Iranian Journal of Numerical Analysis and Optimization

Volume 6 , Number 1

Winter 2016

Ferdowsi University of Mashhad, Iran

In the Name of God

Iranian Journal of Numerical Analysis and Optimization (IJNAO)

This journal is authorized under the registration No. 174/853 dated 1386/2/26, by the Ministry of Culture and Islamic Guidance.

Volume 6, Number 1, Winter 2016

ISSN: 2423-6977

Publisher: Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

Published by: Ferdowsi University of Mashhad Press

Circulation: 100

Address: Iranian Journal of Numerical Analysis and Optimization

Faculty of Mathematical Sciences, Ferdowsi University of Mashhad

P.O. Box 1159, Mashhad 91775, Iran.

Tel. : +98-51-38806222 , **Fax:** +98-51-38807358

E-mail: mjms@um.ac.ir

Website: http://ijnao.um.ac.ir

This journal is indexed by:

• Zentralblatt

• ISC

• SID

به اطلاع کلیه محققان، پژوهشگران، اساتید ارجمند، دانشجویان تحصیلات تکمیلی و نویسندگان محترم می رساند که نشریه ایرانی آنالیز عددی و بهینه سازی - IJNAO - طبق مجوز شماره ۳/۱۸/۵۴۸۹۱۳/ مورخه ۱۳۹۲/۱۰/۲۵ مدیر کل محترم سیاست گذاری و برنامه ریزی امور پژوهشی وزارت علوم، تحقیقات و فناوری، علمی – پژوهشی میباشد.
بر اساس نامه شماره ۹۲/۱۵۳۶/پ مورخه ۱۳۹۲/۱۱/۲۱ سرپرست محترم معاونت پژوهشی وفناوری پایگاه استنادی علوم جهان اسلام، نشریه ایرانی آنالیز عددی و بهینه سازی در پایگاه ISC نیز نمایه می شود.

Iranian Journal of Numerical Analysis and Optimization

Volume 6, Number 1, Winter 2016

Ferdowsi University of Mashhad - Iran

Iranian Journal of Numerical Analysis and Optimization

This journal is published under the auspices of Ferdowsi University of Mashhad

---

* Full Professor

** Associate Professor

We would like to acknowledge the help of Narjes khatoon Zohorian in the preparation of this issue.

Letter from the Editor in Chief

I would like to welcome you to the Iranian Journal of Numerical Analysis and Optimization (IJNAO). This journal is published biannually and supported by the Faculty of Mathematical Sciences at the Ferdowsi University of Mashhad. Faculty of Mathematical Sciences with three centers of excellence and three research centers is well-known in mathematical communities in Iran.

The main aim of the journal is to facilitate discussions and collaborations between specialists in applied mathematics, especially in the fields of numerical analysis and optimization, in the region and worldwide.

Our vision is that scholars from different applied mathematical research disciplines, pool their insight, knowledge and efforts by communicating via this international journal.

In order to assure high quality of the journal, each article is reviewed by subject-qualified referees.

Our expectations for IJNAO are as high as any well-known applied mathematical journal in the world. We trust that by publishing quality research and creative work, the possibility of more collaborations between researchers would be provided. We invite all applied mathematicians especially in the fields of numerical analysis and optimization to join us by submitting their original work to the Iranian Journal of Numerical Analysis and Optimization.

Mohammad Hadi Farahi

Contents

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 1-29

A linearization technique for optimal design of the damping set with internal dissipation

A. Fakharzadeh J.*, H. Alimorad D. and A. Beiranvand

Abstract

Considering a damped wave system defined on a two-dimensional domain, with a dissipative term localized in an unknown subset with an unknown damping parameter, we address the ill-posed shape design problem which consists of optimizing the shape of the unknown subset in order to minimize the energy of the system at a given time. By using a new approach based on the embedding process, first the system is formulated in variational form. Then, by transferring the problem into polar coordinates and defining two positive Radon measures, we represent the problem in a space of measures. Hence, the shape design problem is changed into an infinite linear one whose solution is guaranteed. In this stage, by applying two subsequent approximation steps, the optimal solution (optimal control, optimal region, optimal damping parameter and optimal energy) is identified by a three-phase optimization search technique. Numerical simulations are also given in order to compare this new method with level set algorithm.

Keywords: Damped wave equation; Dissipation control; Radon measure; Search technique; Shape optimization.

1 Introduction and Problem statement

In many technological situations, a given structure whose optimal position is at rest (for instance), starts to vibrate due to uncontrolled disturbances

* Corresponding author
Received 11 February 2015; revised 12 April 2015; accepted 17 June 2015
A. Fakharzadeh J.
Department of Mathematics, Faculty of Mathematics, Shiraz University of Technology, Shiraz, Iran. e-mail: a_fakharzadeh@sutech.ac.ir

H. Alimorad D.
Department of Mathematics, Jahrom University, Jahrom, Iran, P. O. Box: 74135-111
Dept. of Mathematics, Shiraz University of Technology, Shiraz , Iran. e-mail: h.alimorad@sutech.ac.ir

A. Beiranvand
Department of Mathematics, Faculty of Mathematics, Shiraz University of Technology, Shiraz , Iran. e-mail: a.beiranvand@sutech.ac.ir

which we would like to stop. One possibility, although under ideal conditions, is described in [16] through damping mechanisms. In the literature, the problem of optimal stabilization for the 2-D wave equation has been extensively studied from different perspectives (see for instance, [4], [9] and [13]). The analysis performed by Hebrard et al. highlights the effect of the over-damping phenomenon characteristic of this damped wave equation [2]. Freitas [9] and Lopez [13] solved the mentioned problem in which the dissipation vanishes for large values of the constant damping coefficient. In 2006, Munch et al. used Young measures to solve a similar problem and presented a solution method. In that study, the damping coefficient was fixed and the best unknown internal region was determined by the use of the gradient descend method [15]. In sequence, the best damping coefficient and damping set were determined at different times using the level set method [16].

In this paper, we solve the problem of finding an optimal observation domain $\omega \subset \Omega \subset \mathbb{R}^2$ for general damping wave equations in a new way. We optimize not only the placement but also the shape of ω, over all possible measurable subsets of Ω having a certain prescribed measure. Such questions are frequently encountered in engineering applications but have rarely been treated from the mathematical point of view. In this regard, for the first time, we consider a shape optimization problem to find the optimal shape and place of a sensor, modeled by a two-dimensional wave equation. The objective is to find the shape of the damping set that minimizes the energy at some given end time (see [18] and [19]).

2 Optimal wave damping problem

Let $\Omega \subset \mathbb{R}^2$ be a domain with piecewise smooth boundary and consider the two-dimensional damping wave equation with Dirichlet boundary conditions. Consider additionally that ω is a subset of Ω with positive Lebesgue measure which is independent of time $t \in (0,T)$. The resulting equation for the displacement of the sensor is then ([2] and [16])

$$\begin{array}{ll} \ddot{y}_{\omega,a} - \Delta y_{\omega,a} + a(x)\dot{y}_{\omega,a} = 0, & (x,t) \in \Omega \times (0,T), \\ y_{\omega,a} = 0, & (x,t) \in \partial\Omega \times (0,T), \\ y_{\omega,a}(x,0) = y_0(x), \quad \dot{y}_{\omega,a}(x,0) = y_1(x), & x \in \Omega; \end{array} \tag{1}$$

here, $a(x) = a\chi_\omega(x) \in L^\infty(\Omega, \mathbb{R}^+)$ is a damping function where $a \in \mathbb{R}^+$ is unknown, $\varnothing \neq \omega$ is an unknown region in Ω, $\partial\omega$ is a smooth and simple closed curve boundary which must be identified, χ_ω is a characteristic function of ω and y_0 and y_1 also indicate the initial position and velocity, respectively. In addition, regarding the initial conditions, we assume that:

$$(y_0(x), y_1(x)) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times H_0^1(\Omega),$$

where $H_0^1(\Omega)$ (the Sobolev space of order 1 on Ω whose functions are zero on the boundary of Ω [22]). System (1) is well-posed [12], whose unique solution satisfies:

$$y_{\omega,a} \in C((0,T); H^2(\Omega) \cap H_0^1(\Omega) \cap C^1((0,T); H_0^1(\Omega) \cap C^2((0,T); L^2(\Omega)).$$

For every $t > 0$, the energy of system (1) is obtained through ([16])

$$E(\omega, a, t) = \frac{1}{2} \int_\Omega (|\dot{y}_{\omega,a}(x,t)|^2 + |\nabla y_{\omega,a}(x,t)|^2 dx, \tag{2}$$

which satisfies the following dissipation law (see [23]):

$$\dot{E}(\omega, a, t) = -\int_\Omega a(x)|\dot{y}_{\omega,a}(x,t)|^2 dx \leq 0.$$

Here, $y_{\omega,a}$ denotes the transversal displacement at point x in time t.

We attempt to find the unknown region ω and damping function $a(x)$, simultaneously, through a three-phase optimization procedure which is based on an embedding technique. To apply this method, first, we present the problem in variational form; next, it is transferred into a new theoretical measure problem in which two unknown positive Radon measures in a product space of measures are sought. Then, the solution procedure is explained and finally, by a three-phase optimization technique, a nearly optimal shape together with the optimal damping function as well as the minimized value of system energy are constructed.

The paper is organized as follows: the next section is devoted to the basic deformation in variational form. The aim of Section 4 is to state the problem in a polar system. Section 5 deals with the embedding process and approximation schemes. In Section 6, based on the previous discussions, we present the solution algorithm. Then, two numerical simulations are presented in Section 7. Finally, concluding remarks are offered in Section 8.

3 Basic Deformation

In general, it is difficult to identify a calssical solution for problem (1); thus attempts have usually been made to find a weak (or generalized) solution of the problem, which is more applicable in our work. The main idea in this replacement is to convert the problem into the variational form. To this end, by multiplying the first equation of system (1) with a function $\varphi \in H_0^1(\Omega \times (0,T))$ and using Green's theorem, to the initial conditions, for each t, one obtains:

$$\int_\Omega y\Delta\varphi dx - \int_\Omega \varphi\Delta y dx = \int_{\partial\Omega} (y\frac{\partial\varphi}{\partial n} - \varphi\frac{\partial y}{\partial n} ds) = 0,$$

therefore

$$\int_\Omega y\Delta\varphi dx = \int_\Omega \varphi\Delta y dx;$$

then, we have:

$$\int_\Omega \ddot{y}\varphi dx - \int_\Omega y\Delta\varphi dx + \int_\Omega a(x)\dot{y}\varphi dx = 0. \tag{3}$$

Integrating both sides of (3) with respect to t over $[0,T]$ gives:

$$\int_0^T \int_\Omega \ddot{y}\varphi dxdt - \int_0^T \int_\Omega y\Delta\varphi dxdt + \int_0^T \int_\Omega a(x)\dot{y}\varphi dxdt = 0, \tag{4}$$

Double integrating by parts with respect to t from the first left expression and integrating the third expression on the left-hand side of (4), we conclude:

$$\begin{aligned} &\textstyle\int_0^T \int_\Omega \ddot{y}\varphi dxdt = \int_\Omega [\dot{y}(T)\varphi(T) - \dot{y}(0)\varphi(0) - y(T)\dot{\varphi}(T) + y(0)\dot{\varphi}(0)]dx \\ &\qquad\qquad + \textstyle\int_0^T \int_\Omega y\ddot{\varphi}dxdt; \\ &\textstyle\int_0^T \int_\Omega a(x)\dot{y}\varphi dxdt = \int_\Omega a(x)[y(T)\varphi(T) - y(0)\varphi(0) - \int_0^T y\dot{\varphi}]dx \\ &\qquad\qquad = \textstyle\int_\Omega a(x)[y(T)\varphi(T) - y(0)\varphi(0)]dx - \int_0^T \int_\Omega a(x)y\dot{\varphi}dxdt. \end{aligned} \tag{5}$$

Now, by substituting the initial conditions of system (1) in (5), we have:

$$\begin{aligned} &\textstyle\int_0^T \int_\Omega \ddot{y}\varphi dxdt = \int_\Omega [\dot{y}(T)\varphi(T) - y_1(x)\varphi(0) - y(T)\dot{\varphi}(T) + y_0(x)\dot{\varphi}(0)]dx \\ &\qquad\qquad + \textstyle\int_0^T \int_\Omega y\ddot{\varphi}dxdt; \\ &\textstyle\int_0^T \int_\Omega a(x)\dot{y}\varphi dxdt = \int_\Omega a(x)[y(T)\varphi(T) - y_0(x)\varphi(0)]dx - \int_0^T \int_\Omega a(x)y\dot{\varphi}dxdt. \end{aligned} \tag{6}$$

By applying (6), equation (4) is changed to:

$$\begin{aligned} &\textstyle\int_\Omega \dot{y}(T)\varphi(T)dx - \int_\Omega y(T)\dot{\varphi}(T)dx - \int_0^T \int_\Omega y\Delta\varphi dxdt + \int_\Omega a(x)y(T)\varphi(T)dx \\ &\qquad - \textstyle\int_\Omega a(x)y_0(x)\varphi(0)dx - \int_0^T \int_\Omega a(x)y\dot{\varphi}dxdt + \int_0^T \int_\Omega y\ddot{\varphi}dxdt \\ &\qquad = \textstyle\int_\Omega [y_1(x)\varphi(0) - y_0(x)\dot{\varphi}(0)]dx. \end{aligned} \tag{7}$$

Moreover, for all $(x,t) \in \partial\Omega \times [0,T]$ by the initial condition, we have $y(x,t) = 0$; to apply this condition and using Green's theorem, we have:

$$\int_{\partial\Omega} y(x,t)\varphi(x,t).nd\sigma = \int_\Omega div(y(x,t)\varphi(x,t))dx = 0. \tag{8}$$

Since the unknown region ω must lie in Ω and the measure of this unknown region must be non-zero, the set of admissible shapes for problem (1) can be shown as:

$$V_L = \{\omega \subset \Omega : |\omega| = L|\Omega| \ \}, \qquad 0 < L < 1 \tag{9}$$

in which $|\omega|$ indicates the measure of ω and L is a fixed number. This constraint can be shown by the following integral relation:

$$\int_{\omega} dx = L \int_{\Omega} dx. \tag{10}$$

4 Expressing the problem in polar system

The mentioned optimal shape design (OSD) problem is defined based on the unknown geometrical pair $(\omega, \partial\omega)$. This pair consists of a measurable set that can be regarded as a nonempty region, and a simple closed curve which is its boundary. Based on the simplicity property of the curve, our OSD problem depends on the geometry which is used. We prefer to solve the appropriate problems in polar coordinates since where $0 \leq \theta \leq 2\pi$ and $r \geqslant 0$, the curve $r = r(\theta)$ is simple. This simple fact is an essential part in our calculations and also in numerical simulations. Hence, let $x_1 = rcos(\theta)$ and $x_2 = rsin(\theta)$; then, we have:

$$|\nabla y|_T = [(\frac{\partial y}{\partial r})^2 + \frac{1}{r^2}(\frac{\partial y}{\partial \theta})^2], \quad \Delta y|_T = \frac{\partial^2 y}{\partial r^2} + \frac{1}{r}\frac{\partial y}{\partial r} + \frac{1}{r^2}\frac{\partial^2 y}{\partial \theta^2},$$

and therefore:

$$E(\omega, a, T) = \frac{1}{2}\int_{\Omega} [|\dot{y}(r, \theta, T)|^2 + (\frac{\partial y}{\partial r})^2 + \frac{1}{r^2}(\frac{\partial y}{\partial \theta})^2] r dr d\theta.$$

Since the nature of Ω has not changed, but rather its representation has changed, we use the same symbol and, in the end, the optimal shape is shown in polar coordinates.

Additionally, for every $\varphi \in H_0^1(\Omega \times (0, T))$, the mentioned constraint in (7) can be represented as:

$$\begin{aligned} \int_{\Omega} \dot{y}(T)\varphi(T) r dr d\theta - \int_{\Omega} y(T)\dot{\varphi}(T) r dr d\theta - \int_0^T \int_{\Omega} y \Delta\varphi r dr d\theta dt \\ + \int_{\Omega} a(r,\theta) y(T)\varphi(T) r dr d\theta - \int_{\Omega} a(r,\theta) y_0(r,\theta)\varphi(0) r dr d\theta \\ - \int_0^T \int_{\Omega} a(r,\theta) y \dot{\varphi} r dr d\theta dt + \int_0^T \int_{\Omega} y \ddot{\varphi} r dr d\theta dt = \Phi, \end{aligned} \tag{11}$$

in which

$$\Phi = \int_{\Omega} [y_1(r,\theta)\varphi(0) - y_0(r,\theta)\dot{\varphi}(0)] r dr d\theta.$$

Moreover, equations (8) and (9) can be represented in polar coordinates as:

$$\begin{aligned} &\int_\Omega div(y(r,\theta,t)\varphi(r,\theta,t))rdrd\theta = 0,\\ &\int_\omega rdrd\theta = L\int_\Omega rdrd\theta. \end{aligned} \tag{12}$$

Therefore, the problem of obtaining the optimal shape and damping coefficient for minimizing energy in polar coordinates has the following presentation:

$$\begin{aligned} &Min: E(\omega,a,T) = \tfrac{1}{2}\int_\Omega [|\dot{y}(r,\theta,T)|^2 + (\tfrac{\partial y}{\partial r})^2 + \tfrac{1}{r^2}(\tfrac{\partial y}{\partial \theta})^2]rdrd\theta,\\ &S.\ to: \int_\Omega \dot{y}(T)\varphi(T)rdrd\theta - \int_\Omega y(T)\dot{\varphi}(T)rdrd\theta - \int_0^T\int_\Omega y\Delta\varphi rdrd\theta dt\\ &\quad + \int_\Omega a(r,\theta)y(T)\varphi(T)rdrd\theta - \int_\Omega a(r,\theta)y_0(r,\theta)\varphi(0)rdrd\theta\\ &\quad - \int_0^T\int_\Omega a(r,\theta)y\dot{\varphi}rdrd\theta dt + \int_0^T\int_\Omega y\ddot{\varphi}rdrd\theta dt = \Phi;\\ &\quad \int_\Omega div(y(r,\theta,t)\varphi(r,\theta,t))rdrd\theta = 0,\ \forall\varphi\in H_0^1(\Omega\times(0,T));\\ &\quad \int_\omega rdrd\theta = L\int_\Omega rdrd\theta. \end{aligned} \tag{13}$$

To solve (12), we change the problem and consider a new one with a different formulation. By applying this method, we show how one can obtain the optimal region ω, optimal damping function $a(x)$ and the amount of minimized energy simultaneously.

5 Embedding the solution space: metamorphosis

The solution method which is based on an embedding process involves several stages to set up a linear programming problem whose solution converges to the solution of the original problem (see [20]). This is one of the outstanding advantages of this method even for strongly nonlinear problems. Hence, we present a new version of shape measure method to solve the optimal shape design (12). First, by defining a new variational formulation, an optimal control problem equivalent to the original problem is obtained. Then, a measure theoretical approach and a two-stage approximation are used to convert the optimal control problem to a finite dimensional LP. The solution of this LP is used to construct an approximate solution to the original control problem. The proposed approach is practical and accurate enough and its accuracy can be improved as far as desired (see [8]).

5.1 Step 1: Displaying the problem in variational form

In order to transform the optimal shape design into variational form, we need to define some fundamental concepts. The conditions imposed on the functions and sets will serve two important purposes. First, they are reasonable

conditions which are usually met when considering classical problems. Second, they will allow the modification of these classical problems into ones which appear to have some advantages over the classical formulation.

Suppose that $J = [0, 2\pi]$ and $A_1 = [0, r_\Omega]$ are the domains of variables θ and r_1, respectively. For $\Omega = J \times A_1$, let $V \subseteq \mathbb{R}$ be a given bounded and closed set, and $A_2 = [0, r_\omega]$.

Definition 1. Consider the variable $r : J^0 \to A_2$ as an absolutely continuous trajectory function, where $J^0 = (0, 2\pi)$; then, we denote the boundary curve of the unknown region ω by $\partial\omega$ which is introduced by:

$$\partial\omega : r = r(\theta),\ \theta \in J^o. \tag{14}$$

Definition 2. By supposing that r is a simple and closed curve and U is a bounded closed subset in $\mathbb{R}$, we introduce an artificial control function u as follows:

$$u : J^o \to U,$$
$$u(\theta) = \dot{r}(\theta) \equiv g(\theta, r, u),$$

where $r(\theta) \in [0, r_\omega]$ and the boundary $\partial\omega$ is determined by this variable.

Definition 3. Let $S^{'}$ be a bounded closed subset of $\mathbb{R}$ and function $Y : \Omega \times [0, T] \to S'$ be defined in the following way:

$$Y = \frac{\partial y}{\partial \theta}. \tag{15}$$

Since $y = y(r_1, \theta)$ and $r_1 = r_1(\theta) \in \Omega$, we can write $y = y(\theta)$ and $\theta = \theta(r_1)$. In this case, we have $\partial y/\partial r_1 = (\partial y/\partial\theta)(d\theta/dr_1)$ and by introducing $v : J^o \to V$ where $v = dr_1/d\theta$, we can write:

$$\frac{\partial y}{\partial r_1} = \frac{\partial y}{\partial \theta}\frac{1}{v},$$

where $r_1(\theta) \in [0, r_\Omega]$ and $\partial\Omega$ is determined by this variable. From now on, for simiplicity we denote r_1 also with r by regarding that $r_1 = r \in \Omega$; hence the derivative of r in ω and Ω were shown above by u and v respectively.

To identify the relationship between r and v as variables, suppose $E' = J \times A_1 \times V$ and consider h in $C^1(E')$, then:

$$\frac{\partial h}{\partial \theta}(\theta, r, v) = \frac{\partial h}{\partial r}(\theta, r, v)v. \tag{16}$$

In the same way, to display the relationship between functions y and Y as two variables, we define $E = J \times S \times S'$(where S is the range of y) and consider $f \in C^1(E)$; on the basis of $y = y(r, \theta)$, it is possible to express θ according to y implicitly. Therefore, function $f(\theta, y, Y)$ can basically be displayed as $f(\theta, Y)$. That is , y is not an independent variable of f; hence,

$\partial f/\partial y(\theta, Y) = (\partial f/\partial \theta(\theta, Y)).(\partial \theta/\partial y)$. Now, by regarding (14), we have $\partial f/\partial y(\theta, Y) = (\partial f/\partial \theta(\theta, Y)).(1/Y)$ [3]. Since this relationship is satisfied for every $f \in C^1(E)$, it can be concluded that:

$$\int_\Omega \frac{\partial f}{\partial y}(\theta, Y) r dr d\theta = \int_\Omega \frac{\partial f}{\partial \theta}(\theta, Y) \frac{1}{Y} r dr d\theta, \qquad \forall f \in C^1(E). \tag{17}$$

We add this set of constraints to (12) in order to specify the relationship between y and Y when they are considered as variables in the problem.

Since our aim is to identify ω and its unknown boundary, we prefer to display the constraints of (12) as integrals on the boundary of ω as much as we can. Thus, the term on the right-hand side of (6), we have:

$$\int_\Omega a(x) y(T) \varphi(T) dx - \int_\Omega a(x) y_0(x) \varphi(0) dx = \int_\omega a(y(T)\varphi(T) - y_0(x)\varphi(0)) dx;$$

Regarding the Green's theorem :

$$\oint_{\partial\omega} M dx_1 + N dx_2 = \int_\omega (\frac{\partial N}{\partial x_1} - \frac{\partial M}{\partial x_2}) dA,$$

suppose $M = 0$ and $N = \int_0^{x_1} a(y(T)\varphi(T) - y_0(x)\varphi(0)) dx_1$. Considering the fact that $a(x) = a\chi_\omega(x)$, we have:

$$\int_{\partial\omega} a \int_0^{rcos\theta} [y(T)\varphi(T) - y_0(r,\tau)\varphi(0)](\dot{r}cos\tau - rsin\tau)(\dot{r}sin\tau + rcos\tau) d\tau d\tau. \tag{18}$$

Also, the area constraint (9) can be presented as:

$$\frac{1}{2}\int_{\partial\omega} r^2 dr d\theta = L \int_\Omega r dr d\theta = L \frac{1}{T} \int_0^T \int_\Omega r dr d\theta.$$

The independency of the objective function from t results in:

$$E(\omega, a, T) = \frac{1}{2T} \int_0^T \int_\Omega [|\dot{y}(r,\theta,T)|^2 + (\frac{1}{v^2} + \frac{1}{r^2}) Y^2] r dr d\theta dt. \tag{19}$$

Now, by substituting (17) into (10) for every $\varphi \in H_0^1(\Omega \times (0,T))$, we have:

$\frac{1}{T}\int_0^T \int_\Omega \dot{y}(T)\varphi(T) r dr d\theta dt - \frac{1}{T}\int_0^T \int_\Omega y(T)\dot{\varphi}(T) r dr d\theta dt - \int_0^T \int_\Omega y \Delta\varphi r dr d\theta dt$
$+ \int_{\partial\omega} a \int_0^{rcos\theta} [y(T)\varphi(T) - y_0(r,\tau)\varphi(0)](\dot{r}cos\tau - rsin\tau)(\dot{r}sin\tau + rcos\tau) d\tau d\tau$
$- \int_0^T \int_\Omega a(r,\theta) y \dot{\varphi} r dr d\theta dt + \int_0^T \int_\Omega y \ddot{\varphi} r dr d\theta dt = \Phi.$

Integrating (11), (15) and (16) over $[0, T]$ implies that:

$$\int_0^T \int_\Omega div(y(r,\theta,t)\varphi(r,\theta,t))rdrd\theta dt = 0,\ \forall \varphi \in H_0^1(\Omega \times (0,T));$$
$$\frac{1}{T}\int_0^T \int_\Omega \frac{\partial f}{\partial y} rdrd\theta dt = \frac{1}{T}\int_0^T \int_\Omega \frac{\partial f}{\partial \theta}\frac{1}{Y} rdrd\theta dt,\ \forall f \in C^1(E);$$
$$\frac{1}{T}\int_0^T \int_\Omega \frac{\partial h}{\partial \theta}(\theta,r,v) rdrd\theta dt = \frac{1}{T}\int_0^T \int_\Omega \frac{\partial h}{\partial r}(\theta,r,v) vrdrd\theta dt,\ \forall h \in C^1(E').$$

Therefore, problem (12) can be displayed in a new variational form as follows:

$$\begin{aligned}
&Min: E(\omega,a,T) = \frac{1}{2T}\int_0^T \int_\Omega [|\dot{y}(r,\theta,T)|^2 + (\frac{1}{v^2}+\frac{1}{r^2})Y^2] rdrd\theta dt\\
&S.\ to: \frac{1}{2}\int_{\partial\omega} r^2 drd\theta = L\frac{1}{T}\int_0^T \int_\Omega rdrd\theta dt,\\
&\quad \int_0^T \int_\Omega div(y(r,\theta,t)\varphi(r,\theta,t))rdrd\theta dt = 0,\ \forall \varphi \in H_0^1(\Omega \times (0,T));\\
&\quad \frac{1}{T}\int_0^T \int_\Omega \frac{\partial f}{\partial y}(\theta,Y) rdrd\theta dt = \frac{1}{T}\int_0^T \int_\Omega \frac{\partial f}{\partial \theta}\frac{1}{Y}(\theta,Y) rdrd\theta dt,\ \forall f \in C^1(E);\\
&\quad \frac{1}{T}\int_0^T \int_\Omega \frac{\partial h}{\partial \theta}(\theta,r,v) rdrd\theta dt = \frac{1}{T}\int_0^T \int_\Omega \frac{\partial h}{\partial r}(\theta,r,v) vrdrd\theta dt,\ \forall h \in C^1(E');\\
&\quad \frac{1}{T}\int_0^T \int_\Omega \dot{y}(T)\varphi(T) rdrd\theta dt - \frac{1}{T}\int_0^T \int_\Omega y(T)\dot{\varphi}(T) rdrd\theta dt\\
&\quad - \int_0^T \int_\Omega y\Delta\varphi rdrd\theta dt - \int_0^T \int_\Omega a(r,\theta) y\dot{\varphi} rdrd\theta dt;\\
&\quad + \int_{\partial\omega} a \int_0^{rcos\theta} [y(T)\varphi(T) - y_0\varphi(0)](\dot{r}cos\tau - rsin\tau)(\dot{r}sin\tau + rcos\tau) d\tau d\tau\\
&\quad + \int_0^T \int_\Omega y\ddot{\varphi} rdrd\theta dt = \Phi,\ \forall \varphi \in H_0^1(\Omega \times (0,T)).
\end{aligned} \tag{20}$$

Now, by determining a suitable control function, we rewrite the problem in the form of an optimal control problem.

5.2 Step 2: Embedding into measure space

By considering the pair of functions (r, y) as the trajectory and the triple (u, v, Y) as the control vector, problem (25) can be considered as an optimal control problem. In this manner, we need to present the following definition:

Definition 4. Quintuplet $p = (r, u, v, y, Y)$ is called admissible when it satisfies the following conditions:

1. The control functions u, ν and Y are bounded and continuous and take their values on compact sets U, V and $S^{'}$;
2. r is a differentiable function and $r(0) = r(2\pi)$;
3. y is the bounded solution of the linear damped wave system (1);
4. The relations (15) and (16) are satisfied.

The set of all admissible quintuplets is denoted by P. We also suppose that P is nonempty; in other words, we suppose that the system is controllable (This can be seen in [20], for instance).

Let $D = [0,T]\times J\times A_1\times S\times S'\times V$ and $D' = J\times A_2\times U$; for any admissible quintuplets in P, we define the linear, positive and bounded functionals Λ_P and Γ_P on $C(D)$ and $C(D')$ in the following way:

$$\begin{aligned}\Gamma(F) &= \int_0^T \int_\Omega F(t,\theta,r,v,y,Y)drd\theta dt, \qquad \forall F\in C(D);\\ \Lambda(G) &= \int_{\partial\omega} G(\theta,r,u)d\theta, \qquad \forall G\in C(D^{'});\end{aligned} \tag{21}$$

Since $\mathbb{R}^6$ is a locally compact space, by the Heine-Borel theorem ([21]), $D\subseteq\mathbb{R}^6$ is a compact Hausdorff space. Also, for the same reason, D' is a Hausdorff compact space. Therefore, for every given p, Riesz representation theorem ([22]) indicates uniquely two positive Radon measures , μ_P and λ_P, so that:

$$\begin{aligned}\Gamma_P(F) &= \int_D F d\mu_P \equiv \mu_P(F), \qquad \forall F\in C(D);\\ \Lambda_P(G) &= \int_{D'} G d\lambda_P \equiv \lambda_P(G), \qquad \forall G\in C(D^{'});\end{aligned} \tag{22}$$

Consequently, any admissible quintuplets can be displayed as (27) by a unique pair of measures, say (μ_P,λ_P), in a subset F of $\mathcal{M}^+(D)\times\mathcal{M}^+(D')$, where $\mathcal{M}^+(X)$ is the set of all positive Radon measures on X. Therefore, one can transfer problem (25) into a measure space by:

$$(r,u,v,y,Y)\in P \longmapsto (\mu_P,\lambda_P)\in\mathcal{M}(D)\times\mathcal{M}(D').$$

It was proved by Rubio (1986) that such a transformation is an injection. To achieve something new, we enlarge the underlying space and consider the problem of finding a minimizer pair of measures, say (μ^*,λ^*), on the space of all positive related Radon measures which are just satisfied to the conditions of (25) (Not just those that are induced from Riesz Representation theorem). Therefore, our method is somehow global.

We now characterize some properties of admissible pairs. Suppose B is an open disc in $\mathbb{R}^2$ that includes $J\times A_2$; consider $C'(B)$ as the space of real-valued continuously differentiable functions on B. Then, for every ϕ in $C'(B)$, we define:

$$\phi^g(\theta,r,u) = \phi_r(\theta,r)u + \phi_\theta(\theta,r),\ \forall(\theta,r,u)\in D'.$$

Then, since the boundary $\partial\omega$ is a closed and simple curve, we have:

$$\int_J \phi^g(\theta,r,u)d\theta = \int_J \dot{\phi}(\theta,r)d\theta = d_\phi, \quad \forall\phi\in C'(B), \tag{23}$$

where $d_\phi = \phi(2\pi,r_d)-\phi(0,r_d)$, is still unknown since r_d in $(0,r_d)=(2\pi,r_d)$, which is the initial and final point of the closed curve ω, is unknown. We will explain later that it would be characteristic (see Section (6)).

Let $D(J^0)$ be the space of infinitely differentiable real-valued functions with compact support in J^0. Define

$$\psi^g(\theta,r,u)=r(\theta)\dot{\psi}(\theta)+u(\theta)\psi(\theta),\ \forall(\theta,r,u)\in D';$$

then,

$$\int_J \psi^g(\theta,r,u)d\theta=r(2\pi)\psi(2\pi)-r(0)\psi(0)=0,\quad \forall\psi\in D(J^o). \tag{24}$$

The same situation arises for another special choice of functions in $C'(B)$ which are only dependent on variable θ, denoted by $C_1(D^{'})$. Thus, for $\phi(\theta,r,u)\equiv\nu(\theta)$, we can have:

$$\int_J \nu(\theta)d\theta=a_\nu,\ \forall\nu\in C_1(D'), \tag{25}$$

where a_ν is the Lebesgue integral of ν over J.

Regarding the famous properties of admissible quintuplets in P which are looked at in (28), (29) and (25), and the definitions of the pair of measures (μ,λ) in (21), problem (25) can now be displayed as follows in which the measures λ and μ are its unknown variables:

$$\begin{aligned}
&Min:\quad E(\mu,\lambda)=\mu(\tfrac{r}{2T}[|\dot{y}(r,\theta,T)|^2+(\tfrac{1}{v^2}+\tfrac{1}{r^2})Y^2]\\
&S.\ to:\ \lambda(\phi^g(\theta,r,u))=d_\phi,\ \forall\phi\in C'(B);\\
&\lambda(\psi^g(\theta,r,u))=0,\ \forall\psi\in D(J^o);\\
&\lambda(\nu(\theta))=a_\nu,\ \forall\nu\in C_1(D');\\
&\lambda(\tfrac{1}{2}r^2)=L\mu(\tfrac{1}{T}r),\\
&\mu(\tfrac{1}{T}\tfrac{\partial f}{\partial y}r)=\mu(\tfrac{1}{T}\tfrac{\partial f}{\partial\theta}\tfrac{1}{Y}r),\ \forall f\in C^1(E);\\
&\mu(\tfrac{1}{T}\tfrac{\partial h}{\partial\theta}(\theta,r,v)r)=\mu(\tfrac{1}{T}\tfrac{\partial h}{\partial r}(\theta,r,v)vr),\ \forall h\in C^1(E');\\
&\mu(div(y(r,\theta,t)\varphi(r,\theta,t))r)=0,\ \forall\varphi\in H_0^1(\Omega\times(0,T));\\
&\mu(\tfrac{1}{T}\dot{y}(T)\varphi(T)r)-\mu(\tfrac{1}{T}y(T)\dot{\varphi}(T)r)-\mu(y\Delta\varphi r)-\mu(a(r,\theta)y\dot{\varphi}r)+\mu(y\ddot{\varphi}r)\\
&+\lambda(a\int_0^{rcos\theta}[y(T)\varphi(T)-y_0\varphi(0)](\dot{r}cos\tau-rsin\tau)(\dot{r}sin\tau+rcos\tau)d\tau)=\Phi;
\end{aligned} \tag{26}$$

We remind that the theoretical measure problem (26) is linear even though the initial problem is highly nonlinear.

The space $M^+(D)\times M^+(D')$ is a linear space which will become a locally convex topological vector space when it gives the weak*topology. This can be defined by the family of semi-norms $(\mu,\lambda)\mapsto|\mu(F)|+|\lambda(G)|$ for $F\in C(D)$, $G\in C(D')$ and $\epsilon>0$, which can be on the basis of a family of neighborhoods of zero for $M^+(D)\times M^+(D')$. This family is defined by:

$$U_\epsilon=\{(\mu,\lambda)\in M^+(D)\times M^+(D'):|\mu(F_j)|+|\lambda(G_j)|<\epsilon;j=1,2,...,r\}$$

and makes a basis for a weak*topology on the space $M^+(D)\times M^+(D')$ (Many properties of this topology can be found in the literature such as [24]). In this way, $M^+(D) \times M^+(D')$ under this topology is a Hausdorff space ([21]).

The proof of the following theorems can be found in [6], [7] and [20].

Theorem 1.
a) $Q \subseteq M^+(D) \times M^+(D')$ *is compact under the weak* topology on* $M^+(D) \times M^+(D')$
b) The objective function $E(\mu, \lambda)$ *in problem (26) is continuous.*
c) There exists a pair of measures (μ^*, λ^*) *which are optimal for (26) in the set* $Q \subset M^+(D) \times M^+(D')$*; that is, for every* $(\mu, \lambda) \in Q$*, we have:*

$$E(\mu^*, \lambda^*) \leq E(\mu, \lambda).$$

Even though (26) has an optimal solution in Q, it is still very difficult to obtain the exact solution since the underlying spaces are not finite-dimensional, the number of equations is not finite and the unknowns are measures. Therefore, it is completely acceptable to seek for a suboptimal solution. Thus, first, by choosing suitable dense subsets in the appropriate spaces and then, by choosing a finite number of them, the problem is approximated by a semi-finite linear programming.

5.3 Identifying a nearly optimal solution

It is possible to approximate the solution of problem (26) by the solution of a finite-dimensional linear one of sufficiently large dimensions. Besides, by increasing the dimension of the problem, the accuracy of the approximation can be increased. First, we consider the minimization of (26) not only over set Q, but also over its subset called $Q(M_1, M_2, ..., M_7)$ and defined by only a finite number of constraints to be satisfied. This will be achieved by choosing countable sets of functions whose linear combinations are dense in appropriate spaces and then by selecting a finite number of constraints. Let $\{\phi_i : i \in N\}$, $\{\psi_i : i \in N\}$,$\{\nu_i : i \in N\}$,$\{\varphi_i : i \in N\}$, $\{f_i : i \in N\}$ and $\{h_i : i \in N\}$ be countable dense (in the topological convergence sense) sets in spaces $C'(B)$, $D(J^o)$, $C_1(D')$, $H_0^1(\Omega \times (0,T))$, $C^1(E)$ and $C^1(E')$, respectively. By choosing a finite number of functions in each set, the solution of (29) can be approximated by the solution of the following one:

$$
\begin{aligned}
Min: \;& E(\mu,\lambda)=\mu(\tfrac{r}{2T}[|\dot{y}(r,\theta,T)|^2+(\tfrac{1}{v^2}+\tfrac{1}{r^2})Y^2])\\
S.\;\; to: \;& \lambda(\phi_k^g(\theta,r,u))=d_{\phi_k},\;\; k=1,2,...,M_1;\\
& \lambda(\psi_l^g(\theta,r,u))=0,\;\; l=1,2,...,M_2;\\
& \lambda(\nu_s(\theta))=a_s,\;\; s=1,2,...,M_3;\\
& \lambda(\tfrac{1}{2}r^2)-L\mu(\tfrac{1}{T}r)=0,\\
& \mu(F_i)=0,\;\; i=1,2,...,M_4;\\
& \mu(G_j)-\mu(H_j)=0,\;\; j=1,2,...,M_5;\\
& \mu(I_j)-\mu(K_j)=0,\;\; j=1,2,...,M_6;\\
& \mu(L_i)-\mu(P_i)-\mu(Q_i)-\mu(R_i)+\lambda(T_i)+\mu(N_i)=\Phi_i,\\
& \qquad\qquad i=1,2,...,M_7,
\end{aligned}
\tag{27}
$$

where

$$
\begin{cases}
F_i=div(y(r,\theta,t)\varphi_i(r,\theta,t))r, \quad G_j=\frac{1}{T}\frac{\partial f_j}{\partial y}r,\\
H_j=\frac{1}{T}\frac{\partial f_j}{\partial \theta}\frac{1}{Y}r, \quad I_j=\frac{1}{T}\frac{\partial h_j}{\partial \theta}(\theta,r,v)r,\\
K_j=\frac{1}{T}\frac{\partial h_j}{\partial r}(\theta,r,v)vr, \quad L_i=\frac{1}{T}\dot{y}(T)\varphi_i(T)r,\\
P_i=\frac{1}{T}y(T)\dot{\varphi}_i(T)r, \quad Q_i=y\Delta\varphi_i r,\\
R_i=a(r,\theta)y\dot{\varphi}_i r, \quad N_i=y\ddot{\varphi}r,\\
T_i=a\int_0^{rcos\theta}[y(T)\varphi_i(T)-y_0(r,\tau)\varphi_i(0)](\dot{r}cos\tau-rsin\tau)(\dot{r}sin\tau+rcos\tau)d\tau.
\end{cases}
$$

The density property of the selected sets in (27) causes its solution to tend to the solution of (26) when $M_1, M_2, ..., M_7 \to \infty$; thus, if numbers $M_1, ..., M_7$ are selected large enough, (27) is a good approximation of our main problem. Now, the number of constraints of the problem is finite, but the problem is still infinite since the underlying space is a subspace of measures. It would be more convenient if we could approximate the solution just by a solution of a simple finite LP. This is precisely our main attention.

Fakharzadeh et al. (1999) presented that the pair of the optimal measures of (25) are in the form of $\lambda^* = \sum_{m=1}^{M} \beta_m^* \delta(z_m^*)$ and $\mu^* = \sum_{n=1}^{N} \alpha_n^* \delta(Z_n^*)$ in which Z_n^* and z_m^* belong to dense subsets of D and D', respectively; moreover, $\delta(t)$ is a unitary atomic measure with support at the singleton set t. Substituting these forms in (27), it might seem that the problem has been made even more difficult, since, it is transformed into a non-linear one. But, if function $E(\mu,\lambda)$ can be minimized only with respect to the coefficients α_n^* and β_m^*, it will be turned to a linear programming problem. In other words, the solution can be obtained approximately by solving just the simple finite linear programming like below. If one chooses the points Z_n^* and z_m^* from the dense subsets of D and D', this fact could be achieved in the second step of our approximation. (see [6] for more details):

$$
\begin{aligned}
Min:\ & E(\alpha,\beta,a,r_{d_a}) = \textstyle\sum_{n=1}^{n=N} \alpha_n \Theta(Z_n) \\
S.\ to:\ & \textstyle\sum_{m=1}^{m=M} \beta_m \phi_k^g(z_m) = d_{\phi_k}, && k = 1,2,...,M_1; \\
& \textstyle\sum_{m=1}^{m=M} \beta_m \psi_l^g(z_m) = 0, && l = 1,2,...,M_2; \\
& \textstyle\sum_{m=1}^{m=M} \beta_m \nu_s(z_m) = a_s, && s = 1,2,...,M_3; \\
& \textstyle\sum_{m=1}^{m=M} \beta_m \frac{1}{2} r_m^2 - (\frac{1}{T}) L \sum_{n=1}^{n=N} \alpha_n r_n = 0, \\
& \textstyle\sum_{n=1}^{n=N} \alpha_n F_i(Z_n) = 0, && i = 1,2,...,M_4; \\
& \textstyle\sum_{n=1}^{n=N} \alpha_n G_j(Z_n) - \sum_{n=1}^{n=N} \alpha_n H_j(Z_n) = 0, && j = 1,2,...,M_5; \\
& \textstyle\sum_{n=1}^{n=N} \alpha_n I_j(Z_n) - \sum_{n=1}^{n=N} \alpha_n K_j(Z_n) = 0, && j = 1,2,...,M_6; \\
& \textstyle\sum_{n=1}^{n=N} \alpha_n [L_i(Z_n) - P_i(Z_n) - Q_i(Z_n) - R_i(Z_n) + N_i(Z_n)] \\
& + \textstyle\sum_{m=1}^{m=M} \beta_m T_i(z_m) = \Phi_i, && i = 1,2,...,M_7; \\
& \alpha_n \geq 0, \qquad n = 1,2,...,N; \\
& \beta_m \geq 0, \qquad m = 1,2,...,M.
\end{aligned}
\tag{28}
$$

Here, we defined $\Theta(Z) = (r/2T)[|\dot{y}(r,\theta,T)|^2 + (1/v^2 + 1/r^2)Y^2]$. Problem (28) is still non-linear because R_i and T_i are functions of the damping coefficient and d_{φ_k} is unknown since the constant point $(0, r_{d_a})$ of ω is unknown. Now, by using simultaneous three-phase search techniques for (28), the optimal damping coefficient, $r(0) = r(2\pi) = r_{d_a}$, and the optimal coefficients $\alpha_1^*, ..., \alpha_N^*, \beta_1^*, ..., \beta_M^*$ would be found as explained in next section. Thus, one is able to construct the pair of optimal shape and control function in the manner which has been explained in ([8], [7]).

6 Algorithm

To apply the mentioned method for solving problem (1) practically, here we present an algorithmic path for the solution procedure. Regarding the previous statements, we are able to identify the optimal control and optimal region by using the following 4 steps algorithm:

Step 1: The given sets $[0,T], J, A_1, S, S^{'}$ and V which form Ω are divided into n_1, n_2, n_3, n_4, n_5 and n_6 equal parts, and the sets J, A_2 and U which form ω are divided into m_1, m_2 and m_3 equal parts, respectively; so that, the $N = n_1.n_2.n_3.n_4.n_5.n_6$, the number of 6-dimensional cells, and the $M = m_1.m_2.m_3$, the number of 3-dimensional cells in the related spaces are obtained. Then, in each of these 6-dimensional and 3-dimensional cells arbitrary points $Z_i = (t_i, \theta_i, r_i, y_i, Y_i, \nu_i)$ and $z_j = (\theta_j, r_j, u_j)$ are selected respectively.

Step 2: For fixed numbers M_1, M_2, M_3, we select M_1 number of $\phi_k^g(z)$, M_2 of $\psi_l^g(z)$, M_3 of $\nu_s(z)$, and for fixed numbers M_4, M_5, M_6, M_7, we define M_4 number of $F_i(Z)$, M_5 of $G_j(Z)$ and $H_j(Z)$, M_6 of I_j and K_j and M_7 of L_i, P_i, Q_i, R_i, N_i and T_i functions, respectively. Now, one is able to set up the finite linear programming (28) with $N+M$ variables and $M_1+M_2+...+M_7$ constraints, which is dependent on the variables a and r_d.

Step 3: To solve problem (1), we use an iterative method with two inner loops and apply a three-phase optimization approach. In the first loop by giving a fixing the amount of a, the function $J_{1a} : [0,1] \longrightarrow \mathbb{R}$ defined by $J_{1a}(r_d) = E^*(\alpha, \beta, r_d)$ is set up. Then, in the second loop, this function is minimized by the use of a standard minimization technique (like a line search method) as one of the optimization approaches. We remind that in each function, in calculating the standard minimization technique, its related LP (28) should be solved (one of the optimization phases).

If the minimizer of J_{1a} is called $r_{d_a}^*$ with the optimal value $J_{1a}^*(r_{d_a}) \equiv E^*(\alpha, \beta, a, r_{d_a}^*)$, one is able to set up the function $J_2 : [0,1] \to \mathbb{R}$ by $J_2(a) = J_{1a}^*(r_{d_a}^*)$ in the first loop, use a search technique as the last phase of the optimization approach, determine the optimal damping coefficient, say a^*, with the optimal value of energy $J_2^*(a^*) \equiv E^*(\alpha, \beta, a^*, r_{d_a}^*)$. In this manner, the damping coefficient, the value of the energy of the system and the coefficients α's and β's are simultaneously and optimally determined.

Remark 1. In each stage where alternative optimal cases happen, it suffices to select one arbitrarily.

Step 4: Regarding [20], [8] and [7], for the optimal values $\alpha_1^*, \alpha_2^*, ..., \alpha_N^*$, $\beta_1^*, \beta_2^*, ..., \beta_n^*$ obtained from Step 3, the optimal control and the optimal region are determined through the following instructions:

i) Let $\theta_0 = 0$ and $\theta_i = \theta_{i-1} + \beta_i^*$ for $i = 1, 2, ..., M$.
ii) For $\theta \in [\theta_{i-1}, \theta_i)$, set $u^*(\theta) = u_i$, where, u_i is the related component associated with point z_i. In this manner, according to [20], the nearly optimal control can be constructed as a piecewise constant function.
iii) Let $r_0 = r_{2\pi} = r_d$, using the differential equation $u(\theta) = \dot{r}(\theta)$, we take the following difference equation:

$$r_i = r_{i-1} + (\theta_i - \theta_{i-1}).u_i, \qquad i = 1, 2, .., M.$$

Therefore, M number points (θ_i, r_i), $i = 1, 2, ..., M$ of the nearly optimal region are determined. Using curve fitting or connecting them by line segments, we demonstrate the approximated optimal region.

Theorem 2. *If the used minimization techniques used in Step 3 of the*

above algorithm are convergent, then the algorithm converges to the optimal solution of (1) when $M, N, M_1, M_2, ..., M_7$ *tend to infinity.*

Proof. The proof of this theorem is given in Appendix A.

7 Examples

Now, to show the efficiency of our method and to explain how it works, we solve two numerical examples. It is worth mentioning that these examples are taken from [17] and [15] as well as from other studies cited by them in order for the readers to be able to compare the two methods.

Example 1: By defining $\Omega = (0,1) \times (0,1)$ and selecting $a = 10$ (constant), problem (1) was solved by Munch (2009) by the level set method. To apply this method, the author used the gradient descend method and also supplied some necessity relations by applying the finite difference method. In this manner, he used an initial shape to determine the optimal solution. We must mention that this approach is very time-consuming and the resulting optimal shape is also dependent on the number of iterations as shown on Page 25 of [17]. As it is also mentioned in Sub-section 5.1.2 Page 24, the results of the problem for variable a, depend on the initial shape. Moreover, in this case, the local minima have been obtained which are also completely dependent on the initial shape (Page 34 of [17]).

For the chosen initial conditions:

$$\begin{cases} y_0(\theta, r) = 100 sin(\pi r cos\theta) sin(\pi r sin\theta), \\ y_1(\theta, r) = 0, \quad (\theta, r) \in \Omega = J \times A_1, \end{cases}$$

The optimal value obtained by Munch (2009) for $a = 10$ and $T = 2$ was mentioned as $E(\omega, a, T) = 88.17$, and for $a = 10$ and $T = 1$, as $E(\omega, a, T) = 249.10$.

We considered the same condition as above, and additional conditions that are needed for our method as follows:

$$y(T = 1) = \dot{y}(T = 1) = 1.$$

We supposed $L = 0.11164$, the area of the unknown region ω was equal to 0.7 and $(0, r_{d_a})$ was a boundary point of ω which was determined optimally in domain Ω as mentioned in Step 3. Then, by selecting the following functions and setting them in (28) for $M_1 = 2, M_2 = 10, M_3 = 10, M_4 = M_5 = M_6 = M_7 = 2$, we set up the corresponding LP with:

$$\phi_1^g(\theta, r, u) = 2r\theta u + r^2; \qquad \phi_2^g(\theta, r, u) = 2ru;$$

$$\psi_l(\theta) = sin(l\theta),\ l = 1, 2, ..., 5;\ \psi_{l'}(\theta) = (1 - \cos(l'\theta)),\ l' = 1, 2, ..., 5;$$
$$J_s = [\frac{2\pi(s-1)}{10}, \frac{2\pi s}{10}];\ a_s = \int_{J_s} d\theta = \frac{2\pi}{10},\ s = 1, 2, ..., 10.$$

The functions introduced in (27), which expressed the relationship between y and Y and also between r and ν, were determined as below:

$$F_1(\theta, y, Y) = \theta y, \quad F_2(\theta, y, Y) = \theta^2 yY, \quad G_1(\theta, y, Y) = r\theta,$$
$$G_2(\theta, y, Y) = \theta^2 rY, \quad H_1(\theta, y, Y) = \frac{ry}{Y}, \quad H_2(\theta, y, Y) = 2r\theta y,$$
$$h_1(\theta, r, v) = \theta r, \quad h_2(\theta, r, v) = \theta rv, \quad I_1(\theta, r, v) = r^2,$$
$$I_2(\theta, r, v) = r^2 v, \quad K_1(\theta, r, v) = \theta rv, \quad K_2(\theta, r, v) = r\theta v^2.$$

Also, for $i = 1, 2$ we selected $\varphi_i = r^i sin(i\theta)t$; therefore:

$$\begin{cases} \varphi_i(T) = r^i sin(i\theta), \dot{\varphi}_i = r^i sin(i\theta), \\ \ddot{\varphi}_i = 0, \\ \Delta\varphi_i = \frac{\partial^2\varphi_i}{\partial r^2} + \frac{1}{r}\frac{\partial\varphi_i}{\partial r} + \frac{1}{r^2}\frac{\partial^2\varphi_i}{\partial\theta^2} \\ \qquad = (i(i-1) + i - i^2)r^{i-2} sin(i\theta)t = 0. \end{cases}$$

Thus the functions in (27) were illustrated as:

$$\begin{cases} L_i = r^{i+1} sin(i\theta), \\ P_i = r^{i+1} sin(i\theta), \\ Q_i = 0, \\ R_i = ayr^{i+1} sin(i\theta) \\ T_i = a[\frac{(-1)(r^{i+1}u)}{2}(\frac{2i}{i^2-4} - (\frac{\cos((i+2)rcos\theta)}{i+2} + \frac{\cos((i-2)rcos\theta)}{i-2}))+ \\ \frac{(r^{i+2}+u^2r^i)}{4}(\frac{sin(i-2)rcos\theta}{i-2} - \frac{sin(i+2)rcos\theta}{i+2})] \\ N_i = 0; \end{cases}$$

since:

$$\begin{aligned} T_i &= a\int_0^{rcos\theta}[y(T)\varphi_i(T) - y_0(r,\tau)\varphi_i(0)](\dot{r}sin\tau + rcos\tau)(\dot{r}cos\tau - rsin\tau)d\tau \\ &= a\int_0^{rcos\theta}[r^i sin(i\tau)](u\sin\tau + r\cos\tau)(u\cos\tau - r\sin\tau)d\tau \\ &= a\int_0^{rcos\theta}[r^i sin(i\tau)](ru(\cos 2\tau) + (u^2 - r^2)(\cos\tau)(\sin\tau))d\tau \\ &= a\int_0^{rcos\theta}[(r^{i+1}u)sin(i\tau)\cos 2\tau + (u^2 r^i - r^{i+2})(\sin(i\tau)\cos\tau\sin\tau)]d\tau \\ &= a([(\frac{-1}{2})r^{i+1}u(\frac{\cos((i+2)\tau)}{i+2} + \frac{\cos((i-2)\tau)}{i-2})] + \frac{((ur)^i - r^{i+2})}{4}[\frac{\sin(i-2)\tau}{i-2} - \frac{\sin(i+2)\tau}{i+2}])|_0^{r\cos\theta}. \end{aligned}$$

By dividing each of intervals $[0, 1]$ and $[0, 2\pi]$ into ten, $A_1 = [0, \sqrt{2}]$ and $S = [-2, 2]$ into five, $S' = [-2, 2]$ into four, $V = [-1, 1]$ and $A_2 = [0, 1]$ into ten and $U = [-0.6, 0.6]$ into eleven equal parts, we selected $N = 10^5$ points

of Z_i and $M = 1100$ points of z_i; thus, LP (28) was set up with $M + N$ variables and 30 constraints:

$$
\begin{aligned}
&Min : E(\alpha, \beta, a, r_{d_a}) = \textstyle\sum_{n=1}^{n=N} \alpha_n \frac{r_n}{2}[1 + (\frac{1}{v_n} + \frac{1}{r_n})Y_n^2] \\
&S.\ to : \textstyle\sum_{m=1}^{m=M} \beta_m(2r_m\theta_m u_m + r_m^2) = 2\pi(r_{d_a}^2); \\
&\textstyle\sum_{m=1}^{m=M} \beta_m 2r_m u_m = r_{d_a}^2; \\
&\textstyle\sum_{m=1}^{m=M} \beta_m(lr_m cos(l\theta_m) + u_m sin(l\theta_m)) = 0,\ l = 1, 2, ..., 5; \\
&\textstyle\sum_{m=1}^{m=M} \beta_m(l' r_m sin(l'\theta_m) + u_m(1 - cos(l'\theta_m))) = 0,\ l' = 1, 2, ..., 5; \\
&\textstyle\sum_{m=1}^{m=M} \frac{1}{2} r_m^2 \beta_m = 0.7; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(r_n^2 - \theta_n r_n v_n) = 0; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(r_n^2 v_n - \theta_n r_n v_n^2) = 0; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(r_n\theta_n - \frac{r_n y_n}{Y_n}) = 0; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(r_n\theta_n^2 Y_n - 2_n r_n y_n) = 0; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(-a y_n r_n^4 sin(3\theta_n)) + \sum_{m=1}^{m=M} \beta_m(\frac{-1}{2} a r_m^4 u_m)[\frac{6}{5} - (\frac{\cos(5r_m cos\theta_m)}{5} \\
&\qquad + \cos(r_m cos\theta_m))]\tfrac{a(r_m^5 + u_m^2 r_m^3)}{4}[(sin(r_m cos\theta_m) - \tfrac{(sin(6r_m cos\theta_m)}{6}] \\
&\qquad = \textstyle\int_0^{2\pi}\int_0^{\sqrt{2}} (-r^2 \sin(2\theta)\sin(\pi r cos\theta)\sin(\pi r sin\theta) dr d\theta; \\
&\textstyle\sum_{n=1}^{n=N} \alpha_n(-a y_n r_n^5 sin(4\theta_n)) + \sum_{m=1}^{m=M} \beta_m(\frac{-1}{2} a r_m^5 u_m)[\frac{2}{3} - \frac{(\cos(6r_m cos\theta_m)}{6} \\
&\qquad + \tfrac{\cos(2r_m cos\theta_m)}{2})] + \tfrac{a(r_m^6 + u_m^2 r_m^4)}{4}(\tfrac{1}{2}[sin(2r_m cos\theta_m) - \tfrac{(sin(5r_m cos\theta_m)}{5}] \\
&\qquad = \textstyle\int_0^{2\pi}\int_0^{\sqrt{2}} (r^3 \sin(3\theta)\sin(\pi r cos\theta)\sin(\pi r sin\theta) dr d\theta; \\
&\beta_1 + \beta_2 + ... + \beta_{110} = \tfrac{2\pi}{10}; \\
&\vdots \\
&\beta_{991} + \beta_{992} + ... + \beta_{1100} = \tfrac{2\pi}{10}; \\
&\alpha_n \geq 0, \quad n = 1, 2, ..., N;\ \beta_m \geq 0, \quad m = 1, 2, ..., M.
\end{aligned}
\tag{29}
$$

For a fixed $a = 10$, using HBM (Honey-Bee-Method) (see [1]) and the modified Simplex method from MATLAB 7.6, we obtained the nearly optimal artificial control (see Figure 1), shape (Figure 2), point $(0, r_{d_a}^* = 0.35)$ and the also energy value as 175.9.

Comparing with [17], we found that the optimal value of energy obtained by using shape measure method was less, while the obtained optimal regions were mostly approximately the same. Additionally, our method took less time and the obtained optimal region was independent from the number of iterations.

Example 2: In this example, we obtained the optimal value of a and ω simultaneously. As mentioned by Munch et al. (2006), in spite of the initial conditions being symmetrical, for large values of a, the obtained optimal domain might be non-symmetrical but the value of energy would be less.

Consider the two-dimensional damped wave equation (1) which is expressed on a cyclic domain with center at origin and radius as $\sqrt{2}$ in time interval $[0, 1]$. The aim was to obtain the optimal region ω, with a known area in the circle such that energy of the system was minimized in the final time. The conditions were given same as Example 1, while a was variable and supposed to be optimally determined.

With the same action as Example 1 to set up the related linear programming (29), the random search method was applied to obtain the optimal value of J_{1a}. Also, the optimal value of function J_2 was determined by using the HBM. In this manner, the obtained results were as follows:

The optimal damping coefficient $a^* = 24.3624$, $r^*_{d_a} = 0.5469$ and the value of the optimal objective function was 0.101. The optimal (artificial) control and the optimal region ω are shown in Figures 3 and 4.

In Section 5.1.4 of [17] for $T = 1$, $a = 29.09$ and by 2000 iterations, the optimal value of energy was given as 12.56. As emphasized there, for variable a, the optimal domain is completely dependent on the initial shape. In our method, despite $a's$ being variable, the obtained optimal region was independent of the initial shape and the amount of optimal energy was considerably less, while the time consumed also decreased.

8 Conclusion

By doing an embedding process and using the property of positive Radon measures, we presented a new and very useful technique for solving the problem of minimizing the energy of a damped wave system in an unknown region. In this method, the problem was solved by a three-phase optimization search technique where the unknown damping coefficient, the region and a point of its boundary were found optimally. This method has some advantages in comparison to the method used by Munch (2009), since we did not face the difficulties mentioned there; such as level set functions being flat, divergence of the systems with respect to dispersion, and the tendency of time toward infinity when damping of numerical waves approaches zero. The most important characteristic of our shape measure method is its simplicity and its independence from the solution of the initial shape. To obtain the optimal domain, we just need to use three search techniques while solving linear programming problems. Additionally, it is necessary to emphasize that this method is much easier, linear and less time-consuming.

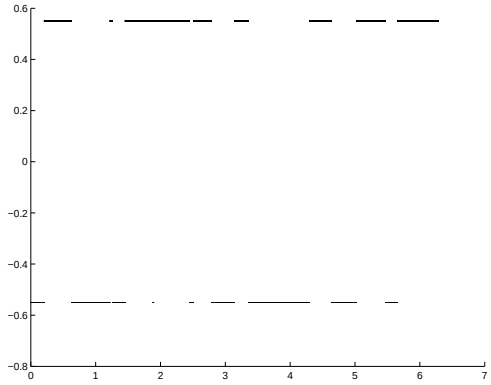

Figure 1: Optimal control function in Example 1

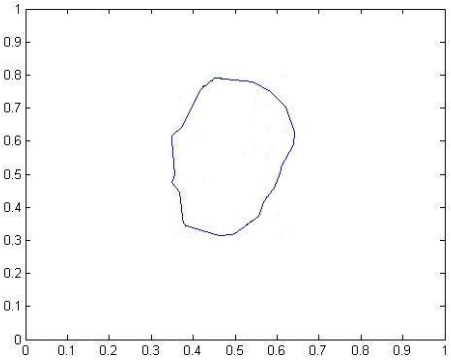

Figure 2: Optimal domain with constant damping coefficient $a = 10$ in Example 1

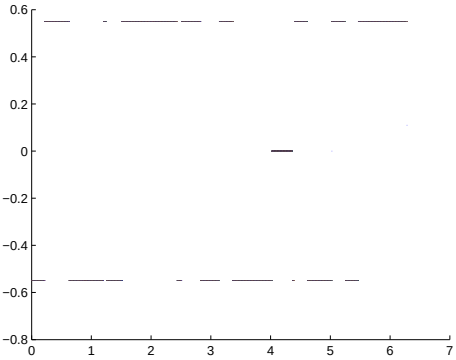

Figure 3: Optimal control function in Example 2

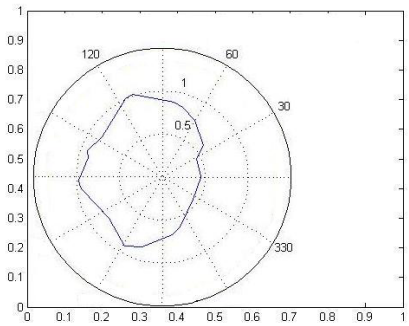

Figure 4: Optimal region in the given domain Ω in Example 2

Appendix A. Proof of Theorem 2

To prove this theorem, first, we present the two following lemmas.

Lemma 1. Consider the linear program (27) consisting of minimizing the function $\mu \to \mu(\Theta) \equiv E(\mu, \lambda)$ over the set $Q(M_1, ..., M_7)$ of measures in $M^+(D) \times M^+(D^{'})$ satisfying conditions (27). When $M_1, M_2, ..., M_7$ tend to infinity,

$$\eta(M_1, ..., M_7) \equiv \inf_{Q(M_1,...,M_7)} E(\mu, \lambda)$$

tends to $\eta = \inf_Q \ E(\mu, \lambda)$. (This lemma is an extension of Proposition III.1 by Rubio (1986)).

Proof. (i) We prove, first, that the sequence $\{\eta(M_1, M_2, ..., M_7)\}$ is convergent when $M_1, ..., M_7$ tend to infinity; consider, first, the subsequence of $\eta(M_1, M_2, ..., M_7)$ in form $\{\eta(M_1, M_1, ..., M_1) : M_1 = 1, 2, ...\}$. Since $Q(M_1, ..., M_1)$ is a subset of positive Radon measures that satisfies in constraint (26). Therefore, since

$$Q(1, 1, ..., 1) \supset Q(2, 2, ..., 2) \supset Q(3, 3, ..., 3) \supset ... \supset Q(M_1, ..., M_1) \supset ... \supset Q,$$
$$then, \eta(1, 1, ..., 1) \leq \eta(2, 2, ..., 2) \leq ... \leq \eta(M_1, ..., M_1) \leq ... \leq \eta.$$

This sequence is non decreasing and bounded above and hence it converges to a number $\zeta \leq \eta$; thus, if $\epsilon > 0$, for $M_1 > N(\epsilon)$, we have:

$$|\eta(M_1, M_1, ..., M_1) - \zeta| < \epsilon \tag{30}$$

consider now $\eta(M_1, M_2, M_1, ..., M_1)$ for both M_1 and M_2 larger than $N(\epsilon)$. Without loss of generality, assume that $M_1 > M_2$. Then:

$$\eta(M_2, M_2, ..., M_2) \leq \eta(M_1, M_2, M_1, ..., M_1) \leq \eta(M_1, M_1, ..., M_1);$$

therefore,

$$\eta(M_2, M_2, ..., M_2) - \zeta \leq \eta(M_1, M_2, M_1, ..., M_1) - \zeta \leq \eta(M_1, M_1, ..., M_1) - \zeta,$$

and according to (30), we have:

$$|\eta(M_1, M_2, ..., M_1) - \zeta| \leq \epsilon. \tag{31}$$

Now consider $\eta(M_1, M_2, M_3, M_1, ..., M_1)$ for $M_1 \geqslant M_2 \geqslant M_3 \geqslant N(\epsilon)$. Then, by the same procedure, one could show that:

$$|\eta(M_1, M_2, M_3, M_1, ..., M_1) - \zeta| \leq \epsilon. \tag{32}$$

In a similar manner, for $M_1 > M_2 > M_3 > M_4 \geq N(\epsilon)$, we have:

$$\begin{aligned}
&\eta(M_4, M_4, ..., M_4) \leq \eta(M_1, M_4, M_4, M_4, M_1, ..., M_1) \\
&\leq \eta(M_1, M_2, M_4, M_4, M_1, ..., M_1) \leq \eta(M_1, M_2, M_3, M_4, M_1, ..., M_1) \\
&\leq \eta(M_1, M_2, M_3, M_1, M_1, ..., M_1),
\end{aligned}$$

and by using (30) and (32), we have:

$$|\eta(M_1, M_2, M_3, M_4, M_1, ..., M_1) - \zeta| \leq \epsilon.$$

Finally, in a similar way, for $M_1 \geq M_2 \geq ... \geq M_7 \geq N(\epsilon)$, one can show that:

$$\eta(M_7, M_7, ..., M_7) \leq \eta(M_1, M_2, ..., M_7) \leq \eta(M_1, M_2, ..., M_6, M_1),$$

and hence:

$$|\eta(M_1, M_2, ..., M_7) - \zeta| \leq \epsilon.$$

Thus, the sequence $\{\eta(M_1, M_2, ..., M_7), M_1 = 1, 2, ..., M_7 = 1, 2, ...\}$ converges to the number ζ as $M_1, ..., M_7$ tend to infinity.
(ii) We must prove now that the limit ζ equals $\eta = \inf_Q \quad E(\mu, \lambda)$.
We, first, show that this limit ζ can be computed sequentially. It is known that

$$\zeta = \lim_{M_1 \to \infty}[\lim_{M_2 \to \infty}[...[\lim_{M_7 \to \infty} \eta(M_1, ..., M_7)]]...],$$

provided that $\lim_{M_7 \to \infty} \eta(M_1, ..., M_7)$ exist since ζ is a finite number. To show the existence of this, we fix $M_1, M_2, ..., M_6$ and vary M_7; since

$$Q(M_1, ..., M_6, 1) \supset Q(M_1, ..., M_6, 2) \supset ... \supset Q(M_1, ..., M_7) \supset ... \supset Q;$$

thus,

$$\eta(M_1, ..., M_6, 1) \leq \eta(M_1, ..., M_6, 2) \leq ... \leq \eta(M_1, ..., M_7) \leq ... \leq \eta.$$

For $M_6 = 1, 2, ...$ the non decreasing and bounded above sequence $\{\eta(M_1, ..., M_7), M_7 = 1, 2, ...\}$ converges to number $\zeta(M_1, ..., M_6)$. Hence, the double limit $\lim_{M_6\to\infty} \lim_{M_7\to\infty}$ can be computed sequentially.

Now we define

$$Q(M_1, ..., M_6) \equiv \cap_{m_7=1}^{\infty} Q(M_1, ..., M_6, M_7) \equiv Q(M_1, ..., M6, \infty).$$

For the fixed numbers $M_1, M_2, ..., M5$, since $\zeta(M_1, ..., M_6) = \lim_{M_7\to\infty} \eta(M_1, ..., M_7) = \inf_{Q(M_1,...,M_6)} \quad E(\mu, \lambda)$, and

$$Q(M_1, ..., M_5, 1) \supset ... \supset Q(M_1, ..., M_5, M_6) \supset Q,$$

we have:

$$\zeta(M_1, ..., M_5, 1) \leq \zeta(M_1, ..., M_5, 2) \leq ... \leq \zeta(M_1, ..., M_5, M_6) \leq \eta,$$

thus $\zeta(M_1, ..., M_5, M_6)$ is convergent where $M_6 \to \infty$ and we have:

$$\zeta(M_1, ..., M_5) = \lim_{M_6\to\infty} \zeta(M_1, ..., M_5, M_6)$$

$$= \lim_{M_6\to\infty} [\lim_{M_7\to\infty} \eta(M_1, ..., M_6, M_7)].$$

In a similar manner, by defining

$$Q(M_1, ..., M_5) \equiv \cap_{M_6=1}^{\infty} Q(M_1, ..., M_6),$$

we would have:

$$\zeta(M_1, ..., M_4) = \lim_{M_5\to\infty} \zeta(M_1, ..., M_4, M_5)$$
$$= \lim_{M_5\to\infty} [\lim_{M_6\to\infty} [\lim_{M_7\to\infty} \eta(M_1, ..., M_6, M_7)]].$$

therefore, in the last stage, we obtain $\lim_{M_1\to\infty} \zeta(M_1) = \zeta$.

(iv) Regarding (i) and (ii), now we can prove $\zeta = \eta$. Let

$$P \equiv \cap_{M_1=1}^{\infty} \cap_{M_2=1}^{\infty} ... \cap_{M_6=1}^{\infty} \cap_{M_7=1}^{\infty} Q(M_1, ..., M_7),$$

then $P \supseteq Q$, since $Q(M_1, ..., M_7) \supset Q$ for all $M_1, M_2, ..., M_7$. We can show that under the conditions of the problem, $Q \supset P$; thus $Q = P$, that will finally imply

$$\zeta = \lim_{M_1\to\infty} ...[\lim_{M_6\to\infty} \zeta(M_1, ..., M_6)] = \inf_{Q} \quad E(\mu, \lambda),$$

which is the contention in the theorem.

For this purpose, we prove that if $(\mu, \lambda) \in P$, then they are also in Q. For a set of total functions such as ϕ_k, $k = 1, 2, ...$, we have $\lambda(\phi_k^g) = d_{\phi_k}$, according to the definition of P. Based on the definition of total functions,

since $\phi^g = \phi_r u + \phi_\theta$ on $D^{'}$, $sup_{D'}|\phi_r - \phi_{r_k}|$, $sup_{D'}|\phi_\theta - \phi_{\theta_k}|$ tend to zero as k tends to infinity, where ϕ_k^g is defined in relationship (27). Therefore:

$$
\begin{aligned}
&|\lambda(\phi^g) - d_\phi| = |\lambda(\phi^g) - d_\phi - \lambda(\phi_k^g) + d_{\phi_k}| \\
&= |\int_{\partial\omega}(\phi_r u + \phi_\theta)d\theta - \int_{\partial\omega}(\phi_{r_k} u + \phi_{\theta_k})d\theta - (d_\phi - d_{\phi_k})| \\
&\leq \int_{\partial\omega}|\phi_r - \phi_{r_k}||u|d\theta + \int_{\partial\omega}|\phi_\theta - \phi_{\theta_k}|d\theta + |d_\phi - d_{\phi_k}| \\
&\leq K_1 sup_{D'}|\phi_r - \phi_{r_k}| + K_2 sup_{D'}|\phi_\theta - \phi_{\theta_k}| + K_3 sup_{D'}|\phi(r,\theta) - \phi_k(r,\theta)|,
\end{aligned}
$$

tend to zero and hence $\lambda(\phi^g) = d_\phi$, i.e. $\lambda \in Q$, where K_1, K_2 and K_3 are constant numbers. Using the similar method, for functions ψ^g and υ_s, we prove the above relationship. Let $\mu \in P$, thus $\mu(F_i) = 0$ (F_i was defined in (27)), based on the definition of total functions, for every given $\varphi \in H_0^1(D)$ and $\epsilon > 0$, there are integer $N > 0$ and scalars γ_i, so that

$$
sup_{[0,T]} sup_\Omega \|div(y\varphi r) - \sum_{i=1}^{N} \gamma_i div(y\varphi_i r)\| < \epsilon;
$$

therefore, for every $F \in C(\Omega \times [0,T])$, we have

$$
\begin{aligned}
&|\mu(F)| = |\mu(F) - \textstyle\sum_{i=1}^{N} \gamma_i \mu(F_i)| \\
&\quad = |\int_0^T \int_\Omega div(y\phi r)drd\theta dt - \textstyle\sum_{i=1}^{N} \int_0^T \int_\Omega \gamma_i div(y\varphi_i r)drd\theta dt| \\
&\quad \leq E_1 sup_{[0,T]} sup_\Omega \|div(y\varphi r) - \textstyle\sum_{i=1}^{N} \gamma_i div(y\varphi_i r)\| \leq E_1\epsilon;\ (E_1 constant).
\end{aligned}
$$

Because ϵ is arbitrary, $|\mu(F)| \to 0$. Sequentially, by considering the density of the functions $f_i \in C(E)$, $h_i \in C(E^{'})$ and $\varphi_i \in H_0^1(D)$, we use the same method in the case of functions G_j, H_j, I_j, K_j, L_i, P_i, Q_i, R_i, N_i and T_i to prove that $\mu \in Q$. Therefore, $P \subset Q$ and the proof is finished.□

Lemma 2. For every $\epsilon > 0$, the problem of minimizing the function $\sum_{n=1}^{N} \alpha_n \Theta(Z_n)$ on the set $P(M_1, M_2, ..., M_7)^\epsilon$ described by the inequalities (34) has a solution for sufficiently large $N = N(\epsilon)$. The solution satisfies:

$$
\eta(M_1, ..., M_7) + \rho(\epsilon) \leq \sum_{n=1}^{N} \alpha_n \Theta(Z_n) \leq \eta(M_1, ..., M_7) + \epsilon, \tag{33}
$$

where $\rho(\epsilon)$ tends to zero as ϵ tends to zero.

$$
\begin{aligned}
&-\epsilon \le \textstyle\sum_{m=1}^{m=M} \beta_m \phi_k^g(z_m) - d_{\phi_k} \le \epsilon, && k = 1, 2, ..., M_1;\\
&-\epsilon \le \textstyle\sum_{m=1}^{m=M} \beta_m \psi_l^g(z_m) - 0 \le \epsilon, && l = 1, 2, ..., M_2;\\
&-\epsilon \le \textstyle\sum_{m=1}^{m=M} \beta_m \nu_s(z_m) - a_s \le \epsilon, && s = 1, 2, ..., M_3;\\
&-\epsilon \le \textstyle\sum_{m=1}^{m=M} \beta_m \frac{1}{2} r_m^2 - (\frac{1}{T}) L \textstyle\sum_{n=1}^{n=N} \alpha_n r_n \le \epsilon,\\
&-\epsilon \le \textstyle\sum_{n=1}^{n=N} \alpha_n F_i(Z_n) - 0 \le \epsilon, && i = 1, 2, ..., M_4;\\
&-\epsilon \le \textstyle\sum_{n=1}^{n=N} \alpha_n G_j(Z_n) - \textstyle\sum_{n=1}^{n=N} \alpha_n H_j(Z_n) - 0 \le \epsilon, && j = 1, 2, ..., M_5;\\
&-\epsilon \le \textstyle\sum_{n=1}^{n=N} \alpha_n I_j(Z_n) - \textstyle\sum_{n=1}^{n=N} \alpha_n K_j(Z_n) - 0 \le \epsilon, && j = 1, 2, ..., M_6;\\
&-\epsilon \le \textstyle\sum_{n=1}^{n=N} \alpha_n [L_i(Z_n) - P_i(Z_n) - Q_i(Z_n) - R_i(Z_n) + N_i(Z_n)]\\
&\qquad\qquad + \textstyle\sum_{m=1}^{m=M} \beta_m T_i(z_m) - \Phi_i \le \epsilon, && i = 1, 2, ..., M_7;\\
&\alpha_n \ge 0, \quad n = 1, 2, ..., N; \quad \beta_m \ge 0, \quad m = 1, 2, ..., M.
\end{aligned}
\tag{34}
$$

This lemma is the developed Theorem III.1 from Rubio (1986).

Proof. (i) Given $\epsilon > 0$, a set

$$\{z^k, k = 1, 2, ..., M_1 + M_2 + M_3, Z^h, h = 1, 2, ..., M_4 + M_5 + M_6 + M_7\} \tag{35}$$

can be introduced, as in the proof of Proposition III.3 of [20], so that inequalities (34) are satisfied. For sufficiently large N and M, the set

$$w_{N,M} = \{z_i : i = 1, 2, ..., N, Z_j : j = 1, 2, ..., M\} \subset w$$

will contain (37); thus, the set $P(M_1, M_2, ..., M_7)^\epsilon$ is nonempty for such values of N and M, since the N-tuple $\{\beta_k^*, k = 1, 2, ..., M_1 + M_2 + M_3, 0, 0, ..., 0\}$ and M-tuple $\{\alpha_l^*, l = 1, 2, ..., M_4 + ... + M_7, 0, 0, ..., 0\}$ are in this set. From the first set of inequalities of (34), by $z_m = 1$ for all $m = 1, 2, ..., M$, we have $-\epsilon \le \sum_{m=1}^{M} \beta_m - \Delta t \le \epsilon$ and from the fifth set of inequalities for $F_i(Z_n) = 1$, we have $-\epsilon \le \sum_{n=1}^{N} \alpha_n \le \epsilon$; this set of N and M-tuples with nonnegative entries is bounded and also closed; thus, it is compact and the linear function $\sum_{n=1}^{N} \alpha_n \Theta(Z_n)$ attains its minimum over this set. Hence, we have:

$$min \quad \sum_{n=1}^{N} \alpha_n \Theta(Z_n) \le \sum_{k=1}^{M_1+...+M_7} \alpha_k \Theta(Z^k) \le \eta(M_1, ..., M_7) + \epsilon; \tag{36}$$

therefore, one of the inequalities of (33) has been proved. For the other, let us define $Q(M_1, ..., M_7)^\epsilon$, by using equations (26) as follow:

$$
\begin{aligned}
&Q(M_1, ..., M_7)^\epsilon = \{\mu \in M^+(D), \lambda \in M^+(D^{'}) || \lambda(\phi_k^g(\theta, r, u)) - d_{\phi_k}| \le \epsilon,\\
&k = 1, 2, ..., M_1; \quad |\lambda(\psi_l^g(\theta, r, u)) - 0| \le \epsilon, \ l = 1, 2, ..., M_2; ...,\\
&|\mu(L_i) - \mu(P_i) - \mu(Q_i) - \mu(R_i) + \lambda(T_i) + \mu(N_i) - \Phi_i| \le \epsilon, \quad i = 1, 2, ..., M_7\}.
\end{aligned}
$$

Then, the set of measures of the type $\mu = \sum_{n=1}^{N} \alpha_n \delta(Z_n)$ and $\lambda = \sum_{m=1}^{M} \beta_m \delta(z_m)$ with the coefficients α_n and β_m in the set $P(M_1, ..., M_7)^\epsilon$, is a subset of $Q(M_1, ..., M_7)^\epsilon$. Thus,

$$min \sum_{n=1}^{N} \alpha_n \Theta(Z_n) \geq min\mu(\Theta), \tag{37}$$

where the minimum in the left-hand side of this inequality is over the set $P(M_1, ..., M_7)^\epsilon$ and the one in right- hand side is over $Q(M_1, ..., M_7)^\epsilon$. Also, $Q(M_1, ..., M_7) = \cap_{\epsilon>0} Q(M_1, ..., M_7)^\epsilon$, and

$$Q(M_1, ..., M_7)^{\epsilon_1} \supset Q(M_1, ..., M_7)^{\epsilon_2} \quad if \quad \epsilon_1 > \epsilon_2. \tag{38}$$

Let $\eta(M_1, ..., M_7, \epsilon)$ be the infimum of $\mu(\Theta)$ over the set of measures $Q(M_1, ..., M_7)^\epsilon$. Then, by (40), we have $\eta(M_1, ..., M_7, \epsilon_1) \leq \eta(M_1, ..., M_7, \epsilon_2)$, if $\epsilon_1 > \epsilon_2$.
It is sufficient for our purposes to consider a sequence of values of $\epsilon = 1/p$ where $p = 1, 2, ...$ Then,

$$\eta(M_1, ..., M_7, 1) \leq \eta(M_1, ..., M_7, \frac{1}{2}) \leq ... \leq \eta(M_1, ..., M_7, \frac{1}{p}) \leq ... \leq \eta,$$

the sequence $\{\eta(M_1, ..., M_7, \frac{1}{p})\}$ is non decreasing and bounded above. Therefore, it converges to a number $\gamma(M_1, ..., M_7)$ satisfying

$$\gamma(M_1, ..., M_7) = \lim_{p\to\infty} \eta(M_1, ..., M_7, \frac{1}{p}) = \inf_{Q(M_1,...,M_7)} \mu(\Theta) = \eta(M_1, ..., M_7).$$

Thus

$$\rho(\epsilon) \equiv \eta(M_1, ..., M_7, \epsilon) - \eta(M_1, ..., M_7) \tag{39}$$

tends to zero as ϵ tends to zero; it follows from (39) and (41) that

$$min \sum_{n=1}^{N} \alpha_n \Theta(Z_n) \geq min\mu(\Theta) = \eta(M_1, ..., M_7) + \rho(\epsilon),$$

where the left-hand minimum is over the set $P(M_1, ..., M_7)^\epsilon$ and the right-hand one is over $Q(M_1, ..., M_7)^\epsilon$. Now we prove Theorem 2 as follow:

Theorem 2. *If the used minimization techniques in Step 3 of the above algorithm are convergent, then, the algorithm converges to the optimal solution of (1) when $M, N, M_1, M_2, ..., M_7$ tend to infinity.*

Proof. To demonstrate the proof, we have used the proof by contradiction. Let $(\alpha^*, \beta^*, a^*, r_d^*)$ be the minimizer of $E(\alpha, \beta, a, r_d)$ but (ω^*, a^*) is not the minimizer of $E(\omega, a, t)$, this means that the algorithm does not converge to

the solution of (1). Thus, there is $(\omega^{'}, a^{'})$ such that

$$E(\omega^{'}, a^{'}, T) < E(\omega^{*}, a^{*}, T). \tag{40}$$

According to the Riesz representation theorem and considering one to one transformation of problem (1), with objective function defined in relationship (2), to problem (26), there are unique measures $\mu^{'}(\Theta)$ and $\mu^{*}(\Theta)$ corresponding to $E(\omega^{'}, a^{'}, T)$ and $E(\omega^{*}, a^{*}, T)$, where:

$$\mu^{'}(\Theta) \equiv E(\mu^{'}, \lambda^{'}) < \mu^{*}(\Theta) \equiv E(\mu^{*}, \lambda^{*}).$$

According to Lemma 1, we have:

$$\eta(M_1, ..., M_7) = \inf_{Q(M_1,...,M_7)} \quad \mu(\Theta) \to \eta = \inf_{Q} \quad \mu(\Theta) = \inf E(\mu, \lambda),$$

and according to Lemma 2,

$$\eta(M_1, ..., M_7) + \rho(\epsilon) \leq \sum_{n=1}^{N} \alpha_n \Theta(Z_n) \leq \eta(M_1, ..., M_7) + \epsilon,$$

therefore,

$$\mu^{'}(\Theta) = \inf_{Q} \ \ \mu(\Theta) \equiv inf \ \ E(\mu, \lambda) = E(\mu^{'}, \lambda^{'}) < E(\mu^{*}, \lambda^{*})$$

and

$$\eta^{'} + \rho(\epsilon) \leq \sum_{n=1}^{N} \alpha_n^{'} \Theta(Z_n) \equiv E(\alpha^{'}, \beta^{'}, a^{'}, r_{d_a}^{'}) \leq \eta^{'} + \epsilon,$$

according to the above relationships:

$$\sum_{n=1}^{N} \alpha_n^{'} \Theta(Z_n) < \sum_{n=1}^{N} \alpha_n^{*} \Theta(Z_n)$$

thus according to (42), we have:

$$E(\alpha^{'}, \beta^{'}, a^{'}, r_d^{'}) < E(\alpha^{*}, \beta^{*}, a^{*}, r_d^{*}).$$

This is in contradiction with what we supposed at the beginning, thus, (ω^{*}, a^{*}) is minimizer of problem (1).

References

1. Abbas, H. A. *Marriage in Honey Bees Optimization (MBO): A Haplomelrosis Polyggnous Swarming Approach*, The Congress on Evolutionary Computation. (2001) 207-214.

2. Alexander, G. R. *Spectral and scattering theory*, United state of America, 1998.

3. Apostol, T. *Mathematical Analysis*, (2nd edition). Addison Wesley, 1974.

4. Cox, S. J. and Zuazua, E. *The rate at which energy decays in a damped string*, Partial Differential Equations. 19 (1994) 213243.

5. Debnath, L. and Mikusinski, P. *Hilbert spaces with Applications*, Elsevier academic Press, 3 rd Edition , 2005.

6. Fakharzadeh, J. A. and Rubio, J. E. *Global Solution of Optimal Shape Design Problems*, J. Anal. and Appl. 18 (1999) 143-155.

7. Fakharzadeh, J. A. and Rubio, J. E. *Best Domain for an Elliptic Problem in Cartesian Coordinates by Means of Shape-measure*, J of control. 11 (2009) 536-554.

8. Fakharzadeh, J. A., Alimorad, D. H. and Rafiei, Z. *Using Linearization and Penalty Approach to Solve Optimal Shape Design Problem with an Obstacle*, J. mathematics and computer Science. 7 (2013) 43-53.

9. Freitas, P. *Optimizing the Rate of Decay of Solution of the Wave Equation Using Genetic Algorithms: A Counterex Ample to the Constant Damping Conjecture*, SIAM J Control Optim. 37 (1999) 376-387.

10. George, B. T., Ross, L. F. and Maurice, D. W. *Calculus and Analytic Geometry* , 10 Edition, Wesley, 2001.

11. Hebrard, P. and Henrot, A. *Optimal Shape and Position of the Actuators for the Stabilization of the String*, J. syst. Control Letters 48 (2003) 199-209.

12. Lions, J. and Magenes, E. *Problem aux Limites Non Homogenes A pplications*, Dunod, Paris, 1968.

13. Lopez-Gomez, J. *On the Linear Damped Wave Equation*, J. Differ. Equations 134 (1997) 26-45.

14. Muller, S. *Variational Model for Microstructure and Phase Transitions*, Lecture at the C.I.M.E summer school 'calculus of variations and geometric evolution problems', 1998.

15. Munch, A., Pedregal, P. and Periago, F. *Optimal Design of the Damping set for the Stabilization of the Wave Equation*, J. Differ. Equations 231 (2006) 331-358.

16. Munch, A., Pedregal, P. and Periago, F. *Optimal Internal Stabilization of the Linear System of Elasticity*, Archive Rational Mechanical Analysis, 2009.

17. Munch, A. *Optimal Internal Dissipation of Damped Wave Equation using a Topological Approach*, Int. J. Appl. Math. Comp. Sci. 19 (2009) 15-37.

18. Privat, Y. and Trelat, E. *Optimal Design of Sensors for a Damped Wave Equation*, Madrid, Discrete Cont. Dynam. Syst. Proceedings of AIMS 2014 workshop.

19. Privat, Y., Trelat, E. and Zuazua, E. *Optimal Shape and Location of Sensors for Parabolic Equations with Random Initial Data*, to appear in Arch. Rat. Mech. Anal. Preprint (2014).

20. Rubio, J. E. *Control and Optimization: the linear treatment of nonlinear problems*, Manchester university Press, Manchester, 1986.

21. Rudin, W. *Pronciples of Mathematical Analysis* , third Edition, McGraw-Hill, new York, 1976.

22. Rudin, W. *Real and Complex Analysis* , second Edition, Tata McGraw-Hill Publishing Co Ltd., new Dehli, 1983.

23. Tcheugoue-Tebou, L. R. *Stabilization of the Wave Equation with Localized Nonlinear Damping*, J. Differ. Equations 145 (1998), 502-524.

24. Treves, F. *Topology Vector Space, Distributions and Kernels* , Academic Press, Now York and London, 1967.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 31-42

Chebyshev Galerkin method for integro-differential equations of the second kind

J. Biazar* and F. Salehi

Abstract

In this paper, we propose an efficient implementation of the Chebyshev Galerkin method for first order Volterra and Fredholm integro-differential equations of the second kind. Some numerical examples are presented to show the accuracy of the method.

Keywords: Volterra integro-differential equations; Galerkin method; Chebyshev polynomials.

1 Introduction

Integro-differential equations occur in various areas. These equations arise in mathematical modeling of many scientific phenomena, such as fluid dynamics, solid state physics, plasma physics, mathematical biology viscoelasticity [33], heat transfer [6], economics [26], chemostat [41], HIV models [4], biotissues [15], static analysis of wind towers or chimneys [35], and chemical kinetics [34]. Integro-differential equations contain both integral and differential operators. The derivatives of the unknown functions may appear to any order [2, 40].

The concepts of integral equations have motivated a large amount of research work in recent years. Many numerical methods have been applied to solve these equations such as: El-gendi and Galerkin [11, 12, 27], Euler-Chebyshev [37], Variational iteration [39], Homotopy perturbation [10, 32], Chebyshev and Taylor collocation [1, 3, 8, 13, 20, 28], Chebyshev Wavelets [5],

*Corresponding author
Received 20 July 2014; revised 15 October 2014; accepted 1 July 2015
J. Biazar
Department of Applied Mathematics, Faculty of Mathematical Science, University of Guilan, Rasht, Iran. e-mail: biazar@guilan.ac.ir

F. Salehi
Department of Mathematics, Darab Branch, Islamic Azad University, Darab, Iran. e-mail: f.salehi630@gmail.com

Spline collocation [7], finite element [9], sinc collocation [43], Bessel polynomials [42], Legendre polynomials [23], Bernstein polynomials [21] and Lagrange polynomials [31] and etc. [27, 36].

Galerkin method is a powerful tool for solving many kinds of equations in various fields of science and engineering. It is one of the most important weighted residual methods invented by the Russian mathematician Boris Grigoryevich Galerkin. Recently, various Galerkin algorithms have been applied in numerical solution of integral equations and integro differential equations. We can mention the following methods that are based on the Galerkin idea: Galerkin finite element [22], iterated Galerkin [38], Galerkin with hybrid functions [27], Crank–Nicolson least–squares Galerkin [18], Wavelet–Galerkin [16], Discrete Galerkin [30], Petrov–Galerkin [25], pseudo–spectral Legendre–Galerkin [14] and etc. There are many different families of orthogonal functions, which can be used. Chebyshev polynomials are considerably useful to solve integro-differential equations.

In this paper, the solution is approximated by a linear combination of the first $N+1$ Chebyshev polynomials, with $\{a_i\}_{i=0}^{N}$, as coefficients. Approximate solution will be simplified as a polynomial in x. This approximation will be substituted in the equation. To determine a_j, one can consider inner product of both sides of the equation, by $T_j(x)$. This procedure reduces the problem to a system of equations. The generated system, which considering the type of the equation will be either linear or nonlinear, can be solved through various methods and the unknown coefficients can be found. Practically, all orthogonal polynomials, on a closed finite interval, can also be applied for approximating functions. But convergence of the partial sums of the first-kind Chebyshev expansion, of a continuous function on $[-1, 1]$, is faster than the partial sums of an expansion in any other orthogonal polynomials [3].

The outline of this paper is as follows: Section 2 presents the method for solving Volterra integro-differential equations. Numerical examples are given in Section 3. Finally, conclusion will be presented in Section 4.

2 Chebyshev Galerkin method

Consider the following Volterra integro-differential equation

$$u'(x) = f(x) + \lambda \int_a^x K(x,t)u(t)\,dt\,, \quad a \le x \le b, \tag{1}$$

$$u(a) = \alpha. \tag{2}$$

where $u(x)$ is the unknown function, $K(x,t)$ is a known continuous and square integrable function, $f(x)$ is a known function, and λ is a real known parameter.

The method under study uses Chebyshev polynomials, well addressed in [29],

as a basis polynomial to approximate the solution on a closed finite interval. Assume that

$$u(x) \approx u_N(x) = \sum_{i=0}^{N} a_i T_i \left(\frac{2x-(b-a)}{b-a} \right), \tag{3}$$

where $T_i\left(\frac{2x-(b-a)}{b-a}\right)$ is shifted Chebyshev polynomial at $[a,b]$. So we have

$$u'(x) \approx u'_N(x) = \sum_{i=0}^{N} \frac{2}{b-a} a_i T'_i \left(\frac{2x-(b-a)}{b-a} \right). \tag{4}$$

Substituting (3) and (4) into (1), results in

$$\begin{aligned} &\sum_{i=0}^{N} \frac{2}{b-a} a_i T'_i \left(\frac{2x-(b-a)}{b-a} \right) \\ &\quad = f(x) + \lambda \sum_{i=0}^{n} a_i \int_a^x K(x,t) T_i \left(\frac{2t-(b-a)}{b-a} \right) dt, \qquad a \le x \le b. \end{aligned} \tag{5}$$

To determine unknown coefficients a_i, we use the Galerkin idea by multiplying both sides of (5) by $T_j\left(\frac{2x-(b-a)}{b-a}\right)$ and then integrating with respect to x from -1 to 1. So we have

$$\begin{aligned} &\sum_{i=0}^{N} \frac{2}{b-a} a_i \int_{-1}^{1} T'_i \left(\frac{2x-(b-a)}{b-a} \right) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx = \\ &\qquad \int_{-1}^{1} f(x) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx + \\ &\int_{-1}^{1} \left(\lambda \sum_{i=0}^{n} a_i \int_a^x K(x,t) T_i \left(\frac{2t-(b-a)}{b-a} \right) dt \right) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx, \end{aligned} \tag{6}$$

for $j = 0, 1, \ldots, N$, or equivalently

$$\begin{aligned} &\sum_{i=0}^{N} \frac{2}{b-a} a_i \int_{-1}^{1} T'_i \left(\frac{2x-(b-a)}{b-a} \right) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx = \\ &\qquad \int_{-1}^{1} f(x) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx + \\ &\lambda \sum_{i=0}^{n} a_i \int_{-1}^{1} \left(\int_a^x K(x,t) T_i \left(\frac{2t-(b-a)}{b-a} \right) dt \right) T_j \left(\frac{2x-(b-a)}{b-a} \right) dx. \end{aligned} \tag{7}$$

If needed the integrals can be calculated by numerical methods. This procedure generates a system of linear equations for the unknown $\{a_i\}_{i=0}^{N}$. Many researchers substitute initial condition

$$u(a) = \alpha \Rightarrow \sum_{i=0}^{N} a_i T_i \left(\frac{2a - (b-a)}{b-a} \right) = \sum_{i=0}^{N} a_i T_i(-1) = \alpha. \tag{8}$$

for the same number of equations in the foregoing linear system.

The unknown parameters are determined by solving the system of equations (7) and (8). Substituting these values in (3) gives the approximate solution of the integro-differential equation (1). Similarly one can apply this approach for a Fredholm integro-differential equation in the following general form:

$$\begin{aligned} u'(x) &= f(x) + \lambda \int_a^b K(x,t) u(t)\, dt, \quad a \le x \le b, \\ u(a) &= \alpha \end{aligned}$$

3 Numerical Examples

In this section, we intend to show the efficiency of the Galerkin method for solving Volterra integro-differential equations of the second kind by Chebyshev polynomials by presenting three illustrative examples. The absolute error for this formulation is defined by

$$E(x) = |u(x) - u_N(x)|.$$

Example 1. Consider the following Fredholm integro-differential equations of the second kind [17]

$$\begin{aligned} u'(x) &= u(x) - \frac{1}{2}x + \frac{1}{(x+1)} - \ln(x+1) + \frac{1}{(\ln 2)^2} \int_0^1 \frac{x}{t+1} u(t)\, dt, \\ u(0) &= 0, \end{aligned} \tag{9}$$

with the exact solution $u(x) = \ln(x+1)$.

To solve Equation (9) we approximate $u(x)$ and $u'(x)$ as follows:

$$\begin{aligned} u_4(x) = \sum_{i=0}^{4} a_i T_i \left(\frac{2x - (b-a)}{b-a} \right) &= a_0 + a_1(2x-1) \\ &+ a_2(8x^2 - 8x + 1) + a_3(32x^3 - 48x^2 + 18x - 1) \\ &+ a_4(128x^4 - 256x^3 + 160x^2 - 32x + 1), \end{aligned} \tag{10}$$

and

$$u_4'(x) = \sum_{i=0}^{4} \frac{2}{b-a} a_i T_i' \left(\frac{2x-(b-a)}{b-a} \right) = 2a_1 + a_2(16x-8) + a_3(96x^2-96x+18) + a_4(512x^3-768x^2+320x-32). \quad (11)$$

Substituting (10) and (11) into (9), results in

$$\begin{aligned} &2a_1 + a_2(16x-8) + a_3(96x^2-96x+18) + a_4(512x^3-768x^2+320x-32) \\ &= a_0 + a_1(2x-1) + a_2(8x^2-8x+1) + a_3(32x^3-48x^2+18x-1) \\ &+ a_4(128x^4-256x^3+160x^2-32x+1) - \frac{1}{2}x + \frac{1}{(x+1)} - \ln(x+1) \\ &+ \frac{x}{(\ln 2)^2} \int_0^1 \frac{1}{t+1} \big(a_0 + a_1(2t-1) + a_2(8t^2-8t+1) \\ &+ a_3(32t^3-48t^2+18t-1) + a_4(128t^4-256t^3+160t^2-32t+1) \big)\, dt, \end{aligned} \quad (12)$$

By multiplying both sides of (12) by $T_j\left(\frac{2x-(b-a)}{b-a}\right)$ and then integrating it with respect to x from -1 to 1, we obtain a system of linear equations which one of them is replaced by the equation

$$u(0) = 0 \Rightarrow \alpha_0 - \alpha_2 + \alpha_4 = 0 \quad (13)$$

Now the unknown coefficients $\{a_i\}_{i=0}^4$ are determined by solving this system. Substituting these values in (3) gives the approximate solution of the integro-differential equation (1). The results have been shown in Table 1, for $N = 4, 8, 12$, and Error is plotted in Figure 1, for $N = 12$.

Example 2. Consider the following Volterra integro-differential equations of the second kind [40]

$$u'(x) = 1 - 2x\sin(x) + \int_0^x u(t)\, dt, \quad u(0) = 0.$$

The exact solution is $y = x\cos(x)$.
Table 2 shows the results for $N = 4, 8, 12$. Also Figure 2 shows absolute error for $N = 12$.

Example 3. Consider the following Volterra integro-differential equations [40]:

$$u'(x) = -1 + \frac{1}{2}x^2 - xe^x - \int_0^x tu(t)\, dt, \quad u(0) = 0.$$

The exact solution is $y = 1 - e^x$. Results have been shown in Table 3, for $N = 4, 8, 12$, and Error plotted in Figure 3, for $N = 12$.

Table 1: Absolute Error for Example 1

	N=4		*N=8*		*N=12*	
x	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*
0.0	0	0	0	0	0	0
0.1	0.0949	4.233e-4	0.0953	4.556e-7	0.0953	4.998e-10
0.2	0.1817	6.317e-4	0.1823	7.690e-7	0.1823	7.938e-10
0.3	0.2615	8.544e-4	0.2624	1.082e-6	0.2624	1.146e-09
0.4	0.3353	1.161e-3	0.3365	1.416e-6	0.3365	1.508e-09
0.5	0.4039	1.543e-3	0.4055	1.860e-6	0.4055	1.964e-09
0.6	0.4680	1.971e-3	0.4700	2.361e-6	0.4700	2.481e-09
0.7	0.5282	2.430e-3	0.5306	2.909e-6	0.5306	3.088e-09
0.8	0.5848	2.951e-3	0.5878	3.614e-6	0.5878	3.821e-09
0.9	0.6382	3.626e-3	0.6418	4.438e-6	0.6419	4.687e-09
1.0	0.6885	4.629e-3	0.6931	5.611e-6	0.6931	5.930e-09

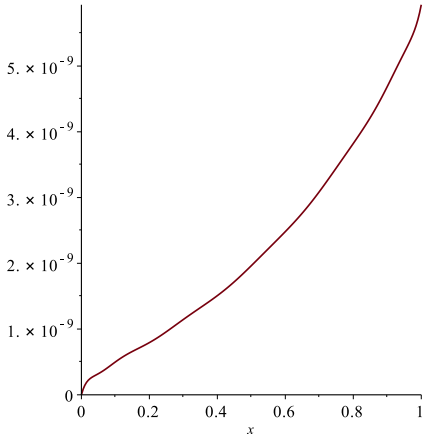

Figure 1: Absolute Error for Example 1

Table 2: Absolute Error for Example 2

	N=4		*N=8*		*N=12*	
x	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*
0	0	0	0	0	0	0
0.1	0.0991	4.308e-4	0.0995	2.308e-9	0.0995	1.383e-15
0.2	0.1954	5.890e-4	0.1960	3.450e-9	0.1960	1.987e-15
0.3	0.2859	6.954e-4	0.2866	4.423e-9	0.2866	2.570e-15
0.4	0.3676	8.387e-4	0.3684	5.136e-9	0.3684	3.046e-15
0.5	0.4378	1.023e-3	0.4388	6.131e-9	0.4388	3.586e-15
0.6	0.4940	1.214e-3	0.4952	7.178e-9	0.4952	4.160e-15
0.7	0.5340	1.381e-3	0.5354	8.072e-9	0.5354	4.747e-15
0.8	0.5558	1.537e-3	0.5574	9.368e-9	0.5574	5.509e-15
0.9	0.5577	1.776e-3	0.5594	1.094e-8	0.5594	6.371e-15
1.0	0.5380	2.300e-3	0.5403	1.382e-8	0.5403	8.087e-15

Table 3: Absolute Error for Example 3

	N=4		*N=8*		*N=12*	
x	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*	*Approx. solution*	*Abs. Error*
0	0	0	0	0	0	0
0.1	-0.1050	1.750e-4	-0.1052	5.037e-10	-0.1052	2.053e-16
0.2	-0.2212	2.406e-4	-0.2214	7.428e-10	-0.2214	2.930e-16
0.3	-0.3496	2.785e-4	-0.3499	9.426e-10	-0.3499	3.730e-16
0.4	-0.4915	3.251e-4	-0.4918	1.069e-9	-0.4918	4.332e-16
0.5	-0.6483	3.850e-4	-0.6487	1.240e-9	-0.6487	4.955e-16
0.6	-0.8217	4.448e-4	-0.8221	1.407e-9	-0.8221	5.561e-16
0.7	-1.0130	4.891e-4	-1.0140	1.516e-9	-1.0140	6.081e-16
0.8	-1.2250	5.177e-4	-1.2260	1.679e-9	-1.2260	6.757e-16
0.9	-1.4590	5.648e-4	-1.4600	1.874e-9	-1.4600	7.438e-16
1.0	-1.7180	7.204e-4	-1.7180	2.311e-9	-1.7180	9.235e-16

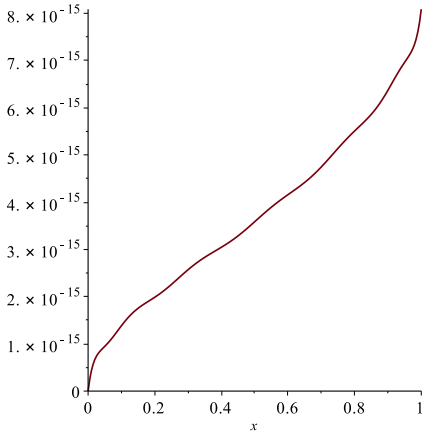

Figure 2: Absolute Error for Example 2

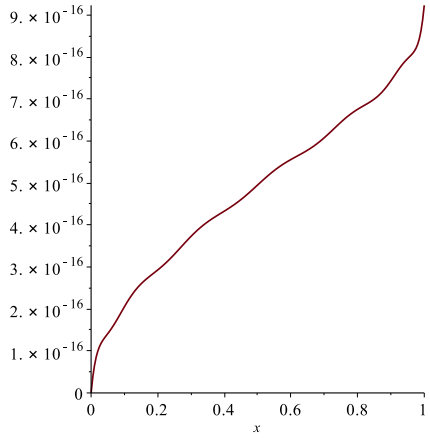

Figure 3: Absolute Error for Example 3

4 Conclusion

This article deals with the numerical solution of the first order Volterra integro-differential equations of the second kind, using Galerkin method by Chebyshev Polynomials. This technique is tested on three examples and the results are satisfactory. In addition this method is portable to high order Volterra integro-differential equations of the second kind and easy to program.

Acknowledgment

Authors would like to thank anonymous reviewers for their valuable, useful, and constructive comments, that without their suggestion this paper was not competent enough to be published in this journal.

Authors also expressed their gratitude to the University of Guilan for financial support.

References

1. Akyuz, A. *Chebyshev Collocation method for solution of linear integro-differential equations*, M.Sc. Thesis, Dokuz Eylul University, Graduate School of Natural and Applied Sciences, 1997.

2. Atkinson, K. *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge University Press, 1997.

3. Avazzadeh, Z. and Heydari, M. *Chebyshev polynomials for solving two dimensional linear and nonlinear integral equations of the second kind*, Comp. Appl. Math. 31(1) (2012), 127–142.

4. Banks, H., Bortz, D. and Holt, S. *Incorporation of variability into the modeling of viral delays in HIV infection dynamics*, Math. Biosci 183 (2003) 63-91.

5. Barzkar, A., Oshagh, M., Assari, P. and Mehrpouya, M. *Numerical Solution of the Nonlinear Fredholm Integral Equation and the Fredholm Integro-differential Equation of Second Kind using Chebyshev Wavelets*, World Appl. Sci. J. 18(12) (2012) 1774-1782.

6. Boland, J. *The analytic solution of the differential equations describing heat flow in houses*, Build. Environ. 37(2002), 1027–1035.

7. Brunner, H. *The approximate solution of initial-value problems for general Volterra integro-differential equations*, Comput. 40(40) (1988) 125-137.

8. Cerdik-Yaslan, H. and Akyuz-Dascioglu, A. *Chebyshev polynomial solution of nonlinear Fredholm-Volterra integro-differential equations*, J. Arts. Sci. 5 (2006) 89-101.

9. Chen, C. and Shih ,T. *Finite Element Methods for Integro differential Equations*, Singapore: World Scienti?c Publishing Co Ltd, 1997.

10. Dehghan, M. and Shakeri, F. *Solution Of An Integro-Differential Equation Arising In Oscillating Magnetic Fields Using He's Homotopy Perturbation Method*, Progr. In Electromagnetics Res., PIER 78 (2008) 361–376.

11. Delves, L. and Mohamed, J. *Computional Methods for Integral Equations*, Cambridge: Cambridge University Press, 1985.

12. El-gendi, S. *Chebyshev solution of differential, integral and integro- differential equations*, Comput. J. 12 (1969) 282-287.

13. Ezzati, R. and Najafalizadeh, S. *Numerical solution of nonlinear Volterra-Fredholm integral equation by using Chebyshev polynomials*, Mathematical Sciences 5(1) (2011) 1-12.

14. Fakhar-Izadi, F. and Dehghan, M. *An efficient pseudo-spectral Legendre-Galerkin method for solving a nonlinear partial integro- differential equation arising in population dynamics*, Math. Meth. Appl. Sci. 36(12) (2012) 1485-1511.

15. Fung, B.Y.C. *Mechanical Properties of Living Tissues, Springer-Verlag, New York*, Springer-Verlag, New York, 1993.

16. Ghasemi, M., Tavassoli kajani, M. and Babolian, E. *Numerical solutions of the nonlinear integro-differential equations: Wavelet-Galerkin method and homotopy perturbation method*, Appl. Math. Comp. 188(1) (2007) 450-455.

17. Golbabai, A. and Seifollahi, S. *Radial basis function networks in the numerical solution of linear integro-differential equations*, Appl. Math. Comput. 188 (2007) 427–432.

18. Guo, H. and Rui, H. *Crank–Nicolson least-squares Galerkin procedures for parabolic integro-differential equations*, Appl. Math. Comp. 180 (2) (2006) 622-634.

19. Holmaker, K. *Global asymptotic stability for a stationary solution of a system of integro-differential equations describing the formation of liver zones*, SIAM J. Math. Anal. 24(1) (1993) 116–128.

20. Huang, Y. and Li, X. *Approximate solution of a class of linear integro-differential equations by Taylor expansion method*, Int. J. Comput. Math 15(3) (2009) 1–12.

21. Isik, O., Sezer, M. and Gney, Z. *Bernstein series solution of a class of linear integro-differential equations with weakly singular kernel*, Appl. Math,Comput. 217 (2011) 7009–7020.

22. Jangveladze, T., Kiguradze, Z. and Neta, B. *Galerkin finite element method for one nonlinear integro-differential model*, Appl. Math. Comp. 217(16) (2011) 6883-6892.

23. Khater, A., Shamardan, A., Callebaut, D. and Sakran, M. *Numerical solutions of integral and integro-differential equations using Legendre polynomials*, Numer. Alg. 46 (2007) 195–218.

24. Lewis, B.A., *On the numerical solution of Fredholm integral equations of the first kind*, J. Inst. Math. Appl. 16(2) (1975) 207–220.

25. Lin, T., Lin, Y., Rao, M. and Zhang, S. *Petrov-Galerkin Methods for Linear Volterra Integro-Differential Equations*, SIAM J. Numer. Anal. 38(3) (2001) 937-963.

26. Makroglou, A. *Numerical solution of some second order integro-differential equations arising in ruin theory, in: Proceedings of the Third Conference in*, in: Proceedings of the Third Conference in Actuarial Science and Finance Held in Samos, Greece, p. pp. 2–5, September 2004.

27. Maleknejad, K. and Tavassoli Kajani, M. *Solving linear integro-differential equation system by Galerkin methods with hybrid functions*, Appl. Math. Comput. 159 (2004) 603–612.

28. Maleknejad, K., Sohrabi, S. and Rostami, Y. *Numerical solution of nonlinear Volterra integral equation of the second kind by using Chebyshev polynomials*, Appl. Math. Cmput. 188 (2007) 123-128.

29. Mason, J.C. and Handscomb, D. *Chebyshev Polynomials*, Boca Raton: Chapman & Hall/CRC, 2003.

30. Pedas, A. and Tamme, E. *Discrete Galerkin method for Fredholm integro-differential equations with weakly singular kernels*, Journal of Computational and Applied Mathematics 213(1) (2008) 111-126.

31. Rashed, M. *Lagrange interpolation to compute the numerical solutions of differential, integral and integro-differential equation*, Appl. Math. Comput. 151 (2004) 869–878.

32. Saberi-Nadjafi, J. and Ghorbani, A. *He's homotopy perturbation method: an effective tool for solving nonlinear integral and integro-differential equations*, Comput. Math. Appl. 58 (2009) 2379-2390.

33. Sergeev, V. *On stability of viscoelastic plate equilibrium*, Automat. Remote Control 68(9) (2007) 1544–1550.

34. Thieme, H.R. *A model for the spatial spread of an epidemic*, J. Math. Biol. 4 (1977) 337–351.

35. Tomasiello, S. *A note on three numerical procedures to solve Volterra integro-differential equations in structural analysis*, Comput. Math. Appl. 62(8) (2011) 3183–3193.

36. Tomasiello, S. *Some remarks on a new DQ-based method for solving a class of Volterra integro-differential equations*, Appl. Math. Comput. 219 (2012) 399–407.

37. Van der Houwen, P. and Sommeijer, B. *Euler–Chebyshev methods for integro-differential equations*, Appl. Numer. Math. 24 (1997) 203–218.

38. Volk, W. *The iterated Galerkin method for linear integro-differential equations*, J. Comp. Appl. Math. 21(1) (1988) 63-74.

39. Wang, S. and He, J. *Variational iteration method for solving integro-differential equations*, Physics letters A 367 (2007) 188–191.

40. Wazwaz, A. *Linear And Nonlinear Integral Equations, Methods And Applications*, Springer, 2011.

41. Wolkowicz, G., Xia, H. and Ruan, S. *Competition in the chemostat: a distributed delay model and its global asymptotic behavior*, SIAM J. Appl. Math. 57 (1997) 1281–1310.

42. Yüzbaşı, S., Sahidin, N. and Sezer, M. *Bessel polynomial solutions of high-order linear Volterra integro-differential equations*, Comput. Math. Appl 62(4) (2011) 1940–1956.

43. Zarebnia, M. *Sinc numerical solution for the Volterra integro-differential equation*, Commun. Nonlinear Sci. Numer. Simulat 15(3) (2010) 700–706.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 43-52

Kudryashov method for exact solutions of isothermal magnetostatic atmospheres

N. Kadkhoda* and H. Jafari

Abstract

The Kudryashov method to look for the exact solutions of the nonlinear differential equations is presented. The Kudryashov method is applied to search for the exact solutions of the Liouville equation and the Sinh-Poisson equation. The equations of magnetohydrostatic equilibria for a plasma in a gravitational field are investigated analytically. An investigation of a family of isothermal magnetostatic atmospheres with one ignorable coordinate corresponding to a uniform gravitational field in a plane geometry is carried out. The distributed current in the model J is directed along the x-axis where x is the horizontal ignorable coordinate. These equations transform to a single nonlinear elliptic equation for the magnetic vector potential u. This equation depends on an arbitrary function of u that must be specified.

Keywords: Kudryashov method; magnetostatic equilibria; nonlinear evolution equations; traveling waves.

1 Introduction

The equations of magnetostatic equilibria have been used extensively to model the solar magnetic structure [1, 4, 9, 11]. An investigation of a family of isothermal magnetostatic atmospheres with one ignorable coordinate corresponding to a uniform gravitational field in a plane geometry is carried out. The force balance consists of the force between $J \wedge B$ (B, magnetic field induction, J is the electric current density), the gravitational force, and gas pressure gradient force. However, in many models, the temperature distribution is specified a priori and direct reference to the energy equations is

*Corresponding author
Received 30 March 2015; revised 17 June 2015; accepted 1 August 2015
N. Kadkhoda
Department of Mathematics, Faculty of Basic Sciences, Bozorgmehr University Of Qaenat, Qaenat, Iran. e-mail: kadkhoda@buqaen.ac.ir

H. Jafari
Department of Mathematics and Computer Science,University of Mazandaran, Babolsar, Iran. e-mail: jafari@umz.ac.ir

eliminated. In solar physics, the equations of magnetostatic have been used to model diverse phenomena, such as the slow evolution stage of solar flares, or the magnetostatic support of prominences [20]. The nonlinear equilibrium problem has been solved in several cases [3, 8, 17, 18]. In this paper, we obtain the exact analytical solutions for the Liouville and sinh-Poisson equations using the Kudryashov method. Because these two models will be special cases of magnetostatic atmospheres model. Also here there is force balance between different forces. The Kudryashov method was developed by Kudryashov on the basis of a procedure analogous to the first step of the test for the Painlev property [2, 7, 7, 9, 10]. The paper is organized as follows :
In Section 2, we describe the methodology of Kudryashov method for solving nonlinear evolution equations when the Riccati equation is used as the simplest equation. We describe the Basic equations in Section 3. We apply this methodology and obtain exact solutions of the Liouville and sinh-Poisson equations in Section 4. Finally, the concluding remarks are presented in Section 5.

2 Analysis of the Kudryashov method

We consider a partial differential equation and we assume that by means of an appropriate transformation this partial differential equation is transformed to a nonlinear ordinary differential equation in the form

$$P(u, u', u'', u''', ...) = 0. \tag{1}$$

Exact solution of this equation can be constructed as finite series

$$u(\xi) = \sum_{i=0}^{n} A_i (G(\xi))^i, \tag{2}$$

where $G(\xi)$ is a solution of some ordinary differential equation referred to as the simplest equation. The simplest equation has two properties:

1. the order of simplest equation should be less than the order of equation (1).
2. we should know the general solution of the simplest equation or at least exact analytical particular solution(s) of the simplest equation.

In this paper, we use the equation of Riccati, as the simplest equation. This equation is a well-known nonlinear ordinary differential equation which possesses the exact solution constructed by elementary function. In this paper for the Riccati equation

$$G'(\xi) = c\, G(\xi) + d\, G(\xi)^2, \tag{3}$$

we use the solution

$$G(\xi) = \frac{c\,\exp[c(\xi+\xi_0)]}{1-d\exp[c(\xi+\xi_0)]}; \qquad d<0, c>0, \tag{4}$$

and

$$G(\xi) = -\frac{c\,\exp[c(\xi+\xi_0)]}{1+d\exp[c(\xi+\xi_0)]}; \qquad d>0, c<0. \tag{5}$$

Here ξ_0 is a constant of integration. Now $u(\xi)$ can be determined explicitly by using the following three steps:

- Step (1). By considering the homogeneous balance between the highest nonlinear terms and the highest order derivatives of $u(\xi)$ in equation (1), the positive integer n in (2) is determined.
- Step (2). By substituting equation (2) with equation (3) into equation (1) and collecting all terms with the same powers of G together, the left hand side of equation (1) is converted into a polynomial. After setting each coefficient of this polynomial to zero, we obtain a set of algebraic equations in terms of $A_i\,(i=0,1,2,...,n)$, c, d.
- Step (3). Solving the system of algebraic equations and then substituting the results and the general solutions of (4) or (5) into (2) gives solutions of (1).

3 Basic equations

The relevant of magnetohydrostatic equations consisting of the equilibrium equation with force balance will be as:

$$J \wedge B - \rho\,\nabla\,\Phi - \nabla\,P = 0, \tag{6}$$

which is coupled with Maxwells equations:

$$J = \frac{\nabla \wedge B}{\mu}, \tag{7}$$

$$\nabla \cdot B = 0, \tag{8}$$

where P, ρ , μ and Φ are the gas pressure, the mass density, the magnetic permeability and the gravitational potential, respectively. It is assumed that the temperature is uniform in space and that the plasma is an ideal gas with equation of state $p = \rho\,R_0\,T_0$, where R_0 is the gas constant and T_0 is the temperature. Then the magnetic field B can be written by the following:

$$B = \nabla\,u \wedge e_x + B_x\,e_x = (B_x, \frac{\partial u}{\partial z}, \frac{-\partial u}{\partial y}). \tag{9}$$

The form of (9) for B ensures that $\nabla \cdot B = 0$, and there is no mono pole or defect structure. equation (6) requires the pressure and density be of the form [11]:

$$P(y,z) = P(u)\, e^{\frac{-z}{h}}, \quad \rho(y,z) = \frac{1}{(gh)} P(u)\, e^{\frac{-z}{h}}, \tag{10}$$

where $h = \frac{R_0\, T_0}{g}$ is the scale height. Substituting equations (7-10) in equation (6), we obtain

$$\nabla^2 u + f(u)\, e^{\frac{-z}{h}} = 0, \tag{11}$$

where

$$f(u) = \mu \frac{dP}{du}. \tag{12}$$

Equation (12) gives

$$P(u) = P_0 + \frac{1}{\mu} \int f(u) du \tag{13}$$

Substituting equation (13) into equation (10), we obtain

$$P(y,z) = (P_0 + \frac{1}{\mu} \int f(u) du)\, e^{\frac{-z}{h}}, \tag{14}$$

$$\rho(y,z) = \frac{1}{gh}(P_0 + \frac{1}{\mu} \int f(u) du)\, e^{\frac{-z}{h}}, \tag{15}$$

where P_0 is constant. Taking transformation

$$x_1 + i\, x_2 = e^{\frac{-z}{l}}\, e^{\frac{iy}{l}} \tag{16}$$

equation (12) reduces to

$$\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + l^2\, f(u)\, e^{(\frac{2}{l} - \frac{1}{h})z} = 0. \tag{17}$$

These equations have been given in Khater et al. (2000).

4 Application of the Kudryashov method

In this section, we will investigate the Kudryashov method for solving specific forms of $f(u)$.

4.1 Liouville equation

We first consider Liouville equation and the following equation will be special case of equation (17). Let us assume f(u) has the form (Dungey, 1953; Low, 1975):

$$f(u) = -\alpha^2 A_0 e^{-\frac{A}{A_0}}, \tag{18}$$

where A_0 and α^2 are constants. Hence

$$P(y,z) = (P_0 + \frac{\alpha^2 A_0^2}{2\mu} e^{\frac{-2A}{A_0}}) e^{\frac{-z}{h}}. \tag{19}$$

Inserting equation (18) into equation (17) we obtain

$$\nabla^2 A/A_0 = l^2\alpha^2 e^{\frac{-2A}{A_0}+(\frac{2}{l}-\frac{1}{h})z}, \tag{20}$$

where $\nabla^2 = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$. Let us set

$$\frac{A}{A_0} = \frac{z}{L} + w(y,z), \tag{21}$$

where L is a constant. Then equation (20) becomes

$$\nabla^2 w - l^2\alpha^2 e^{-2w-(\frac{2}{L}+\frac{1}{h}-\frac{2}{l})z}. \tag{22}$$

Let us identify l by

$$\frac{2}{l} = \frac{2}{L} + \frac{1}{h}, \tag{23}$$

and inserting equation (23) into equation (22) we obtain a Liouville type

$$\phi_{xx} + \phi_{tt} - \alpha^2 l^2 e^{-2\phi} = 0. \tag{24}$$

In order to apply the Kudryashov method, we use the wave transformation $\xi = x - kt$ and change equation (24) into the form

$$(1+k^2)\phi'' = \alpha^2 l^2 e^{-2\phi}, \tag{25}$$

we next use the transformation

$$v = e^{-2\phi}, \tag{26}$$

we obtain

$$(1+k^2)vv'' - (1+k^2)(v')^2 + 2\alpha^2 l^2 u^3 = 0, \tag{27}$$

with balancing according step (1) we get $n = 2$, therefore the solution of (27) can be expressed as follow:

$$v(\xi) = \sum_{i=0}^{2} A_i (G(\xi))^i. \tag{28}$$

Substituting equation (28) along with (3) into (27) and setting the coefficients of all powers of G to zero, we obtain a system of nonlinear algebraic equations for A_0, A_1, A_2 . Solving the resulting system with the help of mathematica, we have the following sets of solutions:

$$\begin{cases} A_0 = 0, \\ A_1 = -\frac{cd(1+k^2)}{l^2\alpha^2}, \\ A_2 = -\frac{d^2(1+k^2)}{l^2\alpha^2}, \end{cases} \tag{29}$$

where $\xi = x - kt$, λ,α,l are constants. Therefore, substituting (29) in (28) and general solution (3) according to (4), we obtain solution of (27) as follows:

$$v_1(\xi) = -\frac{c^2 d(1+k^2)}{l^2\alpha^2} \frac{e^{(c(\xi+\xi_0))}}{(1-de^{(c(\xi+\xi_0))})^2}, \tag{30}$$

where $d < 0$, $c > 0$, $\xi = x - kt$. Using transformation

$$v = e^{-2\phi}, \tag{31}$$

we get solution of (24) as follows:

$$\phi_1(\xi) = -\frac{1}{2}\ln[-\frac{c^2 d(1+k^2)}{l^2\alpha^2} \frac{e^{(c(\xi+\xi_0))}}{(1-de^{(c(\xi+\xi_0))})^2}]. \tag{32}$$

Now substituting (29) in (28) and general solution (3) according to (5), we obtain solution of (27) as follows:

$$v_2(\xi) = \frac{c^2 d(1+k^2)}{l^2\alpha^2} \frac{e^{(c(\xi+\xi_0))}}{(1+de^{(c(\xi+\xi_0))})^2}, \tag{33}$$

where $d > 0, c < 0$, $\xi = x - kt$. Using transformation

$$v = e^{-2\phi}, \tag{34}$$

we get solution of (24) as follows:

$$\phi_2 = -\frac{1}{2}\ln[\frac{c^2 d(1+k^2)}{l^2\alpha^2} \frac{e^{(c(\xi+\xi_0))}}{(1+de^{(c(\xi+\xi_0))})^2}]. \tag{35}$$

4.2 Sinh-Poisson equation

In this section, we consider sinh-Poisson equation which plays an important role in the soliton model with BPS bound [4, 6]. Also, this equation will be special case of equation (17). If we assume

$$f(u) = -\frac{\beta^2}{4}(\frac{A_0}{h})\sinh(\phi). \tag{36}$$

The same as above we have

$$\phi_{xx} + \phi_{tt} = \beta^2 \sinh(\phi). \tag{37}$$

In order to apply the Kudryashov method, we use the wave transformation $\xi = x - kt$ and change equation (37) into the form

$$(1 + k^2)\phi'' = \beta^2 \sinh(\phi), \tag{38}$$

we next use the transformation

$$\begin{cases} v & = e^{\phi}, \\ \sinh(\phi) & = \frac{e^{\phi} - e^{-\phi}}{2}, \end{cases} \tag{39}$$

we obtain

$$2(1 + k^2)vv'' - 2(1 + k^2)\left(v'\right)^2 - \beta^2(v^3 - v) = 0. \tag{40}$$

With balancing according to step (1) we get $n = 2$, therefore the solution of (40) can be expressed as follows:

$$v(\xi) = \sum_{i=0}^{2} A_i (G(\xi))^i. \tag{41}$$

Substituting equation (41) along with (3) into (40) and setting the coefficients of all powers of G to zero, we obtain a system of nonlinear algebraic equations for A_0, A_1, A_2 . Solving the resulting system with the help of mathematica, we have the following sets of solutions:

$$\begin{cases} A_0 = 1, \\ A_1 = \frac{4d}{c}, \\ A_2 = \frac{4d^2}{c^2}, \\ c = \pm\frac{\beta}{\sqrt{1+k^2}}, \end{cases} \tag{42}$$

where $\xi = x - kt$, λ,β are constants. Therefore, using Substituting (42) in (41) and general solution (3) according to (4), we obtain solution of (40) as follows:

$$v_1(\xi) = \frac{\left(1 + de^{c(\xi+\xi 0)}\right)^2}{\left(1 - de^{c(\xi+\xi 0)}\right)^2}; \qquad d < 0, \quad c > 0, \tag{43}$$

where $c = \frac{\beta}{\sqrt{1+k^2}}$ when $\beta > 0$ or $c = -\frac{\beta}{\sqrt{1+k^2}}$ when $\beta < 0$ and $\xi = x - kt$. Using transformation

$$v = e^{\phi}, \tag{44}$$

we get solution of (37) as follows:

$$\phi_1(\xi) = \ln[\frac{\left(1 + de^{c(\xi+\xi 0)}\right)^2}{\left(1 - de^{c(\xi+\xi 0)}\right)^2}], \tag{45}$$

where $c = \frac{\beta}{\sqrt{1+k^2}}$ when $\beta > 0$ or $c = -\frac{\beta}{\sqrt{1+k^2}}$ when $\beta < 0$. Now with Substituting (42) in (41) and general solution (3) according to (5), we obtain solution of (40) as follows:

$$v_2(\xi) = \frac{\left(1 - de^{c(\xi+\xi 0)}\right)^2}{\left(1 + de^{c(\xi+\xi 0)}\right)^2}; \qquad d > 0, \quad c < 0, \tag{46}$$

where $c = \frac{\beta}{\sqrt{1+k^2}}$ when $\beta < 0$ or $c = -\frac{\beta}{\sqrt{1+k^2}}$ when $\beta > 0$ and $\xi = x - kt$. Using transformation

$$v = e^{\phi}, \tag{47}$$

we get solution of (37) as follows:

$$\phi_2 = \ln[\frac{\left(1 - de^{c(\xi+\xi 0)}\right)^2}{\left(1 + de^{c(\xi+\xi 0)}\right)^2}], \tag{48}$$

where $c = \frac{\beta}{\sqrt{1+k^2}}$ when $\beta < 0$ or $c = -\frac{\beta}{\sqrt{1+k^2}}$ when $\beta > 0$.

5 Conclusion

This study shows that the Kudryashov method is quite efficient and practical and is well suited for use in finding exact solutions for the Liouville and Sinh-Poisson equations. The reliability of the method and the reduction in the size of computational domain give this method a wider applicability. In this paper, the Kudryashov method has been successfully used to obtain some exact travelling wave solutions for the Liouville and Sinh-Poisson equations.

Acknowledgements

The authors are very grateful to the referees for their comments.

References

1. Heyvaerts, J., Larsy, J.M., Schatzman, M. and Witomsky, P. *Blowing up of twodimensional magnetohydrostatic equilibria by an increase of electric current or pressure*, Astronomy & Astrophysics 111 (1982) 104-112.

2. Jafari, H., Kadkhoda, N. and Khalique, C.M. *Travelling wave solutions of nonlinear evolution equations using the simplest equation method*, Comp. Math. Appl 64 (2012) 2084-2088.

3. Khater, A.H. *Analytical solutions for some nonlinear two-dimensional magnetostatic equilibria*, Astr. Space. Sci 162 (1989) 151-157.

4. Khater, A.H., El-Attary, M.A., El-Sabbagh, M.F. and Callebaut D.K. *Two-dimensional magnetohydrodynamic equilibria*, Astr. Space. Sci 149 (1988), 217-223.

5. Khater, A.H., El-Kalaawy, O.H and Callebaut D.K. *Backlund transformations and exact solutions for a nonlinear elliptic equation modelling isothermal magentostatic atmosphere*, IMA J. Appl. Math 65 (2000) 97-108.

6. Kudryashov, N.A. *One method for finding exact solutions of nonlinear differential equations*, Commun. Nonlin. Sci. Numer. SimuL 17 (2012) 2248-2253.

7. Kudryashov, N.A. *Simplest equation method to look for exact solutions of nonlinear differential equations*, Chaos. Solit. Fract 24 (2005) 1217-1231.

8. Lerche, I. and Low, B.C. *On the equilibrium of a cylindrical plasma supported horizontally by magnetic elds in uniform gravity*, Solar. Phys 67 (1980) 229-243.

9. Lerche, I. and Low, B.C. *Some nonlinear problems in astrophysics*, Physica D: Nonlinear Phenomena 4 (1982) 293-318.

10. Low, B.C. *Nonlinear Force-Free Magnetic Fields*, Rev. Geophys. Sp. Phys 20 (1982) 145-159.

11. Low, B.C. *Evolving force-free magnetic fields. III- Stages of nonequilibrium and the preflare stage*, The Astrophys. J 239 (1980) 377-388.

12. Martinov, N.k. and Vitanov, N.k. *On the correspondence between the self-consistent 2D Poisson- Boltzmann structures and the sine-Gordon waves*, J. Phys. A: Math. Gen. 25 (1992) L51.

13. Martinov, N.K. and Vitanov, N.K. *On self-consistent thermal equilibrium structures in two-dimensional negative-temperature systems*, Canadian Journal of Physics 72 (1994) 618-624.

14. Salehpour, E., Jafari, H. and Kadkhoda, N. *Application of* $(\frac{G'}{G})$*-expansion method to nonlinear Lienard equation*, Indian Journal of Sci. and Tech 5(2012) 2454-2456 .

15. Vitanov, N.K., Dimitrova, Z.I. and Kantz, H. *Modified method of simplest equation and its application to nonlinear PDEs*, Appl. Math. Comput. 216 (2010) 2587-2595.

16. Vitanov, N.K. *Modified method of simplest equation: Powerful tool for obtaining exact and approximate traveling-wave solutions of nonlinear PDEs*, Commun. Nonlin. Sci. Numer. SimuL 16 (2011) 1176-1185.

17. Webb, G.M. *Isothermal magnetostatic atmospheres. II - Similarity solutions with current proportional to the magnetic potential cubed*, Astrophys. J 327 (1988) 933-949.

18. Webb, G.M. and Zank, G.P. *Application of the sine-Poisson equation in solar magnetostatics*, Solar. Phys 127 (1990) 229-252.

19. Wazwaz, A.M. *Partial Differential Equations and Solitary Waves Theory*, Springer Verlag, (2009).

20. Zwingmann, W. *Theoretical study of onset conditions for solar eruptive processes*, Solar. Phys 111 (1987) 309-331.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 53-78

A nonstandard finite difference scheme for solving three-species food chain with fractional-order Lotka-Volterra model

S. Zibaei and M. Namjoo*

Abstract

In this paper, we introduce fractional-order for a model of tritrophic food chain Lotka-Volterra. Moreover, we discuss the stability analysis of fractional system. The nonstandard finite difference (NSFD) scheme is implemented to study the dynamic behaviors in the fractional-order Lotka-Volterra system. Numerical results show that the NSFD approach is easy to implement and accurate when applied to fractional-order Lotka-Volterra system.

Keywords: Fractional differential equations; Lotka-Volterra model; prey-predator system; Nonstandard finite difference scheme; Stability.

1 Introduction

Biological systems have been studied for many years. In these systems, it is common that state variables represent nonnegative quantities, such as concentrations, physical properties, the size of populations and the amount of chemical compounds [15]. These biological models are commonly based on the systems of ordinary differential equations (ODEs). Exact solutions of these systems are rarely in access and usually complicated; hence good approximations are required. Numerical methods are often the method of choice. They should describe the dynamic behavior of the systems, produce the nonnegative solutions, and reproduce the real dynamics of the biological systems. The interspecies interaction is among the most intensively explored fields of biology. The existance of many mathematical models in that area

*Corresponding author
Received 23 November 2014; revised 6 June 2015; accepted 1 August 2015
S. Zibaei
Department of Mathematics, School of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran. e-mail: s.zibaei@stu.vru.ac.ir

M. Namjoo
Department of Mathematics, School of Mathematical Sciences, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran. email: namjoo@vru.ac.ir

help us understand the population dynamics of analyzed biological systems. Mathematical models of predator-prey systems, characterized by decreasing growth rate of one of the interacting populations and increasing growth rate of the other, consist of systems of ODEs. In most of the modeled interactions, all rates of change are assumed to be time independent, which makes the corresponding systems autonomous. It is not always possible to find the exact solutions of the nonlinear models that have at least two ODEs. It is sometimes more useful to find numerical solutions of these types of systems in order to programme easily and visualize the results. By applying a numerical method on a continuous differential equation system, it becomes a difference equation system, i.e., discrete time system. While applying these numerical methods, it is necessary that the new difference equation system provide the positivity conditions and exhibit the same quantitative behaviours of continuous systems such as stability, bifurcation and chaos. It is well known that some traditional and explicit schemes such as forward Euler and Runge-Kutta are unsuccessful at generating oscillation, bifurcations, chaos and false steady states, despite using adaptative step size [13,17,18]. For forward Euler method, if the step size h is chosen small enough and the positivity conditions are satisfied, the local asymptotic stability for a fixed point is saved while in some special cases Hopf bifurcation cannot be seen. Instead of classical methods, NSFD schemes can alternatively be used to obtain more qualitative results and remove numerical instabilities. These schemes are developed for compensating the weaknesses, such as numerical instabilities that may be caused by standard finite difference methods. Also, the dynamic consistency can be represented by NSFD schemes [10]. The most important advantage of this scheme is that by choosing a convenient denominator function instead of the step size h, better results can be obtained. If the step size h is chosen small enough, the obtained results do not change significantly, but if the step size h gets larger this advantage comes into focus.

As it is well known, in the field of mathematical biology, the traditional Lotka-Volterra systems are very important mathematical models which describe multispecies population dynamics in a nonautonomous environment. Many important and interesting results of the dynamic behaviors for the Lotka-Volterra systems have been found in [3, 19, 20], such as the existence and uniqueness of solutions, the permanence, extinction, global asymptotic behavior and bifurcation. Because of the good memory and hereditary properties of fractional derivatives, it is often necessary to study the corresponding fractional systems. Therefore, the dynamical analysis of the fractional Lotka-Volterra systems has attracted a great deal of attention due to its theoretical and practical significance.

Many important results regarding stability of fractional systems have been obtained. For instance, the stability, existence, uniqueness and numerical solution of the fractional logistic equation are investigated in [7]. The stability and solutions of fractional predator-prey and rabies models are discussed in [1]. In addition, bifurcation properties of fractional systems have been

studied in some papers. For example, conditions for the occurrence of Hopf 's bifurcation are explored based on numerical simulations in [29]. The critical values of the fractional order are identified for which Hopf 's bifurcation may occur based on the stability analysis in [29]. Thus, it is significant to study the dynamical behaviors in the fractional population systems.

Analysis of fractional Lotka-Volterra equations which are obtained from the classical Lotka-Volterra equations in mathematical modeling by the replacing first order derivatives by fractional derivative of order α $(0 < \alpha \leq 1)$ have been the focus of recent research in this field. Lots of universal phenomena can be modeled to a greater degree of accuracy by using the property of these evolution equations. The fractional differential equations have gained much attention recently due to the fact that fractional order system response ultimately converges to the integer order system response.

The current technological advance has made it possible for humans to disturb the environmental balance in nature that may cause immense damages, such as species extinction or starvation. Therefore, understanding the behaviour of the interaction between the species may help biologists and other related parties to prevent those events from happening. The real interaction of prey-predator in nature is complex and comprises both interspecies and external environmental factors. Therefore, several simplifications are usually assumed so that a basic model can be constructed and then developed or modified to approach the real system.

The Lotka-Volterra equations are a system of ODEs in the following form:

$$x' = ax - bxy,$$

$$y' = -cy + dxy,$$
$$x(0) = x_0, \qquad y(0) = y_0,$$

where x and y are prey and predactor, respectively. Here a is the prey growth rate in the absence of the predators, b is the capture rate of prey per predator, d is the rate at which each predator converts captured prey into predator births and c is the constant rate at which death occurs in the absence of prey. They show that ditrophic food chains (i.e. prey-predator systems) permanently oscillate for any initial conditions if the prey growth rate is constant and the predator functional response is linear.

The classical food chain models with only two trophic levels are shown to be insufficient to produce realistic dynamics [5]. Therefore, in this paper, by modifying the classical Lotka-Volterra model, we analyse and simulate the dynamics of a three-species food chain interaction. With non-dimensionalisation, the system of three-species food chain can be written as

$$\begin{aligned} x' &= ax - bxy, \\ y' &= dxy - cy - eyz, \\ z' &= gzy - fz, \\ x(0) &= x_0, \qquad y(0) = y_0, \qquad z(0) = z_0, \end{aligned} \tag{1}$$

where x, y and z denote the non-dimensional population density of the prey, predator and top predator, respectively. The predator y preys on x and the predator z preys on y. Furthermore a, b, c, d, e, f and g are the intrinsic growth rate of the prey, the death rate of the predator, the death rate of the top predator, predation rate of the predator, the conversion rate, predation rate of the top predator and the conversion rate, respectively.

This paper is organized as follows: In the next section, we give some basic definitions and properties of the Grünwald-Letnikov (GL) approximation and provide a brief overview of the important feature of the procedures for constructing NSFD schemes for ODEs. In Section 3, we introduce fractional order into the model that describes Lotka-Volterra system and also stability theorem and fractional Routh-Hurwitz stability conditions are given for the local asymptotic stability of the fractional systems. In Section 4, we will discuss the stability analysis of fractional system. In Section 5, we present the idea of NSFD scheme for solving the fractional order Lotka-Volterra model. Finally in the last section, numerical results show that the NSFD approach is easy to be implemented and accurated when applied to fractional-order Lotka-Volterra system.

2 Preliminaries and notations

In this section, some basic definitions and properties of the fractional calculus theory and nonstandard discretization are discussed.

2.1 Fundamentals of fractional-order

Fractional differential equations (FDEs) have gained considerable importance due to their application in various sciences, such as physics, mechanics, chemistry and engineering [16]. In the recent years, the dynamic behaviors of fractional-order differential systems have received increasing attention. Although the concept of the fractional calculus was discussed in the same time interval of integer-order calculus, the complexity and the lack of applications postponed its progress till a few decades ago. Recently, most of the dynamical systems based on the integer-order calculus have been modified into the fractional order domain due to the extra degrees of freedom and the flexibility which can be used to precisely fit the experimental data much better

than the integer-order modeling. For example, new fundamentals have been investigated in the fractional-order domain for the first time and do not exist in the integer-order systems such as those presented in [9, 16].

2.2 GL approximation

The GL method of approximation for the one-dimensional fractional derivative is as follows [16]:

$$D^{\alpha}x(t) = f(t, x(t)), \qquad x(0) = x_0, \qquad t \in [0, t_f], \tag{2}$$

$$D^{\alpha}x(t) = \lim_{h \to 0} h^{-\alpha} \sum_{j=0}^{[\frac{t_f}{h}]} (-1)^j \binom{\alpha}{j} x(t - jh),$$

where $0 < \alpha < 1$, D^{α} denotes the fractional derivative and h is the step size and $[\frac{t_f}{h}]$ denotes the integer part of $\frac{t_f}{h}$. Therefore, Eq. (2) is discretized as follows:

$$\sum_{j=0}^{n} c_j^{\alpha} x_{n-j} = f(t_n, x_n), \qquad n = 1, 2, 3, ...$$

where $t_n = nh$ and c_j^{α} are the GL coefficients defined as:

$$c_j^{\alpha} = (1 - \frac{1+\alpha}{j}) c_{j-1}^{\alpha}, \qquad c_0^{\alpha} = h^{-\alpha}, \qquad j = 1, 2, 3, ...$$

2.3 NSFD discretization

The initial foundation of NSFD schemes came from the exact finite difference schemes. These schemes are well developed by Mickens [13, 14] in the past decades. These schemes are developed for compensating the weaknesses such as numerical instabilities that may be caused by standard finite difference methods. Regarding the positivity, boundedness and monotonicity of solutions, NSFD schemes have a better performance over the standard finite difference schemes, due to flexibility to construct a NSFD scheme that can preserve certain properties and structures, which are obeyed by the original equations.

The advantages of NSFD schemes have been shown in many numerical applications. Gonzalez-Parra et al. [4] developed NSFD schemes to solve population and biological models. Jordan [8] constructed NSFD schemes for heat transfer problems.

We now give an outline of the critical points which will allow the construction of NSFD discretizations for ODEs.

Consider the autonomous ODE given by

$$x' = f(x), \qquad x(0) = x_0, \qquad t \in [0, t_f],$$

where $f(x)$ is, in general, a nonlinear function of x. For a discrete-time grid with step size, $\triangle t = h$, we replace the independent variable t by

$$t \approx t_n = nh, \qquad n = 0, 1, 2, \ldots, N$$

where $h = \frac{t_f}{N}$. The dependent variable $x(t)$ is replaced by

$$x(t) \approx x_n,$$

where x_n is the approximation of $x(t_n)$.

The first NSFD requirement is that the dependent functions should be modeled on the discrete-time computational grid. Particular examples of this include the following functions [13, 14].

$$\begin{cases} xy \approx 2x_{n+1}y_n - x_{n+1}y_{n+1}, \\ x^2 \approx x_{n+1}x_n, \\ x^3 \approx (\dfrac{x_{n+1} + x_{n-1}}{2})x_n^2. \end{cases}$$

A standard way for representing a discrete first-derivative is given by

$$x' \cong \frac{x_{n+1} - x_n}{h}.$$

However, the NSFD scheme requires that x' has a more general representation

$$x' \cong \frac{x_{n+1} - x_n}{\phi},$$

where the denominator function, i.e. ϕ has the following properties:

(i) $\phi(h) = h + O(h^2)$,

(ii) $\phi(h)$ is an increasing function of h,

(iii) $\phi(h)$ may depend on the parameters appearing in the differential equations.

The paper by Mickens [14] gives a general procedure for determining $\phi(h)$ for systems of ODEs. An example of the NSFD discretization process is its application to the decay equation

$$x' = -\lambda x,$$

where λ is a constant. The discretization scheme is as follows [14]

$$\frac{x_{n+1} - x_n}{\phi} = -\lambda x_n, \qquad \phi(h, \lambda) = \frac{1 - e^{-\lambda h}}{\lambda}.$$

Another example is given by

$$x' = \lambda_1 x - \lambda_2 x^2,$$

where the NSFD scheme is

$$\frac{x_{n+1} - x_n}{\phi} = \lambda_1 x_n - \lambda_2 x_{n+1} x_n, \qquad \phi(h, \lambda_1) = \frac{e^{\lambda_1 h} - 1}{\lambda_1}.$$

It should be noted that the NSFD schemes for these two ODEs are exact in the sense that $x_n = x(t_n)$ for all applicable values of $h > 0$. In general, for an ODE with polynomial terms,

$$x' = ax + (NL) \qquad NL \equiv \text{Nonlinear terms},$$

the NSFD discretization for the linear expressions is given by Mickens [14]

$$\frac{x_{n+1} - x_n}{\phi} = ax_n + (NL)_n,$$

where the denominator function is

$$\phi(h, a) = \frac{e^{ah} - 1}{a}.$$

It follows that if x' is a function of x which does not have a linear term, then the denominator function is just h, i.e. $\phi(h) = h$.

By applying this technique and using the GL discretization method, the following relations are yielded:

$$x_{n+1} = \frac{-\sum_{j=1}^{n+1} c_j^{\alpha} x_{n+1-j} + f(t_{n+1}, x_{n+1})}{c_0^{\alpha}}, \qquad n = 0, 1, 2, ...$$

where $c_0^{\alpha} = \phi(h)^{-\alpha}$.

3 Fractional-order Lotka-Volterra model

Now we introduce fractional-order into the model (1) of Lotka-Volterra chaotic system. The new system is described by the following set of fractional ODEs of order $\alpha_1, \alpha_2, \alpha_3 > 0$, in the following form

$$\begin{aligned} &D^{\alpha_1} x(t) = ax - bxy, \\ &D^{\alpha_2} y(t) = dxy - cy - eyz, \\ &D^{\alpha_3} z(t) = gzy - fz, \\ &x(0) = x_0, \qquad y(0) = y_0, \qquad z(0) = z_0, \\ &0 < \alpha_i \leq 1, \qquad i = 1, 2, 3. \end{aligned} \tag{3}$$

Now, stability theorem on fractional-order systems, fractional Routh-Hurwitz stability conditions and their related results are introduced. The first stability theorem has been given for incommensurate fractional-order systems.

Theorem 1. *([12]) Consider the incommensurate fractional-order system*

$$D^{\alpha} x(t) = f(x(t)), \qquad x(0) = x_0, \tag{4}$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\alpha_i \in (0, 1]$ for $i = 1, 2, \ldots, n$ and $x \in \mathbb{R}^n$. The equilibrium points of (4), are calculated by solving the equations:

$$f(x) = 0.$$

These points are locally asymptotically stable if all eigenvalues λ of the Jacobian matrix $J \equiv \frac{\partial f}{\partial x}$ evaluated at the equilibrium points satisfy:

$$|arg(\lambda)| > \frac{\alpha^{\star}\pi}{2}, \qquad \alpha^{\star} = \max(\alpha_1, \ldots, \alpha_n).$$

Theorem 2. *([11]) Consider the commensurate fractional-order system (4), i.e., $\alpha_1 = \alpha_2 = \cdots = \alpha_n = \alpha^{\star}$. If all eigenvalues of the Jacobian matrix of an equilibrium point satisfy:*

$$|arg(\lambda)| > \frac{\alpha^{\star}\pi}{2},$$

then, the fractional system is locally asymptotically stable at the equilibrium point.

Consider the system of ODEs given by

$$\begin{aligned} X' &= F(X, Y, Z), \\ Y' &= G(X, Y, Z), \\ Z' &= H(X, Y, Z), \end{aligned} \tag{5}$$

where F, G and H are nonlinear functions. Let $\bar{X}$, $\bar{Y}$ and $\bar{Z}$ be the steady-state solution, i.e.,

$$F(\bar{X}, \bar{Y}, \bar{Z}) = G(\bar{X}, \bar{Y}, \bar{Z}) = H(\bar{X}, \bar{Y}, \bar{Z}) = 0.$$

Now consider small perturbations to the steady-state solutions

$$\begin{aligned} X(t) &= \bar{X} + x(t), \\ Y(t) &= \bar{Y} + y(t), \\ Z(t) &= \bar{Z} + z(t). \end{aligned}$$

Frequently these are called perturbations of the steady-state. Substituting, we arrive at

$$\begin{aligned} (\bar{X} + x)' &= F(\bar{X} + x, \bar{Y} + y, \bar{Z} + z), \\ (\bar{Y} + y)' &= G(\bar{X} + x, \bar{Y} + y, \bar{Z} + z), \\ (\bar{Z} + z)' &= H(\bar{X} + x, \bar{Y} + y, \bar{Z} + z). \end{aligned}$$

On the left-hand side we expand the derivatives and that by definition

$$\bar{X}' = \bar{Y}' = \bar{Z}' = 0.$$

On the right-hand side we now expand F, G and H in a Taylor series about the point $(\bar{X}, \bar{Y}, \bar{Z})$. The result is

$$\begin{aligned} x' &= F(\bar{X}, \bar{Y}, \bar{Z}) + F_x(\bar{X}, \bar{Y}, \bar{Z})x + F_y(\bar{X}, \bar{Y}, \bar{Z})y \\ &+ F_z(\bar{X}, \bar{Y}, \bar{Z})z + \text{terms of order } x^2, y^2, z^2, xy, \\ &yz, xz, \text{and higher}, \end{aligned}$$

$$\begin{aligned} y' &= G(\bar{X}, \bar{Y}, \bar{Z}) + G_x(\bar{X}, \bar{Y}, \bar{Z})x + G_y(\bar{X}, \bar{Y}, \bar{Z})y \\ &+ G_z(\bar{X}, \bar{Y}, \bar{Z})z + \text{terms of order } x^2, y^2, z^2, xy, \\ &yz, xz, \text{and higher}, \end{aligned}$$

$$\begin{aligned} z' &= H(\bar{X}, \bar{Y}, \bar{Z}) + H_x(\bar{X}, \bar{Y}, \bar{Z})x + H_y(\bar{X}, \bar{Y}, \bar{Z})y \\ &+ H_z(\bar{X}, \bar{Y}, \bar{Z})z + \text{terms of order } x^2, y^2, z^2, xy, \\ &yz, xz, \text{and higher}. \end{aligned}$$

Again by definition,

$$F(\bar{X}, \bar{Y}, \bar{Z}) = G(\bar{X}, \bar{Y}, \bar{Z}) = H(\bar{X}, \bar{Y}, \bar{Z}) = 0,$$

so we are left with

$$\begin{aligned} x' &= a_{11}x + a_{12}y + a_{13}z, \\ y' &= a_{21}x + a_{22}y + a_{23}z, \\ z' &= a_{31}x + a_{32}y + a_{33}z, \end{aligned}$$

where the matrix of coefficients

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$
$$= \begin{pmatrix} F_x(\bar{X},\bar{Y},\bar{Z}) & F_y(\bar{X},\bar{Y},\bar{Z}) & F_z(\bar{X},\bar{Y},\bar{Z}) \\ G_x(\bar{X},\bar{Y},\bar{Z}) & G_y(\bar{X},\bar{Y},\bar{Z}) & G_z(\bar{X},\bar{Y},\bar{Z}) \\ H_x(\bar{X},\bar{Y},\bar{Z}) & H_y(\bar{X},\bar{Y},\bar{Z}) & H_z(\bar{X},\bar{Y},\bar{Z}) \end{pmatrix},$$

is the Jacobian of the system (5). Hence, the problem has been reduced to a linear system, i.e., $w' = Aw$ with $w = (x, y, z)^T$, for states that are in proximity to the steady-state $(\bar{X}, \bar{Y}, \bar{Z})$.

The Jacobian matrix J of the system (3) at the equilibrium point $E = (x^*, y^*, z^*)$ is computed as

$$J(E) = \begin{pmatrix} a - by^* & -bx^* & 0 \\ dy^* & -c + dx^* - ez^* & -ey^* \\ 0 & gz^* & -f + gy^* \end{pmatrix}. \tag{6}$$

The existence and local stability conditions of these equilibrium points are as follows:

Let $D(P)$ denotes the discriminant of a polynomial P

$$P(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0, \tag{7}$$

and

$$D(P) = 18a_1a_2a_3 + (a_1a_2)^2 - 4a_3(a_1)^3 - 4(a_2)^3 - 27(a_3)^2,$$

using the results of [2], we have the following Routh-Hurwitz stability conditions for FDEs:

(i) If $D(P) > 0$, then the necessary and sufficient condition for the equilibrium point E to be locally asymptotically stable is $a_1 > 0, a_3 > 0, a_1a_2 - a_3 > 0$.

(ii) If $D(P) < 0$, $a_1 \geq 0, a_2 \geq 0, a_3 > 0$, then the equilibrium point E is locally asymptotically stable for $\alpha < 2/3$. However, if $D(P) < 0, a_1 < 0, a_2 < 0, \alpha > 2/3$, then all roots of polynomial (7) satisfy the condition $|arg(\lambda)| < \frac{\alpha\pi}{2}$.

(iii) If $D(P) < 0$, $a_1 > 0, a_2 > 0, a_1a_2 - a_3 = 0$, then the equilibrium point E is locally asymptotically stable for all $\alpha \in [0, 1)$.

(iv) The necessary condition for the equilibrium point E to be locally asymptotically stable is $a_3 > 0$.

In the next section, we discuss the asymptotic stability of the equilibrium point E of the system (3).

4 Stability analysis of the model

To evaluate the equilibrium points of the system (3), let

$$\begin{aligned} ax - bxy &= 0, \\ dxy - cy - eyz &= 0, \\ gzy - fz &= 0, \end{aligned}$$

then the equilibrium points are $E_0 = (0,0,0)$, $E_1 = (0, \frac{f}{g}, -\frac{c}{e})$ and $E_2 = (\frac{c}{d}, \frac{a}{b}, 0)$. All calculations were performed by MAPLE. The local stability conditions of these equilibrium points are as follows:

(i) The Jacobian matrix (6) at the equilibrium point $E_0 = (0,0,0)$ is

$$J(0,0,0) = \begin{pmatrix} a & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & -f \end{pmatrix}, \tag{8}$$

with the characteristic equation

$$P(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0,$$

where

$$a_1 = f + c - a, \qquad a_2 = cf - af - ac, \qquad a_3 = -fac,$$

and $D(P)$ in the above equation is

$$D(P) = (c-f)^2(a+f)^2(a+c)^2.$$

Therefore, the eigenvalues of the Jacobian matrix (8) corresponding to the equilibrium point E_0 are $\lambda_1 = a$, $\lambda_2 = -c$ and $\lambda_3 = -f$.

Clearly, if $c \neq f$ then $D(P) > 0$. Now, since $a_3 < 0$; therefore, based on part (i) in Routh-Hurwitz stability conditions, the equilibrium point E_0 is unstable.

(ii) The Jacobian (6) at the equilibrium point $E_1 = (0, \frac{f}{g}, -\frac{c}{e})$ is

$$J(0, \frac{f}{g}, -\frac{c}{e}) = \begin{pmatrix} \frac{ag - bf}{g} & 0 & 0 \\ \frac{fd}{g} & 0 & -\frac{ef}{g} \\ 0 & -\frac{gc}{e} & 0 \end{pmatrix},$$

where the characteristic equation is

$$P(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0,$$

with

$$a_1 = \frac{bf - ag}{g}, \qquad a_2 = -cf, \qquad a_3 = \frac{cf(ag - bf)}{g},$$

and again

$$D(P) = \frac{4cf(g^2(a^2 - cf) + bf(bf - 2ag))^2}{g^4}.$$

Here, the corresponding eigenvalues are

$$\lambda_1 = \frac{ag - bf}{g}, \qquad \lambda_2 = \sqrt{cf}, \qquad \lambda_3 = -\sqrt{cf}.$$

Obviously, if $g^2(a^2 - cf) + bf(bf - 2ag) \neq 0$ then $D(P) > 0$. Then as $a_1a_2 - a_3 = 0$; therefore, based on part (i) in Routh-Hurwitz stability conditions, the equilibrium point E_1 is an unstable point.

(iii) The Jacobian (6) at the equilibrium point $E_2 = (\frac{c}{d}, \frac{a}{b}, 0)$ is

$$J(\frac{c}{d}, \frac{a}{b}, 0) = \begin{pmatrix} 0 & -\frac{bc}{d} & 0 \\ \frac{ad}{b} & 0 & -\frac{ea}{b} \\ 0 & 0 & \frac{ag - bf}{b} \end{pmatrix}, \tag{9}$$

In this case, the characteristic equation is also

$$P(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3 = 0,$$

where

$$a_1 = -\frac{ag - bf}{b}, \qquad a_2 = ac, \qquad a_3 = -\frac{(ag - bf)ac}{b},$$

and

$$D(P) = -\frac{4ac(bf(bf - 2ag) + g^2a^2 + b^2ca)^2}{b^4}.$$

Therefore, the eigenvalues of the Jacobian matrix (9) corresponding to the equilibrium point E_2 are

$$\lambda_1 = \frac{ag - bf}{b}, \qquad \lambda_2 = i\sqrt{ac}, \qquad \lambda_2 = -i\sqrt{ac}.$$

Clearly, if $bf(bf-2ag)+g^2a^2+b^2ca \neq 0$ then $D(P)<0$. Now if $bf>ag$ then $a_1>0, a_2>0, a_1a_2-a_3=0$ and based on part (iii) in Routh-Hurwitz stability conditions the equilibrium point E_2 is locally asymptotically stable for all $\alpha \in [0,1)$.

5 NSFD for fractional-order Lotka-Volterra model

For system (3) and applying Mickens scheme by replacing the step size h by a function $\phi(h)$ and using the GL discretization method, the following equations are obtained:

$$\begin{aligned}
\sum_{j=0}^{n+1} c_j^{\alpha_1} x_{n+1-j} &= ax_n - bx_{n+1}y_n,\\
\sum_{j=0}^{n+1} c_j^{\alpha_2} y_{n+1-j} &= -cy_{n+1} + dx_{n+1}y_n - ey_{n+1}z_n,\\
\sum_{j=0}^{n+1} c_j^{\alpha_3} z_{n+1-j} &= -fz_{n+1} + gz_ny_{n+1}.
\end{aligned} \tag{10}$$

Comparing equations (10) with system (3), we note the following:
1. The linear and nonlinear terms on the right-hand side of the first equation in system (3) are in the forms

$$x \approx x_n, \qquad -xy \approx -x_{n+1}y_n.$$

2. The linear and nonlinear terms on the right-hand side of the second equation in (3) are

$$-y \approx -y_{n+1}, \qquad xy \approx x_{n+1}y_n, \qquad -yz \approx -y_{n+1}z_n.$$

3. The linear and nonlinear terms on the right-hand side of the third equation in (3) are

$$-z \approx -z_{n+1}, \qquad zy \approx z_ny_{n+1}.$$

Doing some algebraic manipulations to equations (10) yields the following relations:

$$
\begin{aligned}
x_{n+1} &= \frac{-\sum_{j=1}^{n+1} c_j^{\alpha_1} x_{n+1-j} + a x_n}{c_0^{\alpha_1} + b y_n}, \\
y_{n+1} &= \frac{-\sum_{j=1}^{n+1} c_j^{\alpha_2} y_{n+1-j} + d x_{n+1} y_n}{c_0^{\alpha_2} + c + e z_n}, \\
z_{n+1} &= \frac{-\sum_{j=1}^{n+1} c_j^{\alpha_3} z_{n+1-j} + g z_n y_{n+1}}{c_0^{\alpha_3} + f},
\end{aligned}
\tag{11}
$$

where

$$
c_0^{\alpha_1} = \phi_1(h)^{-\alpha_1}, \qquad c_0^{\alpha_2} = \phi_2(h)^{-\alpha_2}, \qquad c_0^{\alpha_3} = \phi_3(h)^{-\alpha_3},
$$

with [21]

$$
\phi_1(h) = \frac{e^{ah} - 1}{a}, \qquad \phi_2(h) = \frac{e^{ch} - 1}{c}, \qquad \phi_3(h) = \frac{e^{fh} - 1}{f}.
$$

Proposition 1. *The numerical solutions obtained from system (11) for case* $0 < \alpha_i \leq 1, \quad i = 1, 2, 3$ *satisfy*

$$
\begin{array}{ll}
x_n > 0 & x_{n+1} > 0 \\
y_n > 0 \Rightarrow & y_{n+1} > 0 \\
z_n > 0 & z_{n+1} > 0
\end{array}
\tag{12}
$$

for all the relevant values of n.

Proof. Since $c_0^{\alpha_i} > 0$ and by recursive relation

$$
c_j^{\alpha_i} = (1 - \frac{1 + \alpha_i}{j}) c_{j-1}^{\alpha_i}, \qquad j = 1, 2, 3, ...
$$

we have $c_j^{\alpha_i} < 0, \quad j > 0$. Now system (11) shows that relations (12) is established. For case $\alpha_i = 1, \quad i = 1, 2, 3$ we should consider the following system:

$$
\begin{aligned}
\frac{x_{n+1} - x_n}{\phi_1} &= a x_n - b x_{n+1} y_n, \\
\frac{y_{n+1} - y_n}{\phi_2} &= -c y_{n+1} + d x_{n+1} y_n - e y_{n+1} z_n, \\
\frac{z_{n+1} - z_n}{\phi_3} &= -f z_{n+1} + g z_n y_{n+1}.
\end{aligned}
$$

By solving this system for x_{n+1}, y_{n+1} and z_{n+1} we conclude that relation (12) holds. □

6 Numerical results

Analytical studies always remain incomplete without numerical verification of the results. In this section, we present numerical simulation to illustrate the results obtained in the previous sections. The numerical experiments are designed to show the dynamical behaviour of the system in three main different sets of parameters and initial conditions:

(i) The case where $bf = ag$,

(ii) The case where $bf > ag$,

(iii) The case where $bf < ag$.

To show the dynamics of the system (3), set the parameter $a = b = c = d = e = f = 1$ given as fixed parameters and g as a varied parameter.

(i) The case where $bf = ag$
For the case $bf = ag$ the equilibrium point E_2 has three eigenvalues with zero real part corresponding with stable centre point in xy plane . We consider the case $\alpha_1 = \alpha_2 = \alpha_3 = 1$ which corresponds to the classical Lotka-Volterra system. Figures 1 and 2 represents the phase portrait for solutions where parameter $g = 1$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 40s and step size $h = 0.1$ and $h = 0.5$. In this case, prey x, predator y and top predator z persist and have populations that vary periodically over time in a common period.

Once again an equilibrium is achieved within the system, such that each predator population increases as the population of its respective prey increases. Each predator population also peaks and then begins to decrease shortly after its respective prey population peaks and begins to decrease. The plots of populations x and y are essentially the same as they were in the $2D$ system, and the new predator population z behaves similarly with respect to y as y behaves with respect to x. All three populations share a common period.

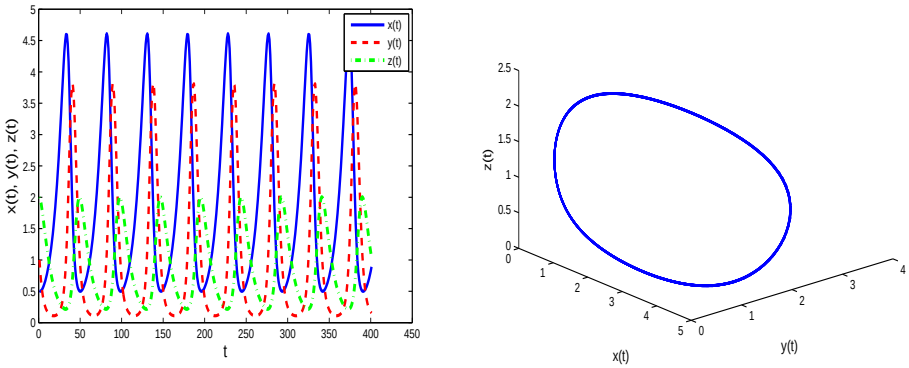

Figure 1: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and $h = 0.1$.

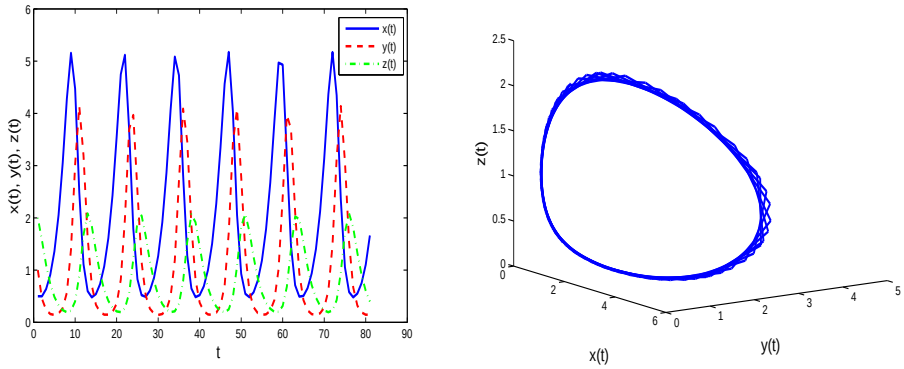

Figure 2: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and $h = 0.5$.

Figures 3 and 4 depict the phase trajectory of the fractional-order Lotka-Volterra chaotic system (3) for commensurate order $\alpha_1 = \alpha_2 = \alpha_3 = 0.90$ and parameters $g = 1$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 40s and step size $h = 0.1$ and $h = 0.5$.

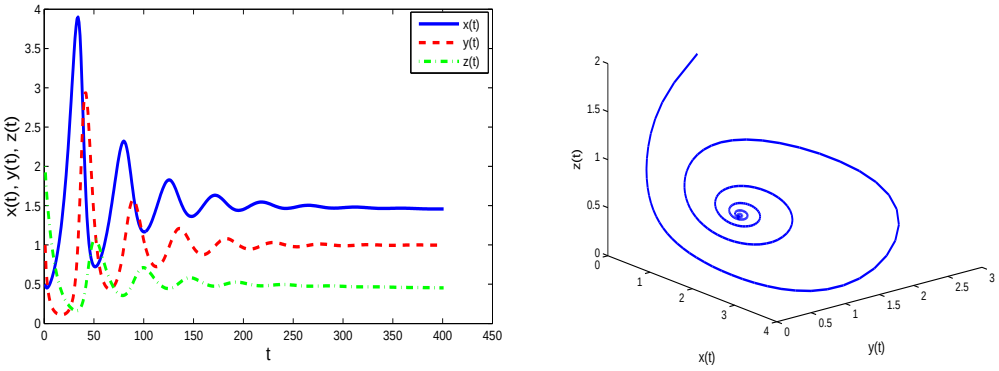

Figure 3: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 0.90$ and $h = 0.1$.

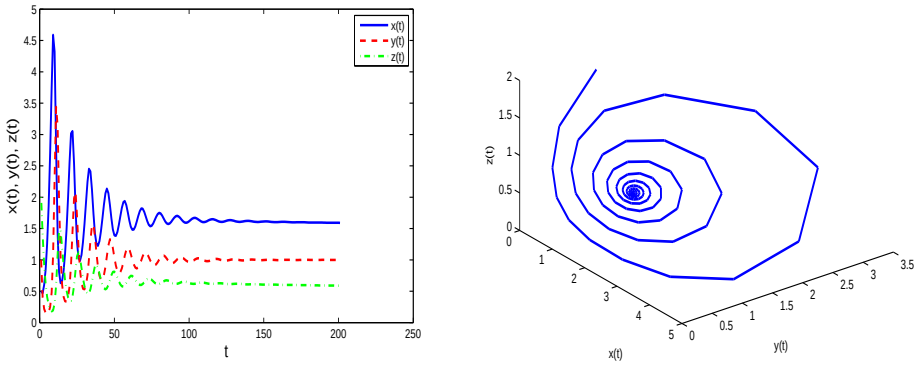

Figure 4: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 0.90$ and $h = 0.5$.

Figures 5 and 6 depict the phase trajectory of the fractional-order Lotka-Volterra chaotic system for incommensurate order and parameters $g = 1$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 40s and step size $h = 0.1$ and $h = 0.5$.

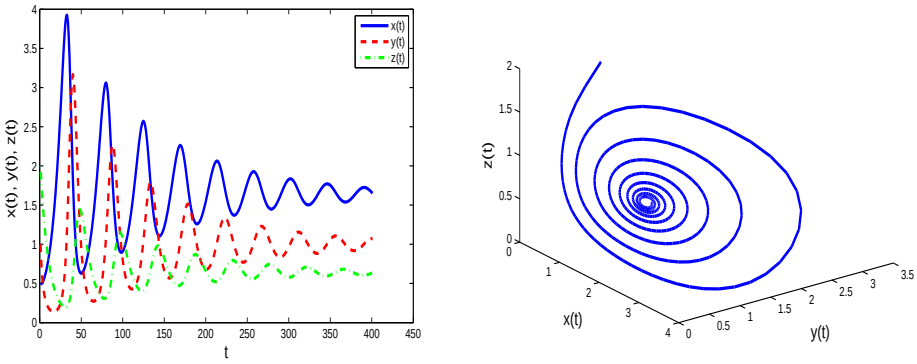

Figure 5: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = 0.99, \alpha_2 = 0.95, \alpha_3 = 0.90$ and $h = 0.1$.

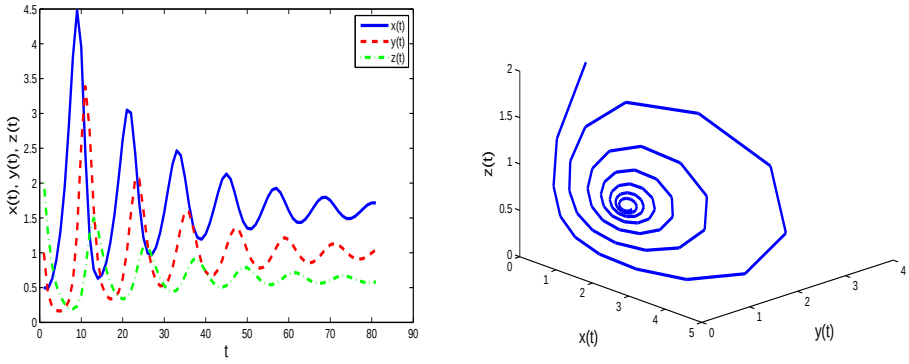

Figure 6: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = 0.95, \alpha_2 = 0.90, \alpha_3 = 0.80$ and $h = 0.5$.

In Figure 7, the phase trajectory of the fractional-order Lotka-Volterra chaotic system is depicted for incommensurate order and parameters $a = 1, b = 2, c = 5, d = 4, e = 3, f = 3, g = 6$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 40s and step size $h = 0.1$.

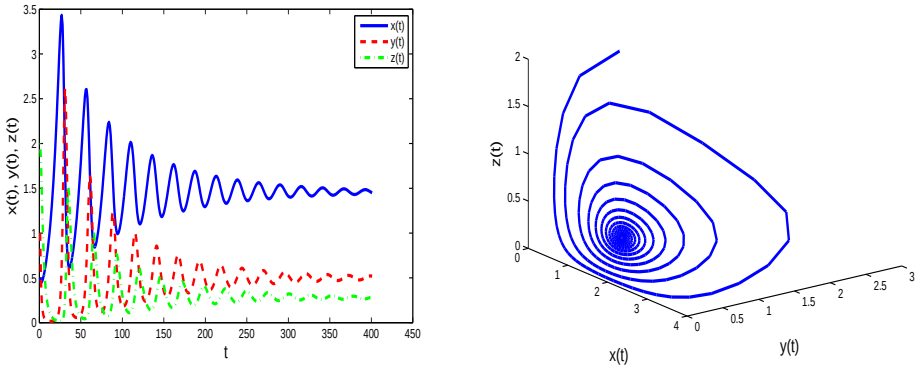

Figure 7: Plot of populations x, y and z over time for the case $bf = ag$ with $\alpha_1 = 0.99, \alpha_2 = 0.95, \alpha_3 = 0.90$ and $h = 0.1$.

(ii) The case where $bf > ag$

For the case where $bf > ag$, two eigenvalues for E_2 are pure imaginary initially-spiral stability corresponding with centre manifold in xy plane and one negative real eigenvalue corresponding with stable one-dimensional invariant curve in z axis. Hence, the equilibrium point E_2 is locally stable spiral sink. On the other hand, prey x and predator y persist and has populations that vary periodically over time with a common period. The solutions are plotted in Figures 8 and 9 for commensurate order $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and parameters $g = 0.88$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 100s and step size $h = 0.1$ and $h = 0.5$.

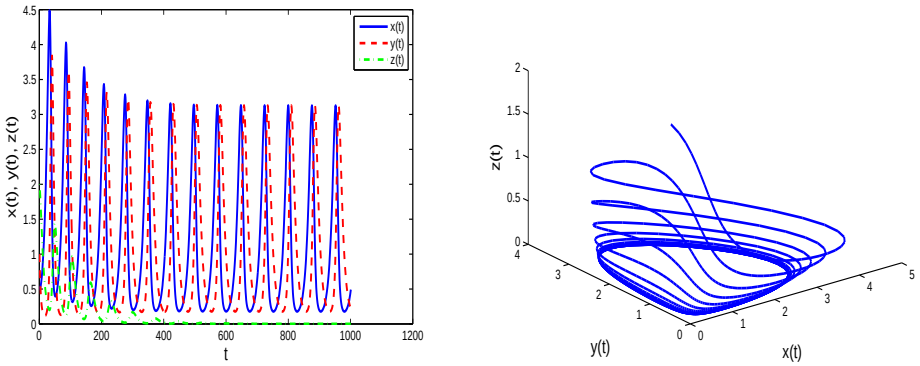

Figure 8: Plot of populations x, y and z over time for the case $bf > ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and $h = 0.1$.

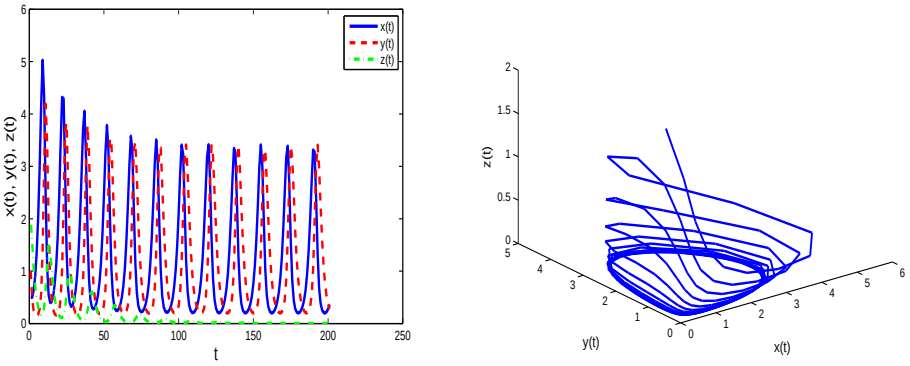

Figure 9: Plot of populations x, y and z over time for the case $bf > ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and $h = 0.5$.

Figures 10 and 11 depict the phase trajectory of the fractional-order Lotka-Volterra chaotic system (3) for incommensurate order $\alpha_1 = 0.90$, $\alpha_2 = 0.80$, $\alpha_3 = 0.70$ and parameters $g = 0.88$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 100s and step size $h = 0.1$ and $h = 0.5$.

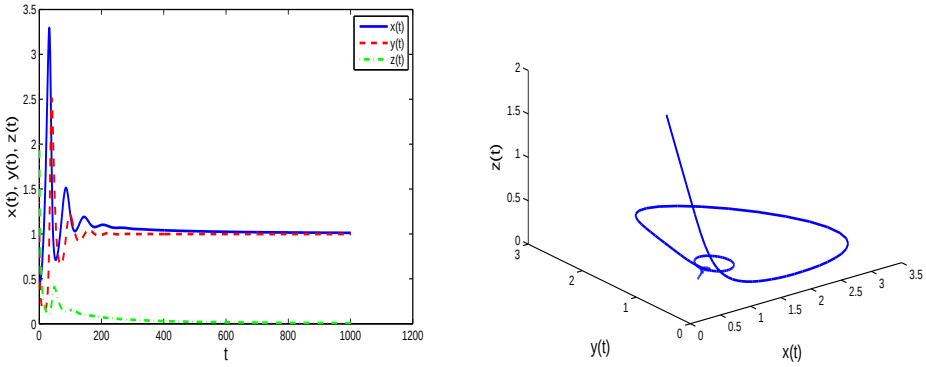

Figure 10: Plot of populations x, y and z over time for the case $bf > ag$ with $\alpha_1 = 0.90, \alpha_2 = 0.80, \alpha_3 = 0.70$ and $h = 0.1$.

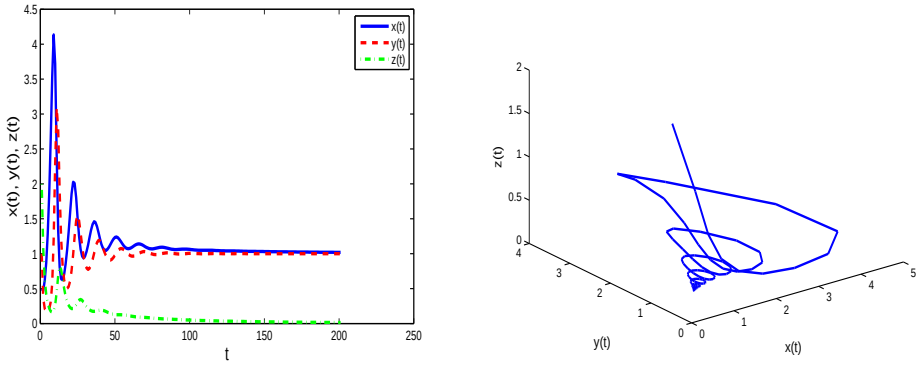

Figure 11: Plot of populations x, y and z over time for the case $bf > ag$ with $\alpha_1 = 0.90, \alpha_2 = 0.80, \alpha_3 = 0.70$ and $h = 0.5$.

(iii) The case where $bf < ag$

For the case where $bf < ag$, two eigenvalues for E_2 is pure imaginary initially-spiral stability corresponding with centre manifold in xy plane and one positive real eigenvalue corresponding to unstable one-dimensional invariant curve in z axes. Hence the equilibrium point E_2 is a locally unstable spiral source. In this case, the prey x and top predator z can survive, growing periodically unstable. On the other hand, predator y persists and has populations that vary periodically stable. The solutions for this case are shown in Figure 12 for commensurate order $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and parameters $g = 1.6$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 50s and step size $h = 0.1$.

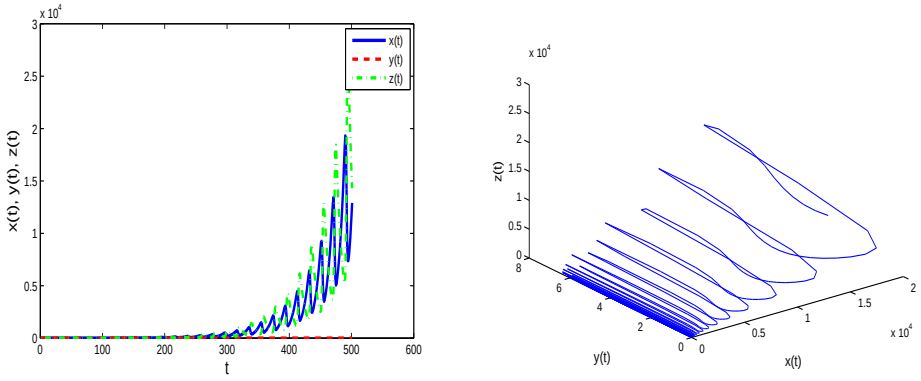

Figure 12: Plot of populations x, y and z over time for the case $bf < ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 1$ and $h = 0.1$.

In Figure 13 the phase trajectory of the fractional-order Lotka-Volterra chaotic system (3) is depicted for commensurate order $\alpha_1 = \alpha_2 = \alpha_3 =$

0.50 and parameters $g = 1.6$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 50s and step size $h = 0.5$.

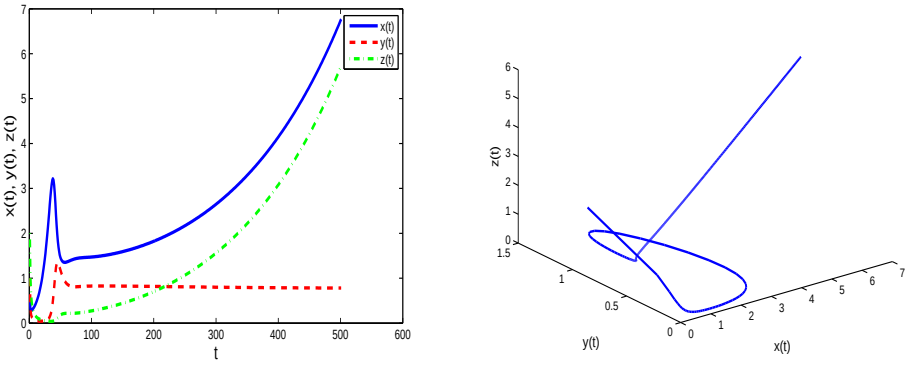

Figure 13: Plot of populations x, y and z over time for the case $bf < ag$ with $\alpha_1 = \alpha_2 = \alpha_3 = 0.50$ and $h = 0.5$.

The solutions for this case are shown in Figure 14 for incommensurate order and parameters $g = 1.6$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 50s and step size $h = 0.1$.

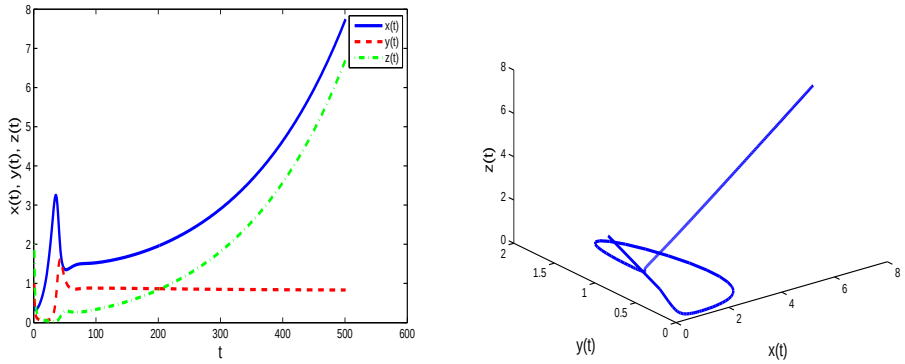

Figure 14: Plot of populations x, y and z over time for the case $bf < ag$ with $\alpha_1 = 0.60, \alpha_2 = 0.50, \alpha_3 = 0.40$ and $h = 0.1$.

In Figure 15 the phase trajectory of the fractional-order Lotka-Volterra chaotic system is depicted for incommensurate order and parameters $g = 1.6$ with the initial conditions $(x(0), y(0), z(0)) = (0.5, 1, 2)$, for simulation time 50s and step size $h = 0.5$.

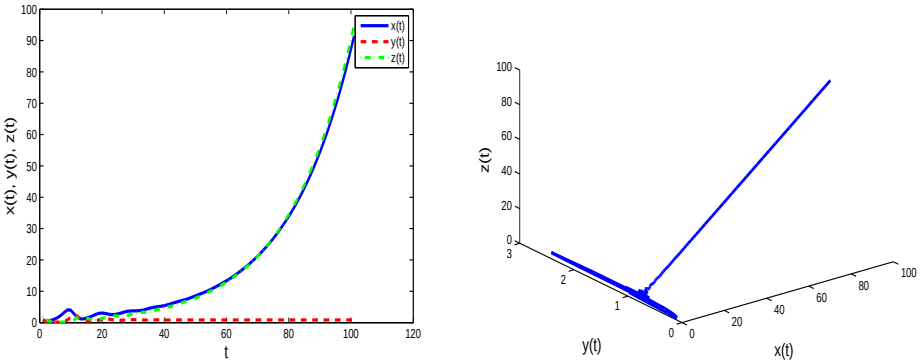

Figure 15: Plot of populations x, y and z over time for the case $bf < ag$ with $\alpha_1 = 0.8, \alpha_2 = 0.6, \alpha_3 = 0.5$ and $h = 0.5$.

In Table 1 for different step size h, the qualitative results, obtained by NSFD scheme, of the fixed point E_2 are respectively compared to classical methods such as forward Euler and 4th order Runge-Kutta. From Table 1, it follows that the CPU time of the method NSFD is less than the CPU time of the forward Euler and Runge-Kutta methods. Also if step size h is chosen small enough, the results of the proposed NSFD scheme are similar with the results of the other two numerical methods. But if the step size h is chosen larger, the efficiency of NSFD scheme is clearly seen.

Table 1: Qualitative results of the equilibrium point E_2 for different time step sizes, t= 0-200 for the case where $bf = ag$

h	Euler	CPU time	Runge-Kutta	CPU time	NSFD	CPU time
0.001	Convergence	0.016342	Convergence	0.032029	Convergence	0.000206
0.01	Convergence	0.014760	Convergence	0.028096	Convergence	0.000205
0.1	Convergence	0.013917	Convergence	0.027959	Convergence	0.000203
0.2	Divergence	–	Convergence	0.025959	Convergence	0.000202
2	Divergence	-	Divergence	–	Convergence	0.000201
10	Divergence	-	Divergence	-	Convergence	0.000201

In Figure 16 the numerical solution of forward Euler and fourth order Runge-Kutta methods are compared with NSFD scheme graphically.

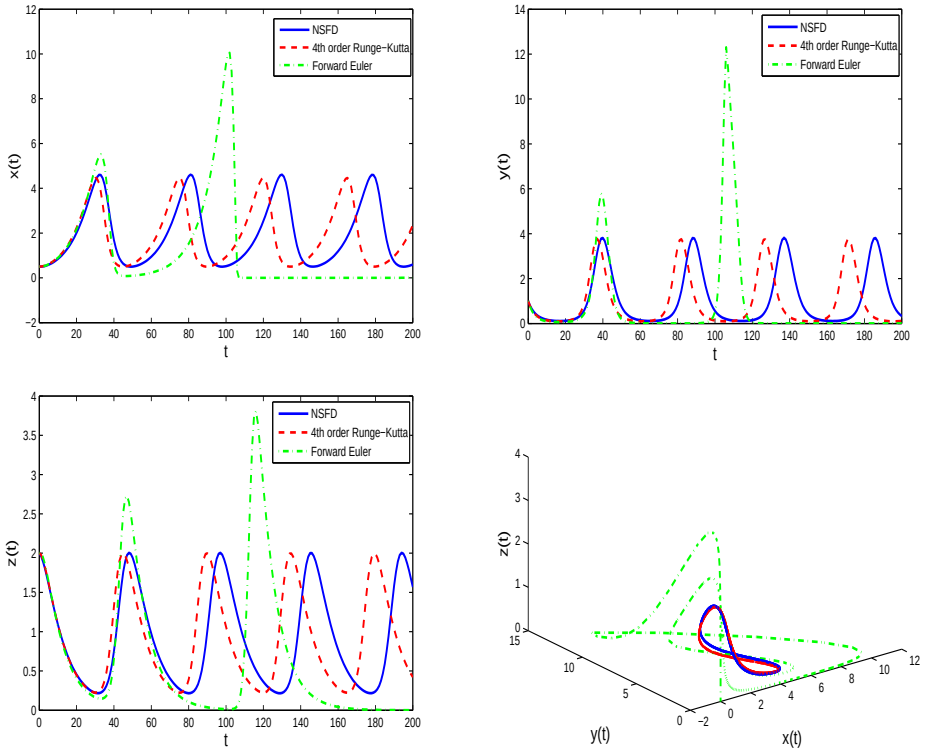

Figure 16: Numerical solutions for forward Euler and fourth order Runge-Kutta and NSFD methods with $h = 0.1$ for the case $bf = ag$ and $\alpha_1 = \alpha_2 = \alpha_3 = 1$.

7 Conclusion

In this paper, we study the fractional-order Lotka-Volterra model. The stability of equilibrium points is studied. Numerical solutions of these models are given. The reason for considering a fractional order system instead of its integer order counterpart is that fractional order differential equations are generalizations of integer order differential equations. Also using fractional order differential equations can help us to reduce the errors arising from the neglected parameters in modelling real life phenomena.

We argue that the fractional order models are at least as good as integer order ones in modeling biological, economic and social systems (generally complex adaptive systems) where memory effects are important.

References

1. Ahmed, E., El-Sayed, A.M.A. and El-Saka, H.A.A. *Equilibrium points, stability and numerical solutions of fractionalorder predator-prey and rabies models*, J. Math. Anal. Appl. 325 (1) (2007) 542-553.

2. Ahmed, E., El-Sayed, A.M.A. and El-Saka, H.A.A. *On some Routh-Hurwitz conditions for fractional order differential equations and their applications in Lorenz, Rössler, Chua and Chen systems*, Phys. Lett. A 358 (2006) 1-4.

3. Ahmad, S. and Lazer, A.C. *Average conditions for global asymptotic stability in a nonautonomous Lotka-Volterra system*, Nonlinear Anal. 40 (1) (2000) 37-49.

4. Arenas, A.J., Gonzalez-Parra, G. and Chen-Charpentier, B.M. *A nonstandard numerical scheme of predictor-corrector type for epidemic models*, Comput. Math. Appl. 59 (12) (2010) 3740-3749.

5. Chauvet, E., Paullet, J.E., Previte, J.P. and Walls, Z. *A Lotka-Volterra three species food chain*, Math. Mag. (75) (2002) 243-255.

6. El-Saka, H.A., Ahmed, E., Shehata, M.I. and El-Sayed, A.M.A. *On stability, persistence, and Hopf bifurcation in fractional order dynamical systems*, Nonlinear Dynam. 56 (1-2) (2009) 121-126.

7. El-Sayed, A.M.A., El-Mesiry, A.E.M. and El-Saka, H.A.A. *On the fractional-order logistic equation*, Appl. Math. Lett. 20 (7) (2007) 817-823.

8. Jordan, P.M. *A nonstandard finite difference scheme for nonlinear heat transfer in a thin finite rod*, J. Difference Equ. Appl. 9 (11) (2003) 1015-1021.

9. Kaslik, E. and Sivasundaram, S. *Nonlinear dynamics and chaos in fractional-order neural networks*, Neural Networks 32 (2012) 245-256.

10. Liao, C. and Ding X. *Nonstandard finite difference variational integrators for multisymplectic PDEs*, J. Appl. Math. 2012 (2012) Article ID 705179, 22 pages.

11. Matignon, D. *Stability result on fractional differential equations with applications to control processing*, Computational engineering in systems applications (1996) 963-968.

12. Mehmat, A.A., Secer, A. and Bayram, M. *Stability, synchronization control and numerical solution of fractional Shimizu-Morioka dynamical system*, Appl. Math. Inf. Sci. 8 (14) (2014) 1699-1705.

13. Mickens, R.E. *Advances in the Applications of Nonstandard Finite Difference Schemes*, Wiley-Interscience, Singapore, 2005.

14. Mickens, R.E. *Calculation of denominator functions for nonstandard finite difference schemes for differential equations satisfying a positivity condition*, Numer. Methods Partial Differential Equations 23 (3) (2007) 672-691.

15. Murray, J.D. *Mathematical Biology I, II, Third edition*, Springer, 2003.

16. Podlubny, I. *Fractional Differential Equations*, Academic Press, New York, 1999.

17. Roeger, L.W. *Dynamically consistent discrete Lotka-Volterra competition models derived from nonstandard finite difference schemes*, Discrete Contin. Dyn. Syst. Ser. B 9 (2) (2008) 415-429.

18. Roeger, L.W. *Local stability of Eulers and Kahans methods*, J. Difference Equ. Appl. 10 (6) (2004) 601-614.

19. Van Den Driessche, P. and Zeeman, M.L. *Three-dimensional competitive Lotka-Volterra systems with no periodic orbits,* SIAM J. Appl. Math. 58 (1) (1998) 227-234.

20. Yan, X.P. and Li, W.T. *Stability and Hopf bifurcation for a delayed cooperative system with diffusion effects*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 18 (2) (2008) 441-453.

21. Zibaei, S. and Namjoo, M. *A NSFD Scheme for Lotka-Volterra Food Web Model*, Iran. J. Sci. Technol. Trans. A Sci. 38 (4) (2014) 399-414.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 79-99

An interactive algorithm for solving multiobjective optimization problems based on a general scalarization technique

M. Ghaznavi*, M. Ilati and E. Khorram

Abstract

The wide variety of available interactive methods brings the need for creating general interactive algorithms enabling the decision maker (DM) to apply freely several convenient methods which best fit his/her preferences. To this end, in this paper, we propose a general scalarizing problem for multiobjective programming problems. The relation between optimal solutions of the introduced scalarizing problem and (weakly) efficient as well as properly efficient solutions of the main multiobjective optimization problem (MOP) is discussed. It is shown that some of the scalarizing problems used in different interactive methods can be obtained from proposed formulation by selecting suitable transformations. Based on the suggested scalarizing problem, we propose a general interactive algorithm (GIA) that enables the DM to specify his/her preferences in six different ways with capability to change his/her preferences any time during the iterations of the algorithm. Finally, a numerical example demonstrating the applicability of the algorithm is provided.

Keywords: Multiobjective optimization; Interactive method; Scalarizing problem; Proper efficiency; Preference information.

---

*Corresponding author
Received 22 February 2015; revised 4 July 2015; accepted 21 October 2015
M. Ghaznavi
Faculty of Mathematical Sciences, Shahrood University of Technology, Shahrood, Iran. e-mail: ghaznavi@shahroodut.ac.ir

M. Ilati
Faculty of Mathematics and Computer Science, Amirkabir University of Technology, 424, Hafez Avenue, 15914 Tehran, Iran. e-mail: m.ilati@aut.ac.ir

E. Khorram
Faculty of Mathematics and Computer Science, Amirkabir University of Technology, 424, Hafez Avenue, 15914 Tehran, Iran. e-mail: eskhor@aut.ac.ir

1 Introduction

The general goal of solving a multiobjective optimization problem (MOP) is to support the decision maker (DM) seeking the most preferred solution of many Pareto optimal solutions as the final one. Inasmuch as finding a most preferred solution needs some extra information from the DM, interactive approaches, based on the participation of the DM, have become popular.

In interactive methods, an iterative algorithm is proposed. Then, the steps of the algorithm are repeated where at each iteration, some information is given to the DM and he/she specifies his/her preferences. The process is repeated until the DM is satisfied with regard to the obtained solution. The benefits of using interactive approaches are that, the DM (i) does not need to have any global preference structure, (ii) has the possibility of learning about the interrelationship between the objectives, (iii) can learn about the feasibility of solutions during the solving process.

Heretofore, many interactive methods have been suggested in the literature [1, 13, 19, 23, 26, 30, 31]. As pointed out already, interactive methods are very useful and realistic to solve an MOP. However, since there have been many interactive methods available, it is not easy to choose an appropriate method conveniently. Therefore, creating global algorithms with an ability to accommodate different methods will be useful. By creating a global algorithm, it is possible for the DM to select freely an appropriate method (and the way of specifying preference information) as well as to switch between methods. To this end, it is necessary to design a general scalarizing problem yielding scalarizing problems used in different interactive methods.

Until now, some global algorithms have been proposed. For example, Gardiner and Steuer [7, 8] proposed a unified algorithm including nine to thirteen different methods. Romero [27] presented another general optimization structure, called extended lexicographic goal programming. Moreover, Vassileva [32] suggested a general scalarizing problem which incorporates different scalarizing problems. More recently, based on a global formulation (GLIDE), Luque et al. [21] proposed a global procedure which accommodates eight interactive methods of different types. Nevertheless, their formulation is unlikely to consider the computational efficiency, therefore Ruiz et al. [28] improved the computational efficiency of GLIDE by reformulating it.

In some of the mentioned publications, the authors have provided theorems concerning (weak) efficiency of the optimal solutions of their proposed general scalarizing problems [21, 28, 32] and as far as we know few results related to proper efficiency have been provided. Now, in this paper we suggest a general scalarizing problem which not only considers computational efficiency by reducing the number of added constraints, but also provides theorems concerning (weak) efficiency as well as proper efficiency of its optimal solutions. The provided results are established without any convexity assumption. Also, by setting suitable values for parameters and index sets of the proposed general scalarizing problem, we obtain many known scalarizing

problems. Based on the mentioned problem, we propose a general interactive algorithm (GIA) to solve a given MOP, subsequently. In this algorithm, the DM has the ability to specify his/her preference information in six different ways.

The rest of this paper is organized as follows: Section 2 contains some preliminaries and basic definitions. In Section 3, we propose our general formulation and obtain some theorems. Section 4 gives some scalarizing problems used in different interactive methods which can be obtained from our general formulation. Section 5 contains our proposed interactive algorithm. In Section 6, some computational and theoretical advantages are mentioned. An example is presented in Section 7 and finally, in Section 8 conclusions are given.

2 Preliminaries and basic definitions

A general multiobjective optimization problem can be written as:

$$\begin{aligned} (MOP) \qquad & \min f(\mathbf{x}) \\ & s.t. \ \ \mathbf{x} \in \mathcal{X}, \end{aligned} \tag{1}$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ is a nonempty compact set, and $f(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), ..., f_p(\mathbf{x}))^T : \mathcal{X} \to \mathbb{R}^p$ is a vector-valued function.

The set of all attainable outcomes or objective vectors is defined as the image of the feasible solutions $\mathbf{x} \in \mathcal{X}$ under f. In fact $\mathcal{Y} := f(\mathcal{X}) \subset \mathbb{R}^p$. For $\mathbf{y^1}$ and $\mathbf{y^2} \in \mathbb{R}^p$, $\mathbf{y^1} \leqq \mathbf{y^2}$ means that $y_i^1 \leq y_i^2$, for each $i = 1, \cdots, p$, also $\mathbf{y^1} \leq \mathbf{y^2}$ stands for $\mathbf{y^1} \leqq \mathbf{y^2}$ and $\mathbf{y^1} \neq \mathbf{y^2}$. Furthermore, $\mathbf{y^1} < \mathbf{y^2}$ means that $y_i^1 < y_i^2$, for each $i = 1, ..., p$. The Pareto cone is defined as $\mathbb{R}^p_{\leqq} = \{\mathbf{y} \in \mathbb{R}^p : \mathbf{y} \leqq 0\}$. $\mathbb{R}^p_{\leq}$ and $\mathbb{R}^p_{>}$ are defined, similarly. In this paper, we shall assume that $\mathcal{Y} := f(\mathcal{X})$ is bounded.

Definition 1. *A feasible solution $\hat{x} \in \mathcal{X}$ is called:*

(i) weakly efficient (weakly Pareto optimal) solution to MOP (1) if there is no other $\mathbf{x} \in \mathcal{X}$ such that $f(\mathbf{x}) < f(\hat{x})$,

(ii) efficient (Pareto optimal) solution to MOP (1) if there is no other $\mathbf{x} \in \mathcal{X}$ such that $f(\mathbf{x}) \leq f(\hat{x})$,

(iii) properly efficient (properly Pareto optimal) solution to MOP (1) if it is efficient and there exists a real positive number M such that for each $i \in \{1, 2, \ldots, p\}$ and each $\mathbf{x} \in \mathcal{X}$ satisfying $f_i(\mathbf{x}) < f_i(\hat{\mathbf{x}})$, there exists an index $j \in \{1, 2, \ldots, p\}$ with $f_j(\hat{\mathbf{x}}) < f_j(\mathbf{x})$ and

$$\frac{f_i(\hat{\mathbf{x}}) - f_i(\mathbf{x})}{f_j(\mathbf{x}) - f_j(\hat{\mathbf{x}})} \leq M.$$

The set of all weakly efficient, efficient, and properly efficient solutions of MOP (1) will be denoted by $\mathcal{X}_{WE}$, $\mathcal{X}_E$ and $\mathcal{X}_{PE}$ respectively. The image $f(\mathbf{x}) \in \mathcal{Y}$ of an (weakly, properly) efficient solution $\mathbf{x} \in \mathcal{X}$, is called (weakly, properly) nondominated point.

Remark 1. Obviously, $\mathcal{X}_{PE} \subseteq \mathcal{X}_E \subseteq \mathcal{X}_{WE}$.

Remark 2. In this paper, we use definition of proper efficiency in the sense of Geoffrion [9]. There are other definitions of proper efficiency which are almost the same when using the Pareto cone as the order cone. For considering relationships between different definitions of proper efficiency one can refer to [4].

Definition 2. *The ideal point* $\mathbf{y^I} = (y_1^I, \ldots, y_p^I)$ *of MOP (1) is defined by* $y_i^I := min_{\mathbf{x}\in\mathcal{X}} f_i(\mathbf{x}), i = 1, \cdots, p$.

Definition 3. *The point* $\mathbf{y^U} := \mathbf{y^I} - \alpha$, *where* $\alpha \in \mathbb{R}^p_>$ *is a vector with small positive components, is called the utopia point of MOP (1).*

Definition 4. *The nadir point* $\mathbf{y^N} = (y_1^N, \ldots, y_p^N)$ *of MOP (1) is defined by* $y_i^N := max_{\mathbf{x}\in\mathcal{X}_E} f_i(\mathbf{x}), i = 1, \cdots, p$.

Definition 5. *The vector* $\bar{\mathbf{y}} = (\bar{y}_1, \ldots, \bar{y}_p) \in \mathbb{R}^p$, *consisting of the desired or aspiration values to the DM, is called a reference point. It should be noted that reference point may be achievable or not.*

One of the most popular approaches to solve a given MOP is scalarization, which involves formulating a single objective problem associated with the given MOP. Let us consider a single objective programming problem as follows:

$$\begin{aligned} &\min g(\mathbf{x}) \\ &s.t. \quad \mathbf{x} \in \mathcal{S}, \end{aligned} \tag{2}$$

where $g : \mathcal{S} \to \mathbb{R}$.

Definition 6. *A feasible solution* $\hat{\mathbf{x}} \in \mathcal{S}$ *is said to be*
(i) an optimal solution of problem (2) if $g(\hat{\mathbf{x}}) \leq g(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{S}$,
(ii) a strictly optimal solution of problem (2) if $g(\hat{\mathbf{x}}) < g(\mathbf{x})$ *for all* $\mathbf{x} \in \mathcal{S}\backslash\{\hat{\mathbf{x}}\}$.

3 A general scalarizing problem

In this section, we propose a general scalarizing problem associated with MOP (1), which is defined such that many scalarizing problems, used in different interactive methods, can be deduced from it by selecting suitable values of

parameters and index sets. The general scalarizing problem is proposed as follows:

$$\min \max_{i\in I_1^k} \quad \lambda_i^k\Big(f_i(\mathbf{x}) - r_i^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\mathbf{x}) - r_t^k)\Big) \\ s.t. \begin{cases} f_i(\mathbf{x}) \le \delta_i^k & \forall i \in I_2^k, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \tag{3}$$

where $\lambda_i^k \ge 0,\ \rho \ge 0,\ \delta_i^k,\ r_i^k$, and $w_t^k \ge 0$ are parameters specified depending on the information given by DM. Also, $I_1^k \ne \emptyset$ and I_2^k are index sets, which are subsets of $\{1, \cdots, p\}$. Notice that hereafter, we make the assumption that on the proposed scalarizing problem, the parameters $\delta_i^k, i \in I_2^k$ are selected such that problem (3) remains feasible. Let k be the current iteration. Then, the optimal solution obtained from scalarizing problem (3) is defined by $\hat{\mathbf{x}}^{k+1}$ and the corresponding objective vector by $f(\hat{\mathbf{x}}^{k+1})$.

According to [14, p. 305], or [23, p. 97], if we replace the max term by a new variable $z \in \mathbb{R}$, then problem (3) is equivalent to the following scalarizing optimization problem:

$$\min \quad z \\ s.t. \begin{cases} \lambda_i^k\Big(f_i(\mathbf{x}) - r_i^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\mathbf{x}) - r_t^k)\Big) \le z & \forall i \in I_1^k, \\ f_i(\mathbf{x}) \le \delta_i^k & \forall i \in I_2^k, \\ \mathbf{x} \in \mathcal{X}. \end{cases} \tag{4}$$

Notice that the scalarizing problem (3) is nondifferentiable, even if the main MOP (1) is differentiable (i.e., the objective functions and constraint functions are differentiable). Therefore, if the original MOP is the differentiable we propose to use formulation (4) since it preserves differentiability. In this case, the scalar optimization problem (4) can be solved with standard methods of (non)linear constraint optimization or using available single objective solvers. However, if the original MOP (1) is nondifferentiable, both scalarized problems (3) and (4) are nondifferentiable, too. In this case, the scalarized problem (3) is recommended since it has a reduced number of constraints.

It should be noted that, unlike the formulations proposed in [21, 28], the bounds on trade-offs generated by the suggested formulation are independent of parameters λ_i. For more details about bounds on trade-offs see [17, 18]. So far, many authors have provided theorems concerning weak efficiency and efficiency of the optimal solutions of the scalarized problems used in the interactive methods. Now, we prove some general theorems concerning weak efficiency, efficiency, as well as proper efficiency of (strictly) optimal solutions of problems (3) and (4). It is important to point out that the following

theorems are general and many theorems concerning (weak, proper) efficiency [4, 23] can be resulted from them. Moreover, the theorems are provided with no convexity assumption. Since problems (3) and (4) are equivalent, we only provide theorems for the first one.

Theorem 3. *Let $\lambda_i^k > 0 \ \forall i \in I_1^k$. If $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}$ is an optimal solution of problem (3), then $\hat{\mathbf{x}}^{k+1}$ is a weakly efficient solution of MOP (1).*

Proof. Let $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}$ be an optimal solution of problem (3) and suppose that $\hat{\mathbf{x}}^{k+1} \notin \mathcal{X}_{WE}$. Then, there exists $\mathbf{x} \in \mathcal{X}$ such that $f(\mathbf{x}) < f(\hat{\mathbf{x}}^{k+1})$. Therefore, $f_i(\mathbf{x}) < f_i(\hat{\mathbf{x}}^{k+1}) \leq \delta_i^k \ \ \forall i \in I_2^k$, which means $\mathbf{x} \in \mathcal{X}$ is a feasible solution for problem (3). Also, we have

$$f_i(\mathbf{x}) - r_i^k < f_i(\hat{\mathbf{x}}^{k+1}) - r_i^k \ \ \forall i \in I_1^k,$$

and

$$\rho \sum_{t=1}^{p} w_t^k (f_t(\mathbf{x}) - r_t^k) \leq \rho \sum_{t=1}^{p} w_t^k (f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k).$$

Therefore,

$$\max_{i \in I_1^k} \lambda_i^k \Big(f_i(\hat{\mathbf{x}}^{k+1}) - r_i^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k) \Big) >$$

$$\max_{i \in I_1^k} \lambda_i^k \Big(f_i(\mathbf{x}) - r_i^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\mathbf{x}) - r_t^k) \Big),$$

which is a contradiction with optimality of $\hat{x}^{k+1}$. Thus, $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}_{WE}$. □

In the following theorem, utilizing the general formulation (3), a sufficient condition for efficiency is provided.

Theorem 4. *If $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}$ is a strictly optimal solution of problem (3), then $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}_E$.*

Proof. The proof is similar to that of Theorem 3. □

It is found out from part (*ii*) of Definition 1, that in an efficient solution it is not possible to improve any criterion without deterioration of at least one other criterion. Sometimes, these trade-offs may be unbounded and it is obvious that efficient solutions with bounded trade-offs (called properly efficient) are desirable. Until now, many scholars have considered relationships between optimal solutions of the scalarizing problem used in their proposed interactive methods and (weakly) efficient solutions of the related MOP [21, 28, 32], but there are fewer results concerning proper efficiency. In the following theorem, we provide a sufficient condition concerning properly efficient solutions of MOP (1).

Theorem 5. *If $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}$ is an optimal solution for problem (3) with $\lambda_i^k > 0\ \forall i \in I_1^k$, $\rho > 0$ and $w^k \in \mathbb{R}^p_>$, then $\hat{\mathbf{x}}^{k+1} \in \mathcal{X}_{PE}$.*

Proof. We show that $\hat{x}^{k+1} \in \mathcal{X}_E$. Let $\hat{\mathbf{x}}^{k+1} \notin \mathcal{X}_E$. Then, there exists $\mathbf{x} \in \mathcal{X}$ with $f_i(\mathbf{x}) \leq f_i(\hat{\mathbf{x}}^{k+1}),\ \forall i \in \{1, \cdots, p\}$ and $f_j(\mathbf{x}) < f_j(\hat{\mathbf{x}}^{k+1})$ for some $j \in \{1, \cdots, p\}$. Hence, $f_i(\mathbf{x}) \leq f_i(\hat{\mathbf{x}}^{k+1}) \leq \delta_i^k\ \forall i \in I_2^k$. Thus, $\mathbf{x} \in \mathcal{X}$ is a feasible solution of (3). Using the assumptions and the definition of efficiency, it follows that:

$$\max_{i \in I_1^k} \lambda_i^k \Big(f_i(\hat{\mathbf{x}}^{k+1}) - r_i^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k) \Big) >$$

$$\max_{i \in I_1^k} \lambda_i^k \Big(f_i(\mathbf{x}) - r_i^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\mathbf{x}) - r_t^k) \Big).$$

This is a contradiction with optimality of $\hat{x}^{k+1}$ and therefore $\hat{x}^{k+1} \in \mathcal{X}_E$. Now, we show that $\hat{x}^{k+1}$ is a properly efficient solution to MOP (1). To this end, we define:

$$M = \max_{i \in \{1, \cdots, p\}} \{ \frac{1 + \rho \sum_{t=1}^{p} w_t^k}{\rho w_i^k} \},$$

and consider an index $i \in \{1, \cdots, p\}$ and $\overline{\mathbf{x}} \in \mathcal{X}$ such that $f_i(\overline{\mathbf{x}}) < f_i(\hat{x}^{k+1})$. To prove the proper efficiency of $\hat{x}^{k+1}$, we must show that there exists an index $j \in \{1, 2, \ldots, p\}$ with $f_j(\hat{\mathbf{x}}^{k+1}) < f_j(\overline{\mathbf{x}})$ such that

$$\frac{f_i(\hat{\mathbf{x}}^{k+1}) - f_i(\overline{\mathbf{x}})}{f_j(\overline{\mathbf{x}}) - f_j(\hat{\mathbf{x}}^{k+1})} \leq M.$$

From efficiency of $\hat{x}^{k+1}$, we conclude that there exists an index $t \in \{1, \cdots, p\}$ such that $f_t(\hat{\mathbf{x}}^{k+1}) < f_t(\overline{\mathbf{x}})$. We define

$$f_j(\hat{\mathbf{x}}^{k+1}) - f_j(\overline{\mathbf{x}}) = \min_{m \in \{1, \cdots, p\}} (f_m(\hat{\mathbf{x}}^{k+1}) - f_m(\overline{\mathbf{x}})). \tag{5}$$

It is obvious that $f_j(\hat{\mathbf{x}}^{k+1}) - f_j(\overline{\mathbf{x}}) < 0$.
Moreover, optimality of $\hat{x}^{k+1}$ for problem (3), concludes

$$\max_{m \in I_1^k} \lambda_m \Big(f_m(\overline{\mathbf{x}}) - r_m^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\overline{\mathbf{x}}) - r_t^k) \Big) \geq$$

$$\max_{m \in I_1^k} \lambda_m \Big(f_m(\hat{\mathbf{x}}^{k+1}) - r_m^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k) \Big).$$

Now, let

$$\lambda_l \Big(f_l(\overline{\mathbf{x}}) - r_l^k + \rho \sum_{t=1}^{p} w_t^k (f_t(\overline{\mathbf{x}}) - r_t^k) \Big) =$$

$$\max_{m\in I_1^k} \lambda_m\Big(f_m(\overline{\mathbf{x}}) - r_m^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\overline{\mathbf{x}}) - r_t^k)\Big).$$

Hence,

$$\lambda_l\Big(f_l(\overline{\mathbf{x}}) - r_l^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\overline{\mathbf{x}}) - r_t^k)\Big) \geq$$

$$\max_{m\in I_1^k} \lambda_m\Big(f_m(\hat{\mathbf{x}}^{k+1}) - r_m^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k)\Big) \geq$$

$$\lambda_l\Big(f_l(\hat{\mathbf{x}}^{k+1}) - r_l^k + \rho\sum_{t=1}^{p} w_t^k(f_t(\hat{\mathbf{x}}^{k+1}) - r_t^k)\Big).$$

Then,

$$0 \geq (f_l(\hat{\mathbf{x}}^{k+1}) - f_l(\overline{\mathbf{x}})) + \rho\sum_{t=1}^{p} w_t^k(f_t(\hat{\mathbf{x}}^{k+1}) - f_t(\overline{\mathbf{x}})). \tag{6}$$

Now, from (5) and (6), we have:

$$0 \geq (f_j(\hat{\mathbf{x}}^{k+1}) - f_j(\overline{\mathbf{x}})) + \rho\sum_{t=1}^{p} w_t^k(f_t(\hat{\mathbf{x}}^{k+1}) - f_t(\overline{\mathbf{x}})).$$

That is,

$$\rho w_i^k(f_i(\hat{\mathbf{x}}^{k+1}) - f_i(\overline{\mathbf{x}})) \leq f_j(\overline{\mathbf{x}}) - f_j(\hat{\mathbf{x}}^{k+1}) + \rho\sum_{\substack{t=1\\ t\neq i}}^{p} w_t^k(f_t(\overline{\mathbf{x}}) - f_t(\hat{\mathbf{x}}^{k+1})) \leq$$

$$(1 + \rho\sum_{\substack{t=1\\ t\neq i}}^{p} w_t^k)(f_j(\overline{\mathbf{x}}) - f_j(\hat{\mathbf{x}}^{k+1})).$$

Hence

$$\frac{f_i(\hat{\mathbf{x}}^{k+1}) - f_i(\overline{\mathbf{x}})}{f_j(\overline{\mathbf{x}}) - f_j(\hat{\mathbf{x}}^{k+1})} \leq \frac{1 + \rho\sum_{\substack{t=1\\ t\neq i}}^{p} w_t^k}{\rho w_i^k} \leq M,$$

which completes the proof. □

It should be noted that, using suitable values for parameters in (3), we can provide necessary conditions related to (weakly, properly) efficient solutions of MOP (1). For example, if we choose $I_1^k = \{1, \cdots, p\}$, $I_2^k = \emptyset$, $r_i^k = y_i^U$ and $w_i^k = 1\ \forall i \in \{1, \ldots, p\}$, then we have the modified weighted Tchebycheff method [15] and, by Theorem 4.2 in [16], for every properly efficient solution of MOP (1) we can find suitable parameters $\lambda_i^k > 0,\ \forall i \in \{1, \cdots, p\}$ and $\rho > 0$ such that this properly efficient solution be an optimal solution of (3).

4 Achieving different scalarizing problems from the general formulation

The general formulations (3) and (4) are generalizations of already known scalarizing problems. In this section, we are going to show that how many famous scalarizing problems (used in different interactive methods) can be attained from (3) and (4) by choosing appropriate values of parameters and index sets. We obtain the scalarizing problems from (3). By a similar method it is possible to obtain them from (4).

4.1 GUESS method and STOM

The GUESS method is one of the simplest interactive methods, proposed by Buchanan [2]. In this method, the DM has to determine the components of the reference point ($\bar{y}_i^k$) as preference information. At the kth iteration, the scalarizing problem used in this method is formulated as follows:

$$\begin{aligned} \min \max_{i=1,\ldots,p} \quad & \frac{f_i(\mathbf{x}) - \bar{y}_i^k}{y_i^N - \bar{y}_i^k} \\ s.t. \quad & \mathbf{x} \in \mathcal{X}. \end{aligned} \tag{7}$$

Notice that the reference vector specified by the DM, must be strictly lower than the nadir objective vector, that is, $\bar{\mathbf{y}} < \mathbf{y^N}$. This scalarizing problem can be achieved from (3) by considering the following replacements:

(1) $I_1^k = \{1, \ldots, p\}$ and $I_2^k = \emptyset$;

(2) $w_i^k = 0$, $\lambda_i^k = \frac{1}{y_i^N - \bar{y}_i^k}$, and $\rho = 0$;

(3) $r_i^k = \bar{y}_i^k$ and $i = 1, \ldots, p$.

The satisficing trade-off method (STOM) [24] can be obtained from (3), similar to the GUESS method, by setting $\lambda_i^k = \frac{1}{\bar{y}_i^k - y_i^U}$ and $r_i^k = y_i^U$ $(i = 1, \ldots, p)$. Other parameter values and index sets are the same as those of GUESS method. In this method, $\bar{\mathbf{y}}$ must be chosen such that $\bar{\mathbf{y}} > \mathbf{y}^U$.

4.2 Reference direction approach

In this method, a vector from the current iteration point to the reference point (a reference direction) is projected onto the efficient set [20]. To obtain the points along the reference direction at the kth iteration, the following scalarizing problem needs to be solved:

$$\min \max_{i=1,\ldots,p} \quad \frac{f_i(\mathbf{x}) - (f_i^k + td_i^k)}{\mu_i} \qquad s.t. \quad \mathbf{x} \in \mathcal{X}, \tag{8}$$

where, f^k is the current nondominated objective vector, $d^k = \bar{\mathbf{y}}^k - f^k$, t has different discrete nonnegative values, and μ is a weighting vector that can be either a reference point presented by the DM or defined as $\mathbf{y^N} - \mathbf{y^U}$. This problem can be obtained from (3) by considering the following replacements:

(1) $I_1^k = \{1, \ldots, p\}$ and $I_2^k = \emptyset$;

(2) $w_i^k = 0$, $\lambda_i^k = \frac{1}{\mu_i}$ and $\rho = 0$;

(3) $r_i^k = f_i^k + td_i^k$ and $i = 1, \ldots, p$.

4.3 Step method

The step method is one of the first known interactive methods [1]. Eschenauer et al. [6] extended this method to nonlinear problems. In this method, based on the current objective vector (f^k), the DM can improve some unacceptable objective functions f_i ($i \in J_1^k$) by relaxing some other objective function(s) f_i ($i \in J_2^k$) such that $J_1^k \cup J_2^k = \{1, \ldots, p\}$. In this regard, the DM must specify upper bounds $\varepsilon_i^k > f_i^k$ for functions f_i ($i \in J_2^k$). In this case, the scalarizing problem is formulated as follows:

$$\min \max_{i=1,\ldots,p} \quad \left(\frac{e_i}{\sum_{j=1}^p e_j}(f_i(\mathbf{x}) - y_i^I)\right) \qquad s.t. \quad \begin{cases} f_i(\mathbf{x}) \le f_i^k & \forall i \in J_1^k, \\ f_i(\mathbf{x}) \le \varepsilon_i^k & \forall i \in J_2^k, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \tag{9}$$

where $e_i = \frac{1}{y_i^I}(\frac{y_i^N - y_i^I}{y_i^N})$, $i = 1, \ldots, p$ (the denominators are not allowed to be zero). We can obtain (9) from (3) using the following replacements:

(1) $I_1^k = \{1, \ldots, p\}$ and $I_2^k = J_1^k \cup J_2^k$;

(2) $w_i^k = 0$, $\forall i \in \{1, \ldots, p\}$, $\lambda_i^k = \frac{e_i}{\sum_{j=1}^p e_j}$ $\forall i \in \{1, \ldots, p\}$ and $\rho = 0$;

(3) $r_i^k = y_i^I$, $\forall i \in \{1, \ldots, p\}$;

(4) $\delta_i^k = f_i^k$ for $i \in J_1^k$ and $\delta_i^k = \varepsilon_i^k$ for $i \in J_2^k$.

4.4 SPOT method

In the SPOT method, given the current objective vector f^k, the DM is asked to select a reference objective function f_l and then compare each objective function f_i $(i = 1, \ldots, p,\ i \neq l)$ with f_l by providing the marginal rates of substitutions (MRSs) m_{li}^k $(i = 1, \ldots, p,\ i \neq l)$ [29]. The MRSs can be approximated as $m_{li}^k \simeq \frac{\Delta f_l^k}{\Delta f_i^k},\ i = 1, \ldots, p$, where Δf_i^k is the amount of improvement, provided by the DM, on the value of the objective function f_i that can exactly compensate for the given amount Δf_l^k to be deteriorated of the reference objective f_l. The intermediate single objective optimization problem, used in this method, can be formulated as follows:

$$s.t. \begin{cases} \min \quad f_l(\mathbf{x}) \\ f_i(\mathbf{x}) \leq f_i^k + \alpha(\mu_{li}^k - m_{li}^k) \quad \forall i \in \{1, \ldots, p\},\ i \neq l, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \tag{10}$$

where $\mu_{li}^k, i \neq l$ are K.K.T multipliers, corresponding to the current non-dominated objective vector [21] and several values for α are set and in this way different solutions are obtained. This problem is achieved from (3), by considering the following transformations:

(1) $I_1^k = \{l\}$ and $I_2^k = \{1, \ldots, p\} \backslash \{l\}$;

(2) $\lambda_l^k = 1, \rho = 0$ and $r_l^k = 0$;

(3) $\delta_i^k = f_i^k + \alpha(\mu_{li}^k - m_{li}^k), i = 1, \ldots, p$ and $i \neq l$.

4.5 Modified reference point method

This method is an interactive reference direction method for solving convex nonlinear integer problems [31]. Here, the DM is asked to set his/her preferences as aspiration levels of the objective functions at each iteration. Let J_1^k be the set of indices of the objective functions which the DM wants to improve and J_2^k denotes the set of indices which can worsen and J_3^k contains the indices that are satisfactory to the DM. The scalarizing problem used in this method is formulated as follows:

$$\begin{aligned} \min \max_{i \in J_1^k,\ j \in J_2^k} \left\{ \frac{f_i(\mathbf{x}) - \bar{y}_i^k}{f_i^k - \bar{y}_i^k}, \quad \frac{f_j(\mathbf{x}) - f_j^k}{\bar{y}_j^k - f_j^k} \right\} \\ s.t. \begin{cases} f_i(\mathbf{x}) \leq f_i^k \quad \forall i \in J_3^k, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \end{aligned} \tag{11}$$

where the denominators must be positive. By the following replacements, problem (11) can be resulted from (3):

(1) $I_1^k = J_1^k \cup J_2^k$ and $I_2^k = J_3^k$;

(2) $w_i^k = 0$, $\lambda_i^k = \frac{1}{f_i^k - \bar{y}_i^k}$ $\forall i \in J_1^k$, $\lambda_i^k = \frac{1}{\bar{y}_i^k - f_i^k}$ $\forall i \in J_2^k$, $\rho = 0$ and $i = 1, \ldots, p$;

(3) $r_i^k = \bar{y}_i^k$ $\forall i \in J_1^k$, $r_i^k = f_i^k$ $\forall i \in J_2^k$ and $\delta_i^k = f_i^k$ $\forall i \in J_3^k$.

4.6 RD method

The reference direction (RD) method was proposed in [25]. At the kth iteration, the DM is asked to specify a reference point $\bar{\mathbf{y}}^k$. Specifying a reference point is equivalent to classifying the objective functions in three classes J_1^k , J_2^k and J_3^k, where these index sets are the same as those defined before. The scalarizing problem related to the RD method is as follows:

$$\min \max_{i \in J_1^k} \quad \frac{f_i(\mathbf{x}) - f_i^k}{f_i^k - \bar{y}_i^k}$$
$$s.t. \begin{cases} f_i(\mathbf{x}) \leq f_i^k & \forall i \in J_3^k, \\ f_i(\mathbf{x}) \leq \bar{y}_i^k + \alpha(f_i^k - \bar{y}_i^k) & \forall i \in J_2^k, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \tag{12}$$

where $0 \leq \alpha < 1$ and the denominators must be positive. The general formulation (3) can be transformed to RD problem (12) by the following replacements:

(1) $I_1^k = J_1^k$ and $I_2^k = J_2^k \cup J_3^k$;

(2) $w_i^k = 0$ $\forall i \in \{1, \ldots, p\}$, $\lambda_i^k = \frac{1}{f_i^k - \bar{y}_i^k}$ $\forall i \in J_1^k$ and $\rho = 0$;

(3) $r_i^k = f_i^k \forall i \in J_1^k$, $\delta_i^k = f_i^k$ $\forall i \in J_3^k$ and $\delta_i^k = \bar{y}_i^k + \alpha(f_i^k - \bar{y}_i^k)$ $\forall i \in J_2^k$.

4.7 $\epsilon-$Constraint method

In this method, one of the objective functions is minimized, while the other objectives are transformed into constraints by setting an upper bound [4,23]. The problem to be solved has the following form:

$$\min \quad f_l(\mathbf{x})$$
$$s.t. \begin{cases} f_j(\mathbf{x}) \leq \varepsilon_j^k & \forall j \in \{1, \ldots, p\}, j \neq l, \\ \mathbf{x} \in \mathcal{X}. \end{cases} \tag{13}$$

By the following transformations, problem (13) can be attained from (3):

(1) $I_1^k = \{l\}$ and $I_2^k = \{1, \ldots, p\} \backslash \{l\}$;

(2) $\lambda_l^k = 1,\ \rho = 0,\ r_l^k = 0,\ \delta_j^k = \varepsilon_j^k,\ j = 1, \ldots, p$ and $j \neq l$.

4.8 The weighted sum method

In this method, a weighting coefficient is associated with each objective function and then the weighted sum of the objectives is minimized [4, 23]. Accordingly, solutions are obtained by solving the following problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{p} \mu_i^k f_i(\mathbf{x}) \\ & s.t. \quad \mathbf{x} \in \mathcal{X}, \end{aligned} \tag{14}$$

with $\mu_i^k \geq 0\ \forall i \in \{1, \ldots, p\}$ and $\sum_{i=1}^{p} \mu_i^k = 1$. By the following replacements, we can obtain this problem from (3):

(1) $I_1^k = \{l\}$, where l is an index with $\mu_l^k \neq 0$ and $I_2^k = \emptyset$;

(2) $\lambda_l^k = \frac{\mu_l^k}{2}$, $\rho = \frac{2}{\mu_l^k}$, $w_i^k = \mu_i^k\ \forall i \neq l,\ w_l^k = \frac{\mu_l^k}{2}$ and $r_i^k = 0\ \forall i \in \{1, \ldots, p\}$.

4.9 Hybrid method

The hybrid method is a combination of the weighted sum method and the $\epsilon-$constraint method [4, 23]. This problem has the following form:

$$\begin{aligned} & \min \sum_{i=1}^{p} \mu_i^k f_i(\mathbf{x}) \\ & s.t. \begin{cases} f_j(\mathbf{x}) \leq \varepsilon_j^k, & \forall j \in \{1, \ldots, p\}, \\ \mathbf{x} \in \mathcal{X}, \end{cases} \end{aligned} \tag{15}$$

where $\mu_i^k \geq 0\ \forall i,\ \sum_{i=1}^{p} \mu_i^k = 1$ and $\varepsilon^k = (\varepsilon_1^k, \ldots, \varepsilon_p^k)$ is an upper bound vector. One can find this problem from (3) by the following transformations:

(1) $I_1^k = \{l\}$, where l is an index with $\mu_l^k \neq 0$ and $I_2^k = \{1, \ldots, p\}$;

(2) $\lambda_l^k = \frac{\mu_l^k}{2}$, $\rho = \frac{2}{\mu_l^k}$, $w_i^k = \mu_i^k\ \forall i \neq l$ and $w_l^k = \frac{\mu_l^k}{2}$;

(3) $r_i^k = 0$ and $\delta_i^k = \varepsilon_i^k,\ \forall i \in \{1, \ldots, p\}$.

Remark 3. Using similar procedure, we can obtain some other single objective problems used in different interactive approaches. For example, the intermediate problems of the interactive surrogate worth trade-off (ISWT) method [3] and the PROJECT method [22] can be obtained easily from our formulations. In addition, the weighted Tchebycheff scalarizing problem [4] and the modified weighted Tchebycheff problem [15] are resulted from the proposed general scalarizing problem.

5 General interactive algorithm

Based on the general formulations given in Section 3, we present a general interactive algorithm (GIA). The proposed algorithm allows the DM to specify his/her preference information in six different ways. Moreover, he/she will be able to change his/her preference information in each iteration. In addition to widely used ways (reference point specification, classification of the objective functions, and specification of marginal rate of substitution) for specifying preference information ([21, 28]), GIA allows the DM to specify his/her preferences as criteria weights, $\varepsilon-$constraint (choosing a reference objective function and setting upper bounds for the other objective functions), or criteria weights and upper bounds for objective functions, simultaneously. The main steps of the GIA are given in Algorithm 1.

As pointed out in Step 4, the values of parameters and index sets depend on the type of preference information given by DM in Step 3. For example, if DM specifies his/her preferences as reference point, we should set parameters and index sets in (3) or (4) so that one of the reference based on scalarizing points problems (see, for example, (7) and (8)) be attained.

6 Computational and theoretical advantages

The GIA and the proposed scalarizing formulation has a number of potential advantages both in theoretical and computational points of view. Here, we indicate only some key potential advantages, with special attention to those not shared by other competing algorithms.

(a) Taking the special characteristics of the problem into account, GIA allows using more efficient optimization methods. In the first step of the GIA, the type of problem (differentiable or nondifferentiable) is specified. This step provides some advantages. For example, if the proposed problem is differentiable, corresponding scalarizing formulation preserves differentiability and can be solved using available single objective solvers. On the other hand, for a nondifferentiable problem,

Algorithm 1. General interactive algorithm (GIA)

Step 1- Determine type of the MOP being solved (differentiable or nondifferentiable).

Step 2- Compute ideal and nadir points. Set $k = 1$. Determine an initial solution (can be specified by the DM or by solving an arbitrary scalarizing problem). Denote this initial solution by $\hat{\mathbf{x}}^k$ and corresponding objective vector by $f(\hat{\mathbf{x}}^k)$. If the DM is satisfied with this solution, go to Step 6.

Step 3- Ask the DM to provide his/her preference information based on $f(\hat{\mathbf{x}}^k)$. The DM can specify his/her preference information in one of the following ways:

3.1. Specifying the desired objective function values as components of the reference point $(\bar{y}_i^k,\ i = 1, \ldots, p)$;

3.2. Classifying the objective functions into two classes J_1^k and J_2^k or three classes J_1^k , J_2^k and J_3^k, described in the text;

3.3. Specifying the marginal rates of substitutions (MRSs);

3.4. Determining the criteria weights;

3.5. Providing preferences with the help of $\epsilon-$constraint;

3.6. Defining preferences with the help of criteria weights and selecting the upper bounds for all objective functions, simultaneously.

Step 4- Based on the preference information, given by the DM in Step 3, set appropriate values for parameters and index sets in formulation (3) (for nondifferentiable MOP) or formulation (4) (for differentiable MOP), and solve it.

Step 5- Present the obtained (weakly, properly) efficient solution(s) and the corresponding objective function vector(s) to the DM. Let DM chooses one of them. In this case, different states can occur:

5.1. If the DM approves this solution as the most preferred one, denote this solution by $\hat{\mathbf{x}}^{k+1}$ and go to Step 6.

5.2. If the DM wants to obtain other solutions with the same preference information, go to Step 4. Note that, in this case, Step 4 should be executed with other values for parameters and index sets.

5.3. If the DM wants to provide new preference information, denote this solution by $\hat{\mathbf{x}}^{k+1}$, set $k := k + 1$ and go to Step 3.

Step 6- Stop.

corresponding scalarizing formulation (3) has a reduced number of constraints which causes a decrease in solving time.

(b) Unlike the algorithms proposed in [21,28,32], the GIA allows the DM to specify his/her preference information in six different ways. Since the satisfaction of DM is an important factor in the interactive algorithms, this aspect of the GIA will increase the satisfaction of DM.

(c) To propose a general algorithm for solving an MOP, it is necessary to convert the MOP problem to a general scalarized problem with perhaps some additional constraints. It is obvious that the number of constraints added to the general scalarized problem has a major effect on the computational time. In Table 1 (for nondifferentiable MOPs) we compare the number of constraints added to the suggested general formulation (3) with those added to some general formulations as GLIDE [21], GLIDE-II [28] and GENWS [32].

Table 1: Number of additional constraints in each formulation (in nondifferentiable case)

Methods	GENWS	GLIDE	GLIDE-II nondiff	Our formulation
GUESS	$card(J_2^k)$	$2p$	0	0
Reference direction approach	–	–	0	0
STOM	p	–	0	0
SPOT	–	$2p$	$p-1$	$p-1$
Modified reference point	$card(J_2^k)+card(J_3^k)$	–	$card(J_3^k)$	$card(J_3^k)$
RD	–	–	–	$card(J_2^k)+card(J_3^k)$
$\epsilon-$ constraint	–	–	–	$p-1$
Weighted sum	–	–	0	0
Hybrid	–	–	–	p
ISWT	–	–	$p-1$	$p-1$
PROJECT	–	$2p$	0	0
Weighted Tchebycheff	–	–	0	0

(d) One of the most important theoretical advantages of the proposed general formulations is that, Theorem 5 enables us to provide results concerning proper efficiency. Unboundedness of the trade-offs means, practically, ignoring at least one of the objective functions when the DM wants to improve another objective function, which is not satisfactory to the DM. Since properly efficient solutions have bounded trade-offs, the DM can improve some unacceptable objective functions with no concern.

(e) All the provided theorems where established without convexity assumptions. In fact, the main MOP can be convex or nonconvex.

7 A numerical example

In this section, we illustrate the procedure mentioned in the GIA on an engineering example of designing a four-bar plane truss, studied in [5]. This

problem, has two conflicting objective functions. We should minimize the volume of the truss (f_1), and its joint displacement (f_2), subject to given physical restrictions on the feasible cross-sectional areas x_1, x_2, x_3, and x_4 of the four bars. The stress on the truss structure is caused by several forces of magnitude F, and $2F$. The length L of each bar and the elasticity constants E and σ of the materials involved are modelled as constants. The mathematical model of this example is as follows:

$$\begin{aligned} \textit{Minimize}\Big\{ & f_1(\mathbf{x}) = L(2x_1 + \sqrt{2}x_2 + \sqrt{2}x_3 + x_4), \\ & f_2(\mathbf{x}) = \frac{FL}{E}\left(\frac{2}{x_1} + \frac{2\sqrt{2}}{x_2} - \frac{2\sqrt{2}}{x_3} + \frac{1}{x_4}\right)\Big\} \\ s.t. \quad & \begin{cases} \frac{F}{\sigma} \leq x_1 \leq \frac{3F}{\sigma}, \\ \sqrt{2}(\frac{F}{\sigma}) \leq x_2 \leq \frac{3F}{\sigma}, \\ \sqrt{2}(\frac{F}{\sigma}) \leq x_3 \leq \frac{3F}{\sigma}, \\ \frac{F}{\sigma} \leq x_4 \leq \frac{3F}{\sigma}, \end{cases} \end{aligned}$$

where, the constant parameters are chosen as $F = 10\ kN$, $E = 2 \times 10^5\ kN/cm^2$, $L = 200\ cm$ and $\sigma = 10\ kN/cm^2$.

The ideal and nadir values for objective functions of this problem are obtained as $\mathbf{y^I} = (y_1^I, y_2^I) = (1400, -5.7191 \times 10^{-4})$ and $\mathbf{y^N} = (y_1^N, y_2^N) = (3.4971 \times 10^3, 0.0406)$. Now, based on the GIA, at first, we should find an initial solution. To this end, the $\epsilon-$constraint scalarizing problem (13) is used, which can be obtained from the proposed formulations by parameters and index sets given in Subsection 4.7, with $l = 2$ and $\epsilon_1 = 1800$. By solving the obtained problem, we find $(1.3906, 1.9963, 1.4142, 1.3957)$ for variables, and $(1800, 0.0157)$ for objective functions. As it can be seen, the values of the objective functions are between ideal and nadir values. Let $\hat{x}^1 = (1.3906, 1.9963, 1.4142, 1.3957)$ and $f(\hat{x}^1) = (1800, 0.0157)$ are shown to DM.

Suppose, the DM wishes to express his/her preference information as the reference point $\bar{\mathbf{y}}^1 = (\bar{y}_1^1, \bar{y}_2^1) = (1600, 0.01)$. Based on this preference given by DM, one of the reference point based scalarizing problems can be selected. Here, we set parameters in our formulation, such that the GUESS scalarizing problem is obtained, and by solving it, $(1.4613, 2.0666, 1.4142, 1.4613)$ is obtained for variables and $(1861.3, 0.0142)$ is attained for the objective function values. At this iteration, the volume of the truss has increased and its joint displacement has slightly decreased. Now, assume that the DM wants to change the type of his/her preference information. According to Step 5 of the GIA, set $\hat{x}^2 = (1.4613, 2.0666, 1.4142, 1.4613)$, $f(\hat{x}^2) = (1861.3, 0.0142)$ and $k = 2$. Now, Step 3 is executed.

Assume that the DM wants to classify the objective functions in two classes $J_1^2 = \{1\}$, and $J_2^2 = \{2\}$. This means, the DM wants to improve f_1 by somewhat relaxing f_2. Assume that the DM gives us $\bar{\mathbf{y}}^2 = (1500, 0.03)$. Now,

the RD scalarizing problem is used with $\alpha = 0.5$. By solving the problem, $(1.1876, 1.6796, 1.4142, 1.1876)$ is obtained for variables and $(1587.6, 0.0221)$ is attained for the objective functions. Assume that the DM wants to provide new preferences by selecting weights for the objective functions. To this end, let $\hat{x}^3 = (1.1876, 1.6796, 1.4142, 1.1876)$, $f(\hat{x}^3) = (1587.6, 0.0221)$, and $k = 3$. Then, Step 4 is executed by $(w_1^3, w_2^3) = (\frac{3}{4}, \frac{1}{4})$ as weights given by the DM. By solving the weighted sum problem (14), $(1, 1.4142, 1.4142, 1)$ is obtained for variables and $(1400, 0.03)$ is obtained for the objective values. As it can be seen, the volume of the truss is in its ideal value, and this satisfies the DM. It is important to point out that by Theorem 5, the obtained objective vector is a properly nondominated point.

8 Conclusions

In this article, we suggested a general scalarizing formulation to obtain a global interactive algorithm for multiobjective optimization problems. We proposed the formulation in two versions; one of them for differentiable and the other for nondifferentiable MOPs. By selecting suitable values for parameters, we proved that optimal solutions of the suggested general scalarizing problem are (weakly, properly) efficient solutions for the main multiobjective problem. Moreover, it was shown that many scalarizing problems used in different interactive methods as GUESS, reference direction approach, Step, STOM, SPOT, modified reference point, and RD methods can be obtained from the proposed general formulation, by selecting suitable transformations. Some of the popular scalarizing problems such as, weighted sum, $\epsilon-$constraint, and hybrid problems derived from our general scalarizing problem. In addition, we proposed a general interactive algorithm. In the proposed algorithm, the DM could express his/her preference information in six different ways, and based on the kind of information given by the DM, a suitable scalarizing problem, by selecting appropriate values for parameters and index sets in the general formulation, was selected. Finally, by a numerical example we illustrated that how the proposed general interactive algorithm can be used.

However, for the future investigation, developing a software based on the suggested general interactive algorithm can be worthwhile. Also, proposing a general interactive procedure for approximate efficient solutions of an MOP can be worth studying. To this end, studying three recently published papers by Ghaznavi-ghosoni[1] and Khorram [10], Ghaznavi-ghosoni et al. [11] and Ghaznavi [12] is recommended.

[1] Previous name of the first author

Acknowledgments:

The authors would like to express their heartfelt thanks to the editors and anonymous referees for their useful suggestions which improved the quality of the paper.

References

1. Benayoun, R., Montgolfier, J., Tergny, J. and Laritchev, O. *Linear programming with multiple objective functions: Step method (STEM)*, Mathematical Programming 1(3) (1971) 366-375.

2. Buchanan, J.T. *A naive approach for solving MCDM problems: the GUESS method*, Journal of the Operational Research Society 48 (1997) 202-206.

3. Chankong, V. and Haimes, Y.Y. *The interactive surrogate worth trade-off (ISWT) method for multiobjective decision-making*, in: S. Zionts (Eds.), Multiple Criteria Problem Solving, Berlin: Springer, 1978, pp. 42-67.

4. Ehrgott, M. *Multicriteria Optimization*, Springer, Berlin, 2005.

5. Engau, A. and Wiecek, M.M. *Generating $\varepsilon-$efficient solutions in multiobjective programming*, European Journal of Operational Research 177 (2007) 1566-1579.

6. Eschenauer, H.A., Osyczka, A. and Schafer, E. *Interactive multicriteria optimization in design process*, in: Eschenauer, H., Koski, J., Osyczka A., (Eds.), Multicriteria Design Optimization Procedures and Applications, Berlin: Springer, 1990, pp. 71-114.

7. Gardiner, L. and Steuer, R.E. *Unified interactive multiple objective programming*, European Journal of Operational Research 74(3) (1994) 391-406.

8. Gardiner, L. and Steuer, R.E. *Unified interactive multiple objective programming: an open architecture for accommodating new procedures*, Journal of the Operational Research Society 45(12) (1994) 1456-1466.

9. Geoffrion, A. *Proper efficiency and the theory of vector maximization*, Journal of Mathematical Analysis and Applications 22 (1968) 618-630.

10. Ghaznavi-ghosoni, B.A. and Khorram, E. *On approximating weakly/properly efficient solutions in multiobjective programming*, Mathematical and Computer Modelling 54 (2011) 3172-3181.

11. Ghaznavi-ghosoni, B.A., Khorram, E. and Soleimani-damaneh, M. *Scalarization for characterization of approximate strong/weak/proper efficiency in multiobjective optimization*, Optimization, 62 (6) (2013) 703-720.

12. Ghaznavi, M. *Optimality conditions via scalarization for approximate quasi efficiency in multiobjective optimization*, Filomat, accepted.

13. Hosseinzadeh Lotfi, F., Jahanshahloo, G.R., Ebrahimnejad, A., Soltanifar, M. and Mansourzadeh, S.M. *Target setting in the general combined-oriented CCR model using an interactive MOLP method*, Journal of Computational and Applied Mathematics 234(1) (2010) 1-9.

14. Jahn, J. *Vector Optimization: Theory, Applications, and Extensions*, Springer-Verlag, Berlin, Germany, 2004.

15. Kaliszewski, I. *A modified weighted Tchebycheff metric for multiple objective programming*, Computers and Operations Research 14 (1987) 315-323.

16. Kaliszewski, I. *A theorem on nonconvex functions and its application to vector optimization*, European Journal of Operational Research 80 (1995) 439–449.

17. Kaliszewski, I. *Using trade-off information in decision making algorithms*, Computers and Operations Research 27 (2000) 161–182.

18. Kaliszewski, I. and Michalowski, W. *Efficient solutions and bounds on tradeoffs*, Journal of Optimization Theory and Applications 94 (1997) 381-394.

19. Kaliszewski, I., Miroforidis, J. and Podkopaev, D. *Interactive multiple criteria decision making based on preference driven evolutionary multiobjective optimization with controllable accuracy*, European Journal of Operational Research 216 (2012) 188-199.

20. Korhonen, P. *Reference direction approach to multiple objective linear programming: Historical overview*, in: Karwan, M. H., Spronk,J., Wallenius, J., (Eds.), Essays in Decision Making: A Volume in Honour of Stanley Zionts, Springer-Verlag, Berlin, Heidelberg, 1997, pp. 74-92.

21. Luque, M., Ruiz, F. and Miettinen, K. *Global formulation for interactive multiobjective optimization*, OR Spectrum 33(1) (2011) 27–48.

22. Luque, M., Yang, J.B. and Wong, B.Y.H. *PROJECT method for multiobjective optimization based on the gradient projection and reference point*, IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans 39(4) (2009) 864-879.

23. Miettinen, K. *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Dordrecht, 1999.

24. Nakayama, H. and Sawaragi, Y. *Satisficing trade-off method for multiobjective programming*, in: M. Grauer, A.P. Wierzbick (Eds.), Interactive Decision Analysis, Berlin: Springer, 1984, pp. 113-122.

25. Narula, S.C., Kirilov, L. and Vassilev, V. *Reference direction approach for solving multiple objective nonlinear programming problems*, IEEE Transactions on Systems, Man, and Cybernetics 24 (1994) 804-806.

26. Park, K.S. and Shin, D.E. *Interactive multiobjective optimization approach to the inputoutput design of opening new branches*, European Journal of Operational Research 220 (2012) 530-538.

27. Romero, C. *Extended lexicographic goal programming: a unified approach*, Omega 29 (2001) 63-71.

28. Ruiz, F., Luque, M. and Miettinen, K. *Improving the computational efficiency in a global formulation (GLIDE) for interactive multiobjective optimization*, Annals of Operations Research 197(1) (2012) 47-70.

29. Sakawa, M. *Interactive multiobjective decision making by the sequential proxy optimization technique: SPOT*, European Journal of Operational Research 9 (4) (1982) 386-396.

30. Taras, S. and Woinaroschy, A. *An interactive multiobjective optimization framework for sustainable design of bioprocesses*, Computers & Chemical Engineering 43 (2012) 10-22.

31. Vassilev, V., Narula, S.C. and Gouljashki, V.G. *An interactive reference direction algorithm for solving multiobjective convex nonlinear integer programming problems*, International Transactions in Operational Research 8 (4) (2001) 367-380.

32. Vassileva, M., Miettinen, K. and Vassilev, V. *Generalized scalarizing problem for multicriteria optimization*, IIT Working Papers IIT/WP-205, Institute of Information Technologies, Bulgaria, 2005.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 101-119

A nonmonotone trust-region-approach with nonmonotone adaptive radius for solving nonlinear systems

K. Amini, H. Esmaeili and M. Kimiaei

Abstract

This paper presents a trust-region procedure for solving systems of nonlinear equations. The proposed approach takes advantages of an effective adaptive trust-region radius and a nonmonotone strategy by combining both of them appropriately. It is believed that selecting an appropriate adaptive radius based on a suitable nonmonotone strategy can improve the efficiency and robustness of the trust-region framework as well as can decrease the computational cost of the algorithm by decreasing the number of subproblems that must be solved. The global convergence to first order stationary points as well as the local q-quadratic convergence of the proposed approach are proved. Numerical experiments show that the new algorithm is promising and attractive for solving nonlinear systems.

Keywords: Nonlinear equations; Trust-region framework; Adaptive radius; Nonmonotone technique.

1 Introduction

In this paper, we consider the nonlinear system of equations

$$F(x) = 0, \quad x \in \mathbb{R}^n, \tag{1}$$

where $F : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable mapping in the form $F(x) := (F_1(x), F_2(x), \cdots, F_n(x))^T$. If $F(x)$ has a zero, then the nonlinear

*Corresponding author
Received 5 April 2015; revised 4 July 2014; accepted 21 October 2015
K. Amini
Department of Mathematics, Faculty of Science, Razi University, Kermanshah, Iran. e-mail: kamini@razi.ac.ir

H. Esmaeili
Department of Mathematics, Bu-Ali Sina University, Hamedan, Iran. e-mail: esmaeili@basu.ac.ir

M. Kimiaei
Department of Mathematics, Asadabad Branch, Islamic Azad University, Asadabad, Iran. e-mail: morteza.kimiaei@gmail.com

system (1) is equivalent to the following nonlinear unconstrained least-squares problem

$$\begin{array}{ll} \min & f(x) := \frac{1}{2}\|F(x)\|^2 \\ \text{s.t.} & x \in \mathbb{R}^n, \end{array} \tag{2}$$

where $\|.\|$ denotes the Euclidean norm. The trust-region frameworks for solving system of nonlinear equations (1) are a popular class of iterative procedures that in each iteration generate a trial step d_k by computing an exact or approximate solution of the following subproblem

$$\begin{array}{l} \min\ m_k(x_k+d) := \frac{1}{2}\|F_k + J_k d\|^2 = f_k + d^T J_k^T F_k + \frac{1}{2} d^T J_k^T J_k d \\ \text{s.t} \quad d \in \mathbb{R}^n \text{ and } \|d\| \le \Delta_k. \end{array} \tag{3}$$

where $f_k := f(x_k)$, $F_k := F(x_k)$, $J_k := F'(x_k)$, Jacobian of $F(x)$, and $\Delta_k > 0$ is trust-region radius. The ratio

$$r_k = \frac{f_k - f(x_k + d_k)}{m_k(x_k) - m_k(x_k + d_k)}. \tag{4}$$

plays a main role in trust region frameworks. Obviously, the model matches the original problem at the current iteration x_k whenever r_k is sufficiently close to 1. Then the agreement is weak or there is no agreement whenever r_k is near zero and there is not agreement when r_k is negative. Generally, if r_k is greater than a positive constant μ, the trial step d_k will be accepted and leading to $x_{k+1} := x_k + d_k$. In this case, it is safe to increase trust region radius in the next iteration. Otherwise, the trust-region radius should be shrunk and the subproblem (3) will be solved again to possibly find an acceptable trial point in the sequel of the process.

It is known that the traditional trust-region framework has some drawbacks: a) the very small trust-region radius Δ_k, increases the total number of the iterates, b) the remarkably large trust-region radius Δ_k increases the total number of solving subproblems, c) ratio (4) does not suffice to create the agreement between the quadratic model and the objective function in trust-region methods leading to increasing computational cost. Using the adaptive radius is an appropriate idea to overcome drawbacks (a) and (b). As a result, many researchers have investigated on finding the best trust-region radius Δ_k, but no one has actually claimed a general rule for generating the trust-region radius. Therefore, in order to decrease the total number of solving subproblems for an arbitrary problem, some adaptive processes determining the radius have been proposed, see [3, 33, 45]. For example, Zhang and Wang [44] proposed an adaptive radius by

$$\Delta_k = c^{p_k}\|F_k\|^{\delta},$$

where $0 < c < 1$ and $0.5 < \delta < 1$ are constant, also p_k is a non-negative integer starting from zero. The major advantage of this method is that the radius does not stay very large and therefore it is possible to prevent resolving

the trust region subproblem. This proposal has some disadvantages: Firstly, the sequence generated by this method is superlinearly convergent with the convergence order 2δ. Secondly, the efficiency of the numerical results is largely dependent on the choice of δ. Finally, this method cannot adequately prevent the very small trust-region radius. To overcome these drawbacks, Fan and Pan in [15] suggest that

$$\Delta_k = c^{p_k} M \|F_k\|,$$

where it is also another satisfactory radius with a constant M, an integer p_k and $c \in (0, 1)$. This choice for the trust region radius plays an important role in proving the quadratic convergence and also prevents some deal from introducing the intensely small trust-region radius. In this method, if $\|F_k\|$ is very small, then the constant M must be chosen so large that the radius is not too small. But for some problems in which $\|F_k\|$ is large, $M\|F_k\|$ will be very large and the number of solving subproblems may be increased. Thus, the amount of computation and the cost of solving problem will be increased.

One of the convenient ways to overcome the drawback (c), is the nonmonotone techniques that can improve the efficiency and the robustness of trust region algorithms, especially when it is applied to highly nonlinear problems, in the presence of narrow curved valley, see for examples [1,2,4,18–20,43,46]. Therefore, a nonmonotone strategy can be employed to increase the efficiency of the proposed procedures.

Contribution. The primary goal in the design of the new method is decreasing computational cost by combining two nonmonotone techniques and adaptive radius trust region. We hope that combining these two techniques can improve numerical performance and efficiency of algorithm. We attain this designed goal by building a new adaptive radius based on nonmonotone technique. The global convergence to first-order critical points along with q-quadratic convergence are being established. The numerical experiments confirm the efficiency and the robustness of the proposed method for solving systems of nonlinear equations.

Organization. This paper is organized as follows: The structure of algorithm will be described after a new adaptive trust-region radius and a nonmonotone technique are proposed in Section 2. In Section 3, we will investigate the global convergence and the quadratic convergence rate of the new algorithm under some suitable assumptions. Numerical results are reported in Section 4. Finally, we end up the paper by some conclusive remarks given in Section 5.

2 Motivation and algorithmic structure

It is well-known that the traditional approaches in unconstrained optimization generally use a globalization technique, like line search or trust-region, to guarantee the global convergence of the algorithm. These globalization techniques mostly enforce a monotonicity to the produced sequence of the objective function values which usually causes a short step to be produced and so a slow numerical convergence encountering highly nonlinear problems in the presence of a narrow curved valley, see [1, 2, 4, 9, 18, 43]. For example, the traditional trust-region framework exploits the ratio (4) which leads to

$$f_k - f_{k+1} \geq m_k(x_k) - m_k(x_k + d_k) > 0.$$

This condition clearly implies that the sequence $\{f_k\}$ should be monotone. In order to avoid this drawback of the Armijo-type line search globalization techniques, a nonmonotone strategy was introduced by Grippo, Lampariello and Lucidi in [18] for unconstrained optimization problems while they modified the Armijo condition by the following condition

$$f(x_k + \alpha_k d_k) \leq f_{l(k)} + \delta \alpha_k g_k^T d_k, \tag{5}$$

where $\delta \in (0, 1)$, $g_k := \nabla f(x_k)$, the gradient of $f(x)$ in x_k, and

$$f_{l(k)} = \max_{0 \leq j \leq m(k)} \{f_{k-j}\}, \quad k \in \mathbb{N} \cup \{0\}, \tag{6}$$

in which $m(0) := 0$ and $0 \leq m(k) \leq \min\{m(k-1) + 1, N\}$ with $N \geq 0$. The theoretical and numerical results have shown that the proposed technique has some remarkable effects and improves both the possibility of finding the global optimum and the rate of convergence for algorithms. Motivated by their work, Deng et al. in [9] made some changes in the ratio (4) which assesses the agreement between the quadratic model and the objective function in trust-region methods. In addition thy introduced the first nonmonotone trust region algorithm. This idea was developed further by Zhou and Xiao [38, 46], Xiao and Chu [37] and Toint [35, 36]. The most common nonmonotone ratio is defined as follows:

$$\widehat{r}_k := \frac{f_{l(k)} - f(x_k + d_k)}{m_k(x_k) - m_k(x_k + d_k)}. \tag{7}$$

To overcome disadvantages (a) and (b), according to the proposed method by Esmaeili and Kimiaei [11]. We define the new adaptive radius by

$$\Delta_k := c^{p_k} N F_{l(k)}, \tag{8}$$

in which $0 < c < 1$, p_k is the smallest nonnegative integer p such that $\widehat{r}_k \geq \mu$ and

$$NF_{l(k)} := \max_{0\le j\le m(k)} \{\|F_{k-j}\|\}, \quad k \in \mathbb{N} \cup \{0\}, \tag{9}$$

in which $m(0) := 0$ and $0 \le m(k) \le \min\{m(k-1)+1, N\}$ with $N \ge 0$. The proposed adaptive trust region radius has some benefits. First, since the sequence $\{NF_{l(k)}\}$ is reduced slowly and is greater than the sequence $\{\|F_k\|\}$ (see (11)), it prevents introducing the intensely small trust-region radius as possible and thus prevents increasing the total number of iterates. Second, Due to the decreasing sequence $\{NF_{l(k)}\}$, Δ_k will not stay too large and it prevents increasing the number of solving subproblems. Hence, using controlling the radius of trust-region, the new method can prevent the production of larger trial step near the optimizer and smaller trial step far from the optimizer.

Our assumptions are identical to those utilized for the proposed approach:

(H1) The level set $L(x_0) := \{x \in \mathbb{R}^n \mid f(x) \le f(x_0)\}$ is bounded for any given $x_0 \in \mathbb{R}^n$ and $F(x)$ is continuously differentiable on compact convex set Ω containing the level set $L(x_0)$.

(H2) The matrix $\{J_k\}$ is bounded and uniformly nonsingular on Ω, i.e. there exists constants $0 < M_0 \le 1 \le M_1$ such that

$$\|J_k\| \le M_1 \quad \text{and} \quad M_0\|F_k\| \le \|J_k^T F_k\| \quad \forall k \in \mathbb{N} \cup \{0\}, \tag{10}$$

.

(H3) The decrease on the model m_k is at least as much as a fraction of that obtained by the Cauchy point, i.e. there exists a constant $\beta \in (0,1)$ such that

$$m_k(x_k) - m_k(x_k + d_k) \ge \beta \ \|J_k^T F_k\| \ \min\left\{\Delta_k, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}, \tag{11}$$

for all $k \in \mathbb{N} \cup \{0\}$.

(H4) $J(x)$ is Lipschitz continuous in $L(x_0)$, with Lipschitz constant γ_L.

We now incorporate both of the two nonmonotone and adaptive radius terms into trust-region and outline the subsequent algorithm:

Algorithm 1 NATR (Nonmonotone Adaptive Trust-Region Algorithm)

Input: An initial point $x_0 \in \mathbb{R}^n$, $c, \mu \in (0,1)$, $N > 0$, $\epsilon > 0$ and k_{max}.
Output: x_b, f_b;

1 **Begin**
2 $\Delta_0 := \|F_0\|$; $f_{l(0)} := 1/2\|F_{l(0)}\|^2$; $NF_{l(0)} := \|F_0\|$; $m(0) := 0$; $k := 0$;
3 **While** $\|F_k\| \geq \epsilon$ && $k \leq k_{max}$ **do**
4 $p := 0$; $\widehat{r}_k := 0$;
5 **While** $\widehat{r}_k < \mu$ **do**
6 specify the trial point d_k by solving the subproblem (3) ;
7 compute $F(x_k + d_k)$;
8 $f(x_k + d_k) := 1/2 \ \|F(x_k + d_k)\|^2$;
9 determine $\widehat{r}_k$ using (7);
10 **If** $\widehat{r}_k < \mu$ **then**
11 $p \leftarrow p + 1$;
12 determine Δ_k using (8);
13 **end**
14 **end**
15 $x_{k+1} := x_k + d_k$; $F_{k+1} := F(x_{k+1})$; $f_{k+1} := f(x_{k+1})$; $J_{k+1} := J(x_{k+1})$;
16 compute J_{k+1} and let $m(k+1) := \min\{m(k)+1, N\}$;
17 calculate $NF_{l(k+1)}$ by (9) and set $f_{l(k+1)} := 1/2NF^2_{l(k+1)}$;
18 select $\Delta_{k+1} := NF_{l(k+1)}$;
19 $k \leftarrow k + 1$;
20 **end**
21 $x_b := x_k$; $f_b := f_k$;
22 **end**

In Algorithm 1, the cycle starting from Line 3 to Line 25 is called the outer cycle, and the cycle starting from Line 5 to Line 14 is called the inner cycle.

Remark 1. The inequality (11) is called the sufficient reduction condition, see [32] and has been investigated by many authors when they extended some inexact methods for solving subproblem (3), for example see [11,13–15]. For global convergence purpose, it is enough to find a vector d_k such that it gives a sufficient reduction in the quadratic model m_k. Well-known convergence results [31] show that the trial step d_k is required to give a reduction in the model m_k that is at least some fixed multiple of the decrease attained by the Cauchy step at each iteration.

Lemma 1. *Suppose that (H4) holds, the sequence $\{x_k\}$ is generated by Algorithm 1 and d_k is a solution of the subproblem (3) such that $\|F(x_k) + J(x_k)d_k\| \leq \|F(x_k)\|$. Then, we have*

$$|f(x_k + d_k) - m_k(x_k + d_k))| \leq O(\|d_k\|^2). \tag{12}$$

Proof. See [9]. □

The following lemma indicates that the inner cycle of Algorithm 1 terminates in a finite number of inner iterates.

Lemma 2. *Suppose that (H2)-(H4) hold and the sequence $\{x_k\}$ is generated by Algorithm 1. Then, the inner cycle of Algorithm 1 is well-defined.*

Proof. Assume that the inner cycle of Algorithm 1 cycles infinitely in the iteration k, i.e., $\Delta_k^p = c^p NF_{l(k)} \to 0$ as $p \to \infty$, equivalently, for any $\eta > 0$, we have $\Delta_k^p < \eta$ for sufficiently large p. Using the fact that x_k is not the optimum of (2), we can conclude that there exists a constant $\epsilon > 0$ such that $\|F_k\| \geq \epsilon$. Without loss of generality, let $\eta := \frac{M_0\epsilon}{M_1^2}$. This fact along with (H2) and (11) imply

$$\begin{aligned} m_k(x_k) - m_k(x_k + d_k^p) &\geq \beta \|J_k^T F_k\| \ \min\left\{\Delta_k^p, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} \\ &\geq \beta M_0 \|F_k\| \ \min\left\{\Delta_k^p, \frac{M_0\epsilon}{M_1^2}\right\} \\ &\geq \beta M_0 \epsilon \ \min\{\Delta_k^p, \eta\} \\ &= \beta M_0 \epsilon \Delta_k^p, \end{aligned} \tag{13}$$

where d_k^p is a solution of subproblem (3) corresponding to p in k-th iterate. Now, Lemma 1 and (13) leads to

$$\begin{aligned} \left|\frac{f_k - f(x_k + d_k^p)}{m_k(x_k) - m_k(x_k + d_k^p)} - 1\right| &= \left|\frac{f(x_k + d_k^p) - m_k(x_k + d_k^p)}{m_k(x_k) - m_k(x_k + d_k^p)}\right| \\ &\leq \frac{O(\|d_k^p\|^2)}{\beta M_0 \epsilon \Delta_k^p} \leq \frac{O((\Delta_k^p)^2)}{\beta M_0 \epsilon \Delta_k^p} \to 0, \quad \text{as} \quad p \to \infty. \end{aligned}$$

Therefore, there exists a sufficiently large p_k such that

$$r_k = \frac{f_k - f(x_k + d_k^{p_k})}{m_k(x_k) - m_k(x_k + d_k^{p_k})} \geq \mu.$$

Besides, from the definition $f_{l(k)}$, it is clear that $f_{l(k)} \geq f_k$. This fact along with the previous inequality immediately implies $\widehat{r}_k \geq r_k \geq \mu$ which means that the inner cycle of Algorithm 1 stops and so Algorithm 1 is well-defined. □

Lemma 3. *Suppose that (H1) holds and the sequence $\{x_k\}$ is generated by Algorithm 1. Then, for all $k \in \mathbb{N} \cup \{0\}$, we have $x_k \in L(x_0)$ and the sequence $\{NF_{l(k)}\}$ is decreasing and convergent.*

Proof. Using the definition of $NF_{l(k)}$, we have

$$\|F_k\| \leq NF_{l(k)}.$$

By induction, the result evidently holds for $k = 0$ because $NF_{l(0)} = \|F_0\|$. Assuming $x_i \in L(x_0)$ for $i = 1, 2, \ldots, k$, we show that $x_{k+1} \in L(x_0)$, for all $k \in \mathbb{N}$. It can be seen

$$\frac{NF_{l(k)}^2}{2} - \frac{\|F_{k+1}\|^2}{2} = f_{l(k)} - f_{k+1} \geq \mu(m_k(x_k) - m_k(x_k + d_k)) > 0,$$

so

$$\|F_{k+1}\| \leq NF_{l(k)} \leq \|F_0\|.$$

Thus, the sequence $\{x_k\}$ is contained in $L(x_0)$. It will be proved that the sequence $\{NF_{l(k)}\}$ is a decreasing sequence. We consider two following cases:

i) $k \geq N$. In this case, we have $m(k) = N$. So, the definition of $NF_{l(k)}$ along with this fact that $\|F_{k+1}\| \leq NF_{l(k)}$ implies that

$$\begin{aligned} NF_{l(k+1)} &= \max_{0 \leq j \leq N} \{\|F_{k+1-j}\|\} \leq \max\{\max_{0 \leq j \leq N} \{\|F_{k-j}\|\}, \|F_{k+1}\|\} \\ &= \max\{NF_{l(k)}, \|F_{k+1}\|\} = NF_{l(k)}. \end{aligned}$$

ii) $k < N$. In this case, we have $m(k) = k$. For any k, $\|F_k\| \leq \|F_0\|$,

$$NF_{l(k)} = F_0, \quad \forall k.$$

These cases show that the sequence $\{NF_{l(k)}\}$ is a decreasing sequence. According to assumption H_1 and $x_k \in L(x_0)$, one can see that the sequence $\{NF_{l(k)}\}$ is convergent. □

By Lemma 3 and since $f(x_k) = \frac{1}{2}\|F(x_k)\|^2$, we can conclude that the sequence $\{f_{l(k)}\}$ is also decreasing and convergent.

3 Convergence theory

In this section, we provide the global convergence and q-quadratic rate of results of the proposed algorithm.

Lemma 4. *Suppose that $\{x_k\}$ is the sequence generated by Algorithm 1. Then, we have*

$$\lim_{k \to \infty} NF_{l(k)} = \lim_{k \to \infty} \|F(x_k)\|.$$

Proof. By Lemma 3.2 in [1] and $f(x_k) = \frac{1}{2}\|F(x_k)\|^2$, we have

$$\lim_{k \to \infty} f_{l(k)} = \lim_{k \to \infty} f(x_k).$$

This implies that

$$\lim_{k \to \infty} NF_{l(k)} = \lim_{k \to \infty} \|F(x_k)\|.$$

□

In order to establish the global convergence of Algorithm 1, one needs to establish the following results.

Lemma 5. *Suppose that assumptions (H2) and (H3) hold, the sequence $\{x_k\}$ is generated by Algorithm 1 and d_k is a solution of the subproblem (3). Then, we have*

$$m_k(x_k) - m_k(x_k + d_k) \geq L_k \ \|F_k\|^2, \tag{14}$$

where $L_k := \beta M_0 \min\left\{c^{p_k}, \frac{M_0}{M_1^2}\right\}$.

Proof. Using (H2) and (11), we have

$$\begin{aligned} m_k(x_k) - m_k(x_k + d_k) &\geq \beta \|J_k^T F_k\| \min\left\{\Delta_k, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} \\ &= \beta \|J_k^T F_k\| \min\left\{c^{p_k} F_{l(k)}, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\} \\ &\geq \beta M_0 \|F_k\| \min\left\{c^{p_k} \|F_k\|, \frac{M_0 \|F_k\|}{M_1^2}\right\} \\ &\geq \beta M_0 \|F_k\|^2 \min\left\{c^{p_k}, \frac{M_0}{M_1^2}\right\} \\ &= \ L_k \|F_k\|^2, \end{aligned}$$

where $L_k = \beta M_0 \min\left\{c^{p_k}, \frac{M_0}{M_1^2}\right\}$. Therefore, the proof is complete. □

At this point, the global convergence of Algorithm 1 based on the mentioned assumptions can be investigated.

Theorem 6. *Suppose that Assumptions (H1)-(H4) hold. Then, Algorithm 1 either stops at a stationary point of $f(x)$ or generates an infinite sequence $\{x_k\}$ such that*

$$\lim_{k \to \infty} \|F_k\| = 0. \tag{15}$$

Proof. By contradiction, let there exists a constant $\epsilon > 0$ and an infinite subset $K \subseteq \mathbb{N}$ satisfying

$$\|F_k\| > \epsilon, \quad \text{for all} \ \ k \in K. \tag{16}$$

Using (16), $\widehat{r}_k > \mu$ and Lemma 5, we can conclude that

$$f_{l(k)} - f_{k+1} = f_{l(k)} - f(x_k + d_k) \geq \mu[m_k(x_k) - m_k(x_k + d_k)] \geq \mu \|F_k\|^2 L_k \geq \mu \epsilon^2 L_k.$$

The left-hand side of above inequality tends to become zero when k goes to infinity and therefore L_k tends to 0. This means that $p_k \to \infty$ that clearly is a contradiction with Lemma 2. Therefore, the hypothesis (16) is not true and the proof is complete. □

To establish the quadratic convergence rate of the sequence generated by Algorithm 1, an additional assumption is required as follows (see [11, 14, 15, 40, 44]).

(H5) There exist constants $c_1 \geq 1$ and $\rho_1 \in (0,1)$ such that

$$c_1\|x - x_*\| \leq \|F(x)\| = \|F(x) - F(x_*)\|, \quad \forall x \in N(x_*, \rho_1).$$

where x_* is a solution of (1) and $N(x_*, \rho_1) := \{x| \ \|x - x_*\| \leq \rho_1\}$.

Remark 2. By (H1) and (H4), the objective function $F(x)$ is continuously differentiable and $J(x)$ is Lipschitz continuous. So, there exist two constants $\gamma_L > 0$ and $\rho_2 \in (0,1)$ such that

$$\|F(x) - F(y) + J(x)(x-y)\| \leq \gamma_L\|x-y\|^2, \quad \text{for all} \ \ x, y \in N(x_*, \rho_2).$$

For the purpose of our q-quadratic convergence, we simply choose $\rho := \min[\rho_1, \rho_2]$.

Theorem 7. *Suppose that Assumptions (H1)-(H5) hold and the sequence $\{x_k\}$ generated by Algorithm 1 converges to x_*. Then, for k sufficiently large, we have*

$$x_{k+1} = x_k + d_k^0,$$

where d_k^0 is the solution of (3) corresponding to $p_k = 0$. Furthermore, the sequence $\{x_k\}$ converges to x_* q-quadratically.

Proof. Let d_k^0 be a solution corresponding to $p_k = 0$ of the subproblem (3) , so d_k^0 is a feasible solution for (3). This along with Lemma 3 and Theorem 1 imply

$$\|d_k^0\| \leq \Delta_k^0 = NF_{l(k)} \to 0, \quad as \ \ k \to \infty. \tag{17}$$

On the other hand, since $p_k = 0$ and $M_0 \leq 1 \leq M_1$, we obtain

$$L_k := \frac{\beta M_0^2}{M_1^2}. \tag{18}$$

Because of the fact that Algorithm 1 is not stopped, it is clear that we have $\|F_k\| \geq \epsilon$. This fact together with Lemma 2, (17) and (18) suggests that

$$\begin{aligned}\left|\frac{f_k - f(x_k + d_k^0)}{m_k(x_k) - m_k(x_k + d_k^0)} - 1\right| &= \left|\frac{m_k(x_k + d_k^0) - f(x_k + d_k^0)}{m_k(x_k) - m_k(x_k + d_k^0)}\right| \\ &\leq \frac{O(\|d_k^0\|^2)}{L_k\|F_k\|^2} \leq \frac{O((\Delta_k^0)^2)}{\frac{\beta M_0^2}{M_1^2}\epsilon^2} \to 0, \quad \text{as} \ \ k \to \infty.\end{aligned}$$

This along with the fact $f_{l(k)} \geq f_k$, for sufficiently large k, implies

$$\widehat{r}_k \geq \mu.$$

Thus, for all sufficiently large k, the trial point d_k^0 is accepted by Algorithm 1, i.e. $x_{k+1} = x_k + d_k^0$.

At this point, the quadratic convergence of the sequence $\{x_k\}$ generated by Algorithm 1 is investigated. Regarding (H2), there exists a constant $M_1 > 0$ such that

$$\|J_k\| \leq M_1, \quad \text{for all } x \in \Omega. \tag{19}$$

Using (19) along with the mean value theorem, for all $x_k \in N(x_*, \rho)$, we can easily see that

$$\|F_k\| = \|F_k - F(x_*)\| = \|J(\xi)\| \, \|x_k - x_*\| \leq M_1 \|x_k - x_*\|, \tag{20}$$

for some $\xi \in [x_k, x_*]$. As a result of this fact and Lemma 4, for any sufficiently large k, it can be concluded that

$$F_{l(k)} \leq M_1 \|x_k - x_*\|,$$

and so

$$\|d_k^0\| \leq N F_{l(k)} \leq M_1 \|x_k - x_*\|. \tag{21}$$

From (H5), it is clear that

$$\|x_k - x_*\| \leq \frac{1}{c_1}\|F_k\| \leq \frac{1}{c_1} N F_{l(k)} \leq N F_{l(k)} = \Delta_k^0.$$

This fact directly implies that $x_k - x_*$ is a feasible point for the subproblem (3). Now, it is straightforwardly followed from Remark 2 and (21) that

$$\begin{aligned} \frac{1}{2}\|F_k + J_k d_k^0\|^2 = m_k(x_k + d_k^0) &\leq m_k(x_k + (x_* - x_k)) \\ &= \frac{1}{2}\|F(x_k + J_k(x_k - x_*))\|^2 \\ &= \frac{1}{2}\|F_k - F_* + J_k(x_k - x_*)\|^2 \\ &\leq \frac{\gamma_L^2}{2}\|x_k - x_*\|^4. \end{aligned} \tag{22}$$

Also (H6), (21) and (22), give us

$$\begin{aligned} c_1\|x_{k+1} - x_*\| \leq \|F(x_{k+1})\| &= \|F(x_k + d_k^0)\| \\ &\leq \|F_k + J_k d_k^0\| + O(\|d_k^0\|^2) \\ &\leq \gamma_L \|x_k - x_*\|^2, \end{aligned}$$

for any sufficiently large k. So

$$\|x_{k+1} - x_*\| = O(\|x_k - x_*\|^2).$$

Hence, the sequence $\{x_k\}$ generated by the Algorithm 1 is q-quadratically convergent. Therefore, the proof is completed. □

4 Numerical experiments

In this section, we report some numerical experiments obtained by running Algorithm 1 (NATR) in comparison with the nonmonotone trust-region algorithm (NTR), the adaptive trust-region algorithm from Zhang et al. in [44] (ATRZ), the nonmonotone version of it (NATRZ), the adaptive trust-region algorithm of Fan and Pan in [15] (ATRF) and the nonmonotone version of it (NATRF) on a set of nonlinear systems of equations with the dimension from 100 to 504 that are selected from the wide range of literatures . The problems 1-36 are chosen from Cruz et al. in [25] and the problems 37-42 are chosen from Lukšan and Vlček in [28]. For all of these codes, the trust-region subproblems are solved by Steihaug-Toint procedure, see [8]. The Steihaug-Toint algorithm terminates at $x_k + d$ when

$$\|\nabla m_k(x_k + d)\| \leq \min\left\{\frac{1}{10}, \|\nabla m_k(x_k + d)\|^{\frac{1}{2}}\right\}\|\nabla m_k(x_k + d)\|.$$

The Jacobian matrix J_k can be either evaluated analytically by a user-supplied function or approximated using finite-differences formula provided by the code. Since the exact computation cannot be appropriate for large scale problems, similar to [5], we used the following finite-differences formula to approximate the Jacobian matrix J_k

$$[J_k]_{\cdot j} \sim \frac{1}{h_j}(F(x_k + h_j e_j) - F_k),$$

where $[J_k]_{\cdot j}$ denotes the j-th column of J_k, e_j is the j-th vector of the canonic basis and

$$h_j := \begin{cases} \sqrt{\epsilon_m} & \text{if } x_{k_j} = 0, \\ \sqrt{\epsilon_m} sign(x_{k_j}) \max\{|x_{k_j}|, \frac{\|x_k\|_1}{n}\} & \text{otherwise.} \end{cases}$$

All codes are written in MATLAB 9 programming environment with double precision format in the same subroutine. In our numerical experiments, the algorithms were stopped whenever

$$\|F_k\| \leq 10^{-5},$$

or when the total number of iterates exceeded 1000. During implementations, It is checked that the codes be converged to the same point and only provided data for problems that all algorithms converged to the identical point while less than of 1 percent of problem was ignored.

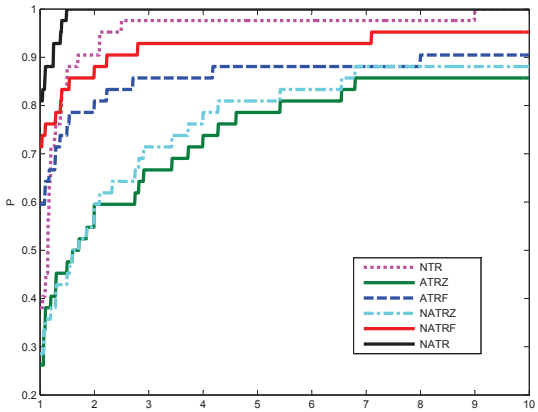

Figure 1: Iterates performance pròfile for the presented algorithms

While NATR algorithm takes advantages of the parameters $\mu = 10^{-6}$, $c = 0.5$. The NTR algorithm employs the parameters $\mu_1 = 0.1$, $\mu_2 = 0.9$ and updates trust-region radius like [8] by the following formula

$$\Delta_{k+1} := \begin{cases} c_1\|d_k\| & \text{if } r_k < \mu_1, \\ \Delta_k & \text{if } \mu_1 \leq r_k \leq \mu_2, \\ c_2\Delta_k & \text{if } r_k \geq \mu_2, \end{cases}$$

where $c_1 = 0.25$ and $c_2 = 0.3$. We also decide to follow the literature [34] in exploiting $\Delta_0 = 1$ as an initial trust-region radius for NTR. The parameters of ATRZ and ATRF have been chosen the same as in articles [44] and [15], respectively. Table 1 indicates the names and dimensions of the test problems considered. Figures 1 and 2 give the performance profiles for all of the algorithms with the choice of finite-differences Jacobian matrix for total number of iterations and total number of function evaluations, respectively. Performance profile gives, for every $\tau \geq 1$, the fraction of the number of problems for which the algorithm is within a factor of τ of the best [10].

Figure 1 clearly indicates that NATR outperforms NTR, ATRZ, ATRF, NATRZ and NATRF regarding the total number of iterates. In particular, NATR has the most wins in nearly 81% of the test problems with the greatest efficiency. Meanwhile, in the sense of the ability of completing a run successfully, it is the best among considered algorithms because it grows up faster than the others and reaches 1 more rapidly. However, as illustrated in Figure 2, NATR implements are remarkably better than the others where it has the most wins for almost 77% of performed tests concerning the total number of function evaluations. Furthermore, Figures 1 and 2 show similar patterns in

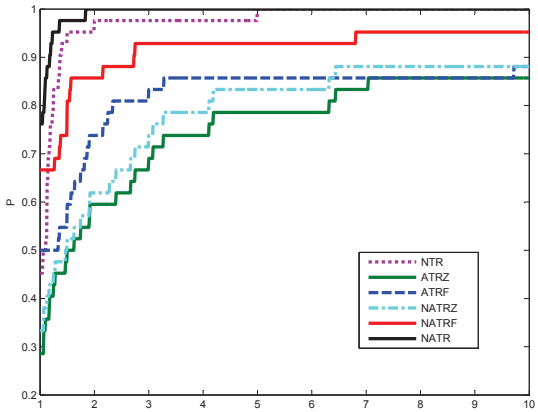

Figure 2: Function evaluations performance profile for the presented algorithms

the sense of the ability of completing a run successfully. As a result, this fact directly implies that the total number of solving the trust-region subproblems is the notably decreased thanks to using the NATR algorithm. These results imply that the proposed algorithm is an efficient and robust approach for solving systems of nonlinear equations.

5 Concluding remarks

It is well-known that trust-region methods for solving systems of nonlinear equations have a remarkable numerical reliability as well as strong theoretical convergence properties. Practical experiments of the trust-region framework indicate that applying nonmonotone adaptive techniques for determining trust-region radius declines the number of solving subproblems and employing nonmonotone strategies increases the efficiency and robustness of the algorithm. In this paper, by exploiting an effective adaptive trust region radius based on a reliable nonmonotone strategy, a new nonmonotone trust region algorithm is introduced for solving systems of nonlinear equations. Nevertheless, these modifications in the traditional trust-region procedure are favorably encouraging so that the global and q-quadratic convergence properties of the proposed algorithms are established. Numerical results on a set of nonlinear systems indicate that the number of iterates and the number of function evaluations are so close to each other that, by

Table 1: List of test functions

Problem name	Dim	Problem name	Dim
Exponential 1	500	Geometric	100
Exponential 2	500	Function 27	500
Extended Rosenbrock	500	Tridimensional valley	501
Chandrasekhar's H-equation	500	Complementary	500
Singular	500	Hanbook	500
Logarithmic	500	Tridiagnal system	500
Broyden tridiagonal	500	Five-diagonal system	500
Trigexp	500	Seven-diagonal system	504
Variable band 1	500	Extended cragg and levy	500
Variable band 2	500	Extended Wood	500
Function 15	500	Triadiagnal exponential	500
Strictly convex 1	500	Brent	500
Strictly convex 2	500	Thorech	500
Function 18	501	Broyden banded	500
Zero Jacobian	500	Discrete integral equation	500
Geometric programming	100	Countercurrent reactors 1	504
Function 21	501	Singular Broyden	500
Linear function-full rank 1	500	Structured Jacobian	500
Linear function-full rank 2	500	Extended Powell Singular	500
Brown almost linear	500	Generalized Broyden banded	500
Variable dimensioned	500	Extended powell badly scaled	500

applying the proposed algorithm, significant profits in computational costs can be obtained.

References

1. Ahookhosh, M. and Amini K. *A nonmonotone trust-region method with adaptive radius for unconstrained optimization*, Computers and Mathematics with applications 60 (2010) 411–422.

2. Ahookhosh, M., Amini K. and Peyghami, M.R. *A nonmonotone trust-region line search method for large-scale unconstrained optimization*, Applied Mathematical Modelling 36 (2012) 478–487.

3. Ahookhosh, M., Amini, K. and Kimiaei, M. *A globally convergent trust-region method for large-scale symmetric nonlinear systems*, Numerical Functional Analysis and Optimization 36 (2015) 830–855.

4. Ahookhosh, M., Esmaeili, H. and Kimiaei, M. *An effective trust-region-based approach for symmetric nonlinear systems*, International Journal of Computer Mathematics 90 (2013) 671–690.

5. Bellavia, S., Macconi, M. and Morini, B. *STRSCNE: A scaled trust-region solver for constrained nonlinear equations*, Computational Optimization and Applications 28, (2004) 31–50.

6. Bouaricha, A. and Schnabel, R.B. *Tensor methods for large sparse systems of nonlinear equations*, Mathematical Programming 82 (1998) 377–400.

7. Broyden, C.G. *The convergence of an algorithm for solving sparse nonlinear systems*, Mathematics of Computation 25 (114) (1971) 285–294.

8. Conn, A.R., Gould, N.I.M. and Toint, Ph.L. *Trust-Region Methods*, Society for Industrial and Applied Mathematics SIAM, Philadelphia, 2000.

9. Deng, N.Y., Xiao, Y. and Zhou, F.J. *Nonmonotonic trust region algorithm*, Journal of Optimization Theory and Applications 26 (1993) 259–285.

10. Dolan, E.D. and Moré, J.J. *Benchmarking optimization software with performance profiles*, Mathematical Programming 91 (2002) 201–213.

11. Esmaeili, H. and Kimiaei, M. *A new adaptive trust-region method for system of nonlinear equations*, Appl. Math. Model. 38 (11–12) (2014) 3003–3015.

12. Esmaeili, H. and Kimiaei, M. *An efficient adaptive trust-region method for systems of nonlinear equations,* International Journal of Computer Mathematics 92 (2015) 151–166.

13. Fan, J.Y. *Convergence rate of the trust region method for nonlinear equations under local error bound condition*, Computational Optimization and Applications. 34 (2005) 215–227.

14. Fan, J.Y. *An improved trust region algorithm for nonlinear equations*, Computational Optimization and Applications 48(1) (2011) 59–70.

15. Fan, J.Y. and Pan, J.Y *A modified trust region algorithm for nonlinear equations with new updating rule of trust region radius*, International Journal of Computer Mathematics 87 (14) (2010) 3186–3195.

16. Fasano, G., Lampariello, F. and Sciandrone, M. *A truncated nonmonotone Gauss-Newton method for large-scale nonlinear least-squares problems*, Computational Optimization and Applications 34 (3) (2006) 343–358.

17. Fischer, A., Shukla, P.K. and Wang, M. *On the inexactness level of robust Levenberg-Marquardt methods*, Optimization 59 (2) (2010) 273–287.

18. Grippo, L., Lampariello, F. and Lucidi, S. *A nonmonotone line search technique for Newton's method*, SIAM Journal on Numerical Analysis 23 (1986) 707–716.

19. Grippo, L., Lampariello, F. and Lucidi, S. *A truncated Newton method with nonmonotone linesearch for unconstrained optimization*, Journal of Optimization Theory and Applications 60 (3) (1989) 401–419.

20. Grippo, L., Lampariello, F. and Lucidi, S. *A class of nonmonotone stabilization method in unconstrained optimization*, Numer. Math. 59 (1991) 779–805.

21. Grippo, L. and Sciandrone, M. *Nonmonotone derivative-free methods for nonlinear equations*, Computational Optimization and Applications 37 (2007) 297–328.

22. Kanzow, C., Yamashita, N. and Fukushima, M. *Levenberg-Marquardt methods for constrained nonlinear equations with strong local convergence properties*, J Comput Appl Math. 172 (2004) 375–397.

23. Kanzow, C. and Kimiaei, M. *A new class of nonmonotone adaptive trust-region methods for nonlinear equations with box constraints*, Preprint 330, Institute of Mathematics, University of Wurzburg, Wurzburg, June 2015. http://www.mathematik.uni-wuerzburg.de/ kanzow/paper/NonmonotoneTR_FinalP.pdf

24. La Cruz, W. and Raydan, M. *Nonmonotone spectral methods for large-scale nonlinear systems*, Optimization Methods and Software 18 (5) (2003) 583–599.

25. La Cruz, W., Venezuela, C., Martínez, J.M. and Raydan, M. *Spectral residual method without gradient information for solving large-scale nonlinear systems of equations: Theory and experiments*, Technical Report RT–04–08, July 2004.

26. Li, D.H. and Fukushima, M. *A derivative-free line search and global convergence of Broyden-like method for nonlinear equations*, Optimization Methods and Software. 13 (2000) 181–201.

27. Lukšan, L. and Vlček, J. *Sparse and partially separable test problems for unconstrained and equality constrained optimization*, Technical Report, No. 767 (1999).

28. Lukšan, L. and Vlček, J. *Truncated trust region methods based on preconditioned iterative subalgorithms for large sparse systems of nonlinear equations*, Journal of Optimization Theory and Applications 95 (3) (1997) 637–658.

29. Macconi, M., Morini, B. and Porcelli, M. *A Gauss–Newton method for solving bound–constrained underdetermined nonlinear systems*, Optimization Methods and Software 24 (2) (2009) 219–235.

30. Nesterov, YU. *Modified Gauss–Newton scheme with worst case guarantees for global performance*, Optimization Methods and Software 22 (3) (2007) 469–483.

31. Nocedal, J. and Wright, S.J. *Numerical Optimization*, Springer, NewYork (2006).

32. Powell, M.J.D. *Convergence properties of a class of minimization algorithms, in Nonlinear Programming*, O.L. Mangasarian, R.R. Meyer, and S.M. Robinson, eds., Academic Press, NewYork, 1975, pp. 1–27.

33. Shi, Z.J. and Guo, J.H. *A new trust region method with adaptive radius*, Computational Optimization and Applications, 213 (2008) 509–520.

34. Toint, Ph.L.*Numerical solution of large sets of algebraic nonlinear equations*, Mathematics of Computation 46 (173) (1986) 175–189.

35. Toint, Ph.L. *Non-monotone trust region algorithm for nonlinear optimization subject to convex constraints*, Math. Program. 77 (1997) 69–94.

36. Toint, Ph.L. *An assessment of nonmonotone line search technique for unconstrained optimization*, SIAM J. Sci. Comput. 17 (1996) 725–739.

37. Xiao Y. and Chu, E.K.W. *Nonmonotone trust region methods*, Technical Report 95/17, Monash University, Clayton, Australia, 1995.

38. Xiao, Y. and Zhou, F.J. *Nonmonotone trust region methods with curvilinear path in unconstrained optimization*, Computing 48 (1992) 303–317.

39. Yamashita, N. and Fukushima, M. *On the rate of convergence of the Levenberg-Marquardt method*, Computing 15 (2001) 239–249.

40. Yuan, G., Lu, S. and Wei, Z. *A new trust-region method with line search for solving symmetric nonlinear equations*, International Journal of Computer Mathematics 88 (10) (2011) 2109–2123.

41. Yuan, G.L., Wei, Z.X. and Lu, X.W. *A BFGS trust-region method for nonlinear equations*, Computing 92 (4) (2011) 317–333.

42. Yuan, Y. *Trust region algorithm for nonlinear equations*, Information 1 (1998) 7–21.

43. Zhang, H.C. and Hager, W.W. *A nonmonotone line search technique for unconstrained optimization*, SIAM journal on Optimization 14 (4) (2004) 1043–1056.

44. Zhang, J. and Wang, Y. *A new trust region method for nonlinear equations*, Mathematical Methods of Operation Research 58 (2003) 283–298.

45. Zhang, X.S., Zhang, J.L. and Liao, L.Z. *An adaptive trust region method and its convergence*, Science in China 45 (2002) 620–631.

46. Zhou, F. and Xiao, Y. *A class of nonmonotone sta bilization trust region methods*, Computing 53 (2) (1994) 119–136.

Iranian Journal of Numerical Analysis and Optimization
Vol. 6, No. 1, (2016), pp 121-136

A contractive mapping on fuzzy normed linear spaces

M. Saheli*

Abstract

In this paper, we use the definition of fuzzy normed spaces given by Bag and Samanta and provide four types of fuzzy versions of contraction. We show that these mappings necessarily have unique fixed points in fuzzy normed linear spaces. Moreover we prove that the presented theorems are indeed fuzzy extensions of their classical counterparts.

Keywords: Fuzzy norm; Fuzzy normed linear space; Fixed point; $\alpha-$ seminorm; Contractive conditions.

1 Introduction

Banach contraction mapping principle is one of the fundamental consequences of analysis. This contraction mapping is an important object in metric fixed point theory. Also its emphasis lies on its wide applicability in branches of mathematics. Some contractive conditions have been introduced in [4, 6, 7, 9, 10].

A natural question is whether we can provide contractive conditions which imply existence of fixed point in a fuzzy Banach space. Recently, Shukla and Chauhan [11] defined the concept of cyclic representation and proved some fixed point results for operators on complete fuzzy metric spaces. In [1], AL-Mayahi and Hadi proved that α-η-φ-contraction functions have a fixed point on fuzzy metric space. Das and Saha [5] considered uniformly locally contractive mappings on a fuzzy metric space and showed that these functions have a unique fixed point. Manro and Tomar [8] focused on the compatibility and non-compatibility of pair of self-maps and established existence of fixed point of the compatible maps on fuzzy metric space.
In this paper, we use the definition of fuzzy normed spaces given in [2] and discuss four types of fuzzy versions of contraction and some corollaries.
We give below some basic preliminaries required for this paper.

*Corresponding author
Received 23 April 2014; revised 6 August 2014; accepted 18 March 2015
M. Saheli
Department of Mathematics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran. e-mail: saheli@vru.ac.ir

Definition 1. [2] Let X be a linear space over R (real number) and N be A fuzzy subset of $X \times R$ such that for all $x, u \in X$ and $c \in R$
$(N1)$ $N(x,t) = 0$ for all $t \leq 0$,
$(N2)$ $x = 0$ if and only if $N(x,t) = 1$ for all $t > 0$,
$(N3)$ If $c \neq 0$ then $N(cx,t) = N(x, t/|c|)$ for all $t \in R$,
$(N4)$ $N(x+u, s+t) \geq min\{N(x,s), N(u,t)\}$ for all $s, t \in R$,
$(N5)$ $N(x,.)$ is a nondecreasing function of R and $\lim_{t\to\infty} N(x,t) = 1$.
Then N is called a fuzzy norm on X.

Assume that
$(N6)$ $N(x,t) > 0$ for all $t > 0$ implies $x = 0$,
$(N7)$ For $x \neq 0$, $N(x,.)$ is a continuous function of R and strictly increasing on the subset $\{t : 0 < N(x,t) < 1\}$ of R.

Definition 2. [3] Let (X, N) be a fuzzy normed linear space.
i) A sequence $\{x_n\} \subseteq X$ is said to converge to $x \in X$ ($\lim\limits_{n\to\infty} x_n = x$), if $\lim\limits_{n\to\infty} N(x_n - x, t) = 1$, for all $t > 0$.
ii) A sequence $\{x_n\} \subseteq X$ is called Cauchy, if $\lim\limits_{n,m\to\infty} N(x_n - x_m, t) = 1$, for all $t > 0$.

Definition 3. If X is a vector space over R, a seminorm is a function $p : X \longrightarrow [0, \infty)$ having the properties:
(i) $p(c(x)) = |c|p(x)$ for all $c \in R$ and $x \in X$.
(ii) $p(x+y) \leq p(x) + p(y)$ for all $x, y \in X$.

Theorem 1. *Let (X, N) be a fuzzy normed linear space. Define*

$$\|x\|_\alpha = \inf\{t > 0 : N(x,t) \geq \alpha\},\ \alpha \in (0,1).$$

Then $\{\|.\|_\alpha : \alpha \in (0,1)\}$ is an ascending family of seminorms on X and they are called α-seminorms on X corresponding to the fuzzy norm N on X.

Proof. (i) Let $x \in X$, $c \in R$ and $\alpha \in (0,1)$, we have

$$\begin{aligned} \|cx\|_\alpha &= \wedge\{t > 0 : N(cx,t) \geq \alpha\} \\ &= \wedge\{t > 0 : N(x, t/|c|) \geq \alpha\} \\ &= \wedge\{|c|t > 0 : N(x,t) \geq \alpha\} \\ &= |c|\|x\|_\alpha. \end{aligned}$$

(ii) Let $x, y \in X$ and $\alpha \in (0,1)$, we obtain that

$$N(x+y, \|x\|_\alpha + \|y\|_\alpha + \epsilon) \geq min\{N(x, \|x\|_\alpha + \epsilon/2), N(y, \|y\|_\alpha + \epsilon/2)\} \geq \alpha,$$

hence $\|x+y\|_\alpha \leq \|x\|_\alpha + \|y\|_\alpha + \epsilon$, as $\epsilon \longrightarrow 0$ then $\|x+y\|_\alpha \leq \|x\|_\alpha + \|y\|_\alpha$. $\square$

2 Fixed point theorems

At first we introduce the following notation.
Denote Ψ to be the set of functions $\psi : [0, +\infty) \longrightarrow [0, +\infty)$ satisfying the following hypotheses:
(i) ψ is continuous and nondecreasing,
(ii) $\psi(t) = 0$ if and only if $t = 0$.
We denote by Φ the set of functions $\phi : [0, +\infty) \longrightarrow [0, +\infty)$ satisfying the following hypotheses:
(i) ϕ is continuous and strictly increasing,
(ii) $\phi(t) = 0$ if and only if $t = 0$.

Theorem 2. *Let (X, N) be a fuzzy Banach space and $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$, $t \in \mathbb{R}$ and $\alpha \in (0, 1]$,*

$$N(x - y, t) \geq \alpha \text{ implies that } N(f(x) - f(y), t - \psi(t)) \geq \alpha,$$

where $\psi \in \Psi$. Then f has a unique fixed point in X.

Proof. Let $x_0 \in X$ and $x_{n+1} = f(x_n)$, for all $n \in \mathbb{N}$. Suppose that $t > 0$, we have

$$N(x - y, t) \leq N(f(x) - f(y), t - \psi(t)), \text{ for all } x, y \in X.$$

Therefore

$$N(x_{n+1} - x_n, t) \leq N(x_{n+2} - x_{n+1}, t - \psi(t)) \leq N(x_{n+2} - x_{n+1}, t),$$

for all $n \in \mathbb{N}$. Hence $\{N(x_{n+1} - x_n, t)\}$ is a bounded and nondecreasing sequence. Thus $\lim_{n\to\infty} N(x_{n+1} - x_n, t)$ exists. Now we have

$$N(x_1 - x_0, t + \psi(t)) \leq N(x_2 - x_1, t + \psi(t) - \psi(t + \psi(t))) \leq N(x_2 - x_1, t),$$

by induction on n, we obtain that

$$N(x_1 - x_0, t + n\psi(t)) \leq N(x_{n+1} - x_n, t), \text{ for all } n \in \mathbb{N}.$$

As $n \longrightarrow \infty$, (N5) implies $\lim_{n\to\infty} N(x_{n+1} - x_n, t) = 1$, for all $t > 0$.
Let $t > 0$, $\epsilon > 0$ and $N \in \mathbb{N}$ such that $1 - \epsilon \leq N(x_{N+1} - x_N, t/2)$ and $1 - \epsilon \leq N(x_{N+1} - x_N, \psi(t/2))$. If $1 - \epsilon \leq N(x - x_N, t/2)$ then

$$\begin{aligned} N(f(x) - x_N, t/2) &\geq \min\{N(f(x) - f(x_N), t/2 - \psi(t/2)), \\ &\qquad N(f(x_N) - x_N, \psi(t/2))\} \\ &\geq \min\{N(x - x_N, t/2), N(x_{N+1} - x_N, \psi(t/2))\} \\ &\geq 1 - \epsilon. \end{aligned}$$

Therefore $1-\epsilon \leq N(x_n - x_N, t/2)$, for all $n \geq N$, so

$$N(x_n - x_m, t) \geq \min\{N(x_n - x_N, t/2), N(x_m - x_N, t/2)\} \geq 1-\epsilon,$$

for all $n, m \geq N$. Since ϵ is arbitrary, $\{x_n\}$ is Cauchy, hence it is convergent. Assume that $\lim_{n\to\infty} x_n = x$. Let $\varepsilon > 0$ and $t > 0$. Then there exists $N_0 > 0$ such that $1-\varepsilon \leq N(x_n - x, t)$ and $1-\varepsilon \leq N(x - x_n, \psi(t))$, for all $n \geq N_0$. Hence

$$\begin{aligned} N(f(x) - x, t) &\geq \min\{N(f(x) - x_{n+1}, t - \psi(t)), N(x_{n+1} - x, \psi(t))\} \\ &\geq \min\{N(x - x_n, t), N(x_{n+1} - x, \psi(t))\} \\ &\geq 1-\epsilon, \text{ for all } n \geq N_0. \end{aligned}$$

Therefore $N(f(x) - x, t) = 1$, for all $t > 0$. Hence $f(x) = x$.
To prove the uniqueness of the fixed point, we let y be any other fixed point of f in X. Suppose that $t > 0$. Similarly, we have

$$N(x - y, t + n\psi(t)) \leq N(f(x) - f(y), t) = N(x - y, t), \text{ for all } n \in \mathbb{N}.$$

As $n \longrightarrow \infty$, we obtain that $N(x - y, t) = 1$, for all $t > 0$, hence $x = y$. □

Corollary 1. *Let (X, N) be a fuzzy Banach space and $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$ and $t \in \mathbb{R}$,*

$$N(f(x) - f(y), t - \psi(t)) \geq N(x - y, t),$$

where $\psi \in \Psi$. Then f has a unique fixed point in X.
Example 1. Let $(X, \|.\|)$ be a Banach space and $f : X \longrightarrow X$ be a function such that

$$\|f(x) - f(y)\| \leq \|x - y\| - \psi(\|x - y\|), \text{ for all } x, y \in X,$$

where $\psi \in \Psi$. Assume that $I - \psi$ is a nondecreasing function and $\psi(\beta t) \leq \beta\psi(t)$, for all $t \in [0, +\infty)$, $\beta \in [0, 1]$. Define a fuzzy norm N as follows:

$$N(x, t) = \begin{cases} t/\|x\| & , \quad 0 < t \leq \|x\| \\ 1 & , \quad \|x\| < t \\ 0 & , \quad t \leq 0. \end{cases}$$

Suppose that $x, y \in X$, $t > 0$, $\alpha \in (0, 1]$ and $N(x - y, t) \geq \alpha$.
Case1: Let $\|x - y\| < t$. Since $I - \psi$ is nondecreasing,

$$\|f(x) - f(y)\| \leq \|x - y\| - \psi(\|x - y\|) \leq t - \psi(t).$$

So $N(f(x) - f(y), t - \psi(t)) = 1 \geq \alpha$.
Case2: Let $0 < t \leq \|x - y\|$. So $t/\|x - y\| = N(x - y, t) \geq \alpha$. Hence $\alpha\|x - y\| \leq t$. Therefore

$$\alpha\|f(x)-f(y)\| \leq \alpha\|x-y\|-\alpha\psi(\|x-y\|) \leq \alpha\|x-y\|-\psi(\alpha\|x-y\|) \leq t-\psi(t).$$

Thus $N(f(x)-f(y), t-\psi(t)) = (t-\psi(t))/(\|f(x)-f(y)\|) \geq \alpha$. By Theorem 2, f has a unique fixed point in X.

Example 2. Let $[0,1] = X$ and $\|x-y\| = |x-y|$, for all $x, y \in X$. Also let $f : X \longrightarrow X$ and $\psi : [0,+\infty) \longrightarrow [0,+\infty)$ be defined as

$$f(x) = x - (1/2)x^2, \text{ for all } x \in X,$$

$$\psi(t) = (1/2)t^2, \text{ for all } t \geq 0.$$

It is clear that $I - \psi$ is a nondecreasing function and $\psi(\beta t) \leq \beta\psi(t)$, for all $t \in [0,+\infty)$, $\beta \in [0,1]$. Define a fuzzy norm N as follows:

$$N(x,t) = \begin{cases} t/\|x\| & , \quad 0 < t \leq \|x\| \\ 1 & , \quad \|x\| < t \\ 0 & , \quad t \leq 0. \end{cases}$$

Suppose that $x, y \in X$. Without loss of generality, we assume that $x > y$. Then

$$\begin{aligned} \|f(x) - f(y)\| &= (x - (1/2)x^2) - (y - (1/2)y^2) \\ &= (x-y) - (1/2)(x-y)(x+y) \\ &\leq (x-y) - (1/2)(x-y)^2 \\ &\leq \|x-y\| - \psi(\|x-y\|). \end{aligned}$$

By Example 1, f has a unique fixed point in X.

Theorem 3. *Let (X, N) be a fuzzy Banach space such that N satisfies (N7) and $\gamma : (0,+\infty) \longrightarrow [0,1)$ be a decreasing function, also $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$, $t > 0$ and $\alpha \in (0,1]$,*

$$N(x-y,t) \geq \alpha \text{ implies that } N(f(x)-f(y), \phi^{-1}(\gamma(t)\phi(t)) \geq \alpha,$$

where $\phi \in \Phi$. Then f has a unique fixed point in X.

Proof. Let $x_0 \in X$ and $x_{n+1} = f(x_n)$, for all $n \in \mathbb{N}$. Suppose that $t > 0$, we have

$$N(x-y,t) \leq N(f(x)-f(y), \phi^{-1}(\gamma(t)\phi(t)).$$

Therefore

$$N(x_{n+1}-x_n,t) \leq N(x_{n+2}-x_{n+1}, \phi^{-1}(\gamma(t)\phi(t))) \leq N(x_{n+2}-x_{n+1},t),$$

for all $n \in \mathbb{N}$. Hence $\{N(x_{n+1} - x_n, t)\}$ is a bounded and nondecreasing sequence. Thus $\lim_{n\to\infty} N(x_{n+1} - x_n, t)$ exists.

Assume that there exists $t > 0$ such that $\lim_{n\to\infty} N(x_{n+1}-x_n, t) < \beta < 1$. Since $N(x_{n+2} - x_{n+1}, s) \geq N(x_{n+1} - x_n, s)$, for all $s > 0$, it follows that

$$0 < t \leq \|x_{n+2} - x_{n+1}\|_\beta \leq \|x_{n+1} - x_n\|_\beta, \text{ for all } n \in \mathbb{N}.$$

Hence $\lim_{n\to\infty} \|x_{n+1} - x_n\|_\beta$ exists. Let $\lim_{n\to\infty} \|x_{n+1} - x_n\|_\beta = b \geq t > 0$. If $N(x_{n+1} - x_n, s) \geq \beta$ then

$$N(x_{n+2} - x_{n+1}, \phi^{-1}(\gamma(s)\phi(s))) \geq N(x_{n+1} - x_n, s) \geq \beta.$$

Therefore $\|x_{n+2} - x_{n+1}\|_\beta \leq \phi^{-1}(\gamma(s)\phi(s))$. Thus

$$\phi(\|x_{n+2} - x_{n+1}\|_\beta) \leq \gamma(s)\phi(s) \leq \gamma(\|x_{n+1} - x_n\|_\beta)\phi(s) \leq \gamma(b)\phi(s).$$

As $s \longrightarrow \|x_{n+1} - x_n\|_\beta$, we get $\phi(\|x_{n+2} - x_{n+1}\|_\beta) \leq \gamma(b)\phi(\|x_{n+1} - x_n\|_\beta)$. As $n \longrightarrow \infty$, one can obtain that $0 < \phi(t) \leq \phi(b) \leq \gamma(b)\phi(b)$. So $1 \leq \gamma(b)$, which is a contradiction. Hence $\lim_{n\to\infty} N(x_{n+1} - x_n, t) = 1$, for all $t > 0$. Let $t > 0$, $\epsilon > 0$ and $N \in \mathbb{N}$ such that

$$1 - \epsilon \leq N(x_{N+1} - x_N, t/2 - \phi^{-1}(\gamma(t/2)\phi(t/2)).$$

If $1 - \epsilon \leq N(x - x_N, t/2)$ then

$$\begin{aligned} N(f(x) - x_N, t/2) &\geq \min\{N(f(x) - x_{N+1}, \phi^{-1}(\gamma(t/2)\phi(t/2))), \\ &\quad N(x_{N+1} - x_N, t/2 - \phi^{-1}(\gamma(t/2)\phi(t/2)))\} \\ &\geq \min\{N(x - x_N, t/2), \\ &\quad N(x_{N+1} - x_N, t/2 - \phi^{-1}(\gamma(t/2)\phi(t/2)))\} \\ &\geq 1 - \epsilon \end{aligned}$$

Therefore $1 - \epsilon \leq N(x_n - x_N, t/2)$, for all $n \geq N$, so

$$N(x_n - x_m, t) \geq \min\{N(x_n - x_N, t/2), N(x_m - x_N, t/2)\} \geq 1 - \epsilon,$$

for all $n, m \geq N$. Since ϵ is arbitrary, $\{x_n\}$ is Cauchy, hence it is convergent. Assume that $\lim_{n\to\infty} x_n = x$. Let $\varepsilon > 0$ and $t > 0$, then there exists $N_0 > 0$ such that $1 - \varepsilon \leq N(x_n - x, t - \phi^{-1}(\gamma(t)\phi(t)))$, for all $n \geq N_0$. Hence

$$\begin{aligned} N(f(x) - x, t) &\geq \min\{N(f(x) - x_{N+1}, \phi^{-1}(\gamma(t)\phi(t))), \\ &\quad N(x_{N+1} - x, t - \phi^{-1}(\gamma(t)\phi(t)))\} \\ &\geq \min\{N(x - x_N, t), N(x_{N+1} - x, t - \phi^{-1}(\gamma(t)\phi(t)))\} \\ &\geq 1 - \epsilon, \text{ for all } n \geq N_0. \end{aligned}$$

Therefore $N(f(x) - x, t) = 1$, for all $t > 0$. Hence $f(x) = x$.
To prove the uniqueness of the fixed point, we let y be any other fixed point

of f in X. If there exists $t > 0$ such that $0 < N(x-y,t) < 1$, then

$$\begin{aligned} N(x-y,t) &\leq N(f(x)-f(y), \phi^{-1}(\gamma(t)\phi(t))) \\ &\leq N(f(x)-f(y),t) \\ &= N(x-y,t), \end{aligned}$$

therefore $N(x-y, \phi^{-1}(\gamma(t)\phi(t))) = N(x-y,t)$. By $(N7)$, $t = \phi^{-1}(\gamma(t)\phi(t))$, we get $\phi(t) = \gamma(t)\phi(t)$. Hence $\gamma(t) = 1$, which is a contradiction. Thus $N(x-y,t) = 1$, for all $t > 0$, so $x = y$. □

Corollary 2. *Let (X,N) be a fuzzy Banach space such that N satisfying $(N7)$ and $\gamma : (0,+\infty) \longrightarrow [0,1)$ be a decreasing function, also $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$ and $t > 0$,*

$$N(f(x)-f(y), \phi^{-1}(\gamma(t)\phi(t)) \geq N(x-y,t),$$

where $\phi \in \Phi$. Then f has a unique fixed point in X.

Example 3. Let $(X, \|.\|)$ be a Banach space and $\gamma : (0,+\infty) \longrightarrow [0,1)$ be a decreasing function and $f : X \longrightarrow X$ be a function such that

$$\phi(\|f(x)-f(y)\|) \leq \gamma(\|x-y\|)\phi(\|x-y\|), \text{ for all } x, y \in X,$$

where $\phi \in \Phi$.
Assume that $\gamma\phi$ is a nondecreasing function and

$$\beta(\phi^{-1}(\gamma(t)\phi(t))) \leq \phi^{-1}(\gamma(\beta t)\phi(\beta t)), \text{ for all } t \in [0,+\infty),\ \beta \in [0,1].$$

Define a fuzzy norm N as follows:

$$N(x,t) = \begin{cases} t/\|x\| & , \quad 0 < t \leq \|x\| \\ 1 & , \quad \|x\| < t \\ 0 & , \quad t \leq 0. \end{cases}$$

Suppose that $x, y \in X$, $t > 0$, $\alpha \in (0,1]$ and $N(x-y,t) \geq \alpha$.
Case1: Let $\|x-y\| < t$, since $\gamma\phi$ is nondecreasing,

$$\phi(\|f(x)-f(y)\|) \leq \gamma(\|x-y\|)\phi(\|x-y\|) \leq \gamma(t)\phi(t).$$

So $\|f(x)-f(y)\| \leq \phi^{-1}(\gamma(t)\phi(t))$. Hence

$$N(f(x)-f(y), \phi^{-1}(\gamma(t)\phi(t))) = 1 \geq \alpha.$$

Case2: Let $0 < t \leq \|x-y\|$. So $t/\|x-y\| = N(x-y,t) \geq \alpha$. This implies that $\alpha\|x-y\| \leq t$. Therefore

$$\begin{aligned}\alpha\|f(x)-f(y)\| &\leq \alpha(\phi^{-1}(\gamma(\|x-y\|)\psi(\|x-y\|)))\\ &\leq \phi^{-1}(\gamma(\alpha\|x-y\|)\phi(\alpha\|x-y\|))\\ &\leq \phi^{-1}(\gamma(t)\phi(t)).\end{aligned}$$

Thus $N(f(x)-f(y),\phi^{-1}(\gamma(t)\phi(t))) = (\phi^{-1}(\gamma(t)\phi(t)))/(\|f(x)-f(y)\|) \geq \alpha$. By Theorem 3, f has a unique fixed point in X.

Example 4. Let $[0,1] = X$ and $\|x-y\| = |x-y|$, for all $x, y \in X$. Also let $f : X \longrightarrow X$, $\gamma : (0,+\infty) \longrightarrow [0,1)$ and $\phi : [0,+\infty) \longrightarrow [0,+\infty)$ be defined as

$$f(x) = x^3, \text{ for all } x \in X,$$

$$\phi(t) = (1/2)t^2, \text{ for all } t \geq 0,$$

$$\gamma(t) = 1/t \text{ for all } t > 0.$$

It is clear that $\gamma\phi$ is a nondecreasing function and $\beta(\phi^{-1}(\gamma(t)\phi(t))) \leq \phi^{-1}(\gamma(\beta t)\phi(\beta t))$, for all $t \in (0,+\infty)$, $\beta \in [0,1]$. Define a fuzzy norm N as follows:

$$N(x,t) = \begin{cases} t/\|x\| & , \quad 0 < t \leq \|x\| \\ 1 & , \quad \|x\| < t \\ 0 & , \quad t \leq 0. \end{cases}$$

Suppose that $x, y \in X$. Without loss of generality, we assume that $x > y$. Then

$$\begin{aligned}\|f(x)-f(y)\| &= x^3 - y^3\\ &= (x-y)(x^2+xy+y^2)\\ &\leq (x-y)\\ &\leq \sqrt{x-y}\\ &= \phi^{-1}(\gamma(\|x-y\|)\phi(\|x-y\|)).\end{aligned}$$

By Example 3, f has a unique fixed point in X.

Theorem 4. *Let (X,N) be fuzzy Banach space such that N satisfying (N7) and $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$, $t > 0$ and $\alpha \in (0,1]$,*

$$N(x-y,t) \geq \alpha \text{ implies that } N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t))) \geq \alpha,$$

where $\phi, \varphi \in \Phi$ and $\varphi(t) \geq \phi(t)$, for all $t > 0$. Then f has a unique fixed point in X.

Proof. Let $x_0 \in X$ and $x_{n+1} = f(x_n)$, for all $n \in \mathbb{N}$. Suppose that $t > 0$, we have

$$N(x-y,t) \leq N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t))).$$

Therefore

$$N(x_{n+1}-x_n,t)\leq N(x_{n+2}-x_{n+1},\varphi^{-1}(\varphi(t)-\phi(t)))\leq N(x_{n+2}-x_{n+1},t),$$

for all $n\in\mathbb{N}$. Hence $\{N(x_{n+1}-x_n,t)\}$ is a bounded and nondecreasing sequence, and $\lim_{n\to\infty}N(x_{n+1}-x_n,t)$ exists.
Assume that there exists $t>0$ such that $\lim_{n\to\infty}N(x_{n+1}-x_n,t)<\beta<1$. Since $N(x_{n+2}-x_{n+1},s)\geq N(x_{n+1}-x_n,s)$, for all $s>0$, then

$$0<t\leq\|x_{n+2}-x_{n+1}\|_\beta\leq\|x_{n+1}-x_n\|_\beta,\ \text{for all } n\in\mathbb{N}.$$

Therefore $\lim_{n\to\infty}\|x_{n+1}-x_n\|_\beta$ exists. Let $\lim_{n\to\infty}\|x_{n+1}-x_n\|_\beta=b\geq t>0$. If $N(x_{n+1}-x_n,s)\geq\beta$ then

$$N(x_{n+2}-x_{n+1},\varphi^{-1}(\varphi(s)-\phi(s)))\geq N(x_{n+1}-x_n,s)\geq\beta,$$

and $\|x_{n+2}-x_{n+1}\|_\beta\leq\varphi^{-1}(\varphi(s)-\phi(s))$. Thus

$$\varphi(\|x_{n+2}-x_{n+1}\|_\beta)\leq\varphi(s)-\phi(s).$$

As $s\longrightarrow\|x_{n+1}-x_n\|_\beta$, one can get

$$\varphi(\|x_{n+2}-x_{n+1}\|_\beta)\leq\varphi(\|x_{n+1}-x_n\|_\beta)-\phi(\|x_{n+1}-x_n\|_\beta).$$

As $n\longrightarrow\infty$, we obtain that $0<\varphi(t)\leq\varphi(b)\leq\varphi(b)-\phi(b)<\varphi(b)$, which is a contradiction. Hence $\lim_{n\to\infty}N(x_{n+1}-x_n,t)=1$, for all $t>0$.
Let $t>0$, $\epsilon>0$ and $N\in\mathbb{N}$ such that

$$1-\epsilon\leq N(x_{N+1}-x_N,t/2-\varphi^{-1}(\varphi(t/2)-\phi(t/2))).$$

If $1-\epsilon\leq N(x-x_N,t/2)$, then

$$\begin{aligned}N(f(x)-x_N,t/2)&\geq\min\{N(f(x)-x_{N+1},\varphi^{-1}(\varphi(t/2)-\phi(t/2))),\\&\quad N(x_{N+1}-x_N,t/2-\varphi^{-1}(\varphi(t/2)-\phi(t/2)))\}\\&\geq\min\{N(x-x_N,t/2),\\&\quad N(x_{N+1}-x_N,t/2-\varphi^{-1}(\varphi(t/2)-\phi(t/2)))\}\\&\geq 1-\epsilon.\end{aligned}$$

Therefore $1-\epsilon\leq N(x_n-x_N,t/2)$, for all $n\geq N$, so

$$N(x_n-x_m,t)\geq\min\{N(x_n-x_N,t/2),N(x_m-x_N,t/2)\}\geq 1-\epsilon,$$

for all $n,m\geq N$. Since ϵ is arbitrary, $\{x_n\}$ is Cauchy, hence it is convergent. Assume that $\lim_{n\to\infty}x_n=x$. Let $\varepsilon>0$ and $t>0$, then there exists $N_0>0$ such that $1-\varepsilon\leq N(x_n-x,t-\varphi^{-1}(\varphi(t)-\phi(t)))$, for all $n\geq N_0$. Hence

$$\begin{aligned} N(f(x)-x,t) &\geq \min\{N(f(x)-x_{N+1},\varphi^{-1}(\varphi(t)-\phi(t))),\\ &\quad N(x_{N+1}-x,t-\varphi^{-1}(\varphi(t)-\phi(t)))\}\\ &\geq \min\{N(x-x_N,t),N(x_{N+1}-x,t-\phi^{-1}(\gamma(t)\phi(t)))\}\\ &\geq 1-\epsilon, \text{ for all } n\geq N_0. \end{aligned}$$

Therefore $N(f(x)-x,t)=1$, for all $t>0$, so $f(x)=x$.
To prove the uniqueness of the fixed point, we let y be any other fixed point of f in X. If there exists $t>0$ such that $0<N(x-y,t)<1$ then

$$\begin{aligned} N(x-y,t) &\leq N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t)))\\ &\leq N(f(x)-f(y),t)\\ &= N(x-y,t), \end{aligned}$$

therefore $N(x-y,\phi^{-1}(\gamma(t)\phi(t)))=N(x-y,t)$. By $(N7)$, we obtain that $t=\varphi^{-1}(\varphi(t)-\phi(t))$, then $\varphi(t)=\varphi(t)-\phi(t)$. Hence $\phi(t)=0$, and $t=0$, which is a contradiction. Thus $N(x-y,t)=1$, for all $t>0$, so $x=y$. □

Corollary 3. *Let (X,N) be fuzzy Banach space such that N satisfying $(N7)$ and $f:X\longrightarrow X$ be a selfmap such that for all $x,y\in X$ and $t>0$,*

$$N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t)))\geq N(x-y,t),$$

where $\phi,\varphi\in\Phi$ and $\varphi(t)\geq\phi(t)$, for all $t>0$. Then f has a unique fixed point in X.

Example 5. Let $(X,\|.\|)$ be a Banach space and $f:X\longrightarrow X$ be a function such that

$$\varphi(\|f(x)-f(y)\|)\leq\varphi(\|x-y\|)-\phi(\|x-y\|), \text{ for all } x,y\in X,$$

where $\varphi,\phi\in\Phi$. Assume that $\varphi-\phi$ is a nondecreasing function and

$$\beta(\varphi^{-1}(\varphi(t)-\phi(t)))\leq\varphi^{-1}(\varphi(\beta t)-\phi(\beta t)), \text{ for all } t\in[0,+\infty),\ \beta\in[0,1].$$

Define a fuzzy norm N as follows:

$$N(x,t)=\begin{cases} t/\|x\| & , \quad 0<t\leq\|x\| \\ 1 & , \quad \|x\|<t \\ 0 & , \quad t\leq 0. \end{cases}$$

Suppose that $x,y\in X$, $t>0$, $\alpha\in(0,1]$ and $N(x-y,t)\geq\alpha$.
Case1: Let $\|x-y\|<t$. Since $\varphi-\phi$ is nondecreasing,

$$\varphi(\|f(x)-f(y)\|)\leq\varphi(\|x-y\|)-\phi(\|x-y\|)\leq\varphi(t)-\phi(t).$$

So $\|f(x)-f(y)\|\leq\varphi^{-1}(\varphi(t)-\phi(t))$. Hence

$$N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t)))=1\geq\alpha.$$

Case2: Let $0<t\leq\|x-y\|$. So $t/\|x-y\|=N(x-y,t)\geq\alpha$. Hence $\alpha\|x-y\|\leq t$, and

$$\begin{aligned}\alpha\|f(x)-f(y)\|&\leq\alpha(\varphi^{-1}(\varphi(\|x-y\|)-\psi(\|x-y\|)))\\&\leq\varphi^{-1}(\varphi(\alpha\|x-y\|)-\phi(\alpha\|x-y\|))\\&\leq\varphi^{-1}(\varphi(t)-\phi(t))\end{aligned}$$

Thus

$$N(f(x)-f(y),\varphi^{-1}(\varphi(t)-\phi(t)))=(\varphi^{-1}(\varphi(t)-\phi(t)))/(\|f(x)-f(y)\|)\geq\alpha$$

By Theorem 4, f has a unique fixed point in X.

Theorem 5. *Let (X,N) be fuzzy Banach space such that N satisfying (N7) and $f:X\longrightarrow X$ be a selfmap such that for all $x,y\in X$, $s,t>0$ and $\alpha\in(0,1]$,*

$$N(x-f(y),t)\geq\alpha \text{ and } N(f(x)-y,s)\geq\alpha \text{ implies that}$$

$$N(f(x)-f(y),1/2(t+s)-\theta(t,s))\geq\alpha,$$

where $\theta:[0,+\infty)^2\longrightarrow[0,+\infty)$ is a continuous mapping such that $\theta(x,y)=0$ if and only if $x=y=0$. Then f has a unique fixed point in X.

Proof. Let $x_0\in X$ and $x_{n+1}=f(x_n)$, for all $n\in\mathbb{N}$. Suppose that $t>0$, so

$$N(f(x)-f(y),1/2(t+s)-\theta(t,s))\geq\min\{N(x-f(y),t),N(f(x)-y,s)\},$$

for all $s>0$. Therefore

$$\begin{aligned}N(x_{n+1}-x_n,1/2(2t+s)-\theta(2t,s))&\geq\min\{N(x_n-x_n,s),\\&\quad N(x_{n+1}-x_{n-1},2t)\}\\&=N(x_{n+1}-x_{n-1},2t),\end{aligned}$$

for all $s>0$. As $s\longrightarrow0$, we obtain that

$$\begin{aligned}N(x_{n+1}-x_n,t-\theta(2t,0))&\geq N(x_{n+1}-x_{n-1},2t)\\&\geq\min\{N(x_{n+1}-x_n,t),N(x_n-x_{n-1},t)\},\end{aligned}$$

for all $n\in\mathbb{N}$ and all $t>0$. Let there exists $t_0>0$ and $n_0\in\mathbb{N}$ such that $N(x_{n_0+1}-x_{n_0},t_0)<N(x_{n_0}-x_{n_0-1},t_0)$. By $(N7)$ and $(N5)$, there is $t_1>0$ such that

$$0\leq N(x_{n_0+1}-x_{n_0},t_0)<N(x_{n_0+1}-x_{n_0},t_0+t_1)<N(x_{n_0}-x_{n_0-1},t_0)\leq1$$

Hence

$$\begin{aligned} N(x_{n+1}-x_n,(t_0+t_1)) &\geq N(x_{n+1}-x_n,(t_0+t_1)-\theta(2(t_0+t_1),0)) \\ &\geq N(x_{n+1}-x_n,(t_0+t_1)). \end{aligned}$$

Thus $N(x_{n+1}-x_n,(t_0+t_1)) = N(x_{n+1}-x_n,(t_0+t_1)-\theta(2(t_0+t_1),0))$. By $(N7)$, $(t_0+t_1)-\theta(2(t_0+t_1),0) = t_0+t_1$. So $\theta(2(t_0+t_1),0)=0$ which is a contradiction. In addition $N(x_{n+1}-x_n,t) \geq N(x_n-x_{n-1},t)$, for all $t>0$ and all $n\in\mathbb{N}$. Therefore $\lim_{n\to\infty} N(x_{n+1}-x_n,t)$ exists.
Assume that there exists $t>0$ such that $\lim_{n\to\infty} N(x_{n+1}-x_n,t)<\beta<1$. Since $N(x_{n+2}-x_{n+1},s) \geq N(x_{n+1}-x_n,s)$, for all $s>0$, it follows that

$$0<t\leq \|x_{n+2}-x_{n+1}\|_\beta \leq \|x_{n+1}-x_n\|_\beta, \text{ for all } n\in\mathbb{N}.$$

Thus $\lim_{n\to\infty}\|x_{n+1}-x_n\|_\beta$ exists. Let $\lim_{n\to\infty}\|x_{n+1}-x_n\|_\beta = b \geq t>0$. If $N(x_{n+1}-x_n,s)\geq\beta$, then $N(x_{n+2}-x_{n+1},s-\theta(2s,0)) \geq N(x_{n+1}-x_n,s)\geq\beta$. So $\|x_{n+2}-x_{n+1}\|_\beta \leq s-\theta(2s,0)$. As $s\longrightarrow \|x_{n+1}-x_n\|_\beta$, we obtain that $\|x_{n+2}-x_{n+1}\|_\beta \leq \|x_{n+1}-x_n\|_\beta - \theta(2\|x_{n+1}-x_n\|_\beta,0)$. As $n\longrightarrow\infty$, we get $0<t\leq b\leq b-\theta(2b,0)\leq b$ and $\theta(2b,0)=0$, which is a contradiction. Hence $\lim_{n\to\infty} N(x_{n+1}-x_n,t)=1$, for all $t>0$.
Next we show that $\{x_n\}$ is a cauchy sequence. If otherwise, then there exist $t_0>0$, $\beta\in(0,1)$ and increasing sequences of integers $\{m_k\}$ and $\{n_k\}$ such that $N(x_{n_k}-x_{m_k},t_0)<\beta$ and $N(x_{n_k-1}-x_{m_k},t_0)\geq\beta$, for all $k\in\mathbb{N}$. Since $\lim_{n\to\infty} N(x_{n+1}-x_n,t)=1$, for all $t>0$, it follows that

$$\lim_{n\to\infty}\|x_{n_k}-x_{n_k-1}\|_\beta = 0 = \lim_{n\to\infty}\|x_{m_k}-x_{m_k-1}\|_\beta.$$

Moreover

$$\begin{aligned} 0<t_0 &\leq \|x_{n_k}-x_{m_k}\|_\beta \\ &\leq \|x_{n_k-1}-x_{n_k}\|_\beta + \|x_{m_k}-x_{n_k-1}\|_\beta \\ &\leq t_0 + \|x_{n_k}-x_{n_k-1}\|_\beta, \end{aligned}$$

for all $k\in\mathbb{N}$. As $k\longrightarrow\infty$, which leads to

$$\lim_{n\to\infty}\|x_{n_k}-x_{m_k}\|_\beta = t_0 = \lim_{n\to\infty}\|x_{m_k}-x_{n_k-1}\|_\beta.$$

Now we have

$$\|x_{m_k}-x_{n_k}\|_\beta \leq \|x_{m_k-1}-x_{m_k}\|_\beta + \|x_{m_k-1}-x_{n_k}\|_\beta + \|x_{n_k}-x_{n_k-1}\|_\beta$$

and

$$\|x_{m_k-1}-x_{n_k}\|_\beta \leq \|x_{m_k-1}-x_{m_k}\|_\beta + \|x_{m_k}-x_{n_k}\|_\beta,$$

for all $k \in \mathbb{N}$. As $k \longrightarrow \infty$, we obtain that $\lim_{n\to\infty} \|x_{n_k} - x_{m_k-1}\|_\beta = t_0$. If $N(x_{n_k-1} - x_{m_k}, t) \geq \beta$ and $N(x_{n_k} - x_{m_k-1}, s) \geq \beta$ then

$$N(x_{n_k} - x_{m_k}, 1/2(t+s) - \theta(t,s)) \geq \beta.$$

Hence

$$\|x_{m_k} - x_{n_k}\|_\beta \leq 1/2(t+s) - \theta(t,s).$$

As $t \longrightarrow \|x_{m_k} - x_{n_k-1}\|_\beta$ and $s \longrightarrow \|x_{m_k-1} - x_{n_k}\|_\beta$, we get

$$\|x_{m_k} - x_{n_k}\|_\beta \leq$$

$$1/2(\|x_{m_k} - x_{n_k-1}\|_\beta + \|x_{m_k-1} - x_{n_k}\|_\beta) - \theta(\|x_{m_k} - x_{n_k-1}\|_\beta, \|x_{m_k-1} - x_{n_k}\|_\beta),$$

for all $k \in \mathbb{N}$. As $k \longrightarrow \infty$, we have $t_0 \leq 1/2(t_0 + t_0) - \theta(t_0, t_0)$. Therefore $\theta(t_0, t_0) = 0$ which is a contradiction. Thus $\{x_n\}$ is Cauchy, hence convergent. Assume that $\lim_{n\to\infty} x_n = x$.

Next we show that $\lim_{n\to\infty} N(f(x) - x_n, t) = 1$, for all $t > 0$. If otherwise, then there exist $t_0 > 0$, $\beta \in (0,1)$ and increasing sequences of integers $\{n_k\}$ such that $N(f(x) - x_{n_k}, t_0) < \beta$, for all $k \in \mathbb{N}$. Since $\lim_{n\to\infty} N(x - x_n, t) = 1$, for all $t > 0$, it follows that $\lim_{n\to\infty} \|x_n - x\|_\beta = 0$. We have

$$\|f(x) - x_n\|_\beta \leq \|x - x_n\|_\beta + \|f(x) - x\|_\beta$$

and

$$\|f(x) - x\|_\beta \leq \|f(x) - x_n\|_\beta + \|x_n - x\|_\beta,$$

for all $n \in \mathbb{N}$. As $n \longrightarrow \infty$, we obtain that $\lim_{n\to\infty} \|f(x) - x_n\|_\beta = \|f(x) - x\|_\beta$. If $N(x_{n_k-1} - f(x), t) \geq \beta$ and $N(x_{n_k} - x, s) \geq \beta$, then

$$\begin{aligned} N(f(x) - x_{n_k}, 1/2(t+s) - \theta(t,s)) &\geq \min\{N(x - x_{n_k}, t), N(f(x) - x_{n_k-1}, s)\} \\ &\geq \beta. \end{aligned}$$

Therefore $\|f(x) - x_{n_k}\|_\beta \leq 1/2(t+s) - \theta(t,s)$. As $t \longrightarrow \|f(x) - x_{n_k-1}\|_\beta$ and $s \longrightarrow \|x - x_{n_k}\|_\beta$, we get

$$0 < t_0 \leq \|f(x) - x_{n_k}\|_\beta \leq$$

$$1/2(\|f(x) - x_{n_k-1}\|_\beta + \|x - x_{n_k}\|_\beta) - \theta(\|f(x) - x_{n_k-1}\|_\beta, \|x - x_{n_k}\|_\beta),$$

for all $k \in \mathbb{N}$. As $k \longrightarrow \infty$, one can obtain

$$0 < t_0 \leq \|f(x) - x\|_\beta \leq 1/2\|f(x) - x\|_\beta - \theta(\|f(x) - x\|_\beta, 0).$$

So $\theta(\|f(x) - x\|_\beta, 0) = 0$, which is a contradiction. This implies that $\lim_{n\to\infty} N(f(x) - x_n, t) = 1$, for all $t > 0$. By $(N4)$, we have

$$N(f(x) - x, t) \geq \min\{N(f(x) - x_n, t/2), N(x - x_n, t/2)\}, \text{ for all } t > 0.$$

As $n \longrightarrow \infty$, we obtain that $N(f(x)-x, t) = 1$, for all $t > 0$. Hence $f(x) = x$. To prove the uniqueness of the fixed point, let y be any other fixed point of f in X. If there exists $t > 0$ such that $0 < N(x - y, t) < 1$, then

$$\begin{aligned} N(x - y, t) &= N(f(x) - f(y), t) \\ &\geq N(f(x) - f(y), t - \theta(t, t)) \\ &\geq \min\{N(x - f(y), t), N(f(x) - y, t)\} \\ &= \min\{N(x - y, t), N(x - y, t)\} \\ &= N(x - y, t). \end{aligned}$$

Therefore $N(x - y, t - \theta(t, t)) = N(x - y, t)$. By $(N7)$, $\theta(t, t) = 0$ which is a contradiction. Thus $N(x - y, t) = 1$, for all $t > 0$. So $x = y$. □

Corollary 4. *Let (X, N) be fuzzy Banach space such that N satisfying $(N7)$. Also $f : X \longrightarrow X$ be a selfmap such that for all $x, y \in X$ and $s, t > 0$,*

$$N(f(x) - f(y), 1/2(t + s) - \theta(t, s)) \geq \min\{N(x - f(y), t), N(f(x) - y, s)\},$$

where $\theta : [0, +\infty)^2 \longrightarrow [0, +\infty)$ is a continuous mapping such that $\theta(x, y) = 0$ if and only if $x = y = 0$. Then f has a unique fixed point in X.

Example 6. Let $(X, \|.\|)$ be a Banach space and $f : X \longrightarrow X$ be a function such that

$$\|f(x) - f(y)\| \leq 1/2(\|x - f(y)\| + \|f(x) - y\|) - \theta(\|x - f(y)\|, \|f(x) - y\|),$$

for all $x, y \in X$, where $\theta : [0, +\infty)^2 \longrightarrow [0, +\infty)$ is a continuous mapping such that $\theta(x, y) = 0$ if and only if $x = y = 0$.
If $s_1 \leq s_2$ and $t_1 \leq t_2$ then $1/2(t_1 + s_1) - \theta(t_1, s_1) \leq 1/2(t_2 + s_2) - \theta(t_2, s_2)$, for all $s_1, t_1, s_2, t_2 > 0$ and $\theta(\beta t, \beta s) \leq \beta\theta(t, s)$, for all $t, s > 0$, $\beta \in [0, 1]$. Define a fuzzy norm N as follows:

$$N(x, t) = \begin{cases} t/\|x\| & , \quad 0 < t \leq \|x\| \\ 1 & , \quad \|x\| < t \\ 0 & , \quad t \leq 0. \end{cases}$$

Suppose that $x, y \in X$, $t > 0$, $\alpha \in (0, 1]$, $N(x - f(y), t) \geq \alpha$ and $N(f(x) - y, s) \geq \alpha$.
Case1: Let $\|x - f(y)\| < t$ and $\|f(x) - y\| < s$. Then

$$\begin{aligned}\|f(x)-f(y)\| &\leq 1/2(\|x-f(y)\|+\|f(x)-y\|)-\theta(\|x-f(y)\|,\|f(x)-y\|)\\ &\leq 1/2(t+s)-\theta(t,s).\end{aligned}$$

Thus $N(f(x)-f(y),1/2(t+s)-\theta(t,s))=1\geq\alpha$.
Case2: Let $0<t\leq\|x-f(y)\|$ and $\|f(x)-y\|<s$. So $t/\|x-f(y)\|=N(x-y,t)\geq\alpha$ and $\|x-f(y)\|<s$. Hence $\alpha\|x-f(y)\|\leq t$ and $\|f(x)-y\|<s$. Therefore

$$\begin{aligned}\alpha\|f(x)-f(y)\| &\leq \alpha(1/2(\|x-f(y)\|+\|f(x)-y\|)-\theta(\|x-f(y)\|,\\ &\quad \|f(x)-y\|))\\ &\leq 1/2(\alpha\|x-f(y)\|+\alpha\|f(x)-y\|)-\theta(\alpha\|x-f(y)\|,\\ &\quad \alpha\|f(x)-y\|)\\ &\leq 1/2(t+s)-\theta(t,s).\end{aligned}$$

Thus

$$N(f(x)-f(y),1/2(t+s)-\theta(t,s))=(1/2(t+s)-\theta(t,s))/(\|f(x)-f(y)\|)\geq\alpha.$$

Case3: Let $0<t\leq\|x-f(y)\|$ and $0<s\leq\|f(x)-y\|$. So $t/\|x-y\|=N(x-y,t)\geq\alpha$ and $s/\|f(x)-y\|=N(f(x)-y,s)\geq\alpha$. Hence $\alpha\|x-f(y)\|\leq t$ and $\alpha\|f(x)-y\|<s$. Therefore

$$\begin{aligned}\alpha\|f(x)-f(y)\| &\leq \alpha(1/2(\|x-f(y)\|+\|f(x)-y\|)-\theta(\|x-f(y)\|,\\ &\quad \|f(x)-y\|))\\ &\leq 1/2(\alpha\|x-f(y)\|+\alpha\|f(x)-y\|)-\theta(\alpha\|x-f(y)\|,\\ &\quad \alpha\|f(x)-y\|)\\ &\leq 1/2(t+s)-\theta(t,s).\end{aligned}$$

Thus

$$N(f(x)-f(y),1/2(t+s)-\theta(t,s))=(1/2(t+s)-\theta(t,s))/(\|f(x)-f(y)\|)\geq\alpha.$$

Case4: Let $\|x-f(y)\|<t$ and $0<s\leq\|f(x)-y\|$. Similar to case2, we obtain that $N(f(x)-f(y),1/2(t+s)-\theta(t,s))\geq\alpha$.
By Theorem 5, f has a unique fixed point in X.

3 Conclusion

We have introduced four contractive conditions in fuzzy normed linear spaces and proved some results about fixed point theorem. In fact, the established properties are the extended fuzzy forms of some classical contractive properties. To reveal this fact, some examples have been studied.

Acknowledgements

Authors are grateful to their anonymous referees and the editor for their constructive comments.

References

1. AL-Mayahi, N.F. and Hadi, S.H. *On α-η-φ contraction in fuzzy metric space and its application*, General Mathematics Notes 26 (2) (2015), 104-118.

2. Bag, T. and Samanta, S.K. *Finite dimensional fuzzy normed linear spaces*, J. Fuzzy Math. 11 (3) (2003), 687-705.

3. Bag, T. and Samanta, S.K. *Fuzzy bounded linear operators*, Fuzzy Sets and Systems 151 (2005), 513-547.

4. Choudhury, B.S. *Unique fixed point theorem for weakly c-contractive mappings*, Kathmandu University Journal of Science, Engineering and Technology 5 (2009), 6-13.

5. Das, N.R. and Saha, M.L. *On fixed points in fuzzy metric spaces*, Annals of Fuzzy Mathematics and Informatics 7 (2) (2014), 313-318.

6. Kadelburg, Z. and Radenovic, S. *On generalized metric spaces: a survey*, TWMS J. Pure Appl. Math. 5 (1) (2014), 3-13.

7. Khan, M.S., Swaleh, M. and Sessa, S. *Fixed point theorems by altering distances between the points*, Bulletin of the Australian Mathematical Society 30 (1) (1984), 1-9.

8. Manro, S. and Tomar, A. *Faintly compatible maps and existence of common fixed points in fuzzy metric space*, Annals of Fuzzy Mathematics and Informatics 8 (2) (2014), 223-230.

9. Rhoades, B.E. *Some theorems on weakly contractive maps*, Nonlinear Analysis: Theory, Methods and Applications 47 (4) (2001), 2683-2693.

10. Rosa, V.L. and Vetro, P. *Common fixed points for α-ψ-φ-contractions in generalized metric spaces*, Nonlinear Analysis: Modelling and Control 19 (1) (2014), 43-54.

11. Shukla, S. and Chauhan, S. *Fuzzy cyclic contraction and fixed point theorems*, Journal of the Egyptian Mathematical Society 23 (1) (2015), 139-143.

Persian Translation of Abstracts

یک روش خطی برای طراحی بهینه چشمه میرایی دارای اتلاف داخلی

علیرضا فخارزاده جهرمی[1] ، *هاجر علیمراد دست خضر*[2,1] *و علی بیرانوند*[1]

[1] دانشگاه صنعتی شیراز، دانشکده ریاضی
[2] دانشگاه جهرم، گروه ریاضی

دریافت مقاله ۲۲ بهمن ۱۳۹۳، دریافت مقاله اصلاح شده ۲۳ فروردین ۱۳۹۴، پذیرش مقاله ۲۷ خرداد ۱۳۹۴

چکیده : یک سیستم موج میرای دو بعدی را در نظر بگیرید، که دارای اتلاف انرژی درزیرمجموعه‌ای مجهول از دامنه موج با پارامتر میرایی مجهول می‌باشد. هدف حل مساله بد وضع طراحی شکل، شامل بهینه‌سازی شکل این زیرمجموعه جهت کمینه کردن انرژی سیستم در یک زمان معین می‌باشد. با استفاده از یک الگوریتم جدید براساس روش نشاندن، برای اولین بار، معادلات سیستم را در قالب تغییراتی نوشته؛ پس از آن، با انتقال مساله به مختصات قطبی و تعریف دو اندازه رادون مثبت، مساله را در فضای اندازه‌ها نمایش می‌دهیم. در این روش، مساله طراحی شکل بهینه به یک مساله برنامه‌ریزی خطی نامتناهی تبدیل می‌شود که وجود جواب آن تضمین شده است. دراین مرحله، با استفاده از دو گام تقریب، جواب بهینه (کنترل بهینه، ناحیه بهینه، پارامتر میرایی بهینه و انرژی بهینه) با یک روش جستجوی سه مرحله‌ای مشخص می‌گردد. به منظور مقایسه این روش جدید با دیگر روش ها شبیه سازی عددی نیزآورده شده است. .

کلمات کلیدی : معادله موج میرا؛ کنترل اتلاف؛ اندازه رادون؛ روش جستجو؛ بهینه‌سازی شکل.

روش گالرکین چبیشف برای معادلات انتگرال-دیفرانسیل ازمرتبه دوم

جعفر بی‌آزار [۱] *و فریده صالحی* [۲]

[۱] گروه ریاضی،دانشکده علوم ریاضی، دانشگاه گیلان، رشت، ایران
[۲] گروه ریاضی، واحد داراب، دانشگاه آزاد اسلامی، داراب، ایران

دریافت مقاله ۲۹ تیر ۱۳۹۳، دریافت مقاله اصلاح شده ۲۳ مهر ۱۳۹۳، پذیرش مقاله ۱۰ تیر ۱۳۹۴

چکیده : در این مقاله، یک اجرای موثر از روش چبیشف گالرکین، برای معادلات انتگرال-دیفرانسیل مرتبه اول ولترا و فردهلم نوع دوم پیشنهاد می شود. چندین مثال عددی برای نشان دادن دقت روش ارائه شده است.

کلمات کلیدی : معادلات انتگرال-دیفرانسیل ولترا؛ روش گالرکین؛ چندجمله ای های چبیشف.

روش ساده ترین معادله برای جواب های دقیق اتمسفر مغناطیسی هم دما

نعمت اله کدخدا [۱] *و حسین جعفری* [۲]

[۱] دانشگاه بزرگمهر قائنات، دانشکده علوم پایه، گروه ریاضی
[۲] دانشگاه مازندران، دانشکده ریاضی و علوم کامپیوتر

دریافت مقاله ۲۹ اسفند ۱۳۹۳، دریافت مقاله اصلاح شده ۲۷ خرداد ۱۳۹۴، پذیرش مقاله ۱۰ مرداد ۱۳۹۴

چکیده : روش کودریاشف برای جستجوی جواب های دقیق معادلات دیفرانسیل غیر خطی ارائه شده است. روش کودریاشف برای جستجو جوابهای دقیق معادله لیوویل و معادله سینویس هایپربولیک-پواسون به کار برده شده است. معادلات تعادل هیدرواستاتیک مغناطیس برای پلاسما در یک میدان گرانشی به صورت تحلیلی مورد بررسی قرار گرفته اند. بررسی یک خانواده از اتمسفر مغناطیسی هم دما با یک مختصات متناظر به یک میدان گرانشی یکنواخت در هندسه مسطحه انجام شده است. جریان توزیع شده در مدل J در امتداد محور x که در آن x مختصات افقی می باشد هدایت شده است. این معادلات تبدیل به یک معادله بیضی غیر خطی برای پتانسیل برداری مغناطیسی u می شوند. این معادله به یک تابع دلخواه u وابسته است که باید مشخص شود.

کلمات کلیدی : روش کودریاشف؛ تعادل مغناطیسی؛ معادلات تکامل غیر خطی؛ امواج تراولینگ.

یک طرح تفاضلی متناهی غیر استاندارد برای حل شبکه‌ی غذایی سه بعدی با مدل لوتکا-ولترا مرتبه‌ی کسری

صادق زیبایی و مهران نامجو

دانشگاه ولی عصر رفسنجان، دانشکده علوم ریاضی، گروه ریاضی

دریافت مقاله ۲ آذر ۱۳۹۳، دریافت مقاله اصلاح شده ۱۶ خرداد ۱۳۹۴، پذیرش مقاله ۱۰ مرداد ۱۳۹۴

چکیده : در این مقاله یک مدل مرتبه‌ی کسری از شبکه‌ی غذایی سه بعدی لوتکا-ولترا را معرفی می‌کنیم. تحلیل پایداری سیستم کسری را شرح می‌دهیم. طرح تفاضلی متناهی غیراستانداردی بیان می‌شود که رفتار دینامیکی سیستم لوتکا-ولترا مرتبه‌ی کسری را مورد مطالعه قرار می‌دهد. نتایج عددی نشان می‌دهند که تقریبات طرح تفاضلی متناهی غیر استاندارد زمانی که برای سیستم لوتکا-ولترا مرتبه‌ی کسری استفاده می‌شوند، بسیار دقیق هستند.

کلمات کلیدی : معادلات دیفرانسیل کسری؛ مدل لوتکا-ولترا؛ سیستم شکار-شکارچی؛ طرح تفاضلی متناهی غیراستاندارد؛ پایداری.

یک الگوریتم تعاملی برای حل مسائل بهینه‌سازی چندهدفه بر اساس یک تکنیک اسکالرسازی عمومی

مهرداد غزنوی[۱] ، *محمد ایلاتی*[۲] *و اسماعیل خرم* [۲]

[۱] دانشگاه شاهرود، دانشکده علوم ریاضی، گروه ریاضی کاربردی
[۲] دانشگاه صنعتی امیرکبیر، دانشکده ریاضی و علوم کامپیوتر

دریافت مقاله ۳ اسفند ۱۳۹۳، دریافت مقاله اصلاح شده ۱۳ تیر ۱۳۹۴، پذیرش مقاله ۲۹ مهر ۱۳۹۴

چکیده : تنوع روش‌های تعاملی موجود، نیاز به ارایه الگوریتم‌های تعاملی عمومی که تصمیم گیرنده را قادر می‌سازند که آزادانه چندین روش مناسب را که مورد ترجیح او است را انتخاب کند را نشان می‌دهد. برای این منظور، در این مقاله، یک مسئله اسکالرسازی عمومی برای مسائل برنامه‌ریزی چند هدفه پیشنهاد می‌دهیم. رابطه بین جواب های بهینه مسئله اسکالرسازی معرفی شده و جواب های کارا (ضعیف) و کارای سره مسئله بهینه‌سازی چند هدفه اصلی بررسی می‌شود. نشان می‌دهیم که با انتخاب تبدیل‌های مناسب، برخی از مسائل اسکالرسازی به کار رفته در روش‌های تعاملی مختلف می‌توانند از فرمول پیشنهاد شده بدست آیند. بر اساس مسئله اسکالرسازی پیشنهاد شده، یک الگوریتم تعاملی عمومی پیشنهاد می‌دهیم که تصمیم گیرنده را قادر می کند ترجیحاتش را با شش روش مختلف و با قابلیت تغییر در ترجیحات در هر زمان در طول تکرارهای الگوریتم مشخص کند. سرانجام، یک مثال عددی که بیانگر کاربردی بودن الگوریتم است ارایه می‌گردد.

کلمات کلیدی : بهینه‌سازی چندهدفه؛ روش تعاملی؛ مسئله اسکالرسازی؛ کارایی سره؛ اطلاعات ترجیحی.

یک الگوریتم ناحیه اطمینان غیریکنوا با شعاع تطبیقی غیریکنوا برای حل دستگاه معادلات غیر خطی

کیوان امینی[۱]، *حمید اسمعیلی*[۲] و *مرتضی کیمیایی*[۳]

[۱] دانشگاه رازی، دانشکده علوم، گروه ریاضی
[۲] دانشگاه بوعلی سینا، گروه ریاضی
[۳] دانشگاه آزاد اسلامی، واحد اسدآباد، گروه ریاضی

دریافت مقاله ۱۶ فروردین ۱۳۹۴، دریافت مقاله اصلاح شده ۱۳ تیر ۱۳۹۴، پذیرش مقاله ۲۹ مهر ۱۳۹۴

چکیده : در این مقاله یک روش ناحیه اطمینان غیریکنوا برای حل دستگاه های معادلات غیر خطی معرفی می گردد که از یک شعاع تطبیقی مناسب استفاده می کند. استفاده همزمان از تکنیک های غیر یکنوا و یک شعاع اطمینان مناسب می تواند کارایی روش های ناحیه اطمینان را به طرز قابل ملاحظه ای افزایش دهد جایی که هزینه محاسباتی روش نیز به دلیل کاهش تعداد زیر مسائل حل شده کاهش می یابد. همگرایی سراسری و $q-$ مجذوری روش تحت شرایط مناسب اثبات گردیده است. نتایج عددی ارائه شده نمایانگر کارایی و سرعت مناسب الگوریتم جدید در مقایسه با الگوریتم های مشابه می باشد.

کلمات کلیدی : دستگاه معادلات غیرخطی؛ الگوریتم ناحیه اطمینان؛ شعاع تطبیقی؛ تکنیک های غیریکنوا.

توابع انقباض روی فضاهای نرم دار فازی

مرتضی ساحلی

دانشگاه ولی عصر رفسنجان، گروه ریاضی

دریافت مقاله ۴ آذر ۱۳۹۳، دریافت مقاله اصلاح شده ۱۹ مهر ۱۳۹۴، پذیرش مقاله ۵ آذر ۱۳۹۴

چکیده : در این مقاله، ما از تعریف فضای نرم دار فازی که توسط بگ و سامانتا ارایه شده است استفاده کرده ایم و چهار نوع از توابع انقباض فازی را معرفی کرده ایم . ما نشان داده ایم که این توابع روی فضاهای نرم دار فازی لزوما دارای نقطه ثابت منحصر به فرد می باشد و نشان خواهیم داد که این قضایای ارایه شده در واقع توسیع فازی قضایای کلاسیک می باشند

کلمات کلیدی : نرم فازی؛ فضای نرم دار فازی؛ نقطه ثابت؛ α-نیم نرم؛ شرایط انقباض.

Aims and scope

Iranian Journal of Numerical Analysis and Optimization (IJNAO) is published twice a year by the Department of Applied Mathematics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad. Papers dealing with different aspects of numerical analysis and optimization, theories and their applications in engineering and industry are considered for publication.

Journal Policy

After receiving an article, the editorial committee will assign referees. Refereeing process can be followed via the web site of the Journal.
The manuscripts are accepted for review with the understanding that the work has not been published and also it is not under consideration for publication by any other journal. All submissions should be accompanied by a written declaration signed by the author(s) that the paper has not been published before and has not been submitted for consideration elsewhere.

Instruction for Authors

The Journal publishes all papers in the fields of numerical analysis and optimization. Articles must be written in English.
All submitted papers will be refereed and the authors may be asked to revise their manuscripts according to the referee's reports. The Editorial Board of the Journal keeps the right to accept or reject the papers for publication.
The papers with more than one authors, should determine the corresponding author. The e-mail address of the corresponding author must appear at the end of the manuscript or as a footnote of the first page.
It is strongly recommended to set up the manuscript by Latex or Tex, using the template provided in the web site of the Journal. Manuscripts should be typed double-spaced with wide margins to provide enough room for editorial remarks.
References should be arranged in alphabetical order by the surname of the first author as examples below:
[1] Stoer, J. and Bulirsch, R. *Introduction to Numerical Analysis*, Springer-verlag, New York, 2002.
[2] Brunner, H. *A survey of recent advances in the numerical treatment of Volterra integral and integro-differential equations*, J. Comput. Appl. Math. 8 (1982), 213-229.

Submission of Manuscripts

Authors may submit their manuscripts by either of the following ways:

a) Online submission (pdf or dvi files) via the website of the Journal at:

http://ijnao.um.ac.ir

b) Via journal's email mjms@um.ac.ir

Copyright Agreement

Upon the acceptance of an article by the Journal, the corresponding author will be asked to sign a "Copyright Transfer Agreement" (see the web site) and send it to the Journal address. This will permit the publisher to publish and distribute the work.

Reviewers Acknowledgment

Peer reviewers are key to advancing scholarship and contributing to the quality of scholarly journals. We would like to sincerely thank the following reviewers who have taken part in the peer-review process for Iranian Journal of Numerical Analysis and Optimization during the last three years.

Abbasbandy, S.	Abdi, A.
Abolfazli Asfahani, J.	Adibi, H.
Afsharnejad, Z.	Akbarian, D.
Aminataee, A.	Amini, K.
Aminikhah, H.	Arab Ameri, M.
Avazzadeh, Z.	Aziz, I.
Babolian, E.	Baghani, O.
Barfeie, M.	Barid Loghmani, Gh.
Benner, P.	Biazar, J.
Cattani, C.	Chen, W.
Dadi, Z.	Darvishi, M.
Davoodi, N.	Deldari, H.
Dereli, Y.	Ding, X.
Ebadi, M.	Effati, S.
Eghbali, N.	Erjaee, Gh.H
Ezzati, R.	Fakharzadeh Jahromi, A.
Farahi, M.H.	Farhadinia, B.
Fatemi, M.	Fathali, J.
Gachpazan, M.	Garshasbi, M.
Ghaemi, M.B.	Ghanbari, R.
Ghane, F.H.	Ghatee, M.
Ghesmati, A.	Ghomanjani, F.
Ghorbani, A.	Ghoreshi, F.
Ghovatmand, M.	Golbabai, A.
Hadizadeh Yazdi, M	Hajarian, M.
Hashemi Borzabadi, A.	Hashemi Mehne, H.
Hassanpur, H.	Hatami, M.
Heidari, A.	Hesameddini, E.
Hojjati, Gh.	Hoseinzadeh Lotfi, F.
Hossieni, M.M.	Irandoust Pakchin, S.
Jafari, H.	Javadi, Sh.
Javidmanesh, E.	Jeiani Rezaei, H.
Karimi Delameh, M.	Ketabi, A.
Khaki Sedigh, A.	Khodabin, M.
Khojasteh Salkuyeh, D.	Kianpour, M.
Kweeng, E.	Liang, H.
Liu, H.	Maalek, F.
Mahdizade Kalsaraee, M.	Maleknejad, Kh.
Marzban, H.R.	Matinfar, M.

Mesforush, A.
Mirdehghan, M.
Mirzapour, F.
Mirzavaziri, M.
Mirzazadeh, M.
Mohammadinejad, M.
Mohebbi, A.
Mokhtari, R.
Namjoo, M.
Nazemi, A.R.
Panjeh Ali Beik, F.
Pariz, N.
Pourgholi, R.
Raayatpanah, M.A.
Rabieimotlagh, O.
Rafiei, A.
Rang, J.
Rashidinia, J.
Rathsfeld, A.
Rivaz, A.
Saadatmandi, A.
Saberi Najafi, J.
Saeed, U.
Saeidian, J.
Safdari, A.
Salahi, M.
Sarra Scott, A.
Serjaee, Gh.H.
Shamsi, M.
Shanazari, K.
Shirazian, M.
Shores, Y.
Soheili, A.R.
Solaymani Fard, O.
Soleimani, M.
Soleymani, F.
Taghizadeh Kakhki, H.
Tareghian, H.R.
Tatari, M
Tohidi, E.
Toutounian Mashhad, F.
Vahidi, J.
Vitanov, N.K.
Yaghouti, M.R.
Yakit Ongun, M.
Yao, G.
Yousefi, S.A.
Zakeri, A.
Zayed, E.M.E.
Zhang, H.

Iranian Journal of Numerical Analysis and Optimization

Former : MRJMS

CONTENTS Vol. 6, No.1, pp 1-136, 2016

web site : http://ijnao.um.ac.ir

Email : mjms@um.ac.ir

ISSN : 2423-6977