12} + 28.675213X^8Y^{20} + 315.1460474X^6Y^{20} + 0.0045644374X^4Y^2 \\
& + 0.092592592X^{12}Y^{18} + 0.006613756X^{14}Y^{16} + 0.0003674309X^{16}Y^{14} + 0.00001587664X^{18}Y^{12} + 5.292214940 \times 10^{-7}X^{20}Y^{10} \\
& + 1.336417914 \times 10^{-8}X^{22}Y^8 + 2.474847 \times 10^{-10}X^{24}Y^6 + 3.172882 \times 10^{-12}X^{26}Y^4 + 2.5181603 \times 10^{-14}X^{28}Y^2 + 8.18181X^8Y^{22} \\
& + 50.38118603X^8Y^{18} + 5.498574166X^{10}Y^{16} + 0.4252971960X^{12}Y^{14} + 0.02370065210X^{14}Y^{12} + 0.0009522201517X^{16}Y^{10}
\end{aligned}
$$

$$
\begin{aligned}
& + 0.000027173X^{18}Y^8 + 5.3189583 \times 10^{-7}X^{20}Y^6 + 6.6770234 \times 10^{-9}X^{22}Y^4 + 4.6958509 \times 10^{-11}X^{24}Y^2 + 1404.755245X^2Y^{26} \\
& + 639.8601399X^4Y^{24} + 167.1794872X^6Y^{22} + 0.3045531286X^{12}Y^{16} + 0.01997316812X^{14}Y^{14} + 0.0009823165116X^{16}Y^{12} \\
& + 0.00003608054X^{18}Y^{10} + 9.7375749 \times 10^{-7}X^{20}Y^8 + 1.869039 \times 10^{-8}X^{22}Y^6 + 2.4031643 \times 10^{-10}X^{24}Y^4 + 228.664704X^4Y^{14} \\
& + 1.213268590X^8Y^{10} + 0.04085683593X^{10}Y^8 + 0.0008978517124X^{12}Y^6 + 3.227565252 \times 10^{-6}X^{14}Y^4 + 561.7687957X^4Y^{16} \\
& + 0.4335693326X^{10}Y^{10} + 0.01385455661X^{12}Y^8 + 0.00451340562X^6Y^2 + 42.40788600X^2Y^8 + 2.93732553X^4Y^6 + 0.10523302X^6Y^4 \\
& + 0.001674303423X^8Y^2 + 199.03978X^2Y^{10} + 13.9872257X^4Y^8 + 0.0058447316X^8Y^4 + 0.00001964256X^{10}Y^2 + 5.144669X^2Y^6 \\
& + 0.2627577054X^4Y^4 + 3.458372903X^{10}Y^{18} - 0.03268649660X^2Y^4 + 459.4842460X^2Y^{12} + 0.3377108057X^6Y^8 - 0.02226484725X^8Y^6 \\
& + 0.00004428331617 + 0.01349109633Y^2 + 0.0002754352545X^2Y^2 + 5516.840083Y^{20} + 1816.649047Y^{14} - 1.049499602 \times 10^{-7}X^{14} \\
& + 3163.515422Y^{16} + 1.0220873 \times 10^{-8}X^{16} - 5.8399748 \times 10^{-10}X^{18} + 2.3768798 \times 10^{-11}X^{20} + 4431.776673Y^{18} + 0.0002872289X^2 \\
& + 9.326519804 \times 10^{-17}X^{30} + 1262.337662Y^{28} + 629.2267732Y^{30} + 1.106191562 \times 10^{-12}X^{24} + 1.316593712 \times 10^{-13}X^{26} \\
& + 6.269493868 \times 10^{-15}X^{28} + 4962.184097Y^{22} + 2197.762238Y^{26} + 3531.667098Y^{24} + 1.774822906 \times 10^{-12}X^{22},
\end{aligned}
$$

$$
\begin{aligned}
p_3 =& \left(-0.00001195292 + 0.0000260308Y^2\right) X^{10} + \left(-7.114837962 \times 10^{-6} + 0.0005020229666Y^2 + 0.001639941Y^4\right) X^8 + 0.12009277Y^4 \\
&+ \left(-0.00005610557 + 0.00183278225Y^2 + 0.02280299875Y^4 + 0.01311953353Y^6\right) X^6 + 0.1020408163Y^{10} + 0.018551431Y^2 + b_{0,6}Y^{12} \\
&+ \left(0.0005311734742 + 0.0044738101Y^2 + 0.06303177Y^4 - 0.04060807Y^6 - 0.1275510204Y^8\right) X^4 + 0.01816845Y^6 - 0.023427738Y^8 \\
&+ \left(-0.002322407 - 0.5832391266Y^6 - 0.09992647621Y^4 - 0.0127896Y^2 - 0.9548104Y^8 - 1.285714286Y^{10}\right) X^2 - 9.296721604 \times 10^{-7}X^{12}, \\
q_3 =& 3.549504802 \times 10^{-8}X^{12} + \left(3.75233349 \times 10^{-7} - 8.9447521 \times 10^{-6}Y^2\right) X^{10} - 0.17710609Y^6 - 0.4236830911Y^8 - 0.05497827Y^4 \\
&+ \left(-5.976207065 \times 10^{-7} - 0.0001327514796Y^2 - 0.0001739257353Y^4\right) X^8 - 0.8017837258Y^{10} - 1.870828694Y^{12} - 0.01168191703Y^2 \\
&+ \left(4.377377643 \times 10^{-6} - 0.0002695595362Y^2 - 0.001729318740Y^4 + 0.003506342824Y^6\right) X^6 - 0.00001357167638 \\
&+ \left(-0.00003294675194 - 0.0004837467973Y^2 - 0.001043554412Y^4 + 0.07597076118Y^6 + 0.08590539918Y^8\right) X^4 \\
&+ \left(0.0001652692900 + 0.2963694530Y^6 + 0.02645090979Y^4 + 0.003497711008Y^2 + 0.3681659965Y^8 + 0.2672612419Y^{10}\right) X^2,
\end{aligned}
$$

$$
\begin{aligned}
p_4 =& -0.00002962184Y^2 - 3.8297467 \times 10^{-6}X^2 + 2.381933471 \times 10^{-6} + 0.08663759Y^8 - 7.207715497 \times 10^{-7}X^8 + 0.2120251263Y^{12} \\
&+ 0.26345327Y^{10} + 7.916018913 \times 10^{-8}X^{10} - 1.1317264X^2Y^{12} + 0.15039865X^4Y^{10} + 0.01616212X^6Y^8 + 0.000536729X^8Y^6 \\
&- 0.000030812X^{10}Y^4 - 2.750838239 \times 10^{-7}X^{12}Y^2 - 1.7005368X^2Y^{14} - 0.007820717X^4Y^{12} + 0.078315111X^6Y^{10} + 0.00426584X^8Y^8 \\
&- 0.0000459X^{10}Y^6 - 7.6568853 \times 10^{-6}X^{12}Y^4 - 6.038425402 \times 10^{-8}X^{14}Y^2 - 2.205089558 \times 10^{-8}X^{16}Y^4 - 1.512407105 \times 10^{-10}X^{18}Y^2 \\
&- 0.00003112X^4Y^2 - 0.033907392X^2Y^6 + 0.00095977X^4Y^4 + 0.0000306X^6Y^2 - 0.007609983X^2Y^4 - 0.00035455X^2Y^2 - 0.10447491X^2Y^8 \\
&+ 0.007629963X^4Y^6 + 0.0001072028X^6Y^4 - 4.453235195 \times 10^{-6}X^8Y^2 - 0.35665131X^2Y^{10} + 0.03136338026X^4Y^8 + 0.002147290911X^6Y^6 \\
&- 0.0000441753X^8Y^4 + 3.864095268 \times 10^{-8}X^{10}Y^2 - 1.841563786X^2Y^{16} - 0.1973784484X^4Y^{14} + 0.069645548X^6Y^{12} + 0.0103551X^8Y^{10} \\
&+ 0.0003097579140X^{10}Y^8 - 0.00001291474675X^{12}Y^6 - 7.368447413 \times 10^{-7}X^{14}Y^4 - 5.696733427 \times 10^{-9}X^{16}Y^2 - 1.666666667X^2Y^{18} \\
&- 0.23148145X^4Y^{16} + 0.0260631X^6Y^{14} + 0.006572936X^8Y^{12} + 0.000412792X^{10}Y^{10} + 3.822155235 \times 10^{-6}X^{12}Y^8 - 5.880238822 \times 10^{-7}X^{14}Y^6 \\
&+ 7.897557947 \times 10^{-6}X^4 - 0.000578944Y^4 + 3.275726908 \times 10^{-6}X^6 + 0.01580221Y^6 - 6.567857534 \times 10^{-9}X^{12} + Y^{20} + 0.4320987654Y^{18} \\
&+ 0.14895035Y^{14} + 8.697329 \times 10^{-10}X^{14} + 0.2120484682Y^{16} + 4.432971 \times 10^{-10}X^{16} + 7.188601670 \times 10^{-11}X^{18} + 2.520678508 \times 10^{-12}X^{20}, \\
q_4 =& 0.0005338465Y^2 - 0.000010277X^2 + 0.0000104657 + 2.974668188Y^8 - 1.1010522 \times 10^{-6}X^8 + 17.293968Y^{12} + 7.975337364Y^{10} \\
&+ 1.104326262 \times 10^{-7}X^{10} - 22.89759875X^2Y^{12} - 0.53298939X^4Y^{10} + 0.083871760X^6Y^8 + 0.008443891X^8Y^6 - 0.0000164785X^{10}Y^4 \\
&- 9.533224927 \times 10^{-7}X^{12}Y^2 - 30.54205821X^2Y^{14} - 6.8568943X^4Y^{12} + 0.2107788253X^6Y^{10} + 0.045167346X^8Y^8 + 0.00138430466X^{10}Y^6 \\
&- 0.0000119485X^{12}Y^4 - 3.202572626 \times 10^{-7}X^{14}Y^2 - 5.688410704 \times 10^{-8}X^{16}Y^4 - 1.264091268 \times 10^{-9}X^{18}Y^2 - 0.0009784105X^4Y^2 \\
&- 0.2852765843X^2Y^6 - 0.000457594119X^4Y^4 + 0.0002402136X^6Y^2 - 0.0390521167X^2Y^4 - 0.003038713X^2Y^2 - 1.302499848X^2Y^8 \\
&+ 0.01410899843X^4Y^6 + 0.000837180X^6Y^4 - 0.00002797447X^8Y^2 - 5.059446422X^2Y^{10} - 0.006682688713X^4Y^8 + 0.01851339X^6Y^6 \\
&+ 6.9110688 \times 10^{-6}X^8Y^4 + 2.250381 \times 10^{-6}X^{10}Y^2 - 25.02283122X^2Y^{16} - 9.344150034X^4Y^{14} - 0.40871367X^6Y^{12} + 0.05784709X^8Y^{10} \\
&+ 0.00514905X^{10}Y^8 + 0.00009641X^{12}Y^6 - 1.574214992 \times 10^{-6}X^{14}Y^4 - 3.9748647 \times 10^{-8}X^{16}Y^2 - 13.93023594X^2Y^{18} - 6.2686061X^4Y^{16} \\
&- 0.5159345X^6Y^{14} + 0.010350541X^8Y^{12} + 0.003450168X^{10}Y^{10} + 0.000169563X^{12}Y^8 + 2.075132225 \times 10^{-6}X^{14}Y^6 + 7.712339527 \times 10^{-6}X^4 \\
&+ 0.016657986Y^4 + 5.169052059 \times 10^{-6}X^6 + 0.66907988Y^6 - 7.227676311 \times 10^{-9}X^{12} + 75.22327Y^{20} + 57.26874773Y^{18} + 30.62462663Y^{14}+ \\
&2.538500949 \times 10^{-10}X^{14} + 43.92610406Y^{16} + 2.3365747 \times 10^{-10}X^{16} + 5.9823249 \times 10^{-11}X^{18} + 2.340909755 \times 10^{-12}X^{20}.
\end{aligned}
$$

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–242, pp. 554–563

Lump Wave Dynamics and Interaction Analysis for an Extended $(2+1)$-Dimensional Kadomtsev–Petviashvili Equation

Majid Madadi[a,*], Esmaeel Asadi[a,b]

[a] Department of Mathematics, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137–66731, Iran
[b] School of Mathematics, Institute for Research in Fundamental Sciences (IPM), P.O. Box 19395-5746, Tehran, Iran

Article Info

Keywords:

Hirota Bilinear;
Soliton Solution;
Lump Wave;
Breather Wave;

2020 MSC:
37K40;
70Kxx;

Abstract

In this paper, we introduce novel constrained conditions into $N-$soliton solutions for a $(2+1)$-dimensional extended Kadomtsev–Petviashvili equation. This integration results in the derivation of lump waves. To investigate the interaction between higher-order lumps and soliton waves, as well as breather waves, we employ the long wave limit method. We analyze the trajectory equations governing the motion before and after the collision of lumps and other waves, and identify conditions under which the lump wave avoids collision with other waves. Several figures are included to illustrate the physical behavior of these solutions.

1. Introduction

The Hirota bilinear method [1] is a powerful technique for solving nonlinear evolution equations. It has gained popularity among scholars due to its directness and simplicity in constructing multiple soliton solutions for nonlinear partial differential equations [2–4].
Lump waves, which are rational function waves localized in all directions in space and time, have attracted significant attention. Various methods such as inverse scattering transformation [5], Grammian determinant method [6], Darboux transformation [7], and the long wave limit method [8] have been developed to study lump waves.
The interaction between lumps and other nonlinear waves, including soliton and breather [9–11], is an active research area, particularly focusing on elastic and inelastic collisions. However, there is still much to understand about the movement of lumps before and after collisions, phase shifts in lumps, and the trajectory equations governing their motion. Further exploration is needed. Additionally, it is worth investigating other forms of interaction between lump waves and other waves, such as scenarios where they never collide or always collide and never separate.
Reference [12] has explored an extended (2 + 1)-dimensional Kadomtsev–Petviashvili (eKP) equation as follows:

$$\gamma_1 U_{xt} + \gamma_2 U_{xx}^2 + \gamma_3 U_{xxxx} + \gamma_4 U_{yy} + \alpha_1 U_{xx} = 0, \tag{1}$$

*Talker
Email addresses: mmadadi@iasbs.ac.ir. (Majid Madadi), easadi@iasbs.ac.ir. (Esmaeel Asadi)

where $U = U(x, y, t)$ is an unknown differentiable function, and γ_i, $i = 1, \ldots, 4$, are real arbitrary parameters. The eKP equation (1) finds applications in nonlinear wave phenomena, plasma physics, optical fiber communications, fluid dynamics, and mathematical physics.
It is noteworthy that Eq. (1) simplifies to the KP equation [1] when $\alpha_1 = 0$. Ref. [12] explores the integrability and soliton solutions for Eq. (1). The Painlevé test is utilized to assess their integrability, and it is found that the equation successfully passes the Painlevé test.
The focus of this study lies in obtaining exact solutions for Eq.(1) using the Hirota bilinear technique. This approach yields Soliton and Lump solutions, as well as hybrid solutions containing solito-lump and breather-lump configurations. Additionally, the research introduces a novel method for tracking the trajectory of a lump wave before and after collisions with other wave types. Specifically, the study establishes and elucidates the conditions under which the entire process ensures that a lump wave either avoids collisions with other waves or, if collisions occur, remains unchanged in its state.
The paper is organized as follows: Section 2 delves into the Hirota bilinear representation and soliton solution of Eq. (1). In Section 3, the focus shifts to the N-lump solution and their dynamics. Moving on to Section 4, we examine hybrid solutions of lump waves and soliton and breather waves, along with the trajectory of the lump wave before and after interaction. Finally, Section 5 presents the conclusions.

2. $N-$soliton solution

In this section, we aim to construct an $N-$soliton solution to Eq. (1). To achieve this, we first bilinearize the equation using the logarithmic transformation $U = \frac{6\gamma_3}{\gamma_2} \ln(f)_{xx}$, which can be expressed as

$$\left(\gamma_1 D_x D_t + \gamma_3 D_x^4 + \gamma_4 D_y^2 + \alpha_1 D_x^2\right) f.f = 0. \tag{2}$$

Then, we choose the function f in the form

$$f = \sum_{\mu\in\{0,1\}^N} \exp\left(\sum_{i=1}^{N} \mu_i \chi_i + \sum_{i<j}^{N} \mu_i \mu_j \ln \Upsilon_{ij}\right), \tag{3}$$

where $\sum_{\mu\in\{0,1\}^N}$ represents the summation over all possible combinations of $\mu_j, \mu_s = 0, 1$, $j, s = 1, 2, \ldots, N$, and the wave variables are

$$\chi_i = k_i x + p_i y + w_i t + \phi_i. \tag{4}$$

As a result, the dispersion relation and the phase shifts are listed below, respectively.

$$w_i = -\frac{\gamma_3 k_i^4 + k_i^2 \alpha_1 + \gamma_4 p_i^2}{k_i \gamma_1}, \tag{5}$$

$$\Upsilon_{i,j} = \frac{3k_i^2 k_j^4 \gamma_3 - 6k_i^3 k_j^3 \gamma_3 + \left(3k_i^4 \gamma_3 - p_i^2 \gamma_4\right) k_j^2 + 2k_i k_j p_i p_j \gamma_4 - k_i^2 p_j^2 \gamma_4}{3k_i^2 k_j^4 \gamma_3 + 6k_i^3 k_j^3 \gamma_3 + \left(3k_i^4 \gamma_3 - p_i^2 \gamma_4\right) k_j^2 + 2k_i k_j p_i p_j \gamma_4 - k_i^2 p_j^2 \gamma_4}. \tag{6}$$

By incorporating formulas (5) and (6) into the function f and then applying it to the logarithmic transformation, an $N-$soliton solution to Eq. (1) can be derived.

3. Lump waves

The objective of this section is to offer a comprehensive explanation of how the asymptotic behavior of the soliton solution (3) results in the formation of lump waves. By meticulously choosing appropriate values for these parameters, one can create wave functions that demonstrate the desired lump-like behavior. These parameters significantly influence the characteristics of the resulting lump solution, including its amplitude, velocity, and position.

Theorem 3.1. *To obtain Nth-order lump solutions, we apply the long wave limit while considering the following conditions in N-soliton solution* (3).

$$N = 2m, \quad k_i = K_i\epsilon, \quad p_i = P_i\epsilon, \quad e^{\phi_i} = -1, \quad \epsilon \to 0,$$
$$K_1 = K_2^*, \ldots, K_{2m-1} = K_{2m}^* \quad P_1 = P_2^*, \ldots, P_{2m-1} = P_{2m}^*. \tag{7}$$

The notation $\epsilon \to 0$ indicates that, after substituting the revised values of the parameters k_i and p_i into expression (3), an evaluation is performed to determine the limiting behavior of the resulting function.
Put more simply, the comprehensive solutions for the N-th order lump can be expressed as the subsequent form.

$$U_L^{(N)} = \frac{6\gamma_3}{\gamma_2}\ln(f_N)_{xx}. \tag{8}$$

Where the function f_N is represented by the following expression.

$$f_N = \prod_{s=1}^{N}\theta_s + \frac{1}{2}\sum_{j,s}^{N} B_{js}\prod_{p\neq j,s}^{N}\theta_p + \ldots + \frac{1}{m!2^m}\sum_{l,s,..,m,n}^{N}\overbrace{B_{ls}B_{jk}\ldots B_{mn}}^{m}\prod_{q\neq l,s,..,m,n}^{N}\theta_q + \ldots, \tag{9}$$

where

$$\theta_i = -\frac{\left(K_i^2\alpha_1 + P_i^2\gamma_4\right)t}{K_i\gamma_1} + xK_i + yP_i, \quad B_{ij} = \frac{12K_i^3K_j^3\gamma_3}{\gamma_4\left(K_iP_j - K_jP_i\right)^2}. \tag{10}$$

The m-th lump wave in the solution (8) *is characterized by the following properties:*

- *The trajectory of the m-th lump wave is given by*

$$y = \frac{\gamma_4\left(K_{2m-1}P_{2m} + K_{2m}P_{2m-1}\right)}{K_{2m-1}K_{2m}\alpha_1 - P_{2m-1}P_{2m}\gamma_4}x, \tag{11}$$

- *The amplitude formula is*

$$A_m^{[N]} = \frac{\gamma_4\left(K_{2m-1}P_{2m} - K_{2m}P_{2m-1}\right)^2}{\gamma_2K_{2m-1}^2K_{2m}^2}, \tag{12}$$

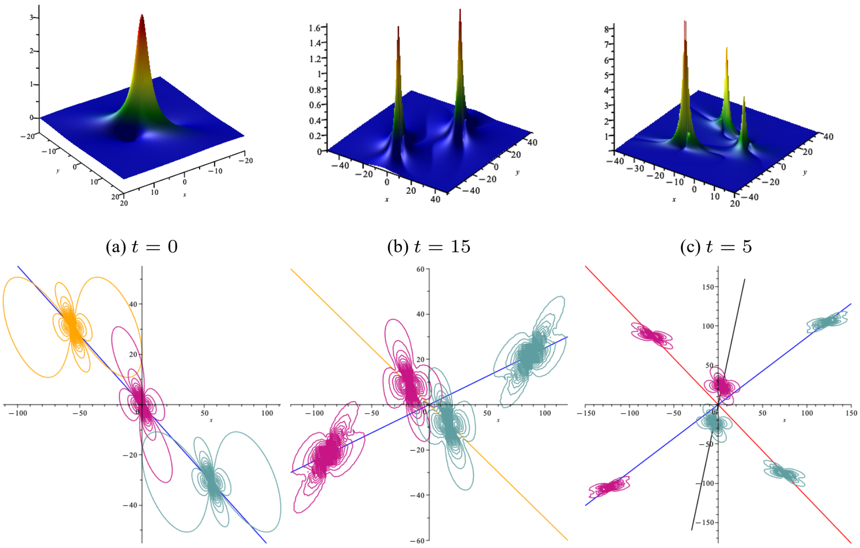
- *Velocity of the lump wave are determined by*

$$V_m^{[N]} = \sqrt{\frac{\left(K_{2m-1}K_{2m}\alpha_1 - P_{2m-1}P_{2m}\gamma_4\right)^2}{K_{2m}^2\gamma_1^2K_{2m-1}^2} + \frac{\gamma_4^2\left(K_{2m-1}P_{2m} + K_{2m}P_{2m-1}\right)^2}{K_{2m}^2\gamma_1^2K_{2m-1}^2}}. \tag{13}$$

To achieve a single lump, one can substitute $N = 2$ into the formula (9). This substitution results in the expression of U_1 as shown below:

$$U_1 = \frac{6\gamma_3}{\gamma_2}\ln(\theta_1\theta_2 + B_{1,2})_{xx}, \tag{14}$$

where, θ_i and $B_{i,j}$ are represented in (10). Figure 1 (a) illustrates the dynamics of the one-lump solution (14). With specific parameter selections, the amplitude and velocity of the wave are $A_1^{[1]} = 3.43.408284$ and $V_1^{[1]} = 0.6369421$, respectively. Furthermore, Figure 1 (d) displays the trajectory of the wave over time. For $N = 4, 6$, the two and three-lump solutions are presented in Figure 1 (b,c). As observed in the figure, the two-lump solution exhibits two separate peaks that move away from each other as time increases. The amplitude and velocity of the first wave are $V_1^{[2]} = 5.4589$ and $A_1^{[2]} = 1.583$, respectively, while for the second wave, these values are $V_2^{[2]} = 1.36$ and $A_2^{[2]} = 1.498$, respectively. Moreover, in the three-lump solution, there are three separate peaks with the following amplitudes and velocities: $V_1^{[3]} = 0.66899$, $V_2^{[3]} = 4.60977$, $V_3^{[3]} = 3.2811$ and $A_1^{[3]} = 3.07$, $A_2^{[3]} = 6.25$, $A_3^{[3]} = 8.134$. Additionally, the trajectory of each peak in the aforementioned solution is clearly represented in Figure 1 (e,f).



(a) $t = 0$ (b) $t = 15$ (c) $t = 5$

(d) $t = -100, 0, 100$ (e) $t = -20, 20$ (f) $t = -35, 35$

Fig. 1. **Panels (a, d):** One-lump soution with, $K_1 = K_2^* = 2 + 3\,\mathrm{I}$, $P_1 = P_2^* = 4$. Trajectory of lump $y = -\frac{16}{29}x$ (blue line). **Panels (b, e):** Two-lump solution with, $K_1 = K_2^* = 2 + 4\,\mathrm{I}$, $K_3 = K_4^* = 1 + 4\,\mathrm{I}$, $P_1 = P_2^* = 3$, $P_3 = P_4^* = 3$. Trajectory of the first lump $y = -\frac{24}{13}x$ (orange line) and the second lump $y = \frac{15}{8}x$ (blue line). **Panels (c, f):** Three-lump solution with, $K_1 = K_2^* = 1.5 + 4.0\,\mathrm{I}$, $K_3 = K_4^* = -1 + \mathrm{I}$, $K_5 = K_6^* = 2 + 2\,\mathrm{I}$, $P_1 = P_2^* = 4$, $P_3 = P_4^* = 3$, $P_5 = P_6^* = 5$. Trajectory of the first lump $y = \frac{6}{7}x$ (blue line), the second lump $y = -\frac{20}{17}x$ (red line) and the third lump $y = \frac{19}{5}x$ (black line). For the same selection of $\alpha_1 = 1, \gamma_1 = 1, \gamma_2 = 1, \gamma_3 = -1, \gamma_4 = 1$.

4. Interaction between lump and soliton waves

In this section, we delve into the development of a unique hybrid solution that merges both lump and other wave forms. This hybrid solution is attained through the utilization of the long wave limit method [8] based on the N-soliton solution (3). Employing this method allows us to derive a range of semi-rational solutions that incorporate blends of lumps and soliton lines, as well as combenation of lumps and breathers.

Proposition 4.1. *To obtain a hybrid solution comprising $\mathcal{L}$ lump waves and $\mathcal{S}$ soliton waves, we can take the long wave limit with the following restrictions in Eq.* (3).

$$\begin{aligned} &N = 2\mathcal{L} + \mathcal{S}, \quad m = 1, 2, ..., 2\mathcal{L}, \quad k_m = K_m\epsilon, \quad p_m = P_m\epsilon, \quad \phi_m = \pi\boldsymbol{i}, \quad \epsilon \to 0, \\ &K_1 = K_2^*,\ K_3 = K_4^*, \dots, K_{2\mathcal{L}-1} = K_{2\mathcal{L}}^*, \quad P_1 = P_2^*,\ P_3 = P_4^*, \dots, P_{2\mathcal{L}-1} = P_{2\mathcal{L}}^*. \end{aligned} \tag{15}$$

Studying the trajectory equations of lump waves and soliton lines before and after collision [10], enhances our understanding of their complex interaction. In the following, we discuss this issue further.

Theorem 4.2. *For the mixed solution comprising $\mathcal{L}$-lump waves and $\mathcal{S}$-soliton lines, under the condition $\lambda_3, \lambda_4, \dots, \lambda_{2+\mathcal{S}} \neq 0$, and based on conditions* (15)*, the trajectory equations of an arbitrary lump wave are provided below:*

$$\begin{aligned} (x_b, y_b) &= \left(X + \sum_{s=2\mathcal{L}+1}^{N} h_b(\lambda_s)\kappa_s,\ Y + \sum_{s=2\mathcal{L}+1}^{N} h_b(\lambda_s)\vartheta_s \right), \\ (x_a, y_a) &= \left(X + \sum_{s=2\mathcal{L}+1}^{N} h_a(\lambda_s)\kappa_s,\ Y + \sum_{s=2\mathcal{L}+1}^{N} h_a(\lambda_s)\vartheta_s \right), \end{aligned} \tag{16}$$

where

$$(X,Y)=\left(\frac{(K_{2m-1}K_{2m}\alpha_1-P_{2m-1}P_{2m}\gamma_4)}{K_{2m}\gamma_1K_{2m-1}}t,\ \frac{\gamma_4\,(K_{2m-1}P_{2m}+K_{2m}P_{2m-1})}{K_{2m}\gamma_1K_{2m-1}}t\right), \tag{17}$$

$$\kappa_s=\frac{-B_{2m-1,s}P_{2m}+B_{2m,s}P_{2m-1}}{K_{2m-1}P_{2m}-K_{2m}P_{2m-1}},\ \vartheta_s=\frac{B_{2m-1,s}K_{2m}-B_{2m,s}K_{2m-1}}{K_{2m-1}P_{2m}-K_{2m}P_{2m-1}}. \tag{18}$$

$$\lambda_s=\frac{-K_{2m-1}K_{2m}k_s^4\gamma_3-P_{2m-1}P_{2m}k_s^2\gamma_4+p_s\gamma_4\,(K_{2m-1}P_{2m}+K_{2m}P_{2m-1})\,k_s-K_{2m-1}K_{2m}p_s^2\gamma_4}{K_{2m}\gamma_1K_{2m-1}k_s}, \tag{19}$$

$$h_b(x)=\begin{cases}1, & x<0\\ 0, & x\geq 0\end{cases},\quad h_a(x)=\begin{cases}0, & x\leq 0\\ 1, & x>0\end{cases}, \tag{20}$$

For $1\leq i<j\leq 2\mathcal{L}$, $B_{i,j}$ *are presented in (10) and for* $1\leq i\leq 2\mathcal{L}$ *and* $j>2\mathcal{L}$ *it is as fallows:*

$$B_{i,j}=\frac{12K_i^3k_j^3\gamma_3}{3K_i^2k_j^4\gamma_3-K_i^2p_j^2\gamma_4+2K_iP_ik_jp_j\gamma_4-P_i^2k_j^2\gamma_4}. \tag{21}$$

The change in the phase of the lump wave before and after the collision can be expressed as

$$\Delta_b=\sum_{s=2\mathcal{L}+1}^{N} sign(\lambda_s)\Delta_{bs}, \tag{22}$$

where

$$\Delta_{bs}=\frac{-B_{2m,s}K_{2m-1}^2K_{2m}\alpha_1+B_{2m-1,s}\left(K_{2m}^2\alpha_1+P_{2m}^2\gamma_4\right)K_{2m-1}-B_{2m,s}K_2P_{2m-1}^2\gamma_4}{(K_{2m-1}P_{2m}-K_{2m}P_1)\,(K_{2m-1}K_{2m}\alpha_1-P_{2m-1}P_{2m}\gamma_4)}. \tag{23}$$

However, the amplitude and velocity of the peak do not change before and after the collision, and they are represented by formula (12) *and* (13)*, respectively.*

Proof. To prove this theorem, we consider solutions that consist of a lump wave and a soliton line. The proof for the other situations remains the same, and we will omit their proof for brevity.
First, let us consider a mixed solution consisting of a lump wave and a soliton wave based on the conditions (15) as follows:

$$U_{LS}^{(1)}=\frac{6\gamma_3}{\gamma_2}\ln(f_{LS}^{(1)})_{xx}. \tag{24}$$

with

$$f_{LS}^{(1)}=\theta_1\theta_2+B_{1,2}+(B_{1,3}B_{2,3}+\theta_2B_{1,3}+\theta_1B_{2,3}+\theta_1\theta_2+B_{1,2})\,\mathrm{e}^{\chi_3}. \tag{25}$$

Regarding the lump wave's path in (11), we assume it follows a straight line before and after collision. Hence, the function $f_{LS}^{(1)}$ is bound by the subsequent conditions.

$$(x,y)=\left(\frac{(K_1K_2\alpha_1-P_1P_2\gamma_4)}{K_2\gamma_1K_1}t+c_1,\ \frac{\gamma_4\,(K_1P_2+K_2P_1)}{K_2\gamma_1K_1}t+c_2\right). \tag{26}$$

Substituting (26) into (25) yields

$$f_{LS}^{(1)}=\mathrm{e}^{t\,\lambda_3+\beta_3}\,(B_{1,3}B_{2,3}+B_{1,3}\zeta_2+B_{2,3}\zeta_1+\zeta_1\zeta_2+B_{1,2})+B_{1,2}+\zeta_1\zeta_2, \tag{27}$$

where

$$\zeta_1 = K_1 c_1 + P_1 c_2, \;\; \zeta_2 = K_2 c_1 + P_2 c_2, \;\; \beta_3 = c_1 k_3 + c_2 p_3 + \phi_3.$$

When $\lim_{t\to\pm\infty} f_{LS}^{(1)}$ we can derive the following approximate expressions.
Case I: For $\lambda_3 > 0$.

$$f_b = B_{1,2} + \zeta_1\zeta_2, \quad f_a = B_{1,3}B_{2,3} + B_{1,3}\zeta_2 + B_{2,3}\zeta_1 + \zeta_1\zeta_2 + B_{1,2}. \tag{28}$$

By substituting the values of c_1 and c_2 from (26) into (28), we obtain the following expressions:

$$f_b = \theta_1\theta_2 + B_{1,2}, \quad f_a = B_{1,3}B_{2,3} + B_{1,3}\theta_2 + B_{2,3}\theta_1 + \theta_1\theta_2 + B_{1,2}. \tag{29}$$

These expressions feature the functions f_b and f_a, representing the states of the lump peak and soliton before and after the collision, respectively. It is crucial to highlight that these functions adhere to the bilinear Eq. (2).
Substituting (29) into (24) and equating the derivatives of the solutions with respect to x and y to zero allows us to ascertain the trajectories of the peak before and after the collision:

$$\begin{aligned}
(x_b, y_b) &= \left(\frac{(K_1K_2\alpha_1 - P_1P_2\gamma_4)}{K_2\gamma_1K_1}t, \frac{\gamma_4\,(K_1P_2 + K_2P_1)}{K_2\gamma_1K_1}t\right),\\
(x_a, y_a) &= \left(\frac{(K_1K_2\alpha_1 - P_1P_2\gamma_4)}{K_2\gamma_1K_1}t + \frac{-B_{1,3}P_2 + B_{2,3}P_1}{K_1P_2 - K_2P_1}, \frac{\gamma_4\,(K_1P_2 + K_2P_1)}{K_2\gamma_1K_1}t + \frac{B_{1,3}K_2 - B_{2,3}K_1}{K_1P_2 - K_2P_1}\right).
\end{aligned} \tag{30}$$

In this scenario, the validity of expressions (35) is confirmed. Additionally, the phase change (22) can be easily determined by comparing the peak's trajectories before and after the collision, without involving the time parameter. When substituting the values of x_b, y_b, x_a, and y_a into (24), it becomes apparent that the peak's amplitude remains unchanged following the collision.
Case II: Let us consider the scenario where $\lambda_3 < 0$. The proof follows a similar approach to the previous one, but with a notable distinction: $\lim_{t\to\pm\infty} f_{LS}^{(1)}$ yields the following expressions:

$$f_b = B_{1,3}B_{2,3} + B_{1,3}\zeta_2 + B_{2,3}\zeta_1 + \zeta_1\zeta_2 + B_{1,2}, \quad f_a = B_{1,2} + \zeta_1\zeta_2. \tag{31}$$

In this particular case, the trajectory after the collision precisely mirrors the trajectory before the collision in the previous case. Conversely, the trajectory before the collision in this case aligns with the trajectory after the collision in Case I. □

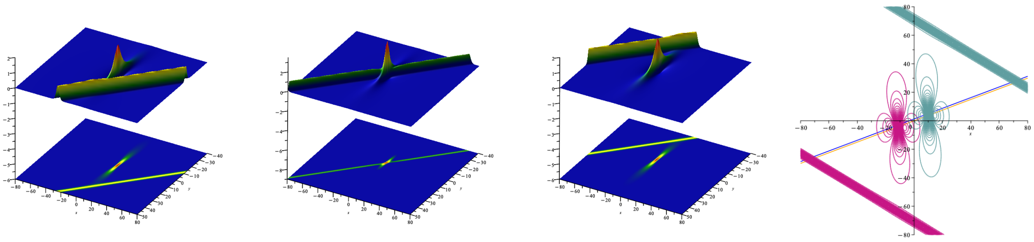
Corollary 4.3. *The condition for the avoidance of collision or the preservation of wave states during the interaction between a lump wave and a soliton wave arises when $\lambda_s = 0$. Simply put, this condition is met when the velocity of the lump wave matches the velocity of the soliton line, given by:*

$$V_{\mathcal{L}} = V_{\mathcal{S}}, \tag{32}$$

where $V_{\mathcal{S}} = \left[-\frac{w_ik_i}{k_i^2+p_i^2}, -\frac{w_ip_i}{k_i^2+p_i^2}\right]$, and w_i, $V_{\mathcal{L}}$, and λ_s are defined in (5), (13), and (19), respectively.

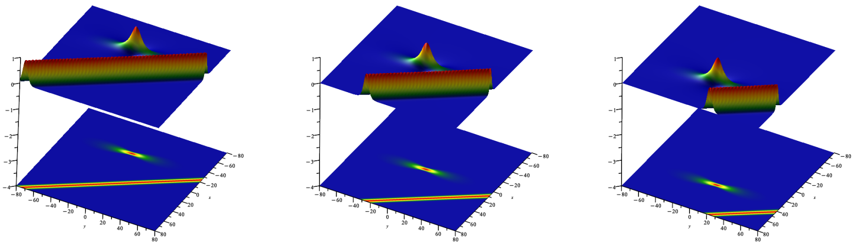
We investigate three scenarios of wave collisions for clearer insight. Firstly, we analyze the collision between a soliton wave and a lump wave. Secondly, we explore the collision of two soliton waves and a lump wave. Lastly, we examine the collision between one soliton wave and two lump waves.

Example 4.4. The interaction between a lump wave and a soliton wave solution (24), is illustrated in Figure 5. Upon evaluating Eq. (19), we ascertain $\lambda_3 = -4.0771635 < 0$. Consequently, we deduce $h_b(\lambda_3) = 1$ and $h_a(\lambda_3) = 0$. Consequently, the lump wave initially follows the trajectory $y = 0.3741x - 0.7167$ before colliding with the soliton wave. However, post-collision, the lump wave alters its course, shifting to $y = 0.3741x$. Additionally, the change in phase, denoted by $\Delta_{b3} = 0.7167$. Notably, the lump wave maintains consistent velocity and amplitude pre and post collision, with $V_1^{[1]} = 0.5386$ and $A_1^{[1]} = 1.964$. Moreover, based on Corollary 4.3, specific conditions are established to ensure that the interaction between lump and soliton waves never results in a collision. This phenomenon is illustrated in Figure 3.



(a) $t = -10$ (b) $t = 0$ (c) $t = 10$ (d)

Fig. 2. Superposition of a lump and a soliton wave with $K_1 = K_2^* = \frac{2}{7} - 3\,\mathrm{I}, P_1 = P_2^* = \frac{3}{2}, \alpha_1 = 1, k_3 = \frac{4}{5}, p_3 = \frac{4}{3}, \phi_3 = 0, \gamma_1 = 1, \gamma_2 = -1, \gamma_3 = -1, \gamma_4 = 2$. **Panel c:** Trajectory of lump before the interaction $y = 0.3741x - 0.7167$ (orange line) and after the interaction $y = 0.3741x$ (blue line) for $t = -20$ (crimson) and $t = 20$ (cadet blue).



(a) $t = -40$ (b) $t = 0$ (c) $t = 40$

Fig. 3. In the superposition of a lump and a soliton wave these waves nevere colid if: $K_1 = K_2^* = \frac{2}{7} - 3\,\mathrm{I}, P_1 = P_2^* = \frac{3}{2}, \alpha_1 = 1, k_3 = \frac{4}{5}, p_3 = 0.5402102468\phi_3 = -50, \gamma_1 = 1, \gamma_2 = -1, \gamma_3 = -1, \gamma_4 = 1$.

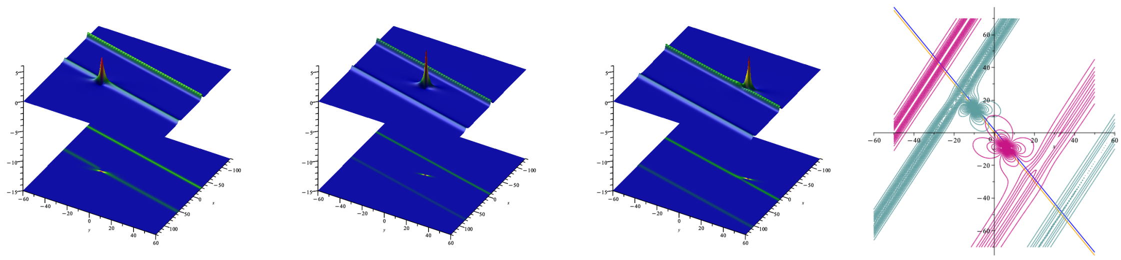
Example 4.5. The solution represented by $U_{LS}^{(2)} = \frac{6\gamma_3}{\gamma_2}\ln(f_{LS}^{(2)})_{xx}$, where

$$\begin{aligned} f_{LS}^{(2)} &= \mathrm{e}^{\chi_3}\left(B_{1,3}B_{2,3} + B_{1,3}\theta_2 + B_{2,3}\theta_1 + \theta_1\theta_2 + B_{1,2}\right) + \mathrm{e}^{\chi_4}\left(B_{1,4}B_{2,4} + B_{1,4}\varpi_2 + B_{2,4}\theta_1 + \theta_1\theta_2 + B_{1,2}\right) \\ &+ \Upsilon_{3,4}\mathrm{e}^{\chi_3+\chi_4}(B_{1,3}B_{2,3} + B_{1,3}B_{2,4} + B_{1,3}\theta_2 + B_{1,4}B_{2,3} + B_{1,4}B_{2,4} + B_{1,4}\theta_2 + B_{2,3}\theta_1 + B_{2,4}\theta_1 + \theta_1\theta_2 + B_{1,2}) \\ &+ \theta_1\theta_2 + B_{1,2}, \end{aligned} \tag{33}$$

combines one lump wave and two solitons. This solution is shown in Figure 4 for various times. After computation, we confirm that $\lambda_3 = -1.834791667 < 0$ and $\lambda_4 = -1.37944 < 0$. Following Theorem 4.2, the trajectories of the lump wave before and after the interaction are visually illustrated in Figure 4 (d), with the phase shift indicated by $\Delta_{b_3} = 0.3986457101$. Notably, the velocity and amplitude of the lump wave remain constant before and after the collision, with values of $V_1^{[2]} = 1.4422$ and $A_1^{[2]} = 6.48$, respectively.

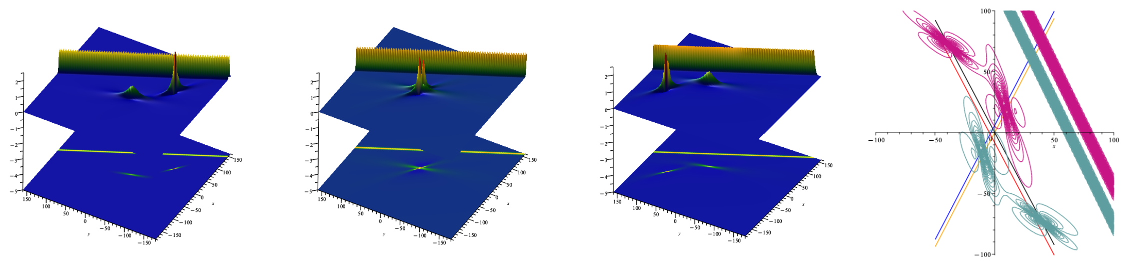
Example 4.6. To achieve a composite solution involving two lump waves and a line soliton we set $N = 6, \mathcal{S} = 1$, and $\mathcal{L} = 2$ according to conditions (15). This results in the emergence of a novel solution denoted as $U_{LS}^{(3)}$, as depicted in Figure 5. For the first lump wave, we determine $\lambda_5 = \frac{-7}{10}$, while for the second lump wave, $\lambda_5 = \frac{11}{2}$. As per Theorem 4.2, the trajectories before and after the collision of the first and second lump waves are delineated in Figure 5 (c). The velocity and amplitude of the first lump wave are quantified as $V_1^{[3]} = 5.4589$ and $A_1^{[3]} = 2.88$, respectively. Correspondingly, for the second lump wave, we compute $A_2^{[3]} = 0.72$ and $V_2^{[3]} = 1.36$.

In this part, we present a method for examining the interaction between lump waves and breather waves. To obtain a combined solution comprising $\mathcal{L}$ lump waves and $\mathcal{B}$ breather waves for Eq. (1), we start by setting $N = 2(\mathcal{B} + \mathcal{L})$



(a) $t = -20$ (b) $t = 0$ (c) $t = 20$ (d)

Fig. 4. Superposition of a lump and two soliton waves with $K_1 = K_2^* = \frac{1}{2} - \frac{3\mathrm{I}}{2}, P_1 = P_2^* = \frac{3}{2}, \alpha_1 = 1, k_3 = \frac{3}{4}, k_4 = \frac{1}{2}, p_3 = -\frac{2}{5}, \phi_3 = 30, \phi_4 = -20, \gamma_1 = 1, \gamma_2 = -1, \gamma_3 = -1, \gamma_4 = 2$. **Panel c:** Trajectory of the lump before the interaction $y = -1.5x + 0.3986$ (orange line) and after the interaction $y = -1.5x$ (blue line). for $t = -10$ (crimson) and $t = 10$ (cadet blue).



(a) $t = -20$ (b) $t = 0$ (c) $t = 20$ (d)

Fig. 5. Superposition of two lumps and a soliton wave with $K_1 = K_2^* = 2 - \mathrm{I}, K_3 = K_4^* = 1 - \mathrm{I}, P_1 = P_2^* = 3, P_3 = \frac{3}{5}, P_4 = \frac{3}{5}, \alpha_1 = 1, k_5 = 1, p_5 = \frac{1}{2}, \phi_5 = -70, \gamma_1 = 1, \gamma_2 = -1, \gamma_3 = -1, \gamma_4 = 2$. **Panel c:** Trajectory of the first lump before the interaction $y = 1.875x + 5.9343$ (orange line) and after the interaction $y = 1.875x$ (blue line). Trajectory of the second lump before the interaction $y = -1.8461x - 8.405$ (black line) and after the interaction $y = -1.8461x$ (red line) for $t = -15$ (crimson) and $t = 15$ (cadet blue).

and then impose specific constraints on the N-soliton solution (3) as follows:

$$\begin{aligned}
&1 \le m \le 2\mathcal{L}, \quad k_{2m-1} = k_{2m}^* = K_{2m-1}\epsilon, \quad p_{2m-1} = p_{2m}^* = P_{2m-1}\epsilon, \quad \phi_{2m-1} = \phi_{2m} = \pi\mathbf{i}, \quad \epsilon \to 0,\\
&k_{2\mathcal{L}+1} = k_{2\mathcal{L}+2}^*, \dots, k_{2\mathcal{L}+2\mathcal{B}-1} = k_{2\mathcal{L}+2\mathcal{B}}^*, \quad p_{2\mathcal{L}+1} = p_{2\mathcal{L}+2}^*, \dots, p_{2\mathcal{L}+2\mathcal{B}-1} = q_{2\mathcal{L}+2\mathcal{B}}^*,\\
&\phi_{2\mathcal{L}+1} = \phi_{2\mathcal{L}+2}^*, \dots, \phi_{2\mathcal{L}+2\mathcal{B}-1} = \phi_{2\mathcal{L}+2\mathcal{B}}^*.
\end{aligned} \tag{34}$$

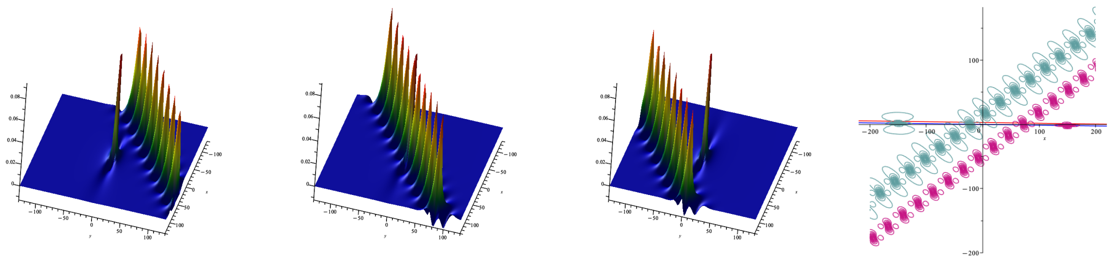
In a similar manner as previously, we can formulate a theorem delineating the path followed by the lump wave before and after its interaction with the breather wave.

Proposition 4.7. *The equations governing the trajectory of a lump wave before and after colliding with breather waves for $\lambda_s \neq 0$ are listed as follows:*

$$\begin{aligned}
(x_b, y_b) &= \left(X + \sum_{s=2\mathcal{L}+1}^{N} h_b(Re(\lambda_s))\kappa_s,\ Y + \sum_{s=2\mathcal{L}+1}^{N} h_b(Re(\lambda_s))\vartheta_s \right),\\
(x_a, y_a) &= \left(X + \sum_{s=2\mathcal{L}+1}^{N} h_a(Re(\lambda)_s)\kappa_s,\ Y + \sum_{s=2\mathcal{L}+1}^{N} h_a(Re(\lambda_s))\vartheta_s \right),
\end{aligned} \tag{35}$$

where X, Y, κ_s, ϑ_s, λ_s, h_b and h_a are given by Eqs. (17)-(20) and $Re(\lambda_s)$ denotes the real part of λ_s.

Example 4.8. For $N = 4$, which corresponds to $\mathcal{L} = 1$ and $\mathcal{B} = 1$ according to conditions (34), a hybrid solution comprising a breather wave and a lump wave can be derived. To visually illustrate the collision between these waves, Figure 6 depicts the physical behavior of this interaction.



(a) $t = -10$ (b) $t = 0$ (c) $t = 10$ (d)

Fig. 6. Superposition of a lump and a breather wave with $K_1 = K_2^* = \frac{3}{7} + \frac{1}{5}, P_1 = P_2^* = 1 - 2\,\mathrm{I}, \alpha_1 = 1, k_3 = k_4^* = \frac{1}{12} - \frac{1}{8}, p_3 = p_4^* = -\frac{1}{8} - \frac{1}{6}, \gamma_1 = 1, \gamma_2 = 3, \gamma_3 = 1, \gamma_4 = 1, \phi_3 = \phi_4^* = 0$. **Panel c:** Trajectory of the lump before the interaction $y = -0.01196x - 0.01509$ (red line) and after the interaction $y = -0.01196x$ (blue line). for $t = -7$ (crimson) and $t = 7$ (cadet blue),

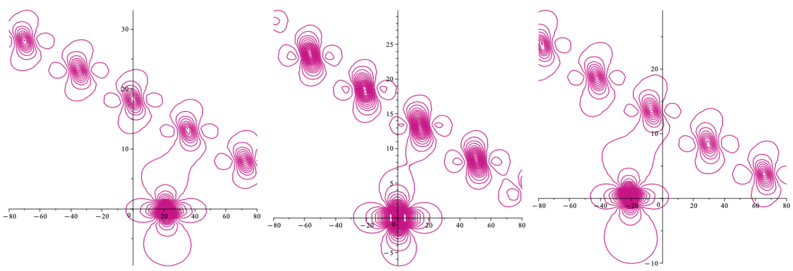
The collision behavior observed in the superposition of the lump and soliton, as explored in Corollary 4.3, similarly occurs here. That is, the lump wave and the breather wave either do not collide, or if they do, they remain in the same state.

Corollary 4.9. *When the condition $Re(\lambda_s) = 0$ is satisfied, lump and breather waves either do not collide or remain in a collided state. This condition signifies that the velocity of the lump wave equals the velocity of the breather wave, expressed as*

$$V_{\mathcal{L}} = V_{\mathcal{B}}, \tag{36}$$

where $V_{\mathcal{B}} = \left[-\frac{Re(w_i)Re(k_i)}{(Re(k_i))^2+(Re(p_i)^2}, -\frac{Re(w_i)Re(p_i)}{(Re(p_i))^2+(Re(p_i))^2}\right]$, and w_i and $V_{\mathcal{L}}$ are defined in (5) and (13).

The collision between the lump wave and the breather wave, illustrated in Figure 7, occurs in such a way that the two waves pass through each other without any interference or interaction. This behavior is achieved through carefully chosen parameters that meet the conditions outlined in Corollary 4.9.



(a) $t = -1$ (b) $t = 0$ (c) $t = 1$

Fig. 7. In the superposition of a lump and a breather wave these waves nevere colid if: $K_1 = K_2^* = \frac{3}{7} + \frac{1}{5}, P_1 = P_2^* = 1 - 2\,\mathrm{I}, \alpha_1 = 1, k_3 = k_4^* = \frac{1}{12} - \frac{1}{8}, p_3 = p_4^* = 0.6014984038 + 0.3775372501\,\mathrm{I}, \phi_3 = \phi_4^* = -10, \gamma_1 = 1, \gamma_2 = 3, \gamma_3 = 1, \gamma_4 = 1.$

5. Conclusion

In this study, we obtained $N-$soliton solution for Eq. (1) to investigate the dynamics of lump waves. By leveraging the asymptotic behavior of soliton solutions and employing the long wave limit method, we successfully derived multiple lump solutions. We also examined the interactions between lump waves and other wave types, including soliton and breather waves. Notably, we calculated the trajectory of the peak before and after each collision, and identified conditions under which the lump wave avoids collision with other waves. Furthermore, we demonstrated that if a collision occurs, the lump wave remains unchanged. Our research includes a comprehensive graphical analysis of the solutions, accompanied by detailed explanations of key parameters such as velocity, amplitude, and peak location for each wave.

References

[1] Hirota, Ryogo, Exact solution of the Korteweg—de Vries equation for multiple collisions of solitons, Physical Review Letters, 27 (1971) 19-28.

[2] Wang, Shaofu, Novel soliton solutions of CNLSEs with Hirota bilinear method, Journal of Optics, 54 (2023) 1-6.

[3] Kumar, Sachin and Mohan, Brij, A generalized nonlinear fifth-order KdV-type equation with multiple soliton solutions: Painlevé analysis and Hirota Bilinear technique, Physica Scripta, 97 (2022) 123-132.

[4] Yuan, Feng and Ghanbari, Behzad, A study of interaction soliton solutions for the $(2+1)$-dimensional Hirota–Satsuma–Ito equation, Nonlinear Dynamics, 18 (2024) 1-9.

[5] Ma, Wen-Xiu and Huang, Yehui and Wang, Fudong, Inverse scattering transforms and soliton solutions of nonlocal reverse-space nonlinear Schrödinger hierarchies, Studies in Applied Mathematics, 145 (2020) 563-585.

[6] Ohta, Yasuhiro and Yang, Jianke, General high-order rogue waves and their dynamics in the nonlinear Schrödinger equation, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 468 (2012) 1716-1740.

[7] Wu, Xi-Hu and Gao, Yi-Tian, Generalized Darboux transformation and solitons for the Ablowitz–Ladik equation in an electrical lattice, Applied Mathematics Letters, 137 (2023) 108-122.

[8] Ablowitz, MJ and Satsuma, J, Solitons and rational solutions of nonlinear evolution equations, Journal of Mathematical Physics, 19 (1978) 2180-2186.

[9] Qi, Zequn and Chen, Qingqing and Wang, Miaomiao and Li, Biao, New mixed solutions generated by velocity resonance in the (2+ 1)-dimensional Sawada–Kotera equation, Nonlinear Dynamics, 108 (2022) 1617-1626.

[10] Zhang, Zhao and Yang, Xiangyu and Li, Wentao and Li, Biao, Trajectory equation of a lump before and after collision with line, lump, and breather waves for (2+ 1)-dimensional Kadomtsev–Petviashvili equation, Chinese Physics B, 28, (2019), 1-18.

[11] Chen, Si-Jia and Yin, Yu-Hang and Lü, Xing, Elastic collision between one lump wave and multiple stripe waves of nonlinear evolution equations, Communications in Nonlinear Science and Numerical Simulation, 130 (2024) 107205.

[12] Akinyemi, Lanre, Shallow ocean soliton and localized waves in extended (2+ 1)-dimensional nonlinear evolution equations, Physics Letters A 463 (2023) 1-24.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–263, pp. 564–569

Four weak solutions for nonlinear equations

Saeid Shokooh [1]*, Amir Ali Shamlou [2] and Mehdi Shahini [3]

Department of Mathematics, Faculty of Basic Sciences, Gonbad Kavous University, Gonbad Kavous, Iran.

Article Info

Keywords:
boundary value problem
weak solutions
variational method

2020 MSC:
58E05
47J30

Abstract

In this paper, we investigate the existence of weak solutions for a second-order boundary value problem. We show that problems in the following form

$$\begin{cases} -v'' + p(x)v' + q(x)v = \lambda s_1(x)f_1(v) + s_2(x)f_2(v), & \alpha < x < \beta, \\ v(\alpha) = 0 = v(\beta) \end{cases}$$

have at least four weak solutions.

1. Introduction

In this paper we study the existence of weak solutions for the following boundary value problem

$$\begin{cases} -v'' + p(x)v' + q(x)v = \lambda s_1(x)f_1(v) + s_2(x)f_2(v), & \alpha < x < \beta, \\ v(\alpha) = 0 = v(\beta) \end{cases} \tag{1}$$

where $0 \leq \alpha < \beta, \lambda \in \mathbb{R}^+$, $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ are non-constant continuous functions, $s_1, s_2 \in L^1(\mathbb{R})$, $p, q \in L^\infty([\alpha, \beta])$ such that

$$\text{ess inf}_{x\in[\alpha,\beta]} q(x) > -\frac{\pi^2}{(\beta-\alpha)^2}.$$

The boundary value problem discussed in this paper finds applications in diverse scientific and engineering fields, including biology, mechanical engineering, chemical engineering, and physics. To address the problem, we utilize a combination of functional analysis methods and variational techniques. Our approach to studying problem (1) builds upon the multiplicity result previously established in [4, 5]. For the convenience of the reader, we provide this result below.

[1]Email address: shokooh@gonbad.ac.ir
[2]Email address: amiralishamlou1381@gmail.com
[3]Email address: mehdi.shahini@gmail.com
*Talker

Theorem 1.1 ([4, Theorem 1]). *Assume that Y be a real reflexive Banach space; $J : Y \to \mathbb{R}$ be a coercive and sequentially weakly lower semicontinuous C^1 functional whose derivative admits a continuous inverse on Y^*, $I_1, I_2 : Y \to \mathbb{R}$ are C^1 functionals with compact derivative. Let there exist two points $u^\diamond, v^\diamond \in Y$ with the following properties:*

(i) $u^\diamond$ is a strict local minimum of J and $J(u^\diamond) = I_1(u^\diamond) = I_2(u^\diamond) = 0$;

(ii) $\Phi(v^\diamond) \leq \Psi_1(v^\diamond)$ and $\Psi_2(v^\diamond) > 0$.

Also, let for some $\sigma \in \mathbb{R}$ either

$$\begin{aligned} &\sup_{\lambda>0} \inf_{v\in Y}(\lambda(J(v) - I_1(v) - \sigma) - I_2(v)) \\ &< \inf_{v\in Y} \sup_{\lambda>0}(\lambda(J(v) - I_1(v) - \sigma) - I_2(v)) \end{aligned} \tag{2}$$

or

$$\begin{aligned} &\sup_{\lambda>0} \inf_{v\in Y}(J(v) - I_1(v) - \lambda(\sigma + I_2(v))) \\ &< \inf_{v\in Y} \sup_{\lambda>0}(J(v) - I_1(v) - \lambda(\sigma + I_2(v))), \end{aligned} \tag{3}$$

and suppose that

$$\max\left\{ \limsup_{\|v\|\to+\infty} \frac{I_1(v)}{J(v)}, \limsup_{v\to u^\diamond} \frac{I_1(v)}{I(v)} \right\} < 1 \tag{4}$$

and

$$\max\left\{ \limsup_{\|v\|\to+\infty} \frac{I_2(v)}{J(v)}, \limsup_{v\to u^\diamond} \frac{I_2(v)}{J(v)} \right\} \leq 0. \tag{5}$$

Under such hypotheses, there exists $\lambda^* > 0$ such that the equation $J' = I_1' + \lambda^* I_2'$ has at least four solutions in Y. More precisely, among them, one is $u^\diamond$ as a strict local, not global minimum and two are global minima of the functional $J - I_1 - \lambda^* I_2$.

Remark 1.2. It is important to remark that, in view of Theorem 1 of [3], condition (2) is equivalent to the existence of $u^{\diamond\diamond}, v^{\diamond\diamond} \in X$ satisfying

$$J(u^{\diamond\diamond}) - I_1(u^{\diamond\diamond}) < \sigma < J(v^{\diamond\diamond}) - I_1(v^{\diamond\diamond})$$

and

$$\frac{\sup_{(J-I_1)^{-1}(]-\infty,\sigma])} I_2 - I_2(u^{\diamond\diamond})}{\sigma - J(u^{\diamond\diamond}) + I_1(u^{\diamond\diamond})} < \frac{\sup_{(J-I_1)^{-1}(]-\infty,\sigma])} I_2 - I_2(v^{\diamond\diamond})}{\sigma - J(v^{\diamond\diamond}) + I_1(v^{\diamond\diamond})}.$$

Likewise, condition (3) is equivalent to the existence of $u_1, v_1 \in X$ satisfying

$$I_2(v^{\diamond\diamond}) < \sigma < J_2(u^{\diamond\diamond})$$

and

$$\frac{J(u^{\diamond\diamond}) - I_1(u^{\diamond\diamond}) - \inf_{I_2^{-1}([\sigma,+\infty[)}(J - I_1)}{I_2(u^{\diamond\diamond}) - \sigma} < \frac{J(v^{\diamond\diamond}) - I_1(v^{\diamond\diamond}) - \inf_{I_2^{-1}([\sigma,+\infty[)}(J - I_1)}{I_2(v^{\diamond\diamond}) - \sigma}.$$

In the Sobolev space $Y = W_0^{1,2}([\alpha, \beta])$, we adopt a norm defined as:

$$\|v\| = \left(\int_\alpha^\beta e^{-P(x)} |v'(x)|^2 dx + \int_\alpha^\beta e^{-P(x)} q(x) |v(x)|^2 dx \right)^{\frac{1}{2}},$$

where $P(x) = \int_0^x p(t)dt$ for all $x \in [\alpha, \beta]$.

Remark 1.3. In view of Proposition 2.1 in [1], we have

$$c_1\|v\|_Y \leq \|v\| \leq c_2\|v\|_Y$$

where

$$c_1 = \begin{cases} \left(\min_{x\in[\alpha,\beta]} e^{-P(x)}\right)^{\frac{1}{2}} & \text{if ess inf}_{x\in[\alpha,\beta]} q(x) \geq 0, \\ \left[\min_{x\in[\alpha,\beta]} e^{-P(x)} \left(1 + \text{ess inf}_{x\in[\alpha,\beta]} q(x) \frac{(\beta-\alpha)^2}{\pi^2}\right)^2\right]^{\frac{1}{2}} & \text{if ess inf}_{x\in[\alpha,\beta]} q(x) < 0, \end{cases}$$

$$c_2 = \begin{cases} \left[\max_{x\in[\alpha,\beta]} e^{-P(x)} \left(1 + \text{ess sup}_{x\in[\alpha,\beta]} q(x) \frac{(\beta-\alpha)^2}{\pi^2}\right)^2\right]^{\frac{1}{2}} & \text{if ess sup}_{x\in[\alpha,\beta]} q(x) \geq 0, \\ \left(\max_{x\in[\alpha,\beta]} e^{-P(x)}\right)^{\frac{1}{2}} & \text{if ess sup}_{x\in[\alpha,\beta]} q(x) < 0, \end{cases}$$

and $\|v\|_Y = \left(\int_\alpha^\beta |v'(x)|^2 dx\right)^{\frac{1}{2}}$. Also, for all $v \in Y$

$$\max_{x\in[\alpha,\beta]} |v(x)| \leq \frac{\sqrt{\beta-\alpha}}{2c_1}\|v\|.$$

The following lemma of [2, 6] is useful in the proof of our results.

Lemma 1.4. *Let $s : [\alpha,\beta] \to [0,+\infty[$ be a non-zero function in $L^1([\alpha,\beta])$ and let $f : \mathbb{R} \to \mathbb{R}$ be a continuous non-zero function, we define the functional $Z_{s,f}$ on Y as*

$$Z_{s,f}(v) = \int_\alpha^\beta s(x) f(v(x)) dx$$

for all $v \in Y$. Then, the following hold:

$$\limsup_{\|v\|\to 0} \frac{Z_{s,f}(v)}{\|v\|^2} \leq \frac{\beta-\alpha}{2c_1^2}\|s\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{t\to 0} \frac{f(t)}{|t|^2}\right\} \tag{6}$$

and

$$\limsup_{\|v\|\to+\infty} \frac{Z_{s,f}(v)}{\|v\|^2} \leq \frac{\beta-\alpha}{2c_1^2}\|s\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{|t|\to+\infty} \frac{f(t)}{|t|^2}\right\}. \tag{7}$$

Furthermore, if we define the functional $Z_{s,F}$ on Y by $Z_{s,F}(v) = \int_\alpha^\beta s(x)F(v(x))dx$, where $F(t) = \int_0^t f(\xi)\, d\xi$ for all $t \in \mathbb{R}$, then $Z_{s,F}$ is in $C^1(Y,\mathbb{R})$ and its derivative is given by

$$Z'_{s,F}(v)(u) = \int_\alpha^\beta s(x) f(v(x))\, u(x) dx$$

for all $u, v \in Y$. Also, due to the compact embedding of Y in $C^0([\alpha,\beta])$, we can infer that the mapping $Z'_{s,F} : Y \to Y^*$ is a compact operator.

The weak solution to problem (1) can be characterized by the critical points of the functional

$$\frac{1}{2}\|v\|^2 - \lambda Z_{s_1,F_1}(v) - Z_{s_2,F_2}(v),$$

where $v \in Y$. In other words, a function v belongs to Y as a weak solution to problem (1) if and only if it is a critical point of the given functional.

2. Main result

Here, you will find the main result stated.

Theorem 2.1. *Suppose $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ are two continuous non-constant functions, and $s_1, s_2 : [\alpha, \beta] \to [0, +\infty[$ are two non-constant functions in $L^1([\alpha, \beta])$. Let the following conditions hold:*

$$(f1)\ \ \max\left\{ \limsup_{|t|\to+\infty} \frac{\int_0^t f_1(\xi)\,d\xi}{|t|^2}, \limsup_{t\to 0} \frac{\int_0^t f_1(\xi)\,d\xi}{|t|^2} \right\} \leq 0,$$

$$(f2)\ \ \sup_{t\in\mathbb{R}} \int_0^t f_2(\xi)\,d\xi < +\infty, \quad \limsup_{t\to 0} \frac{\int_0^t f_2(\xi)\,d\xi}{|t|^2} < \frac{4c_1^2}{(\beta-\alpha)\|s_2\|_{L^1([\alpha,\beta])}}.$$

Additionally, suppose there are $\rho > \frac{\sqrt{\beta-\alpha}}{2c_1} \max\left\{1, (\|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2)^{\frac{1}{2}}\right\}$ and $a \in \mathbb{R}$ satisfying the following conditions:

$$(f3)\ \ 0 < \int_0^a f_1(\xi)\,d\xi = \sup_{|t|\leq\rho} \int_0^t f_1(\xi)\,d\xi < \sup_{t\in\mathbb{R}} \int_0^t f_1(\xi)\,d\xi,$$

$$(f4)\ \ |a|^2 \leq \frac{\beta-\alpha}{5c_2^2} \|s_2\|_{L^1([\alpha,\beta])} \int_0^a f_2(\xi)\,d\xi.$$

Under these assumptions, there exists $\lambda^* > 0$ such that the problem

$$\begin{cases} -v'' + p(x)v' + q(x)v = \lambda^* s_1(x) f_1(v) + s_2(x) f_2(v), & \alpha < x < \beta, \\ v(\alpha) = 0 = v(\beta) \end{cases}$$

has at least four weak solutions.

Proof. Our objective is to apply Theorem 1.1. Let I_1, I_2 equal, respectively, to Z_{s_2,F_2}, Z_{s_1,F_1} as defined in Section 1. For any $v \in Y$, let the functional $J : Y \to \mathbb{R}$ we define the functional $J : Y \to \mathbb{R}$ as follows

$$J(v) := \frac{1}{2}\|v\|^2.$$

The functional J is continuously differentiable, and its differential at any point $v \in Y$ is given by:

$$J'(v)(u) = \int_\alpha^\beta e^{-P(x)}(v'(x)u'(x) + q(x)v(x)u(x))dx$$

for all $u \in Y$. Also, J is sequentially weakly lower semicontinuous. Consider $u^\diamond = 0$ and define $v^\diamond(x)$ as follows:

$$v^\diamond(x) = \begin{cases} 5a\frac{x-\alpha}{\beta-\alpha}, & \alpha \leq x < \alpha + \frac{\beta-\alpha}{5}, \\ a, & \alpha + \frac{\beta-\alpha}{5} \leq x \leq \beta - \frac{\beta-\alpha}{5}, \\ 5a\frac{\beta-x}{\beta-\alpha}, & \beta - \frac{\beta-\alpha}{5} < y \leq \beta. \end{cases}$$

So,

$$\|v^\diamond\|_Y^2 = \int_\alpha^\beta |v^{\diamond\prime}(x)|^2 dx = \frac{10a^2}{\beta-\alpha}.$$

It is evident that assumption (i) of Theorem 1.1 holds.

Furthermore, in view of $(f3)$ and $(f4)$, we observe that

$$J(v^\diamond) \le \frac{5c_2^2 a^2}{\beta-\alpha} \le \|s_2\|_{L^1([\alpha,\beta])} \int_0^a f_2(\xi)\, d\xi = Z_{s_2,F_2}(v^\diamond)$$

and

$$Z_{s_1,F_1}(v^\diamond) = \|s_1\|_{L^1([\alpha,\beta])} \int_0^a f_1(\xi)\, d\xi > 0,$$

thus condition (ii) of Theorem 1.1 holds. For any v belonging to Y, it can be observed that

$$\frac{Z_{s_2,F_2}(v)}{J(v)} = \frac{2Z_{s_2,F_2}(v)}{\|v\|^2},$$

and according to Lemma 1.4, we obtain

$$\limsup_{\|v\|\to+\infty} \frac{Z_{s_2,F_2}(v)}{J(v)} \le \frac{\beta-\alpha}{2c_1^2}\|s_2\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{|t|\to+\infty} \frac{\int_0^t f_2(\xi)\, d\xi}{|t|^2}\right\} \le 0.$$

Furthermore, considering Lemma 1.4 and $(f2)$, one has

$$\limsup_{\|v\|\to 0} \frac{Z_{s_2,F_2}(v)}{J(v)} \le \frac{\beta-\alpha}{2c_1^2}\|s_2\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{\xi\to 0} \frac{\int_0^t f_2(\xi)\, d\xi}{|t|^2}\right\} < 1.$$

Thus, (4) is satisfied. Similarly, based on Lemma 1.4 and (f_1), we have

$$\limsup_{\|v\|\to+\infty} \frac{Z_{s_1,F_1}(v)}{J(v)} \le \frac{\beta-\alpha}{2c_1^2}\|s_1\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{|t|\to+\infty} \frac{\int_0^t f_1(\xi)\, d\xi}{|t|^2}\right\} \le 0$$

and

$$\limsup_{\|v\|\to 0} \frac{Z_{s_1,F_1}(v)}{J(v)} \le \frac{\beta-\alpha}{2c_1^2}\|s_1\|_{L^1([\alpha,\beta])} \max\left\{0, \limsup_{t\to 0} \frac{\int_0^t f_1(\xi)\, d\xi}{|t|^2}\right\} \le 0.$$

Now, we verify the satisfaction of (2). With $\rho > \frac{\sqrt{\beta-\alpha}}{2c_1}$, it is evident that

$$1 - \|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2 < \frac{2\rho^2 c_1^2}{\beta-\alpha} - \|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2.$$

Thus, there exists $\sigma \in \mathbb{R}$ such that

$$\max\left\{0, 1 - \|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2\right\} < \sigma < \frac{2\rho^2 c_1^2}{\beta-\alpha} - \|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2\}.$$

Suppose v belongs to Y such that $J(v) - Z_{s_2,F_2}(v) \le \sigma$. Consequently, we obtain

$$\|v\| \le \left(2\sigma + \|s_2\|_{L^1([\alpha,\beta])} \sup_{\mathbb{R}} F_2\right)^{\frac{1}{2}} \le \frac{2\rho c_1}{\sqrt{\beta-\alpha}}.$$

Therefore, due to the fact that Y is embedded in $C^0([\alpha,\beta])$, one has

$$\left\{v \in Y \,:\, J(v) - Z_{s_2,F_2}(v) \le \sigma\right\} \subseteq \left\{v \in Y \,:\, \sup_{x\in[\alpha,\beta]} |v(x)| \le \rho\right\}. \tag{8}$$

To satisfy the equivalent form of (2) mentioned in Remark 1.2, select $u^{\diamond\diamond} = v^\diamond$ and set $v^{\diamond\diamond}$ any constant d satisfying $F_1(d) > \sup_{[-\rho,\rho]} F_1$. The existence of such a d is guaranteed by $(f3)$. By virtue of $(f4)$, we obtain

$$J(u^{\diamond\diamond}) - Z_{s_2,F_2}(u^{\diamond\diamond}) \le \frac{5c_2^2 a^2}{\beta-\alpha} - \|s_2\|_{L^1([\alpha,\beta])} \int_0^a f_2(\xi)\, d\xi \le 0 < \sigma.$$

Furthermore, it is necessary that $J(v^{\diamond\diamond}) - Z_{s_2,F_2}(v^{\diamond\diamond})$ is strictly greater than σ. Otherwise, from (8), we would have $|d| \le \rho$ and $F_1(d) \le \sup_{[-\rho,\rho]} F_1$, which leads to a contradiction. Therefore, by virtue of $(f3)$ and the selection of d, we can easily deduce that

$$\sup_{(J-Z_{s_2,F_2})^{-1}(]-\infty,\rho])} Z_{s_1,F_1} \le Z_{s_1,F_1}(u^{\diamond\diamond})$$

and

$$\sup_{(J-Z_{s_2,F_2})^{-1}(]-\infty,\sigma])} Z_{s_1,F_1} \le Z_{s_1,F_1}(v^{\diamond\diamond}).$$

Thus, the following inequalities are valid:

$$\frac{\sup_{(J-Z_{s_2,F_2})^{-1}(]-\infty,\sigma])} Z_{s_1,F_1} - Z_{s_1,F_1}(u^{\diamond\diamond})}{\sigma - J(u^{\diamond\diamond}) + Z_{s_2,F_2}(u^{\diamond\diamond})} < 0 < \frac{\sup_{(J-Z_{s_2,F_2})^{-1}(]-\infty,\sigma])} Z_{s_1,F_1} - Z_{s_1,F_1}(v^{\diamond\diamond})}{\rho - \Phi(v_1) + T_{b,G}(v_1)}$$

furthermore, with each assumption of Theorem 1.1 being met, the proof is now finished □

.

References

[1] E. Amoroso, G. Bonanno, G. D'Aguì, S. Foti, Multiple solutions for nonlinear Sturm–Liouville differential equations with possibly negative variable coefficients, *Nonlinear Anal. Real World Appl.,* **69** (2023), 103730.

[2] F. Cammaroto and L. Vilasi, Multiplicity results for a Neumann boundary value problem involving the $p(x)$-Laplacian, *Taiwanese J. Math.,* **16** (2012), 621-634.

[3] G. Cordaro, On a minimax problem of Ricceri, *J. Inequal. Appl.,* **6** (2001), 261-285.

[4] B. Ricceri, A class of nonlinear eigenvalue problems with four solutions, *J. Nonlinear Convex Anal.,* **11** (2010), 503-511.

[5] B. Ricceri, Another four critical points theorem, *Nonlinear Analysis and Convex Analysis,* **2** (2011), 163-171.

[6] S. Shokooh, A class of fourth-order Kirchhoff type elliptic problems with four solutions, *Miskolc Math. Notes,* **21** (2020), 1019-1030.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–302, pp. 570–575

On the existence of solutions for a class of integro-differential equations

Saeid Shokooh [1]*, Fatemeh Nohtani [2] and Mohammad Fozouni [3]

Department of Mathematics, Faculty of Basic Sciences, Gonbad Kavous University, Gonbad Kavous, Iran.

Article Info

Keywords:
integro-differential equation
critical points
variational method

2020 MSC:
58E05
47J30

Abstract

This paper examines an integro-differential equation that incorporates a positive parameter. By employing the variational method and critical point theory, we demonstrate that when the control parameter is within a suitable range, our problem possesses a nontrivial weak solution. It is important to emphasize that the outcomes of our findings do not rely on any symmetry assumptions.

1. Introduction

In this study, we investigate the existence of weak solutions for the following Dirichlet boundary value problem

$$\begin{cases} -(a(x)u'(x))' + \sigma b(x) \int_\alpha^\beta b(x)u(x)dx = \lambda g(x, u(x)) & \text{in } (\alpha, \beta), \\ u(\alpha) = 0 = u(\beta). \end{cases} \tag{1}$$

Here, we consider a positive parameter λ, $0 \leq \alpha < \beta$, and an L^1-Carath'eodory function $g : [\alpha, \beta] \times \mathbb{R} \to \mathbb{R}$. Furthermore, the coefficient a belongs to the class $C^1([\alpha, \beta])$ and satisfies $a^- :=$ ess, $\inf_{x\in[\alpha,\beta]} a(x) > 0$. Similarly, the coefficient b is an L^2 function defined on the interval $[\alpha, \beta]$, with the condition that b is not identically zero. Finally, we introduce a real constant σ that satisfies the inequality:

$$\sigma > -\frac{\pi^2 a^-}{(\beta - \alpha)^2 \|b\|^2_{L^2}}. \tag{2}$$

Our aim is to establish the existence of weak solutions for the given boundary value problem, taking into account the specified conditions and constraints. For this purpose, we will rely on the variational principles proposed by Ricceri

[1]Email address: shokooh@gonbad.ac.ir
[2]Email address: fatemehnohtani210@gmail.com
[3]Email address: m.foozooni@gmail.com
*Talker

and Bonanno [8, Theorem 2.5] (also see [3, Theorem 5.1 and Proposition 2.1]). With the reader's convenience in consideration, we would like to present the following theorem, derived from Theorem 5.1 and Proposition 2.1 in the paper [3], as the main tool for establishing the results in the next section.
For $X \neq \emptyset$, $\Phi, \Psi : X \to \mathbb{R}$, we define the functions

$$\beta(r_1, r_2) := \inf_{v \in \Phi^{-1}(]r_1, r_2[)} \frac{\sup_{u \in \Phi^{-1}(]r_1, r_2[)} \Psi(u) - \Psi(v)}{r_2 - \Phi(v)} \tag{3}$$

and

$$\rho(r_1, r_2) := \sup_{v \in \Phi^{-1}(]r_1, r_2[)} \frac{\Psi(v) - \sup_{u \in \Phi^{-1}(]-\infty, r_1])} \Psi(u)}{\Phi(v) - r_1} \tag{4}$$

for all $r_1, r_2 \in \mathbb{R}$, with $r_1 < r_2$.

Theorem 1.1 ([3, Theorem 5.1]). *Let X be a real Banach space and let $\Phi, \Psi : X \to \mathbb{R}$ be two continuously Gâteaux differentiable functions. Assume that there are $r_1, r_2 \in \mathbb{R}$, with $r_1 < r_2$, such that*

$$\beta(r_1, r_2) < \rho(r_1, r_2),$$

where β and ρ are given by (3) and (4), and for each $\lambda \in \left]\frac{1}{\rho(r_1,r_2)}, \frac{1}{\beta(r_1,r_2)}\right[$ the functional $I_\lambda = \Phi - \lambda\Psi$ satisfies ${}^{[r_1]}(\mathrm{PS})^{[r_2]}$-condition. Then, for each $\lambda \in \left]\frac{1}{\rho(r_1,r_2)}, \frac{1}{\beta(r_1,r_2)}\right[$ there is $u_{0,\lambda} \in \Phi^{-1}(]r_1, r_2[)$ such that $I_\lambda(u_{0,\lambda}) \leq I_\lambda(u)$ for all $u \in \Phi^{-1}(]r_1, r_2[)$ and $I'_\lambda(u_{0,\lambda}) = 0$.

To facilitate comprehension in the next section, we will now introduce some notations that will be utilized later in the paper. Put

$$H_0^1([\alpha, \beta]) \quad := \quad \Big\{u \in L^2([0,1]) : u' \in L^2([\alpha, \beta]),\ u(\alpha) = u(\beta) = 0\Big\}.$$

Consider $X = H_0^1([\alpha, \beta])$ equipped with the standard norm defined as follows:

$$\|u\| := \left(\int_\alpha^\beta |u'(x)|^2\, dx\right)^{1/2}.$$

Let λ_1 represent the first eigenvalue of the problem

$$\begin{cases} -u'' = \lambda u & \text{in } (\alpha, \beta), \\ u(\alpha) = 0 = u(\beta), \end{cases}$$

Based on a well-known result, we have

$$\lambda_1 = \min_{u \in X,\ u \neq 0,} \frac{\|u'\|_{L^2}}{\|u\|_{L^2}} = \frac{\pi^2}{(\beta - \alpha)^2}. \tag{5}$$

Proposition 1.2. *Consider (2). Then, the norm $\|u\|_\sigma$ defined by*

$$\|u\|_\sigma = \left[\int_\alpha^\beta a(x)|u'(x)|^2 + \sigma\left(\int_\alpha^\beta b(x)u(x)\right)^2\right]^{1/2}$$

serves as a norm on X and is equivalent to the standard norm. Specifically, one has

$$c_1\|u\| \leq \|u\|_\sigma \leq c_2\|u\|$$

for all $u \in X$, where c_1, c_2 with $0 < c_1 \leq c_2$, are given by

$$c_1 = \left(a^- + \min\left\{0, \frac{\sigma(\beta-\alpha)^2\|b\|_{L^2}^2}{\pi^2}\right\}\right)^{1/2},$$

$$c_2 = \left(a^+ + \max\left\{0, \frac{\sigma(\beta-\alpha)^2\|b\|_{L^2}^2}{\pi^2}\right\}\right)^{1/2},$$

where $a^+ = \operatorname{ess\,sup}_{x\in[\alpha,\beta]} a(x)$.

The article [9] contains the proof for the proposition stated above.

Remark 1.3. The scalar product

$$\langle u, v\rangle_\sigma = \int_\alpha^\beta a(x)u'(x)v'(x)dx + \sigma \int_\alpha^\beta b(x)u(x)dx \int_\alpha^\beta b(x)v(x)dx$$

clearly defines the norm $\|.\|_\sigma$.

Remark 1.4. By utilizing Proposition 2.1 of [1] and Proposition (1.2), we can derive the following inequality:

$$\max_{x\in[\alpha,\beta]} |u(x)| \leq \frac{\sqrt{\beta-\alpha}}{2c_1}\|u\|_\sigma$$

for all $u \in X$.

A function $u \in X$ is considered a *weak solution* of problem (1) if it meets the following criterion:

$$\int_\alpha^\beta a(x)u'(x)v'(x)\,dx + \sigma\int_\alpha^\beta b(x)u(x)dx\int_\alpha^\beta b(x)v(x)dx - \lambda\int_\alpha^\beta g(x,u(x))v(x)\,dx = 0$$

for all $v \in X$.

2. Main result

This section focuses on the main outcomes.
Let ν be a given non-negative constant and τ be a positive constant such that $c_1\nu \neq 2c_2\tau$. We define

$$T_\tau(\nu) := \frac{\int_\alpha^\beta \sup_{|t|\leq\nu} G(x,t)\,dx - \int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx}{\frac{2c_1^2\nu^2}{\beta-\alpha} - \frac{4c_2^2\tau^2}{\beta-\alpha}}.$$

Theorem 2.1. *Suppose that condition (2) holds, and let there exist a non-negative constant ν_1 and two positive constants ν_2 and τ satisfying $\nu_1 < \sqrt{2}\,\tau$ and $\tau < \frac{c_1}{\sqrt{2}c_2}\nu_2$. Additionally, assume the following properties:*

(A1) $G(x,t) \geq 0$ *for all* $(x,t) \in ([\alpha, \alpha+\frac{\beta-\alpha}{4}] \cup [\beta-\frac{\beta-\alpha}{4}, \beta]) \times [0,\tau]$*;*

(A2) $T_\tau(\nu_2) < T_\tau(\nu_1)$.

Then, for each $\lambda \in]\frac{1}{T_\tau(\nu_1)}, \frac{1}{T_\tau(\nu_2)}[$, problem (1) possesses at least one nontrivial weak solution $u_0 \in X$ such that

$$\frac{2c_1\nu_1}{\sqrt{\beta-\alpha}} < \|u_0\|_\sigma < \frac{2c_1\nu_2}{\sqrt{\beta-\alpha}}.$$

Proof. To utilize Theorem 1.1, we define the functionals $\Phi, \Psi : X \to \mathbb{R}$ for any $u \in X$ as follows

$$\Phi(u) := \frac{1}{2}\|u\|_\sigma^2, \qquad \Psi(u) := \int_\alpha^\beta G(x, u(x))\, dx.$$

The functionals Φ and Ψ are well-defined and continuously differentiable. At the point $u \in X$, their derivatives $\Phi'(u)$ and $\Psi'(u)$ are operators defined as:

$$\Phi'(u)(v) = \int_\alpha^\beta a(x)u'(x)v'(x)\, dx + \sigma \int_\alpha^\beta b(x)u(x)dx \int_\alpha^\beta b(x)v(x)dx,$$

$$\Psi'(u)(v) = \int_\alpha^\beta g(x, u(x))v(x)\, dx$$

for any $v \in X$, respectively. In addition, the derivative Φ' has a continuous inverse on X^*.
Also, since Ψ is sequentially weakly upper semicontinuous, Ψ' is a compact operator. Hence, according to Proposition (4), $\Phi - \lambda\Psi$ satisfies ${}^{[r_1]}(\text{PS})^{[r_2]}$-condition where $r_1 = \frac{2c_1^2\nu_1^2}{\beta-\alpha}$ and $r_2 = \frac{2c_1^2\nu_2^2}{\beta-\alpha}$.
Now, we define $\overline{v} \in X$ as follows:

$$\overline{v}(x) = \begin{cases} 4\tau\frac{x-\alpha}{\beta-\alpha}, & x \in \left[\alpha, \alpha + \frac{\beta-\alpha}{4}\right[, \\ \tau, & x \in \left[\alpha + \frac{\beta-\alpha}{4}, \beta - \frac{\beta-\alpha}{4}\right], \\ 4\tau\frac{\beta-x}{\beta-\alpha}, & x \in \left]\beta - \frac{\beta-\alpha}{4}, \beta\right]. \end{cases}$$

The following inequalities hold, which can be easily verified:

$$\frac{4c_1^2\tau^2}{\beta-\alpha} \le \Phi(\overline{v}) \le \frac{4c_2^2\tau^2}{\beta-\alpha}. \tag{6}$$

From the conditions $\nu_1 < \sqrt{2}\,\tau$ and $\tau < \frac{c_1}{\sqrt{2}c_2}\,\nu_2$ it follows $r_1 < \Phi(w) < r_2$.
Considering the definition of Φ, we note that

$$\begin{aligned}\Phi^{-1}(]-\infty, r[) &= \{u \in X : \Phi(u) < r_2\} \\ &\subseteq \{u \in X : \|u\|_\sigma^2 < 2r_2\} \\ &\subseteq \{u \in X : |u(x)| < \nu_2 \quad \text{for all } x \in [\alpha, \beta]\}.\end{aligned}$$

Hence, we can conclude that

$$\begin{aligned}\sup_{u\in\Phi^{-1}(]-\infty,r_2[)} \Psi(u) &= \sup_{u\in\Phi^{-1}(]-\infty,r_2[)} \int_\alpha^\beta G(x, u(x))\, dx \\ &\le \int_\alpha^\beta \sup_{|t|\le\nu_2} G(x, t)\, dx.\end{aligned}$$

By employing the same reasoning as earlier, we obtain

$$\sup_{u\in\Phi^{-1}(]-\infty,r_1[)} \Psi(u) \le \int_\alpha^\beta \sup_{|t|\le\nu_1} G(x, t)\, dx.$$

Given that $0 \le \overline{v}(x) \le \tau$ for each $x \in [\alpha, \beta]$, the condition (A1) guarantees that

$$\int_\alpha^{\alpha+\frac{\beta-\alpha}{4}} G(x, \overline{v})\, dx + \int_{\beta-\frac{\beta-\alpha}{4}}^\beta G(x, \overline{v})\, dx \ge 0,$$

and so,

$$\Psi(w) \geq \int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx.$$

Consequently, we deduce that

$$\begin{aligned}\beta(r_1,r_2) &\leq \frac{\sup_{u\in\Phi^{-1}(]-\infty,r_2[)}\Psi(u)-\Psi(\overline{v})}{r_2-\Phi(\overline{v})}\\ &\frac{\int_\alpha^\beta \sup_{|t|\leq\nu_2} G(x,t)\,dx - \int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx}{\frac{2c_1^2\nu_2^2}{\beta-\alpha}-\frac{4c_2^2\tau^2}{\beta-\alpha}} = T_\tau(\nu_2).\end{aligned}$$

In addition, one has

$$\begin{aligned}\rho(r_1,r_2) &\geq \frac{\Psi(\overline{v})-\sup_{u\in\Phi^{-1}(]-\infty,r_1])}\Psi(u)}{\Phi(\overline{v})-r_1}\\ &\geq \frac{\int_\alpha^\beta \sup_{|t|\leq\nu_1} G(x,t)\,dx - \int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx}{\frac{2c_1^2\nu_1^2}{\beta-\alpha}-\frac{4c_2^2\tau^2}{\beta-\alpha}} = T_\tau(\nu_1).\end{aligned}$$

Therefore, considering assumption (A2), we can infer that $\beta(r_1,r_2) < \rho(r_1,r_2)$. Consequently, by applying Theorem 1.1, we can conclude that for every $\lambda \in]\frac{1}{T_\tau(\nu_1)}, \frac{1}{T_\tau(\nu_2)}[$, considering that the weak solutions of the problem (1) correspond precisely to the solutions of the equation $\Phi'(u)-\lambda\Psi'(u)=0$, the desired conclusion follows. □

We now present the following result, which directly follows from Theorem 2.1:

Theorem 2.2. *Let ν and d be positive constants satisfying $\tau < \frac{c_1}{\sqrt{2}c_2}\nu$, and assume that the conditions (A1) and (2) are fulfilled. Additionally, suppose the following assumptions hold:*

(A3) $\frac{\int_\alpha^\beta \sup_{|t|\leq\nu} G(x,t)\,dx}{\nu^2} < \frac{c_1^2}{2c_2^2}\frac{\int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dt}{\tau^2}$,

(A4) $G(x,0)=0$ *for every* $x\in[\alpha,\beta]$.

Then, for each

$$\lambda \in \left]\frac{\frac{4c_2^2\tau^2}{\beta-\alpha}}{\int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx}, \frac{\frac{2c_1^2\nu^2}{\beta-\alpha}}{\int_\alpha^\beta \sup_{|t|\leq\nu} G(x,t)\,dx}\right[,$$

there exists at least one nontrivial weak solution $u_0\in X$ to problem (1) such that $|u_0(x)|<\nu$ for all $x\in[\alpha,\beta]$.

Proof. By setting $\nu_1=0$ and $\nu_2=\nu$, we can establish the desired result by utilizing Theorem 2.1. Specifically, considering assumption (A3), we have:

$$\begin{aligned}T_\tau(\nu) &< \frac{\left(1-\frac{\frac{4c_2^2\tau^2}{\beta-\alpha}}{\frac{2c_1^2\nu^2}{\beta-\alpha}}\right)\int_\alpha^\beta \sup_{|t|\leq\nu} G(x,t)\,dx}{\frac{2c_1^2\nu^2}{\beta-\alpha}-\frac{4c_2^2\tau^2}{\beta-\alpha}}\\ &= \frac{\int_\alpha^\beta \sup_{|\nu|\leq\nu} G(x,t)\,dx}{\frac{2c_1^2\nu^2}{\beta-\alpha}}.\end{aligned}$$

In addition, considering assumption (A4), we have the following:

$$T_\tau(0) = \frac{\int_{\alpha+\frac{\beta-\alpha}{4}}^{\beta-\frac{\beta-\alpha}{4}} G(x,\tau)\,dx}{\frac{2c_1^2\nu^2}{\beta-\alpha}}.$$

Therefore, the conclusion can be derived directly from Theorem 2.1. □

References

[1] E. Amoroso, G. Bonanno, G. D'Aguì, S.D. Caro, S. Foti, D. O'regan, A. Testa, Second-order differential equations for the power converters dynamical performance analysis, *Math Meth Appl Sci.,* **49** (2022), 5573–5591.

[2] E. Amoroso, G. Bonanno, G. D'Aguì, S. Foti, Multiple solutions for nonlinear Sturm–Liouville differential equations with possibly negative variable coefficients, *Nonlinear Anal. Real World Appl.,* **69** (2023), 103730.

[3] G. Bonanno, A critical point theorem via the Ekeland variational principle, *Nonlinear Anal.,* **75** (2012), 2992–3007.

[4] F. Cammaroto and L. Vilasi, Multiplicity results for a Neumann boundary value problem involving the $p(x)$-Laplacian, *Taiwanese J. Math.,* **16** (2012), 621-634.

[5] G. Cordaro, On a minimax problem of Ricceri, *J. Inequal. Appl.,* **6** (2001), 261-285.

[6] B. Ricceri, A class of nonlinear eigenvalue problems with four solutions, *J. Nonlinear Convex Anal.,* **11** (2010), 503-511.

[7] B. Ricceri, Another four critical points theorem, *Nonlinear Analysis and Convex Analysis,* **2** (2011), 163-171.

[8] B. Ricceri; A general variational principle and some of its applications, *J. Comput. Appl. Math.,* **113** (2000), 401–410.

[9] S. Shokooh, A class of fourth-order Kirchhoff type elliptic problems with four solutions, *Miskolc Math. Notes,* **21** (2020), 1019-1030.

[10] S. Shokooh, F. Nohtani, M. Fozouni, Existence of weak solutions for an integro-differential equation by variational method, submitted.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, OR–226, pp. 576–584

Control of fractional discrete-time linear systems by state feedback matrix

Marzie Sepehr[a,*], Hojjat Ahsani Tehrani[b]

[a] *Ph.D. student of Applied Mathematics, Operations Research,Faculty of Mathematical Sciences, Shahed University of Technology, Shahed, Iran*
[b] *Professor of Applied Mathematics, Numerical Analysis, Faculty of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran*

Article Info

Keywords:
Assign eigenvalue
Discrete-time
Fractional
Linear system
Practical
Stability
State feedback

2020 MSC:
49J15
49K15

Abstract

Practical stability of the fractional discrete-time linear systems by the state feedback matrix is investigated. First our system by the definitions of fractional order and a new state vector is simplified to a standard system, and it is one of the advantages of our method, because working with standard systems is easier than. Then by putting a state feedback in standard system, closed-loop of standard system is determined. It is enough to assign desired eigenvalues to the last system by similarity transformation. At the end it is showed which inputs and states converge to zero by an example.

1. Introduction

Fractional discrete-time and continuous-time linear systems have been of great interest recently among researches. Especially in the last few decades there has been an explosion of research activities on the application of fractional calculus in various fields of science such as physics, mechanics, chemistry, engineering, hydrology application, polymer theology, system biology and etc [1–4].
The problem of stability of positive fractional discrete-time linear systems is addressed in the papers [9–12]. But in many applications, just stability of controlled object is not enough, and it is required that the poles of the closed-loop system should lie in a certain restricted region of stability. Some recent contributions to the theory of fractional differential equations can be found in [5–8].
In this paper we investigate a method for finding the solution of fractional discrete-time linear systems with the state feedback. In section2 the fractional linear system is converted to a standard linear system by the definition of fractional order system and defining a new state vector and then by a state feedback a closed-loop of standard system is defined. In section 3 a method based on similarity transformation for assigning of eigenvalues to standard system is displayed.

*Talker
Email addresses: Marzie.sepehr68@gmail.com (Marzie Sepehr), hasani@shahroodut.ac.ir (Hojjat Ahsani Tehrani)

In section 4 for more intuitive steps is written an algorithm and a numerical example.
The method in this paper has some superiority that we do not have in papers [9, 11, 17] like, no have difficult equations to take our time and make some errors. Also we deal with standard system which working with this system is surely easier than fractional system. In last section we bring a numerical example and show that our method is applicable and all states and inputs converge to the balanced point, i.e., zero.

2. Positive fractional systems

In this paper the following definition of the fractional discrete derivative

$$\Delta^{\alpha} x_k = \sum_{j=0}^{k} (-1)^j \binom{\alpha}{j} x_{k-j} \quad , \quad 0 < \alpha < 1 \,, \tag{1}$$

will be used [8], where $\alpha \in \mathbb{R}$ is the order of the fractional difference, and

$$\binom{\alpha}{j} = \begin{cases} 1 & j = 0 \\ \frac{\alpha(\alpha-1)\cdots(\alpha-j+1)}{j!} & j = 1, 2, \cdots \end{cases} \tag{2}$$

Consider the fractional discrete linear system, described by the state-space equation

$$\Delta^{\alpha} x_{k+1} = A x_k + B u_k \,, \quad k \in \mathbb{Z}_+ \tag{3}$$

where $x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$ are respectively the state and input vectors, A, B are real matrixes of appropriate dimensions.
Using the definition (1) we may write the equation (3) in the form

$$x_{k+1} + \sum_{j=1}^{k+1} (-1)^j \binom{\alpha}{j} x_{k-j+1} = A x_k + B u_k \tag{4}$$

From (2) it follows that the coefficients

$$c_j = c_j(\alpha) = (-1)^j \binom{\alpha}{j+1} \,, \quad j = 1, 2, \cdots \tag{5}$$

Strongly decrease for increasing j and they are positive for $0 < \alpha < 1$. In practical problems it is assumed that j is bounded by some natural number h.

In this case the equation (4) takes the form

$$x_{k+1} = A_{\alpha} x_k + \sum_{j=1}^{h} c_j x_{k-j} + B u_k \,, \quad k \in \mathbb{Z}_+ \tag{6}$$

where

$$A_{\alpha} = A + I_n \alpha. \tag{7}$$

Note that the equation (6) describe a linear discrete-time system with h delays in state.

Definition 2.1. [8] The positive fractional system (4) is called practically stable if and only if the system (6) is asymptotically stable.

Defining the new state vector

$$\bar{x}_k = \begin{bmatrix} x_k \\ x_{k-1} \\ \vdots \\ x_{k-h} \end{bmatrix} \tag{8}$$

we may write the equation (6) in the form

$$\bar{x}_{k+1} = \bar{A}\bar{x}_k + \bar{B}u_k\,, \quad k \in \mathbb{Z}_+ \tag{9}$$

where

$$\bar{A} = \begin{bmatrix} A_\alpha & c_1 I_n & \cdots & c_{h-1} I_n & c_h I_n \\ I_n & 0 & \cdots & 0 & 0 \\ 0 & I_n & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & I_n & 0 \end{bmatrix} \in \mathbb{R}_+^{\bar{n}\times\bar{n}}$$

$$\bar{B} = \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}_+^{\bar{n}\times m}\,, \quad \bar{n} = (1+h)n. \tag{10}$$

3. Problem statement

Consider the linear system (9) with $\operatorname{rank}(\bar{B}) = m$, and with the state feedback

$$u_k = Kx_k \tag{11}$$

where $K \in \mathbb{R}_+^{m\times n}$ is a gain matrix.

By substituting (11) in (9) we obtain the closed-loop system

$$\bar{x}_{k+1} = (\bar{A} + \bar{B}K)\bar{x}_k = \Gamma\bar{x}_k \tag{12}$$

In fact K is chosen, such that the closed-loop system eigenvalues placed inside the unit circle then the system (12) is asymptotically stable.

4. Similarity transformation

Consider the state transformation

$$\bar{x}_k = T\tilde{x}_k \tag{13}$$

where T can be obtained by elementary similarity operations as described in [9]. By substituting (13) in (9) vector companion equation is

$$\tilde{x} = \tilde{A}\tilde{x}_k + \tilde{B}u_k.$$

In this way, $\tilde{A} = T^{-1}\bar{A}T$ and $\tilde{B} = T^{-1}\bar{B}$ are in a compact canonical form known as vector companion form:

$$\tilde{A} = \begin{bmatrix} & G_0 & \\ & \cdots\cdots\cdots & \\ I_{\bar{n}-m} & & O_{\bar{n}-m\times m} \end{bmatrix},\ \tilde{B} = \begin{bmatrix} B_0 \\ \cdots\cdots \\ O_{\bar{n}-m\times m} \end{bmatrix}$$

Here G_0 is an $m \times \bar{n}$ matrix and B_0 is an $m \times m$ upper triangular matrix. Note that if the Kronecker invariants of the pair $(\bar{B}, \bar{A})$ are regular, then $\tilde{A}$ and $\tilde{B}$ are always in the above form [9]. In the case of irregular Kronecker invariants, some rows of $I_{\bar{n}-m}$ in $\tilde{A}$ are displaced [10]. It may also be concluded that if the vector companion form of $\tilde{A}$ obtained from similarity operations has the above structure, then the Kronecker invariants associated with the pair $(\bar{B}, \bar{A})$ are regular [9].
The state feedback matrix which assigns all the eigenvalues to zero, for the transformed pair $(\tilde{B}, \tilde{A})$, is then chosen as

$$u = -B_0^{-1}G_0\tilde{x} = \tilde{F}\tilde{x} \tag{14}$$

which results in the primary state feedback matrix for the pair $(\bar{B}, \bar{A})$ defined as

$$F_p = \tilde{F}T^{-1} \tag{15}$$

Therfore the transformed closed-loop matrix $\tilde{\Gamma}_0 = \tilde{A} + \tilde{B}\tilde{F}$ assumes a compact Jordan form with zero eigenvalues

$$\tilde{\Gamma}_0 = \begin{bmatrix} & O_{m\times n} & \\ & \cdots\cdots & \\ I_{n-m} & & O_{n-m\times m} \end{bmatrix}$$

Theorem 4.1. *Let D be a block diagonal matrix in the form*

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{16}$$

If such block diagonal matrix D with self conjugate eigenvalue spectrum $\Lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_n\}$ is added to the transformed closed-loop matrix $\tilde{\Gamma}_0$, then the eigenvalues of the resulting matrix is the eigenvalues in the spectrum.

Proof. The sum of $\tilde{\Gamma}_0$ with D has the form:

$$A_\lambda = \tilde{\Gamma}_0 + D = \begin{bmatrix} & O_{m\times n} & \\ & \cdots\cdots & \\ I_{n-m} & & O_{n-m\times m} \end{bmatrix} + \begin{bmatrix} \lambda_1 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & 0 & 0 & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_k & 0 & \cdots & 0 \\ I_1 & 0 & \cdots & 0 & \lambda_{k+1} & \cdots & 0 \\ \vdots & \ddots & \cdots & 0 & 0 & \ddots & \vdots \\ 0 & \cdots & I_r & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

By expanding $\det(A_\lambda - \lambda I)$ along the first row it is obvious that the eigenvalues of A_λ are the same as the eigenvalues of D.
For the case of irregular Kronecker invariants [10] only some of the unit columns of I_{n-m} are displaced, since the unit elements are always below the main diagonal, the proof applies in the same manner. □

Hence, the matrix $\tilde{A}_\lambda$ thus obtained will be in primary vector companion form such that:

$$\tilde{A}_\lambda = \begin{bmatrix} & G_\lambda & \\ & \cdots\cdots & \\ I_{n-m} & & O \end{bmatrix}$$

where G_0 is an $m \times n$ matrix.

Let

$$\tilde{K}_\lambda = B_0^{-1}G_\lambda$$

and

$$\tilde{K} = \tilde{F} + \tilde{K}_\lambda = B_0^{-1}(-G_0 + G_\lambda)$$

.
then the feedback matrix of the pair $(\tilde{A}, \tilde{B})$ is defined by:

$$K = \tilde{K}T^{-1} = B_0^{-1}(-G_0 + G_\lambda)T^{-1} \tag{17}$$

If the above conditions are satisfied then the problem of stabilization can be solved by using the following algorithm:

5. Algorithm

In this section we first give an algorithm for finding a state feedback matrix which assigns zero eigenvalues to the closedloop system. Then we determine a state feedback matrix which assigns the closed-loop eigenvalues in specified spectrum.

Input: The controllable pair (A, B) and the primary state feedback $F_p, B_)^{-1}$ and T^{-1} which are calculated by the algorithm proposed by Karbassi and Bell [9, 10].

Step 1. Construct the block diagonal matrix D in the form (16).

Step 2. Find set $\quad A_\lambda = \tilde{\Gamma}_0 + D$

Step 3. Transform A_λ to primary vector companion form $\tilde{A}_\lambda$ as in (18) using elementary similarity operations as specified in corollary of theorem (4.1) .

Step 4. Using the formula (17) compute the state feedback matrix K.

Example 5.1. Consider the fractional system

$$\Delta^\alpha x_{k+1} = Ax_k + Bu_k \ , \quad K \in \mathbb{Z}_+$$

with $\alpha = 0.9, h = 2$ and matrixes

$$A = \begin{bmatrix} 1 & 2 & 3 \\ -2 & 1 & -1 \\ 4 & -1 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 0 \end{bmatrix}.$$

the state feedback is chosen so that so that the eigenvalues of the closed loop system are placed in the spectrum.

Since $m, n = 3$, using (5), (7) and (10) we obtain

$$c_1 = \frac{-\alpha(\alpha-1)}{2} = 0.45\,, \ c_2 = 0.0165\,,$$
$$\bar{n} = (1+h).n = 9$$

and

$$\bar{A} = \begin{bmatrix} 1.9 & 2 & 3 & 0.045 & 0 & 0 & 0.0165 & 0 & 0 \\ -2 & 1.9 & -1 & 0 & 0.045 & 0 & 0 & 0.0165 & 0 \\ 4 & -1 & 2.9 & 0 & 0 & 0.045 & 0 & 0 & 0.0165 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}_{9\times 9}$$

$$\bar{B} = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}_{3\times 9}$$

The open loop eigenvalues are

$$\{5.5393, 1.7027, -0.5272, 0.1904, -0.1644, -0.0043 \pm 0.0544i, -0.0161 \pm 0.0971i\}$$

which some of them are outside unit circle then system is unstable. In order to locate them inside the unit circle, first the feedback matrix where assigns zero to the eigenvalues of the closed loop system we obtain

$$F_p = \begin{bmatrix} -4.0333 & 2.3 & -2.2333 & 0.015 & 0.015 & -0.045 & 0.0055 & 0.0055 & -0.0165 \\ 0.0333 & -1.3 & -0.6667 & -0.015 & -0.015 & 0 & -0.0055 & -0.0055 & 0 \\ 2.0667 & -1.7 & 0.5667 & -0.03 & 0.015 & 0.045 & -0.011 & 0.0055 & 0.0165 \end{bmatrix}$$

Now we obtain the feedback matrix so that the eigenvalues of the closed loop system are placed in the spectrum

$$\Lambda = \{0.1 - 0.2\,i,\ 0.3\,i,\ -0.5 + 0.1\,i,\ 0.1 + 0.2\,i,\ -0.3\,i,\ -0.5 - 0.1\,i,\ 0.6,\ -0.4,\ 0.9\}$$

Using (36) we obtain the state feedback matrix

$$K = \begin{bmatrix} -4.3 & 2.0333 & -1.4333 & 0.0717 & 0.0717 & -0.2150 & -0.0045 & -0.0045 & 0.0135 \\ -0.1 & -1.4333 & -0.6667 & -0.045 & -0.045 & 0 & -0.0175 & -0.0175 & 0 \\ 2 & -1.6667 & 0.6667 & 0.3967 & -0.1983 & -0.595 & 0.1450 & -0.0725 & -0.2175 \end{bmatrix}$$

It can be verified that the eigenvalues of the closed-loop system are inside the unit circle the specified spectrum. The following forms state and input vectors to move your balance point.

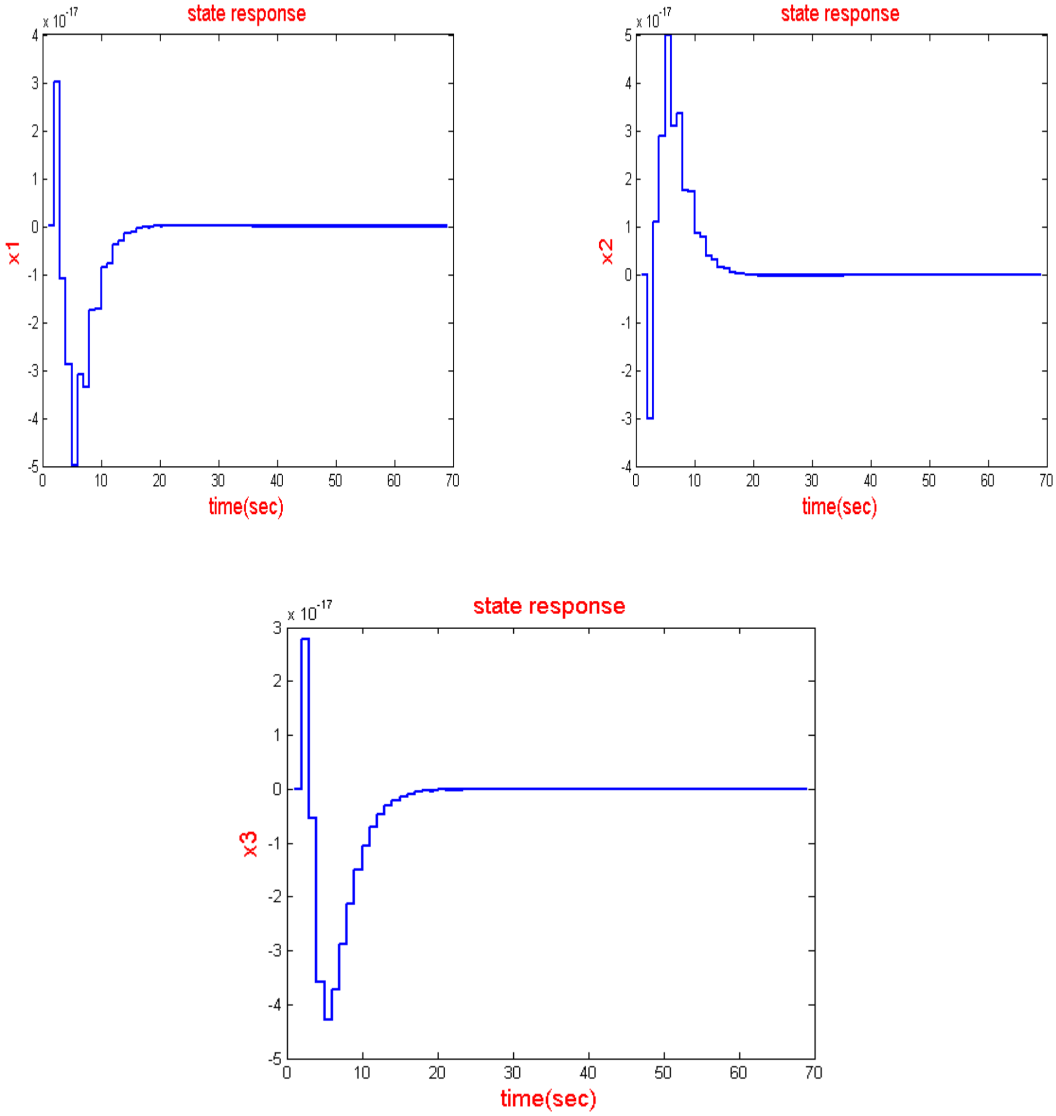


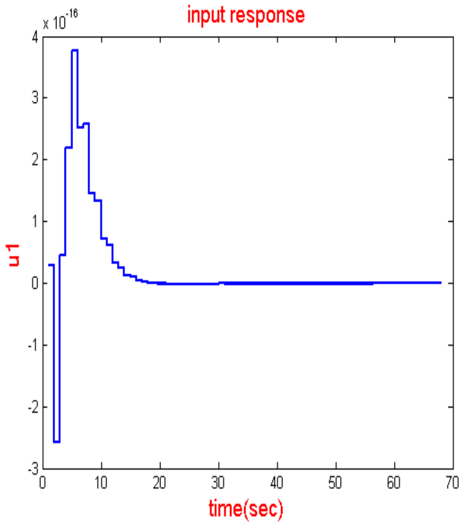
Fig. 1. State vectors to move your balance point

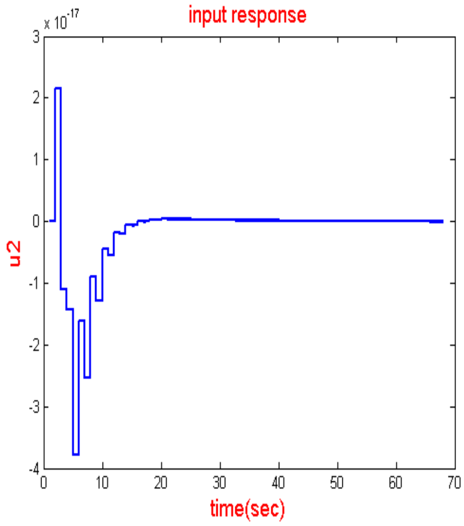
6. Concluding remarks

A method for finding the solution of fractional discrete-time linear systems with the state feedback is investigated. First by the definition of fractional order system and defining a new state vector, the fractional linear system is converted to a standard linear system. This is one of the advantages of this method, because we deal with standard system and it is more comfortable than using fractional systems. Second a method based on similarity transformation for assigning of eigenvalues to standard system is used. Finally an algorithm and an example for more intuitive is written and convergence of all inputs and states to zero is showed.

References

[1] K. Balachandran, S. Kiruthika and J.J. Trujillo, *On fractional impulsive equations of Sobolev type with nonlocal condition in Banach spaces, Computers & Mathematics with Applications, Vol.62, 1157–1165, 2011.*





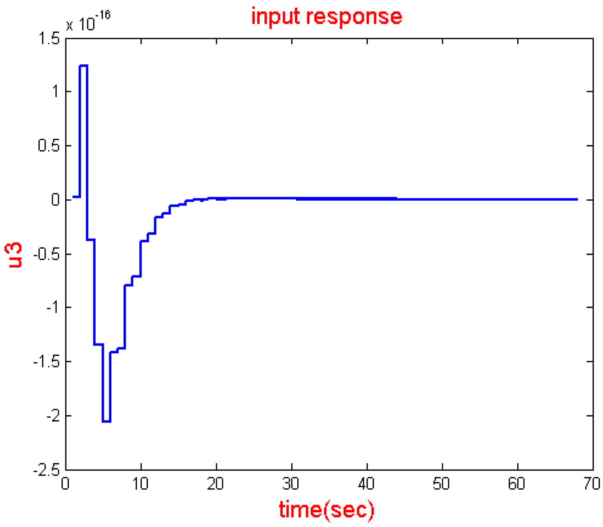


Fig. 2. Input vectors to move your balance point

[2] *K. Balachandran and S. Kiruthika,* Existence results for fractional integrodifferential equations with nonlocal condition via resolvent operators, Computers & Mathematics with Applications, Vol. 62, 1350–1358, 2011.

[3] K. Balachandran, S. Kiruthika and J.J. Trujillo, *Existence results for fractional impulsive integrodifferential equations in Banach spaces, Communications in Nonlinear Science and Numerical Simulation, Vol. 16, 1970–1977, 2011.*

[4] *M. Buslowicz,* Practical robust stability of positive fractional scalar discrete-time systems, Zeszyty Naukowe Politechniki Sl askiej: Automatyka, Vol. 151, 25–30, 2008.

[5] M. Buslowicz and T. Kaczorek, *Simple conditions for practical stability of positive fractional discrete-time linear systems, International Journal of Applied Mathematics and Computer Science, Vol. 19, 263–269, 2009*

[6] *T. Kaczorek,* Practical stability of positive fractional discretetime systems, Bull. Pol. Ac.: Tech, Vol. 56, 313–318, 2008.

[7] T. Kaczorek, *Stabilization of fractional discrete-time linear systems using state-feedback, Logistyka, CD–CD, 2009.*

[8] *T. Kaczorek,* Stability and stabilization of positive fractional linear systems by state-feedbacks, Bulletin of the Polish Academy of Sciences: Technical Sciences, Vol. 58, 537–554, 2010

[9] S.M. Karbassi and D.J. Bell, *Parametric time-optimal control of linear discrete-time systems by state feedback Part 1. Regular Kronecker invariants, International journal of control, Vol. 57, 817–830, 1993.*

[10] *S.M. Karbassi and D.J. Bell,* Parametric time-optimal control of linear discrete-time systems by state feedback Part 2. Irregular Kronecker invariants, International Journal of Control, Vol. 57, 831–839, 1993.

[11] F. Liu, V. Anh and I. Turner, *Numerical solution of the space fractional Fokker–Planck equation, Journal of Computational and Applied*

Mathematics, Vol. 166, No. 1 , 209-219, 2004.
[12] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Physics reports, Vol. 339, No. 1, 1–77, 2000.
[13] K.S. Miller and B. Ross, *An introduction to the fractional calculus and fractional differential equations, John Wiley & Sons, 1993.*
[14] K. Nishimoto, Fractional Calculus, Decartess Press, Koriama, 1984.
[15] K.B. Oldham and J. Spanier, *The fractional calculus, Academic Press, New York, 1974.*
[16] L. Podlubny, Fractional Differential Equations, Academic Press, San Diego, 1999.
[17] Y.A. Rossikhin and M.V. Shitikova, *Analysis of dynamic behaviour of viscoelastic rods whose rheological models contain fractional derivatives of two different orders, Zeitschrift für Angewandte Mathematik und Mechanik, Vol. 81, 363–376, 2001.*
[18] Z.B. Wang, G.Y. Cao and X. J. Zhu, Application of fractional calculus in system modeling, Journal of Shanghai Jiaotong University of Chinese edition, Vol. 38, 802–805, 2004.
[19] Y. Zhou and F. Jiao, *Nonlocal Cauchy problem for fractional evolution equations, Nonlinear Analysis: Real World Applications, Vol. 11, 4465–4475, 2010.*

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–126, pp. 585–587

Stanley's Conjecture on the cohen macaulay simplicial complexes of codimension 2

Seyed Mohammad Ajdani*

[a] *Department of Mathematics, Zanjan Branch, Islamic Azad University, Zanjan, Iran.*

Article Info

Keywords:
Stanley depth
Cohen-Macaulay
vertex decomposable

2020 MSC:
13A30
13C12, 13F55

Abstract

Let Δ be a simplicial complex on vertex set $[n]$. It is shown that if Δ is cohen macaulay simplicial complexes of codimension 2, then Δ is vertex decomposable and Stanley's conjecture holds for $K[\Delta]$. As a consequence we show that if Δ is a quasi-forest simplicial complex,Then $\Delta^{\vee}$ is vertex decomposable.

1. Introduction

Let Δ be a simplicial complex on vertex set $[n] = \{1, \cdots, n\}$, i.e. Δ is a collection of subsets of $[n]$ with the the property that if $F \in \Delta$, then all subsets of F are also in Δ. An element of Δ is called a *face* of Δ, and the maximal faces of Δ under inclusion are called *facets*. We denote by $\mathcal{F}(\Delta)$ the set of facets of Δ. The *dimension* of a face F is defined as $\dim F = |F| - 1$, where $|F|$ is the number of vertices of F. The dimension of the simplicial complex Δ is the maximum dimension of its facets. A simplicial complex Δ is called *pure* if all facets of Δ have the same dimension. Otherwise it is called non-pure. We denote the simplicial complex Δ with facets $F_1, \dots, F_t$ by $\Delta = \langle F_1, \dots, F_t\rangle$. A simplex is a simplicial complex with only one facet. For the simplicial complexes Δ_1 and Δ_2 defined on disjoint vertex sets, the join of Δ_1 and Δ_2 is $\Delta_1 * \Delta_2 = \{F \cup G \ : \ F \in \Delta_1,\ G \in \Delta_2\}$.

For the face F in Δ, the link, deletion and star of F in Δ are respectively, denoted by $\text{link}_{\Delta} F$, $\Delta \setminus F$ and $\text{star}_{\Delta} F$ and are defined by $\text{link}_{\Delta} F = \{G \in \Delta \ : \ F \cap G = \varnothing,\ F \cup G \in \Delta\}$ and $\Delta \setminus F = \{G \in \Delta \ : \ F \nsubseteq G\}$ and $\text{star}_{\Delta} F = \langle F\rangle * \text{link}_{\Delta} F$.

Let $R = K[x_1, \dots, x_n]$ be the polynomial ring in n indeterminates over a field K. To a given simplicial complex Δ on the vertex set $[n]$, the Stanley–Reisner ideal is the squarefree monomial ideal whose generators correspond to the non-faces of Δ. We say the simplicial complex Δ is Cohen–Macaulay, if $K[x_1, \dots, x_n]/I_{\Delta}$ is Cohen–Macaulay .

The facet ideal of Δ is the squarefree monomial ideal generated by monomials $x_F = \prod_{i\in F} x_i$ where F is a facet of Δ and is denoted by $I(\Delta)$. The complement of a face F is $[n] \setminus F$ and is denoted by F^c. Also, the complement of the simplicial complex $\Delta = \langle F_1, \dots, F_r\rangle$ is $\Delta^c = \langle F_1^c, \dots, F_r^c\rangle$. The Alexander dual of Δ is defined by $\Delta^{\vee} =$

*Seyed Mohammad Ajdani
Email address: majdani2@yahoo.com (Seyed Mohammad Ajdani)

$\{F^c\ : F \notin \Delta\}$. It is known that for the complex Δ one has $I_{\Delta^\vee} = I(\Delta^c)$. Also we call $K[\Delta] := S/I_\Delta$ the *Stanley-Reisner ring* of Δ. One of interesting problems in combinatorial commutative algebra is the Stanley's conjectures. The Stanley's conjectures are studied by many researchers. Let R be a $\mathbb{N}^n$- graded ring and M a $\mathbb{Z}^n$- graded R- module. Then Stanley [2] conjectured that

$$\text{depth}\,(M) \leq \text{sdepth}\,(M).$$

He also conjectured in [3] that each Cohen-Macaulay simplicial complex is partitionable. Herzog, Soleyman Jahan and Yassemi in [5] showed that the conjecture about partitionability is a special case of the Stanley's first conjecture. In this paper, we show that if Δ is cohen macaulay simplicial complexes of codimension 2, then Δ is vertex decomposable and Stanley's conjecture holds for $K[\Delta]$.

2. Preliminaries

In this section we fix some notation and recall some definitions. For a monomial $u = x_1^{a_1} \dots x_n^{a_n}$ in R, we denote the support of u by supp (u) and it is the set of those variables x_i that $a_i \neq 0$. Let m be another monomial in R. If for all $x_i \in$ supp (u), $x_i^{a_i} \nmid m$ then we set $[u, m] = 1$, otherwise we set $[u, m] \neq 1$.
For a monomial ideal $I \subset R$ we set $I^u = (m_i \in G(I)\ :\ [u, m_i] \neq 1)$ and
$I_u = (m_i \in G(I)\ :\ [u, m_i] = 1)$.
The concept of shedding monomial and k-decomposable monomial ideals was first introduced by Rahmati and Yassemi in [7].

Definition 2.1. Let I be a monomial ideal and $G(I) = \{m_1, \dots, m_r\}$. The monomial $u = x_1^{a_1} \dots x_n^{a_n}$ is called a shedding monomial of I if $I_u \neq 0$ and for each $m_i \in G(I_u)$ and each $x_l \in$ supp (u) there exists $m_j \in G(I^u)$ such that $\langle m_j : m_i\rangle = \langle x_l\rangle$.

Definition 2.2. Let I be a monomial ideal and $G(I) = \{m_1, \dots, m_r\}$. Then I is a *k-decomposable ideal if $r = 1$ or else has a shedding monomial u with $\mid$ supp (u) $\mid\leq k+1$ such that the ideals I^u and I_u are k-decomposable. Note that since $\mid G(I)\mid$ is finite, the recursion procedure will stop.*

A 0-decomposable ideal is called *variable decomposable*. Also, a monomial ideal is decomposable if it is k-decomposable for some $k \geq 0$.

Definition 2.3. A simplicial complex Δ is recursively defined to be *vertex decomposable*, if it is either a simplex, or else has some vertex v so that,

(a) Both $\Delta \setminus v$ and $\text{link}_\Delta(v)$ are vertex decomposable, and

(b) No face of $\text{link}_\Delta(v)$ is a facet of $\Delta \setminus v$.

A vertex v which satisfies in condition (b) is called a *shedding vertex*.

A monomial ideal $I \subset R = K[x_1, \dots, x_n]$ generated in a single degree is called polymatroidal if for any $u, v \in G(I)$ such that $\deg_{x_i}(u) > \deg_{x_i}(v)$ there an index j with $\deg_{x_j}(u) < \deg_{x_j}(v)$ such that $x_j(u/x_i) \in G(I)$. A squarefree polymatroidal ideal is called matroidal. Also, a monomial ideal I is called weakly polymatroidal if for every two monomials $u = x_1^{a_1} \dots x_n^{a_n} > v = x_1^{b_1} \dots x_n^{b_n}$ in $G(I)$ such that $a_1 = b_1, \dots, a_{t-1} = b_{t-1}$ and $a_t > b_t$, there exists $j > t$ such that $x_t(v/x_j) \in I$. It is clear from the definition that a polymatroidal ideal is weakly polymatroidal.
The following results from [7] are crucial in this paper.

Theorem 2.4. *[7, Theorem 2.10] Let Δ be a (not necessarily pure) d-dimensional simplicial complex on vertex set $[n]$. Then Δ is k-decomposable if and only if $I_{\Delta^\vee}$ is k-decomposable, where $k \leq d$.*

Proposition 2.5. *[7, Lemma 3.8] If I is an squarefree monomial ideal generated in degree 2 which has a linear resolution, then after suitable renumbering of the variables, I is weakly polymatroidal.*

Theorem 2.6. *[7, Theorem 3.5] Let $I \subset R$ be a weakly polymatroidal ideal. Then I is 0-decomposable.*

3. Vertex decomposability of cohen macaulay simplicial complexes of codimension 2

As the main result of this section, it is shown that if Δ is cohen macaulay simplicial complexes of codimension 2, then Δ is vertex decomposable and Stanley's conjecture holds for $K[\Delta]$.

Theorem 3.1. *If Δ is a Cohen-Macaulay simplicial complex of codimension 2, then Δ is vertex decomposable.*

Proof. Since Δ is Cohen-Macaulay simplicial complex of codimension 2, by a result of Eagon and Reiner [6], $I_{\Delta^\vee}$ is a squarefree monomial ideal which has 2-linear resolution. Hence by Proposition 2.5 and Theorem 2.6, $I_{\Delta^\vee}$ is 0-decomposable. It follows from Theorem 2.4 that Δ is vertex decomposable. $\square$

As an immediate consequence we have the following:

Corollary 3.2. *Let Δ be a quasi-forest simplicial complex which is not a simplex. Then $\Delta^\vee$ is vertex decomposable.*

Proof. It is proved in [8] that each quasi-forest is a flag complex. So I_Δ is generated by quadratic monomials and hence $ht(I_{\Delta^\vee}) = 2$. Since Δ is quasi-forest by [8, Corollary 5.5], we have $\operatorname{pd}(K[\Delta^\vee]) = 2$. Therefore $\Delta^\vee$ is Cohen-Macaulay of codimension 2 and by Theorem 3.1, $\Delta^\vee$ is vertex decomposable. $\square$

Stanley conjectured in [2] the upper bound for the depth of $K[\Delta]$ as the following:

$$\operatorname{depth}(K[\Delta]) \leq \operatorname{sdepth}(K[\Delta]).$$

Also we recall another conjecture of Stanley. Let Δ be again a simplicial complex on $\{x_1, \ldots, x_n\}$ with facets $G_1, \ldots, G_t$. The complex Δ is called partitionable if there exists a partition $\Delta = \bigcup_{i=1}^{t}[F_i, G_i]$ where $F_i \subseteq G_i$ are suitable faces of Δ. Here the interval $[F_i, G_i]$ is the set of faces $\{H \in \Delta : F_i \subseteq H \subseteq G_i\}$. In [3] and [4] respectively Stanley conjectured each Cohen-Macaulay simplicial complex is partitionable. This conjecture is a special case of the previous conjecture. Indeed, Herzog, Soleyman Jahan and Yassemi [5] proved that for Cohen-Macaulay simplicial complex Δ on $\{x_1, \ldots, x_n\}$ we have that $\operatorname{depth}(K[\Delta]) \leq \operatorname{sdepth}(K[\Delta])$ if and only if Δ is partitionable. Since each vertex decomposable simplicial complex is shellable and each shellable complex is partitionable. Then as a consequence of our results we obtain :

Corollary 3.3. *Let Δ be a cohen macaulay simplicial complexes of codimension 2 on vertex set $[n]$. Then Δ is partitionable and Stanley's conjecture holds for $K[\Delta]$.*

Proof. Since each vertex decomposable simplicial complex is shellable and each shellable complex is partitionable.Therefore by theorem 3.1 proof is completed. $\square$

References

[1] A Conca, E. De Negri, M-sequences, graph ideals and ladder ideals of linear type. J. Algebra **211** (1999), no. 2, 599-624
[2] R. P. Stanley, Linear Diophantine Equations and Local Cohomology, Invent. Math. **68**(1982), 175-193.
[3] R. P. Stanley, Combinatorics and commutative algebra. Second edition. Progress in Mathematics **41**, Birkhauser Boston 1996.
[4] R. p. Stanley, Positivity problems and conjectures in algebraic combinatorics. In Mathematics: frontiers and perspectives, 295-319, Amer. Math. Soc., Providence, RI 2000.
[5] J. Herzog, A. Soleyman Jahan, S. Yassemi, Stanley decompositions and partitionable simplicial complexes. J. Algebr. Comb. **27**(2008), 113-125
[6] J. Eagon and V. Reiner, Resolutions of Stanley–Reisner rings and Alexander duality, J. Pure Appl. Algebra **130** (1998), 265–275.
[7] R. Rahmati Asghar, S. Yassemi, k-decomposable monomial ideals. to appear in .
[8] X. Zheng, Homological properties of monomial ideals associated to quasi-trees and lattices. ph.d thesis.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–269, pp. 588–590

Fully cancellation and fully cocancellation S-acts on monoids

Zohreh Habibi[a,*], Masoomeh Hezarjaribi[b]

[a] *Department of Mathematics, Payame Noor University, Tehran, Iran*
[b] *Department of Mathematics, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
fully cancellation
fully cancellation
S-act

2020 MSC:
msc1
msc2

Abstract

In this paper, the notions Fully cancellation and fully cocancellation S-act of monoid S-acts is studied. The behaviour of these is investigate. Also, the relation between fully cancellation and fully cocancellaetion S-acts is studied.

1. Introduction and preliminaries

Acts over monoids, as universal algebras with unary operations, appear as basic mathematical models of some important notions in theoretical computer science and physics like automata, dynamical systems, etc. In [1] and [2] the notions fully cancellation and fully cocancellation modules is studied. In this study, we study these notions , in category of S-acts on monoids and investigate some properties of them.

First we give some preliminary needed in the following.

Suppose that S be a monoid. By a *(right) S-act* or *act over S*, we mean a non-empty set A jointly a map $A \times S \to A, (a, s) \mapsto as$, so that for all $a \in A, s, r \in S$, $(as)r = a(sr)$ and $a1 = a$.

A subset $B \subseteq A$ that is non-empty is called a *subact* of A if for all $b \in B$ and $s \in S$, $bs \in B$. Let A and B be two S-acts. A mapping $f : A \to B$ is called a *homomorphism* if for all $a \in A, s \in S$, $f(as) = f(a)s$. The category of all S-acts as well as all homomorphisms between them is denoted by **Act**-S. In this category, monomorphisms are exactly injective homomorphisms.

A non-empty subset I of a monoid S is called a *right ideal* of S if $xs \in I$ for any $x \in I$ and $s \in S$. An element $\theta \in A$ for which $\theta s = \theta$ for all $s \in S$ is said to be a *zero* or *fixed element* of A. A *congruence* on an S-act A is an equivalence relation ρ on A for which $a\rho a'$ implies that $(as)\rho(a's)$ for $a, a' \in A$ and $s \in S$. An S-act A is called *decomposable* if there exist subacts B and C of A such that $A = B \cup C$ and $B \cap C = \emptyset$. Otherwise, A is called *indecomposable*.

A *left zero semigroup* is a semigroup S with $sr = s$ for all $s, r \in S$. A *right zero semigroup* is defined similarly. Throughout, S stands for a monoid unless otherwise stated. For more see [3].

*Zohreh Habibi
Email addresses: Z_habibi@pnu.ac.ir (Zohreh Habibi), Masoomeh. hezarjaribi@pnu.ac.ir (Masoomeh Hezarjaribi)

2. Fully cancellation and fully cocancellation S-act

In this section, we define two notion fully cancellation and fully cocancellation and study some properties of these two notions. Also, we prove the condition which under it, two notion fully cancellation and fully cocancellation S-act is same.

Definition 2.1. Let A be an S-act. A is called fully cancellation if for any non-zero ideal I of S and for every subact B and C of A, if $BI = CI$, then we have $B = C$.

Example 2.2. The S-act $\mathbb{Z}$ on monoid $(\mathbb{Z}, .)$ with usual multiplication is fully cancellation. But $\mathbb{Z}$-act $\mathbb{Z}_4$ is not fully cancellation.

Remark 2.3. Any subact of fully cancellation S-act is fully cancellation.
The homomorphism image of the fully cancellation S-act is not fully cancellation. For example, consider the homomorphism $\pi : \mathbb{Z} \longrightarrow \mathbb{Z}_4$ on monoid $(\mathbb{Z}, .)$. Clearly, $\mathbb{Z}$ is fully cancellation and $\mathbb{Z}_4$ is not fully cancellation.
Suppos that $f : A \longrightarrow B$ be an one-to-one homomorphism.The inverse image of fully cancellation is fully cancellation. For this, consider an one-to-one homomorphism $f : A \longrightarrow B$ and fully cancellation subact B' of B. We show that $F^{-1}(B')$ is fully cancellation. Let A_1, A_2 be subacts of $f^{-1}(B')$ and I be a non-zero ideal of S so that $A_1I = A_2I$. We have $f(A_1I) = f(A_2I)$. So, we have $f(A_1)I = f(A_2)I$. Since B' is fully cancellation, we have $f(A_1) = F(A_2)$ and therefore $A_1 = A_2$.
Let A be a simple S-act. If A is fully cancellation, then it is faithful. For this, consider $0 \in Ann(A)$. We have $As = \theta$. Suppose that B ba proper subact of A. So, We have $Bs = \theta$. Therefore $As = Bs$. Since A is fully cancellation, we have $A = B$, which is contradiction.

Proposition 2.4. *Let A be a fully cancellation and cancellation S-act. Then every non-zero ideal of S is a cancellation ideal.*

Proof. Suppos that I be a non-zero ideal of S, J_1 and J_2 be subacts of I such that $IJ_1 = IJ_2$. We have $I(J_1A) = I(J_2A)$. Since A is fully cancellation, J_1A and J_2A is fully cancellation, so $J_1A = J_2A$. Now, we have $J_1 = J_2$, since A is cancellation. □

Corollary 2.5. *Suppos that A be a multiplication cancellation S-act. Then A is fully cancellation if and only if any non-zero ideal of S is a cancellation ideal.*

Proof. Sufficieny is follows from Proposition 2.4. For conversely, let A_1, A_2 be two subacts of A and I be a non-zero ideal of S such that $A_1I = A_2I$. We show that $A_1 = A_2$. We have $A_1 = (A_1 : A)A = (A_2 : A)A = A_2$, as the result. □

Theorem 2.6. *Suppos that $(A_j)_{j\in J}$ be a family of S-acts. Then $\prod_{j\in J} A_j$ is fully cancellation if and only if for any $j \in J$, A_j is fully cancellation.*

The next corollary is clear by Theorem 2.6.

Corollary 2.7. *Suppose that $(B_j)_{j\in J}$ be a family of S-acts. Then $\bigoplus_{j\in J} B_j$ is fully cancellation if and only if B_j is fully cancellation, for any $j \in J$.*

Definition 2.8. Let B be an S-act. B is a fully cocancellation if for subacts C_1 and C_2 of B and non-zero ideal J of S, if $(C_1 :_B J) = (C_2 :_B J)$, implies $C_1 = C_2$.

Remark 2.9. Any homomorphism image of fully cocancellation S-act is fully cocancellation.

Theorem 2.10. *Let $(B_j)_{j\in J}$ be a family of S-acts. If $\prod_{j\in J} B_j$ is fully cocancellation S-act then for any $j \in J$, B_j is fully cocancellation.*

Lemma 2.11. *Suppos that A be an S-act and B be an subact of A, C and D be subacts of B and J be a non-zero idempotent ideal of S so that $(C :_B J) = (D :_B J)$. So, we have $(C :_A J) = (D :_A J)$.*

Proposition 2.12. *Suppos that S be a regular monoid. Then any subact of fully cancellation S-act is fully cocancellation.*

Theorem 2.13. *Suppose that A be an S-act. Then A is a fully cocancellation if and only if $(B :_A I) \subseteq (C :_A I)$, implies $B \subseteq C$, for any subacts B, C of S-act A and non-zero ideal I of S.*

Theorem 2.14. *Suppose that S be a group. Then any S-act is fully cocancellation.*

Lemma 2.15. *Suppose that A is a comultiplication S-act. If $(\theta :_A I) \subseteq (\theta :_A J)$ for two ideals I and J of S, then $JA \subseteq IA$.*

We have the duality of the pervious lemma for multiplication S-act, i.e, if A is multiplication S-act and $(IA \subseteq JA)$, then $(\theta :_A J) \subseteq (\theta :_A I)$.

Theorem 2.16. *Let A be a multiplication and comultiplication S-act. Then A is a fully cancellation S-act if and only if A is a fully cocancellation S-act.*

Proof. Suppose that A is a fully cancellation S-act and $(B :_A I) = (C :_A I)$, for subacts B, C of A and non-zero ideal I of S. Since A is comultiplication S-act, there exist ideals J and K of S such that $B = (\theta :_A J)$ and $C = (\theta :_A K)$. Therefore, $(\theta :_A JI) = (\theta :_A KI)$ and so by Lemma 2.15, we have $JIA = KIA$. Since A is fully cancellation S-act, we can conclude $JA = KA$. Now by duality of Lemma 2.15, we have $B = (\theta_A J) = (\theta :_A K) = C$. For conversely, suppose that $BI = CI$ for subacts B and C of A and ideal I of S. Since A is multiplication S-act, there exist ideals J and K of S so that $B = AJ$ and $C = AK$. Now by Lemma 2.15, we have $(\theta :_A JI) = (\theta :_A KI)$. Since A is a fully cocancellation S-act, we have $(\theta :_A J) = (\theta :_A K)$. Therefore by Lemma 2.15, we have $B = C$. □

References

[1] F. Farshadifar and H. Ansari. The dual notion of fully cancellation modules. *Journal of discrete mathematical Sciences and Cryptography*, pages 1–8, 2021.

[2] I. M. A. Hadi and A. A. Elewi. Fully cancellation and naturally cancellation modules. *Aeronautical Research Council Reports & Memoranda*, 17(3):178–184, 2018.

[3] U. Knauer M. Kilp and A. V. Mikhalev. *Monoids, Acts and Categories*. de Gruyter, Berlin, 2000.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AL–270, pp. 591–593

Comultiplication S-acts on monoids

Masoomeh Hezarjaribi[a,*], Zohreh Habibi[b]

[a]*Department of Mathematics, Payame Noor University, Tehran, Iran*
[b]*Department of Mathematics, Payame Noor University, Tehran, Iran*

Article Info

Keywords:
comultiplication
monoid
S-act

2020 MSC:
msc1
msc2

Abstract

In this paper, the notion comultiplication of monoid S-acts is studied. The behaviour of comultiplication is investigate. Also, it is shown that the product of a family of comultiplication S-acts is comultiplication.

1. Introduction and preliminaries

Acts over monoids, as universal algebras with unary operations, appear as basic mathematical models of some important notions in theoretical computer science and physics like automata, dynamical systems, etc. In [2] , Ansari-toroghy and Farshadifar study the notion comultiplication modules. In this research we study new notion comultiplication, in category of S-acts on monoids and investigate some properties of them.
First, we give some preliminary needed in the following.

Let S be a monoid. A (*right*) *S-act* is a non-empty set A together with a map $A \times S \to A, (a, s) \mapsto as$, such that for all $a \in A, s, t \in S$, $(as)t = a(st)$ and $a1 = a$. A non-empty subset $B \subseteq A$ is called a *subact* of A if $bs \in B$ for all $b \in B$ and $s \in S$. An element θ in an S-act A is said to be a *zero* or *fixed element* if $\theta s = \theta$ for all $s \in S$. Let A and B be two S-acts. A mapping $f : A \to B$ is called an *S-homomorphism* if $f(as) = f(a)s$, for all $a \in A, s \in S$. The category of all S-acts and homomorphisms between them is denoted by **Act**-S. We recall category $\mathbf{Act}_0$**-S** in which all monoids contain zero 0. In this category, monomorphisms are exactly injective homomorphisms.
A non-empty subset I of a monoid S is called a *right ideal* of S if $xs \in I$ for any $x \in I$ and $s \in S$.
An element $\theta \in A$ for which $\theta s = \theta$ for all $s \in S$ is said to be a *zero* or *fixed element* of A. A *congruence* on an S-act A is an equivalence relation ρ on A for which $a\rho a'$ implies that $(as)\rho(a's)$ for $a, a' \in A$ and $s \in S$.
A *left zero semigroup* is a semigroup S with $sr = s$ for all $s, r \in S$. A *right zero semigroup* is defined similarly. Throughout, S stands for a monoid unless otherwise stated. For undefined terms and notions about S-acts used here, we refer to [3]. In the all of this paper S is a monoid (with zero) and all S-acts are centered.

*Zohreh Habibi
Email addresses: Masoomeh.hezarjaribi@pnu.ac.ir (Masoomeh Hezarjaribi), Z_habibi@pnu.ac.ir (Zohreh Habibi)

2. Comultiplication S-act

In this section we define the notion of comultiplication S-act and study the behaviour of this notion with respect to product. Also, we results some properties of this notion.

We recall that an S-act A is said to be a multiplication S-act if for any subact B of A, there exists a two sided ideal I of S such that, $B = AI$. For more, see[1].

Definition 2.1. An S-act A is called comultiplication S-act , if for any subact B of A, there exists a two sided ideal I of S such that $B = (\theta :_A I)$.

Example 2.2. Consider $\mathbb{Z}$-act $\mathbb{Z}(p^\infty)$ on monoid $(\mathbb{Z}, .)$ with usual multiplication, which clearly is comultiplication S-act. But $\mathbb{Z}$-act $\mathbb{Z}$ is not comultiplication.

Remark 2.3. We recall that , in the category of S-acts, subacts of multiplication S-act is not necessarily multiplication and it is true for pure subacts of multiplication S-act are multiplication.
Now, we claim subact of comultiplication S-ct is comultiplication.
For this, let A be a comultiplication S-act and B be subact of A. Then, since A is comultiplication, there exists a two sided ideal $I$of S such that $B = (\theta :_A I)$. Now, we have $B = (\theta :_A I) = (\theta :_A I)$. So, B is a comultiplication S-act.

Remark 2.4. Consider homomorphism $f : A \longrightarrow B$. If A is comultiplication S-act then $f(A)$ is comultiplication. For this, consider subact C of $f(A)$. We show that there exists a two sided ideal I of S such that $C = (\theta_B :_{f(A)} I)$. Since A is comultiplication S-act, there is a two sided ideal I' of S so that $f^{-1}(C) = (\theta_A :_A I')$. So, we have $C = (\theta_B :_{f(A)} I')$.

We recall that in the **Act-**S, the product of a family of S-acts is their cartesian product with the componentwise action, i.e, if $(A_i)_{i\in I}$ is a family of S-acts, then the product $\prod_{i\in I} A_i$ is cartesian product of this family with projections morphisms $\rho_i : \prod_{i\in I} A_i \longrightarrow A_i$ by $\rho_j((x_i)_{i\in I}) = X_j, j \in I, (x_i)_{i\in I} \in \prod_{i\in I} A_i$.

Theorem 2.5. *Suppose that $(B_j)_{j\in J}$ is a family of S-acts. Then B_j is comultiplication, for any $j \in J$ if $\prod_{j\in J} B_j$ is comultiplication. The conversely for finite family of S-acts is correct.*

Proof. Suppose that $\prod_{j\in J} B_j$ is comultiplication S-act. We show that for any $j \in J$, B_i is a comultiplication S-act. Consider C_j as a subact of S-act B_j.
Let $C = (\theta_1, \theta_2, \cdots, \theta_{j-1}, c_j, \theta_{j+1}, \cdots) \leq \prod_{j\in J} B_i$. Since $\prod_{j\in J} B_i$ is comultiplication S-act, there exists a two sided ideal I of S such that $C = (\theta_{\prod_{j\in J} B_j} :_{\prod_{j\in J} B_j} I)$. So, we have $C_j I = \theta_j$. Hence, $C_j = (\theta_j :_{B_j} I)$. For conversely, consider subact $D = \{(d_1, d_2, d_3, \cdots, d_n) | d_j \in B_j\}$ of $\prod_{j\in J} B_i$. Clearly, for any $j \in J$, we have $\langle d_i \rangle = (\theta_j :_{B_j} J_j)$. Now, let $K = \{j_1 j_2 \cdots j_n | j_j \in J_j\}$. Clearly, J is an ideal of S and $C = (\theta_{\prod_{j\in J} B_j} :_{\prod_{j\in J} B_j} J)$. □

We recall that for any subact B of A, the notion annihilator of B is denoted by $Ann_S(B) = \{s \in S | Bs = \theta\}$.

Proposition 2.6. *Let A be an S-act . Then A is a comultiplication S-act if and only if for any subact B of A, $B = (\theta_A :_A Ann(B))$.*

Proof. Sufficieny is clear.
For conversely, Since A is a comultiplication S-act, there is a two sided ideal I of S such that $B = (\theta_A :_A I)$. So, $I \subseteq Ann(B)$ and we have $(\theta_A :_A Ann(B)) \subseteq (\theta_A :_A I) = B$. □

Proposition 2.7. *Let B be a comultiplication S-act and $\{C_j\}_{j\in J}$ be a family of subacts of S-act B so that $\bigcap_{j\in J} C_j = \theta$. For any subact D of A, we have $D = \bigcap_{j\in J}(D \cup C_j)$.*

Proposition 2.8. *Let A be an S-act. Then if A is a comultiplication prime S-act, the A is a simple S-act.*

Proposition 2.9. *Let A be an S-act. Then the following are equivalent.*
(i) *A is a comultiplication S-act.*
(ii) *For any subact B of A and any two sided ideal J of S with $B \subseteq (\theta :_A J)$, there exists a two sided ideal K of S such that $J \subset K$ and $B \subseteq (\theta :_A K)$.*

Lemma 2.10. *Let A be a comultiplication S-act. Let J be a minimal two sided ideal of S such that $(\theta :_A J) = \theta$, then A is cyclic.*

References

[1] A. A. Estaji and M. Shabani. Multiplication s-acts. In *38 th Annual Iranian Mathematics Conference*, Zanjan, 2017.
[2] H.Ansari-Toroghy and F. Farshadifar. The dual notion of multiplication modules. *Taiwanese. J. Math*, 11:1189–1201, 2007.
[3] U. Knauer M. Kilp and A. V. Mikhalev. *Monoids, Acts and Categories*. de Gruyter, Berlin, 2000.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, DE–304, pp. 594–599

New Soliton Solutions To The Shallow Water Waves With Coupled Time Fractional Boussinesq Equation

A. Sarli[a], A. Sharif[a,*], N. Taghinezhad[a]

[a] *Department of Mathematics, Faculty of Sciences, Gonbad Kavous University,Gonbad, Iran,*

Article Info

Keywords:
Soliton solutions
coupled fractional Boussinesq equation
three wave method

2020 MSC:
34B15
47E05

Abstract

Finding the simplest modeling for extracting the exact analytical solutions of nonlinear partial differential equations has become one of the most important topics in the field. In this study we obtain the soliton solutions of the fractional Boussinesq equation by using the three wave method. To show the accuracy of our results, we discuss some special cases by adjusting some potential parameters and also compute the graphical simulation of the results. Our results agree well with the results obtained via other methods.

1. Introduction

Exact or approximate solutions for the NLPDEs and Schrödinger equation play a fundamental role in many branches of physics and chemistry [1, 2]. Even though it was propounded about a century ago, the equation remains very challenging to solve analytically (even more so since the beginning of quantum mechanics). Theoretical physicists have strived to obtain exact or approximate solutions for the Schrödinger equation for various potentials of physical interest [3-5]. This is because the solution contains all the necessary information needed for a full description of a quantum state, including the probability density and entropy of the system [6-8]. Different investigations have been considered by researchers using this potential, some of which include the recent one by Durmus [9-10]. This importance has made the traces to such equations tangible in many branches of science, including mathematics, physics [1-3], electrical engineering, astronomy, mechanics, economics, and many other existing disciplines [4]-[6]. Based on these remarkable effects, several analytical methods have been successfully applied to obtain exact solutions of such equations. Some of these methods are homotopy analysis method [7], the variational iteration method [8], the exp-function method [9], Logistic function method [10], the generalized G'/G-expansion [11], the elliptic finder method [12]-[14], the exponential rational function idea [15], the modified Kudryashov technique [16] and sub-equation method [17]. To see more methods, please refer to [18]-[20]. These authors calculated the normalized radial wave functions and obtained an analytical formula of phase shifts. They also investigated the corresponding bound states by studying the analytical properties of the scattering amplitude.

*Talker
Email address: sharif@gonbad.ac.ir (A. Sharif)

A brief description of fractional coupled Boussinesq equation is provided in the second section of this paper. In section three application of the method[21] to the FBE and graphical behavior of solutions introduced. Finally, conclusions are presented in the last section of the article.

2. Basic structure of the fractional coupled Boussinesq equation

we consider the following time fractional coupled Boussinesq equation (BE)

$$\begin{cases} \frac{\partial^\alpha u}{\partial t^\alpha} + uu_x + v_x + qu_{xx} = 0 \\ \frac{\partial^\alpha v}{\partial t^\alpha} + (uv)_x + pu_{xxx} - qv_{xx} = 0 \end{cases} ,0 \leq \alpha < 1. \tag{1}$$

Where$p, q \in R$. Using the transformation$u\left(x,t\right) = U\left(\xi\right), v\left(x,t\right) = V\left(\xi\right)$; where $\xi = kx + \omega\frac{t^\alpha}{\alpha}$and once integrating respect toξ, Eq. (1) becomes an following ordinary differential equation,

$$\begin{array}{l} \omega U + \frac{k}{2}U^2 + kV + qk^2U' = R_1, \\ \omega V + kUV + pk^3U'' - qk^2V' = R_2, \end{array} \tag{2}$$

Where $R_1$and R_2 are the integration constants of first- and second-equation of system (2), respectively. From first-equation of system (2), we get

$$V = \frac{1}{k}\left(R_1 - \omega U - \frac{k}{2}U^2 - qkU'\right), \tag{3}$$

By substituting Eq. (3) into the second-equation of system (2), and for simplifying we set $R_1 = 0$and$R_2 = 0$, we get the following covering equation

$$-\frac{\omega^2}{k}U - \frac{3}{2}\omega U^2 - \frac{k}{2}U^3 + k^3\left(p + q^2\right)U'' = 0, \tag{4}$$

A brief application of the method to the FBE is provided in the second section of this paper. In section three graphical behavior of solutions introduced. Finally, conclusions are presented in the last section of the article.

3. Three wave approache to the FBE

Primis, we suppose that Eq. (**??**) has the following three-wave solutions

$$U\left(\xi\right) = \gamma_1 e^{\delta\xi} + \gamma_2\cos\left(\lambda_1\xi\right) + \gamma_3 e^{-\delta\xi} + 2\gamma_4\cosh\left(\lambda_2\xi\right). \tag{5}$$

Where $\gamma_1, .., \gamma_4, \delta, \lambda_1, \lambda_2$are unfamiliar constants to be determined later. With substituting (5) into (4) and collect coefficients of $e^{i\delta\xi}, \cos\left(\lambda_1\xi\right), \cosh\left(\lambda_2\xi\right), \sin\left(\lambda_1\xi\right), \sinh\left(\lambda_2\xi\right), i = -2, -1, 0, 1, 2$ and let them equal to zero. So we obtain the algebraic equations and by solving these equations we have:

Set 1: $\gamma_1 = 0, \gamma_3 = 0, \gamma_2 \neq 0, \gamma_4 \neq 0$ then by solving algebraic equation we have

$$\gamma_2 = \frac{1}{2}, \gamma_4 = -\frac{1}{2}, w = -\frac{1}{2}\sqrt{q^2 + pk^2}, k = k \tag{6}$$

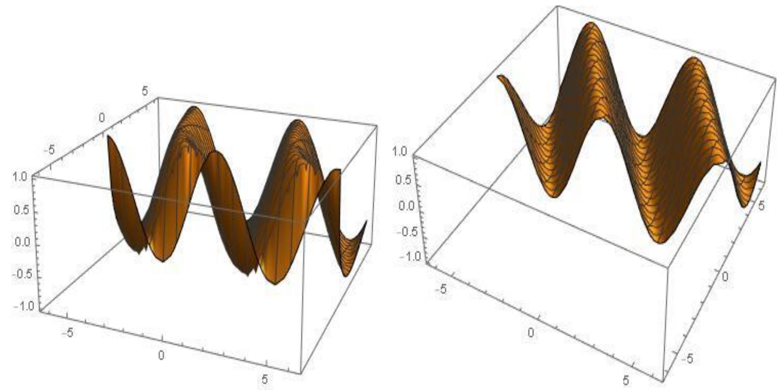
So we have general solutions of eq. (1) as follows

$$u_1\left(x,t\right) = \frac{1}{2}\cos\left(\lambda_1\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right) - \cosh\left(\lambda_2\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right)$$

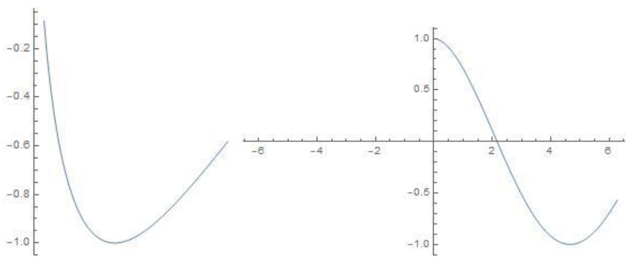
So from (3) we directly obtain

$$\begin{array}{l} v_1\left(x,t\right) = \left[R_1 - \frac{w}{2}\cos\left(\lambda_1\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right) + w\cosh\left(\lambda_2\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right) - \right. \\ \frac{k}{2}\left(\frac{1}{2}\cos\left(\lambda_1\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right) - \cosh\left(\lambda_2\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right)\right)^2 + \\ \left. \frac{1}{2}qk\sin\left(\lambda_1\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right) + qk\sinh\left(\lambda_2\left(kx - \frac{1}{2}\sqrt{q^2 + pk^2}\frac{t^\alpha}{\alpha}\right)\right)\right] \end{array}$$

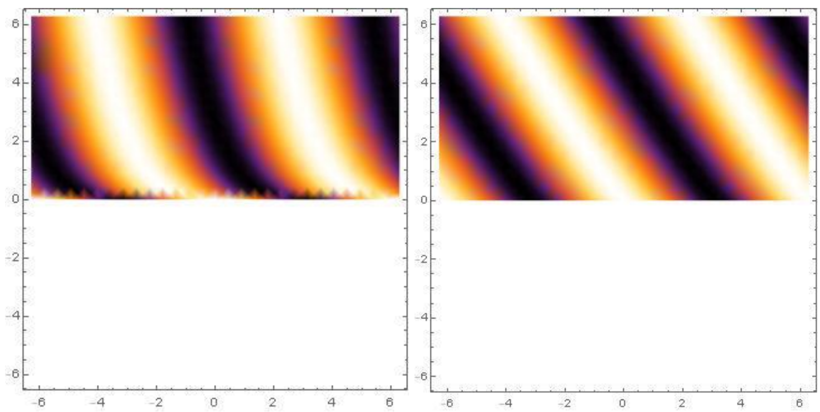
Dynamical Analysis For Real Part



3D-Plot for ($\alpha = 0.3$) 3D-Plot for ($\alpha = 0.9$)



2D-Plot for ($\alpha = 0.3$) 2D-Plot for ($\alpha = 0.9$)



Density Plot for ($\alpha = 0.3$) Density Plot for ($\alpha = 0.9$)

Fig. 1: geraphical behavior of $u_1(x,t)$for , $\alpha = 0.3$ (A,C and E) and $\alpha = 0.9$ (B,D and F)

Set 2: $\gamma_2 = 0, \gamma_4 = 0, \gamma_1 \neq 0, \gamma_3 \neq 0$ then by solving algebraic equation we have

$$\gamma_1 = \frac{2}{3}\frac{\delta^2 k^4 q^2 + \delta^2 k^4 p - w^2}{k^2 \gamma_3}, \gamma_3 = \gamma_3, w = \sqrt{q^2+p}k^2\delta, k = k \tag{7}$$

So we have

$$u_2\left(x,t\right) = \frac{2}{3}\frac{\delta^2 k^4 q^2 + \delta^2 k^4 p - w^2}{k^2 \gamma_3} e^{\delta\left(kx + \sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)} + \gamma_3 e^{-\delta\left(kx + \sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}$$

From (3) we have

$$
\begin{aligned}
&v_2(x,t)=\tfrac{2}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^3\gamma_3}e^{\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}+\tfrac{\gamma_3}{k}e^{-\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}-\\
&-\tfrac{1}{2}\left(\tfrac{2}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2\gamma_3}e^{\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}+\gamma_3e^{-\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}\right)^2-\\
&q\left(\tfrac{2}{3}\delta\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2\gamma_3}e^{\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}-\gamma_3\delta e^{-\delta\left(kx+\sqrt{q^2+p}k^2\delta\frac{t^\alpha}{\alpha}\right)}\right)
\end{aligned}
$$

Set 3: $\gamma_1=0,\gamma_2=0,\gamma_3\neq0,\gamma_4\neq0$ then by solving algebraic equation we have

$$
\gamma_4=\frac{2}{3}\frac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{kw},\gamma_3=\gamma_3,w=w,k=\frac{\sqrt{\sqrt{q^2+p}\delta w}}{\sqrt{q^2+p}\delta} \tag{8}
$$

So

$$
u_3(x,t)=\gamma_3e^{-\delta\left(kx+\omega\frac{t^\alpha}{\alpha}\right)}+\frac{4}{3}\frac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{kw}\cosh\left(\lambda_2\left(kx+\omega\frac{t^\alpha}{\alpha}\right)\right)
$$

and

$$
\begin{aligned}
&v_3(x,t)=\tfrac{1}{k}\left(R_1-\omega U-\tfrac{k}{2}U^2-qkU'\right),\\
&\tfrac{R_1}{k}-\tfrac{\gamma_3}{k}e^{-\delta\left(kx+\omega\frac{t^\alpha}{\alpha}\right)}-\tfrac{4}{3}\tfrac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{k^2w}\cosh\left(\lambda_2\left(kx+\omega\tfrac{t^\alpha}{\alpha}\right)\right)-\\
&w\gamma_3e^{-\delta\left(kx+\omega\frac{t^\alpha}{\alpha}\right)}-\tfrac{4}{3}\tfrac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{k}\cosh\left(\lambda_2\left(kx+\omega\tfrac{t^\alpha}{\alpha}\right)\right)-\\
&\tfrac{1}{2}\left(\gamma_3e^{-\delta\left(kx+\omega\frac{t^\alpha}{\alpha}\right)}+\tfrac{4}{3}\tfrac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{kw}\cosh\left(\lambda_2\left(kx+\omega\tfrac{t^\alpha}{\alpha}\right)\right)\right)^2+\\
&\tfrac{\gamma_3}{k}\delta e^{-\delta\left(kx+\omega\frac{t^\alpha}{\alpha}\right)}-\tfrac{4\lambda_2}{3}\tfrac{\lambda_2^2k^4q^2+\lambda_2^2k^4p-w^2}{k^2w}\sinh\left(\lambda_2\left(kx+\omega\tfrac{t^\alpha}{\alpha}\right)\right)
\end{aligned}
$$

Set 4: $\gamma_1\neq0,\gamma_2=0,\gamma_3\neq0,\gamma_4\neq0$ then by solving algebraic equation we have

$$
\begin{aligned}
&\gamma_4=\tfrac{4}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2},\gamma_3=\tfrac{1}{2},\gamma_1=\gamma_1,w=w,\\
&k=\tfrac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}},
\end{aligned} \tag{9}
$$

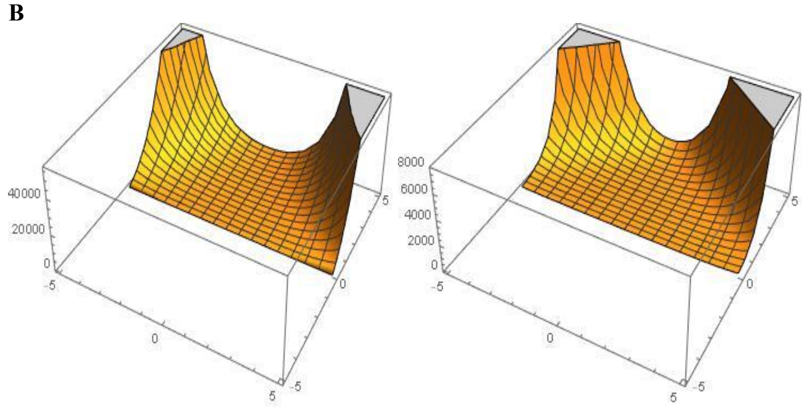
$$
\begin{aligned}
&u_4(x,t)=\gamma_1e^{\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}+\tfrac{1}{2}e^{-\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}+\\
&\tfrac{8}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2}\cosh\left(\lambda_2\left(\tfrac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\tfrac{t^\alpha}{\alpha}\right)\right)
\end{aligned}
$$

Now from (3) we have

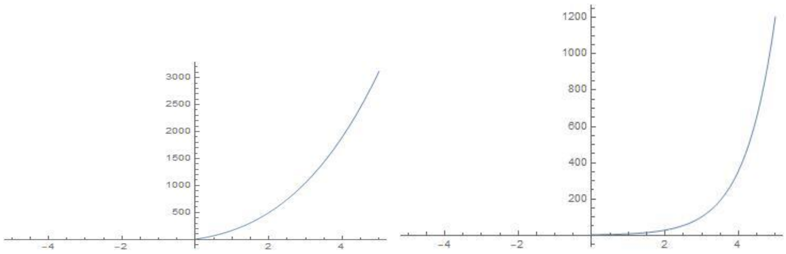
$$
\begin{aligned}
&v_4(x,t)=\tfrac{1}{k}R_1-\tfrac{1}{k}\gamma_1we^{\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}-\tfrac{1}{2k}we^{-\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}-\\
&\tfrac{8w}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^3}\cosh\left(\lambda_2\left(\tfrac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\tfrac{t^\alpha}{\alpha}\right)\right)-\\
&\tfrac{1}{2}\left[\gamma_1e^{\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}+\tfrac{1}{2}e^{-\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}+\right.\\
&\left.\tfrac{8}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2}\cosh\left(\lambda_2\left(\tfrac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\tfrac{t^\alpha}{\alpha}\right)\right)\right]^2-\\
&\gamma_1\delta qe^{\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}+\tfrac{q\delta}{2}e^{-\delta\left(\frac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\frac{t^\alpha}{\alpha}\right)}-\\
&\tfrac{8q\lambda_2}{3}\tfrac{\delta^2k^4q^2+\delta^2k^4p-w^2}{k^2}\cosh\left(\lambda_2\left(\tfrac{\sqrt{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2}w}}{\sqrt{2\delta^2q^2-q^2\lambda_2^2+2\delta^2p-p\lambda_2^2w}}x+w\tfrac{t^\alpha}{\alpha}\right)\right)
\end{aligned}
$$

Dynamical Analysis For Real Part

B

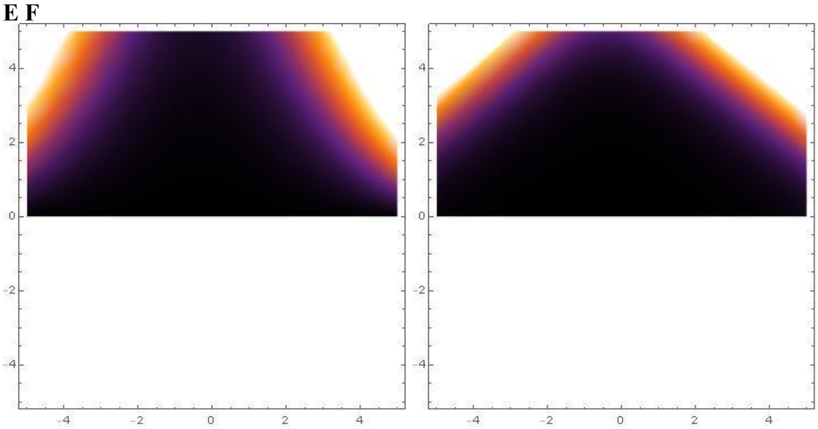


3D-Plot for **(**$\alpha = 0.3$**)** 3D-Plot for **(**$\alpha = 0.9$**)**



2D-Plot for **(**$\alpha = 0.3$**)** 2D-Plot for **(**$\alpha = 0.9$**)**

E F



Density Plot for ($\alpha = 0.3$) Density Plot for ($\alpha = 0.9$)

Fiq. 2: geraphical behavior of $u_3(x,t)$for , $\alpha = 0.3$ (A,C and E) and $\alpha = 0.6$ (B,D and F)

4. Concluding remarks

In this article, we solved the coupled fractional Boussinesq equation using three wave method. We also presented the graphical behaviour of the solutions. It was found that our results agree with the existing literature. Finally, this study has many applications in different areas of physics and chemistry, such as atomic physics, molecular physics, and chemistry. The structure considered for the equation consists of a series of arbitrary parameters that lead to many well-known models by considering certain options for them. One of the main advantages of this method is the determination of different categories of solutions for the equation in a single framework; This means that the method can determine different types of solutions for the equation in a single process.

References

[1] J.G. Liu, W.H. Zhu, , Zhou, L., Y.K. Xiong, Multi-waves, breather wave and lump-stripe interaction solutions in a (2+1)-dimensional variable-coefficient Korteweg–de Vries equation. Nonlinear Dyn. 97, 2127–2134 (2019)

[2] J.G.Liu, , He, Y.: Abundant lump and lump-kink solutions for the new (3+1)-dimensional generalized Kadomtsev–Petviashvili equation. Nonlinear Dyn. 92(3), 1103–1108 (2018)

[3] J.G. Liu, M. Eslami, , H., Rezazadeh, Mirzazadeh, M.: Rational solutions and lump solutions to a non-isospectral and generalized variable-coefficient Kadomtsev–Petviashvili equation. Nonlinear Dyn. 95, 1027–1033 (2019).

[4] S. Kumar, Kumar, A.: Lie symmetry reductions and group invariant solutions of (2+1)-dimensional modified Veronese web equation. Nonlinear Dyn. 98(3), 1891–1903 (2019)

[5] Y.L. Ma, Interaction and energy transition between the breather and rogue wave for a generalized nonlinear Schrödinger system with two higher-order dispersion operators in optical fibers. Nonlinear Dyn. 97, 95–105 (2019)

[6] Zhao, D., Zhaqilao: On two new types of modified short pulse equation. Nonlinear Dyn. 100, 615–627 (2020)

[7] S.J. Liano, The Proposed Homotophy Analysis Technique for the Solution of Nonlinear Problems (Ph.D. thesis), Shanghai Jiao Tong University, (1992).

[8] A.M. Wazwaz, The Variational iteration method for solving linear and nonlinear system of PDEs, Comput. Math. Appl. 54 895-902 (2007).

[9] J.H. He, X.H. Wu, Exp-function method for nonlinear wave equations, Chaos Solitons Fractals 30 700-708 (2006).

[10] N.A. Kudryashov, P. N. Ryabov, Analytical and numerical solutions of the generalized dispersive Swift Hohenberg equation, appl. Math. Comput. 286 171-177(2016).

[11] J. Manafian, M. Fazli Aghdaei, M. Khalilian, R. S. Jeddi, Application of the generalized G/G-expansion method for nonlinear PDEs to obtaining soliton wave solution, Optik 135 395-406 (2017).

[12] Biswas A, Ekici M, Sonmezoglu A, Alshomrani AS, Belic MR. Optical solitons with Kudryashov's equation by extended trial function. Optik;202:163290(2020).

[13] Sajid N, Akram G. Novel solutions of Biswas-Arshed equation by newly -model expansion method. Optik. 1;211:164564(2020). 6

[14] Ghanbari B, Kumar S, Niwas M, Baleanu D. The Lie symmetry analysis and exact Jacobi elliptic solutions for the Kawahara-KdV type equations." Results in Physics: 104006(2021).

[15] K. Hosseini, Z. Ayati, R. Ansari, New exact traveling wave solutions of the TzitzÃ©ica type equations using a novel exponential rational function method, Optik 148- 85-89(2017).

[16] Ege SM, Misirli E. The modified Kudryashov method for solving some fractional-order nonlinear equations. Advances in Difference Equations. 2014 ; (1):1-3(2014).

[17] H Rezazadeh, A Neirameh, M Eslami, A Bekir, A Korkmaz, A sub-equation method for solving the cubicâquartic NLSE with the Kerr law nonlinearity, Modern Physics Letters B 33 (**??**), 1950197(2019).

[18] H Rezazadeh, M Mirzazadeh, SM Mirhosseini-Alizamini, A Neirameh , Optical solitons of Lakshmanan Porsezianâ Daniel model with a couple of nonlinearities, Optik 164, 414-423(2019).

[19] H Rezazadeh, D Kumar, A Neirameh, M Eslami, M Mirzazadeh, Applications of three methods for obtaining optical soliton solutions for the Lakshmanan Porsezian Daniel model with Kerr law nonlinearity, Pramana 94 (1), 1-11(2018).

[20] Z, Zhao, Z, Dai, C, Wang, Extend three-wave method for the (1+2)-dimensional Ito equation, Applied Mathematics and Computation, 2010, 217, n5 pp 2295-2300

[21] T A Sulaiman, H Bulut, A Yokus & H M Baskonus, On the exact and numerical solutions to the coupled Boussinesq equation arising in ocean engineering, Indian Journal of Physics volume 93, pages647–656 (2019)

[22] J.L. Bona, V.A. Dougalis and D.E. Mitsotakis, Numerical solution of Boussinesq systems of KdV-KdV type: II. Evolution of radiating solitary waves, Nonlinearity 21 (2008) 2825- 2848.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, OR–217, pp. 600–607

Joint Propagation Delay and Load Balancing Aware Controller Mapping for Distributed SDN Control Planes

Ahmad Jalili[a]

[a] *Department of Computer Engineering, Faculty of Basic Sciences and Engineering, Gonbad Kavous University, Gonbad Kavous, Iran*

Article Info

Keywords:
Facility Location Problem (FLP), Heuristic Algorithms, Multi-Objective Combinatorial Optimization (MOCO), Software Defined Networking (SDN).

Abstract

Software-defined networking (SDN) enables flexible network management by decoupling the control plane from the data forwarding plane. However, as network scale and traffic demands grow, effective multi-controller deployment becomes crucial for improving control plane scalability and reliability. Existing schemes for controller placement mainly focus on minimizing propagation delays or maximizing resilience, failing to adequately consider controller load balancing requirements. In this paper, we propose a dynamic load balancing scheme for multi-controller SDN deployment based on traffic propagation delays and controller processing capacities. We formulate an optimization problem to jointly minimize intra-domain and inter-domain communication costs subject to delay and capacity constraints. We model the switch-controller interactions as M/M/1 queues and use propagation clustering algorithm to determine the optimal number of controllers and their mapping to switches. Extensive simulations on Internet OS3E topologies demonstrate that our scheme achieves superior balanced multi-controller deployment compared to affinity propagation and genetic algorithms, while maintaining low control plane latencies under fluctuating traffic conditions through dynamic load balancing across controllers.

1. Introduction

The software-defined networking (SDN) paradigm has emerged as a promising architecture to revolutionize traditional network infrastructure. By decoupling the control plane that governs network intelligence from the underlying data forwarding plane, SDN facilitates flexible programmability, simplified management, and fostered innovation through logically centralized control [1]. However, the logically centralized control plane designed around a single controller raises significant concerns regarding resilience, scalability, and reliability as networks continue to expand in scale and traffic demands [2]. Consequently, multi-controller deployment has become an integral part of effective SDN control plane design for large-scale production networks.

Deploying multiple distributed controllers partitions the control plane into distinct control domains, with each controller managing a subset of network switches. This distribution of intelligent control functionalities aims to enhance network resilience through elimination of single points of failure [3]. Furthermore, multi-controller architectures alleviate the load on any individual controller, thereby improving control plane scalability and responsiveness. However,

*Talker
Email address: jalili@gonbad.ac.ir (Ahmad Jalili)

the performance benefits of multi-controller SDN critically depend on the strategic placement of controllers and the mapping of switches to respective control domains.

Inappropriate controller placement resulting in suboptimal mappings between switches and controllers can lead to excessive propagation delays, load imbalances across controllers, and increased communication overhead between control domains for state synchronization [4]. Therefore, substantial research efforts have been dedicated to developing efficient algorithms for optimized multi-controller deployment. Existing schemes have predominantly focused on minimizing worst-case or average propagation latencies between switches and their assigned controllers. Other work has aimed to maximize resilience through backup controller mapping or ensuring path diversity. However, these schemes largely overlook the load balancing requirements across controllers, which is a pivotal consideration for scalable SDN control planes handling dynamic traffic patterns.

Suboptimal load distribution across controllers can severely degrade control plane performance. An overloaded controller processing an excessive number of flow setup requests experiences increased queuing delays, resulting in higher latencies for path establishment and flow rule installation. This not only impacts overall network throughput but can also lead to transient loops or packet losses during the delay window before new flows are programmed [5]. Conversely, underutilized controllers represent inefficient use of provisioned resources. Therefore, balanced load distribution is critical for responsive and efficient SDN operations.

Furthermore, existing approaches predominantly focus on optimizing static controller deployments based on predicted traffic patterns. However, the increasingly dynamic nature of modern networks, with fluctuating demands and evolving application requirements, necessitates adaptive load balancing schemes that can dynamically adjust control domains based on the current network state. Static controller mappings inevitably lead to load imbalances under shifts in spatio-temporal traffic distribution. Dynamically reassigning switch-to-controller mappings becomes essential to ensure load distribution equilibrium and consistent control plane performance.

In this context, we present a novel load balancing scheme for multi-controller SDN deployment that holistically accounts for traffic propagation latencies and controller processing capacities under dynamic network conditions. We formulate an optimization problem to minimize the combined communication costs of intra-domain interactions between switches and controllers, as well as inter-domain coordination overhead between controllers. Our objective is subject to crucial constraints on maximum tolerable propagation delays and controller capacity limits. Specifically, our major contributions are as follows:

1) We develop a comprehensive model that jointly considers traffic propagation delays and controller processing capacities as pivotal factors influencing the multi-controller deployment problem. We transform the flow setup requests into an M/M/1 queuing model to quantify the experienced latencies at each controller.

2) For the initial static deployment phase, we propose a modified affinity propagation clustering algorithm (PSOAP) that leverages particle swarm optimization to intelligently adjust the clustering parameters and determine the appropriate number of control domains.

The rest of the paper is organized as follow. Related works of the CPP are presented in following section. In Section 3, Model and Formulation are presented. Algorithm design and a thorough presentation of its operators are brought in Section 4. Section 5 explores the simulation and evaluation. Conclusion and future works are explained in final section.

2. Related Works

The placement of SDN controllers is an essential consideration in optimizing network performance. Several papers propose different approaches to address this issue. Jeya et al. introduce the use of SDN architecture in smart grids and propose a method to find the optimal solution for controller placement in a smart grid using SDN architecture [6]. The paper discusses the challenges of placing controllers in a smart grid implementing SDN architecture, but it does not specifically mention the use of latency and other metrics for controller placement.

Authors in [7] propose a deep reinforcement learning-based model to dynamically adjust the controller placement in a virtualized environment to minimize OpenFlow latency. In this article , a novel deep reinforcement learning model is introduced to dynamically adjust the controller's position, aiming to minimize delay within a virtualized environment. Outcomes reveal that the introduced method outperforms both random and generic strategies. The primary objective of the model is to enhance performance in a virtualized setting, but it did not explore other latencies metrics.

In [8], a thorough investigation of the MCPP in Software-Defined Networking is conducted, focusing on aspects such as load balancing, reliability, and dynamic techniques. The study encompasses a categorization of these techniques, simplifying researchers' grasp of their underlying principles. The paper contributes to addressing the network management issues faced by traditional networking in the context of increasing Internet usage and evolving technologies. The authors in [9] propose a linear programming method to attain a trade-off between setting-up cost and delay in the placement of SDN controllers in large scale networks. The introduced method specifies the location, and the minimum number of demanded controllers based on the topology, processing capacity of controllers, and setup cost. The result shows that proposed method attains the lowest setting-up cost and average delay of controlling traffic compared to previous methods. It also identifies the switches associated with each suggested controller in actual network topologies. However, the study did not consider other factors such as reliability, energy consumption, or load balancing.

3. Model and Formulation

In this section, we explore the issue of controller load imbalance during the deployment of multiple controllers in a distributed network. We establish a mathematical model that encompasses various performance parameters influencing the controller deployment process. We develop a comprehensive model for the multi-controller deployment problem in software-defined networks that jointly considers traffic propagation latencies and controller processing capacities. Our objective is to determine an optimized mapping of switches to controllers that minimizes the combined communication overhead while ensuring delay guarantees and balanced load distribution across controllers.

3.1. Network Model

We model the SDN network topology as an undirected graph*G = (V, E)*, where*V*represents the set of nodes comprising switches and controllers, and*E*denotes the set of bidirectional links interconnecting the nodes. Let*M*be the number of controllers in the network with the set of controllers denoted as*C = {C1, C2, ..., CM}*. Similarly, we define N as the number of switches and*S = {S1, S2, ..., SN}*as the set of switches in the network. Therefore, the total number of nodes is|*V*| = *M* + *N*.

We assume the traffic follows a dynamic model whereλ_i^trepresents the flow setup request rate generated by the ith switch*Si*during the time slot*t*. The variable*dij*denotes the shortest path distance between switch*Si*and controller*Cj*, calculated as the sum of the weights of the links traversed along the minimum cost path. We define a binary variablex_{ij}^tsuch that$x_{ij}^t = 1$if the*ith*switch is assigned to the jth controller during time t, and$x_{ij}^t = 0$otherwise.

3.2. Controller Capacity Model

The capacity of each controller, representing its ability to process incoming flow setup requests, is determined by factors such as CPU performance, memory, and network bandwidth. We denote the processing capacity of the*jth*controller as*Aj*. However, controllers must also reserve a fraction of their resources for critical control plane operations such as state synchronization with other controllers for resilience. We account for this overhead by introducing a redundancy factor*LM ϵ [0, 1]*that is common across all controllers.

To characterize the load on each controller, we model the switch-controller interactions as an M/M/1 queuing system, which is an appropriate representation for independent and memoryless arrival of flow setup requests [10] The total flow request rate that the*jth*controller needs to process during time t is given by:

$$\phi_j^t = \sum\nolimits_{i=1}^{N} \lambda_i^t x_{ij}^t$$

Assuming the flow request processing times at each controller follow an exponential distribution with a mean of 1/Aj, we can apply queuing theory to derive the average sojourn time ωtj experienced by a flow request at the jth controller:

$$\omega_j^t = \frac{1}{A_j - \phi_j^t}$$

Furthermore, considering the compute time for route computation is proportional to the network size |V|, we can express the average response time Δt_j for the jth controller as:

$$\Delta t_j = \omega_j^t . |v|^2$$

3.3. Communication Cost Model

The communication cost for the multi-controller deployment comprises two key components: the intra-domain cost arising from interactions between switches and their assigned controllers, and the inter-domain cost due to coordination between controllers for state synchronization.
1) Intra-domain Communication Cost: When a new flow arrives at a switch, it generates a packet-in message to its managing controller. The controller computes the routing path, installs the corresponding flow rules on the switches along the path, and issues packet-out messages.
2) Inter-domain Communication Cost: In multi-controller SDN deployments, controllers synchronize network state information among themselves to maintain a consistent global view.

3.4. Optimization Problem Formulation

Our goal is to determine an optimized mapping of switches to controllers that minimizes the total communication cost, while ensuring the maximum propagation delay between any switch-controller pair remains below an acceptable threshold δ, and no controller is overloaded beyond its redundancy-adjusted capacity. Formally, we define the following optimization problem:

$$\begin{aligned} \min \mathit{Total} = \gamma D_{req} + (1-\gamma)\, D_{syn} \\ s.t. \sum_{j\in M} x_{ij}^{t} = 1 \\ \Phi\,(t) \le L_j A_j \\ D\,(t) \le \delta \\ x_{ij}^{t} \in \{0,1\} \end{aligned}$$

The weight $\gamma \epsilon$ [0, 1] allows adjusting the tradeoff between prioritizing intra-domain vs. inter-domain communication costs. The constraints ensure each switch is mapped to exactly one controller, no controller load exceeds its redundancy-adjusted capacity, and the overall network-wide average delay is below the threshold δ.
This formulation coherently captures the crucial factors impacting multi-controller SDN deployments, including traffic dynamics, controller capacities, propagation latencies, and communication overheads. However, solving this NP-hard problem exactly is computationally intractable for large networks. Therefore, we propose efficient heuristic algorithms.

4. Algorithm Design

We propose two complementary algorithms to address the multi-controller deployment problem in SDNs - one for the initial static scenario and another for dynamic traffic-aware load balancing. For the former, we present PSOAP, a modified affinity propagation clustering algorithm that leverages particle swarm optimization to determine the appropriate number of control domains and optimize the switch-to-controller mapping.

4.1. PSOAP: Particle Swarm Optimization based Affinity Propagation

Affinity Propagation (AP) is a well-established clustering algorithm that identifies representative exemplars among the data points themselves, rather than requiring the number of clusters to be pre-specified [11]. AP operates by exchanging real-valued responsibility and availability messages between data points until a set of exemplars and corresponding clusters gradually emerges.
Let*S = {s1, s2, ..., sN}*represent the set of switches in the network, treated as the data points for clustering. The similarity*s(i,j)*between switches*si*and*sj*is defined based on the communication cost$\lambda\,(t)_i\, d\,(i,j)\, x_{ij}$, which captures the traffic rateλ*ti*generated at*si*and the shortest path distance*dij*to the potential controller at*sj*. The higher this cost, the more dissimilar the switches.
AP aims to maximize the net similarity summed over all data points, subject to self-similarities*s(i,i) = p*serving as preferable clustering biases. Initially, availabilities*a(i,j)*are set to zero, while responsibilities*r(i,j)*measure how well sj

serves as an exemplar for*si*, compared to other candidates. Messages are iteratively updated as:

$$R\left(i,j\right)=S\left(i,j\right)-max\left\{A\left(i,j^{`}\right)+S(i,j^{`})\right\}$$
$$A\left(i,j\right)=\min\left(0,R\left(j,j\right)\right)+\sum max\left\{0,R(i^{`},j)\right\}$$

These ”responsibilities” (1) and ”availabilities” (2) are combined into a single matrix indicating the clustering affiliations. However, AP’s convergence is sensitive to the preference values p and the damping factor λ that prevents numerical oscillations during updates.

Our PSOAP algorithm treats p and λ as particle positions in a multi-dimensional search space, which are iteratively tuned via particle swarm optimization (PSO) [12] to intelligently identify the near-optimal preferences and damping leading to superior clustering quality. Each particle maintains a velocity v that governs its movement, which gets updated based on the particle’s previous best position pbest and the global best gbest across all particles as:

$$V_{id}=\omega V_{id}+\eta_1 rand()\left(P_{id}-X_{id}\right)+\eta_2 rand()(P_{gd}-X_{id})$$

Here, ω is the inertia weight controlling exploration versus exploitation, while $\eta 1$ and $\eta 2$ are the cognitive and social learning rates. r1 and r2 are uniformly random numbers in [0,1].

The fitness of each particle position (p, λ) is evaluated by executing the AP clustering procedure using those preference and damping parameters, then computing the overall communication cost as per the objective function defined in Section 2. PSOAP continually explores the search space, identifying the (p, λ) parameters minimizing this cost while adhering to the delay and capacity constraints outlined earlier. This yields the optimal number of control domains as well as the switch-to-controller mapping by treating the identified cluster exemplars as controller locations.

Algorithm 1 summarizes the steps of PSOAP, integrating PSO and AP until convergence. In each PSO iteration, AP is executed for every particle position using the specified (p, λ) preferences and damping. Particle fitnesses are evaluated based on the resulting communication costs, and personal/global bests are updated. Velocities are then calculated as per (3) to adjust particle positions for the next iteration until the desired solution quality is attained or the maximum number of iterations is reached.

Here is the pseudocode for the Particle Swarm Optimization based Affinity Propagation (PSOAP) algorithm:

Input: G(V,E) - Network graph with switches V and links E, λ - Traffic rates of switches, A - Controller processing capacities, L - Controller redundancy factors

Output: X - Switch-to-controller mapping matrix

Initialize particle swarm with random positions (p, λ) and velocities
while termination condition not met do
for each particle (p, λ) do
// Run Affinity Propagation (AP) with current p, λ
Compute similarity matrix S based on λ, link costs
Initialize responsibilities R and availabilities A
repeat
Update R using equation (1)
Update A using equation (2)
until convergence
Get exemplar switches as controller locations from R, A
Obtain switch-controller mapping X
Evaluate particle fitness = Communication Cost(X)
end for
Update personal best positions pbest
Update global best position gbest
// Update particle positions and velocities
for each particle (p, λ) do
Calculate new velocity v using equation (3)
Update particle position (p, λ) using equation (4)
end for

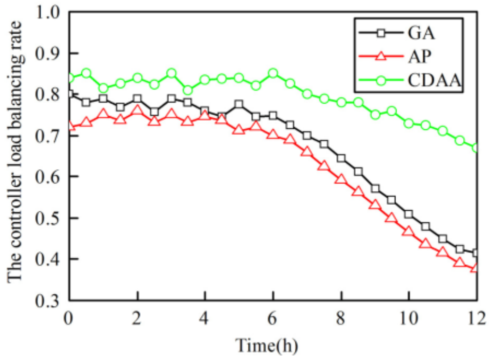


Fig. 1. Load balancing rate

22: end while
23: Return X with minimum Communication Cost
The algorithm takes the network graph, switch traffic rates, controller capacities and redundancy factors as inputs (line 1). It initializes a swarm of particles, where each particle represents a candidate solution with preference p and damping λ parameters for Affinity Propagation (AP) clustering (line 1). The main loop (lines 2-22) runs until a termination condition is met, e.g., maximum iterations or convergence. For each particle (p, λ) (line 3): It runs the Affinity Propagation clustering procedure by first computing the similarity matrix S based on traffic and link costs (line 5). It initializes the responsibility R and availability A matrices (line 6). It iteratively updates R and A using the standard AP equations (1) and (2) until convergence (lines 7-10). Finally, it returns the switch-controller mapping X with the minimum communication cost (line 23).

5. Simulation and Evaluation

To comprehensively evaluate the efficacy of our proposed multi-controller deployment scheme, we conduct extensive simulations across diverse scenarios and benchmark against state-of-the-art approaches. This section delineates the simulation environment, parameter configurations, and quantitative performance analyses carried out to validate our algorithms.
Network Topologies: We evaluate our algorithms on two topologies representing contrasting network scales and characteristics:
1) Internet OS3E Topology [12]: This topology models a regional IP/MPLS backbone network comprising 34 nodes and 42 bidirectional links spanning the western United States. Its moderate size allows assessing our schemes across a wide range of scenarios.
2) GScaleBB Topology [13]: To analyze the performance on large-scale networks, we use this national backbone topology consisting of 774 nodes and 1016 links distributed across the United States. Its substantial size poses significant computational challenges for controller placement algorithms.
Controller Implementation: Our simulations incorporate a real SDN controller platform by integrating the widely-adopted OpenDaylight controller and instrumenting its application layer with custom modules implementing our PSOAP and CDAA algorithms.
Benchmarks: We compare the performance of our PSOAP and CDAA algorithms against the following state-of-the-art multi-controller deployment schemes:
1) Affinity Propagation (AP): The standard AP clustering algorithm that minimizes propagation delays.
2) Genetic Algorithm (GA): A heuristic evolutionary algorithm for the multi-objective controller placement problem considering propagation delays and load balancing.
3) Static Mapping (SM): A baseline approach that statically maps switches to their closest controllers without any dynamic remapping, representing traditional SDN with a fixed control plane partitioning.
For all simulations, we implement each algorithm in Java and execute them on an Intel Xeon E5-2680 v2 server with 128GB RAM, averaging results over 20 independent runs.

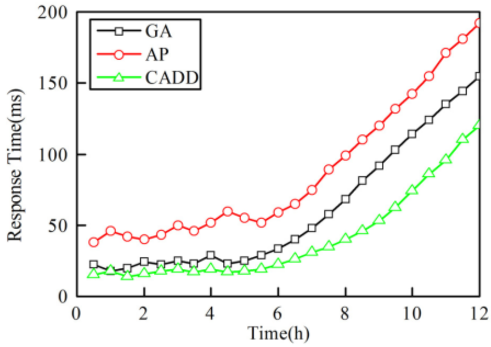


Fig. 2. Response time under different traffic loads in OS3E

Performance Metrics: We quantify the efficacy using the following key metrics:
1) Communication Cost: The combined intra-domain and inter-domain communication overheads as per our objective function, indicative of the overall control plane efficiency.
2) Load Balancing Ratio: Ratio of the maximum controller load to the minimum, lower values indicating better load distribution across controllers.
3) Control Plane Latency: The average and 99th percentile response times experienced by flow setup requests at the controllers, directly impacting application performance.
4) Reassignment Overheads: For dynamic schemes, the frequency and volume of switch remappings between controllers, capturing consistency and stability factors.
Through meticulous simulations spanning these diverse scenarios, metrics, benchmarks, and sensitivity analyses, we comprehensively evaluate our schemes to validate their effectiveness in achieving optimized yet resilient multi-controller deployments for large-scale SDN control planes.

6. CONCLUSION

This paper proposed a novel multi-controller deployment framework for software-defined networks (SDNs) that jointly optimizes traffic propagation delays and dynamic load balancing across controllers. We developed PSOAP, a modified affinity propagation clustering algorithm using particle swarm optimization for initial controller placement, and CDAA for dynamic remapping of switches to balance loads based on current traffic conditions and controller capacities modeled as M/M/1 queues. Extensive simulations across diverse topologies demonstrated the superiority of our load-aware scheme over existing affinity propagation, genetic algorithms, and static mapping approaches, achieving lower communication costs, balanced controller utilization, and sub-10ms control plane latencies even under fluctuating traffic. The holistic framework ensures responsive and scalable SDN control planes for large production networks. Future work includes extending to cloud-edge infrastructures and online learning for proactive traffic-aware adaptation.

References

[1] . Maleh, Y., Qasmaoui, Y., El Gholami, K., Sadqi, Y., & Mounir, S. (2023). A comprehensive survey on SDN security: threats, mitigations, and future directions. Journal of Reliable Intelligent Environments, 9(2), 201-239.

[2] . Jalili, A., Keshtgari, M., & Akbari, R. (2018). A new set covering controller placement problem model for large scale SDNs. Information Systems & Telecommunication, 25, 6.

[3] . Isong, B., Molose, R. R. S., Abu-Mahfouz, A. M., & Dladlu, N. (2020). Comprehensive review of SDN controller placement strategies. IEEE Access, 8, 170070-170092.

[4] . Jalili, A., & Keshtgari, M. (2020). A new reliable controller placement model for software-defined WANs. Journal of AI and Data Mining, 8(2), 269-277.

[5] . Almadani, B., Beg, A., & Mahmoud, A. (2021). DSF: A distributed sdn control plane framework for the east/west interface. IEEE Access, 9, 26735-26754.

[6] . G. Ishigaki and N. Shinomiya, “Controller placement algorithm to alleviate burdens on communication nodes,” in Proc. Int. Conf. Comput., Netw. Commun., Feb. 2016, pp. 1–5.

[7] . Y. Jimenez, C. Cervelló-Pastor, and A. J. Garcia, ‘‘On the controller placement for designing a distributed SDN control layer,’’ in Proc. Netw. Conf., Jun. 2014, pp. 1–9.
[8] . Hock, D., Gebert, S., Hartmann, M., Zinner, T., & Tran-Gia, P. (2014, May). POCO-framework for Pareto-optimal resilient controller placement in SDN-based core networks. In 2014 IEEE Network Operations and Management Symposium (NOMS) (pp. 1-2). IEEE.
[9] Jeya, R., Venkatakrishnan, G. R., & Nagarajan, V. (2023). Placing Controllers Using Latency Metrics in a Smart Grid Implementing Software-Defined Networking Architecture. Advances in Science and Technology, 124, 828-835.
[10] L. F. Müller, R. R. Oliveira, M. C. Luizelli, L. P. Gaspary, and M. P. Barcellos, ‘‘Survivor: An enhanced controller placement strategy for improving SDN survivability,’’ in Proc. IEEE Global Commun. Conf., Dec. 2014, pp. 1909–1915.
[11] Xu, C., Xu, C., Li, B., Li, S., & Li, T. (2023). Load-Aware Dynamic Controller Placement Based on Deep Reinforcement Learning in SDN-Enabled Mobile Cloud-Edge Computing Networks. Computer Networks, 109900.
[12] A. Jalili, V. Ahmadi, M. Keshtgari, and M. Kazemi, ‘‘Controller placement in software-defined wan using multi objective genetic algorithm,’’ in Proc. 2nd Int. Conf. Knowl.-Based Eng. Innov. (KBEI), Nov. 2015, pp. 656–662.
[13] V. Ahmadi, A. Jalili, S. M. Khorramizadeh, and M. Keshtgari, ‘‘A hybrid NSGA-II for solving multiobjective controller placement in SDN,’’ in Proc. 2nd Int. Conf. Knowl.-Based Eng. Innov., Nov. 2015, pp. 663–669.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, AN–296, pp. 608–620

Existence of Three Solutions for a $2n$-th-Order Impulsive Differential Equation

Osman Halakoo[a,*]

[a] *Department of Mathematics, Islamic Azad University, Aliabad Katoul Branch, Aliabad Katoul, Iran*

Article Info

Keywords:
Impulsive differential equations
weak solution
classical solution
critical point

2020 MSC:
34A37
34B37

Abstract

In this paper, we prove the existence of at least three solutions for some $2n$-th-order impulsive equations. A particular case and an example are then presented.

1. Introduction

Let $n \in \mathbb{N} - \{1\}$. In this paper, we consider the $2n$-th-order impulsive boundary-value problem

$$\begin{cases} (-1)^n u^{(2n)}(t) + \cdots + u^{(4)}(t) - (p(t)u'(t))' + q(t)u(t) = \lambda f(t, u(t)), \quad t \neq t_j,\ t \in [0,1] \\ u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0 = u^{(n)}(0) = u^{(n)}(1) \\ (-1)^{i-1} \Delta\left(u^{(n+i-1)}(t_j)\right) = \mu I_{ij}\left(u^{(i-1)}(t_j)\right), \quad i = 1, 2, \ldots, n,\ j = 1, 2, \ldots, m \end{cases} \tag{1}$$

where $p \in C^1([0,1] \times [0,+\infty))$, $q \in L^\infty([0,1])$, $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ is an L^1-Carathéodory function, λ, μ are positive constants, $0 = t_0 < t_1 < t_2 < \cdots < t_n < t_{n+1} = 1$, $\Delta(u(t_j)) = u(t_j^+) - u(t_j^-) = \lim_{t \to t_j^+} u(t) - \lim_{t \to t_j^-} u(t)$ and $I_{ij} : \mathbb{R} \to \mathbb{R}$ are continuous for every $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.
Let $p^- = \operatorname{ess\,inf}_{t \in [0,1]} p(t)$ and $q^- = \operatorname{ess\,inf}_{t \in [0,1]} q(t)$ and suppose that the condition

$$\min\left\{\frac{p^-}{\pi^2}, \frac{q^-}{\pi^4}, \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}\right\} > -1 \tag{2}$$

is satisfied in problem (1). We set

$$\sigma = \min\left\{\frac{p^-}{\pi^2}, \frac{q^-}{\pi^4}, \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}, 0\right\} \in (-1, 0], \quad \delta = \sqrt{1+\sigma} \in (0,1]$$

*Talker
Email address: osman.halakoo@yahoo.com, halakoo@aliabadiau.ac.ir (Osman Halakoo)

The study of impulsive boundary-value problems is important due to its various applications in which abrupt changes at certain times in the evolution process appear. The dynamics of evolving processes is often subjected to abrupt changes such as shocks, harvesting and natural disasters. Often these short-term perturbations are treated as having acted instantaneously or in the form of impulses. Such problems arise in physics, population dynamics, biotechnology, pharmacokinetics and industrial robotics. Many researchers pay their attention to impulsive differential equations by variational method and critical point theory, for example in [2], the existence of at least one non-trivial classical solution to the nonlinear Dirichlet impulsive boundary-value problem

$$\begin{cases} -(p(t)u'(t))' + q(t)u(t) = \lambda f(t, u(t)), \quad t \in [0, T],\ t \neq t_j \\ u(0) = u(T) = 0 \\ \Delta u'(t_j) = \lambda I_j(u(t_j)), \quad j = 1, 2, \ldots, n \end{cases}$$

is established. Also in [1], the existence of three classical solutions for the fourth-order impulsive bandary-value problem

$$\begin{cases} u^{(iv)}(t) + Au''(t) + Bu(t) = \lambda f(t, u(t)) + \mu g(t, u(t)), \quad t \neq t_j,\ t \in [0, 1] \\ \Delta(u''(t_j)) = I_{1j}(u'(t_j)), \quad j = 1, 2, \ldots, n \\ -\Delta(u'''(t_j)) = I_{2j}(u(t_j)), \quad j = 1, 2, \ldots, n \\ u(0) = u(1) = u''(0) = u''(1) = 0 \end{cases}$$

is proved, where $A \leq 0 \leq B$ are real constants.

2. PRELIMINARIES

We now state two critical point theorems established by Bonanno and coauthors [3, 4] which are the main tools for the proofs of our results.

Theorem 2.1 ([4, Theorem 2.6]). *Let X be a reflexive real Banach space; $\Phi : X \to \mathbb{R}$ be a sequentially weakly lower semicontinuous, coercive and continuously Gâteaux differentiable functional whose Gâteaux derivative admits a continuous inverse on X^*, $\Psi : X \to \mathbb{R}$ be a sequentially weakly upper semicountinuous, continuously Gâteaux differentiable functional whose Gâteaux derivative is compact, such that $\Phi(0) = \Psi(0) = 0$.*
Assume that there exist $r > 0$ and $\bar{x} \in X$, with $r < \Phi(\bar{x})$ such that

(i) $\sup_{\Phi(x)\leq r} \Psi(x) < \frac{r\Psi(\bar{x})}{\Phi(\bar{x})}$,
(ii) for each λ in,

$$\Lambda_r := \left(\frac{\Phi(\bar{x})}{\Psi(\bar{x})}, \frac{r}{\sup_{\Phi(x)\leq r} \Psi(x)} \right),$$

the functional $\Phi - \lambda\Psi$ is coercive.

Then, for each $\lambda \in \Lambda_r$ the functional $\Phi - \lambda\Psi$ has at least three distinct critical points in X.

Theorem 2.2 ([3, Theorem 3.2]). *Let X be a reflexive real Banach space; $\Phi : X \to \mathbb{R}$ be a convex, coercive and continuously Gâteaux differentiable functional whose Gâteaux derivative admits a continuous inverse on X^*, $\Psi : X \to \mathbb{R}$ be a continuously Gâteaux differentiable functional whose Gâteaux derivative is compact, such that*

$$\inf_{x\in X} \Phi(x) = \Phi(0) = \Psi(0) = 0.$$

Assume that there exist two positive constants $r_1, r_2 > 0$ and $\bar{x} \in X$, with $2r_1 < \Phi(\bar{x}) < \frac{r_2}{2}$, such that

(j) $\frac{\sup_{\Phi(x)<r_1} \Psi(x)}{r_1} < \left(\frac{2}{3}\right) \frac{\Psi(\bar{x})}{\Phi(\bar{x})}$,

(jj) $\frac{\sup_{\Phi(x)<r_2}\Psi(x)}{r_2} < \left(\frac{1}{3}\right)\frac{\Psi(\bar{x})}{\Phi(\bar{x})}$,

(jjj) *for each* λ *in,*

$$\Lambda_{r_1,r_2} := \left(\frac{3}{2}\frac{\Phi(\bar{x})}{\Psi(\bar{x})}, \min\left\{\frac{r_1}{\sup_{\Phi(x)<r_1}\Psi(x)}, \frac{r_2}{2\sup_{\Phi(x)<r_2}\Psi(x)}\right\}\right)$$

and for every $x_1, x_2 \in X$, *which are local minima for the functional* $\Phi - \lambda\Psi$, *and such that* $\Psi(x_1) \geq 0$ *and* $\Psi(x_2) \geq 0$, *one has* $\inf_{t\in[0,1]}\Psi(tx_1 + (1-t)x_2) \geq 0$.

Then, for each $\lambda \in \Lambda_{r_1,r_2}$ *the functional* $\Phi - \lambda\Psi$ *has at least three distinct critical points which lie in* $\Phi^{-1}((-\infty, r_2))$.

Let us introduce some notations which will be used later. Consider the Sobolev spaces

$$\begin{aligned} W^{n,2}([0,1]) =& H^n([0,1]) := \left\{u \in L^2([0,1]) : u', u'', \dots, u^{(n)} \in L^2([0,1])\right\}, \\ W_0^{n-1,2}([0,1]) =& H_0^{n-1}([0,1]) := \Big\{u \in L^2([0,1]) : u', u'', \dots, u^{(n-1)} \in L^2([0,1]), \\ & u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0\Big\}. \end{aligned}$$

Take

$$\begin{aligned} X =& H^n([0,1]) \cap H_0^{n-1}([0,1]) := \Big\{u \in L^2([0,1]) : u', u'', \dots, u^{(n)} \in L^2([0,1]), \\ & u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0\Big\}. \end{aligned}$$

Define the inner product

$$\begin{aligned} (u, v) =& \int_0^1 u^{(n)}(t)v^{(n)}(t)dt + \int_0^1 u^{(n-1)}(t)v^{(n-1)}(t)dt + \cdots \\ &+ \int_0^1 p(t)u'(t)v'(t)dt + \int_0^1 q(t)u(t)v(t)dt, \end{aligned}$$

which induces the norm

$$\|u\|_X = \left(\int_0^1 \left(u^{(n)}(t)\right)^2 dt + \int_0^1 \left(u^{(n-1)}(t)\right)^2 dt + \cdots + \int_0^1 p(t)\left(u'(t)\right)^2 dt + \int_0^1 q(t)(u(t))^2 dt\right)^{\frac{1}{2}}.$$

Also, we introduce the norm:

$$|||u||| = \left(\|u^{(n)}\|_2^2 + \|u^{(n-1)}\|_2^2 + \cdots + \|u''\|_2^2\right)^{\frac{1}{2}}.$$

Then by [6, Lemma 2.3], the following Poincaré-inequalities hold for every $u \in X$:

$$\|u\|_2 \leq \frac{1}{\pi^2}\|u''\|_2 \tag{3}$$

$$\|u'\|_2 \leq \frac{1}{\pi}\|u''\|_2. \tag{4}$$

Now, we have the following useful proposition.

Proposition 2.3. *Let* $u \in X$ *and* $D = \left(1 + \frac{\|p\|_\infty}{\pi^2} + \frac{\|q\|_\infty}{\pi^4}\right)^{\frac{1}{2}}$. *Then*

$$\delta|||u||| \leq \|u\|_X \tag{5}$$

$$\|u\|_X \leq D|||u|||. \tag{6}$$

Proof. To prove of (5), we have four cases.

1. If $p^- \geq 0, q^- \geq 0$, then $\sigma = 0$ and $\delta = 1$. Then we have

$$\begin{aligned}\|u\|_X^2 &= \int_0^1 \left(|u^{(n)}(t)|^2 + \cdots + |u''(t)|^2 + p(t)|u'(t)|^2 + q(t)|u(t)|^2\right) dt \\ &\geq \left(\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2 + p^-\|u'\|_2^2 + q^-\|u\|_2^2\right) \\ &\geq \left(\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2\right) = |||u|||^2.\end{aligned}$$

So in this case we have $\delta|||u||| \leq \|u\|_X$.

2. If $q^- < 0 < p^-$ then $-1 < \sigma = \frac{q^-}{\pi^4} < 0$ and $0 < \delta = \sqrt{1 + \frac{q^-}{\pi^4}} < 1$. Since $q^- < 0$, by (3) we have

$$\begin{cases} \frac{q^-}{\pi^4}\|u''\|_2^2 \leq q^-\|u\|_2^2 \leq \int_0^1 q(t)|u(t)|^2 dt \\ \|u''\|_2^2 \leq \int_0^1 |u''(t)|^2 dt \end{cases}$$

Then by summing up inequalities, we have

$$\begin{aligned}\left(1 + \frac{q^-}{\pi^4}\right)\|u''\|_2^2 \leq& \|u''\|_2^2 + \int_0^1 q(t)|u(t)|^2 dt \\ \Rightarrow \delta^2|||u|||^2 =& \left(1 + \frac{q^-}{\pi^4}\right)\left(\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2\right) \\ \leq& \left(1 + \frac{q^-}{\pi^4}\right)\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ \leq& \int_0^1 q(t)|u(t)|^2 dt + \|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ \leq& \int_0^1 q(t)|u(t)|^2 dt + \int_0^1 p(t)|u'(t)|^2 dt + \|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ =& \|u\|_X^2.\end{aligned}$$

So in this case we have $\delta|||u||| \leq \|u\|_X$.

3. If $p^- < 0 < q^-$ then $-1 < \sigma = \frac{p^-}{\pi^2} < 0$ and $0 < \delta = \sqrt{1 + \frac{p^-}{\pi^2}} < 1$. Since $p^- < 0$, by (4) we have

$$\begin{cases} \frac{p^-}{\pi^2}\|u''\|_2^2 \leq p^-\|u'\|_2^2 \leq \int_0^1 p(t)|u'(t)|^2 dt \\ \|u''\|_2^2 \leq \int_0^1 |u''(t)|^2 dt \end{cases}$$

Then by summing up inequalities, we have

$$\begin{aligned}\left(1 + \frac{p^-}{\pi^2}\right)\|u''\|_2^2 \leq& \|u''\|_2^2 + \int_0^1 p(t)|u'(t)|^2 dt \\ \Rightarrow \delta^2|||u|||^2 =& \left(1 + \frac{p^-}{\pi^2}\right)\left(\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2\right) \\ \leq& \left(1 + \frac{p^-}{\pi^2}\right)\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ \leq& \int_0^1 p(t)|u'(t)|^2 dt + \|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ \leq& \int_0^1 q(t)|u(t)|^2 dt + \int_0^1 p(t)|u'(t)|^2 dt + \|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ =& \|u\|_X^2.\end{aligned}$$

So in this case we have $\delta|||u||| \leq \|u\|_X$.

4. If $p < 0, q < 0$ then $-1 < \sigma = \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4} < 0$ and $0 < \delta = \sqrt{1 + \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}} < 1$. So by (3) and (4) we have

$$\begin{cases} \frac{q^-}{\pi^4}\|u''\|_2^2 \leq q^-\|u\|_2^2 \leq \int_0^1 q(t)|u(t)|^2dt \\ \frac{p^-}{\pi^2}\|u''\|_2^2 \leq p^-\|u'\|_2^2 \leq \int_0^1 p(t)|u'(t)|^2dt \\ \|u''\|_2^2 \leq \int_0^1 |u''(t)|^2dt \end{cases}$$

Then by summing up inequalities, we have

$$\begin{aligned} &\left(1 + \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}\right)\|u''\|_2^2 \leq \int_0^1 q(t)|u(t)|^2dt + \int_0^1 p(t)|u'(t)|^2dt + \|u''\|_2^2 \\ \Rightarrow &\delta^2|||u|||^2 = \left(1 + \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}\right)\left(\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2\right) \\ &\leq \left(1 + \frac{p^-}{\pi^2} + \frac{q^-}{\pi^4}\right)\|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ &\leq \int_0^1 q(t)|u(t)|^2dt + \int_0^1 p(t)|u'(t)|^2dt + \|u''\|_2^2 + \|u'''\|_2^2 + \cdots + \|u^{(n)}\|_2^2 \\ &= \|u\|_X^2. \end{aligned}$$

So in this case we have $\delta|||u||| \leq \|u\|_X$. Then (5) is proved. To prove of (6), taking (3) and (4) into account, we have

$$\begin{aligned} \|u\|_X^2 =&\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2 + \int_0^1 p(t)(u'(t))^2dt + \int_0^1 q(t)(u(t))^2dt \\ \leq&\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2 + \int_0^1 |p(t)|(u'(t))^2dt + \int_0^1 |q(t)|(u(t))^2dt \\ \leq&\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2 + \|p\|_\infty \int_0^1 (u'(t))^2dt + \|q\|_\infty \int_0^1 (u(t))^2dt \\ \leq&\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2 + \frac{\|p\|_\infty}{\pi^2}\|u''\|_2^2 + \frac{\|q\|_\infty^2}{\pi^4}\|u''\|_2^2 \\ \leq&\left(1 + \frac{\|p\|_\infty}{\pi^2} + \frac{\|q\|_\infty}{\pi^4}\right)\left(\|u^{(n)}\|_2^2 + \cdots + \|u''\|_2^2\right) \\ =&D^2|||u|||^2. \end{aligned}$$

Then we have $\|u\|_X \leq D|||u|||$ and (6) is proved.

$\square$

Remark 2.4. By (5) and (6) we have $\delta|||u||| \leq \|u\|_X \leq D|||u|||$, i.e. the norms $|||\cdot|||$ and $\|\cdot\|_X$ are equivalent. On the other, for the norm in $C^n([0,1])$,

$$\|u\|_\infty = \max\left\{\max_{t\in[0,1]}|u(t)|, \max_{t\in[0,1]}|u'(t)|, \ldots, \max_{t\in[0,1]}|u^{(n)}(t)|\right\},$$

by [7] we have $\|u\|_\infty \leq \frac{1}{2}\|u'\|_2$ then for every $u \in X$:

$$\max_{t\in[0,1]}|u(t)| \leq \|u\|_\infty \leq \frac{1}{2}\|u'\|_2 \leq \frac{1}{2\pi}\|u''\|_2 \leq \frac{1}{2\pi}|||u||| \leq \frac{1}{2\pi\delta}\|u\|_X. \tag{7}$$

Here and in the sequel $f : [0,1]\times\mathbb{R} \to \mathbb{R}$ is an L^1-Carathéodory function. We recall that $f : [0,1]\times\mathbb{R} \to \mathbb{R}$ is an L^1-Carathéodory function if

(a) the mapping $t \mapsto f(t,x)$ is measurable for every $x \in \mathbb{R}$;
(b) the mapping $x \mapsto f(t,x)$ is continuous for almost every $t \in [0,1]$;
(c) for every $\rho > 0$ there exists a function $l_\rho \in L^1([0,1])$ such that

$$\sup_{|x|\le\rho} |f(t,x)| \le l_\rho(t)$$

for almost every $t \in [0,1]$;

Corresponding to f we introduce the function F as follow

$$F : [0,1]\times\mathbb{R} \to \mathbb{R}$$
$$(t,x) \mapsto F(t,x) := \int_0^x f(t,\varepsilon)d\varepsilon.$$

We say that $u \in C([0,1])$ is a classical solution of problem (1), if it satisfies the equation in (1) a.e. on $[0,1] \setminus \{t_1, t_2, \ldots, t_m\}$, the limits $u^{(i)}(t_j^+)$ and $u^{(i)}(t_j^-)$, for every $1 \le i \le n$, $1 \le j \le m$, exist, satisfy n impulsive conditions in (1) and the boundary conditions $u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0 = u^{(n)}(0) = u^{(n)}(1)$.
On the other hand if the parts of equation in (1) multiplied by an arbitrary function $v \in X$ and then integrated in $x \in [0,1]$, then by n times integration by parts we have

$$\int_0^1 \left(u^{(n)}(t)v^{(n)}(t) + \cdots + u''(t)v''(t) + p(t)u'(t)v'(t) + q(t)u(t)v(t)\right) dt =$$
$$-\mu\sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}\left(u^{(i-1)}(t_j)\right) v^{(n-i)}(t_j) + \lambda\int_0^1 f(t,u(t))v(t)dt$$

for all $v \in X$. Then we say that function $u \in X$ in the above equality, is a weak solution of (1).

Lemma 2.5 ([5, Lemma 2.2]). *The function $u \in X$ is a weak solution of problem (1) if and if u is a classical solution of problem (1).*

Lemma 2.6. *Assume that:*

(H_1) *There exist positive constants α_i, β_i and $\sigma_i \in [0,1)$, $i = 1,2\ldots,n$ such that for all $x \in \mathbb{R}$, $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,m$:*

$$|I_{ij}(x)| \le \alpha_i + \beta_i|x|^{\sigma_i}.$$

Then for any $u \in X$, we have

$$\left|\sum_{j=1}^{m}\int_0^{u^{(i-1)}(t_j)} I_{ij}(x)dx\right| \le m\left(\alpha_i\|u\|_\infty + \frac{\beta_i}{\sigma_i+1}\|u\|_\infty^{\sigma_i+1}\right),\ \forall i = 1,2,\ldots,n.$$

Proof. By (H_1), for any $i = 1, 2, \dots, n$ we deduce

$$\begin{aligned}
\left|\sum_{j=1}^{m}\int_0^{u^{(i-1)}(t_j)} I_{ij}(x)dx\right| \leq & \sum_{j=1}^{m}\left|\int_0^{u^{(i-1)}(t_j)} I_{ij}(x)dx\right| \\
\leq & \sum_{j=1}^{m}\int_0^{u^{(i-1)}(t_j)} |I_{ij}(x)|\,dx \\
\leq & \sum_{j=1}^{m}\int_0^{u^{(i-1)}(t_j)} (\alpha_i + \beta_i|x|^{\sigma_i})\,dx \\
\leq & \sum_{j=1}^{m}\left(\alpha_i|x| + \frac{\beta_i}{\sigma_i+1}|x|^{\sigma_i+1}\right)_{x=0}^{u^{(i-1)}(t_j)} \\
= & \sum_{j=1}^{m}\left(\alpha_i\left|u^{(i-1)}(t_j)\right| + \frac{\beta_i}{\sigma_i+1}\left|u^{(i-1)}(t_j)\right|^{\sigma_i+1}\right) \\
\leq & \sum_{j=1}^{m}\left(\alpha_i\|u\|_\infty + \frac{\beta_i}{\sigma_i+1}\|u\|_\infty^{\sigma_i+1}\right) \\
= & m\left(\alpha_i\|u\|_\infty + \frac{\beta_i}{\sigma_i+1}\|u\|_\infty^{\sigma_i+1}\right).
\end{aligned}$$

Thus the conclusion is achived. □

3. MAIN RESULTS

Put

$$k = \frac{4\delta^2\pi^2}{(n-1)\left(\frac{(2n-2)!}{(n-2)!}\right)^2\left(1+\frac{\|p\|_\infty}{\pi^2}+\frac{\|q\|_\infty}{\pi^4}\right)\left[\frac{1}{\alpha^{2n-1}}+\frac{1}{(1-\beta)^{2n-1}}\right]},$$

$$\Gamma_{i,c} = \frac{\alpha_i}{c} + \left(\frac{\beta_i}{\sigma_i+1}\right)c^{\sigma_i-1},$$

where $0 < \alpha < \beta < 1$ and c is a positive constant and α_i, β_i and σ_i, $(i = 1, 2, \dots, n)$, are given by (H_1).
We state our main result as follows:

Theorem 3.1. *Suppose that (H_1) is satisfied and there exist two positive constants c, d with $c < (n-1)\frac{d}{\pi}\left[\frac{1}{\alpha^3}+\frac{1}{(1-\beta)^3}\right]^{\frac{1}{2}}$, such that*

(A_1) $F(t,\varepsilon) \geq 0$ *for all* $(t,\varepsilon) \in ([0,\alpha]\cup[\beta,1])\times[0,d]$,

(A_2) $\frac{\int_0^1 \max_{|\varepsilon|\leq c} F(t,\varepsilon)dt}{c^2} < \frac{k\int_\alpha^\beta F(t,d)dt}{d^2}$,

(A_3) $\limsup_{|\varepsilon|\to+\infty}\frac{\sup_{t\in[0,1]}F(t,\varepsilon)}{\varepsilon^2} \leq \frac{\pi^2\int_0^1\max_{|\varepsilon|\leq c}F(t,\varepsilon)dt}{4c^2}$.

Then for every

$$\lambda \in \Lambda = \left(\frac{2\delta^2\pi^2 d^2}{k\int_\alpha^\beta F(t,d)dt}, \frac{2\delta^2\pi^2c^2}{\int_0^1\max_{|\varepsilon|\leq c}F(t,\varepsilon)dt}\right),$$

there exists

$$\rho = \frac{1}{2m}\min_{1\leq i\leq n}\left\{\frac{2\delta^2\pi^2c^2 - \lambda\int_0^1\max_{|\varepsilon|\leq c}F(t,\varepsilon)dt}{c^2\Gamma_{i,c}}, \frac{k\lambda\int_\alpha^\beta F(t,d)dt - 2\delta^2\pi^2d^2}{d^2\Gamma_{i,\left(\frac{d}{\sqrt{k}}\right)}}\right\}$$

such that for each $\mu \in [0,\rho)$ the problem (1) has at least three distinct classical solutions.

Proof. First, we observe that due to (A_2) the interval Λ is non-empty and consequently, one has $\rho > 0$. Now, fix λ and μ as in the conclusion. Our aim is to apply Theorem 2.1. We define the functionals Φ and Ψ as follows

$$\begin{aligned} &\Phi : X \to \mathbb{R} \\ &\qquad u \mapsto \Phi(u) = \frac{1}{2}\|u\|_X^2, \\ &\Psi : X \to \mathbb{R} \\ &\qquad u \mapsto \Psi(u) = \int_0^1 F(t,u(t))dt - \frac{\mu}{\lambda}\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{u^{(i-1)}(t_j)} I_{ij}(x)dx, \end{aligned}$$

and for the fix $\lambda \in \Lambda$ put

$$\begin{aligned} &E_\lambda : X \to \mathbb{R} \\ &\qquad u \mapsto E_\lambda(u) = \Phi(u) - \lambda\Psi(u). \end{aligned} \tag{8}$$

Using the property of f and the continuity of I_{ij}, $i = 1, 2, \dots, n$ and $j = 1, 2, \dots, m$, we obtain that $\Phi, \Psi \in C^1(X, \mathbb{R})$ and for any $v \in X$, we have

$$\Phi'(u)(v) = \int_0^1 \Big(u^{(n)}(t)v^{(n)}(t) + \cdots + u''(t)v''(t) + p(t)u'(t)v'(t) + q(t)u(t)v(t)\Big)\,dt$$

and

$$\Psi'(u)(v) = \int_0^1 f(t,u(t))v(t)dt - \frac{\mu}{\lambda}\sum_{i=1}^{n}\sum_{j=1}^{m} I_{ij}\left(u^{(i-1)}(t_j)\right)v^{(n-i)}(t_j).$$

Since any norm in a Banach space is a sequentially weakly lower semicontinuous functional then it is obvious that the functional Φ is nonnegative, coercive, continuously Gâteaux differentiable and sequentially weakly lower semicontinuous and Φ' has a continuous inverse on X^*, also Ψ is continuously Gâteaux differentiable and Ψ' is compact. With standard arguments, we deduce that the critical points of the functional E_λ are the weak solutions of problem (1) and so they are classical. We will verify (i) and (ii) of Theorem 2.1. Put $r = 2(\delta\pi c)^2$. Taking (7) into account, for every $u \in X$ such that $\Phi(u) \le r$, one has $\max_{t\in[0,1]} |u(t)| \le c$. Consequently, from Lemma 2.6 it follows that

$$\sup_{\Phi(u)\le r} \Psi(u) \le \int_0^1 \max_{|\varepsilon|\le c} F(t,\varepsilon)dt + \frac{\mu}{\lambda}m\sum_{i=1}^{n}\left(\alpha_i c + \frac{\beta_i}{\sigma_i+1}c^{\sigma_i+1}\right),$$

that is,

$$\frac{\sup_{\Phi(u)\le r}\Psi(u)}{r} \le \frac{1}{2\delta^2\pi^2}\left[\frac{\int_0^1 \max_{|\varepsilon|\le c} F(t,\varepsilon)dt}{c^2} + \frac{\mu}{\lambda}m\sum_{i=1}^{n}\Gamma_{i,c}\right].$$

Hence, bearing in mind that $\mu < \rho$, one has

$$\frac{\sup_{\Phi(u)\le r}\Psi(u)}{r} < \frac{1}{\lambda}. \tag{9}$$

Put

$$w(t) = \begin{cases} d\sum_{i=0}^{n-1}(-1)^{n-1-i}\binom{2n-3-i}{n-1-i}\binom{2n-2}{i}\left(\frac{x}{\alpha}\right)^{2n-2-i}, & x \in [0,\alpha) \\ d, & x \in [\alpha,\beta] \\ \frac{d}{(1-\beta)^{2n-2}}\Big[\binom{2n-3}{n-1}(2n-2)\sum_{i=0}^{2n-3}\frac{(-1)^{n-1-i}}{2n-2-i}\Big(\sum_{j=\max\{0,-n+1+i\}}^{\min\{i,n-2\}}\binom{n-2}{n-2-j} & \\ \qquad\binom{n-1}{n-1-i+j}\beta^{i-j}\Big)x^{2n-2-i} + \sum_{i=n-1}^{2n-2}(-1)^i\binom{2n-2}{i}\beta^{2n-2-i}\Big], & x \in (\beta,1] \end{cases}$$

Clearly $w \in X$. Moreover, taking (3), (4) and (5) into account, one has

$$\begin{aligned} &4(n-1)^2\delta^2 d^2\left[\frac{1}{\alpha^3}+\frac{1}{(1-\beta)^3}\right] \leq \|w\|_X^2 \\ \leq &(n-1)d^2\left(\frac{(2n-2)!}{(n-2)!}\right)^2\left[\frac{1}{\alpha^{2n-1}}+\frac{1}{(1-\beta)^{2n-1}}\right]\left(1+\frac{\|p\|_\infty}{\pi^2}+\frac{\|q\|_\infty}{\pi^4}\right) \\ = &\frac{4\delta^2\pi^2 d^2}{k}, \end{aligned} \tag{10}$$

and for all $t \in [0,1]$ we have $0 \leq w(t) \leq d$.
So from $c < (n-1)\frac{d}{\pi}\left[\frac{1}{\alpha^3}+\frac{1}{(1-\beta)^3}\right]^{\frac{1}{2}}$, we obtain

$$r = 2(\delta\pi c)^2 < 2\delta^2(n-1)^2 d^2\left[\frac{1}{\alpha^3}+\frac{1}{(1-\beta)^3}\right] \leq \frac{1}{2}\|w\|_X^2 = \Phi(w).$$

Moreover, again from (10), we have $\Phi(w) < \frac{2\delta^2\pi^2 d^2}{k}$.
Now, due to Lemma (2.6), (A_1), (7) and (10) one has

$$\begin{aligned} \Psi(w) &= \int_0^1 F(t,w(t))dt - \frac{\mu}{\lambda}\sum_{i=1}^{n}\sum_{j=1}^{m}\int_0^{w^{(i-1)}(t_j)} I_{ij}(x)dx \\ &\geq \int_\alpha^\beta F(t,d)dt - \frac{\mu}{\lambda}m\sum_{i=1}^{n}\left(\alpha_i\|w\|_\infty + \frac{\beta_i}{\sigma_i+1}\|w\|_\infty^{\sigma_i+1}\right) \\ &\geq \int_\alpha^\beta F(t,d)dt - \frac{\mu}{\lambda}m\frac{d^2}{k}\sum_{i=1}^{n}\Gamma_{i,\left(\frac{d}{\sqrt{k}}\right)}. \end{aligned}$$

So, we obtain

$$\frac{\Psi(w)}{\Phi(w)} \geq \frac{k\int_\alpha^\beta F(t,d)dt - \frac{\mu}{\lambda}md^2\sum_{i=1}^{n}\Gamma_{i,\left(\frac{d}{\sqrt{k}}\right)}}{2(\delta\pi d)^2}.$$

Since $\mu < \rho$, one has

$$\frac{\Psi(w)}{\Phi(w)} > \frac{1}{\lambda}. \tag{11}$$

Therefore, from (9) and (11), condition (i) of Theorem (3) is fulfilled, i.e.

$$\sup_{\Phi(u)\leq r}\Psi(u) < r\frac{\Psi(w)}{\Phi(w)}.$$

Now, to prove the coercivity of the functional $E_\lambda = \Phi - \lambda\Psi$, due to (A_3) and the choice of λ in the conclusion, we have

$$\limsup_{|\varepsilon|\to+\infty}\frac{\sup_{t\in[0,1]}F(t,\varepsilon)}{\varepsilon^2} < \left(\frac{\pi^4 d^2}{2}\right)\frac{1}{\lambda}.$$

So, we can fix η satisfying

$$\limsup_{|\varepsilon|\to+\infty}\frac{\sup_{t\in[0,1]}F(t,\varepsilon)}{\varepsilon^2} < \eta < \left(\frac{\pi^4 d^2}{2}\right)\frac{1}{\lambda}.$$

Then, there exists a constant h such that

$$F(t,\varepsilon) \leq \eta|\varepsilon|^2 + h, \quad \forall t \in [0,1], \ \ \forall \varepsilon \in \mathbb{R}.$$

Taking into account Lemma (2.6), (3), (5) and (7), it follows that for all $u \in X$, we have

$$\begin{aligned}
E_\lambda(u) =& \Phi(u) - \lambda\Psi(u) = \frac{1}{2}\|u\|_X^2 - \lambda\int_0^1 F(t,u(t))dt + \mu\sum_{i=1}^n\sum_{j=1}^m\int_0^{u^{(i-1)}(t_j)} I_{ij}(x)dx \\
\geq& \frac{1}{2}\|u\|_X^2 - \lambda\eta\|u\|_2^2 - \lambda h - \mu m\sum_{i=1}^n\left[\frac{\alpha_i}{2\delta\pi}\|u\|_X + \frac{\beta_i}{\sigma_i+1}\left(\frac{1}{2\delta\pi}\right)^{\sigma_i+1}\|u\|_X^{\sigma_i+1}\right] \\
\geq& \frac{1}{2}\|u\|_X^2 - \frac{\lambda\eta}{\pi^4}\|u''\|_2^2 - \lambda h - \mu m\sum_{i=1}^n\left[\frac{\alpha_i}{2\delta\pi}\|u\|_X + \frac{\beta_i}{\sigma_i+1}\left(\frac{\|u\|_X}{2\delta\pi}\right)^{\sigma_i+1}\right] \\
\geq& \frac{1}{2}\|u\|_X^2 - \frac{\lambda\eta}{\pi^4}|||u|||^2 - \lambda h - \mu m\sum_{i=1}^n\left[\frac{\alpha_i}{2\delta\pi}\|u\|_X + \frac{\beta_i}{\sigma_i+1}\left(\frac{\|u\|_X}{2\delta\pi}\right)^{\sigma_i+1}\right] \\
\geq& \left(\frac{1}{2} - \frac{\lambda\eta}{\delta^2\pi^4}\right)\|u\|_X^2 - \lambda h - \mu m\sum_{i=1}^n\left[\frac{\alpha_i}{2\delta\pi}\|u\|_X + \frac{\beta_i}{\sigma_i+1}\left(\frac{\|u\|_X}{2\delta\pi}\right)^{\sigma_i+1}\right].
\end{aligned}$$

So, the functional $E_\lambda = \Phi - \lambda\Psi$ is coercive. Thus condition (ii) of Theorem 2.1 is satisfied. Now, the conclusion of Theorem 2.1 can be used. It follows that, for every

$$\lambda \in \Lambda \subseteq \left(\frac{\Phi(w)}{\Psi(w)}, \frac{r}{\sup_{\Phi(u)\leq r}\Psi(u)}\right),$$

the functional $E_\lambda = \Phi - \lambda\Psi$ has at least three distinct critical points in X, which are the weak solutions and so classical solutions of the problem (1). This completes the proof. □

Corollary 3.2. *Suppose that* (H_1) *holds. Let* $\Theta \in L^1([0,1])$ *be a non-negative and non-zero function and let* $l : \mathbb{R} \to \mathbb{R}$ *be a continuous function. Put:* $\theta_0 = \int_\alpha^\beta \theta(t)dt$, $\|\theta\|_1 = \int_0^1 \theta(t)dt$, $L(\varepsilon) = \int_0^\varepsilon l(x)dx$, $\forall\varepsilon \in \mathbb{R}$. *Assume that there exist two positive constants* c, d, *with* $c < (n-1)\frac{d}{\pi}\left[\frac{1}{\alpha^3} + \frac{1}{(1-\beta)^3}\right]^{\frac{1}{2}}$, *such that*

(A_1') $L(\varepsilon) \geq 0$, $\forall\varepsilon \in [0,d]$;

(A_2') $\frac{\max_{|\varepsilon|\leq c} L(\varepsilon)}{c^2} \leq \frac{4\pi^2\theta_0 L(d)}{(n-1)\left(\frac{(2n-2)!}{(n-2)!}\right)^2\|\theta\|_1 d^2\left[\frac{1}{\alpha^{2n-1}} + \frac{1}{(1-\beta)^{2n-1}}\right]}$;

(A_3') $\limsup_{|\varepsilon|\to+\infty}\frac{L(\varepsilon)}{\varepsilon^2} \leq 0$.

Then for every

$$\lambda \in \left(\frac{(n-1)\left(\frac{(2n-2)!}{(n-2)!}\right)^2\left[\frac{1}{\alpha^{2n-1}} + \frac{1}{(1-\beta)^{2n-1}}\right]d^2}{2\theta_0 L(d)}, \frac{2\pi^2 c^2}{\|\theta\|_1\max_{|\varepsilon|\leq c}L(\varepsilon)}\right)$$

there exists

$$\rho = \frac{1}{2m}\min_{1\leq i\leq n}\left\{\frac{2\pi^2c^2 - \lambda\|\theta\|_1\max_{|\varepsilon|\leq c}L(\varepsilon)}{c^2\Gamma_{i,c}}, \frac{\frac{4\lambda\theta_0\pi^2 L(d)}{(n-1)\left(\frac{(2n-2)!}{(n-2)!}\right)^2\left[\frac{1}{\alpha^{2n-1}} + \frac{1}{(1-\beta)^{2n-1}}\right]} - 2\pi^2 d^2}{d^2\Gamma_{i,\left(\frac{d}{\sqrt{k}}\right)}}\right\}$$

such that, for each $\mu \in [0,\rho)$ *the problem*

$$\begin{cases}
(-1)^n u^{(2n)}(t) + \cdots + u^{(4)}(t) = \lambda\theta(t)l(u(t)), \quad t \neq t_j, \ t \in [0,1] \\
u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0 = u^{(n)}(0) = u^{(n)}(1) \\
(-1)^{i-1}\Delta\left(u^{(n+i-1)}(t_j)\right) = \mu I_{ij}\left(u^{(i-1)}(t_j)\right), \quad i = i, 2, \ldots, n, \ j = 1, 2, \ldots, m
\end{cases}$$

has at least three classical solutions.

Proof. The proof follows from Theorem 3.1 by choosing $f(t,x) = \theta(t)l(x)$ and $p(t) = 0 = q(t)$. Then we have $\sigma = 0$ and $\delta = 1$. $\square$

The following lemma will be crucial in our arguments.

Lemma 3.3. *Suppose that $f(t,x) \geq 0$ for all $(t,x) \in [0,1] \times \mathbb{R}$ and $I_{ij}(x) \leq 0$ for all $x \in \mathbb{R}$, $i = 1,2,\ldots,n$ and $j = 1,2,\ldots,m$. If u is a classical solution of (1), then $u(t) \geq 0$ for all $t \in [0,1]$.*

Proof. If u is a classical solution of (1), then for all $v \in X$ we have

$$\begin{aligned}
&(-1)^n \int_0^1 u^{(2n)}(t)v(t)dt + \cdots + \int_0^1 u^{(4)}(t)v(t)dt - \int_0^1 (p(t)u'(t))'v(t)dt \\
&+ \int_0^1 q(t)u(t)v(t)dt - \lambda \int_0^1 f(t,u(t))v(t)dt = 0.
\end{aligned}$$

Let $v(t) = \max\{-u(t), 0\}$ for all $t \in [0,1]$, clearly $v(t) \geq 0$, $v \in X$ and by n times integration by parts we have

$$\begin{aligned}
0 =&(-1)^n \sum_{j=1}^{m} \int_{t_j}^{t_{j+1}} u^{(2n)}(t)v(t)dt + \cdots + \sum_{j=1}^{m} \int_{t_j}^{t_{j+1}} u^{(4)}(t)v(t)dt - \int_0^1 (p(t)u'(t))'v(t)dt \\
&+ \int_0^1 q(t)u(t)v(t)dt - \lambda \int_0^1 f(t,u(t))v(t)dt \\
=&(-1)^n \sum_{j=1}^{m} u^{(2n-1)}(t)v(t)|_{t_j}^{t_{j+1}} + (-1)^{n-1} \sum_{j=1}^{m} u^{(2n-2)}(t)v'(t)|_{t_j}^{t_{j+1}} + \cdots - \sum_{j=1}^{m} u^{(n)}(t)v^{(n-1)}(t)|_{t_j}^{t_{j+1}} \\
&+ \int_0^1 u^{(n)}(t)v^{(n)}(t)dt + \cdots + \int_0^1 u''(t)v''(t)dt + \int_0^1 p(t)u'(t)v'(t)dt \\
&+ \int_0^1 q(t)u(t)v(t)dt - \lambda \int_0^1 f(t,u(t))v(t)dt \\
=&(-1)^{n-1} \sum_{j=1}^{m} \Delta u^{(2n-1)}(t_j)v(t_j) + (-1)^{n-2} \sum_{j=1}^{m} \Delta u^{(2n-2)}(t_j)v'(t_j) + \cdots + \sum_{j=1}^{m} \Delta u^{(n)}(t_j)v^{(n-1)}(t_j) \\
&+ \int_0^1 u^{(n)}(t)v^{(n)}(t)dt + \cdots + \int_0^1 u''(t)v''(t)dt + \int_0^1 p(t)u'(t)v'(t)dt \\
&+ \int_0^1 q(t)u(t)v(t)dt - \lambda \int_0^1 f(t,u(t))v(t)dt \\
\leq& \sum_{j=1}^{m} \mu I_{nj}(u^{(n-1)}(t_j)) + \sum_{j=1}^{m} \mu I_{(n-1)j}(u^{(n-2)}(t_j)) + \cdots + \sum_{j=1}^{m} \mu I_{1j}(u(t_j)) \\
&- \int_0^1 (v^{(n)}(t))^2 dt - \cdots - \int_0^1 (v''(t))^2 dt - \int_0^1 p(t)(v'(t))^2 dt \\
&- \int_0^1 q(t)(v(t))^2 dt - \lambda \int_0^1 f(t,u(t))v(t)dt \\
\leq& - \|v\|_X^2 \leq 0.
\end{aligned}$$

So $v(t) = 0$ for $t \in [0,1]$. $\square$

Put

$$G_{i,c} = \sum_{j=1}^{m} \min_{|\varepsilon| \leq c} \int_0^{\varepsilon} I_{ij}(x)dx, \quad \forall c > 0, \ i = 1,2,\ldots,n.$$

Our other main result is as follows.

Theorem 3.4. *Assume that there exist three positive constants* c_1, c_2, d *with* $\frac{\pi c_1}{(n-1)\left[\frac{1}{\alpha^3}+\frac{1}{(1-\beta)^3}\right]^{\frac{1}{2}}} < d < \sqrt{\frac{k}{2}}c_2$ *such that*

(B_1) $f(t,x) \geq 0$, $\forall (t,x) \in [0,1] \times [0,c_2]$;

(B_2) $\frac{\int_0^1 F(t,c_1)dt}{c_1^2} < \frac{2}{3}k\frac{\int_\alpha^\beta F(t,d)dt}{d^2}$;

(B_3) $\frac{\int_0^1 F(t,c_2)dt}{c_2^2} < \frac{k}{3}\frac{\int_\alpha^\beta F(t,d)dt}{d^2}$.

Then for every

$$\lambda \in \Lambda' = \left(\frac{3\delta^2\pi^2 d^2}{k\int_\alpha^\beta F(t,d)dt}, \delta^2\pi^2 \min\left\{\frac{2c_1^2}{\int_0^1 F(t,c_1)dt}, \frac{c_2^2}{\int_0^1 F(t,c_2)dt}\right\}\right)$$

and for every non-positive continuous function I_{ij}, $i = 1,2,\ldots,n$, $j = 1,2,\ldots,m$, *there exists*

$$\rho^* = \frac{1}{2}\min_{1\leq i\leq n}\left\{\frac{\lambda\int_0^1 F(t,c_1)dt - 2\delta^2\pi^2c_1^2}{G_{i,c_1}}, \frac{\lambda\int_0^1 F(t,c_2)dt - \delta^2\pi^2c_2^2}{G_{i,c_2}}\right\}$$

such that, for each $\mu \in (0,\rho^*)$ *the problem* (1) *has at least three classical solutions* u_q, $q = 1,2,3$, *such that* $0 \leq \|u_q\|_\infty \leq c_2$.

Proof. Without loss of generality, we can assume $f(t,x) \geq 0$ for all $(t,x) \in [0,1] \times \mathbb{R}$. Fix λ, I_{ij} and μ as in the conclusion and take X, Φ, Ψ and w as in the proof of Theorem 3.1. Put $r_1 = 2\delta^2\pi^2c_1^2$ and $r_2 = 2\delta^2\pi^2c_2^2$. Therefore, one has $2r_1 < \Phi(w) < \frac{r_2}{2}$ and since $\mu < \rho^*$, one has

$$\begin{aligned}\frac{1}{r_1}\sup_{\Phi(u)<r_1}\Psi(u) \leq& \frac{1}{2\delta^2\pi^2c_1^2}\left(\int_0^1 F(t,c_1)dt - \frac{\mu}{\lambda}G_{1,c_1} - \cdots - \frac{\mu}{\lambda}G_{n,c_1}\right)\\ <&\frac{1}{\lambda} < \frac{k}{3\delta^2\pi^2}\frac{\int_\alpha^\beta F(t,d)dt}{d^2} \leq \frac{2}{3}\frac{\Psi(w)}{\Phi(w)},\end{aligned}$$

and

$$\begin{aligned}\frac{2}{r_2}\sup_{\Phi(u)<r_2}\Psi(u) \leq& \frac{1}{\delta^2\pi^2c_2^2}\left(\int_0^1 F(t,c_2)dt - \frac{\mu}{\lambda}G_{1,c_2} - \cdots - \frac{\mu}{\lambda}G_{n,c_2}\right)\\ <&\frac{1}{\lambda} < \frac{k}{3\delta^2\pi^2}\frac{\int_\alpha^\beta F(t,d)dt}{d^2} \leq \frac{2}{3}\frac{\Psi(w)}{\Phi(w)}.\end{aligned}$$

Therefore, conditions (j) and (jj) of Theorem 2.2 are satisfied. Finally, let u_1 and u_2 be two local minima for functional $\Phi - \lambda\Psi$. Then, u_1 and u_2 are critical points for $\Phi - \lambda\Psi$, and so, they are weak solutions and so classical solutions for the problem (1). Hence, owing to Lemma 3.3, we obtain $u_1(t) \geq 0$ and $u_2(t) \geq 0$ for all $t \in [0,1]$. So, one has $\Psi(su_1 + (1-s)u_2) \geq 0$ for all $s \in [0,1]$. From Theorem 2.2 the functional $\Phi - \lambda\Psi$ has at least three distinct critical points which are weak solutions of (1). This completes the proof. □

Example 3.5. Consider the following problem

$$\begin{cases}(-1)^n u^{(2n)}(t) + \cdots + u^{(4)}(t) - u''(t) + (t-1)^2u(t) = \lambda u^2(3-4u)\sin(\pi t), \quad t \neq t_1, t \in (0,1)\\ \Delta(u^{(n)}(t_1)) = \mu(1 - \sqrt[3]{u(t_1)})\\ -\Delta(u^{(n+1)}(t_1)) = \mu(1 + \sqrt{u'(t_1)})\\ (-1)^i\Delta(u^{(n+i)}(t_1)) = \mu(1), \quad i = 2,3,\ldots,n-1\\ u(0) = u(1) = u'(0) = u'(1) = \cdots = u^{(n-2)}(0) = u^{(n-2)}(1) = 0 = u^{(n)}(0) = u^{(n)}(1).\end{cases}$$

It is sufficient to apply Theorem 3.1 by choosing $c = \frac{1}{64}$ and $d = \frac{1}{2}$ we have

$$\delta = 1, \ \ \|p\|_\infty = 1, \ \ \|q\|_\infty = 1, \ \ k = \frac{4\pi^2}{(n-1)\left(\frac{(2n-2)!}{(n-2)!}\right)^2 \left(1 + \frac{1}{\pi^2} + \frac{1}{\pi^4}\right)\left[\frac{1}{\alpha^{2n-1}} + \frac{1}{(1-\beta)^{2n-1}}\right]}.$$

Then for each

$$\lambda \in \Lambda = \left(\frac{(2n-2)\left(\frac{(2n-2)!}{(n-2)!}\right)^2 \left(\pi + \frac{1}{\pi} + \frac{1}{\pi^3}\right)\left[\frac{1}{\alpha^{2n-1}} + \frac{1}{(1-\beta)^{2n-1}}\right]}{\cos(\alpha\pi) - \cos(\beta\pi)}, 2015\right]$$

and for each $0 < \mu < \frac{9.8696 - 0.00489\lambda}{76}$, the above problem admits at least three non-trivial solutions.

References

[1] Ghasem A Afrouzi, Armin Hadjian, and Vicenţiu D Rădulescu. Variational approach to fourth-order impulsive differential equations with two control parameters. *Results in Mathematics*, 65:371–384, 2014.

[2] Ghasem A Afrouzi, Saeid Shokooh, and Armin Hadjian. Existence of solutions to dirichlet impulsive differential equations through a local minimization principle. *Electronic Journal of Differential Equations*, 2014(147):1–11, 2014.

[3] Gabriele Bonanno and Pasquale Candito. Non-differentiable functionals and applications to elliptic problems with discontinuous nonlinearities. *Journal of Differential Equations*, 244(12):3031–3059, 2008.

[4] Gabriele Bonanno and Salvatore A Marano. On the structure of the critical set of non-differentiable functions with a weak compactness condition. *Applicable Analysis*, 89(1):1–10, 2010.

[5] Gabriele Bonanno, Beatrice Di Bella, and Johnny Henderson. Existence of solutions to second-order boundary-value problems with small perturbations of impulses. *Electron. J. Differ. Equ*, 126:2013, 2013.

[6] Lambertus A Peletier, William C Troy, and RCAM Van der Vorst. Stationary solutions of a fourth order nonlinear diffusion equation. *Differential equations*, 31(2):301–314, 1995.

[7] Giorgio Talenti. Some inequalities of sobolev type on two-dimensional spheres. In *General Inequalities 5: 5th International Conference on General Inequalities, Oberwolfach, May 4–10, 1986*, pages 401–408. Springer, 1987.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, OR–295, pp. 621–625

Solving Time-Delay optimal control problems via Artificial Neural Networks

Tahereh Jabbari Khanbehbin[a,*]

[a] *Department of Mathematics and Statistics, Faculty of Basic Sciences, Gonbad Kavous University, Gonbad Kavous, Iran*

Article Info

Keywords:
Time-delay optimal control problems
Artificial neural networks
Unconstrained optimization problem

2020 MSC:
93-10

Abstract

This article presents a new approach for solving the Optimal Controls of linear time delay systems with a quadratic cost functional. In this study, the Artificial Neural Networks are employed for convert delay optimal control problem to a unconstrained optimization problem. Then by using an optimization algorithm, the optimal control law is obtained. Finally, Illustrative examples are included to demonstrate the validity and applicability of the technique.

1. Introduction

The control of systems with time-delay has been of considerable concern. Delays occur frequently in biological, chemical, electronic, engineering, transportation systems and so on. Therefore, there are many attempts available in the literature to approximately solve this problem.
Oğuztöreli [6], in1963, was one of the pioneers in the analytical-based approach for time-delay Optimal Control Problems (OCPs). For the first time, Kharatishvili [4] generalized the Pontryagin maximum principle for this type of problems. The system resulting his work is a two-point boundary value problem (TPBVP) involving both advance and delay terms whose exact solution, except in very special cases, is very difficult. Therefore, the main object of all computational aspects of optimal time-delay systems has been to devise a methodology to avoid the solution of the mentioned TPBVP.
Also, Artificial Neural networks (ANNs) are considerable as a effective tools for function approximation, recently. For the first time, in 1944, two researchers from Chicago University named McCullough and Walter Pitts presented the first model of neural networks. The perceptron was the first trainable neural network proposed by Cornell University psychologist Frank Rosenblatt in 1957. Mathematicians proved that continuous functions can be approximated by a multi-layer perceptron on the basis of a compact set of R^n. In a theorem in [1], Gybenko proves that a Neural Network approximation with a sigmoid active function can approximate continuous functions succesfully. For the first time, in order to solve PDEs and ODEs, Lagaris et al. proposed using neural networks [2]. Effati and Pakdaman in [8] used

*Talker
Email address: jabbari.tahere@gmail.com (Tahereh Jabbari Khanbehbin)

ANNs for approximating the state, co-state, and control functions for optimal control problems. Sabouri et al. in [3] can solve fractional optimal control problems with Neural Networks and etc. In another work Effati et al. Recently, Bhagya and Dash in [7] discussed a variety of applications of ANN to the modeling of nonlinear problems in food engineering.
Here, we attempt to implement the ability of neural networks to approximate time-delay OCPs. This paper is organized as follow:
The main results are discussed in 2, in this section we design a new a new neural network for solving time-delay OCPs. In 3, we give a numerical example to demonstrate the effectiveness and accuracy of the proposed technique. Finally, with the conclusion in 4, we end the article.

2. Main results

Consider the linear system with delay in the state variable

$$\begin{cases} \dot{x}(t) = Ax(t) + A_1x(t-\tau) + Bu(t), & t_0 \leq t \leq t_f, \\ x(t) = \phi(t), & t_0 - \tau \leq t \leq t_0, \end{cases} \tag{1}$$

where $u(t)$ in $PC([t_0, t_f], \mathbb{R}^n)$ and $x(t)$ in $PC^1([t_0 - \tau, t_f], \mathbb{R}^n)$ are the control and state variables, respectively. In fact, the parameter $\tau > 0$ is nonnegative and indicates the time delay. Furthermore, the initial state function $\phi(t)$ is continuous in $C([t_0 - \tau, t_0], \mathbb{R}^n)$, and finally, the matrices A, B, and A_1 are real constants with appropriate dimensions. For $t \in [t_0, t_f]$, our aim is to obtain, $u^*(t)$, the optimal control law minimizing the quadratic cost function

$$J = \frac{1}{2}\int_{t_0}^{t_f} (u^T(t)Ru(t) + x^T(t)Qx(t))dt + \frac{1}{2}x^T(t_f)Q_f x(t_f), \tag{2}$$

in which $R \in \mathbb{R}^{m\times n}$ is a positive definite matrix and Q and $Q_f \in \mathbb{R}^{n\times n}$ are positive semi-definite matrices.
For time-delay OCPs, it follows from [4] that the pontryagin maximum principle provides necessary conditions of optimality for the problem (1) and (2) as follows:

$$\begin{cases} \dot{x}(t) = Ax(t) + A_1x(t-\tau) - BR^{-1}B^T\lambda(t), & t_0 \leq t \leq t_f, \\ \dot{\lambda}(t) = \begin{cases} -Qx(t) - A^T\lambda(t) - A_1^T\lambda(t+\tau), & t_0 \leq t \leq t_f - \tau, \\ -Qx(t) - A^T\lambda(t), & t_f - \tau < t \leq t_f, \end{cases} \\ x(t) = \phi(t), \quad t_0 - \tau \leq t \leq t_0, \\ \lambda(t_f) = Q_f x(t_f). \end{cases} \tag{3}$$

The Hamiltonian function from which the above conditions are derived is

$$H(x,u,\lambda,t) = \lambda^T(t)[Ax(t) + Bu(t) + A_1x(t-\tau) + \frac{1}{2}x^T(t)Qx(t) + \frac{1}{2}u^T(t)Ru(t)], \tag{4}$$

where $\lambda(t) \in PC^1([t_0, t_f], \mathbb{R}^n)$ is called co-state vector. Moreover,

$$u^*(t) = -R^{-1}B^T\lambda(t), \tag{5}$$

for $t_0 \leq t \leq t_f$, is the optimal control law. We recall that the system (3) is a TPBVP with both time-advance and time-delay terms. Unfortunately, in general, this problem does not have any analytical solution. Therefore, providing an efficient method for solving this difficult problem numerically is very important.

For solving TPBVP (3), we suggested following approximations based on ANN for state and co-state variables:

$$x_N(t, W_x) = \begin{cases} \phi(t) + (\psi(t) - \psi(t_0))N_x(t, W_x), & t \geq t_0 \\ \phi(t), & t \leq t_0 \end{cases}$$

$$\lambda_N(t, W_\lambda) = \begin{cases} N_\lambda(t, W_\lambda), & \text{اگر} \quad Q_f = 0 \\ (t - t_f)N_\lambda(t, W_\lambda) + Q_f x(t_f), & \text{اگر} \quad Q_f \neq 0 \end{cases} \tag{6}$$

(7)

where x_N and λ_N are satisfying in initial and final conditions. Also,

$$N(t, W) = \sum_{i=1}^{k} v^i \sigma(\theta^i), \qquad \theta^i = w^i t + b^i, \tag{8}$$

is a perceptron ANN with two layers. W_x and W_λ are weight vectors according to input, output and bias weightes for $x(t)$ and $\lambda(t)$. σ is considered as an arbitrary activation function, here we implemented the sigmoid function in the numerical examples, as follows:

$$\sigma(x) = \frac{1}{1 + e^{-x}}, \tag{9}$$

With proposing this approximation functions for $x(t)$ and $\lambda(t)$ and substituting them in TPBVP (3), we have

$$\begin{cases} \dot{x}_N(t, W_x) = Ax_N(t, W_x) + A_1 x_N(t - \tau, W_x) - S\lambda_N(t, W_\lambda), & t_0 \leqslant t \leqslant t_f, \\ \dot{\lambda}_N(t, W_\lambda) = \begin{cases} -Qx_N(t, W_x) - A^T\lambda_N(t, W_\lambda) - A_1^T\lambda_N(t + \tau, W_\lambda), & t_0 \leqslant t < t_f - \tau, \\ -Qx_N(t, W_x) - A^T\lambda_N(t, W_\lambda), & t_f - \tau \leqslant t \leqslant t_f, \end{cases} \\ x_N(t) = \phi(t), & t_0 - \tau \leq t \leqslant t_0, \\ \lambda_N(t_f, W_\lambda) = Q_f x_N(t_f, W_x), \end{cases} \tag{10}$$

For solving (10), we introduce the following error function

$$\begin{cases} E_x(t, W) = \Big(\dot{x}_N(t, W_x) - \big(Ax_N(t, W_x) + A_1 x_N(t - \tau, W_x) - S\lambda_N(t, W_\lambda)\big)\Big)^2, \\ t_0 \leqslant t \leqslant t_f, \\ E_\lambda(t, W) = \begin{cases} \Big(\dot{\lambda}_N(t, W_\lambda) - \big(-Qx_N(t, W_x) - A^T\lambda_N(t, W_\lambda) - A_1^T\lambda_N(t + \tau, W_\lambda)\big)\Big)^2, \\ t_0 \leqslant t < t_f - \tau, \\ \Big(\dot{\lambda}_N(t, W_\lambda) - \big(-Qx_N(t, W_x) - A^T\lambda_N(t, W_\lambda)\big)\Big)^2, & t_f - \tau \leqslant t \leqslant t_f, \end{cases} \end{cases} \tag{11}$$

where $W = (W_x, W_\lambda)$ contains all the weights of the approximate functions. Finally, we write the neural network error function as

$$R(t, W) = E_x(t, W) + E_\lambda(t, W) \tag{12}$$

Now, in order to minimize the weights of the neural network, discretize the interval $[t_0, t_f]$ with m points $t_k, k = 1, 2, \ldots, m$, then We are solving the following unconstrained optimization problem

$$\min R(W) = \sum_{k=1}^{m} E_x(t_k, W) + E_\lambda(t_k, W) \tag{13}$$

Any classical mathematical optimization algorithm such as the fastest reduction, Newton, conjugate gradient,... and heuristic approaches such as Genetic or Ant algorithms, can be used to solve this problem. We have used of matlab optimization packages.

3. Numerical results

Example 3.1. Consider the time-delay system

$$\begin{cases} \dot{x} = u(t) - x(t-1), & 0 \leq t \leq 1, \\ x(t) = 1, & -1 \leq t \leq 0, \end{cases} \tag{14}$$

to minimize this quadratic cost functional

$$J = \int_0^1 [\frac{1}{2}x^2(t) + \frac{1}{2}u^2(t)]dt. \tag{15}$$

Now, our aim is to obtain the optimal control, $u(t)$, subject to (14) that minimizes (15).
the necessary conditions of optimality for the problem (14) and (15) are as follow

$$\begin{cases} \dot{x}(t) = -x(t-1) + u(t) \\ \dot{\lambda}(t) = \begin{cases} -x(t) + \lambda(t+1), & t = 0, \\ -x(t) \quad 0 < t \leq 1, \end{cases} \\ x(t) = 1, \quad -1 \leq t \leq 0, \\ \lambda(1) = 0, \end{cases}$$

and the optimal control law is

$$u^*(t) = -\lambda(t)$$

With choosing $\psi(t) = t.$ we have following approximation functions for $x(t)$ and $\lambda(t)$,

$$x_N(t, W_x) = \begin{cases} 1 + tN_x(t, W_x), & t \geq 0 \\ 1, \quad t \leq 0, \end{cases}$$

$$\lambda_N(t, W_\lambda) = N_\lambda(t, W_\lambda), \quad 0 \leq t \leq 1$$

To solve this problem, we trained the neural network in the interval $[0, 1]$ with 100 training points and $k = 10$ neurons. The exact solutions for $u(t)$ and $x(t)$ are, respectively, obtained as follows:

$$u^*(t) = 1 + \frac{1}{\cosh(1)}(\sinh(t-1) - \cosh(t)), \tag{16}$$

$$x^*(t) = \frac{1}{\cosh(1)}(\cosh(t-1) - \sinh(t)). \tag{17}$$

Moreover, it follows from [5] that the optimal value of cost functional is $J^* = 0.1480542786$. It can be shown that the approximate value of the cost functional calculated by the proposed ANN method is equal to $J = 0.1480543001$. It is clear that the approximate value of J is very close to the optimal value. Also, we depict the simulation curves of the trajectory of $x(t)$, control variable $u(t)$, and their exact values in Figs. 1 and 2.

4. Conclusion

In this paper, we propose approximation functions based on neural network model for stste and co-state variables, to solve time-delay OCPs. This technique convert time-delay OCPs to a unconstrained optimization problem that can be easily solved using an optimization algorithm. The numerical results were presented to illustrate the high accuracy and efficiency of our proposed approach. Further research can be done on the extension of using Neural networks for solving time-delay OCPs with time dependent delays in the control and state.

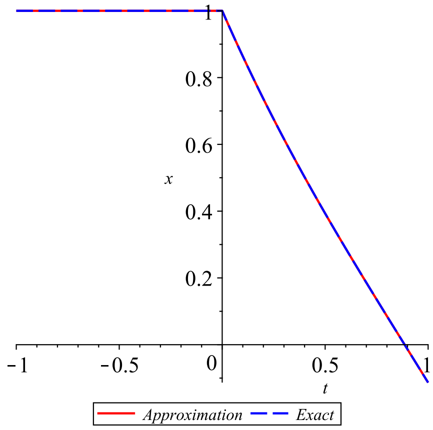


Fig. 1. Approximation and exact values of state variable for Example 3.1

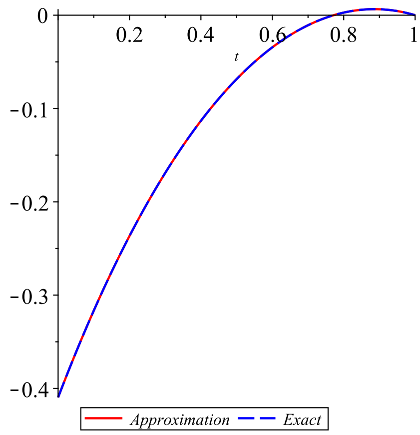


Fig. 2. Approximation and exact values of control variable for Example 3.1

References

[1] G. Gybenko. Approximation by superposition of sigmoidal functions. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989.

[2] D. Fotiadis I. Lagaris, A. Likas. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9(5):987–1000, 1998. doi: 10.1109/72.712178.

[3] M. Pakdaman J. Sabouri, S. Effati. A neural network approach for solving a class of fractional optimal control problems. *Neural Processing Letters*, 45(1):59–74, 2017. doi: 10.1007/s11063-016-9510-5.

[4] G. Kharatishvili. The maximum principle in the theory of optimal processes involving delay. *Dokl. Akad. Nauk*, 136:39–42, 1961.

[5] T. Mueller. Optimal control of linear systems with time lag. In *Third Annual Allerton Conf. on Circuit and System Theory*, volume I11, pages 339–345, Univ. Illinois, Urbana, 1965.

[6] M.N. Oğuztöreli. A time optimal control problem for systems described by differential difference equations,. *SIAM J Appl Math, Series A: Control.*, 1(3):290–310, 1963. doi: 10.1137/0301017.

[7] G. Bhagya Raj and K.K. Kshirod K Dash. Comprehensive study on applications of artificial neural network in food process modeling. *Critical Reviews in Food Science and Nutrition*, 62(10):2756–2783, 2022. doi: 10.1080/10408398.2020.1858398.

[8] M. Pakdaman S. Effati. Optimal control problem via neural networks. *Neural Computing and Applications*, 23(7-8):2093–2100, 2013. doi: 10.1007/s00521-012-1156-2.

Gonbad Kavous University

The 4th National Congress on Mathematics and Statistics

https://conf.gonbad.ac.ir/msc1402

The 4th National Congress on Mathematics and Statistics, NA–305, pp. 626–630

Spectral Method for solving Wave Equation

Mehdi Shahini[a,*]

[a]*Department of Mathematics and Statistics, Gonbad Kavous University, Gonbad Kavous, Iran*

Article Info

Keywords:
Partial Differential Equation
Spectral Methods
Wave Equation
Chebyshev polynomials

2020 MSC:
35L02
35L04
35L05
35P05

Abstract

The wave equation is a fundamental partial differential equation that describes the propagation of waves, such as sound or light waves, through a medium. In this paper, a spectral collocation method based on shifted Chebyshev polynomials and Gaussian nodes is applied to solve wave equation. The obtained results show efficiency and performance of the method.

1. Introduction

Partial differential equations (PDEs) are useful tools for describing the natural phenomena of science and engineering models. Because of this connection with phenomena in the physical world, PDEs are of widespread interest. For instance, the heat flow and wave propagation are some phenomena in physics which are described via PDEs with initial and boundary conditions. The diffusion of neutrons in nuclear reactor dynamics, population models, the dispersion of a chemically reactive material, and many physical phenomena of fluid dynamics, quantum mechanics, electricity, and so on are also governed by PDEs.

Scientists and mathematicians have become actively involved in the study of countless problems offered by PDEs. The primary reason for this research was that it plays a vital role in modern mathematical sciences, mainly in applied physics, mathematical modeling, and engineering. With the development of PDEs, several methods such as the characteristics method, spectral methods, and perturbation techniques have been employed to get the solution of problems. However, there is no general method to find analytical solutions of some PDEs. So, finding new numerical techniques to obtain the solutions of equations is of concern. Traditional methods like finite differences or finite elements have been widely used to solve PDEs. However, spectral methods offer a powerful alternative, particularly when dealing with smooth problems or requiring high precision. Among spectral methods, those based on orthogonal functions stand out for their accuracy and efficiency [1–3].

In this paper, we investigate a new numerical scheme for solving PDEs. The scheme is based on collocation method based on Chebyshev polynomials.

*Talker
Email address: mehdi.shahini@gonbad.ac.ir (Mehdi Shahini)

2. Governing equations

The wave equation, a fundamental partial differential equation (PDE) in physics, models how waveforms such as sound or light propagate through different media. The prototype for all hyperbolic partial differential equations is the one-way wave equation. It is an one-dimensional first-order linear PDE given by:

$$u_t + cu_x = 0, \tag{1}$$

where $u(x,t)$ represents the wave function, c is a constant which is called the speed of propagation along the characteristic, t is time, and x is the spatial variable. The subscript denotes differentiation, i.e., $u_t = \frac{\partial u}{\partial t}$ and $u_x = \frac{\partial u}{\partial x}$.
The equation comes with appropriate initial and boundary conditions:

$$u(x,0) = u_0(x), \quad a \leq x \leq b \tag{2}$$

$$u(a,t) = u(b,t) = v(t), \quad t \geq 0 \tag{3}$$

The exact solution of Eq. (1) is:

$$u(x,t) = u_0(x - ct) \tag{4}$$

The more general hyperbolic equation is:

$$u_t + cu_x = f(x,t,u), \tag{5}$$

3. Proposed method

In this paper, a new spectral collocation method is applied to solve PDEs. In traditional spectral methods, the unknown function is expanded over spatial variable and other methods are applied for time parameter.
In this method, orthogonal functions are used in both time and spatial to approximate the function $u(x,t)$. So, we have

$$u(x,t) \approx u_{MN}(x,t) = \sum_{m=0}^{M} \sum_{n=0}^{N} a_{m,n} \varphi_n(t) \phi_m(x). \tag{6}$$

Here, $\varphi_n(t)$ and $\phi_m(x)$ are appropriate orthogonal functions defined on the problem domain.
By choosing collocation nodes and substituting them on the governing equations, we have a set of algebraic equations which can be solved by proper methods.

4. Numerical results

In this section, some numerical examples are given to show the efficiency of the method. In the examples, shifted Chebyshev polynomials, are chosen as the both spatial and time trial functions in (6):

$$\phi_m(x) = T_m^S(x), \tag{7}$$

$$\varphi_n(t) = T_n^S(t). \tag{8}$$

Also, shifted Chebyshev Gauss Lobatto and shifted Chebyshev Gauss Radau nodes are used for spatial collocation and time collocation nodes, respectively:

$$s_m \text{ :roots of } T_{M+1}^S(x) - T_{M-1}^S(x), \quad m = 0, 1, \ldots, M, \tag{9}$$

$$\tau_n \text{ :roots of } T_{N+1}^S(t) + T_N^S(t), \quad n = 0, 1, \ldots, N, \tag{10}$$

To solve the problems, at first, $u(x,t)$ is substituted by the approximated function, $u_{MN}(x,t)$, in the governing equations. Then, the equations are evaluated at collocation nodes to construct the system of algebraic equations.

Example 4.1. Consider the following problem:

$$u_t + 2\pi u_x = 0, \quad 0 \le x \le 2\pi, \quad t \ge 0, \tag{11}$$

$$u(x,0) = \exp(\sin x) \quad 0 \le x \le 2\pi, \tag{12}$$

$$u(0,t) = u(2\pi,t) = \exp(\sin(-2\pi t)), \quad t \ge 0. \tag{13}$$

The exact solution of the problem is:

$$u(x,t) = \exp(\sin(x - 2\pi t)). \tag{14}$$

The results for $M = 32$ and $N = 16$ are shown in Figs. 1 and 2.

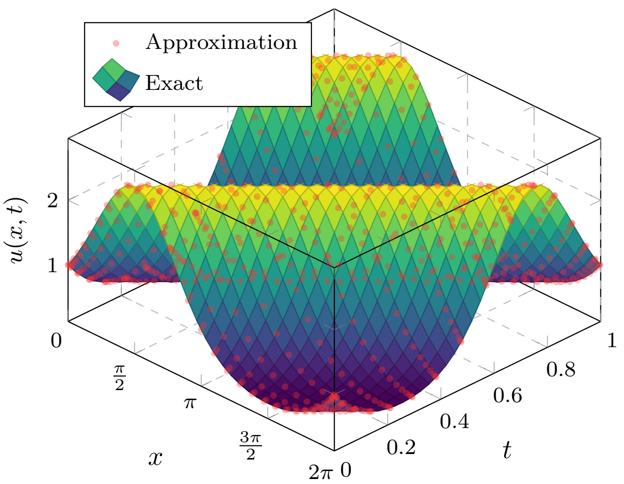
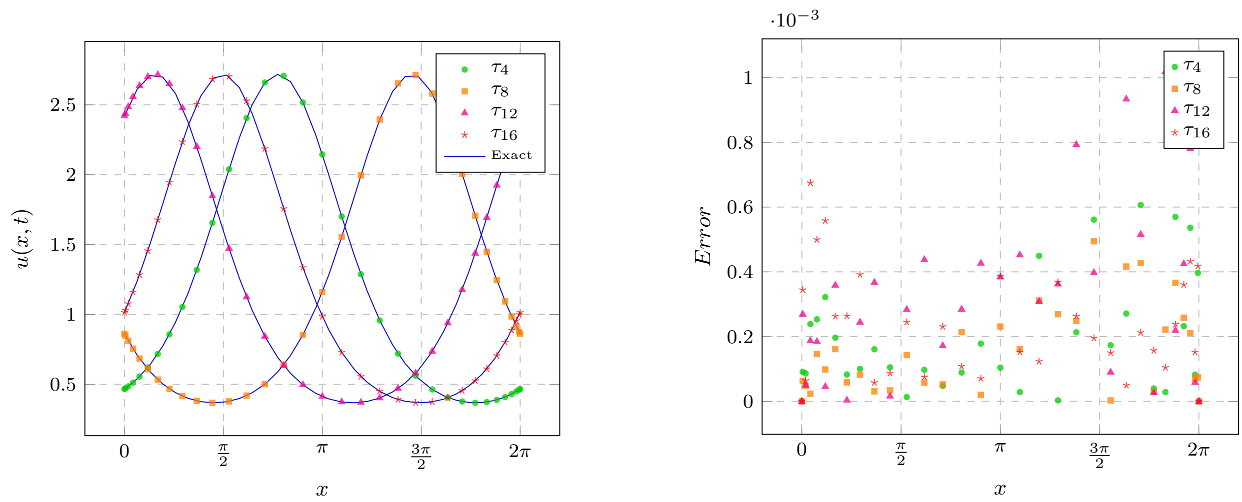


Fig. 1. Solution of Example 4.1.



(a) Solution of the problem for specified values of t.

(b) Error between obtained solution and the exact one, $|u - u_{M,N}|$.

Fig. 2. Obtained results for Example 4.1.

Example 4.2. Consider the following problem:

$$u_t + 2\pi u_x = x + t, \quad 0 \le x \le \pi, \quad t \ge 0, \tag{15}$$

$$u(x,0) = 0 \quad 0 \le x \le \pi, \tag{16}$$

$$u(0,t) = 0, \quad u(\pi,t) = \pi t, \quad t \ge 0. \tag{17}$$

The exact solution of the problem is:

$$u(x,t) = xt. \tag{18}$$

The results for $M = 32$ and $N = 16$ are shown in Figs. 3 and 4.

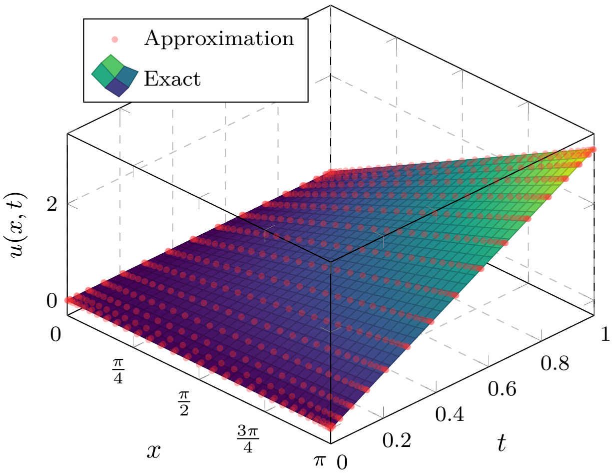
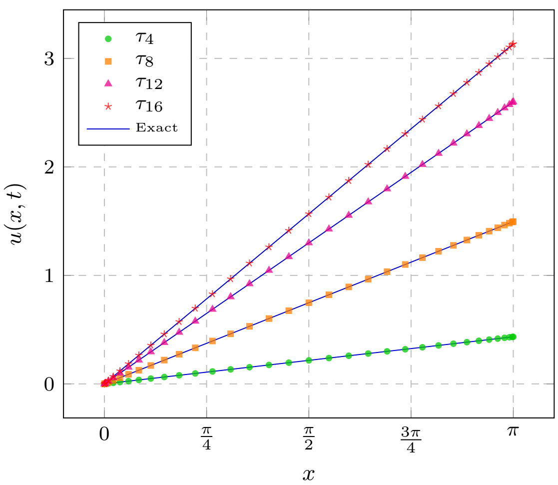
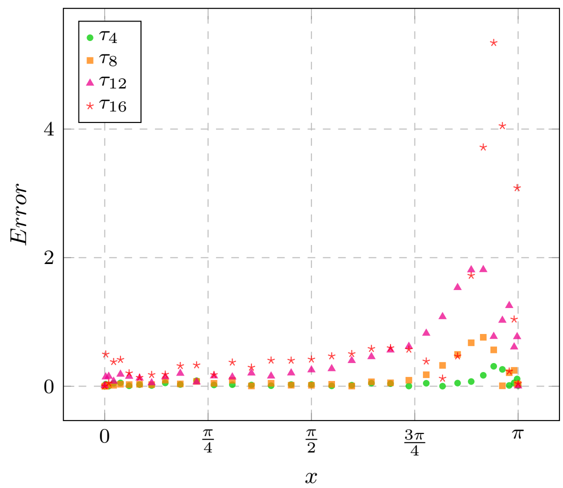


Fig. 3. Solution of Example 4.2.



(a) Solution of the problem for specified values of t.



(b) Error between obtained solution and the exact one, $|u - u_{M,N}|$.

Fig. 4. Obtained results for Example 4.2.

5. Conclusion

In this article, we explored the application of a new numerical scheme for solving partial differential equations using a collocation method based on Chebyshev polynomials.

The proposed spectral collocation method utilizes orthogonal functions in both spatial and temporal dimensions to approximate the solution $u(x, t)$. By selecting appropriate collocation nodes and solving the resulting algebraic equations, accurate solutions to PDEs are obtained.
Numerical examples demonstrate the efficiency of this method, with shifted Chebyshev polynomials serving as trial functions. Overall, spectral methods based on orthogonal functions provide a powerful tool for tackling complex PDEs in diverse fields.

References

[1] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang. *Spectral Methods in Fluid Dynamic*. Springer-Verlag, New York, 1987.
[2] R. A. Frazer, W. P. Jones, and S. W. Skan. Approximations to functions and to the solutions of differential equations. *Aeronautical Research Council Reports & Memoranda*, 1799:1937–1943, 1937.
[3] John C Strikwerda. *Finite difference schemes and partial differential equations*. SIAM, 2004.