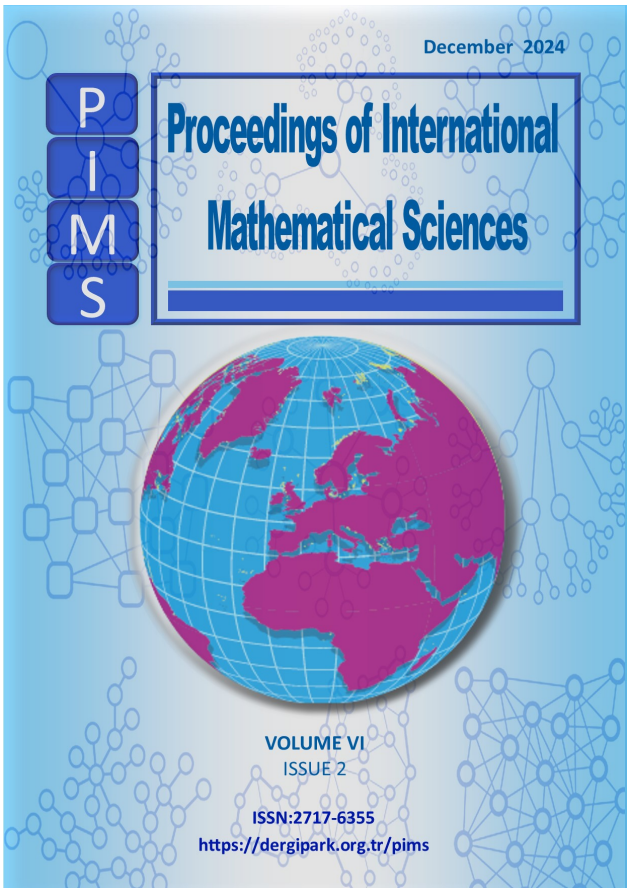
VOLUME VI ISSUE 2 December 2024
https://dergipark.org.tr/tr/pub/pims
ISSN:2717-6355

PROCEEDINGS OF INTERNATIONAL MATHEMATICAL SCIENCES



Publishing Board

İçindekiler

Proceedings of International Mathematical Sciences
ISSN: 2717-6355, URL: https://dergipark.org.tr/tr/pub/pims
Volume 6 Issue 2 (2024), Pages 44-53.
Doi: https://doi.org/10.47086/pims.1535676

ON SUBCLASS OF ANALYTIC FUNCTIONS DEFINED BY q-ANALOGUE OF MODIFIED TREMBLAY FRACTIONAL DERIVATIVE OPERATOR

ABDULLAH ALSOBOH*, AND WAGGAS GALIB ATSHAN**
*PHILADELPHIA UNIVERSITY, AMMAN, JORDAN. ORCID NUMBER: 0000-0002-5183-2654
**UNIVERSITY OF AL-QADISIYAH, DIWANIYAH, IRAQ. ORCID NUMBER: 0000-0002-7033-8993.

Abstract. In this research, by using the principle of quantum calculus, we introduce a modified fractional derivative operator $\mathcal{T}^{\xi,F}_{\mathfrak{q},\varsigma}$ of the analytic functions in the open unit disc $\diamond = \{\varsigma : \varsigma \in \mathbb{C}, |\varsigma| < 1\}$. The operator $\mathcal{T}^{\xi,F}_{\mathfrak{q},\varsigma}$ can then be used to introduce a new subclass of analytic functions $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. We present the necessary conditions for functions belonging to the subclass $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. Furthermore, we discuss a growth and distortion bounds, the convolution condition, and the radii of starlikeness. In addition, we present neighbourhoods problems involving the q-analogue of a modified Tremblay operator for functions in the introduced class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$.

1. Introduction

Let $\diamond = \{\varsigma : \varsigma \in \mathbb{C}, |\varsigma| < 1\}$ denote the open unit disc and $\mathcal{A}$ the class of functions $\hbar(\varsigma)$ of the form

$$\hbar(\varsigma) = \varsigma + \sum_{\varkappa=2}^{\infty} a_\varkappa \varsigma^\varkappa, \quad (\varsigma \in \diamond) \tag{1.1}$$

that are analytic in the open unit disc $\diamond$. Furthermore, let $\mathcal{S}$ be the subset of $\mathcal{A}$ consisting of one-to-one (univalent) functions in $\diamond$.

The convolution of functions $\hbar$ as in (1.1) and the function

$$y(\varsigma) = \varsigma + \sum_{\varkappa=2}^{\infty} \gamma_\varkappa \varsigma^\varkappa,$$

is defined by:

$$(\hbar * y)(\varsigma) \quad = \quad \hbar(\varsigma) * y(\varsigma) = \varsigma + \sum_{\varkappa=2}^{\infty} a_\varkappa \gamma_\varkappa \varsigma^\varkappa.$$

2020 *Mathematics Subject Classification.* Primary: 30C45.
Key words and phrases. Analytic functions; starlikeness; Radii of starlikeness; neighbourhoods problems; fractional operator; q-calculus.

Submitted on 20.08.2024, Accepted on 09.10.2024.
Communicated by Ayhan Esi and Nazlım Deniz Aral.

The fractional q-calculus is an extension of ordinary fractional calculus, and it has become increasingly popular in recent decades due to its wide range of applications in various fields of science and engineering, particularly mathematics ([12], [20]). The concept of fractional q-calculus was introduced by Al-Salam and Verma [4], Al-Salam [5], and Agrawal [1]. They also explored some basic properties of fractional q-derivatives. In addition, Isogawa et al. [14] investigated some fundamental properties of fractional q-derivatives. Several problems involving fractional q-calculus operators have recently been recognized ([2, 15, 16, 17, 21, 24, 25]). In 2011, Garg and Chanchlani [13] defined a q-analog of Saigo's fractional integrals. Two authors, Exton [10] and Gasper [11], have written books about q-calculus.

The following are the notations and definitions again for main terms in q-calculus, which may be found in Gasper and Rahman [11] and Purohit and Rania [18], as follows:

1) The q-shifted factorial $(\vartheta, \mathfrak{q})_\varkappa$ is defined for $\vartheta \in \mathbb{C}$ and $0 < \mathfrak{q} < 1$ by:

$$(\vartheta; \mathfrak{q})_\varkappa = (\mathfrak{q}^\vartheta; \mathfrak{q})_\varkappa = \begin{cases} \prod_{\iota=0}^{\varkappa-1} (1 - \vartheta \mathfrak{q}^\iota), & \varkappa > 0 \\ \prod_{\iota=0}^{\infty} (1 - \vartheta \mathfrak{q}^\iota), & \varkappa \to \infty. \end{cases} \tag{1.2}$$

Equivalently,

$$(\vartheta; \mathfrak{q})_\varkappa = \frac{\Gamma_\mathfrak{q}(\vartheta + \varkappa)(1 - \mathfrak{q})^\varkappa}{\Gamma_\mathfrak{q}(\vartheta)}, \tag{1.3}$$

where the q-gamma function (see for example Gasper and Rahman [11]), is given by

$$\Gamma_\mathfrak{q}(\vartheta) = \frac{(\mathfrak{q}, \mathfrak{q})_\infty}{(\mathfrak{q}^\vartheta, \mathfrak{q})_\infty (1 - \mathfrak{q})^{\vartheta - 1}}, \quad \vartheta \neq 0, -1, -2, \ldots . \tag{1.4}$$

2) [**?**] For $0 < \mathfrak{q} < 1$. The q-derivative, also known as the q-difference operator, of a function $\hbar$ is defined by

$$\partial_\mathfrak{q} \hbar(\varsigma) = \begin{cases} \frac{\hbar(\varsigma) - \hbar(\mathfrak{q}\varsigma)}{\varsigma - \mathfrak{q}\varsigma}, & \text{if} \quad \varsigma \neq 0, \\ \hbar'(0), & \text{if} \quad \varsigma = 0, \\ \hbar'(\varsigma), & \text{if} \quad \mathfrak{q} \to 1^-, \ \varsigma \neq 0. \end{cases} . \tag{1.5}$$

3) The q-Jackson's integral of a function $\hbar$ is defined by:

$$\int_0^\varsigma \hbar(\text{J}) \partial_\mathfrak{q} \text{J} \quad = \quad \varsigma (1 - \mathfrak{q}) \sum_{\varkappa=0}^{\infty} \mathfrak{q}^\varkappa \hbar(\mathfrak{q}^\varkappa \varsigma),$$

provided that the series converges.

In 2010, Purhot and Yadav [18] introduced fractional integral operator and fractional derivative operator by

Definition 1.1. [18] *The fractional integral operator $I^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma)$, which operates on a function $\hbar(\varsigma)$ of order ϑ $(\vartheta > 0)$, is defined as follows:*

$$I^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = \frac{1}{\Gamma_{\mathfrak{q}}(\vartheta)}\int_0^{\varsigma}(\varsigma - \tau\mathfrak{q})_{\vartheta-1}\hbar(\tau)\partial_{\mathfrak{q}}\tau,$$

Here, $\hbar(\varsigma)$ is an analytic function in a simply-connected region of the ς-plane that includes the origin.

Definition 1.2. [18] *The fractional derivative operator $D^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma)$ of a function $\hbar(\varsigma)$ of order ϑ $(0 \le \vartheta < 1)$ is defined as*

$$D^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = \partial_{\mathfrak{q}}I^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = \frac{1}{\Gamma_{\mathfrak{q}}(1-\vartheta)}\partial_{\mathfrak{q}}\int_0^{\varsigma}(\varsigma - \tau\mathfrak{q})_{\vartheta-1}\hbar(\tau)\partial_{\mathfrak{q}}\,\tau.$$

Definition 1.3. [18]*(Extended Fractional $\mathfrak{q}$-Derivative Operator) Under the hypotheses of Definition 2, the fractional $\mathfrak{q}$-derivative for a function $f(\varsigma)$ of order ϑ is defined by*

$$D^{\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = D^{m}_{\mathfrak{q},\varsigma}I^{m-\vartheta}_{\mathfrak{q},\varsigma}\hbar(\varsigma), \quad (m-1 \le \vartheta < m), \quad m \in \mathbb{N}_0 = \mathbb{N}\cup\{0\}.$$

By virtue of Definitions 1.1, 1.2 and 1.3, we have

$$I^{\vartheta}_{\mathfrak{q},\varsigma}\varsigma^{\varkappa} = \frac{\Gamma_{\mathfrak{q}}(\varkappa+1)}{\Gamma_{\mathfrak{q}}(\varkappa+\vartheta+1)}\varsigma^{\varkappa+\vartheta}, \quad (\varkappa \in \mathbb{N}, \vartheta > 0),$$

and

$$D^{\vartheta}_{\mathfrak{q},\varsigma}\varsigma^{\varkappa} = \frac{\Gamma_{\mathfrak{q}}(\varkappa+1)}{\Gamma_{\mathfrak{q}}(\varkappa-\vartheta+1)}\varsigma^{\varkappa-\vartheta}, \quad (\varkappa \in \mathbb{N}, 0 \le \vartheta < 1).$$

Now, let us define the $\mathfrak{q}$-analogue of the Tremblay operator. The modified $\mathfrak{q}$-Tremblay operator ofor analytic functions in the complex domain is then given by:

Definition 1.4. *For $0 < \gamma \le 1, 0 < \xi \le 1, 0 \le \xi - \gamma < 1, \xi \ge \gamma$ and $\hbar \in \mathcal{A}$. The $\mathfrak{q}$-analouge of Tremblay derivative operator can be defined by*

$$\Psi^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = \frac{\Gamma_{\mathfrak{q}}(\gamma)}{\Gamma_{\mathfrak{q}}(\xi)}\varsigma^{1-\gamma}D^{\xi-\gamma}_{\mathfrak{q},\varsigma}(\varsigma^{\xi-1}\hbar(\varsigma)).$$

Definition 1.5. *Let $\hbar \in \mathcal{A}$, the $\mathfrak{q}$-analouge of modified Tremblay operator denoted by $\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma} : \mathcal{A} \to \mathcal{A}$ and defined as*

$$\begin{aligned}\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma) &= \frac{[\gamma]_{\mathfrak{q}}}{[\xi]_{\mathfrak{q}}}\Psi^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\\ &= \frac{\Gamma_{\mathfrak{q}}(\gamma+1)}{\Gamma_{\mathfrak{q}}(\xi+1)}\varsigma^{1-\gamma}D^{\xi-\gamma}_{\mathfrak{q},\varsigma}(\varsigma^{\xi-1}\hbar(\varsigma))\\ &= \varsigma + \sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}a_{\varkappa}\varsigma^{\varkappa},\end{aligned}$$

where $0 < \gamma \le 1$, $0 < \xi \le 1$, $0 \le \xi - \gamma < 1$ and $\xi \ge \gamma$.

Remark. *We can conclude that, when we choose the parameters q, ξ and γ, the operator $\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}$ can lead to other operators results. Examples are presented for further illustration.*

1) *For $\xi = 1$ and $\gamma = 1 - \vartheta$, we get the operator $\Omega^{\vartheta}_{\mathfrak{q},\varsigma}$ studied by Purohit and Rania [18].*
2) *For $\mathfrak{q} \to 1^{-}$, then $\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma) = \mathcal{T}^{\xi,\gamma}_{\varsigma}f$ the modified Tremblay operator studied by Esa et.al [8].*

3) For $\xi = 1$, $\gamma = 1$ and $\mathfrak{q} \to 1^-$ we get the Tremblay operator $\mathcal{T}_{\varsigma}^{\xi,\gamma} f$ syudied by Tremblay [23].

Various authors, such as Alb Lupaș and Oros [3], Purohit and Rania [17], Atshan et al. [6], Seoudy and Aouf [22], Frasin and Darus [9], Ramadan and Darus [19], Elhaddad and Darus [7], and others, have conducted studies on different subfamilies of normalized analytic functions. These publications have introduced a novel subclass $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$ of $\mathcal{A}$. This subclass incorporates the operator $\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)$ and is represented as follows:

Definition 1.6. *The class of functions $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$ is denoted by $\hbar \in \mathcal{A}$ and satisfies the inequality:*

$$\left|\frac{1}{d}\left(\frac{\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right) + \vartheta\varsigma^2\partial_{\mathfrak{q}}^2\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right)}{(1-\vartheta)(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma) + \vartheta\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right)} - 1\right)\right| < F, \tag{1.6}$$

where $\varsigma \in \diamond$, $d \in \mathbb{C} \setminus \{0\}$, $0 < F \leq 1$, $0 \leq \vartheta \leq 1$, $0 < \gamma \leq 1$, $0 < \xi \leq 1$, $0 \leq \xi - \gamma < 1$, and $\xi \geq \gamma$.

2. Main Results

This section examines the conditions that must be met for equation (1.6) to yield the function $\hbar$ in the class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. It also highlights the significance of these criteria for functions in this class. Furthermore, it presents growth and distortion bounds, q-raddi of stralikness of order λ $(0 \leq \lambda < 1)$, and the neighborhood problems for the class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. The necessary and sufficient conditions for functions $\hbar \in \mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$ are first discussed in our theorem.

Theorem 2.1. *Let the function $\hbar$ as is in (1.1) belong to the class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$ if and only if the following inequality holds:*

$$\sum_{\varkappa=2}^{\infty} \frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta\,([\varkappa]_{\mathfrak{q}}-\mathfrak{q}) + \mathfrak{q}\,(1+\vartheta F|d|)\,) + F|d|\Big)|a_{\varkappa}| \leq F|d|. \tag{2.1}$$

Proof. Suppose $\hbar$ belongs to the set $\mathcal{A}$ and that inequality (2.1) is satisfied. Consequently, we arrive at the following expression:

$$\begin{aligned}
&\left|\frac{\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right) + \vartheta\varsigma^2\partial_{\mathfrak{q}}^2\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right)}{(1-\vartheta)\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right) + \vartheta\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}_{\mathfrak{q},\varsigma}^{\xi,\gamma}\hbar(\varsigma)\right)} - 1\right| \\
&= \left|\frac{\begin{array}{l}\varsigma + \sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa]_{\mathfrak{q}}a_{\varkappa}\varsigma^{\varkappa} + \\ \qquad \vartheta\left(\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa]_{\mathfrak{q}}[\varkappa-1]_{\mathfrak{q}}a_{\varkappa}\varsigma^{\varkappa}\right)\end{array}}{\begin{array}{l}(1-\vartheta)\left(\varsigma + \sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}a_{\varkappa}\varsigma^{\varkappa}\right) + \\ \qquad \vartheta\left(\varsigma + \sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa]_{\mathfrak{q}}a_{\varkappa}\varsigma^{\varkappa}\right)\end{array}} - 1\right| \\
&= \left|\frac{\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa-1]_{\mathfrak{q}}(\mathfrak{q}+\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q}))a_{\varkappa}\varsigma^{\varkappa}}{\varsigma + \sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}(1+\mathfrak{q}\vartheta[\varkappa-1]_{\mathfrak{q}})a_{\varkappa}\varsigma^{\varkappa}}\right|
\end{aligned}$$

$$\leq \frac{\sum_{\varkappa=2}^{\infty} \frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa-1]_{\mathfrak{q}}(\mathfrak{q}+\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q}))|a_{\varkappa}||\varsigma|^{\varkappa-1}}{1-\sum_{\varkappa=2}^{\infty} \frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}(1+\mathfrak{q}\vartheta[\varkappa-1]_{\mathfrak{q}})|a_{\varkappa}||\varsigma|^{\varkappa-1}} < F|d|.$$

When we consider values of ς on the real axis and let $\varsigma \to 1^{-}$, we obtain

$$\sum_{\varkappa=2}^{\infty} \frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)|a_{\varkappa}| < F|d|. \tag{2.2}$$

Conversely, suppose $\hbar \in \mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$, we obtain the following inequality

$$\left|\frac{1}{d}\left(\frac{\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)+\vartheta\varsigma^{2}\partial^{2}_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)}{(1-\vartheta)(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)+\vartheta\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)}-1\right)\right| > -F, \tag{2.3}$$

$$\mathfrak{R}e\left\{\frac{\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)+\vartheta\varsigma^{2}\partial^{2}_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)}{(1-\vartheta)\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)+\vartheta\varsigma\partial_{\mathfrak{q}}\left(\mathcal{T}^{\xi,\gamma}_{\mathfrak{q},\varsigma}\hbar(\varsigma)\right)}-1+F|d|\right\} > 0$$

This need to complete

$$\mathfrak{R}e\left\{\frac{\varsigma+\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}[\varkappa]_{\mathfrak{q}}(1+\vartheta[\varkappa-1]_{\mathfrak{q}})a_{\varkappa}\varsigma^{\varkappa}}{\varsigma+\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}(1+\vartheta\mathfrak{q}[\varkappa-1]_{\mathfrak{q}})a_{\varkappa}\varsigma^{\varkappa}}-1+F|d|\right\} > 0$$

or

$$\mathfrak{R}e\left\{\frac{F|d|\varsigma+\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)a_{\varkappa}\varsigma^{\varkappa}}{\varsigma+\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}(1+\vartheta\mathfrak{q}[\varkappa-1]_{\mathfrak{q}})a_{\varkappa}\varsigma^{\varkappa}}\right\} > 0.$$

The inequality can be expressed as follows, taking into account the real part of the expression $-e^{i\theta}$: $\mathfrak{R}e\left\{-e^{i\theta}\right\} \geq |e^{i\theta}| = -1$.

$$\frac{F|d|\mathsf{r}-\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)a_{\varkappa}\mathsf{r}^{\varkappa}}{\mathsf{r}-\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}(1+\vartheta\mathfrak{q}[\varkappa-1]_{\mathfrak{q}})a_{\varkappa}\mathsf{r}^{\varkappa}} > 0.$$

By employing the mean value theorem for the limit as r approaches 1^{-}, we derive the inequality 2.1. Thus, we have concluded the proof of Theorem 2.1. □

Corollary 2.2. *Assuming that the function $\hbar$ is of the form (1.1) and belongs to the class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$, then the following inequality can be expresse*

$$|a_{\varkappa}| \leq \frac{F|d|\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)}{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}, \tag{2.4}$$

for $\varkappa \geq 2$.

The following result will provide bounds on the growth and distortion of functions in the class $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$.

Theorem 2.3. *The following inequalities hold true for any function $\hbar$ in the class $\mathcal{D}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ when $|\varsigma|=\mathsf{r}<1$:*

$$\mathsf{r}-\frac{F|d|(1+\gamma)\mathsf{r}^2}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}\leq|\hbar(\varsigma)|\leq\mathsf{r}+\frac{F|d|(1+\gamma)\mathsf{r}^2}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)},\tag{2.5}$$

and

$$1-\frac{F|d|(1+\mathfrak{q})(1+\gamma)\mathsf{r}}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}\leq|\partial_{\mathfrak{q}}\hbar(\varsigma)|\leq1+\frac{F|d|(1+\mathfrak{q})(1+\gamma)\mathsf{r}}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}.\tag{2.6}$$

These inequalities are sharp by the function

$$\hbar(\varsigma)\quad=\quad\varsigma+\frac{F|d|(1+\gamma)}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}\varsigma^2.$$

Proof. Given $\hbar\in\mathcal{D}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ from (2.1) and since

$$\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}\Big(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|)\Big)+F|d|\Big)$$

is increasing and positive for $\varkappa\geq2$, then we have

$$\begin{aligned}&\frac{1+\xi}{1+\gamma}\Big(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|\Big)\sum_{\varkappa=2}^{\infty}a_\varkappa\leq\\&\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}\Big(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|)\Big)+F|d|\Big)\sum_{\varkappa=2}^{\infty}a_\varkappa\\&\leq F|d|,\end{aligned}$$

which is equivalent to,

$$\sum_{\varkappa=2}^{\infty}a_\varkappa\leq\frac{F|d|(1+\gamma)}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}.\tag{2.7}$$

We can acquire this through the utilization of the properties of the modulus function

$$\begin{aligned}|\hbar(\varsigma)|\quad&=\quad\Big|\varsigma+\sum_{\varkappa=2}^{\infty}a_\varkappa\varsigma^\varkappa\Big|\\&\leq\quad|\varsigma|+\sum_{\varkappa=2}^{\infty}|a_\varkappa||\varsigma|^\varkappa\\&\leq\quad\mathsf{r}+\mathsf{r}^2\sum_{\varkappa=2}^{\infty}|a_\varkappa|\\&\leq\quad\mathsf{r}+\frac{F|d|(1+\gamma)\mathsf{r}^2}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)},\quad\text{by (2.7)}.\end{aligned}$$

and

$$\begin{aligned}|\hbar(\varsigma)|\quad&=\quad\Big|\varsigma+\sum_{\varkappa=2}^{\infty}a_\varkappa\varsigma^\varkappa\Big|\geq|\varsigma|-\sum_{\varkappa=2}^{\infty}|a_\varkappa||\varsigma|^\varkappa\\&\geq\quad\mathsf{r}-\mathsf{r}^2\sum_{\varkappa=2}^{\infty}|a_\varkappa|\geq\mathsf{r}-\frac{F|d|(1+\gamma)\mathsf{r}^2}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)},\quad\text{by (2.7)}.\end{aligned}$$

Now, by applying the Jackson's derivative of (1.5) with respect to ς, we get:

$$\begin{aligned}|\partial_{\mathfrak{q}}\hbar(\varsigma)| &= \left|1+\sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}a_{\varkappa}\varsigma^{\varkappa-1}\right| \leq 1+\sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}|a_{\varkappa}||\varsigma|^{\varkappa}\\ &\leq \mathsf{r}+[2]_{\mathfrak{q}}\mathsf{r}^2\sum_{s=2}^{\infty}|a_{\varkappa}| \leq \mathsf{r}+\frac{F|d|(1+\mathfrak{q})(1+\gamma)}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}\mathsf{r}^2.\end{aligned}$$

In other hand,

$$\begin{aligned}|\partial_{\mathfrak{q}}\hbar(\varsigma)| &= \left|1+\sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}a_{\varkappa}\varsigma^{\varkappa-1}\right| \geq 1-\sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}|a_{\varkappa}||\varsigma|^{\varkappa}\\ &\geq \mathsf{r}-[2]_{\mathfrak{q}}\mathsf{r}^2\sum_{\varkappa=2}^{\infty}|a_{\varkappa}| \geq \mathsf{r}-\frac{F|d|(1+\mathfrak{q})(1+\gamma)}{(1+\xi)(\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|+F|d|)}\mathsf{r}^2.\end{aligned}$$

□

The neighbourhoods problems of the class $\mathcal{D}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ will now be determined.

Definition 2.1. *Let $\hbar\in\mathcal{A}$ and $\sigma>0$. We define the $(m,\sigma,\mathfrak{q})$-neighbourhood of $\hbar$ as follows:*

$$\mathcal{N}_{\sigma,\mathfrak{q}}(\hbar)=\left\{g\in\mathcal{A}:g(\varsigma)=\varsigma+\sum_{\varkappa=2}^{\infty}b_{\varkappa}\varsigma^{\varkappa}\ \text{and}\ \sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}|a_{\varkappa}-b_{\varkappa}|\leq\sigma\right\}. \tag{2.8}$$

In particular, for the identity functions $e(\varsigma)=z$, we have

$$\mathcal{N}_{\sigma,\mathfrak{q}}(e)=\left\{g\in\mathcal{A}:g(\varsigma)=\varsigma+\sum_{\varkappa=2}^{\infty}b_{\varkappa}\varsigma^{\varkappa}\ \text{and}\ \sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}|b_{\varkappa}|\leq\sigma\right\}. \tag{2.9}$$

Definition 2.2. *A function $\hbar\in\mathcal{A}$ belong to the class $\mathcal{D}^{\nu}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ if there exists a function $\pounds\in\mathcal{D}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ such that*

$$\left|\frac{\hbar(\varsigma)}{\pounds(\varsigma)}-1\right|\leq 1-\nu,\quad 0\leq\psi<1,\quad(\varsigma\in\diamond). \tag{2.10}$$

Theorem 2.4. *$f\,\pounds\in\mathcal{D}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q})$ and*

$$\nu = 1-\frac{[1+\xi]_{\mathfrak{q}}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)}{[1+\xi]_{\mathfrak{q}}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)-F|d|[1+\gamma]_{\mathfrak{q}}},$$

then

$$\mathcal{N}_{\sigma,q}(\pounds)\subseteq\mathcal{D}^{\nu}\bigoplus(\vartheta,F,d,\xi,\gamma;\mathfrak{q}).$$

Proof. Let $\hbar\in\mathcal{N}_{\sigma,q}(\pounds)$, we find from (2.8) that

$$\sum_{\varkappa=2}^{\infty}[\varkappa]_{\mathfrak{q}}|a_{\varkappa}-b_{\varkappa}| \leq \sigma,$$

which implies the coefficient inequality

$$\sum_{\varkappa=2}^{\infty}|a_{\varkappa}-b_{\varkappa}| \leq \frac{\sigma}{1+\mathfrak{q}}.$$

Since $\pounds \in \mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$, and Using relation (2.1) of Theorem **??**, we have

$$\begin{aligned}
&\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(2+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(2+\gamma)}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)\sum_{\varkappa=2}^{\infty}|b_{\varkappa}| \\
\leq &\sum_{\varkappa=2}^{\infty}\frac{\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)}{\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)|b_{\varkappa}| \leq F|d|,
\end{aligned}$$

for $(\varkappa \geq 2)$, which implies

$$\sum_{\varkappa=2}^{\infty}|b_{\varkappa}| \leq \frac{F|d|[1+\gamma]_{\mathfrak{q}}}{[1+\xi]_{\mathfrak{q}}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)}, \tag{2.11}$$

and so

$$\begin{aligned}
\left|\frac{\hbar(\varsigma)}{\pounds(\varsigma)}-1\right| &< \frac{\sum_{\varkappa=2}^{\infty}|a_{\varkappa}-b_{\varkappa}|}{1-\sum_{\varkappa=2}^{\infty}b_{\varkappa}} \\
&\leq \frac{\sigma}{1+\mathfrak{q}}\cdot\left\{\frac{[1+\xi]_{\mathfrak{q}}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)}{[1+\xi]_{\mathfrak{q}}\Big((\vartheta+\mathfrak{q}+\mathfrak{q}\vartheta F|d|)+F|d|\Big)-F|d|[1+\gamma]_{\mathfrak{q}}}\right\} \\
&= 1-\nu.
\end{aligned}$$

Thus, for given ν and by Definition 2.1, we have $\hbar \in \mathcal{D}^{\nu}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$.

Finally, we establish the radii of starlikeness of order λ for functions in the class $\mathcal{D}^{\lambda}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. □

Theorem 2.5. *Let $\hbar \in \mathcal{A}$ from the class $\mathcal{D}^{\lambda}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$. The function $\hbar$ univalent starlike of order λ, $0 \leq \lambda < 1$ and $|\varsigma| < \mathfrak{r}_0$, where*

$$\mathfrak{r}_0 = \inf_{k}\left\{\frac{(1-\lambda)\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)}{F|d|([2]_{\mathfrak{q}}-\lambda)\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)}\right\}^{\frac{1}{n-1}}. \tag{2.12}$$

Proof. We show that

$$\left|\frac{\varsigma\partial_{\mathfrak{q}}(\hbar(\varsigma))}{\hbar(\varsigma)}-1\right| \leq 1-\lambda, \quad (|\varsigma| < r_0).$$

Considering that

$$\left|\frac{\varsigma\partial_{\mathfrak{q}}(\hbar(\varsigma))}{\hbar(\varsigma)}-1\right| = \left|\frac{\sum_{\varkappa=2}^{\infty}([\varkappa]_{\mathfrak{q}}-1)a_{\varkappa}\varsigma^{\varkappa-1}}{1+\sum_{\varkappa=2}^{\infty}a_{\varkappa}\varsigma^{\varkappa-1}}\right| \leq \frac{\sum_{\varkappa=2}^{\infty}([\varkappa]_{\mathfrak{q}}-1)a_{\varkappa}|\varsigma|^{\varkappa-1}}{1-\sum_{\varkappa=2}^{\infty}a_{\varkappa}|\varsigma|^{\varkappa-1}},$$

to prove the theorem, we must show that

$$\frac{\sum_{\varkappa=2}^{\infty}([\varkappa]_{\mathfrak{q}}-1)a_{\varkappa}|\varsigma|^{\varkappa-1}}{1-\sum_{\varkappa=2}^{\infty}a_{\varkappa}|\varsigma|^{\varkappa-1}} \leq 1-\lambda,$$

which equivalent to

$$\sum_{\varkappa=2}^{\infty}([\varkappa]_{\mathfrak{q}}-\lambda)a_{\varkappa}|\varsigma|^{\varkappa-1} \leq 1-\lambda,$$

and applying Theorem **??**, we have

$$|\varsigma| \leq \left\{ \frac{(1-\lambda)\Gamma_{\mathfrak{q}}(\gamma+1)\Gamma_{\mathfrak{q}}(\varkappa+\xi)\Big([\varkappa-1]_{\mathfrak{q}}(\vartheta([\varkappa]_{\mathfrak{q}}-\mathfrak{q})+\mathfrak{q}(1+\vartheta F|d|))+F|d|\Big)}{F|d|([2]_{\mathfrak{q}}-\lambda)\Gamma_{\mathfrak{q}}(\xi+1)\Gamma_{\mathfrak{q}}(\varkappa+\gamma)} \right\}^{\frac{1}{\varkappa-1}}.$$

Hence, the proof is complete. □

3. Conclusion

In this article, we introduce a new class of normalized analytic functions called $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$, which is associated with the modified $\mathfrak{q}$-Tremblay operator on the open unit disk $\diamond$. We investigate the necessary conditions for functions belonging to the subclass $\mathcal{D}\bigoplus(\vartheta, F, d, \xi, \gamma; \mathfrak{q})$, as well as the growth and distortion bounds, the convolution condition, the radii of starlikeness, and the neighborhood problems involving the $\mathfrak{q}$-analogue of a modified Tremblay operator for functions in this class.

Our results extend and generalize some of the known results in the literature on analytic functions. We believe that our findings will have useful applications in various areas of mathematics, such as complex analysis, geometric function theory, and applied mathematics.

In summary, this article contributes to the ongoing research in the field of complex analysis by providing a more profound understanding of the theory and applications of analytic functions. The results obtained in this article have the potential for future generalization through the utilization of post-quantum calculus and other q-analogues of the fractional derivative operator. Additionally, further research may be conducted to explore additional subclasses and their respective properties.

References

[1] P.R. Agarwal, Certain fractional q-integrals and q-derivatives, Math. Proc. Camb. Philos. Soc. **66**(2)(1969) 2365–370.

[2] I. Aldawish and M. Darus, tarlikeness of q-differential operator involving quantum calculus, The Kangwon-Kyungki Math. Soc. **22**(4) (2020) 699–709.

[3] A. Alb Lupaş, and G. Oros, Fractional integral of a confluent hypergeometric function applied to defining a new class of analytic functions. Symmetry **14** (2022) 427.

[4] W. Al-Salam, A. Verma, A fractional Leibniz q-formula, Pacific Journal Of Mathematics **60** (1975) 1–9.

[5] W. Al-Salam, Some fractional q-integrals and q-derivatives, Proceedings Of The Edinburgh Mathematical Society **15** (1966) 135-140.

[6] W. Atshan, A. Battor, and N. Abbas, On a new subclass of univalent functions with positive coefficients defined by Ruscheweyh Derivative, Eur. Jour. Sci. Res. **140** (2016) 85–92.

[7] S. Elhaddad, M. Darus, On meromorphic functions defined by a new operator containing the Mittag–Leffler function. Symmetry, **11**(2) (2019), 210.

[8] Z. Esa, A. Kilicman, R. Ibrahim, M. Ismail and S. Husain, Application of modified complex tremblay operator. AIP Conference Proceedings **1739** (2016), 020059.

[9] B. Frasin and M. Darus, Subclass of analytic functions defined by q-derivative operator associated with Pascal distribution series. AIMS Math, **6** (2021) 5008-5019

[10] Exton, H. q-Hypergeometric functions and applications, E. Horwood Halsted Press, (1983).

[11] G. Gasper, and Rahman, M. Basic hypergeometric series, Cambridge university press (2011).

[12] A. Kilbas, H. Srivastava and J. Trujillo, Theory and applications of fractional differential equations. Cambridge university press (2011).

[13] M. Garg and L. Chanchlani, q-analogues of Saigo's fractional calculus operators, Bull. Math. Anal. Appl. **3** (2011) 169–179.

[14] S. Isogawa, N. Kobachi and S. Hamada, A q-analogue of Riemann-Liouville fractional derivative, Res. Rep. Yatsushiro Nat. Coll. Tech. **29** (2007) 59–68.

[15] I. Podlubny, An introduction to fractional derivatives, fractional differential equations to methods of their solution and some of their applications, Math. Sci. Eng. **198** (1999), 0924-34008

[16] J. Prajapat, R. Raina and H. Srivastava, Some inclusion properties for certain subclasses of strongly starlike and strongly convex functions involving a family of fractional integral operators, Int. Tran. Spec. Fun. **18** (2007) 639–651.
[17] S. Purohit and R. Raina, Certain subclasses of analytic functions associated with fractional q-calculus operators, Math. Scand., **109**, (2011) 55–70
[18] S. Purohit, and R. Yadav, On generalized fractional q-integral operators involving the q-Gauss hypergeometric function, Bull. Math. Anal. Appl. **2** (2010) 35–44.
[19] S. Ramadan and M. Darus, Univalence of an integral operator defined by generalized operators, Int. J. Math. Comput. Sci. **4** (2010) 1117–1119.
[20] S. Samko, *Fractional integrals and derivatives: Theory And Applications*, Gordon and Breach, (1993).
[21] R. Saxena, R. Yadav, S. Purohit and S. Kalla, Kober fractional q-integral operator of the basic analogue of the H-function, Rev. Téc. Fac. Ing. Univ. Zulia. **28** (2005) 154-158.
[22] T. Seoudy and M. Aouf, Coefficient estimates of new classes of q-starlike and q-convex functions of complex order, J. Math. Inequal **10** (2016) 135-145.
[23] R. Tremblay, Une Contribution a la Theorie de la Derivee Fractionnaire. Ph. D. Thesis, (1974).
[24] R. Yadav, S. Purohit, and S. Kalla, On generalized Weyl fractional q-integral operator involving generalized basic hypergeometric functions,Fract. Calc. Appl. Anal. **11** (2008) 129–142.
[25] R. Yadav, and S. Purohit, On applications of Weyl fractional q-integral operator to generalized basic hypergeometric functions, Kyungpook Math. J. **46** (2006) 235-245.

Abdullah Alsoboh,
Philadelphia University, Amman, Jordan, phone: 00962770959697
Email address: aalsoboh@philadelphia.edu.jo

Waggas Galib Atshan,
University of Al-Qadisiyah, Diwaniyah, Iraq, phone: 009647830724519
Email address: Waggas.galib@qu.edu.iq

Proceedings of International Mathematical Sciences
ISSN: 2717-6355, URL: https://dergipark.org.tr/tr/pub/pims
Volume 6 Issue 2 (2024), Pages 54-67.
Doi: https://doi.org/10.47086/pims.1503458

THE NEUTROSOPHIZE OF NEW CONTINUITY SPECIES

AHU ACIKGOZ*, AND FERHAT ESENBEL**
*BALIKESIR UNIVERSITY, 10145 BALIKESIR, TURKEY. ORCID NUMBER: 0000-0003-1468-8240
**BALIKESIR UNIVERSITY, 10145 BALIKESIR, TURKEY. ORCID NUMBER: 0000-0001-5581-3621.

Abstract. In this study, after explaining the process that makes the study necessary in the introduction and giving the necessary definitons in the preliminaries, in the third section, some types of open set that were previously defined in general topology and various non-standard topological spaces are presented and the relationships between them are explained with the help of a diagram. Then, the concept of neutrosophic af-open set is defined and its relations with other open set types are examined and their properties are investigated in neutrosophic topology. In the following sections, the concept of af-open set is generalized and different types of continuities are introduced using these new concepts of open set, and the connections between them are illustrated with examples and diagrams.

1. Introduction

The concept of open set has always been one of the indispensable characters of the world of topology. This concept has been divided into many different types in itself as a result of the constant change of social life and the constant change of people's needs. For example, M. H. Alqahtani [4, 5] presented the concepts of F-open set and C-open set in 2023. Later, new ones continued to be added to these open set varieties. These open set types provided scientists with the opportunity to re-approach many concepts from different perspectives that had been previously introduced in topology and to examine their properties in general topology and some other non-standard topologies as in [6, 7, 8, 9, 17, 18]. Furthermore, these open set variants led to the introduction of many types of functions, mappings and continuities as in [13, 21, 22, 25].
Smarandache's introduction of the concept of neutrosophic set in [27] created facilities to make contributions to some other disiplicines as in [10, 11, 12, 14, 16, 19, 20, 24, 28] and allowed the introduction of different non-standard topological spaces. Like uncultivated fields, these new non-standard topological spaces enabled scientists to give products to the world of science unlike anything done until then as in [1, 2, 3, 23]. In this study, we aimed to join these scientists by introducing the concept of neutrosophic af-open set. Also, topological properties of af-interior, af-closure operators are presented by using the

2020 *Mathematics Subject Classification.* Primary: 54A05, 54C10, 54D30, 54D101.
Key words and phrases. Neutrosophic topological space; neutrosophic af-interior operator; neutrosophic af-neighbourhood; neutrosophic af-continous function.

Submitted on 22.06.2024, Accepted on 10.09.2024.
Communicated by Nazlım Deniz Aral.

concept of neutrosophic af-open sets. Moreover, the notions of neutrosophic af-continuous functions and some other continuity types are introduced and the connections between them are illustrated with diagrams.

2. Preliminaries

In this section, we present the basic definitions related to neutrosophic set theory.

Definition 2.1. [27] *A neutrosophic set A on the universe set X is defined as:*

$$A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle : x \in X\},$$

where T, I, F $: X \to \,]^{-}0, 1^{+}[$ *and* $^{-}0 \leq T_A(x) + I_A(x) + F_A(x) \leq 3^{+}$.

Scientifically, membership functions, indeterminacy functions and non-membership functions of a neutrosophic set take value from real standart or nonstandart subsets of $]^{-}0, 1^{+}[$. However, these subsets are sometimes inconvenient to be used in real life applications such as economical and engineering problems. On account of this fact, we consider the neutrosophic sets, whose membership function, indeterminacy functions and non-membership functions take values from subsets of $[0, 1]$.

Definition 2.2. [15] *Let X be a nonempty set. If r, t, s are real standard or non standard subsets of* $]^{-}0, 1^{+}[$ *then the neutrosophic set* $x_{r,t,s}$ *is called a neutrosophic point in X given by*

$$x_{r,t,s}(x_p) = \begin{cases} (r,t,s), & \text{if } x = x_p \\ (0,0,1), & \text{if } x \neq x_p \end{cases}$$

For $x_p \in X$*, it is called the support of* $x_{r,t,s}$*, where r denotes the degree of membership value, t denotes the degree of indeterminacy and s is the degree of non-membership value of* $x_{r,t,s}$.

Definition 2.3. [26] *Let A be a neutrosophic set over the universe set X. The complement of A is denoted by* A^c *and is defined by:*

$A^c = \left\{\left\langle x, F_{\tilde{F}(e)}(x), 1 - I_{\tilde{F}(e)}(x), T_{\tilde{F}(e)}(x)\right\rangle : x \in X\right\}$.

It is obvious that $[A^c]^c = A$.

Definition 2.4. [26] *Let A and B be two neutrosophic sets over the universe set X. A is said to be a neutrosophic subset of B if* $T_A(x) \leq T_B(x), I_A(x) \leq I_B(x), F_A(x) \geq F_B(x)$*, every xinX. It is denoted by* $A \subseteq B$*. A is said to be neutrosophic equal to B if* $A \subseteq B$ *and* $B \subseteq A$*. It is denoted by* $A = B$.

Definition 2.5. [26] *Let* F_1 *and* F_2 *be two neutrosophic sets over the universe set X. Then their union is denoted by* $F_1 \cup F_2 = F_3$ *is defined by:*

$$F_3 = \{\langle x, T_{F_3}(x), I_{F_3}(x), F_{F_3}(x) : x \in X\rangle\},$$

where

$$T_{F_3(x)} = \max\{T_{F_1(x)}, T_{F_2}(x)\},$$

$$I_{F_3(x)} = \max\{I_{F_1(x)}, I_{F_2}(x)\},$$

$$F_{F_3(x)} = \min\{F_{F_1(x)}, F_{F_2}(x)\}.$$

Definition 2.6. [26] *Let* F_1 *and* F_2 *be two neutrosophic sets over the universe set X. Then their intersection is denoted by* $F_1 \cap F_2 = F_4$ *is defined by:*

$$F_4 = \{\langle x, T_{F_4}(x), I_{F_4}(x), F_{F_4}(x) : x \in X\rangle\},$$

where

$$T_{F_4(x)} = \min\{T_{F_1(x)}, T_{F_2}(x)\},$$

$$I_{F_4(x)} = \min\{I_{F_1(x)}, I_{F_2}(x)\},$$

$$F_{F_4(x)} = \max\{F_{F_1(x)}, F_{F_2}(x)\}.$$

Definition 2.7. [26] *A neutrosophic set F over the universe set X is said to be a null neutrosophic set if $T_F(x) = 0$, $I_F(x) = 0$, $F_F(x) = 1$, every $x \in X$. It is denoted by 0_X.*

Definition 2.8. [26] *A neutrosophic set F over the universe set X is said to be an absolute neutrosophic set if $T_F(x) = 1$, $I_F(x) = 1$, $F_F(x) = 0$, every $x \in X$. It is denoted by 1_X.*

Clearly $0_X^c = 1_X$ and $1_X^c = 0_X$.

Definition 2.9. [26] *Let $NS(X)$ be the family of all neutrosophic sets over the universe the set X and $\tau \subset NS(X)$. Then τ is said to be a neutrosophic topology on X if:*

1) 0_X and 1_X belong to τ;

2) The union of any number of neutrosophic sets in τ belongs to τ;

3) The intersection of a finite number of neutrosophic sets in τ belongs to τ.

Then (X, τ) is said to be a neutrosophic topological space over X. Each member of τ is said to be a neutrosophic open set [26].

Definition 2.10. [21] *Let (X, τ) be a neutrosophic topological space over X and F be a neutrosophic set over X. Then F is said to be a neutrosophic closed set iff its complement is a neutrosophic open set.*

Definition 2.11. [1] *A neutrosophic point $x_{r,t,s}$ is said to be neutrosophic quasi-coincident (neutrosophic q-coincident, for short) with F, denoted by $x_{r,t,s}\ q\ F$ if and only if $x_{r,t,s} \not\subseteq F^c$. If $x_{r,t,s}$ is not neutrosophic quasi-coincident with F, we denote by $x_{r,t,s}\ \tilde{q}\ F$.*

Definition 2.12. [1] *A neutrosophic set F in a neutrosophic topological space (X, τ) is said to be a neutrosophic q-neighborhood of a neutrosophic point $x_{r,t,s}$ if and only if there exists a neutrosophic open set G such that*
$x_{r,t,s}\ q\ G \subset F$.

Definition 2.13. [1]*) A neutrosophic set G is said to be neutrosophic quasi-coincident (neutrosophic q-coincident, for short) with F, denoted by $G\ q\ F$ if and only if $G \not\subseteq F^c$. If G is not neutrosophic quasi-coincident with F, we denote by $G\ \tilde{q}\ F$.*

Definition 2.14. [3] *A neutrosophic point $x_{r,t,s}$ is said to be a neutrosophic interior point of a neutrosophic set F if and only if there exists a neutrosophic open q-neighborhood G of $x_{r,t,s}$ such that $G \subset F$. The union of all neutrosophic interior points of F is called the neutrosophic interior of F and denoted by F°.*

Definition 2.15. [1] *A neutrosophic point $x_{r,t,s}$ is said to be a neurosophic cluster point of a neutrosophic set F if and only if every neutrosophic open q-neighborhood G of $x_{r,t,s}$ is q-coincident with F. The union of all neutrosophic cluster points of F is called the neutrosophic closure of F and denoted by $\overline{F}$.*

Definition 2.16. [1] *Let f be a function from X to Y. Let B be a neutrosophic set in Y with members hip function $T_B(y)$, indeterminacy function $I_B(y)$ and non-membership function $F_B(y)$. Then, the inverse image of B under f, written as $f^{-1}(B)$, is a neutrosophic subset of X whose membership function, indeterminacy function and non-membership function are*

defined as $T_{f^{-1}(B)}(x) = T_B(f(x))$, $I_{f^{-1}(B)}(x) = I_B(f(x))$ *and* $F_{f^{-1}(B)}(x) = F_B(f(x))$ *for all* x *in* X, *respectively.*
Conversely, let A *be a neutrosophic set in* X *with membership function* $T_A(x)$, *indeterminacy function* $I_A(x)$ *and non-membership function* $F_A(x)$. *The image of* A *under* f, *written as* $f(A)$, *is a neutrosophic subset of* Y *whose membership function, indeterminacy function and non-membership function are defined as*

$$T_{f(A)}(y) = \begin{cases} sup_{z\in f^{-1}(y)}\{T_A(z)\}, & if\ f^{-1}(y)\ is\ not\ empty, \\ 0, & if\ f^{-1}(y)\ is\ empty, \end{cases}$$

$$I_{f(A)}(y) = \begin{cases} sup_{z\in f^{-1}(y)}\{I_A(z)\}, & if\ f^{-1}(y)\ is\ not\ empty, \\ 0, & if\ f^{-1}(y)\ is\ empty, \end{cases}$$

$$F_{f(A)}(y) = \begin{cases} sup_{z\in f^{-1}(y)}\{F_A(z)\}, & if\ f^{-1}(y)\ is\ not\ empty, \\ 0, & if\ f^{-1}(y)\ is\ empty, \end{cases}$$

for all y *in* Y, *where* $f^{-1}(y) = \{x : f(x) = y\}$, *respectively.*

3. Neutrosophic af-open sets

This section provides some new definitions that form the cornerstones of the sections that follow.

Definition 3.1. *A neutrosophic set* F *in a neutrosophic topological space* (X, τ) *is said to be*

a) Neutrosophic semiopen, if $F \subseteq \overline{F^\circ}$,
b) Neutrosophic preopen, $F \subseteq (\overline{F})^\circ$,
c) Neutrosophic β*-open,* $F \subseteq \overline{(\overline{F})^\circ}$,
d) Neutrosophic α*-open, if* $F \subseteq (\overline{(F^\circ)})^\circ$.

By Definition 17, the following diagram is obtained:

$$\begin{array}{ccc} neutrosophic\ open \to neutrosophic\ \alpha - open & \to & neutrosophic\ pre - open \\ & \downarrow & \downarrow \\ & neutrosophic\ semi - open \to & neutrosophic\ \beta - open \end{array}$$

Diagram I

Definition 3.2. *If,* F *be a neutrosophic set in neutrosophic topological space* (X, τ) *then,* $\overline{F}_s = \bigcap\{F : F \subseteq A, Aisneutrosophicsemiclosed\}$ *(resp.* $F^\circ_s = \bigcup\{F : F \subseteq A, Aisneutrosophicsemiopen\}$*) is called a neutrosophic semiclosure of* F *(resp. neutrosophic semi-interior of* F*).*

Definition 3.3. *If,* F *be a neutrosophic set in neutrosophic topological space* (X, τ) *then,* $\overline{F}_p = \bigcap\{F : F \subseteq A, Aisneutrosophicpreclosed\}$ *(resp.* $F^\circ_p = \bigcup\{F : F \subseteq A, Aisneutrosophicpreopen\}$*) is called a neutrosophic preclosure of* F *(resp. neutrosophic pre interior of* F*).*

Definition 3.4. *If,* F *be a neutrosophic set in neutrosophic topological space* (X, τ) *then,* $\overline{F}_\beta = \bigcap\{F : F \subseteq A, Aisneutrosophic\beta closed\}$ *(resp.* $F^\circ_s = \bigcup\{F : F \subseteq A, Aisneutrosophic\beta open\}$*) is called a neutrosophic* β *closure of* F *(resp. neutrosophic* β *interior of* F*).*

Definition 3.5. *If,* F *be a neutrosophic set in neutrosophic topological space* (X, τ) *then,* $\overline{F}_\alpha = \bigcap\{F : F \subseteq A, Aisneutrosophic\alpha closed\}$ *(resp.* $F^\circ_\alpha = \bigcup\{F : F \subseteq A, Aisneutrosophic\alpha open\}$*) is called a neutrosophic* α *closure of* F *(resp. neutrosophic* α *interior of* F*).*

Definition 3.6. *Let (X,τ) be a neutrosophic topological space. A neutrosophic set λ is af-open set if $\lambda \subseteq (\lambda \cup \mu)^\circ$ for every μ is neutrosophic open set such that $0_X \neq \mu \neq 1_X$. The complement of the neutrosophic af -open set is called neutrosophic af-closed. We denote the family of all neutrosophic af-open (resp. neutrosophic af-closed) sets of a neutrosophic topological spece (X,τ) by $NafO(X,\tau)$(resp. $NafC(X,\tau)$).*

Problem 3.1. *Let (X,τ) be a neutrosophic topological space. In Definition 22, for every $\mu \in \tau$ such that $0_X \neq \mu \neq 1_X$, can we obtain a new type of neutrosophic af-open sets by taking the neutrosophic closure of μ instead of μ?*

Theorem 3.2. *Every neutrosophic open set in a neutrosophic topological space (X,τ) is neutrosophic af-open set.*

Proof. Let (X,τ) be any neutrosophic topological space and let $\lambda \subseteq X$ be any neutrosophic open set. Therefore, $\lambda = \lambda^\circ \subseteq (\lambda \cup \mu)^\circ$ is neutrosophic open set such that $0_X \neq \mu \neq 1_X$. Thus, $\lambda$is neutrosophic af-open set. Then for the collection of $NafO(X,\tau)$, $\tau \subseteq NafO(X,\tau)$. □

Remark. *The converse of Theorem 3.2. is not always true as shown by the following example.*

Example 3.1. *Let (X,τ) be a neutrosophic topological space, with $X = \{a,b,c\}$, $\tau = \{0_X, \lambda, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.5, 0.5, 0.5\rangle, \langle b, 0.7, 0.7, 0.3\rangle, \langle c, 0.9, 0.9, 0.1\rangle\}$ and $\mu = \{\langle a, 0.4, 0.4, 0.6\rangle, \langle b, 0.3, 0.3, 0.7\rangle, \langle c, 0.9, 0.9, 0.1\rangle\}$. Then, $\mu \in NafO(X,\tau)$, and but the set μ is not neutrosophic open.*

Theorem 3.3. *Let (X,τ) be any neutrosophic topological space and λ, μ be two neutrosophic af-open sets. Then, the following properties are hold:*

(1) $\lambda \cap \mu$ is neutrosophic af-open set.
(2) $\lambda \cup \mu$ is neutrosophic af-open set.

Proof. (1) Let λ and μ be two neutrosophic af-open sets. Then from Definition 3.6, $\lambda \subseteq (\lambda \cup \beta)^\circ$ and $\mu \subseteq (\mu \cup \beta)^\circ$ for every β is neutrosophic open set and $0_X \neq \beta \neq 1_X$. Then, $\lambda \cap \mu \subseteq (\lambda \cup \beta)^\circ \cap (\mu \cup \beta)^\circ = ((\lambda \cup \beta) \cap (\mu \cup \beta))^\circ \subseteq ((\lambda \cap \mu) \cup \beta)^\circ$.
(2) Let λ and μ be two neutrosophic af-open sets. Then from Definition 3.6, $\lambda \subseteq (\lambda \cup \beta)^\circ$ and $\mu \subseteq (\mu \cup \beta)^\circ$ for every β is neutrosophic open set and $0_X \neq \beta \neq 1_X$. Then, $\lambda \cup \mu \subseteq (\lambda \cup \beta)^\circ \cup (\mu \cup \beta)^\circ = ((\lambda \cup \beta) \cup (\mu \cup \beta))^\circ \subseteq ((\lambda \cup \mu) \cup \beta)^\circ$. □

Proposition 3.4. *Let (X,τ) be any neutrosophic topological space. If, for every $\alpha \in \Delta$, $\lambda_\alpha \in NafO(X,\tau)$, then $\bigcup_{\alpha\in\Delta} \lambda_\alpha \in NafO(X,\tau)$.*

Proof. Let $\lambda_\alpha \in NafO(X,\tau)$ for every $\alpha \in \Delta$. Then, $\lambda_\alpha \subseteq \bigcup_{\alpha\in\Delta} \lambda_\alpha$, for every $\alpha \in \Delta$. For any β is neutrosophic open ($0_X \neq \beta \neq 1_X$) and each $\alpha \in \Delta$, we get $\lambda_\alpha \subseteq (\lambda_\alpha \cup \beta)^\circ \subseteq ((\bigcup_{\alpha\in\Delta} \lambda_\alpha) \cup \beta)^\circ$. Hence we have $\bigcup_{\alpha\in\Delta} \lambda_\alpha \subseteq ((\bigcup_{\alpha\in\Delta} \lambda_\alpha) \cup \beta)^\circ$. Therefore $\bigcup_{\alpha\in\Delta} \lambda_\alpha \in NafO(X,\tau)$. □

Theorem 3.5. *Let (X,τ) be any neutrosophic topological space and $\tau_{NafO} = \{\lambda : \lambda \in NafO(X,\tau)\}$. Then is τ_{NafO} a neutrosophic topology such that $\tau \subseteq \tau_{NafO}$.*

Proof. According to Theorem 3.2, we have $\tau \subseteq \tau_{NafO}$. We show that τ_{NafO} is a neutrosophic topology.

(1) It is clear that $0_X, 1_X \in \tau_{NafO}$.
(2) and (3) are seen that from Theorem 3.3 and Proposition 3.4.

□

4. Generalizations of Neutrosophic af-open sets

Definition 4.1. *A subset λ of a neutrosophic topological space (X, τ) is said to be*

(i) neutrosophic $af\alpha-$ open if $\lambda \subseteq (\lambda \cup \mu)^{\circ}_{\alpha}$, for every μ is neutrosophic open and $0_X \neq \mu \neq 1_X$,
(ii) neutrosophic $afp-$ open if $\lambda \subseteq (\lambda \cup \mu)^{\circ}_{p}$, for every μ is neutrosophic open and $0_X \neq \mu \neq 1_X$,
(iii) neutrosophic $afs-$ open if $\lambda \subseteq (\lambda \cup \mu)^{\circ}_{s}$, for every μ is neutrosophic open and $0_X \neq \mu \neq 1_X$,
(iv) neutrosophic $af\beta-$ open if $\lambda \subseteq (\lambda \cup \mu)^{\circ}_{\beta}$, for every μ is neutrosophic open and $0_X \neq \mu \neq 1_X$.

The complement of a neutrosophic afα-open (resp. neutrosophic afp-open, neutrosophic afs-open, neutrosophic afβ-open) set is said to be neutrosophic afα-closed (resp. neutrosophic afp-closed, neutrosophic afs-closed, neutrosophic afβ-closed). The family of all f neutrosophic afα-open (neutrosophic afα-closed) (resp. neutrosophic afp-open (neutrosophic afp-closed), neutrosophic afs-open (neutrosophic afs-closed), neutrosophic afβ-open (neutrosophic afβ-closed)) sets in a neutrosophic topological space (X, τ) is denoted by $NafαO(X, \tau)(NafαC(X, \tau))(resp.NafPO(X, \tau)$ $(NafPC(X, \tau)), NafSO(X, \tau)(NafSC(X, \tau)), NafβO(X, \tau)(NafβC(X, \tau)))$.

From Definition 4.1, we have the following diagram:

$$\begin{array}{ccccc} & & \text{neutrosophic open} & & \\ & & \downarrow & & \\ \text{neutrosophic } af- \text{open} & \rightarrow & \text{neutrosophic } af\alpha- \text{open} & \rightarrow & \text{neutrosophic } afp- \text{open} \\ & & \downarrow & & \downarrow \\ & & \text{neutrosophic } afs- \text{open} & \rightarrow & \text{neutrosophic } af\beta- \text{open} \end{array}$$

Diagram II

Problem 4.1. *In the above definition, for every $\mu \in \tau$ such that $0_X \neq \mu \neq 1_X$, can a new types of neutrosophic af-open set be given by taking the neutrosophic closure of μ instead of μ?*

Remark. *The inverses of the requirements in the diagram above may not always be true.*

Example 4.1. *It can be seen from Example 3.1 that not every neutrosophic af-open set is a neutrosophic open set.*

Example 4.2. *Let (X, τ) be a neutrosophic topological space, with $X = \{a, b, c\}$, $\tau = \{0_X, \lambda, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.2, 0.2, 0.8\rangle,$ $\langle b, 0.7, 0.7, 0.3\rangle, \langle c, 0.4, 0.4, 0.6\rangle\}$ and $\mu = \{\langle a, 0.7, 0.7, 0.3\rangle, \langle b, 0.9, 0.9, 0.1\rangle, \langle c, 0.1, 0.1, 0.9\rangle\}$. Then, $\mu \in NafαO(X, \tau)$, and but the set μ is not neutrosophic af-open.*

Example 4.3. *Let (X, τ) be a neutrosophic topological space, with $X = \{a, b, c\}$, $\tau = \{0_X, \mu, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.2, 0.2, 0.8\rangle,$ $\langle b, 0.3, 0.3, 0.7\rangle, \langle c, 0.7, 0.7, 0.3\rangle\}$ and $\mu = \{\langle a, 0.1, 0.1, 0.9\rangle, \langle b, 0.2, 0.2, 0.8\rangle, \langle c, 0.2, 0.2, 0.8\rangle\}$. Then, $\lambda \in NafSO(X, \tau)$, and but the set λ is neither neutrosophic afα-open nor neutrosophic afp-open.*

Example 4.4. *Let (X, τ) be a neutrosophic topological space, with $X = \{a, b, c\}$, $\tau = \{0_X, \mu, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.3, 0.3, 0.7\rangle,$ $\langle b, 0.8, 0.8, 0.2\rangle, \langle c, 0.7, 0.7, 0.3\rangle\}$ and $\mu = \{\langle a, 0.1, 0.1, 0.9\rangle, \langle b, 0.3, 0.3, 0.7\rangle, \langle c, 0.4, 0.4, 0.6\rangle\}$.*

Then, $\lambda \in NafPO(X,\tau)$, and but the set λ is neither neutrosophic afα-open nor neutrosophic afs-open.

Remark. *From Example 4.3 and Example 4.4, neutrosophic afp-open sets and neutrosophic afs-open sets are independent of each other.*

Example 4.5. *Let (X,τ) be a neutrosophic topological space, with $X = \{a,b,c\}$, $\tau = \{0_X, \lambda, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.1, 0.1, 0.9\rangle, \langle b, 0.3, 0.3, 0.7\rangle, \langle c, 0.1, 0.1, 0.9\rangle\}$ and $\mu = \{\langle a, 0.3, 0.3, 0.7\rangle, \langle b, 0.5, 0.5, 0.5\rangle, \langle c, 0.7, 0.7, 0.3\rangle\}$. Then, $\mu \in Naf\beta O(X,\tau)$, and but the set μ is not neutrosophic afp-open.*

Example 4.6. *Let (X,τ) be a neutrosophic topological space, with $X = \{a,b,c\}$, $\tau = \{0_X, \lambda, 1_X\}$, where λ, μ are two neutrosophic sets defined as $\lambda = \{\langle a, 0.2, 0.2, 0.8\rangle, \langle b, 0.8, 0.8, 0.2\rangle, \langle c, 0.5, 0.5, 0.5\rangle\}$ and $\mu = \{\langle a, 0.6, 0.6, 0.4\rangle, \langle b, 0.5, 0.5, 0.5\rangle, \langle c, 0.4, 0.4, 0.6\rangle\}$. Then, $\mu \in Naf\beta O(X,\tau)$, and but the set μ is not neutrosophic afs-open.*

5. Neutrosophic af-interior and neutrosophic af-closure operators

Definition 5.1. *Let (X,τ) be a neutrosophic topological space and λ a neutrosophic subset of X. The neutrosophic af-interior, λ°_{af}, is defined as follows : $\lambda^\circ_{af} = \bigcup\{\mu : \mu \in NafO(X,\tau), \mu \subseteq \lambda\}$.*

Theorem 5.1. *Let (X,τ) be a neutrosophic topological space and λ, μ neutrosophic subsets of X. Then the following statements are hold:*

(1) λ°_{af} is neutrosophic af-open set,
(2) $\lambda^\circ_{af} \subseteq \lambda$,
(3) λ°_{af} is the largest neutrosophic af-open subset contained in the set λ,
(4) $(\lambda^\circ_{af})^\circ_{af} = \lambda^\circ_{af}$,
(5) If $\lambda \subseteq \mu$, $\lambda^\circ_{af} \subseteq \mu^\circ_{af}$,
(6) $\lambda^\circ_{af} \cup \mu^\circ_{af} \subseteq (\lambda \cup \mu)^\circ_{af}$,
(7) $\lambda^\circ_{af} \cap \mu^\circ_{af} = (\lambda \cap \mu)^\circ_{af}$.

Proof. 1) λ°_{af} is neutrosophic af-open set. Indeed, the union of neutrosophic af-open sets belonging to the neutrosophic topological space τ is neutrosophic af-open from the Proposition 3.4.
2) It is clear from Definition 5.1.
3) Let's assume the opposite, that is, a neutrosophic af-open set β that is larger than the set λ°_{af} that the set λ contains. That is, $\lambda^\circ_{af} \subseteq \beta\lambda$. On the other hand, for every $\mu \subseteq \lambda$ neutrosophic af-open set from Definition 5.1, $\mu \subseteq \lambda^\circ_{af}$. If we take $\mu = \beta$ specifically, we find $\beta \subseteq \lambda^\circ_{af}$. Then $\beta = \lambda^\circ_{af}$ is obtained. Thus, the neutrosophic set λ°_{af} is the largest neutrosophic af-open subset contained in the set *lambda*.
4) Let $\beta = \lambda^\circ_{af}$. By (2) and Definition 5.1, $\beta = \beta^\circ_{af}$. Then, $\lambda^\circ_{af} = (\lambda^\circ_{af})^\circ_{af}$.
5) Since $\lambda \subseteq \mu$ and $\lambda^\circ_{af} \subseteq \lambda$, $\lambda^\circ_{af} \subseteq \mu$. By (2), $\mu^\circ_{af} \subseteq \mu$. From (3), since μ°_{af} is the largest neutrosophic open set contained in μ neutrosophic sets, $\lambda^\circ_{af} \subseteq \mu^\circ_{af} \subseteq \mu$. In that case $\lambda^\circ_{af} \subseteq \mu^\circ_{af}$.
6) $\lambda \subseteq \lambda \cup \mu$ and $\mu \subseteq \lambda \cup \mu$ always hold. From (5), $\lambda^\circ_{af} \subseteq (\lambda \cup \mu)^\circ_{af}$ and $\mu^\circ_{af} \subseteq (\lambda \cup \mu)^\circ_{af}$, respectively. Therefore, $\lambda^\circ_{af} \cup \mu^\circ_{af} \subseteq (\lambda \cup \mu)^\circ_{af}$.
7) It is always hold that $\lambda \cap \mu \subseteq \lambda$ and $\lambda \cap \mu \subseteq \mu$. From (5), we obtain $(\lambda \cap \mu)^\circ_{af} \subseteq \lambda^\circ_{af}$ and $(\lambda \cap \mu)^\circ_{af} \subseteq \mu^\circ_{af}$, respectively. Hence, $(\lambda \cap \mu)^\circ_{af} \subseteq \lambda^\circ_{af} \cap \mu^\circ_{af}$. On the other hand, $\lambda^\circ_{af} \subseteq \lambda$ and $\mu^\circ_{af} \subseteq \mu$. From here $\lambda^\circ_{af} \cap \mu^\circ_{af} \subseteq \lambda \cap \mu$. Since $\lambda^\circ_{af} \cap \mu^\circ_{af}$ are neutrosophic af-open sets and

$(\lambda \cap \mu)^{\circ}_{af}$ is the largest neutrosophic af-open set contained in the $\lambda \cap \mu$ neutrosophic set, we have $\lambda^{\circ}_{af} \cap \mu^{\circ}_{af} \subseteq (\lambda \cap \mu)^{\circ}_{af} \subseteq \lambda \cap \mu$. Thus, $\lambda^{\circ}_{af} \cap \mu^{\circ}_{af} = (\lambda \cap \mu)^{\circ}_{af}$. □

Theorem 5.2. *Let (X, τ) be a neutrosophic topological space and and a neutrosophic subset λ of X. Then, λ neutrosophic set to be af-open set if and only if, $\lambda^{\circ}_{af} = \lambda$.*

Proof. $\Rightarrow$ Let λ be a neutrosophic af-open set. From Theorem 5.1 (2), $\lambda^{\circ}_{af} \subseteq \lambda$. On the other hand, since λ is a neutrosophic af-open set, $\lambda \subseteq \lambda$ and by Definition 5.1, $\lambda \subseteq \lambda^{\circ}_{af}$. In that case $\lambda = \lambda^{\circ}_{af}$.
$\Leftarrow$ According to the hypothesis, let's take $\lambda = \lambda^{\circ}_{af}$. Since λ°_{af} is a neutrosophic af-open set and $\lambda = \lambda^{\circ}_{af}$, so λ is a neutrosophic af-open set. □

Lemma 5.3. *For 1_X and 0_X neutrosophic af-open sets, then $(1_X)^{\circ}_{af} = 1_X$ and $(0_X)^{\circ}_{af} = 0_X$.*

Definition 5.2. *Let (X, τ) be a neutrosophic topological space and a neutrosophic subset λ of X. The neutrosophic af-closure of λ, $\overline{\lambda}_{af}$, is defined as follows : $\overline{\lambda}_{af} = \bigcap\{\beta : \lambda \subseteq \beta, \beta \in NafC(X, \tau)\}$.*

Theorem 5.4. *Let (X, τ) be a neutrosophic topological space and λ, μ neutrosophic subsets of X. Then the following statements are hold:*

(1) $\overline{\lambda}_{af}$ is neutrosophic af-closed set,
(2) $\lambda \subseteq \overline{\lambda}_{af}$,
(3) $\overline{\lambda}_{af}$ is the smallest neutrosophic af-closed set containing λ,
(4) $\overline{(\overline{\lambda}_{af})}_{af} = \overline{\lambda}_{af}$,
(5) If $\lambda \subseteq \mu$, $\overline{\lambda}_{af} \subseteq \overline{\mu}_{af}$,
(6) $\overline{(\lambda \cap \mu)}_{af} \subseteq \overline{\lambda}_{af} \cap \overline{\mu}_{af}$,
(7) $\overline{(\lambda \cup \mu)}_{af} = \overline{\lambda}_{af} \cup \overline{\mu}_{af}$,
(8) $\overline{(1_X)}_{af} = 1_X$ and $\overline{(0_X)}_{af} = 0_X$.

Theorem 5.5. *Let λ be any neutrosophic set in a neutrosophic topological space (X, τ). Then, $\overline{(\lambda^c)}_{af} = (\lambda^{\circ}_{af})^c$ and $(\lambda^c)^{\circ}_{af} = (\overline{\lambda}_{af})^c$.*

Proof. We see that a neutrosophic af-open set $\beta \subseteq \lambda$ is precisely the complement of a neutrosophic af-closed set $\nu = \beta^c \supseteq \lambda^c$. Thus
$(\lambda)^{\circ}_{af} = \bigcup\{\nu^c : \nu \text{ is neutrosophic } af - \text{closed and } \nu \supseteq \lambda^c\}$
$= \bigcap(\{\nu : \nu \text{ is neutrosophic } af - \text{closed and } \nu \supseteq \lambda^c\})^c$
$= (\overline{(\lambda^c)}_{af})^c$
whence
$\overline{(\lambda^c)}_{af} = (\lambda^{\circ}_{af})^c$.

Next let β be any neutrosophic af-open set. Then for a neutrosophic af-closed set $\mu \supseteq \lambda$, $\beta = \mu^c \subseteq \lambda^c$. $\overline{\lambda}_{af} = \bigcap\{\beta^c : \beta \text{ is neutrosophic } af - \text{open and } \beta \subseteq \lambda^c\}$
$= \bigcup(\{\beta : \beta \text{ is neutrosophic } af - \text{open and } \beta \subseteq \lambda^c\})^c$
$= ((\lambda^c)^{\circ}_{af})^c$.
As a result
$(\lambda^c)^{\circ}_{af} = (\overline{\lambda}_{af})^c$. □

Definition 5.3. *Let β be a neutrosophic set in a neutrosophic topological space (X, τ) and $x_{r,t,s}$ is a neutrosophic point of X. β is called:*

(i) af-neighbourhood of $x_{r,t,s}$, if there exista a neutrosophic af-open set μ such that $x_{r,t,s} \in \mu \subseteq \beta$.
(ii) af-q-neighbourhood of $x_{r,t,s}$ if there exista a neutrosophic af-open set μ such that $x_{r,t,s} \in q\mu \subseteq \beta$.

Theorem 5.6. *A neutrosophic set β is neutrosophic af-open set if and only if, for each neutrosophic point $x_{r,t,s} \in \beta$, β is a af-neighbourhood of $x_{r,t,s}$.*

Proof. Straightforward. □

Definition 5.4. *Let (X, τ) be the neutrosophic topological space, λ be a neutrosophic set in (X, τ) and $x_{r,t,s}$ be a neutrosophic point. If, every af-q-neighborhood of $x_{r,t,s}$ is quasi-coincident with λ, then $x_{r,t,s}$ is said to be a af-cluster point of λ.*

Theorem 5.7. *Let β be a neutrosophic set and $x_{r,t,s}$ a neutrosophic point in a neutrosophic topological space (X, τ). Then, $x_{r,t,s} \in \overline{\beta}_{af}$ if and only if, every af-q-neighbourhood of $x_{r,t,s}$ is quasi-coincident with β.*

6. Neutrosophic af-continous functions

Definition 6.1. *A function $f : (X, \tau) \to (Y, \sigma)$ is said to be neutrosophic af-continuous, if, for each $\lambda \in \sigma$, $f^{-1}(\lambda)$ is neutrosophic af-open in (X, τ).*

Theorem 6.1. *Every neutrosophic continuous function is neutrosophic af-continuous.*

Proof. By Theorem 3.2, every neutrosophic open set is neutrosophic af-open and the proof is obvious. □

Example 6.1. *Let (X, τ), (Y, σ) be a neutrosophic topological spaces, with $X = \{a, b\}$, $Y = \{0.1, 0.4\}$, $\tau = \{0_X, \mu, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where β, μ are two neutrosophic sets defined as $\mu = \{\langle a, 0.3, 0.3, 0.7\rangle, \langle b, 0.7, 0.7, 0.3\rangle\}$ and $\beta = \{\langle 0.1, 0.2, 0.2, 0.8\rangle,$ $\langle 0.4, 0.2, 0.2, 0.8\rangle\}$ in neutrosophic topological spaces (X, τ), (Y, σ), respectively. Then, a function $f : (X, \tau) \to (Y, \sigma)$ defined as $f(a)=0.1$, $f(b)=0.4$ is neutrosophic af-continuous but not neutrosophic continuous.*

Definition 6.2. *A function $f : (X, \tau) \to (Y, \sigma)$ is said to be neutrosophic $af\alpha-$ continuous (resp. neutrosophic afp-continuous, neutrosophic afs-continuous, neutrosophic afβ-continuous) if for each $\lambda \in \sigma$, $f^{-1}(\lambda)$ is neutrosophic afα-open (resp. neutrosophic afp-open, neutrosophic afs-open, neutrosophic afβ-open) in (X, τ).*

By Definitions 6.1 and 6.2, the following implications hold:

$$
\begin{array}{ccccc}
 & & neutrosophic\ -cont & & \\
 & & \downarrow & & \\
neutrosophic\ af-cont & \to & neutrosophic\ af\alpha-cont & \to & neutrosophic\ afp-cont \\
 & & & & \downarrow \qquad \downarrow \\
 & & & & neutrosophic\ afs-cont \to \\
 & & neutrosophic\ af\beta-cont & &
\end{array}
$$

Diagram III

Remark. *None of the implications in Diagram III is reversible as shown by examples stated below.*

Example 6.2. *It can be seen from Example 6.1 that not every neutrosophic af-continuous function is a neutrosophic continuous.*

Example 6.3. *Let (X,τ), (Y,σ) be a neutrosophic topological spaces, with $X = \{a,b\}$, $Y = \{0.2, 0.5\}$, $\tau = \{0_X, \lambda, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where λ, β are two neutrosophic sets defined as $\lambda = \{\langle a, 0.7, 0.7, 0.3\rangle, \langle b, 0.4, 0.4, 0.6\rangle\}$ and $\beta = \{\langle 0.2, 0.9, 0.9, 0.1\rangle,$ $\langle 0.5, 0.1, 0.1, 0.9\rangle\}$ in neutrosophic topological spaces (X,τ), (Y,σ), respectively. Then, a function $f : (X,\tau) \to (Y,\sigma)$ defined by $f(a)=0.2$, $f(b)=0.5$ neutrosophic afα-continuous but not neutrosophic af-continuous.*

Example 6.4. *Let (X,τ), (Y,σ) be a neutrosophic topological spaces, with $X = \{a,b\}$, $Y = \{0.1, 0.4\}$, $\tau = \{0_X, \mu, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where μ, β are two neutrosophic sets defined as $\mu = \{\langle a, 0.2, 0.2, 0.8\rangle, \langle b, 0.2, 0.2, 0.8\rangle\}$ and $\beta = \{\langle 0.1, 0.3, 0.3, 0.7\rangle,$ $\langle 0.4, 0.7, 0.7, 0.3\rangle\}$ in neutrosophic topological spaces (X,τ), (Y,σ), respectively. Then, a function $f : (X,\tau) \to (Y,\sigma)$ defined as $f(a)=0.1$ and $f(b)=0.4$ is neutrosophic afs-continuous but neither neutrosophic afα-continuous nor neutrosophic afp-continuous.*

Example 6.5. *Let (X,τ), (Y,σ) be a neutrosophic topological spaces, with $X = \{a,b,c\}$, $Y = \{0.1, 0.3, 0.5\}$, $\tau = \{0_X, \mu, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where μ, β are two neutrosophic sets defined as $\mu = \{\langle a, 0.2, 0.2, 0.8\rangle, \langle b, 0.4, 0.4, 0.6\rangle, \langle c, 0.5, 0.5, 0.5\rangle\}$ and* $\beta = \{\langle 0.1, 0.4, 0.4, 0.6\rangle, \langle 0.3, 0.9, 0.9, 0.1\rangle, \langle 0.5, 0.8, 0.8, 0.2\rangle\}$ *in neutrosophic topological spaces (X,τ), (Y,σ), respectively. Then, a function $f : (X,\tau) \to (Y,\sigma)$ defined as $f(a)=0.1$, $f(b)=0.3$ and $f(c)=0.5$ is neutrosophic afp-continuous but neither neutrosophic afα-continuous nor neutrosophic afs-continuous.*

Example 6.6. *Let (X,τ), (Y,σ) be a neutrosophic topological spaces, with $X = \{a,b,c\}$, $Y = \{0.2, 0.5, 0.6\}$, $\tau = \{0_X, \lambda, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where λ, β are two neutrosophic sets defined as $\lambda = \{\langle a, 0.1, 0.1, 0.9\rangle, \langle b, 0.4, 0.4, 0.6\rangle, \langle c, 0.1, 0.1, 0.9\rangle\}$ and* $\beta = \{\langle 0.2, 0.3, 0.3, 0.7\rangle, \langle 0.5, 0.5, 0.5, 0.5\rangle, \langle 0.6, 0.8, 0.8, 0.2\rangle\}$ *in neutrosophic topological spaces (X,τ), (Y,σ), respectively. Then, a function $f : (X,\tau) \to (Y,\sigma)$ defined as $f(a)=0.2$, $f(b)=0.5$ and $f(c)=0.6$ is neutrosophic afβ-continuous but not neutrosophic afp-continuous.*

Example 6.7. *Let (X,τ), (Y,σ) be a neutrosophic topological spaces, with $X = \{a,b,c\}$, $Y = \{0.3, 0.5, 0.7\}$, $\tau = \{0_X, \lambda, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where λ, β are two neutrosophic sets defined as $\lambda = \{\langle a, 0.2, 0.2, 0.8\rangle, \langle b, 0.8, 0.8, 0.2\rangle, \langle c, 0.5, 0.5, 0.5\rangle\}$ and* $\beta = \{\langle 0.3, 0.6, 0.6, 0.4\rangle, \langle 0.5, 0.5, 0.5, 0.5\rangle, \langle 0.7, 0.4, 0.4, 0.6\rangle\}$ *in neutrosophic topological spaces (X,τ), (Y,σ), respectively. Then, a function $f : (X,\tau) \to (Y,\sigma)$ defined as $f(a)=0.3$, $f(b)=0.5$ and $f(c)=0.7$ is neutrosophic afβ-continuous but not neutrosophic afs-continuous.*

Corollary 6.2. *A function $f : (X,\tau) \to (Y,\sigma)$ is neutrosophic af-continuous if and only if, $f : (X,\tau) \to (Y,\sigma)$ is neutrosophic continuous.*

Proof. This is an immediate consequence of Theorem 3.5. □

Theorem 6.3. *A function $f : (X,\tau) \to (Y,\sigma)$ is neutrosophic af-continuous and $g : (Y,\sigma) \to (Z,\eta)$ is neutrosophic continuous, then $gof : (X,\tau) \to (Z,\eta)$ is neutrosophic af-continuous.*

Proof. It is clear. □

By using neutrosophic af-neighborhood, neutrosophic af-open sets, neutrosophic af-closed sets, neutrosophic af-interior and neutrosophic af-closure, we obtain characterizations of neutrosophic af-continuous functions.

Lemma 6.4. *Let (X,τ) be a neutrosophic topological space. A neutrosophic subset μ is neutrosophic af-closed if and only if $\overline{(\mu\cap\beta)}\subseteq\mu$ for every neutrosophic closed set β of X such that $0_X\neq\beta\neq 1_X$.*

Proof. μ is neutrosophic af-closed if and only if μ^c is neutrosophic af-open. By Definition 3.6, $\mu^c\subseteq(\mu^c\cup\alpha)^\circ$ for every $\alpha\in\tau$ such that $0_X\neq\alpha\neq 1_X$.
This is equivalent to $((\mu^c\cup\alpha)^\circ)^c\subseteq\mu$. Now, we have $((\mu^c\cup\alpha)^\circ)^c=\overline{((\mu^c\cup\alpha)^c)}=\overline{(\mu\cap\alpha^c)}$. Therefore, we obtain $\overline{(\mu\cap\beta)}\subseteq\mu$ for every neutrosophic closed set β of X such that $0_X\neq\alpha\neq 1_X$. □

Theorem 6.5. *For a function $f:(X,\tau)\to(Y,\sigma)$, the following properties are equivalent:*

(1) f is neutrosophic af-continuous;
(2) For each point $x_{r,t,s}\in X$ and each neutrosophic open set $\mu\in Y$ containing $f(x_{r,t,s})$, there exists $\alpha\in NafO(X)$ such that $x\in\alpha$, $f(\alpha)\subseteq\mu$;
(3) For each point $x_{r,t,s}\in X$ and each neutrosophic open set μ of Y containing $f(x_{r,t,s})$, there exists a neutrosophic af-neighorhood λ of $x_{r,t,s}$ such that $f(\lambda)\subseteq\mu$;
(4) The inverse image of each neutrosophic closed set in Y is neutrosophic af-closed;
(5) For each neutrosophic closed set μ of Y, $\overline{(f^{-1}(\mu)\cap\beta)}\subseteq f^{-1}(\mu)$ for every closed set in X such that $0_X\neq\beta\neq 1_X$;
(6) For each neutrosophic subset μ of Y, $\overline{(f^{-1}(\overline{(\mu)})\cap\beta)}\subseteq f^{-1}(\overline{(\mu)})$ for every neutrosophic closed set β in X such that $0_X\neq\beta\neq 1_X$;
(7) For each neutrosophic subset λ of X, $f(\overline{(\lambda\cap\beta)})\subseteq\overline{(f(\lambda))}$ for every neutrosophic closed set β in X such that $0_X\neq\beta\neq 1_X$;
(8) For each neutrosophic subset μ of Y, $\overline{(f^{-1}(\mu))}_{af}\subseteq f^{-1}(\overline{(\mu)})$;
(9) For each neutrosophic subset μ of Y, $f^{-1}((\mu)^\circ)\subseteq(f^{-1}(\mu))^\circ_{af}$.

Proof. (1) ⇒ (2): Let $x_{r,t,s}\in X$ and μ be any neutrosophic open set of Y containing $f(x_{r,t,s})$. Set $\alpha=f^{-1}(\mu)$, then by Definition 5.4, α is a neutrosophic af-open set containing $x_{r,t,s}$ and $f(\alpha)\subseteq\mu$.
(2) ⇒ (3): Every neutrosophic af-open set containing $x_{r,t,s}$ is a neutrosophic af-neighborhood of $x_{r,t,s}$ and the proof is obvious.
(3) ⇒ (1): Let μ be any neutrosophic open set in Y. For each $x_{r,t,s}\in f^{-1}(\mu)$, $f(x_{r,t,s})\in\mu\in\sigma$. By (3) there exists a neutrosophic af- neighborhood ν of $x_{r,t,s}$ such that $f(\nu)\subseteq\mu$; hence $x_{r,t,s}\in\nu\subseteq f^{-1}(\mu)$. There exists $\alpha_{x_{r,t,s}}\in NafO(X)$ such that $x_{r,t,s}\in\alpha_{x_{r,t,s}}\subseteq\nu\subseteq f^{-1}(\mu)$. Hence $f^{-1}(\mu)=\bigcup\{\alpha_{x_{r,t,s}}:x_{r,t,s}\in f^{-1}(\mu)\}\in NafO(X)$. This shows that f is neutrosophic af-continuous.
(1) ⇒ (4) ⇒ (5) ⇒ (1): By Lemma 6.4, the proof is obvious.
(5) ⇒ (6): For each neutrosophic subset μ of Y, $\overline{(\mu)}$ is neutrosophic closed in Y and the proof is obvious.
(6) ⇒ (7): Let λ be any neutrosophic subset of X. Set $\mu=f(\lambda)$, then by (6) $\overline{(\lambda\cap\beta)}\subseteq\overline{(f^{-1}(\overline{(f(\lambda))})\cap\beta)}\subseteq f^{-1}(\overline{(f(\lambda))})$ for every neutrosophic closed set β in X such that $0_X\neq\beta\neq 1_X$. Therefore, we obtain for each neutrosophic subset λ of X, $f(\overline{(\lambda\cap\beta)})\subseteq\overline{(f(\lambda))}$ for every neutrosophic closed set β in X such that $0_X\neq\beta\neq 1_X$.
(7) ⇒ (1): Let μ be any open set of Y. Then μ^c is neutrosophic closed in Y. Set $\alpha=f^{-1}(\mu^c)$, then by (7) $f(\overline{(f^{-1}(\mu^c)\cap\beta)})\subseteq\overline{(f(f^{-1}(\mu^c)))}=\mu^c$ for every neutrosophic closed set β in X such that $0_X\neq\beta\neq 1_X$. Therefore, we have

$\overline{(f^{-1}(\mu^c) \cap \beta)}$
$\subseteq f^{-1}(f(\overline{(f^{-1}(\mu^c) \cap \beta)})$
$\subseteq f^{-1}(\mu^c) = (f^{-1}(\mu))^c$.
Therefore, $f^{-1}(\mu) \subseteq \overline{((f^{-1}(\mu^c) \cap \beta))}^c$
$= ((f^{-1}(\mu^c) \cap \beta)^c)^\circ$
$= (f^{-1}(\mu) \cup \beta^c)^\circ$
$= (f^{-1}(\mu) \cup \alpha)^\circ$
for every neutrosophic open set α of X such that $0_X \neq \beta \neq 1_X$.
(4) $\Rightarrow$ (8): Let μ be any neutrosophic subset of Y. By (4) $f^{-1}(\overline{\mu})$ is neutrosophic af-closed in X and
$f^{-1}(\mu) \subseteq f^{-1}(\overline{\mu})$. Therefore, $\overline{(f^{-1}(\mu))}_{af} \subseteq f^{-1}(\overline{\mu})$.
(8) $\Rightarrow$ (9): Let μ be any neutrosophic subset of Y. Then,
$f^{-1}(\mu^\circ) = f^{-1}((\overline{\mu^c})^c)$
$=(f^{-1}(\overline{\mu^c}))^c \subseteq (\overline{((f^{-1}(\mu))^c)}_{af})^c$
$= (f^{-1}(\mu))^\circ_{af}$
(9) $\Rightarrow$ (1): Let μ be any neutrosophic open set of Y. By (9), $f^{-1}(\mu) \subseteq (f^{-1}(\mu))^\circ_{af} \subseteq f^{-1}(\mu)$. Therefore, we have $(f^{-1}(\mu))^\circ_{af} = f^{-1}(\mu)$ and hence f is neutrosophic af-continuous. □

Definition 6.3. *A function $f : (X, \tau) \to (Y, \sigma)$ is said to be neutrosophic af-irresolute if for each neutrosophic af-open set μ in (Y, σ), $f^{-1}(\mu)$ is neutrosophic af-open in (X, τ).*

Theorem 6.6. *If a function $f : (X, \tau) \to (Y, \sigma)$ is neutrosophic af-irresolute, then f is neutrosophic af-continuous.*

The converse of Theorem 6.6 is not always true as shown by the following example.

Example 6.8. *Let (X, τ), (Y, σ) be a neutrosophic topological spaces, with $X = \{a, b, c\}$, $Y = \{0.1, 0.7, 0.5\}$, $\tau = \{0_X, \lambda, 1_X\}$, $\sigma = \{0_Y, \beta, 1_Y\}$, where λ, β are two neutrosophic sets defined as $\lambda = \{\langle a, 0.3, 0.3, 0.7\rangle, \langle b, 0.2, 0.2, 0.8\rangle, \langle c, 0.5, 0.5, 0.5\rangle\}$ and*
$\beta = \{\langle 0.1, 0.3, 0.3, 0.7\rangle, \langle 0.7, 0.2, 0.2, 0.8\rangle, \langle 0.5, 0.5, 0.5, 0.5\rangle\}$
in neutrosophic topological spaces (X, τ), (Y, σ), respectively. Then, a function $f : (X, \tau) \to (Y, \sigma)$ defined as $f(a)=0.1$, $f(b)=0.7$ and $f(c)=0.5$ is neutrosophic af-continuous but not neutrosophic af-irresolute.

Definition 6.4. *A function $f : (X, \tau) \to (Y, \sigma)$ is said to be neutrosophic $af-open$ (resp. neutrosophic $af\alpha-open$, neutrosophic $afp-open$, neutrosophic $afs-open$, neutrosophic $af\beta-open$), if $f(\lambda)$ is neutrosophic af-open (resp. neutrosophic afα-open, neutrosophic afp-open, neutrosophic afs-open, neutrosophic afβ-open) in (Y, σ) for every neutrosophic open set λ in (X, τ).*

Proposition 6.7. *Every neutrosophic open function is neutrosophic af-open.*

Proof. It is obvious. □

Remark. *As can be seen from Example 3.1, the converse of Proposition 6.7 may not always be true.*

Theorem 6.8. *A function $f : (X, \tau) \to (Y, \sigma)$ is neutrosophic af-open if and only if for each neutrosophic subset μ in (Y, σ) each neutrosophic closed set β in (X, τ) containing $f^{-1}(\mu)$, there exists a neutrosophic af-closed set ν in (Y, σ) containing μ such that $f^{-1}(\nu) \subseteq \beta$.*

Proof. Necessity. Let $\nu = (f(\beta^c))^c$. Since $f^{-1}(\mu) \subseteq \beta$, we have $f(\beta^c) \subseteq \mu^c$. Since f is neutrosophic af-open, then ν is neutrosophic af-closed and $f^{-1}(\nu) = (f^{-1}(f(\beta^c)))^c \subseteq (\beta^c)^c = \beta$ Sufficieny. Let α be any neutrosophic open set in (X, τ) and $\mu = (f(\alpha))^c$. Then, $f^{-1}(\mu) = (f^{-1}(f(U)))^c \subseteq \alpha^c$ and α^c is neutrosophic closed. By the hypothesis, there exists a neutrosophic af-closed set ν in (Y, σ) containing μ such that $f^{-1}(\nu) \subseteq \alpha^c$. Then, we have $\nu \subseteq (f(\alpha))^c$. Therefore, we obtain $(f(\alpha))^c \subseteq \nu \subseteq (f(\alpha))^c$ and $f(\alpha)$ is neutrosophic af-open in (Y, σ). This shows that f is neutrosophic af-open. □

Proposition 6.9. *A function $f : (X, \tau) \to (Y, \sigma)$ is neutrosophic open and $g : (Y, \sigma) \to (Z, \eta)$ is neutrosophic af-open, then $gof : (X, \tau) \to (Z, \eta)$ is neutrosophic af-open.*

7. Conclusion

Our main aim when starting this study was to offer a new alternative to the open set types that were previously introduced in mathematics and formed the basis of many studies. In the preliminaries section of our study, some definitions that are necessary to introduce this new open set type and that we have used in our previous studies are included. In the third subheading, we redefined some open set types that have been used for a long time in topological spaces from a new perspective, and after illustrating the relationship between these open set types with the help of a diagram, we introduced the new open set type. By examining the properties of this new type of open set that we have introduced, we tried to eliminate the question marks that may arise in the minds of scientists who will conduct future research, with the help of examples, which we hope will inspire our study. In the fourth subheading of our study, we introduced open set types, which we can call sub-types of our new open set type, and after examining their properties and giving examples of these properties, we illustraed the relationship between them with the help of a diagram. In the fifth subheading, we introduced different interior and closure operators and neighborhood types with these operators with the help of our new set types. In the sixth subheading, which is the last subheading of our study, we examined new types of continuity.
Our expectation is that this study will pave the way for new research in topology and other sub-branches of mathematics. In addition, one of our primary goals is to help create new works that will contribute to human life in different branches of science.

Acknowledgments. The authors would like to thank the referees for the valuable remarks and suggestions which improved the paper.

References

[1] A. Acikgoz, H. Çakalli, F. Esenbel, and L. D. Kočinac, A quest of G-continuity in neutrosophic spaces, Math. Methods Appl. Sci. **44**(9) (2021), 7834-7844.
[2] A. Acikgöz, and F. Esenbel, Neutrosophic soft semi-regularization and neutrosophic soft sub-maximality, Filomat **36**(2) (2022), 651668.
[3] A. Acikgoz, and F Esenbel, Neutrosophic Connected topological spaces, Fasc. Math. **66** (2023), 5-22.
[4] M.H. Alqahtani, $F-$ open and $F-$ closed sets in Topological Spaces, Eur. J. Pure Appl. Math. **16**(2) (2023), 819-832.
[5] M.H. Alqahtani, and H.Y. Saleh, A Novel Class of Separation Axioms, Compactness, and Continuity via C-Open Sets, Math. **11**(23) (2023), 4729.
[6] T.M. Al-shami, and A. Mhemdi, A weak form of soft α-open sets and its applications via soft topologies, AIMS Math. **8**(5) (2023), 11373-11396.
[7] T.M. Al-shami, A. Mhemdi, and R. Abu-Gdairi, A Novel framework for generalizations of soft open sets and its applications via soft topologies, Math. **11**(4) (2023), 840.
[8] T.M. Al-shami, A. Mhemdi, A. Rawshdeh, and H.H. Al-Jarrah, On weakly soft somewhat open sets, Rocky Mountain J. Math. **54**(1) (2024), 13-30.

[9] S. Alzahran, A.I. El-Maghrabi, M.A. Al-Juhani, and M. S. Badr, New approach of soft M-open sets in soft topological spaces, J. King Saud Univ. Sci. **35**(1) (2023), 102414.
[10] T. Bera, and N.K. Mahapatra, Ranking of thermal power plants focusing on air pollution: a neutrosophic assessment, J. Clean. Prod. **316** (2021), 128250.
[11] T. Bera, and N.K. Mahapatra, On solving linear programming problem by duality approach in neutrosophic environment, Int. J. Math. Oper. Res. **18**(3) (2021), 310-335.
[12] T. Bera, and N.K. Mahapatra, Selection of waste water treatment plans alternative: a neutrosophy based MCDM approach, Opsearch, (2023), 1-24.
[13] H. Çakalli, A. Acikgoz, and F. Esenbel, A Note on Strong Forms Of Fuzzy Soft Pre-Continuity, Bol. Soc. Parana. Mat. **41** (2023), 1-13.
[14] S. Dey, and G.C. Ray, Redefined neutrosophic composite relation and its application in medical diagnosis, Int. J. Nonlinear Anal. Appl. (Special Issue for selected papers of ICDACT-2021.) **13** (2022), 43-52.
[15] R. Dhavaseelan, S. Jafari, C. Ozel, and M.A. Al-Shumrani, Generalized neutrosophic contra-continuity. In: F Smarandache, S Pramanik, eds, New Trends Neutrosophic Theory Appl. (2018) Vol. II, 365-380.
[16] V. B. Dimitri, M.G. De Jesús, and P. A. Enrique, Neutrosophic cognitive maps for analysis of social problems analysis, Int. J. Neutrosophic Sci. **19** (1) (2022).
[17] G. Carlos, On binary α-open sets and binary $\alpha - \omega$-open sets in binary topological spaces, South Asian J. Math. **11**(1) (2021), 1-11.
[18] S. Jayalakshmi, and A. Manonmani, Binary regular beta closed sets and binary regular beta open sets in binary topological spaces, Int. J. Anal. Exp. Modal Anal. **12**(4) (2020), 494-497.
[19] K. Mondal, and S. Pramanik, A study on problems of Hijras in West Bengal based on neutrosophic cognitive maps, Neutrosophic Sets Syst. **5** (2014), 21-26.
[20] K. Mondal, and S. Pramanik, Neutrosophic decision making model of school choice, Neutrosophic Sets Syst. **7**(1) (2015), 62-68.
[21] O. Mucuk, and H. Çakalli, G-Connectedness in Topological Groups with Operations, Filomat **32**(3) (2018), 1079-1089.
[22] O. Mucuk, and H. Çakalli, On G-Compactness of Topological Groups with Operations, Filomat **36**(20) (2022), 7113-7121.
[23] T.Y. Ozturk, Some Structures on Neutrosophic Topological Spaces, Appl. Math. Nonlinear Sci. **6**((1) (2021), 467-478.
[24] S. Pramanik, and S.N. Chackrabarti, A study on problems of construction workers in West Bengal based on neutrosophic cognitive maps, Int. J. Innov. Res. Sci. Eng. Tech. **2**(11) (2013), 6387-6394.
[25] H. Sahin, M. Aslantas, and A.A. Nasir, Some Extended Results for Multivalued F-Contraction Mappings, Axioms **12**(2) (2023).
[26] A.A. Salama, and S.A. Alblowi, Neutrosophic set and neutrosophic topological spaces, IOSR J. Math. **3** (2012), 31-35.
[27] F. Smarandache, Neutrosophic set, a generalisation of the intuitionistic fuzzy sets, Int. J. Pure Appl. Math. **24** (2005), 287-297.
[28] W.B.V. Kandasamy, and F. Smarandache, Analysis of Social aspects of Migrant laborers living with HIV/AIDS using Fuzzy Theory and Neutrosophic Cognitive Maps, Xiquan, Phoenix, (2004).

Ahu Acikgoz,
Department of Mathematics, Balikesir University, 10145 Balikesir, Turkey, 0000-0003-1468-8240
Email address: ahuacikgoz@gmail.com

Ferhat Esenbel.
Department of Mathematics, Balikesir University, 10145 Balikesir, Turkey, 0000-0001-5581-3621
Email address: fesenbel@gmail.com

Proceedings of International Mathematical Sciences
ISSN: 2717-6355, URL: https://dergipark.org.tr/tr/pub/pims
Volume 6 Issue 2 (2024), Pages 68-76.
Doi: https://doi.org/10.47086/pims.1577951

SOME BEST PROXIMITY POINT RESULTS ON b-METRIC SPACES WITH AN APPLICATION

MUSTAFA ASLANTAS*, AND HAKAN SAHIN**
*ÇANKIRI KARATEKIN UNIVERSITY, ÇANKIRI, TÜRKIYE. ORCID: 0000-0003-4338-3518
**BURSA TECHNICAL UNIVERSITY, BURSA, TÜRKIYE. ORCID: 0000-0002-4671-7950.

Abstract. In this paper, we introduce the concept of ϑ-p-proximal contraction mapping on b-metric spaces. Then, we obtain some best proximity results for these mappings. Also, an example to support the validity and superiority of our result has been given. Lastly, for the existence of solutions of nonlinear fractional differential equations of Caputo type we provide an application.

1. Introduction

In nonlinear analysis, game theory, approximation theory, differential equations, and control systems, fixed point theory is a crucial tool for resolving a variety of issues. As a result, numerous authors have enhanced fixed point theory. The Banach contraction principle [4], which is considered as the foundation of fixed point theory on metric spaces, was presented in this context. Let (Π, η) be a complete metric space and $\kappa : \Pi \to \Pi$ be a contraction mapping, then κ has a unique fixed point. The existence and uniqueness of fixed points in this field have been showed by numerous results [7, 11, 12]. Lately, Popescu [17] extended Banach contraction by introducing a new type of contractive condition called p-contraction. Let (Π, η) be a metric space and $\kappa : \Pi \to \Pi$ be a mapping. If there exists a ϱ in $[0, 1)$ such that

$$\eta(\kappa\check{r}, \kappa\hat{s}) \leq \varrho[\eta(\check{r}, \hat{s}) + |\eta(\check{r}, \kappa\check{r}) - \eta(\hat{s}, \kappa\hat{s})|]$$

for all $\check{r}, \hat{s} \in \Pi$, then κ is said to be a p-contraction mapping. Then, Popescu [17] proved that every p-contraction on a complete metric space has a unique fixed point.

Taking into account nonself mappings $\kappa : A \to B$ where A, B are nonempty subsets of a metric space (Π, η), the fixed point theory has recently been improved. A solution to the equation $\kappa\check{r} = \check{r}$ cannot exist if the intersection of A and B is empty. Then, it is natural to search if there is a point $\check{r}$ in A such that $\eta(\check{r}, \kappa\check{r}) = \eta(A, B)$ which is called a best proximity point of κ [6]. Numerous authors have written about this subject due to the fact that each best proximity point turns into a fixed point in the case of $A = B = \Pi$ [2, 5, 13, 14, 15].

The relevant basic definitions and symbols of best proximity point theory are now restated.

2020 *Mathematics Subject Classification.* Primary: 54H25; 47H10.
Key words and phrases. b-metric spaces; best proximity point; nonlinear fractional differential equations.

Submitted on 01.11.2024, Accepted on 09.12.2024.
Communicated by Mehmet Dik.

Let (Π, η) be a metric space and $\emptyset \neq A, B \subseteq \Pi$. We will use the subsets of A and B, respectively:

$$A_0 = \{\hat{s} \in A : \eta(\hat{s}, \check{r}) = \eta(A, B) \text{ for some } \check{r} \in B\}$$

and

$$B_0 = \{\check{r} \in B : \eta(\hat{s}, \check{r}) = \eta(A, B) \text{ for some } \hat{s} \in A\}$$

where $\eta(A, B) = \inf\{\eta(\hat{s}, \check{r}) : \hat{s} \in A \text{ and } \check{r} \in B\}$.

Definition 1.1. [5] *Let (Π, η) be a metric space and A, B be nonempty subsets of Π. A mapping $\kappa : A \to B$ is said to be proximal contraction, if there exists a real number $\varrho \in [0, 1)$ such that*

$$\left.\begin{array}{c} \eta(\hat{w}_1, \kappa\check{r}_1) = \eta(A, B) \\ \eta(\hat{w}_2, \kappa\check{r}_2) = \eta(A, B) \end{array}\right\} \Rightarrow \eta(\hat{w}_1, \hat{w}_2) \leq \varrho\eta(\check{r}_1, \check{r}_2)$$

for all $\hat{w}_1, \hat{w}_2, \check{r}_1, \check{r}_2 \in A$.

Definition 1.2. [9] *Let (Π, η) be a metric space and $\emptyset \neq A, B \subseteq \Pi$. Assume that $\vartheta : A \times A \to [0, \infty)$ is a function and $\kappa : A \to B$ is a mapping. If the following condition holds, we say that κ is ϑ-proximal admissible*

$$\left.\begin{array}{c} \vartheta(\check{r}_0, \check{r}_1) \geq 1 \\ \eta(\check{r}_1, \kappa\check{r}_0) = \eta(A, B) \\ \eta(\check{r}_2, \kappa\check{r}_1) = \eta(A, B) \end{array}\right\} \Longrightarrow \vartheta(\check{r}_1, \check{r}_2) \geq 1$$

for all $\check{r}_0, \check{r}_1, \check{r}_2 \in A$.

On the other hand, Czerwik [8, 10] established an extansion of the famous principle in a different approach than the results found in the literature by introducing a pleasant concept of a b-metric.

Definition 1.3. [10] *Let Π be a non-empty set and $\eta : \Pi \times \Pi \to [0, \infty)$ be a function satisfying for all $\hat{s}, \check{r}, z \in \Pi$,*

b1) *$\hat{s} = \check{r}$ if and only if $\eta(\hat{s}, \check{r}) = 0$,*
b2) *$\eta(\hat{s}, \check{r}) = \eta(\check{r}, \hat{s})$,*
b3) *$\eta(\hat{s}, z) \leq s[\eta(\hat{s}, \check{r}) + \eta(\check{r}, z)]$ where $s \geq 1$.*

Then, η is called a b-metric on Π with coefficient s. Also, (Π, η) is said to be a b-metric space.

Each metric space is obviously a b-metric space. The opposite might not be right, though. In fact, let $\Pi = \mathbb{R}$ and $\eta : \Pi \times \Pi \to [0, \infty)$ be a function defined by $\eta(\hat{s}, \check{r}) = (\hat{s} - \check{r})^2$ for all $\hat{s}, \check{r} \in \Pi$. Then (Π, η) is a b-metric space with the coefficient $s = 2$. Choose $\hat{s} = 7$, $\check{r} = 4$ and $z = 5$, then

$$\eta(7, 4) = 9 > 5 = \eta(7, 5) + \eta(5, 4).$$

Hence, it is not a metric space.

Let (Π, η) be a b-metric space with the coefficient $s \geq 1$. Let $\{\check{r}_n\}$ be sequence in Π and $\hat{s} \in \Pi$. Then, the sequence $\{\check{r}_n\}$ converges to $\hat{s}$ with respect to τ_η if and only if

$$\lim_{n \to \infty} \eta(\check{r}_n, \check{r}) = 0.$$

The sequence $\{\check{r}_n\}$ is called a Cauchy sequence if for all $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ satisfying $\eta(\check{r}_n, \check{r}_m) < \varepsilon$ for all $m, n \geq n_0$. (Π, η) is called a complete b-metric space if each Cauchy sequence converges to $\check{r} \in \Pi$ with respect to τ_η.

Any b-metric might not be continuous, in contrast to the regular metric. The following definition, which is crucial to our primary findings, helps us get beyond this drawback.

Definition 1.4. [3] *Let* (Π, η) *be a b-metric space with the coefficient* $s \geq 1$ *and* $\emptyset \neq A, B \subseteq \Pi$ *with* $A_0 \neq \emptyset$. *The pair* (A, B) *holds the property* (M_C) *if for every sequences* $\{\check{r}_n\}$ *in* A_0, $\{\hat{s}_n\}$ *in* B_0 *and* $\check{r} \in A$, $\hat{s} \in B$, *we have*

$$\lim_{n\to\infty} \eta(\check{r}_n, \check{r}) = \lim_{n\to\infty} \eta(\hat{s}_n, \hat{s}) = 0 \implies \lim_{n\to\infty} \eta(\check{r}_n, \hat{s}_n) = \eta(\check{r}, \hat{s}).$$

Now, we recall the following definition.

Definition 1.5. [3]*Let* (Π, η) *be a b-metric space with the coefficient* $s \geq 1$. *If each sequence* $\{\hat{s}_n\}$ *in* B *such that* $\eta(\check{r}, B) \leq \lim_{n\to\infty} \eta(\check{r}, \hat{s}_n) \leq s\eta(\check{r}, B)$ *for some* $\check{r} \in A$ *has a convergent subsequence in* B, *then* B *is called an* s*-approximately compact with respect to* A.

In this paper, we obtain some best proximity results on b-metric spaces by introducing the concept of ϑ-p-proximal contraction mapping. Also, we give an example to support the validity and superiority of our results. Finally, an application to an existence of the solution of nonlinear fractional differential equations for Caputo type is given.

2. Main Results

We begin this section with the following definition.

Definition 2.1. *Let* (Π, η) *be a b-metric space with* $s \geq 1$, $A, B \subseteq \Pi$ *with* $A_0 \neq \emptyset$. *Assume that* $\vartheta : A \times A \to [0, \infty)$ *is a function and* $\kappa : A \to B$ *is a mapping. If there exist* $\varrho \in \left[0, \frac{1}{2s-1}\right)$ *such that*

$$\left.\begin{array}{l} \eta(\hat{w}, \kappa\check{r}) = \eta(A, B) \\ \\ \eta(v, \kappa\hat{s}) = \eta(A, B) \end{array}\right\} \vartheta(\check{r}, \hat{s})\eta(\hat{w}, v) \leq \varrho\left\{\eta(\check{r}, \hat{s}) + |\eta(\check{r}, \hat{w}) - \eta(\hat{s}, v)|\right\}$$

for all $\check{r}, \hat{s}, \hat{w}, v \in A$, *then we say* κ *is an* ϑ-p*-contraction mapping.*

Now, we give an important condition for our main result.

(H) If $\{\check{r}_n\} \subseteq A_0$ is a sequence satisfying $\vartheta(\check{r}_n, \check{r}_{n+1}) \geq 1$ and $\check{r}_n \to \check{r} \in A$, then there is a subsequence $\{\check{r}_{n_k}\}$ of $\{\check{r}_n\}$ satisfying $\vartheta(\check{r}_{n_k}, \check{r}) \geq 1$ for all $k \in \mathbb{N}$.

Theorem 2.1. *Let* (Π, η) *be a complete b-metric space with* $s \geq 1$, $A, B \subseteq \Pi$ *with* $A_0 \neq \emptyset$. *Assume that the following conditions hold:*

i) *the condition* (H) *holds and* $\kappa : A \to B$ *is* ϑ*-p-proximal contraction mapping with* $\kappa(A_0) \subseteq B_0$,
ii) κ *is an* ϑ*-proximal admissible,*
iii) *the pair* (A, B) *satisfies the property* (M_C),
iv) *there are* $\check{r}_0, \check{r}_1 \in A_0$ *such that* $\eta(\check{r}_1, \kappa\check{r}_0) = \eta(A, B)$ *and* $\vartheta(\check{r}_0, \check{r}_1) \geq 1$,
v) B *is an* s*-approximately compact with respect to* A.

Then, κ *has a best proximity point* $\check{r}^*$ *in* A.*If for another best proximity point* $\hat{s}^* \in A$, $\vartheta(\check{r}^*, \hat{s}^*) \geq 1$, *then* $\check{r}^* = \hat{s}^*$.

Proof. From the condition (iv), there are $\check{r}_0, \check{r}_1 \in A_0$ such that $\eta(\check{r}_1, \kappa\check{r}_0) = \eta(A, B)$ and $\vartheta(\check{r}_0, \check{r}_1) \geq 1$. Since $\kappa\check{r}_1 \in \kappa(A_0) \subseteq B_0$, there exists $\check{r}_2 \in A_0$ such that

$$\eta(\check{r}_2, \kappa\check{r}_1) = \eta(A, B).$$

Since κ is an ϑ-proximal admissible, we get $\vartheta(\check{r}_1, \check{r}_2) \geq 1$. Similarly, since $\kappa\check{r}_2 \in \kappa(A_0) \subseteq B_0$, there exists $\check{r}_3 \in A_0$ such that

$$\eta(\check{r}_3, \kappa\check{r}_2) = \eta(A, B).$$

Since κ is an ϑ-proximal admissible, we get $\vartheta(\check{r}_2, \check{r}_3) \geq 1$. Repeating this process, we can construct a sequence $\{\check{r}_n\}$ in A such that

$$\eta(\check{r}_{n+1}, \kappa\check{r}_n) = \eta(A, B) \text{ and } \vartheta(\check{r}_n, \check{r}_{n+1}) \geq 1 \tag{2.1}$$

for all $n \geq 1$. Then, we have

$$\begin{aligned} \eta(\check{r}_n, \check{r}_{n+1}) &\leq \vartheta(\check{r}_n, \check{r}_{n+1})\eta(\check{r}_n, \check{r}_{n+1}) \\ &\leq \varrho\left\{\eta(\check{r}_{n-1}, \check{r}_n) + |\eta(\check{r}_{n-1}, \check{r}_n) - \eta(\check{r}_n, \check{r}_{n+1})|\right\} \end{aligned}$$

Suppose that there exists $n_0 \in \mathbb{N}$ such that $\eta(\check{r}_{n_0-1}, \check{r}_{n_0}) \leq \eta(\check{r}_{n_0}, \check{r}_{n_0+1})$, then we have

$$\begin{aligned} \eta(\check{r}_{n_0}, \check{r}_{n_0+1}) &\leq \vartheta(\check{r}_{n_0}, \check{r}_{n_0+1})\eta(\check{r}_{n_0}, \check{r}_{n_0+1}) \\ &\leq \varrho\left\{\eta(\check{r}_{n_0-1}, \check{r}_{n_0}) + \left|\eta(\check{r}_{n_0-1}, \check{r}_{n_0}) - \eta(\check{r}_{n_0}, \check{r}_{n_0+1})\right|\right\} \\ &= \varrho\left\{\eta(\check{r}_{n_0-1}, \check{r}_{n_0}) + \eta(\check{r}_{n_0}, \check{r}_{n_0+1}) - \eta(\check{r}_{n_0-1}, \check{r}_{n_0})\right\} \\ &= \varrho\eta(\check{r}_{n_0}, \check{r}_{n_0+1}) \\ &< \eta(\check{r}_{n_0}, \check{r}_{n_0+1}). \end{aligned}$$

This is a contradiction. Then, we assume that $\eta(\check{r}_n, \check{r}_{n+1}) < \eta(\check{r}_{n-1}, \check{r}_n)$ for all $n \geq 1$. Therefore, we get

$$\begin{aligned} \eta(\check{r}_n, \check{r}_{n+1}) &\leq \vartheta(\check{r}_n, \check{r}_{n+1})\eta(\check{r}_n, \check{r}_{n+1}) \\ &\leq \varrho\left\{\eta(\check{r}_{n-1}, \check{r}_n) + \eta(\check{r}_n, \check{r}_{n-1}) - \eta(\check{r}_n, \check{r}_{n+1})\right\} \\ &= 2\varrho\eta(\check{r}_{n-1}, \check{r}_n) - \varrho\eta(\check{r}_n, \check{r}_{n+1}) \end{aligned}$$

and so,

$$\eta(\check{r}_n, \check{r}_{n+1}) \leq \left(\frac{2\varrho}{\varrho+1}\right)\eta(\check{r}_{n-1}, \check{r}_n)$$

for all $n \geq 1$. Using the last inequality, we have

$$\begin{aligned} \eta(\check{r}_n, \check{r}_{n+1}) &\leq \left(\frac{2\varrho}{\varrho+1}\right)\eta(\check{r}_{n-1}, \check{r}_n) \\ &\leq \left(\frac{2\varrho}{\varrho+1}\right)^2\eta(\check{r}_{n-2}, \check{r}_{n-1}) \\ &\vdots \\ &\leq \left(\frac{2\varrho}{\varrho+1}\right)^n\eta(\check{r}_0, \check{r}_1) \end{aligned}$$

for all $n \in \mathbb{N}$. Now, assume $n \in \mathbb{N}$ and $p \in \mathbb{N}$. Then, we have

$$\begin{aligned} \eta(\check{r}_n, \check{r}_{n+p}) &\leq s\eta(\check{r}_n, \check{r}_{n+1}) + s^2\eta(\check{r}_{n+1}, \check{r}_{n+2}) + ... + s^p\eta(\check{r}_{n+p-1}, \check{r}_{n+p}) \\ &\leq \frac{1}{s^{n-1}}\left\{\begin{array}{c}\left(\frac{2\varrho s}{\varrho+1}\right)^n\eta(\check{r}_0, \check{r}_1) + \left(\frac{2\varrho s}{\varrho+1}\right)^{n+1}\eta(\check{r}_0, \check{r}_1) + \cdots \\ + \left(\frac{2\varrho s}{\varrho+1}\right)^{n+p-1}\eta(\check{r}_0, \check{r}_1)\end{array}\right\} \\ &= \left(\frac{2\varrho s}{\varrho+1}\right)^n\left\{1 + \frac{2\varrho s}{\varrho+1} + \cdots + \left(\frac{2\varrho s}{\varrho+1}\right)^{p-1}\right\}\eta(\check{r}_0, \check{r}_1) \\ &\leq \frac{\left(\frac{2\varrho s}{\varrho+1}\right)^n}{1 - \frac{2\varrho s}{\varrho+1}}\eta(\check{r}_0, \check{r}_1). \end{aligned}$$

Thus, $\{\check{r}_n\}$ is a Cauchy sequence in A. Since (Π, η) is a complete b-metric space and A is a closed subset of Π, there exists a $\check{r}^* \in A$ such that $\check{r}_n \to \check{r}^*$. Moreover, we have

$$\begin{aligned} \eta(\check{r}^*, B) &\leq \eta(\check{r}^*, \kappa\check{r}_n) \\ &\leq s\eta(\check{r}^*, \check{r}_{n+1}) + s\eta(\check{r}_{n+1}, \kappa\check{r}_n) \\ &= s\eta(\check{r}^*, \check{r}_{n+1}) + s\eta(A, B) \\ &\leq s\eta(\check{r}^*, \check{r}_{n+1}) + s\eta(\check{r}^*, B). \end{aligned}$$

Therefore, we get

$$\eta(\check{r}^*, B) \leq \lim_{n\to\infty} \eta(\check{r}^*, \kappa\check{r}_n) \leq s\eta(\check{r}^*, B).$$

Since B is s-approximately compact with respect to A, there exists a subsequence $\{\kappa\check{r}_{n_k}\}$ of $\{\kappa\check{r}_n\}$ such that $\kappa\check{r}_{n_k} \to \hat{s}^* \in B$ as $k \to \infty$. Therefore by taking $k \to \infty$ in $\eta(\check{r}_{n_k+1}, \kappa\check{r}_{n_k}) = \eta(A, B)$, since the pair (A, B) satisfies the property (M_C), we have $\eta(\check{r}^*, \hat{s}^*) = \eta(A, B)$, and so $\check{r}^* \in A_0$. Also, since $\kappa\check{r}^* \in \kappa(A_0) \subseteq B_0$, there exists $z \in A_0$ such that

$$\eta(z, \kappa\check{r}^*) = \eta(A, B). \tag{2.2}$$

On the other hand, using the condition (H) we can say that there exists a subsequence $\{\check{r}_{n_r}\}$ of $\{\check{r}_n\}$ such that $\vartheta(\check{r}_{n_r}, \check{r}^*) \geq 1$ for all $r \in \mathbb{N}$. Also, since κ is an ϑ-proximal admissible mapping, we have $\vartheta(\check{r}_{n_r+1}, z) \geq 1$ for all $r \in \mathbb{N}$. Now, from (2.1), (2.2) and the condition of ϑ-p-proximal contraction, we obtain

$$\begin{aligned} \eta(\check{r}_{n_r+1}, z) &\leq \vartheta(\check{r}_{n_r+1}, z)\eta(\check{r}_{n_r+1}, z) \\ &\leq \varrho\left(\eta(\check{r}_{n_r}, \check{r}^*) + \left|\eta(\check{r}_{n_r}, \check{r}_{n_r+1}) - \eta(\check{r}^*, z)\right|\right) \end{aligned}$$

for all $r \in \mathbb{N}$. Thus, taking limit as $r \to \infty$ we have

$$\eta(\check{r}^*, z) \leq \varrho\eta(\check{r}^*, z),$$

which gives $\check{r}^* = z$. From (2.2), the point $\check{r}^*$ is best proximity point of the mapping κ. Now, assume that $\check{r}^*$ and $\hat{s}^*$ are different best proximity points of κ in A and $\vartheta(\check{r}^*, \hat{s}^*) \geq 1$. Then, we get

$$\eta(\check{r}^*, \kappa\check{r}^*) = \eta(A, B)$$

and

$$\eta(\hat{s}^*, \kappa\hat{s}^*) = \eta(A, B).$$

Since the mapping κ is ϑ-p-proximal contraction, we have

$$\begin{aligned} \eta(\check{r}^*, \hat{s}^*) &\leq \vartheta(\check{r}^*, \hat{s}^*)\,\eta(\check{r}^*, \hat{s}^*) \\ &\leq \varrho\left(\eta(\check{r}^*, \hat{s}^*) + |\eta(\check{r}^*, \check{r}^*) - \eta(\hat{s}^*, \hat{s}^*)|\right) \\ &= \varrho\eta(\check{r}^*, \hat{s}^*) \end{aligned}$$

which gives $\check{r}^* = \hat{s}^*$. This is contradiction. Hence, κ has only one best proximity point. □

Example 2.1. *Let* $\Pi = \mathbb{R}^2$ *and* $\eta : \Pi \times \Pi \to \mathbb{R}$ *be a function defined by*

$$\eta((\check{r}_1, \check{r}_2), (\hat{s}_1, \hat{s}_2)) = \max\{\check{r}_1, \hat{s}_1\} + (\check{r}_2 - \hat{s}_2)^2$$

for all $(\check{r}_1, \check{r}_2), (\hat{s}_1, \hat{s}_2) \in \Pi$. *Then,* (Π, η) *is a b-metric space with coefficient* $s = 2$. *Now, consider the sets* $A = [0, 1] \times \{0\}$ *and* $B = [0, 1] \times \{1\}$. *We get* $\eta(A, B) = 1$, A *is closed,* $A_0 = \{(0, 0)\}$ *and* $B_0 = \{(0, 1)\}$. *Also, it can be seen that* B *is an* s*-approximately compact with respect to* A *and the pair* (A, B) *satisfies the property* (M_C). *Define a function* $\vartheta : A \times A \to [0, \infty)$ *and a mapping* $\kappa : A \to B$ *as*

$$\vartheta(\check{r},\hat{s}) = \begin{cases} 1 & , \quad \check{r},\hat{s} \in A_0 \\ \\ 0 & , \quad otherwise \end{cases}$$

and

$$\kappa(\check{r}_1, 0) = \left(\frac{\check{r}_1}{2}, 1\right),$$

respectively. Then, we can choose $\check{r}_0 = (0,0) = \check{r}_1 \in A_0$ *such that* $\eta(\check{r}_1, \kappa\check{r}_0) = \eta(A,B)$ *and* $\vartheta(\check{r}_0, \check{r}_1) \geq 1$*, and so the condition (iv) hold. Also, it is clear that* $\kappa : A \to B$ *is an* ϑ*-proximal admissible and* ϑ*-p-proximal contraction mapping with* $\kappa(A_0) \subseteq B_0$*. Hence, all hypotheses of Theorem 2.1 are satisfied, and so there is a unique best proximity point* $\check{r} = (0,0)$ *of* κ*.*

If we take $A = B = \Pi$ in Theorem 2.1, then we obtain the following fixed point result.

Corollary 2.2. *Let* (Π, η) *be a complete b-metric space with the coefficient* $s \geq 1$ *and* $\kappa : \Pi \to \Pi$ *be a continuous mapping. If the following conditions hold,*

i) *for all* $\hat{s}, \check{r} \in \Pi$*, it is satisfied*

$$\vartheta(\check{r},\hat{s})\eta(\kappa\check{r},\kappa\hat{s}) \leq \varrho\left\{\eta(\check{r},\hat{s}) + |\eta(\check{r},\kappa\check{r}) - \eta(\hat{s},\kappa\hat{s})|\right\},$$

ii) *If* $\{\check{r}_n\} \subseteq \Pi$ *is a sequence such that* $\vartheta(\check{r}_n, \check{r}_{n+1}) \geq 1$ *and* $\check{r}_n \to \check{r} \in \Pi$*, then there exists a subsequence* $\{\check{r}_{n_k}\}$ *of* $\{\check{r}_n\}$ *such that* $\vartheta(\check{r}_{n_k}, \check{r}) \geq 1$ *for all* $k \in \mathbb{N}$*.*

iii) κ *is an* ϑ*-admissible,*

iv) *there is* $\check{r}_0 \in \Pi$ *such that* $\vartheta(\check{r}_0, \kappa\check{r}_0) \geq 1$,

then, κ *has a fixed point* $\check{r}^*$ *in* Π*.If for another fixed point* $\hat{s}^* \in \Pi$*,* $\vartheta(\check{r}^*, \hat{s}^*) \geq 1$*, then* $\check{r}^* = \hat{s}^*$*.*

If we take $s = 1$ in Corollary 2.2, then we obtain the following fixed point result.

Corollary 2.3. *Let* (Π, η) *be a complete metric space and* $\kappa : \Pi \to \Pi$ *be a mapping. If the following conditions hold,*

i) *for all* $\hat{s}, \check{r} \in \Pi$*, it is satisfied*

$$\vartheta(\check{r},\hat{s})\eta(\kappa\check{r},\kappa\hat{s}) \leq \varrho\left\{\eta(\check{r},\hat{s}) + |\eta(\check{r},\kappa\check{r}) - \eta(\hat{s},\kappa\hat{s})|\right\}, \tag{2.3}$$

ii) *If* $\{\check{r}_n\} \subseteq \Pi$ *is a sequence such that* $\vartheta(\check{r}_n, \check{r}_{n+1}) \geq 1$ *and* $\check{r}_n \to \check{r} \in \Pi$*, then there exists a subsequence* $\{\check{r}_{n_k}\}$ *of* $\{\check{r}_n\}$ *such that* $\vartheta(\check{r}_{n_k}, \check{r}) \geq 1$ *for all* $k \in \mathbb{N}$*.*

iii) κ *is an* ϑ*-admissible,*

iv) *there is* $\check{r}_0 \in \Pi$ *such that* $\vartheta(\check{r}_0, \kappa\check{r}_0) \geq 1$,

then, κ *has a fixed point* $\check{r}^*$ *in* Π*. If for another fixed point* $\hat{s}^* \in \Pi$*,* $\vartheta(\check{r}^*, \hat{s}^*) \geq 1$*, then* $\check{r}^* = \hat{s}^*$*.*

3. Application

For nonlinear fractional differential equations of Caputo type, we provide adequate requirements for their existence and uniqueness in this section. For a continuous function $g : [0,\infty) \to \mathbb{R}$ of order $\alpha > 0$, the Caputo derivative is defined as

$${}^{C}D^{\alpha}(g(\gamma)) = \frac{1}{\Gamma(n-\alpha)}\int_0^{\gamma}(\gamma - s)^{n-\alpha-1}g^{(n)}(s)ds, \alpha > 0, n-1 < \alpha < n$$

where Γ is the gamma function and n is an integer.

The following nonlinear fractional differential equation of Caputo type

$${}^{C}D^{\alpha}(g(\gamma)) = f(\gamma, \check{r}(\gamma)) \tag{3.1}$$

with integral boundary conditions

$$\check{r}(0) = 0 \text{ and } \check{r}(1) = \int_0^{\theta} \check{r}(\hat{w})d\hat{w}$$

where $1 < \alpha \leq 2$, $0 < \theta < 1$, $\check{r} \in C[0,1]$ which is the space of all continuous real-valued functions defined on $[0,1]$ and $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ is a continuous function. Since f is a continuous, the equation (3.1) is equivalent to the following integral equation [1, 16]:

$$\begin{aligned} \check{r}(\gamma) &= \frac{1}{\Gamma(\alpha)} \int_0^{\gamma} (\gamma - \hat{w})^{\alpha-1} f(\hat{w}, \check{r}(\hat{w}))d\hat{w} \\ &\quad - \frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)} \int_0^{1} (1 - \hat{w})^{\alpha-1} f(\hat{w}, \check{r}(\hat{w}))d\hat{w} \\ &\quad + \frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)} \int_0^{\theta} \left(\int_0^{\hat{w}} (\hat{w} - r)^{\alpha-1} f(r, \check{r}(r))dr \right) d\hat{w}. \end{aligned} \tag{3.2}$$

Theorem 3.1. *Let $\chi : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a function. Suppose the following conditions hold:*

i) *the mapping $\kappa : C[0,1] \to C[0,1]$*

$$\begin{aligned} \kappa\check{r}(\gamma) &= \frac{1}{\Gamma(\alpha)} \int_0^{\gamma} (\gamma - \hat{w})^{\alpha-1} f(\hat{w}, \check{r}(\hat{w}))d\hat{w} \\ &\quad - \frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)} \int_0^{1} (1 - \hat{w})^{\alpha-1} f(\hat{w}, \check{r}(\hat{w}))d\hat{w} \\ &\quad + \frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)} \int_0^{\theta} \left(\int_0^{\hat{w}} (\hat{w} - r)^{\alpha-1} f(r, \check{r}(r))dr \right) d\hat{w} \end{aligned}$$

for all $\check{r} \in C[0,1]$ and $\gamma \in [0,1]$, is a continuous mapping where $1 < \alpha \leq 2$, $0 < \theta < 1$.

ii) *there exists $\check{r}_0 \in C[0,1]$ such that $\chi(\check{r}_0(\gamma), \kappa\check{r}_0(\gamma)) \geq 0$ for all $\gamma \in [0,1]$.*

iii) *if for each $\gamma \in [0,1]$ and $\check{r}, \hat{s} \in C[0,1]$, $\chi(\check{r}(\gamma), \hat{s}(\gamma)) \geq 0$, then $\chi(\kappa\check{r}(\gamma), \kappa\hat{s}(\gamma)) \geq 0$.*

iv) *for each sequence $\{\check{r}_n\} \subseteq C[0,1]$ such that for all $\gamma \in [0,1]$, $\{\check{r}_n(\gamma)\}$ converges to $\check{r}(\gamma)$ for some $\check{r} \in C[0,1]$ and $\chi(\check{r}_n(\gamma), \check{r}_{n+1}(\gamma)) \geq 0$ for all $n \geq 1$, then there exists a subsequence $\{\check{r}_{n_k}\}$ of $\{\check{r}_n\}$ such that $\chi(\check{r}_{n_k}(\gamma), \check{r}(\gamma)) \geq 0$ for all $\gamma \in [0,1]$ and $k \geq 1$.*

v) *there exists q in $[0,1)$ such that*

$$|f(\hat{w}, \check{r}(\hat{w})) - f(\hat{w}, \hat{s}(\hat{w}))| \leq \frac{\Gamma(\alpha+1)}{5} \left\{ \varrho \left(|\check{r}(\hat{w}) - \hat{s}(\hat{w})|^2 \right) + N(\check{r}, \kappa) \right\}^{\frac{1}{2}}$$

where $\varrho \in [0,1)$ and $N(\check{r}, \kappa) = \left| \sup_{\hat{w}\in[0,1]} |\check{r}(\hat{w}) - \kappa\check{r}(\hat{w})|^2 - \sup_{\hat{w}\in[0,1]} |\hat{s}(\hat{w}) - \kappa\hat{s}(\hat{w})|^2 \right|$. Then, the problem (3.1) has a unique solution.

Proof. Let $\Pi = C[0,1]$ and $\eta : \Pi \times \Pi \to [0,\infty)$ a function defined by

$$\eta(\hat{w}, v) = \sup_{\gamma\in[0,1]} |\hat{w}(\gamma) - v(\gamma)|^2$$

for all $\gamma \in [0,1]$ and $\hat{w}, v \in \Pi$. Hence, (Π, η) is a complete b-metric space with $s = 2$. We shall show that κ satisfies the inequality (2.3). Let's take $\check{r}, \hat{s} \in \Pi$ with $\chi(\check{r}(\gamma), \hat{s}(\gamma)) \geq 0$

for all $\gamma \in [0,1]$.Then, we have

$$
\begin{aligned}
|\kappa\check{r}(\gamma)-\kappa\hat{s}(\gamma)| &= \left|\begin{array}{c}
\frac{1}{\Gamma(\alpha)}\int_0^{\gamma}(\gamma-\hat{w})^{\alpha-1}f(\hat{w},\check{r}(\hat{w}))d\hat{w}\\
-\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\int_0^{1}(1-\hat{w})^{\alpha-1}f(\hat{w},\check{r}(\hat{w}))d\hat{w}\\
+\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\int_0^{\theta}\left(\int_0^{\hat{w}}(\hat{w}-r)^{\alpha-1}f(r,\check{r}(r))dr\right)d\hat{w}\\
-\frac{1}{\Gamma(\alpha)}\int_0^{\gamma}(\gamma-\hat{w})^{\alpha-1}f(\hat{w},\hat{s}(\hat{w}))d\hat{w}\\
+\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\int_0^{1}(1-\hat{w})^{\alpha-1}f(\hat{w},\hat{s}(\hat{w}))d\hat{w}\\
-\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\int_0^{\theta}\left(\int_0^{\hat{w}}(\hat{w}-r)^{\alpha-1}f(r,\hat{s}(r))dr\right)d\hat{w}
\end{array}\right|\\
&\leq \frac{1}{\Gamma(\alpha)}\left\{\int_0^{\gamma}|\gamma-\hat{w}|^{\alpha-1}\left(|f(\hat{w},\check{r}(\hat{w}))-f(\hat{w},\hat{s}(\hat{w}))|\right)d\hat{w}\right\}\\
&\quad+\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\left\{\int_0^{1}(1-\hat{w})^{\alpha-1}\left(|f(\hat{w},\check{r}(\hat{w}))-f(\hat{w},\hat{s}(\hat{w}))|\right)d\hat{w}\right\}\\
&\quad+\frac{2\gamma}{(2-\theta^2)\Gamma(\alpha)}\left\{\int_0^{\theta}\left(\int_0^{\hat{w}}|\hat{w}-r|^{\alpha-1}\left(|f(r,\check{r}(r))-f(r,\hat{s}(r))|\right)dr\right)d\hat{w}\right\}\\
&\leq \int_0^{\gamma}\left(\begin{array}{c}\frac{|\gamma-\hat{w}|^{\alpha-1}}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{5}\\ \times\left\{\ \varrho\left(|\check{r}(\hat{w})-\hat{s}(\hat{w})|^2+N(\check{r},\kappa)\right)\ \right\}^{\frac{1}{2}}\end{array}\right)d\hat{w}\\
&\quad+\frac{2\gamma}{(2-\theta^2)}\int_0^{1}\left(\begin{array}{c}\frac{(1-\hat{w})^{\alpha-1}}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{5}\\ \times\left\{\ \varrho\left(|\check{r}(\hat{w})-\hat{s}(\hat{w})|^2+N(\check{r},\kappa)\right)\ \right\}^{\frac{1}{2}}\end{array}\right)d\hat{w}\\
&\quad+\frac{2\gamma}{(2-\theta^2)}\int_0^{\theta}\left(\int_0^{\hat{w}}\left[\begin{array}{c}\frac{|\hat{w}-r|^{\alpha-1}}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{5}\\ \times\left\{\ \varrho\left(|\check{r}(\hat{w})-\hat{s}(\hat{w})|^2+N(\check{r},\kappa)\right)\ \right\}^{\frac{1}{2}}\end{array}\right]dr\right)d\hat{w}\\
&\leq \frac{\Gamma(\alpha+1)}{5}\left\{\varrho\left(\eta(\check{r},\hat{s})+|\eta(\check{r},\kappa\check{r})-\eta(\hat{s},\kappa\hat{s})|\right)\right\}^{\frac{1}{2}}\\
&\quad\times\sup_{\gamma\in[0,1]}\left\{\frac{1}{\Gamma(\alpha+1)}+\frac{2\gamma}{(2-\theta^2)}\left(\frac{1}{\Gamma(\alpha+1)}+\frac{1}{\Gamma(\alpha+1)}\right)\right\}\\
&\leq \left\{\varrho\left(\eta(\check{r},\hat{s})+|\eta(\check{r},\kappa\check{r})-\eta(\hat{s},\kappa\hat{s})|\right)\right\}^{\frac{1}{2}}
\end{aligned}
$$

which implies that

$$\eta(\kappa\check{r},\kappa\hat{s})\leq\varrho\left(\eta(\check{r},\hat{s})+|\eta(\check{r},\kappa\check{r})-\eta(\hat{s},\kappa\hat{s})|\right) \tag{3.3}$$

$\check{r},\hat{s}\in\Pi$ with $\chi\left(\check{r}(\gamma),\hat{s}(\gamma)\right)\geq 0$ for all $\gamma\in[0,1]$. Now, consider the mapping $\alpha:\Pi\times\Pi\to[0,\infty)$ defined by

$$\alpha(\check{r},\hat{s})=\begin{cases}1 & ,\quad \chi\left(\check{r}(\gamma),\hat{s}(\gamma)\right)\geq 0 \text{ for all } \gamma\in[0,1]\\ 0 & ,\quad \text{otherwise}\end{cases}.$$

Then, the inequality (3.3) is satisfied for all $\check{r},\hat{s}\in\Pi$ with $\alpha(\check{r},\hat{s})\geq 1$, that is, for all $\check{r},\hat{s}\in\Pi$,

$$\alpha(\check{r},\hat{s})\eta(\kappa\check{r},\kappa\hat{s})\leq\varrho\left(\eta(\check{r},\hat{s})+|\eta(\check{r},\kappa\check{r})-\eta(\hat{s},\kappa\hat{s})|\right)$$

is satisfied. Also, from (iii), the mapping κ is α-admissible mapping. Using the conditions (ii) and (iv) we say that all conditions of Corollary 2.3 are met, and so κ has a fixed point. Therefore, there is a solution to the nonlinear fractional differential equation of Caputo type (3.1) □

References

[1] R. P. Agarwal, M. Benchohra and S. Hamani, A survey on existence results for boundary value problems of nonlinear fractional differential equations and inclusions, Acta Applicandae Mathematicae, **109**(3) (2010), 973–1033.

[2] I. Altun, M. Aslantas and H. Sahin, Best proximity point results for *p*-proximal contractions, Acta Math. Hungar., **162** (2020), 393–402.

[3] M: Aslantaş, Best proximity point theorems for proximal b-cyclic contractions on b-metric spaces. Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics, **70**(1) (2021), 483–496.

[4] S. Banach, Sur les opérations dans les ensembles abstraits et leur applications aux équations intégrales, Fund. Math, **3** (1922), 133–181.

[5] S. S. Basha, Extensions of Banach's contraction principle, Numer. Funct. Anal. Optim., **31** (2010), 569–576.

[6] S. S. Basha and P. Veeramani, Best approximations and best proximity pairs, Acta Sci. Math., **63** (1997), 289–300.

[7] V. Berinde, Contracţii generalizate şi aplicaţii, Editura Club Press 22, Baia Mare, 1997.

[8] N. Hussain, Z. D. Mitrović and S. Radenović, A common fixed point theorem of isher in *b*-metric spaces, Revista de la Real Academia de Ciencias Exactas, Físicasy Naturales. Serie A. Matemáticas, **113** (2019), 949–956.

[9] M. Jleli, B. Samet, Best proximity points for α–ψ-proximal contractive type mappings and applications. Bulletin des Sciences Math´ematiques **137**(8) (2013), 977–995.

[10] P. D. Proinov, Fixed point theorems for generalized contractive mappings in metric spaces, Journal of Fixed Point Theory and Applications, **22** (2020).

[11] S. Reich and A. J. Zaslavski, Existence of a unique fixed point for nonlinear contractive mappings, Mathematics, **8** (2020).

[12] I. A. Rus, Generalized contractions and applications, Cluj University Press, Cluj-Napoca, (2001).

[13] H. Sahin, M. Aslantas and I. Altun, Feng-Liu type approach to best proximity point results for multivalued mappings, J. Fixed Point Theory Appl, **22** (2020), https://doi.org/10.1007/s11784-019-0740-9.

[14] V. Sankar Raj, A best proximity point theorem for weakly contractive non-self-mappings,Nonlinear Anal. TMA, **74** (2011), 4804–4808.

[15] V. Sankar Raj, Best proximity point theorems for non-self mappings, Fixed Point Theory, **14** (2013), 447–454.

[16] W. Shatanawi, E. Karapinar, H. Aydi, and A. Fulga, Wardowski type contractions with applications on Caputo type nonlinear fractional differential equations, University Politechnica of Bucharest Scientfic Bulletin, Series A, **82**(2) (2020), 157-170.

[17] O. Popescu, A new type of contractive mappings in complete metric spaces, Bull. Transilv. Univ. Braşov, Ser. III, Math. Inform. Phys., **1**(50) (2008), 479–482.

Mustafa Aslantas,

Department of Mathematics, Faculty of Science, Çankiri Karatekin University, 18100 Çankiri, Türkiye, Phone: +90(539) 2721158 Orcid:0000-0003-4338-3518

Email address: maslantas@karatekin.edu.tr

Hakan Sahin,

Department of Mathematics, Bursa Technical University, 16310 Bursa, Türkiye, Phone:+90(542)7308137 Orcid:0000-0002-4671-7950

Email address: hakan.sahin@btu.edu.tr

Proceedings of International Mathematical Sciences
ISSN: 2717-6355, URL: https://dergipark.org.tr/tr/pub/pims
Volume 6 Issue 2 (2024), Pages 77-99.
Doi: https://doi.org/ 10.47086/pims.1579364

HOW ANALYSIS CAN TEACH US THE OPTIMAL WAY TO DESIGN NEURAL OPERATORS

VU ANH LE*, AND MEHMET DIK**
*BELOIT COLLEGE, USA. ORCID: 0009-0000-1904-5186
**ROCKFORD UNIVERSITY, USA. ORCID: 0000-0003-0643-2771.

Abstract. This paper presents a mathematics-informed approach to neural operator design, building upon the theoretical framework established in our prior work [1]. By integrating rigorous mathematical analysis with practical design strategies, we aim to enhance the stability, convergence, generalization, and computational efficiency of neural operators. We revisit key theoretical insights, including stability in high dimensions, exponential convergence, and universality of neural operators. Based on these insights, we provide detailed design recommendations, each supported by mathematical proofs and citations. Our contributions offer a systematic methodology for developing next-gen neural operators with improved performance and reliability.

1. Introduction

Neural operators have changed the way we approach problems involving mappings between infinite-dimensional function spaces, particularly in solving partial differential equations (PDEs) [2, 3, 4]. By extending the capabilities of neural networks from finite-dimensional data to function spaces, architectures such as the Fourier Neural Operator (FNO) and Deep Operator Network (DeepONet) have demonstrated significant success in approximating solution operators with significantly reduced computational costs.

In our prior work [1], we developed a mathematical framework for analyzing neural operators, proving their stability, convergence properties, and capacity for universal approximation between function spaces. We also established probabilistic bounds on generalization error, linking it to sample size and network capacity. Building upon this foundation, the primary objective of this paper is to translate these theoretical insights into actionable design recommendations for neural operators. By doing so, we aim to bridge the gap between theory and practice, suggesting better neural operator architectures and saving time in design.

The remainder of the paper is organized as follows. In Section 2, we reinstate the theoretical results from the prior paper [1], including the definitions of neural operators and the key theorems related to the behaviors of neural operators. In Section 3, we present our detailed design recommendations, illustrating how each recommendation enhances neural

2020 *Mathematics Subject Classification.* Primary: 65N12; 68T05.
Key words and phrases. Neural operators, functional analysis, convergence, stability, generalization.

Submitted on 04.11.2024, Accepted on 12.11.2024.
Communicated by Huseyin Cakalli.

operator performance. Finally, Appendix A contains the full proofs of all the theorems, lemmas, and propositions presented in this paper.

2. Theoretical Framework Reminder

In this section, we provide a concise summary of the key theoretical results established in our previous work [1]. These results form the foundation upon which we build our design recommendations for neural operators.

2.1. **Stability in High-Dimensional PDEs.**

Theorem 2.1 (Stability of Neural Operators in High-Dimensional PDEs)**.** *Let $\mathcal{G}_\theta : H^s(D) \to H^t(D)$ be a neural operator parameterized by θ, mapping between Sobolev spaces over a domain $D \subset \mathbb{R}^d$. Suppose $\mathcal{G}_\theta$ satisfies a Lipschitz continuity condition:*

$$\|\mathcal{G}_\theta(u) - \mathcal{G}_\theta(v)\|_{H^t(D)} \leq L\|u - v\|_{H^s(D)}$$

for all $u, v \in H^s(D)$ and some Lipschitz constant $L > 0$. Then, for any $u \in H^s(D)$, the neural operator produces stable approximations in high-dimensional D:

$$\|\mathcal{G}_\theta(u)\|_{H^t(D)} \leq L\|u\|_{H^s(D)} + C,$$

where $C = \|\mathcal{G}_\theta(0)\|_{H^t(D)}$ is a constant depending on θ and the domain D.

2.2. **Exponential Convergence.**

Theorem 2.2 (Exponential Convergence of Neural Operator Approximations)**.** *Let $\mathcal{G}_\theta$ be a contraction mapping on $H^t(D)$ with contraction constant $0 < q < 1$. Then, for any $u \in H^t(D)$, the iterated application $\mathcal{G}_\theta^n(u)$ converges exponentially to the fixed point u^* of $\mathcal{G}_\theta$:*

$$\|\mathcal{G}_\theta^n(u) - u^*\|_{H^t(D)} \leq q^n\|u - u^*\|_{H^t(D)}.$$

2.3. **Universality and Generalization.**

Theorem 2.3 (Universality of Neural Operators for PDE Solvers)**.** *Let $\mathcal{T} : H^s(D) \to H^t(D)$ be a continuous operator. Then, for any $\epsilon > 0$, there exists a neural operator $\mathcal{G}_\theta$ such that:*

$$\|\mathcal{G}_\theta(u) - \mathcal{T}(u)\|_{H^t(D)} \leq \epsilon,$$

for all u in a compact subset of $H^s(D)$.

Theorem 2.4 (Generalization Error of Neural Operators)**.** *Let $\mathcal{G}_\theta$ be a neural operator trained on N samples $\{(u_i, \mathcal{T}(u_i))\}$ drawn i.i.d. from a distribution $\mathcal{D}$. Suppose $\mathcal{G}_\theta$ has Lipschitz constant L with respect to θ, and the loss function ℓ is Lipschitz and bounded. Then, with probability at least $1 - \delta$, the generalization error satisfies:*

$$\mathbb{E}_{u\sim\mathcal{D}}[\ell(\mathcal{G}_\theta(u), \mathcal{T}(u))] \leq \frac{1}{N}\sum_{i=1}^{N} \ell(\mathcal{G}_\theta(u_i), \mathcal{T}(u_i)) + L\sqrt{\frac{\ln(1/\delta)}{2N}}.$$

3. Design Recommendations for Neural Operators

Based on the theoretical insights from the previous section, we propose several design recommendations to enhance neural operator performance. Each recommendation is supported by detailed theorems, lemmas, and proofs, either directly or in Appendix A, to examine their impacts.

3.1. **Design Neural Operators as Contraction Mappings.** Ensuring that the neural operator $\mathcal{G}_\theta$ satisfies the contraction property guarantees stability and exponential convergence. By designing $\mathcal{G}_\theta$ as a contraction mapping, we leverage the Banach Fixed Point Theorem [5] to ensure the existence and uniqueness of a fixed point, as well as exponential convergence to that fixed point. This approach enhances both the stability and efficiency of the neural operator when approximating solutions to partial differential equations (PDEs).

To ensure that $\mathcal{G}_\theta$ is a contraction mapping, we must design the neural network components to satisfy certain Lipschitz conditions. Specifically, we have the following theorem:

Theorem 3.1 (Lipschitz Condition for Neural Networks)**.** *Suppose each layer of the neural operator $\mathcal{G}_\theta$ is Lipschitz continuous with Lipschitz constant L_i, and the activation functions are Lipschitz continuous with Lipschitz constant L_σ. Then the overall Lipschitz constant L of $\mathcal{G}_\theta$ satisfies:*

$$L \leq \left(\prod_{i=1}^{N} L_i \right) L_\sigma^N,$$

where N is the number of layers.

Proof. See Appendix A.1. □

By constraining the spectral norm of each weight matrix W_i to be less than or equal to $q^{1/N}$, where $q \in (0, 1)$ is the desired contraction constant, and choosing activation functions with Lipschitz constant $L_\sigma \leq 1$, we can ensure that $\mathcal{G}_\theta$ becomes a contraction mapping with contraction constant $L \leq q$. This is formalized in the following corollary:

Corollary 3.2 (Ensuring Contraction via Spectral Normalization)**.** *By constraining $\|W_i\| \leq q^{1/N}$ and choosing $L_\sigma \leq 1$, the overall Lipschitz constant satisfies $L \leq q$, ensuring that $\mathcal{G}_\theta$ is a contraction mapping with contraction constant q.*

Proof. See Appendix A.1. □

Designing $\mathcal{G}_\theta$ as a contraction mapping enhances stability by ensuring that small perturbations in the input lead to proportionally smaller changes in the output. Specifically, we have:

Lemma 3.3 (Stability of Contraction Mappings)**.** *A contraction mapping $\mathcal{G}_\theta$ on a metric space $(X, \|\cdot\|)$ satisfies:*

$$\|\mathcal{G}_\theta(u + \delta u) - \mathcal{G}_\theta(u)\| \leq q\|\delta u\|,$$

where δu is a small perturbation in the input.

Proof. We aim to show that if $\mathcal{G}_\theta$ is a contraction mapping on a metric space $(X, \|\cdot\|)$ with contraction constant $q \in [0, 1)$, then for any $u \in X$ and any perturbation $\delta u \in X$, the following inequality holds:

$$\|\mathcal{G}_\theta(u + \delta u) - \mathcal{G}_\theta(u)\| \leq q\|\delta u\|.$$

We start by defining the contraction mapping. A mapping $\mathcal{G}_\theta : X \to X$ is called a contraction mapping if there exists a constant $q \in [0, 1)$ such that for all $x, y \in X$,

$$\|\mathcal{G}_\theta(x) - \mathcal{G}_\theta(y)\| \leq q\|x - y\|.$$

Let $u \in X$ be any point in the metric space, and let $\delta u \in X$ be a perturbation. We consider the images of u and $u + \delta u$ under the mapping $\mathcal{G}_\theta$.

Applying the contraction property to $x = u + \delta u$ and $y = u$, we have:

$$\|\mathcal{G}_\theta(u + \delta u) - \mathcal{G}_\theta(u)\| \leq q\|(u + \delta u) - u\| = q\|\delta u\|.$$

This inequality directly shows that the change in the output of $\mathcal{G}_\theta$ due to the perturbation δu is at most q times the magnitude of the perturbation. Since $q < 1$, the mapping $\mathcal{G}_\theta$ attenuates the effect of the perturbation.

So far, the lemma demonstrates that $\mathcal{G}_\theta$ is Lipschitz continuous with Lipschitz constant q:

$$\|\mathcal{G}_\theta(u + \delta u) - \mathcal{G}_\theta(u)\| \leq q\|\delta u\|.$$

This property implies stability with respect to input perturbations, meaning that small changes in the input u result in proportionally smaller changes in the output $\mathcal{G}_\theta(u)$. This is crucial for ensuring that errors or uncertainties in the input do not amplify through the mapping, which is particularly important in iterative methods and numerical computations. □

Moreover, the exponential convergence to the fixed point reduces computational effort by potentially decreasing the number of iterations or layers required to achieve a desired level of accuracy.

Theorem 3.4 (Reduction in Iterations Needed for Convergence)**.** *Let $\epsilon > 0$ be the desired accuracy. The number of iterations n required to achieve $\|\mathcal{G}_\theta^n(u) - u^*\| \leq \epsilon$ is bounded by:*

$$n \geq \frac{\ln\left(\frac{\|u-u^*\|}{\epsilon}\right)}{\ln\left(\frac{1}{q}\right)}.$$

Proof. We aim to determine a bound on the number of iterations n required for the iterated mapping $\mathcal{G}_\theta^n(u)$ to approximate the fixed point u^* within a desired accuracy $\epsilon > 0$, i.e.,

$$\|\mathcal{G}_\theta^n(u) - u^*\| \leq \epsilon.$$

Recall that a contraction mapping $\mathcal{G}_\theta$ on a complete metric space $(X, \|\cdot\|)$ has a unique fixed point $u^* \in X$ satisfying $\mathcal{G}_\theta(u^*) = u^*$. Moreover, the sequence $\{\mathcal{G}_\theta^n(u)\}_{n=0}^\infty$, where $\mathcal{G}_\theta^n$ denotes the n-fold composition of $\mathcal{G}_\theta$, converges to u^* for any initial point $u \in X$.

The contraction property ensures that:

$$\|\mathcal{G}_\theta(u) - \mathcal{G}_\theta(v)\| \leq q\|u - v\|, \quad \text{for all } u, v \in X,$$

where $q \in [0, 1)$ is the contraction constant.

We first establish the rate at which the iterates $\mathcal{G}_\theta^n(u)$ converge to u^*. Using the contraction property repeatedly, we have:

$$\begin{aligned}
\|\mathcal{G}_\theta^n(u) - u^*\| &= \|\mathcal{G}_\theta^n(u) - \mathcal{G}_\theta^n(u^*)\| \\
&\leq q\|\mathcal{G}_\theta^{n-1}(u) - u^*\| \\
&\leq q^2\|\mathcal{G}_\theta^{n-2}(u) - u^*\| \\
&\vdots \\
&\leq q^n\|u - u^*\|.
\end{aligned}$$

To achieve the desired accuracy ϵ, we require:

$$\|\mathcal{G}_\theta^n(u) - u^*\| \leq q^n\|u - u^*\| \leq \epsilon.$$

Rewriting the inequality:

$$q^n \leq \frac{\epsilon}{\|u - u^*\|}.$$

Taking the natural logarithm on both sides:

$$\ln q^n \leq \ln\left(\frac{\epsilon}{\|u - u^*\|}\right).$$

Simplifying:

$$n \ln q \leq \ln \epsilon - \ln \|u - u^*\|.$$

Since $\ln q < 0$ (because $0 \leq q < 1$), we multiply both sides by -1 (which reverses the inequality direction):

$$-n \ln q \geq \ln \|u - u^*\| - \ln \epsilon.$$

Recognizing that $-\ln q = \ln\left(\frac{1}{q}\right)$, we have:

$$n \ln\left(\frac{1}{q}\right) \geq \ln\left(\frac{\|u - u^*\|}{\epsilon}\right).$$

Solving for n, we obtain:

$$n \geq \frac{\ln\left(\frac{\|u-u^*\|}{\epsilon}\right)}{\ln\left(\frac{1}{q}\right)}.$$

Generally, this inequality provides a lower bound on the number of iterations n required to achieve an approximation error less than or equal to ϵ. The bound depends logarithmically on the ratio $\frac{\|u-u^*\|}{\epsilon}$ and inversely on $\ln\left(\frac{1}{q}\right)$. A smaller contraction constant q (i.e., closer to zero) results in a larger denominator, thus reducing the required number of iterations n. □

This shows that a smaller contraction constant q leads to fewer iterations needed for convergence.

3.2. **Integrate Multi-Scale Representations.** Combining global (Fourier) and local (wavelet) representations allows the neural operator to capture features at multiple scales, enhancing its ability to approximate complex functions with varying spatial frequencies.

Employing both Fourier and wavelet transforms enables efficient representation of functions with features spanning various spatial frequencies [6]. This multi-scale approach aligns with the clustering behavior in function space and enhances the operator's capacity to approximate complex solution mappings.

We formalize this with the following theorem:

Theorem 3.5 (Approximation Using Combined Bases)**.** *Any function $f \in L^2(D)$ can be approximated arbitrarily well using a finite combination of Fourier and wavelet basis functions.*

Proof. See Appendix A.2. □

Implementing spectral convolution layers utilizing the Fast Fourier Transform (FFT) for global feature extraction [3], and incorporating wavelet transform layers to capture local irregularities and singularities [7], allows for efficient computation.

Lemma 3.6 (Efficient Computation with Multi-Scale Layers)**.** *The integration of Fourier and wavelet layers allows for efficient computation by leveraging the FFT and Discrete Wavelet Transform (DWT), both of which have computational complexity $O(N \log N)$.*

Proof. See Appendix A.2. □

Integrating multi-scale representations enhances the neural operator's ability to model functions with sharp transitions or localized features, leading to improved approximation accuracy.

Theorem 3.7 (Improved Approximation Error with Multi-Scale Representations)**.** *Let $f \in L^2(D)$ be a function with both smooth and localized features. A neural operator employing multi-scale representations can approximate f with an error ϵ that decreases exponentially with the number of basis functions used.*

Proof. See Appendix A.2. □

This demonstrates that multi-scale representations can achieve lower approximation errors more efficiently than single-scale methods.

3.3. **Ensure Universal Approximation Capability.** Increasing the network capacity appropriately ensures sufficient depth and width for approximating complex operators. The Universal Approximation Theorem for operators indicates that a neural operator with sufficient capacity can approximate any continuous operator to arbitrary precision on compact subsets of the input space [8, 4].

By increasing the depth and width of the neural network, we enhance its capacity to approximate complex functions. Specifically, we have:

Theorem 3.8 (Capacity Growth with Network Size)**.** *The expressive capacity of a neural network grows exponentially with depth and polynomially with width* [9].

Proof. See Appendix A.3. □

Using activation functions capable of representing complex mappings, such as ReLU or Tanh, facilitates universal approximation [10].

Enhancing the network's capacity allows the neural operator to approximate more complex solution mappings with higher precision. However, increasing capacity improves approximation accuracy but also increases the risk of overfitting. Regularization techniques must be employed to mitigate this risk.

Lemma 3.9 (Trade-off Between Capacity and Overfitting)**.** *While increasing capacity improves approximation accuracy, it also increases the risk of overfitting. Regularization techniques must be employed to mitigate this risk.*

Proof. See Appendix A.3. □

Balancing network capacity with appropriate regularization leads to better performance.

3.4. **Enhance Generalization through Regularization.** Applying regularization techniques such as weight decay, dropout, or spectral normalization controls the complexity of the neural operator and prevents overfitting.

Regularization techniques constrain the effective capacity of the neural operator, mitigating overfitting and improving generalization to unseen data [11].

Implement weight decay by adding a penalty term to the loss function:

$$L_{\text{total}} = L_{\text{data}} + \lambda \sum_i \|W_i\|_F^2, \tag{3.1}$$

where L_{data} is the original loss, λ is the regularization parameter, and $\|W_i\|_F$ is the Frobenius norm of weight matrix W_i.

Theorem 3.10 (Effectiveness of Weight Decay)**.** *Weight decay reduces the effective capacity of the neural network by penalizing large weights, which helps prevent overfitting* [12].

Proof. See Appendix A.4. □

Apply dropout by randomly setting a fraction of the neurons' outputs to zero during training [13].

Lemma 3.11 (Dropout Prevents Co-adaptation)**.** *Dropout reduces overfitting by preventing neurons from co-adapting on training data, leading to more robust features.*

Proof. See Appendix A.4. □

Additionally, apply spectral normalization to limit the spectral norm of weight matrices, ensuring controlled Lipschitz constants [14].

Regularization techniques lead to reduced overfitting, enhancing the neural operator's performance on unseen data.

Theorem 3.12 (Improved Generalization with Regularization)**.** *Regularized neural operators exhibit lower generalization error bounds compared to unregularized models.*

Proof. Follows from standard results in statistical learning theory [15]. □

Controlling the Lipschitz constant via spectral normalization also contributes to stability.

3.5. **Optimize Computational Efficiency.** Implementing spectral methods and parallelization reduces computational complexity and exploits hardware capabilities. Efficient computational methods allow the neural operator to handle larger problem sizes and higher-dimensional PDEs without incurring prohibitive computational costs.

Ensure that the neural operator architecture is compatible with GPU acceleration and distributed computing frameworks. Under ideal conditions, parallel computing can achieve a speedup proportional to the number of processing units, up to the limits imposed by Amdahl's Law.

Theorem 3.13 (Speedup with Parallel Computing)**.** *Under ideal conditions, the speedup S achievable by parallel computing is:*

$$S = N,$$

where N is the number of processors, assuming perfect parallelization.

Proof. See Appendix A.5. □

However, Amdahl's Law imposes a limit due to the serial portion of the code.

Lemma 3.14 (Amdahl's Law)**.** *The maximum speedup S achievable by parallelization is:*

$$S = \frac{1}{(1-P) + \frac{P}{N}},$$

where P is the fraction of the program that can be parallelized, and N is the number of processors.

Proof. See Appendix A.5. □

Implement spectral convolution layers using FFT algorithms, which are highly optimized for parallel execution [17].

Optimizing computational efficiency enables the neural operator to scale to larger datasets, higher resolutions, and more complex PDEs.

Theorem 3.15 (Feasibility of High-Dimensional Problems)**.** *Efficient computational implementations make it feasible to apply neural operators to high-dimensional problems that were previously intractable due to computational limitations.*

Proof. See Appendix A.5. □

4. Conclusion

So far, we have translated theoretical insights into practical design recommendations for neural operators, each supported by rigorous mathematical proofs and relevant citations. By using contraction mappings, multi-scale representations, sufficient network capacity, regularization, and computational optimizations, we enhance the stability, convergence, generalization, and efficiency of neural operators.

Looking forward, future work should include exploring adaptive architectures that dynamically adjust their structure based on input complexity, incorporating probabilistic methods to quantify prediction uncertainty, and integrating neural operators with classical numerical methods to achieve enhanced performance.

Acknowledgments. We want to thank Google Research for providing support and mentorships for student Vu-Anh Le, as well as the Mathematics and Computer Science Department at Beloit College.

Appendix A: Proofs of Theorems and Lemmas

In this appendix, we provide detailed proofs of the theorems and lemmas referenced in the main text.

Appendix A.1. Proofs for Section 3.1: Design Neural Operators as Contraction Mappings

A.1.1. **Proof of Theorem 3.1 (Lipschitz Condition for Neural Networks).**

Proof. We aim to show that the overall Lipschitz constant L of the neural operator $\mathcal{G}_\theta$, composed of N layers and activation functions, satisfies:

$$L \leq \left(\prod_{i=1}^{N} L_i \right) L_\sigma^N,$$

where each layer f_i has a Lipschitz constant L_i, and the activation function σ has a Lipschitz constant L_σ.

We begin by representing the neural operator $\mathcal{G}_\theta$ as a composition of affine transformations (layers) and activation functions. Specifically, for an input u, we have:

$$\mathcal{G}_\theta(u) = f_N \circ \sigma \circ f_{N-1} \circ \sigma \circ \cdots \circ \sigma \circ f_1(u).$$

Each layer f_i is defined as an affine transformation:

$$f_i(x) = W_i x + b_i,$$

where W_i is the weight matrix and b_i is the bias vector for layer i.

An affine transformation f_i is Lipschitz continuous with Lipschitz constant $L_i = \|W_i\|$, where $\|W_i\|$ denotes the induced matrix norm (operator norm) of W_i. For any $x, y \in \mathbb{R}^n$, we have:

$$\|f_i(x) - f_i(y)\| = \|W_i x + b_i - (W_i y + b_i)\| = \|W_i(x - y)\| \leq \|W_i\| \cdot \|x - y\|.$$

This inequality shows that the Lipschitz constant of f_i is $\|W_i\|$. The operator norm $\|W_i\|$ can be explicitly calculated or bounded. For example, if W_i is a matrix, its operator norm induced by the Euclidean norm is the largest singular value of W_i.

In addition, the activation function $\sigma : \mathbb{R} \to \mathbb{R}$ is assumed to be Lipschitz continuous with Lipschitz constant L_σ. Common activation functions satisfy this property. For example, the *ReLU* activation function ($\sigma(x) = \max(0, x)$) has $L_\sigma = 1$; the *Sigmoid* function ($\sigma(x) = \dfrac{1}{1 + e^{-x}}$) has $L_\sigma = \dfrac{1}{4}$; and the *Tanh* function ($\sigma(x) = \tanh(x)$) has $L_\sigma = 1$.

For any $x, y \in \mathbb{R}$, we have:

$$|\sigma(x) - \sigma(y)| \leq L_\sigma |x - y|.$$

When extending to vector inputs, since activation functions are applied element-wise, the Lipschitz constant remains the same:

$$\|\sigma(x) - \sigma(y)\| \leq L_\sigma \|x - y\|.$$

It is a fundamental property that the composition of two Lipschitz continuous functions is Lipschitz continuous, with the Lipschitz constant of the composition being at most the product of the individual Lipschitz constants. Specifically, let $f : X \to Y$ and $g : Y \to Z$ be Lipschitz continuous functions with constants L_f and L_g, respectively. Then, the composition $h = g \circ f$ is Lipschitz continuous with Lipschitz constant $L_h \leq L_g L_f$. For a proof of this property, see standard analysis texts such as Rudin [18].

We now apply this property recursively to the layers and activation functions of the neural operator.

Starting with $x_0 = u$, the output after the first layer and activation function is:

$$x_1 = \sigma(f_1(x_0)).$$

By applying the composition property, the Lipschitz constant from x_0 to x_1 is:

$$L_1^{(1)} = L_\sigma L_1.$$

Similarly, for the subsequent layers ($i = 2$ to N), we have:

$$x_i = \sigma(f_i(x_{i-1})),$$

with Lipschitz constant from x_{i-1} to x_i given by:

$$L_i^{(i)} = L_\sigma L_i.$$

The Lipschitz constant from the input u to the output x_N is then the product of the individual Lipschitz constants:

$$L_{\text{total}} = \prod_{i=1}^{N} L_i^{(i)} = \left(\prod_{i=1}^{N} L_\sigma L_i \right) = L_\sigma^N \left(\prod_{i=1}^{N} L_i \right).$$

Therefore, the overall Lipschitz constant L of the neural operator $\mathcal{G}_\theta$ satisfies:

$$L \leq L_\sigma^N \left(\prod_{i=1}^{N} L_i \right).$$

This result shows that the neural operator is Lipschitz continuous, and its Lipschitz constant depends on the product of the Lipschitz constants of the layers and activation functions.

To control the Lipschitz constant of the layers, one can apply spectral normalization [14], which scales the weight matrices so that their spectral norms are bounded. This helps in ensuring that the neural operator is a contraction mapping if desired. The choice of activation function also affects the overall Lipschitz constant. Using activation functions with smaller Lipschitz constants can aid in controlling the Lipschitz constant of the entire network.

Moreover, increasing the depth N of the network can lead to an exponential increase in the Lipschitz constant due to the term L_σ^N. Care must be taken to balance depth with the desired Lipschitz properties. For discussions on the impact of depth on Lipschitz constants, see Bartlett et al. (2017) [19].

Thus, we have shown that the neural operator $\mathcal{G}_\theta$ is Lipschitz continuous with Lipschitz constant bounded by $L \leq \left(\prod_{i=1}^N L_i\right) L_\sigma^N$. □

A.1.2. **Proof of Corollary 3.2 (Ensuring Contraction via Spectral Normalization).**

Proof. We aim to show that by constraining the spectral norm of each weight matrix W_i such that $\|W_i\| \leq q^{1/N}$ and choosing the activation function σ with Lipschitz constant $L_\sigma \leq 1$, the overall Lipschitz constant L of the neural operator $\mathcal{G}_\theta$ satisfies $L \leq q$. This ensures that $\mathcal{G}_\theta$ is a contraction mapping with contraction constant q.

From Theorem 3.1, we know that the overall Lipschitz constant of the neural operator is bounded by:

$$L \leq L_\sigma^N \left(\prod_{i=1}^N L_i \right),$$

where $L_i = \|W_i\|$ is the Lipschitz constant of layer i. By constraining the spectral norm of each weight matrix to $\|W_i\| \leq q^{1/N}$, it follows that:

$$L_i \leq q^{1/N}.$$

Substituting this into the expression for L, we obtain:

$$L \leq L_\sigma^N \left(\prod_{i=1}^N q^{1/N} \right) = L_\sigma^N \left(q^{1/N} \right)^N = L_\sigma^N q.$$

Since we have chosen $L_\sigma \leq 1$, it follows that $L_\sigma^N \leq 1$. Therefore:

$$L \leq q.$$

This result shows that the neural operator $\mathcal{G}_\theta$ has a Lipschitz constant bounded by q, ensuring it is a contraction mapping.

By ensuring the spectral norms of the weight matrices are appropriately bounded, we control the Lipschitz constants of the layers. Spectral normalization [14] is a technique that rescales the weight matrices to have a desired spectral norm, effectively controlling the Lipschitz constant of each layer. This is crucial for ensuring the overall network satisfies the contraction condition.

Choosing an activation function with Lipschitz constant $L_\sigma \leq 1$ is also essential. Common activation functions like ReLU ($L_\sigma = 1$) and Tanh ($L_\sigma = 1$) satisfy this condition. Functions like the Sigmoid have $L_\sigma = \frac{1}{4}$, which also meets the requirement.

Ensuring that the neural operator is a contraction mapping allows us to invoke the Banach Fixed Point Theorem [5], guaranteeing the existence and uniqueness of fixed points

and the convergence of iterative processes. This is particularly important in the context of solving equations and iterative methods within neural networks.

Thus, by constraining the spectral norms of the weight matrices and choosing suitable activation functions, we ensure that $\mathcal{G}_\theta$ is a contraction mapping with contraction constant q.

□

Appendix A.2. Proofs for Section 3.2: Integrate Multi-Scale Representations

A.2.1. **Proof of Theorem 3.5 (Approximation Using Combined Bases).**

Proof. We aim to demonstrate that any function $f \in L^2(D)$ can be approximated arbitrarily well using a finite combination of Fourier and wavelet basis functions.

Firstly, recall that the set of complex exponentials $\{e^{ikx}\}_{k\in\mathbb{Z}}$ forms an orthonormal basis for L^2 functions defined on a compact domain with periodic boundary conditions. This is the foundation of Fourier series, which effectively capture the global behavior of functions [20].

Wavelet bases, constructed from dilations and translations of a mother wavelet $\psi(x)$, provide an orthonormal basis for $L^2(\mathbb{R})$ and are adept at representing local features due to their time-frequency localization [6]. They allow for multiresolution analysis, capturing both coarse and fine details of a function.

By combining these bases, we leverage the strengths of both global and local representations. Specifically, for a function $f \in L^2(D)$, we can express it as:

$$f(x) = \sum_{k\in\mathbb{Z}} c_k e^{ikx} + \sum_{j\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} d_{j,m}\psi_{j,m}(x),$$

where $\psi_{j,m}(x) = 2^{j/2}\psi(2^j x - m)$ are the wavelet functions at scale j and position m, and c_k, $d_{j,m}$ are the Fourier and wavelet coefficients, respectively.

In practice, we approximate $f(x)$ using finite sums:

$$f_N(x) = \sum_{k=-K}^{K} c_k e^{ikx} + \sum_{j=J_0}^{J} \sum_{m=0}^{M_j} d_{j,m}\psi_{j,m}(x),$$

where K, J, and M_j are finite truncation limits.

The approximation error is given by:

$$\|f - f_N\|_{L^2(D)} = \left\| \sum_{|k|>K} c_k e^{ikx} + \sum_{j>J} \sum_{m} d_{j,m}\psi_{j,m}(x) \right\|_{L^2(D)}.$$

As both the Fourier series and wavelet series converge in $L^2(D)$, increasing K and J allows the approximation error to be made arbitrarily small.

For functions with smooth global behavior and localized irregularities, the Fourier basis efficiently captures the global smooth components, while the wavelet basis captures local features and discontinuities [21]. This combined approach often leads to faster convergence and better approximation with fewer terms than using either basis alone.

Therefore, any function $f \in L^2(D)$ can be approximated arbitrarily well by a finite combination of Fourier and wavelet basis functions, as the sum of two complete bases is still complete in $L^2(D)$.

□

A.2.2. **Proof of Lemma 3.6 (Efficient Computation with Multi-Scale Layers).**

Proof. We aim to show that integrating Fourier and wavelet transforms into neural network layers allows for efficient computation with computational complexity $O(N \log N)$, where N is the number of data points.

Consider a discrete signal $x = [x_0, x_1, \ldots, x_{N-1}] \in \mathbb{R}^N$. The Discrete Fourier Transform (DFT) of x is defined as:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i2\pi kn/N}, \quad k = 0, 1, \ldots, N-1.$$

Computing the DFT directly requires $O(N^2)$ operations due to the nested summations.

The Fast Fourier Transform (FFT) algorithm reduces this complexity to $O(N \log N)$ by recursively decomposing the DFT into smaller DFTs and exploiting symmetries in the complex exponentials [17].

In neural networks, convolution operations are essential. The convolution of two discrete signals x and h is defined as:

$$(y)_n = (x * h)_n = \sum_{m=0}^{N-1} x_m h_{(n-m) \bmod N}.$$

Computing this convolution directly has a complexity of $O(N^2)$.

However, the Convolution Theorem states that convolution in the time domain corresponds to pointwise multiplication in the frequency domain:

$$\mathcal{F}\{x * h\} = \mathcal{F}\{x\} \cdot \mathcal{F}\{h\},$$

where $\mathcal{F}\{\cdot\}$ denotes the Fourier Transform, and $\cdot$ represents element-wise multiplication.

Therefore, we can compute the convolution efficiently by:

(1) Computing $\mathcal{F}\{x\}$ and $\mathcal{F}\{h\}$ using the FFT, each requiring $O(N \log N)$ operations.
(2) Performing element-wise multiplication: $Y_k = X_k \cdot H_k$, which requires $O(N)$ operations.
(3) Computing the inverse FFT of Y_k to obtain y_n, requiring $O(N \log N)$ operations.

The total computational complexity is $O(N \log N)$.

In neural networks, spectral convolution layers utilize this approach to perform convolution operations efficiently [3]. By transforming inputs and filters to the frequency domain, convolutions become element-wise multiplications, significantly reducing computational cost.

The Discrete Wavelet Transform (DWT) also provides a time-frequency representation of a signal, capturing both location and scale information. For a signal x, the DWT decomposes it into approximation coefficients $a_j[n]$ and detail coefficients $d_j[n]$ at different scales j.

At each level j, the approximation coefficients are computed by convolution with a scaling filter (low-pass filter) $h[n]$, followed by downsampling:

$$a_j[n] = \sum_k a_{j-1}[k]\, h[2n - k].$$

The detail coefficients are computed using a wavelet filter (high-pass filter) $g[n]$:

$$d_j[n] = \sum_k a_{j-1}[k]\, g[2n - k].$$

Here, $a_{j-1}[k]$ are the approximation coefficients from the previous level, and the downsampling by 2 reduces the number of samples by half at each level.

The overall computational complexity for computing all levels of the DWT is $O(N)$, as the amount of computation halves at each subsequent level [22].

In neural networks, wavelet transform layers can be integrated to capture features at multiple scales efficiently. By applying the DWT within the network, we can extract localized features with reduced computational cost.

By integrating both the FFT and DWT into neural network layers, we achieve efficient computation for both global and local feature extraction.

- **FFT-based Convolution:** Allows for efficient computation of convolutional layers with complexity $O(N \log N)$.
- **DWT-based Feature Extraction:** Provides multiresolution analysis with complexity $O(N)$.

When combined, the overall computational complexity remains $O(N \log N)$, dominated by the FFT operations.

In general, integrating these efficient algorithms enables neural operators to handle high-dimensional inputs and large datasets without prohibitive computational costs. This is essential for practical applications involving partial differential equations and other complex systems where computational efficiency is critical.

Therefore, by utilizing the computational efficiencies of the FFT and DWT within neural network architectures, we can perform the necessary operations in neural operators with $O(N \log N)$ complexity or better, enabling scalable and efficient computation.

□

A.2.3. **Proof of Theorem 3.7 (Improved Approximation Error with Multi-Scale Representations).**

Proof. We aim to demonstrate that for a function $f \in L^2(D)$ with both smooth and localized features, a neural operator employing multi-scale representations can approximate f with an error ϵ that decreases exponentially with the number of basis functions used.

Consider approximating f using a finite combination of Fourier and wavelet basis functions:

$$f_N(x) = \sum_{k=-K}^{K} c_k e^{ikx} + \sum_{j=J_0}^{J} \sum_{m=0}^{M_j} d_{j,m} \psi_{j,m}(x),$$

where:

- c_k are the Fourier coefficients given by $c_k = \frac{1}{2\pi} \int_D f(x) e^{-ikx}\, dx$.
- $\psi_{j,m}(x)$ are wavelet basis functions at scale j and translation m.
- $d_{j,m}$ are the wavelet coefficients given by $d_{j,m} = \int_D f(x) \psi_{j,m}(x)\, dx$.

The approximation error in the L^2 norm is given as:

$$\epsilon^2 = \|f - f_N\|^2_{L^2(D)} = \int_D |f(x) - f_N(x)|^2\, dx.$$

Expanding this, we have:

$$\epsilon^2 = \left\| \sum_{|k|>K} c_k e^{ikx} + \sum_{j>J} \sum_{m} d_{j,m} \psi_{j,m}(x) \right\|^2_{L^2(D)}.$$

By Parseval's identity, the squared L^2 norm of a function equals the sum of the squares of its coefficients:

$$\epsilon^2 = \sum_{|k|>K} |c_k|^2 + \sum_{j>J} \sum_{m} |d_{j,m}|^2.$$

Now, the decay of the Fourier coefficients $|c_k|$ is directly related to the smoothness of f. If f is s times continuously differentiable over D, then by standard results in Fourier analysis [23]:

$$|c_k| \leq \frac{C}{|k|^s},$$

for some constant $C > 0$. This implies that the tail of the Fourier series (coefficients with $|k| > K$) decreases rapidly with K, and the error from truncating the Fourier series decreases as:

$$\sum_{|k|>K} |c_k|^2 \leq C' K^{-(2s-1)},$$

where C' is another constant depending on f.

Similarly, the decay of wavelet coefficients $|d_{j,m}|$ depends on the regularity of f. For functions in the Besov space $B^s_{p,q}$, wavelet coefficients satisfy [6, 24]:

$$|d_{j,m}| \leq C 2^{-j(s+\frac{1}{2}-\frac{1}{p})},$$

where s is the smoothness parameter, and p, q relate to the integrability and summability of the coefficients.

The sum of the squared wavelet coefficients for scales $j > J$ is then bounded by:

$$\sum_{j>J} \sum_{m} |d_{j,m}|^2 \leq C'' 2^{-2J(s-\frac{1}{2})},$$

with C'' depending on f and the wavelet basis.

Combining these decay estimates, the total approximation error is bounded by:

$$\epsilon^2 \leq C' K^{-(2s-1)} + C'' 2^{-2J(s-\frac{1}{2})}.$$

By selecting K and J such that:

$$K = K_0 N^{\alpha}, \quad 2^J = J_0 N^{\beta},$$

for some $\alpha, \beta > 0$, and constants K_0, J_0, we can make ϵ decrease exponentially with N, the total number of basis functions used.

To optimize the approximation, we balance the contributions of the Fourier and wavelet terms. For functions that are smooth overall but have localized irregularities, the Fourier coefficients decay rapidly except near discontinuities, where wavelet coefficients capture the localized features efficiently.

By choosing α and β appropriately, we ensure that both terms in the error bound decrease at similar rates, minimizing the total error. This balancing act leverages the strengths of both bases.

In the context of neural operators, incorporating multi-scale representations allows the network to approximate functions with both global smoothness and local irregularities effectively. The neural network learns to represent f using a combination of global (Fourier) and local (wavelet) features.

The exponential decay in approximation error implies that the number of neurons (or parameters) required to achieve a desired accuracy ϵ grows only logarithmically with $1/\epsilon$. This is a significant improvement over methods that do not exploit multi-scale structures.

Conclusion

Therefore, the multi-scale representation enhances the approximation capabilities of the neural operator, achieving an approximation error ϵ that decreases exponentially with the number of basis functions used. This approach aligns with the principles of sparse representation and compressed sensing [25], where functions are represented using a small number of significant coefficients.

□

Appendix A.3. Proofs for Section 3.3: Ensure Universal Approximation Capability

A.3.1. **Proof of Theorem 3.8 (Capacity Growth with Network Size).**

Proof. We aim to demonstrate that for a feedforward neural network using ReLU activation functions, the number of linear regions represented by the network grows exponentially with the depth of the network and polynomially with its width.

A ReLU activation function $\sigma(x) = \max(0, x)$ introduces piecewise linearity into the network. Each neuron with a ReLU activation divides its input space into two regions: one where the neuron is active ($x > 0$) and one where it is inactive ($x \leq 0$). The combination of these regions across all neurons leads to a partitioning of the input space into linear regions, within which the neural network behaves as a linear function.

Consider a feedforward ReLU network with L layers. Let n_l denote the number of neurons in layer l, for $l = 1, 2, \ldots, L$. The input dimension is n_0. The total number of neurons is $N = \sum_{l=1}^{L} n_l$.

Montúfar et al. [26] have shown that the maximal number of linear regions R that such a network can represent satisfies:

$$R \geq \prod_{l=1}^{L} \left(\frac{n_l}{n_l - n_{l-1}} \right)^{n_{l-1}}.$$

When all layers have the same width n (i.e., $n_l = n$ for all l) and $n \geq n_0$, this simplifies to:

$$R \geq \left(\frac{n}{n - n_0} \right)^{n_0} \left(\frac{n}{n - n} \right)^{(L-1)n}.$$

Since $n - n = 0$, the expression becomes undefined. To address this, we consider the more accurate lower bound provided by Serra et al. [27], which refines the estimate of linear regions:

$$R \geq 2^{\sum_{l=1}^{L} n_l}.$$

This indicates that the number of linear regions grows exponentially with the total number of neurons in the network.

Alternatively, Montúfar et al. [26] provide a simpler lower bound for fully connected networks with ReLU activations:

$$R \geq \left(\frac{n}{n_0} \right)^{n_0} n^{(L-1)n_0}.$$

This expression shows that R grows exponentially with the depth L and polynomially with the width n.

Let's take an example calculation. For a network where $n = n_0$ (constant width equal to input dimension), the lower bound simplifies to:

$$R \geq n^{(L-1)n_0}.$$

Since $n = n_0$, we have:

$$R \geq n_0^{(L-1)n_0} = \left(n_0^{n_0} \right)^{L-1}.$$

This clearly demonstrates exponential growth with respect to the depth L.

With regard to the implications for expressive capacity, The exponential growth of the number of linear regions with depth implies that deeper networks can represent more complex functions by partitioning the input space into a greater number of linear regions. Each region corresponds to a different linear function, and the network's overall function is piecewise linear.

Raghu et al. [9] analyzed the trajectory length through the network as a measure of expressivity and found that depth contributes exponentially to expressivity measures, while width contributes polynomially.

Therefore, we conclude that the expressive capacity of ReLU neural networks grows exponentially with the network's depth and polynomially with its width, as evidenced by the number of linear regions they can represent. This result supports the assertion that deeper networks have greater expressive power.

□

A.3.2. **Proof of Lemma 3.9 (Trade-off Between Capacity and Overfitting).**

Proof. We aim to demonstrate that increasing the capacity of a neural network can lead to overfitting, highlighting the trade-off between model complexity and generalization ability.

Let $\mathcal{H}$ denote the hypothesis space of functions that the neural network can represent. Increasing the network's capacity expands $\mathcal{H}$, allowing the model to approximate more complex functions. Specifically, a higher-capacity network can achieve a smaller empirical risk (training error) R_{emp} by fitting the training data more precisely.

However, the true risk (generalization error) R depends on how well the model performs on unseen data. According to the bias-variance decomposition [28], the expected generalization error can be expressed as:

$$\mathbb{E}_{\mathcal{D}}[R] = \text{Bias}^2 + \text{Variance} + \sigma^2,$$

where:

- Bias measures the error due to simplifying assumptions made by the model;
- Variance measures the sensitivity of the model to fluctuations in the training set;
- σ^2 is the irreducible error inherent in the data.

As the capacity of the network increases, the bias tends to decrease because the model can fit the training data more closely. However, the variance tends to increase because the model becomes more sensitive to small fluctuations or noise in the training data. This increased variance can lead to overfitting, where the model captures noise and irrelevant patterns, resulting in a decrease in generalization performance.

Overfitting is characterized by a situation where:

$$R_{\text{emp}} \downarrow, \quad R \uparrow,$$

meaning that while the training error decreases, the validation or test error increases.

To prevent overfitting, regularization techniques are employed to constrain the complexity of the hypothesis space $\mathcal{H}$. Regularization can be introduced by adding a penalty term $\Omega(\theta)$ to the loss function $L(\theta)$, leading to the regularized loss:

$$L_{\text{reg}}(\theta) = L(\theta) + \lambda\Omega(\theta),$$

where θ represents the network parameters, and $\lambda > 0$ controls the strength of the regularization.

Common regularization methods include:

(1) **Weight Decay (L2 Regularization)**: Penalizes large weights by setting $\Omega(\theta) = \frac{1}{2}\|\theta\|_2^2$.
(2) **L1 Regularization**: Encourages sparsity by setting $\Omega(\theta) = \|\theta\|_1$.
(3) **Dropout**: Randomly sets a fraction of activations to zero during training to prevent co-adaptation [13].

By constraining $\mathcal{H}$, regularization reduces variance at the expense of a slight increase in bias, ultimately improving the generalization error R.

Therefore, there exists a trade-off between model capacity and overfitting: increasing capacity enhances the ability to fit complex functions but may lead to overfitting if not properly regularized. Effective regularization techniques are essential to balance this trade-off and achieve optimal generalization performance [11].

□

Appendix A.4. Proofs for Section 3.4: Enhance Generalization through Regularization

A.4.1. **Proof of Theorem 3.10 (Effectiveness of Weight Decay).**

Proof. We aim to show that weight decay (L2 regularization) effectively reduces overfitting by penalizing large weights, thereby constraining the model complexity and improving generalization.

Consider a neural network with parameters (weights) θ. The standard loss function $L(\theta)$ measures the discrepancy between the network's predictions and the training data. Weight decay modifies the loss function by adding a regularization term:

$$L_{\text{reg}}(\theta) = L(\theta) + \lambda \frac{1}{2}\|\theta\|_2^2,$$

where $\|\theta\|_2^2 = \sum_i \theta_i^2$ is the squared L2 norm of the weights, and $\lambda > 0$ is the regularization coefficient.

The gradient of the regularized loss with respect to the weights is:

$$\nabla_\theta L_{\text{reg}}(\theta) = \nabla_\theta L(\theta) + \lambda\theta.$$

During training with gradient descent, the weight update rule becomes:

$$\theta^{(t+1)} = \theta^{(t)} - \eta\left(\nabla_\theta L(\theta^{(t)}) + \lambda\theta^{(t)}\right),$$

where η is the learning rate.

The term $\lambda\theta^{(t)}$ acts as a force that drives the weights toward zero. This discourages the model from assigning excessive importance to any particular feature, effectively reducing the complexity of the model.

By penalizing large weights, weight decay reduces the variance component of the generalization error. According to the bias-variance decomposition, the expected generalization error can be written as:

$$\mathbb{E}_{\mathcal{D}}[R] = \text{Bias}^2 + \text{Variance} + \sigma^2.$$

Weight decay increases the bias slightly due to the added constraint but decreases the variance more significantly, leading to a net reduction in generalization error.

Moreover, in linear models, weight decay corresponds to Ridge Regression [29], where the regularization term stabilizes the inversion of ill-conditioned matrices, leading to more robust solutions.

Therefore, weight decay effectively prevents overfitting by constraining the magnitude of the weights, promoting simpler models that generalize better to unseen data [12].

□

A.4.2. **Proof of Lemma 3.11 (Dropout Prevents Co-adaptation).**

Proof. We aim to demonstrate that dropout regularization reduces overfitting by preventing co-adaptation of neurons and encouraging the network to learn robust feature representations.

During training, dropout randomly deactivates a fraction p of the neurons in each layer. For a neuron with activation h_i, the modified activation $\tilde{h}_i$ during training is:

$$\tilde{h}_i = h_i \cdot \zeta_i,$$

where ζ_i is a Bernoulli random variable:

$$\zeta_i = \begin{cases} 1, & \text{with probability } q = 1 - p, \\ 0, & \text{with probability } p. \end{cases}$$

This random deactivation forces the network to learn redundancies because any neuron could be dropped out at any time. As a result, neurons cannot rely on specific other neurons being present and must learn features that are useful in conjunction with many different subsets of other neurons.

By preventing co-adaptation, where neurons adjust to rely on outputs from specific other neurons, dropout reduces the risk of overfitting. The network becomes less sensitive to the noise and variations in the training data, improving generalization to unseen data [13].

At test time, to compensate for the dropped activations during training, the weights are scaled by a factor of q (or equivalently, activations are multiplied by q):

$$h_i^{\text{test}} = q h_i.$$

This ensures that the expected output of each neuron remains the same between training and testing:

$$\mathbb{E}[\tilde{h}_i] = q h_i.$$

Therefore, dropout effectively prevents co-adaptation by encouraging neurons to learn individually useful features, reducing overfitting and enhancing the robustness of the network's predictions.

□

Appendix A.5. Proofs for Section 3.5: Optimize Computational Efficiency

A.5.1. **Proof of Theorem 3.13 (Speedup with Parallel Computing).**

Proof. We aim to show that parallel computing can provide a speedup in computation time for parallelizable tasks but that the overall speedup is limited by the serial portion of the computation, as described by Amdahl's Law.

Let:

- T_1 be the total execution time on a single processor;
- T_N be the total execution time using N processors;
- P be the fraction of the program that can be parallelized ($0 \leq P \leq 1$);
- S be the speedup achieved: $S = \dfrac{T_1}{T_N}$.

The single-processor execution time is:

$$T_1 = T_{\text{serial}} + T_{\text{parallel}},$$

where T_{serial} and T_{parallel} are the times spent on serial and parallelizable portions, respectively.

When using N processors, the parallel portion ideally scales perfectly, so its execution time becomes T_{parallel}/N. The total execution time on N processors is:

$$T_N = T_{\text{serial}} + \frac{T_{\text{parallel}}}{N}.$$

Substituting $T_{\text{serial}} = (1 - P)T_1$ and $T_{\text{parallel}} = PT_1$, we have:

$$T_N = (1 - P)T_1 + \frac{PT_1}{N}.$$

Therefore, the speedup S is:

$$S = \frac{T_1}{T_N} = \frac{T_1}{(1 - P)T_1 + \frac{PT_1}{N}} = \frac{1}{(1 - P) + \frac{P}{N}}.$$

As $N \to \infty$, the speedup approaches its theoretical maximum:

$$S_{\max} = \lim_{N\to\infty} S = \frac{1}{1 - P}.$$

This demonstrates that the speedup is limited by the serial portion of the code. Even with an infinite number of processors, the execution time cannot be reduced below $(1 - P)T_1$.

Thus, while parallel computing significantly reduces computation time for parallelizable tasks, Amdahl's Law shows that the overall speedup is constrained by the fraction of the code that must be executed serially [30].

□

A.5.2. **Proof of Lemma 3.14 (Amdahl's Law).**

Proof. We aim to derive Amdahl's Law, which quantifies the theoretical speedup in latency of the execution of a task when a portion of it is parallelized.

Let:

- T_1 be the execution time on a single processor;
- T_N be the execution time on N processors;
- P be the fraction of the execution time that is parallelizable.

The execution time on N processors is:

$$T_N = T_{\text{serial}} + T'_{\text{parallel}},$$

where:

$$T_{\text{serial}} = (1 - P)T_1,$$
$$T'_{\text{parallel}} = \frac{PT_1}{N}.$$

Therefore:

$$T_N = (1 - P)T_1 + \frac{PT_1}{N}.$$

The speedup S is given by:

$$S = \frac{T_1}{T_N} = \frac{T_1}{(1 - P)T_1 + \frac{PT_1}{N}} = \frac{1}{(1 - P) + \frac{P}{N}}.$$

This equation represents Amdahl's Law, showing how the speedup S depends on the number of processors N and the parallelizable fraction P. It illustrates that as N increases, the speedup asymptotically approaches $1/(1-P)$, emphasizing the diminishing returns due to the serial portion of the computation.

Therefore, Amdahl's Law captures the fundamental limitation of parallel computing: the speedup is constrained by the serial fraction of the task, regardless of the number of processors [30].

□

A.5.3. **Proof of Theorem 3.15 (Feasibility of High-Dimensional Problems).**

Proof. We aim to demonstrate that optimizing computational efficiency through spectral methods and parallel computing enables neural operators to effectively handle high-dimensional problems.

Let $\Omega \subset \mathbb{R}^d$ be a d-dimensional domain, and let $u : \Omega \to \mathbb{R}$ be a function of interest. Traditional numerical methods for solving partial differential equations (PDEs), such as finite difference or finite element methods, require discretizing each dimension into n points. This results in a total of $N = n^d$ grid points. Operations like matrix-vector multiplication or convolution over this grid have computational complexities that scale at least linearly with N, and often worse, leading to $O(N^2)$ operations for certain tasks.

The exponential growth of N with respect to the dimension d is known as the "curse of dimensionality." It renders computations infeasible for large d using traditional methods.

Spectral methods, such as the Fourier Transform, provide an alternative by transforming differential operators into algebraic ones in the frequency domain. The d-dimensional Discrete Fourier Transform (DFT) of a function u sampled on a regular grid is defined as:

$$\hat{u}_{\mathbf{k}} = \sum_{\mathbf{n} \in \mathbb{Z}_n^d} u_{\mathbf{n}} e^{-i2\pi \frac{\mathbf{k}\cdot\mathbf{n}}{n}}, \quad \mathbf{k} \in \mathbb{Z}_n^d,$$

where $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$, and $\mathbf{n}$ and $\mathbf{k}$ are d-dimensional index vectors.

Computing the DFT directly requires $O(N^2)$ operations. However, the Fast Fourier Transform (FFT) algorithm reduces this to $O(N \log N)$ by exploiting symmetries and redundancies in the computation [17].

In neural operators, convolution operations are essential. Consider the convolution of two functions $u, v : \Omega \to \mathbb{R}$:

$$(w)(\mathbf{x}) = (u * v)(\mathbf{x}) = \int_\Omega u(\mathbf{y}) v(\mathbf{x} - \mathbf{y})\, d\mathbf{y}.$$

Computing this convolution directly requires $O(N^2)$ operations due to the nested summations over all grid points.

Applying the Convolution Theorem, the Fourier Transform converts convolution into pointwise multiplication:

$$\hat{w}_{\mathbf{k}} = \hat{u}_{\mathbf{k}} \cdot \hat{v}_{\mathbf{k}}.$$

This reduces the convolution computation to:

(1) Compute $\hat{u}_{\mathbf{k}}$ and $\hat{v}_{\mathbf{k}}$ using the FFT: $O(N \log N)$ operations each;
(2) Perform pointwise multiplication: $O(N)$ operations;
(3) Compute the inverse FFT to obtain $w(\mathbf{x})$: $O(N \log N)$ operations.

The total computational complexity becomes $O(N \log N)$, a significant reduction from $O(N^2)$.

In high-dimensional problems, many functions of interest exhibit sparsity or low-rank structures in the spectral domain. If $\hat{u}_{\mathbf{k}}$ is sparse, meaning that significant energy is concentrated in a subset $\mathcal{K} \subset \mathbb{Z}_n^d$ with $|\mathcal{K}| = s \ll N$, we can approximate u using:

$$u(\mathbf{x}) \approx \sum_{\mathbf{k} \in \mathcal{K}} \hat{u}_{\mathbf{k}} e^{i2\pi \frac{\mathbf{k}\cdot\mathbf{x}}{n}}.$$

Computations then involve s significant coefficients instead of N, reducing complexity to $O(s \log N)$.

We now consider the FFT algorithm. We regard that this algorithm is highly parallelizable. In a parallel computing environment with P processors, we can divide the data equally among processors. Each processor performs FFT computations on its subset of data:

$$T_{\text{compute}} = O\left(\frac{N \log(N/P)}{P}\right).$$

Communication between processors is required to combine results, but for large N, the computation time dominates, and communication overhead can be minimized with efficient algorithms and network architectures [31].

Assuming ideal parallel efficiency, the total computational complexity per processor is reduced to approximately $O\left(\frac{N \log N}{P}\right)$.

We now consider the behaviors of **neural Operators in High Dimensions**. Neural operators, such as the Fourier Neural Operator [3], leverage spectral convolutions to learn mappings between function spaces. By representing integral kernel operations in the frequency domain, neural operators can efficiently handle high-dimensional inputs.

Consider a neural operator layer defined as:

$$(u_{\text{next}})(\mathbf{x}) = \sigma\left(Wu(\mathbf{x}) + \mathcal{F}^{-1}\left(R \cdot \mathcal{F}[u]\right)(\mathbf{x})\right),$$

where:

- W is a linear transformation;
- σ is a nonlinear activation function;
- $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier and inverse Fourier transforms, respectively;
- R is a learned filter in the frequency domain.

Computing this layer involves FFTs and pointwise operations, all of which have computational complexities that scale as $O(N \log N)$ and are amenable to parallelization.

By combining spectral methods that reduce per-processor computational complexity to $O(N \log N)$ and parallel computing that reduces wall-clock time by distributing computations across P processors, the overall computational effort becomes manageable even in high-dimensional settings.

Therefore, optimizing computational efficiency through spectral methods and parallel computing enables neural operators to handle high-dimensional problems effectively, mitigating the curse of dimensionality and making practical solutions feasible for complex, real-world applications.

□

REFERENCES

[1] V.-A. Le and M. Dik, "A mathematical analysis of neural operator behaviors," arXiv preprint arXiv:2410.21481, 2024.

[2] N. Kovachki, Z. Li, B. Liu, K. Azizzadenesheli, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Neural operator: Learning maps between function spaces," SIAM J. Sci. Comput. **43**5 (2021), A3172–A3192.

[3] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Fourier neural operator for parametric partial differential equations," in Proceedings of the International Conference on Learning Representations (ICLR), 2021. Available at https://openreview.net/forum?id=c8P9NQVtmnO.

[4] L. Lu, P. Jin, and G. E. Karniadakis, "DeepONet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators," arXiv preprint arXiv:1910.03193, 2019.

[5] S. Banach, "Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales," Fund. Math. **3** (1922), 133–181.

[6] I. Daubechies, "Ten Lectures on Wavelets," SIAM, Philadelphia, PA, 1992.

[7] Y. Zhao and S. Sun, "Wavelet neural operator: A neural operator based on the wavelet transform," arXiv preprint arXiv:2201.12086, 2022.

[8] T.-S. Chen and H.-Y. Chen, "Universal approximation to nonlinear operators by neural networks with arbitrary activation functions and its application to dynamical systems," IEEE Trans. Neural Networks, **6**(4) (1995), 911–917.

[9] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein, "On the expressive power of deep neural networks," in Proceedings of the 34th International Conference on Machine Learning, **70**, PMLR, 2017, pp. 2847–2854.

[10] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," Neural Networks **2**(5) (1989), 359–366.

[11] I. Goodfellow, Y. Bengio, and A. Courville, "Deep Learning," MIT Press, 2016. Available at http://www.deeplearningbook.org.

[12] A. Krogh and J. A. Hertz, "A simple weight decay can improve generalization," in Advances in Neural Information Processing Systems 4, Morgan-Kaufmann, 1992, pp. 950–957.

[13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," J. Mach. Learn. Res., **15** (2014), 1929–1958.

[14] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida, "Spectral normalization for generative adversarial networks," arXiv preprint arXiv:1802.05957, 2018.

[15] M. Mohri, A. Rostamizadeh, and A. Talwalkar, "Foundations of Machine Learning," MIT Press, 2018.

[16] N. Cohen, O. Sharir, and A. Shashua, "On the expressive power of deep learning: A tensor analysis," in Proceedings of the 29th Annual Conference on Learning Theory, **49**, PMLR, 2016, pp. 698–728.

[17] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex Fourier series," Math. Comp. **19**(90) (1965), 297–301.

[18] W. Rudin, "Principles of Mathematical Analysis," 3rd ed., McGraw-Hill, New York, 1976.

[19] P. L. Bartlett, D. J. Foster, and M. J. Telgarsky, "Spectrally-normalized margin bounds for neural networks," arXiv preprint arXiv:1706.08498, 2017.

[20] D. Jackson, "The Theory of Approximation," American Mathematical Society, Providence, RI, 1930.

[21] S. Mallat, "A Wavelet Tour of Signal Processing," 2nd ed., Academic Press, San Diego, CA, 1999.

[22] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," IEEE Trans. Pattern Anal. Mach. Intell., **11**(7) (1989), 674–693.

[23] E. C. Titchmarsh, "Introduction to the Theory of Fourier Integrals," Oxford University Press, Oxford, 1948.

[24] A. Cohen, "Numerical Analysis of Wavelet Methods," Elsevier, 2003.

[25] D. L. Donoho, "Compressed sensing," IEEE Trans. Inform. Theory, **52**(4) (2006), 1289–1306.

[26] G. F. Montúfar, R. Pascanu, K. Cho, and Y. Bengio, "On the number of linear regions of deep neural networks," in Advances in Neural Information Processing Systems 27, Curran Associates, Inc., 2014, pp. 2924–2932.

[27] T. Serra, C. Tjandraatmadja, and S. Ramalingam, "Bounding and counting linear regions of deep neural networks," in Proceedings of the 35th International Conference on Machine Learning, 2018, pp. 4558–4566.

[28] S. Geman, E. Bienenstock, and R. Doursat, "Neural networks and the bias/variance dilemma," Neural Comput., **4**(1) (1992), 1–58.

[29] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," Technometrics, **12**(1) (1970), 55–67.
[30] G. M. Amdahl, "Validity of the single processor approach to achieving large scale computing capabilities," in Proceedings of the Spring Joint Computer Conference, ACM, 1967, pp. 483–485.
[31] M. Frigo and S. G. Johnson, "FFTW: An adaptive software architecture for the FFT," in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, **3**, 1998, pp. 1381–1384.

Vu Anh Le,
Department of Mathematics and Computer Science, Beloit College, USA
Email address: csplevuanh@gmail.com

Mehmet Dik,
Department of Mathematics, Computer Science & Physics, Rockford University, USA
Email address: mdik@rockford.edu

Proceedings of International Mathematical Sciences
ISSN: 2717-6355, URL: https://dergipark.org.tr/tr/pub/pims
Volume 6 Issue 2 (2024), Pages 100-108.
Doi: https://doi.org/10.47086/pims.1567406

MODIFIED PELL MATRIX TECHNIQUE FOR SOLVING OPTIMAL CONTROL PROBLEMS

SHAYMAA HUSSAIN*, FARAH J. AL-ZAHID**, SUHA SHIHAB***, AND NADIA M.G. AL-SAIDI****
*UNIVERSITY OF TECHNOLOGY, BAGHDAD, IRAQ. ORCID: 0000-0003-4275-0400
**UNIVERSITY OF BAGHDAD, BAGHDAD, IRAQ. ORCID: 0000-0001-5547-1422
***UNIVERSITY OF TECHNOLOGY, BAGHDAD, IRAQ. ORCID: 0000-0003-3730-0149
****UNIVERSITY OF TECHNOLOGY, BAGHDAD, IRAQ. ORCID: 0000-0002-7255-5246.

Abstract. The orthogonal polynomial basis functions are used to solve different mathematical problems, especially for optimal control and many other engineering problems, which attract many researchers to work on. In this study, the modified Pell polynomials (MPPs) are presented and their new properties are investigated to be used for solution approximation of optimal control problems. Some formulas for MPPs are derived by matrices. A new exact formula expressing the derivatives of MPPs explicitly of any degree is constructed. The main advantage of the presented formulas is that the new properties of MPPs greatly simplify the original problems and the result will lead to easy calculation of the coefficients of expansion, it also increases the accuracy and reduces the computational time. A new computational method along with the MPPs is proposed to solve one of the optimal control problems. Numerical results are included to demonstrate the validity of this new technique. It shows an important improvement in error approximation when the polynomial degree is increased. The contribution in this work is based on the idea of the approximate algorithm in terms of MPPs and their new properties to treat the optimal control problem numerically with less number of terms and unknown parameters with a satisfactory accuracy.

1. Introduction

Recently, there has been increased interest among scientists and engineers to use orthogonal polynomial basis functions along with approximate solutions to solve different problems. The main advantage of such a technique is the ability to reduce a complicated problem to another simple one [1, 2, 3, 4]. The authors of [5] extended using orthogonal polynomials to solve problems in the calculus of variation numerically. They applied the generalized Vieta- Pell polynomials for numerical treatment of variation calculus problems while improved Chebyshev polynomials were applied in [6] for solving optimal control

2020 *Mathematics Subject Classification.* Primary: 49M05; 33F05.
Key words and phrases. Modified Pell polynomials; orthogonal polynomial; optimal control problem; oscillatory systems; numerical approximation.

Submitted on 15.10.2024, Accepted on 26.11.2024.
Communicated by Mehmet Dik.

problems. Furthermore; Boubaker polynomials [7, 8] and Hermite functions [9] were utilized to solve approximately optimal control problems and fractional calculus of variation respectively.

Stabilizing nonlinear chaotic and dynamical systems represents the main core for controlling such systems [10, 11, 12, 13, 14]. The modified Pell polynomial particularly can be used to perform such stabilization. The utilization of this modified polynomial by offering good computational efficiency and high accuracy follows the same scenario in solving the control problem.

The motivation in this work deals with the new application of a modified Pell polynomial for numerical solutions of an optimal control problem. The study of numerical solutions of special classes in optimal control has provided an interesting field for mathematical sciences researchers. For some work concerned with both Vieta-Pell and Vieta Pell Lucas polynomial basis functions can be found in [15, 16].

Motivated by the above presentation, we are interested in suggesting a new iterative algorithm to solve optimal control problems numerically together with modified Pell basis functions to parameterize the state variables. We aim to obtain the accuracy and efficiency simultaneously. Hence the first goal of this work is to introduce MPP with some important properties and then use it to perform the parameterization of state variables in order to solve some problems in optimal control.The work in this article is organized as follows: modified Pell polynomials definition is presented first in section 2. Their new properties are also included in section 2. The new contribution of the modified Pell polynomials is discussed in section 3 to solve a special application in optimal control problems. The conclusion is listed in section 4.

2. Definition of Modified Pell Polynomials

The following modified Pell basis polynomials are obtained
The modified Pell basis polynomials of degree n are defined by:

$$q_n(x) = \frac{n}{2}\sum_{r=0}^{\lfloor n/2\rfloor}\frac{(n-r-1)!}{(n-2r)!r!}(2x)^{n-2r} \tag{1.1}$$

with special values
$q_n(1) = q_n(-1)$ for even n and $q_n(1) = -q_n(-1)$ for odd n
Define $\Psi(x) = \begin{bmatrix} q_0(x) & q_1(x) & \dots & q_n(x) \end{bmatrix}^T$, where one can get:

$$\Psi(x) = AT_n(x). \tag{1.2}$$

The matrix A is a lower triangular matrix of order $(n+1)\times(n+1)$. The element of A can be listed as below:

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 2 & \cdots & 0 \\ 0 & 3 & 0 & 4 & 0 \\ 1 & 0 & 8 & 0 & 8 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \tag{1.3}$$

and

$$a_{ij} = \begin{cases} 1, & \text{if } j = 1, \\ 2a_{i-1,j-1} + a_{i-2,j}, & \text{if } j \geq i, \text{i is odd}, i, j = 2, 3, \cdots \\ 0, & j > i. \end{cases} \tag{1.4}$$

and the vector $T_n(x)$ of dimension $(n+1) \times 1$ is defined by:

$$T_n(\mathrm{x}) = \begin{bmatrix} 1 & x & x^2 & \dots & x^n \end{bmatrix}^T \tag{1.5}$$

2.1 Function Approximation. A function $u(x)$ as is square integrable in $[-1, 1]$, It can be expanded in terms of modified Pell basis as below:

$$u(x) = \sum_{i=0}^{n} a_i q_i(x) = a^T q(x) \tag{1.6}$$

where
$a^T = \begin{bmatrix} a_0 & a_1 & \dots & a_n \end{bmatrix}$, then $a = Q^{-1}\langle u(x), q(x)\rangle$, then the matrix Q of order $n \times n$ is called a dual matrix which is derived later.

2.2 Some properties of Modified Pell Sequence: Theorem 1. For $n \geq 1$, the following formulas can be obtained
(i) $\sum_{i=1}^{n} q_i(x) = \frac{1}{2x}\left[q_{n+1}(x) + q_n(x) - (q_1(x) + q_0(x))\right]$,
(ii) $\sum_{i=1}^{n} q_{2i}(x) = \frac{1}{2x}\left(q_{2n+1}(x) + q_1(x)\right)$,
(iii) $\sum_{i=1}^{n} q_{2i+1}(x) = \frac{1}{2x}\left(q_{2n}(x) + q_0(x)\right)$.

Proof. The mathematical induction is suggested to prove (i)
Take $n = 1$, then $q_1(x) = x = 1/2x\left[2x^2 + 1 + x - x - 1\right]$.
Let the result in (i) be valid for $n = k$, then

$$\sum_{i=1}^{k} q_i(x) = \frac{1}{2x}\left[q_{k+1}(x) + q_k(x) - (q_1(x) + q_0(x))\right]$$

$$\text{or } 2\sum_{i=1}^{k+1} q_i(x) = q_{k+1}(x) + q_k(x) - (q_1(x) + q_0(x)) + 2xq_{k+1}(x)$$

Apply the recursive relation related with modified Pell polynomials to obtain
$q_{n+1}(x) = 2xq_n(x) + q_{n-1}(x)$,
Therefore; $2\sum_{i=1}^{k+1} q_i(x) = q_{k+2}(x) + q_{k+1}(x) - (q_1(x) + q_0(x))$
or $\sum_{i=1}^{k+1} q_i(x) = \frac{1}{2x}\left[q_{k+2}(x) + q_{k+1}(x) - (q_1(x) + q_0(x))\right]$
This is the same result in Eq.(1.6).
Note that the identities (ii) and (iii) are a direct result of i.

2.3 Dual Operational Matrix of Modified Pell Polynomials. This section illustrates the building of a modified Pell dual operational matrix. The cross-product integration of two modified Pell basis vectors is taken as below

$$\begin{aligned} Q &= \int_{-1}^{1} \left\langle AT_n(x), (AT_n(x))^T \right\rangle dx \\ &= A\int_{-1}^{1} \left(T_n(x)T_n(x)^T\right) dx = AHA^T \end{aligned}$$

The matrix A is defined in Eq.(1.3) and the matrix H is given for the particular case with $n = 3$ as below

$$H = \begin{pmatrix} \int_{-1}^{1} q_0(x)q_0(x)dx & \int_{-1}^{1} q_0(x)q_1(x)dx & \int_{-1}^{1} q_0(x)q_2(x)dx \\ \int_{-1}^{1} q_1(x)q_0(x)dx & \int_{-1}^{1} q_1(x)q_1(x)dx & \int_{-1}^{1} q_1(x)q_2(x)dx \\ \int_{-1}^{1} q_2(x)q_0(x)dx & \int_{-1}^{1} q_2(x)q_1(x)dx & \int_{-1}^{1} q_2(x)q_2(x)dx \end{pmatrix}$$

Hence

$$H = \begin{pmatrix} 2 & 0 & \frac{10}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{10}{3} & 0 & \frac{94}{15} \end{pmatrix}$$

In general, the element of the constant matrix h_{ij} can be determined as follows

$$H_{ij} = \begin{cases} \sum_{r=0}^{\lfloor j/2 \rfloor} 2^{j-2r-1} \frac{(j-r-1)!}{(j-2r+1)!r!}, & \text{if } |i-j| \text{is even} \\ 0, & \text{otherwise.} \end{cases}$$

2.4 Derivative for Modified Pell Polynomials Matrix: consider the vector $\Psi(x)$ in Eq.(1.2) can be written in matrix form as follows
$\Psi(x) = AT(x)$
One can get
$\dot{\Psi}(x) = A\dot{T}_n(x)$
where $\dot{T}_n(x) = \begin{bmatrix} 0 & 1 & 2x & \dots & nx^{n-1} \end{bmatrix}^T$, this equation can be reformulated as:
$\dot{T}_n(x) = A_1 T_n(x)$
Then
$\dot{\Psi}(x) = AA_1 T_n(x)$,
where: $A_1 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & n \end{pmatrix} \begin{pmatrix} 1 \\ x \\ x^2 \\ \vdots \\ x^{n-1} \end{pmatrix}$.
and A is given in Eq.(1.3).

2.5 Operational Matrix of Integration for Modified Pell Polynomials. Let M be an operational matrix of integration of order $(n+1) \times (n+1)$, then

$$\int_{-1}^{x} q(t)dt = MX$$

where

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ \frac{-1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{5}{3} & 1 & 0 & \frac{2}{3} & 0 & 0 \\ \frac{-10}{4} & 0 & \frac{3}{2} & 0 & 1 & 0 \\ 79 & 1 & 0 & \frac{8}{3} & 0 & \frac{8}{5} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

and $X = \begin{bmatrix} x & x^2 & \dots & x^{n+1} \end{bmatrix}^T$.
This can be evaluated below

$$m_{ij} = i \begin{cases} \sum_{r=0}^{\lfloor i/2 \rfloor} \frac{(i-r-1)!2^{i-2r-1}}{r!(i-2r)!(i-2r+1)}, & j = 1, \text{ even } i \\ \sum_{r=0}^{ii/2\rfloor} \frac{(i-r-1)!2^{i-2r-1}}{r!(i-2r)!(i-2r+1)}, & j = 1, \text{ odd } i \\ \frac{1}{i}\left(m_{i-1,j} + 2m_{i-1,j-1}\right), & j > 1 \end{cases}$$

2.6. Initial and Final Values. Lemma 1. Let $\sigma_1 = n \sum_{r=0}^{\lfloor n/2 \rfloor} \frac{1}{n-r}\binom{n-r}{r} 2^{n-2r-1}$, Then

$$q_n(1) = \sigma_1, \text{ and } q_n(-1) = \begin{cases} \sigma_1, & \text{if } n \text{ even} \\ -\sigma_1, & \text{if } n \text{ odd} \end{cases}$$

Lemma 2. We have $q_{n+1}(-1)\,(q_n(1) - q_n(-1)) \neq 0$,
If n is even, then
$q_{2n+1}(-1)q_{2n}(1) - q_{2n+1}(1)q_{2n}(-1) = \sigma_1 \cdot (-\sigma_1) \cdot \sigma_1 = -2{\sigma_1}^2 \neq 0$.
If n is odd, then

$$q_{2(n+1)}(-1)q_{2n+1}(1) - q_{2(n+1)}(1)q_{2n+1}(-1) = \sigma_1 \cdot \sigma_1 - (-\sigma_1 \cdot \sigma_1) = 2\sigma_1^2 \neq 0$$

3. Application of Modified Pell Polynomial in optimal control

Consider the following optimal control problem which describes many important oscillating phenomena in some dynamic [12, 13, 14, 15, 16], engineering and physical systems $J = \int_{-\tau}^{0} \frac{1}{2}u^2(t)dt$, governed by the following system $\ddot{x}(t) + \sigma\dot{x}(t) + \omega^2 x(t) + \in x^3(t) = f\cos(\alpha t) - u(t), -\tau \leq t \leq 0$,
where τ is known, $\sigma \geq 0$, is the viscous damping coefficient, f and α are the amplitude ω is the stiffness parameter, and frequency of the external input, respectively.

The initial and boundary conditions can be given as follows:

$$x(-\tau) = \alpha, x(0) = 0, \dot{x}(-\tau) = \beta, \dot{x}(0) = 0$$

To obtain the optimal performance index $J(\cdot)$, the following steps are suggested:
In order to use Modified Pell polynomials, the transformation $\tau = \frac{1}{2}\tau(t-1)$ is used to obtain the following restated optimal control problem

$$J = \frac{\tau}{2}\int_{-1}^{1} \frac{1}{2}u^2(t)dt \tag{1.7}$$

governed by the following system

$$\ddot{x}(t) = \frac{1}{2}\tau^2\left[-\sigma\dot{x}(t) - \omega^2 x(t) - \in x^3(t) + f\cos(\alpha t) - u(t)\right], t \in [-1, 1] \tag{1.8}$$

The initial and boundary conditions can be given as follows:

$$x(-1) = \alpha, x(1) = 0, \dot{x}(-1) = \beta, \dot{x}(1) = 0 \tag{1.9}$$

To illustrate the present method for obtaining the optimal performance index $J(\cdot)$, the following steps are suggested:

Step 1. Assume that the approximate solution of the state variables $x(t)$ in terms of MPPs that satisfy the conditions given in Eq.(1.9) below:
$x^1(t) = a_0q_0(t) + a_1q_1(t) + a_2q_2(t) + a_3q_3(t) + a_4q_4(t)$
It is worth mentioning that the function $x^1(t)$ is chosen to satisfy the conditions in Eq.(1.9). Therefore, the following equations are obtained:
$a_0 + a_1 + 3a_2 + 7a_3 + 17a_4 = 0$

$a_0 - a_1 + 3a_2 - 7a_3 + 17a_4 = \alpha$
$a_1 + 4a_2 + 15a_3 + 48a_4 = 0$
$a_1 - 4a_2 + 15a_3 - 48a_4 = \beta$
Step 2. Eliminate the unknown a_0, a_1, a_2 and a_3 to get
$a_0 = \frac{1}{2}\left[\left(\alpha + \frac{3}{4}\beta\right) - 33.5a_4\right], a_1 = \frac{1}{2}(-2\alpha - \beta), a_2 = \frac{1}{8}(-\beta - 96a_4),$
$a_3 = \frac{1}{14}(\alpha + \beta)$

As a result $x^1(t) = \frac{1}{2}\left[\left(\alpha + \frac{3}{4}\beta\right) - 33.5a_4\right] q_0(t) + \frac{1}{2}(-2\alpha - \beta)q_1(t)$

$$+\frac{1}{8}(-\beta - 96a_4)\, q_2(t) + \frac{1}{14}(\alpha + \beta)q_3(t) + a_4 q_4(t)$$

and then obtain the first approximation to $u(t)$ using Eq.(1.8)

$$u^1(t) = f\left(x^1(t), \ddot{x}^1(t)\right)$$

Step 3. Obtain J as a function of the unknown a_4 by determining

$$J^1(a_4) = \int_{-1}^{1} F(u_1(t))\, dt$$

Step4. Minimize $J_1(a^*)$ is the solution to the problem in Eq's.(1.7-1.9).
Step 5. Calculate $x^1(t)$ and $u^1(t)$ from a^* approximately. The procedure is repeated until an acceptable accuracy is obtained.

Note that the approximate solution in the n step is given by
$\mathrm{x_n}(t) = x_{n-1}(t) + \sum_{i=n-1}^{n+1} a_i q_i(t)$,
with

$$a_{n-1} = \frac{q_{n+1}(-1)q_{n+2}(1) - q_{n+1}(1)q_{n+2}(-1)}{q_n(-1)q_{n+1}(1) - q_n(1)q_{n+1}(-1)} a_{n+1}$$
$$a_n = \frac{q_n(-1)q_{n+2}(1) - q_n(1)q_{n+2}(-1)}{q_n(1)q_{n+1}(-1) - q_n(-1)q_{n+1}(1)} a_{n+1}$$

The optimal control problem in Eq's.(1.7-1.9) is solved with the following choice of numerical values of parameters in a certain standard case: $\omega = 1, \sigma = 1, \in = 1, f = 0, \tau = 2, \alpha = 0.5, \beta = -0.5$. Figures 1-5 illustrate the values of the state and the control for different values of n while the relative absolute errors are plotted in Figure 6.

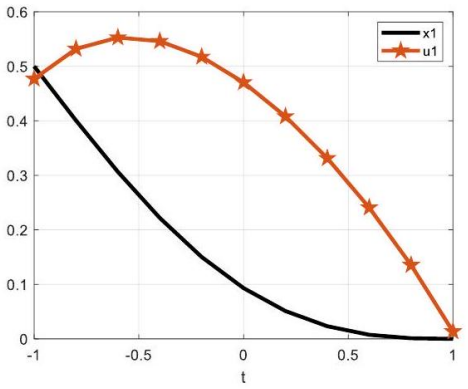


FIGURE 1. The approximate $x(t)$ and $u(t)$ with $n = 5$.

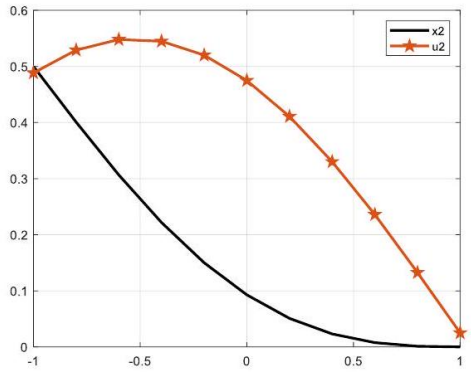


FIGURE 2. The approximate state $x(t)$ and $u(t)$ with $n = 6$.

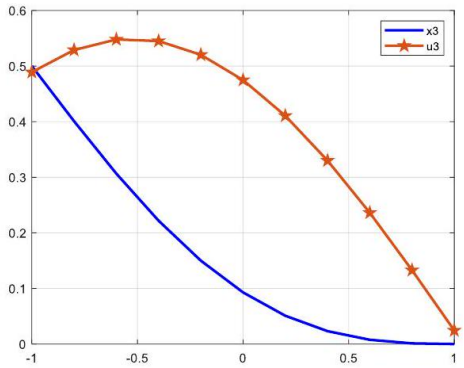


FIGURE 3. The approximate $x(t)$ and $u(t)$ with $n = 7$.

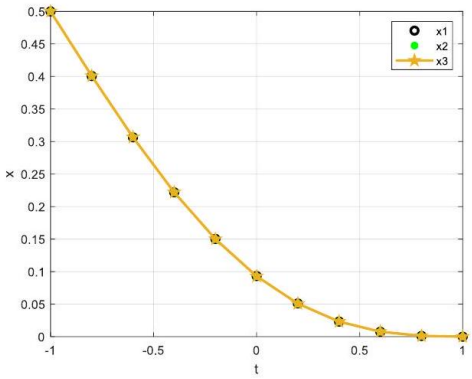


FIGURE 4. The approximate state $x(t)$ with different n.

The relative errors of the optimal cost functional J for $n = 5, 6, 7$ are respectively 00008106925, 0.00000017319, 0.00000001325.

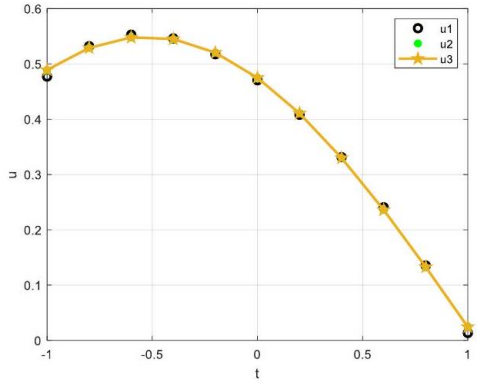


FIGURE 5. The approximate control $u(t)$

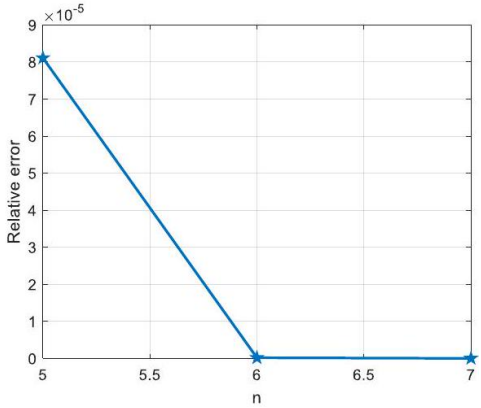


FIGURE 6. The Relative Errors of J with different n.

4. CONCLUSION

This paper proposed the modified Pell polynomial method for the optimal control problem. The new modified parameterization technique has been investigated for the approximate solution based on MPPs, which was our one important highlight. Numerical results are provided to prove the effectiveness of the suggested method. The obtained results show that as the order of the modified Pell polynomial increases the error in the approximate solution will be decreased and exactly close to the exactly one with satisfactory decimal places. This is the main modification of the approach and this small contribution to the assumption of iterative method in terms of MPPs results when obtaining the approximate solution with the minimum number of MPPs terms and satisfactory accuracy.

REFERENCES

[1] M. A. Bayrak, A. Demir, E. Ozbilge, Numerical solution of fractional diffusion equation by Chebyshev collocation method and residual power series method. Alexandria Engineering Journal, (2020).

[2] J. S. Olsen, J. Mortensen, A. S. Telyakovskiy, Polynomial approximate solutions of an unconfined Forchheimer groundwater flow equation. Advances in Water Resources, **123** (2019), 189-200.

[3] M. Sarhan, S. Shihab, B. Kashem, M. Rasheed, New Exact Operational Shifted Pell Matrices and Their Application in Astrophysics, Journal of Physics: Conference Series,**1879**(2) (2021), 022122.
[4] B. Kafash, A. Delavarkhalafi, S. M. Karbassi, Application of Chebyshev polynomials to derive efficient algorithms for the solution of optimal control problems. Scientia Iranica, **19**(3) (2012), 795-805.
[5] I. A. U. Khaleel, S. Shihab, On Generalized Vieta-Pell functions and their associated operational matrices. In AIP Conference Proceedings, **2820**(1) (2023, June). AIP Publishing.
[6] S. H. Ismaeil, S. Shihab, Some results for new modified Chebyshev functions with application. In AIP Conference Proceedings, **2820**(1) (2023, June). AIP Publishing.
[7] K. Rabiei, Y. Ordokhani, Boubaker hybrid functions and their application to solve fractional optimal control and fractional variational problems. Applications of Mathematics, **63**(5) (2018), 541-567.
[8] A. Bencheikh, L. Chiter, A new operational matrix based on Boubaker polynomials for solving fractional Emden-Fowler problem, arXiv preprint arXiv:2202.09787, (2022).
[9] A. Noori Dalawi, M. Lakestani, E. Ashpazzadeh, An Efficient Algorithm for the Multi-Scale Solution of Nonlinear Fractional Optimal Control Problems, Mathematics, **10**(20) (2022), 1-16.
[10] S. H. Salih, N. Al-Saidi, R. A. Zboon, A reliable numerical algorithm for stabilizing of the 2-dimensional logistic hyperchaotic trajectory. Al-Mustansiriyah Journal of Science, **33**(1) (2022), 51-56.
[11] N. M. Al-Saidi, S. H. Salih, A new convex controller for stabilizing of two symmetrical logistic maps. In Journal of Physics: Conference Series, **2322**(1) (2022, August). IOP Publishing.
[12] S. H. Salih, N. M. Al-Saidi, 3D-Chaotic discrete system of vector borne diseases using environment factor with deep analysis. AIMS Mathematics, **7**(3) (2022), 3972-3987.
[13] Z. K. Alabacy, S. H. Salih, S. Khlil, N. M. Al-Saidi, Stabilizing ecosystem dynamics: The toxicity effect on an ecosystem in the presence of self-defence. Commun. Math. Biol. Neurosci., **2024**(2024), ArticleID 62.
[14] S. H. Salih, N. M. Al-Saidi, A. Malik, S. J. Obaiys, Analyzing Stability and Controlling Chaos in a Unique Host Parasitoid Discrete Dynamical System with Applications in Artificial Intelligence. Mathematical Modelling of Engineering Problems, **11**(8) (2024).
[15] B. Kuloğlu, E. Özkan, A. G. Shannon, Incomplete generalized Vieta–Pell and Vieta–Pell–Lucas polynomials. Notes on Number Theory and Discrete Mathematics, **27**(4) (2021), 245-256.
[16] S. Uygun, H. Karatas, H. Aytar, Notes on Generalization of Vieta-Pell and Vieta-Pell Lucas polynomials. International Journal of Mathematics Research, **12**(1) (2020), 5-22.

Shaymaa Hussain,
Applied Science Department, University of Technology, Baghdad, Iraq, https://orcid.org/0000-0003-4275-0400
Email address: Shaymaa.h.salih@uotechnology.edu.iq

Farah J. AL-Zahid,
University of Baghdad, College of Science for Women, Department of Mathematics https://orcid.org/0000-0001-5547-1422
Email address: farahj_math@csw.uobaghdad.edu.iq

Suha Shihab,
Applied Science Department, University of Technology, Baghdad, Iraq https://orcid.org/0000-0003-3730-0149
Email address: suha.n.ahihab@uotechnology.edu.iq

Nadia M.G. Al-Saidi
Applied Science Department, University of Technology, Baghdad, Iraq https://orcid.org/0000-0002-7255-5246
Email address: nadia.m.ghanim@uotechnology.edu.iq