ISSN 2319-3786
VOLUME 11, ISSUE 1, JANUARY 2023

M
J
M

Malaya Journal of Matematik

an international journal of mathematical sciences

MKD PUBLISHING HOUSE
Dawn To Researchers
5, Venus Garden, Sappanimadai Road, Karunya Nagar (Post),
Coimbatore- 641114, Tamil Nadu, India.
www.mkdpress.com | www.malayajournal.org

The Malaya Journal of Matematik is published quarterly in single volume annually and four issues constitute one volume appearing in the months of January, April, July and October.

Subscription

The subscription fee is as follows:

USD 350.00 For USA and Canada

Euro 190.00 For rest of the world

Rs. 4000.00 In India. (For Indian Institutions in India only)

Prices are inclusive of handling and postage; and issues will be delivered by Registered Air-Mail for subscribers outside India.

Subscription Order

Subscription orders should be sent along with payment by Cheque/ D.D. favoring "Malaya Journal of Matematik" payable at COIMBATORE at the following address:

MKD Publishing House

5, Venus Garden, Sappanimadai Road, Karunya Nagar (Post),

Coimbatore- 641114, Tamil Nadu, India.

Contact No. : +91-9585408402

E-mail : info@mkdpress.com; editorinchief@malayajournal.org; publishingeditor@malayajournal.org

Website : https://mkdpress.com/index.php/index/index

Vol. 11 No. 01 (2023): Malaya Journal of Matematik (MJM)

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 1–10.
http://doi.org/10.26637/mjm1101/001

Hypersphere and the fourth Laplace-Beltrami operator in 4-space

ERHAN GÜLER[*1]

[1] *Faculty of Sciences, Department of Mathematics, Bartın University, 74100 Bartın, Turkey.*

***R**eceived 07 April 2022; **A**ccepted 16 November 2022*

Abstract. We consider hypersphere $\mathbf{x}(u,v,w)$ in the four dimensional Euclidean space $\mathbb{E}^4$. We compute the fourth Laplace-Beltrami operator of the hypersphere satisfying $\Delta^{IV}\mathbf{x} = \mathcal{A}\mathbf{x}$, where $\mathcal{A} \in Mat\,(4,4)$.

AMS Subject Classifications: Primary 53A07; Secondary 53C42.

Keywords: Four space, Euclidean space, the fourth Laplace-Beltrami operator, hypersphere, Gauss map, curvatures.

Contents

1. Introduction

In differential geometry, hyper-surfaces theory have been worked by many mathematicians for a long time. For example, Obata worked [41] certain conditions for a Riemannian manifold to be isometric with a sphere; Takahashi [44] proved that a connected Euclidean submanifold is of 1-type, iff it is either minimal in $\mathbb{E}^m$ or minimal in some hypersphere of $\mathbb{E}^m$; Chern, do Carmo, and Kobayashi [15] gave minimal submanifolds of a sphere with second fundamental form of constant length; Cheng and Yau considered hypersurfaces with constant scalar curvature; Lawson [37] gave minimal submanifolds in his book.

Chen [9–12] studied submanifolds of finite type whose immersion into $\mathbb{E}^m$ (or $\mathbb{E}^m_\nu$) by using a finite number of eigenfunctions of their Laplacian. Some results of 2-type spherical closed submanifolds were given by [6, 7, 10]; Garay researched [25] an extension of Takahashi's theorem in $\mathbb{E}^m$. Chen and Piccinni [13] focused submanifolds with finite type Gauss map in $\mathbb{E}^m$. Dursun [20] considered hypersurfaces with pointwise 1-type Gauss map in $\mathbb{E}^{n+1}$.

[*] **Corresponding author.** Email address: **eguler@bartin.edu.tr** (Erhan GÜLER)
ORCID: https://orcid.org/0000-0003-3264-6239

In $\mathbb{E}^3$; Takahashi [44] proved that minimal surfaces and spheres are the only surfaces satisfying the condition $\Delta r = \lambda r$, $\lambda \in \mathbb{R}$; Ferrandez, Garay, and Lucas [22] gave that the surfaces satisfying $\Delta H = AH$, $A \in Mat(3,3)$ are either minimal, or an open piece of sphere or of a right circular cylinder; Choi and Kim [17] studied the minimal helicoid in terms of pointwise 1-type Gauss map of the first kind; Garay [24] classified a certain class of finite type surfaces of revolution; Dillen, Pas and Verstraelen [18] focused that the only surfaces satisfying $\Delta r = Ar + B$, $A \in Mat(3,3)$, $B \in Mat(3,1)$ are the minimal surfaces, the spheres and the circular cylinders; Stamatakis and Zoubi [43] obtained surfaces of revolution satisfying $\Delta^{III} x = Ax$; Senoussi and Bekkar [42] introduced helicoidal surfaces M^2 which are of finite type with respect to the fundamental forms I, II and III, i.e., their position vector field $r(u,v)$ satisfies the condition $\Delta^J r = Ar$, $J = I, II, III$, where $A \in Mat(3,3)$; Kim, Kim and Kim [34] gave Cheng-Yau operator and Gauss map of surfaces of revolution.

In $\mathbb{E}^4$; Moore [39, 40] worked general rotational surfaces; Hasanis and Vlachos [31] considered hypersurfaces with harmonic mean curvature vector field; Cheng and Wan [14] gave complete hypersurfaces with CMC; Kim and Turgay [35] introduced surfaces with L_1-pointwise 1-type Gauss map; Arslan et al [2] worked Vranceanu surface with pointwise 1-type Gauss map; Arslan et al [3] studied generalized rotational surfaces; Aksoyak and Yaylı [32] worked flat rotational surfaces with pointwise 1-type Gauss map; Güler, Magid and Yaylı [29] introduced helicoidal hypersurfaces; Güler, Hacısalihoğlu, and Kim [28] studied Gauss map and the third Laplace-Beltrami operator of the rotational hypersurface; Güler and Turgay [30] focused Cheng-Yau operator and Gauss map of rotational hypersurfaces; Güler [27] found rotational hypersurfaces satisfying $\Delta^I R = AR$, where $A \in Mat(4,4)$. He [26] also studied fundamental form *IV* and curvature formulas of the hypersphere.

In Minkowski 4-space $\mathbb{E}^4_1$; Ganchev and Milousheva [23] indicated analogue of surfaces of [39, 40]; Arvanitoyeorgos, Kaimakamais, and Magid [5] studied that if the mean curvature vector field of M^3_1 satisfies the equation $\Delta H = \alpha H$ (α a constant), then M^3_1 has CMC; Arslan and Milousheva considered meridian surfaces of elliptic or hyperbolic type with pointwise 1-type Gauss map; Turgay introduced some classifications of Lorentzian surfaces with finite type Gauss map; Dursun and Turgay gave space-like surfaces in with pointwise 1-type Gauss map. Aksoyak and Yaylı [33] obtained general rotational surfaces with pointwise 1-type Gauss map in $\mathbb{E}^4_2$. Bektaş, Canfes, and Dursun [8] worked surfaces in a pseudo-sphere with 2-type pseudo-spherical Gauss map in $\mathbb{E}^5_2$.

We consider hypersphere in the four dimensional Euclidean space $\mathbb{E}^4$. In Section 2, we give some notions of four space. We give curvature formulas of any hypersurface in Section 3. Finally, we define hypersphere in Section 4. We compute hypersphere satisfying $\Delta^{IV}\mathbf{x} = \mathcal{A}\mathbf{x}$ for some 4×4 matrix $\mathcal{A}$ in the last section.

2. Preliminaries

In this section, we give some of basic facts and definitions, then describe notations used in this paper. Let $\mathbb{E}^m$ denote the Euclidean m-space with the canonical Euclidean metric tensor given by $\widetilde{g} = \langle\ ,\ \rangle = \sum_{i=1}^{m} dx_i^2$, where $(x_1, x_2, \dots, x_m)$ is a rectangular coordinate system in $\mathbb{E}^m$. Consider an m-dimensional Riemannian submanifold of the space $\mathbb{E}^m$. We denote the Levi-Civita connections of $\mathbb{E}^m$ and M by $\widetilde{\nabla}$ and ∇, respectively. We shall use letters X, Y, Z, W (resp., ξ, η) to denote vectors fields tangent (resp., normal) to M. The Gauss and Weingarten formulas are given, respectively, by

$$\widetilde{\nabla}_X Y = \nabla_X Y + h(X,Y), \tag{2.1}$$

$$\widetilde{\nabla}_X \xi = -A_\xi(X) + D_X \xi, \tag{2.2}$$

where h, D and A are the second fundamental form, the normal connection and the shape operator of M, respectively.

For each $\xi \in T_p^\perp M$, the shape operator A_ξ is a symmetric endomorphism of the tangent space T_pM at

$p \in M$. The shape operator and the second fundamental form are related by

$$\langle h(X,Y),\xi\rangle = \langle A_\xi X, Y\rangle .$$

The Gauss and Codazzi equations are given, respectively, by

$$\langle R(X,Y,)Z,W\rangle = \langle h(Y,Z),h(X,W)\rangle - \langle h(X,Z),h(Y,W)\rangle, \tag{2.3}$$
$$(\bar{\nabla}_X h)(Y,Z) = (\bar{\nabla}_Y h)(X,Z), \tag{2.4}$$

where $R,\ R^D$ are the curvature tensors associated with connections ∇ and D, respectively, and $\bar{\nabla}h$ is defined by

$$(\bar{\nabla}_X h)(Y,Z) = D_X h(Y,Z) - h(\nabla_X Y, Z) - h(Y, \nabla_X Z).$$

2.1. Hypersurfaces of Euclidean space

Now, let M be an oriented hypersurface in the Euclidean space $\mathbb{E}^{n+1}$, $\mathbf{S}$ its shape operator (i.e. Weingarten map) and x its position vector. We consider a local orthonormal frame field $\{e_1, e_2, \ldots, e_n\}$ of consisting of principal directions of M corresponding from the principal curvature k_i for $i = 1,2,\ldots n$. Let the dual basis of this frame field be $\{\theta_1,\theta_2,\ldots,\theta_n\}$. Then the first structural equation of Cartan is

$$d\theta_i = \sum_{i=1}^{n} \theta_j \wedge \omega_{ij}, \quad i,j = 1,2,\ldots,n, \tag{2.5}$$

where ω_{ij} denotes the connection forms corresponding to the chosen frame field. We denote the Levi-Civita connection of M and $\mathbb{E}^{n+1}$ by ∇ and $\widetilde{\nabla}$, respectively. Then, from the Codazzi equation (2.3), we have

$$e_i(k_j) = \omega_{ij}(e_j)(k_i - k_j), \tag{2.6}$$
$$\omega_{ij}(e_l)(k_i - k_j) = \omega_{il}(e_j)(k_i - k_l) \tag{2.7}$$

for distinct $i,j,l = 1,2,\ldots,n$.

We put $s_j = \sigma_j(k_1,k_2,\ldots,k_n)$, where σ_j is the j-th elementary symmetric function given by

$$\sigma_j(a_1,a_2,\ldots,a_n) = \sum_{1\le i_1<i_2<\ldots<i_j\le n} a_{i_1}a_{i_2}\ldots a_{i_j}.$$

We use following notation

$$r_i^j = \sigma_j(k_1,k_2,\ldots,k_{i-1},k_{i+1},k_{i+2},\ldots,k_n).$$

By the definition, we have $r_i^0 = 1$ and $s_{n+1} = s_{n+2} = \cdots = 0$. We call the function s_k as the k-th mean curvature of M. We would like to note that functions $H = \frac{1}{n}s_1$ and $K = s_n$ are called the mean curvature and Gauss-Kronecker curvature of M, respectively. In particular, M is said to be j-minimal if $s_j \equiv 0$ on M.

In $\mathbb{E}^{n+1}$, to find the i-th curvature formulas $\mathfrak{C}_i$ (Curvature formulas sometimes are represented as mean curvature H_i, and sometimes as Gaussian curvature K_i by different writers, such as [1] and [36]. We will call it just i-th curvature $\mathfrak{C}_i$ in this paper.), where $i = 0,...,n$, firstly, we use the characteristic polynomial of **S:**

$$P_{\mathbf{S}}(\lambda) = 0 = \det(\mathbf{S} - \lambda I_n) = \sum_{k=0}^{n} (-1)^k s_k \lambda^{n-k}, \tag{2.8}$$

where $i = 0,...,n$, I_n denotes the identity matrix of order n. Then, we get curvature formulas $\binom{n}{i}\mathfrak{C}_i = s_i$. That is, $\binom{n}{0}\mathfrak{C}_0 = s_0 = 1$ (by definition), $\binom{n}{1}\mathfrak{C}_1 = s_1, \ldots, \binom{n}{n}\mathfrak{C}_n = s_n = K$.

k-th fundamental form of M is defined by $I\left(\mathbf{S}^{k-1}(X), Y\right) = \left\langle \mathbf{S}^{k-1}(X), Y \right\rangle$. So, we have

$$\sum_{i=0}^{n} (-1)^i \binom{n}{i} \mathfrak{C}_i I\left(\mathbf{S}^{n-i}(X), Y\right) = 0. \tag{2.9}$$

In particular, one can get classical result $\mathfrak{C}_0 III - 2\mathfrak{C}_1 II + \mathfrak{C}_2 I = 0$ of surface theory for $n = 2$. See [36] for details.

For a Euclidean submanifold $x\colon M \longrightarrow \mathbb{E}^m$, the immersion (M, x) is called *finite type,* if x can be expressed as a finite sum of eigenfunctions of the Laplacian Δ of (M, x), i.e. $x = x_0 + \sum_{i=1}^{k} x_i$, where x_0 is a constant map, $x_1, \ldots, x_k$ non-constant maps, and $\Delta x_i = \lambda_i x_i$, $\lambda_i \in \mathbb{R}$, $i = 1, \ldots, k$. If λ_i are different, M is called *k-type*. See [10] for details.

2.2. Rotational hypersurfaces

We obtain a rotational hypersurface (rot-hypface for short) in Euclidean 4-space. Before we proceed, we would like to note that the definition of rot-hypfaces in Riemannian space forms were defined in [19]. A rot-hypface $M \subset \mathbb{E}^{n+1}$ generated by a curve $\mathcal{C}$ around an axis $\mathcal{C}$ that does not meet $\mathcal{C}$ is obtained by taking the orbit of $\mathcal{C}$ under those orthogonal transformations of $\mathbb{E}^{n+1}$ that leaves $\mathfrak{r}$ pointwise fixed (See [19, Remark 2.3]).

Throughout the paper, we identify a vector (a, b, c, d) with its transpose. Consider the case $n = 3$, and let $\mathcal{C}$ be the curve parametrized by

$$\gamma(w) = (\xi(w), 0, 0, \varphi(w)), \tag{2.10}$$

where ξ, φ are differentiable functions. If $\mathfrak{r}$ is the x_4-axis, then an orthogonal transformations of $\mathbb{E}^{n+1}$ that leaves $\mathfrak{r}$ pointwise fixed has the form

$$\mathbf{Z}(v, w) = \begin{pmatrix} \cos u \cos v & -\sin u & -\cos u \sin v & 0 \\ \sin u \cos v & \cos u & -\sin u \sin v & 0 \\ \sin v & 0 & \cos v & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad u, v \in \mathbb{R}.$$

Therefore, the parametrization of the rot-hypface generated by a curve $\mathcal{C}$ around an axis $\mathfrak{r}$ is given by $\mathbf{x}(u, v, w) = \mathbf{Z}(u, v)\gamma(w)$.

Definition 2.1. *Let $\mathbf{x} = \mathbf{x}(u, v, w)$ be an immersion from $M^3 \subset \mathbb{E}^3$ to $\mathbb{E}^4$. In 4-space, inner product is given by*

$$\langle \overrightarrow{x}, \overrightarrow{y} \rangle = x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4,$$

and triple vector product is defined by

$$\overrightarrow{x} \times \overrightarrow{y} \times \overrightarrow{z} = \det \begin{pmatrix} e_1 & e_2 & e_3 & e_4 \\ x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \end{pmatrix},$$

where $\overrightarrow{x} = (x_1, x_2, x_3, x_4)$, $\overrightarrow{y} = (y_1, y_2, y_3, y_4)$, $\overrightarrow{z} = (z_1, z_2, z_3, z_4)$.

Definition 2.2. *Definition For a hypface $\mathbf{x}$ in 4-space, we have*

$$(g_{ij})_{3\times3}, \ (h_{ij})_{3\times3}, \ (t_{ij})_{3\times3}, \tag{2.11}$$

where (g_{ij}), (h_{ij}), and (t_{ij}) are the first, second, and the third fundamental form matrices (or I,II, and III), respectively, where $g_{11} = \langle \mathbf{x}_u, \mathbf{x}_u \rangle$, $g_{12} = \langle \mathbf{x}_u, \mathbf{x}_v \rangle$, $g_{13} = \langle \mathbf{x}_u, \mathbf{x}_w \rangle$, $g_{22} = \langle \mathbf{x}_v, \mathbf{x}_v \rangle$, $g_{23} = \langle \mathbf{x}_v, \mathbf{x}_w \rangle$, $g_{33} =$

$\langle \mathbf{x}_w, \mathbf{x}_w \rangle$, $h_{11} = \langle \mathbf{x}_{uu}, e \rangle$, $h_{12} = \langle \mathbf{x}_{uv}, e \rangle$, $h_{13} = \langle \mathbf{x}_{uw}, e \rangle$, $h_{22} = \langle \mathbf{x}_{vv}, e \rangle$, $h_{23} = \langle \mathbf{x}_{vw}, e \rangle$, $h_{33} = \langle \mathbf{x}_{ww}, e \rangle$, $e_{11} = \langle e_u, e_u \rangle$, $e_{12} = \langle e_u, e_v \rangle$, $e_{13} = \langle e_u, e_w \rangle$, $e_{22} = \langle e_v, e_v \rangle$, $e_{23} = \langle e_v, e_w \rangle$, $e_{33} = \langle e_w, e_w \rangle$. *Here,*

$$e = \frac{\mathbf{x}_u \times \mathbf{x}_v \times \mathbf{x}_w}{\|\mathbf{x}_u \times \mathbf{x}_v \times \mathbf{x}_w\|} \tag{2.12}$$

is unit normal (i.e. the Gauss map) of hypface $\mathbf{x}$*.*

Product matrices $(g_{ij})^{-1} \cdot (h_{ij})$ gives the matrix of the shape operator $\mathbf{S}$ of hypface $\mathbf{x}$ in 4-space. See [28–30] for details.

3. i-th Curvatures

To compute the i-th mean curvature formula $\mathfrak{C}_i$, where $i = 0, .., 3$, we use characteristic polynomial $P_{\mathbf{S}}(\lambda) = a\lambda^3 + b\lambda^2 + c\lambda + d = 0$:

$$P_{\mathbf{S}}(\lambda) = \det(\mathbf{S} - \lambda I_3) = 0.$$

Then, obtain $\mathfrak{C}_0 = 1$ (by definition), $\binom{3}{1}\mathfrak{C}_1 = \binom{3}{1}H = -\frac{b}{a}$, $\binom{3}{2}\mathfrak{C}_2 = \frac{c}{a}$, $\binom{3}{3}\mathfrak{C}_3 = K = -\frac{d}{a}$.

Therefore, we find i-th curvature folmulas depends on the coefficients of the fundamental forms (g_{ij}) and (h_{ij}) in 4-space. See [26] for details.

Theorem 3.1. *Any hypersurface* $\mathbf{x}$ *in* $\mathbb{E}_2^4$ *has following curvature formulas,* $\mathfrak{C}_0 = 1$ *(by definition),*

$$\mathfrak{C}_1 = \frac{\left\{\begin{array}{c} (g_{11}h_{22} + g_{22}h_{11} - 2g_{12}h_{12})g_{33} + (g_{11}g_{22} - g_{12}^2)h_{33} \\ -2(g_{13}h_{13}g_{22} - g_{23}h_{13}g_{12} - g_{13}h_{23}g_{12} \\ +g_{11}g_{23}h_{23} - g_{13}g_{23}h_{12}) - g_{23}^2h_{11} - g_{13}^2h_{22} \end{array}\right\}}{3\left[(g_{11}g_{22} - g_{12}^2)g_{33} - g_{11}g_{23}^2 + 2g_{12}g_{13}g_{23} - g_{22}g_{13}^2\right]}, \tag{3.1}$$

$$\mathfrak{C}_2 = \frac{\left\{\begin{array}{c} (g_{11}h_{22} + g_{22}h_{11} - 2g_{12}h_{12})\,h_{33} + \left(h_{11}h_{22} - g_{12}^2\right)g_{33} \\ -2(g_{13}h_{13}h_{22} - g_{23}h_{13}h_{12} - g_{13}h_{23}h_{12} \\ +g_{23}h_{23}h_{11} - h_{13}h_{23}g_{12}) - g_{11}h_{23}^2 - g_{22}h_{13}^2 \end{array}\right\}}{3\left[(g_{11}g_{22} - g_{12}^2)g_{33} - g_{11}g_{23}^2 + 2g_{12}g_{13}g_{23} - g_{22}g_{13}^2\right]}, \tag{3.2}$$

$$\mathfrak{C}_3 = \frac{\left(h_{11}h_{22} - h_{12}^2\right)h_{33} - h_{11}h_{23}^2 + 2h_{12}h_{13}h_{23} - h_{22}h_{13}^2}{(g_{11}g_{22} - g_{12}^2)g_{33} - g_{11}g_{23}^2 + 2g_{12}g_{13}g_{23} - g_{22}g_{13}^2}. \tag{3.3}$$

Proof. See [26] for details. ■

A hypersurface $\mathbf{x}$ in $\mathbb{E}^4$ is i-minimal, when $\mathfrak{C}_i = 0$ identically on $\mathbf{x}$.

4. Hypersphere

In this section, we define hypersphere, then find its differential geometric properties in $\mathbb{E}^4$.

For an open interval $I \subset \mathbb{R}$, let $\gamma : I \longrightarrow \Pi$ be a curve in a plane Π in $\mathbb{E}^4$, and let ℓ be a straight line in Π.

Definition. A rotational hypersurface in $\mathbb{E}^4$ is called hypersphere, when a profile curve

$$\gamma(w) = (r \cos w, 0, 0, r \sin w)$$

rotates around a axis $\ell = (0, 0, 0, 1)$ for hyperradius $r > 0$.

So, in 4-space, the hypersphere which is spanned by the vector ℓ, is as follows

$$\mathbf{x}(u, v, w) = Z(u, v)\gamma(w). \tag{4.1}$$

Therefore, more clear form of (4.1) is as follows

$$\mathbf{x}(u,v,w)=\begin{pmatrix} r\cos u\cos v\cos w \\ r\sin u\cos v\cos w \\ r\sin v\cos w \\ r\sin w \end{pmatrix}, \tag{4.2}$$

where $r>0$ and $0\leq u,v,w\leq 2\pi$. When $w=0$, we have a sphere in $\mathbb{E}^4$.

Next, we will obtain the Gauss map and the curvatures of the hypersphere (4.2). The first quantities of (4.2) are as follows

$$(g_{ij})=\text{diag}\left(r^2\cos^2 v\cos^2 w,\ r^2\cos^2 w,\ r^2\right). \tag{4.3}$$

We have $\det(g_{ij})=r^6\cos^2 v\cos^4 w$. Using (2.12), we get the Gauss map of the hypersphere (4.2) as follows

$$e=\begin{pmatrix} \cos u\cos v\cos w \\ \sin u\cos v\cos w \\ \sin v\cos w \\ \sin w \end{pmatrix}. \tag{4.4}$$

Using the second differentials of (4.2) with respect to u,v,w, and the Gauss map (4.4) of the hypersphere (4.2), we have the second quantities as follows

$$(h_{ij})=\text{diag}\left(-r\cos^2 v\cos^2 w,\ -r\cos^2 w,\ -r\right). \tag{4.5}$$

So, we get $\det(h_{ij})=-r^3\cos^2 v\cos^4 w$. Using $(g_{ij})^{-1}\cdot(h_{ij})$, we calculate the shape operator matrix of the hypersphere (4.2): $\mathbf{S}=-\frac{1}{r}I_3$. Differentiating (4.4) with respect to u,v,w, we find the third quantities as follows

$$(t_{ij})=\text{diag}\left(\cos^2 v\cos^2 w,\ \cos^2 w,\ 1\right). \tag{4.6}$$

Here, $\det(t_{ij})=\cos^2 v\cos^4 w$. Computing (3.1), (3.2) and (3.3), with (4.3), (4.5), respectively, we find the curvatures of the hypersphere (4.2) as follows ($\mathfrak{C}_0=1$ by definition)

$$\mathfrak{C}_1=-\frac{1}{r},\ \mathfrak{C}_2=\frac{1}{r^2},\ \mathfrak{C}_3=-\frac{1}{r^3}.$$

Using $(f_{ij})=(t_{ij})\cdot\mathbf{S}=(h_{ij})\cdot\mathbf{S^2}=(g_{ij})\cdot\mathbf{S}^3$, we obtain the fourth fundamental form matrix $(f_{ij})_{3\times 3}$ of hypersphere (4.2) as follows

$$(f_{ij})=\text{diag}\left(-\tfrac{1}{r}\cos^2 v\cos^2 w,\ -\tfrac{1}{r}\cos^2 w,\ -\tfrac{1}{r}\right). \tag{4.7}$$

See [26] for details.

5. Hypersphere Satisfying $\Delta^{IV}\mathbf{x}=\mathcal{A}\mathbf{x}$

In this section, we give the fourth Laplace-Beltrami operator of a smooth function, then calculate it using hypersphere.

The inverse of the fourth fundamental form matrix $IV=(f_{ij})$ of any hypersurface is as follows

$$\frac{1}{f}\begin{pmatrix} f_{22}f_{33}-f_{23}f_{32} & -(f_{12}f_{33}-f_{13}f_{32}) & f_{12}f_{23}-f_{13}f_{22} \\ -(f_{21}f_{33}-f_{31}f_{23}) & f_{11}f_{33}-f_{13}f_{31} & -(f_{11}f_{23}-f_{21}f_{13}) \\ f_{21}f_{32}-f_{22}f_{31} & -(f_{11}f_{32}-f_{12}f_{31}) & f_{11}f_{22}-f_{12}f_{21} \end{pmatrix},$$

where

$$\begin{aligned} f &= \det(f_{ij}) \\ &= f_{11}f_{22}f_{33}-f_{11}f_{23}f_{32}+f_{12}f_{31}f_{23}-f_{12}f_{21}f_{33}+f_{21}f_{13}f_{32}-f_{13}f_{22}f_{31}. \end{aligned}$$

Definition 5.1. *The fourth Laplace-Beltrami operator of a smooth function $\phi = \phi(x^1, x^2, x^3) \mid_{\mathbf{D}} (\mathbf{D} \subset \mathbb{R}^3)$ of class C^3 with respect to the fourth fundamental form of a hypersurface $\mathbf{M}$ is the operator Δ^{IV} which is defined by as follows*

$$\Delta^{IV}\phi = \frac{1}{|f|^{1/2}} \sum_{i,j=1}^{3} \frac{\partial}{\partial x^i} \left(|f|^{1/2} f^{ij} \frac{\partial \phi}{\partial x^j} \right). \tag{5.1}$$

where $\left(f^{ij}\right) = \left(f_{kl}\right)^{-1}$ *and* $f = \det\left(f_{ij}\right)$.

Clearly, we can write (5.1) as follows

$$\frac{1}{|f|^{1/2}} \left\{ \begin{array}{l} \frac{\partial}{\partial x^1}\left(|f|^{1/2} f^{11} \frac{\partial \phi}{\partial x^1}\right) - \frac{\partial}{\partial x^1}\left(|f|^{1/2} t^{12} \frac{\partial \phi}{\partial x^2}\right) + \frac{\partial}{\partial x^1}\left(|f|^{1/2} t^{13} \frac{\partial \phi}{\partial x^3}\right) \\ -\frac{\partial}{\partial x^2}\left(|f|^{1/2} f^{21} \frac{\partial \phi}{\partial x^1}\right) + \frac{\partial}{\partial x^2}\left(|f|^{1/2} t^{22} \frac{\partial \phi}{\partial x^2}\right) - \frac{\partial}{\partial x^2}\left(|f|^{1/2} t^{23} \frac{\partial \phi}{\partial x^3}\right) \\ +\frac{\partial}{\partial x^3}\left(|f|^{1/2} f^{31} \frac{\partial \phi}{\partial x^1}\right) - \frac{\partial}{\partial x^3}\left(|f|^{1/2} t^{32} \frac{\partial \phi}{\partial x^2}\right) + \frac{\partial}{\partial x^3}\left(|f|^{1/2} t^{33} \frac{\partial \phi}{\partial x^3}\right) \end{array} \right\}. \tag{5.2}$$

Hence, using a smooth function $\phi = \phi(u, v, w)$, we re-write (5.2) as follows

$$\frac{1}{|f|^{1/2}} \left\{ \begin{array}{l} \frac{\partial}{\partial x^1}\left(\frac{(f_{22}f_{33}-f_{23}f_{32})\phi_u-(f_{13}f_{32}-f_{12}f_{33})\phi_v+(f_{12}f_{23}-f_{13}f_{22})\phi_w}{|f|^{1/2}}\right) \\ -\frac{\partial}{\partial x^2}\left(\frac{(f_{12}f_{33}-f_{13}f_{32})\phi_u-(f_{11}f_{33}-f_{13}f_{31})\phi_v+(f_{21}f_{13}-f_{11}f_{23})\phi_w}{|f|^{1/2}}\right) \\ +\frac{\partial}{\partial x^3}\left(\frac{(f_{12}f_{23}-f_{13}f_{22})\phi_u-(f_{21}f_{13}-f_{11}f_{23})\phi_v+(f_{11}f_{22}-f_{12}f_{21})\phi_w}{|f|^{1/2}}\right) \end{array} \right\}. \tag{5.3}$$

Therefore, the fourth Laplace-Beltrami operator of the hypersphere (4.2) is given by

$$\Delta^{IV}\mathbf{x} = \frac{1}{|f|^{1/2}} \left\{ \frac{\partial}{\partial u}\left(\frac{f_{22}f_{33}\mathbf{x}_u}{|f|^{1/2}}\right) + \frac{\partial}{\partial v}\left(\frac{f_{11}f_{33}\mathbf{x}_v}{|f|^{1/2}}\right) + \frac{\partial}{\partial w}\left(\frac{f_{11}f_{22}\mathbf{x}_w}{|f|^{1/2}}\right) \right\}, \tag{5.4}$$

Getting more clear form of the fourth Laplace-Beltrami operator $\Delta^{IV}\mathbf{x}$ of the hypersphere (4.2), we use (4.7) and (5.4). Differentiating $\frac{f_{22}f_{33}}{|f|^{1/2}}\mathbf{x}_u$, $\frac{f_{11}f_{33}}{|f|^{1/2}}\mathbf{x}_v$, $\frac{f_{11}f_{22}}{|f|^{1/2}}\mathbf{x}_w$, with respect to u, v, w, respectively, and substituting them into (5.4), we get following relations between the fourth Laplace-Beltrami operator, Gauss map, and the curvatures of the hypersphere (4.2).

Corollary 5.2. *Let $\mathbf{x} : M^3 \longrightarrow \mathbb{E}^4$ be an immersion given by (4.2). Then $\mathbf{x}$ has*

$$\Delta^{IV}\mathbf{x} = -3r^2 e,$$

where e is the Gauss map of the hypersphere $\mathbf{x}$.

Corollary 5.3. *Let $\mathbf{x} : M^3 \longrightarrow \mathbb{E}^4$ be an immersion given by (4.2). Then $\mathbf{x}$ has $\Delta^{IV}\mathbf{x} = \mathcal{A}\mathbf{x}$, where*

$$\mathcal{A} = -3r\mathfrak{C}_0 I_4 = 3r^2\mathfrak{C}_1 I_4 = -3r^3\mathfrak{C}_2 I_4 = 3r^4\mathfrak{C}_3 I_4,$$

$\mathcal{A} \in Mat(4,4)$, and I_4 is identity matrix.

6. Acknowledgement

The author would like to thank the editor and the referees for their valuable suggestions and critical remarks for improving this paper.

References

[1] ALIAS, L.J., GÜRBÜZ, N.: An extension of Takashi theorem for the linearized operators of the highest order mean curvatures, *Geom. Dedicata*, **121**(2006), 113–127.

[2] ARSLAN, K., BAYRAM, B.K., BULCA, B., KIM, Y.H., MURATHAN, C., ÖZTÜRK, G.: Vranceanu surface in $\mathbb{E}^4$ with pointwise 1-type Gauss map, *Indian J. Pure Appl. Math.,* **42**(1)(2011), 41–51.

[3] ARSLAN, K., BAYRAM, B.K., BULCA, B., ÖZTÜRK, G.: Generalized rotation surfaces in $\mathbb{E}^4$, *Results in Math.,* **61**(3)(2012), 315–327.

[4] ARSLAN, K., MILOUSHEVA, V.: Meridian surfaces of elliptic or hyperbolic type with pointwise 1-type Gauss map in Minkowski 4-space, *Taiwanese J. Math.,* **20**(2)(2016), 311–332.

[5] Arvanitoyeorgos, A., KAIMAKAMIS, G., MAGID, M.: Lorentz hypersurfaces in $\mathbb{E}_1^4$ satisfying $\Delta H = \alpha H$, *Illinois J. Math.,* **53**(2)(2009), 581–590.

[6] BARROS, M., CHEN, B.Y.: Stationary 2-type surfaces in a hypersphere, *J. Math. Soc. Japan*, **39**(4)(1987), 627–648.

[7] BARROS, M., Garay, O.J.: 2-type surfaces in S^3, *Geom. Dedicata*, **24**(3)(1987), 329–336.

[8] BEKTAŞ, B.; CANFES, E.Ö; DURSUN, U.: Classification of surfaces in a pseudo-sphere with 2-type pseudo-spherical Gauss map, *Math. Nachr.,* **290**(16)(2017), 2512–2523.

[9] CHEN, B.Y.: On submanifolds of finite type, *Soochow J. Math.,* **9**(1983), 65–81.

[10] CHEN, B.Y.: *Total Mean Curvature and Submanifolds of Finite Type*, World Scientific, Singapore, 1984.

[11] CHEN, B.Y.: *Finite Type Submanifolds and Generalizations*, University of Rome, 1985.

[12] CHEN, B.Y.: Finite type submanifolds in pseudo-Euclidean spaces and applications, *Kodai Math. J.,* **8**(3)(1985), 358–374.

[13] CHEN, B.Y., PICCINNI, P.: Submanifolds with finite type Gauss map, *Bull. Austral. Math. Soc.,* **35**(1987), 161–186.

[14] CHENG, Q.M., WAN, Q.R.: Complete hypersurfaces of $\mathbb{R}^4$ with constant mean curvature, *Monatsh. Math.,* **118**(1994), 171–204.

[15] CHERN, S.S., DO CARMO, M., KOBAYASHI, S.: *Minimal Submanifolds of a Sphere with Second Fundamental Form of Constant Length*, Functional Analysis and Related Fields, Springer, Berlin, 1970.

[16] CHENG, S.Y., YAU, S.T.: Hypersurfaces with constant scalar curvature, *Math. Ann.,* **225**(1977), 195–204.

[17] CHOI, M., KIM, Y.H.: Characterization of the helicoid as ruled surfaces with pointwise 1-type Gauss map, *Bull. Korean Math. Soc.,* **38**(2001), 753–761.

[18] DILLEN, F., PAS, J., VERSTRAELEN, L.: On surfaces of finite type in Euclidean 3-space, *Kodai Math. J.,* **13**(1990), 10–21.

[19] DO CARMO, M., DAJCZER, M.: Rotation Hypersurfaces in Spaces of Constant Curvature, *Trans. Amer. Math. Soc.,* **277**(1983), 685–709.

[20] DURSUN, U.: Hypersurfaces with pointwise 1-type Gauss map, *Taiwanese J. Math.,* **11**(5)(2007), 1407–1416.

[21] DURSUN, U., TURGAY, N.C.: Space-like surfaces in Minkowski space $\mathbb{E}_1^4$ with pointwise 1-type Gauss map, *Ukrainian Math. J.,* **71**(1)(2019), 64–80.

[22] FERRANDEZ, A., GARAY, O.J., LUCAS, P.: On a certain class of conformally at Euclidean hypersurfaces. In Global Analysis and Global Differential Geometry; Springer: Berlin, Germany 48–54, 1990.

[23] GANCHEV, G., MILOUSHEVA, V.: General rotational surfaces in the 4-dimensional Minkowski space, *Turkish J. Math.,* **38**(2014), 883–895.

[24] GARAY, O.J.: On a certain class of finite type surfaces of revolution, *Kodai Math. J.,* **11**(1988), 25–31.

[25] GARAY, O.J.: An extension of Takahashi's theorem, *Geom. Dedicata*, **34**(1990), 105–112.

[26] GÜLER, E.: Fundamental form *IV* and curvature formulas of the hypersphere, *Malaya J. Mat.,* **8**(4)(2020), 2008–2011.

[27] GÜLER, E.: Rotational hypersurfaces satisfying $\Delta^I R = AR$ in the four-dimensional Euclidean space, *J. Polytech.,* **24**(2)(2021), 517-520.

[28] GÜLER, E., HACISALIHOĞLU, H.H., KIM, Y.H.: The Gauss map and the third Laplace-Beltrami operator of the rotational hypersurface in 4-space, *Symmetry*, **10**(9)(2018), 1–12.

[29] GÜLER, E., MAGID, M., YAYLI, Y.: Laplace-Beltrami operator of a helicoidal hypersurface in four-space, *J. Geom. Symmetry Phys.,* **41**(2016), 77–95.

[30] GÜLER, E., TURGAY, N.C.: Cheng-Yau operator and Gauss map of rotational hypersurfaces in 4-space, *Mediterr. J. Math.,* **16**(3)(2019), 1–16.

[31] HASANIS, TH., VLACHOS, TH.: Hypersurfaces in $\mathbb{E}^4$ with harmonic mean curvature vector field, *Math. Nachr.,* **172**(1995), 145–169.

[32] KAHRAMAN AKSOYAK, F., YAYLI, Y.: Flat rotational surfaces with pointwise 1-type Gauss map in $\mathbb{E}^4$, *Honam Math. J.,* **38**(2)(2016), 305–316.

[33] KAHRAMAN AKSOYAK, F., YAYLI, Y.: General rotational surfaces with pointwise 1-type Gauss map in pseudo-Euclidean space $\mathbb{E}_2^4$, *Indian J. Pure Appl. Math.,* **46**(1)(2015), 107–118.

[34] KIM, D.S., KIM, J.R., KIM, Y.H.: Cheng-Yau operator and Gauss map of surfaces of revolution, *Bull. Malays. Math. Sci. Soc.,* **39**(4)(2016), 1319–1327.

[35] KIM, Y.H., TURGAY, N.C.: Surfaces in $\mathbb{E}^4$ with L_1-pointwise 1-type Gauss map, *Bull. Korean Math. Soc.,* **50**(3)(2013), 935–949.

[36] KÜHNEL, W.: *Differential Geometry, Curves-surfaces-manifolds*, Third ed. Translated from the 2013 German ed. AMS, Providence, RI, 2015.

[37] LAWSON, H.B.: *Lectures on Minimal Submanifolds*, 2nd ed. Mathematics Lecture Series 9; Publish or Perish, Inc., Wilmington, DE, USA, 1980.

[38] LEVI-CIVITA, T.: Famiglie di superficie isoparametriche nellordinario spacio euclideo, *Rend. Acad. Lincei.,* **26**(1937), 355–362.

[39] MOORE, C.: Surfaces of rotation in a space of four dimensions, *Ann. Math.,* **21**(1919), 81–93.

[40] MOORE, C.: Rotation surfaces of constant curvature in space of four dimensions, *Bull. Amer. Math. Soc.,* **26**(1920), 454–460.

[41] OBATA, M.: Certain conditions for a Riemannian manifold to be isometric with a sphere, *J. Math. Soc. Jpn.,* **14**(1962), 333–340.

[42] SENOUSSI, B., BEKKAR, M.: Helicoidal surfaces with $\Delta^J r = Ar$ in 3-dimensional Euclidean space, *Stud. Univ. Babeş-Bolyai Math.,* **60**(3)(2015), 437–448.

[43] STAMATAKIS, S., ZOUBI, H.: Surfaces of revolution satisfying $\Delta^{III} x = Ax$, *J. Geom. Graph,* **14**(2)(2010), 181–186.

[44] TAKAHASHI, T.: Minimal immersions of Riemannian manifolds, *J. Math. Soc. Jpn.,* **18**(1966), 380–385.

[45] TURGAY, N.C.: Some classifications of Lorentzian surfaces with finite type Gauss map in the Minkowski 4-space, *J. Aust. Math. Soc.,* **99**(3)(2015), 415–427.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 11–24.
http://doi.org/10.26637/mjm1101/02

A new hybrid algorithm for maximum likelihood estimation in a model of accident frequencies

Issa Cherif GERALDO[*1], **Tchilabalo Abozou KPANZOU**[2] and **Edoh KATCHEKPELE**[2]

[1] *Laboratoire d'Analyse, de Modélisations Mathématiques et Applications (LAMMA), Département de Mathématiques, Faculté des Sciences, Université de Lomé, 1 B.P. 1515 Lomé 1, Togo.*
[2] *Laboratoire de Modélisation Mathématique et d'Analyse Statistique Décisionnelle (LaMMASD), Département de Mathématiques, Faculté des Sciences et Techniques, Université de Kara, Kara, Togo.*

*R*eceived 15 July 2022; *A*ccepted 17 October 2022

Abstract. In this paper, we are interested in the numerical computation of the constrained maximum likelihood estimator (MLE) of the parameter vector of a discrete statistical model used in statistics applied to road safety. The parameter vector is divided into two blocks: one block with the parameter of interest and the second block with secondary parameters. The MLE is the solution to a system of non-linear implicit equations difficult to solve in closed-form. To overcome this difficulty, we propose a hybrid algorithm (HA) mixing the use of a one-dimensional Newton-Raphson (NR) algorithm for the first equation of the system and a fixed-point strategy for the remaining equations. Our proposed algorithm involves no matrix inversion but it partially enjoys the quadratic convergence rate of the one-dimensional NR algorithm. We illustrate its performance on simulated data and we compare it to Newton-Raphson (NR) and quasi-Newton algorithms which are two of the most used optimization algorithms. The results suggest that our HA outperforms NR and quasi-Newton algorithms. It is accurate and converges quickly for all the starting values.

AMS Subject Classifications: 62F10, 62F30, 62H10, 62H12, 62P99.

Keywords: Numerical optimization, Constrained optimization, Maximum likelihood, Multinomial distribution, Cyclic algorithm, Hybrid algorithm, Road safety measure.

Contents

[*] **Corresponding author.** Email address: **cherifgera@gmail.com** (I. C. GERALDO), **t.kpanzou@univkara.net** (T. A. KPANZOU), **edohkatchekpele@gmail.com** (E. KATCHEKPELE)

1. Introduction

The statistical evaluation of road safety measures (for example, increase or decrease of speed limit, modification of ground markings, installation of roundabouts, etc...) is one of the very important issues in statistics applied to road safety. Assume that a road safety measure has been implemented at s $(s > 0)$ geographical sites (hereinafter referred to as *treatment sites*) where crashes are categorized in r $(r > 0)$ severity levels. Also assume that each treatment site is paired with a given site (hereinafter called *control site*) with the same characteristics as the treatment site (traffic flow, accident exposure, road conditions, etc...) except that the measure was not implemented. For all $k = 1, \dots, s$, let

$$\boldsymbol{X}_k = (X_{11k}, X_{12k}, \dots, X_{1rk}, X_{21k}, X_{22k}, \dots, X_{2rk})^\top,$$

where X_{ijk} is the random number of crashes of severity level j $(j = 1, \dots, r)$ occurred on treatment site k in period i ($i = 1$ corresponds to the period before implementation of the measure and $i = 2$ corresponds to the period after). Also let

$$\boldsymbol{z}_k = (z_{1k}, \dots, z_{rk})^\top,$$

where z_{jk} is a (non-random) coefficient equal to the number of crashes of severity level j in the "after" period divided by the number of crashes of the same severity level in the "before" period on the control site paired with treatment site k.

Let $\alpha > 0$ be the overall mean effect of the measure (for example, $\alpha = 0.7 < 1$ means that the measure allowed $(1 - \alpha) \times 100\% = 30\%$ reduction in crashes) and $\beta_{jk} \in [0, 1]$ $(j = 1, \dots, r;\ k = 1, \dots, s)$ be the probability that a crash occurring on treatment site k has severity level j. Thus, for all $k = 1, \dots, s$, the vector $\boldsymbol{\beta}_k = (\beta_{1k}, \dots, \beta_{rk})^\top$ belongs to $\mathbb{S}_{r-1}$, where

$$\mathbb{S}_{r-1} = \left\{ (p_1, \dots, p_r)^\top \in [0,1]^r,\ \sum_{j=1}^{r} p_j = 1 \right\}. \tag{1.1}$$

Let $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ be the vector composed of the $1 + sr$ parameters, where $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_s^\top)^\top \in (\mathbb{S}_{r-1})^s$. The question here is how to define the probability distribution of random vectors $\boldsymbol{X}_1, \dots, \boldsymbol{X}_s$ in such a way that it can make it possible to estimate the vector $\boldsymbol{\theta}$.

There exist different models in the literature [28, 33, 34]. For all of them, the computation of the maximum likelihood estimate (MLE) $\hat{\boldsymbol{\theta}}$ of vector $\boldsymbol{\theta}$ requires the numerical maximization a log-likelihood function $L(\boldsymbol{\theta})$. A comprehensive review of modern numerical optimization algorithms is available in reference papers and books such as [1, 4, 11, 18, 19, 21, 26, 35]. The very first algorithm that comes to mind is the Newton-Raphson (NR) algorithm. It starts from an initial guess $\boldsymbol{\theta}^{(0)}$ and iterates according to the scheme

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} - [\nabla^2 L(\boldsymbol{\theta}^{(m)})]^{-1} \nabla L(\boldsymbol{\theta}^{(m)}),$$

where $\nabla^2 L(\boldsymbol{\theta}^{(m)})$ and $\nabla L(\boldsymbol{\theta}^{(m)})$ are respectively the Hessian matrix and the gradient vector evaluated at $\boldsymbol{\theta}^{(m)}$. The convergence of NR algorithm is guaranteed only if the starting guess $\boldsymbol{\theta}^{(0)}$ is close to the true solution $\hat{\boldsymbol{\theta}}$ that is unknown in practice [11]. So, the success of NR algorithm strongly depends on the appropriate choice of $\boldsymbol{\theta}^{(0)}$ in a neighbourhood of the unknown $\hat{\boldsymbol{\theta}}$. In computational terms, NR algorithm can be very costly and heavy for high-dimensional problems or simply impossible to implement because of the numerical inversion of the Hessian matrix at each iteration. When NR algorithm is unsuccessful, alternatives such as quasi-Newton algorithms [35] (which compute approximations of $[\nabla^2 L(\boldsymbol{\theta}^{(m)})]^{-1}$ at each iteration) and Derivative-Free Optimization (DFO) algorithms [2, 23, 37] may be used but they are considered as effective in solving small to mid-size problems. In practice, it is known that no optimization algorithm is perfect and one must find the most suitable algorithm for each problem.

In this paper, we build a hybrid algorithm (HA) which combines a one-dimensional NR method for estimating α and a fixed-point strategy for estimating $\boldsymbol{\beta}$. Our proposed HA alternates between computing α from $\boldsymbol{\beta}$ using the one-dimensional NR algorithm and computing $\boldsymbol{\beta}$ from α using the fixed-point based strategy. Starting from an initial guess $\boldsymbol{\theta}^{(0)}$, our proposed hybrid algorithm for estimating $\boldsymbol{\theta}$ updates the successive iterates in the following manner: at step $m+1$, $\alpha^{(m+1)}$ is updated from $\boldsymbol{\beta}^{(m)}$ afterwards $\boldsymbol{\beta}^{(m+1)}$ is updated from $\alpha^{(m+1)}$ and $\boldsymbol{\beta}^{(m)}$. Our proposed algorithm is thus a cyclic algorithm (because it cycles through the components of $\boldsymbol{\theta}$ updating one from the other rather than updating the whole parameter vector at once) and it may be easily implemented.

The remainder of this paper is organized as follows. In Section 2, we describe the statistical model and give the likelihood function. In Section 3, we present our new hybrid algorithm for estimating the parameter vector $\boldsymbol{\theta}$. In Section 4, we present the results of the numerical study of the proposed algorithm and the results of its comparison with NR and quasi-Newton algorithms which are two of the most used optimization algorithms. The paper ends with Section 5 dedicated to the conclusion and some discussions.

2. Statistical model

As mentioned earlier, there exist different models in the literature [28, 33, 34]. The model described in [33] has been the subject of several research works in the particular case $s = 1$ [9, 10, 27, 29, 32] and in the general case $s \geqslant 1$ [6, 30, 31]. These results can be adapted to the model of [28] which is a reparametrization of the one in [33]. The models in [33] and [34] have been compared in the case $s = 1$ by [8] who demonstrated theoretical results on the measure of divergence between the two models and showed through real data that both models are very competitive.

In this paper, we are interested in the model of [34] which assumes that for all $k = 1, \ldots, s$,

$$\boldsymbol{X}_k \rightsquigarrow \mathcal{M}(n_k; \boldsymbol{\pi}_k(\boldsymbol{\theta}|\boldsymbol{z}_k)), \tag{2.1}$$

where $\mathcal{M}$ denotes the multinomial distribution, n_k is the total number of crashes recorded on treatment site k,

$$\boldsymbol{\pi}_k(\boldsymbol{\theta}|\boldsymbol{z}_k) = (\pi_{11k}(\boldsymbol{\theta}|\boldsymbol{z}_k), \ldots, \pi_{1rk}(\boldsymbol{\theta}|\boldsymbol{z}_k), \pi_{21k}(\boldsymbol{\theta}|\boldsymbol{z}_k), \ldots, \pi_{2rk}(\boldsymbol{\theta}|\boldsymbol{z}_k))^\mathsf{T} \tag{2.2}$$

$$\pi_{1jk}(\boldsymbol{\theta}|\boldsymbol{z}_k) = \frac{\beta_{jk}}{1 + \alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle}, \quad j = 1, \ldots, r, \tag{2.3}$$

$$\pi_{2jk}(\boldsymbol{\theta}|\boldsymbol{z}_k) = \frac{\alpha\beta_{jk}\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle}{1 + \alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle}, \quad j = 1, \ldots, r, \tag{2.4}$$

and $\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle = \sum_{j'=1}^{r} z_{j'k}\beta_{j'k}$. The log-likelihood (see [34, p. 1275]) of observed data $\boldsymbol{x}_k = (x_{11k}, \ldots, x_{1rk}, x_{21k}, \ldots, x_{2rk})$ such as $\sum_{i=1}^{2}\sum_{j=1}^{r} x_{ijk} = n_k$, $k = 1, \ldots, s$, is given to one additive constant by:

$$L(\boldsymbol{\theta}) = \sum_{k=1}^{s}\sum_{j=1}^{r}\left\{x_{\bullet jk}\log\beta_{jk} + x_{2jk}\log\alpha - x_{\bullet jk}\log\Big(1 + \alpha\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle\Big) + x_{2jk}\log\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k\rangle\right\}, \tag{2.5}$$

where $x_{\bullet jk} = x_{1jk} + x_{2jk}$. The maximum likelihood estimation problem is the following constrained maximization problem:

$$\begin{cases} \text{maximize } L(\boldsymbol{\theta}) & \text{(2.6a)} \\ \text{subject to} & \\ \alpha > 0 \quad \text{and} \quad \beta_{jk} > 0, \quad j = 1, \ldots, r; \quad k = 1, \ldots, s, & \text{(2.6b)} \\ \displaystyle\sum_{j=1}^{r}\beta_{jk} = 1, \quad k = 1, \ldots, s. & \text{(2.6c)} \end{cases}$$

The maximum likelihood estimate (MLE) of $\boldsymbol{\theta}$ is denoted $\hat{\boldsymbol{\theta}} = (\hat{\alpha}, \hat{\boldsymbol{\beta}}^\mathsf{T})^\mathsf{T}$, where $\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\beta}}_1^\mathsf{T}, \ldots, \hat{\boldsymbol{\beta}}_s^\mathsf{T})^\mathsf{T}$ and for all $k = 1, \ldots, s$, $\hat{\boldsymbol{\beta}}_k = (\hat{\beta}_{1k}, \ldots, \hat{\beta}_{rk})^\mathsf{T}$.

In the case $s = 1$, [7] has obtained the closed-form expression of the MLE. In the next section, we build a hybrid algorithm for computing the MLE $\hat{\boldsymbol{\theta}}$ in the case $s \geqslant 1$.

3. A hybrid algorithm for computing the MLE of the vector parameter

Let us begin with the following result.

Theorem 3.1. *For all $k = 1, \dots, s$, let*

$$x_{1\bullet k} = \sum_{j=1}^{r} x_{1jk}, \quad x_{2\bullet k} = \sum_{j=1}^{r} x_{2jk} \quad \textit{and} \quad x_{1\bullet\bullet} = \sum_{k=1}^{s} x_{1\bullet k}.$$

Then, the MLE $\hat{\boldsymbol{\theta}}$ is solution to the following system of non-linear equations:

$$\begin{cases} \displaystyle\sum_{k=1}^{s} \frac{n_k}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} - x_{1\bullet\bullet} = 0 & \text{(3.1a)} \\ \\ \displaystyle x_{\bullet jk} - \frac{n_k \hat{\beta}_{jk}(\hat{\alpha} z_{jk}+1)}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} - \frac{x_{2\bullet k}\hat{\beta}_{jk}(\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle - z_{jk})}{\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0, & \\ \qquad\qquad\qquad j = 1, \dots, r, \quad k = 1, \dots, s. & \text{(3.1b)} \end{cases}$$

Proof. In [34, p. 1276], it is proved that $\hat{\boldsymbol{\theta}}$ is solution to the following system of non-linear equations:

$$\begin{cases} \displaystyle\sum_{k=1}^{s}\sum_{j=1}^{r} \frac{x_{2jk} - \hat{\alpha} x_{1jk}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0 & \text{(3.2a)} \\ \\ \displaystyle x_{\bullet jk} - \frac{n_k \hat{\beta}_{jk}(\hat{\alpha} z_{jk}+1)}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} - \frac{x_{2\bullet k}\hat{\beta}_{jk}(\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle - z_{jk})}{\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0, & \\ \qquad\qquad\qquad j = 1, \dots, r, \quad k = 1, \dots, s. & \text{(3.2b)} \end{cases}$$

The non-obvious part of the proof consists in proving Equation (3.1a) using Equation (3.2a). Passing the summation over index j to the numerator, we obtain Equation (3.2a) under the equivalent form

$$\sum_{k=1}^{s} \frac{x_{2\bullet k} - \hat{\alpha} x_{1\bullet k}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0.$$

Since $x_{2\bullet k} = n_k - x_{1\bullet k}$, we also have

$$\sum_{k=1}^{s} \frac{n_k - x_{1\bullet k} - \hat{\alpha} x_{1\bullet k}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} = 0$$

which is equivalent to

$$\sum_{k=1}^{s} \frac{n_k}{1+\hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k\rangle} - \sum_{k=1}^{s} x_{1\bullet k} = 0.$$

The proof is thus completed. ■

In the case $s = 1$, [7] obtained the exact expression of $\hat{\boldsymbol{\theta}}$. But in the general case $s \geqslant 1$, non-linear Equations (3.1a) and (3.1b) cannot be solved in closed-form. There are implicit relations between parameters $\hat{\alpha}$ and $\hat{\beta}_{jk}$'s which can make it possible to compute $\hat{\alpha}$ from the $\hat{\beta}_{jk}$'s and conversely, compute the $\hat{\beta}_{jk}$'s from $\hat{\alpha}$. So, we propose the following strategy for a fast and efficient computation of $\hat{\boldsymbol{\theta}}$: for a given estimate $\boldsymbol{\theta}^{(m)} = (\alpha^{(m)}, (\boldsymbol{\beta}^{(m)})^\top)^\top$ obtained after m iterations,

(a) use Equation (3.1a) to compute $\alpha^{(m+1)}$ from $\boldsymbol{\beta}^{(m)}$ and

(b) use Equation (3.1b) to compute the $\beta_{jk}^{(m+1)}$'s (the components of $\boldsymbol{\beta}^{(m+1)}$) from $\alpha^{(m+1)}$ and $\boldsymbol{\beta}^{(m)}$.

We propose a one-dimensional Newton-Raphson (NR) algorithm to perform Step (a) and a fixed point strategy for Step (b).

3.1. Computation of $\alpha^{(m+1)}$ via a NR algorithm

The following lemma states that for all estimate $\boldsymbol{\beta}^{(m)}$ of $\boldsymbol{\beta}$, Equation (3.1a) (considered as an equation of unknown α) has a unique solution $\alpha^{(m+1)}$.

Lemma 3.2. *Let*

$$x_{2\bullet\bullet} = \sum_{k=1}^{s} x_{2\bullet k}$$

and let $\boldsymbol{\beta}^{(m)}$ be the estimate of $\boldsymbol{\beta}$ after m iterations. Let us rewrite Equation (3.1a) *under the form $\Psi_{\boldsymbol{\beta}^{(m)}}(u) = 0$ where $\Psi_{\boldsymbol{\beta}^{(m)}}$ is the function from $[0, +\infty[$ to $]-x_{1\bullet\bullet}, x_{2\bullet\bullet}]$ defined by:*

$$\Psi_{\boldsymbol{\beta}^{(m)}}(u) = -x_{1\bullet\bullet} + \sum_{k=1}^{s} \frac{n_k}{1 + u\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle}. \tag{3.3}$$

Then, Equation $\Psi_{\boldsymbol{\beta}^{(m)}}(u) = 0$ has a unique solution $\alpha^{(m+1)}$.

Proof. The function $\Psi_{\boldsymbol{\beta}^{(m)}}$ is differentiable and its derivative $\Psi'_{\boldsymbol{\beta}^{(m)}}$, defined on $[0, +\infty[$ by

$$\Psi'_{\boldsymbol{\beta}^{(m)}}(u) = -\sum_{k=1}^{s} \frac{n_k \langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle}{(1 + u\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle)^2},$$

is strictly negative. Thus, $\Psi_{\boldsymbol{\beta}^{(m)}}$ is a continuous and strictly decreasing function. Therefore,

$$\Psi_{\boldsymbol{\beta}^{(m)}}\Big([0, +\infty[\Big) = \left] \lim_{u \to +\infty} \Psi_{\boldsymbol{\beta}^{(m)}}(u), \Psi_{\boldsymbol{\beta}^{(m)}}(0) \right],$$

where

$$\Psi_{\boldsymbol{\beta}^{(m)}}(0) = -x_{1\bullet\bullet} + \sum_{k=1}^{s} n_k = \sum_{k=1}^{s} (-x_{1\bullet k} + n_k) = \sum_{k=1}^{s} x_{2\bullet k} = x_{2\bullet\bullet} > 0$$

and

$$\lim_{u \to +\infty} \Psi_{\boldsymbol{\beta}^{(m)}}(u) = -x_{1\bullet\bullet} < 0.$$

We can conclude that $\Psi_{\boldsymbol{\beta}^{(m)}}$ is bijective from $[0, +\infty[$ to $]-x_{1\bullet\bullet}, x_{2\bullet\bullet}]$ and, since $-x_{1\bullet\bullet} < 0 < x_{2\bullet\bullet}$, the equation $\Psi_{\boldsymbol{\beta}^{(m)}}(u) = 0$ has a unique solution. ■

For a given $\boldsymbol{\beta}^{(m)}$, computing $\alpha^{(m+1)}$ (the unique solution to Equation $\Psi_{\boldsymbol{\beta}^{(m)}}(u) = 0$) in closed-form is complicated when $s > 1$. Thus, a numerical root finding algorithm is required. There certainly exist different algorithms for this purpose (see for example [12, Chapter 3]). Here, we propose the following one-dimensional Newton-Raphson (NR) algorithm:

$$u^{(\ell+1)} = u^{(\ell)} - \frac{\Psi_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}{\Psi'_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}, \qquad \ell = 1, 2, \ldots \tag{3.4}$$

where the starting guess $u^{(0)}$ is positive. The main advantage of NR algorithm is that it converges quadratically to the solution (the number of correct significant digits doubles with each iteration) if the starting guess $u^{(0)}$ is close to the unknown solution. And there lies the problem: the choice of $u^{(0)}$. Fortunately, we prove hereafter that, if we set $u^{(0)} = 0$, then, the convergence of NR iterations (3.4) is guaranteed.

Lemma 3.3. *Let $\boldsymbol{\beta}^{(m)}$ be the estimate of $\boldsymbol{\beta}$ after m iterations. If $u^{(0)} = 0$, then, the NR iterations*

$$u^{(\ell+1)} = u^{(\ell)} - \frac{\Psi_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}{\Psi'_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}, \qquad \ell = 1, 2, \ldots \tag{3.5}$$

converge to the next iterate $\alpha^{(m+1)}$ such that $\Psi_{\boldsymbol{\beta}^{(m)}}(\alpha^{(m+1)}) = 0$.

Proof. The proof uses a result from [3, Section 4.5] which states that if $f : [a, b] \to \mathbb{R}$ is a function twice differentiable such that $f(a)f(b) < 0$, and $f'(u)$ and $f''(u)$ are non-zero and preserve signs over $[a, b]$, then, proceeding from an initial approximation $u^{(0)} \in [a, b]$ such that $f(u^{(0)})f''(u^{(0)}) > 0$, the NR sequence $(u^{(\ell)})$ defined by

$$u^{(\ell+1)} = u^{(\ell)} - \frac{f(u^{(\ell)})}{f'(u^{(\ell)})}, \qquad \ell = 1, 2, \ldots,$$

converges to the unique root of f in $[a, b]$.

Let $u_M = \alpha^{(m+1)} + 1$. As $\Psi_{\boldsymbol{\beta}^{(m)}}$ is a strictly decreasing function and $\Psi_{\boldsymbol{\beta}^{(m)}}(\alpha^{(m+1)}) = 0$, we have $\Psi_{\boldsymbol{\beta}^{(m)}}(0) \times \Psi_{\boldsymbol{\beta}^{(m)}}(u_M) < 0$. Moreover, the function $\Psi_{\boldsymbol{\beta}^{(m)}}$ is twice differentiable and, for all $u \in [0, +\infty[$,

$$\Psi'_{\boldsymbol{\beta}^{(m)}}(u) = -\sum_{k=1}^{s} \frac{n_k \langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle}{(1 + u\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle)^2} < 0 \tag{3.6}$$

and

$$\Psi''_{\boldsymbol{\beta}^{(m)}}(u) = \sum_{k=1}^{s} \frac{2n_k (\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle)^2}{(1 + u\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle)^3} > 0. \tag{3.7}$$

We also have $u^{(0)} = 0 \in [0, u_M]$ and $\Psi_{\boldsymbol{\beta}^{(m)}}(0) \times \Psi''_{\boldsymbol{\beta}^{(m)}}(0) > 0$. We can conclude that the sequence $(u^{(\ell)})$ defined by (3.4) converges to the desired value $\alpha^{(m+1)}$. ■

Remark 3.4. *Actually, since $\Psi''_{\boldsymbol{\beta}^{(m)}}(u) > 0$ for all $u > 0$, any starting guess $u^{(0)}$ such that $\Psi_{\boldsymbol{\beta}^{(m)}}(\alpha^{(0)}) > 0$ could be suitable to guarantee convergence of the iterative scheme* (3.4)*. But the search of such an $u^{(0)}$ seems doomed because when $u \neq 0$, the denominators contained in the expression of $\Psi_{\boldsymbol{\beta}^{(m)}}(u)$ do not vanish any more.*

3.2. A fixed-point scheme for computing $\boldsymbol{\beta}^{(m+1)}$

From Equation (3.1b), we have for all $k = 1, \ldots, s$ and $j = 1, \ldots, r$,

$$\hat{\beta}_{jk} \left(\frac{n_k(\hat{\alpha} z_{jk} + 1)}{1 + \hat{\alpha}\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k \rangle} + x_{2\bullet k} - \frac{x_{2\bullet k} z_{jk}}{\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k \rangle} \right) = x_{\bullet jk}. \tag{3.8}$$

In this case too, it is complicated to get $\hat{\beta}_{jk}$ in closed-form because it also appears in the denominator as a component of the weighted sum $\langle \boldsymbol{z}_k, \hat{\boldsymbol{\beta}}_k \rangle$. For all $k = 1, \ldots, s$ and $j = 1, \ldots, r$, we propose to compute $\beta_{jk}^{(m+1)}$ using the following fixed-point scheme:

$$\beta_{jk}^{(m+1)} = x_{\bullet jk} \Bigg/ \left(\frac{n_k(\alpha^{(m+1)} z_{jk} + 1)}{1 + \alpha^{(m+1)}\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle} + x_{2\bullet k} - \frac{x_{2\bullet k} z_{jk}}{\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle} \right). \tag{3.9}$$

3.3. The hybrid algorithm (HA)

Our proposed hybrid algorithm is Algorithm 3.5. It starts from $\boldsymbol{\theta}^{(0)} = (\alpha^{(0)}, (\boldsymbol{\beta}^{(0)})^\top)^\top$, where $\alpha^{(0)} > 0$ is randomly set and $\boldsymbol{\beta}^{(0)} = ((\boldsymbol{\beta}_1^{(0)})^\top, \dots, (\boldsymbol{\beta}_s^{(0)})^\top)^\top$ is also randomly set such that for all $k = 1, \dots, s$, $\boldsymbol{\beta}_k^{(0)} = (\beta_{1k}^{(0)}, \dots, \beta_{rk}^{(0)})^\top \in \mathbb{S}_{r-1}$. At the $(m+1)-$iteration, the update $\alpha^{(m+1)}$ is computed from $\boldsymbol{\beta}^{(m)}$ using the one-dimensional NR iterations (3.4), afterwards $\boldsymbol{\beta}^{(m+1)}$ is updated from $\alpha^{(m+1)}$ and $\boldsymbol{\beta}^{(m)}$ using Equation (3.9) and so on. This process is repeated until a convergence criterion is satisfied.

Algorithm 3.5.
Input: $\boldsymbol{x}_1, \dots, \boldsymbol{x}_s, \boldsymbol{z}_1, \dots, \boldsymbol{z}_s, \epsilon_1 > 0$ *and* $\epsilon_2 > 0$.
Output: MLE $\hat{\boldsymbol{\theta}}$.

1. (a) *Set* $m = 0$
   (b) *Initialize* $\boldsymbol{\theta}^{(0)} = (\alpha^{(0)}, (\boldsymbol{\beta}_1^{(0)})^\top, \dots, (\boldsymbol{\beta}_s^{(0)})^\top)^\top$ *randomly such that* $\alpha^{(0)} > 0$, *and for all* $k = 1, \dots, s$, $\boldsymbol{\beta}_k^{(0)} = (\beta_{1k}^{(0)}, \dots, \beta_{rk}^{(0)})^\top \in \mathbb{S}_{r-1}$.
2. (a) *Update* $\alpha^{(m+1)}$ *as follows:*
   i. *Set* $\ell = 0$
   ii. *Set* $u^{(\ell)} = 0$
   iii. *Set*
   $$u^{(\ell+1)} = u^{(\ell)} - \frac{\Psi_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}{\Psi'_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)})}$$
   iv. *Set* $\ell \leftarrow \ell + 1$
   v. *If* $|\Psi_{\boldsymbol{\beta}^{(m)}}(u^{(\ell)}) - \Psi_{\boldsymbol{\beta}^{(m)}}(u^{(\ell-1)})| \geqslant \epsilon_1$, *go back to Step (ii).*
   vi. *Set* $\alpha^{(m+1)} = u^{(\ell)}$.

   (b) *For all* $k = 1, \dots, s$, *compute* $\boldsymbol{\beta}_k^{(m+1)} = (\beta_{1k}^{(m+1)}, \dots, \beta_{rk}^{(m+1)})^\top$ *where, for all* $j = 1, \dots, r$,
   $$\beta_{jk}^{(m+1)} = x_{\bullet jk} \Bigg/ \left(\frac{n_k(\alpha^{(m+1)} z_{jk} + 1)}{1 + \alpha^{(m+1)} \langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle} + x_{2\bullet k} - \frac{x_{2\bullet k} z_{jk}}{\langle \boldsymbol{z}_k, \boldsymbol{\beta}_k^{(m)} \rangle} \right)$$
   (c) *Update* $\boldsymbol{\beta}^{(m+1)} = ((\boldsymbol{\beta}_1^{(m+1)})^\top, \dots, (\boldsymbol{\beta}_s^{(m+1)})^\top)^\top$.
   (d) *Set* $\boldsymbol{\theta}^{(m+1)} = (\alpha^{(m+1)}, (\boldsymbol{\beta}^{(m+1)})^\top)^\top$
   (e) *Set* $m \leftarrow m + 1$.
3. *If* $|L(\boldsymbol{\theta}^{(m)}) - L(\boldsymbol{\theta}^{(m-1)})| \geqslant \epsilon_2$, *go back to Step 2.*
4. *Set* $\hat{\boldsymbol{\theta}} \leftarrow \boldsymbol{\theta}^{(m)}$.

4. Simulation study

In this section, we compare our proposed HA to some of the best optimization algorithms available in R software [36] in terms of accuracy, robustness (the ability of each algorithm to perform well regardless of the starting guess) and computation time. The selected algorithms are the Newton-Raphson (NR) algorithm and the quasi-Newton BFGS algorithm (named after its authors Broyden, Fletcher, Goldfarb and Shanno) [35, Section 6.1]. The NR and BFGS algorithms are implemented using respectively the R packages **nleqslv** [13] and **alabama** [38]. Accuracy is measured with the Mean Squared Error (MSE)

$$\mathrm{MSE}(\hat{\boldsymbol{\theta}}|\boldsymbol{\theta}^0) = \frac{1}{1+sr} \|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}^0\|^2,$$

where $\boldsymbol{\theta}^0 = (\alpha^0, (\boldsymbol{\beta}_1^0)^\top, \ldots, (\boldsymbol{\beta}_s^0)^\top)^\top$ is the true value and $\|\cdot\|$ represents the usual Euclidean norm. Robustness is measured by comparing the convergence proportions (the ratio of the number of times each algorithm converged to the total number of replications) and the numbers of iterations for different starting guesses.

4.1. Data generation

The data are generated using Formulas (2.1), (2.2), (2.3) and (2.4) where the true value $\boldsymbol{\theta}^0 = (\alpha^0, (\boldsymbol{\beta}_1^0)^\top, \ldots, (\boldsymbol{\beta}_s^0)^\top)^\top$ of $\boldsymbol{\theta}$ is presented under five scenarios described below:

- **Scenario 1**: $s = 2$, $r = 2$,

$$\alpha^0 = 0.8, \qquad \boldsymbol{\beta}_1^0 = (0.85, 0.15)^\top, \qquad \boldsymbol{\beta}_2^0 = (0.40, 0.60)^\top.$$

- **Scenario 2**: $s = 5$, $r = 3$,

$$\begin{aligned} \alpha^0 = 1,\ \boldsymbol{\beta}_1^0 &= (0.80, 0.15, 0.05)^\top, \\ \boldsymbol{\beta}_2^0 &= (0.10, 0.30, 0.60)^\top, \\ \boldsymbol{\beta}_3^0 &= (0.35, 0.30, 0.35)^\top, \\ \boldsymbol{\beta}_4^0 &= (0.70, 0.20, 0.10)^\top, \\ \boldsymbol{\beta}_5^0 &= (0.30, 0.40, 0.30)^\top. \end{aligned}$$

- **Scenario 3**: $s = 10$, $r = 5$,

$$\begin{aligned} \alpha^0 = 1,\ \boldsymbol{\beta}_k^0 &= (0.40, 0.10, 0.05, 0.25, 0.20)^\top, \quad k \in \{1, 3, 5, 9\}, \\ \boldsymbol{\beta}_k^0 &= (0.30, 0.15, 0.10, 0.25, 0.20)^\top, \quad k \in \{2, 4, 7\}, \\ \boldsymbol{\beta}_k^0 &= (\underbrace{0.20, \ldots, 0.20}_{5})^\top, \quad i \in \{6, 8, 10\}. \end{aligned}$$

- **Scenario 4**: $s = 10$, $r = 10$,

$$\begin{aligned} \alpha^0 = 1.2,\ \boldsymbol{\beta}_k^0 &= (0.40, 0.10, 0.05, \underbrace{0.10}_{2}, \underbrace{0.05}_{5})^\top, \quad k \in \{1, 5, 7, 10\}; \\ \boldsymbol{\beta}_k^0 &= (\underbrace{0.10}_{3}, \underbrace{0.05}_{2}, 0.10, 0.25, \underbrace{0.05}_{2}, 0.15)^\top, \quad k \in \{2, 3, 6\}; \\ \boldsymbol{\beta}_k^0 &= (\underbrace{0.10, \ldots, 0.10}_{10})^\top, \quad k \in \{4, 8, 9\}. \end{aligned}$$

- **Scenario 5**: $s = 20$, $r = 10$,

$$\begin{aligned} \alpha^0 = 1.2,\ \boldsymbol{\beta}_k^0 &= (0.40, 0.10, 0.05, \underbrace{0.10}_{2}, \underbrace{0.05}_{5})^\top, \quad k \in \{1, 5, 7, 10, 11, 15, 17, 20\}, \\ \boldsymbol{\beta}_k^0 &= (\underbrace{0.10}_{3}, \underbrace{0.05}_{2}, 0.10, 0.25, \underbrace{0.05}_{2}, 0.15)^\top, \quad k \in \{2, 3, 6, 12, 13, 16\}, \\ \boldsymbol{\beta}_k^0 &= (\underbrace{0.10, \ldots, 0.10}_{10})^\top, \quad k \in \{4, 8, 9, 14, 18, 19\}. \end{aligned}$$

These scenarios allow to have low and large values of the number of parameters $(1 + sr)$. The values of the number of parameters for these scenarios are given in Table 1.

For each scenario, we set $n_1 = \cdots = n_s = n$, where n had two different values: a low value ($n = 50$) and a great value ($n = 5000$). In order to explore a plethora of starting guesses in the parameter space, we have considered a random initialization scheme for setting the starting guess $\boldsymbol{\theta}^{(0)} = (\alpha^{(0)}, (\boldsymbol{\beta}^{(0)})^\top)^\top$. The parameter $\alpha^{(0)}$ is randomly generated from an uniform distribution and each $\boldsymbol{\beta}_k^{(0)} = (\beta_{1k}^{(0)}, \ldots, \beta_{rk}^{(0)})^\top$ is randomly generated as $\boldsymbol{\beta}_k^{(0)} = \boldsymbol{U}_k / \sum_{j=1}^r u_{jk}$ where $\boldsymbol{U}_k = (u_{1k}, \ldots, u_{rk})^\top$ is a $r-$dimensional vector whose components are randomly generated from a uniform distribution $\mathcal{U}[0.05; 0.95]$.

Table 1: Number of parameters for the different scenarios

Scenario	1	2	3	4	5
s	2	5	10	10	20
r	2	3	5	10	10
Number of parameters $(1 + sr)$	5	16	51	101	201

Table 2: Results for Scenario 1 ($s = 2$ and $r = 2$). Values in brackets are standard deviations.

		TRUTH	HA	NR	BFGS
	α	0.800	0.817 (0.172)	0.817 (0.172)	0.811 (0.170)
	β_{11}	0.850	0.848 (0.049)	0.848 (0.049)	0.848 (0.048)
	β_{21}	0.150	0.152 (0.049)	0.152 (0.049)	0.152 (0.048)
	β_{12}	0.400	0.400 (0.071)	0.400 (0.071)	0.399 (0.071)
	β_{22}	0.600	0.600 (0.071)	0.600 (0.071)	0.601 (0.071)
$n = 50$	Convergence proportion (%)	-	100	99.6	77.1
	Iterations	-	7.1 (1.6)	5.7 (1.5)	13 (0)
	CPU time (secs)	-	0.004	0.002	0.100
	Time ratio	-	1	0.6	27.7
	Log-likelihood	-	-121.37	-121.37	-121.37
	MSE	-	9e-03	8.9e-03	8.7e-03
	α	0.800	0.800 (0.016)	0.800 (0.016)	0.800 (0.016)
	β_{11}	0.850	0.850 (0.005)	0.850 (0.005)	0.850 (0.005)
	β_{21}	0.150	0.150 (0.005)	0.150 (0.005)	0.150 (0.005)
	β_{12}	0.400	0.400 (0.007)	0.400 (0.007)	0.400 (0.007)
	β_{22}	0.600	0.600 (0.007)	0.600 (0.007)	0.600 (0.007)
$n = 5000$	Convergence proportion (%)	-	100	99.8	77.7
	Iterations	-	6.5 (1.1)	6 (1.5)	16 (0)
	CPU time (secs)	-	0.004	0.003	0.121
	Time ratio	-	1	0.8	33.3
	Log-likelihood	-	-12245.17	-12245.17	-12245.17
	MSE	-	8.1e-05	8.1e-05	8.2e-05

Table 3: Results for Scenario 2 ($s = 5$ and $r = 3$). Values in brackets are standard deviations.

		HA	NR	BFGS
	Convergence proportion (%)	100	99.9	51.6
	Iterations	9.5 (1.8)	8 (2.8)	14 (0.2)
$n = 50$	CPU time (secs)	0.007	0.017	0.248
	Time ratio	1	2.5	37.1
	Log-likelihood	-390.89	-390.89	-390.89
	MSE	4.3e-03	4.3e-03	4.2e-03
	Convergence proportion (%)	100	99.8	56.1
	Iterations	8.1 (1.1)	8.3 (2.9)	18 (4)
$n = 5000$	CPU time (secs)	0.005	0.013	0.405
	Time ratio	1	2.9	89.4
	Log-likelihood	-39171.57	-39171.57	-40218.56
	MSE	4.4e-05	4.4e-05	1.8e-02

Table 4: Results for Scenario 3 ($s = 10$ and $r = 5$). Values in brackets are standard deviations.

		HA	NR	BFGS
$n = 50$	Convergence proportion (%)	100	99.8	21.4
	Iterations	10.8 (1.8)	10.6 (3.2)	13.9 (0.3)
	CPU time (secs)	0.006	0.070	0.321
	Time ratio	1	11.9	54.1
	Log-likelihood	-1083.73	-1083.73	-1083.73
	MSE	3e-03	3e-03	3e-03
$n = 5000$	Convergence proportion (%)	100	99.9	24
	Iterations	8.6 (0.9)	10.5 (2.8)	16.7 (4.4)
	CPU time (secs)	0.005	0.062	0.452
	Time ratio	1	13.4	98.3
	Log-likelihood	-109133.14	-109133.14	-111075.03
	MSE	3.2e-05	3.2e-05	5.4e-03

Table 5: Results for Scenario 4 ($s = 10$ and $r = 10$). Values in brackets are standard deviations.

		HA	NR	BFGS
$n = 50$	Convergence proportion (%)	100	97.1	20.1
	Iterations	10.1 (1.5)	14.2 (4.9)	13.9 (0.3)
	CPU time (secs)	0.008	0.162	0.635
	Time ratio	1	19.5	76.4
	Log-likelihood	-1406.16	-1406.16	-1406.16
	MSE	1.9e-03	1.9e-03	1.9e-03
$n = 5000$	Convergence proportion (%)	100	93.4	19.6
	Iterations	8.4 (0.7)	15 (5.3)	17.8 (2)
	CPU time (secs)	0.005	0.181	1.012
	Time ratio	1	34.4	193
	Log-likelihood	-138384.13	-138384.13	-138673.98
	MSE	1.8e-05	1.8e-05	9.6e-04

Table 6: Results for Scenario 5 ($s = 20$ and $r = 10$). Values in brackets are standard deviations.

		HA	NR	BFGS
$n = 50$	Convergence proportion (%)	100	95.2	3.5
	Iterations	11 (1.6)	16.6 (5.1)	14 (0)
	CPU time (secs)	0.008	0.644	2.079
	Time ratio	1	80.1	258.6
	Log-likelihood	-2816.68	-2816.68	-2816.68
	MSE	1.7e-03	1.7e-03	1.6e-03
$n = 5000$	Convergence proportion (%)	100	88	5
	Iterations	8.8 (0.7)	17.9 (5.7)	11 (8)
	CPU time (secs)	0.006	0.732	2.254
	Time ratio	1	123.9	381.6
	Log-likelihood	-276918.57	-276918.57	-282868.47
	MSE	1.7e-05	1.7e-05	8.6e-03

4.2. Results

Tables 2 to 6 present the average estimates for 1000 replicates for each scenario and each value of n. In these tables, CPU (central process unit) times are given in seconds and time ratios are calculated as the ratio between the mean CPU time of a given algorithm and the mean CPU time of the HA. Thus, the time ratio of the HA is always 1. To save space, the estimate $\hat{\boldsymbol{\theta}}$ has been included only for Scenario 1 (Table 2).

Ideally, any iterative algorithm should converge to a solution close to the true value in a relatively reasonable computation time. By analysing Tables 2 to 6, we see that the HA has a convergence proportion of 100%. The NR algorithm has a convergence proportion between 88% and 99.6% so for this model, the convergence proportion of the NR algorithm does not reach 100% even for a small number of parameters. The convergence proportion of the BFGS algorithm decreases drastically (from 77.7% down to 3.5%) when the number of parameters increases. Concerning precision, the general trend observed is the decrease of the MSE when the sample size n increases. But the MSE of BFGS algorithm remain relatively high when n goes from 50 to 5000.

For computation times, it is noticed that the computation time of HA is stable (between 0.004 and 0.008 seconds) despite the increase in the number of parameters to be estimated. This is not the case for NR and BFGS algorithms. The analysis of the CPU time ratios shows that the computation times required by NR and BFGS algorithms increase with the number of parameters. Our proposed hybrid algorithm is up to 123 times faster than NR and up to 381 times faster than BFGS.

As far as the number of iterations is concerned, the HA has a low average number of iterations. The BFGS and NR algorithms look much more sensitive to the starting guesses. The average number of iterations of NR algorithm and its standard deviation increase with the number of parameters while the number of iterations of BFGS algorithm remains very high.

Overall, our hybrid algorithm outperforms NR and BFGS algorithms in terms of convergence proportion, convergence speed, computation time and accuracy.

5. Discussion and conclusion

For numerical optimization in general, and computation of maximum likelihood estimates (MLE) specifically, Newton-Raphson (NR) algorithm is the very first to be considered because of its fast convergence when the starting guess is close to the unknown solution. But because it requires matrix inversion at each iteration, it becomes tricky in computation terms for high-dimensional problems. Alternatives such as the quasi-Newton BFGS algorithm are usually considered as good remedies when NR is unsuccessful.

In this paper, we presented a new hybrid algorithm (HA) for estimating the parameter vector $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^{\mathsf{T}})^{\mathsf{T}}$ of a statistical model used in road safety. The parameter $\alpha > 0$ which is also the parameter of interest represents the mean effect (in the multiplicative sense) of a road safety measure and the vector $\boldsymbol{\beta}$ is a vector of sr probabilities where s is the number of sites where the measure has been applied and r is the number of accidents severity levels. Our HA mixes a one-dimensional NR approach for computing the parameter α and a fixed-point strategy for computing $\boldsymbol{\beta}$. It cycles through the parameter vector updating α from $\boldsymbol{\beta}$ and $\boldsymbol{\beta}$ from both α and $\boldsymbol{\beta}$ until a convergence criterion is satisfied. It thus partially enjoys (for the estimation of α) the fast convergence property of NR while avoiding its defects (it requires no matrix inversion and the starting value is automated in order to guarantee convergence). Since it cycles through the parameters, our HA is also a cyclic algorithm and, therefore, as claimed by [16], it enjoys overall fast convergence because the log-likelihood function is always driven in the right direction. The numerical studies performed in this paper suggest that our HA outperforms NR and BFGS algorithms in terms of convergence proportion, convergence speed, computation time and accuracy. They also suggest that the HA is globally convergent (it converges to the MLE for all starting guesses). A future work may be devoted to the study of the theoretical properties of our proposed hybrid algorithm such as the global convergence (the convergence to the MLE whatever the starting point $\boldsymbol{\theta}^{(0)}$).

In this paper, we have not mentioned MM (Minorization-Maximization) algorithms [16, 20, 22] which have become very popular over the years. Strictly speaking, the MM method is not a simple box in which we insert the function $L(\theta)$ while waiting for the output but it is rather a principle of construction of an optimization algorithm for a specific function. For example, the MM principle for building an algorithm for maximization of a log-likelihood $L(\boldsymbol{\theta})$ consists in constructing a minorizing function $g(\theta)$ for $L(\boldsymbol{\theta})$ such that the maximization of $g(\theta)$ is equivalent to that of $L(\theta)$ and afterwards construct an algorithm for maximizing $g(\theta)$ (see, for example, [16, 20, 22] for the definition of a minorizing function). The design of an MM algorithm is often not easy and for each $L(\boldsymbol{\theta})$ it is necessary to build an adequate MM algorithm. This explains why the design of an MM algorithm is often done by model [5, 14, 15, 17, 24, 39]. In this line of thought, in [25], the authors constructed an MM algorithm for estimating the parameters of the road safety model of [33]. However this algorithm cannot be used for the model [34] considered in the present paper. In a future work, it would be interesting to build an MM algorithm for the model of [34] and compare it with our hybrid algorithm.

References

[1] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal and C. Sagastizábal, *Numerical Optimization – Theoretical and Practical Aspects*, Springer Verlag, Berlin, (2006).

[2] A. R. Conn, K. Scheinberg and L. N. Vicente, *Introduction to Derivative-Free Optimization*, MPS-SIAM Book Series on Optimization. SIAM, Philadelphia, (2009).

[3] B. P. Demidovich and I. A. Maron, *Computational Mathematics*. MIR Publishers, Moscow, (1981).

[4] J. E. Dennis Jr and R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*, SIAM, Philadelphia, (1996).

[5] J. Ding, G. L. Tian and K. C. Yuen, A new MM algorithm for constrained estimation in the proportional hazards model, *Comput. Statist. Data Anal.*, **84**(2015), 135–151.

[6] I. C. Geraldo, An Automated Profile-Likelihood-Based Algorithm for Fast Computation of the Maximum Likelihood Estimate in a Statistical Model for Crash Data, *J. Appl. Math.*, **2022**(2022), Article ID 6974166, 11 p.

[7] I. C. Geraldo, On the maximum likelihood estimator for a discrete multivariate crash frequencies model, *Afr. Stat.*, **15(2)**(2020), 2335–2348.

[8] I. C. Geraldo and T. A. Kpanzou, Theoretical and empirical comparison of two discrete statistical models of crash frequencies, *Malaya J. Mat.*, **10(04)**(2022), 322–335.

[9] I. C. Geraldo, A. N'Guessan and K. E. Gneyou, A note on the strong consistency of a constrained maximum likelihood estimator used in crash data modeling, *C. R. Math. Acad. Sci. Paris*, **353(12)**(2015), 1147–1152.

[10] I. C. Geraldo, A. N'Guessan and K. E. Gneyou, Global convergence and ascent property of a cyclic algorithm used for statistical analysis of crash data, *Afr. Stat.*, **13(2)**(2018), 1631–1643.

[11] I. Griva, S. G. Nash, and A. Sofer, *Linear and Nonlinear Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 2nd edition, (2009).

[12] R. K. Gupta, *Numerical methods: Fundamentals and applications*, Cambridge University Press, New Delhi, (2019).

[13] B. Hasselman, *nleqslv: Solve Systems of Nonlinear Equations*, URL `https://CRAN.R-project.org/package=nleqslv`, R package version 3.3.2., (2018).

[14] D. R. HUNTER, MM algorithms for generalized Bradley-Terry models, *Ann. Statist.*, **32(1)**(2004), 384–406.

[15] D. R. HUNTER AND R. LI, Variable Selection using MM Algorithms, *Ann. Statist.*, **33(4)**(2005), 1617–1642.

[16] D. R. HUNTER AND K. LANGE, A tutorial on MM algorithms, *Amer. Statist.*, **58(1)**(2004), 30–37.

[17] J. JAMES, MM algorithm for general mixed multinomial logit models, *J. Appl. Econometrics*, **32(4)**(2017), 841–857.

[18] M. J. KOCHENDERFER AND T. A. WHEELER, *Algorithms for optimization*, The MIT Press, Cambridge (Massachusetts), (2019).

[19] K. LANGE, *Optimization*, Springer, New York, 2nd edition, (2013).

[20] K. LANGE, *MM optimization algorithms*, SIAM, Philadelphia, (2016).

[21] K. LANGE, E. C. CHI, AND H. ZHOU, A brief survey of modern optimization for statisticians, *Int. Stat. Rev.*, **82(1)**(2014), 46–70.

[22] K. LANGE, D. R. HUNTER AND I. YANG, Optimization transfer using surrogate objective functions, *J. Comput. Graph. Statist.*, **9(1)**(2000), 1–20.

[23] J. LARSON, M. MENICKELLY, AND S. M. WILD, Derivative-free optimization methods, *Acta Numer.*, **28**(2019), 287–404.

[24] S. LEE AND J. Z. HUANG, A coordinate descent MM algorithm for fast computation of sparse logistic PCA, *Comput. Statist. Data Anal.*, **62**(2013), 26–38.

[25] A. MKHADRI, A. N'GUESSAN AND B. HAFIDI, An MM algorithm for constrained estimation in a road safety measure modeling, *Comm. Statist. Simulation Comput.*, **39(5)**(2010), 1057–1071.

[26] J. F. MONAHAN, *Numerical Methods of Statistics*, Cambridge University Press, 2nd edition, (2011).

[27] A. N'GUESSAN, Analytical existence of solutions to a system of nonlinear equations with application, *J. Comput. Appl. Math.*, **234(1)**(2010), 297–304.

[28] A. N'GUESSAN AND F. BELLAVANCE, A confidence interval estimation problem using Schur complement approach, *C. R. Math. Rep. Acad. Sci. Canada*, **27(3)**(2005), 84–91.

[29] A. N'GUESSAN AND I. C. GERALDO, A cyclic algorithm for maximum likelihood estimation using Schur complement, *Numer. Linear Algebra Appl.*, **22(6)**(2015), 1161–1179.

[30] A. N'GUESSAN AND C. LANGRAND, A covariance components estimation procedure when modelling a road safety measure in terms of linear constraints, *Statistics*, **39(4)**(2005), 303–314.

[31] A. N'GUESSAN AND C. LANGRAND, A Schur complement approach for computing subcovariance matrices arising in a road safety measure modelling. *J. Comput. Appl. Math.*, **177(2)**(2005), 331–345.

[32] A. N'GUESSAN AND M. TRUFFIER, Impact d'un aménagement de sécurité routière sur la gravité des accidents de la route, *J. SFdS*, **49**(2008), 23–41.

[33] A. N'GUESSAN, A. ESSAI AND C. LANGRAND, Estimation multidimensionnelle des contrôles et de l'effet moyen d'une mesure de sécurité routière, *Revue de statistique appliquée*, **49(2)**(2001), 85–102.

[34] A. N'GUESSAN, A. ESSAI AND M. N'ZI, An estimation method of the average effect and the different accident risks when modelling a road safety measure: A simulation study, *Comput. Statist. Data Anal.*, **51(2)**(2006), 1260–1277.

[35] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer, second edition, (2006).

[36] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, URL `https://www.R-project.org/`, (2022).

[37] L. M. Rios and N. V. Sahinidis, Derivative-free optimization: a review of algorithms and comparison of software implementations, *J. Global Optim.*, **56(3)**(2013), 1247–1293.

[38] R. Varadhan, *alabama: Constrained Nonlinear Optimization*, URL `https://CRAN.R-project.org/package=alabama`, R package version 2015.3-1, (2015).

[39] H. Zhou and K. Lange, MM algorithms for some discrete multivariate distributions, *J. Comput. Graph. Statist.*, **19(3)**(2010), 645–665.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 25–42.
http://doi.org/10.26637/mjm1101/003

Approximation of solution for generalized Basset equation with finite delay using Rothe's approach

RAKSHA DEVI*[1] AND DWIJENDRA N. PANDEY[2]

[1,2] *Department of Mathematics, Indian Institute of Technology Roorkee, Uttarakhand, India.*

Received 27 August 2022; Accepted 02 November 2022

Abstract. This study focuses on the use of the Riemann-Liouville fractional (R-L) derivative to address an initial boundary value problem for a fractional order differential equation with finite delay (FDDE). Rothe's methodology is used to prove the existence and uniqueness of the strong solution and classical solution to the restated abstract FDDE. Some examples based on abstract theory and numerical solutions of FDDEs arising in fluid dynamics are presented.

AMS Subject Classifications: 34G20, 34K37, 12H20.

Keywords: Accretive operator, strong solution, classical solution, delay differential equation, Rothe's method.

Contents

1. Introduction

Fractional differential equations are now being utilized to describe real-world issues in engineering, science, finance, and other fields. Numerous methods based on integer order derivatives do not adequately capture the complexity of real-world occurences[17, 18]. There are various definitions for fractional derivatives in contrast to integer order derivatives. The R-L derivative is dealt in this analysis as it is seen in the study by Li et al. [28] that the R-L derivative is more realistic compared to other derivatives and the Riemann derivative is quite helpful in characterizing anomalous diffusion, Levy flights, and traps [24, 29, 30].

It is seen in the literature that differential equations with R-L fractional derivatives are difficult to study because of initial conditions as there is a singularity at $t = 0$. While R-L FDEs with homogeneous initial conditions are treated similarly to FDEs with Caputo derivatives [17]. Heymans et al. [25] and Hristova et al.

*Corresponding author. Email address: **raksha_d@ma.iitr.ac.in** (Raksha Devi), **dwijendra.pandey@ma.iitr.ac.in** (Dwijendra N. Pandey)

[26] provide an excellent summary of the study of initial conditions for fractional differential equations with R-L derivatives.

Furthermore, many real-world processes and phenomena are defined by the influence of the state variable's past values, which gives rise to delays in differential equations. Caini et al. [15] studied the effect of delay in Mars to Earth communications through orbiters. Kyrychko et al.[27] studied the models for high-speed milling: which is a very common cutting process in the industry, a moving conveyor belt loaded with two oscillating connected masses and digital control contains handles by finite delay differential equation. There is a vast application of delay differential equations in population dynamics [16]. This motivates us to study the existence of solutions to FDDEs. Apart from this, to the best of our knowledge in the literature, there is no study on the strong and classical solutions of delay differential equations with the R-L fractional derivative. This study is concerned with the following fractional differential equation with finite delay in a Banach space X having uniformly convex dual X^*

$$\begin{cases} \frac{du(t)}{dt} + D_\alpha u(t) + Au(t) = f(t, u(t), u(t-\tau)), \quad t \in (0, T] \\ u(t) = \phi(t), \quad t \in [-\tau, 0) \\ {}_0I_t^{1-\alpha} u(t)|_{t=0} = \phi(0). \end{cases} \tag{1.1}$$

where D_α and ${}_0I_t^{1-\alpha}$ denote the R-L derivative and R-L integral of fractional order, $\alpha, 0 < \alpha < 1$ respectively and $\phi \in \mathcal{C}_0 := C([-\tau, 0]; X)$ i.e ϕ is a continuous $X-$ valued function on $[-\tau, 0]$. $-A$ generates an analytic semigroup of contractions in X, and $\tau > 0$ and $T < \infty$ are constants. Here, the considered equation (1.1) is also known as a generalized Basset equation with finite delay. In particular, if we consider $\alpha = \frac{1}{2}$ and $f(t, u(t), u(t-\tau)) \equiv f(t)$, equation (1.1) becomes Basset equation. In [11], Ashyralyev proved the well-posedness of the Basset equation in a Banach space X.

This analysis uses Rothe's methodology to demonstrate the existence of the unique solution of FDDE since it may also be used to determine a numerical solution. E. Rothe proposed Rothe's approach in 1931 to solve a second-order scalar parabolic initial value problem [1]. In [1], a parabolic boundary value problem of second order in two variables was converted in the system of ordinary differential equations to get approximate solutions. Later the method of lines was used to solve various partial differential equations of higher orders. In 1956, Lady$\hat{z}$enskaja [2] applied the Rothe approach to equations higher than the second order. Then, a number of other authors see, e.g., J. Ne$\hat{c}$as [3], J. Ka$\hat{c}$ur [4] applied this technique to demonstrate a few a priori estimates, based on which questions about existence and convergence are easily resolved. Rothe's approach was used by Bahuguna and Raghavendra [5] to demonstrate that nonlinear Schrodinger-type problems have a strong solution.

Several authors, including Agarwal and Bahuguna [7], Bahuguna and Raghvendra [6], S. Abbas et al. [8], Shruti [9], and Darshana et al. [10], used Rothe's approach to demonstrate the existence of the unique strong and weak solutions to integer order differential equations.

In 2019, motivated by Ashyralyav [11], Bahuguna and Anjali [12] proved the existence of the unique, strong solution to the following initial value problem for an FDE

$$\begin{cases} \frac{du(t)}{dt} + D_{0+}^\alpha u(t) + Au(t) = f(t), \quad t \in (0, T), \quad \alpha \in (0, 1) \\ u(0) = 0 \end{cases} \tag{1.2}$$

in a Banach space, X whose dual X^* is uniformly convex, and $-A$ generates an analytic semigroup of contractions in X. Here D_{0+}^α is the R-L derivative.

Rothe's approach was used by Chaoui et al. [14] to demonstrate the existence of the one and only solution as well as a few regularity findings for fractional diffusion integrodifferential with the fractional integral condition. This approach was then used by Bahuguna and Anjali [13] to prove the existence of the unique strong solution to the abstract fractional integrodifferential equations.

The article is structured as follows; section 2 contains some fundamental definitions, notations, and presumptions. Section 3, outlines the Rothe's methodology based on which some apriori estimates are proved. The major result is stated and proven in section 4, and the application of previous section is shown in section 5.

2. Basics and Assumptions

This section contains some basic definitions, preliminary information, and assumptions that will be utilised to demonstrate the main theorem.

Throughout the work, assume that X is a Banach space with uniformly convex dual X^* and $\|.\|$, $\|.\|_{X^*}$ are the norms of X and X^*. Here $\mathcal{C}_t := C([-\tau, t]; X)$ for $t \in [0, T]$ is the Banach space of all continuous functions from $[-\tau, t]$ into X endowed with the supremum norm

$$\|\phi\|_t := \sup_{-\tau \leq \zeta \leq t} \|\phi(\zeta)\|, \quad \phi \in \mathcal{C}_t$$

Definition 2.1. *[17] Let $I = (a, b)$ and $f(x) \in AC^n(a, b)$ and $n - 1 < \alpha < n, n \in \mathbb{N}_0$. The R-L derivative of function f of order α is defined as*

$$\mathcal{D}^{\alpha}_{a+} f(x) = \left(\frac{d}{dx}\right)^n I^{n-\alpha}_{a+} f(x) = \frac{1}{\Gamma(n-\alpha)} \left(\frac{d}{dx}\right)^n \int_a^x (x-t)^{n-\alpha-1} f(t) dt$$

where $AC^n[a, b] = \{f : [a, b] \to \mathbb{R} : f^{(n-1)} \in AC[a, b]\}$ and $I^{n-\alpha}_{a+}$ is known as the R-L integral of fractional order.

Definition 2.2. *[19] "Let $\Delta = \{z : \varphi_1 < \arg z < \varphi_2, \varphi_1 < 0 < \varphi_2\}$ and for $z \in \Delta$, let $T(z)$ be a bounded linear operator. The family $T(z)$, z $\in \Delta$ is an analytic semigroup in Δ if*

1. *$z \to T(z)$ is analytic in Δ.*
2. *$T(0) = I$ and $\lim_{\substack{z \to 0 \\ z \in \Delta}} T(z)x = x$ for every $x \in X$.*
3. *$T(z_1 + z_2) = T(z_1) T(z_2)$ for $z_1, z_2 \in \Delta$.*

A semigroup $T(t)$ will be called analytic in some sector Δ containing the nonnegative real axis.

Definition 2.3 (Chapter-3 [23])**.** *For a given 'Gauge function' (A continuous and strictly increasing function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ such that $\phi(0) = 0$ and $\lim_{t \to \infty} \phi(t) = \infty$) the mapping $J_\phi : X \to 2^{X^*}$ defined by*

$$J_\phi x := \{u^* \in X^* : \langle x, u^* \rangle = \|x\| \|u^*\|_{X^*}; \|u^*\|_{X^*} = \phi(\|x\|)\}$$

is called the 'duality mapping' with Gauge function ϕ. This mapping is single-valued as X^* is uniformly convex.

Definition 2.4. *[19] An operator A is "m-accretive" if*

$$\langle Au, J(u) \rangle \geq 0, \quad \forall \quad u \in D(A),$$

where J is the duality mapping and $R(I + \lambda A) = X$ for $\lambda > 0$.

Remark 2.5.

1. *If $-A$ generates C_0 semigroup, then A is m-accretive [19].*
2. *For a linear operator A, its domain of definition is given by*

$$D(A) := \left\{ v \in X : \lim_{t \to 0^+} \frac{T(t)v - v}{t} exists \right\}$$

where $T(t)$ is a semigroup of bounded linear operators.

Definition 2.6. *A function u is said to be strong solution of problem (1.1) if it satisfies following properties:*

1. $u \in C(I; X)$ *and* $u \in D(A)$.
2. D_α *exists and is continuous on* I, *where* $0 < \alpha < 1$.
3. u *satisfies given equation (1.1) a.e. on* I *with initial condition* $u(t) = \phi(t)$.

The form of strong solution [19] of problem (1.1) is given by

$$u(t) = T(t)\phi(0) - \int_0^t T(t-s)D_\alpha u(s)ds + \int_0^t T(t-s)f(s,u(s),u(s-\tau))ds \quad in \quad X$$

Lemma 2.7. *[13] If* $(p_n), (q_n)$ *and* (r_n) *are nonnegative sequences and*

$$p_n \le q_n + \sum_{0\le k<n} r_k p_k \text{ for } n \ge 0,$$

$$\& \text{ then } \quad p_n \le q_n + \sum_{0\le k<n} r_k q_k \exp\left(\sum_{k<j<n} r_j\right) \text{ for } n \ge 0.$$

This is known as 'Discrete Gronwall's' lemma.

Lemma 2.8. *[13] Let* $y(t)$ *is a non-negative continuous function on* $(0, T]$ *and* $g(t) > 0$ *be continuous, increasing function on* $[0, T]$. *If there are positive constants* A, B *such that*

$$y(t) \le Ag(t) + B\int_0^t \frac{y(s)}{(t-s)^\alpha}ds, \quad 0 \le t \le T,$$

then there exists a constant C *such that,*

$$y(t) \le Cg(t).$$

where $0 < \alpha < 1$ *and this reult also holds if* $y(t)$ *is piecewise continuous function.*

Lemma (3) is given to make us understand that the initial condition in the integral sense in equation (1.1) can also be written differently:

Lemma 2.9. *[18] Let* $\alpha \in (0, 1)$ *and* $r > 0, u : [0, r] \to \mathbb{R}$ *be a Lebesgue measurable function.*

1. *If* $\exists$ *a.e. a limit* $\lim_{t\to 0+}\left[t^{1-\alpha}u(t)\right] = e \in \mathbb{R}$, *then there also exists a limit*

$$\begin{aligned} {}_0I_t^{1-\alpha}u(t)\big|_{t=0} &:= \lim_{t\to 0+} \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{u(s)}{(t-s)^\alpha}ds = e\Gamma(\alpha) \\ &= \Gamma(\alpha)\lim_{t\to 0+}\left[t^{1-\alpha}u(t)\right]. \end{aligned}$$

2. *If* $\exists$ *a.e. a limit* $\lim_{t\to 0+} {}_0I_t^{1-\alpha}u(t) = e \in \mathbb{R}$, *and if* $\exists$ *the limit* $\lim_{t\to 0+}\left[t^{1-\alpha}u(t)\right]$, *then*

$$\lim_{t\to 0+}\left[t^{1-\alpha}u(t)\right] = \frac{e}{\Gamma(\alpha)} = \frac{1}{\Gamma(\alpha)}\lim_{t\to 0+} {}_0I_t^{1-\alpha}u(t)$$

Now we consider the following assumptions for proving the main result:

- **(B1)** $-A$ generates an analytic semigroup of contractions in X.

- **(**B2) The function f defined from $[0,T]\times X\times X\to X$ satisfies a local-Lipschitz like condition

$$\|f(t,u_1,\tilde{u_1})-f(s,u_2,\tilde{u_2}\|\leq L_f(r)[|t-s|+\|u_1-u_2\|+\|\tilde{u_1}-\tilde{u_2}\|],$$

for all u_i, $\tilde{u_i}\in B_r(X,\phi(0))$ and $L_f(r)$ is a non-decreasing function and, for $r>0$,

$$B_r(X,\phi(0))=\{u\in X:\|u-\phi(0)\|\leq r\}.$$

- **(**B3) $(I+A)^{-1}$ is compact.

3. Discretization and apriori estimates

We divide the interval $[0,T]$ into n-subintervals of lengths $h_n=\frac{T}{n}$ and at each of the division points $t_j^n=jh,\quad j=1,2,3,...,n$ in order to apply Rothe's approach of temporal discretization.

Discretization scheme for fractional derivative $D_\alpha u(t)$:
At $t=t_j^n$, the L_1 [20] approach is used to estimate the R-L fractional derivative:

$$D_\alpha u(t_j^n)\approx\frac{\phi(0)}{(t_j^n)^\alpha}+\frac{1}{\Gamma(2-\alpha)}\sum_{i=1}^{j}b_{j-i}\frac{\left(u_i^n-u_{i-1}^n\right)}{h_n}h_n^{1-\alpha},\quad\text{where}\quad j=1,2,3,..,n \tag{3.1}$$

$$=\frac{\phi(0)}{(t_j^n)^\alpha}+\sum_{i=1}^{j}\left(u_i^n-u_{i-1}^n\right)d_i^{j,n},\quad\text{where}\quad j=1,2,3..,n \tag{3.2}$$

where $b_k=(k+1)^{1-\alpha}-k^{1-\alpha}$ and $d_i^{j,n}=b_{j-i}\frac{h_n^{-\alpha}}{\Gamma(2-\alpha)}$.
We discretize $\frac{du}{dt}$ by forward difference scheme

$$\frac{du}{dt}=\frac{u_i^n-u_{i-1}^n}{h_n}$$

We let that $\phi(0)\in D(A)$. For a fix $r>0$, we choose r_0 such that

$$M=\|f(0,\phi(0),\phi(0))\|+\|Au_0\|+r_0d$$

where d is a positive constant. We consider $u_0^n=\phi(0)\geq 0\ \forall,\ n\in\mathbb{N}$ and now we establish a discretization of our problem in the direction of time-axis, at $t=t_j^n\quad j=1,2,3..,n$ problem (1.1) becomes

$$\begin{cases}\frac{u_j^n-u_{j-1}^n}{h_n}+Au_j^n+D_\alpha u_j^n=f_j^n,\\ u(t_j^n)=\phi(t_j^n)\\ {}_0I_{t_j^n}^{1-\alpha}u(t_j^n)|_{t_j^n\to 0}=\phi(0)\quad\forall j=1,2,..,n\quad\text{whenever}\quad n\to\infty\end{cases} \tag{3.3}$$

where, $f_j^n=f(t_j^n,u_j^n,u(t_j^n-\tau))$.

Lemma 3.1. *If conditions **(B1)-(B3)** hold then for $n\in N$, $j=1,2,...,n$,*

1. $\|u_j^n-u_0\|\leq C_1$ *where* $u_0=\phi(0)$.
2. $\|\delta u_j^n\|\leq C_2$.

where C_1, C_2 are positive constants, independent of j,h and n.

Proof. We may establish this claim by using two methods: first, discretizing the relevant fractional integral equation; and second, discretizing the fractional derivative in direct form. Here, the first half of the lemma is demonstrated by discretizing the fractional derivative, and the second part of the lemma is demonstrated by discretizing the fractional integral.

Firstly using definition of R-L derivative

$$D_\alpha u(t) = \frac{d}{dt}(I^{1-\alpha}u(t))$$

where $I^{1-\alpha}$ is R-L integral, then equation (3.3) becomes

$$\frac{1}{h_n}(u_j^n - u_{j-1}^n) + Au_j^n + \frac{1}{h_n}(I^{1-\alpha}u_j^n - I^{1-\alpha}u_{j-1}^n) = f_j^n \tag{3.4}$$

Putting $j = 1$ in (3.4) and subtracting Au_0 from both sides of obtained equation

$$\frac{1}{h_n}(u_1^n - u_0^n) + Au_1^n - Au_0^n + \frac{1}{h_n}(I^{1-\alpha}u_1^n - I^{1-\alpha}u_0^n) = f_1^n - Au_0^n \tag{3.5}$$

Here for simplicity, we write $u_j^n = u_j$

By applying $J(u_1 - u_0)$ on both sides and using the definition of J for gauge function $\phi(t) = t$, we obtain

$$\langle u_1 - u_0, J(u_1 - u_0)\rangle + h_n\langle A(u_1 - u_0), J(u_1 - u_0) + \langle I^{1-\alpha}(u_1 - u_0), J(u_1 - u_0)\rangle = h_n\langle f_1^n - Au_0, J(u_1 - u_0)\rangle \tag{3.6}$$

By using m-accretivity of A and the definition of duality map J

$$\|u_1 - u_0\| \le h_n \left[\|f(t_1, u_1, \tilde{u}_1) - f(0, u_0, \tilde{u}(0))\| + \|f(0, u_0, \tilde{u}(0))\| + \|Au_0\|\right]$$

considering $u(t-\tau) = \tilde{u}$ and using inequality $\langle I^{1-\alpha}(u_1 - u_0), (u_1 - u_0)\rangle \ge 0$ (see [14])

$$\|u_1 - u_0\| \le h_n \left[|t_1| + \|u_1 - \phi(0)\| + \|\tilde{u}_1 - \phi(0)\| + \|f(0, \phi(0), \phi(0))\| + \|Au_0\|\right] \le h_n M \le C$$

We will prove this result by induction; for this, we assume that

$$\|u_i^n - u_0\| \le C \quad \forall i < j. \tag{3.7}$$

Now we show that

$$\|u_j^n - u_0\| \le C \tag{3.8}$$

Subtracting Au_0 from both sides of (3.4)

$$\frac{1}{h_n}(u_j^n - u_{j-1}^n) + Au_j^n - Au_0 + \frac{1}{h_n}(I^{1-\alpha}u_j^n - I^{1-\alpha}u_{j-1}^n) = f_j^n - Au_0$$

$$\frac{1}{h_n}((u_j^n - u_0) - (u_{j-1}^n - u_0)) + A(u_j^n - u_0) + \frac{1}{h_n}(I^{1-\alpha}u_j^n - I^{1-\alpha}u_{j-1}^n) = f_j^n - Au_0$$

By applying $J(u_j^n - u_0)$ on both sides, we get

$$\begin{aligned}\langle u_j - u_0, J(u_j - u_0)\rangle + h_n\langle A(u_j - u_0), J(u_j - u_0)\rangle + \langle I^{1-\alpha}(u_j - u_{j-1}), J(u_j - u_0)\rangle \\ = \langle (u_{j-1} - u_0), J(u_j - u_0)\rangle + h_n\langle f_j - Au_0, J(u_j - u_0)\rangle\end{aligned}$$

Considering,

$$\langle I^{1-\alpha}(u_j - u_{j-1}), J(u_j - u_0)\rangle = \langle I^{1-\alpha}(u_j - u_{j-1} + u_0 - u_0), J(u_j - u_0)\rangle \ge 0$$

Accretivity of A and the previous condition implies that

$$\|u_j - u_0\| \leq \|u_{j-1} - u_0\| + Mh_n$$

Now using equation (3.7) we get required result i.e. $\|u_j - u_0\| \leq C_1 \quad \forall\ j = 1, 2, 3, ...$ Now, for proving inequality (2), we consider the discretized form of an equation (1.3) by discretizing the R-L integral of fractional order

$$\frac{1}{h_n}(u_j^n - u_{j-1}^n) + Au_j^n + \frac{\phi(0)}{(t_j^n)^\alpha} + \sum_{i=1}^{j}\left(u_i^n - u_{i-1}^n\right) d_i^{j,n} = f_j^n \tag{3.9}$$

where b_k and d_i are defined same as previously.
For $j = 1$

$$\frac{1}{h_n}(u_1^n - u_0^n) + Au_1^n + \frac{\phi(0)}{(t_j^n)^\alpha} + \frac{1}{h_n\Gamma(2-\alpha)}(u_1^n - u_0^n)h_n^{1-\alpha} = f_1^n$$

i.e.

$$\frac{1}{h_n}\left(1 + \frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)(u_1^n - u_0^n) + \frac{\phi(0)}{t^\alpha} + Au_1^n = f_1^n.$$

Due to accretivity of A and $(1 + \frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}) > 1$, we have

$$\|\delta u_1^n\| \leq \|f_1^n\|.$$

where $\delta u_j^n = \frac{u_j^n - u_{j-1}^n}{h_n}, \forall\ j = 1, 2, 3, ...$

For values of $j \geq 2$, subtracting the equation (3.3) $j-1$ from the equation (3.3) for j

$$\frac{1}{h_n}(u_j^n - u_{j-1}^n) + Au_j^n - Au_{j-1}^n + D_\alpha u_j^n - D_\alpha u_{j-1}^n = \frac{1}{h_n}(u_{j-1}^n - u_{j-2}^n) + f_j^n - f_{j-1}^n.$$

Using discretized value of D_α from equation (3.1), we obtain

$$\tfrac{1}{h_n}(u_j^n - u_{j-1}^n) + Au_j^n - Au_{j-1}^n + \tfrac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\delta u_j^n$$

$$= \delta u_{j-1}^n + \tfrac{1}{\Gamma(2-\alpha)}\textstyle\sum_{i=1}^{j-1} b_{j-1-i}\delta u_i^n h_n^{1-\alpha} - \tfrac{1}{\Gamma(2-\alpha)}\textstyle\sum_{i=1}^{j-1} b_{j-i}\delta u_i^n h_n^{1-\alpha} + f_j^n - f_{j-1}^n.$$

Hence

$$\left(1 + \tfrac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\delta u_j^n + Au_j^n - Au_{j-1}^n$$
$$= \left(1 + (b_0 - b_1)\tfrac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\delta u_{j-1}^n + f_j^n - f_{j-1}^n + \tfrac{1}{\Gamma(2-\alpha)}\textstyle\sum_{i=1}^{j-2}\left(b_{j-1-i} - b_{j-i}\right)\delta u_i^n h_n^{1-\alpha}$$

Since $(1 + \frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}) > 1$,

$$\delta u_j^n + Au_j^n - Au_{j-1}^n \leq \left(1 + (b_0 - b_1)\frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\delta u_{j-1}^n + f_j^n - f_{j-1}^n + \frac{1}{\Gamma(2-\alpha)}\sum_{i=1}^{j-2}\left(b_{j-1-i} - b_{j-i}\right)\delta u_i^n h_n^{1-\alpha} \tag{3.10}$$

Taking the duality product with $J(u_j^n - u_{j-1}^n)$ owing Accretivity of A and hypotheses $(B1), (B2), (B3)$ and Lemma $3(i)$ we have

$$\left\|\delta u_j^n\right\| \leq \left(1 + (b_0 - b_1)\frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\left\|\delta u_{j-1}^n\right\| + \left\|f_j^n - f_{j-1}^n\right\| + \frac{1}{\Gamma(2-\alpha)}\sum_{i=1}^{j-2}\left(b_{j-1-i} - b_{j-i}\right)\left\|\delta u_i^n\right\| h_n^{1-\alpha} \tag{3.11}$$

Using Lipschitz condition on f

$$\|f(t_j^n, u_j^n, \tilde{u}_{j-1}^n) - f(t_{j-1}^n, u_{j-1}^n, \tilde{u}_{j-1}^n\| \le L_f(r)[|t_j^n - t_{j-1}^n| + \|u_j^n - u_{j-1}^n\| + \|\tilde{u}_j^n - \tilde{u}_{j-1}^n\|],$$

$$\|f(t_j^n, u_j^n, \tilde{u}_{j-1}^n) - f(t_{j-1}^n, u_{j-1}^n, \tilde{u}_{j-1}^n\| \le C\left[h_n + 3h_n \delta u_j^n\right]$$

Now equation (3.10) becomes,

$$\|\delta u_j^n\| \le \left(1 + (b_0 - b_1)\frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\|\delta u_{j-1}^n\| + C\left[h_n + 3h_n\delta u_j^n\right] + \frac{1}{\Gamma(2-\alpha)}\sum_{i=1}^{j-2}(b_{j-1-i} - b_{j-i})\|\delta u_i^n\| h_n^{1-\alpha} \tag{3.12}$$

Rearranging equation (3.11)

$$\|\delta u_j^n\| \le D_1\left(1 + (b_0 - b_1)\frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}\right)\|\delta u_{j-1}^n\| + D_2 h_n + \frac{D_3}{\Gamma(2-\alpha)}\sum_{i=1}^{j-2}(b_{j-1-i} - b_{j-i})\|\delta u_i^n\| h_n^{1-\alpha}$$

where D_1, D_2, D_3 are positive constants depending on C and h_n.
Now using Discrete Gronwall's inequality, we will get

$$\|\delta u_j^n\| \le C_2.$$

■

Lemma 3.2. *If conditions **(B1)-(B3)** hold then for $n \in N$, $j = 1, 2, .., n$,*

$$\|D^\alpha u_j^n\| \le C_3, \quad \|Au_j^n\| \le C_4$$

where C_3, C_4 are positive constants, independent of j,h and n.

Proof. Since

$$D_\alpha u_j^n = \frac{\phi(0)}{t^\alpha} + \frac{1}{\Gamma(2-\alpha)}\sum_{i=1}^{j} b_{j-i}\delta u_i^n h_n^{1-\alpha}$$

Using Lemma1(ii) and

$$b_{j-i} = (j-i+1)^{1-\alpha} - (j-i)^{1-\alpha} = (j-i)^{1-\alpha}\left[\left(1 + \frac{1}{(j-i)}\right)^{1-\alpha} - 1\right]$$

Using binomial expansion

$$b_{j-i} = (j-i)^{1-\alpha}\left(\frac{1-\alpha}{j-i} - \frac{\alpha(1-\alpha)}{2!(j-i)^2} + \frac{\alpha(1-\alpha)(1+\alpha)}{3!(j-i)^3} + ...\right)$$

since $0 < \alpha < 1$, so

$$b_{j-i} = \frac{1-\alpha}{(j-i)^\alpha} - \frac{\alpha(1-\alpha)}{2!(j-i)^{1+\alpha}} + \frac{\alpha(1-\alpha)(1+\alpha)}{3!(j-i)^{2+\alpha}} + ...$$

using the fact that $0 < \alpha < 1$ so $\frac{1}{(j-i)^n} < \frac{1}{(j-i)^\alpha}$ for each $n \in N$ and here $i < j$.
We select d_α as the maximum of all the numerators in the previous expression of b_{j-i}.

$$b_{j-i} \le d_\alpha\left(\frac{1}{j^\alpha} + \frac{1}{(j-1)^\alpha} + \frac{1}{(j-2)^\alpha} + \cdots\right)$$

Now, we can express b_{j-i} by the following inequality

$$b_{j-i} \leq d_\alpha \sum_{i=1}^{j} \frac{1}{(j-i+1)^\alpha}$$

Hence $D_\alpha u_j^n$ becomes

$$\|D_\alpha u_j^n\| \leq d_\alpha C_2 \sum_{i=1}^{j} \frac{h_n}{[(j-i+1)h_n]^\alpha}$$

Now proceeding similarly as [Lemma 7 [13]], we will get required result $\|D_\alpha u_j^n\| \leq C_3$.

Now using triangle inequality,

$$\|Au_j^n\| \leq \|f_j^n\| + \|D_\alpha u_j^n\| + \|\delta u_j^n\|$$
$$\|Au_j^n\| \leq \|f_j^n - f_0^n\| + \|f_0^n\| + \|D_\alpha u_j^n\| + \|\delta u_j^n\|$$

Using lemma(3.1), and the first part of the Lemma and local-Lipschitz condition on f, we get the required result

$$\|Au_j^n\| \leq C_4.$$

■

We next define a sequence of step functions. $X^n : [-h_n, T] \to D(A)$ by

$$X^n(t) = \begin{cases} \phi(0), & t \in [-h_n, 0], \\ u_j^n, & t \in \left[t_{j-1}^n, t_j^n\right]. \end{cases}$$

Further we introduce sequence U^n of polygonal functions from $[-\tau, T] \to D(A)$, given by

$$U^n(t) = \begin{cases} \phi(t), & t \in [-\tau, 0], \\ u_{j-1}^n + \frac{t-t_{j-1}^n}{h_n}\left(u_j^n - u_{j-1}^n\right), & t \in \left(t_{j-1}^n, t_j^n\right]. \end{cases}$$

Remark 3.3. *From Lemma (3.1), we observe that the function $U^n(t)$ are Lipschitz continuous on $[-\tau, T]$ with a uniform Lipschitz constant. The sequence $U^n(t) - X^n(t) \to 0$ in X as $n \to \infty$ on $(0, T]$. Furthermore the sequences $X^n(t)$ and $U^n(t)$ are uniformly bounded in X. By definition of $X^n(t)$, the boundedness of this sequence is clear using lemma(3.1).*

To prove boundedness of Rothe's sequence $U^n(t)$ in X.
Since,

$$0 \leq \frac{t - t_{j-1}^n}{h_n} \leq 1 \quad in \quad [t_j^n - t_{j-1}^n],$$

in order that by definition of $U^n(t)$, for an arbitrary $t \in [t_j^n, t_{j-1}^n]$,

$$\begin{aligned} \|U^n(t)\| &= \left\| u_{j-1}^n \left(1 - \frac{t - t_{j-1}^n}{h_n}\right) + u_j^n \frac{t - t_{j-1}^n}{h_n} \right\| \\ &\leq \left\| u_{j-1}^n \left(1 - \frac{t - t_{j-1}^n}{h_n}\right) \right\| + \left\| u_j^n \frac{t - t_{j-1}^n}{h_n} \right\| \\ &\leq \left(1 - \frac{t - t_{j-1}^n}{h_n}\right) C_1 + \frac{t - t_{j-1}^n}{h_n} C_1 \leq C_1. \end{aligned}$$

Now, we define a sequence of step functions $\tilde{D}_\alpha U^n : [0, T] \to X$ by

$$\tilde{D}_\alpha U^n(t) = \begin{cases} 0, & t = 0, \\ \frac{\phi(0)}{t^\alpha} + \sum_{i=1}^{j} a_{j-i} \frac{u_i^n - u_{i-1}^n}{h_n} \frac{h_n^{1-\alpha}}{\Gamma(2-\alpha)}, & t \in \left(t_{j-1}^n, t_j^n\right]. \end{cases}$$

Lemma 3.4. $\|\tilde{D}_\alpha U^n(t) - D_\alpha U^n(t)\| \to 0$ *as* $n \to \infty$ *uniformly on* $(0,T)$.

Proof. See the proof of Lemma 8 in [13] and the calculation for $D_\alpha U^n(t)$ in [12]. ■

Lemma 3.5. *There exists a subsequence* $\{U^{n_k}\}$ *of* $\{U^n\}$ *and a function* $u : [0,T] \to D(A)$ *such that* $U^{n_k} \to u$ *in* $C([0,T],X)$ *and* $AX^{n_k}(t) \rightharpoonup Au(t)$ *uniformly in* X *as* $n \to \infty$*. Furthermore,* $Au(t)$ *is weakly continuous.*

Proof. [21] For proving this define $Y^n = (I+A)X^n$. $\{Y^n(t)\}$ is uniformly bounded by using Remark 1. Since $X^n = (I+A)^{-1}Y^n$, assumption $(B2)$ implies that a subsequence $\{X^{n_k}(t)\}$ of $\{X^n(t)\}$ converges strongly in X (using compact criterion). Let $u(t)$ be the limit of $X^{n_k}(t)$. $U^m(t) \to u(t)$ as $n \to \infty$ by using Remark 1.The sequence $U^{n_k} \in C(I,X)$ is equicontinuous and for $t \in I$, $\{U^{n_k}(t)\}$ is relatively compact in X. Therefore, the Ascoli-Arzela theorem implies that $U^{n_k} \to u$ in $C(I,X)$ i.e. $U^n(t)$ converges to $u(t)$ uniformly on every compact subinterval. Since each U^{n_k} is Lipschitz continuous with uniform Lipschitz constant M_0, so

$$\|u(t_1) - u(t_2)\| = lim_{n\to\infty}\|U^n(t_1) - U^n(t_2)\| \le M_0|t_1 - t_2|$$

and

$$\|u(t)\| = lim_{n\to\infty}\|U^n(t)\| \le C, \quad u \in C([0,T],X).$$

u is Lipschitz continuous.
We use Lemma 4 the uniform boundedness of $\{AX^n\}$. For each $t \in I$, $X^{n_k} \to u(t)$ as $n \to \infty$, uniformly on I. Since X^* is uniformly convex, using Lemma (2.5) from [22] we can say that $u(t) \in D(A)$ for $t \in I$ and $AX^{n_k} \rightharpoonup Au(t)$ uniformly on I as $n \to \infty$. To prove the weak continuity of $Au(t)$, let $t_k \in I$ be such that $t_k \to t$ as $t_k \to t$ as $k \to \infty$. As u is Lipschitz continuous so $u(t_k) \to u(t)$ as $k \to \infty$. Since $\{Au(t_k)\}$ is uniformly bounded so $Au(t_k) \rightharpoonup Au(t)$ as $k \to \infty$ ([22], using Lemma (2.5)). ■

Lemma 3.6. *There exists a subsequence* $\{U^{n_k}\}$ *of* $\{U^n\}$ *such that* $\frac{d^- U^{n_k}}{dt} \rightharpoonup \frac{du}{dt}$ *and* $D_\alpha U^{n_k} \rightharpoonup D_\alpha u$ *in* $L^2([0,T],X)$*, as* $n \to \infty$.

Proof. The proof of this Lemma is based on Lemma (10) from [13].
Since $D_\alpha U^n(t)$ is uniformly bounded in $L^2([0,T],X)$. Every bounded sequence in L^2 has a weakly convergent subsequence so there exist a subsequence $D_\alpha U^{n_k}(t)$ and a function $\zeta \in L^2([0,T],X)$ such that

$$D_\alpha U^{n_k}(t) \rightharpoonup \zeta \in L^2([0,T],X).$$

Define $Y^n : [0,T] \to X$ by

$$Y^n(t) = I^{1-\alpha}U^n(t) = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{U^n(s)}{(t-s)^\alpha}ds$$

For $t \in (0, t_1^n)$
Using Rothe's sequence $U^n(t) = u_0^n + \frac{(t-t_0^n)}{h_n}(u_1^n - u_0^n) = \phi(0) + \frac{t}{h_n}(u_1^n - \phi(0))$
As ϕ is continuous function on closed interval $[-\tau, 0]$ so $|\phi(0)| \le a$, where a is positive constant. Hence

$$Y^n(t) = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{\phi(0) + \frac{s(u_1^n - \phi(0))}{h_n}}{(t-s)^\alpha}ds$$

$$= \frac{\phi(0)}{\Gamma(1-\alpha)}\int_0^t \frac{1}{(t-s)^\alpha}ds + \frac{(u_1^n - \phi(0))}{h_n\Gamma(1-\alpha)}\int_0^t \frac{s}{(t-s)^\alpha}ds$$

$$= \frac{-t^{1-\alpha}}{\Gamma(2-\alpha)}\phi(0) + \frac{t^{2-\alpha}}{h_n}\frac{(\delta u_1^n - \phi(0))}{\Gamma(3-\alpha)} \tag{3.13}$$

Since δu_1^n, $\phi(0)$ are bounded and $Y^n(0_+) = 0$ for every n.
For $x^* \in X^*$, we can see

$$\langle Y^{n_k}(t), x^* \rangle = \int_0^t \langle D_\alpha U^{n_k}(s), x^* \rangle ds.$$

Since $U^{n_k} \to u \in C([0,T], X)$, $Y^{n_k}(t) \to \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{u(s)}{(t-s)^\alpha} ds$ as $k \to \infty$. Thus passing the limit $k \to \infty$, we get

$$\left\langle \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{u(s)}{(t-s)^\alpha} ds, x^* \right\rangle = \int_0^t \langle \zeta(s), x^* \rangle ds$$

Hence

$$\zeta(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{u(s)}{(t-s)^\alpha} ds.$$

Similarly, we can show $\frac{d^- U^{n_k}}{dt} \rightharpoonup \frac{du}{dt}$ in $L^2([0,T], X)$. ■

Remark 3.7.

$$\tilde{D}_\alpha U^n \rightharpoonup D_\alpha u \in L^2([0,T], X), as \quad n \to \infty.$$

Lemma 3.8. *$Au(t)$ is Bochner integrable on $[0,T]$.*

Proof. The proof of Bochner's integrability of $Au(t)$ is based on the Lemma 4.6 in [22]. ■

Remark 3.9. *Equation (1.3) can be written as*

$$\frac{d^- U^n(t)}{dt} + AX^n(t) + \tilde{D}_\alpha U^n(t) = f^n(t), t \in [0,T] \tag{3.14}$$

4. Main Results

Before proving the main theorem, let us recall well-known results from real analysis, which can be found in any standard analysis book.

Remark 4.1. ***Lebesgue bounded convergence theorem:*** *Let g_n be a sequence of measurable functions on a set of finite measure ω. Suppose g_n is uniformly pointwise bounded on ω, that is , there is a number $m \geq 0$ for which $|g_n| \leq m$ for all n. If $g_n \to g$ pointwise on ω, then $\lim_{n\to\infty} \int_\omega g_n = \int_\omega g$.*

Theorem 4.2. *Let $-A$ generates an analytic semigroup of contractions in X such that $(B1) - (B3)$ hold. Then the fractional differential equation with finite delay (1.1) has a unique strong and classical solution.*

Proof. **Existence of strong solution-** Integrating equation (3.14) from 0 to t, we obtain

$$U^n(t) - \phi(0) + \int_0^t AX^n(s)ds + \int_0^t \tilde{D}_\alpha U^n(s)ds = \int_0^t f^n(s)ds \quad in \quad X$$

For each $\psi \in X^*$, we get

$$\langle U^n(t), \psi \rangle - \langle \phi(0), \psi \rangle + \int_0^t \langle AX^n(s), \psi \rangle ds + \int_0^t \langle \tilde{D}_\alpha U^n(s), \psi \rangle ds = \int_0^t \langle f^n(s), \psi \rangle ds \quad in \quad X^*$$

Rewriting the above equation for the subsequence n_k of n, we have

$$\langle U^{n_k}(t), \psi \rangle - \langle \phi(0), \psi \rangle + \int_0^t \langle AX^{n_k}(s), \psi \rangle ds + \int_0^t \langle \tilde{D}_\alpha U^{n_k}(s), \psi \rangle ds = \int_0^t \langle f^{n_k}(s), \psi \rangle ds$$

Owing to Lebesgue bounded convergence theorem, apriori estimates and remarks, as $k \to \infty$

$$\langle u(t), \psi\rangle - \langle \phi(0), \psi\rangle + \int_0^t \langle Au(s), \psi\rangle ds + \int_0^t \langle D_\alpha u(s), \psi\rangle ds = \int_0^t \langle f(s, u(s), u(s-\tau)), \psi\rangle ds \tag{4.1}$$

Using $\int_0^t \langle D_\alpha u(s), \psi\rangle ds = \langle I_{0+}^{1-\alpha} u(t), \psi\rangle$ in equation (4.1), we obtain

$$\langle u(t) + I_{0+}^{1-\alpha} u(t), \psi\rangle - \langle \phi(0), \psi\rangle = -\int_0^t \langle Au(s), \psi\rangle ds + \int_0^t \langle f(s, u(s), u(s-\tau)), \psi\rangle ds$$

The continuous differentiability of $\langle u(t) + I_{0+}^{1-\alpha} u(t), \psi\rangle$ is provided by the continuity of the integrands on the RHS. Now, owing to Bochner integrability of $Au(t)$, the strong derivative of $u(t) + I_{0+}^{1-\alpha} u(t)$ exists a.e. on $[0, T]$, implies that

$$\frac{du(t)}{dt} + D_\alpha u(t) + Au(t) = f(t, u(t), u(t-\tau)), \quad for \quad a.e. \quad t \in (0, T]$$

Clearly, u is Lipschitz continuous on $[0, T]$ and $u(t) \in D(A)$ for $t \in [0, T]$.

Uniqueness of strong solution- If there are two strong solutions, u_1 and u_2, they will both satisfy the differential equation (1.1), then

$$\begin{cases} \frac{dy(t)}{dt} + D_\alpha y(t) + Ay(t) = f(t, u_1(t), u_1(t-\tau)) - f(t, u_2(t), u_2(t-\tau)) \quad t > 0 \\ y(t) = 0, \quad t \in [-\tau, 0) \\ {}_{0t}I^{1-\alpha} y(t)|_{t=0} = 0; \end{cases} \tag{4.2}$$

where $y(t) = u_1(t) - u_2(t)$ and assume $H(t) = f(t, u_1(t), u_1(t-\tau)) - f(t, u_2(t), u_2(t-\tau))$

Now equation (4.2) becomes

$$\frac{dy(t)}{dt} + D_\alpha y(t) + Ay(t) = H(t)$$

$S(t)$ is the semigroup generated by $-A$. Then

$$y(t) = -\int_0^t S(t-s) D_\alpha y(s) ds + \int_0^t S(t-s) H(s) ds \tag{4.3}$$

Since $y(t)$ is differentiable almost everywhere and $-A$ generates an analytic semigroup, differentiating equation (4.3), we get

$$y'(t) = -D_\alpha y(t) + H(t) + \int_0^t AS(t-s) D_\alpha y(s) ds - \int_0^t AS(t-s) H(s) ds \tag{4.4}$$

For homogeneous initial condition, [17], Caputo fractional derivative and R-L derivative becomes the same. Hence for $0 < \alpha < 1$

$$D_\alpha y(t) = \frac{1}{\Gamma(1-\alpha)} \frac{d}{dt} \int_0^t \frac{y(s)}{(t-s)^\alpha} ds = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{y'(s)}{(t-s)^\alpha} ds \tag{4.5}$$

Using (4.4) in (4.5), we obtain

$$\begin{cases} D_\alpha y(t) = & \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{(-D_\alpha y(s) + H(s))}{(t-s)^\alpha} ds \\ & + \frac{1}{\Gamma(1-\alpha)} \int_0^t \int_s^t \frac{1}{(t-r)^\alpha} AS(r-s) dr D_\alpha y(s) ds \\ & - \frac{1}{\Gamma(1-\alpha)} \int_0^t \int_s^t \frac{1}{(t-r)^\alpha} AS(r-s) dr H(s) ds. \end{cases} \tag{4.6}$$

We can easily show the following estimate, using properties of analytic semigroup [Chapter-2 [19]] and using estimate derived in [11].

$$\left\| \frac{1}{\Gamma(1-\alpha)} \int_r^t \frac{1}{(t-r)^\alpha} AS(r-s)dr \right\| \leq \frac{C}{(t-r)^\alpha} \tag{4.7}$$

where C is a positive constant. Hence equation (4.6) can be written as

$$\|D_\alpha y(t)\| \leq C \int_0^t \frac{\|D_\alpha y(s)\|}{(t-s)^\alpha} ds + C \int_0^t \frac{\|H(s)\|}{(t-s)^\alpha} ds$$

Now we find the estimate of $\|H(t)\|$ using hypothesis $(B2)$

$$\|H(t)\| = \|f(t, u_1(t), u_1(t-\tau)) - f(t, u_2(t), u_2(t-\tau))\| \tag{4.8}$$

$$\leq L_f \left[\|u_1(t) - u_2(t)\| + \|\tilde{u}_1(t) - \tilde{u}_2(t)\|\right] = L_f \left[\|y(t)\| + \|\tilde{y}(t)\|\right] \tag{4.9}$$

For second term in R.H.S, when $-\tau \leq t - \tau \leq 0$, then $y(t-\tau) = 0$ using condition (4.2) and when $0 \leq t - \tau \leq T - \tau \leq T$ then we can write $\|y(t-\tau)\| \leq \|y(t)\|_t$
Hence, equation(4.9) can be written as

$$\|H(t)\| \leq g_f \|y\|_t \tag{4.10}$$

where g_f is dependent on L_f and $\|y(s)\|_t = sup_{0\leq s\leq t}\|y(s)\|$
Now following similarly as in [13], we get

$$\|y\|_t^2 \leq 2C \int_0^t \|y(s)\|_s^2 ds$$

using Grownwall's inequality, we have

$$\|y(t)\|_t = 0 \quad t \in [0, T]$$

This implies $y(t) = u_1(t) - u_2(t) = 0$. This proves the uniqueness of strong solution.

Existence, uniqueness of classical solution:
The function $\bar{f}(t) = f(t, u(t), u(t-\tau))$ is local-Lipschitz continuous and X is reflexive Banach space so $\bar{f}(t)$ is differentiable a.e. and $\bar{f}'$ is in $L^1((0,T), X)$ by using Generalized Rademacher's theorem in reflexive Banach space.
Given that $-A$ is an analytical semigroup in X, the corollary 3.3 in [19] states that the existence of the unique strong solution u entails the existence of the unique classical solution to (1.1). ■

5. Application

The conclusion established in the previous section is applied in this section, and the Rothe's temporal discretization is used to determine a numerical solution.

Example 5.1. *Consider the following fractional differential equation with finite delay in $X = L^2(\Omega)$, where Ω is a bounded domain in $\mathbb{R}^n$ with smooth boundary*

$$\frac{\partial u(t,x)}{\partial t} + \frac{\partial^\alpha u(t,x)}{\partial t^\alpha} + A(x,D)u(t,x) = F(t,x,u(t,x),u(t-\tau,x)) \quad t \in (0,T] \tag{5.1}$$

with conditions

$$u(t,x) = 0, \quad x \in \partial\Omega, \quad t \in (0,T] \tag{5.2}$$

$$I^{1-\alpha}u(0)|_{t=0} = \phi(x), \quad in \quad \Omega \tag{5.3}$$

$$u(t,x) = \phi(t,x), \quad t \in [-\tau,0) \quad x \in \Omega \tag{5.4}$$

where $0 < \alpha < 1$ *and* $\frac{\partial^\alpha}{\partial t^\alpha}$ *is R-L derivative of order* α *and* $x \in \Omega$.
$\phi \in \mathcal{C}_0 := C([-\tau,0];X)$*, so there are many choices for* $\phi(t,x)$*. Here we consider* $\phi(t,x) = exp(x)t^{2+\alpha}$,$0 < \alpha < 1$,$x \in \Omega$ *and* $t \in [-\tau,0]$*, where* $\tau > 0$*. It is easy to check that* $\phi(t,x)$ *is a continuous function.*

Let $A(x,D)u = \sum_{|\beta|\le 2m} a_\beta(x)D^\beta u$ *be a strongly elliptic operator in* Ω*. The coefficients* $a_\beta(x)$ *of* $A(x,D)$ *are assumed to be smooth enough. Let* $A = A(x,D)$ *be a strongly elliptic operator of order* $2m$ *on a bounded domain* Ω *in* $\mathbb{R}^n$.
Set $D(A_2) = W^{2m,2}(\Omega) \cap W_0^{m,2}(\Omega)$ *and* $A_2(u) = A(x,D)u \quad for \quad u \in D(A_2)$*. By using theorem [[19], Chapter 7], we say that the operator* $-A_2$ *generates an analytic semigroup of contractions on* $X = L^2(\Omega)$ *and then by using theorem* $(I + A_2)^{-1}$ *is a compact operator.*
Let $F(t,x,u(t,x),u(t-\tau,x)) = f(t,x) + g(u(t,x),u(t-\tau,x))$*, we assume that* f *and* g *satisfy local-Lipschitz like condition such that*

$$\|f(t_1,x_1) - f(t_2,x_2)\| \le q[|t_1 - t_2| + |x_1 - x_2|]$$

$$\|g(u(t_1),\tilde{u}(t_1)) - g(u(t_2),\tilde{u}(t_2))\| \le L_g(r)[\|u(t_1) - u(t_2)\| + \|\tilde{u}(t_1) - \tilde{u}(t_2)\|]$$

where $\tilde{u} = u(t-\tau)$*,* q *is a positive constant and* $L_g(r)$ *is a non-decreasing function.*
In particular, we may consider:

1. $f(t,x) = exp(-t)\sqrt(t) + x$ *satisfies Lipschitz condition for* $t \in (0,T]$ *and* $x \in X$.
2. $g(u(t),\tilde{u}(t)) = sin(u(t))\,(\tilde{u}(t)) + sin(\tilde{u}(t)))$ *and clearly,* g *satisfies local-Lipschitz like condition for* $L_g(r) = 2(r+a)$*,* $r > 0$ *and* $|\phi(0)| \le a$ *as* $u(t),\tilde{u}(t) \in B_r(X,\phi(0))$ *, which is defined in section* 2.

Now, we can rewrite equations $(5.1) - (5.4)$ *as*

$$\frac{du(t)}{dt} + D_\alpha u(t) + A_2 u(t) = F(t,u(t),u(t-\tau)) \quad t \in (0,T] \tag{5.5}$$

with conditions

$$\begin{cases} u(t) = \phi(t), & t \in [-\tau,0) \\ {}_{0t}I^{1-\alpha}u(t)|_{t=0} = \phi(0) \end{cases} \tag{5.6}$$

The analysis met all of theorem 4.2's assumptions, hence the existence of the unique strong and classical solutions to $(5.5) - (5.6)$ *implies the existence of the unique strong and classical solutions to* $(5.1) - (5.4)$.

Numerical Example: If we know the existence of a solution for FDDEs then we can go for the numerical solution of FDDEs.

Example 5.2. *Using Rothe's methodology, we find out numerical solution of following FDDE.*

$$\frac{\partial}{\partial t}u(t,x) + \frac{\partial^\alpha}{\partial t^\alpha}u(t,x) - \frac{\partial^2}{\partial x^2}u(t,x) = -u(x,t-s) + h(x,t) \quad t \in (0,T] \quad x \in [0,2] \tag{5.7}$$

with conditions

$$\begin{cases} u(x,t) = \phi(x,t), \quad t \in [-s,0], x \in [0,2] \\ u(0,t) = u(2,t) = 0 \quad t \in [0,1] \end{cases} \tag{5.8}$$

assume $\phi(x,t) = t^2(2x - x^2) = u(x,t)$ *for this exact solution* $h(x,t) = \frac{\Gamma(3)}{\Gamma(3-\alpha)}(2x-x^2)t^{2-\alpha} + 2t(t+2x-x^2) + x(2-x)(t-s)^2$.

As we have checked the existence of the unique classical solution for the previous example similarly we can do this here by converting the differential equation into abstract form.

Let $X = L^2([0,2],\mathbb{R})$ *and the operator* A *defined on* X *by* $Au = -u''$ *with domain*

$$D(A) = \{u \in L^2([0,2]); u', u'' \in L^2[0,2], u(0) = u(2) = 0\}$$

It is well known that A *generates an analytic semigroup on* $L^2[0,2]$ *and it satisfies all the assumptions of the theorem (4.2). So there exist classical solution for considered example by using theorem(4.2).*

Since the classical solution for a given fractional differential equation is not always known, in this case we apply Rothe's approach to get the numerical solution for a given FDDEs.

To solve numerically for the delay s*=0.5 and order of the fractional derivative* $alpha$*=0.2, we follow a few steps.*

- *The article uses a time-stepping strategy to solve the delay problem* $(5.7) - (5.8)$.
- *The relevant difference quotients from the theory are used to replace the time derivatives.*
- *Now, at* $t = t_n$*, build a discretization of our problem* (5.7) *along the time axis.*
- *We get a set of differential equations in the variable* x.
- *There are other additional ways [31–34] to solve the system of ordinary differential equations, however this article uses further discretization in space variables.*

Figures 1 *and* 2 *are solution surfaces for numerical and analytic solutions and figure* 3 *shows the parabolic behavior of the solution for different values of the time. If we increase the time steps, we get a numerical solution approaching the classical solution.*

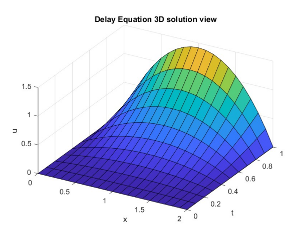

Figure 1: Numerical Solution surface for dx=dt=1/10

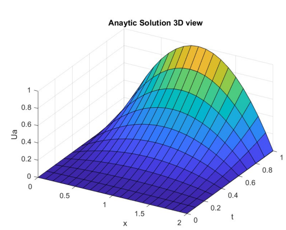

Figure 2: Analytic Solution surface for dx=dt=1/10

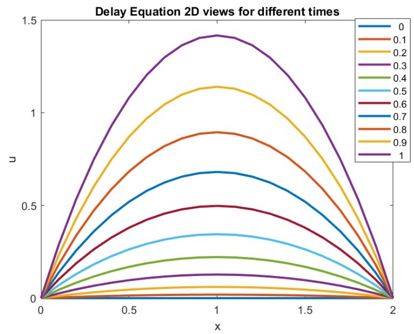

Figure 3: Numerical solution view for different values of time

6. Conclusion

In this study, we present some theoretical conclusions and numerical solution concerning the existence of the unique classical solution to the initial boundary value problem of fractional order differential equations with finite delay. The considered problem is a generalisation of the Basset problem with a finite delay, which occurs in fluid dynamics when an unstable particle accelerates in a viscous fluid due to the force of gravity. To demonstrate the existence of the unique considered problem, we calculated certain apriori results using semigroup theory and some hypotheses on the source function.

7. Acknowledgement

The author is thankful to the Ministry of Human Resource Development, Government of India for its financial support.

References

[1] E. ROTHE, Zweidimensional parabolische Randwertaufgaben als Grenzfall eindimensionaler Randwertaufgaben, *Math. Ann., **102**(1930), 650-670 (in German)*

[2] LADYẑENSKAJA, O. A, On the Solutions of Nonstationary operator Equations, *Mat. Sbornik **39**(1956) (In Russian).*

[3] NEĉAS, J, Application of Rothe's Method to abstract Parabolic Equations, *Czech. Math. J. 24(1974), 496-500.*

[4] KAĉUR, J, Application of Rothe's method to Nonlinear Equations, *Nonlinear Evolution Equations and Potential Theory, 89–93.*

[5] D. BAHUGUNA, V. RAGHVENDRA, Application of Rothe's method to nonlinear Schrodinger type equations, *Applicable Analysis, **31**(1994), 149-160.*

[6] D. BAHUGUNA, V. RAGHVENDRA, Application of Rothe's method to nonlinear integrodifferential equations in Hilbert spaces, *Nonlinear Analysis: Theory, Methods and Applications, **23**(1)(1994), 75–81.*

[7] S. AGARWAL AND D. BAHUGUNA, *Method of semidiscretization in time to nonlinear retarded differential equations with nonlocal history conditions, IJMMS 2004:**37**, 1943-1956.*

[8] D. BAHUGUNA, S ABBAS, J DABAS, *Partial functional differential equation with an integral condition and applications to population dynamics, Nonlinear Analysis: Theory, Methods, and Applications Volume **69**, Issue 8, Pages 2623-2635*

[9] SHRUTI A. DUBEY, The method of lines applied to nonlinear nonlocal functional differential equations, *J. Math. Anal. Appl., **376**(2011), 275-281.*

[10] DARSHANA DEVI, DURANTA, RAJIB HALOI, Rothe's Method For Solving Semi-linear Differential Equations With Deviating Arguments*; Electronic Journal of Differential Equations, 2020(2020), No. 120, 1-10.*

[11] A. ASHYRALYEV, Well-posedness of the Basset problem in space of smooth functions, *Applied Mathematics Letters, **24**(2011), 1176-1180.*

[12] D. BAHUGUNA AND ANJALI JAISWAL, Application of Rothe's Method to fractional differential equations, *Malaya Journal of Matemik, **7**(3)(2019), 399-407.*

[13] D. BAHUGUNA AND ANJALI JAISWAL, Rothe time discretization method for fractional integro-differential equations, *International Journal for Computational Methods in Engineering Science and Mechanics,* **20***(6)(2019), 540–547.*

[14] ABDERRAZEK CHAOUI AND AHMED HALLACI, On the solution of a fractional diffusion integrodifferential equation with Rothe time discretization, *Numerical Functional Analysis and Optimization,* **39***(6)(2018), 643-654.*

[15] *C. Caini, R. Firrincieli1, T. de Cola, I. Bisio3, M. Cello and G. Acar,* Mars to Earth communications through orbiters: Delay-Tolerant/Disruption-Tolerant Networkin performance analysis, *Int. J. Satell. Commun. Network, **32**(2014), 127–140.*

[16] YANG KUANG, *Delay Differential Equations: With Application in Population Dynamics, Mathematica in Science and Engineering, Volume **191**.*

[17] I. PODLUBNY, *Fractional Differential Equations, Academic Press, New York, 1999.*

[18] A. KILBAS, H. SRIVASTAVA, J. TRUJILLO, *Theory and Applications of Fractional Differential Equations, Elsevier Science and Technology, 2006, J. Phys. Chem. 1964, **68**, 5, 1084–1091.*

[19] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations, New York: Springer Verlag, 1983.*

[20] C. LI AND Z. FANHAI, Finite difference methods for fractional differential equations, *Int. J. Bifurcation Chaos,* **22***(04)(2012), 1–28.*

[21] D. BAHUGUNA AND V. RAGHAVENDRA, Rothe's method to parabolic integral-differential equations via abstract integra-differential equations, *Applicable Analysis,* **33***(3-4)(1989), 153–167.*

[22] TOSIO KATO, Nonlinear semigroups and evolution equations, *J. Math. Soc. Japan. Vol. **19**, No 4, 1967.*

[23] C. CHIDUME, Geometric Properties of Banach Spaces and Nonlinear Iterations, *Lecture Notes in Mathematics 1965, Springer-Verlag London Limited 2009.*

[24] C. LI, D. CHEN, Y.Q, On Riemann-Liouville and Caputo Derivatives, *Discret. Dyn. Nat. Soc., 2011, 562494.*

[25] NICOLE HEYMANS AND IGOR PODLUBNY, Physical interpretation of initial conditions for fractional differential equations with Riemann-Liouville fractional derivatives, *Rheol Acta, **45**(2006), 765–771.*

[26] SNEZHANA HRISTOVA, RAVI AGARWAL, AND DONAL O'REGAN, Explicit Solutions of Initial Value Problems for Linear Scalar Riemann-Liouville Fractional Differential Equations With a Constant Delay, *Advances in Difference Equations 2020(1).*

[27] YULIYA KYRYCHKO AND STEPHEN JOHN HOGAN, On the Use of Delay Equations in Engineering Applications, *Journal of Vibration and Control,* ***16****(7-8).*

[28] CHANGPIN LI AND WEIHUA DENG, Remarks on fractional derivatives, *Applied Mathematics and Computation,* ***187****(2)(2007), 777–784.*

[29] V. J. ERVIN, N. HEUER, AND J. P. ROOP, Numerical approximation of a time-dependent, nonlinear, space fractional diffusion equation, *SIAM Journal on Numerical Analysis,* ***45****(2)(2007), 572–591.*

[30] P. ZHUANG, F. LIU, V. ANH, AND I. TURNER, New solution and analytical techniques of the implicit numerical method for the anomalous subdiffusion equation, *SIAM Journal on Numerical Analysis,* ***46****(2)(2008), 1079–1095.*

[31] AKBAR MOHEBBI, Finite difference and spectral collocation methods for the solution of semilinear time fractional convection-reaction-diffusion equations with time delay, *Journal of Applied Mathematics and Computing,* ***61****(2019), 635–656.*

[32] SARITA NANDAL AND DWIJENDRA N. PANDEY, Numerical solution of non-linear fourth order fractional sub-diffusion wave equation with time delay, *Applied Mathematics and Computation,* ***369****(3)(2019), 124900.*

[33] DEVENDRA KUMAR, PARVIN KUMARI, A parameter-uniform collocation scheme for singularly perturbed delay problems with integral boundary condition, *Journal of Applied Mathematics and Computing,* ***63****(2020), 813–828.*

[34] MAHMOUD SHERIF, IBRAHIM ABOUELFARAG, TAREK AMER, Numerical solution of Fractional delay differential equations using Spline functions, *International Journal of Pure and Applied Mathematics* ***90****(1)(2014), 73–83.*

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 43–65.
http://doi.org/10.26637/mjm1101/004

Existence and controllability of impulsive stochastic integro-differential equations with state-dependent delay

MBARACK FALL[1], FIRMIN BODJRENOU[2], CARLOS OGOUYANDJOU[2] AND MAMADOU ABDOUL DIOP[*1]

[1] *Université Gaston Berger de Saint-Louis, UFR SAT, Département de Mathématiques, B.P. 234, Saint-Louis, Sénégal.*
[2] *Institut de Mathématiques et de Sciences Physiques, URMPM B.P. 613, Porto-Novo, Bénin.*

***R*eceived 27 July 2022; *A*ccepted 16 November 2022**

Abstract. This study aims to investigate the existence of mild solutions for a class of impulsive stochastic integrodifferential equations with state-dependent delay in a real separable Hilbert space, as well as the controllability of these solutions. We offer Sufficient conditions for the existence and controllability results using the fixed point techniques combined with the theory of resolvent operator in Grimmer and analysis stochastic. Finally, we provide an example to illustrate the obtained results.

AMS Subject Classifications: 34K50, 34K45, 93B05.

Keywords: Existence, controllability, resolvent operator, impulsive stochastic integro-differential equations, semigroup theory, fixed point theorems, state-dependent delay.

Contents

1. Introduction

Integrodifferential equations represent a wide range of natural phenomena, including biological models, chemical kinetics, electronics, and fluid dynamics. The theory of the integrodifferential equations was generalized to a stochastic functional integrodifferential equations by considering disturbances. As a result, many mathematicians have studied the theory of integrodifferential equations with resolution operators in recent decades (see [22, 23, 35, 39] and the references therein). The resolvent operator is analogous to the semigroup

[*]**Corresponding author.** Email address: **fall.mbarack@ugb.edu.sn** (Mbarack FALL), **firminbodjrenou06@gmail.com** (Firmin BODJRENOU), **ogouyandjou@imsp-uac.org** (Carlos OGOUYANDJOU), **mamadou-abdoul.diop@ugb.edu.sn** (Mamadou Abdoul DIOP)

for differential equations in a Banach space. However, it will not be a semigroup since it does not satisfy the semigroup properties. The existence, uniqueness, stability, controllability, and other quantitative and qualitative aspects of stochastic integrodifferential equations have recently attracted much attention (see [6, 12–14, 37]). In many cases, deterministic models will fluctuate due to random or seemingly random environmental noise. As a result, we have to move from deterministic to stochastic situations

Delay differential equations are an essential branch of nonlinear analysis with numerous applications in almost every field. Usually, the deviation of the arguments depends only on time (see [4, 15, 16]); however, when the deviation of the arguments depends on both the state variable x and the time t, this form of the equation is known as self-reference or state-dependent equations. Equations with state-dependent delays have piqued the interest of specialists because they have numerous application models, such as the two-body problem of classical electrodynamics, and they also have numerous applications within the class of problems with memories, such as in hereditary phenomena, see [42, 43]. Several articles (see [1, 18, 30, 31]) investigated this equation.

Many phenomena and evolution processes in physics, chemical technology, population dynamics, and natural sciences can change state abruptly or be perturbed quickly. We can see these disruptions as impulses in the system. In addition to communications, mechanics (jump discontinuities in velocity), electrical engineering, medicine, and biology, impulsive issues emerge in various other areas. The monographs by Benchohra et al. [7], Graef et al. [20, 21], Laskshmikantham et al. [3], and Samoilenko and Perestyuk [40] provide a thorough introduction to basic theory. On the other hand, Milman and Myshkis [34] studied differential equations with impulses for the first time, followed by a period of active research culminating in the monograph by Halanay and Wexler [24].

In the field of mathematical control theory, the idea of controllability plays an essential role. It makes it possible to use a control that is admissible to guide the system from its initial state to its final state within a specified amount of time. The concept of controllability is crucial to the study of finite-dimensional control theory. Therefore, it is only natural to attempt to generalize it to an infinite number of dimensions. The controllability of nonlinear systems with different types of nonlinearities has been studied using fixed point concepts[5]. Several authors have investigated the controllability of semilinear and nonlinear systems, represented by differential and integrodifferential equations in finite or infinite dimensional Banach spaces, respectively[9, 36, 44, 46].

Recently, Ma and Liu [32] studied the exact controllability and continuous dependence of fractional neutral integrodifferential equations with state-dependent delay in Banach spaces. Slama and Boudaoui [41], authors proved a new set of sufficient conditions for a class of fractional nonlinear stochastic differential inclusions using fractional calculus, stochastic analysis theory, semigroup theory, and Bohnenblust-Karlin's fixed point theorem.To the best of our knowledge, the literature related to stochastic impulsive integrodifferential remains limited.

Inspired by the works mentioned above, the main objective of this manuscript is to investigate the existence and controllability results for the following model

$$\begin{cases} dz(t) = \Big[Az(t) + \displaystyle\int_0^t \Gamma(t-s)z(s)ds + g\left(t, z_t, \int_0^t h(t,s,z_s)ds\right)\Big]dt \\ \qquad\qquad +\xi(t, z_{\sigma(t,z_t)})dw(t),\ t \in J = [0,a], t \neq t_i, \\ \Delta z(t_i) = I_i(z_{t_i}), i = 1, \dots, m \\ z_0 = \varphi \in \mathcal{B}, \end{cases} \tag{1.1}$$

where the state $z(\cdot)$ takes values in a real separable Hilbert space $\mathbb{X}$, the operator A is an infinitesimal generator of a C_0-semigroup $\big(T(t)\big)_{t\geq 0}$ on $\mathbb{X}$, $(\Gamma(t))_{t\geq 0}$ is a family of closed linear operators on $\mathbb{X}$ with domain $D(\Gamma(t)) \supset D(A)$, the history $z_t : (-\infty, 0] \longrightarrow \mathbb{X}, z_t(\alpha) = z(t+\alpha)$, for $t \leq 0$, belongs to the phase space $\mathcal{B}$, which will be described axiomatically later, the mappings $g : J \times \mathcal{B} \times \mathbb{X} \longrightarrow \mathbb{X}, h : J \times J \times \mathcal{B} \longrightarrow \mathbb{X}, \xi :$ $J \times \mathcal{B} \longrightarrow \mathcal{L}_Q(\mathbb{Y}, \mathbb{X}), \sigma : J \times \mathcal{B} \times (-\infty, a]$ are appropriate functions that will be specified later, for $i = 1, \dots, m$, the functions $I_i : \mathcal{B} \longrightarrow \mathbb{X}$ represent the impulses, let $0 = t_0 < t_1 < \cdots < t_m < t_{m+1} = a$ be prefixed points, and $\Delta z(t_i)$ represents the jump of the function z at t_i, which is defined by $\Delta z(t_i) = z(t_i^+) - x(t_z^-)$, where $z(t_i^+)$ and $z(t_i^-)$ denote the right and left limits of $z(t)$ at $t = t_i$, respectively. It should be emphasized that the existence and controllability results for impulsive stochastic evolution integral differential equations with state dependent delay in the form (1.1) have not yet been explored.

Here, we discuss the existence and controllability of the system(1.1) by using resolvent operator in the sense of Grimmer and stochastic analysis tools combined with fixed point theory.

We will proceed as follows: Definitions and Lemma, which are necessary to derive the main results, are outlined in Section 2. In Section 3, we prove the existence results using the Krasnoselskii-Schaefer fixed point theorem's implication. Section 4 is devoted to controllability. As a concluding point, an example is provided in Section 5 to illustrate the theoretical outcomes.

2. Preliminaries

In this section, we briefly review some basic definitions and notations that will be used in the subsequent sections.

2.1. Brownian motion

Let $(\mathbb{X}, \|\cdot\|, \langle\cdot,\cdot\rangle)$ and $(\mathbb{Y}, \|\|_{\mathbb{Y}}, \langle\cdot,\cdot\rangle_{\mathbb{Y}})$ be two real separable Hilbert spaces, $\{\eta_k\}_{k=1}^{\infty}$ denote a complete orthonormal basis of $\mathbb{Y}$ and $\{w(t) : t \geq 0\}$ be a cylindrical $\mathbb{Y}-$value $Q-$Wiener process in which Q is a finite nuclear covariance operator. Denote $Tr(Q) = \sum_{k=1}^{\infty} \gamma_k$, which satisfies $Q\eta_k = \gamma_k \eta_k$, $(\gamma_k \geq 0, k = 1, 2, \ldots)$. Set

$$w(t) = \sum_{k=1}^{\infty} \sqrt{\gamma_k}\beta_k(t)\eta_k, t \geq 0,$$

where $\{\beta_k(t)\}_{k=1}^{\infty}$ is a sequence of real-values independant one-dimensional standards Brownian motions over a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

It is assumed that $\mathcal{F}_t = \sigma\{w(s) : 0 \leq s \leq t\}$ is the $\sigma-$ algebra generated by w and $\mathcal{F}_t = \{\mathcal{F}_t\}_{t\geq 0}$. Let $\psi \in \mathcal{L}(\mathbb{Y}, \mathbb{X})$ and define

$$\|\psi\|_Q^2 = Tr(\psi Q \psi^*) = \sum_{k=1}^{\infty} \|\sqrt{\gamma_k}\psi\eta_k\|^2,$$

where ψ^* is the adjoint of the operator ψ, and $\mathcal{L}(\mathbb{Y}, \mathbb{X})$ denotes the space of all bounded linear operators from K into H endowed with the same norm $\|\cdot\|$. if $\|\psi\|_Q^2 < \infty$, the ψ is called a $Q-$Hilbert- Schmidt operator. The completion $\mathcal{L}_Q(\mathbb{Y}, \mathbb{X})$ of $\mathcal{L}(\mathbb{Y}, \mathbb{X})$ with respect to the topology induced by the norm $\|\cdot\|_Q$, is a Hilbert space with the above norm topology, where $\|\psi\|_Q = \langle\psi, \psi\rangle^{\frac{1}{2}}$. The collection of all strongly measurable, square-integrable, $\mathbb{X}-$ valued random variables, denoted by $\mathcal{L}_2(\Omega, \mathbb{X})$, is a Banach space equipped with the norm $\|z\|_{\mathcal{L}_2} = \left(\mathbb{E}\|z\|^2\right)^{\frac{1}{2}}$, where the expectation E is defined by $\mathbb{E}z = \int_{\Omega} z(w)d\mathbb{P}$.

Let $C(J, \mathcal{L}_2(\Omega, \mathbb{X}))$ be the Banach space of all continuous maps from J into $\mathcal{L}_2(\Omega, \mathbb{X})$ satisfying the condition $\sup_{0\leq t\leq a} \mathbb{E}\|z(t)\|^2 < \infty$. An important subspace $\mathcal{L}_2^0(\Omega, \mathbb{X})$ of $\mathcal{L}_2(\Omega, \mathbb{X})$ is given by

$$\mathcal{L}_2^0(\Omega, \mathbb{X}) = \{z \in \mathcal{L}_2(\Omega, \mathbb{X}) : z \, is \, \mathcal{F}_0\text{-measurable}\}.$$

For more details, we refer the reader to Da Patro and Zabczyk, LesZek Gawarecki and Vidyadhar Mandrekar [19].

2.2. Integrodifferential equations in Banach spaces

Here we recall some knowledge on partial integrodifferential equations and the related resolvent operators. Let $\mathcal{H}$ be the Banach space $D(A)$ equipped with the graph norm defined by

$$\|\theta\|_{\mathcal{H}} := \|A\theta\| + \|\theta\| \text{ for } \theta \in \mathcal{H}.$$

We denote by $\mathcal{C}(\mathbb{R}^+, \mathcal{D})$, the space of all functions from $\mathbb{R}^+$ into $\mathcal{D}$ which are continuous. Let us consider the following system:

$$\begin{cases} \theta'(t) = A\theta(t) + \int_0^t \Gamma(t-s)\theta(s)ds \quad \text{for } \ t \in [0, a], \\ \theta(0) = \theta_0 \in \mathcal{D}, \end{cases} \tag{2.1}$$

where A and $\Gamma(t)$ are closed linear operators on a Banach space $\mathcal{D}$.

Definition 2.1 ([22])**.** *A resolvent operator for Eq.* (2.1) *is a bounded linear operator valued function $R(t) \in \mathcal{L}(\mathcal{D})$ for $t \in [0, a]$, having the following properties :*

(i) *$R(0) = I$ (the identity map of $\mathcal{D}$) and $\|R(t)\| \leq Ne^{\beta t}$ for some constants $N > 0$ and $\beta \in \mathbb{R}$.*

(ii) *For each $\theta \in \mathcal{D}$, $R(t)\theta$ is strongly continuous for $t \in [0, a]$.*

(iii) *For $\theta \in \mathcal{H}$, $R(\cdot)\theta \in \mathcal{C}^1(\mathbb{R}^+; \mathcal{D}) \cap \mathcal{C}(\mathbb{R}^+; \mathcal{H})$ and*

$$\begin{aligned} R'(t)\theta &= AR(t)\theta + \int_0^t \Gamma(t-s)R(s)\theta ds \\ &= R(t)A\theta + \int_0^t R(t-s)\Gamma(s)\theta ds, \quad \textit{for } t \in [0, a]. \end{aligned}$$

In what follows, we make the following assumptions.

($\mathbf{R_1}$) The operator A is the infinitesimal generator of a strongly continuous semigroup $(\mathsf{T}(t))_{t\geq 0}$ on $\mathcal{D}$.

($\mathbf{R_2}$) $(\Gamma(t))_{t\geq 0}$ is is a family of linear operators on $\mathcal{D}$ such that $\Gamma(t)$ is continuous when regarded as a linear map from $\mathcal{H}$ into $\mathcal{D}$ for almost all $t \geq 0$. For any $\theta \in \mathcal{D}$, the map $t \mapsto \Gamma(t)\theta$ is bounded, differentiable and the derivative $t \mapsto \Gamma'(t)\theta$ is bounded and uniformly continuous for $t \geq 0$.

Theorem 2.2. [22] *Assume that* ($\mathbf{R_1}$)-($\mathbf{R_2}$) *hold. Then there exists a unique resolvent operator to the Cauchy problem* (2.1)*.*

We have the following useful results.

Lemma 2.3. *[11] Let the assumptions* ($\mathbf{R_1}$) *and* ($\mathbf{R_2}$) *be satisfied. Then, there exists a constant $\Delta = \Delta(a)$ such that*

$$\|R(t+\epsilon) - R(\epsilon)R(t)\|_{\mathcal{L}(\mathcal{D})} \leq \Delta\epsilon, \qquad \forall\, 0 < \epsilon \leq t \leq a.$$

Theorem 2.4 (Theorem 6, [17])**.** *Let A be the infinitesimal generator of a C_0-semigroup $(T(t))_{t\geq 0}$ and let $(\mathbf{\Gamma}(t))_{t\geq 0}$ satisfy* ($\mathbf{R_2}$)*. Then the resolvent operator $(R(t))_{t\geq 0}$ for Eq.* (2.1) *is compact for $t > 0$ if and only if $(T(t))_{t\geq 0}$ is compact for $t > 0$.*

In the sequel, we recall some results on the existence of solutions for the following integro-differential equation:

$$\begin{cases} \theta'(t) = A\theta(t) + \int_0^t \Gamma(t-s)\theta(s)ds + l(t) \quad \text{for } \ t \in [0, a], \\ \theta(0) = \theta_0 \in \mathcal{D}. \end{cases} \tag{2.2}$$

where l is a continuou function.

Definition 2.5. [23] *A continuous function $\theta :\ [0, +\infty) \to \mathcal{D}$ is said to be a strict solution of Eq.* (2.2) *if*

1. *$\theta \in \mathcal{C}^1([0, +\infty), \mathcal{D}) \cap \mathcal{C}([0, +\infty), \mathcal{H})$,*
2. *θ satisfies Eq.* (2.2) *for $t \geq 0$.*

Theorem 2.6. [23] *Assume that* $(\mathbf{R_1})$ *and* $(\mathbf{R_2})$ *hold. If* θ *is a strict solution of Eq.* (2.2), *then the following variation of constant formula holds*

$$\theta(t) = R(t)\theta_0 + \int_0^t R(t-s)l(s)ds \quad for\ t \geq 0. \tag{2.3}$$

Accordingly, we can establish the following definition.

Definition 2.7. *[23] A function* $\theta : [0,\infty) \longrightarrow \mathcal{D}$ *is said a mild solution of Eq.* (2.2) *for* $\theta_0 \in \mathcal{D}$*, if* θ *satisfies the variation of constants formula* (2.3).

In what follows, we say a function $z : [b,c] \longrightarrow \mathbb{X}$ is a normalized piecewise continuous function on $[b,c]$ if z is piecewise continuous and left continuous on $(b,c]$. We denote by $\mathcal{PC}([b,c],\mathbb{X})$ the space formed by the normalized piecewise continuous, $\mathcal{F}_t$-adapted measurable process from $[b,c]$ into $\mathbb{X}$. Particularly, we introduce the space $\mathcal{PC}$ formed by all $\mathcal{F}_t$-adapted measurable, $\mathbb{X}$ valued stochastic process $\{z(t) : t \in [0,a]\}$ such that z is continuous at $t \neq t_i$, $z(t_i^-) = z(t_i)$ and $z(t_i^+)$ exists, for $i = 1,2,\cdots,m$. Then $(\mathcal{PC}, \|\cdot\|_{\mathcal{PC}})$ is a Banach space with norm given by

$$\|z\|_{\mathcal{PC}} = \sup_{s\in J}\left(E\|z(s)\|^2\right)^{\frac{1}{2}}.$$

For $z \in \mathcal{PC}$, we denote $\tilde{z}_i, i = 1,2,\ldots,m$, the function $\tilde{z}_i \in C\big([t_i,t_{i+1}];\mathcal{L}_2(\Omega,\mathbb{X})\big)$ given by

$$\tilde{z}_i(t) = \begin{cases} z(t), & for\, t \in (t_i,t_{i+1}], \\ z(t_i^+), & if\, t = t_i. \end{cases}$$

Moreover, for $\mathcal{E} \subseteq \mathcal{PC}$, we denote by $\tilde{\mathcal{E}}_i, i = 0,1,2,\ldots,m$, the set $\tilde{\mathcal{E}}_i = \{\tilde{z}_i : z \in \mathcal{E}\}$.

Lemma 2.8 ([26], [28]). .
A set $\mathcal{E} \subseteq \mathcal{PC}$*, is relatively compact in* $\mathcal{PC}$*, if and only if the set* $\tilde{\mathcal{E}}_i$ *is relatively compact in* $C([t_i,t_{i+1}];\mathcal{L}_2(\Omega,\mathbb{X}))$ *for every* $i = 0,1,\ldots,m$.

In order to deal with the infinite delay, we will consider the phase space $\mathcal{B}$ which was described by Hale and Kato in [25]. More precisely, $\mathcal{B}$ will be a seminormed linear space of $\mathcal{F}_0-$measurable functions defined from $(-\infty,0]$ into $\mathbb{X}$, and satisfying the following axioms:
A: If $z : (-\infty,\rho+a] \longrightarrow \mathbb{X}, a > 0$ is such that $z_\rho \in \mathcal{B}$ and $x|_{[\rho,\rho+a]} \in \mathcal{PC}([\rho,\rho+a],\mathbb{X})$, then, for every $t \in [\rho,\rho+a)$, : the following conditions hold:

(i) $z_t \in \mathcal{B}$,

(ii) $\mathbb{E}\|z(t)\| \leq H\|z_t\|_{\mathcal{B}}$,

(iii) $\|z_t\|_{\mathcal{B}} \leq K_1(t-\rho)\sup_{\rho\leq s\leq t}\mathbb{E}\|z(s)\| + K_2(t-\rho)\|z_\rho\|_{\mathcal{B}}$,
where $H > 0$ is a constant, $K_1(\cdot), K_2(\cdot) : [0,+\infty) \longrightarrow [1,+\infty)$, $K_1(\cdot)$ is continuous, $K_2(\cdot)$ is locally bounded, and $H, K_1(\cdot), K_2(\cdot)$ are independent of $z(\cdot)$.

B: The space $\mathcal{B}$ is complete.
The following results will be required in computation.

Lemma 2.9. *[45]. Let* $z : (-\infty,a] \longrightarrow \mathbb{X}$ *be can an* $\mathcal{F}_0$*-adapted process* $z_0 = \varphi(t) \in \mathcal{L}_2^0(\Omega,\mathcal{B})$ *and* $z|_J \in \mathcal{PC}(J,\mathbb{X})$*, then*

$$\|z_s\|_{\mathcal{B}} \leq \tilde{K}_2\|\varphi\|_{\mathcal{B}} + \tilde{K}_1\sup_{0\leq s\leq a}\mathbb{E}\|x(s)\|,$$

where $\tilde{K}_1 = \sup_{t\in J}K_1(t)$ *and* $\tilde{K}_2 = \sup_{t\in J}K_2(t)$.

In order to handle the delay function σ, the next result is a very useful.

Lemma 2.10. [27] *Let $z : (-\infty, a] \longrightarrow \mathbb{X}$ be a function such that $z_0 = \varphi$, and $z|_J \in \mathcal{PC}$. Then*

$$\|z_s\|_{\mathcal{B}} \leq (\tilde{K}_2 + J_0^{\varphi})\|\varphi\|_{\mathcal{B}} + \tilde{K}_1 \sup \mathbb{E}\{\|z(\theta)\| : \theta \in [0, \max\{0, s\}]\}, s \in \mathcal{Z}(\sigma^-) \cup J,$$

where $J_0^{\varphi} = \sup\{J^{\varphi}(t) : t \in \mathcal{Z}(\sigma^-)\}$.

Now, we give two important fixed point theorems and Burkholder- Davis-Gundy's inequality which are used in the proof of the main results.

Lemma 2.11. *[38]. Let Ψ be a condensing operator on a Banach space H, i.e., Ψ is continuous and takes bounded sets into bounded sets, and $\nu(\Psi(C)) \leq (C)$ for every bounded set C of $\mathbb{H}$ with $\nu(C) > 0$, where $\nu(\cdot)$ denotes the Kuratowskii measure of noncompactness. If $\Psi(F) \subset F$ for a convex, closed and bounded set F of $\mathbb{H}$, then Ψ has a fixed point in $\mathbb{H}$.*

Lemma 2.12. *[8] . Let Ψ_1 and Ψ_2 be two operators of a Banach space $\mathbb{H}$ such that*

(a) *Ψ_1 is a contraction, and*

(b) *Ψ_2 is completely continuous.*

Then, either

(i) *the operator equation $\Psi_1 z + \Psi_2 z = z$ has a solution, or*

(ii) *the set $\mathbb{M} = \{z \in \mathbb{H} : \alpha\Psi_1(\frac{z}{\alpha}) + \alpha\Psi_2(z) = z\}$ is unbounded for $\alpha \in (0, 1)$.*

Lemma 2.13. *[10] For any $p \geq 1$ and for arbitrary $L_Q(\mathbb{Y}, \mathbb{X})$-valued predictable process $\Psi(\cdot)$,*

$$\sup_{s\in[0,t]} \mathbb{E}\Big\| \int_0^s \Psi(r)dw(r)\Big\|_{\mathbb{X}}^{2p} \leq (p(2p-1))^p \Big(\int_0^t (\mathbb{E}\|\Psi(s)\|_Q^{2p})^{\frac{1}{p}} \Big)^p. \tag{2.4}$$

We now end this part by stating the definition of mild solution for Eq. (1.1).

Definition 2.14. *An $\mathcal{F}_t-$adapted stochastic process $z : (-\infty, a] \longrightarrow \mathbb{X}$ is said to be a mild solution of Eq.* (1.1) *if $z_0 = \varphi \in \mathcal{B}$, $z_\sigma(s, z_s) \in \mathcal{B}$ satisfying $z_0 \in \mathcal{L}_2^0(\Omega, \mathbb{X}), z|_J \in \mathcal{PC}$. The function $R(t-s)g(s, z_s, \int_0^s h(s, \tau, z_\tau)d\tau)$ is integrable for each $s \in [0, a]$ and the following conditions hold:*

(i) *$\{z_t : t \in J\}$ is $\mathcal{B}-$valued and the restriction of $z(\cdot)$ to the interval $(t_i, t_{i+1}], i = 1, 2, \ldots, m$ is continuous ;*

(ii) *$\Delta z(t_i) = I_i(z_{t_i}), i = 1, 2, \ldots, m;$*

(iii) *for each $t \in J$, $z(t)$ satisfies the following integral equation*

$$z(t) = R(t)\varphi(0) + \int_0^t R(t-s)g(s, z_s, \int_0^s h(s, \tau, z_\tau)d\tau)ds$$
$$+ \int_0^t R(t-s)\xi(s, z_{\sigma(s,z_s)})dw(s) + \sum_{0<t_i<t} R(t-t_i)I_i(z_{t_i}).$$

3. Existence results

This section is devoted to the study of existence of mild solutions for Eq. (1.1). Throughout this work, we assume that $\sigma : J\times\mathcal{B} \longrightarrow (-\infty, a]$ is continuous and $M = \sup_{t\in J} \|R(t)\|$. In the following, we firstly introduce the subsequent hypotheses:

(A_1) Let $\mathcal{Z}(\sigma^-) = \{\sigma(s,\varphi) \leq 0, \sigma(s,\varphi) : (s,\varphi) \in J\times\mathcal{B}\}$. The function $t \longrightarrow \varphi_t$ is well defined from $\mathcal{Z}(\sigma^-)$ into $\mathcal{B}$ and there exists a continuous and bounded function $J^\varphi : \mathcal{Z}(\sigma^-) \longrightarrow (0,\infty)$ such that $\|\varphi_t\|_\mathcal{B} \leq J^\varphi(t)\|\varphi\|_\mathcal{B}$ for every $t \in \mathcal{Z}(\sigma^-)$.

(A_2) The resolvent operator $\big(R(t)\big)_{t\geq 0}$ is compact for $t > 0$.

(A_3) The function $\xi : J\times\mathcal{B} \longrightarrow \mathcal{L}_Q(\mathbb{Y},\mathbb{X})$ satisfies the following properties:

- (i) The function $\xi(\cdot, z) : J \longrightarrow \mathcal{L}_Q(\mathbb{Y},\mathbb{X})$ is strongly measurable for every $z \in \mathcal{B}$,
- (ii) The function $\xi(t,\cdot) : \mathcal{B} \longrightarrow \mathcal{L}_Q(\mathbb{Y},\mathbb{X})$ is continuous on $\mathcal{Z}(\sigma^-) \cup J$,
- (iii) There exist an integrable function $l : J \longrightarrow [0,\infty)$ and a non-decreasing function $\mu_l \in \mathcal{C}([0,\infty); (0,\infty))$ such that, for every $(t,z) \in J\times\mathcal{B}$,

$$\mathbb{E}\|\xi(t,z)\|^2 \leq l(t)\mu_l(\|z\|^2_\mathcal{B}), \lim_{\delta\longrightarrow\infty} \inf \frac{\mu_l(\delta)}{\delta} = \Theta < \infty.$$

(A_4) There exist constants $d_1 > 0$ and $d_1^* > 0$ for all $\varphi, \psi \in \mathcal{B}, t, s \in J$, such that

$$\mathbb{E}\left\|\int_0^t [h(t,s,\varphi) - h(t,s,\psi)]ds\right\|^2 \leq d_1\|\varphi-\psi\|^2$$

and $d_1^* = a\sup_{0\leq s\leq t\leq a} \mathbb{E}\|h(t,s,0)\|^2$.

(A_5) The function $g : J\times\mathcal{B}\times\mathbb{X} \longrightarrow \mathbb{X}$ is continuous and there exist constants $d_2 > 0$ and $d_2^* > 0$ for all $t \in J, \varphi_1, \varphi_2 \in \mathcal{B}, z_1, z_2 \in \mathbb{X}$ such that

$$\mathbb{E}\|g(t,\varphi_1,z_1) - g(t,\varphi_2,z_2)\|^2 \leq d_2\big(\|\varphi_1-\varphi_2\|^2_\mathcal{B} + \mathbb{E}\|z_1-z_2\|^2_\mathbb{X}\big)$$

and $d_2^* = \sup_{t\in J} E\|g(t,0,0)\|^2$.

(A_6) The maps I_i, are completely continuous and there exist positive constant $\lambda_i, i = 1,2,\ldots,m$, such that $\mathbb{E}\|I_i(z)\|^2 \leq \lambda_i\|z\|^2_\mathcal{B}$ for all $z \in \mathcal{B}$.

Remark 3.1. *Let $\varphi \in \mathcal{B}$ and $t \leq 0$. The notation φ_t represents the function defined by $\varphi_t(\theta) = \varphi(t+\theta)$. Consequently, if the function $z(\cdot)$ in the Axiom A is such that $z_0 = \varphi$, then $z_t = \varphi_t$. We observe that φ_t is well defined for $t < 0$, since the domain of φ is $(-\infty, 0]$.*

Theorem 3.2. *Assume that $(\mathbf{R_1})$-$(\mathbf{R_2})$, $(A_1)-(A_6)$ are satisfied and $z_0 \in \mathcal{L}^0_2(\Omega,\mathbb{X}), \varphi \in \mathbb{X}$. If*

$$T_0 = 1 - 4\bigg(4M^2a^2d_2\tilde{K}_1^2 + 8d_2d_1M^2a^2\tilde{K}_1^2 + 2M^2m\tilde{K}_1^2\sum_{i=1}^m \lambda_i\bigg) > 0, \tag{3.1}$$

and

$$\frac{8\tilde{K}_1^2M^2Tr(Q)}{T_0}\int_0^a l(s)ds \leq \int_{T^*}^\infty \frac{ds}{\mu_l(s)},$$

where

$$T^* = C + \frac{8\tilde{K}_1^2}{T_0}\bigg[]M^2Hd^2\|\varphi\|^2_\mathcal{B} + 2M^2a^2\left(d_2C + 2d_2d_1C + 2d_2d_1^* + d_2^*\right) + M^2m\sum_{i=1}^m \lambda_iC\bigg]$$

with $C = 2(\tilde{K}_2 + J_0^\varphi)^2\|\varphi\|^2_\mathcal{B}$, then there exists a mild solution of Eq (1.1).

Proof. Let $\mathbb{F} = \{z \in \mathcal{PC} : z(0) = \varphi(0)\}$ be a space endowed with the uniform convergence topology. We define the operator $\Psi : \mathbb{F} \longrightarrow \mathbb{F}$ by

$$(\Psi z)(t) = \begin{cases} 0, \quad t \in (-\infty, 0] \\ R(t)\varphi(0) + \displaystyle\int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds \\ + \displaystyle\int_0^t R(t-s)\xi(s, z_{\sigma(s,\tilde{z}_s)})dw(s) + \sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i}), \quad t \in J, \end{cases}$$

where $\tilde{z} : (-\infty, a] \longrightarrow \mathbb{X}$ is such that $\tilde{z}_0 = \varphi$ and $\tilde{z} = z$ on J. In view of hypotheses (A_2),(A_4) and (A_5), we have the following inequality

$$\begin{aligned} &\mathbb{E}\left\|\int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds\right\|^2 \\ &\quad \leq 2M^2 a \int_0^t \left(d_2[\|\tilde{z}_s\|^2_{\mathcal{B}} + 2d_1\|\tilde{z}_s\|^2_{\mathcal{B}} + 2d_1^*] + d_2^*\right) ds. \end{aligned}$$

Then from the Bochner theorem [33], it follows that $R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)$ is integrable on $[0,t)$, which allows us to conclude that Ψ is a well-defined operator from $\mathbb{F}$ into $\mathbb{F}$. We prove that Ψ has a fixed point, which is a mild solution of the Eq.(1.1). Now, we decompose Ψ as $\Psi_1 + \Psi_2$, where

$$\begin{aligned} (\Psi_1 z)(t) &= R(t)\varphi(0) + \int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds \\ (\Psi_2 z)(t) &= \int_0^t R(t-s)\xi(s, z_{\sigma(s,\tilde{z}_s)})dw(s) + \sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i}), \ t \in J. \end{aligned}$$

Firstly, we show that Ψ_1 is a contraction. Next, we prove that Ψ_2 is a completely continuous. In order to apply Lemma 2.12 we give the proof in several steps.

Step 1: We will show the set $\mathbb{S} = \{z \in \mathbb{F} : \epsilon\Psi_1(\frac{z}{\epsilon}) + \epsilon\Psi_2(z) = z\}$ is bounded on J for some $\epsilon \in (0,1)$. Consider the following nonlinear operator equation

$$z(t) = \epsilon\Psi z(t), \quad 0 < \epsilon < 1, \tag{3.2}$$

where the operator Ψ has already been defined. Next we give a priori estimate for the solutions of the above equation. Let $z \in \mathbb{F}$ be a possible solution of $z(t) = \epsilon\Psi z(t)$ for some $0 < \epsilon < 1$, we have

$$\begin{aligned} z(t) &= \epsilon R(t)\varphi(0) + \epsilon\int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds \\ &\quad + \epsilon\int_0^t R(t-s)\xi(s, z_{\sigma(s,\tilde{z}_s)})dw(s) + \epsilon\sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i}), \quad t \in J. \end{aligned} \tag{3.3}$$

Using (3.3), hypotheses $(A_2)-(A_6)$, Hölder and Burkholder-Davis-Gundy's inequalities, we have

$$\begin{aligned}
&\mathbb{E}\|z(t)\|^2 \qquad (3.4)\\
&\quad\leq 4\mathbb{E}\|\epsilon R(t)\varphi(0\|^2 + 4\mathbb{E}\|\epsilon\int_0^t R(t-s)g(s,\tilde{z}_s,\int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds\|^2\\
&\qquad + 4\|\epsilon\int_0^t R(t-s)\xi(s,z_{\sigma(s,\tilde{z}_s)})dw(s)\|^2 + 4\|\epsilon\sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i})\|^2\\
&\quad\leq 4M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 8M^2a\int_0^t\Bigg\{d_2\Bigg(2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\\
&\qquad + 2d_1\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\right] + 2d_1^*\Bigg) + d_2^*\Bigg\}ds\\
&\qquad + 4M^2Tr(Q)\int_0^t l(s)\mu_l\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq \tau\leq a}\mathbb{E}\|z(\tau)\|^2\right]ds\\
&\qquad + 4M^2m\sum_{i=1}^m\lambda_i\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\right]\\
&\quad\leq 4M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 8M^2a^2\Bigg[d_2\Bigg(2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\\
&\qquad + 2d_1\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\right] + 2d_1^*\Bigg) + d_2^*\Bigg]\\
&\qquad + 4M^2Tr(Q)\int_0^t l(s)\mu_l\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq s\leq a}\mathbb{E}\|z(s)\|^2\right]ds\\
&\qquad + 4M^2m\sum_{i=1}^m\lambda_i\left[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2\sup_{0\leq t\leq a}\mathbb{E}\|z(t)\|^2\right]. \qquad (3.5)
\end{aligned}$$

Let $\vartheta(s)=\sup_{0\leq s\leq a}\mathbb{E}\|z(s)\|^2$ and $C=2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2$. From (3.4), we have

$$\begin{aligned}
&\vartheta(t)\\
&\quad\leq 4\Bigg\{M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 2M^2a^2\Bigg\{d_2\left[C+2\tilde{K}_1^2\vartheta(t)+2d_1\left[C+2\tilde{K}_1^2\vartheta(t)\right]+2d_1^*\right]+d_2^*\Bigg\}\\
&\qquad + M^2Tr(Q)\int_0^t l(s)\mu_l\left[C+2\tilde{K}_1^2\vartheta(s)\right]ds + M^2m\sum_{i=1}^m\lambda_i\left[C+2\tilde{K}_1^2\vartheta(t)\right]\Bigg\}\\
&\quad\leq 4\Bigg\{M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 2M^2a^2\left[d_2C+2d_2d_1C+2d_2d_1^*+d_2^*\right]\\
&\qquad + M^2Tr(Q)\int_0^t l(s)\mu_l\left[C+2\tilde{K}_1^2\vartheta(s)\right]ds + M^2m\sum_{i=1}^m\lambda_iC\\
&\qquad + \vartheta(t)\left[4M^2a^2d_2\tilde{K}_1^2 + 8d_2d_1M^2a^2\tilde{K}_1^2 + 2M^2m\tilde{K}_1^2\sum_{i=1}^m\lambda_i\right]\Bigg\}.
\end{aligned}$$

It follows that

$$\vartheta(t) \leq \frac{4}{T_0}\Big\{M^2H^2\|\varphi\|^2_{\mathcal{B}} + 2M^2a^2\left[d_2C + 2d_2d_1C + 2d_2d_1^* + d_2^*\right] + M^2Tr(Q)\int_0^t l(s)\mu_l\left[C + 2\tilde{K}_1^2\vartheta(s)\right]ds + M^2m\sum_{i=1}^{m}\lambda_iC\Big\}.$$

Let $\omega(t) = C + 2\tilde{K}_1^2\vartheta(t)$. Since $\sigma(s;\tilde{z}_s) \leq s$ for every $s \in [0,a]$, we have

$$\omega(t) \leq T^* + \frac{8\tilde{K}_1^2M^2Tr(Q)}{T_0}\int_0^t l(s)\mu_l(\omega(s))ds.$$

Denoting by $\nu(t)$ the right-hand side of the last inequality, we find that

$$\nu^{'}(t) \leq \frac{8\tilde{K}_1^2M^2Tr(Q)}{T_0}l(t)\mu_l(\nu(t))$$

and

$$\int_{T^*}^{\nu(t)}\frac{ds}{\mu_l(s)} \leq \frac{8\tilde{K}_1^2M^2Tr(Q)}{T_0}\int_0^a l(s)ds \leq \int_{T^*}^{\infty}\frac{ds}{\mu_l(s)}.$$

Consequently, we see that $\nu(t)$ is bounded, which proves that z is bounded in $\mathbb{F}$ for any $z \in \mathbb{S}$. Hence $\mathbb{S}$ is bounded on J for $\epsilon \in (0,1)$.

Step 2: Ψ_1 is a contraction operator on $\mathbb{F}$.

Let $t \in J$ and $z, y \in \mathbb{F}$. Then, by assumptions (A_2),(A_4) and (A_5) and Lemma 2.10 , we have

$$\begin{aligned}
&\mathbb{E}\|(\Psi_1z)(t) - (\Psi_1y)\|^2\\
&\leq \mathbb{E}\left\|\int_0^t R(t-s)\big[g\left(s,\tilde{z}_s,\int_0^s h(s,\tau,\tilde{z}_\tau)d\tau\right) - g\left(s,\tilde{y}_s,\int_0^s h(s,\tau,\tilde{y}_\tau)d\tau\right)\big]ds\right\|^2\\
&\leq M^2a\int_0^t\mathbb{E}\left\|g\left(s,\tilde{z}_s,\int_0^s h(s,\tau,\tilde{z}_\tau)d\tau\right) - g\left(s,\tilde{y}_s,\int_0^s h(s,\tau,\tilde{y}_\tau)d\tau\right)\right\|^2ds\\
&\leq M^2a^2(d_2+d_1d_2)\|\tilde{z}_s - \tilde{y}_s\|^2_{\mathcal{B}}\\
&\leq M^2a^2(d_2+d_1d_2)\tilde{K}_1^2\sup_{0\leq s\leq a}\|\tilde{z}(s) - \tilde{y}(s)\|^2\\
&\leq M^2a^2(d_2+d_1d_2)\tilde{K}_1^2\|\tilde{z} - \tilde{y}\|_{\mathcal{PC}}\\
&= L_0\|x-y\|^2_{\mathcal{PC}},
\end{aligned}$$

where $L_0 = M^2a^2(d_2+d_1d_2)\tilde{K}_1^2$. By (3.1), we see that $L_0 < 1$. As a consequence Ψ_1 is a contraction operator on $\mathbb{F}$.

Step 3: Ψ_2 is a completely continuous operator on $\mathbb{F}$. We will do it into several steps.

(a) $\Psi_2 : \mathbb{F} \longrightarrow \mathbb{F}$ is continuous.
Let $\{z^n\}_{n=0}^{\infty} \subseteq \mathbb{F}$, with $z^n \longrightarrow z$ in $\mathbb{F}$. Then, there is a number $q > 0$ such that $\mathbb{E}\|z^n\|^2 \leq q$ for all n and a.e. $t \in J$, so $z^n \in B_q(0,\mathbb{F}) = \{z \in \mathbb{F} : \mathbb{E}\|z\|^2 \leq q\}$ and $z \in B_q(0,\mathbb{F})$. From Axiom A, it is not hard to see that $(\tilde{z}^n)_s \longrightarrow \tilde{z}_s$ uniformly for $s \in (-\infty,a]$ as $n \longrightarrow \infty$. By hypotheses (A_1) and (A_3), we obtain

$$\xi(s,\tilde{z}^n_{\sigma(s,(\tilde{z}^n)_s)}) \longrightarrow \xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)}) \quad \text{as} \quad n \longrightarrow \infty$$

for each $s \in [0,t]$, and

$\mathbb{E}\left\|\xi(s,\tilde{z}^n_{\sigma(s,(\tilde{z}^n)_s)})-\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})\right\|^2 \leq 2l(t)\mu_l[2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|^2_{\mathcal{B}}+2\tilde{K}_1^2 q].$
Then, by the dominated convergence theorem and the continuity of $I_i,\ i=1,2,\ldots,m$, we get

$$\begin{aligned}
&\|\Psi_2 z^n-\Psi_2 z\|^2_{\mathcal{PC}}\\
&\quad\leq 2\mathbb{E}\left\|\int_0^t R(t-s)\left[\xi(s,\tilde{z}^n_{\sigma(s,(\tilde{z}^n)_s)})-\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})\right]dw(s)\right\|^2\\
&\qquad+2\mathbb{E}\left\|\sum_{0<t_i<t}R(t-t_i)[I_i(\tilde{z}^n_{t_i})-I_i(\tilde{z}_{t_i})]\right\|^2\\
&\quad\leq 2M^2Tr(Q)\int_0^t\mathbb{E}\left\|\xi(s,\tilde{z}^n_{\sigma(s,(\tilde{z}^n)_s)})-\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})\right\|^2\\
&\qquad+2M^2m\sum_{i=1}^m\mathbb{E}\left\|I_i(\tilde{z}^n_{t_i})-I_i(\tilde{z}_{t_i})\right\|^2\to 0\quad\text{as}\quad n\to\infty.
\end{aligned}$$

Consequently, we get

$$\lim_{n\to\infty}\|\Psi_2 z^n-\Psi_2 z\|^2_{\mathcal{PC}}=0,$$

and this proves that Ψ_2 is continuous.

(b) Ψ_2 maps bounded sets into bounded sets in $\mathbb{F}$.
For each $q>0$, let $B_q(0,\mathbb{F})=\{z\in\mathbb{F}:\mathbb{E}\|z\|^2\leq q\}$. Then, $B_q(0,\mathbb{F})$ is a bounded closed convex subset of $\mathbb{F}$. In fact, it suffices to show that there is a positive constant N_0 such that $\mathbb{E}\|\Psi_2 x\|^2\leq N_0$ for each $z\in B_q(0,\mathbb{F})$.
We set $q^*=2(\tilde{K}_2+J_0^{\varphi})^2\|\varphi\|^2_{\mathcal{B}}+2\tilde{K}_1^2 q$. From Lemma 2.10, (A_3) and (A_6), we have

$$\begin{aligned}
&\mathbb{E}\|(\Psi_2 z)(t)\|^2\\
&\quad\leq 2\mathbb{E}\left\|\int_0^t R(t-s)\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)\right\|^2+2\mathbb{E}\left\|\sum_{0<t_i<t}R(t-t_i)I_i(\tilde{z}_{t_i})\right\|^2\\
&\quad\leq 2M^2Tr(Q)\int_0^a\mathbb{E}\|\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})\|ds+2M^2m\sum_{i=0}^m\mathbb{E}\|I_i(\tilde{z}_{t_i})\|^2\\
&\quad\leq 2M^2Tr(Q)\int_0^a l(s)\mu_l(q^*)ds+2M^2m\sum_{i=1}^m\lambda_i q^*\\
&\quad\leq N_0,
\end{aligned}$$

which gives that, for each $z\in B_q(0,\mathbb{F})$, $E\|\Psi_2 z\|^2\leq N_0$.

Now it remains to show that $\Psi_2(B_q(0,\mathbb{F}))$ is equicontinuous and $\Psi_2(B_q(0,\mathbb{F}))(t)$ is precompact in $\mathbb{F}$. For this purpose, we decompose Ψ_2 as $\Upsilon_1+\Upsilon_2$, where Υ_1 and Υ_2 are the operators on $B_q(0,\mathbb{F})$ defined, respectively, by

$$(\Upsilon_1 z)(t)=\int_0^t R(t-s)\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)$$

and

$$(\Upsilon_2 z)(t)=\sum_{0<t_i<t}R(t-t_i)I_i(\tilde{z}_{t_i}).$$

(c) First, we show that $\Upsilon_1(B_q(0,\mathbb{F}))$ is equicontinuous and $\Upsilon_1(B_q(0,\mathbb{F}))(t)$ is relatively compact in $\mathbb{F}$. Let $0 < t_1 < t_2 \leq a$, for each $z \in B_q(0,\mathbb{F})$. Using (A_2) and (A_3), we obtain

$$\begin{aligned}
&\mathbb{E}\|(\Upsilon_1 z)(t_2) - (\Upsilon_1 z)(t_1)\|^2 \\
&\quad = \mathbb{E}\left\|\int_0^{t_2} R(t_2-s)\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s) - \int_0^{t_1} R(t_1-s)\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)\right\|^2 \\
&\quad \leq 2\mathbb{E}\left\|\int_0^{t_1}\big[R(t_2-s)-R(t_1-s)\big]\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)\right\|^2 \\
&\qquad + 2\mathbb{E}\left\|\int_{t_1}^{t_2} R(t_2-s)\xi(s,\tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)\right\|^2 \\
&\quad \leq 2Tr(Q)\int_0^{t_1} l(s)\mu_l(q^*)\|R(t_2-s)-R(t_1-s)\|^2 ds \\
&\qquad + 2Tr(Q)\int_{t_1}^{t_2} l(s)\mu_l(q^*)\|R(t_2-s)\|^2 ds \\
&\quad \leq 2Tr(Q)\mu_l(q^*)\int_0^{a} l(s)\|R(t_2-s)-R(t_1-s)\|^2 ds \\
&\qquad + 2Tr(Q)M^2\mu_l(q^*)\int_{t_1}^{t_2} l(s)ds.
\end{aligned}$$

Since $R(t)$ is continuous in the uniform operator topology, it follows that the right-hand side of the above inequality tends to zero and hence $\mathbb{E}\|(\Upsilon_1 z)(t_2) - (\Upsilon_1 z)(t_1)\|^2$ converges to zero independent of $z \in B_q(0,\mathbb{F})$ as $t_2 - t_1 \longrightarrow 0$. Thus the set $\{\Upsilon_1 z : z \in B_q(0,\mathbb{F})\}$ is equicontinuous. The equicontinuity for the other cases $t_1 < t_2 < 0$ or $t_1 \leq 0 \leq t_2 \leq a$ are very simple.

Next, we show the precompactness of $\Upsilon_1(B_q(0,\mathbb{F}))(t)$ in $\mathbb{F}$. By the virtue of the compactness of the resolvent operator $R(t)$ for $t > 0$ and the continuity of ξ, we see that the set

$$\{R(t-s)\xi(s,\theta) : s \in [0,a], \|\theta\|_{\mathcal{B}}^2 \leq q^*\}$$

is relatively compact in $\mathbb{X}$. Then, applying the mean value theorem for the Bochner integral, we get

$$(\Upsilon_1 z)(t) \in t\overline{conv}(\{R(t-s)\xi(s,\theta) : s \in [0,a], \|\theta\|_{\mathcal{B}}^2 \leq q^*\}),$$

which implies that $\{(\Upsilon_1 z)(t) : z \in B_q(0,\mathbb{F})\}$ is relatively compact in $\mathbb{F}$.

(d) Υ_2 is completely continuous.

We prove that Υ_2 is completely continuous. We can conclude that Υ_2 is continuous based on the proof in Step 3 (a). From the definition of Υ_2, for $q > 0,\ t \in [t_i, t_{i+1}], i = 1,2,\ldots,m$, and $z \in B_q(0,\mathbb{F})$, we find that

$$\overline{\Upsilon_2} z(t) \in \begin{cases} \sum_{j=1}^{i} R(t-t_j)I_j(B_{q^*}(0,\mathbb{X})), & t \in (t_i,t_{i+1}), \\ \sum_{j=1}^{i} R(t_{i+1}-t_j)I_j(B_{q^*}(0,\mathbb{X})) & \text{if} \quad t = t_{i+1}, \\ \sum_{j=1}^{i-1} R(t_i-t_j)I_j(B_{q^*}(0,\mathbb{X})) + I_i(B_{q^*}(0,\mathbb{X})) & \text{if} \quad t = t_i, \end{cases}$$

where $q^* = 2(\tilde{K}_2 + J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2 q$, which proves that $[\overline{\Upsilon_2(B_q)}]_i(t)$ is relatively compact in $\mathbb{F}$, for every $t \in [t_i, t_{i+1}]$, since the maps I_i are completely continuous for all $i = 1,2,\ldots,m.$ Moreover, using the compactness of the operators I_i and the assumption (A_2), we can prove that $[\overline{\Upsilon_2(B_q)}]_i$ is equicontinuous at t, for every $t \in [t_i, t_{i+1}]$. According to Lemma 2.8, we know that Υ_2 is completely continuous. As a result Ψ_2 is completely continuous. Hence, by Krasnoselskii-Schaefer fixed point theorem, we realize that Ψ has a fixed point on $\mathbb{F}$, which is a mild solution of Eq. (1.1). This completes the proof.

■

Instead of the assumption (A_6) discussed in Theorem 3.2, assume that the maps I_i satisfy some Lipschitz conditions. In this instance, we can also prove the existence of mild solutions. In addition, let us introduce the following condition:

(A_7) The maps I_i are completely continuous and there are positive constants b_i, c_i such that

$$\mathbb{E}\|I_i(z) - I_i(y)\|^2 \leq b_i\|z - y\|^2_{\mathcal{B}},$$

and $c_i = \sup_{t\in J} \mathbb{E}\|I_i(0)\|^2$, for $z, y \in \mathcal{B}, i = 1, 2, \ldots, m$.

Theorem 3.3. *Assume that* $(\mathbf{R_1})$-$(\mathbf{R_2})$,(A_1)-(A_5) *and* (A_7) *hold and* $z_0 \in \mathcal{L}^0_2(\Omega, \mathbb{X})$. *Then there exists a mild solution of Eq.* (1.1) *provided that*

$$8M^2\tilde{K}_1^2\bigg(2a^2d_2 + 4a^2d_2d_1 + 2m\sum_{i=1}^{m} b_i + Tr(Q)\Theta\int_0^a l(s)ds\bigg) < 1. \tag{3.6}$$

Proof. Let Ψ be the map defined as in the proof of Theorem 3.2. For better readability, we split the proof into two steps.

Step 1: $\Psi(B_q(0, \mathbb{F})) \subseteq B_q(0, \mathbb{F})$ for some $r > 0$.
We affirm that there exist a positive constant $r > 0$ such that $\Psi(B_q(0, \mathbb{F})) \subseteq B_q(0, \mathbb{F})$. We proceed by contradiction. Suppose that it is not true. Then for each $r > 0$, there exists a function $z^q(t^q) \in B_q(0, \mathbb{F})$ such that $\Psi(z^q) \notin B_q(0, \mathbb{F})$, i.e., $q < \mathbb{E}\|(\Psi z^q)(t^q)\|^2$ for some $t^q \in J$. Thus, from the assumptions we have

$$\begin{aligned}
q &< \mathbb{E}\|(\Psi z^q)(t^q)\|^2\\
&\leq 4\mathbb{E}\|R(t^q)\varphi(0)\|^2 + 4\mathbb{E}\bigg\|\int_0^{t^q} R(t^q - s)g(s, \tilde{z}^q_s, \int_0^{t^q} h(s, \tau, \tilde{z}^q_\tau)d\tau)ds\bigg\|^2\\
&+ 4\mathbb{E}\bigg\|\int_0^{t^q} R(t^q - s)\xi(s, \tilde{z}^q_{\sigma(s,\tilde{z}^q_s)})dw(s)\bigg\|^2 + 4\mathbb{E}\bigg\|\sum_{0<t_i<t} R(t^q - s)I_i(\tilde{z}^q_{t_i})\bigg\|^2\\
&\leq 4M^2H^2\|\varphi\|^2 + 8M^2a\int_0^{t^q} \big\{d_2\big[\|\tilde{z}^q_s\|^2_{\mathcal{B}} + 2d_1(\|\tilde{z}^q_s\|^2_{\mathcal{B}}) + 2d_1^*\big] + d_2^*\big\}\\
&+ 4M^2Tr(Q)\int_0^{t^q} \mathbb{E}\|\xi(s, \tilde{z}^q_{\sigma(s,\tilde{z}^q_s)})\|^2ds + 4M^2m\sum_{i=1}^{m}\mathbb{E}\|I_i(\tilde{z}^q_{t_i})\|^2\\
&\leq 4M^2H^2\|\varphi\|^2 + 8M^2a^2\big\{d_2\big[(C + 2\tilde{K}_1^2q) + 2d_1(C + 2\tilde{K}_1^2q) + 2d_1^*\big] + d_2^*\big\}\\
&+ 4M^2Tr(Q)\int_0^a l(s)\mu_l(C + 2\tilde{K}_1^2q)ds + 4M^2m\sum_{i=1}^{m}\big\{2b_i(C + 2\tilde{K}_1^2q) + 2c_i\big\}
\end{aligned}$$

where $C = 2(\tilde{K}_2 + J_0^\varphi)^2\|\varphi\|^2$. Dividing both sides by q and taking the limit as $q \to \infty$, we obtain

$$1 \leq 8M^2\tilde{K}_1^2\bigg(2a^2d_2 + 4a^2d_2d_1 + 2m\sum_{i=1}^{m} b_i + Tr(Q)\Theta\int_0^a l(s)ds\bigg)$$

which contradicts (3.6). Hence, for some positive number q, we have $\Psi(B_q(0, \mathbb{F})) \subseteq B_q(0, \mathbb{F})$.

Step 2: Ψ is a condensing map. Let $\Psi = \Psi_1 + \Psi_2$, where

$$(\Psi_1 x)(t) = R(t)\varphi(0) + \int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau)ds + \sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i})$$

$$(\Psi_2 z)(t) = \int_0^t R(t-s)\xi(s, \tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s),\ t \in J.$$

From the proof of Theorem 3.2, Ψ_2 is completely continuous on $B_q(0,\mathbb{F})$. Next, we have to show that Ψ_1 is a contraction map. Let $z, y \in B_q(0,\mathbb{F})$. Then, using hypotheses (A_2), (A_4), (A_5) and (A_7) we get

$$\begin{aligned}
&\mathbb{E}\|(\Psi_1 z)(t) - (\Psi_1 y)(t)\|^2 \\
&\leq \mathbb{E}\Big\| \int_0^t R(t-s)\big[g(s,\tilde{z}_s, \int_0^s h(s,\tau,\tilde{z}_\tau)d\tau) - g(s,\tilde{y}_s, \int_0^s h(s,\tau,\tilde{y}_\tau)d\tau)\big]ds \\
&\quad + \sum_{0<t_i<t} R(t-t_i)\big[I_i(\tilde{z}_{t_i}) - I_i(\tilde{y}_{t_i})\big]\Big\|^2 \\
&\leq 2M^2a^2\,(d_2 + d_1 d_2)\,\tilde{K}_1^2 \sup_{0\leq s\leq a} \mathbb{E}\|\tilde{z}(s) - \tilde{y}(s)\|^2 \\
&\quad + 2M^2 m \sum_{i=1}^m b_i \tilde{K}_1^2 \sup_{0\leq s\leq a} \mathbb{E}\|\tilde{z}(s) - \tilde{y}(s)\|^2.
\end{aligned}$$

It follows that

$$\begin{aligned}
\|\Psi_1 z - \Psi_1 y\|^2_{\mathcal{PC}} &\leq 2\tilde{K}_1^2\Big[M^2a^2(d_2 + d_1 d_2) + mM^2\sum_{i=1}^m b_i\Big]\|z-y\|_{\mathcal{PC}} \\
&= \kappa\|z-y\|_{\mathcal{PC}},
\end{aligned}$$

where $\kappa = 2\tilde{K}_1^2\left[M^2a^2(d_2 + d_1 d_2) + mM^2\sum_{i=1}^m b_i\right]$. By (3.6), we deduce that $\kappa < 1$, which yields that Ψ_1 is a contraction map. Considering Sadovskii's fixed point theorem, we conclude that there exists a fixed point for Ψ on $B_q(0,\mathbb{F})$, which is a mild solution for Eq. (1.1). ■

4. Controllability results

In this section, we examine the controllability of the following impulsive stochastic integro-differential equation with state-dependent delay:

$$\begin{cases} dz(t) = \Big[Az(t) + \int_0^t \Gamma(t-s)z(s)ds + B\vartheta(t) + g\left(t, z_t, \int_0^t h(t,s,z_s)ds\right)\Big]dt \\ \qquad\qquad +\xi(t, z_{\sigma(t,z_t)})dw(t),\ t \in J = [0,a], t \neq t_i, \\ \Delta z(t_i) = I_i(z_{t_i}), i = 1,\ldots,m, \\ z_0 = \varphi \in \mathcal{B}, \end{cases} \tag{4.1}$$

where g, h, A, ξ, I_i, are the same as in the Eq. (1.1). The control function $\vartheta(\cdot)$ takes its values in $L_2(J,\mathcal{U})$ of admissible control functions for a separable Hilbert space $\mathcal{U}$, and B is a bounded linear operator from $\mathcal{U}$ into $\mathbb{X}$. First, we give the definitions of mild solution and controllability for the system (4.1).

Definition 4.1. *A $\mathcal{F}_t$−adapted stochastic process $z : (-\infty, a] \longrightarrow \mathbb{X}$ is called a mild solution of the system* (4.1) *if $z_0 = \varphi \in \mathcal{B}$, $z_{\sigma(s,z_s)} \in \mathcal{B}$ satisfying $z_0 \in \mathcal{L}_2^0(\Omega,\mathbb{X})$, $z|_J \in \mathcal{PC}$. The function $R(t-s)g(s, z_s, \int_0^s h(s,\tau,z_\tau)d\tau)$ is integrable for each $s \in [0,a]$ and the following conditions hold:*

(i) *$\{z_t : t \in J\}$ is $\mathcal{B}-$valued and the restriction of $z(\cdot)$ to the interval $(t_i, t_{i+1}], i = 1, 2, \ldots, m$ is continuous ;*

(ii) $\Delta z(t_i) = I_i(z_{t_i}), i = 1, 2, \ldots, m;$

(iii) *for each $t \in J, z(t)$ satisfies the following integral equation*

$$z(t) = R(t)\varphi(0) + \int_0^t R(t-s)g(s, z_s, \int_0^s h(s, \tau, z_\tau)d\tau)ds + \int_0^t R(t-s)B\vartheta(s)ds + \int_0^t R(t-s)\xi(s, z_{\sigma(s,z_s)})dw(s) + \sum_{0<t_i<t} R(t-t_i)I_i(z_{t_i}).$$

Definition 4.2. *The system* (4.1) *is said to be controllable on the interval J, if for every initial function $z_0 = \varphi \in \mathcal{B}$, there exists a stochastic control $\vartheta \in L^2(J,\mathcal{U})$ that is adapted to the filtration $\{\mathcal{F}\}_{t\geq 0}$ such that the mild solution of the system* (4.1) *satisfies $z(a) = z_1$.*

We aim to transfer system (4.1) from $z(0)$ to $z(a) = z_1$. To achieve that purpose, we must assume:

(A_8) The linear operator $\mathcal{W} : L^2(J,\mathcal{U}) \longrightarrow \mathbb{X}$, defined by

$$\mathcal{W}_\vartheta = \int_0^a R(t-s)B\vartheta(s)ds,$$

has a bounded invertible operator $\mathcal{W}^{-1}$ which takes values in $L^2(J,\mathcal{U})/\ker \mathcal{W}$ and there exist positive constants M_1, M_2 such that $\|B\|^2 \leq M_1$ and $\|\mathcal{W}^{-1}\|^2 \leq M_2$.

(A_9) The function $\xi : J \times \mathcal{B} \longrightarrow \mathcal{L}_Q(\mathbb{Y}, \mathbb{X})$ is continuous and there exists constant $M_\xi > 0, \tilde{M}_\xi > 0$ for $z, y \in \mathcal{B}$ such that

$$\mathbb{E}\|\xi(t, z) - \xi(t, y)\|^2 \leq M_\xi\|z - y\|^2_{\mathcal{B}}$$

and $\tilde{M}_\xi = \sup_{t\in J} \mathbb{E}\|\xi(t, 0)\|^2$.

Theorem 4.3. *Assume that $(\mathbf{R_1}) - (\mathbf{R_2})$, $(A_4) - (A_5)$ and $(A_7) - (A_9)$ hold and $z_0 \in \mathcal{L}^0_2(\Omega, \mathbb{X})$. Then the system* (4.1) *is controllable provided that*

$$5G\tilde{K}_1^2 \left(1 + 5M^2M_1M_2a^2\right) \leq 1 \tag{4.2}$$

where $G = 4M^2a^2d^2 + 8d_1d_2M^2a^2 + 4M^2aTr(Q)M_\xi + 4M^2m\sum_{i=1}^m b_i$ *and* $M = \sup_{0\leq t\leq a} \|R(t)\|$.

Proof. Define the control process with terminal state $z_1 = z(a)$.

$$\vartheta^a_z(t) = \mathcal{W}^{-1}\Big\{z_1 - R(a)\varphi(0) - \int_0^a R(a-s)g(s, \tilde{z}_s, \int_0^s h(s, \tau, \tilde{z}_\tau)d\tau)ds - \int_0^a R(a-s)\xi(s, \tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s) - \sum_{i=1}^m R(a-t_i)I_i(\tilde{z}_{t_i})\Big\}(t).$$

Using this control, we define the operator $\varXi : \mathbb{F} \longrightarrow \mathbb{F}$ by

$$(\varXi z)(t) = R(t)\varphi(0) + \int_0^t R(t-s)g(s, \tilde{z}_s, \int_0^s h(s, \tau, \tilde{z}_\tau)d\tau)ds + \int_0^t R(t-s)B\vartheta^a_z(s)ds + \int_0^t R(t-s)\xi(s, \tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s) + \sum_{0<t_i<t} R(t-t_i)I_i(\tilde{z}_{t_i}).$$

where $\tilde{z} : (-\infty, a] \longrightarrow \mathbb{X}$ is such that $\tilde{z}_0 = \varphi$ and $\tilde{z} = z$ on J. From the assumptions, we know that the map Ξ is well defined and continuous.

Now, we prove that the operator Ξ has a fixed point in $\mathbb{F}$, which is a mild system solution (4.1). Observe that $(\Xi z)(a) = z_1$. This means that the control ϑ_z^a steers the system from φ to z_1 in finite time a, implying that the system (4.1) is controllable.

Let $q^* = 2(\tilde{K}_2 + J_0^{\varphi})^2\|\varphi\|_{\mathcal{B}}^2 + 2\tilde{K}_1^2 q$ for each $q > 0$. For any $z \in B_q(0, \mathbb{F})$ and $q > 0$, from the assumptions $(A_4) - (A_5)$ and $(A_7) - (A_9)$, we have

$$\begin{aligned}
&\mathbb{E}\|\vartheta_z^a(t)\|^2 \\
&\quad \leq 5\mathbb{E}\|\mathcal{W}^{-1}\|\bigg\{\mathbb{E}\|z_1\|^2 + \mathbb{E}\|R(t)\varphi(0)\|^2 \\
&\qquad + \mathbb{E}\Big\|\int_0^a R(a-s)g(s, \tilde{z}_s, \int_0^s h(s, \tau, \tilde{z}_\tau)d\tau)ds\Big\|^2 \\
&\qquad + \Big\|\int_0^a R(a-s)\xi(s, \tilde{z}_{\sigma(s,\tilde{z}_s)})dw(s)\Big\|^2 + \mathbb{E}\Big\|\sum_{0<t_i<t} R(a-t_i)I_i(\tilde{z}_{t_i})\Big\|^2\bigg\} \\
&\quad \leq 5M_2\bigg\{\mathbb{E}\|z_1\|^2 + M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 2M^2a^2\left(d_2q^* + 2d_1d_2q^* + 2d_1^*d_2 + d_2^*\right) \\
&\qquad + M^2aTr(Q)\left(2M_\xi q^* + 2\tilde{M}_\xi\right) + M^2m\sum_{i=1}^m (2b_iq^* + 2c_i)\bigg\} = \Omega.
\end{aligned}$$

Furthermore, for any $z, y \in B_q(0, \mathbb{F})$, we obtain

$$\begin{aligned}
&\mathbb{E}\|\vartheta_z^a(t) - \vartheta_y^a(t)\|^2 \\
&\quad \leq 3\tilde{K}_1^2M_2\bigg\{M^2a^2(d_2 + d_1d_2) + M^2aTr(Q)M_\xi + M^2m\sum_{i=1}^m b_i\bigg\}\|z-y\|_{\mathcal{PC}}^2.
\end{aligned}$$

For the sake of convenience, we break the proof into two steps.

Step 1: We show that Ξ maps $B_q(0, \mathbb{F})$ into itself.
It is enough to show that there exists a positive constant $q > 0$ such that $\Xi(B_q(0, \mathbb{F})) \subseteq B_q(0, \mathbb{F})$. Suppose that this assertion is false. Then for each > 0, there exists a function $z^q(t^q) \in B_q(0, \mathbb{F})$, such that $\Xi(z^q) \notin B_q(0, \mathbb{F})$, that is $q < \mathbb{E}\|(\Xi z^q)(t^q)\|^2$ for some $t^q \in J$. Thus, using hypotheses $(A_4) - (A_5)$ and $(A_7) - (A_9)$, we obtain

$$\begin{aligned}
q &< \mathbb{E}\|(\Xi z^q)(t^q)\|^2 \\
&\leq 5M^2H^2\|\varphi\|_{\mathcal{B}}^2 + 10M^2a^2\left(d_2q^* + 2d_1d_2q^* + 2d_1^*d_2 + d_2^*\right) \\
&\quad + 5M^2a^2M_1\Omega + 10M^2aTr(Q)\left(M_\xi q^* + M_\xi^*\right) + 10M^2m\sum_{i=1}^m (b_iq^* + c_i).
\end{aligned}$$

Dividing both sides by q and taking the limit as $q \to \infty$, we get

$$1 < 5G\tilde{K}_1^2\left(1 + 5M^2M_1a^2M_2\right),$$

where $G = 4M^2a^2d_2 + 8M^2a^2d_2d_1 + 4M^2aTr(Q)M_\xi + 4M^2m\sum_{i=1}^m b_i$, which is contrary to (4.2). Hence, for some positive number q, we have $\Xi(B_q(0, \mathbb{F})) \subseteq B_q(0, \mathbb{F})$.

Step 2: We prove that Ξ is a contraction operator. Let $z, y \in B_q(0, \mathbb{F})$, we obtain

$$
\begin{aligned}
&\mathbb{E}\|(\Xi z)(t) - (\Xi y)(t)\|^2 \\
&\quad \leq 4M^2a^2\left(d_2 + d_1d_2\right)\|z-y\|_{\mathcal{B}}^2 + 4M^2aTr(Q)M_\xi\|z-y\|_{\mathcal{B}}^2 \\
&\qquad + 4M^2m\sum_{i=1}^{m} b_i\|z-y\|_{\mathcal{B}}^2 \\
&\qquad + 12M^2M_1M_2a^2\big\{M^2a^2(d_2+d_1d_2) + M^2aTr(Q)M_\xi \\
&\qquad + M^2m\sum_{i=1}^{m} b_i\big\}\|z-y\|_{\mathcal{B}}^2 \\
&\quad \leq 4M^2a^2\tilde{K}_1^2\left(d_2 + d_1d_2\right)\|z-y\|_{\mathcal{PC}}^2 + 4M^2aTr(Q)\tilde{K}_1^2M_\xi\|z-y\|_{\mathcal{PC}}^2 \\
&\qquad + 4M^2m\tilde{K}_1^2\sum_{i=1}^{m} b_i\|z-y\|_{\mathcal{PC}}^2 \\
&\qquad + 12M^2M_1M_2a^2\tilde{K}_1^2\big\{M^2a^2(d_2+d_1d_2) + M^2aTr(Q)M_\xi \\
&\qquad + M^2m\sum_{i=1}^{m} b_i\big\}\|z-y\|_{\mathcal{PC}}^2 \\
&\qquad = 4G'\tilde{K}_1^2\left(1 + 3M_1M^2a^2M_2\right)\|z-y\|_{\mathcal{PC}}^2
\end{aligned}
$$

where $G' = M^2a^2(d_2 + d_1d_2) + M^2aTr(Q)M_\xi + M^2m\sum_{i=1}^{m} b_i$. Thanks to (4.2), we see that $4G'\tilde{K}_1^2\left(1 + 3M_1M^2a^2M_2\right) < 1$. Therefore Ξ is a contraction operator, and according to Banach's fixed point theorem it has a unique fixed point in $\mathbb{F}$, which is a mild solution Eq. (4.1). Thus, Eq. (4.1) is controllable. This completes the proof. ■

5. An example

To illustrate our obtained results, we consider the following impulsive stochastic integrodifferential equation with state-dependent delay of the form

$$
\begin{cases}
\dfrac{\partial}{\partial t}y(t,x) = \dfrac{\partial^2}{\partial x^2}y(t,x) + f\left(t, y(t-\tau,x), \displaystyle\int_0^t k(t,s,y(s-\tau,x))ds\right) \\
\qquad\qquad + \displaystyle\int_0^t u(t-s)\dfrac{\partial^2}{\partial x^2}y(s,x)ds \\
\qquad\qquad + \displaystyle\int_{-\infty}^t v(t,x,s-t)\mathcal{N}\left[y(s-\sigma_1(t)\sigma_2(\|y(t,x)\|),x)\right]dw(s), \\
0 \leq x \leq \pi, \tau > 0, t \in J = [0,a], \\
y(t,0) = y(t,\pi) = 0, t \in J, \\
y(\theta,x) = \varphi(\theta,x), \theta \in (-\infty,0], 0 \leq x \leq \pi, \\
\Delta y(t_i)(x) = \displaystyle\int_{-\infty}^{t_i} \alpha_i(t_i-s)y(s,x)ds, i = 1,2,\ldots,m, 0 \leq x \leq \pi,
\end{cases}
\tag{5.1}
$$

where $0 < t_1 < \cdots < t_m < a$ are prefixed numbers and $u : [0,\infty) \longrightarrow [0,\infty)$ is bounded and C^1-function such that u' is bounded and uniformly continuous,
$\sigma_1 : [0,\infty) \longrightarrow [0,\infty), \sigma_2 : [0,\infty) \longrightarrow [0,\infty)$ are continuous functions; $a > 0$; $f, k, v, \mathcal{N}$,
$I_i, (i = 1,2,\ldots,m)$, and φ are appropriate functions, which will be specified later.

To study this system, we consider the space $\mathbb{X} = \mathbb{Y} = L^2([0,a],\mathbb{R})$ and define the operator $A : D(A) \subset \mathbb{X} \longrightarrow \mathbb{X}$ as

$$\begin{cases} D(A) = H^2([0,\pi]) \cap H_0^1([0,\pi]) \\ A\kappa = \dfrac{\partial^2}{\partial x^2}\kappa. \end{cases}$$

Then, $A\kappa = -\sum_{n=1}^{\infty} n^2 \langle \kappa, s_n\rangle s_n,\ \kappa \in D(A)$, where $s_n(x) = \sqrt{\dfrac{2}{\pi}}\sin nx,\ n = 1,2,\ldots$ is the orthogonal basis of eigenvectors of A.
It is known that A is the infinitesimal generator of an analytic semigroup $\big(T(t)\big)_{t\geq 0}$ on $\mathbb{X}$, which is given by

$$T(t)z = \sum_{n=1}^{\infty} e^{-n^2 t}\langle z, s_n\rangle s_n \text{ for all } z \in \mathbb{X} \text{ and every } t \geq 0.$$

Therefore $(\mathbf{R_1})$ holds. In addition, the semigroup $\big(T(t)\big)_{t\geq 0}$ genered by A is compact for $t > 0$. Then by Theorem 2.4, the corresponding resolvent operator is also compact. Hence, (A_2) holds. Let $\varphi(t)(x) = \varphi(\theta,x), (\theta,x) \in (-\infty,0]\times[0,\pi], y(t)(x) = y(t,x)$. Assume that $q : (-\infty,0] \longrightarrow (0,\infty)$ is a Lebesgue integrable function with $l = \int_{-\infty}^{0}\tilde{q}(t)dt < \infty$. For any $a > 0$, define

$$\mathcal{B} = \big\{\zeta : (-\infty,0] \longrightarrow \mathbb{X} |\ (\mathbb{E}\|\zeta(\theta)\|^2)^{\frac{1}{2}} \quad \text{is a bounded and measurable function on } [-b,0] \\ \text{and} \quad \int_{-\infty}^{0} q(s)(\mathbb{E}\|\zeta(s)\|^2)^{\frac{1}{2}}ds < \infty\big\}.$$

Now, we take $q(t) = e^{2t}, t < 0$, then we get $p = \int_{-\infty}^{0} q(t)dt = \frac{1}{2}$ and

$$\|\zeta\|_{\mathcal{B}} = \int_{-\infty}^{0} q(s) \sup_{s\leq\theta\leq 0}\left(\mathbb{E}\|\zeta(\theta)\|^2\right)^{\frac{1}{2}} ds.$$

It is easy to verify that $(\mathcal{B}, \|\cdot\|_{\mathcal{B}})$ is a Banach space. In order to represent the system (5.1) to the abstract form (1.1), we define the functions $g : J\times\mathcal{B}\times\mathbb{X} \longrightarrow \mathbb{X}$, $\xi : J\times\mathcal{B} \longrightarrow \mathcal{L}_Q(\mathbb{Y},\mathbb{X})$, $\sigma : J\times\mathcal{B} \longrightarrow (-\infty,0]$, $I_k : \mathcal{B} \longrightarrow \mathbb{X}$ respectively by

$$\begin{aligned} g\left(t,\zeta,\int_0^t h(t,s,\zeta)ds\right)(x) &= f\left(t,\zeta(\theta,x),\int_0^t k(t,s,\zeta(\theta,x)ds\right) \\ &= \int_{-\infty}^{0}\gamma(\theta)\zeta(\theta)(x)d\theta + \int_0^t\int_{-\infty}^{0}\beta_1(s)\beta_2(\tau)\zeta(\tau,x)d\tau ds, \\ \xi(t,\zeta)(x) &= \int_{-\infty}^{0} v(t,x,\theta)\mathcal{N}(\zeta(\theta,x))d\theta, \\ \sigma(\theta,\zeta) &= \theta - \sigma_1(\theta)\sigma_2(\|\zeta(0)\|), \\ I_i(\zeta)(x) &= \int_{-\infty}^{0}\alpha_i(-\theta)\zeta(\theta,x)d\theta, i = 1,2,\ldots,m. \end{aligned}$$

On the other hand, let $\Gamma : D(A) \subset \mathbb{X} \to \mathbb{X}$ be the operator defined by

$$\Gamma(t)x = u(t)Ax, \text{ for } t \geq 0 \text{ and } x \in D(A).$$

Under these definition, system (5.1) is then rewriten in the following form

$$\begin{cases} dz(t) = \Big[Az(t) + \displaystyle\int_0^t \Gamma(t-s)z(s)ds + g\left(t,z_t,\int_0^t h(t,s,z_s)ds\right)\Big]dt \\ \qquad\qquad +\xi(t,z_{\sigma(t,z_t)})dw(t),\ t\in J = [0,a], t\neq t_i, \\ \Delta z(t_i) = I_i(z_{t_i}), i = 1,\ldots,m \\ z_0 = \varphi \in \mathcal{B}. \end{cases} \tag{5.2}$$

Since u is bounded and C^1-function such that $u^{'}$ is bounded and uniformly continuous, ($\mathbf{R_2}$) is fulfilled. Hence, by Theorem 2.2, Eq. (2.1) has a unique resolvent operator $\big(R(t)\big)_{t\geq 0}$ on $\mathbb{X}$, which is also operator-norm continuous for $t \geq 0$ thanks to Theorem 2.4.

To establish the existence result for the mild solution of (5.1), we need the following conditions:

(i) The function $\gamma(\theta) \geq 0$ is continuous in $(-\infty, 0]$ satisfying

$$\int_{-\infty}^{0} \gamma^2(\theta)d\theta < \infty, \gamma_g = \left(\int_{-\infty}^{0} \frac{(\gamma(s))^2}{q(s)} ds\right)^{\frac{1}{2}} < \infty.$$

(ii) $\beta_1, \beta_2 : \mathbb{R} \longrightarrow \mathbb{R}$ are continuous and

$$\gamma_g^* = \left(\int_{-\infty}^{0} \frac{(\beta_2(s))^2}{q(s)} ds\right)^{\frac{1}{2}} < \infty.$$

(iii) the functions $\alpha_i \in C(\mathbb{R}, \mathbb{R})$ and

$$c_i = \left(\int_{-\infty}^{0} \frac{\alpha_i^2(-\theta)}{q(s)} d\theta\right)^{\frac{1}{2}} < \infty, i = 1, 2, \ldots, m,$$

(iv) The function $v_2(t, x, \theta) \leq 0$ is continuous on $J \times [0, 2\pi] \times (-\infty, 0]$ and satisfies

$$\int_{-\infty}^{0} v(t, x, \theta)d\theta = \delta(t, x) < \infty.$$

(v) The function $\mathcal{N}(\cdot)$ is continuous and satisfies $0 \leq \mathcal{N}(y(\theta, x)) \leq \mu_l \left(\int_{-\infty}^{0} e^{2s} \|y(s, \cdot)\|_{\mathcal{L}_2} ds\right)$ for $(\theta, x) \in (-\infty, 0] \times [0, 2\pi]$, where $\mu_l(\cdot)$ is positive, continuous and nondecreasing in $[0, \infty)$.

Under the above assumptions, we obtain

$$\begin{aligned}
\|I_i(\zeta)\|_{\mathcal{L}_2} &= \left[\int_0^{\pi} \mathbb{E}\Big\| \int_{-\infty}^{0} \alpha_i(-\theta)\zeta(\theta, x)d\theta\Big\|^2 dx\right]^{\frac{1}{2}} \\
&= \left[\int_0^{\pi} \mathbb{E}\Big\| \int_{-\infty}^{0} \frac{\alpha_i(-\theta)}{q^{\frac{1}{2}}(\theta)} q^{\frac{1}{2}}(\theta)\zeta(\theta, x)d\theta\Big\|^2 dx\right]^{\frac{1}{2}} \\
&\leq \left[\int_0^{\pi} \Big(\int_{-\infty}^{0} \frac{\alpha_i^2(-\theta)}{q(\theta)} d\theta\Big)\Big(\int_{-\infty}^{0} q(\theta)\mathbb{E}\|\zeta(\theta, x)\|^2 d\theta\Big) dx\right]^{\frac{1}{2}} \\
&= \Big(\int_{-\infty}^{0} \frac{\alpha_i^2(-\theta)}{q(\theta)} d\theta\Big)^{\frac{1}{2}} \Big[\int_0^{\pi}\int_{-\infty}^{0} q(\theta)\mathbb{E}\|\zeta(\theta, x)\|^2 d\theta dx\Big]^{\frac{1}{2}} \\
&\leq \Big(\int_{-\infty}^{0} \frac{\alpha_i^2(-\theta)}{q(\theta)} d\theta\Big)^{\frac{1}{2}} \Big[\int_{-\infty}^{0} q(\theta) \int_0^{\pi} \mathbb{E}\|\zeta(\theta, x)\|^2 dx d\theta\Big]^{\frac{1}{2}} \\
&\leq c_i \Big[\int_{-\infty}^{0} q(\theta) \sup_{s\leq\theta\leq 0} \mathbb{E}\|\zeta(\theta)\|^2 d\theta\Big]^{\frac{1}{2}} \\
&\leq c_i \|\zeta\|_{\mathcal{B}},
\end{aligned}$$

$$
\begin{aligned}
&\|g(t,\zeta,w)\|_{\mathcal{L}_2}\\
&\quad= \Big[\int_0^{\pi}\mathbb{E}\Big\|\int_{-\infty}^{0}\gamma(\theta)\zeta(\theta)(x)d\theta+\int_0^t\int_{-\infty}^{0}\beta_1(s)\beta_2(\tau)\zeta(\tau,x)d\tau ds\Big\|^2dx\Big]^{\frac{1}{2}}\\
&\quad\le \Big[\int_0^{\pi}\mathbb{E}\Big\|\int_{-\infty}^{0}\gamma(\theta)\zeta(\theta)(x)d\theta\Big\|^2dx\Big]^{\frac{1}{2}}\\
&\qquad+\Big[\int_0^{\pi}\mathbb{E}\Big\|\int_0^t\int_{-\infty}^{0}\beta_1(s)\beta_2(\tau)\zeta(\tau,x)d\tau ds\Big\|^2dx\Big]^{\frac{1}{2}}\\
&\quad\le \gamma_g\|\zeta\|_{\mathcal{B}}+a\Big[\int_0^{2\pi}\big(\int_0^t\beta_1^2(s)ds\big)\big(\mathbb{E}\Big\|\int_{-\infty}^{0}\beta_2(\tau)\zeta(\tau,x)d\tau\Big\|^2\big)dx\Big]^{\frac{1}{2}}\\
&\quad\le \gamma_g\|\zeta\|_{\mathcal{B}}+a\big(\int_0^t\beta_1^2(s)ds\big)^{\frac{1}{2}}\big[\int_0^{\pi}\mathbb{E}\Big\|\int_{-\infty}^{0}\beta_2(\tau)\zeta(\tau,x)d\tau\Big\|^2dx\big]^{\frac{1}{2}}\\
&\quad\le \gamma_g\|\zeta\|_{\mathcal{B}}+a\big(\int_0^t\beta_1^2(s)ds\big)^{\frac{1}{2}}\gamma_g^*\|\zeta\|_{\mathcal{B}}\\
&\quad= \big[\gamma_g+a\|\beta_1\|_\infty\gamma_g^*\big]\|\zeta\|_{\mathcal{B}},
\end{aligned}
$$

$$
\begin{aligned}
\|\xi(t,\zeta)\|_{\mathcal{L}_2} &= \Big[\int_0^{\pi}\mathbb{E}\Big\|\int_{-\infty}^{0}v(t,x,\theta)\mathcal{N}(\zeta(\theta,x))d\theta\Big\|^2dx\Big]^{\frac{1}{2}}\\
&\le \Big[\int_0^{\pi}\mathbb{E}\Big\|\int_{-\infty}^{0}v(t,x,\theta)\mu_l\left(\int_{-\infty}^{0}e^{2s}\|\zeta(s)(\cdot)\|_{\mathcal{L}_2}ds\right)d\theta\Big\|^2dx\Big]^{\frac{1}{2}}\\
&\le \Big[\int_0^{\pi}\big(\int_{-\infty}^{0}v(t,x,\theta)\mu_l\left(\int_{-\infty}^{0}e^{2s}\sup\|\zeta(s)(\cdot)\|_{\mathcal{L}_2}ds\right)d\theta\big)^2dx\Big]^{\frac{1}{2}}\\
&\le \Big[\int_0^{\pi}\big(\int_{-\infty}^{0}v(t,x,\theta)d\theta\big)^2dx\Big]^{\frac{1}{2}}\mu_l\left(\|\zeta\|_{\mathcal{B}}\right)\\
&= \Big[\int_0^{\pi}\delta^2(t,x)dx\Big]^{\frac{1}{2}}\mu_l(\|\zeta\|_{\mathcal{B}})\\
&= l(t)\mu_l(\|\zeta\|_{\mathcal{B}}),
\end{aligned}
$$

where $l(t)=\left(\int_0^{\pi}\delta^2(t,x)dx\right)^{\frac{1}{2}}$. Therefore, $g, I_i(i=1,2,\cdots,m)$ are bounded by $\mathbb{E}\|g\|_{\mathbb{X}}^2\le L_2, \mathbb{E}\|I_i\|_{\mathbb{X}}^2\le c_i^2$, where $L_2=[\gamma_g+a\|\beta_1\|_\infty\gamma_g^*]^2$. In addition, from the estimation of $\xi(t,\zeta)$, it is easy to see that the function ξ satisfy the hypothesis (A_3). Hence by Theorem 3.3, the system (5.1) has a mild solution on J.

Conclusion

This article focuses on a new kind of state-dependent delay neutrality of impulsive stochastic integrodifferential equations in a real separable Hilbert space. We obtained the existence and controllability of mild solutions using the fixed point theorems and resolvent operator theory in the sense of Grimmer. We provided an example to show the effectiveness of the main results.In addition, to obtain the immediate results discussed in this paper, the Krasnoselskii–Schaefer fixed point theorem, the Sadovskii fixed point theorem, and the Banach fixed point theorem were all successfully applied under a variety of distinct conditions. In upcoming research, we will investigate the controllability and stability of solutions for impulsive stochastic integrodifferential systems that either have jumps in their dynamics or are driven by the Rosenblatt process. This research will take place shortly.

6. Acknowledgement

The author is thankful to the editors and the referees for his valuable suggestions which improved the presentation of the paper.

References

[1] F. Andrade, C. Cuevas, H.R. Henríquez, Periodic solutions of abstract functional differential equations with state-dependent delay, *Math Methods Appl Sci.*, **39** (2016), 3897–3909.

[2] K. Balachandran, J.P. Dauer, Controllability of nonlinear systems via Fixed Point Theorems, *J. Optim. Theory Appl.*, **53** (1987), 345–352.

[3] D.D. Bainov, V. Lakshmikantham, and P. S. Simeonov, *Theory of Impulsive Differential Equations*, World Scientic, Singapore, 1989.

[4] C.T.H. Baker, C.A.H. Paul; *Pitfalls in parameter estimation for delay differential equations*. SIAM J. Sei. Comp. 18(1997), 305–314.

[5] K. Balachandran, J.P. Dauer, Controllability of nonlinear systems in Banach spaces, *A survey. J. Optim. Theory Appl.*, **115** (2002), 7–28.

[6] I.M. Barka, M.A. Diop, K. Ezzinbi and M.H.M. Hamit, Controllability for nonlocal stochastic integrodifferential evolution equations with the lack of compactness, *Stochastic Analysis and Applications*, (2021).

[7] M.Benchohra, J. Henderson, and S.K. Ntouyas, *Impulsive Differential Equations and Inclusions*, Hindawi Publishing Corporation, New York, Vol. 2, 2006.

[8] T.A. Burton, C. Kirk, A fixed point theorem of Krasnoselskii-Schaefer type, *Math. Nachr.*, **189** (1998), 23–31.

[9] E.N. Chukwu, S.M. Lenhart Controllability questions for nonlinear systems in abstract spaces, *J. Optim. Theory Appl.*, **68** (1991), 437–462.

[10] G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge 1992.

[11] W. Desh, R.C. Grimmer and W. Schappacher, Some considerations for linear integrodifferential equations, *Journal of Mathematical analysis and Applications*, **104** (1984), 219–234.

[12] M.A. Diop, M.M. Zene, On the asymptotic stability of impulsive neutral stochastic partial integro-differential equations with variable delays and Poisson jumps, *Afrika Matematika*, **27(1-2)** (2016), 215–228.

[13] M.A. Diop, A.T.G. Amoussou, C. Ogouyandjou, M.S. Mohamed, Existence results for an neutral integrodifferential stochastic functional differential impulsive equations with state-dependent delay, *Transactions of A. Razmadze Mathematical Institute*, **173** (2019), 17–31.

[14] M.A. Diop, P.A.D. Guindo, M. Fall, A. Diakhaby, Optimal controls for stochastic functional integrodifferential equations, *Elec. J. of Math. Anal. and Appl.*, **9** (2021), 241–261.

[15] R.D. DRIVER, Existence Theory for a Delay Differential System, *Contrib. Diff.* **1** (1963), 317–336.

[16] R.D. DRIVER, *Ordinary and Delay Differential Equations*, Applied Mathematical Sciences 20, Springer: New York, NY, USA, 1977.

[17] K. EZZINBI, G. DEGLA, P. NDAMBOMVE, Controllability for some Partial Functional Integrodifferential Equations with Nonlocal Conditions in Banach Spaces, *Discussiones Mathematicae, Differential Inclusions, Control and Optimization*, **35(1)** (2015),1–22.

[18] M. FALL, M.D. DIDIYA, A.W. GNONLONFOUN, M. A. DIOP, On Impulsive Integrodifferential Equations With State Dependent Delay, *Journal of Numerical Mathematics and Stochastics*, **13(1)** (2022), 29–54.

[19] L. GAWARECKI, V. MANDREKAR, Stochastic Differential Equations in Infinite Dimensions with Applications to Stochastic Partial Differential Equations, *Spinger, Berlin, Heidelberg*, (2011).

[20] J.R. GRAEF, J. HENDERSON, A. OUAHAB, *Impulsive Differential Inclusions. A Fixed Point Approach*, De Gruyter Series in Nonlinear Analysis and Applications 20, Berlin: de Gruyter, (2013).

[21] J.R. GRAEF, J. HENDERSON, A. OUAHAB, *Topological Methods for Differential Equations and Inclusions*, Monographs and Research Notes in Mathematics, Boca Raton, FL: CRC Press, (2019).

[22] R.C GRIMMER, Resolvent operators for integral equations in a Banach space, *Transactions of the American Mathematical Society*, **273**, (1982), 333–349.

[23] R. GRIMMER, A.J. PRITCHARD, Analytic resolvent operators for integral equations in a Banach space, *Journal of Differential Equations*, **50** (1983), 234–259.

[24] A. HALANAY, D. WEXLER, Teoria Calitativa a sistemelor Impulsuri, *(in Romanian), Editura Academiei Republicii Socialiste România, Bucharest, Romanian*, (1968).

[25] J.K. HALE, J. KATO, Phase spaces for retarded equations with infinite delay, *Funkc. Ekvacioj*, **21** (1978), 11–41

[26] E. HERNANDEZ, M. PIERRI, G. GONCALVES, Existence results for a impulsive abstract partial differential equation with state-dependent delay, *Comput. Math. Appl.*, **52** (2006), 411–420.

[27] E. HERNÁNDEZ, A. PROKOPCZYK, L. LADEIRA A note on partial functional differential equations with state-dependent delay, *Nonlinear Analysis:RWA*, 7(4)(2006), 510–519.

[28] E. HERNÁNDEZ, R. SAKTHIVEL, S.T. AKI, Existence results for impulsive evolution differential equations with state-dependent delay, *Electron. J. Differ. Equ.*, **2008** (2008), 1–11.

[29] H.R. HENRIQUEZ, Existence of solutions of non-autonomous second order functional differential equations with infinite delay, *Nonlinear Anal.*, **74** (2011), 3333–3352.

[30] E. HERNÁNDEZ, M. PIERRI, S-asymptotically periodic solutions for abstract equations with state-dependent delay, *Bull Aust Math Soc.*, **98(3)** (2018), 456–464.

[31] E. HERNÁNDEZ, J. WU, Existence, Uniqueness and Qualitative Properties of Global Solutions of Abstract Differential Equations with State-Dependent Delay, *Proceedings of the Edinburgh Mathematical Society*, **62(3)** (2019), 771–788.

[32] H. MA, B. LIU, Exact controllability and continuous dependence of fractional neutral integro-differential equations with state dependent delay, *Acta Mathematica Scientia. Series B*, **37(1)** (2017), 235–258.

[33] C.M. MARLE, Measures et Probabilités, *Hermann, Paris* (1974).

[34] V.D. MILMAN, A.A. MYSHKIS, On the stability of motion in the presence of impulses, *Sib. Math. J., (in Russian)* **1**, (1960), 233–237.

[35] H. OKA, Integrated resolvent operators, *Journal of Integral Equations*, **7** (1995) 193–232.

[36] A. PAZY, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer, New York, NY, USA, (1983).

[37] K. RAVIKUMAR, A.T.G. AMOUSSOU, C. OGOUYANDJOU, M.A. DIOP, Existence and uniqueness of mild solutions of stochastic partial integro-differential impulsive equations with infinite delay via resolvent operator, *Advances in Dynamical Systems and Applications*, **14(1)** (2019), 83–118.

[38] B.N. SADOVSKII, On a fixed point principle, *Funct. Anal. Appl.*, **1** (1967) , 71–74.

[39] R. SAKTHIVEL, Q.H. CHOI, S.M. ANTHONI, *Controllability result for nonlinear evolution integrodifferential systems*, *Applied Mathematics Letters*, **17** (2004) 1015–1023.

[40] A.M. SAMOILENKO, N.A. PERESTYUK, *Impulsive Differential Equations*, World Scientic, Singapore, (1995).

[41] A. SLAMA, A. BOUDAOUI, Approximate controllability of fractional nonlinear neutral stochastic differential inclusion with nonlocal conditions and infinite delay, *Arab. J. Math.*, **6** (2017) 31–54.

[42] N. TUAN, L. NGUYEN, On solutions of a system of hereditary and self-referred partial-differential equations, *Numer. Algorithms*, **55** (2010), 101–113.

[43] U. VAN LE, L. NGUYEN, Existence of solutions for systems of self-referred and hereditary differential equations, *Electron. J. Differ. Equ.,* **51** (2008), 1–7.

[44] Z. YAN, Controllability of semilinear integro-differential systems with nonlocal conditions, *Int. J. Comput. Appl. Math.*, **3** (2007), 221–236.

[45] Z. YAN, X. YAN , Existence of solutions for impulsive partial stochastic neutral integrodifferential equations with state-dependent delay, *Collect. Math*. **64** (2013), 235–250.

[46] X. ZHANG, Exact controllability of semilinear evolution systems and its application, *J. Optim. Theory Appl.*, **107** (2000), 415–432.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 66–79.
http://doi.org/10.26637/mjm1101/005

Existence and uniqueness of classical and mild solutions of fractional Cauchy problem with impulses

VISHANT SHAH*[1], JAITA SHARMA[1] AND RAJU K. GEORGE[2]

[1] *Faculty of Techonology & Engineering, Department of Applied Mathematics, The Maharaja Sayajirao University of Baroda, Vadodara, Gujarat, India.*
[2] *Department of Mathematics, Indian Institute of Space Science and Technology, Thiruvananthapuram, Kerala, India.*

***R**eceived 04 April 2022; **A**ccepted 11 November 2022*

Abstract. In this manuscript, we have established conditions for the existence and uniqueness of mild and classical solutions to the fractional order Cauchy problem by including and without including impulses over the completed norm linear space (Banach space). Conditions are established using the concept of generators and the generalised Banach fixed point theorem, which are weaker conditions than the previously derived conditions. We have also established the conditions under which a mild solution to the problem gives rise to a classical solution to the given problem. Finally, illustrations of the existence and uniqueness of the solution are provided to validate our derived results.

AMS Subject Classifications: 47D06, 34A12, 34A08, 35A01, 31C25.

Keywords: Existence of solution, Fractional evolution equation, Fixed point theorem, Impulsive conditions.

Contents

1. Introduction

The various problems in physics, engineering, and biological sciences that have abrupt changes for a small amount of time are well explained in terms of impulses. Therefore, problems like removal of insertion of biomass, populations of species with abrupt changes, abrupt harvesting, and various problems containing abrupt changes are modelled into impulsive differential equations[3, 9, 13, 19, 23–25, 28, 34]. Many researchers have studied the qualitative properties like existence, uniqueness, and asymptotic behaviour of impulsive differential equations using various techniques. These studies are found in the articles cited [1, 2, 15, 26, 35, 37] and reference their in.

*Corresponding author. Email address: **vishantmsu83@gmail.com** (Vishant Shah)

On the other hand, due to the inherited property of the fractional derivative operator [6, 8], many nonlinear complicated problems, such as seepage flow in porous media, anomalous diffusion, wave and transport, and many other problems, are now being remodelled into fractional differential equations [12, 14, 16–18, 27, 30, 31, 33, 38]. Fractional calculus developed become one of the most well-liked areas of applied mathematics as a result of the numerous uses of fractional differential equations. This draws a lot of academics interested in differential equations and fractional calculus. Numerous scholars, including [10, 11, 22, 36], have examined the qualitative properties, such as existence and uniqueness of mild solutions to fractional equations using diverse methodologies. Researchers have looked into the existence and originality of impulsive fractional differential equations. Benchohra and Slimani[5] investigated the presence and distinctiveness of a mild solution to impulsive differential equations in one dimension. To find adequate criteria for the existence and uniqueness of the mild solution, they employed the fixed point theorems of Banach, Schaefer, and Leray-Schauder. With the use of Banach contraction principle and semigroup theory, Mohphu [29] researched the existence and uniqueness of mild solutions. By assuming the sectroial property of the linear operator A, Ravichandran and Arjunan [32] investigated the existence and uniqueness of the classical and mild solutions of impulsive fractional integro-differential equations on Banach space. Balachandra et al. By omitting the semigroup property from Mohphu's work, al. [4] examined the existence and uniqueness of mild solutions to impulsive fractional integro-differential equations on a Banach space. The classical solution of a fractional order differential equation of the Caputo type is described by Kataria and Patel [20], who also examine the congruence between the classical and mild solutions of more extended impulsive fractional equations on a Banach space.

Krasnoselskii's fixed point was utilised by Borah and Bora [7] and Kataria et al. [21] to demonstrate the necessary conditions for the existence of mild solutions for the non-local fractional differential equations with non-instantaneous impulses.

In this paper, we develop necessary criteria for mild solution and classical solution of the impulsive fractional evolution problem,

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)) \quad t \neq t_k,\ k = 1, 2, \cdots, p \\ \Delta u(t_k) &= I_k(u(t_k)), \quad t = t_k,\ k = 1, 2, \cdots, p \\ u(t_0) &= u_0 \end{aligned} \tag{1.1}$$

over the interval $[0, T_0]$ on a Banach space $\mathcal{U}$. Here, ${}^cD^\alpha$ denotes Caputo fractional differential operator of order $0 < \alpha \leq 1$, $A : \mathcal{U} \to \mathcal{U}$ is linear operator and $f : [0, T_0] \times \mathcal{U} \to$ is nonlinear function. $I_k : \mathcal{U} \to \mathcal{U}$ are impulse operator at time $t = t_k$, fro $k = 1, 2, \cdots, p$ and their existence and uniqueness. We also developed conditions underwhich classical solution and mild solution of (1.1) are coincide.

The outline of the article is as follows: In section-2, we discussed some preliminaries from fractional calculus followed by motivation to study in section-3. Section-4 and section-5 discusses the existence and uniqueness results of fractional evolution equation without and with impulses followed by conclusion in section-6.

2. Prelimnaries

In this section we introduce notations, definitions, assumptions preliminary facts which are used throughout this paper.

Definition 2.1. *([22, 30]) The Riemann-Liouville fractional integral operator of $\alpha > 0$, of function $f \in L_1(\mathbb{R}_+)$ is defined as*

$$I^\alpha_{t_0+} f(t) = \frac{1}{\Gamma(\alpha)} \int_{t_0}^t (t-s)^{\alpha-1} f(s) ds,$$

provided the integral on right side exist. Where $\Gamma(\cdot)$ is gamma function.

Definition 2.2. *([22, 30]) The Caputo fractional derivative of order $\alpha > 0$, $n-1 < \alpha < n$, $n \in \mathbb{N}$, is defined as*

$${}^cD^\alpha_{t_0+} f(t) = \frac{1}{\Gamma(n-\alpha)} \int_0^t (t-s)^{n-\alpha-1} \frac{d^n f(s)}{ds^n} ds,$$

provided the integral on the right exist and $n = [\alpha] + 1$.

Definition 2.3. *One and two parameter Mittag-Lefflar function is defined as:*

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}, \quad E_{\alpha,\beta}(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + \beta)}$$

for all $\alpha, \beta > 0$ *and* $z \in \mathbb{C}$ *respectively.*

Definition 2.4. *[37] Let* X *be Banach space. Then the set*
$PC([t_0, T], X) = \Big\{u : [t_0, T] \to X; u \ is \ continuous \ at \ t \neq t_k, \ left \ continuous \ at \ t = t_k, \ right \ limit \ at \ t = t_k \ exist \ for \ all \ k = 1, 2, \cdots, p\Big\}.$
This set $PC([t_0, T], X)$ *is Banach space under the norm defined by* $||u||_{PC} = sup\big\{||u(t)||; \ t \in [t_0, T]\big\}$.

3. Motivation

This section is devoted to motivation behind studying the existence and uniqueness of solution for the Caputo Cauchy problem. Consider the non-homogeneous diffusion equation without impulses

$$\begin{aligned} {}^cD^\alpha u(t,x) &= u_{xx}(t,x) + F(t,x), \\ u(t,0) &= u(t,\pi) = 0, \\ u(0,x) &= u_0(x) \end{aligned} \tag{3.1}$$

over the rectangle $[0, T_0] \times [0, \pi]$. The solution of this equation (3.1) using the Laplace transform and Fourier series at any time $t \in [0, T_0]$ is given by

$$u(t,x) = T_\alpha(t)u_0(x) + \int_0^t (t-s)^{\alpha-1} T_{\alpha,\alpha}(t-s)F(s,x)ds \tag{3.2}$$

where, the families of operators $T_\alpha(t), T_{\alpha,\beta}(t) : \mathcal{U} \to \mathcal{U}$ for all $t \in [0, T_0]$ are defined as

$$T_\alpha(t)z = \sum_{n=1}^{\infty} E_\alpha(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

and

$$T_{\alpha,\beta}(t)z = \sum_{n=1}^{\infty} E_{\alpha,\beta}(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

in the space

$$\mathcal{U} = \big\{z : [0, \pi] \to \mathbb{R} : z'' \ exists \ and \ z(0) = z(\pi) = 0\big\}$$

the functions $E_\alpha(\cdot)$ and $E_{\alpha,\beta}(\cdot)$ are Mittag-Leffler functions of one and two parameter family respectively and $\phi_n(x)$ are orthonormal Fourier basis corresponding to eigenvalues.

In view of the equation (3.2) we can define mild solution of semi-linear diffusion equation

$$\begin{aligned} {}^cD^\alpha u(t,x) &= u_{xx}(t,x) + F(t,u), \\ u(t,0) &= u(t,\pi) = 0, \\ u(0,x) &= u_0(x) \end{aligned} \tag{3.3}$$

as a function u satisfy the equation

$$u(t,x) = T_\alpha(t)u_0(x) + \int_0^t (t-s)^{\alpha-1} T_{\alpha,\alpha}(t-s)F(s,u)ds \tag{3.4}$$

where, the families of operators $T_\alpha(t)$, $T_{\alpha,\beta}(t)$ are defined above.

Observe that the operator A in equation (3.3) is neither bounded nor semigroup property but solutions of equation (3.3) exist under certain conditions (derived in the Section -4). From this we can say that the function u is the mild solution of diffusion equation (3.3) if u satisfies the integral equation (3.4). Using this concept we can easily study the various qualitative properties like existence and uniqueness of solution, various types of stability and controllability of the Caputo fractional evolution system (1.1) with and without impulses. This motivates to study existence and uniqueness of solutions of Caputo fractional evolution equation (1.1).

4. Mild and classical solutions without impulses

In this section, we are going to discuss existence and uniqueness of classical and mild solutions of the fractional order evolution equation (1.1) without impulses by using the concept of generators, motivated from the previous section.
Consider the fractional order evolution equation without impulses over the interval $[0, T_0]$ of the form:

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(0) &= u_0 \end{aligned} \tag{4.1}$$

in the general Banach space $\mathcal{U}$, where $A : \mathcal{U} \to \mathcal{U}$ is linear operator, ${}^cD^\alpha$ is fractional differential operator of Caputo type for $0 < \alpha \leq 1$ and $F : [0, T_0] \times \mathcal{U} \to \mathcal{U}$ is nonlinear function.
We define the operators which is generated by the linear operator A.

Definition 4.1. *The families of operators $T_\alpha(t), T_{\alpha,\beta}(t) : \mathcal{U} \to \mathcal{U},\ t \geq 0$ are generated by a linear operator $A : \mathcal{U} \to \mathcal{U}$ satisfies the following properties:*

(1) $T_\alpha(0) = I$ and $T_{\alpha,\beta}(0) = I$ where, I is identity operator

(2) $T(t)$ satisfies the linear fractional equation ${}^cD^\alpha u(t) = A(t)u(t)$ in Banach space $\mathcal{U}$

(3) $\lim_{\beta\to 1} T_{\alpha,\beta}(t) = T_\beta(t)$

Example 4.2. *The operators $T_\alpha(t), T_{\alpha,\beta}(t) : \mathcal{U} \to \mathcal{U}$ for all $t \in [0, T]$ are defined as*

$$T_\alpha(t)z = \sum_{n=1}^{\infty} E_\alpha(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

and

$$T_{\alpha,\beta}(t)z = \sum_{n=1}^{\infty} E_{\alpha,\beta}(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

defined on the space

$$\mathcal{U} = \left\{ z : [0,\pi] \to \mathbb{R} : z'' \ exists \ and \ z(0) = z(\pi) = 0 \right\}$$

are generated by the linear operator $A = \frac{\partial^2}{\partial x^2}$ satisfies the above properties.

With the operators $T_\alpha(\cdot)$ and $T_{\alpha,\beta}$,the mild and classical solutions of Caputo fractional evolution equation (4.1) is defined as follows

Definition 4.3. *The function $u \in \mathcal{U}$ is called mild solution of Caputo fractional order $(0 < \alpha \leq 1)$ evolution equations* (4.1) *over the interval $[0, T_0]$ if u satisfies the equation of the form:*

$$u(t) = T_\alpha(t)u_0 + \int_0^t (t-s)^{\alpha-1} T_{\alpha,\alpha}(t-s)F(s,u)ds \tag{4.2}$$

where, $T(t)$ and $T_\alpha(t)$ are generated by the linear operator A.

Definition 4.4. *The solution $u \in \mathcal{U}$ is classical solution of semi-linear fractional order evolution equation* (4.1) *of α order Caputo fractional derivative with respect to t exists and continuous.*

Theorem 4.5. *The fractional order Caputo fractional evolution equation* (4.1) *has a unique mild solution over the interval $[0, T_0]$ if following properties are satisfied.*

(1) The families of operators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ generated by the operator $A(t)$ are continuous and bounded over $[0, T_0]$. That is, there exist positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$ for all $t \in [0, T_0]$.

(2) The nonlinear function F is continuous with respect to t and there exist r_0 such that F Lipchitz continuous with respect to u in $B_{r_0} = \{u \in \mathcal{U}; ||u|| \leq r_0\}$. That is, there exist positive constant L such that $||F(t,u) - F(t,v)|| \leq L||u - v||$ for all $t \in [0, T_0]$ and $u \in B_{r_0}$.

Proof. Define the operator $\mathcal{F} : \mathcal{U} \to \mathcal{U}$ as:

$$\mathcal{F}u(t) = T_\alpha(t)u_0 + \int_0^t (t-s)^{\alpha-1} T_{\alpha,\alpha}(t-s)F(s,u)ds.$$

To show (4.1) has unique mild solution it is sufficient to show $\mathcal{F}^{(m)}$ is contraction for some $m \in \mathbb{N}$.
For any $u, v \in B_{r_0}$ and $n \in \mathbb{N}$, we have

$$\begin{aligned} &||\mathcal{F}^{(n)}u(t) - \mathcal{F}^{(n)}v(t)|| \\ &\leq M_\alpha L \int_0^t (t-s)^{\alpha-1} ||\mathcal{F}^{(n-1)}u(s) - \mathcal{F}^{(n-1)}v(s)||ds \\ &\leq M_\alpha^2 L^2 \int_0^t \int_0^{s_1} (t-s_1)^{\alpha-1}(s_1-s)^{\alpha-1} ||\mathcal{F}^{(n-2)}u(s) - \mathcal{F}^{(n-2)}v(s)||dsds_1 \end{aligned}$$

Continuing this process to get

$$\begin{aligned} &||\mathcal{F}^{(n)}u(t) - \mathcal{F}^{(n)}v(t)|| \\ &\leq M_\alpha^n L^\alpha \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} (t-s_1)^{\alpha-1}(s_1-s)^{\alpha-1} \cdots (s_{n-1}-s)^{\alpha-1} ||u(s) - v(s)||dsds_{n-1}\cdots ds_1 \\ &\leq M_\alpha^n L^n \int_0^{T_0} \int_0^{T_0} \cdots \int_0^{T_0} (T_0-s_1)^{\alpha-1}(T_0-s_2)^{\alpha-1} \cdots (T_0-s)^{alpha-1} ||u-v||dsds_{n-1}\cdots ds_1 \\ &\leq M_\alpha^n L^n \int_0^{T_0} (T_0-s)^{n(\alpha-1)} \frac{(T_0-s)^n}{(n-1)!} ||u-v||ds \\ &\leq \frac{(M_\alpha L)^n T_0^{n\alpha}}{n!\alpha} ||u-v|| \end{aligned}$$

Therefore, for any fixed T_0 and sufficiently large integer n say m the operator $\mathcal{F}^{(m)}$ is contraction therefore by generalized Banach fixed point theorem $\mathcal{F}$ has unique fixed point. Hence, (4.1) has unique mild solution given by (4.2). ∎

Example 4.6. *The operators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ generated for the equation* (3.4) *are continuous and bounded. Hence, there exist positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$. Therefore, the equation* (3.3) *has unique mild solution given by* (3.4) *since F is continuous with respect to t and Lipchitz continuous with respect to u in a given Banach space over the interval* $[0, T_0]$.

Remark 4.7. *We have the following observations from the theorem-4.5.*

(1) Conditions derived in the Theorem-4.5 are more liberal than previously derived conditions by the author for the similar system.

(2) The conditions obtained in Theorem-4.5 are sufficient but not necessary.

Now we consider a system in which the initial time is taken $t = t_0$ instead of $t = 0$. Thus the theorem-4.5 can be extended as follows:

Corollary 4.8. *The fractional evolution equation*

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(t_0) &= u_0 \end{aligned} \tag{4.3}$$

has unique mild solution over interval $[t_0, T_0]$ given by

$$u(t) = T_\alpha(t - t_0)u_0 + \int_{t_0}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds \tag{4.4}$$

if following conditions are satisfied:

(1) The families of operators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ generated by the operator A are continuous and bounded over $[t_0, T_0]$. That is, there exist positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$ for all $t \in [t_0, T_0]$

(2) The nonlinear function F is continuous with respect to t and Lipchitz continuous with respect to u. That is, there exist positive constant L such that $||F(t, u) - F(t, v)|| \leq L||u - v||$ for all $t \in [t_0, T_0]$ for $u, v \in B_{r_0}$.

Condition for the classical solution of the system (4.1) is given by the following theorem:

Theorem 4.9. *The mild solution of* (4.1) *is the classical solution if*

(1) $u_0 \in \mathcal{D}(A)$ (Domain of A)

(2) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ are continuously differentiable for all $t > 0$.

(3) The function F is differentiable with respect to t and continuous with respect to u.

Proof. Let $u(t)$ be the mild solution of (4.1). Therefore $u(t)$ satisfies the corresponding integral equation (4.2). Assuming conditions (1),(2) and (3) of the hypothesis, the fractional Caputo derivative of $u(t)$ in equation (4.2) exists and is continuous. Moreover for all $t \in [0, T_0]$ the function $u(t) \in \mathcal{D}(A)$. Hence the mild solution $u(t)$ defined by (4.2) is classical solution of the equation (4.1). This completes the proof of the theorem. ■

Similarly one have the classical solution for the system (4.4).

Corollary 4.10. *The mild solution given by* (4.4) *of* (4.3) *is the classical solution if*

(1) $u_0 \in \mathcal{D}(A)$ (Domain of A)

(2) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ are continuously differentiable for all $t \in [t_0, T_0]$

(3) The function F is differentiable with respect to t and continuous with respect to u.

The following theorems gives the uniqueness of the classical solution of both the systems.

Theorem 4.11. *Equation* (4.1) *has unique classical solution over the interval* $[0, T_0]$ *if*

(1) $u_0 \in \mathcal{D}(A)$ (Domain of A).

(2) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ of the linear operator A are continuously differentiable and bounded over the interval $[0, T_0]$.

(3) The function F is differentiable with respect to t and Lipchitz continuous with respect to u in B_{r_0}.

Proof. Using the condition (2) the generators are continuously differentiable and bounded over $[0, T_0]$ so, they are continuous and bounded over $[0, T_0]$. This means there exist positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$ and condition (3) the function F is continuous with respect t and Lipchitz continuous with respect to u and applying theorem-4.5 the equation (4.1) and has unique mild solution given by (4.2). Assuming (1), (2) and (3) this mild solution becomes classical solution of the equation (4.1). Since mild solution is unique, the classical solution is also unique. ■

Corollary 4.12. *Equation* (4.3) *has unique classical solution over the interval* $[t_0, T_0]$ *if*

(1) $u_0 \in \mathcal{D}(A)$ (Domain of A).

(2) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ of the linear operator A are continuously differentiable and bounded over the interval $[t_0, T_0]$.

(3) The function F is differentiable with respect to t and Lipchitz continuous with respect to u in B_{r_0}.

Example 4.13. *Consider the fractional order equation*

$$^{c}D^{\alpha}w(t,x) + w\frac{\partial w}{\partial x}(t,x) + \frac{\partial^2 w}{\partial x^2}(t,x) = f(t, w(t,x)) \tag{4.5}$$

on the domain $[0, T_0]$ boundary conditions

$$w(t,0) = w(t,2\pi) = 0 \tag{4.6}$$

with initial condition $w(0,x) = w_0$. The domain of the operator $Aw = -\frac{\partial^2 w}{\partial x^2}$ is $\mathcal{D}(A) = \{z \in L^2[0,2\pi] : z'' \; continuous \; and \; satisfies \; boundary \; conditions \; \}$. Then the mild solution in the interval $[0, T_0]$ of the equation (4.5) *with conditions* (4.6) *is given by*

$$w(t,x) = T_\alpha(t)w_0 + \int_0^t (t-s)^{\alpha-1} T_{\alpha,\alpha}(t-s)\{\frac{1}{2}\frac{\partial w^2}{\partial x} + f(s,w)\}ds \tag{4.7}$$

where,

$$T_\alpha(t)z = \sum_{n=1}^{\infty} E_\alpha(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

and

$$T_{\alpha,\beta}(t)z = \sum_{n=1}^{\infty} E_{\alpha,\beta}(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

are the generators of the linear operator A. $\phi_n(x)$ are orthogonal Fourier basis functions in $L^2[0, 2\pi]$.
We have following observation:

(1) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ are defined in equation (4.7) *are continuously differentiable with respect to t. Therefore there exists positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$ respectively.*

(2) The first non linear term in (4.5) $\frac{1}{2}\frac{\partial w^2}{\partial x}$ *is composition of two continuous operators $Pw = \frac{1}{2}\frac{\partial w}{\partial x}$ and $Qw = w^2$ which are continuous with respect to t and Lipchitz continuous with respect to w in finite closed ball B_{r_0} as the operator P is linear and the partial derivative of Q with respect to w exist for every w. Moreover P and Q are differentiable with respect to arguments t and w.*

Therefore equation (4.5) *has unique mild solution given by* (4.7) *if the second term $f(t, w)$ is continous with respect to t and Lipchitz continous with respect to w The mild solution* (4.7) *is unique classical solution of* (4.5) *if $f(t, w)$ is differentiable and $w_0 \in D(A)$.*

5. Mild and classical solutions with impulses

In this section we are going to derive set of sufficient conditions for the existence and uniqueness of classical as well mild solution of impulsive fractional evolution equation (1.1). We are also deriving the conditions in which the classical and mild solutions are coincides.

Definition 5.1. ***Classical Solution****[20]*
A solution $u(t)$ is a classical solution of the equation (1.1) *for $0 < \alpha < 1$ if $u(t) \in PC([0, T_0], \mathcal{U}) \cap C^\alpha(J', \mathcal{U})$ where, $J' = [0, T_0] - \{t_1, t_2, \cdots, t_p\}$ and $C^\alpha(J', \mathcal{U}) = \big\{u : J' \to \mathcal{U} :^c D^\alpha u(t)$ exist and continuous at each $t \in J'\big\}$, $u(t) \in D(A)$ (Domain of A) for $t \in J'$ and satisfies* (1.1) *on* $[0, T_0]$.

Definition 5.2. ***Mild Solution***
A function $u(t) \in PC([0, T_0], \mathcal{U})$ is a mild solution of the equation (1.1) *if it satisfies*

$$u(t) = \begin{cases} T_\alpha(t-t_i)\big(\prod_{k=i}^{1} T_\alpha(t_k - t_{k-1})\big)u_0 + T_\alpha(t-t_i)\sum_{j=1}^{i}\big(\prod_{k=j}^{2} T_\alpha(t_k - t_{k-1}\big) \\ \int_{t_{j-1}}^{t_j}(t_j - s)^{\alpha-1}T_{\alpha,\alpha}(t_j - s)F(s, u(s))ds + \int_{t_i}^{t}(t-s)^{\alpha-1}T_{\alpha,\alpha}(t-s)F(s, u(s))ds \\ +T_\alpha(t-t_i)\sum_{j=1}^{i}\big(\prod_{k=i}^{3} T_\alpha(t_k - t_{k-1})\big)I_k u(t_k) \end{cases} \tag{5.1}$$

for each $t \in [t_i, t_{i+1})$.

Here, the families of operators $T(t)$ and $T_\alpha(t)$ are generated by the linear operator A.

Theorem 5.3. *The fractional order semi-linear impulsive evolution equation* (1.1) *has unique mild solution over the interval $[0, T_0]$ if following properties are satisfied.*

(1) The families of operators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ generated by the operator A are continuous and bounded over $[0, T_0]$. That is there exist positive constants M and M_α such that $||T_\alpha(t)|| \leq M$ and $||T_{\alpha,\beta}(t)|| \leq M_\alpha$ for all $t \in [0, T_0]$.

(2) The nonlinear function F is continuous with respect to t and Lipchitz continuous with respect to u in B_{r_0}. That is there exist positive constant L such that $||F(t, u) - F(t, v)|| \leq L||u - v||$ for all $t \in [0, T_0]$ and $u, v \in B_{r_0}$.

(3) Impulses I_k at $t = t_k$ for $k = 1, 2, \cdots, k$ are continuous and bounded.

Proof. Over the interval $[0, t_1]$ the equation (1.1) becomes,

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(t_0) &= u_0 \end{aligned} \tag{5.2}$$

Assuming conditions (1) and (2) of the hypotheses and using theorem-4.5 the equation (5.2) has unique mild solution over the interval $[0, t_1)$ given by

$$u(t) = T_\alpha(t - t_0)u_0 + \int_{t_0}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds. \tag{5.3}$$

At $t = t_1$ the mild solution $u(t_1^-)$ becomes:

$$u(t_1^-) = T_\alpha(t_1 - t_0)u_0 + \int_{t_0}^{t_1} (t_1 - s)^{\alpha-1} T_{\alpha,\alpha}(t_1 - s)F(s, u)ds.$$

Over the interval $[t_1, t_2)$ the equation (1.1) becomes:

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(t_1^+) &= u_1 = u(t_1^-) + I_1 u(t_1) \end{aligned} \tag{5.4}$$

Here, impulse operator I_1 is continuous and bounded. Assuming conditions (1) and (2) and applying corollary 4.8, the equation (5.4) has unique mild solution over the interval $[t_1, t_2)$ given by

$$u(t) = T_\alpha(t - t_1)u_1 + \int_{t_1}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds. \tag{5.5}$$

Continuing in this way the equation (1.1) over the interval $[t_i, t_{i+1})$ becomes

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(t_i^+) &= u_i = u(t_i^-) + I_i u(t_i). \end{aligned} \tag{5.6}$$

Assuming condition (1) and (2) of the hypotheses and applying corollary-4.8 the equation (5.6) has unique mild solution over the interval $[t_i, t_{i+1})$ given by

$$u(t) = T_\alpha(t - t_i)u_i + \int_{t_i}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds. \tag{5.7}$$

Finally over the interval $[t_p, T_0]$ the equation (1.1) becomes:

$$\begin{aligned} {}^cD^\alpha u(t) &= Au(t) + F(t, u(t)), \\ u(t_p^+) &= u_1 = u(t_p^-) + I_p u(t_p). \end{aligned} \tag{5.8}$$

Assuming condition (1) and (2) of the hypotheses and applying corollary-4.8 the equation (5.8) has unique mild solution over the interval $[t_p, T_0]$ given by

$$u(t) = T_\alpha(t - t_p)u_p + \int_{t_p}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds. \tag{5.9}$$

Therefore for any $t \in [t_i, t_{i+1})$ for $i = 1, 2, \cdots, p$ the equation (1.1) has unique mild solution given by

$$\begin{aligned} u(t) &= T_\alpha(t - t_i)u_i + \int_{t_i}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds \\ &= T_\alpha(t - t_i)[u(t_i^-) + I_i u(t_i)] + \int_{t_i}^{t} (t - s)^{\alpha-1} T_{\alpha,\alpha}(t - s)F(s, u)ds \end{aligned}$$

Substituting the values of u_k's for $k = 1, 2, \cdots, i$ we obtained,

$$u(t) = T_\alpha(t - t_i)\big(\prod_{k=i}^{1} T_\alpha(t_k - t_{k-1})\big)u_0 + T_\alpha(t - t_i)\sum_{j=1}^{i}\big(\prod_{k=j}^{2} T_\alpha(t_k - t_{k-1}\big)$$
$$\int_{t_{j-1}}^{t_j}(t_j - s)^{\alpha-1}T_{\alpha,\alpha}(t_j - s)F(s, u(s))ds$$
$$+ \int_{t_i}^{t}(t - s)^{\alpha-1}T_{\alpha,\alpha}(t - s)F(s, u(s))ds + T_\alpha(t - t_i)\sum_{j=1}^{i}\big(\prod_{k=i}^{3} T_\alpha(t_k - t_{k-1})\big)I_k u(t_k).$$

We complete the proof by showing $u(t) \in PC([0, T_0], \mathcal{U})$ for all $t \in [0, T_0]$.
If $t \in [0, T_0]$ for all $j = 1, 2, \cdots, p$ then $t \in [t_i, t_{i+1})$ for atleast one i. Assuming conditions (1), (2) and (3) we get the continuity of u at $t \neq t_i$ and left continuous at $t = t_i$ and right limit exist at $t = t_i$. Therefore $u(t) \in PC([0, T_0], \mathcal{U})$. Hence, equation (1.1) has unique mild solution in $PC([0, T_0], \mathcal{U})$. ■

Theorem 5.4. *The mild solution* (5.1) *of* (1.1) *is the classical solution if*
(1) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ are continuously differentiable for all $t > 0$.
(2) The function F is diffrentiable with respect to t and continuous with respect to u.
(3) Impulses I_k at $t = t_k$ are for $k = 1, 2, \cdots, k$ differentiable and bounded.
(4) u_0 and $I_k u(t_k)$ are in $\mathcal{D}(A)$ (Domain of A).

Proof. Over the interval $[0, t_1)$ the equation (1.1) becomes (5.2) which is evolution equation without impulses. Applying theorem-4.9 the mild solution (5.3) becomes classical solution of (1.1) over the interval $[0, t_1)$ by assuming the conditions (1), (2) and (4).
In the interval $[t_1, t_2)$ the equation (1.1) becomes (5.3) and I_1 is differentiable and bounded with $I_1 u(t_1) \in \mathcal{D}(A)$ therefore, $u_1 \in \mathcal{D}(A)$. Again assuming the conditions (1), (2) and (4) and using corollary- 4.10 the mild solution (5.5) becomes a classical solution of (1.1) over the interval $[t_1, t_2)$.
Continuing in same manner the mild solution (5.7) of equation (1.1) over the interval $[t_i, t_{i+1})$ becomes classical solution of (1.1).
Finally, the mild solution (5.9) of the equation (1.1) becomes classical solution of equation (1.1) over the interval $[t_p, T_0]$ proceeding in similar manner.
Hence the mild solution (5.1) of equation (1.1) becomes classical solution of (1.1) over the whole interval $[0, T_0]$. This completes the proof. ■

Now we discuss the uniqueness of classical solution of impulsive evolution equation (1.1).

Theorem 5.5. *Classical solution of* (1.1) *is unique if*

(1) The generators $T_\alpha(t)$ and $T_{\alpha,\beta}(t)$ are continuously differentiable for all $t > 0$.

(2) The function F is differentiable with respect to t and Lipschitz continuous with respect to u on B_{r_0}.

(3) Impulses I_k at $t = t_k$ are for $k = 1, 2, \cdots, k$ differentiable and bounded.

(4) u_0 and $I_k u(t_k)$ are in $\mathcal{D}(A)$ (Domain of A).

Proof. Under the assumption (1), (2), (3) and (4) the mild solution (5.1) of equation (1.1) becomes a classical solution. Lipchitz continuity of F with respect to u leads to uniqueness of mild solution. Since mild solution of (1.1) is unique therefore classical solution of (1.1) is unique. ■

Example 5.6. *Consider the semi-linear fractional order impulsive heat equation*

$$\begin{aligned} {}^cD_t^\alpha u(t,x) &= \frac{\partial^2 u(t,x)}{\partial x^2} + u\frac{\partial u}{\partial x}(t,x), \quad t \neq t_1, t_2 \cdots, t_p \\ u(t,0) &= u(t,\pi) = 0 \\ u(0,t) &= u_0 = x(\pi - x) \\ \Delta u(t_k) &= I_k(t_k) = a_k u(t_k^-), \quad t = t_k, \quad (a_k\text{'s are constants}) \quad k = 1, 2, \cdots, p \end{aligned} \tag{5.10}$$

over the interval $[0, T_0]$. *Here* t_k*'s are time points where impulses are applied.*
We have following observations:

(1) The operator $A = \frac{\partial^2}{\partial x^2}$ *over the domain* $\mathcal{D}(A) = \{z : [0,\pi] \to \mathbb{R} : z'' \text{ exists and } z(0) = z(\pi) = 0\}$ *generates the continuously differentiable and bounded families of operators* $T_\alpha(t)$ *and* $T_{\alpha,\beta}(t)$ *defined by*

$$T(t)z = \sum_{n=1}^{\infty} E_\alpha(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

and

$$T_{\alpha,\beta}(t)z = \sum_{n=1}^{\infty} E_{\alpha,\beta}(-n^2 t^\alpha) < z, \phi_n > \phi_n$$

respectively.

(2) The nonlinear function $F(t,u) = u\frac{\partial u}{\partial x} = \frac{\partial}{\partial x}u^2$ *is differentiable with respect to* t *and Lipchitz continuous with respect to* u *on* B_{r_0}.

(3) Impulses $I_k u(t_k) = a_k u(t_k^-)$ *are differentiable such that* $I_k u(t_k) \in \mathcal{D}(A)$.

(4) $u_0 \in \mathcal{D}(A)$.

Therefore by Theorem 5.3, 5.4 and 5.5 the equation (5.10) *has unique mild solution given by*

$$\begin{aligned} u(t) &= T_\alpha(t - t_i)\Big(\prod_{k=i}^{1} T_\alpha(t_k - t_{k-1})\Big)u_0 \\ &+ T_\alpha(t - t_i)\sum_{j=1}^{i}\Big(\prod_{k=j}^{2} T_\alpha(t_k - t_{k-1})\Big)\int_{t_{j-1}}^{t_j}(t_j - s)^{\alpha-1}T_{\alpha,\alpha}(t_j - s)\frac{\partial}{\partial x}u^2 ds \\ &+ \int_{t_i}^{t}(t - s)^{\alpha-1}T_{\alpha,\alpha}(t - s)\frac{\partial}{\partial x}u^2 ds + T_\alpha(t - t_i)\sum_{j=1}^{i}\Big(\prod_{k=i}^{3} T_\alpha(t_k - t_{k-1})\Big)a_k u(t_k) \end{aligned} \tag{5.11}$$

for all $t \in [0, T_0]$. *Moreover this mild solution* (5.11) *becomes classical solution of* (5.10). *Since mild solution is unique therefore classical solution is unique.*

6. Conclusion

The fractional semi-linear evolution equation over general Banach space without and with impulses has a set of mild and classical solutions, which are deduced in this article. We developed the novel notion of generators and derived the adequate requirements—which are more lax criteria and apply to a broader class of fractional evolution equations using the generalised Banach fixed point theorem.

7. Acknowledgement

The author is thankful to the referee for his valuable suggestions which improved the presentation of the paper.

References

[1] M. U. AKHMET, On the general problem of stability for impulsive differential equations, *J. Math. Anal. Appl*, **288** (2003), 182-196.

[2] A. ANGURAJ AND M. M. ARJUNAN, Existance and uniqueness of mild and classical solutions of impusive evolution equations, *Electronic Journal of Differential Equations*, **111** (2005), 1-8.

[3] D. D. BAINOV AND A. B. DISHLIEV, Population dynamics control in regard to minimizing the time necessary for the regeneration of a biomass taken away from the population,*Appl.Math. Comput.*, **39**(1990), 37-48.

[4] K. BALACHANDRAN,K. S. KIRUTHIKA, J. J. TRUJILLO, Existence results for fractional impulsive integro-differential equations in Banach spaces, *Commun Nonlinear Sci Numer Simulat*. **16**(2011), 1970-1977.

[5] M. BENCHOHRA, AND B. A. SLIMANI, Existence and uniqueness of solutions to impulsive fractional differential equations, *Eel. J. Of Diff. Eq.*, **2009**(2009), 1-11.

[6] B. BONILLA, M. RIVERO, L. RODRIGUEZ-GERMA, AND J. J. TRUJILLO, Fractional differential equations as alternative models to nonlinear differential equations, *Applied Mathematics and Computation*, **187**(2007), 79–88.

[7] J. BORAH AND N. BORA, Existence of Mild Solution of A Class of Nonlocal Fracitonal Order Differential Equations with Non Instaneous Impulses, *Fractional Calculus and Applied Analysis*, **22**(2019), 495-508.

[8] Y. CHENG, AND G. GUOZHU, On the solution of nonlinear fractional order differential equations, *Nonlinear Analysis: Theory, Methods and Applications*, **310**(2005), 26-29.

[9] K. DISHLIEVA, Impulsive Differential Equations and Applications, *J Applied Computat Math*, **2016**(2016), 1-6.

[10] D. DELBOSCO, AND L. RODINO, Existence and uniqueness for nonlinear fractional differential equations, *Journal of Mathematical Analysis and Applications*, **204**(1996), 609-625.

[11] M. M. EL-BORAI, Semigroups and some nonlinear fractional differential equations, *Applied Mathematics and Computations*, **149**(2004), 823-831.

[12] A. M. A. EL-SAYED, Fractional order wave equation, *International Journal of Theoretical Physics*,**35**(1996), 311-322.

[13] L. H. ERBE, H. I. FREEDMAN, X. Z. LIU, AND J. H. WU, Comparison principles for impulsive parabolic equations with applications to models of single species growth, *J. Austral. Math. Soc.*,*32*(1991), 382-400.

[14] V. GAFIYCHUK, B. DATSAN, AND V. MELESHKO, Mathematical modeling of time fractional reaction-diffusion system, *Journal of Computational and Applied Mathematics*, **220** (2008), 215-225.

[15] P. GONZALEZ, AND M. PINTO, Asymptotic behavior of impulsive differential equations, *Rokey Mountain Journal of Mathematics*, **26**(1996), 165-173.

[16] J. H. HE, Approximate analytical solution for seepage flow with fractional derivatives in porous media, *Computer Methods in Applied Mechanics and Engineering*, **167**(1998), 57-68.

[17] J. H. HE, Some applications of nonlinear fractional differential equations and their approximations, *Bulletin of Science and Technology*, **15**(1999), 86-90.

[18] O. S. IYIOLA, AND G. OLAY, The fractional Rosenau-Hyman model and its approximate solution, *Alexandria Engineering Journal*, **55**(2016), 1655-1659.

[19] E. JOELIANTO, Linear Impulsive differential equations for hybird system modelling, *2003 Europian Control Conference (EEC) Cambridge, UK, 1-4 Sept*, (2003), 3335-3340. .

[20] H. R. KATARIA, AND P. H. PATEL, Congruency between classical and mild solutions of Caputo fractional impulsive evolution equation on Banach Space,*International Journal of Advance Science and Technology*, **29**(2020), 1923-1936.

[21] H. R. KATARIA, P. H. PATEL, AND V. SHAH,Existence results of noninstantaneous impulsive fractional integro-differential equation, *Demonstratio Mathematica*, **53**(2022), 373-384.

[22] A. A. KILBAS, H. M. SRIVASTAVA, AND J. J. TRUJILLO, Theory and Applications of Fractional Differential Equations, *Elsevier Science*, (2006).

[23] M. KIRANE, AND Y. V. ROGOVCHENKO, Comparison results for systems of impulse parabolic equations with applications to population dynamics,*Nonlinear Analysis*, **28**(1997), 263-277.

[24] V. LAKSHMIKANTHAM, D. D. BAINOV, AND P. S. SIMEONOV, Theory of Impulsive Differential Equations, *World Scientific*,**Singapore**, 1989.

[25] Z. G. LI, Y. WEN, AND Y. C. SOH, Analysis and Design of Impulsive Control, *IEEE Transaction on Automatic Control*, **46**(2001), 894-897.

[26] J. H. LIU, Nonlinear impulsive evolution equations, *Dynamics Contin. Diser.Impulsive System*, **6**(1999), 77-85.

[27] R. METZLER, AND J. KLAFTER, The restaurant at the end of random walk, the recent developments in description of anomalous transport by fractional dynamics, *Journal of Physics A: A Mathematical and General*, **37**(2004), 161-208.

[28] V. MILMAN, AND A. MYSHKIS, On the stability of motion in the presence of impulses, *Siberian Math. J.*, **1**(1996), 233-237.

[29] G. MOPHOU, Existence and uniqueness of mild solutions to impulsive fractional differential equations, *Nonlinear Analysis*, **72**(2010),1604–1615.

[30] I. PODLUBNY, *Fractional Differential Equation*, Academic Press, San Diego, (1999).

[31] J. C. PRAJAPATI, K. B. KACHHIA, AND S. P. KOSTA, Fractional calculus approach to study temperature distribution within a spinning satellite, *Alexandria Engineering Journal*, **55**(2016), 2345-2350.

[32] C. RAVICHANDRAN, AND M. ARJUNAN, Existence and Uniqueness results for impulsive fractional integro-differential equations in Banach spaces, *International Journal of Nonlinear Science*, **11**(2011), 427-439.

[33] M.RENARDY, W. J. HRUSA, AND J. A. NOHEL, Mathematical problems in viscoelasticity, *Longman Scientific and technical, Newyork*, (1987).

[34] Y. V. ROGOVCHENKO, Nonlinear Impulse Evolution Systems and Applications to Population Models, *J. Math. Anal Appl.*, **207**(1997), 300-315.

[35] Y. ROGOVCHENKO, Impusive evolution systems: main results and new trends, *Dynamics Contin. Diser. Impulsive Sys.*, **3**(1997), 57-88.

[36] S. G. SAMKO, A. A. KILBAS, AND O. I. MARICHEV, Fractional Integrals and Derivatives; Theory and Applications,*Gorden and Breach Publications*, (1993).

[37] V. SHAH, R. K. GEORGE, J. SHARMA, AND P. MUTHUKUMAR, Existence and Uniqueness of Classical and Mild Solutions of Generalized Impulsive Evoution Equations, *International Journal of Nonlinear Sciences and Numerical Solutions*, **19**(2018), 775-780.

[38] M. TAMSIR, AND V. K. SRIVASTAVA, Revisiting the approximate analytical solution of fractional order gas dynamics equation, *Alexandria Engineering Journal*, **55** (2016), 867-874.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 80–93.
http://doi.org/10.26637/mjm1101/006

On a generalized fractional differential Cauchy problem

Mesfin Etefa[1], **Gaston M. N'Guérékata**[*2], **Pierre Ngnepieba**[3] **and Olaniyi S. Iyiola**[4]

[1] *Department of Mathematics, NEERLab Laboratory, Bowie State University, 14000 Jericho Park Rd., Bowie, MD 20715, USA.*
[2] *University Distinguished Professor, Department of Mathematics, NEERLab Laboratory, Morgan State University, 1700 E. Cold Spring Lane, Baltimore, M.D. 21251, USA.*
[3] *College of Science and Technology, Florida A& M University, Tallahassee, FL USA.*
[4] *Department of Mathematics, Clarkson University, Potsdam, NY, USA.*

***R*eceived 01 September 2022; *A*ccepted 30 December 2022**

Abstract. Qualitative results for abstract problems are very important in understanding mathematical analysis on which any application is possible. The focus of this paper is twofold: first, we investigate the existence and uniqueness of mild solutions to a generalized Cauchy problem for the nonlinear differential equation with non-local conditions in a Banach space X. This is achievable using some fixed point theorems in infinite dimensional spaces. Secondly, we study the stability results of the system in the sense of Ulam-Hyers-Rassias. Our results improve and generalize most recent related results in the literature.

AMS Subject Classifications: 26A33, 34A12, 34G20.

Keywords: κ-Hilfer operator; Cauchy problem; mild solution; existence theory; Krasnoselski theorem; Ulam-Hyers-Rassias stability.

Contents

1. Introduction

During the last three decades, Fractional Calculus has become a significant research topic in mathematics on account of its wide range of applications in solving real world problems. These applications are found in different fields of studies including science, finance, mathematical biology, engineering and social sciences. Nonlocality nature of fractional order derivatives has become great tool for modeling complex phenomena for which the structures have inherent non-local properties. Examples of the applications are seen specifically in anomalous transport, and anomalous diffusion, biological modeling including pattern formation and cancer

***Corresponding author.** Email address: Gaston.NGuerekata@morgan.edu

treatment, financial modeling including chaos and long memory, dynamical system control theory, random walk, viscoelasticity, and nanotechnology, see [5, 6, 8–10, 19, 21, 22]. Despite many advantages of using fractional order derivatives in modeling, rigorous theoretical results on existence and uniqueness of solutions including stability are needed to drive any meaningful analysis on the subject. This is our profound concern here.

In 2018, da C. Sousa and Olivera [11] presented for the first time, the concept of κ-Hilfer fractional derivative. This has influenced several papers recently. Our work here is based on the so-called κ-Hilfer fractional operator. The new definition is a generalization of several well-known and well-studied fractional derivatives in the literature. However, several theoretical results are lacking. Our focus in this paper is to provide some important and useful results needed to advance research on generalized fractional calculus. The main motivation of this work is based on the work done by N'Guérékata in 2009 [12], in which he proved the existence and uniqueness of mild solutions to the fractional differential equation involving nonlocal conditions:

$$D^{\alpha}u(t) = f(t, u(t)), \quad t \in I := [0, T]$$

$$u(0) + g(u) = u_0.$$

Here D^{α} represents the Caputo fractional derivative of order $0 < \alpha < 1$, $f : I \times X \to E$ is a given function with some sufficient conditions to be specified below. Furthermore, X is a (complex) Banach space with norm $\|\cdot\|$, and $I := [0, T], T > 0$. The nonlocal condition is defined as

$$g(u) = \sum_{i=1}^{n} c_i u(t_i)$$

where $c_i,\ i = 1, 2, ..., n$ are some given constants and $0 < t_1 < t_2 < ... < t_n \leq T$ a partition of I. Let's recall that the concept of "nonlocal conditions" were first introduced by K. Deng in his pioneer paper [7]. In his work, he demonstrated that using the nonlocal condition $u(0) + g(u) = u_0$ to describe for example the diffusion phenomenon of a small amount of gas in a transparent tube can provide a more efficient insight than if one consider the classical Cauchy problem $u(0) = u_0$. We observe also that since Deng's paper, such problems have attracted numerous researchers: see [4, 6, 17] and many references therein.

Recently, F. Norouzi and G. M. N'Guérékata [24] studied some existence and uniqueness results of mild solutions to the κ- Hilfer semilinear neutral fractional differential equations in a general Banach space X involving an infinite delay

$$\begin{cases} [u(t) - h(t, u_t)] = B\, u(t) + g(t, u(t), u_t),\ t \in [0, T], \quad T > 0 \\ \qquad u(t) = \psi(t), \qquad\qquad t \in (-\infty, 0] \end{cases} \tag{1.1}$$

using some classical fixed point theorems. The history function $\psi(t)$ taking values in an abstract space, and the linear operator B generates a semigroup of linear operators $(S(t)_{t\geq 0}$ which are uniformly bounded on X.

The present work is motivated by the papers mentioned above. Basically we will focus on the qualitative results (existence and stability) of mild solutions to the nonlinear abstract Cauchy problem involving nonlocal condition

$$\begin{cases} {}^{\mathbb{H}}D_{0^+}^{\alpha,\beta;\kappa} u(t) = f(t, u(t)),\ t \in I \\ I_{0^+}^{1-\sigma;\kappa} u(0) = u_0 - g(u), \end{cases} \tag{1.2}$$

Here ${}^{\mathbb{H}}D^{\alpha,\beta;\kappa}$ represents the κ-Hilfer operator, $I^{1-\sigma;\kappa}$ is the left sided $\kappa-$Riemann-Liouville fractional integral operator, α and β are the order and the type of the derivative respectively, and $\sigma = \alpha + \beta\,(1-\alpha)$. The results are new even in the case of a finite dimensional space as mentioned below.

2. Preliminaries

Useful results involving fractional operators and their definitions are presented in this Section. We denote by $\mathcal{C} := C([a,b],E)$ the usual Banach space of all continuous functions from $[a,b]$ to E a (complex) Banach space equipped with the topology of the uniform convergence induced by the norm $\|x\|_{\mathcal{C}}. = \sup_{t\in[a,b]} \|x(t)\|$ Furthermore, we assume that $\kappa \in C^1(I,\mathbb{R})$ is a monotonically increasing function and $\kappa'(x) > 0$.

Definition 2.1. *[11] Let (a,b) be an interval such that $-\infty \le a < b \le \infty$. We assume that $\kappa(x)$ is a positive and monotonically increasing function defined on $(a,b]$ where $\kappa'(x)$ is continuous on (a,b). Then the fractional integral (left-sided) of a function w defined on on $[a,b]$ with respect to a function κ is given as*

$$I_{a^+}^{\alpha;\kappa} w(t) = \frac{1}{\Gamma(\alpha)} \int_{a^+}^{t} \kappa'(s)(\kappa(t) - \kappa(s))^{\alpha-1} w(s)ds.$$

Definition 2.2. *[11] Let $\alpha > 0$, $n \in \mathbb{N}$, (a,b) be an interval such that $-\infty \le a < b \le \infty$ and let $\kappa'(x) \ne 0$. The Riemann-Liouville (left-sided) fractional derivative of a function w of order α with respect to a function κ is defined by*

$$\begin{aligned} D_{a^+}^{\alpha;\kappa} w(t) &= \left(\frac{1}{\kappa'(t)}\frac{d}{dt}\right) I^{n-\alpha;\kappa} w(t) \\ &= \frac{1}{\Gamma(n-\alpha)} \left(\frac{1}{\kappa'(t)}\frac{d}{dt}\right)^n \int_a^t \kappa'(s)(\kappa(t) - \kappa(s))^{n-\alpha-1} w(s)ds. \end{aligned}$$

Definition 2.3. *[11] Let $n-1 < \alpha < n$ with $n+1,2,...,$ $I = [a,b]$ is the interval such that $-\infty \le a < b \le \infty)$, $w, \kappa \in C^n(I,\mathbb{R})$ two functions such that κ is increasing and $\kappa'(t) \ne 0$, for all $t \in I$. The left sided and right sided $\kappa-$Hilfer fractional derivative $D_{a^+}^{\alpha,\beta;\kappa}(\cdot)$ of function of order α and type $0 \le \beta \le 1$ are defined respectively by*

$${}^{\mathbb{H}}D_{a^+}^{\alpha,\beta;\kappa} w(t) = I_{a^+}^{\beta(n-\alpha);\kappa} \left(\frac{1}{\kappa'(t)}\frac{d}{dt}\right)^n I_{a^+}^{(1-\beta)(n-\alpha);\kappa} w(t),$$

and

$${}^{\mathbb{H}}D_{b^-}^{\alpha,\beta;\kappa} w(t) = I_{b^-}^{\beta(n-\alpha);\kappa} \left(-\frac{1}{\kappa'(t)}\frac{d}{dt}\right)^n I_{b^-}^{(1-\beta)(n-\alpha);\kappa} w(t).$$

Theorem 2.4. *[11] If $0 \le \beta \le 1$, $n-1 < \alpha < 1$ such that $\gamma = \alpha + \beta(n-\alpha)$ and $w \in C^n(I)$, then*

$$I_{a^+}^{\alpha;\kappa}\, {}^{\mathbb{H}}D_{a^+}^{\alpha,\beta;\kappa} w(t) = w(t) - \sum_{k=1}^{n} \frac{(\kappa(x) - \kappa(a))^{\gamma-k}}{\Gamma(\gamma-k+1)} w_\kappa^{[n-k]} I_{a^+}^{(1-\beta)(n-\alpha);\kappa} w(a),$$

and

$$I_{b^-}^{\alpha;\kappa}{}^{\mathbb{H}}D_{b^-}^{\alpha,\beta;\kappa} w(t) = w(t) - \sum_{k=1}^{n} \frac{(-1)^k(\kappa(b) - \kappa(x))^{\gamma-k}}{\Gamma(\gamma-k+1)} w_\kappa^{[n-k]} I_{b^-}^{(1-\beta)(n-\alpha);\kappa} w(a).$$

Theorem 2.5. *[11] Let $w \in C^1[a,b], \alpha > 0$, and $0 \le \beta \le 1$, then*

$${}^{\mathbb{H}}D_{a^+}^{\alpha,\beta;\kappa} I_{a^+}^{\alpha;\kappa} w(t) = w(t) \quad \text{and} \quad {}^{\mathbb{H}}D_{b^-}^{\alpha,\beta;\kappa} I_{b^-}^{\alpha;\kappa} w(t) = w(t).$$

3. Main Results

Let's now discuss our main results. From now on I will be the finite interval $[0,T]$. Firstly, we indicate the following assumptions which are needed in the proofs of results.

(H1): $f : \mathbb{R} \times I \to E$ is a Caratheodory function, meaning for every $u \in E$, $f(t,u)$ is strongly measurable with

respect to first variable and for every $t \in I, f(t,u)$ is continuous with respect to the second variable
(H2): $\|f(t,u) - f(t,v)\| \leq L_f\|u - v\|, \forall t \in \mathbb{R}, \forall u, v \in E, \gamma > 0$.
(H3): $g : \mathcal{C} \to E$ is continuous and

$$\|g(u) - g(v)\| \leq b\|u - v\|, \forall u, v \in \mathcal{C}, b > 0.$$

(H2'): There exists a function $\mu \in L^1(I)$ such that

$$\|f(t,u) - f(t,v)\| \leq \mu(t)\|u - v\|, \forall t \in I, \text{for all} u, v \in E.$$

For $0 < \alpha < 1$ and $0 \leq \beta \leq 1$, the nonlinear system (1.1) is then equivalent to the integral equation

$$u(t) = \frac{(\kappa(t) - \kappa(0))^{\sigma-1}}{\Gamma(\sigma)}(u_0 - g(u)) + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t) - \kappa(s))^{\alpha-1} f(s, u(s))ds, \tag{3.1}$$

for each $t \in I$.

Proof. We do not recall the proof since it is straightforward from [11, 15]. ■

Definition 3.1. *A continuous function $x : I \to E$ is said to be a mild solution of Equations (1.1) if it can be written as*

$$u(t) = \frac{(\kappa(t) - \kappa(0))^{\sigma-1}}{\Gamma(\sigma)}(u_0 - g(u)) + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t) - \kappa(s))^{\alpha-1} f(s, x(s))ds, \quad t \in I.$$

We now present our results.

Theorem 3.2. *Assume that assumptions **(H1-H3)** hold with*

$$b < \frac{1}{2}, \quad \frac{(\kappa(t) - \kappa(0))^{\sigma-1}}{\Gamma(\sigma)} \leq 1, \forall t \in I, \quad \textit{and} \quad \gamma < \frac{\Gamma(\alpha+1)}{2\kappa(T)^\alpha},$$

then there exists a unique mild solution to (1.1).

Proof. Consider the operator $\Omega : \mathcal{C} \to \mathcal{C}$ defined by

$$(\Omega u)(t) := \frac{(\kappa(t) - \kappa(0))^{\sigma-1}}{\Gamma(\sigma)}(u_0 - g(u)) + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t) - \kappa(s))^{\alpha-1} f(s, u(s))ds.$$

It is obvious that this operator Ω is well-defined. From assumptions **(H1-H3)** we can define

$$M := \sup_{t \in I} ||f(t,0)|| \quad \text{and} \quad P := \sup_{x \in \mathcal{C}} ||g(u)||$$

and choose

$$r \geq \left(P + \|u_0\| + \frac{M\kappa(T)^\alpha}{\Gamma(\alpha+1)}\right).$$

For $B_r := \{x \in \mathcal{C} : \, ||u_0|| \leq r\}$, we can show that $FB_r \subset B_r$. Then if

$$\Delta_\kappa^\sigma(t,0) = \frac{(\kappa(t) - \kappa(0))^{\sigma-1}}{\Gamma(\sigma)},$$

we obtain the following estimate for $u \in B_r$

$$\|\Omega u(t)\| \leq \|\Delta_\kappa^\sigma(t,0)u_0\| + \|\Delta_\kappa^\sigma(t,0)g(u)\| + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t) - \kappa(s))^{\alpha-1}\|f(s, u(s))\|ds$$

$$
\begin{aligned}
&\leq \|u_0\| + P + \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}(\|f(s,u(s))-f(s,0)\| + \|f(s,0)\|)ds \\
&\leq \|u_0\| + P + \frac{Lr+M}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}ds \\
&= \|u_0\| + P + \left(\frac{Lr+M}{\Gamma(\alpha)}\right) \frac{(\kappa(t)-\kappa(0))^\alpha}{\alpha} \\
&\leq \|u_0\| + P + \left(\frac{Lr+M}{\Gamma(\alpha+1)}\right) \kappa(T)^\alpha \\
&\leq r,
\end{aligned}
$$

with suitable choices of L_f and r.
Let $u, v \in \mathcal{C}$. Furthermore we have

$$
\begin{aligned}
\|(\Omega u)(t) - (\Omega v)(t)\| &\leq \|g(u)-g(v)\| \\
&\quad + \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u(s))-f(s,v(s)\|ds \\
&\leq b\|u-v\|_{\mathcal{C}} + \frac{1}{\Gamma(\alpha)} L\|u-v\|_{\mathcal{C}} \frac{(\kappa(t)-\kappa(0))^\alpha}{\alpha} \\
&\leq b\|u-v\|_{\mathcal{C}} + \frac{L_f\kappa(T)^\alpha}{\Gamma(\alpha+1)}\|u-v\|_{\mathcal{C}} \\
&= \left(b + \frac{L_f\kappa(T)^\alpha}{\Gamma(\alpha+1)}\right)\|u-v\|_{\mathcal{C}} \\
&\leq \Phi\|u-v\|_{\mathcal{C}},
\end{aligned}
$$

where

$$
\Phi = \Phi_{b,L_f,T,\alpha} := \left(b + \frac{L_f\kappa(T)^\alpha}{\Gamma(\alpha+1)}\right)
$$

Since $0 < \Phi < 1$ and $\|\Omega u - \Omega v\|_{\mathcal{C}} \leq \Phi\|u-v\|_{\mathcal{C}}$, Ω turns out to be a contractive mapping. We deduce that Ω has a unique fixed point, which is the mild solution of the Cauchy problem. The problem is solved. ■

Remark 3.1. *This result generalizes Theorem 2.1 in [12] in the case where $E = \mathbb{R}^n$ and $\kappa(t) = t$.*

Let's now consider the local problem associated to Equation (1.1), that is g is identically zero on J.

Theorem 3.3. *Let's suppose that ((**H1**)-(**H2**)) hold and*

$$
\kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} \leq 1.
$$

Then there exists a unique mild solution to Equation (1.1).

Proof. Consider the operator Ω as in Theorem 3.2. Then we have

$$
\begin{aligned}
\|(\Omega u)(t) - (\Omega v)(t)\| &\leq \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u(s))-f(s,v(s))\|ds \\
&\leq \frac{\|u-v\|_{\mathcal{C}}}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\mu(s)ds.
\end{aligned}
$$

Also we have

$$
\|(\Omega u)^2(t) - (\Omega v)^2(t)\| \leq \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,Fu(s))-f(s,Fv(s))\|ds
$$

$$\leq \frac{\|u-v\|_{\mathcal{C}}}{2\Gamma(\alpha)^2}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\mu(s)\int_0^s \mu(\sigma)d\sigma ds$$
$$\leq \frac{\|u-v\|_{\mathcal{C}}}{2\Gamma(\alpha)^2}\|\mu\|^2_{L^1(I)}.$$

We obtain inductively

$$\|(\Omega u)^n(t)-(\Omega v)^n(t)\| \leq \frac{\|u-v\|_{\mathcal{C}}}{n!(\Gamma(\alpha)^n}\|\mu\|^n_{L^1(I)}.$$

If n is large enough then we can get

$$\frac{\|\mu\|^n_{L^1(I)}}{n!\Gamma(\alpha)^n} < 1$$

By a generalization of the Banach fixed point theorem, Equation (3.2) has a unique mild solution. ■

Remark 3.2. *This result is new even in the case of finite dimensional space for the fractional derivatives in the sense of Caputo or Riemann-Liouville..*

Theorem 3.4. *(Krasnoselskii). Let X be a closed convex and nonempty subset of a Banach space E. Let A_1, A_2 be two mappings such that*

(a) $A_1u + A_2v \in X$ *for every* $u, v \in E$;

(b) A_1 *is a compact and continuous;*

(c) A_2 *is a contraction.*

We further include the following assumption for the result that follows.

(H4): $\|f(t,u)\| \leq \mu(t), \forall (t,u) \in I \times E, \mu \in L^1(I, \mathbb{R}^+)$.

Theorem 3.5. *Suppose assumptions **(H1)**, **(H3)**, and **(H4)**. Let $b < 1$. Then Equation (1.1) has at least one mild solution on I*

Proof. Let's take r such that

$$r \geq \|u_0\| + P + \frac{\kappa(T)^\alpha \|\mu\|_{L^1}}{\Gamma(\alpha+1)}.$$

Then we define on B_r the operators A_1, A_2 by

$$(A_1u)(t) = \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}f(s,u(s))ds \tag{3.2}$$

and

$$(A_2u)(t) = \frac{(\kappa(t)-\kappa(0))^{\sigma-1}}{\Gamma(\sigma)}(u_0 - g(u)). \tag{3.3}$$

Let's prove that if $u, v \in B_r$ implies $A_1u + A_2v \in B_r$.

Indeed we have

$$\|A_1u + A_2v\| \leq \|\Delta^\sigma_\kappa(t,0)(u_0 - g(u))\| + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u(s))\|ds$$
$$\leq \|u_0\| + P + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\mu(s)ds$$

$$\leq \|u_0\| + P + \frac{\|\mu\|_{L^1}}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} ds$$
$$= \|u_0\| + P + \frac{\|\mu\|_{L^1}}{\Gamma(\alpha)} \left(\frac{(\kappa(t)-\kappa(0))^{\alpha}}{\alpha}\right)$$
$$\leq \|u_0\| + P + \frac{\kappa(T)^{\alpha}\|\mu\|_{L^1}}{\Gamma(\alpha+1)}$$
$$\leq r.$$

In view of **(H3)**, B is a contraction mapping since $b < 1$. Continuity of u implies that $(A_1u)(t)$ is continuous based on of **(H1)**.

Let's observe that A_1 is uniformly bounded on B_r. This is because of the inequality

$$\|(A_1u)(t)\| \leq \frac{\kappa(T)^{\alpha}\|\mu\|_{L^1}}{\Gamma(\alpha+1)}.$$

In addition, we proceed to show that $(A_1u)(t)$ is equicontinuous.

Let $t_1 \in I,\ , t_2 \in I$ and $u \in B_r$. Since u is bounded on the compact set $I \times B_r$, then $\sup_{(t,u)\in J\times B_r} \|f(t,u)\| := K < \infty$, we get

$$\begin{aligned}
\|A_1u(t_1) - A_1u(t_2)\| &= \frac{1}{\Gamma(\alpha)} \left\| \int_0^{t_1} \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1} + f(s,u(s))ds \right.\\
&\left. - \frac{1}{\Gamma(\alpha)} \int_0^{t_2} \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1} f(s,u(s))ds \right\|\\
&= \frac{1}{\Gamma(\alpha)} \left\| \int_{t_2}^{t_1} \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1} f(s,u(s))ds \right.\\
&\left. - \frac{1}{\Gamma(\alpha)} \int_0^{t_2} \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1} f(s,u(s))ds \right\|\\
&\leq \frac{1}{\Gamma(\alpha)} \left\| \int_{t_2}^{t_1} \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1} f(s,u(s))ds \right\|\\
&+ \frac{1}{\Gamma(\alpha)} \left(\left\| \int_0^{t_2} [\kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1} - \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1}] f(s,u(s))ds \right\| \right)\\
&\leq \frac{K}{\Gamma(\alpha)} \left\| \frac{(\kappa(t_1)-\kappa(t_2))^{\alpha}}{\alpha} - \frac{(\kappa(t_2)-\kappa(0))^{\alpha}}{\alpha} + \frac{(\kappa(t_1)-\kappa(0))^{\alpha}}{\alpha} \right\|\\
&\leq \frac{K}{\Gamma(\alpha+1)} \|(\kappa(t_1)-\kappa(t_2))^{\alpha} - (\kappa(t_2)-\kappa(0))^{\alpha} + (\kappa(t_1)-\kappa(0))^{\alpha}\|
\end{aligned}$$

which is independent of x and $|A_1u(t_1) - A_1u(t_2)| \to 0$ as $t_2 \to t_1$. Therefore, $A_1u(t)$ is equi-continuous. Using Arzela-Ascoli's theorem, $A_1(B_r)$ is a relatively compact subset of $\mathcal{C}$, which implies that operator A_1 is compact. We conclude the proof using Krasnoselskii's theorem. ■

4. Extension Results

For our next results, we make the following assumptions.
(H5): For every $u \in E$, there exists $c_f > 0$, such that

$$\|f(t,u)\| \leq c_f(1 + \|u\|)$$

and for every $u \in C(I,E)$, there exists a $c_g \in (0,1)$ such that

$$\|g(u)\| \leq c_g(1+\|u\|_{\mathcal{C}}),$$

(H6): For every $t \in I$, the set

$$\Omega = \{\kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} f(s,u(s)) : \ u \in C(J,E), 0 \leq s \leq t\}$$

is relatively compact.

Theorem 4.1. *[15] Let u, v be two integrable functions and g continuous, with domain $[a,b]$. Let $\kappa \in C^1(I)$ be an increasing function such that $\kappa'(t) \neq 0, \forall t \in I$. Assume that*
(1) u, v are nonnegative;
(2) g is nonnegative and nondecreasing.
If

$$u(t) \leq v(t) + g(t) \int_a^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} u(\tau) d\tau,$$

then

$$u(t) \leq v(t) + \int_a^t \sum_{k=1}^{\infty} \frac{[g(t)\Gamma(\alpha)]^k}{\Gamma(\alpha k)} \kappa'(\tau)(\kappa(t)-\kappa(\tau))^{\alpha k-1} v(\tau) d\tau.$$

Corollary 4.1. *[15] Let v be a nondecreasing function on $[a,b]$. Under the hypothesis of the above theorem, we have*

$$u(t) \leq v(t) E_\alpha(g(t)\Gamma(\alpha)[\kappa(t)-\kappa(\tau)]^\alpha), \ \forall t \in [a,b].$$

Lemma 4.2. *For our next result, we first prove that there exists $\xi > 0$ such that $\|u(t)\| \leq \xi, \ \ \forall t \in I$.*

Proof. Using the assumptions of Theorem 3.2 and hypothesis **(H5)**, one derives the estimate

$$\begin{aligned}
\|u(t)\| &\leq \|u_0 - g(u)\| + \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} \|f(s,x(s))\| ds, \quad t \in I. \\
&\leq \|u_0 - g(u)\| + \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} c_f(1+\|u(s)\|) ds \\
&\leq \|u_0\| + c_g + c_g\|u\|_{\mathcal{C}} + \frac{c_f}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} ds \\
&\quad + \frac{c_f}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} \|u(s)\| ds.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
(1-c_g)\|u\|_{\mathcal{C}} \leq \ & \|u_0\| + c_g + \frac{c_f}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} ds \\
& + \frac{c_f}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} \|u(s)\| ds.
\end{aligned}$$

Therefore,

$$\|u(t)\| \leq \frac{\Gamma(\alpha+1)(\|u_0\|+c_g) + c_f \kappa(T)^\alpha}{(1-c_g)\Gamma(\alpha+1)} + \frac{c_f}{(1-c_g)\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} \|u(s)\| ds.$$

Using Theorem 4.2 and Corollary 4.3, we get

$$\|u(t)\| \leq \frac{\Gamma(\alpha+1)(\|u_0\|+c_g) + c_f \kappa(T)^\alpha}{(1-c_g)\Gamma(\alpha+1)} \sum_{n=0}^{\infty} \frac{\left(\frac{c_f \Gamma(\alpha)\kappa(T)^\alpha}{(1-c_g)\Gamma(\alpha)}\right)^n}{\Gamma(n\alpha+1)}$$

$$\leq \frac{\Gamma(\alpha+1)(\|u_0\|+c_g)+c_f\kappa(T)^\alpha}{(1-c_g)\Gamma(\alpha+1)}\sum_{n=0}^{\infty}\frac{(c_f\kappa(T)^\alpha)^n}{(1-c_g)^n\Gamma(n\alpha+1)}$$

Since $\sum_{n=0}^{\infty}\frac{(c_f\kappa(T)^\alpha)^n}{(1-c_g)^n\Gamma(n\alpha+1)}$ is the Mittag-Leffler function, it suffices that
$\|u(t)\| < \xi$ ■

Theorem 4.2. *Suppose **(H1)**, **(H5)**, **(H6)** hold and let*

$$c_g + \frac{c_f\kappa(T)^\alpha}{\Gamma(\alpha+1)} < 1.$$

Then the system has at least one mild solution.

Proof. The proof is given in the following steps:
Let (3) and (4) be defined on $\mathcal{C}_\xi := \{u \in C(I,E) : \|u(t)\| \leq \xi, t \in I\}$.
Step 1: We first show that $Au + Bv \in C_\xi$ for every $u, v \in C_\xi$. Using Equations (3)-(4), we have the following estimate

$$\begin{aligned}
\|(Au)(t)+(Bv)(t)\| &\leq \|u_0\| + \|g(v)\| + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u(s))\|ds \\
&\leq \|u_0\| + c_g(1+\|v\|_{\mathcal{C}}) + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}c_f(1+\|u(s)\|)ds \\
&\leq \|u_0\| + c_g(1+\xi) + \frac{c_f(1+\xi)}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}ds \\
&\leq \|u_0\| + c_g(1+\xi) + \frac{c_f(1+\xi)\kappa(T)^\alpha}{\Gamma(\alpha+1)} \\
&\leq \xi.
\end{aligned}$$

which implies that $Au + Bv \in C_\xi$.
Step 2: B is a contractive operator on C_ξ.
Indeed if we take any $v_1, v_2 \in C_\xi$, then we have

$$\|Bv_1 - Bv_2\|_{\mathcal{C}} = \|g(v_1) - g(v_2)\| \leq L_g\|v_1 - v_2\|_{\mathcal{C}}.$$

Therefore B is a contraction mapping.
Step 3: A is a continuous mapping.
Using **(H1)**, if $\{u_n\}$ is a sequence in C_ξ so that $u_n \to u$ in C_ξ, then,

$$f(s,u_n(s)) \to f(s,u(s)) \quad \text{as} \quad n \to \infty.$$

For $t \in I$,we have

$$\|(Au_n)(t)-(Au)(t)\| \leq \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u_n(s)), f(s,u(s))\|ds$$

Using **(H7)**, for $t \in I$,

$$\|f(s,u_n(s)) - f(s,u(s))\| \leq L_f(\xi)\|u_n(s)-u(s)\| \leq 2\xi L_f(\xi).$$

Also, using the fact that $s \to 2\xi L_f(\xi)\kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}$ is integrable, Lebesgue's Dominated Convergence Theorem gives that

$$\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u_n(s)), f(s,u(s))\|ds \to 0.$$

As $n \to \infty$, $Au_n \to Au$. Thus, A is continuous.
Step 4: A is a compact mapping.
If $\{u_n\}$ is a sequence on C_ξ, then

$$\begin{aligned}\|(Au_n)(t)\| &\leq \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}\|f(s,u_n(s))\|ds \\ &\leq \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1}c_f(1+\|u_n(s)\|)ds \\ &\leq \frac{(1+\xi)\kappa(T)^\alpha c_f}{\Gamma(\alpha+1)}.\end{aligned}$$

Therefore, $\{u_n\}$ is uniformly bounded.
Next, we prove the equicontinuity of $\{Au_n\}$. Let $0 \leq t_1 \leq t_2 \leq T$. Then

$$\begin{aligned}&\|(Au_n)(t_1)-(Au_n)(t_2)\| \\ &\leq \frac{1}{\Gamma(\alpha)} \int_0^{t_1} \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1} - \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1}\|f(s,u_n(s))\|ds \\ &\quad + \frac{1}{\Gamma(\alpha)} \int_{t_1}^{t_2} \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1}\|f(s,u_n(s))\|ds \\ &\leq \frac{c_f}{\Gamma(\alpha)} \int_0^{t_1} \kappa'(s)(\kappa(t_1)-\kappa(s))^{\alpha-1} - \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1}(1+\|u_n(s)\|)ds \\ &\quad + \frac{c_f}{\Gamma(\alpha)} \int_{t_1}^{t_2} \kappa'(s)(\kappa(t_2)-\kappa(s))^{\alpha-1}(1+\|u_n(s)\|)ds \\ &\leq \frac{c_f(1+\xi)}{\Gamma(\alpha+1)}\left((\kappa(t_1)-\kappa(t_0))^\alpha - (\kappa(t_2)-\kappa(0))^\alpha + 2(\kappa(t_2)-\kappa(t_1))^\alpha\right) \\ &\leq \frac{c_f(1+\xi)}{\Gamma(\alpha+1)}(\kappa(t_2)-\kappa(t_1))^\alpha.\end{aligned}$$

As $t_2 \to t_1$, $(\kappa(t_2)-\kappa(t_1))^\alpha \to 0$, and thus $\{Au_n\}$ is equicontinuous. In view of (H6) and theorem 3.4, $\overline{conv}K$ is a compact set. For every $t^* \in I$,

$$\begin{aligned}(Au_n)(t^*) &= \frac{1}{\Gamma(\alpha)} \int_0^t \kappa'(s)(\kappa(t^*)-\kappa(s))^{\alpha-1} f(s,u_n(s))ds \\ &= \frac{1}{\Gamma(\alpha)} \lim_{t\to\infty} \frac{t^*}{k}\left[\kappa'\left(\frac{it^*}{k}\right)\left(\kappa(t^*)-\kappa\left(\frac{it^*}{k}\right)\right)^{\alpha-1} f\left(\frac{it^*}{k}, u_n\left(\frac{it^*}{k}\right)\right)\right] \\ &= \frac{t^*}{k}\eta_n,\end{aligned}$$

where

$$\eta_n = \lim_{t\to\infty} \frac{1}{k}\left[\kappa'\left(\frac{it^*}{k}\right)\left(\kappa(t^*)-\kappa\left(\frac{it^*}{k}\right)\right)^{\alpha-1} f\left(\frac{it^*}{k}, u_n\left(\frac{it^*}{k}\right)\right)\right].$$

Since $\overline{conv}K$ is convex and compact, $\eta_n \in \overline{conv}K$. Hence for every $t^* \in I$, the set $\{Au_n\}$ is relatively compact. From Ascoli-Arzela theorem, every $\{Au_n(t)\}$ contains a uniformly convergent subsequence $\{Au_{n_k}(t)\}(k = 1, 2, ...)$ on I. Thus, the set $\{Au : u \in C_\xi\}$ is relatively compact. Therefore A is a completely continuous mapping. Using Krasnoselskii's Theorem, we conclude that $A + B$ possesses a fixed point on C_ξ which is the mild solution to the system (1.1). The proof is complete. ■

5. Stability Result

In this last section, we establish the stability of (1.1) with regard to Ulam-Hyers stability. We first define the following mapping $\Lambda : C(I,E) \to C(I,E)$ as follows:

$$\Lambda v(t) = ^{\mathbb{H}} D_{0^+}^{\alpha,\beta;\kappa} v(t) - f(t, v(t)), t \in I.$$

Let $\epsilon > 0$ be given and $v(t) \in C(I,E)$ satisfy

$$v(t) = \frac{(\kappa(t)-\kappa(0))^{\sigma-1}}{\Gamma(\sigma)}(v_0 - g(v)) + \frac{1}{\Gamma(\alpha)}\int_0^t \kappa'(s)(\kappa(t)-\kappa(s))^{\alpha-1} f(s, v(s))ds, \quad t \in J.$$

Definition 5.1. *[14] Problem (1) is said to be Ulam-Hyers stable (or stable in the sense of Ulam-Hyers) if for every $\epsilon > 0$*

$$\|\Lambda y\| \le \epsilon,$$

and for every mild solution y of (1.1),there exists $\rho > 0$ and a mild solution $u \in C(I,E)$ to (1.1) such that

$$\|u(t) - v(t)\| \le \rho\epsilon^*,$$

where $\epsilon^* > 0$ and depends on ϵ.

Definition 5.2. *[13] Let $m \in C(R^+, R^+)$ so that for every mild solution y of (1.1), there exists a mild solution $x \in C(I,E)$ of problem (1.1) such that*

$$\|u(t) - v(t)\| \le m\epsilon^*, t \in I.$$

Definition 5.3. *[14] Problem (1.1) is called Ulam-Hyers-Rassias stable with respect to $\Theta \in C(I, R^+)$ if for $\epsilon > 0$,*

$$\|\Lambda v(t)\| \le \epsilon\Theta(t), \ t \in I.$$

and there exists $\rho > 0$ and $v \in C(I,E)$ such that

$$\|u(t) - v(t)\| \le \rho\epsilon_*\Theta(t), t \in I,$$

and $\epsilon_* > 0$ depends on ϵ

Theorem 5.4. *Assume $\|f(t,u(t))\| \le p(t)q(\|u\|)$ where $p \in C(I, R_+)$ and $q : R_+ \to R_+$, and $\Phi < 1$. Then problem (1.1) is both Ulam-Hyers and generalized Ulam-Hyers stable.*

Proof. Let $x \in C(I,E)$ be a solution of (1.1),and let y be any solution satisfying Definition 5.1. We determine that the operators Λ and $F - Id$ (identity operator) are equivalent for all solutions $v \in C(I,E)$ of (1.1) satisfying $\Phi < 1$. Thus , using the fixed point property of operator F, we conclude that

$$\begin{aligned}\|v(t) - u(t)\| &= \|v(t) - Fv(t) + Fv(t) - u(t)\| \\ &= \|v(t) - Fv(t) + Fv(t) - Fu(t)\| \\ &\le \|Fv(t) - Fu(t)\| + \|Fv(t) - v(t)\| \\ &\le \Phi\|u - v\| + \epsilon,\end{aligned}$$

because $\Phi < 1$ and and $\epsilon > 0$,

$$\|u - v\| \le \frac{\epsilon}{1-\Phi}.$$

Fixing $\epsilon_* = \dfrac{\epsilon}{1-\Phi}$, and $\rho = 1$ we get the Ulam-Hyers stability. Using $m\epsilon = \dfrac{\epsilon}{1-\Phi}$, we get the generalized Ulam-Hyers stability. ■

Corollary 5.1. *If definition 5.3 is satisfied for $\Theta \in C(I, R^+)$, and*

$$L < \frac{\Gamma(\alpha+1)(1-b)}{\kappa(T)^\alpha},$$

problem (1) is Ulam-Hyers-Rassias with respect to Θ.

Proof. Directly follows from the proof of Theorem 5.4 where

$$\|u(t) - v(t)\| \le \epsilon_* \Theta(t),\ t \in I,$$

and

$$\epsilon_* = \frac{\epsilon}{1-\Phi}.$$

The proof is complete. ■

6. Conclusion

A generalized Cauchy problem is the central focus of this paper formulated using the generalized fractional operator called κ-Hilfer operator. The existence and uniqueness results of the abstract model are studied with nonlocal conditions using classical fixed point theorems. Furthermore, a stability result for the abstract problem is obtained in the sense of Ulam-Hyers-Rassians. The results obtained in this paper are very useful in many applications where κ-Hilfer operator is being adopted. They generalize many recent results in this field. The arguments used here can be adapted to many other problems in infinite dimensional spaces.

7. Acknowledgement

This research was conducted when the authors were on the Adjoint 2022 program at the Mathematical Science Research Institute (MSRI) at the University of California at Berkeley. The authors are very grateful for the support they received.

References

[1] S. Abbas, M. Benchohra, G.M. N'Guérékata, *Topics in Fractional Differential Equations*, Developments in Mathematics, 27. Springer, New York, 2012.

[2] S. Abbas, M. Benchohra, G.M. N'Guérékata, *Advanced Fractional Differential and Integral Equations*, Mathematics Research Developments, Nova Science Publishers, Inc, New Yorj, 2015.

[3] S. Abbas, M. Benchohra, G.M. N'Guérékata, Instantaneous and noninstantaneous impulsive integro-differential equations in Banach spaces, *Abstr. Appl. Anal.* 2020, 1-8.

[4] L. Byszewski, Theorems about the existence and uniqueness of solutions of a semilinear nonlocal Cauchy problem, *J. Math. Anal. Appl.,* **162**(1991), 494–505.

[5] L. D. Baleanu, G.C. Wu, S.D. Zeng, Chaos analysis and asymptotic stability of generalized Caputo fractional differential equations, *Chaos Solitons Fractals*, **102**(2017), 99–105.

[6] L. D. Baleanu, Z.B. Guvenc, J.T Machado, *New Trends in Nanotechnology and Fractional Calculus Applications* Springer: New York, NY, USA, (2010).

[7] K. Deng, Exponential decay of solutions of semilinear parabolic equations with nonlocal initial conditions, *J. Math. Anal. Appl.,* **179**(1993), 630–637.

[8] O.S. Iyiola, F.D. Zaman, A fractional diffusion equation model for cancer tumor, *Am. Inst. Phys. Adv.,* 2014(4), 107121.

[9] O.S. Iyiola, B. Oduro, T. Zabilowicz, B. Iyiola, D. Kenes, System of time fractional models for COVID-19: modeling, analysis and solutions, *Symmetry*, 2021(13), 787.

[10] D. Kumar, A.R. Seadawy, A.K. Joardar, System of Time Fractional Models for COVID-19: Modeling, Analysis and Solutions, *Chin. J. Phys.,* 2018(56), 75–85

[11] J. Vanterler da C. Sousa, E. Capelas Oliveira, On the $\kappa-$Hilfer fractional derivatives, *Common Nonlinear Sci Numer Simul.,* **60**(2018), 72–91.

[12] G. M. N'Guérékata, A Cauchy problem for some fractional abstract differential equation with non local condition, *Nonlinear Analysis*, **70**(2019), 1873–1879.

[13] Vanterler da C. Sousa, E. Capelas Oliveira, On the Ulam-Hyers-Rassias stability for non-linear fractional differential equations using the κ-Hilfer operator, *J. Fixed Point Theory Appl.,* 20(3)(2018), 96–113.

[14] J. Vanterler da C. Sousa, E. Capelas Oliveira, Ulam-Hyers stability of a nonlinear fractional Volterra integro-differential equation, *Appl. Math. Lett.,* **81**(2018), 50–56.

[15] J. Vanterler da C. Sousa, E. Capelas Oliveira, A Gronwall inequality and the Cauchy-type problem by means of $\kappa-$ Hilfer operator, *Diff. Equ. & Appl.,* **11**(1)(2019), 87–106.

[16] M.S. Abdo, S.K. Panchal, Fractional integrodifferential equations involving κ-Hilfer fractional derivative, *Adv. Appl. Math. Mech.,* **11**(2019), 338–359.

[17] R. C. Koeller , Applications of fractional calculus to the theory of viscoelasticity, *Trans. ASME J. Appl. Mech.,* **51**(2)(1984), 299–307.

[18] R. Hilfer, *Applications of Fractional Calculus in Physics*, World Scientific, River Edge, New Jersey, 2000.

[19] R. Hilfer, L. Anton, Fractional master equations and fractal time random walks, *Phys. Rev.,* (1995), 51, R848–R851.

[20] K.B. Lima, J. Vanterler da C. Souza, E. Capelas de Oliveira, Ulam-Hyers type stability for κ-Hilfer fractional differential equations with impulses and delay, *Computational and Applied Maths,* (2021) 40:293.

[21] F. Mainardi, *Fractional Calculus and Waves in Linear Viscoelasticity*, Imperial College Press: London, UK, 2010.

[22] H. Nasrolahpour, A note on fractional electrodynamics, *Commun. Nonlinear. Sci. Numer. Simul.,* **18**(2013), 2589–2593.

[23] F. Norouzi, G.M. N'Guérékata, A study of κ-Hilfer differential system with applications in financial crisis, *Chaos, Solitons and Fractals:X*, 6(2021), 100056.

[24] F. Norouzi, G.M. N'Guérékata, Existence results to a κ-Hilfer neutral fractional evolution equations with infinite delay, *Nonauton. Dyn. Syst.,* **8**(1)(2021), 101–124.

[25] C.J. SOUSA, CAPELAS, E. DE OLIVEIRA, Ulam-Hyers stability of a nonlinear fractional Volterra integro-differential equation, *Appl. Math. Lett.,* **81**(2018), 50–56.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 94–106.
http://doi.org/10.26637/mjm1101/007

On distinguishing labelling of groups for the conjugation action

Madhu Dadhwal*[1] and Pankaj[2]
[1,2] *Department of Mathematics and Statistics, Himachal Pradesh University, Summer Hill, Shimla-171005, Himachal Pradesh, India.*

Received 09 March 2022; Accepted 24 November 2022

Abstract. In this paper, the conjugation action of various classes of groups on themselves is studied to obtain their distinguishing numbers along with a distinguishing labelling for the said action. An equivalent condition concerning the existence of a 2-distinguishing labelling for the action of a group $\mathcal{G}$ on a $\mathcal{G}$-set $\mathcal{X}$ and a partition of $\mathcal{X}$ into two subsets is established. Also, the distinguishing number for the conjugation action of a group acting on itself is completely characterized.
AMS Subject Classifications: 05C25, 05C78, 20D60.

Keywords: Distinguishing number, distinguishing group actions, conjugation action, distinguishing labelling of sets.

1. Introduction

The concept of distinguishing number originated from an elementary problem known as the Frank Rubin's Key problem [5], which states that:

Professor X has n keys on a circular key ring, but he can not see them. Now, the question arises: How many shapes does Professor X need to use in order to keep n keys on the ring and still be able to select the proper key by feel?

Albertson and Collins [1] popularized the aforementioned problem and took it in the realm of graphs by connecting it with the action of symmetries of a graph on its set of vertices. They define the distinguishing number for a graph G to be the minimum k for which the vertices of G can be labeled from 1 to k such that no non-trivial symmetry (automorphism) of the graph G preserves all of the vertex labels.

The surprise answer to the above problem is that only two different handle shapes are required if six or more keys are there in the key ring; but minimum three different handle shapes are required to distinguish the keys for three, four and five keys in the key ring. This motivated us to study and understand the concept of distinguishing labelling and distinguishing number for the conjugation action of a group on itself.

Further, in [6], Tymoczko generalized the notion of the distinguishing number for an arbitrary group action that is, the distinguishing number for the action of an arbitrary group $\mathcal{G}$ on a $\mathcal{G}$-set $\mathcal{X}$ which is not necessarily the action of the automorphism group of a graph on the set of vertices of that graph.

The main objective of this paper is to compute the distinguishing number and a distinguishing labelling for the conjugation action of some well known classes of groups such as Q_{4n} (dicyclic group), V_{8n}, U_{6n} and SD_{8n} (semi-dihedral group) acting on itself. It is observed that for each of these classes of groups the distinguishing number is 2 and this concurrence of the distinguishing number for the conjugation action for all aforementioned groups arises a natural question as follows:

Does there exist any group $\mathcal{G}$ with distinguishing number more than 2 for its conjugation action on itself?

This question stimulates us to completely characterize the distinguishing number for the conjugation of a group on itself. We answer the above question in the sixth section of this paper. In order to achieve this aim, an equivalent condition for the existence of a 2-distinguishing labelling for the action of a group on a $\mathcal{G}$-set $\mathcal{X}$ is established. In addition, by using this condition and some other results, we completely characterize the distinguishing number for the group action of a group $\mathcal{G}$, acting on itself by the conjugation action.

The main results proved in the present paper are:

*__Corresponding author.__ Email addresses: **mpatial.math@gmail.com** (Madhu Dadhwal), **pankajratramath@gmail.com** (Pankaj)

(1) The distinguishing numbers and a corresponding distinguishing labelling for the conjugation action of some well known classes of groups such as Q_{4n} (Dicyclic group), V_{8n}, U_{6n} and SD_{8n} (semi-dihedral group) are computed.

(2) A relation between a distinguishing labelling for the action of a group $\mathcal{G}$ on a $\mathcal{G}$-set $\mathcal{X}$ and a partition of the underlying set $\mathcal{X}$ has been identified.

(3) The distinguishing number for the conjugation action of a group on itself is completely characterized.

Our notations are as follows: $\mathcal{G}$ is a group and $\mathcal{X}$ is a $\mathcal{G}$-set. The stabilizer of a subset $\mathcal{A} \subseteq \mathcal{X}$ is $Stab_{\mathcal{G}}(\mathcal{A}) = \{g \in \mathcal{G} : ga = a \text{ for all } a \in \mathcal{A}\}$. We begin this article with the following definitions:

Definition 1.1. *[6] Let $\mathcal{G}$ be a group acting on a nonempty set $\mathcal{X}$. A map $\phi : \mathcal{X} \longrightarrow \{1, 2, ..., k\}$ is said to be a k-distinguishing labelling of the action of $\mathcal{G}$ on the set $\mathcal{X}$ if the only group elements that preserve the labelling are in $Stab_{\mathcal{G}}(\mathcal{X})$. Equivalently, the map ϕ is a k-distinguishing labelling if $\{g : \phi \circ g\ (x) = \phi(x)$ for all $x \in \mathcal{X}\} = Stab_{\mathcal{G}}(\mathcal{X})$.*

Definition 1.2. *[6] The distinguishing number $\mathcal{D}_{\mathcal{G}}(\mathcal{X})$ of the set $\mathcal{X}$ with a given group action of $\mathcal{G}$ on $\mathcal{X}$ is the minimum k for which there is a k-distinguishing labelling.*

It is also pertinent to notice that if the $\mathcal{G}$-set $\mathcal{X}$ is equal to $\mathcal{G}$, then under the conjugation action of the group $\mathcal{G}$ on $\mathcal{X}$, the set $Stab_{\mathcal{G}}(\mathcal{X}) = \{g \in \mathcal{G} : ghg^{-1} = h \text{ for all } h \in \mathcal{X}\} = Z(\mathcal{G})$, is the center of the group $\mathcal{G}$. In this case, a map $\phi : \mathcal{X} \longrightarrow \{1, 2, ..., k\}$ is said to be a k-distinguishing labelling for the conjugation action of $\mathcal{G}$ on the set $\mathcal{X}$ if the only group elements that preserve the labelling are in the center of the group $\mathcal{G}$. More precisely, $\{g : \phi(ghg^{-1}) = \phi(h) \text{ for all } h \in \mathcal{G}\} = Z(\mathcal{G})$. Also, the minimum k for which there exists a k-distinguishing labelling ϕ satisfying the above equality is the distinguishing number for the conjugation action of a group on itself.

Moreover, it is not difficult to see that a non-Abelian group $\mathcal{G}$, under the conjugation action, cannot act on itself by fixing each of its elements.

The above observation, in view of Proposition 2.1 [6], leads to the following theorem.

Theorem 1.3. *The distinguishing number for the conjugation action of a non-Abelian group $\mathcal{G}$ acting on itself, is at least* 2.

Throughout this paper, the action of a group $\mathcal{G}$ represents the conjugation action of the group $\mathcal{G}$ on itself, unless stated otherwise.

2. Distinguishing Labelling for the Dicyclic group Q_{4n}

In group theory, a dicyclic group, $\mathcal{G} = Q_{4n} = \langle a, b : a^{2n} = 1, a^n = b^2, ab = ba^{-1} \rangle$ is a non-Abelian group of order $4n$ for $n > 1$, which can be viewed as an extension of the cyclic group of order 2 by a cyclic group of order 2n, described by an exact sequence as follows:

$$1 \longrightarrow \mathbb{Z}_2 \longrightarrow Q_{4n} \longrightarrow \mathbb{Z}_{2n} \longrightarrow 1.$$

The dicyclic group is a subgroup of the unit quaternions generated by the elements a and j; where $a = e^{\frac{i\pi}{n}}$ is an n^{th} root of unity. Further, the number of conjugacy classes in Q_{4n} is $n + 3$ namely;

$\{1\}$,
$\{a^n\}$,
$\{a^r, a^{-r}\}; (1 \leq n \leq n-1)$,
$\{a^{2j}b : 0 \leq j \leq n-1\}$,
$\{a^{2j+1}b : 0 \leq j \leq n-1\}$.

Note that one can realize the set of elements of the group Q_{4n} as $Q_{4n} = \{a^k, a^k b : k = 0, 1, 2, ..., 2n-1\}$ and we begin this section with the following lemma.

Lemma 2.1. *For the dicyclic group Q_{4n}, we have*

(1) $ba^j = a^{-j}b$, for any integer j.

(2) For $n > 1$, $Z(Q_{4n}) = \{1, a^n\}$.

As the group Q_{4n} is non-Abelian for $n \geq 2$, while it is Abelian for $n = 1$, therefore in consideration of Proposition 2.1 [6], we have

Proposition 2.2. *If Q_4 acts on itself by conjugation, then $\mathcal{D}_{Q_4}(Q_4) = 1$.*

Before proceeding further, for convenience sake, we partition the set of elements of Q_{4n} for $n \geq 2$, as follows:

$$Q_{4n} = \mathcal{X}_n^{'} \cup \mathcal{X}_n, \text{ where}$$

$\mathcal{X}_n^{'} = \{a^{n+1}, a^{n+2}, ..., a^{2n-1}, a^2b, a^3b\}$ and $\mathcal{X}_n = Q_{4n} \backslash \mathcal{X}_n^{'} = \{1, a, ..., a^n, b, a^4b, a^6b, ..., a^{2n-2}b, ab, a^5b, a^7b, ..., a^{2n-1}b\}$.

Proposition 2.3. *The stabilizer of the subset $\mathcal{X}_n^{'} = \{a^{n+1}, a^{n+2}, ..., a^{2n-1}, a^2b, a^3b\}$ in Q_{4n} is $Stab_{Q_{4n}}(\mathcal{X}_n^{'}) = \{1, a^n\}$.*

Proof. By Lemma 2.1(2), it is enough to show that $Stab_{Q_{4n}}(\mathcal{X}_n^{'}) \subseteq \{1, a^n\}$. For this, let $g \in Stab_{Q_{4n}}(\mathcal{X}_n^{'}) \cap Q_{4n}$. Then $ghg^{-1} = h$, for all $h \in \mathcal{X}_n^{'}$ and we know that an element in Q_{4n} is of the form a^ib^j, where $0 \leq i \leq 2n-1$, $j = 0, 1$, so is g. We claim that if $g = a^ib^j \in Stab_{Q_{4n}}(\mathcal{X}_n^{'})$, then $j = 0$. If possible, let $j \neq 0$, this implies that $g = a^ib$, in this case there exists an element $h = a^{n+1} \in \mathcal{X}_n^{'}$ satisfying $ghg^{-1} = (a^ib)a^{n+1}(b^{-1}a^{-i}) = a^{-n-1} = a^{n-1} \neq a^{n+1} = h$, as $n \geq 2$. This forces j to be 0. Thus, an element $g \in Stab_{Q_{4n}}(\mathcal{X}_n^{'})$ can be of the form $a^i, 0 \leq i \leq 2n-1$. Now, for $h = a^2b \in \mathcal{X}_n^{'}$, we have $ghg^{-1} = h$ if and only if $(a^i)a^2b(a^{-i}) = a^2b$ if and only if $a^{2i+2}b = a^2b$ if and only if $i = 0$ or n. Thus $Stab_{Q_{4n}}(\mathcal{X}_n^{'}) \subseteq \{1, a^n\}$. ■

The next theorem provides a 2-distinguishing labelling for the conjugation action of Q_{4n} on itself.

Theorem 2.4. *Let Q_{4n} with $n \geq 2$, acts on itself by conjugation. Then the map $\phi : Q_{4n} \longrightarrow \{1, 2\}$ defined by $\phi(x) = \begin{cases} 1\,; & \text{if } x \in \mathcal{X}_n \\ 2\,; & \text{otherwise} \end{cases}$, is a 2-distinguishing labelling of Q_{4n}.*

Proof. Define a labelling, $\phi : Q_{4n} \longrightarrow \{1, 2\}$ for the conjugation action of Q_{4n} on itself by $\phi(x) = 1$, for all $x \in \mathcal{X}_n$ and $\phi(x) = 2$ otherwise. Now, for ϕ to be a 2-distinguishing labelling, it is sufficient to prove that $\{g \in Q_{4n} : \phi(ghg^{-1}) = \phi(h),\ \text{for all } h \in Q_{4n}\} = \{1, a^n\}$. Further, by Proposition 2.3, it is enough to show that if $g \in Q_{4n} \setminus \{1, a^n\}$, there exists an element $h \in \mathcal{X}_n^{'}$ such that $ghg^{-1} \notin \mathcal{X}_n^{'}$, that is $ghg^{-1} \in \mathcal{X}_n$. For this assume that $g \in Q_{4n} \setminus \{1, a^n\}$, if $g = a^i;\ i \in \{0, 1, 2, ..., 2n-1\} \setminus \{0, n\}$, choose $h = a^2b \in \mathcal{X}_n^{'}$, then $ghg^{-1} = a^ia^2ba^{-i} = a^{2i+2}b$ and clearly, for $i \in \{0, 1, 2, ..., 2n-1\} \setminus \{0, n\}$, $a^{2i+2}b \in \mathcal{X}_n$. On the other hand, if $g = a^ib;\ i \in \{0, 1, 2, ..., 2n-1\}$ and if $i \neq 2$ or $n+2$, take $h = a^2b \in \mathcal{X}_n^{'}$, then $ghg^{-1} = (a^ib)a^2b(b^{-1}a^{-i}) = a^{i-2}ba^{-i} = a^{2i-2}b \in \mathcal{X}_n$. Otherwise, when $i = 2$ or $n+2$, choose $h = a^3b \in \mathcal{X}_n^{'}$ and we have $ghg^{-1} = (a^ib)a^3b(b^{-1}a^{-i}) = a^{i-3}ba^{-i} = a^{2i-3}b = ab \in \mathcal{X}_n$. Thus $\{g \in Q_{4n} : \phi(ghg^{-1}) = \phi(h),\ \text{for all } h \in Q_{4n}\} \subseteq \{1, a^n\}$ and the result follows from Lemma 2.1. ■

By combining Theorem 1.3 and Theorem 2.4 we have

Theorem 2.5. *If Q_{4n} $(n > 1)$ acts on itself by conjugation, then $\mathcal{D}_{Q_{4n}}(Q_{4n}) = 2$.*

3. Distinguishing Labelling for V_{8n} action

The group V_{8n} is introduced by James and Liebeck [4], for an odd positive integer n as follows:

$$V_{8n} = \langle a, b : a^{2n} = b^4 = 1, aba = b^{-1}, ab^{-1}a = b\rangle.$$

Later, Darafsheh and Poursalavati [2], observed that with the above presentation, the group V_{8n} can also be defined for an arbitrary n. However, the conjugacy classes of the group V_{8n} differ, depending upon whether n is an even or an odd positive integer. When n is odd, the group V_{8n} has $2n+3$ conjugacy classes precisely;

$\{1\}$,
$\{b^2\}$,
$\{a^{2r+1}, a^{-2r-1}b^2\}; (r = 0, 1, ..., n-1)$,
$\{a^{2s}, a^{-2s}\}$,
$\{a^{2s}b^2, a^{-2s}b^2\}; (s = 1, 2, ..., \frac{n-1}{2})$,
$\{a^j b^k : j \text{ even}, k = 1 \text{ or } 3\}$ and
$\{a^j b^k : j \text{ odd}, k = 1 \text{ or } 3\}$.

While when n is even, the group V_{8n} has $2n+6$ conjugacy classes namely;

$\{1\}$,
$\{b^2\}$,
$\{a^n\}$,
$\{a^n b^2\}$,
$\{a^{2r+1}, a^{-2r-1}b^2\}; (r = 0, 1, ..., n-1)$,
$\{a^{2s}, a^{-2s}\}$,
$\{a^{2s}b^2, a^{-2s}b^2\}; (s = 1, 2, ..., \frac{n}{2} - 1)$,
$\{a^{2k}b^{(-1)^k} : 0 \le k \le n-1\}$,
$\{a^{2k}b^{(-1)^{k+1}} : 0 \le k \le n-1\}$,
$\{a^{2k+1}b^{(-1)^k} : 0 \le k \le n-1\}$ and
$\{a^{2k+1}b^{(-1)^{k+1}} : 0 \le k \le n-1\}$.

Clearly, V_{8n} is a non-Abelian group of order $8n$ and its elements are of the form $a^r, a^r b, a^r b^2, a^r b^3$; where $r = 0, 1, ..., 2n-1$.

Lemma 3.1. *Let $\mathcal{G} = V_{8n} = \langle a, b : a^{2n} = 1 = b^4, aba = b^{-1}, ab^{-1}a = b\rangle$. Then we have*

(1) $ba^j = a^{-j}b^{(-1)^j}$.

(2) $b^2 a^j = a^j b^2$.

(3) $b^3 a^j = a^{-j} b^{(-1)^{j+1}}$.

(4) $Z(V_{8n}) = \begin{cases} \langle a^n, b^2\rangle; & \text{if } 2 \mid n \\ \langle b^2\rangle; & \text{if } 2 \nmid n \end{cases}$.

Proof. (1) We will prove the result by induction on j. By hypothesis, $ba = a^{-1}b^{-1}$. Assume that the result holds for all positive integers up to $j-1$, so we have, $ba^{j-1} = a^{1-j}b^{(-1)^{j-1}}$. Further, $ba^j = (ba^{j-1})a = (a^{1-j}b^{(-1)^{j-1}})a = a^{1-j}(b^{(-1)^{j-1}}a) = a^{-j}b^{(-1)^j}$, as $b^{-1}a = a^{-1}b$ and $b^{-1}a^{-1} = ab$. Thus, $ba^j = a^{-j}b^{(-1)^j}$ holds for every non negative integer j. Similarly, we can prove by induction that $ba^{-j} = a^j b^{(-1)^j}$ holds for any non negative integer j. Therefore, (1) holds for any $j \in \mathbb{Z}$.

(2) and (3) follow by using (1), repeatedly.

(4) Straightforward. ■

Next, we partition the set of elements in the group V_{8n}, n odd, as follows:

$$V_{8n} = \mathcal{X}_n^{'} \cup \mathcal{X}_n, \text{ where}$$

$\mathcal{X}_n^{'} = \{a^{-1}b^2, a^{-3}b^2, ..., a^{1-2n}b^2, a^{-2}, a^{-4}, ..., a^{1-n}, a^{-2}b^2, a^{-4}b^2, ..., a^{1-n}b^2, b^3, ab^3\}$ and $\mathcal{X}_n = V_{8n} \setminus \mathcal{X}_n^{'} = \{1, a, ..., a^n, b, a^4b, a^6b, ..., a^{2n-2}b, ab, a^5b, a^7b, ..., a^{2n-1}b\}$.

Proposition 3.2. *The stabilizer $Stab_{V_{8n}}(\mathcal{X}_n^{'})$ of the subset $\mathcal{X}_n^{'}$ in V_{8n}, n odd, is the set $\{1, b^2\}$.*

Proof. Let $g \in Stab_{V_{8n}}(\mathcal{X}_n^{'}) \cap V_{8n}$. Then, by definition $ghg^{-1} = h$, for all $h \in \mathcal{X}_n^{'}$. Also, an element $g \in V_{8n}$ will be of the form a^ib^j, where $-n < i \leq n$ and $j = 0, 1, 2, 3$. We claim that if $g \in Stab_{V_{8n}}(\mathcal{X}_n^{'})$, then $j \neq 1, 3$. Otherwise, $g = a^ib$ or a^ib^3, where $-n < i < n$. In this case, there exists an element $h = ab^3 \in \mathcal{X}_n^{'}$ satisfying $ghg^{-1} = \begin{cases} a^{2i-1}b; & \text{if } 2 \mid i \\ a^{2i-1}b^3; & \text{if } 2 \nmid i \end{cases}$. Clearly, $ghg^{-1} = h = ab^3$ if and only if $i = 1$. Furthermore, if $i \neq 1$, then $ghg^{-1} \neq h$, for $h = ab^3$ and in the case when $i = 1$, replace $h = ab^3$ with $b^3 \in \mathcal{X}_n^{'}$. Then $ghg^{-1} = a^2b \neq b^3 = h$. Therefore, an element $g \in Stab_{V_{8n}}(\mathcal{X}_n^{'})$ must be of the form a^ib^j, where $-n < i \leq n$ and $j = 0, 2$.

Next, we shall show that if $g \in Stab_{V_{8n}}(\mathcal{X}_n^{'})$, then $i = 0$. For this, set $h = ab^3 \in \mathcal{X}_n^{'}$. Then $ghg^{-1} = \begin{cases} a^{2i+1}b^3; & \text{if } 2 \mid i \\ a^{2i+1}b; & \text{if } 2 \nmid i \end{cases}$. Again, $ghg^{-1} = h = ab^3$ if and only if $i = 0$. Thus, $Stab_{V_{8n}}(\mathcal{X}_n^{'}) \subseteq \{1, b^2\}$. Further, by Lemma 3.1(4), for any subset S of V_{8n}, $\langle b^2 \rangle \subset Stab_{V_{8n}}(S)$. Hence $Stab_{V_{8n}}(\mathcal{X}_n^{'}) = \{1, b^2\}$. ■

In the forthcoming theorem, we establish a 2-distinguishing labelling for the conjugation action of V_{8n}, n odd, acting on itself.

Theorem 3.3. *If V_{8n}, n odd, acts on itself by conjugation, then the map $\phi : V_{8n} \longrightarrow \{1, 2\}$ defined by $\phi(x) = \begin{cases} 1; & \text{if } x \in \mathcal{X}_n \\ 2; & \text{otherwise} \end{cases}$, is a 2-distinguishing labelling.*

Proof. Suppose V_{8n}, n odd, acts on itself by conjugation. Define a labelling, $\phi : V_{8n} \longrightarrow \{1, 2\}$ by $\phi(x) = 1$ for all $x \in \mathcal{X}_n$ and $\phi(x) = 2$ otherwise. Note that for ϕ to be a 2-distinguishing labelling, it is enough to show that

$$\{g \in V_{8n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in V_{8n}\} = \{1, b^2\}.$$

Clearly, by Proposition 3.2, it is equivalent to prove that if $g \in V_{8n} \setminus \{1, b^2\}$, then there exists an element $h \in \mathcal{X}_n^{'}$ such that $ghg^{-1} \notin \mathcal{X}_n^{'}$, that is $ghg^{-1} \in \mathcal{X}_n$. For this, assume that $g \in V_{8n} \setminus \{1, b^2\}$. If the element $g = a^i$ or a^ib^2; $-n < i \leq n, i \neq 0$, then we can choose $h = ab^3 \in \mathcal{X}_n^{'}$, then $ghg^{-1} = \begin{cases} a^{2i+1}b^3; & \text{if } 2 \mid i \\ a^{2i+1}b; & \text{if } 2 \nmid i \end{cases}$. Clearly, $a^{2i+1}b \notin \mathcal{X}_n^{'}$, for $2 \nmid i$. Moreover, if $2 \mid i$, then $a^{2i+1}b^3 \in \mathcal{X}_n^{'}$ if and only if $i = 0$, which is not possible.

On the other hand, the element g should only be of the form $g = a^ib^j$; $-n < i \leq n$, $j = 1$ or 3. Take $h = ab^3 \in \mathcal{X}_n^{'}$, then $ghg^{-1} = \begin{cases} a^{2i-1}b; & \text{if } 2 \mid i \\ a^{2i-1}b^3; & \text{if } 2 \nmid i \end{cases}$. Again, if $2 \mid i$, then $a^{2i-1}b \notin \mathcal{X}_n^{'}$ and observe that in case $2 \nmid i$, then $a^{2i-1}b^3 \in \mathcal{X}_n^{'}$ if and only if $i = 1$ or $i = 1-n$. Note that $i = 1-n$ is not possible, as n is an odd positive integer and $2 \nmid i$. Thus, assume that $i \neq 1$, this implies that $ghg^{-1} \notin \mathcal{X}_n^{'}$. Otherwise, when $i = 1$, the element g will be either ab or ab^3. In any of these cases, for $h = b^3$, we have $ghg^{-1} = abb^3b^{-1}a^{-1} = ab^3a^{-1} = a^2b \notin \mathcal{X}_n^{'}$. This infers that, whenever $g \in V_{8n} \setminus \{1, b^2\}$, there always exists an element $h \in \mathcal{X}_n^{'}$ such that $ghg^{-1} \in \mathcal{X}_n$. Therefore, we conclude that $\{g \in V_{8n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in V_{8n}\} \subseteq \{1, b^2\}$ and hence the result. ■

Now, we turn our attention to examine the distinguishing labelling and distinguishing number for the conjugation action of the group V_{8n}, when n is even. Observe that in this case

$Z(V_{8n}) = Stab_{V_{8n}}(V_{8n}) = \{1, a^n, b^2, a^n b^2\}$.

Assume that $n \in 2\mathbb{Z}^+$ and we partition the set of elements in the group V_{8n} as follows:

$$V_{8n} = \mathcal{Y}_n^{'} \cup \mathcal{Y}_n, \text{ where}$$

$\mathcal{Y}_n^{'} = \{a^{-1}b^2, a^{-3}b^2, ..., a^{1-2n}b^2, a^{-2}, a^{-4}, ..., a^{2-n}, a^{-2}b^2, a^{-4}b^2, ..., a^{2-n}b^2, a^2b, a^3b, a^2b^3, a^3b^3\}$ and $\mathcal{Y}_n = V_{8n} \setminus \mathcal{Y}_n^{'}$. This partition will be used to define a 2-distinguishing labelling for the conjugation action of the group V_{8n}, for even n.

Proposition 3.4. *The stabilizer $Stab_{V_{8n}}(\mathcal{Y}_n^{'})$ of the subset $\mathcal{Y}_n^{'}$ is $\{1, a^n, b^2, a^n b^2\}$.*

Proof. Let $g \in Stab_{V_{8n}}(\mathcal{Y}_n^{'}) \cap V_{8n}$. By hypothesis $ghg^{-1} = h$, for all $h \in \mathcal{Y}_n^{'}$. Also, an element $g \in V_{8n}$ will be of the type a^ib^j, where $-n < i \leq n$ and $j = 0, 1, 2, 3$. We claim that if $g \in Stab_{V_{8n}}(\mathcal{Y}_n^{'})$, then $j \neq 1, 3$. Otherwise, either $g = a^ib$ or a^ib^3, where $-n < i \leq n$ and there exists an element $h = a^2b^3 \in \mathcal{Y}_n^{'}$ satisfying $ghg^{-1} = \begin{cases} a^{2i-2}b^3; & \text{if } 2 \mid i \\ a^{2i-2}b; & \text{if } 2 \nmid i \end{cases}$. Note that $ghg^{-1} = h = a^2b^3$ if and only if $2 \mid i$ and $i = 2$ or $2 - n$. Thus, if $i \neq 2$ or $2 - n$, then $ghg^{-1} \neq h$ for $h = a^2b^3$. Now, if $i = 2$ or $2 - n$, then replace $h = a^2b^3$ with $a^3b^3 \in \mathcal{X}_n^{'}$ and we have $ghg^{-1} = ab \neq a^3b^3 = h$. Therefore, an element $g \in Stab_{V_{8n}}(\mathcal{Y}_n^{'})$ will be of the form a^ib^j, where $-n < i \leq n$ and $j = 0, 2$.

Next, we further show that if $g \in Stab_{V_{8n}}(\mathcal{Y}_n^{'})$, then either $i = 0$ or $i = n$. For this, set $h = a^2b^3 \in \mathcal{Y}_n^{'}$, then $ghg^{-1} = \begin{cases} a^{2i+2}b^3; & \text{if } 2 \mid i \\ a^{2i+2}b; & \text{if } 2 \nmid i \end{cases}$. Clearly, $ghg^{-1} = h = a^2b^3$ if and only if $i = 0$ or n. Hence, an element $g \in Stab_{V_{8n}}(\mathcal{Y}_n^{'})$ should be of the form a^ib^j; where $i = 0, n$ and $j = 0, 2$, so $Stab_{V_{8n}}(\mathcal{Y}_n^{'}) \subseteq \{1, a^n, b^2, a^nb^2\}$. Further, since $Z(V_{8n}) = \{1, a^n, b^2, a^nb^2\}$, so $\{1, a^n, b^2, a^nb^2\} \subseteq Stab_{V_{8n}}(\mathcal{Y}_n^{'})$. Hence, $Stab_{V_{8n}}(\mathcal{Y}_n^{'}) = \{1, a^n, b^2, a^nb^2\}$. ∎

Theorem 3.5. *Let V_{8n}, n even, acts on itself by conjugation. Then there exists a map $\phi : V_{8n} \longrightarrow \{1, 2\}$ defined by $\phi(x) = \begin{cases} 1; & \text{if } x \in \mathcal{Y}_n \\ 2; & \text{otherwise} \end{cases}$, is a 2-distinguishing labelling.*

Proof. Suppose that the group V_{8n}, n even, acts on itself by conjugation. Define a labelling, $\phi : V_{8n} \longrightarrow \{1, 2\}$ for this action by $\phi(x) = 1$ for all $x \in \mathcal{Y}_n$ and $\phi(x) = 2$ otherwise. On imitating the same process as in Theorem 3.3, it suffices to show that

$$\{g \in V_{8n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in V_{8n}\} = \{1, a^n, b^2, a^nb^2\}.$$

Moreover, by Proposition 3.4, it is equivalent to prove that if $g \in V_{8n} \setminus \{1, a^n, b^2, a^nb^2\}$, then there exists an element $h \in \mathcal{Y}_n^{'}$ such that $ghg^{-1} \notin \mathcal{Y}_n^{'}$. For this, assume that $g \in V_{8n} \setminus \{1, a^n, b^2, a^nb^2\}$ and in case the element $g = a^i$ or a^ib^2; $-n < i \leq n$, then it is trivial to see that, i can not take the values 0 and n. Thus, set $h = a^2b^3 \in \mathcal{Y}_n^{'}$ and we have $ghg^{-1} = \begin{cases} a^{2i+2}b^3; & \text{if } 2 \mid i \\ a^{2i+2}b; & \text{if } 2 \nmid i \end{cases}$. Also, note that if $2 \nmid i$, then $a^{2i+2}b \in \mathcal{Y}_n^{'}$, if and only if $i = 0$ or n. Again, if $2 \mid i$, then $a^{2i+2}b^3 \in \mathcal{Y}_n^{'}$ if and only if $i = 0, n$. Thus, $ghg^{-1} \notin \mathcal{Y}_n^{'}$, because i cannot take the values 0 and n.

On the other hand, if $g = a^ib$ or a^ib^3; $-n < i \leq n$. In this case, take $h = a^2b^3 \in \mathcal{Y}_n^{'}$, then $ghg^{-1} = \begin{cases} a^{2i-2}b; & \text{if } 2 \nmid i \\ a^{2i-2}b^3; & \text{if } 2 \mid i \end{cases}$. Note that $a^{2i-2}b \in \mathcal{Y}_n^{'}$ if and only if $2 \nmid i$ and $i = 2$ or $2 - n$, which is not possible, as n is even. Further, if $2 \mid i$, then $a^{2i-2}b^3 \in \mathcal{Y}_n^{'}$ if and only if $i = 2$ or $-n + 2$. Hence, if $i \neq 2, -n + 2$, then $ghg^{-1} \notin \mathcal{Y}_n^{'}$. Moreover, if we choose $h = a^3b^3$ and if $i = 2$ or $-n + 2$, then $ghg^{-1} = ab \notin \mathcal{Y}_n^{'}$. Thus, we have proved that if $g \in V_{8n} \setminus \{1, a^n, b^2, a^n b^2\}$, then there exists an element $h \in \mathcal{Y}_n^{'}$ such that $ghg^{-1} \notin \mathcal{Y}_n^{'}$. Therefore, we conclude that $\{g \in V_{8n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in V_{8n}\} \subseteq \{1, a^n, b^2, a^n b^2\}$ and by using Lemma 3.1 the reverse inclusion is immediate. ■

By combining Theorem 1.3, Theorem 3.3 and Theorem 3.5, we have Theorem 3.6, which provides the distinguishing number for the conjugation action of the group V_{8n} on itself, for an arbitrary n.

Theorem 3.6. *If V_{8n} acts on itself by conjugation action, then $D_{V_{8n}}(V_{8n}) = 2$.*

4. Distinguishing Labelling for action of the group U_{6n}

Recall from [4] that $\mathcal{G} = U_{6n}$ is a non-Abelian group of order $6n$ and it is generated by two elements a and b such that $a^{2n} = 1, b^3 = 1, ba = ab^{-1}$. The group U_{6n} can be represented as follows:

$$U_{6n} = \langle a, b : a^{2n} = 1, b^3 = 1, ba = ab^{-1} \rangle.$$

However, group U_{6n} can be viewed as a group isomorphic to the semi direct product of a cyclic group of order 3 by a cyclic group of order $2n$. Clearly, the subgroup generated by the generator b is a normal subgroup of order 3 and the subgroup generated by a is a cyclic subgroup of order $2n$. In addition, the group U_{6n} has $3n$ conjugacy classes namely:
$\{a^{2r}\}$,
$\{a^{2r}b, a^{2r}b^2\}$,
$\{a^{2r+1}, a^{2r+1}b, a^{2r+1}b^2\}; (r = 0, 1, 2, ..., n-1)$.

Lemma 4.1. *Let $\mathcal{G} = U_{6n} = \langle a, b : a^{2n} = 1, b^3 = 1, ba = ab^{-1} \rangle$ and j be an arbitrary integer. Then we have*

(1) $ba^j = \begin{cases} a^j b^2; & \text{if } 2 \nmid j \\ a^j b; & \text{if } 2 \mid j \end{cases}$

(2) $b^2 a^j = \begin{cases} a^j b; & \text{if } 2 \nmid j \\ a^j b^2; & \text{if } 2 \mid j \end{cases}$

(3) $Z(U_{6n}) = \langle a^2 \rangle$.

Proof. (1) First we shall show by induction that for any non-negative integer j, $ba^j = \begin{cases} a^j b^2; & \text{if } 2 \nmid j \\ a^j b; & \text{if } 2 \mid j \end{cases}$. Clearly, by group relation $ba = a^{-1}b$ and suppose this holds for all positive integers up to $m-1$. Now, consider $ba^m = (ba^{m-1})a = \begin{cases} (a^{m-1}b^2)a; & \text{if } 2 \nmid m-1 \\ (a^{m-1}b)a; & \text{if } 2 \mid m-1 \end{cases} = \begin{cases} a^m b^2; & \text{if } 2 \nmid m \\ a^m b; & \text{if } 2 \mid m \end{cases}$. Hence, the result holds for all non negative integers. Moreover, by imitating the same process we can similarly prove by induction that $ba^j = \begin{cases} a^j b^2; & \text{if } 2 \nmid j \\ a^j b; & \text{if } 2 \mid j \end{cases}$ holds for any negative integer j and this completes the proof.

Results (2) and (3) are straightforward. ■

Moreover, we can write U_{6n} as a disjoint union of two subsets $\mathcal{X}_n^{'}$ and $\mathcal{X}_n$, where $\mathcal{X}_n^{'} = \{a^i b^2 : i = 1, 2, ..., 2n-1\}$ and $\mathcal{X}_n = Q_{4n} \setminus \mathcal{X}_n^{'}$.

Proposition 4.2. *The stabilizer of the subset $\mathcal{X}_n^{'} = \{a^i b^2 : i = 1, 2, ..., 2n-1\}$ is $Stab_{U_{6n}}(\mathcal{X}_n^{'}) = \langle a^2 \rangle$.*

Proof. By Lemma 4.1(3), it is sufficient to show that $Stab_{U_{6n}}(\mathcal{X}_n^{'}) \subseteq \langle a^2 \rangle$. For this, let $g \in Stab_{U_{6n}}(\mathcal{X}_n^{'}) \cap U_{6n}$. Then by hypothesis, $ghg^{-1} = h$, for all $h \in \mathcal{X}_n^{'}$. Also, an element in U_{6n} will be of the form $a^i b^j$, where $0 \leq i \leq 2n-1$ and $j = 0, 1, 2$, and so is g. We claim that if $g = a^i b^j \in Stab_{U_{6n}}(\mathcal{X}_n^{'})$, then $j = 0$. If possible, let $j \neq 0$, then the element g will be of the form $a^i b$ or $a^i b^2$. Now, in case $g = a^i b$ and $2 \nmid i$, then choose $h = b^2$, else in the other case $2 \mid i$, choose $h = ab^2$. Thus by using Lemma 4.1, it follows that $ghg^{-1} \neq h$. On the other hand, when $g = a^i b^2$ and $2 \nmid i$, choose $h = ab^2$ or in case $2 \mid i$, take $h = b^2$. Further, by Lemma 4.1, we observe that in either of these cases $ghg^{-1} \neq h$. Thus, the element g must be of the form a^i, where $0 \leq i \leq 2n-1$. Further, for $h = b^2 \in \mathcal{X}_n^{'}$, we have $ghg^{-1} = h$ if and only if $2 \mid i$ and hence, $Stab_{U_{6n}}(\mathcal{X}_n^{'}) \subseteq \langle a^2 \rangle$. ■

In the upcoming theorem we obtain a 2-distinguishing labelling for the conjugation action of U_{6n} on itself.

Theorem 4.3. *The map* $\phi : U_{6n} \longrightarrow \{1, 2\}$ *defined by* $\phi(x) = \begin{cases} 1\,; & \text{if } x \in \mathcal{X}_n \\ 2\,; & \text{otherwise} \end{cases}$*, is a* 2*-distinguishing labelling of* U_{6n}.

Proof. Define a map, $\phi : U_{6n} \longrightarrow \{1, 2\}$, by $\phi(x) = 1$, for all $x \in \mathcal{X}_n$ and $\phi(x) = 2$, otherwise. In order to prove that ϕ is a 2-distinguishing labelling, it is sufficient to show that $\{g \in U_{6n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in U_{6n}\} = \langle a^2 \rangle$. Now, in light of Proposition 4.2, it is equivalent to prove that for every $g \in U_{6n} \setminus \langle a^2 \rangle$, there exists an element $h \in \mathcal{X}_n^{'}$ such that $ghg^{-1} \in \mathcal{X}_n$. Let $g \in U_{6n} \setminus \langle a^2 \rangle$. Then, g will be of the form $g = a^i;\ i \notin \{0, 2, ..., 2n-2\}$. Fix $h = ab^2 \in \mathcal{X}_n^{'}$ and we have $ghg^{-1} = ab$ and certainly, $ab \in \mathcal{X}_n$. Otherwise, the element g will be of the form $g = a^i b$ or $g = a^i b$; $0 \leq i \leq 2n-1$. In either case, if $2 \nmid i$, set $h = b^2 \in \mathcal{X}_n^{'}$, then by using Lemma 4.1 (2), we get $ghg^{-1} = a^i b^2 a^{-i} = b$, as $2 \nmid i$. Clearly, $ghg^{-1} = b \notin \mathcal{X}_n$. Further, if $2 \mid i$, then for $g = a^i b$ choose $h = ab^2$, we have $ghg^{-1} = (a^i b)ab^2(b^{-1}a^{-i}) = a^i baba^{-i} = a^{i+1}b^{-1}ba^{-i} = a$. Furthermore, for $g = a^i b^2$, take $h = ab^2 \in \mathcal{X}_n^{'}$ and we get $ghg^{-1} = (a^i b^2)ab^2(b^{-1}a^{-i}) = a^i b^2 a^{-i+1} = ab \notin \mathcal{X}_n^{'}$. Thus $\{g \in U_{6n} : \phi(ghg^{-1}) = \phi(h), \text{ for all } h \in U_{6n}\} \subseteq \langle a^2 \rangle$ and the result follows by using Lemma 4.1 (3). ■

The following theorem is a direct consequence of Theorem 1.3 and Theorem 4.3.

Theorem 4.4. *If* U_{6n} *acts on itself by conjugation, then* $\mathcal{D}_{U_{6n}}(U_{6n}) = 2$.

5. Distinguishing Labelling for semi-dihedral group SD_{8n} action

The semi-dihedral group, SD_{8n} [3] is a non-Abelian group of order $8n$. For $n \geq 2$, this group can be presented as follows:

$$SD_{8n} = \langle a, b : a^{4n} = b^2 = 1, bab = a^{2n-1} \rangle.$$

Clearly, the elements of the semi-dihedral group are of the form a^r or $ba^r; r = 0, 1, ..., 4n-1$. As observed in the case of V_{8n}, the group SD_{8n} has $2n+3$ or $2n+6$ conjugacy classes, when n is even or odd respectively. We begin with the following definition

Definition 5.1. *[3] Define* $C^{even} := C_1 \cup C_2^{even} \cup C_3^{even}$ *and* $C^{odd} := C_1 \cup C_2^{odd} \cup C_3^{odd}$*, where* $C_1 := \{0, 2, 4, ..., 2n\}, C_2^{even} := \{1, 3, ..., n-1\}, C_3^{even} := \{2n+1, 2n+3, 2n+5, ..., 3n-1\}$ *and* $C_2^{odd} := \{1, 3, 5, ..., n\}, C_3^{odd} := \{2n+1, 2n+3, 2n+5, ..., 3n\}$*. Also, define* $C_{even}^{\dagger} := C_1 \setminus \{0, 2n\}, C_{odd}^{\dagger} := C_2^{even} \cup C_3^{even}, C_{2,3}^{odd} := C_2^{odd} \cup C_3^{odd}$ *and* $C_*^{even} := C^{even} \setminus \{0, 2n\}, C_*^{odd} := C^{odd} \setminus \{0, n, 2n, 3n\}$.

The next proposition provides us with the conjugacy classes of SD_{8n}.

Proposition 5.2. *[3] The conjugacy classes of* SD_{8n}*,* $n \geq 2$*, are as follows:*

- *If* n *is even, then there are* $2n+3$ *conjugacy classes. Precisely,*
  - 2 *classes of size one being* $[1] = \{1\}$ *and* $[a^{2n}] = a^{2n}$*,*

- $(2n-1)$ *classes of size two being* $[a^r]=\{a^r, a^{(2n-1)r}\}$*, where* $r\in C_*^{even}$ *and*
- 2 *classes of size* $2n$ *being* $[b]=\{ba^{2t}: t=0,1,...,2n-1\}$ *and* $[ba]=\{ba^{2t+1}: t=0,1,...,2n-1\}$.

- *If* n *is odd, then there are* $2n+6$ *conjugacy classes. Precisely,*
  - 4 *classes of size one being* $[1]=\{1\}, [a^n]=\{a^n\}, [a^{2n}]=\{a^{2n}\}$ *and* $[a^{3n}]=\{a^{3n}\}$,
  - $(2n-2)$ *classes of size two being* $[a^r]=\{a^r, a^{(2n-1)r}\}$*, where* $r\in C_*^{odd}$ *and*
  - 4 *classes of size* n *being* $[b]=\{ba^{4t}: t=0,1,...,n-1\}, [ba]=\{ba^{4t+1}: t=0,1,...,n-1\}, [ba^2]=\{ba^{4t+2}: t=0,1,...,n-1\}$*, and* $[ba^3]=\{ba^{4t+3}: t=0,1,...,n-1\}$.

Lemma 5.3. *[3] Let* $\mathcal{G}=SD_{8n}=\langle a,b: a^{4n}=1=b^2, bab=a^{2n-1}\rangle$*. Then*

(1) $a^kb=ba^{(2n-1)k}$.

(2) $Z(V_{8n})=\begin{cases}\langle a^n\rangle; & \text{if } 2\nmid n\\ \langle a^{2n}\rangle; & \text{if } 2\mid n\end{cases}$.

For n **even**, we can partition the set of elements of SD_{8n} as a disjoint union of two subsets $\mathcal{X}_n^{'}$ and $\mathcal{X}_n$ of SD_{8n}, where $\mathcal{X}_n^{'}=\{a^{r(2n-1)}: r=2,4,6,...,2n-2,1,3,5,...,n-1,2n+1,2n+3,...,3n-1\}\cup\{ba^2,ba^3\}$ and $\mathcal{X}_n=SD_{8n}\setminus\mathcal{X}_n^{'}$. This partition will be used to define a 2-distinguishing labelling for the conjugation action of the group SD_{8n}, for an even n.

Proposition 5.4. *The stabilizer* $Stab_{SD_{8n}}(\mathcal{X}_n^{'})$ *of the subset* $\mathcal{X}_n^{'}$ *is* $\{1,a^{2n}\}$.

Proof. Let $g\in Stab_{SD_{8n}}(\mathcal{X}_n^{'})\cap SD_{8n}$. Then by definition $ghg^{-1}=h$, for all $h\in\mathcal{X}_n^{'}$. Also, an element g in the group SD_{8n} will be of the form a^ib^j, where $0\le i\le 4n-1$ and $j=0,1$. We claim that if $g\in Stab_{SD_{8n}}(\mathcal{X}_n^{'})$, then $j\neq 1$. If not, then $g=a^ib$ and we have $h=a^{2n-1}\in\mathcal{X}_n^{'}$ such that $ghg^{-1}=a^iba^{2n-1}b^{-1}a^{-i}=a^i(ba^{2n-1}b)a^{-i}=a^iaa^{-i}=a$. Now, $ghg^{-1}=h=a^{2n-1}$ if and only if $n=1$, which is not possible, as $n\ge 2$. Therefore, an element $g\in Stab_{SD_{8n}}(\mathcal{X}_n^{'})$ should be of the form a^i, for some $0\le i\le 4n-1$. Next, we show that either $i=0$ or $i=2n$. For this, set $h=ba^2\in\mathcal{X}_n^{'}$, then $ghg^{-1}=a^iba^2a^{-i}=a^iba^{2-i}=ba^{(2n-1)i+2-i}=ba^{2i(n-1)+2}$. Again, $ghg^{-1}=h=ba^2$ if and only if $2i(n-1)\equiv 0(mod\,4n)$ if and only if $2i\equiv 0(mod\,4n)$. Therefore, we have $i=0$ or $2n$. Thus, $Stab_{SD_{8n}}(\mathcal{X}_n^{'})\subseteq\{1,a^{2n}\}$. Obviously, $\{1,a^{2n}\}\subseteq Stab_{SD_{8n}}(\mathcal{X}_n^{'})$, as the center $Z(SD_{8n})=\{1,a^{2n}\}$. Hence $Stab_{SD_{8n}}(\mathcal{X}_n^{'})=\{1,a^{2n}\}$. ■

Theorem 5.5. *Let* SD_{8n}*, n even, acts on itself by conjugation. Then the map* $\phi: SD_{8n}\longrightarrow\{1,2\}$ *defined by* $\phi(x)=\begin{cases}1; & \text{if } x\in\mathcal{X}_n\\ 2; & \text{otherwise}\end{cases}$*, is a 2-distinguishing labelling of* SD_{8n}.

Proof. Define a labelling, $\phi: V_{8n}\longrightarrow\{1,2\}$ of this action by $\phi(x)=1$, for all $x\in\mathcal{X}_n$ and $\phi(x)=2$, otherwise. From the Definition 1.1, for ϕ to be a 2-distinguishing labelling, it suffices to show that $\{g\in SD_{8n}: \phi(ghg^{-1})=\phi(h),\ \text{for all } h\in SD_{8n}\}=\{1,a^{2n}\}$. Now, by Proposition 5.4, it is equivalent to prove that if $g\in SD_{8n}\setminus\{1,a^{2n}\}$, then there exists an element $h\in\mathcal{X}_n^{'}$ such that $ghg^{-1}\notin\mathcal{X}_n^{'}$. For this, let $g\in SD_{8n}\setminus\{1,a^{2n}\}$ and $g=a^i; i\in\{0,1,2,...,4n-1\}\setminus\{0,2n\}$. Then choose $h=ab^2\in\mathcal{X}_n^{'}$, so that $ghg^{-1}=a^iba^2a^{-i}=ba^{2i(n-1)+2}$. Now, $ba^{2i(n-1)+2}\in\mathcal{X}_n^{'}$, if and only if $a^{2i(n-1)+2}=a^2$ or a^3 if and only if $2i(n-1)+2\equiv 2$ or 3 $(mod\,4n)$. Clearly, $2i(n-1)+2\equiv 3(mod\,4n)$, is not possible, as n is even. Also, $2i(n-1)+2\equiv 2(mod\,4n)$ holds if and only if $2i(n-1)\equiv 0(mod\,4n)$ if and only if $i=0,2n$, which is not possible. On the other hand, if $g=ba^i$; $0\le i\le 4n-1$, then in this case, fix $h=a^{2n-1}\in\mathcal{X}_n^{'}$, which leads to $ghg^{-1}=a^iba^{2n-1}b^{-1}a^{-i}=a^iaa^{-i}=a$ and certainly $ghg^{-1}=a\notin\mathcal{X}_n^{'}$. Thus $\{g\in SD_{8n}: \phi(ghg^{-1})=\phi(h),\ \text{for all } h\in SD_{8n}\}\subseteq\{1,a^{2n}\}$ and the reverse inclusion follows by using Lemma 5.3 (2). ■

Next, we shall examine the distinguishing labelling for the conjugation action of the semi-dihedral group, when n is an odd positive integer.

For n odd, we can partition the group SD_{8n} as follows:

$$SD_{8n} = \mathcal{Y}_n^{'} \cup \mathcal{Y}_n,$$

where $\mathcal{Y}_n^{'} = \{a^{r(2n-1)} : r = 2, 4, 6, ..., 2n-2, 1, 3, 5, ..., n-2, 2n+1, 2n+3, ..., 3n-2\} \cup \{ba^4, ba^5, ba^6, ba^7\}$ and $\mathcal{Y}_n = SD_{8n} \setminus \mathcal{Y}_n^{'}$. Also, we use this partition to define a 2-distinguishing labelling for the conjugation action of the group SD_{8n} acting on itself.

Proposition 5.6. *The stabilizer $Stab_{SD_{8n}}(\mathcal{Y}_n^{'})$ of the subset $\mathcal{Y}_n^{'}$ is $\{1, a^n, a^{2n}, a^{3n}\}$.*

Proof. Let $g \in Stab_{SD_{8n}}(\mathcal{Y}_n^{'}) \cap SD_{8n}$. Then by definition $ghg^{-1} = h$, for all $h \in \mathcal{Y}_n^{'}$. Also, an element g in SD_{8n} will be of the form a^ib^j, where $0 \leq i \leq 4n-1$ and $j = 0, 1$. We claim that if $g \in Stab_{SD_{8n}}(\mathcal{Y}_n^{'})$, then $j \neq 1$. If not, then $g = a^ib$ and we have $h = a^{2n-1} \in \mathcal{Y}_n^{'}$ such that $ghg^{-1} = a^iba^{2n-1}b^{-1}a^{-i} = a^i(ba^{2n-1}b)a^{-i} = a^iaa^{-i} = a$. Now, $ghg^{-1} = h$ if and only if $a = a^{2n-1}$ if and only if $n = 1$, which is not possible, as $n \geq 2$. This infers that an element $g \in Stab_{SD_{8n}}(\mathcal{X}_n^{'})$ will be of the form a^i, for some $0 \leq i \leq 4n-1$. Next, we show that $i \in \{0, n, 2n, 3n\}$. For this, choose $h = ba^4 \in \mathcal{Y}_n^{'}$ and we have $ghg^{-1} = a^iba^4a^{-i} = a^iba^{4-i} = ba^{(2n-1)i+4-i} = ba^{2i(n-1)+4}$. Again, $ghg^{-1} = h = ba^4$ if and only if $2i(n-1) \equiv 0(mod\, 4n)$ if and only if $i \equiv 0(mod\, n)$, as n is odd. This implies that $i = 0$ or n or $2n$ or $3n$. Thus $Stab_{SD_{8n}}(\mathcal{Y}_n^{'}) \subseteq \{1, a^n, a^{2n}, a^{3n}\}$. Furthermore, $\{1, a^n, a^{2n}, a^{3n}\} \subseteq Stab_{SD_{8n}}(\mathcal{Y}_n^{'})$, since $Z(SD_{8n}) = \{1, a^n, a^{2n}, a^{3n}\}$. Hence $Stab_{SD_{8n}}(\mathcal{Y}_n^{'}) = \{1, a^n, a^{2n}, a^{3n}\}$. ■

Theorem 5.7. *If SD_{8n}, n odd, acts on itself by conjugation, then the map $\phi : SD_{8n} \longrightarrow \{1, 2\}$ defined by* $\phi(x) = \begin{cases} 1\,; & \text{if } x \in \mathcal{Y}_n \\ 2\,; & \text{otherwise} \end{cases}$*, is a 2-distinguishing labelling of SD_{8n}.*

Proof. Suppose that the group SD_{8n}, n odd acts on itself by conjugation. Define a labelling, $\phi : V_{8n} \longrightarrow \{1, 2\}$ of this action by $\phi(x) = 1$, for all $x \in \mathcal{Y}_n$ and $\phi(x) = 2$, otherwise. Again, for ϕ to be a 2-distinguishing labelling, it is required to show that $\{g \in SD_{8n} : \phi(ghg^{-1}) = \phi(h),\ \text{for all } h \in SD_{8n}\} = \{1, a^n, a^{2n}, a^{3n}\}$. Now, by the definition of ϕ and Proposition 5.6, it is equivalent to prove that for $g \in SD_{8n} \setminus \{1, a^n, a^{2n}, a^{3n}\}$, there exists an element $h \in \mathcal{Y}_n^{'}$ such that $ghg^{-1} \notin \mathcal{Y}_n^{'}$. For this, assume that $g \in SD_{8n} \setminus \{1, a^n, a^{2n}, a^{3n}\}$ and $g = a^i; i \in \{0, 1, 2, ..., 4n-1\} \setminus \{0, n, 2n, 3n\}$. One can choose $h = ab^4 \in \mathcal{Y}_n^{'}$ and we get $ghg^{-1} = a^iba^4a^{-i} = ba^{2i(n-1)+4}$. Now, $ba^{2i(n-1)+4} \in \mathcal{Y}_n^{'}$, if and only if $2i(n-1)+4 \equiv 4$ or 5 or 6 or 7 $(mod\, 4n)$. Clearly, $2i(n-1)+4 \equiv 5, 6, 7(mod\, 4n)$, is not possible, as n is an odd positive integer. Also, $2i(n-1)+4 \equiv 4(mod\, 4n)$ holds if and only if $2i(n-1) \equiv 0(mod\, 4n)$ if and only if $i = 0, n, 2n, 3n$, which is not possible. On the other hand, if $g = ba^i$; $0 \leq i \leq 4n-1$, then in this case, we fix $h = a^{2n-1} \in \mathcal{Y}_n^{'}$, then $ghg^{-1} = a^iba^{2n-1}b^{-1}a^{-i} = a^iaa^{-i} = a$ and certainly $ghg^{-1} = a \notin \mathcal{Y}_n^{'}$. Thus $\{g \in SD_{8n} : \phi(ghg^{-1}) = \phi(h),\ \text{for all } h \in SD_{8n}\} \subseteq \{1, a^n, a^{2n}, a^{3n}\}$ and by using Lemma 5.3 (2) the reverse inclusion follows immediately. ■

Finally in light of Theorem 1.3, Theorem 5.5 and Theorem 5.7, we have the following theorem which provides the distinguishing number for the conjugation action of SD_{8n} on itself, for arbitrary n.

Theorem 5.8. *If SD_{8n} acts on itself by conjugation action, then $\mathcal{D}_{SD_{8n}}(SD_{8n}) = 2$.*

6. Some Characterizations

Theorem 2.5 in [6], along with the conclusions from each section discussed so far, arise a natural question that

Whether there exists a non-Abelian group acting via conjugation action on itself with the distinguishing number other than 2?

The main purpose of this section is to find the answer of the aforementioned question and surprisingly, we answer it in negative at the end of this section. Before reaching a conclusion, we need some characterizations for the existence of a 2-distinguishing labelling, in general, for the action of a group $\mathcal{G}$ on a $\mathcal{G}$-set $\mathcal{X}$, not necessarily the conjugation action of $\mathcal{G}$ on itself. First, we give some basic definitions:

Definition 6.1. *Let $\mathcal{G}$ be a group acting on a set $\mathcal{X}$. A subset $\mathcal{A}$ of the $\mathcal{G}$-set $\mathcal{X}$ is called $\mathcal{G}$-invariant, if the subset $\{ga : g \in \mathcal{G}, a \in \mathcal{A}\} \subseteq \mathcal{A}$, under the action of $\mathcal{G}$ restricted to $\mathcal{A}$.*

It is easy to see that the orbits of a $\mathcal{G}$-set $\mathcal{X}$ are $\mathcal{G}$-invariant subsets of $\mathcal{X}$. Moreover, if $\mathcal{A}$ is a $\mathcal{G}$-invariant subset of a $\mathcal{G}$-set $\mathcal{X}$, then it is always a union of orbits of a $\mathcal{G}$-set $\mathcal{X}$.

Definition 6.2. *Let $\mathcal{G}$ be a group acting on a set $\mathcal{X}$ and $g \in \mathcal{G}$. A subset $\mathcal{A}$ of the $\mathcal{G}$-set $\mathcal{X}$ is called g-invariant, if the subset $\{ga : g \in \mathcal{G}, a \in \mathcal{A}\} \subseteq \mathcal{A}$, under the action of $\mathcal{G}$ restricted to $\mathcal{A}$.*

Theorem 6.3. *Let $\mathcal{G}$ be a group acting on a set $\mathcal{X}$. If the $\mathcal{G}$-set $\mathcal{X}$ can be partitioned into a disjoint union of two subsets (say) $\mathcal{X}_1$ and $\mathcal{X}_2$, which are not g-invariant, for every $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, then there exists a 2-distinguishing labelling for the action of $\mathcal{G}$ on the set $\mathcal{X}$.*

Proof. Assume that $\mathcal{X}$ can be partitioned as a disjoint union of two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ such that for each $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, the subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ are not invariant under the action of g. Define a labelling $\phi : \mathcal{X} \longrightarrow \{1, 2\}$ by $\phi(x) = 1$, if $x \in \mathcal{X}_1$ and $\phi(x) = 2$, otherwise. Claim that the labelling $\phi : \mathcal{X} \longrightarrow \{1, 2\}$, is a 2-distinguishing labelling. For this, it suffices to prove that $\{g \in \mathcal{G} : \phi \circ g\ (x) = \phi(x) \text{ for all } x \in \mathcal{X}\} \subseteq Stab_{\mathcal{G}}(\mathcal{X})$. Equivalently, we will prove that if $g \in \mathcal{G} \setminus Stab_{\mathcal{G}}(\mathcal{X})$, then the element g can not be a member of the set $\{g \in \mathcal{G} : \phi \circ g\ (x) = \phi(x), \text{ for all } x \in \mathcal{X}\}$. If not, then clearly by the definition of ϕ and the fact that $\mathcal{X}$ is a disjoint union of $\mathcal{X}_1$ and $\mathcal{X}_2$, we have $g.x \in \mathcal{X}_1$ or $\mathcal{X}_2$ for $x \in \mathcal{X}_1$ or $\mathcal{X}_2$ respectively, for some $g \in \mathcal{G} \setminus Stab_{\mathcal{G}}(\mathcal{X})$. Therefore, for some $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, the subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ are g-invariant, which is a contradiction to the given hypothesis. Thus, $\{g \in \mathcal{G} : \phi \circ g\ (x) = \phi(x) \text{ for all } x \in \mathcal{X}\} \subseteq Stab_{\mathcal{G}}(\mathcal{X})$ and hence ϕ is a 2-distinguishing labelling. ■

The upcoming theorem provides a characterization for the existence of a 2-distinguishing labelling for a group acting on a set. Also, it is observed that the conclusion of the above theorem still remains true if at least one of the partitioning components is not invariant under the action of each element of the group.

Theorem 6.4. *Let $\mathcal{X}$ be a $\mathcal{G}$-set. Then the following statements are equivalent:*

(1) There is a 2-distinguishing labelling for the action of the group $\mathcal{G}$ on the set $\mathcal{X}$.

(2) The $\mathcal{G}$-set $\mathcal{X}$ can be partitioned as a disjoint union of two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ such that for every $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, at least one of them is not g-invariant.

(3) The $\mathcal{G}$-set $\mathcal{X}$ can be partitioned as a disjoint union of two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ that are not g-invariant, for every $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$.

Proof. $(1) \Rightarrow (2)$: Assume that there exists a 2-distinguishing labelling ϕ, for the action of the group $\mathcal{G}$ on the set $\mathcal{X}$. Partition the $\mathcal{G}$-set $\mathcal{X}$ as the disjoint union of two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$, where $\mathcal{X}_i = \{x \in \mathcal{X} : \phi(x) = i\}$, for $i = 1, 2$. Next, we shall show that for every $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, the set $\mathcal{X}_1$ is not g-invariant. On the contrary, let us assume that $\mathcal{X}_1$ is $Stab_{\mathcal{G}}(\mathcal{X})^c$-invariant. Therefore, by definition, for $g \notin Stab_{\mathcal{G}}(\mathcal{X})$ and $x \in \mathcal{X}_1$, we have $gx \in \mathcal{X}_1$. Then clearly, if $x \in \mathcal{X}_2$ we have $gx \in \mathcal{X}_2$. Otherwise, for some $x \in \mathcal{X}_2$, we have $y = gx \in \mathcal{X}_1$. Note that since $Stab_{\mathcal{G}}(\mathcal{X})$ is a group, so if $g \notin Stab_{\mathcal{G}}(\mathcal{X})$, then $g^{-1} \notin Stab_{\mathcal{G}}(\mathcal{X})$. Consequently, we have $y \in \mathcal{X}_1$, while $g^{-1}(y) = g^{-1}gx = x \in \mathcal{X}_2$, which is not possible. Thus, there is an element g outside the set $Stab_{\mathcal{G}}(\mathcal{X})$ that preserves the labelling, which is a contradiction, as ϕ is a 2-distinguishing labelling.

$(2) \Rightarrow (3)$: Suppose that the $\mathcal{G}$-set $\mathcal{X}$ can be partitioned into a disjoint union of two subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ of $\mathcal{X}$ such that for every $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$, at least one of the $\mathcal{X}_i$ is not g-invariant. Without loss of generality, we can assume that $\mathcal{X}_1$ is not g-invariant. Therefore, there exists an element $x \in \mathcal{X}_1$ such that for some $g \notin Stab_{\mathcal{G}}(\mathcal{X})$, we have $x \neq y = gx \notin \mathcal{X}_1$. Thus $y = gx \in \mathcal{X}_2$. Note that $g \notin Stab_{\mathcal{G}}(\mathcal{X})$ if and only if $g^{-1} \notin Stab_{\mathcal{G}}(\mathcal{X})$. Clearly, there is $y \in \mathcal{X}_2$ and $g^{-1} \notin Stab_{\mathcal{G}}(\mathcal{X})$ satisfying $g^{-1}y = x \in \mathcal{X}_1$. Thus $\mathcal{X}_2$ is not g-invariant, for all $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$. Similarly, for all $g \in Stab_{\mathcal{G}}(\mathcal{X})^c$ if $\mathcal{X}_2$ is not g-invariant, then so is $\mathcal{X}_1$. Hence, the $\mathcal{G}$-set $\mathcal{X}$ can be partitioned as a union of two disjoint subsets $\mathcal{X}_1$ and $\mathcal{X}_2$ which are not g-invariant.

$(3) \Rightarrow (1)$: Follows from Theorem 6.3. ■

An immediate consequence of the above theorem is

Corollary 6.5. *Let $\mathcal{G}$ be a group acting on a set $\mathcal{X}$ and $\phi : \mathcal{X} \longrightarrow \{1, 2\}$ be a 2-distinguishing labelling for the action of $\mathcal{G}$ on the set $\mathcal{X}$. Then the subset $\mathcal{X}_i = \{x \in \mathcal{X} : \phi(x) = i\}$, $i = 1, 2$, cannot be a union of orbits.*

Also, we have

Corollary 6.6. *Let $\mathcal{X} = P_1 \sqcup P_2$ be a partition of a $\mathcal{G}$-set $\mathcal{X}$. If the partitioning subsets P_1 and P_2 are a union of orbits of $\mathcal{G}$, then the action of $\mathcal{G}$ on $\mathcal{X}$ can not be 2-distinguishable.*

In the forthcoming results, we completely characterize the distinguishing number for conjugation action. In fact, we find a connection of the distinguishing number for the above specified group action depending on the fact whether the group is an Abelian group or not.

Theorem 6.7. *If a group $\mathcal{G}$ acts on itself by conjugation, then $\mathcal{D}_{\mathcal{G}}(\mathcal{G}) = 1$ if and only if $\mathcal{G}$ is an Abelian group.*

Proof. Note that $\mathcal{G}$ is Abelian if and only if $\mathcal{G} = Z(\mathcal{G}) = \{g : ghg^{-1} = h, \text{ for all } h \in \mathcal{G}\}$ if and only if under the conjugation action, the group $\mathcal{G}$ acts on itself by fixing each of its elements. Thus, the result follows by using Proposition 2.1 [6]. ■

Finally, when a group $\mathcal{G}$ acts on itself by conjugation, we have

Theorem 6.8. *A group $\mathcal{G}$ is non-Abelian if and only if $\mathcal{D}_{\mathcal{G}}(\mathcal{G}) = 2$.*

Proof. Clearly, Theorem 6.7 guarantees that if the distinguishing number for the conjugation action of a group $\mathcal{G}$ on itself is 2, that is, $\mathcal{D}_{\mathcal{G}}(\mathcal{G}) = 2$, then the group $\mathcal{G}$ is non-Abelian. This completes the sufficient part of the present theorem.

For the necessary part, assume that a non-Abelian group $\mathcal{G}$ acts on itself by conjugation. Then, in view of Theorem 1.3, it is sufficient to prove that there exists a 2-distinguishing labelling for the action of the group $\mathcal{G}$. However, using Theorem 6.4, it is enough to prove that the non-Abelian group $\mathcal{G}$ can be partitioned as the disjoint union of two subsets $\mathcal{X}$ and $\mathcal{Y}$ such that for every $g \notin Z(\mathcal{G})$, at least one of the partitioning subsets is not g-invariant. Next, we construct such a partition of $\mathcal{G}$ as follows:

Let $\mathcal{X}$ be the set constructed by taking exactly one element from each conjugacy class and $\mathcal{Y} = \mathcal{G} \setminus \mathcal{X}$. Clearly, $\mathcal{G} = \mathcal{X} \sqcup \mathcal{Y}$. Note that an element of a group belongs to its center if and only if its conjugacy class contains exactly one element. Moreover, a group is non-Abelian if and only if it has a conjugacy class containing at least two elements. Thus, we conclude that the partitioning components $\mathcal{X}$ and $\mathcal{Y}$ are non empty and satisfy the property that $Z(\mathcal{G}) \subsetneq \mathcal{X}$ and $\mathcal{Y} \neq \mathcal{G} \setminus Z(\mathcal{G})$, as $\mathcal{G}$ is non-Abelian.

We claim that the partition $\mathcal{X} \sqcup \mathcal{Y}$ of $\mathcal{G}$ satisfies the condition: for every $g \notin Z(\mathcal{G})$, at least one of the partitioning subset $\mathcal{X}$ or $\mathcal{Y}$ is not g-invariant. If not, then for every $g \notin Z(\mathcal{G})$, both the partitioning subsets $\mathcal{X}$ and $\mathcal{Y}$ are g-invariant. In particular, for every $g \notin Z(\mathcal{G})$ and $x \in \mathcal{X}$, we have $gxg^{-1} \in \mathcal{X}$. Now, we shall show that $\mathcal{X} = \mathcal{G}$. For this, let $g \in \mathcal{G}$. Clearly, if $g \in Z(\mathcal{G})$, then $g \in \mathcal{X}$. However, if $g \in \mathcal{G} \setminus (Z(\mathcal{G}) \cup \mathcal{X})$, then by the construction of $\mathcal{X}$, we can find an element $x \in \mathcal{X}$ in the conjugacy class of g, as the set $\mathcal{X}$ has a non-trivial intersection with each conjugacy class of $\mathcal{G}$. Therefore, there exists an element $k \in \mathcal{G} \setminus Z(\mathcal{G})$ such that $g = k^{-1}xk$ with $k \neq x$. Since, $Z(\mathcal{G})$ is a subgroup of $\mathcal{G}$, so we conclude that $k^{-1} \notin Z(\mathcal{G})$. Finally, one can choose $h = k^{-1} \notin Z(\mathcal{G})$ and $x = kgk^{-1} \in \mathcal{X}$, so that $hxh^{-1} = k^{-1}(kgk^{-1})k = g \in \mathcal{X}$, as the partitioning subset $\mathcal{X}$ is g-invariant. Thus, we conclude that $\mathcal{X} = \mathcal{G}$, which is a contradiction, as $\mathcal{Y} \neq \phi$. On the other hand, if the partitioning subset $\mathcal{Y}$ is g-invariant for every $g \notin Z(\mathcal{G})$, then in a similar way, this assumption leads to a contradiction that $\mathcal{Y} = \mathcal{G} \setminus Z(\mathcal{G})$. Hence, for every $g \notin Z(\mathcal{G})$, at least one of the partitioning subsets $\mathcal{X}$ or $\mathcal{Y}$ is not g-invariant. This completes the proof. ■

7. Acknowledgement

The second author gratefully acknowledges the financial assistance of CSIR-HRDG for JRF.

References

[1] M.O. ALBERTSON AND K.L. COLLINS, Symmetry Breaking in Graphs, *Elect. J. Comb.,* **3**(1996), #R18.

[2] M.R. DARAFSHEH AND N.S. POURSALAVATI, On the existence of the orthogonal basis of the symmetry classes of tensors associated with certain groups, *SUT J. Math.,* **37**(2001), 1–17.

[3] M. HORMOZI AND K. RODTES, Symmetry classes of tensors associated with the semi-dihedral groups SD_{8n}, *Collo. Math.,* **131**(2013), 59-67.

[4] G. JAMES AND M. LIEBECK, *Representations and Characters of Groups,* (Cambridge University Press, New York, 2001).

[5] F. RUBIN, Problem 729, *J. Recreat. Math.,* **11**(1979), pp.128.

[6] J. TYMOCZKO, Distinguishing Number for Graph and Groups, *Elect. J. Comb.,* **11**(2004), #R63.

MALAYA JOURNAL OF MATEMATIK
Malaya J. Mat. **11(01)**(2023), 107–116.
http://doi.org/10.26637/mjm1101/008

The area of the Bézier polygonal region of the BézierCurve and derivatives in E^3

ŞEYDA KILIÇOĞLU*[1] AND SÜLEYMAN ŞENYURT[2]

[1] *Faculty of Science, Department of Mathematics, Başkent University, Ankara, Turkey.*
[2] *Faculty of Science, Department of Mathematics, Ordu University, Ordu, Turkey.*

*R*eceived 18 June 2022; *A*ccepted 11 October 2022

Abstract. In the paper, we have first defined the area of the Bézier polygonal region which contains the n^{th} $order$ Bézier Curve and its first, second and third derivatives based on the control points of n^{th} $order$ Bézier curve in E^3. Further, the area of the Bézier polygonal region containing the 5^{th} $order$ Bézier curve and the corresponding derivatives is examined based on the control points of $5^{th}order$ Bézier Curve in E^3.

AMS Subject Classifications: 53A04, 53A05.

Keywords: Bézier polygon, $5^{th}order$ Bézier Curve, Bézier polygonal region.

Contents

1. Introduction

French engineer Pierre Bézier, who used Bézier curves to design automobile bodies studied with them in 1962. But the study of these curves was first developed in 1959 by mathematician Paul de Casteljau using de Casteljau's algorithm, a numerically stable method to evaluate Bézier curves. A Bézier curve is frequently used in computer graphics and related fields, in vector graphics, used in animation as a tool to control motion. To guarantee smoothness, the control point at which two curves meet must be on the line between the two control points on either side. In animation applications, such as Adobe Flash and Synfig, Bézier curves are used to outline, for example, movement. Users design the wanted path in Bézier curves, and the application creates the needed frames for the object to move along the path. For 3D animation Bézier curves are often used to define 3D paths as well as 2D paths for key frame interpolation. We have been motivated by the following studies. In [2, 6], the use of Bézier curves on object modeling purposes has been given for Computer-Aided Geometric designs. Moreover, Bézier curves with curvature and torsion continuity has been examined in [8]. In [13], Frenet apparatus of the cubic Bézier curves have been examined in E^3. The matrix representations for a given Bézier curve and its derivatives have been contented in [7, 10–12, 17].In addition, the use and the generation method of Bézier curves have other possible applications as given in [1, 3–5, 9]. Recently, the examination of a Bézier curve by means of curve pairs such as involute, Bertrand or Mannheim partner curves has been given in [14–16].

*Corresponding author. Email address: **seyda@baskent.edu.tr** (Şeyda Kılıçoğlu)

2. Preliminaries

A Bézier curve is defined by a set of control points P_0 through P_n, where n is called its order. If $n = 1$ for linear, if $n = 2$ for quadratic, if $n = 3$ for cubic Bézier curve, etc. The first and last control points are always the end points of the curve; however, the intermediate control points (if any) generally do not lie on the curve. Generaly Bézier curve can be defined by $n+1$ control points $P_0, P_1, ..., P_n$ and has the following form, the points P_i are called control points for the Bézier curve. The polygon formed by connecting the Bézier points with lines, starting with P_0 and finishing with P_n, is called the Bézier polygon (or control polygon). Bézier curve with $n+1$ control points $P_0, P_1, ..., P_n$ has the following equation [2, 6]

$$\mathbf{B}(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} (t) [P_i] \ , \quad t \in [0,1]$$

where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ are the binomial coefficients.

Theorem 2.1. *The derivatives of a given Bézier curve* $\mathbf{B}(t)$ *is*

$$\mathbf{B}'(t) = \sum_{i=0}^{n-1} \binom{n-1}{i} t^i (1-t)^{n-i-1} Q_i$$

where $Q_i = n(P_{i+1} - P_i)$ *[2, 6].*

Given points P_0 and P_1, a linear Bézier curve is simply a straight line between those two points. Linear Bézier curve is given by $\alpha(t) = (1-t) P_0 + t P_1$ and also the matrix form of a linear Bézier curve is

$$\alpha(t) = \begin{bmatrix} t & 1 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \end{bmatrix}.$$

A quadratic Bézier curve is the path traced by the function $\alpha(t)$, given points P_0, P_1 and P_2 which can be interpreted as the linear interpolant of corresponding points on the linear Bézier curves from P_0 to P_1 and from P_1 to P_2 respectively.and also a quadratic Bézier curve has the matrix form with control points P_0 , P_1 and P_2

$$\alpha(t) = \begin{bmatrix} t^2 \\ t \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \end{bmatrix}.$$

Four points in the plane or in higher-dimensional space define a cubic Bézier curve with the following equation $\alpha(t) = (1-t)^3 P_0 + 3t(1-t)^2 P_1 + 3t^2(1-t) P_2 + t^3 P_3$ with the matrix form of a cubic Béziercurve with control points P_0, P_1, P_2, and P_3, is

$$\alpha(t) = \begin{bmatrix} t^3 \\ t^2 \\ t \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \end{bmatrix}.$$

We have already examined the cubic Bézier curves and involutes in [13] and [14], respectively.

3. The area of the Bézier polygonal regions

Definition 3.1. *The* P_i *polygon formed by connecting the Bézier control points with lines, starting with* P_0 *and finishing with* P_n*, is called the Bézier polygon (or control polygon). The convex hull of the Bézier polygon contains the Bézier curve.*

Definition 3.2. *The area of the Bézier polygonal region containing the* n^{th} $order$ *Bézier Curve which is given as*

$$\alpha(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{n-i} (t) [P_i] \ , \quad t \in [0,1] .$$

with control points P_0, P_1, ..., P_n *is defined as the sum of the area of the each area of triangles* $\Delta(P_0, P_1, P_2)$, $\Delta(P_0, P_2, P_3)$, $\Delta(P_0, P_3, P_4)$, ...,$\Delta(P_0, P_{n-1}, P_n)$ *as in the following way*

$$A(P_0, P_1, ..., \ P_n) = A(P_0, P_1, P_2) + A(P_0, P_2, P_3) + ... + A(P_0, P_{n-1}, P_n) .$$

Theorem 3.3. *The area of the Bézier polygonal region containing the* 5^{th} $order$ *BézierCurve and derivatives in* E^3 *is*

$$A(P_0, P_1, P_2, P_3, P_4, \ P_5) = \frac{1}{2}\sum_{i=1}^{4} \|P_0 \wedge (P_i + P_{i+1})\|$$

Proof. From the definition the area of the Bézier polygonal region containing the 5^{th} $order$ Bézier Curve

$$\alpha(t) = \sum_{i=0}^{5} \binom{5}{i} t^i (1-t)^{5-i} (t) [P_i] \ , \quad t \in [0,1] .$$

with control points P_0, P_1, P_2, P_3, P_4, and P_5 is defined as the sum of the area of the each area of triangles $\Delta(P_0, P_1, P_2)$, $\Delta(P_0, P_2, P_3)$, $\Delta(P_0, P_3, P_4)$, and $\Delta(P_0, P_4, P_5)$ as in the following way

$$A(P_0, P_1, P_2, P_3, P_4, \ P_5) = A(P_0, P_1, P_2) + A(P_0, P_2, P_3) + A(P_0, P_3, P_4) + A(P_0, P_4, P_5) .$$

The matrix representation of 5^{th} $order$ Bézier curve with control points P_0, P_1, P_2, P_3, P_4, and P_5 is

$$\alpha(t) = \begin{bmatrix} t^5 \\ t^4 \\ t^3 \\ t^2 \\ t \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 & 5 & -10 & 10 & -5 & 1 \\ 5 & -20 & 30 & -20 & 5 & 0 \\ -10 & 30 & -30 & 10 & 0 & 0 \\ 10 & -20 & 10 & 0 & 0 & 0 \\ -5 & 5 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} P_0 \\ P_1 \\ P_2 \\ P_3 \\ P_4 \\ P_5 \end{bmatrix}$$

The area of the Bézier polygonal region that contains the 5^{th} $order$ BézierCurve with control points P_0, P_1, P_2, P_3, P_4, and P_5 is defined as the sum of the area of the

$$\begin{aligned} A(P_0, P_1, P_2, P_3, P_4, \ P_5) &= A(P_0, P_1, P_2) + A(P_0, P_2, P_3) + A(P_0, P_3, P_4) + A(P_0, P_4, P_5) \\ &= \frac{1}{2} \begin{matrix} (\|P_0P_1 \wedge P_0P_2\| + \|P_0P_2 \wedge P_0P_3\| + \|P_0P_3 \wedge P_0P_4\| \\ + \|P_0P_4 \wedge P_0P_5\|) \end{matrix} \\ &= \frac{1}{2}\sum_{i=1}^{4} \|P_0 \wedge (P_i + P_{i+1})\| . \end{aligned}$$

■

We can generalize the above theorem to the n^{th} $order$ of a Bézier curve, hence we get the following theorem;

Theorem 3.4. *The area of the Bézier polygonal region having the* n^{th} *order Bézier Curve and derivatives in* E^3 *is*

$$A(P_0, P_1, P_2, P_3, ..., \ P_n) = \frac{1}{2}\sum_{i=1}^{n-1} \|P_0 \wedge (P_i + P_{i+1})\| .$$

Theorem 3.5. *The area of the Bézier polygonal region having the first derivative of* 5^{th} *order of a Bézier curve as a* $4th$ *order Bézier curve with control points*P_0, P_1, P_2, P_3, P_4, *and* P_5 *of* 5^{th} *order BézierCurve*

$$A\left(Q_0,Q_1,Q_2,Q_3,Q_4\right)=\frac{25}{2}\sum_{i=1}^{3}\left\|\left(P_0-P_1\right)\wedge\left(P_i-P_{i+2}\right)\right\|$$

Proof. The matrix representation of the first derivative of 5^{th} $order$ of a Bézier curve as a $4th$ order Béziercurve with control points Q_0,Q_1,Q_2,Q_3,Q_4

$$\alpha'(t)=\begin{bmatrix} t^4\\ t^3\\ t^2\\ t\\ 1\end{bmatrix}^T\begin{bmatrix} 1 & -4 & 6 & -4 & 1\\ -4 & 12 & -12 & 4 & 0\\ 6 & -12 & 6 & 0 & 0\\ -4 & 4 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix} Q_0\\ Q_1\\ Q_2\\ Q_3\\ Q_4\end{bmatrix}$$

where the control points, $Q_0=5\left(P_1-P_0\right), Q_1=5\left(P_2-P_1\right), Q_2=5\left(P_3-P_2\right), Q_3=5\left(P_4-P_3\right)$, and $Q_4=5\left(P_5-P_4\right)$ respectively. The area of the Bézier polygonal region contains the first derivative of 5^{th} $order$ of a Bézier curve as a $4th$ order Bézier curve with control points Q_0,Q_1,Q_2,Q_3,Q_4 is

$$\begin{aligned}A\left(Q_0,Q_1,Q_2,Q_3,Q_4\right)&=\frac{1}{2}\sum_{i=1}^{3}\left\|Q_0\wedge\left(Q_i+Q_{i+1}\right)\right\|\\
&=A\left(Q_0,Q_1,Q_2\right)+A\left(Q_0,Q_2,Q_3\right)+A\left(Q_0,Q_3,Q_4\right)\\
&=\frac{1}{2}\left(\left\|Q_0Q_1\wedge Q_0Q_2\right\|+\left\|Q_0Q_2\wedge Q_0Q_3\right\|+\left\|Q_0Q_3\wedge Q_0Q_4\right\|+\left\|Q_0Q_4\wedge Q_0Q_5\right\|\right)\\
&=\frac{1}{2}\left(\left\|\left(Q_1+Q_2\right)\wedge\left(-Q_0\right)\right\|+\left\|\left(Q_2+Q_3\right)\wedge\left(-Q_0\right)\right\|+\left\|\left(Q_3+Q_4\right)\wedge\left(-Q_0\right)\right\|\right)\\
&=\frac{1}{2}\left(\left\|Q_0\wedge\left(Q_1+Q_2\right)\right\|+\left\|Q_0\wedge\left(Q_2+Q_3\right)\right\|+\left\|Q_0\wedge\left(Q_3+Q_4\right)\right\|\right)\\
&=\frac{1}{2}\sum_{i=1}^{3}\left\|Q_0\wedge\left(Q_i+Q_{i+1}\right)\right\|\end{aligned}$$

Also using the control points P_0, P_1, P_2, P_3, P_4, and P_5 of 5^{th} $order$ BézierCurve

$$\begin{aligned}2A\left(Q_0,Q_1,Q_2,Q_3,Q_4\right)&=\left\|Q_0\wedge\left(Q_1+Q_2\right)\right\|+\left\|Q_0\wedge\left(Q_2+Q_3\right)\right\|+\left\|Q_0\wedge\left(Q_3+Q_4\right)\right\|\\
&=\left\|5\left(P_1-P_0\right)\wedge\left(5\left(P_2-P_1\right)+5\left(P_3-P_2\right)\right)\right\|\\
&\quad+\left\|5\left(P_1-P_0\right)\wedge\left(5\left(P_3-P_2\right)+5\left(P_4-P_3\right)\right)\right\|\\
&\quad+\left\|5\left(P_1-P_0\right)\wedge\left(5\left(P_4-P_3\right)+5\left(P_5-P_4\right)\right)\right\|\\
&=25\left\|\left(P_1-P_0\right)\wedge\left(\left(P_2-P_1\right)+\left(P_3-P_2\right)\right)\right\|\\
&\quad+25\left\|\left(P_1-P_0\right)\wedge\left(\left(P_3-P_2\right)+\left(P_4-P_3\right)\right)\right\|\\
&\quad+25\left\|\left(P_1-P_0\right)\wedge\left(\left(P_4-P_3\right)+\left(P_5-P_4\right)\right)\right\|\\
&=25\left\|\left(P_0-P_1\right)\wedge\left(P_1-P_3\right)\right\|\\
&\quad+25\left\|\left(P_0-P_1\right)\wedge\left(P_2-P_4\right)\right\|\\
&\quad+25\left\|\left(P_0-P_1\right)\wedge\left(P_3-P_5\right)\right\|.\end{aligned}$$

This complete the proof. ■

If we generalize the above theorem to the n^{th} $order$ of a Bézier curve we get the following theorem;

Theorem 3.6. *The area of the Bézier polygonal region containing the first derivative of* n^{th} *order of a Bézier curve as a* $(n-1)^{th}$ *order Béziercurve with control points* $Q_0, Q_1, Q_2, ..., Q_{n-1}$ *is*

$$A\left(Q_0, Q_1, Q_2, ..., Q_{n-1}\right) = \frac{1}{2}\sum_{i=1}^{n-2}\left\|Q_0 \wedge \left(Q_i + Q_{i+1}\right)\right\|.$$

Also using the control points $P_0,\ P_1,\ ...,\ \ P_n$ *of* n^{th} *order BézierCurve*

$$A\left(Q_0, Q_1, Q_2, ..., Q_{n-1}\right) = \frac{1}{2}n^2\sum_{i=1}^{n-2}\left\|\left(P_0 - P_1\right) \wedge \left(P_i - P_{i+2}\right)\right\|.$$

Theorem 3.7. *The area of the Bézier polygonal region containing the second derivative of* 5^{th} *order of a Bézier curve as a* $3rd$ *order Béziercurve with control points* $P_0,\ P_1,\ P_2,\ P_3,\ P_4,$ *and* $\ P_5$ *of* 5^{th} *order BézierCurve is*

$$A\left(R_0, R_1, ...,\ R_{n-2}\right) = \frac{20^2}{2}\sum_{i=1}^{n-3}\left\|\left(P_0 - 2P_1 + P_2\right) \wedge \left(P_i - P_{i+1} - P_{i+2} + P_{i+3}\right)\right\|.$$

Proof. The matrix representation of the second derivative of 5^{th} $order$ of a Bézier curve with control points R_0, R_1, R_2, R_3 is

$$\alpha^{''}(t) = \begin{bmatrix} t^3 \\ t^2 \\ t \\ 1 \end{bmatrix}^T \begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 3 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} R_0 \\ R_1 \\ R_2 \\ R_3 \end{bmatrix}$$

where $R_0,\ R_1, R_2,\ R_3$ are control points. The area of the Bézier polygonal region having the second derivative of 5^{th} $order$ of a Bézier curve as a $3rd$ order Béziercurve with control points $R_0,\ R_1,\ R_2,$ and R_3 is

$$A\left(R_0, R_1, ...,\ R_{n-2}\right) = \frac{1}{2}\sum_{i=1}^{2}\left\|R_0 \wedge \left(R_i + R_{i+1}\right)\right\|$$

Also using the control points $P_0,\ P_1,\ P_2,\ P_3,\ P_4,$ and $\ P_5$ of 5^{th} $order$ BézierCurve, and

$$R_0 = 20\left(P_0 - 2P_1 + P_2\right), R_1 = 20\left(P_1 - 2P_2 + P_3\right),$$
$$R_2 = 20\left(P_2 - 2P_3 + P_4\right), R_3 = 20\left(P_3 - 2P_4 + P_5\right)$$

and

$$R_1 + R_2 = 20\left(P_1 - P_2 - P_3 + P_4\right), R_2 + R_3 = 20\left(P_2 - P_3 - P_4 + P_5\right)$$

we get the proof as in the following way

$$\begin{aligned} A\left(R_0, R_1, R_2,\ R_3\right) &= \frac{1}{2}\left(\left\|R_0 \wedge \left(R_1 + R_2\right)\right\| + \left\|R_0 \wedge \left(R_2 + R_3\right)\right\|\right) \\ &= \frac{1}{2}\left(\left\|20\left(P_0 - 2P_1 + P_2\right) \wedge \left(R_1 + R_2\right)\right\| + \left\|20\left(P_0 - 2P_1 + P_2\right) \wedge \left(R_2 + R_3\right)\right\|\right) \\ A\left(R_0, R_1, R_2,\ R_3\right) &= \frac{20^2}{2}\sum_{i=1}^{2}\left\|\left(P_0 - 2P_1 + P_2\right) \wedge \left(P_i - P_{i+1} - P_{i+2} + P_{i+3}\right)\right\|. \end{aligned}$$

■

If we generalize the above theorem to the n^{th} $order$ of a Bézier curve we get the following theorem;

Theorem 3.8. *The area of the Bézier polygonal region contains the second derivative of* n^{th} *order of a Bézier curve as a* $(n-2)^{th}$ *order Béziercurve with control points is* $R_0, R_1, ..., R_{n-2}$

$$A\left(R_0, R_1, R_2,\ R_3\right) = \frac{1}{2}\sum_{i=1}^{n-3} \left\| R_0 \wedge \left(R_i + R_{i+1}\right)\right\|.$$

Also using the control points $P_0, P_1, ...,\ P_n$ *of* n^{th} *order BézierCurve*

$$A\left(R_0, R_1, R_2,\ R_3\right) = \frac{1}{2}\left(n(n-1)\right)^2 \sum_{i=1}^{2} \left\| \left(P_0 - 2P_1 + P_2\right) \wedge \left(P_i - P_{i+1} - P_{i+2} + P_{i+3}\right)\right\|.$$

Theorem 3.9. *The area of the Bézier polygonal region containing the third derivative of* 5^{th} *order of a Bézier curve as a* $2nd$ *order Béziercurve with control points* S_0, S_1, S_2 *is*

$$A\left(S_0, S_1, S_2\right) = \frac{1}{2}\left\| S_0 \wedge \left(S_1 + S_2\right)\right\|$$

Also using the control points $P_0, P_1, P_2, P_3, P_4,$ *and* P_5 *of* 5^{th} *order BézierCurve*

$$A\left(S_0, S_1, S_2\right) = 2.60^2 \left\| \left(-P_0 + 3P_1 - 3P_2 + P_3\right) \wedge \left(-2P_0 + 5P_1 + 2P_3 + 5P_4 + P_5\right)\right\|$$

Proof. The matrix representation of the third derivative of 5^{th} $order$ of a Bézier curve with control points S_0, S_1, S_2 is

$$\alpha^{'''}(t) = \begin{bmatrix} t^2 \\ t \\ 1 \end{bmatrix}^T \begin{bmatrix} 1 & -2 & 1 \\ -2 & 2 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} S_0 \\ S_1 \\ S_2 \end{bmatrix}$$

where

$S_0 = 60\left(6P_1 - 2P_0 - 6P_2 + 2P_3\right),\ S_1 = 60\left(2P_1 - P_0 - 2P_3 + P_4\right)$, and
$S_2 = 60\left(3P_1 - P_0 - 4P_2 + 4P_3 - 3P_4 + P_5\right)$
hence

$$S_1 + S_2 = 60\left(5P_1 - 2P_0 + 2P_3 + 5P_4 + P_5\right)$$

The area of the Bézier polygonal region for the third derivative of 5^{th} $order$ of a Bézier curve as a $2nd$ order Béziercurve with control points S_0, S_1, S_2 is

$$A\left(S_0, S_1, S_2\right) = \frac{1}{2}\left\| S_0 \wedge \left(S_1 + S_2\right)\right\|.$$

Hence

$$\begin{aligned} A\left(S_0, S_1, S_2\right) &= \frac{1}{2}\left\| \left(S_1 - S_0\right) \wedge \left(S_2 - S_0\right)\right\| \\ &= \frac{1}{2}\left\| S_0 \wedge \left(S_1 + S_2\right)\right\| \\ &= \frac{60^2}{2}\left\| \left(-2P_0 + 6P_1 - 6P_2 + 2P_3\right) \wedge \left(5P_1 - 2P_0 + 2P_3 + 5P_4 + P_5\right)\right\| \\ A\left(S_0, S_1, S_2\right) &= \frac{60^2}{2}\left\| \left(-P_0 + 3P_1 - 3P_2 + P_3\right) \wedge \left(-2P_0 + 5P_1 + 2P_3 + 5P_4 + P_5\right)\right\|. \end{aligned}$$

We have the proof. ■

If we generalize the above theorem to the n^{th} $order$ of a Bézier curve we get the following theorem;

Theorem 3.10. *The area of the Bézier polygonal region for the third derivative of* n^{th} *order of a Bézier curve as a* $(n-3)^{nd}$ *order Béziercurve with control points* $S_0, S_1, ..., S_{n-3}$ *is*

$$A\left(S_0, S_1, ..., S_{n-3}\right) = \frac{1}{2}\sum_{i=1}^{2} \left\| S_0 \wedge \left(S_i + S_{i+1}\right) \right\|$$

Also using the control points P_0, P_1, P_2, P_3, P_4, *and* P_5 *of* 5^{th} *order BézierCurve*

$$A\left(S_0, S_1, ..., S_{n-3}\right) = \frac{\left({}^{n}\boldsymbol{P}_3\right)^2}{2}\sum_{i=1}^{2} \left\| \left(-P_0 + 3P_1 - 3P_2 + P_3\right) \wedge \left(-2P_{i-1} + 5P_i + 2P_{i+2} + 5P_{i+3} + P_{i+4}\right) \right\|,$$

where ${}^{n}\boldsymbol{P}_3 = n(n-1)(n-2)$ *is permutation.*

Theorem 3.11. *The length of the* T_0T_1, *of the fourth derivative of* 5^{th} *order of a Bézier curve is a linear Béziercurve, with control points* T_0, *and* T_1 *is*

$$\|T_0T_1\| = 5.4.3.2.1 \left\| -P_0 + 5P_1 - 10P_2 + 10P_3 - 5P_4 + P_5 \right\|$$

Proof. The fourth derivative of 5^{th} $order$ of a Bézier curve has the following representation.

$$\alpha'^{v}(t) = \begin{bmatrix} t \\ 1 \end{bmatrix}^{T} \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} T_0 \\ T_1 \end{bmatrix}$$

where

$$\begin{aligned} T_0 &= 120P_0 - 480P_1 + 720P_2 - 480P_3 + 120P_4 \\ T_1 &= 120P_1 - 480P_2 + 720P_3 - 480P_4 + 120P_5 \end{aligned}$$

are the control points of the fourth derivative of $5^{th} order$ of a Bézier curve based on the P_0, P_1, P_2, ..., and P_5.

$$\begin{aligned} \|T_0T_1\| &= \left\| \begin{array}{c} \left(120P_1 - 480P_2 + 720P_3 - 480P_4 + 120P_5\right) \\ -\left(120P_0 - 480P_1 + 720P_2 - 480P_3 + 120P_4\right) \end{array} \right\| \\ &= \left\| 600P_1 - 120P_0 - 1200P_2 + 1200P_3 - 600P_4 + 120P_5 \right\| \\ &= 5.4.3.2.1 \left\| -P_0 + 5P_1 - 10P_2 + 10P_3 - 5P_4 + P_5 \right\| \end{aligned}$$

■

Example 3.12. *Let* $\alpha(t)$ *be a* $5th$ *order Bézier curve given by the following parametrization:*

$$\alpha(t) = \begin{array}{c} \big(74t^5 - 210t^4 + 180t^3 - 50t^2 + 5t + 1, \\ -79t^5 + 185t^4 - 130t^3 + 10t^2 + 10t + 1, \\ -63t^5 + 95t^4 - 30t^3 - 5t + 2\big) \end{array}$$

with control points, $P_0 = (1,1,2), P_1 = (2,3,1), P_2 = (-2,6,0), P_3 = (7,-3,-4), P_4 = (5,0,5)$, $P_5 = (0,-3,-1)$.

The area of the Bézier polygonal region containing the 5^{th} $order$ *Bézier curve is*

$$\begin{aligned}
&A\left(P_0, P_1, P_2, P_3, P_4,\ P_5\right)\\
&=\frac{1}{2}\sum_{i=1}^{4}\|P_0\wedge(P_i+P_{i+1})\|\\
&=\frac{1}{2}\left(\|P_0\wedge(0,9,1)\|+\|P_0\wedge(5,3,-4)\|+\|P_0\wedge(12,-3,1)\|+\|P_0\wedge(5,-3,4)\|\right)\\
&=\frac{1}{2}\left\|\begin{vmatrix} i & j & k\\ 1 & 1 & 2\\ 0 & 9 & 1\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ 1 & 1 & 2\\ 5 & 3 & -4\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ 1 & 1 & 2\\ 12 & -3 & 1\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ 1 & 1 & 2\\ 5 & -3 & 4\end{vmatrix}\right\|\\
&=39.\,531\ unit\ square.
\end{aligned}$$

The area of the Bézier polygonal region containing the first derivative of 5^{th} $order$ *of a Bézier curve is*

$$\begin{aligned}
&A\left(Q_0, Q_1, Q_2, Q_3, Q_4\right)\\
&=\frac{1}{2}5^2\sum_{i=1}^{3}\|(P_0-P_1)\wedge(P_i-P_{i+2})\|\\
&=\frac{1}{2}5^2\left(\|(P_0-P_1)\wedge(P_1-P_3)\|+\|(P_0-P_1)\wedge(P_2-P_4)\|+\|(P_0-P_1)\wedge(P_3-P_5)\|\right)\\
&=\frac{1}{2}5^2\left(\|(-1\ \ -2\ \ \ 1)\wedge(P_1-P_3)\|+\|(P_0-P_1)\wedge(P_2-P_4)\|\right.\\
&\qquad\left.+\|(P_0-P_1)\wedge(P_3-P_5)\|\right)\\
&=\frac{1}{2}25\left(\left\|\begin{vmatrix} i & j & k\\ -1 & -2 & 1\\ -5 & 6 & 5\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ -1 & -2 & 1\\ -7 & 6 & -5\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ -1 & -2 & 1\\ 7 & 0 & -3\end{vmatrix}\right\|\right)\\
&=\frac{1551.\,0}{2}\\
&=775.\,5\ unit\ square
\end{aligned}$$

The area of the Bézier polygon that contains the second derivative of 5^{th} $order$ *of a Bézier curve as a* $3rd$ *order Béziercurve with control points* R_0, R_1, R_2, R_3 *is*

$$\begin{aligned}
A\left(R_0, R_1, R_2,\ R_3\right)&=\frac{1}{2}20^2\sum_{i=1}^{2}\|(P_0-2P_1+P_2)\wedge(P_i-P_{i+1}-P_{i+2}+P_{i+3})\|\\
&=\frac{1}{2}20^2\left(\left\|\begin{matrix}-5 & 1 & 0\end{matrix}\wedge(P_1-P_2-P_3+P_4)\right\|+\left\|\begin{matrix}-5 & 1 & 0\end{matrix}\wedge(P_2-P_3-P_4+P_5)\right\|\right)\\
&=\frac{1}{2}20^2\left(\left\|\begin{matrix}-5 & 1 & 0\end{matrix}\wedge\left(\begin{matrix}2 & 0 & 10\end{matrix}\right)\right\|+\left\|\begin{matrix}-5 & 1 & 0\end{matrix}\wedge\left(\begin{matrix}-14 & 6 & -2\end{matrix}\right)\right\|\right)\\
&=\frac{1}{2}20^2\left\|\begin{vmatrix} i & j & k\\ -5 & 1 & 0\\ 2 & 0 & 10\end{vmatrix}\right\|+\left\|\begin{vmatrix} i & j & k\\ -5 & 1 & 0\\ -14 & 6 & -2\end{vmatrix}\right\|\\
&=\frac{20431}{2}\\
&=10.216\ unit\ square.
\end{aligned}$$

The area of the Bézier polygonal region containing the third derivative of 5^{th} $order$ *of a Bézier curve using the* *control points* $P_0, P_1, P_2, P_3, P_4,$ *and* P_5 *of* 5^{th} *order Bézier Curve is*

$$
\begin{aligned}
A(S_0,S_1,S_2) &= \frac{1}{2}60^2 \left\|(-P_0+3P_1-3P_2+P_3)\wedge(5P_1-2P_0+2P_3+5P_4+P_5)\right\| \\
&= \frac{1}{2}.60^2 \left\| \begin{matrix} \begin{pmatrix} -\begin{pmatrix}1 & 1 & 2\end{pmatrix}+3\begin{pmatrix}2 & 3 & 1\end{pmatrix} \\ -3\begin{pmatrix}-2 & 6 & 0\end{pmatrix}+\begin{pmatrix}7 & -3 & -4\end{pmatrix}\end{pmatrix} \\ \wedge\begin{pmatrix} 5\begin{pmatrix}2 & 3 & 1\end{pmatrix}-2\begin{pmatrix}1 & 1 & 2\end{pmatrix}+2\begin{pmatrix}7 & -3 & -4\end{pmatrix} \\ +5\begin{pmatrix}5 & 0 & 5\end{pmatrix}+\begin{pmatrix}0 & -3 & -1\end{pmatrix}\end{pmatrix}\end{matrix} \right\| \\
&= \frac{1}{2}60^2 \left\|\begin{pmatrix}18 & -13 & -3\end{pmatrix}\wedge\begin{pmatrix}47 & 4 & 17\end{pmatrix}\right\| \\
&= \frac{1}{2}.60^2 \left\|\begin{vmatrix} i & j & k \\ 18 & -13 & -3 \\ 47 & 4 & 17\end{vmatrix}\right\| \\
&= 5696.0/2 \\
&= 2848.0 \ unit\ square
\end{aligned}
$$

4. Acknowledgement

The author is thankful to the referee for her valuable suggestions which improved the presentation of the paper.

References

[1] A. LEVENT AND B. SAHIN, Cubic bezier-like transition curves with new basis function, *Proceedings of the Institute of Mathematics and Mechanics,National Academy of Sciences of Azerbaijan*, **44(2)**(2018), 222–228.

[2] D. MARSH, Applied Geometry for Computer Graphics and CAD. *Springer Science and Business Media.*, 2006.

[3] E. ERKAN, & S. YÜCE, Some Notes on Geometry of Bézier Curves in Euclidean 4-Space, *Journal of Engineering Technology and Applied Sciences*, **5(3)**(2020), 93–101.

[4] H. ERISKIN, AND A. YÜCESAN, Bézier curve with a minimal jerk energy, *Mathematical Sciences and Applications E-Notes*, **4(2)**(2016), 139–148.

[5] F. TAS AND K. ILARSLAN, A new approach to design the ruled surface, *International Journal of Geometric Methods in Modern Physics*, **16(6)**(2019), 1950093.

[6] G. FARIN, Curves and Surfaces for Computer-Aided Geometric Design, *Academic Press*, 1996.

[7] H. ZHANG AND F. JIEQING, Bézier Curves and Surfaces (2). *State Key Lab of CAD&CG*, Zhejiang University, 2006.

[8] H. HAGEN, Bézier curves with curvature and torsion continuity, *Rocky Mountain J. Math.*, **16(3)** (1986),629–638.

[9] D. PÁLEŠ, AND J. RÉDL, Bézier curve and its application, *Mathematics in Education, Research and Applications*, **1(2)**(2015), 49–55.

[10] S. BAYDAS AND B. KARAKAS, Defining a curve as a Bézier curve, *Journal of Taibah University for Science*, **13(1)**(2019), 522–528.

[11] Ş. Kiliçoğlu and S. Şenyurt, On the Matrix Representation of 5th order Bézier Curve and derivatives, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistic,* **71(1)**(2022), 133–152.

[12] Ş. Kiliçoğlu and S. Şenyurt, An examination on to find 5th Order Bézier Curve in E^3, *Journal of New Theory*, **37**(2021), 35–44.

[13] Ş. Kiliçoğlu and S. Şenyurt, On the cubic Bézier curves in E^3 *Ordu University Journal of Science and Technology*, **9(2)**(2019), 83–97.

[14] Ş. Kiliçoğlu and S. Şenyurt, On the Involute of the Cubic Bézier Curve by Using Matrix Representation in E^3. *European Journal of Pure and Applied Mathematics*, **13**(2020), 216–226.

[15] Ş. Kiliçoğlu and S. Şenyurt, On the Mannheim partner of a cubic Bézier curve in E^3, *International Journal of Maps in Mathematics*, **5(2)**(2022), 182–197.

[16] Ş. Kiliçoğlu and S. Şenyurt, On the Bertrand Mate of Cubic Bézier Curve by Using Matrix Representation in $\mathbf{E}^3$, *Turkish Journal of Mathematics and Computer Science*, **14(2)**(2022), 376–383.

[17] T. A. Aydin A matrix presentation of higher order derivatives of Bézier curves and surfaces, *Journal of Science and Arts,* **21**(1), (2021), 77-90.