

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys

Volume 23 Number 5 2023

CONTENTS

Solvability of Nonlinear Elliptic Problems with Degenerate Coercivity in Weighted Sobolev Space 461
Y. Akdim, S. Lalaoui Rhali and Y. Oumouacha

An Algorithm for Solving First-Kind Two-Dimensional Volterra Integral Equations Using Collocation Method 475
F. Birem, A. Boulmerka, H. Laib and C. Hennous

Application of Model Predictive Control (MPC) to Longitudinal Motion of the Aircraft Using Polynomial Chaos 487
K. D. R. Dewi, K. Fahim and S. Subchan

Singular Reaction-Diffusion System Arising from Quenching 499
Samiha Djemai and Salim Mesbahi

Maximum and Anti-Maximum Principles for Boundary Value Problems for Ordinary Differential Equations in Neighborhoods of Simple Eigenvalues 507
P. W. Eloë and J. T. Neugebauer

Performance Comparison of Sliding Mode Control and Sliding PID for Rescue ROV 519
T. Herlambang, A. Suryowinoto, I. Kurniastuti, H. Nurhadi and K. Oktafianto

Class of Nilpotent Distributions and \mathfrak{N}_2 -Distributions 538
Y. Khellaf, A. Chikh-Salah and N. Bensalem

Moore-Spiegel Chaotic Encryption for Digital Images and Voices 545
W. S. Mada Sanjaya, Akhmad Roziqin, Agung Wijaya Temiesela, M. Fauzi Badru Zaman, Aria Dewa Wibiksana and Dyah Anggraeni

Study of a Penalty Method for Nonlinear Optimization Based on a New Approximate Function 561
C. Souli and A. Leulmi

Contents of Volume 23, 2023 571

NONLINEAR DYNAMICS & SYSTEMS THEORY

Volume 23, No. 5, 2023

Nonlinear Dynamics and Systems Theory

An International Journal of Research and Surveys

EDITOR-IN-CHIEF A.A.MARTYNYUK

*S.P.Timoshenko Institute of Mechanics
National Academy of Sciences of Ukraine, Kiev, Ukraine*

MANAGING EDITOR I.P.STAVROULAKIS

Department of Mathematics, University of Ioannina, Greece

REGIONAL EDITORS

P.BORNE, Lille, France
 A.OKNIŃSKI, Kielce, Poland
Europe

M.BOHNER, Rolla, USA
 HAO WANG, Edmonton, Canada
USA and Canada

T.A.BURTON, Port Angeles, USA
 C.CRUZ-HERNANDEZ, Ensenada, Mexico
USA and Latin America

M.ALQURAN, Irbid, Jordan
Jordan and Middle East

T.HERLAMBANG, Surabaya, Indonesia
Indonesia and New Zealand

Nonlinear Dynamics and Systems Theory
An International Journal of Research and Surveys

EDITOR-IN-CHIEF A.A.MARTYNYUK

The S.P.Timoshenko Institute of Mechanics, National Academy of Sciences of Ukraine,
Nesterov Str. 3, 03680 MSP, Kiev-57, UKRAINE / e-mail: journalndst@gmail.com

MANAGING EDITOR I.P.STAVROULAKIS

Department of Mathematics, University of Ioannina
451 10 Ioannina, HELLAS (GREECE) / e-mail: ipstav@cc.uoi.gr

ADVISORY EDITOR A.G.MAZKO,

Institute of Mathematics of NAS of Ukraine, Kiev (Ukraine)
e-mail: mazko@imath.kiev.ua

REGIONAL EDITORS

M.ALQURAN (Jordan), e-mail: marwan04@just.edu.jo
P.BORNE (France), e-mail: Pierre.Borne@ec-lille.fr
M.BOHNER (USA), e-mail: bohner@mst.edu
T.A.BURTON (USA), e-mail: taburton@olypen.com
C.CRUZ-HERNANDEZ (Mexico), e-mail: ccruz@cicese.mx
T.HERLAMBANG (Indonesia), e-mail: teguh@unusa.ac.id
A.OKNINSKI (Poland), e-mail: fizao@tu.kielce.pl
HAO WANG (Canada), e-mail: hao8@ualberta.ca

EDITORIAL BOARD

Adzkiya, D. (Indonesia)	Kloedon, P. (Germany)
Artstein, Z. (Israel)	Kokologiannaki, C. (Greece)
Awrejcewicz, J. (Poland)	Kouzou, A. (Algeria)
Braiek, N.B. (Tunisia)	Krishnan, E.V. (Oman)
Chen Ye-Hwa (USA)	Limarchenko, O.S. (Ukraine)
De Angelis, M. (Italy)	Lopez Gutierrez R.M. (Mexico)
Denton, Z. (USA)	Peterson, A. (USA)
Djemai, M. (France)	Radziszewski, B. (Poland)
Dshalalow, J.H. (USA)	Shi Yan (Japan)
Gajic Z. (USA)	Sivasundaram, S. (USA)
Georgiev, S.G. (France)	Sree Hari Rao, V. (India)
Georgiou, G. (Cyprus)	Staicu V. (Portugal)
Honglei Xu (Australia)	Vatsala, A. (USA)
Jafari, H. (South African Republic)	Zuyev, A.L. (Germany)
Khusainov, D.Ya. (Ukraine)	

ADVISORY COMPUTER SCIENCE EDITORS

A.N.CHERNIENKO and A.S.KHOROSHUN, Kiev, Ukraine

ADVISORY LINGUISTIC EDITOR

S.N.RASSHYVALOVA, Kiev, Ukraine

© 2023, InforMath Publishing Group, ISSN 1562-8353 print, ISSN 1813-7385 online, Printed in Ukraine
No part of this Journal may be reproduced or transmitted in any form or by any means without
permission from InforMath Publishing Group.

INSTRUCTIONS FOR CONTRIBUTORS

(1) General. Nonlinear Dynamics and Systems Theory (ND&ST) is an international journal devoted to publishing peer-refereed, high quality, original papers, brief notes and review articles focusing on nonlinear dynamics and systems theory and their practical applications in engineering, physical and life sciences. Submission of a manuscript is a representation that the submission has been approved by all of the authors and by the institution where the work was carried out. It also represents that the manuscript has not been previously published, has not been copyrighted, is not being submitted for publication elsewhere, and that the authors have agreed that the copyright in the article shall be assigned exclusively to InforMath Publishing Group by signing a transfer of copyright form. Before submission, the authors should visit the website:

<http://www.e-ndst.kiev.ua>

for information on the preparation of accepted manuscripts. Please download the archive Sample_NDST.zip containing example of article file (you can edit only the file Samplefilename.tex).

(2) Manuscript and Correspondence. Manuscripts should be in English and must meet common standards of usage and grammar. To submit a paper, send by e-mail a file in PDF format directly to

*Professor A.A. Martynyuk, Institute of Mechanics,
Nesterov str.3, 03057, MSP 680, Kiev-57, Ukraine
e-mail: journalndst@gmail.com*

or to one of the Regional Editors or to a member of the Editorial Board. Final version of the manuscript must typeset using LaTeX program which is prepared in accordance with the style file of the Journal. Manuscript texts should contain the title of the article, name(s) of the author(s) and complete affiliations. Each article requires an abstract not exceeding 150 words. Formulas and citations should not be included in the abstract. AMS subject classifications and key words must be included in all accepted papers. Each article requires a running head (abbreviated form of the title) of no more than 30 characters. The sizes for regular papers, survey articles, brief notes, letters to editors and book reviews are: (i) 10-14 pages for regular papers, (ii) up to 24 pages for survey articles, and (iii) 2-3 pages for brief notes, letters to the editor and book reviews.

(3) Tables, Graphs and Illustrations. Each figure must be of a quality suitable for direct reproduction and must include a caption. Drawings should include all relevant details and should be drawn professionally in black ink on plain white drawing paper. In addition to a hard copy of the artwork, it is necessary to attach the electronic file of the artwork (preferably in PCX format).

(4) References. Each entry must be cited in the text by author(s) and number or by number alone. All references should be listed in their alphabetic order. Use please the following style:

Journal: [1] H. Poincare, Title of the article. *Title of the Journal* **volume**
(issue) (year) pages. [Language]

Book: [2] A.M. Lyapunov, *Title of the Book*. Name of the Publishers, Town, year.

Proceeding: [3] R. Bellman, Title of the article. In: *Title of the Book*. (Eds.).
Name of the Publishers, Town, year, pages. [Language]

(5) Proofs and Sample Copy. Proofs sent to authors should be returned to the Editorial Office with corrections within three days after receipt. The corresponding author will receive a sample copy of the issue of the Journal for which his/her paper is published.

(6) Editorial Policy. Every submission will undergo a stringent peer review process. An editor will be assigned to handle the review process of the paper. He/she will secure at least two reviewers' reports. The decision on acceptance, rejection or acceptance subject to revision will be made based on these reviewers' reports and the editor's own reading of the paper.

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys
Published by InforMath Publishing Group since 2001

Volume 23

Number 5

2023

CONTENTS

Solvability of Nonlinear Elliptic Problems with Degenerate Coercivity in Weighted Sobolev Space	461
<i>Y. Akdim, S. Lalaoui Rhali and Y. Oumouacha</i>	
An Algorithm for Solving First-Kind Two-Dimensional Volterra Integral Equations Using Collocation Method	475
<i>F. Birem, A. Boulmerka, H. Laib and C. Hennous</i>	
Application of Model Predictive Control (MPC) to Longitudinal Motion of the Aircraft Using Polynomial Chaos	487
<i>K. D. R. Dewi, K. Fahim and S. Subchan</i>	
Singular Reaction-Diffusion System Arising from Quenching	499
<i>Samihha Djemai and Salim Mesbahi</i>	
Maximum and Anti-Maximum Principles for Boundary Value Problems for Ordinary Differential Equations in Neighborhoods of Simple Eigenvalues	507
<i>P. W. Eloë and J. T. Neugebauer</i>	
Performance Comparison of Sliding Mode Control and Sliding PID for Rescue ROV	519
<i>T. Herlambang, A. Suryowinoto, I. Kurniastuti, H. Nurhadi and K. Oktafianto</i>	
Class of Nilpotent Distributions and \mathfrak{N}_2 -Distributions	538
<i>Y. Khellaf, A. Chikh-Salah and N. Bensalem</i>	
Moore-Spiegel Chaotic Encryption for Digital Images and Voices	545
<i>W. S. Mada Sanjaya, Akhmad Roziqin, Agung Wijaya Temiesela, M. Fauzi Badru Zaman, Aria Dewa Wibiksana and Dyah Anggraeni</i>	
Study of a Penalty Method for Nonlinear Optimization Based on a New Approximate Function	561
<i>C. Souli and A. Leulmi</i>	
Contents of Volume 23, 2023	571

Founded by A.A. Martynyuk in 2001.

Registered in Ukraine Number: KB No 5267 / 04.07.2001.

NONLINEAR DYNAMICS AND SYSTEMS THEORY

An International Journal of Research and Surveys

Impact Factor from SCOPUS for 2022: SJR – 0.293, SNIP – 0.784 and CiteScore – 1.5

Nonlinear Dynamics and Systems Theory (ISSN 1562–8353 (Print), ISSN 1813–7385 (Online)) is an international journal published under the auspices of the S.P. Timoshenko Institute of Mechanics of National Academy of Sciences of Ukraine and Curtin University of Technology (Perth, Australia). It aims to publish high quality original scientific papers and surveys in areas of nonlinear dynamics and systems theory and their real world applications.

AIMS AND SCOPE

Nonlinear Dynamics and Systems Theory is a multidisciplinary journal. It publishes papers focusing on proofs of important theorems as well as papers presenting new ideas and new theory, conjectures, numerical algorithms and physical experiments in areas related to nonlinear dynamics and systems theory. Papers that deal with theoretical aspects of nonlinear dynamics and/or systems theory should contain significant mathematical results with an indication of their possible applications. Papers that emphasize applications should contain new mathematical models of real world phenomena and/or description of engineering problems. They should include rigorous analysis of data used and results obtained. Papers that integrate and interrelate ideas and methods of nonlinear dynamics and systems theory will be particularly welcomed. This journal and the individual contributions published therein are protected under the copyright by International InforMath Publishing Group.

PUBLICATION AND SUBSCRIPTION INFORMATION

Nonlinear Dynamics and Systems Theory will have 5 issues in 2024, printed in hard copy (ISSN 1562–8353) and available online (ISSN 1813–7385), by InforMath Publishing Group, Nesterov str., 3, Institute of Mechanics, Kiev, MSP 680, Ukraine, 03057. Subscription prices are available upon request from the Publisher, EBSCO Information Services (<mailto:journals@ebSCO.com>), or website of the Journal: <http://e-ndst.kiev.ua>. Subscriptions are accepted on a calendar year basis. Issues are sent by airmail to all countries of the world. Claims for missing issues should be made within six months of the date of dispatch.

ABSTRACTING AND INDEXING SERVICES

Papers published in this journal are indexed or abstracted in: Mathematical Reviews / MathSciNet, Zentralblatt MATH / Mathematics Abstracts, PASCAL database (INIST–CNRS) and SCOPUS.



Solvability of Nonlinear Elliptic Problems with Degenerate Coercivity in Weighted Sobolev Space

Y. Akdim¹, S. Lalaoui Rhali² and Y. Oumouacha^{2*}

¹ *LAMA Laboratory, Faculty of Sciences Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, Fez, Morocco.*

² *Engineering Sciences Laboratory (LSI), Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, Taza, Morocco.*

Received: July 8, 2023; Revised: November 16, 2023

Abstract: In this paper, we investigate the existence of our entropy solution for the nonlinear elliptic equation

$$-\operatorname{div}[\omega(x)a(x, u, \nabla u)] = f - \operatorname{div} F, \quad \text{in } \Omega,$$

in the setting of the weighted Sobolev space $W_0^{1,p}(\Omega, \omega)$. We focus on the case where the operator has a degenerate coercivity and $f \in L^1(\Omega)$, $F \in [L^{p'}(\Omega, \omega^{1-p'})]^N$.

Keywords: *nonlinear elliptic equations; degenerate coercivity; entropy solutions; weighted Sobolev spaces.*

Mathematics Subject Classification (2010): 35J60; 35J70; 46E35; 70K99; 93-02.

1 Introduction

Partial differential equations have many applications in various areas of engineering, mathematics, physics, and other applied sciences (see for instance [8, 20]). In the last years, there has been an increasing interest in the study of various mathematical problems in weighted Sobolev spaces motivated by many considerations in applications (see [1, 2, 5, 10, 11] and the references therein).

Let Ω be a bounded smooth subset of \mathbb{R}^N with $N \geq 2$ and $1 < p < \infty$. We are interested in proving the existence of entropy solutions to the following elliptic Dirichlet problem:

$$(P) \begin{cases} -\operatorname{div}[\omega(x)a(x, u, \nabla u)] = f - \operatorname{div} F, & \text{in } \Omega, \\ u(x) = 0, & \text{on } \partial\Omega. \end{cases}$$

* Corresponding author: <mailto:youssef.oumouacha@usmba.ac.ma>

Here $a(x, s, \xi) : \Omega \times \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ is a Carathéodory function satisfying the following conditions:

$$a(x, s, \xi) \cdot \xi \geq \frac{\alpha}{(1 + |s|)^{\theta(p-1)}} |\xi|^p \quad (1)$$

for some $\alpha > 0$ and some real number θ such that $0 \leq \theta < 1$.

As far as the datum f and F are concerned, we will assume that f belongs to the space $L^1(\Omega)$, and $F \in [L^{p'}(\Omega, \omega^{1-p'})]^N$ with $(1/p + 1/p' = 1)$.

Problems like (\mathcal{P}) have been studied by many authors in the non-weighted case. In [18], Leone and Porretta studied the nonlinear elliptic problem

$$Bu = f(x) - \operatorname{div}(F) \quad \text{in } \Omega$$

in the setting of Sobolev spaces, where $Bu = -\operatorname{div}(a(x, u, \nabla u))$ is a Leray-Lions operator from $W_0^{1,p}(\Omega)$ to $W^{-1,p'}(\Omega)$, they demonstrated the existence of entropy solutions. In addition, Alvino et al. [4] have proved that the nonlinear elliptic equations $-\operatorname{div}(a(x, u, \nabla u)) = f$ admit the entropy solutions under assumption (1).

Notice that the existence of a weak solution for the Dirichlet problem (\mathcal{P}) has been obtained by Cavalheiro in [12] under the condition

$$a(x, s, \xi) \cdot \xi \geq \alpha |\xi|^p, \quad (2)$$

and by assuming that $f/\omega \in L^{p'}(\Omega, \omega)$. Also, he discussed in [11] the existence of our entropy solution when $f \in L^1(\Omega)$.

Our objective in this work is to study the problem (\mathcal{P}) when the operator satisfies assumption (1) instead of (2). The main difficulties that arise in our study are due, on the one hand, to the fact that the differential operator $A(u) = -\operatorname{div}(\omega(x)a(x, u, \nabla u))$, which is well defined between $W_0^{1,p}(\Omega, \omega)$ and its dual $W^{-1,p'}(\Omega, \omega^{1-p'})$, may not be coercive on $W_0^{1,p}(\Omega, \omega)$ when u is large. This imposes that, even if the datum is extremely regular, the classical methods used to demonstrate the existence of a solution to problem (\mathcal{P}) cannot be used.

On the other hand, f only belongs to $L^1(\Omega)$, making it difficult to show the existence of a weak solution. To get around this difficulty, we will use in this paper the concept of entropy solutions. This concept was introduced in [7], and then used by many authors to study elliptic equations (see [2–4, 6, 10, 18]).

The structure of this paper is as follows. In Section 2, we recall some preliminary results which will be used later. In Section 3, we give the assumptions on the data, then we state the main results which will be proved in Section 4.

2 Preliminaries

In this section, we provide a brief facts about the weighted Sobolev space as well as some \mathcal{A}_p -weight features. Let $\omega = \omega(x)$ be a weight function, that is, $0 < \omega < \infty$, and a locally integrable function on \mathbb{R}^N . By integration, each weight ω generates a measure on the measurable subsets of \mathbb{R}^N . This measure is denoted by μ and defined as follows:

$$\mu(\mathcal{S}) = \int_{\mathcal{S}} \omega(x) dx$$

for a measurable set $\mathcal{S} \subset \mathbb{R}^N$.

Definition 2.1 Let ω be a weight and $1 < p < \infty$. We say that ω belongs to \mathcal{A}_p -weight if there exists a positive constant $C_{\omega,p}$ such that, for every ball $\mathcal{B} \subset \mathbb{R}^N$,

$$\left(\frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} \omega dx\right) \left(\frac{1}{|\mathcal{B}|} \int_{\mathcal{B}} \omega^{1/(1-p)} dx\right)^{p-1} \leq C_{\omega,p} \quad \text{if } p > 1,$$

where $|\cdot|$ denotes the N -dimensional Lebesgue measure in \mathbb{R}^N .

Lemma 2.1 [16]. Let \mathcal{B} be a ball in \mathbb{R}^N and \mathcal{S} be a measurable subset of \mathcal{B} . If $\omega \in \mathcal{A}_p, 1 < p < \infty$, then

$$\left(\frac{|\mathcal{S}|}{|\mathcal{B}|}\right)^p \leq C_{\omega,p} \frac{\mu(\mathcal{S})}{\mu(\mathcal{B})}.$$

Remark 2.1 If $\mu(\mathcal{S}) = 0$, then $|\mathcal{S}| = 0$. Thus, for every sequence (u_n) in \mathcal{B} that converges μ -a.e. to some u , we have $u_n \rightarrow u$ a.e.

The weighted Lebesgue space $L^p(\Omega, \omega)$ is defined for every weight ω and $1 \leq p < \infty$ by

$$L^p(\Omega, \omega) = \left\{ u = u(x) : u\omega^{1/p} \in L^p(\Omega) \right\},$$

and it is endowed with the norm

$$\|u\|_{L^p(\Omega, \omega)} = \left(\int_{\Omega} |u(x)|^p \omega(x) dx \right)^{1/p}.$$

Definition 2.2 Let Ω be a bounded open subset of $\mathbb{R}^N, 1 < p < \infty$ and let ω be an \mathcal{A}_p -weight. The weighted Sobolev space $W^{1,p}(\Omega, \omega)$ is defined as the set of all functions $u \in L^p(\Omega, \omega)$ with weak derivatives $\frac{\partial u}{\partial x_i} \in L^p(\Omega, \omega)$, for all $i = 1, \dots, N$.

The norm of u in $W^{1,p}(\Omega, \omega)$ is given by

$$\|u\|_{W^{1,p}(\Omega, \omega)} = \left(\int_{\Omega} |u|^p \omega(x) dx + \int_{\Omega} |\nabla u|^p \omega(x) dx \right)^{\frac{1}{p}}. \tag{3}$$

The space $W_0^{1,p}(\Omega, \omega)$ is defined as the closure of $C_0^\infty(\Omega)$ with respect to the norm

$$\|u\|_{W_0^{1,p}(\Omega, \omega)} = \left(\int_{\Omega} |\nabla u|^p \omega(x) dx \right)^{\frac{1}{p}}. \tag{4}$$

A compact imbedding is required because we are working with compactness methods to find solutions to nonlinear elliptic equations. As a result, we assume also that the domain Ω is smooth.

Theorem 2.1 [13]. Let Ω be a bounded smooth domain. For $\omega \in \mathcal{A}_p$, we have the compact embedding

$$W_0^{1,p}(\Omega, \omega) \hookrightarrow L^p(\Omega, \omega).$$

Theorem 2.2 [14]. Let Ω be a bounded open subset of \mathbb{R}^N . Take $1 < p < \infty$ and a function $\omega \in \mathcal{A}_p$. There exist positive constants C_Ω and δ such that for all $u \in C_0^\infty(\Omega)$ and all η satisfying $1 \leq \eta \leq \frac{N}{N-1} + \delta$,

$$\|u\|_{L^{\eta p}(\Omega, \omega)} \leq C_\Omega \|\nabla u\|_{L^p(\Omega, \omega)}. \tag{5}$$

Definition 2.3 Let ω be a weight function and let q be a positive real number. The weighted Marcinkiewicz space $\mathcal{M}^q(\Omega, \omega)$ is the set of all measurable functions $f : \Omega \rightarrow \mathbb{R}$ such that the function

$$\Phi_f(k) = \mu(\{x \in \Omega : |f(x)| > k\}) \quad k > 0,$$

satisfies, for some positive constant C , an estimate of the form $\Phi_f(k) \leq Ck^{-q}$.

Remark 2.2 It follows from [17] that if $1 \leq q < p$ and $\Omega \subset \mathbb{R}^N$ is a bounded set, then

$$L^p(\Omega, \omega) \subset \mathcal{M}^p(\Omega, \omega) \text{ and } \mathcal{M}^p(\Omega, \omega) \subset L^q(\Omega, \omega).$$

3 Basic Assumptions and Main Result

3.1 Basic assumptions

Let Ω be an open bounded smooth domain of \mathbb{R}^N ($N \geq 2$), $p > 1$ and $\omega \in \mathcal{A}_p$. Let $a : \Omega \times \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ be a Carathéodry function (that is, $a(\cdot, s, \xi)$ is measurable on Ω for every (t, ξ) in $\mathbb{R} \times \mathbb{R}^N$, and $a(x, \cdot, \cdot)$ is continuous on $\mathbb{R} \times \mathbb{R}^N$ for almost every x in Ω). Assume that

$$a(x, s, \xi) \cdot \xi \geq b(|s|)|\xi|^p \quad (6)$$

for almost every x in Ω and for every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$, where

$$b(s) = \frac{\alpha}{(1+s)^{\theta(p-1)}} \quad (7)$$

for some $0 \leq \theta < 1$ and some $\alpha > 0$;

$$|a(x, s, \xi)| \leq l_0(x) + l_1(x)|s|^{p-1} + l_2(x)|\xi|^{p-1} \quad (8)$$

for almost every x in Ω and for every $(s, \xi) \in \mathbb{R} \times \mathbb{R}^N$, where l_0, l_1 and l_2 are non-negative functions, with $l_0 \in L^{p'}(\Omega, \omega)$ and $l_1, l_2 \in L^\infty(\Omega)$;

$$[a(x, s, \xi) - a(x, s, \xi')] \cdot (\xi - \xi') > 0 \quad (9)$$

for almost every x in Ω and for every $s \in \mathbb{R}$, for every ξ, ξ' for every ξ, ξ' in \mathbb{R}^N with $\xi \neq \xi'$. As regards the source term, we assume that

$$f \in L^1(\Omega) \quad \text{and} \quad F \in [L^{p'}(\Omega, \omega^{1-p'})]^N. \quad (10)$$

3.2 Main result

We first give the definition of an entropy solution of problem (\mathcal{P}) . For a given constant $k > 0$, the truncation function $T_k : \mathbb{R} \rightarrow \mathbb{R}$ is defined by

$$T_k(s) = \max\{-k, \min\{k, s\}\}.$$

We denote by $\mathcal{T}_0^{1,p}(\Omega, \omega)$ the set of all measurable functions $u : \Omega \rightarrow \mathbb{R}$ such that for every $k > 0$, the truncated function $T_k(u)$ belongs to $W_0^{1,p}(\Omega, \omega)$.

Let $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$, then there exists a unique measurable function $v : \Omega \rightarrow \mathbb{R}^N$ such that

$$\nabla T_k(u) = v\chi_{\{|u| < k\}}. \quad (11)$$

If $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$, the weak gradient ∇u of u is defined as the unique function v which satisfies (11).

Definition 3.1 A function $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$ is an entropy solution of the problem (P) if

$$\int_{\Omega} \omega a(x, u, \nabla u) \cdot \nabla T_k(u - v) dx = \int_{\Omega} f T_k(u - v) dx + \int_{\Omega} F \cdot \nabla T_k(u - v) dx \quad (12)$$

for every $k > 0$ and for every $v \in W_0^{1,p}(\Omega, \omega) \cap L^\infty(\Omega)$.

The main result proved in this paper is the following.

Theorem 3.1 Assume that the Carathéodry function a satisfies (6)-(9). Then there exists an entropy solution u of the problem (P).

Proposition 3.1 The entropy solution u of the problem (P) satisfies

$$u \in \mathcal{M}^r(\Omega, \omega) \quad \text{with } r = \eta(p - 1)(1 - \theta), \quad (13)$$

$$|\nabla u| \in \mathcal{M}^s(\Omega, \omega) \quad \text{with } s = \frac{\eta(p - 1)(1 - \theta)}{r + \theta(p - 1) + 1}, \quad (14)$$

where η is a constant such that $1 \leq \eta \leq \frac{N}{N-1} + \delta$.

4 Proof of the Main Result

4.1 Useful lemmas

Lemma 4.1 Let $u \in \mathcal{T}_0^{1,p}(\Omega, \omega)$, where $\omega \in \mathcal{A}_p$, $1 < p < \infty$. Let $1 \leq \lambda < p$, and suppose that u satisfies

$$\int_{\{|u| < k\}} |\nabla u|^p \omega dx \leq M k^\lambda, \quad \forall k > 0. \quad (15)$$

Then u belongs to $\mathcal{M}^r(\Omega, \omega)$ with $r = \eta(p - \lambda)$ (where $1 \leq \eta \leq \frac{N}{N-1} + \delta$). More precisely, there exists $C > 0$ such that

$$\Phi_u(k) \leq C M^\eta k^{-r}.$$

Proof. For $0 < \varepsilon \leq k$, we have $\{x \in \Omega : |u| \geq \varepsilon\} = \{x \in \Omega : |T_k(u)| \geq \varepsilon\}$. Thus

$$\begin{aligned} \mu(\{x \in \Omega : |u| > \varepsilon\}) &= \mu(\{x \in \Omega : |T_k(u)| \geq \varepsilon\}) \\ &= \int_{\{|T_k(u)| \geq \varepsilon\}} \omega(x) dx \\ &= \frac{1}{\varepsilon^{\eta p}} \int_{\{|T_k(u)| \geq \varepsilon\}} \varepsilon^{\eta p} \omega(x) dx \\ &\leq \frac{1}{\varepsilon^{\eta p}} \int_{\{|T_k(u)| \geq \varepsilon\}} |T_k(u)|^{\eta p} \omega(x) dx \\ &\leq \frac{1}{\varepsilon^{\eta p}} \|T_k(u)\|_{L^{\eta p}(\Omega, \omega)}^{\eta p}. \end{aligned}$$

By Theorem 2.2 and inequality (15), we get

$$\|T_k(u)\|_{L^{\eta p}(\Omega, \omega)} \leq C_\Omega \| |\nabla T_k(u)| \|_{L^p(\Omega, \omega)} \leq C_\Omega (M k^\lambda)^{1/p},$$

which implies that

$$\mu(\{x \in \Omega : |u| > \varepsilon\}) \leq \frac{1}{\varepsilon^{\eta p}} C(Mk^\lambda)^\eta.$$

Therefore, by taking $\varepsilon = k$, we have

$$\mu(\{x \in \Omega : |u| > k\}) \leq \frac{1}{k^{\eta p}} C(Mk^\lambda)^\eta = CM^\eta k^{-\eta(p-\lambda)} = CM^\eta k^{-r}.$$

Lemma 4.2 *Assume that the hypothesis in Lemma 4.1 holds true. Then $|\nabla u| \in \mathcal{M}^s(\Omega, \omega)$, where $s = pr/(r + \lambda)$ (with r as in Lemma 4.1). More precisely, there exists $C > 0$ such that*

$$\Phi_{\nabla u}(k) \leq CM^{(r+\eta\lambda)/(r+\lambda)} k^{-s}.$$

Proof. We set for every $k, \rho > 0$, the function

$$\Psi(k, \rho) = \mu(\{x \in \Omega : |\nabla u|^p > \rho, |u| > k\}).$$

It is clear that the function $\rho \rightarrow \Psi(k, \rho)$ is decreasing. Thus for $k > 0$ and $\rho > 0$, we obtain

$$\Psi(0, \rho) \leq \frac{1}{\rho} \int_0^\rho \Psi(0, s) ds \leq \Psi(k, 0) + \frac{1}{\rho} \int_0^\rho (\Psi(0, s) - \Psi(k, s)) ds. \tag{16}$$

From Lemma 4.1, we have

$$\Psi(k, 0) \leq CM^\eta k^{-r}. \tag{17}$$

Now, observe that

$$\begin{aligned} \Psi(0, s) - \Psi(k, s) &= \mu(\{x \in \Omega : |\nabla u|^p > s, |u| > 0\}) - \mu(\{x \in \Omega : |\nabla u|^p > s, |u| > k\}) \\ &= \mu(\{x \in \Omega : |\nabla u|^p > s, |u| \leq k\}). \end{aligned}$$

Hence, we get by using Proposition 6.24 in [15] that

$$\begin{aligned} \int_0^\infty (\Psi(0, s) - \Psi(k, s)) ds &= \int_0^\infty \mu(\{x \in \Omega : |\nabla u|^p > s, |u| \leq k\}) ds \\ &= \int_{\{|u| \leq k\}} |\nabla u(x)|^p \omega(x) dx \\ &\leq Mk^\lambda. \end{aligned} \tag{18}$$

Going back to (16) and using (17) and (18), we obtain

$$\Psi(0, \rho) \leq CM^\eta k^{-r} + \frac{1}{\rho} Mk^\lambda. \tag{19}$$

A minimization of the right-hand side of (19) in k gives

$$\Psi(0, \rho) \leq CM^{(r+\eta\lambda)/(r+\lambda)} \rho^{-r/(r+\lambda)}.$$

Setting $\rho = h^p$, we obtain

$$\Psi(0, h^p) = \mu(\{x \in \Omega : |\nabla u| > h\}) \leq CM^{(r+\eta\lambda)/(r+\lambda)} h^{-rp/(r+\lambda)} = CM^{(r+\eta\lambda)/(r+\lambda)} h^{-s},$$

where $s = rp/(r + \lambda)$.

Let $n \in \mathbb{N}$, and define, for u in $W_0^{1,p}(\Omega, \omega)$, the differential operator

$$A_n(u) = -\operatorname{div}[\omega(x)a(x, T_n(u), \nabla u)].$$

Lemma 4.3 *The operator A_n maps $W_0^{1,p}(\Omega, \omega)$ into its dual $W^{-1,p'}(\Omega, \omega^{1-p'})$. Moreover, A_n is bounded, pseudomonotone and coercive in the following sense:*

$$\frac{\langle A_n v, v \rangle}{\|v\|_{W_0^{1,p}(\Omega, \omega)}} \rightarrow +\infty \quad \text{if} \quad \|v\|_{W_0^{1,p}(\Omega, \omega)} \rightarrow +\infty, v \in W_0^{1,p}(\Omega, \omega).$$

Proof. By (8), we deduce that for every u in $W_0^{1,p}(\Omega, \omega)$,

$$\begin{aligned} \int_{\Omega} \left(\omega |a(x, T_n(u), \nabla u)| \right)^{p'} \omega^{1-p'} dx & \\ & \leq \int_{\Omega} \left(l_0 + l_1 |T_n(u)|^{p/p'} + l_2 |\nabla u|^{p/p'} \right)^{p'} \omega dx \\ & \leq C \left[\|l_0\|_{L^p(\Omega, \omega)}^{p'} + \left(C_{\Omega} \|l_1\|_{L^{\infty}(\Omega)}^{p'} + \|l_2\|_{L^{\infty}(\Omega)}^{p'} \right) \|u\|_{W_0^{1,p}(\Omega, \omega)}^p \right], \end{aligned}$$

which means that $\omega a(x, T_n(u), \nabla u)$ belongs to $(L^{p'}(\Omega, \omega^{1-p'}))^N$. Therefore

$$A_n(u) \in W^{-1,p'}(\Omega, \omega^{1-p'}), \quad \forall n \in \mathbb{N}.$$

Thanks to Hölder’s inequality and (8), we have for all $u, v \in W_0^{1,p}(\Omega, \omega)$,

$$\begin{aligned} |\langle A_n u, v \rangle| & \leq \left(C \left[\|l_0\|_{L^p(\Omega, \omega)}^{p'} + \left(C_{\Omega} \|l_1\|_{L^{\infty}(\Omega)}^{p'} + \|l_2\|_{L^{\infty}(\Omega)}^{p'} \right) \|u\|_{W_0^{1,p}(\Omega, \omega)}^p \right] \right)^{1/p'} \\ & \quad \times \|v\|_{W_0^{1,p}(\Omega, \omega)}. \end{aligned}$$

Thus A_n is bounded from $W_0^{1,p}(\Omega, \omega)$ to $W^{-1,p'}(\Omega, \omega^{1-p'})$.

For the coercivity, by using (6), we get for every $u \in W_0^{1,p}(\Omega, \omega)$,

$$\begin{aligned} \langle A_n u, u \rangle & = \int_{\Omega} \omega a(x, T_n(u), \nabla u) \cdot \nabla u dx \\ & \geq \int_{\Omega} \omega b(|T_n(u)|) |\nabla u|^p dx \\ & \geq b(n) \|u\|_{W_0^{1,p}(\Omega, \omega)}^p. \end{aligned}$$

Hence, the operator A_n is coercive.

It remains to show that A_n is pseudomonotone. We thus consider a sequence u_j in $W_0^{1,p}(\Omega, \omega)$ such that

$$\begin{cases} u_j \rightharpoonup u, & \text{in } W_0^{1,p}(\Omega, \omega), \\ A_n u_j \rightharpoonup \psi^n, & \text{in } W^{-1,p'}(\Omega, \omega^{1-p'}), \\ \limsup_{j \rightarrow \infty} \langle A_n u_j, u_j \rangle \leq \langle \psi^n, u \rangle. \end{cases} \tag{20}$$

We shall prove that

$$\psi^n = A_n u \text{ and } \langle A_n u_j, u_j \rangle \rightarrow \langle A_n u, u \rangle.$$

Firstly, since $W_0^{1,p}(\Omega, \omega) \hookrightarrow L^p(\Omega, \omega)$, one has

$$u_j \rightarrow u \quad \text{in } L^p(\Omega, \omega) \text{ for a subsequence denoted again by } (u_j)_j.$$

Due to the boundedness of the sequence $(u_j)_j$ in $W_0^{1,p}(\Omega, \omega)$, and using the growth assumption (8), we have that $\omega a(x, T_n(u_j), \nabla u_j)$ is bounded in $(L^{p'}(\Omega, \omega^{1-p'}))^N$.

Therefore, there exists a function $\phi \in L^{p'}(\Omega, \omega^{1-p'})$ such that

$$\omega a(x, T_n(u_j), \nabla u_j) \rightharpoonup \phi \text{ weakly in } (L^{p'}(\Omega, \omega^{1-p'}))^N. \quad (21)$$

It is clear that, for all $v \in W_0^{1,p}(\Omega, \omega)$,

$$\begin{aligned} \langle \psi^n, v \rangle &= \lim_{j \rightarrow +\infty} \langle A_n u_j, v \rangle \\ &= \lim_{j \rightarrow +\infty} \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla v dx \\ &= \int_{\Omega} \phi \cdot \nabla v dx. \end{aligned} \quad (22)$$

Hence, by hypotheses, we have

$$\begin{aligned} \limsup_{j \rightarrow +\infty} \langle A_n u_j, u_j \rangle &= \limsup_{j \rightarrow +\infty} \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla u_j dx \\ &\leq \int_{\Omega} \phi \cdot \nabla u dx. \end{aligned} \quad (23)$$

On the other hand, by (9),

$$[a(x, T_n(u_j), \nabla u_j) - a(x, T_n(u_j), \nabla u)] \cdot \nabla(u_j - u) > 0.$$

Hence

$$\begin{aligned} \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla u_j dx &> \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla u dx \\ &+ \int_{\Omega} \omega a(x, T_n(u_j), \nabla u) \cdot (\nabla u_j - \nabla u) dx. \end{aligned} \quad (24)$$

Using (21), we get

$$\liminf_{j \rightarrow +\infty} \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla u_j dx \geq \int_{\Omega} \phi \cdot \nabla u. \quad (25)$$

By using (23) and (25), we get

$$\lim_{j \rightarrow +\infty} \int_{\Omega} \omega a(x, T_n(u_j), \nabla u_j) \cdot \nabla u_j dx = \int_{\Omega} \phi \cdot \nabla u dx. \quad (26)$$

As a result of (22) and (26), we get

$$\langle A_n u_j, u_j \rangle \rightarrow \langle \psi^n, u \rangle \text{ as } j \rightarrow +\infty.$$

Yet, due to (26) and the strong convergence $\omega a(x, T_n(u_j), \nabla u) \rightarrow \omega a(x, T_n(u), \nabla u)$ in $(L^{p'}(\Omega, \omega^{1-p'}))^N$, we deduce that

$$\lim_{j \rightarrow +\infty} \int_{\Omega} (\omega a(x, T_n(u_j), \nabla u_j) - \omega a(x, T_n(u_j), \nabla u)) \cdot (\nabla u_j - \nabla u) dx = 0,$$

and so, by virtue of Lemma 3.2 in [1],

$$\nabla u_j \rightarrow \nabla u \text{ a.e. in } \Omega,$$

we deduce then that $\omega a(x, T_n(u_j), \nabla u_j)$ converges to $\omega a(x, T_n(u), \nabla u)$ weakly in $(L^{p'}(\Omega, \omega^{1-p'}))^N$. This implies that $\psi^n = A_n u$.

4.2 Approximate problem

We consider the sequence of approximate problems

$$(P_n) \begin{cases} A_n(u_n) = f_n - \operatorname{div} F & \text{in } \Omega, \\ u_n = 0 & \text{on } \partial\Omega, \end{cases}$$

where (f_n) is a sequence of functions in $C_0^\infty(\Omega)$ which is strongly convergent to f in $L^1(\Omega)$ such that $\|f_n\|_{L^1(\Omega)} \leq \|f\|_{L^1(\Omega)}$.

Since the source term $f_n - \operatorname{div}(F)$ belongs to the dual space $W^{-1,p'}(\Omega, \omega^{1-p'})$, in view of Lemma 4.3, the operator A_n is pseudomonotone, and by Theorem 2.7 in [19], there exists at least one solution $u_n \in W_0^{1,p}(\Omega, \omega)$ of problem (P_n) in the sense that

$$\int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla v \, dx = \int_{\Omega} f_n v \, dx + \int_{\Omega} F \cdot \nabla v \, dx \tag{27}$$

for every $v \in W_0^{1,p}(\Omega, \omega)$.

By choosing $v = T_k(u_n)$ in (27), we can take $n > k$, then applying (6) in the first term and using Hölder's then Young's inequalities in the last one, we obtain

$$b(k) \int_{\Omega} |\nabla T_k(u_n)|^p \omega \, dx \leq k \|f\|_{L^1(\Omega)} + \frac{b(k)}{2} \int_{\Omega} |\nabla T_k(u_n)|^p \omega \, dx + C_1 \int_{\Omega} \left| \frac{F}{\omega} \right|^{p'} \omega \, dx.$$

Thus we have the estimate

$$\int_{\Omega} |\nabla T_k(u_n)|^p \omega \, dx \leq C_2 (1 + k)^{\theta(p-1)+1}. \tag{28}$$

4.3 Local convergence of u_n in μ -measure

Combining the previous estimation with Lemma 4.1, we conclude that $u_n \in \mathcal{M}^r(\Omega, \omega)$ with $r = \eta(p-1)(1-\theta)$, we also have that $\mu(\{x \in \Omega : |u_n(x)| > k\})$ is bounded uniformly in n for every $k > 0$, that is,

$$\mu(\{x \in \Omega : |u_n(x)| > k\}) \leq C k^{-r}, \tag{29}$$

and by Lemma 4.2, we have that the sequence $(|\nabla u_n|)$ is bounded in $\mathcal{M}^s(\Omega, \omega)$ with $s = \frac{pr}{r+\theta(p-1)+1}$. Our objective now is to prove that $u_n \rightarrow u$ locally in μ -measure. For that, we will use the same reasoning as in the proof of Theorem 2.11 in [10] (see also Theorem 6.1 in [7]). Let $\rho > 0$, we have

$$\{|u_n - u_m| > \rho\} \subset \{|u_n| > k\} \cup \{|u_m| > k\} \cup \{|T_k(u_n) - T_k(u_m)| > \rho\}$$

so that

$$\mu(\{|u_n - u_m| > \rho\}) \leq \mu(\{|u_n| > k\}) + \mu(\{|u_m| > k\}) + \mu(\{|T_k(u_n) - T_k(u_m)| > \rho\}). \quad (30)$$

Fix $\varepsilon > 0$. From (29), there exists $k_\varepsilon = k$ such that

$$\mu(\{|u_n| > k\}) + \mu(\{|u_m| > k\}) \leq \frac{\varepsilon}{2}.$$

Since $(\nabla T_k(u_n))$ is bounded in $L^p_{loc}(\Omega, \omega)$ for all $k > 0$ and $T_k(u_n)$ belongs to $W_0^{1,p}(\Omega, \omega)$, we can assume that $(T_k(u_n))$ is a Cauchy sequence in $L^q(\Omega \cap B_R, \omega)$ for any $q < p\eta = p^*$ and any $R > 0$ and

$$T_k(u_n) \rightarrow T_k(u) \quad \text{in } L^p_{loc}(\Omega, \omega) \text{ and a.e.}$$

Then

$$\begin{aligned} \mu(\{|T_k(u_n) - T_k(u_m)| > \rho\} \cap B_R) &= \int_{\{|T_k(u_n) - T_k(u_m)| > \rho\} \cap B_R} \omega dx \\ &\leq k^{-q} \int_{\Omega \cap B_R} |T_k(u_n) - T_k(u_m)|^q \omega dx \leq \frac{\varepsilon}{2} \end{aligned}$$

for all $n, m \geq n_0(k, \rho, R)$. It follows that (u_n) is a Cauchy sequence in μ -measure. As a consequence, there exist a function u and a subsequence, still denoted by u_n , such that

$$u_n \rightarrow u \quad \mu\text{-a.e. in } \Omega, \quad (31)$$

then by Remark 2.1, one has

$$u_n \rightarrow u \quad \text{a.e. in } \Omega. \quad (32)$$

Using (31) and (28), we have

$$\begin{aligned} T_k(u_n) &\rightharpoonup T_k(u) \text{ weakly in } W_0^{1,p}(\Omega, \omega) \text{ for every } k > 0, \\ T_k(u_n) &\rightarrow T_k(u) \text{ strongly in } L^p(\Omega, \omega) \text{ and } \mu\text{-a.e. in } \Omega \text{ for every } k > 0. \end{aligned} \quad (33)$$

Hence, $T_k(u) \in W_0^{1,p}(\Omega, \omega)$.

Furthermore, by the weak lower semicontinuity of the norm $W_0^{1,p}(\Omega, \omega)$, estimate (28) still holds for u , that is,

$$\int_{\Omega} |\nabla T_k(u)|^p \omega dx \leq C(1+k)^{\theta(p-1)+1}, \quad \forall k > 0.$$

Applying again Lemma 4.1 and Lemma 4.2, we find that $u \in \mathcal{M}^r(\Omega, \omega)$ and $|\nabla u| \in \mathcal{M}^s(\Omega, \omega)$.

4.4 Strong convergence of truncations

Our aim now is prove that

$$T_k(u_n) \rightarrow T_k(u) \quad \text{strongly in } W_0^{1,p}(\Omega, \omega) \text{ for all } k > 0. \quad (34)$$

For $n > k$, we write

$$\begin{aligned} I(n) &= \int_{\Omega} \omega[a(x, T_k(u_n), \nabla T_k(u_n)) - a(x, T_k(u_n), \nabla T_k(u))] \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &= \int_{\Omega} \omega a(x, T_k(u_n), \nabla T_k(u_n)) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &\quad - \int_{\Omega} \omega a(x, T_k(u_n), \nabla T_k(u)) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &= I_1(n) - I_2(n). \end{aligned}$$

Keeping in mind that (6) implies that $a(x, s, 0) = 0$, we get

$$\begin{aligned} I_1(n) &= \int_{\{|u_n| < k\}} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &= \int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &\quad - \int_{\{|u_n| \geq k\}} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx. \end{aligned}$$

Observing that $\nabla T_k(u_n) = 0$ on the set $\{|u_n| \geq k\}$, we obtain

$$\begin{aligned} I(n) &= \int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &\quad + \int_{\{|u_n| \geq k\}} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla T_k(u)dx \\ &\quad - \int_{\Omega} \omega a(x, T_k(u_n), \nabla T_k(u)) \cdot \nabla(T_k(u_n) - T_k(u))dx. \end{aligned}$$

We take $T_k(u_n) - T_k(u)$ as a test function in (27) and we get

$$\begin{aligned} &\int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx \\ &= \int_{\Omega} f_n(T_k(u_n) - T_k(u))dx + \int_{\Omega} F \cdot \nabla(T_k(u_n) - T_k(u))dx. \end{aligned}$$

By the almost convergence of u_n and using the strong convergence of f_n in $L^1(\Omega)$, we obtain

$$\lim_{n \rightarrow \infty} \int_{\Omega} f_n(T_k(u_n) - T_k(u))dx = 0.$$

Also, since F belongs to $(L^{p'}(\Omega, \omega^{1-p'}))^N$ and by (33), we obtain

$$\lim_{n \rightarrow \infty} \int_{\Omega} F \cdot \nabla(T_k(u_n) - T_k(u))dx = 0.$$

Therefore

$$\lim_{n \rightarrow \infty} \int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla(T_k(u_n) - T_k(u))dx = 0. \tag{35}$$

Using the growth assumption (8), for every u in $W_0^{1,p}(\Omega, \omega)$, we have that $\omega|a(x, T_n(u), \nabla u)|$ is bounded $L^{p'}(\Omega, \omega^{1-p'})$. Therefore, it converges weakly to some g in $L^{p'}(\Omega, \omega^{1-p'})$ and we have

$$\lim_{n \rightarrow \infty} \int_{\{|u_n| \geq k\}} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla T_k(u) dx = \int_{\{|u| \geq k\}} g \cdot \nabla T_k(u) = 0. \tag{36}$$

By virtue of Vitali’s theorem, we obtain

$$\omega(x)a(x, T_n(u_n), \nabla T_k(u)) \rightarrow \omega(x)a(x, u, \nabla T_k(u)) \text{ strongly in } \left(L^{p'}(\Omega, \omega^{1-p'})\right)^N.$$

It follows from (33) that

$$\lim_{n \rightarrow \infty} \int_{\Omega} \omega a(x, T_k(u_n), \nabla T_k(u)) \cdot \nabla (T_k(u_n) - T_k(u)) dx = 0. \tag{37}$$

Bringing together (35)-(37), we conclude that

$$\lim_{n \rightarrow \infty} I(n) = 0.$$

Now we can apply Lemma 3.2 in [1] to get (34). Hence, for every fixed $k > 0$, we have

$$\omega a(x, T_k(u_n), \nabla T_k(u_n)) \rightarrow \omega a(x, T_k(u), \nabla T_k(u)) \text{ in } (L^{p'}(\Omega, \omega^{1-p'}))^N. \tag{38}$$

4.5 Passage to the limit

We will now demonstrate that u satisfies (12). Let $v \in W_0^{1,p}(\Omega, \omega) \cap L^\infty(\Omega)$. Testing (27) with $\psi_n = T_k(u_n - v)$, we get

$$\int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla \psi_n dx = \int_{\Omega} f_n \psi_n dx + \int_{\Omega} F \cdot \nabla \psi_n dx.$$

If $M = k + \|v\|_{L^\infty(\Omega)}$ and $n > M$, then

$$\begin{aligned} \int_{\Omega} \omega a(x, T_n(u_n), \nabla u_n) \cdot \nabla T_k(u_n - v) dx &= \int_{\Omega} \omega a(x, T_n(u_n), \nabla T_M(u_n)) \cdot \nabla T_k(u_n - v) dx \\ &= \int_{\Omega} \omega a(x, T_M(u_n), \nabla T_M(u_n)) \cdot \nabla T_k(u_n - v) dx. \end{aligned}$$

Thus, we can write

$$\int_{\Omega} \omega a(x, T_M(u_n), \nabla T_M(u_n)) \cdot \nabla T_k(u_n - v) dx = \int_{\Omega} f_n T_k(u_n - v) dx + \int_{\Omega} F \cdot \nabla T_k(u_n - v) dx. \tag{39}$$

Hence we can pass to the limit as n tends to infinity, using (33) and (38), we obtain

$$\int_{\Omega} \omega a(x, u, \nabla u) \cdot \nabla T_k(u - v) dx = \int_{\Omega} f T_k(u - v) dx + \int_{\Omega} F \cdot \nabla T_k(u - v) dx$$

for every $v \in W_0^{1,p}(\Omega, \omega) \cap L^\infty(\Omega)$ and for every $k > 0$.

Example 4.1 We put ourselves in the situation $N = 2$, $p = 3$. Let $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$, the weight function $\omega(x, y) = (x^2 + y^2)^{-1/2}$ is such that $\omega \in \mathcal{A}_3$. And the function $f(x, y) = \frac{\cos(xy)}{(x^2 + y^2)^{1/3}} \in L^1(\Omega)$ and $F(x, y) = \left((x^2 + y^2) \sin(xy), (x^2 + y^2)^{-1/3} \cos(xy) \right) \in \left[L^{\frac{3}{2}}(\Omega, \omega^{-\frac{1}{2}}) \right]^2$. The Carathéodory function is defined as follows: $a : \Omega \times \mathbb{R} \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$, $a((x, y), s, \xi) = \frac{\xi}{\sqrt{1+|s|}}$. Therefore, by virtue of Theorem 3.1, the problem

$$\begin{cases} -\operatorname{div}[\omega(x, y)a((x, y), u, \nabla u)] = f(x, y) - \operatorname{div} F(x, y) & \text{in } \Omega, \\ u(x, y) = 0, & \text{on } \partial\Omega, \end{cases}$$

has an entropy solution.

5 Conclusion

Through this work, we were able to demonstrate the existence and regularity of solutions for some nonlinear elliptic equations of the form $-\operatorname{div}[\omega(x)a(x, u, \nabla u)] = f - \operatorname{div} F$, in the framework of the weighted Sobolev spaces. The novelty here is that the operator $A(u) = -\operatorname{div}[\omega(x)a(x, u, \nabla u)]$ is a nonlinear degenerate elliptic operator in the sense that the Carathéodory function $a(\cdot, \cdot, \cdot)$ satisfies the degenerate coercivity (6) instead of the case where A is a uniformly elliptic operator, that is, when b is the constant function. Let us point out that this work can be seen as a generalization of the work in [11] and [18].

References

- [1] Y. Akdim, E. Azroul and A. Benkirane. Existence of solutions for quasilinear degenerate elliptic equations. *Electronic Journal of Differential Equations* **2001** (71) (2001) 1–19.
- [2] Y. Akdim, E. Azroul and M. Rhoudaf. Entropy solutions of nonlinear elliptic equations with measurable boundary conditions and without strict monotonicity conditions. *Eur. J. Pure Appl. Math.* **1** (4) (2008) 56–71.
- [3] Y. Akdim, M. Rhoudaf and A. Salmani. Existence of solution for nonlinear anisotropic degenerated elliptic unilateral problems. *Nonlinear Dynamics and Systems Theory* **18** (3) (2018) 213–224.
- [4] A. Alvino, L. Boccardo, V. Ferone, L. Orsina and G. Trombetti. Existence results for nonlinear elliptic equations with degenerate coercivity. *Annali di Matematica* **182** (2003) 53–79.
- [5] K. Ammar and H. Redwane. Renormalized solutions for nonlinear degenerate elliptic problems with L^1 -data. *Rev. Mat. Complut.* **22** (2009) 37–52.
- [6] E. Azroul, M. B. Benboubker and S. Ouaro. The obstacle problem associated with nonlinear elliptic equations in generalized Sobolev spaces. *Nonlinear Dynamics and Systems Theory* **14** (3) (2014) 224–243.
- [7] P. Bélinan, L. Boccardo, T. Gallouët, R. Gariepy, M. Pierre and J.L. Vasquez. An L^1 theory of existence and uniqueness of solutions of nonlinear elliptic equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **22** (2) (1995) 241–273.
- [8] S. Bentouati, M. Islem Soualhi, A. Abdellah El-Hadj and H. Yeklef. Design and Analysis of Continuous Positive Airway Pressure Valve Using a 3D Printing and Computational Fluid Dynamic. *Nonlinear Dynamics and Systems Theory* **21** (3) (2021) 229–237.

- [9] L. Boccardo, F. Murat and J. P. Puel. Existence of bounded solutions for nonlinear elliptic unilateral problems. *Ann. Mat. Pura Appl.* **152** (1988) 183–196.
- [10] A. C. Cavalheiro. The solvability of Dirichlet problem for a class of degenerate elliptic equations with L^1 -data. *Applicable Analysis* **85** (8) (2006) 941–961.
- [11] A. C. Cavalheiro. Existence of entropy solutions for degenerate quasilinear elliptic equations in L^1 . *Communications in Mathematics* **22** (2014) 57–69.
- [12] A. C. Cavalheiro. Existence of solutions for some degenerate quasilinear elliptic equations. *Le Matematiche* **LXIII** (II) (2008) 101–112.
- [13] S.-K. Chua, S. Rodney and R. Wheeden. A compact embedding theorem for generalized Sobolev spaces. *Pacific J. Math.* **265** (1) (2013) 17–57.
- [14] E. Fabes, C. Kenig and R. Serapioni. The local regularity of solutions of degenerate elliptic equations. *Comm. PDEs.* **7** (1982) 77–116.
- [15] G. B. Folland. *Real Analysis*. Wiley-Interscience, New York, 1984.
- [16] J. Heinonen, T. Kilpeläinen and O. Martio *Nonlinear Potential Theory of Degenerate Elliptic Equations*. Oxford Math. Monographs, Clarendon Press, 1993.
- [17] A. Kufner, O. John and S. Fučík. *Function Spaces*. Noordhoff International Publishing, Leyden, 1977.
- [18] C. Leone and A. Porretta. Entropy solutions for nonlinear elliptic equations in L^1 . *Nonlinear Anal.* **32** (1998) 325–334.
- [19] J. L. Lions. *Quelques méthodes de résolution de problèmes aux limites non linéaires*. Dunod, Gauthier-Villars, Paris, 1969.
- [20] K. Oktafianto, A. Z. Arifin, E. F. Kurniawati, T. Tafrikan, T. Herlambang and F. Yudianto. Tsunami Wave Simulation in the Presense of a Barrier. *Nonlinear Dynamics and Systems Theory* **23** (1) (2023) 89–78.



An Algorithm for Solving First-Kind Two-Dimensional Volterra Integral Equations Using Collocation Method

F. Birem, A. Boulmerka, H. Laib* and C. Hennous

*Laboratory of Mathematics and their interactions,
University Center Abdelhafid Boussouf, Mila, Algeria*

Received: July 2, 2023; Revised: November 9, 2023

Abstract: The proposed study presents a collocation method to address two types of two-dimensional Volterra integral equations (2D VIEs): nonlinear first kind and linear second kind. The nonlinear equations of the first kind are transformed into the linear second kind equations. A convergent algorithm using the Taylor polynomials is developed to construct a collocation solution that approximates the solution of 2D VIEs of the second kind. The study includes various numerical examples to compare the results of different methods and demonstrate the proposed approach's accuracy and validity. This validation procedure plays a pivotal role in nonlinear dynamics and systems theory, establishing the reliability and stability of novel methods.

Keywords: *two-dimensional Volterra integral equations of the first and second kind; collocation method; Taylor polynomials; error analysis.*

Mathematics Subject Classification (2010): 45D05, 45L05, 65R20, 70K99, 93A99.

1 Introduction

The nonlinear 2D VIE of the first kind, which includes an unknown function u , can be represented in a standard form as follows:

$$\int_0^\tau \int_0^z \kappa(\tau, z, t, s)H(u(t, s))dsdt = f(\tau, z), \quad (\tau, z) \in D, \quad (1)$$

where D is a subset of \mathbb{R}^2 defined as $[0, T] \times [0, Z]$, f and κ are smooth functions on their corresponding domains. Additionally, H is a continuous inverse function that is

* Corresponding author: <mailto:hafida.laib@gmail.com>

nonlinear with respect to u . To solve equation (1), we substitute $\omega(t, s) = H(u(t, s))$ and obtain the linear equation

$$\int_0^\tau \int_0^z \kappa(\tau, z, t, s) \omega(t, s) ds dt = f(\tau, z), \quad (\tau, z) \in D. \quad (2)$$

To obtain an approximation for ω , we transform the first kind VIE (2) into the second kind VIE (3) by differentiating equation (2) with respect to z and τ . This transformation technique aligns with the strategies used in various nonlinear systems analyses. It allows researchers and practitioners to simplify the problem while retaining essential characteristics, thus aiding the analysis and comprehension of complex nonlinear systems. This conversion technique is effective only under the conditions that $f(\tau, 0) = f(0, z) = 0$ and $\kappa(\tau, z, \tau, z) \neq 0$ for $(\tau, z) \in D$, and results in a linear 2D VIE of the following form:

$$\begin{aligned} \omega(\tau, z) = & g(\tau, z) + \int_0^\tau \kappa_1(\tau, z, t) \omega(t, z) dt + \int_0^z \kappa_2(\tau, z, s) \omega(\tau, s) ds \\ & + \int_0^\tau \int_0^z \kappa_3(\tau, z, t, s) \omega(t, s) ds dt, \quad (\tau, z) \in D, \end{aligned} \quad (3)$$

where the functions g , κ_1 , κ_2 and κ_3 are given smooth functions defined on their corresponding domains by

$$\begin{aligned} \kappa_1(\tau, z, t) &:= -\frac{\partial \kappa}{\partial \tau}(\tau, z, t, z) / \kappa(\tau, z, \tau, z), \quad \kappa_2(\tau, z, s) := -\frac{\partial \kappa}{\partial z}(\tau, z, \tau, s) / \kappa(\tau, z, \tau, z), \\ \kappa_3(\tau, z, t, s) &:= -\frac{\partial^2 \kappa}{\partial \tau \partial z}(\tau, z, t, s) / \kappa(\tau, z, \tau, z), \quad g(\tau, z) := \frac{\partial^2 f}{\partial \tau \partial z}(\tau, z) / \kappa(\tau, z, \tau, z). \end{aligned}$$

The solution for (1) can be approximated as $H^{-1}(\omega(t, s)) = u(t, s)$. The existence and uniqueness of the solution for equation (1), using $H(u(t, s)) = \omega(t, s)$ and equation (3), have been proposed in [1].

The applications of VIEs extend to a diverse range of fields, including physical and engineering domains, population dynamics, economics and finance, fluid dynamics, and heat transfer. By providing a method to efficiently and accurately solve nonlinear integral equations, we contribute to the modeling and analyzing complex nonlinear systems. Our algorithm's ability to handle nonlinearity is directly relevant to nonlinear dynamics, as it provides a means to understand and predict the behaviors of such systems. However, solving these equations has motivated mathematicians to develop reliable methods for their solutions [2–10]. In [2], a method based on applying 2D block-pulse functions was utilized to solve nonlinear 2D VIEs of the first kind. An Euler-type technique was discussed in [1]. The Chelyshkov polynomial strategy for solving (1) was considered in [6]. In [7], the Tau technique was employed to approximate the solution of (2). Nemati and Ordokhani [8] used operational matrices of Legendre polynomials to approximate the solution of a class of (1), specifically when $H = u^n$ and n is a positive integer. In [9], a multi-step method was implemented for the numerical solution of nonlinear 2D VIEs of the first kind. A special case of (3) for $\kappa_1 = \kappa_2 = 0$ is considered in [10].

This paper presents an extension of the collocation method proposed in previous works such as [11–14], to solve equations (1) and (3) by utilizing Taylor's theorem in two variables. Additionally, the method is straightforward to implement, and the iterative formulas used to obtain the approximate solution do not require solving any algebraic equations. This showcases the possibility of our method to address broader challenges in

nonlinear dynamics that involve integral equations, paving the way for its adaptation in various related problem domains.

The remainder of this paper is structured as follows: the next section outlines our approach to approximating the solution of equation (3) through the Taylor polynomials. Section 3 focuses on the convergence analysis. To demonstrate the validity of our theoretical results, we provide several numerical examples in Section 4. Finally, in Section 5, we present our conclusion and offer suggestions for future research.

2 Description of the Method

In this section, we approximate solutions of 2D VIE (3) in the space

$S_{p-1,p-1}^{(-1)}(\Pi_{N,M}) = \{v : v_{n,m} = v|_{D_{n,m}} \in \pi_{p-1,p-1}, n = 0, 1, \dots, N - 1; m = 0, 1, \dots, M - 1\}$ of the real bivariate polynomial spline functions of degree (at most) $p - 1$ in τ and z . Its dimension is NMp^2 . Here, $\Pi_N = \{\tau_i = ih, i = 0, 1, \dots, N\}$ and $\Pi_M = \{z_j = jk, j = 0, 1, \dots, M\}$ denote, respectively, uniform partitions of the intervals $[0, T]$ and $[0, Z]$ with the step sizes given by $h = \frac{T}{N}$ and $k = \frac{Z}{M}$. These partitions defined a grid for D $\Pi_{N,M} = \Pi_N \times \Pi_M = \{(\tau_n, z_m), 0 \leq n \leq N, 0 \leq m \leq M\}$. Set the subintervals $\sigma_n = [\tau_n; \tau_{n+1}), n = 0, 1, \dots, N - 2; \sigma_{N-1} = [\tau_{N-1}, \tau_N]$, $\delta_m = [z_m; z_{m+1}), m = 0, 1, \dots, M - 2; \delta_{M-1} = [z_{M-1}, z_M]$, and $D_{n,m} := \sigma_n \times \delta_m$ for all $n = 0, 1, \dots, N - 1; m = 0, 1, \dots, M - 1$.

To defined the collocation solution, we use the Taylor polynomial on each rectangle $D_{n,m}; n = 0, 1, \dots, N - 1; m = 0, 1, \dots, M - 1$. Note that the solution ω of (3) is known at the point $(0, 0)$: $\omega(0, 0) = g(0, 0)$.

2.1 Taylor collocation solution in $D_{0,0}$

We approximate ω in the rectangle $D_{0,0}$ by the polynomial

$$v_{0,0}(\tau, z) = \sum_{i+j=0}^{p-1} \frac{1}{i!j!} \frac{\partial^{i+j}\omega(0,0)}{\partial\tau^i\partial z^j} \tau^i z^j, \quad (\tau, z) \in D_{0,0}, \tag{4}$$

where $\frac{\partial^{i+j}\omega(0,0)}{\partial\tau^i\partial z^j}$ is the exact value of $\frac{\partial^{i+j}\omega}{\partial\tau^i\partial z^j}$ at the point $(0, 0)$. We differentiate equation (3) j times with respect to z and i times with respect to τ , we obtain

$$\begin{aligned} \frac{\partial^{i+j}\omega(0,0)}{\partial\tau^i\partial z^j} &= \partial_1^{(i)}\partial_2^{(j)}g(0,0) \\ &+ \sum_{l=0}^j \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{j}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial\tau^{q-\eta}} \left[\partial_1^{(i-1-q)}\partial_2^{(j-l)}\kappa_1(\tau, z, \tau) \right]_{z=0}^{\tau=0} \frac{\partial^{\eta+l}\omega(0,0)}{\partial\tau^\eta\partial z^l} \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{\eta=0}^i \binom{r}{l} \binom{i}{\eta} \frac{\partial^{i-\eta}}{\partial\tau^{i-\eta}} \left[\frac{\partial^{r-l}}{\partial z^{r-l}} \left(\partial_2^{(j-1-r)}\kappa_2(\tau, z, z) \right) \right]_{z=0}^{\tau=0} \frac{\partial^{\eta+l}\omega(0,0)}{\partial\tau^\eta\partial z^l} \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{r}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial\tau^{q-\eta}} \left[\frac{\partial^{i-1-q}}{\partial\tau^{i-1-q}} \Big|_{t=\tau} \left(\frac{\partial^{r-l}}{\partial z^{r-l}} \left[\partial_2^{(j-1-r)}\kappa_3(\tau, z, t, z) \right] \right) \right]_{z=0}^{\tau=0} \\ &\times \frac{\partial^{\eta+l}\omega(0,0)}{\partial\tau^\eta\partial z^l}. \end{aligned}$$

2.2 Taylor collocation solution in $D_{n,0}$

We approximate ω in the rectangles $D_{n,0}$, $n = 1, \dots, N - 1$ by the polynomials

$$v_{n,0}(\tau, z) = \sum_{i+j=0}^{p-1} \frac{1}{i!j!} \frac{\partial^{i+j} \hat{v}_{n,0}(\tau_n, 0)}{\partial \tau^i \partial z^j} (\tau - \tau_n)^i z^j, \quad (\tau, z) \in D_{n,0}, \quad (5)$$

where $\hat{v}_{n,0}$ is the exact solution of the integral equation

$$\begin{aligned} \hat{v}_{n,0}(\tau, z) &= g(\tau, z) + \int_0^z \kappa_2(\tau, z, s) \hat{v}_{n,0}(\tau, s) ds \\ &+ \sum_{\xi=0}^{n-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \kappa_1(\tau, z, t) v_{\xi,0}(t, z) dt + \int_{\tau_n}^{\tau} \kappa_1(\tau, z, t) \hat{v}_{n,0}(t, z) dt \\ &+ \sum_{\xi=0}^{n-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \int_0^z \kappa_3(\tau, z, t, s) v_{\xi,0}(t, s) ds dt + \int_{\tau_n}^{\tau} \int_0^z \kappa_3(\tau, z, t, s) \hat{v}_{n,0}(t, s) ds dt. \end{aligned} \quad (6)$$

We differentiate (6) j times with respect to z and i times with respect to τ , we obtain

$$\begin{aligned} \frac{\partial^{i+j} \hat{v}_{n,0}(\tau_n, 0)}{\partial \tau^i \partial z^j} &= \partial_1^{(i)} \partial_2^{(j)} g(\tau_n, 0) \\ &+ \sum_{\xi=0}^{n-1} \sum_{l=0}^j \binom{j}{l} \int_{\tau_\xi}^{\tau_{\xi+1}} \left[\partial_1^{(i)} \partial_2^{(j-l)} \kappa_1(\tau, z, t) \right]_{z=0}^{\tau=\tau_n} \frac{\partial^l v_{\xi,0}(t, 0)}{\partial z^l} dt \\ &+ \sum_{l=0}^j \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{j}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial \tau^{q-\eta}} \left[\partial_1^{(i-1-q)} \partial_2^{(j-l)} \kappa_1(\tau, z, \tau) \right]_{z=0}^{\tau=\tau_n} \frac{\partial^{\eta+l} \hat{v}_{n,0}(\tau_n, 0)}{\partial \tau^\eta \partial z^l} \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{\eta=0}^i \binom{r}{l} \binom{i}{\eta} \frac{\partial^{i-\eta}}{\partial \tau^{i-\eta}} \left[\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_2(\tau, z, z)] \right]_{z=0}^{\tau=\tau_n} \frac{\partial^{\eta+l} \hat{v}_{n,0}(\tau_n, 0)}{\partial \tau^\eta \partial z^l} \\ &+ \sum_{\xi=0}^{n-1} \sum_{r=0}^{j-1} \sum_{l=0}^r \binom{r}{l} \int_{\tau_\xi}^{\tau_{\xi+1}} \frac{\partial^i}{\partial \tau^i} \left[\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_3(\tau, z, t, z)] \right]_{z=0}^{\tau=\tau_n} \frac{\partial^l v_{\xi,0}(t, 0)}{\partial z^l} dt \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{r}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial \tau^{q-\eta}} \left[\frac{\partial^{i-1-q}}{\partial \tau^{i-1-q}} \Big|_{t=\tau} \left(\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_3(\tau, z, t, z)] \right) \right]_{z=0}^{\tau=\tau_n} \\ &\times \frac{\partial^{\eta+l} \hat{v}_{n,0}(\tau_n, 0)}{\partial \tau^\eta \partial z^l}. \end{aligned}$$

2.3 Taylor collocation solution in $D_{n,m}$

We approximate ω by $v_{n,m}$ in $D_{n,m}$, $n = 0, 1, \dots, N - 1$ and $m = 1, \dots, M - 1$, so that

$$v_{n,m}(\tau, z) = \sum_{i+j=0}^{p-1} \frac{1}{i!j!} \frac{\partial^{i+j} \hat{v}_{n,m}(\tau_n, z_m)}{\partial \tau^i \partial z^j} (\tau - \tau_n)^i (z - z_m)^j, \quad (\tau, z) \in D_{n,m}, \quad (7)$$

where $\hat{v}_{n,m}$ is the exact solution of the integral equation

$$\begin{aligned} \hat{v}_{n,m}(\tau, z) &= g(\tau, z) + \sum_{\xi=0}^{n-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \kappa_1(\tau, z, t) v_{\xi,m}(t, z) dt + \int_{\tau_n}^{\tau} \kappa_1(\tau, z, t) \hat{v}_{n,m}(t, z) dt \\ &+ \sum_{\rho=0}^{m-1} \int_{z_\rho}^{z_{\rho+1}} \kappa_2(\tau, z, s) v_{n,\rho}(\tau, s) ds + \int_{z_m}^z \kappa_2(\tau, z, s) \hat{v}_{n,m}(\tau, s) ds \\ &+ \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \int_{z_\rho}^{z_{\rho+1}} \kappa_3(\tau, z, t, s) v_{\xi,\rho}(t, s) ds dt \\ &+ \sum_{\xi=0}^{n-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \int_{z_m}^z \kappa_3(\tau, z, t, s) v_{\xi,m}(t, s) ds dt + \sum_{\rho=0}^{m-1} \int_{\tau_n}^{\tau} \int_{z_\rho}^{z_{\rho+1}} \kappa_3(\tau, z, t, s) v_{n,\rho}(t, s) ds dt \\ &+ \int_{\tau_n}^{\tau} \int_{z_m}^z \kappa_3(\tau, z, t, s) \hat{v}_{n,m}(t, s) ds dt. \end{aligned} \tag{8}$$

We differentiate (8) j times with respect to z and i times with respect to τ , we obtain

$$\begin{aligned} \frac{\partial^{i+j} \hat{v}_{n,m}(\tau_n, z_m)}{\partial \tau^i \partial z^j} &= \partial_1^{(i)} \partial_2^{(j)} g(\tau_n, z_m) \\ &+ \sum_{\xi=0}^{n-1} \sum_{l=0}^j \binom{j}{l} \int_{\tau_\xi}^{\tau_{\xi+1}} \partial_1^{(i)} \partial_2^{(j-l)} \kappa_1(\tau_n, z_m, t) \frac{\partial^l v_{\xi,m}(t, z_m)}{\partial z^l} dt \\ &+ \sum_{l=0}^j \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{j}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial \tau^{q-\eta}} \left[\partial_1^{(i-1-q)} \partial_2^{(j-l)} \kappa_1(\tau, z, \tau) \right]_{z=z_m}^{\tau=\tau_n} \frac{\partial^{\eta+l} \hat{v}_{n,m}(\tau_n, z_m)}{\partial \tau^\eta \partial z^l} \\ &+ \sum_{\rho=0}^{m-1} \sum_{\eta=0}^i \binom{i}{\eta} \int_{z_\rho}^{z_{\rho+1}} \partial_1^{(i-\eta)} \partial_2^{(j)} \kappa_2(\tau_n, z_m, s) \frac{\partial^\eta v_{n,\rho}(\tau_n, s)}{\partial \tau^\eta} ds \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{\eta=0}^i \binom{r}{l} \binom{i}{\eta} \frac{\partial^{i-\eta}}{\partial \tau^{i-\eta}} \left[\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_2(\tau, z, z)] \right]_{z=z_m}^{\tau=\tau_n} \frac{\partial^{\eta+l} \hat{v}_{n,m}(\tau_n, z_m)}{\partial \tau^\eta \partial z^l} \\ &+ \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} \int_{\tau_\xi}^{\tau_{\xi+1}} \int_{z_\rho}^{z_{\rho+1}} \partial_1^{(i)} \partial_2^{(j)} \kappa_3(\tau_n, z_m, t, s) v_{\xi,\rho}(t, s) ds dt \\ &+ \sum_{\xi=0}^{n-1} \sum_{r=0}^{j-1} \sum_{l=0}^r \binom{r}{l} \int_{\tau_\xi}^{\tau_{\xi+1}} \frac{\partial^i}{\partial \tau^i} \left[\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_3(\tau, z, t, z)] \right]_{z=z_m}^{\tau=\tau_n} \frac{\partial^l v_{\xi,m}(t, z_m)}{\partial z^l} dt \\ &+ \sum_{\rho=0}^{m-1} \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{q}{\eta} \int_{z_\rho}^{z_{\rho+1}} \frac{\partial^{q-\eta}}{\partial \tau^{q-\eta}} \left[\partial_1^{(i-1-q)} \partial_2^{(j)} \kappa_3(\tau, z, \tau, s) \right]_{z=z_m}^{\tau=\tau_n} \frac{\partial^\eta v_{n,\rho}(\tau_n, s)}{\partial \tau^\eta} ds \\ &+ \sum_{r=0}^{j-1} \sum_{l=0}^r \sum_{q=0}^{i-1} \sum_{\eta=0}^q \binom{r}{l} \binom{q}{\eta} \frac{\partial^{q-\eta}}{\partial \tau^{q-\eta}} \left[\frac{\partial^{i-1-q}}{\partial \tau^{i-1-q}} \Big|_{t=\tau} \left(\frac{\partial^{r-l}}{\partial z^{r-l}} [\partial_2^{(j-1-r)} \kappa_3(\tau, z, t, z)] \right) \right]_{z=z_m}^{\tau=\tau_n} \\ &\times \frac{\partial^{\eta+l} \hat{v}_{n,m}(\tau_n, z_m)}{\partial \tau^\eta \partial z^l}. \end{aligned}$$

3 Study of Convergence and Error of the Numerical Method

We consider the space $L^\infty(D)$ with the norm

$$\|\varphi\|_{L^\infty(D)} = \inf \{C \in \mathbb{R} : |\varphi(\tau, z)| \leq C \text{ for a.e. } (\tau, z) \in D\} < \infty.$$

The following lemmas will be used in proving the convergence of the presented method.

Lemma 3.1 (*Gronwall-type inequality [2]*) Let $\omega(\tau, z)$ and $p(\tau, z)$ be non-negative continuous functions in $\Omega = [a, b] \times [c, d]$, and let $p(\tau, z)$ be nondecreasing in each of the variables in Ω and satisfy the following inequality:

$$\omega(\tau, z) \leq p(\tau, z) + \kappa \int_a^\tau \omega(t, z) dt + \kappa \int_c^z \omega(\tau, s) ds + \kappa \int_a^\tau \int_c^z \omega(t, s) ds dt, (\tau, z) \in \Omega,$$

where κ is a positive constant. Then there exists a positive constant ν such that

$$\omega(\tau, z) \leq \nu p(\tau, z).$$

Lemma 3.2 (*Discrete Gronwall-type inequality [15]*) Let $\{k_j\}_{j=0}^n$ be a given non-negative sequence and the sequence $\{\varepsilon_n\}$ satisfy $\varepsilon_0 \leq p_0$ and $\varepsilon_n \leq p_0 + \sum_{j=0}^{n-1} k_j \varepsilon_j$, $n \geq 1$, with $p_0 \geq 0$. Then

$$\varepsilon_n \leq p_0 \exp \left(\sum_{j=0}^{n-1} k_j \right), \quad n \geq 1.$$

Lemma 3.3 (*Discrete Gronwall-type inequality of two variables [1]*) Let $\omega_{n,m}$ be a given non-negative sequence, and let b_1, b_2, b_3 and β be independent of h and k and strictly positive. If the sequence $\omega_{n,m}$ satisfies

$$\omega_{n,m} \leq hb_1 \sum_{\xi=0}^{n-1} \omega_{\xi,m} + kb_2 \sum_{\rho=0}^{m-1} \omega_{n,\rho} + hkb_3 \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} \omega_{\xi,\rho} + \beta,$$

for all $n = 0, 1, \dots, N, m = 0, 1, \dots, M$, then

$$\omega_{n,m} \leq \beta \exp(\gamma(Nh + Mk)),$$

where $\gamma = \frac{1}{2} \left(b_1 + b_2 + \sqrt{(b_1 + b_2)^2 + 4b_3} \right)$.

Theorem 3.1 Let g, κ_1, κ_2 and κ_3 be p times continuously differentiable on their respective domains. Then (4), (5), (7) define a unique approximation $v \in S_{p-1, p-1}^{(-1)}(\Pi_{N, M})$, and the resulting error function $e(\tau, z) = \omega(\tau, z) - v(\tau, z)$ satisfies

$$\|e\|_{L^\infty(D)} \leq C(h + k)^p,$$

where C is a finite constant independent of h and k .

Proof. Define the error $e(\tau, z)$ on $D_{n,m}$ by $e_{n,m}(\tau, z) = \omega(\tau, z) - v_{n,m}(\tau, z)$ for all $n = 0, 1, \dots, N - 1$ and $m = 0, 1, \dots, M - 1$.

There exists a constant C independent of h and k such that

$$\|e_{n,m}\|_{L^\infty(D_{n,m})} \leq C(h + k)^p$$

for all $n = 0, 1, \dots, N - 1$ and $m = 1, \dots, M - 1$.

First, let $(\tau, z) \in D_{0,0}$, we obtain from (4),

$$|e_{0,0}(\tau, z)| \leq \sum_{i+j=p} \frac{1}{i!j!} \left\| \frac{\partial^{i+j}\omega}{\partial\tau^i\partial z^j} \right\| h^i k^j.$$

When using a more direct generalization of the procedures utilized in Lemma 3.6 in [10], there exists a positive number $\alpha(p)$ such that $\left\| \frac{\partial^{i+j}\hat{v}_{n,m}}{\partial\tau^i\partial z^j} \right\| \leq \alpha(p)$ for all $n = 0, 1, \dots, N - 1$, $m = 0, \dots, M - 1$ and $i + j = 0, \dots, p$, where $\hat{v}_{0,0}(\tau, z) = \omega(\tau, z)$ for $(\tau, z) \in D_{0,0}$. Hence,

$$|e_{0,0}(\tau, z)| \leq \alpha(p) \sum_{i+j=p} \frac{1}{i!j!} h^i k^j = \underbrace{\frac{\alpha(p)}{p!}}_{C_1} (h + k)^p.$$

Second, let $(\tau, z) \in D_{n,0}$, for all $n = 1, \dots, N - 1$, we have from (6),

$$\begin{aligned} |\omega(\tau, z) - \hat{v}_{n,0}(\tau, z)| &\leq \sum_{\xi=0}^{n-1} h\kappa \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})} + \sum_{\xi=0}^{n-1} hk\kappa \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})} \\ &+ \kappa \int_{\tau_n}^{\tau} |\omega(t, z) - \hat{v}_{n,0}(t, z)| dt + \kappa \int_0^z |\omega(\tau, s) - \hat{v}_{n,0}(\tau, s)| ds \\ &+ \kappa \int_{\tau_n}^{\tau} \int_0^z |\omega(t, s) - \hat{v}_{n,0}(t, s)| ds dt, \end{aligned}$$

where $\kappa = \max\{\|\kappa_i\|_{L^\infty(D)}, i = 1, 2, 3\}$, then by Lemma 3.1

$$\begin{aligned} |\omega(\tau, z) - \hat{v}_{n,0}(\tau, z)| &\leq \left(\sum_{\xi=0}^{n-1} h\kappa \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})} + \sum_{\xi=0}^{n-1} hk\kappa \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})} \right) \nu \\ &\leq \sum_{\xi=0}^{n-1} h \underbrace{\kappa(1 + Z)}_{\lambda_1} \nu \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})}, \end{aligned}$$

which implies that

$$\begin{aligned} \|e_{n,0}\|_{L^\infty(D_{n,0})} &\leq \|\omega - \hat{v}_{n,0}\| + \|\hat{v}_{n,0} - v_{n,0}\| \\ &\leq \sum_{\xi=0}^{n-1} h\lambda_1 \|e_{\xi,0}\|_{L^\infty(D_{\xi,0})} + \frac{\alpha(p)}{p!} (h + k)^p, \end{aligned}$$

then, by Lemma 3.2, we have

$$\|e_{n,0}\|_{L^\infty(D_{n,0})} \leq \underbrace{\frac{\alpha(p)}{p!} \exp(T\lambda_1)}_{C_2} (h + k)^p.$$

Third, let $(\tau, z) \in D_{n,m}$ for all $n = 0, \dots, N - 1$ and $m = 1, \dots, M - 1$, we have from (8),

$$\begin{aligned}
 |\omega(\tau, z) - \hat{v}_{n,m}(\tau, z)| &\leq \sum_{\xi=0}^{n-1} h\kappa \|e_{\xi,m}\| + \sum_{\rho=0}^{m-1} k\kappa \|e_{n,\rho}\| + \kappa \int_{\tau_n}^{\tau} \int_{z_m}^z |\omega(t, s) - \hat{v}_{n,m}(t, s)| ds dt \\
 &+ \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} hk\kappa \|e_{\xi,\rho}\| + \sum_{\xi=0}^{n-1} hk\kappa \|e_{\xi,m}\| + \sum_{\rho=0}^{m-1} hk\kappa \|e_{n,\rho}\| \\
 &+ \kappa \int_{\tau_n}^{\tau} |\omega(t, z) - \hat{v}_{n,m}(t, z)| dt + \kappa \int_{z_m}^z |\omega(\tau, s) - \hat{v}_{n,m}(\tau, s)| ds,
 \end{aligned}$$

then by Lemma 3.1,

$$\begin{aligned}
 |\omega(\tau, z) - \hat{v}_{n,m}(\tau, z)| &\leq \sum_{\xi=0}^{n-1} h \underbrace{\kappa(1+k)\nu}_{\lambda_2} \|e_{\xi,m}\| + \sum_{\rho=0}^{m-1} k \underbrace{\kappa(1+h)\nu}_{\lambda_3} \|e_{n,\rho}\| \\
 &+ \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} hk \underbrace{\kappa\nu}_{\lambda_4} \|e_{\xi,\rho}\|,
 \end{aligned}$$

which implies that

$$\begin{aligned}
 \|e_{n,m}\| &\leq \|\omega - \hat{v}_{n,m}\| + \|\hat{v}_{n,m} - v_{n,m}\| \\
 &\leq \sum_{\xi=0}^{n-1} h\lambda_2 \|e_{\xi,m}\| + \sum_{\rho=0}^{m-1} k\lambda_3 \|e_{n,\rho}\| + \sum_{\xi=0}^{n-1} \sum_{\rho=0}^{m-1} hk\lambda_4 \|e_{\xi,\rho}\| + \frac{\alpha(p)}{p!} (h+k)^p,
 \end{aligned}$$

using Lemma 3.3, we obtain

$$\|e_{n,m}\| \leq \underbrace{\frac{\alpha(p)}{p!} \exp(\gamma_3(T+Z))}_{C_3} (h+k)^p$$

such that $\gamma_3 = \frac{1}{2} (\lambda_2 + \lambda_3 + \sqrt{(\lambda_2 + \lambda_3)^2 + 4\lambda_3})$.

Thus, the proof is completed by taking $C = \max\{C_1, C_2, C_3\}$. \square

4 Numerical examples

In this section, we present numerical experiments that assess the performance of the Taylor collocation method (TCM) for solving problems of the form (1) in Example 4.4 and the form (2) in Examples 4.1–4.3. We also compare the TCM results with those obtained using other methods such as the multi-step method [9], Euler-type method, and Trapezoidal method [1], Chelyshkov polynomials method [6], bivariate shifted Legendre functions method [16], and two-dimensional block-pulse functions method [17]. In each example, we compare the TCM solution with the results obtained from previous references. Our numerical experiments were conducted using Maple version 17 and a PC with Intel Core i7-2630QM CPU @2.00 GHz and 8,00 Go of RAM, running MS Windows 7 operating system. We observed that the TCM produces more accurate results than the previous methods.

(τ, z)	$N = M = 10, p = 3$	$N = M = 20, p = 3$	$N = M = 10, p = 4$
(0.1, 0.1)	1.73e - 06	5.99e - 07	2.59e - 06
(0.2, 0.2)	1.84e - 05	5.29e - 06	2.20e - 05
(0.3, 0.3)	6.23e - 05	1.68e - 05	7.02e - 05
(0.4, 0.4)	1.34e - 04	3.53e - 05	1.47e - 04
(0.5, 0.5)	2.28e - 04	5.85e - 05	2.46e - 04
(0.6, 0.6)	3.29e - 04	8.35e - 05	3.52e - 04
(0.7, 0.7)	4.27e - 04	1.07e - 04	4.55e - 04
(0.8, 0.8)	5.14e - 04	1.28e - 04	5.46e - 04
(0.9, 0.9)	5.85e - 04	1.45e - 04	6.21e - 04
(1.0, 1.0)	1.86e - 03	3.17e - 04	5.93e - 04

Table 1: Absolute errors function for Example 4.1.

Example 4.1 Consider the linear 2D VIE of the first kind

$$\int_0^\tau \int_0^z (\tau z + 1)\omega(t, s)dsdt = f(\tau, z), \quad \tau, z \in [0, 1]. \tag{9}$$

By differentiating both sides of equation (9), we obtain

$$\omega(\tau, z) = g(\tau, z) - \int_0^\tau \frac{z\omega(t, z)}{\tau z + 1} dt - \int_0^z \frac{\tau\omega(\tau, s)}{\tau z + 1} ds - \int_0^\tau \int_0^z \frac{\omega(t, s)}{\tau z + 1} dsdt, \tag{10}$$

where $g(\tau, z) = \frac{-3\tau^2 + (2+3\tau+3\tau z)\tau e^z}{2(1+\tau z)}$, which has the exact solution $\omega(\tau, z) = \tau e^z$.

A comparison between the approximate and exact solutions is shown in Table 1 by applying the TCM to equation (10) at some points with $p = 3, 4$ and $(N, M) = (10, 10), (20, 20)$.

Example 4.2 Consider the linear 2D VIE of the first kind [9]

$$\left(\frac{\tau^2 z^2 + 2 \sin(\tau z) - 2\tau z \cos(\tau z)}{2z^2} \right) \sin(z) = \int_0^\tau \int_0^z (\sin(zt) + 1)\omega(t, s)dsdt$$

for $\tau, z \in [0, 1]$, and the exact solution is $\omega(\tau, z) = \tau \cos(z)$. This equation is equivalent to the following linear 2D VIE of the second kind:

$$\omega(\tau, z) = \tau \cos(z) + \frac{\tau^2 \sin(z) \cos(\tau z)}{\sin(\tau z) + 1} - \int_0^z \frac{\tau \cos(\tau s)}{\sin(\tau z) + 1} \omega(\tau, s)ds.$$

The numerical results for $p = 4$ and $N = M = 15$ of the TCM and the numerical results obtained by using the multi-step method [9] are compared in Table 2.

Example 4.3 Consider the linear 2D VIE of the first kind [1]

$$f(\tau, z) = \int_0^\tau \int_0^z (\sin(z + t) + \sin(\tau + s) + 3)\omega(t, s)dsdt, \quad \tau, z \in [0, 2],$$

where $g(\tau, z)$ is chosen so that the exact solution is $\omega(\tau, z) = \cos(\tau + z)$.

In Table 3, the numerical results for $p = 3$ and $h = k = 0.1, 0.05$ of the present method (TCM) are compared with the numerical results obtained by using the Euler-type method (EM) and Trapezoidal method (TM) [1], Chelyshkov polynomials method (2D-CPs) [6], bivariate shifted Legendre functions method [16] and two-dimensional block-pulse functions method (2D BPFs) [17].

(τ, z)	multi-steps method	present method
$(2^{-7}, 2^{-7})$	$2.38e - 07$	$2.00e - 12$
$(2^{-6}, 2^{-6})$	$1.90e - 06$	$4.00e - 11$
$(2^{-5}, 2^{-5})$	$1.57e - 05$	$1.24e - 09$
$(2^{-4}, 2^{-4})$	$2.25e - 06$	$3.97e - 08$
$(2^{-3}, 2^{-3})$	$1.51e - 07$	$1.88e - 07$
$(2^{-2}, 2^{-2})$	$1.92e - 07$	$2.66e - 07$
$(2^{-1}, 2^{-1})$	$6.16e - 07$	$8.87e - 08$

Table 2: Comparison of the absolute errors of Example 4.2.

(τ, z)	TCM $h = 0.05$	EM [1] $h = 0.05$	TM [1] $h = 0.05$	Method in [16] $M = 4$	2D-BPFs [17] $m = 32$
(1, 1)	$8.57e - 07$	$4.06e - 02$	$9.80e - 04$	$4.96e - 06$	$6.08e - 02$
(1, 2)	$2.29e - 05$	$1.23e - 02$	$5.47e - 04$	$5.98e - 06$	$4.00e - 03$
(2, 1)	$2.29e - 05$	$1.23e - 02$	$5.47e - 04$	$9.87e - 03$	$4.00e - 03$
(2, 2)	$4.27e - 05$	$4.06e - 02$	$2.03e - 03$	$1.22e - 05$	$4.74e - 02$
$\ e\ _\infty$	$4.27e - 05$	$5.49e - 02$	$2.03e - 03$	/	/

(τ, z)	TCM		2D-CPs [6]	
	$h = 0.1$	$h = 0.05$	$N = 2$	$M = 4$
(0.1, 0.1)	$2.77e - 07$	$3.78e - 08$	$7.41e - 03$	$4.77e - 06$
(0.2, 0.2)	$9.38e - 07$	$1.18e - 07$	$4.43e - 04$	$2.10e - 05$
(0.3, 0.3)	$1.63e - 06$	$2.03e - 07$	$5.40e - 03$	$7.20e - 06$
(0.4, 0.4)	$2.14e - 06$	$2.67e - 07$	$6.65e - 03$	$8.20e - 06$
(0.5, 0.5)	$2.33e - 06$	$2.91e - 07$	$4.48e - 03$	$6.40e - 06$
(0.6, 0.6)	$2.09e - 06$	$2.62e - 07$	$5.43e - 05$	$6.90e - 06$
(0.7, 0.7)	$1.28e - 06$	$1.60e - 07$	$4.80e - 03$	$1.20e - 05$
(0.8, 0.8)	$2.85e - 07$	$3.66e - 08$	$7.93e - 03$	$5.00e - 06$
(0.9, 0.9)	$2.80e - 06$	$3.58e - 07$	$7.08e - 03$	$3.00e - 05$
(1, 1)	$6.85e - 06$	$8.59e - 07$	$2.77e - 04$	$1.80e - 06$

Table 3: Comparison of the absolute errors for Example 4.3.

Example 4.4 Consider the nonlinear 2D VIE of the first kind [17]

$$\frac{1}{9}(e^{\tau+z} - e^{\tau+4z} - e^{7\tau+z} + e^{7\tau+4z}) = \int_0^\tau \int_0^z 2e^{\tau+z}\omega^3(t, s)dsdt$$

for $\tau, z \in [0, 1]$, and the exact solution is $\omega(\tau, z) = e^{\tau+2z}$. This equation is equivalent to the following linear 2D VIE of the second kind:

$$u(\tau, z) = g(\tau, z) - \int_0^\tau u(t, z)dt - \int_0^z u(\tau, s)ds - \int_0^\tau \int_0^z u(t, s)dsdt,$$

where $u = \omega^3$. In Table 4, the numerical results for $p = 3$ and $N = M = 64$ of the TCM are compared with the numerical results obtained by using the Chelyshkov polynomials method (2D CPs) [6], bivariate shifted Legendre functions method [16] and two-dimensional block-pulse functions method (2D BPFs) [17].

$(2^{-i}, 2^{-i})$	2D-BPFs [17]	Method in [16]	2D-CPs [6]	Present method
$i = 1$	$1.0e - 1$	$2.6e - 6$	$3.5e - 5$	$6.1e - 6$
$i = 2$	$4.6e - 2$	$4.6e - 6$	$2.0e - 6$	$2.6e - 6$
$i = 3$	$2.9e - 2$	$6.3e - 7$	$1.5e - 5$	$1.3e - 6$
$i = 4$	$2.3e - 2$	$1.2e - 5$	$1.2e - 5$	$7.2e - 7$
$i = 5$	$2.0e - 2$	$3.8e - 6$	$5.9e - 5$	$3.7e - 7$
$i = 6$	$3.1e - 2$	$9.0e - 6$	$9.6e - 5$	$1.9e - 7$

Table 4: Comparison of the absolute errors of Example 4.4.

5 Conclusion

In this paper, the problem expressed in (1) is transformed into a linear 2D VIE of the second kind, which is given by (3). A collocation method using the Taylor polynomials is developed to solve the 2D VIE of the second kind. The convergence and error analysis of this method are investigated, and numerical examples are provided to illustrate its effectiveness and accuracy. The numerical results confirm the theoretical estimates, and comparisons with other methods are presented. This method can be easily generalized and applied to a system of 2D VIEs of the first and second kinds.

References

- [1] S. McKee, T. Tang and T. Diogo. An Euler-type method for two-dimensional Volterra integral equations of the first kind. *IMA Journal of Numerical Analysis* **20** (3) (2000) 423–440.
- [2] K. Maleknejad and M. Shahabi. Numerical solution of two-dimensional nonlinear Volterra integral equations of the first kind using operational matrices of hybrid functions. *International Journal of Computer Mathematics* **99** (10) (2022) 2105–2122.
- [3] D. J. Sabeg, R. Ezzati and K. Maleknejad. Solving two-dimensional integral equations of fractional order by using operational matrix of two-dimensional shifted Legendre polynomials. *Nonlinear Dyn. Syst. Theory* **18** (3) (2018) 297–306.
- [4] M. Asgari, R. Ezzat and H. Jafari. Solution of 2D Fractional Order Integral Equations by Bernstein Polynomials Operational Matrices. *Nonlinear Dyn. Syst. Theory* **19** (2019), 10–20.
- [5] M. Mounni and M. Tilioua. A Neural Network Approximation for a Model of Micromagnetism. *Nonlinear Dyn. Syst. Theory* **22** (4) (2022) 432–446.
- [6] A. M. S. Mahdy, D. Shokry and K. Lotfy. Chelyshkov polynomials strategy for solving 2-dimensional nonlinear Volterra integral equations of the first kind. *Computational and Applied Mathematics* **41** 257(2022).
- [7] A. Tari, M. Rahimi, S. Shahmorad and F. Talati. Solving a class of two-dimensional linear and nonlinear Volterra integral equations by the differential transform method. *Journal of Computational and Applied Mathematics* **228** (1) (2009) 70–76.
- [8] S. Nemati, P. M. Lima and Y. Ordokhani. Numerical solution of a class of two-dimensional nonlinear Volterra integral equations using Legendre polynomials. *Journal of Computational and Applied Mathematics* **242** (2013) 53–69.
- [9] S. M. Torabi and A. Tari Marzabad. Numerical solution of two-dimensional integral equations of the first kind by multi-step methods. *Computational Methods for Differential Equations* **4** (2) (2016) 128–138.

- [10] H. Laib, A. Boulmerka, A. Bellour and F. Birem. Numerical solution of two-dimensional linear and nonlinear Volterra integral equations using Taylor collocation method. *Journal of Computational and Applied Mathematics* **417** (2023) 114537.
- [11] H. Laib, A. Bellour and M. Bousselsal. Numerical solution of high-order linear Volterra integro-differential equations by using Taylor collocation method. *International Journal of Computer Mathematics* **96** (5) (2019) 1066–1085.
- [12] H. Laib, A. Bellour and A. Boulmerka. Taylor collocation method for a system of nonlinear Volterra delay integro-differential equations with application to COVID-19 epidemic. *International Journal of Computer Mathematics* **99** (4) (2022) 852–876.
- [13] A. Bellour, M. Bousselsal and H. Laib. Numerical solution of second-order linear delay differential and integro-differential equations by using Taylor collocation method. *International Journal of Computational Methods* **17** (9) (2020) 1950070.
- [14] H. Laib, A. Bellour and A. Boulmerka. Taylor collocation method for high-order neutral delay Volterra integro-differential equations. *Journal of Innovative Applied Mathematics and Computational Sciences* **2** (1) (2022).
- [15] H. Brunner. *Collocation Methods for Volterra Integral and Related Functional Differential Equations*. Cambridge university press, **15**, 2004.
- [16] S. Nemati and Y. Ordokhani. Solving nonlinear two-dimensional Volterra integral equations of the first-kind using bivariate shifted Legendre functions. *International Journal of Mathematical Modelling* **5** (3) (2015) 219–230.
- [17] K. Maleknejad, S. Sohrabi and B. Baranji. Application of 2d-bpfs to nonlinear integral equations. *Communications in Nonlinear Science and Numerical Simulation* **15** (3) (2010) 527–535.



Application of Model Predictive Control (MPC) to Longitudinal Motion of the Aircraft Using Polynomial Chaos

K. D. R. Dewi, K. Fahim and S. Subchan *

Department of Mathematics, Faculty of Science and Data Analytics, Sepuluh Nopember Institute of Technology, Surabaya, Indonesia

Received: July 25, 2023; Revised: November 5, 2023

Abstract: Dynamical systems can be stochastic or uncertain because of some assumptions or distractions that limit the problem. This occurs when the system is obtained from data using system identifiers with various uncertainties. One example of a system that contains uncertainty parameters is the longitudinal motion of the aircraft model. The longitudinal motion of the aircraft requires control, so in this study, control was applied using the Model Predictive Control (MPC) method. Before applying control to the aircraft model, the Polynomial Chaos expansion will be applied to the state space model to get the deterministic model. The simulation uses different prediction horizons (N_p) and polynomial orders (r). Based on the simulation results, it was found that the pitch rate output can approach the given pitch rate reference.

Keywords: *polynomial chaos; hermite polynomial; model predictive control.*

Mathematics Subject Classification (2010): 33C45, 34H10, 93C15.

1 Introduction

Mathematical models are the representations of phenomena or realities written in mathematical equations. The process of constructing a mathematical model of reality or a problem is called mathematical modelling [1]. Mathematical models can be written in the dynamic system [2]. A dynamic system is a system that changes or experiences dynamics over time. In practice, dynamic systems are not always deterministic. The dynamic system can be stochastic because there are assumptions to limit the problem

* Corresponding author: <mailto:Subchan@matematika.its.ac.id>

or disturbances. Uncertainty in the parameters occurs if the system parameters are not known. This can also happen when the system model is obtained from the data using system identification so that the system's transfer function has an uncertainty range [3].

One of the methods for approximating linear dynamic systems with parameter uncertainty is the Polynomial Chaos method. By using the Polynomial Chaos method, the stochastic system will be transformed into a deterministic system with a larger dimension state space [4]. Research on the Polynomial Chaos method in dynamical systems was conducted by Bhattacharya [5] in 2014. This research develops a control design using the Robust State Feedback Control method with probabilistic system parameters. In this study, the Polynomial Chaos method approach was applied to the F-16 aircraft model. Based on the simulation results, it was found that by using the Polynomial Chaos method, the control design showed good consistency. Another study discussing the Polynomial Chaos method was conducted by Tadiparthi and Bhattacharya [6] in 2020. In this research, the Robust Linear Quadratic Regulator (LQR) algorithm was developed with a control design based on the Polynomial Chaos method.

Model Predictive Control (MPC) is an advanced process control method that is widely applied in industrial processes. Research on the Robust Model Predictive Control (MPC) method was conducted by Asfihani, et al. [7] in 2019. In this study, the Robust MPC method was applied to a linear model of Unmanned Surface Vehicle (USV) motion. Robust MPC is used to control Dubin's track tracking [8, 9]. In this study, sea waves are considered as a disturbance. The simulation results show that the Robust MPC method can guide ships to follow the trajectory and reject disturbances. Another study discussing the MPC method was conducted in 2020 by Asfihani, et al. [10]. The MPC method is applied to control the missile. The MPC aims to minimize the time it takes for the missile to reach a moving target. The simulation results show that the fastest time to reach the target is 20 seconds, when the horizon prediction value is 10 and given a constraint on the state.

Based on the previous research described above, the Polynomial Chaos and Model Predictive Control (MPC) are applied for the longitudinal motion of the F-16 aircraft control. In this paper, the nonlinear aircraft model is taken from Stevens, et al. [11]. The Taylor expansion is employed for linearization so that a linear model is obtained based on the coefficient data of the F-16 aircraft. The longitudinal motion model of the F-16 aircraft contains stochastic uncertainty in the parameter, so the Polynomial Chaos method is applied to the linear model [4,12] to obtain a deterministic model. The previous studies in [4–6] assumed that the random variables Δ were uniformly distributed. In this study, it is assumed that the distribution of the random variables is Gaussian. Then the control will be applied to a deterministic state space with the MPC method.

This paper is constructed as follows. The longitudinal motion model of the aircraft is defined in Section 2. Then Section 3 explains the Polynomial Chaos method to transform stochastic state space into deterministic state space. Section 4 explains the MPC design implemented in the deterministic state space obtained from the Polynomial Chaos method. Next, Section 5 deals with discussions based on the simulation result. And last, Section 6 provides the conclusion.

2 Linear Model of Longitudinal Motion of the Aircraft

Longitudinal motion is the movement of the aircraft in a vertical direction such as climbing or swooping. The control affecting the aircraft's longitudinal motion response is the

elevator deflection [13]. The condition when the elevator angle is negative causes the tail of the plane to go down and the nose to go up (pitch angle is positive) [14].

In this study, the F-16 longitudinal motion model is used in the form of a stochastic state space that contains uncertainty parameters. The state variable $\mathbf{x} = [\alpha \ q \ x_e]^T$ consists of the angle of attack (α), pitch rate (q), and elevator state (x_e) that captures the actuator dynamics. The control input (u) is in the form of an elevator command or an elevator deflection with an angle in degrees (δ_{ec}). So the stochastic state space for the longitudinal motion of the aircraft is written as follows [4, 12]:

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}(\Delta)\mathbf{x} + \mathbf{B}u, \\ \begin{bmatrix} \dot{\alpha} \\ \dot{q} \\ \dot{x}_e \end{bmatrix} &= \begin{bmatrix} -0,6398 & 0,9378 & -0,0014 \\ a_{21}(\Delta) & a_{22}(\Delta) & a_{23}(\Delta) \\ 0 & 0 & -20,2000 \end{bmatrix} \begin{bmatrix} \alpha \\ q \\ x_e \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 20,2 \end{bmatrix} [\delta_{ec}] \end{aligned} \quad (1)$$

with the output in the form of the pitch rate (q) so that the output equation is as follows:

$$y = \mathbf{C}\mathbf{x} = \begin{bmatrix} 0 & \frac{180}{\pi} & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ q \\ x_e \end{bmatrix}, \quad (2)$$

where

$$\begin{aligned} a_{21}(\Delta) &= -1,5679(1 + 0,1\Delta), \\ a_{22}(\Delta) &= -0,8791(1 + 0,1\Delta), \\ a_{23}(\Delta) &= -0,1137(1 + 0,1\Delta). \end{aligned}$$

The stochastic uncertainty parameter in the longitudinal motion model of the F-16 aircraft is shown by $\mathbf{A}(\Delta)$, a matrix function of the random variable Δ . Random variables Δ are assumed to be Gaussian distributed.

The random variable Δ represents the pitch of the plane. This relates to the derived coefficients of the aircraft pitching moment. The pitching moment is the moment that pivots on the pitch axis/ Y axis [14]. The random variables $a_{21}(\Delta)$, $a_{22}(\Delta)$, and $a_{23}(\Delta)$ represent the derived coefficients of pitch stiffness, pitch damping plane, and the power of the elevator control [11].

3 Polynomial Chaos Method

Polynomial Chaos is a deterministic approach used to handle uncertainty evolution when probabilistic uncertainty exists in the system parameters. [15]. Let $\Delta : \Omega \rightarrow \mathbb{R}^d$ be a random variable and $L^2(\Omega, \mathcal{F}, P)$ be the set of all random variables ξ over the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\int_{\Omega} |\xi|^2 d\mathbb{P} < \infty$. The second-order general process of

$X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ can be written in Polynomial Chaos as follows [6]:

$$X(\omega) = \sum_{i=0}^{\infty} x_i \phi_i(\Delta(\omega)), \quad (3)$$

where ω denotes the sample point and $\phi(\Delta)$ represents the general Polynomial Chaos (gPC) basis with degree p over the random variable Δ .

For a random variable Δ with a given distribution, a family of the orthogonal basis functions $\{\phi_i\}$ can be selected such that its weight function $f(x)$ exhibits a similar form as the probability density function, i.e.,

$$\int_{\mathcal{D}_\Delta} \phi_i(x)\phi_j(x)f(x) dx = E[\phi_i(\Delta)\phi_j(\Delta)] = E[\phi_i^2(\Delta)]\delta_{ij} = \langle \phi_i, \phi_j \rangle, \tag{4}$$

where $E[\cdot]$ represents the expectation relative to the probability measure denoted by \mathbb{P} and pdf f .

The relation between the choice of polynomials and given Δ distribution can be seen in Table 1.

Random Variable Δ	Polynomial Choice
Beta	Jacobi
Gamma	Laguerre
Uniform	Legendre
Gaussian	Hermite

Table 1: The Relation between Polynomial Choices and the $\Delta(\omega)$ Distribution.

A stochastic linear system is defined as follows:

$$\dot{\mathbf{x}}(t, \Delta) = \mathbf{A}(\Delta)\mathbf{x}(t, \Delta) + \mathbf{B}(\Delta)\mathbf{u}(t), \tag{5}$$

where $\mathbf{x} \in \mathbb{R}^{n_x}$, $\mathbf{u} \in \mathbb{R}^{n_u}$. The system in Equation (5) contains probability uncertainties in its system parameters.

By applying the general finite-order Polynomial Chaos expansion, we get [6]

$$\hat{x}_i(t, \Delta) = \sum_{k=0}^p x_{i,k}(t)\phi_k(\Delta) = \tilde{\mathbf{x}}_i(t)^T \Phi(\Delta), \tag{6}$$

$$\hat{A}_{ij}(\Delta) = \sum_{k=0}^p a_{ij,k}\phi_k(\Delta) = \tilde{\mathbf{a}}_{ij}^T \Phi(\Delta), \tag{7}$$

$$\hat{B}_{ij}(\Delta) = \sum_{k=0}^p b_{ij,k}\phi_k(\Delta) = \tilde{\mathbf{b}}_{ij}^T \Phi(\Delta), \tag{8}$$

where $\tilde{\mathbf{x}}_i(t), \tilde{\mathbf{a}}_{ij}, \tilde{\mathbf{b}}_{ij}, \Phi(\Delta) \in \mathbb{R}^p$ is defined as

$$\begin{aligned} \tilde{\mathbf{x}}_i(t) &= [x_{i,0}(t) \ \cdots \ x_{i,p}(t)]^T, \\ \tilde{\mathbf{a}}_{ij} &= [a_{ij,0} \ \cdots \ a_{ij,p}]^T, \\ \tilde{\mathbf{b}}_{ij} &= [b_{ij,0} \ \cdots \ b_{ij,p}]^T, \\ \Phi(\Delta) &= [\phi_0(\Delta) \ \cdots \ \phi_p(\Delta)]^T. \end{aligned}$$

The value of p is determined by the dimension of Δ (d) and the order of the orthogonal polynomial $\{\phi_k\}$ (denoted by r), which satisfies $p + 1 = \frac{(d+r)!}{d!r!}$. By using the Galerkin projection onto $\{\phi_k\}_{k=0}^p$, the coefficients $a_{ij,k}$ and $b_{ij,k}$ can be written as follows:

$$a_{ij,k} = \frac{\langle A_{ij}(\Delta), \phi_k(\Delta) \rangle}{\langle \phi_k(\Delta), \phi_k(\Delta) \rangle}, \tag{9}$$

$$b_{ij,k} = \frac{\langle B_{ij}(\Delta), \phi_k(\Delta) \rangle}{\langle \phi_k(\Delta), \phi_k(\Delta) \rangle}. \tag{10}$$

By substituting the Polynomial Chaos expansion (Equations (6), (7), (8)) in the stochastic linear system (5), where $\{\phi_k\}$ is the basis of an orthogonal polynomial, we get

$$\begin{aligned} \sum_{k=0}^p \dot{\mathbf{x}}_{i,k}(t)\phi_k(\Delta) &= \sum_{j=1}^{n_x} \sum_{k=0}^p \sum_{l=0}^p a_{ij,k}x_{i,l}(t)\phi_k(\Delta)\phi_l(\Delta) \\ &+ \sum_{j=1}^{n_u} \sum_{k=0}^p b_{ij,k}\phi_k(\Delta)u_j(t). \end{aligned} \tag{11}$$

Applying the inner product on both sides with ϕ_m for $m = 0, 1, 2, \dots, p$, we get

$$\begin{aligned} \sum_{k=0}^p \dot{\mathbf{x}}_{i,k}(t)\langle \phi_k(\Delta), \phi_m(\Delta) \rangle &= \sum_{j=1}^{n_x} \sum_{k=0}^p \sum_{l=0}^p a_{ij,k}x_{i,l}(t)\langle \phi_k(\Delta)\phi_l(\Delta), \phi_m(\Delta) \rangle \\ &+ \sum_{j=1}^{n_u} \sum_{k=0}^p b_{ij,k}u_j(t)\langle \phi_k(\Delta), \phi_m(\Delta) \rangle. \end{aligned} \tag{12}$$

Since $\{\phi_m\}$ is an orthogonal basis, $\langle \phi_k, \phi_m \rangle = 0$ for $k \neq m$, Equation (12) becomes

$$\begin{aligned} \dot{\mathbf{x}}_{i,m}(t)\|\phi_m\|^2 &= \sum_{j=1}^{n_x} \sum_{k=0}^p \sum_{l=0}^p a_{ij,k}x_{i,l}(t)\langle \phi_k(\Delta)\phi_l(\Delta), \phi_m(\Delta) \rangle \\ &+ \sum_{j=1}^{n_u} b_{ij,k}u_j(t)\|\phi_m\|^2. \end{aligned} \tag{13}$$

Then divide (13) by $\|\phi_m\|^2$ to get

$$\dot{\mathbf{x}}_{i,m}(t) = \sum_{j=1}^{n_x} \sum_{k=0}^p \sum_{l=0}^p a_{ij,k}x_{i,l}(t)C_{klm} + \sum_{j=1}^{n_u} b_{ij,m}u_j(t), \tag{14}$$

where $C_{klm} = \frac{\langle \phi_k\phi_l, \phi_m \rangle}{\|\phi_m\|^2}$.

From (14), the deterministic differential equation is obtained as follows:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \tag{15}$$

where

$$\begin{aligned} \mathbf{x} &= [\mathbf{x}_1^T \quad \mathbf{x}_2^T \quad \dots \quad \mathbf{x}_{n_x}^T]^T, \\ \mathbf{u} &= [\mathbf{u}_1^T \quad \mathbf{u}_2^T \quad \dots \quad \mathbf{u}_{n_u}^T]^T, \\ \mathbf{A} &= [\mathbf{A}_{ij}] = \sum_{k=0}^p a_{ij,k}\mathbf{T}_k, \quad i, j = 1, 2, \dots, n_x, \\ \mathbf{B} &= [\mathbf{B}_{ij}] = \mathbf{b}_{ij}, \quad i = 1, 2, \dots, n_x; j = 1, 2, \dots, n_u. \end{aligned}$$

And

$$\mathbf{T}_k = \begin{bmatrix} C_{k00} & C_{k10} & \dots & C_{kp0} \\ C_{k01} & C_{k11} & \dots & C_{kp1} \\ \vdots & \vdots & \ddots & \vdots \\ C_{k0p} & C_{k1p} & \dots & C_{kpp} \end{bmatrix}$$

with the dimensions of \mathbf{x} , \mathbf{A} , \mathbf{B} , and \mathbf{u} equal to $n_x(p + 1) \times 1$, $n_x(p + 1) \times n_x(p + 1)$, $n_x(p + 1) \times n_u$, and $n_u \times n_u$, respectively.

4 Design of Model Predictive Control

MPC is a control system that can predict the future system output by considering both the present input and output information. The advantages of the MPC method are that it can handle multi-variable system control problems, has the ability to feed-forward control to compensate for measured disturbances, and takes into account input and state constraints [16] [17]. The basic MPC structure is presented in Figure 1 [17].

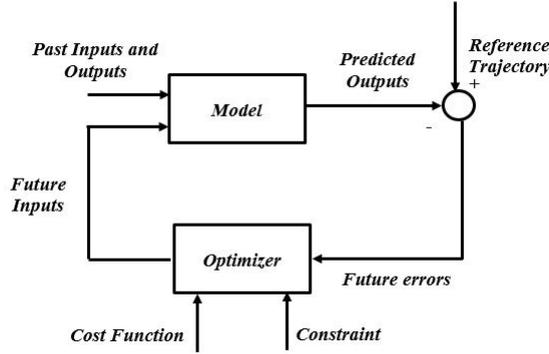


Figure 1: MPC Structure.

In this study, MPC is applied to a discrete deterministic system given in (16) and (17):

$$\mathcal{X}(k+1) = \mathbf{A}_d \mathcal{X}(k) + \mathbf{B}_d \mathbf{U}(k), \quad (16)$$

$$\mathcal{Y}(k) = \mathbf{C} \mathcal{X}(k). \quad (17)$$

In designing the MPC control in this study, the objective function formulation and constraints were determined as prediction constraints using a dynamic system and limit constraints on control inputs. By using $N_p = N_c$, the MPC objective function can be written as

$$J = \sum_{j=1}^{N_p} [(\mathcal{Y}_{ref}(k+j|k) - \mathcal{Y}(k+j|k))^T \mathbf{Q} (\mathcal{Y}_{ref}(k+j|k) - \mathcal{Y}(k+j|k)) + \mathbf{U}(k+j-1|k)^T \mathbf{R} \mathbf{U}(k+j-1|k)], \quad (18)$$

where \mathcal{Y}_{ref} is the reference output and \mathcal{Y} is the system output subject to

$$\mathcal{X}(k+j+1|k) = \mathbf{A}_d \mathcal{X}(k+j|k) + \mathbf{B}_d \mathbf{U}(k+j|k), \quad j = 0, 1, \dots, N_p - 1. \quad (19)$$

$$\mathcal{Y}(k+j|k) = \mathbf{C} \mathcal{X}(k+j|k), \quad j = 1, 2, \dots, N_p, \quad (20)$$

$$\mathbf{U}_{min} \leq \mathbf{U}(k+j|k) \leq \mathbf{U}_{max}, \quad j = 0, 1, \dots, N_p - 1. \quad (21)$$

By transforming the objective function (18) and constraints in equations (19), (20), (21) into quadratic programming form, the problem formulation for MPC is written as follows:

$$\min_{\mathbf{U}} J(\mathbf{U}) = \frac{1}{2} \mathbf{U}^T \mathbf{H} \mathbf{U} + \mathbf{U}^T \mathbf{f} \quad (22)$$

subject to

$$\begin{bmatrix} \mathbf{u}_{min} \\ \mathbf{u}_{min} \\ \mathbf{u}_{min} \\ \vdots \\ \mathbf{u}_{min} \end{bmatrix} \leq \mathbf{U} \leq \begin{bmatrix} \mathbf{u}_{max} \\ \mathbf{u}_{max} \\ \mathbf{u}_{max} \\ \vdots \\ \mathbf{u}_{max} \end{bmatrix}, \tag{23}$$

where

$$\mathbf{f} = 2\Phi^T \bar{\mathbf{Q}}(\mathbf{F}\mathcal{X}(k|k) - \mathbf{Y}_{ref}),$$

$$\mathbf{H} = 2(\Phi^T \bar{\mathbf{Q}}\Phi + \bar{\mathbf{R}}),$$

$$\bar{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{Q} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{Q} \end{bmatrix}_{n_y \cdot N_p \times n_y \cdot N_p},$$

$$\bar{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{R} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{R} \end{bmatrix}_{n_u \cdot N_p \times n_u \cdot N_p},$$

$$\mathbf{F} = \begin{bmatrix} \mathbf{C}\mathbf{A}_d \\ \mathbf{C}\mathbf{A}_d^2 \\ \mathbf{C}\mathbf{A}_d^3 \\ \vdots \\ \mathbf{C}\mathbf{A}_d^{N_p} \end{bmatrix}_{n_y \cdot N_p \times n_x(p+1)},$$

$$\Phi = \begin{bmatrix} \mathbf{C}\mathbf{B}_d & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}\mathbf{A}_d\mathbf{B}_d & \mathbf{C}\mathbf{B}_d & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}\mathbf{A}_d^2\mathbf{B}_d & \mathbf{C}\mathbf{A}_d\mathbf{B}_d & \mathbf{C}\mathbf{B}_d & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{C}\mathbf{A}_d^{N_p-1}\mathbf{B}_d & \mathbf{C}\mathbf{A}_d^{N_p-2}\mathbf{B}_d & \mathbf{C}\mathbf{A}_d^{N_p-3}\mathbf{B}_d & \cdots & \mathbf{C}\mathbf{B}_d \end{bmatrix}_{n_y \cdot N_p \times n_u \cdot N_p}.$$

5 Simulation and Discussion

The MPC simulation aims to make the pitch rate follow the given reference, $y_{ref} = q_{ref} = 0^\circ/s$. The value of the pitch rate of $0^\circ/s$ means that there is no difference or change in the pitch angle every time, so it can be said that there is no movement on the nose of the aircraft or it is in a stable condition. The parameter values used in the simulation are $T_s = 0,05$, $\mathbf{x}(0) = [30^\circ, 10^\circ/s, 15^\circ]^T$, $\mathbf{Q} = 10.000$, $\mathbf{R} = 1$, $\mathbf{u}_{min} = -25^\circ$, and $\mathbf{u}_{max} = 25^\circ$.

This simulation used the value of polynomial order $r = 3$ and prediction horizon $N_p = 10$. The optimal control results (\mathbf{U}^*) are applied to the stochastic state space of the F-16 aircraft by generating the random variable Δ , which is Gaussian distributed in 1000 simulations. The elevator control δ_{ec} and the pitch rate output are obtained as follows.

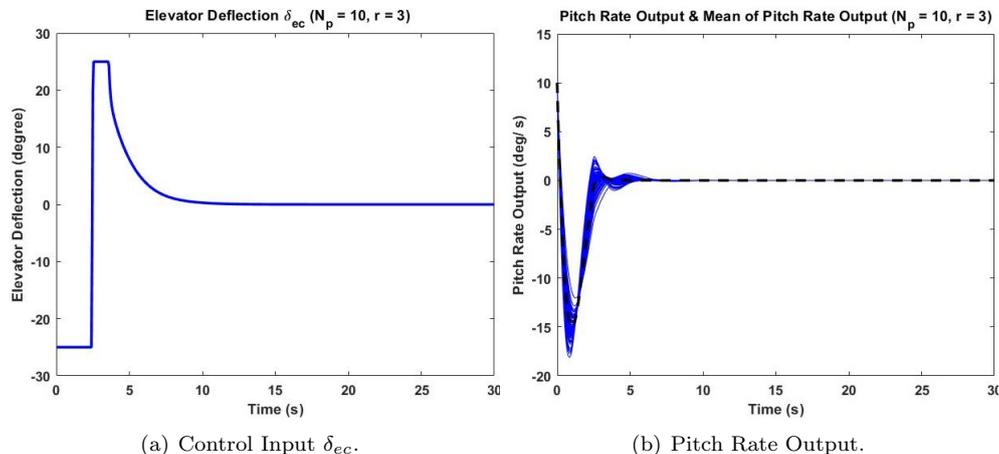


Figure 2: Control Input & Pitch Rate Output ($N_p = 10$, $r = 3$).

From Figure 2(a), it can be seen that the elevator deflection movement is still within the given constrain, $-25^\circ \leq \delta_{ec} \leq 25^\circ$. Changes in the ups and downs of the δ_{ec} value indicate the effort of the control input so that the output approaches the given reference, namely $q_{ref} = 0^\circ/s$.

In Figure 2(b), the blue line is the pitch rate output, while the black dotted line shows the mean of the pitch rate output by simulating 1000 times. From Figure 2(b), it can be seen that the output of the pitch rate by simulating 1000 times and the mean of pitch rate output both converge or approach the given reference, $q = 0^\circ/\text{second}$. From the simulation, the mean of the MAE (mean absolute error) by simulating 1000 times is 0,78171.

5.1 Simulation with various prediction horizon values

The first simulation used the polynomial order value $r = 3$ and different prediction horizon values ($N_p = 5, N_p = 10, N_p = 20$). The simulation results of the elevator deflection control δ_{ec} and the mean of the pitch rate output by doing 1000 simulations are shown in the following figures.

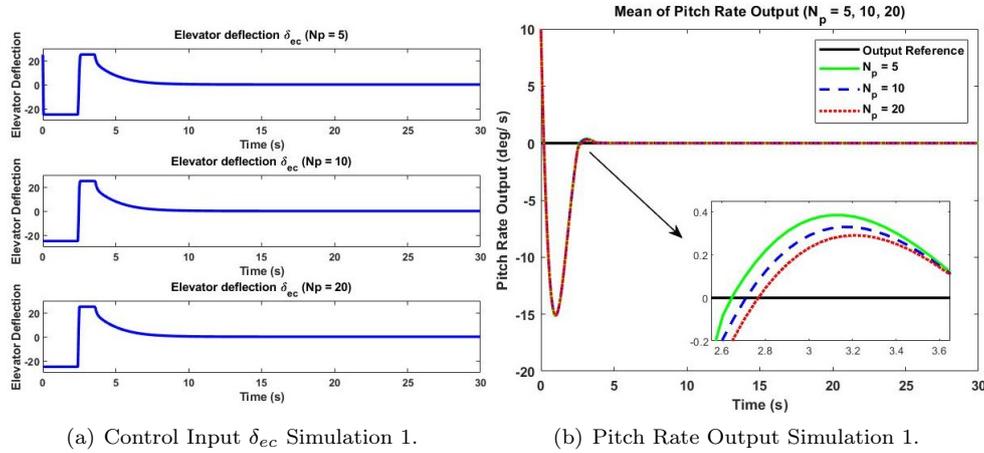


Figure 3: Control Input & Pitch Rate Output for Different Prediction Horizon ($N_p = 5, 10, 20$).

Based on Figure 3(b), different prediction horizon values N_p affect the system’s output response towards the given reference. For a smaller prediction horizon value $N_p = 5$, from the simulation, it looks longer to reach the reference pitch rate. To find out the difference in system response with different prediction horizon values N_p , the mean of MAE for each prediction horizon value is given as follows.

Prediction Horizon (N_p)	MAE (degree/ s)
2	0,79424
5	0,78501
8	0,78310
10	0,78171
15	0,78150
20	0,78134

Table 2: MAE for Different Prediction Horizon N_p .

5.2 Simulation with various orders of the polynomial

The second simulation used different Hermite polynomial orders ($r = 2, r = 5, r = 8$). This simulation used the same prediction horizon value, $N_p = 20$. The control input δ_{ec} and the mean pitch rate output are given by the following figures.

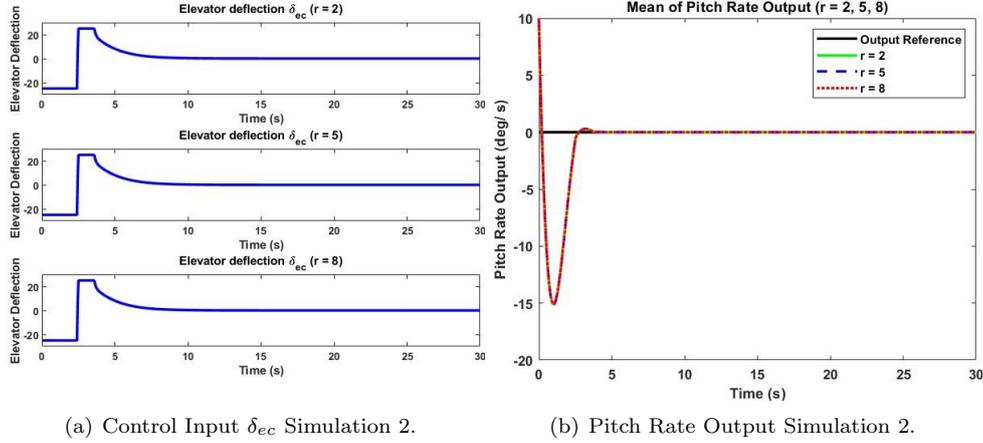


Figure 4: Control Input & Pitch Rate Output for Different Polynomial Orders ($r = 2, 5, 8$).

From Figure 4(b), it can be seen that the pitch rate can approach the given reference value for all polynomial order values. The simulation obtains the same pitch rate output for 30 seconds simulation time by using the polynomial order $r = 2, r = 5, r = 8$. Based on the simulation results, the same mean of the MAE value is obtained for different polynomial orders ($r = 2, r = 5, r = 8$). The mean of MAE and computation time for different polynomial orders are given in the following table.

Polynomial Orders (r)	MAE (degree/s)	Computation Time (s)
2	0, 78134	6, 10266
3	0, 78134	6, 41526
4	0, 78134	7, 65380
5	0, 78134	7, 93255
6	0, 78134	11, 39271
7	0, 78134	26, 17483
8	0, 78134	49, 87485

Table 3: MAE and Computation Time for Different Polynomial Orders r .

6 Conclusion

In this study, Polynomial Chaos and MPC are implemented for controlling the longitudinal motion of the F-16 aircraft, which has uncertainty in the parameter system (Δ). The contribution of this paper is the random variables Δ are assumed to be Gaussian distributed. The simulation results indicated that the Polynomial Chaos and MPC methods could be implemented properly for the linear model of the F-16 aircraft with uncertain stochastic parameters. This can be seen from the pitch rate output, which can follow and satisfy the given reference ($q_{ref} = 0^\circ/s$). Based on the simulation results for different prediction horizon values (N_p), a larger N_p gives a better pitch rate output response and a smaller mean of MAE. Meanwhile, different orders of the Hermite polynomial do not

significantly influence the result of control input and the pitch rate output. However, a higher order of the Hermite Polynomial requires a longer computation time.

Acknowledgment

This research was funded by the Directorate of Research and Community Service (DRPM), Sepuluh Nopember Institute of Technology under contract No. 1691/PKS/ITS/2023.

References

- [1] D. Mooney and R. Swift. *A Course in Mathematical Modeling*. American Mathematical Society, Providence, 2021.
- [2] R. L. Devaney. *A First Course in Chaotic Dynamical Systems: Theory and Experiment*. CRC Press, Florida, 2018.
- [3] R. Bhattacharya. Robust LQR Design for Systems with Probabilistic Uncertainty. *International Journal of Robust and Nonlinear Control* **29** (2019) 3217–3237.
- [4] J. Fisher and R. Bhattacharya. Linear Quadratic Regulation of Systems with Stochastic Parameter Uncertainties. *Automatica* **45** (12) (2009) 2831–2841.
- [5] R. Bhattacharya. Robust State Feedback Control Design with Probabilistic System Parameters. In: *53rd IEEE Conference on Decision and Control*. IEEE, New Jersey, 2014, 2828–2833.
- [6] V. Tadiparthi and R. Bhattacharya. Robust LQR for Uncertain Discrete-time Systems Using Polynomial Chaos. In: *2020 American Control Conference (ACC)*. IEEE, New Jersey, 2020, 4472–4477.
- [7] T. Asfihani, S. Subchan, D. Rosyid and A. Sulisetyono. Dubins Path Tracking Controller of USV using Model Predictive Control in Sea Field. *J. Eng. Appl. Sci.* **14** (20) (2019) 7778–7787.
- [8] E. Apriliani, S. Subchan, F. Yunaini and S. Hartini. Estimation and Control Design of Mobile Robot Position. *Far East Journal of Mathematical Sciences* **77** (1) (2013) 115.
- [9] T. Herlambang, S. Subchan, H. Nurhadi and D. Adzkiya. Motion Control Design of UN-USAITS AUV Using Sliding PID. *Nonlinear Dynamics and Systems Theory* **20** (1) (2020) 51–60.
- [10] T. Asfihani, M. K. Mufidah, S. Subchan and D. Adzkiya. Missile Control Design for Moving Target Using Model Predictive Control. *Journal of Physics: Conference Series*. IOP Publishing, Bristol, 2020, 012069.
- [11] B. L. Stevens, F. L. Lewis and E. N. Johnson. *Aircraft Control and Simulation: Dynamics, Controls Design, and Autonomous Systems*. John Wiley & Sons, New Jersey, 2015.
- [12] L. T. Nguyen. *Simulator Study of Stall/Post-Stall Characteristics of a Fighter Airplane with Relaxed Longitudinal Static Stability*. National Aeronautics and Space Administration, Washington D.C., 1979.
- [13] D. McLean. *Automatic Flight Control Systems*. Prentice Hall, New Jersey, 1990.
- [14] M. V. Cook. *Flight Dynamics Principles: a Linear Systems Approach to Aircraft Stability and Control*. Butterworth-Heinemann, Oxford, 2012.
- [15] S. Hsu and R. Bhattacharya. Design of Linear Parameter Varying Quadratic Regulator in Polynomial Chaos Framework. *International Journal of Robust and Nonlinear Control* **30** (16) (2020) 6661–6682.

- [16] Z. Li and J. Sun. Disturbance Compensating Model Predictive Control with Application to Ship Heading Control. *IEEE Transactions on Control Systems Technology* **20** (1) (2011) 257–265.
- [17] E. Camacho and C. Bordons. *Model Predictive Control (Advanced Textbooks in Control and Signal Processing)*. Springer, London, 1999.



Singular Reaction-Diffusion System Arising from Quenching

Samiha Djemai and Salim Mesbahi*

LMFN Laboratory, Department of Mathematics, Faculty of Sciences, Ferhat Abbas University, Setif, Algeria.

Received: April 26, 2023; Revised: November 25, 2023

Abstract: In this paper, we study a singular parabolic reaction-diffusion system with positive Dirichlet boundary conditions. It is shown that certain conditions are sufficient to guarantee finite-time quenching and global existence of solutions. This system appears in the modeling of the quenching phenomena.

Keywords: *reaction-diffusion system; quenching; singular parabolic equations.*

Mathematics Subject Classification (2010): 35K57, 35K67, 93A30.

1 Introduction

Quenching refers to the process of rapidly cooling a material from a high temperature to a lower temperature. This is done to alter the material's physical or mechanical properties such as hardness or strength. The rapid cooling prevents the material from undergoing a gradual cooling process, which would allow the material to form larger crystals that could weaken the material's structure. Quenching can be accomplished using different methods, including immersion in water, oil, or air, depending on the desired outcome. The study of this important phenomenon began in 1975 with a paper by Kawarada [5], where he studied a model in one space dimension. That paper was an introduction to the large-scale studies of the quenching problem by many researchers in several scientific fields. For a detailed survey, we refer to Chan [3], Levine [7], Rouabah et al. [13], Zouaoui et al. [20].

By using reaction-diffusion models, researchers can simulate the behavior of quenching processes and predict the resulting microstructure and mechanical properties of the metal. This can help in the design of new quenching techniques and in the optimization of

* Corresponding author: <mailto:salimbra@gmail.com>

existing ones. For more research on the phenomenon of quenching via reaction-diffusion systems, we refer the readers to Bonis [2], Ji et al. [4], Mesbahi [8], Mu et al. [11], Pei and Li [12], Salin [14–16], Wang [17], Zheng and Song [18], Zheng and Wang [19] and the references therein, where we will also find, in addition to the results by Mesbahi [9] and [10], many theoretical and numerical methods frequently used to study such problems.

In biology, quenching is a process that involves the rapid cooling of a sample in order to interrupt or halt certain biological processes. This procedure has several uses, including stopping metabolic processes and preserving metabolite profile of a sample in metabolomics. Protein synthesis and degradation can also be stopped for protein level and modification analysis in cells or tissues, while RNA in cells or tissues can be preserved for the analysis of gene expression. Moreover, microbial cultures can be preserved for long-term storage or transport by rapidly cooling them to halt growth and metabolic activity.

Quenching has many applications in medicine. One common medical application of quenching is cryotherapy, where extreme cold to treat disease or injury is used. This can include using liquid nitrogen to freeze and destroy cancerous tissue, or the use of ice packs to reduce swelling and inflammation. Another application is controlling the release of drugs from drug delivery systems. Rapid cooling of the system can halt or slow down drug release, enabling sustained release over time. Furthermore, quenching can aid in the preservation of biological samples such as blood or tissue samples for analysis or storage. Rapid cooling can prevent degradation of the sample and preserve its integrity for later use.

Quenching is also an important process in the manufacture of contact lenses. Typically, after the lenses are shaped, they undergo thermal quenching by being immersed in cold water. This process helps in solidifying their structure and preventing any deformation or distortion during handling and further processing. Furthermore, it enhances the mechanical and optical features of the lenses making them stronger, more resistant to damage, and long-lasting. Chemical quenching is also used by manufacturers to adjust the properties of the lenses. For instance, to crosslink the polymer chains in the lenses or to enhance their strength and flexibility. For better understanding, we refer to Barka et al. [1], Khurshid et al. [6].

In this work, we are interested in the study of the following reaction-diffusion system with general singular terms and positive Dirichlet boundary conditions that can be applied to the quenching phenomenon:

$$\begin{cases} (u_1)_t - \Delta u_1 = -f_1(u_2) & \text{in } (0, T) \times \Omega, \\ \vdots & \vdots \\ (u_{m-1})_t - \Delta u_{m-1} = -f_{m-1}(u_m) & \text{in } (0, T) \times \Omega, \\ (u_m)_t - \Delta u_m = -f_m(u_1) & \text{in } (0, T) \times \Omega, \\ u_1 = u_2 = \dots = u_m = 1 & \text{on } (0, T) \times \partial\Omega, \\ u_1(0, x) = u_{10}(x), \dots, u_m(0, x) = u_{m0}(x) & \text{in } \Omega, \end{cases} \quad (1)$$

where $\Omega \subset \mathbb{R}^N$ ($N \geq 2$) is a bounded domain with smooth boundary. The functions f_j ($1 \leq j \leq m$) are positive on $(0, 1]$. The initial data satisfy

$$\begin{cases} u_{10}, u_{20}, \dots, u_{m0} \in C^2(\Omega) \cap C^1(\overline{\Omega}), \\ u_{j0} = 1, \text{ for all } 1 \leq j \leq m, \text{ on } \partial\Omega, \\ 0 < u_{j0} \leq 1, \text{ for all } 1 \leq j \leq m, \text{ in } \overline{\Omega}. \end{cases} \quad (2)$$

The rest of this paper is organized as follows. In the next section, we state our main results. In the third section, we prove some important preliminary results. The fourth section is devoted to the proof of the main results. The paper ends with a concluding remarks and perspectives.

2 Statement of Main Results

2.1 Assumptions

For this model, the finite-time quenching phenomena are caused by singular nonlinearities in the absorption terms of (1).

Definition 2.1 We say the solution (u_1, \dots, u_m) of problem (1) quenches if (u_1, \dots, u_m) exists in the classical sense and is positive for all $0 \leq t < T$, and also satisfies $\inf_{t \rightarrow T} \min_{x \in [0,1]} \{(u_1(t, x), \dots, u_m(t, x))\} = 0$. In this case, T is called quenching time.

To study problem (1), we also assume that the positive functions $f_j : (0, 1] \rightarrow (0, +\infty)$, $1 \leq j \leq m$, satisfy the following simple assumptions which allow them to be chosen from a wide range:

(H₁) The functions f_j , $1 \leq j \leq m$, are locally Lipschitz on $(0, 1]$,

(H₂) $f'_j(s) < 0$ on $(0, 1]$ for all $1 \leq j \leq m$,

(H₃) $\lim_{s \rightarrow 0^+} f_j(s) = +\infty$ for all $1 \leq j \leq m$.

In order to state our results more conveniently, we denote by φ the first eigenfunction associated with the first eigenvalue λ_1 of the problem

$$\begin{cases} \Delta\varphi + \lambda\varphi = 0 & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega, \end{cases}$$

normalized by $\int_{\Omega} \varphi(x)dx = 1$, with $\varphi(x) > 0$ in Ω .

2.2 The main results

The following theorem gives us a sufficient condition for finite-time quenching.

Theorem 2.1 *Under hypotheses (H₁) – (H₃), the solution of problem (1) quenches in finite time for any initial data provided that λ_1 is small enough.*

Many quenching studies confirm that time-derivatives blow-up while the solution itself remains bounded. We refer, for example, to Chan [3] and Kawarada [5]. Throughout this paper, without any special explanation, we assume that the initial data u_{10}, \dots, u_{m0} satisfy

$$\Delta u_{10} - f_1(u_{20}) < 0, \dots, \Delta u_{m0} - f_m(u_{10}) < 0 \text{ in } \Omega. \tag{3}$$

Thus, the global existence of solutions can be described by the following theorem.

Theorem 2.2 *If the diameter of Ω is small enough and the initial data satisfies $0 < \varepsilon \leq u_{10}, \dots, u_{m0} \leq 1$ in Ω , then under hypotheses (H₁)–(H₃), the solution of problem (1) does not quench in finite time. In this case, we say that the solution (u_1, \dots, u_m) exists globally.*

3 Preliminary Results

We will prove two important lemmas which we will use to prove our main results.

Lemma 3.1 *Assume that the initial data satisfy (3), then $(u_1)_t, \dots, (u_m)_t < 0$ in $(0, T) \times \Omega$.*

Proof. Let $I_j(t, x) = (u_j)_t(t, x)$ for all $1 \leq j \leq m$ and $(t, x) \in (0, T) \times \Omega$. Differentiating system (1) with respect to t , we have

$$\begin{cases} \frac{\partial}{\partial t} I_1 = \Delta (u_1)_t - (u_2)_t f'_1(u_2) & \text{in } (0, T) \times \Omega, \\ \vdots & \vdots \\ \frac{\partial}{\partial t} (I_m)(x, t) = \Delta (u_m)_t - (u_1)_t f'_m(u_1) & \text{in } (0, T) \times \Omega, \\ I_1 = I_2 = \dots = I_m = 0 & \text{on } (0, T) \times \partial\Omega, \\ I_j(0, x) < 0, \text{ for all } 1 \leq j \leq m & \text{in } \Omega, \end{cases}$$

which, after simplification, gives

$$\begin{cases} \frac{\partial}{\partial t} I_1 - \Delta I_1 = -I_2 f'_1(u_2) & \text{in } (0, T) \times \Omega, \\ \vdots & \vdots \\ \frac{\partial}{\partial t} I_m - \Delta I_m = -I_1 f'_m(u_1) & \text{in } (0, T) \times \Omega, \\ I_1 = I_2 = \dots = I_m = 0 & \text{on } (0, T) \times \partial\Omega, \\ I_j(0, x) < 0, \text{ for all } 1 \leq j \leq m & \text{in } \Omega. \end{cases} \tag{4}$$

By the comparison principle, we have, for all $(t, x) \in (0, T) \times \Omega$,

$$I_j(t, x) = (u_j)_t(t, x) < 0 \text{ for all } 1 \leq j \leq m.$$

This shows that u_1, \dots, u_m are strictly decreasing in time.

Now, we consider the radial solutions of problem (1) on $\Omega = B_R = \{x \in \mathbb{R}^N : |x| < R\}$.

Lemma 3.2 *Let (u_1, \dots, u_m) be the global solution of problem (1) with $(u_{10}, \dots, u_{m0}) \equiv (1, \dots, 1)$, $u_1, \dots, u_m \geq b$ in $(0, \infty) \times \overline{B}_R$ for some $b \in (0, 1)$. Then (u_1, \dots, u_m) approaches uniformly from above to a solution (U_1, \dots, U_m) of the steady-state problem*

$$\begin{cases} \Delta U_1 = f(U_2) & \text{in } B_R, \\ \vdots & \vdots \\ \Delta U_{m-1} = f(U_m) & \text{in } B_R, \\ \Delta U_m = f(U_1) & \text{in } B_R, \\ U_1 = U_2 = \dots = U_m = 1 & \text{on } \partial B_R. \end{cases} \tag{5}$$

Proof. Since $(1, \dots, 1)$ is a strict super-solution of problem (1), by Lemma 3.1, we have $(u_1)_t, \dots, (u_m)_t < 0$ in $(0, \infty) \times B_R$. Define the functions

$$Q_j(t, x) = \int_{B_R} G(x, y) u_j(t, y) dy, \text{ in } (0, \infty) \times B_R, \text{ for all } 1 \leq j \leq m,$$

where $G(x, y)$ is Green’s function associated with the operator $-\Delta$ on B_R under Dirichlet boundary conditions. Hence

$$\begin{aligned} \frac{\partial}{\partial t}(Q_1) &= \int_{B_R} G(x, y) (u_1)_t(t, y) dy \\ &= \int_{B_R} G(x, y) \Delta u_1(t, y) dy - \int_{B_R} G(x, y) f_1(u_2(t, y)) dy, \\ &\vdots \\ \frac{\partial}{\partial t}(Q_m) &= \int_{B_R} G(x, y) (u_m)_t(t, y) dy \\ &= \int_{B_R} G(x, y) \Delta u_m(t, y) dy - \int_{B_R} G(x, y) f_m(u_1(t, y)) dy, \end{aligned}$$

this gives us

$$\begin{aligned} \frac{\partial}{\partial t}(Q_1) &= 1 - u_1(x, y) - \int_{B_R} G(x, y) f_1(u_2(t, y)) dy, \\ &\vdots \\ \frac{\partial}{\partial t}(Q_m) &= 1 - u_m(x, y) - \int_{B_R} G(x, y) f_m(u_1(t, y)) dy. \end{aligned}$$

It follows from $(u_j)_t < 0$ for all $1 \leq j \leq m$, that

$$G(x, y) f_1(u_2(t, y)), \dots, G(x, y) f_{m-1}(u_m(t, y)) \text{ and } G(x, y) f_m(u_1(t, y))$$

are nondecreasing with respect to t . According to the monotone convergence theorem with

$$b \leq U_j(x) = \lim_{t \rightarrow 0} u_j(t, x) \text{ for all } 1 \leq j \leq m,$$

we have

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{\partial}{\partial t}(Q_1) &= 1 - U_1(x) - \int_{B_R} G(x, y) f_1(U_2(y)) dy, \\ &\vdots \\ \lim_{t \rightarrow 0} \frac{\partial}{\partial t}(Q_m) &= 1 - U_m(x) - \int_{B_R} G(x, y) f_m(U_1(y)) dy. \end{aligned}$$

Furthermore, since Q_1, \dots, Q_m are bounded, $(Q_1)_t, \dots, (Q_m)_t \leq 0$, and by $(u_1)_t, \dots, (u_m)_t < 0$, we have

$$\lim_{t \rightarrow 0} \frac{\partial}{\partial t}(Q_j) = 0 \text{ for all } 1 \leq j \leq m,$$

which yields

$$\begin{aligned} U_1(x) &= 1 - \int_{B_R} G(x, y) f_1(U_2(y)) dy, \\ &\vdots \\ U_m(x) &= 1 - \int_{B_R} G(x, y) f_m(U_1(y)) dy, \end{aligned}$$

and therefore (U_1, \dots, U_m) is a solution of problem (5), and the uniform convergence is ensured by Dini's theorem.

4 Proofs of the Main Results

Proof. [of Theorem 2.1] Let (u_1, \dots, u_m) be the solution of problem (1) with the maximal existence time T . By the maximum principle, we have $0 \leq u_j \leq 1$ for all $1 \leq j \leq m$, in $(0, T) \times \Omega$. Let

$$\psi_j(t) = \int_{\Omega} (1 - u_j) \varphi dx \quad \text{for all } 1 \leq j \leq m, \quad t \in [0, T) \quad (6)$$

and

$$\Psi(t) = \psi_1(t) + \dots + \psi_m(t), \quad t \in [0, T). \quad (7)$$

By hypotheses $(H_1) - (H_3)$ and the corresponding Taylor expansions, we can easily get

$$f_1(u_2) \geq \delta(1 - u_2) + c_1, \dots, f_m(u_1) \geq \delta(1 - u_1) + c_m, \quad (8)$$

where δ, c_1, \dots, c_m are positive constants determined by $f_1(u_2), \dots, f_m(u_1)$.

By a straight-forward computation and (8), we have

$$\begin{aligned} \psi_1'(t) &= - \int_{\Omega} \Delta u_1 \varphi dx + \int_{\Omega} f_1(u_2) \varphi dx \\ &= \int_{\Omega} \Delta(1 - u_1) \varphi dx + \int_{\Omega} f_1(u_2) \varphi dx \\ &\geq -\lambda_1 \int_{\Omega} (1 - u_1) \varphi dx + \delta \int_{\Omega} (1 - u_2) \varphi dx + c_1 \int_{\Omega} \varphi dx \\ &= -\lambda_1 \psi_1(t) + \delta \psi_2(t) + c_1. \end{aligned}$$

In the same way, with $\psi_2(t), \dots, \psi_m(t)$, we finally get the following inequalities:

$$\begin{aligned} \psi_1'(t) &\geq -\lambda_1 \psi_1(t) + \delta \psi_2(t) + c_1, \\ &\vdots \\ \psi_m'(t) &\geq -\lambda_1 \psi_m(t) + \delta \psi_1(t) + c_m. \end{aligned}$$

Using (7), we get

$$\Psi'(t) \geq (\delta - \lambda_1) \Psi(t) + C, \quad \text{with } C = c_1 + \dots + c_m. \quad (9)$$

Since $0 \leq u_j \leq 1$ in $(0, T) \times \Omega$, then $0 \leq 1 - u_j \leq 1$ in $(0, T) \times \Omega$, which clearly implies by (6) that $0 \leq \psi_j(t) \leq 1$ for all $1 \leq j \leq m$, consequently, $1 \leq \Psi(t) \leq m$. Since λ_1 is small enough, it is obvious that $(\delta - \lambda_1) \Psi(t) + C > 0$. Then, by (9), we have

$$\frac{d\Psi}{(\delta - \lambda_1) \Psi(t) + C} \geq dt, \quad t \in [0, T),$$

which gives, by integration from 0 to T ,

$$t \leq \begin{cases} \frac{1}{\delta - \lambda_1} \log \left(\frac{(\delta - \lambda_1) \Psi(t) + C}{(\delta - \lambda_1) \Psi(0) + C} \right) & \text{if } \delta \neq \lambda_1, \\ \frac{1}{C} (\Psi(t) - \Psi(0)) & \text{if } \delta = \lambda_1. \end{cases} \quad (10)$$

Now, letting $t \rightarrow T^-$ in (10) and combining $\lim_{t \rightarrow T^-} \Psi(t) \leq m$, we get

$$T \leq \begin{cases} \frac{1}{\delta - \lambda_1} \log \left(\frac{m(\delta - \lambda_1) + C}{(\delta - \lambda_1)\Psi(0) + C} \right) & \text{if } \delta \neq \lambda_1, \\ \frac{1}{C} (m - \Psi(0)) & \text{if } \delta = \lambda_1. \end{cases} \tag{11}$$

Since $1 \leq \Psi(t) \leq m$, we can easily arrive at the positivity of the right-hand side of (11), which shows finite time quenching of the solutions in system (1). This ends the proof of Theorem 2.1.

Proof. Consider the auxiliary system

$$\begin{cases} \frac{\partial}{\partial t} \bar{u}_1 = \Delta \bar{u}_1 - f(\bar{u}_2) & \text{in } (0, T) \times \Omega, \\ \vdots & \vdots \\ \frac{\partial}{\partial t} \bar{u}_m = \Delta \bar{u}_m - f(\bar{u}_1) & \text{in } (0, T) \times \Omega, \\ \bar{u}_1 = \dots = \bar{u}_m = 1 & \text{on } (0, T) \times \partial\Omega, \\ \bar{u}_1(0, x) = \dots = \bar{u}_m(0, x) = 1 & \text{in } \bar{\Omega}. \end{cases}$$

By the comparison principle, we have $u_j \leq \bar{u}_j$ for all $1 \leq j \leq m$.

We first consider the following system:

$$\begin{cases} \Delta \bar{u}_1^* = f_1(1) & \text{in } B_R, \\ \vdots & \vdots \\ \Delta \bar{u}_m^* = f_m(1) & \text{in } B_R, \\ \bar{u}_1^* = \dots = \bar{u}_m^* = 1 & \text{on } \partial B_R. \end{cases}$$

By Green’s function, the solution is $(\bar{u}_1^*, \dots, \bar{u}_m^*)$ denoted as follows:

$$\bar{u}_j^* = \frac{f_j(1) \left(|x|^2 - R^2 \right)}{2N} + 1, \quad 1 \leq j \leq m$$

and

$$\min \bar{u}_j^* = \frac{-f_j(1) R^2}{2N} + 1, \quad 1 \leq j \leq m.$$

Clearly, $(\bar{u}_1^*, \dots, \bar{u}_m^*)$ is a super solution of (1). By Lemma 3.2, the solution (u_1, \dots, u_m) of (1) is global only if $\bar{u}_1^*, \dots, \bar{u}_m^* > 0$.

5 Concluding Remarks and Perspectives

This contribution advances mathematical research on quenching phenomena. The results of this study can be used to study other singular reaction-diffusion phenomena. We managed to overcome some difficulties and achieved very important results. This leads us to think more about the problem and do further theoretical and numerical research under other conditions. These efforts will advance quenching technology and modeling in many scientific fields.

References

- [1] B. Barka et al. Thermophysical Behavior of Polycarbonate: Effect of Free Quenching above and below the Glass Transition Temperature. *Advanced Materials Research* **1174** (2022) 123–136.
- [2] I. de Bonis. *Singular elliptic and parabolic problems: existence and regularity of solutions*, PhD Thesis. Sapienza University, Rome, 2015.
- [3] C. Y. Chan. Recent advances in quenching phenomena. *Proceedings of Dynamic Systems and Applications*, Atlanta, GA, **2** (1995) 107–113, Dynamic, Atlanta, GA (1996).
- [4] R. H. Ji, C. Y. Qu and L. D. Wang. Simultaneous and non-simultaneous quenching for coupled parabolic system. *Appl. Anal.* **94** (2015) 234–251.
- [5] H. Kawarada. On solutions of initial-boundary problem for $u_t = u_{xx} + \frac{1}{(1-u)}$. *Publ. Res. Inst. Math. Sci.* **10** (1974/75), 729–736.
- [6] Z. Khurshid et al. *Novel Techniques of Scaffold Fabrication for Bioactive Glasses*. In Woodhead Publishing Series in Biomaterials, Biomedical, Therapeutic and Clinical Applications of Bioactive Glasses, Publisher: Elsevier, UK, (2019) 497–519. ISBN 9780081021965.
- [7] H. A. Levine. *Advances in quenching, Nonlinear diffusion equations and their equilibrium states*. Gregynog, (1989) 319–346, Progr. Nonlinear Differential Equations Appl., Birkhauser Boston, Boston, MA, 1992.
- [8] S. Mesbahi. On the existence of weak solutions for a class of singular reaction-diffusion systems. *Hacet. J. Math. Stat.* **51** (3) (2022) 757–774.
- [9] A. Mesbahi and S. Mesbahi. On the Existence of Periodic Solutions of a Degenerate Parabolic Reaction-Diffusion Model. *Nonlinear Dyn. Syst. Theory* **22** (2) (2022) 197–205.
- [10] S. Mesbahi. *Analyse mathématique de systèmes de réaction-diffusion quasi-linéaires*. Editions Universitaires Européennes, 2019. ISBN-13 : 978-6138492917.
- [11] C. Mu, S. Zhou and D. Liu. Quenching for a reaction-diffusion system with logarithmic singularity. *Nonlinear Anal.* **71** (2009) 5599–5605.
- [12] H. Pei, Z. Li. Quenching for a parabolic system with general singular terms. *J. Nonlinear Sci. Appl.* **9** (8) (2016) 5281–5290.
- [13] F. Rouabah, D. Dadache and N. Haddaoui. Thermophysical and Mechanical Properties of Polystyrene: Influence of Free Quenching. *International Scholarly Research Notices* **2012** (Article ID 161364) (2012) 8 pages.
- [14] T. Salin. *Quenching and blow-up problems for reaction diffusion equations*. Helsinki University of Technology Institute of Mathematics Research Reports **A466**, 2004.
- [15] T. Salin. Quenching-rate estimate for a reaction diffusion equation with weakly singular reaction term, *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.* **11** (4) (2004).
- [16] T. Salin. On quenching with logarithmic singularity. *Nonlinear Anal. TMA* **52** (2003) 261–289.
- [17] Q. Wang. Quenching Phenomenon for a Parabolic MEMS Equation. *Chin. Ann. Math. Ser. B* **39** (1) (2018) 129–144.
- [18] S. N. Zheng and X. F. Song. Quenching rates for heat equations with coupled singular nonlinear boundary flux. *Sci.China Ser. A* **51** (2008) 1631–1643.
- [19] S. N. Zheng and W. Wang. Non-simultaneous versus simultaneous quenching in a coupled nonlinear parabolic system, *Nonlinear Anal.* **69** (2008), 2274–2285.
- [20] F. Zouaoui, F. Rouabah, Y. Nouar and M. Fois. Effect of heat treatment on the thermophysical properties of copper-powder-filled polycarbonate and polycarbonate containing paraffin. *Journal of Polymer Engineering* **39** (8) (2019) 729–735.



Maximum and Anti-Maximum Principles for Boundary Value Problems for Ordinary Differential Equations in Neighborhoods of Simple Eigenvalues

P. W. Eloe^{1*} and J. T. Neugebauer²

¹ *Department of Mathematics, University of Dayton, Dayton, Ohio 45469 USA.*

² *Department of Mathematics and Statistics, Eastern Kentucky University, Richmond, Kentucky 40475 USA.*

Received: October 7, 2022; Revised: October 20, 2023

Abstract: It has been shown that, under suitable hypotheses, for boundary value problems of the form $Ly + \lambda y = f$, $BCy = 0$, where L is a linear differentiable operator and BC denotes the linear boundary operator, there exists $\Lambda > 0$ such that $f \geq 0$ implies $\lambda y \geq 0$ for $\lambda \in [-\Lambda, \Lambda] \setminus \{0\}$, where y is the unique solution of $Ly + \lambda y = f$, $BCy = 0$. So, the boundary value problem satisfies a maximum principle for $\lambda \in [-\Lambda, 0)$ and the boundary value problem satisfies an anti-maximum principle if $\lambda \in (0, \Lambda]$. Moreover, this information is provided in the one inequality, $\lambda y \geq 0$. In this study, we shall provide suitable hypotheses such that for boundary value problems of the form $Ly + \beta y' = f$, $BCy = 0$, where L is an ordinary differentiable operator and BC denotes the boundary operator, there exists $\mathcal{B} > 0$ such that $f \geq 0$ implies $\beta y' \geq 0$ for $\beta \in [-\mathcal{B}, \mathcal{B}] \setminus \{0\}$, where y is the unique solution of $Ly + \beta y' = f$, $BCy = 0$. Under suitable boundary conditions, one obtains sign properties on solutions and derivatives of solutions. Two examples satisfying the suitable hypotheses are provided and one application of monotone methods is provided to illustrate an application of the main result.

Keywords: *maximum principle; anti-maximum principle; ordinary differential equation; boundary value problem.*

Mathematics Subject Classification (2010): 34B08; 34B18; 34B27; 34L15.

* Corresponding author: <mailto:peloe1@udayton.edu>

1 Introduction

The maximum principle is an important tool in the study of differential equations and we refer the reader to the well-known book [14] for many applications. For example, for the specific boundary value problem for a second order ordinary differential equation, $y'' + \lambda y = f$, $y'(0) = 0$, $y'(1) = 0$, if $\lambda < 0$, then this boundary value problem satisfies a maximum principle. In particular, for $f \in C[0, 1]$, the boundary value problem is uniquely solvable and f nonnegative implies y is nonpositive, where y is the unique solution associated with f . In the study of boundary value problems for ordinary differential equations, the maximum principle implies that the associated Green's function is of constant sign, and in this case, the Green's function is nonpositive on $(0, 1) \times (0, 1)$.

Clément and Peletier [8] were the first to discover an anti-maximum principle. They were primarily interested in partial differential equations, but they illustrated the anti-maximum principle with the boundary value problem, $y'' + \lambda y = f$, $y'(0) = 0$, $y'(1) = 0$, $0 < \lambda < \frac{\pi^2}{4}$. For this particular boundary value problem, if $0 < \lambda < \frac{\pi^2}{4}$, if $f \in C[0, 1]$, the boundary value problem is uniquely solvable and f nonnegative implies y is nonnegative, where y is the unique solution associated with f .

Since the publication of [8], there have been many studies of boundary value problems with parameter and the change of behavior from maximum to anti-maximum principles as a function of the parameter. In the case of partial differential equations, we refer to [1, 2, 7, 9, 10, 12, 13, 15]. In the case of ordinary differential equations, we refer to [3–6, 16]. In this paper, we shall continue to study the change in behavior of boundary value problems for ordinary differential equations, with respect to maximum and anti-maximum principles, through simple eigenvalues.

In an interesting study produced in [7], those authors began with a differential equation

$$y''(t) + \lambda y(t) = f(t), \quad t \in [0, 1], \quad (1)$$

and considered either periodic boundary conditions or Neumann boundary conditions. Key to their argument is that for $f = 0$, at $\lambda = 0$, the boundary value problem, (1) with periodic or Neumann boundary conditions, is at resonance since constant functions are nontrivial solutions. Further, $\lambda = 0$ is a simple eigenvalue and the eigenspace is $\langle 1 \rangle$, where $\langle 1 \rangle$ denotes the linear span of the 1 function. Employing the resolvent, the inverse of $(D^2 + \lambda I)$ for $\lambda \neq 0$, under the imposed boundary conditions, if it exists, and the partial resolvent for $\lambda = 0$, and under the assumption that $f \geq 0$ (with $f \in \mathcal{L}[0, 1]$), the authors in [7] obtained sufficient conditions to construct an interval $[-\Lambda, \Lambda]$, $\Lambda > 0$, a constant $K > 0$, independent of f such that

$$\lambda y(t) \geq K|f|_1, \quad \lambda \in [-\Lambda, \Lambda] \setminus \{0\}, \quad 0 \leq t \leq 1,$$

where $|f|_1 = \int_0^1 |f(s)| ds$. With this one inequality, the authors showed that for $\Lambda \leq \lambda < 0$, the boundary value problem, (1) with periodic or Neumann boundary conditions, satisfies a maximum principle and for $0 < \lambda \leq \Lambda$, the boundary value problem (1) with periodic or Neumann boundary conditions, satisfies an anti-maximum principle. They proceeded to produce many nice examples in that paper.

Consider the boundary value problem

$$y''(t) + \beta y'(t) = f(t), \quad t \in [0, 1], \quad (2)$$

$$y(0) = 0, \quad y'(0) = y'(1). \quad (3)$$

For $f = 0$, $\beta = 0$ is a simple eigenvalue and generates the eigenspace $\langle t \rangle$, the linear span of t . For $\beta \neq 0$,

$$G(\beta; t, s) = \begin{cases} \frac{e^{-\beta(1-s)} - e^{-\beta} e^{-\beta(t-s)}}{\beta(1-e^{-\beta})}, & 0 \leq t \leq s \leq 1, \\ \frac{e^{-\beta(1-s)} - e^{-\beta} e^{-\beta(t-s)}}{\beta(1-e^{-\beta})} + \frac{1-e^{-\beta(t-s)}}{\beta}, & 0 \leq s \leq t \leq 1, \end{cases} \tag{4}$$

is the Green’s function for the boundary value problem (2), (3). Note that

$$\beta G(\beta; t, s) > 0, \quad (t, s) \in (0, 1] \times [0, 1], \text{ and } \beta \frac{\partial}{\partial t} G(\beta; t, s) > 0, \quad [t, s] \in [0, 1] \times [0, 1].$$

So, if y denotes the solution of (2), (3), then $f \geq 0$ implies $\beta y' \geq 0$ and $\beta y \geq 0$. This observation indicates that the principle obtained in [7] can be extended to other order derivatives.

Our goal in this paper is to study boundary value problems for ordinary differential equations containing a parameter β such that $\beta = 0$ is a simple eigenvalue generating an eigenspace $\langle t - t_0 \rangle$ for some constant t_0 and modify the methods produced in [7]; in particular, we shall assume $f \geq 0$ and obtain sufficient conditions to construct an interval $[-\mathcal{B}, \mathcal{B}]$, $\mathcal{B} > 0$, a constant $K > 0$, independent of f , and an inequality

$$\beta y'(t) \geq K|f|_1, \quad \beta \in [-\mathcal{B}, \mathcal{B}] \setminus \{0\}, \quad 0 \leq t \leq 1, \tag{5}$$

where y is a unique solution of the boundary value problem associated with f . It will follow that if $0 < |\beta| \leq \mathcal{B}$, and if $f \geq 0$, then $\beta y' \geq 0$.

In Section 2, following the lead of [7], we shall define the concept of a strong signed maximum principle in y' . In Section 3, we shall obtain sufficient conditions for (5) and hence obtain sufficient conditions for adherence to a strong signed maximum principle in y' . In Section 4, we shall illustrate the main result, Theorem 3.1, with two examples. In each example, the boundary conditions are such that (5) generates a natural partial order in $C^1[0, 1]$.

We close in Section 5 with an application of a monotone method applied to a nonlinear problem related to one of the examples produced in Section 4. At $\beta = 0$, the problem is at resonance. The problem is shifted [11] by $\beta y'$ and $\beta > 0$ or $\beta < 0$ is chosen as a function of the monotonicity properties of the nonlinearity.

2 Strong Signed Maximum Principle

Assume \mathcal{A} is a linear operator with $\text{Dom}(\mathcal{A}) \subset C^1[0, 1]$ and $\text{Im}(\mathcal{A}) \subset C[0, 1]$. Let $Dy = y'$ for $y \in C^1[0, 1]$. The following definition is motivated by Definition 1 found in [7].

Definition 2.1 For $\beta \in \mathbb{R} \setminus \{0\}$, the operator $\mathcal{A} + \beta D$ satisfies a **signed maximum principle in Dy** if for each $f \in C[0, 1]$, the equation

$$(\mathcal{A} + \beta D)y = f, \quad y \in \text{Dom}(\mathcal{A}),$$

has a unique solution, y , and $f(t) \geq 0$, $0 \leq t \leq 1$ implies $\beta Dy(t) \geq 0$, $0 \leq t \leq 1$. The operator $\mathcal{A} + \beta D$ satisfies a **strong signed maximum principle in Dy** if $f(t) \geq 0$, $0 \leq t \leq 1$ and $f(t) > 0$ on some interval of positive length, implies $\beta Dy(t) > 0$, $0 < t < 1$.

Remark 2.1 Throughout this study, the phrases “maximum principle” or “anti-maximum principle” may be used loosely. If so, we mean the following. If $f \geq 0$ implies $y \leq 0$ (or $Dy \leq 0$), the phrase, maximum principle, may be used. This is precisely the case for the classical second order differential equation with Dirichlet boundary conditions. If $f \geq 0$ implies $y \geq 0$ (or $Dy \geq 0$), the phrase, anti-maximum principle, may be used. This is the case observed in [8] where the phrase, anti-maximum principle, was coined.

Remark 2.2 As pointed out in the Introduction, for the boundary value problem $y''(t) + \beta y'(t) = f(t)$, with boundary conditions (3), $f(t) \geq 0$, $0 \leq t \leq 1$ implies $\beta Dy(t) \geq 0$, $0 \leq t \leq 1$, and $\beta y(t) \geq 0$, $0 \leq t \leq 1$. In the application of the main theorem, Theorem 3.1, one only concludes (5). In the examples produced in Section 4, the boundary conditions are such that (5) implies further that for some $t_0 \in [0, 1]$, $\beta(t - t_0)y(t) \geq 0$, $0 \leq t \leq 1$. In particular, in each example, a signed maximum principle in Dy will generate a natural partial order on $C^1[a, b]$ in which monotone methods can be applied.

3 The Main Theorem

Let $C[0, 1]$ denote the Banach space of continuous real-valued functions defined on $[0, 1]$ with norm $|y|_0 = \max_{0 \leq t \leq 1} |y(t)|$ and let $C^1[0, 1]$ denote the Banach space of continuously differentiable real-valued functions defined on $[0, 1]$ with

$$\|y\| = \max\{|y|_0, |y'|_0\}.$$

Also, $C[0, 1] \subset \mathcal{L} = L^1[0, 1]$, and so, we shall also have use for $|f|_1 = \int_0^1 |f(s)| ds$. For $f \in \mathcal{L}$, set

$$\bar{f} = \int_0^1 f(t) dt,$$

and define

$$\tilde{C} \subset C[0, 1] = \{f \in C[0, 1] : \bar{f} = 0\}, \quad \tilde{\mathcal{L}} \subset \mathcal{L} = \{f \in L^1[0, 1] : \bar{f} = 0\}.$$

Let $t_0 \in \mathbb{R}$. Assume $\mathcal{A} : \text{Dom}(\mathcal{A}) \rightarrow \mathcal{L}$ denotes a linear operator satisfying

$$\text{Dom}(\mathcal{A}) \subset C^1[a, b] \quad \text{Ker}(\mathcal{A}) = \langle t - t_0 \rangle, \quad \text{Im}(\mathcal{A}) = \tilde{\mathcal{L}}, \quad (6)$$

where $\langle t - t_0 \rangle$ denotes the linear span of $t - t_0$. Assume further that for $\tilde{f} \in \tilde{\mathcal{L}}$, the problem $\mathcal{A}y = \tilde{f}$ is uniquely solvable with solution $y \in \text{Dom}(\mathcal{A})$ and such that $(y') = 0$. In particular, define

$$\text{Dom}(\tilde{\mathcal{A}}) = \{y \in \text{Dom}(\mathcal{A}) : (y') = 0\},$$

and then

$$\mathcal{A}|_{\text{Dom}(\tilde{\mathcal{A}})} : \text{Dom}(\tilde{\mathcal{A}}) \rightarrow \tilde{\mathcal{L}}$$

is one to one and onto. Moreover, if $\mathcal{A}\tilde{y} = \tilde{f}$ for $\tilde{f} \in \tilde{\mathcal{L}}$, $\tilde{y} \in \text{Dom}(\tilde{\mathcal{A}})$, assume there exists a constant $K_1 > 0$ depending only on \mathcal{A} such that

$$|\tilde{y}'|_0 \leq K_1 |\tilde{f}|_1. \quad (7)$$

For $f \in \mathcal{L}$, define

$$\tilde{f} = f - \bar{f},$$

and for $y \in \text{Dom}(\mathcal{A})$, define

$$\tilde{y} = y - \bar{y}'(t - t_0),$$

which implies

$$\tilde{y}' = y' - \bar{y}'.$$

Finally, assume there exists $\mathcal{A}' : \text{Dom}(\mathcal{A}') \rightarrow \mathcal{L}$ such that $\mathcal{A} = \mathcal{A}'D$. In this context, we rewrite

$$\mathcal{A}y + \beta y' = f, \quad y \in \text{Dom}(\mathcal{A}), \tag{8}$$

as

$$(\mathcal{A}' + \beta\mathcal{I})Dy = f, \quad Dy \in \text{Dom}(\mathcal{A}'). \tag{9}$$

Define $\text{Dom}(\tilde{\mathcal{A}}') = \{v \in \text{Dom}(\mathcal{A}') : \bar{v} = 0\} \subset C[0, 1]$ and it follows that

$$\mathcal{A}'|_{\text{Dom}(\tilde{\mathcal{A}}')} : \text{Dom}(\tilde{\mathcal{A}}') \rightarrow \tilde{\mathcal{L}}$$

is one to one and onto.

With the decompositions $\tilde{f} = f - \bar{f}$ and $\tilde{y} = y - \bar{y}'(t - t_0)$, it follows that $\tilde{f} \in \tilde{\mathcal{L}}$ and $\tilde{y} \in \text{Dom}(\tilde{\mathcal{A}}')$, or more appropriately, $D\tilde{y} \in \text{Dom}(\tilde{\mathcal{A}}')$. So, equation (8) or equation (9) decouples as follows:

$$\mathcal{A}'D\tilde{y} + \beta D\tilde{y} = (\mathcal{A}' + \beta\mathcal{I})D\tilde{y} = \tilde{f}, \tag{10}$$

$$\beta D\bar{y}'(t - t_0) = \beta \bar{y}' = \bar{f}. \tag{11}$$

Denote the inverse of $(\mathcal{A}' + \beta\mathcal{I})$, if it exists, by \mathcal{R}_β and denote the inverse of $\mathcal{A}'|_{\text{Dom}(\tilde{\mathcal{A}}')}$ by \mathcal{R}_0 . So, $\mathcal{R}_0 : \tilde{\mathcal{L}} \rightarrow C[0, 1]$ and

$$D\tilde{y} = \mathcal{R}_0\tilde{f} \text{ if, and only if, } \mathcal{A}'(D\tilde{y}) = \tilde{f}. \tag{12}$$

Note that (12) implies

$$D\tilde{y} = \mathcal{R}_0\mathcal{A}'D\tilde{y} \tag{13}$$

since $D\tilde{y} \in \text{Dom}(\tilde{\mathcal{A}}')$,

Since $\tilde{\mathcal{C}} \subset \tilde{\mathcal{L}}$, we can also consider $\mathcal{R}_0 : \tilde{\mathcal{C}} \rightarrow C[0, 1]$. Let

$$\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} = \sup_{|v|_0=1} |\mathcal{R}_0 v|_0, \quad v, \mathcal{R}_0 v \in \tilde{\mathcal{C}},$$

and

$$\|\mathcal{R}_0\|_{\tilde{\mathcal{L}} \rightarrow \tilde{\mathcal{C}}} = \sup_{|v|_1=1} |\mathcal{R}_0 v|_0, \quad v \in \tilde{\mathcal{L}}, \quad \tilde{\mathcal{R}}_0 v \in \tilde{\mathcal{C}}.$$

Since $D\tilde{y} \in \tilde{\mathcal{C}}$, it follows that $|\mathcal{R}_0 D\tilde{y}|_0 \leq \|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} |D\tilde{y}|_0$. Similarly, $\tilde{f} \in \tilde{\mathcal{L}}$ implies $|\mathcal{R}_0 \tilde{f}|_0 \leq \|\mathcal{R}_0\|_{\tilde{\mathcal{L}} \rightarrow \tilde{\mathcal{C}}} |\tilde{f}|_1$.

Theorem 3.1 *Assume $\mathcal{A} : \text{Dom}(\mathcal{A}) \rightarrow C[0, 1]$ denotes a linear operator satisfying (6) and (7), and assume that for $\tilde{f} \in \tilde{\mathcal{L}}$, the problem $\mathcal{A}y = \tilde{f}$ is uniquely solvable with solution $y \in \text{Dom}(\mathcal{A})$ such that $(\bar{y}') = 0$. Further, assume there exists $\mathcal{A}' : \text{Dom}(\mathcal{A}') \rightarrow C[0, 1]$ such that $\mathcal{A} = \mathcal{A}'D$. Assume $\mathcal{A}'|_{\text{Dom}(\tilde{\mathcal{A}}')} : \text{Dom}(\tilde{\mathcal{A}}') \rightarrow \tilde{\mathcal{L}}$ is one to one and*

onto. Then there exists $B_1 > 0$ such that if $0 < |\beta| \leq B_1$, then \mathcal{R}_β , the inverse of $(\mathcal{A}' + \beta\mathcal{I})$, exists. Moreover, if $\tilde{f} \in \tilde{L}$, $B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} < 1$ and $0 < |\beta| \leq B_1$, then

$$|\mathcal{R}_\beta \tilde{f}|_0 \leq \frac{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}{1 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}} |\tilde{f}|_1. \quad (14)$$

Further, there exists $\mathcal{B} \in (0, B_1)$ such that if $0 < |\beta| \leq \mathcal{B}$, then the operator $(\mathcal{A} + \beta D)$ satisfies a strong signed maximum principle in Dy .

Proof. Employ (13) and apply \mathcal{R}_0 to (10) to obtain

$$D\tilde{y} + \beta\mathcal{R}_0 D\tilde{y} = \mathcal{R}_0 \tilde{f}.$$

Note that (7) implies that $\mathcal{R}_0 : \tilde{\mathcal{L}} \rightarrow \tilde{\mathcal{C}}$ is continuous. Assume $|\beta|\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} < 1$. Then $(\mathcal{I} + \beta\mathcal{R}_0) : \tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}$ is invertible and

$$D\tilde{y} = (\mathcal{I} + \beta\mathcal{R}_0)^{-1} \mathcal{R}_0 \tilde{f}.$$

So, assume $0 < B_1 < \frac{1}{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}$ and assume $|\beta| \leq B_1$. Then $\mathcal{R}_\beta = (\mathcal{I} + \beta\mathcal{R}_0)^{-1} \mathcal{R}_0$ exists. Moreover,

$$\begin{aligned} |D\tilde{y}|_0 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} |D\tilde{y}|_0 &\leq |D\tilde{y}|_0 - |\beta|\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} |D\tilde{y}|_0 \\ &\leq |(\mathcal{I} + \beta\mathcal{R}_0)D\tilde{y}|_0 = |\mathcal{R}_0 \tilde{f}|_0 \leq \|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}} |\tilde{f}|_1 \end{aligned}$$

and (14) is proved since $D\tilde{y} = \mathcal{R}_\beta \tilde{f}$.

Now assume $f \in \mathcal{L}$ and assume $f \geq 0$ a.e. Then $\bar{f} = |f|_1$. Let $0 < |\beta| \leq B_1 < \frac{1}{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}$, write $f = \bar{f} + \tilde{f}$ and consider

$$\beta Dy = \beta \mathcal{R}_\beta f = \beta \mathcal{R}_\beta (\bar{f} + \tilde{f}).$$

Note that $\beta \mathcal{R}_\beta \bar{f} = \bar{f}$ since $(\mathcal{A}' + \beta\mathcal{I})\bar{f} = \beta\bar{f}$. So,

$$\begin{aligned} \beta Dy &= \beta \mathcal{R}_\beta f = \beta \mathcal{R}_\beta (\bar{f} + \tilde{f}) \\ &= \bar{f} + \beta \mathcal{R}_\beta \tilde{f} \geq |f|_1 - |\beta|\|\mathcal{R}_\beta \tilde{f}\|_0. \end{aligned}$$

Continue to assume that $0 < |\beta| \leq B_1$; it now follows from (14) that

$$\beta Dy \geq |f|_1 - |\beta| \left(\frac{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}{1 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}} \right) |\tilde{f}|_1.$$

Since $\tilde{f} = f - \bar{f}$, and $|\tilde{f}|_1 \leq |f|_1 + \bar{f} = 2|f|_1$, assume

$$\mathcal{B} < \min \left\{ B_1, \left(\frac{1 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}{2\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}} \right) \right\}.$$

Then

$$\beta Dy \geq \left(1 - 2\mathcal{B} \left(\frac{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}{1 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}} \right) \right) |f|_1$$

and (5) is valid with

$$K = \left(1 - 2\mathcal{B} \left(\frac{\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}}{1 - B_1\|\mathcal{R}_0\|_{\tilde{\mathcal{C}} \rightarrow \tilde{\mathcal{C}}}} \right) \right).$$

□

4 Examples

Example 4.1

Our first example considers boundary conditions that contain (3). Let $t_0 \in [0, 1]$ and consider the boundary value problem

$$y'' + \beta y' = f, \quad 0 \leq t \leq 1, \tag{15}$$

$$y(t_0) = 0, \quad y'(0) = y'(1). \tag{16}$$

So, for the boundary value problem (15), (16), $\mathcal{A} = D^2$, $\mathcal{A}' = D$, $\text{Ker}(\mathcal{A}) = \langle t - t_0 \rangle$ or $\text{Ker}(\mathcal{A}') = \langle 1 \rangle$.

We point out that if $t_0 = 0$ or $t_0 = 1$, the Fredholm alternative will imply that $\text{Im}(\mathcal{A}) = \tilde{\mathcal{L}}$. If $t_0 = 0$, then $f \in \text{Im}(\mathcal{A})$, if, and only if, f is orthogonal to solutions of the adjoint problem

$$y'' = 0, \quad 0 \leq t \leq 1, \quad y(0) = y(1), \quad y'(1) = 0.$$

Thus, f is orthogonal to the constant functions. If $t_0 = 1$, then f is orthogonal to solutions of the adjoint problem

$$y'' = 0, \quad 0 \leq t \leq 1, \quad y(0) = y(1), \quad y'(0) = 0,$$

and again, f is orthogonal to the constant functions.

However, if $t_0 \in [0, 1]$, one can show directly that $\text{Im}(\mathcal{A}) = \tilde{\mathcal{L}}$. If $f \in \text{Im}(\mathcal{A})$, then there exists a solution y of

$$y''(t) = f(t), \quad 0 \leq t \leq 1, \quad y(t_0) = 0, \quad y'(0) = y'(1),$$

which implies

$$0 = y'(1) - y'(0) = \int_0^1 y''(t)dt = \int_0^1 f(t)dt,$$

and $f \in \tilde{\mathcal{L}}$. Likewise, if $f \in \tilde{\mathcal{L}}$, then

$$y(t) = \int_0^t (t - s)f(s)ds - \int_0^{t_0} (t_0 - s)f(s)ds \tag{17}$$

is a solution of

$$y''(t) = f(t), \quad 0 \leq t \leq 1, \quad y(t_0) = 0, \quad y'(0) = y'(1),$$

which implies $f \in \text{Im}(\mathcal{A}')$. Thus, if $t_0 \in [0, 1]$, $\text{Im}(\mathcal{A}) = \tilde{\mathcal{L}}$.

To argue that $\mathcal{A}y = f$ is uniquely solvable with solution $y \in \text{Dom}(\tilde{\mathcal{A}})$, (17) implies the solvability. For uniqueness, if y_1 and y_2 are two such solutions, then $(y_1 - y_2)(t) = c(t - t_0)$ and $y_1 - y_2 \in \text{Dom}(\tilde{\mathcal{A}})$ implies $c = 0$.

Finally, (17) implies (7) is satisfied with $K_1 = 1$.

Theorem 3.1 applies and there exists $\mathcal{B} > 0$ such that if $0 < |\beta| \leq \mathcal{B}$, then $(\mathcal{A} + \beta\mathcal{I})$ has the strong signed maximum principle in Dy . Thus, $f \geq 0$ implies $\beta Dy \geq 0$. Hence, a natural partial order in which to apply the method of upper and lower solutions and monotone methods to a nonlinear boundary value problem is

$$y \in C^1[0, 1] \succeq 0 \iff \beta(t - t_0)y(t) \geq 0, 0 \leq t \leq 1, \text{ and } \beta y'(t) \geq 0, 0 \leq t \leq 1. \tag{18}$$

In Section 5, we shall employ monotone methods with respect to this partial order and obtain sufficient conditions for the existence of maximal and minimal solutions of a nonlinear boundary value problem associated with the boundary conditions (16).

Example 4.2

For the second example, let $h > 0$, and we consider a family of boundary conditions

$$y(0) = hy(1), \quad y'(0) = y'(1). \quad (19)$$

The boundary conditions (19) contain the periodic boundary conditions at $h = 1$. In this example, however, we exclude $h = 1$.

For the boundary value problem (15), (19), $\mathcal{A} = D^2$ and $\mathcal{A}' = D$, $\text{Ker}(\mathcal{A}) = \langle t + \frac{h}{1-h} \rangle$ or $\text{Ker}(\mathcal{A}') = \langle 1 \rangle$. Appealing directly to the Fredholm alternative, $f \in \text{Im}(\mathcal{A})$ if, and only if, f is orthogonal to solutions of the adjoint problem,

$$y'' = 0, \quad 0 \leq t \leq 1, \quad y(0) = y(1), \quad hy'(0) = y'(1).$$

Thus, $\text{Im}(\mathcal{A}) = \tilde{\mathcal{L}}$. Again, $f \in \tilde{\mathcal{L}}$ implies $\text{Dom}(\mathcal{A}) = \{y \in \mathcal{B} : \bar{D}y = 0\}$. Again, K in (7) can be computed since if $\tilde{f} \in \tilde{\mathcal{C}}$, then

$$\tilde{y}(t) = \int_0^t (t-s)\tilde{f}(s)ds + \frac{h}{1-h} \int_0^1 (1-s)\tilde{f}(s)ds.$$

Thus, Theorem 3.1 applies and there exists $\mathcal{B} > 0$ such that if $0 < |\beta| \leq \mathcal{B}$, then $(\mathcal{A}' + \beta D)$ satisfies the strong signed maximum principle in Dy .

To determine sign conditions on βy , four cases arise. If $0 < \beta \leq \mathcal{B}$, one considers the two cases, $1 < h$ or $0 < h < 1$. If $0 < \beta \leq \mathcal{B}$, then βy is increasing, which in turn implies y is increasing. If $h > 1$, then $\frac{y(0)}{y(1)} > 1$, and it follows that $y(t) < 0$ for $0 \leq t < 1$. If $0 < h < 1$, then $0 < \frac{y(0)}{y(1)} < 1$, and it follows that $y(t) > 0$ for $0 < t \leq 1$. Two analogous cases can be analyzed if $0 > \beta \geq -\mathcal{B}$. So, for example, if $\beta > 0$ and $1 < h$, a natural partial order in which to apply the method of upper and lower solutions and monotone methods to a nonlinear problem is

$$y \in C^1[0, 1] \succeq 0 \iff y(t) \geq 0, 0 \leq t \leq 1, \text{ and } y'(t) \geq 0, 0 \leq t \leq 1.$$

5 A Monotone Method

Let $f : [0, 1] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ be continuous. Let $t_0 \in [0, 1]$ and consider the boundary value problem

$$y''(t) = f(t, y(t), y'(t)), \quad t \in [0, 1], \quad (20)$$

$$y(t_0) = 0, \quad y'(0) = y'(1). \quad (21)$$

Assume that f satisfies the following monotonicity properties:

$$\begin{aligned} f(t, y, z_1) &< f(t, y, z_2) \text{ for } (t, y) \in [0, 1] \times \mathbb{R}, \quad z_1 > z_2, \\ f(t, y_1, z) &< f(t, y_2, z) \text{ for } (t, z) \in [t_0, 1] \times \mathbb{R}, \quad y_1 > y_2, \\ f(t, y_1, z) &> f(t, y_2, z) \text{ for } (t, z) \in [0, t_0] \times \mathbb{R}, \quad y_1 < y_2. \end{aligned} \quad (22)$$

So, f is monotone decreasing in the third component; for $t_0 < t \leq 1$, f is monotone decreasing in the second component and for $0 \leq t < t_0$, f is monotone increasing in the second component.

Apply a shift to (20) and consider the equivalent boundary value problem

$$y''(t) + \beta y'(t) = f(t, y(t), y'(t)) + \beta y'(t), \quad 0 \leq t \leq 1,$$

with boundary conditions (21), where $\beta < 0$. Assume $|\beta|$ is small such that $|\beta| \leq \mathcal{B}$, where $\mathcal{B} > 0$ is shown to exist in Theorem 3.1. Note that if $g(t, y, z) = f(t, y, z) + \beta z$ and f satisfies (22), then g satisfies (22).

Assume the existence of solutions, w_1 and v_1 , of the following boundary value problems for differential inequalities

$$\begin{aligned} w_1''(t) &\geq f(t, w_1(t), w_1'(t)), & 0 \leq t \leq 1, & \quad v_1''(t) \leq f(t, v_1(t), v_1'(t)), & 0 \leq t \leq 1, & \quad (23) \\ w_1(t_0) &= 0, & w_1'(0) = w_1'(1), & \quad v_1(t_0) = 0, & v_1'(0) = v_1'(1). & \end{aligned}$$

Assume further that

$$(t - t_0)(v_1(t) - w_1(t)) \geq 0, \quad 0 \leq t \leq 1, \quad (v_1'(t) - w_1'(t)) \geq 0, \quad 0 \leq t \leq 1. \quad (24)$$

Motivated by (18) and noting that $\beta < 0$, define a partial order \succeq on $C^1[0, 1]$ by

$$u \in C^1[0, 1] \succeq 0 \iff (t - t_0)u(t) \leq 0, 0 \leq t \leq 1, \text{ and } u'(t) \leq 0, 0 \leq t \leq 1.$$

Then the assumption (24) implies $w_1 \succeq v_1$.

Define iteratively, the sequences $\{v_k\}_{k=1}^\infty, \{w_k\}_{k=1}^\infty$, where

$$\begin{aligned} v_{k+1}''(t) + \beta v_{k+1}'(t) &= f(t, v_k(t), v_k'(t)) + \beta v_k'(t), & 0 \leq t \leq 1, & \quad (25) \\ v_{k+1}(t_0) &= 0, & v_{k+1}'(0) = v_{k+1}'(1), & \end{aligned}$$

and

$$\begin{aligned} w_{k+1}''(t) + \beta w_{k+1}'(t) &= f(t, w_k(t), w_k'(t)) + \beta w_k'(t), & 0 \leq t \leq 1, & \quad (26) \\ w_{k+1}(t_0) &= 0, & w_{k+1}'(0) = w_{k+1}'(1). & \end{aligned}$$

Theorem 3.1 implies the existence of each v_{k+1}, w_{k+1} since if $0 < |\beta| \leq \mathcal{B}$, the inverse of $(\mathcal{A} + \beta D)$ exists.

Theorem 5.1 *Assume $f : [0, 1] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous and assume f satisfies the monotonicity properties (22). Assume the existence of two times continuously differentiable functions, v_1 and w_1 , satisfying (23). Define the sequences of iterates $\{v_k\}_{k=1}^\infty, \{w_k\}_{k=1}^\infty$ by (25) and (26), respectively. Then, for each $k \in \mathbb{N}_1$,*

$$w_k \succeq w_{k+1} \succeq v_{k+1} \succeq v_k. \quad (27)$$

Moreover, $\{v_k\}_{k=1}^\infty$ converges in $C^1[0, 1]$ to a solution v of (20) and $\{w_k\}_{k=1}^\infty$ converges in $C^1[0, 1]$ to a solution w of (20) satisfying

$$w_k \succeq w_{k+1} \succeq w \succeq v \succeq v_{k+1} \succeq v_k. \quad (28)$$

Proof. Since v_1 satisfies a differential inequality given in (24),

$$v_2''(t) + \beta v_2'(t) = f(t, v_1(t), v_1'(t)) + \beta v_1'(t) \geq v_1''(t) + \beta v_1'(t), \quad 0 \leq t \leq 1.$$

Set $u = v_2 - v_1$ and u satisfies a boundary value problem for a differential inequality

$$u''(t) + \beta u'(t) \geq 0, \quad 0 \leq t \leq 1, \quad u(t_0) = 0, \quad u'(0) = u'(1).$$

The signed maximum principle applies and $u \succeq 0$; in particular, $v_2 \succeq v_1$. Similarly, $w_1 \succeq w_2$. Now, set $u = w_2 - v_2$ and

$$u''(t) + \beta u'(t) = (f(t, w_1(t), w_1'(t)) - f(t, v_1(t), v_1'(t))) + \beta(w_1'(t) - v_1'(t)), \quad 0 \leq t \leq 1, \\ u(t_0) = 0, \quad u'(0) = u'(1).$$

Since f satisfies (22) and $\beta(w_1'(t) - v_1'(t)) \geq 0$, $0 \leq t \leq 1$, it follows that

$$u''(t) + \beta u'(t) \geq 0, \quad 0 \leq t \leq 1,$$

and again, the signed maximum principle applies and $u \succeq 0$. In particular, $w_2 \succeq v_2$. Thus, (27) is proved for $k = 1$. It follows by a straightforward induction that (27) is valid using the arguments presented in this paragraph.

To obtain the existence of limiting solutions v and w satisfying (28), note that the sequence $\{v_k'\}$ is monotone and appropriately bounded. Thus, the sequence $\{v_k'\}$ is converging pointwise on $[0, 1]$. Dini's theorem then implies the uniform convergence of the sequence $\{v_k\}$ on $[0, 1]$ since $\{v_k(t)\}$ is monotone for each t and is appropriately bounded. This argument can be repeated to obtain the uniform convergence of $\{v_k''\}$ on $[0, 1]$. Since $v_{k+1}''(t) = f(t, v_k(t), v_k'(t)) + \beta(v_k'(t) - v_{k+1}'(t))$, the sequence $\{v_k''\}$ is converging pointwise on $[0, 1]$. Now, Dini's theorem implies the uniform convergence of the sequence $\{v_k''\}$ on $[0, 1]$. Again, employ $v_{k+1}''(t) = f(t, v_k(t), v_k'(t)) + \beta(v_k'(t) - v_{k+1}'(t))$, and it follows that the sequence $\{v_k''\}$ converges uniformly on $[0, 1]$. This implies that if $v \in C^1[0, 1]$ is the limit of $\{v_k\}$ (meaning v_k is converging to v uniformly and v_k' is converging to v' uniformly), then $\{v_k''\}$ converges uniformly to v'' on $[0, 1]$ and v is a solution of (20), (21) satisfying (28). Similarly, the solution w of (20), (21) satisfying (28) exists, and the theorem is proved. \square

Suppose now f satisfies the "anti"-inequalities to (22); that is, suppose f satisfies

$$\begin{aligned} f(t, y, z_1) &> f(t, y, z_2) \text{ for } (t, y) \in [0, 1] \times \mathbb{R}, \quad z_1 > z_2, \\ f(t, y_1, z) &> f(t, y_2, z) \text{ for } (t, z) \in [t_0, 1] \times \mathbb{R}, \quad y_1 > y_2, \\ f(t, y_1, z) &< f(t, y_2, z) \text{ for } (t, z) \in [0, t_0] \times \mathbb{R}, \quad y_1 < z_2. \end{aligned} \quad (29)$$

One can appeal to the signed maximum principle and apply a shift to (20) and consider the equivalent boundary value problem, $y''(t) + \beta y'(t) = f(t, y(t), y'(t)) + \beta y'(t)$, $0 \leq t \leq 1$, where $\beta > 0$. Note, if f satisfies (29) and $\beta > 0$, then $g(t, y, z) = f(t, y, z) + \beta z$ satisfies (29).

Now, assume the existence of solutions, w_1 and v_1 , of the following differential inequalities

$$\begin{aligned} w_1''(t) &\leq f(t, w_1(t), w_1'(t)), \quad 0 \leq t \leq 1, \quad v_1''(t) \geq f(t, v_1(t), v_1'(t)), \quad 0 \leq t \leq 1, \\ w_1(t_0) &= 0, \quad w_1'(0) = w_1'(1), \quad v_1(t_0) = 0, \quad v_1'(0) = v_1'(1). \end{aligned} \quad (30)$$

Assume further that

$$(t - t_0)(v_1(t) - w_1(t)) \leq 0, \quad 0 \leq t \leq 1, \quad (v_1'(t) - w_1'(t)) \leq 0, \quad 0 \leq t \leq 1. \quad (31)$$

Noting that $\beta > 0$, define a partial order \succeq_1 on $C^1[0, 1]$ by

$$u \in C^1[0, 1] \succeq_1 0 \iff (t - t_0)u(t) \geq 0, 0 \leq t \leq 1, \text{ and } u'(t) \geq 0, 0 \leq t \leq 1.$$

In particular, assume $v_1 \succeq_1 v_1$.

Theorem 5.2 *Assume $f : [0, 1] \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is continuous and assume f satisfies the monotonicity properties (29). Assume the existence of two times continuously differentiable functions, v_1 and w_1 , satisfying (30) and (31). Define the sequences of iterates $\{v_k\}_{k=1}^\infty$, $\{w_k\}_{k=1}^\infty$ by (25) and (26), respectively. Then, for each $k \in \mathbb{N}_1$,*

$$v_k \succeq_1 v_{k+1} \succeq_1 w_{k+1} \succeq_1 w_k.$$

Moreover, $\{v_k\}_{k=1}^\infty$ converges in $C^1[0, 1]$ to a solution v of (20) and $\{w_k\}_{k=1}^\infty$ converges in $C^1[0, 1]$ to a solution w of (20) satisfying

$$v_k \succeq_1 v_{k+1} \succeq_1 v \succeq_1 w \succeq_1 w_{k+1} \succeq_1 w_k.$$

6 Conclusion

Boundary value problems for ordinary differential equations with dependence on a real parameter β , where $\beta = 0$ is a simple eigenvalue, are studied. The concept of a maximum principle in $\beta y'$ is defined. Sufficient conditions are obtained such that if $\mathcal{A}y + \beta y' = f$ is a representation of the boundary value problem, there exists a punctured neighborhood of $\beta = 0$ such that $f \geq 0$ implies $\beta y' \geq 0$, where y is the unique solution of $\mathcal{A}y + \beta y' = f$. Two examples are provided to illustrate the main theorem and an application of a monotone method is given.

References

- [1] B. Alziary, J. Fleckinger and P. Takáč. An extension of maximum and anti-maximum principles to a Schrödinger equation in \mathbb{R}^2 , *J. Differential Equations* **156** (1999) 122–152.
- [2] D. Arcoya and J.L. Gámez. Bifurcation theory and related problems: anti-maximum principle and resonance, *Comm. Partial Differential Equations* **26** (2001) (9–10) 1879–1911.
- [3] I.V. Barteneva, A. Cabada and A. O. Ignatyev. Maximum and anti-maximum principles for the general operator of second order with variable coefficients, *Appl. Math. Comput.* **134** (2003) 173–184.
- [4] A. Cabada and J.Á Cid. On comparison principles for the periodic Hill’s equation. *J. Lond. Math. Soc.* **86** (1) (2012) 272–290.
- [5] A. Cabada, Alberto, J.Á Cid and L. López-Somoza. *Maximum Principles for the Hill’s Equation*. Academic Press, London, 2018.
- [6] A. Cabada, Alberto, J.Á Cid and M. Tvrdý. A generalized anti-maximum principle for the periodic one-dimensional p -Laplacian with sign changing potential. *Nonlinear Anal.* **72** (2010) (7-8) 3434–3446.
- [7] J. Campos, J. Mawhin and R. Ortega. Maximum principles around an eigenvalue with constant eigenfunctions. *Commun. Contemp. Math.* **10** (6) (2008) 1243–1259.

- [8] Ph. Clément and L.A. Peletier. An anti-maximum principle for second-order elliptic operators. *J. Differential Equations* **34** (1979) 218–229.
- [9] Ph. Clément and G. Sweers. Uniform anti-maximum principles. *J. Differential Equations* **164** (2000) 118–154.
- [10] P. Hess. An antimaximum principle for linear elliptic equations with an indefinite weight function, *J. Differential Equations* **41** (1981) 369–374.
- [11] G. Infante, P. Pietramala and F.A.F. Tojo. Nontivial solutions of local and nonlocal Neumann boundary value problems, *Proc. Roy. Soc. Edinburgh Sect. A* **146** (2) (2016) 337–369.
- [12] J. Mawhin. Partial differential equations also have principles: Maximum and antimaximum, *Contemporary Mathematics* **540** (2011) 1–13.
- [13] Y. Pinchover. Maximum and anti-maximum principles and eigenfunctions estimates via perturbation theory of positive solutions of elliptic equations. *Math. Ann.* **314** (1999) 555–590.
- [14] M. H. Protter and H. Weinberger. *Maximum Principles in Differential Equations*. Prentice Hall, Englewoods Cliffs, N.J., 1967.
- [15] P. Takáč. An abstract form of maximum and anti-maximum principles of Hopf's type. *J. Math. Anal. Appl.* **201** (1996) 339–364.
- [16] M. Zhang. Optimal conditions for maximum and antimaximum principles of the periodic solution problem. *Bound. Value Probl.* (2010), Art. ID 410986, 26 pp.



Performance Comparison of Sliding Mode Control and Sliding PID for Rescue ROV

T. Herlambang¹, A. Suryowinoto^{2*}, I. Kurniastuti¹, H. Nurhadi³
and K. Oktafianto⁴

¹ *Department of Information Systems, Universitas Nahdlatul Ulama Surabaya, Indonesia.*

² *Electrical Engineering Department, Institut Teknologi Adhi Tama Surabaya, Indonesia.*

³ *Department of Industrial Mechanical Engineering, Sepuluh Nopember Institute of Technology, Indonesia.*

⁴ *Department of Mathematics, University of PGRI Ronggolawe, Indonesia.*

Received: July 31, 2022; Revised: October 17, 2023

Abstract: The underwater vehicle has been developed by many countries to be one kind of defense technology. This unmanned underwater vehicle is usually called the Remotely Operated Underwater Vehicle (ROV). The ROV is commonly used for underwater exploration and as a defense vehicle. The ROV can move by using six degrees of freedom (6-DOF) and requires a control system in order for the ROV to move as intended. In this paper, a controller was synthesized in the 6-DOF Evacuation ROV linear model with the Sliding Mode Control (SMC) and Sliding PID (SPID) methods. The contribution of the paper provides an analysis of numerical study and stability analysis by using the Lyapunov function for the performance of both control system methods. The focus of this paper is the comparison between the performance of SMC and SPID on the AUV linear model, of which SPID is a combination of SMC and PID. The simulation results show that the SMC and SPID methods have a good stability and small error of about 0.1% - 4%. Further, the results show that the SPID method is more stable than SMC.

Keywords: *remotely operated underwater vehicle; sliding mode control; sliding proportional integral derivative; linear model; Lyapunov function.*

Mathematics Subject Classification (2010): 93C10, 93D05.

* Corresponding author: <mailto:andysuryo@itats.ac.id>

1 Introduction

Underwater robots are widely developed by many countries as one of their defense technologies. The underwater defense technology is very important for a country in keeping and maintaining its marine wealth. This underwater robot is usually called the Remotely Operated Underwater Vehicle (ROV). The ROV is commonly used for marine security systems, assisting evaluation on at-sea accidents, and underwater exploration [1]. The ROV has two kinds of motions, namely, the translational and rotation motions on the x -, y -, and z -axes. Two parts of the ROV functioning to control the ROV motion are the propulsion system and the fin system. The propulsion system is used to regulate the angular velocity of the ROV, and the fin system is used to adjust the angle of the fin and the rudder position. The development of the ROV was initiated in the 1970s together with the initial investigation of the usefulness of the ROV system, and in 1970-1980, the ROV technology development and experiments were carried out. In 1980-1990, the experiments using prototypes were done, then in 1990-2000, the ICT-based ROV was developed.

This encouraged researches to improve the technology for innovations, in particular on the controller design of unmanned vehicles. The development of the controllers aimed to organize the actuator such that the unmanned vehicle can be stable in its motion as expected. Several studies related to the AUV control system began in the 2000s when Chiu et al. [2] used a fuzzy control sliding mode with a 2-DOF model. In 2002, W. Naeem [3] used a predictive control model for a nonlinear 2-DOF (surge and yaw) model. Kim and Ura [4] used the R-One Robot AUV with a length of 8.3 meters, a diameter of 1.2 meters, and a mass of 4400 kg with PID. Repoulias and Papadopoulos [5] used partial state-feedback in nonlinear 3-DOF (surge, sway and yaw) models. Lapiere and Soesanto [6] used the Infante AUV with a length of 4,215 m and a mass of 23 Kg with path-following control. Akcaya et al. [7] used SMC in a 3-DOF linear model. Rezazadegan et al. [8] used an adaptive nonlinear controller in the 5-DOF model. Oktafianto et al. [9] used SMC in the linear model of AUV. Herlambang et al. [10] used the PID controller method to control the movement of the 6-DOF linear model applied to UNUSAITS AUV, Nurhadi et al. [11] used the SMC method to control the Surge, Heave and Pitch Motions in the 3-DOF nonlinear model, in 2020, Herlambang et al. [12] used SPID Control in the 6-DOF linear model of the AUV, and also used the Linear Quadratic Regulator (LQR) for the linear motion system of the AUV [14].

The focus of this paper was to compare two control methods: SMC and SPID for 6-DOF motion control, of which the SPID controller is a combination of the SMC and PID controllers. This study used the 6-DOF linear model, and the linear model was obtained by linearization of the nonlinear 6-DOF model. The result of this study was the comparison of the stability performance of SMC and SPID shown by numerical simulation on the response results of both methods. The stability obtained by both methods was quite significant with an error of 0.1%-4% with a settling time of 0.1-1 second.

In order to analyze an ROV, it is necessary to consider the axis system which comprises the Earth Fixed Frame (EFF) and Body Fixed Frame (BFF) that are displayed in Figure 1 [14]. The motion of the ROV has 6 degrees of freedom (6 DOF) comprising 3 degrees of freedom in the direction of translational motions on the x -, y -, and z -axes and 3 degrees of freedom in the rotational motions on the x -, y - and z -axes. The general description of an ROV with 6 DOF can be expressed in (1), where η represents the position and orientation vectors over the EFF and v represents the linear and angular

velocities over the BFF. The details have been described in [15].

2 Remotely Operated Underwater Vehicle (ROV)

In order to analyze an ROV, it is necessary to consider the axis system which comprises the Earth Fixed Frame (EFF) and Body Fixed Frame (BFF) that are displayed in Figure 1 [14]. The motion of the ROV has 6 degrees of freedom (6 DOF) comprising 3 degrees of freedom in the direction of translational motions on the x -, y -, and z -axes and 3 degrees of freedom in the rotational motions on the x -, y - and z -axes. The general description of an ROV with 6 DOF can be given as (1) [15]

$$\begin{aligned} \eta &= [\eta_1^T, \eta_2^T]^T, & \eta_1 &= [x, y, z]^T, & \eta_2 &= [\theta, \theta, \Psi]^T, \\ v &= [v_1^T, v_2^T]^T, & v_1 &= [u, v, w]^T, & v_2 &= [p, q, r]^T, \\ \tau &= [\tau_1^T, \tau_2^T]^T, & \tau_1 &= [X, Y, Z]^T, & \tau_2 &= [K, M, N]^T. \end{aligned} \tag{1}$$

This study uses the prototype of the Evacuation ROV, and the specifications of the Evacuation ROV are: the weight is 16 kg, the length is 2 meter and the diameter is 300 mm [15]. The properties of the Evacuation ROV are displayed in Figure 2 and Table 1.

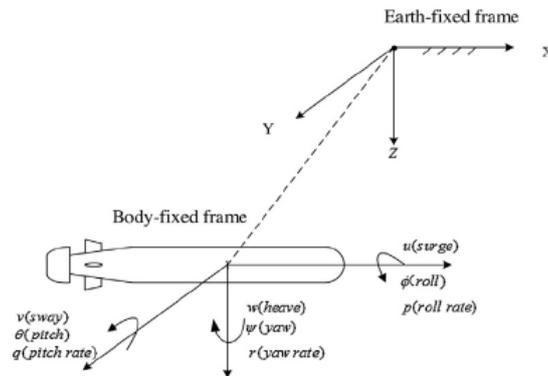


Figure 1: ROV Motion with Six Degrees of Freedom [20].

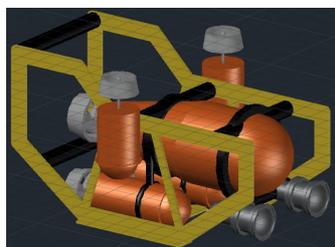


Figure 2: Profile of Evacuation ROV.

Table 1: Specification of the Evacuation ROV.

Weight	15 Kg
Overall Length	900 mm
Beam	300 mm
Controller	Wired Control ArduSUB with Joystick
Sensors	Depth Sensor, Sonar
Camera	TTL Camera
Lighting	1500 LM, 145° Beam Dimmable
Battery	11.8 V Li Po 5200 mAh
Material	Carbon Fiber
Main Propulsion	T200 Motor Thruster Include Propeller
Maneuver Propulsion	T200 Motor Thruster Include Propeller
Service Speed	1, 6 knots
Operation Depth	5 – 10 m

Variable η represents the vector position and orientation w.r.t. EFF, whereas τ represents the force vectors and moments that are working on the ROV w.r.t. BFF, namely surge (u), sway (v), heave (w), roll (p), pitch (q) and yaw (r). The total forces and moments which are working on the ROV can be gained by combining hydrostatic, hydrodynamic and thrust forces. In this case, we assumed that the diagonal inertia tensor (I_o) is zero, in order to obtain the total forces and moments of the whole model as follows.

Surge:

$$m[\dot{u} - vr + wq - x_G(q^2 + r^2) + y_G(pq - \dot{r}) + z_G(pr + \dot{q})] = X_{res} + X_{|u|u}|u| + X_{\dot{u}}\dot{u} + X_{wq}wq + X_{qq}qq + X_{vr}vr + X_{rr}rr + X_{prop}, \quad (2)$$

Sway:

$$m[\dot{v} - wp + ur - y_G(r^2 + p^2) + z_G(qr - \dot{p}) + x_G(pq + \dot{r})] = Y_{res} + Y_{|v|v}|v| + Y_{\dot{v}}\dot{v} + Y_{\dot{r}}\dot{r} + Y_{ur}ur + Y_{wp}wp + Y_{pq}pq + Y_{uv}uv + Y_{uu\delta_s}u^2\delta_s, \quad (3)$$

Heave:

$$m[\dot{w} - uq + vp - z_G(p^2 + q^2) + x_G(rp - \dot{q}) + y_G(rq + \dot{p})] = Z_{res} + Z_{|w|w}|w| + Z_{q|q|q}|q| + Z_{\dot{w}}\dot{w} + Z_{\dot{q}}\dot{q} + Z_{uq}uq + Z_{vp}vp + Z_{rp}rp + Z_{uw}uw + Z_{uu\delta_s}u^2\delta_s, \quad (4)$$

Roll:

$$I_x\dot{p} + (I_z - I_y)qr + m[y_G(\dot{w} - uq + vp) - z_G(\dot{v} - wp + ur)]K_{res} + K_{p|p|p}|p| + K_{\dot{p}}\dot{p} + K_{prop}, \quad (5)$$

Pitch:

$$I_y\dot{q} + (I_x - I_z)rp + m[z_G(\dot{u} - vr + wq) - x_G(\dot{w} - uq + vp)] = M_{res} + M_{|w|w}|w| + M_{q|q|q}|q| + M_{\dot{w}}\dot{w} + M_{\dot{q}}\dot{q} + M_{uq}uq + M_{vp}vp + M_{rp}rp + M_{uw}uw + M_{uu\delta_s}u^2\delta_s. \quad (6)$$

3 Linearization

In general, it is difficult to design a controller for nonlinear systems (2)-(6). Therefore, the non-linear ROV model is linearized by using the Jacobi matrix. The general equations of the non-linear ROV model can be written as

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t), t), \\ y(t) &= g(x(t), u(t), t), \end{aligned} \tag{7}$$

and the Jacobi matrix is defined as

$$\frac{\partial f(\bar{x}, \bar{u}, t)}{\partial x} = \begin{bmatrix} \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial x_1} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial x_2} & \cdots & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial x_n} \\ \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial x_1} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial x_2} & \cdots & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(\bar{x}, \bar{u}, t)}{\partial x_1} & \frac{\partial f_n(\bar{x}, \bar{u}, t)}{\partial x_2} & \cdots & \frac{\partial f_n(\bar{x}, \bar{u}, t)}{\partial x_n} \end{bmatrix}. \tag{8}$$

When the Jacobi matrix shown in (8) is used to linearize the ROV equation of motion for 6-DOF, then we obtain the following equation:

$$\frac{\partial f(\bar{x}, \bar{u}, t)}{\partial x} = \begin{bmatrix} \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_1(\bar{x}, \bar{u}, t)}{\partial r} \\ \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_2(\bar{x}, \bar{u}, t)}{\partial r} \\ \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_3(\bar{x}, \bar{u}, t)}{\partial r} \\ \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_4(\bar{x}, \bar{u}, t)}{\partial r} \\ \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_5(\bar{x}, \bar{u}, t)}{\partial r} \\ \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial u} & \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial v} & \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial w} & \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial p} & \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial q} & \frac{\partial f_6(\bar{x}, \bar{u}, t)}{\partial r} \end{bmatrix}, \tag{9}$$

where the functions f_1, f_2, f_3, f_4, f_5 and f_6 are

$$f_1 = \{X_{res} + X_{|u|u}|u| + X_{wq}wq + X_{qq}qq + X_{vr}vr + X_{rr}rr + X_{prop} - m[-vr + wq - x_G(q^2 + r^2) + pqy_G + prz_G]\}/\{m - X_{\dot{u}}\}, \tag{10}$$

$$f_2 = \{Y_{res} + Y_{|v|v}|v| + Y_{r|r}|r| + Y_{\dot{r}}\dot{r} + Y_{ur}ur + Y_{wp}wp + Y_{pq}pq + Y_{uv}uv + Y_{uu\delta_r}u^2\delta_r - m[-wp + ur - y_G(r^2 + p^2) + qr z_G + pq x_G]\}/\{m - Y_{\dot{v}}\}, \tag{11}$$

$$f_3 = \{Z_{res} + Z_{|w|w}|w| + Z_{q|q}|q| + Z_{\dot{q}}\dot{q} + Z_{uq}uq + Z_{vp}vp + Z_{rp}rp + Z_{uw}uw + Z_{uu\delta_s}u^2\delta_s - m[-uq + vp - z_G(p^2 + q^2) + rp x_G + rq y_G]\}/\{m - Z_{\dot{w}}\}, \tag{12}$$

$$f_4 = \frac{K_{res} + K_{p|p}|p| + K_{prop} - ((I_z - I_y)qr + m[y_G(-uq + vp) - z_G(-wp + ur)])}{I_x - K_{\dot{p}}}, \tag{13}$$

$$f_5 = \{M_{res} + M_{w|w}|w| + M_{q|q}|q| + M_{\dot{w}}\dot{w} + M_{uq}uq + M_{vp}vp + M_{rp}rp + M_{uw}uw + M_{uu\delta_s}u^2\delta_s - ((I_x - I_z)rp + m[z_G(-vr + wq) - x_G(-uq + vp)])\}/\{I_y - M_{\dot{q}}\}, \tag{14}$$

$$f_6 = \{N_{res} + N_{v|v}|v| + N_{r|r}|r| + N_{\dot{v}}\dot{v} + N_{ur}ur + N_{wp}wp + N_{pq}pq + N_{uv}uv + N_{uu\delta_r}u^2\delta_r - ((I_y - I_z)pq + m[x_G(-wp + ur) - y_G(-vr + wq)])\}/\{I_z - N_{\dot{r}}\}. \tag{15}$$

The linear model obtained from the above process is as follows:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t), \end{aligned} \quad (16)$$

where

$$A = J_x = \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{mz_G}{m-X_{\dot{u}}} & \frac{-my_G}{m-X_{\dot{u}}} \\ 0 & 1 & 0 & -\frac{mz_G}{m-Y_{\dot{v}}} & 0 & \frac{(mx_G-Y_{\dot{r}})}{m-Y_{\dot{v}}} \\ 0 & 0 & 1 & \frac{my_G}{m-Z_{\dot{w}}} & -\frac{(mx_G+Z_{\dot{q}})}{m-Z_{\dot{w}}} & 0 \\ 0 & -\frac{mz_G}{I_x-K_{\dot{p}}} & \frac{my_G}{I_x-K_{\dot{p}}} & 1 & 0 & 0 \\ \frac{mz_G}{I_y-M_{\dot{q}}} & 0 & -\frac{(mx_G+M_{\dot{w}})}{I_y-M_{\dot{q}}} & 0 & 1 & 0 \\ -\frac{my_G}{I_z-N_{\dot{r}}} & \frac{(mx_G-N_{\dot{v}})}{I_z-N_{\dot{r}}} & 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} a_1 & b_1 & c_1 & d_1 & e_1 & g_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 & g_2 \\ a_3 & b_3 & c_3 & d_3 & e_3 & g_3 \\ a_4 & b_4 & c_4 & d_4 & e_4 & g_4 \\ a_5 & b_5 & c_5 & d_5 & e_5 & g_5 \\ a_6 & b_6 & c_6 & d_6 & e_6 & g_6 \end{bmatrix},$$

$$B = J_u = \begin{bmatrix} 1 & 0 & 0 & 0 & \frac{mz_G}{m-X_{\dot{u}}} & \frac{-my_G}{m-X_{\dot{u}}} \\ 0 & 1 & 0 & -\frac{mz_G}{m-Y_{\dot{v}}} & 0 & \frac{(mx_G-Y_{\dot{r}})}{m-Y_{\dot{v}}} \\ 0 & 0 & 1 & \frac{my_G}{m-Z_{\dot{w}}} & -\frac{(mx_G+Z_{\dot{q}})}{m-Z_{\dot{w}}} & 0 \\ 0 & -\frac{mz_G}{I_x-K_{\dot{p}}} & \frac{my_G}{I_x-K_{\dot{p}}} & 1 & 0 & 0 \\ \frac{mz_G}{I_y-M_{\dot{q}}} & 0 & -\frac{(mx_G+M_{\dot{w}})}{I_y-M_{\dot{q}}} & 0 & 1 & 0 \\ -\frac{my_G}{I_z-N_{\dot{r}}} & \frac{(mx_G-N_{\dot{v}})}{I_z-N_{\dot{r}}} & 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \times \begin{bmatrix} A_1 & B_1 & C_1 & D_1 & E_1 & G_1 \\ A_2 & B_2 & C_2 & D_2 & E_2 & G_2 \\ A_3 & B_3 & C_3 & D_3 & E_3 & G_3 \\ A_4 & B_4 & C_4 & D_4 & E_4 & G_4 \\ A_5 & B_5 & C_5 & D_5 & E_5 & G_5 \\ A_6 & B_6 & C_6 & D_6 & E_6 & G_6 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, D = 0.$$

4 Sliding Mode Control

SMC is a controller that is robust w.r.t. internal and external disturbances. As such, SMC has been widely used in many applications. In order to apply SMC to a system, we need to proceed according to the algorithm presented in Figure 3.

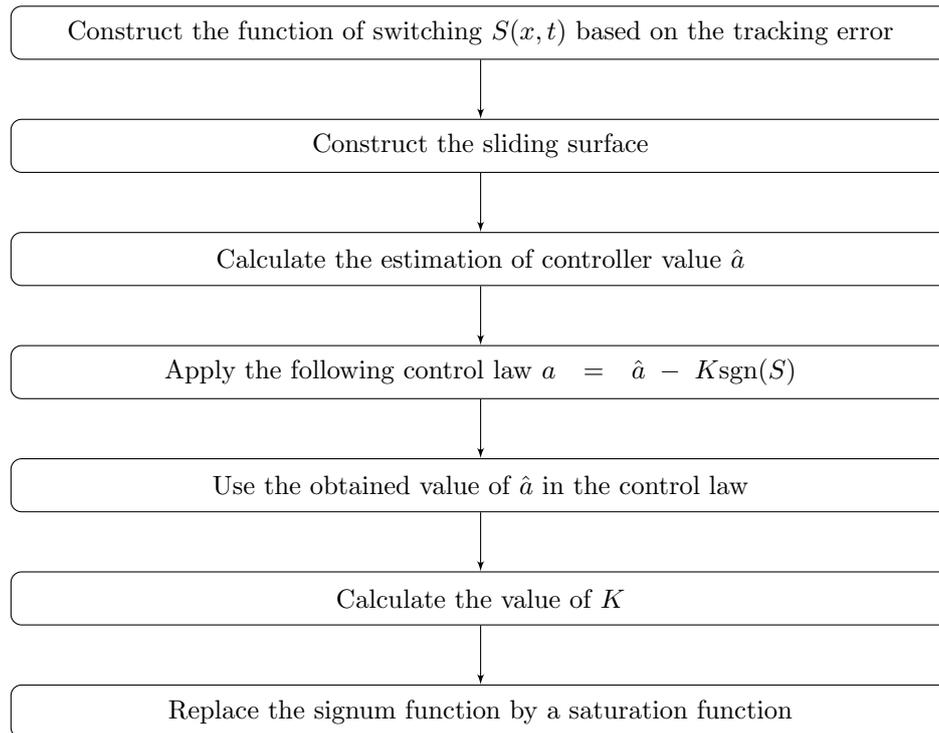


Figure 3: Algorithm of Sliding Mode Control.

5 Sliding PID

The design of the Sliding-PID control system is a combination of SMC and PID. In this study, the system first passes through SMC, then the result is optimized by the PID controller. The process works as follows. The notation e represents the error signal, i.e., the difference between the output and the reference. The error signal is substituted into the sliding surface equation $S = \dot{e} + \lambda e$ for some $\lambda \in \mathbb{R}$. Then the value of S is used by the PID to determine the input signal u . This scheme is shown in the block diagram in Figure 4.

6 Control System Design

Designing the SMC control system for the 6-DOF linear model involves creating control equations for surge, sway, heave, roll, pitch and yaw, obtained by the SMC Algorithm

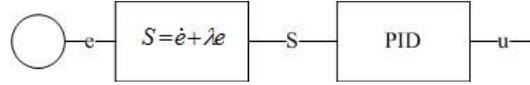


Figure 4: Block Diagram of SPID.

described in the previous section. From (16), the status of surge can be written as follows:

$$\dot{u} = aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1X_{prop} + BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}. \quad (17)$$

To find the control of the surge, the tracking error of surge is determined as follows:

$$\tilde{u} = u - u_d,$$

where u_d is a constant function. Since the system has order 1, the switching function is formed as follows:

$$S(u, t) = \left(\frac{d}{dt} \right)^{n-1} \tilde{u}, \quad \text{with } n = 1,$$

$$S(u, t) = \left(\frac{d}{dt} \right)^{1-1} \tilde{u},$$

$$S(u, t) = \tilde{u} = u - u_d.$$

The derivative of S is

$$\dot{S}(u, t) = \dot{u} - \dot{u}_d. \quad (18)$$

Since u_d is a constant function, one has $\dot{u}_d = 0$. By substituting the equation (17) into (18), we get

$$\dot{S}(u, t) = aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1X_{prop} + BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}. \quad (19)$$

Next, the value of \hat{X}_{prop} is determined from the equation (19), when the value of $\dot{S} = 0$, as follows:

$$aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1X_{prop} + BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2} = 0. \quad (20)$$

Then the obtained \hat{X}_{prop} is

$$\hat{X}_{prop} = - \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r}{AA_1} \right) - \left(\frac{BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}}{AA_1} \right). \quad (21)$$

Since we require that the control law has to satisfy the condition of sliding, we have

$$X_{prop} = \hat{X}_{prop} - K_1 \text{sgn}(S), \quad (22)$$

then from the equation (21) and (22), it is obtained that

$$X_{prop} = - \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r}{AA_1} \right) - \left(\frac{BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}}{AA_1} \right) - K_1\text{sgn}(S). \quad (23)$$

By substituting the equation (23) into (19), it is obtained that

$$\dot{S}(u, t) = -AA_1K_1\text{sgn}(S). \quad (24)$$

Next, the value of K_1 is set up by substituting the equation (24) into the following equation to meet the sliding condition, that is,

$$S\dot{S} \leq -\eta|S|, \quad (25)$$

$$-SAA_1K_1\text{sgn}(S) \leq -\eta|S|,$$

$$-AA_1K_1\text{sgn}(S) \leq -\frac{\eta|S|}{S},$$

$$K_1 \geq \frac{\eta}{AA_1\text{sgn}(S)}. \quad (26)$$

From the equation (26), the value of K_1 is obtained as follows:

$$K_1 = \left| \max \frac{\eta}{AA_1} \right|. \quad (27)$$

Next, in order to minimize chattering, a boundary layer is used by replacing the function of signum (sgn) into the function of saturation (sat) as follows:

$$X_{prop} = \hat{X}_{prop} - K_1\text{sat} \left(\frac{S}{\phi} \right). \quad (28)$$

Thus, the design of the SMC control of surge, obtained by substituting the equations (21) and (27) into the equation (28), is as follows:

$$X_{prop} = - \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2}}{AA_1} \right) - \frac{GG_1\delta_{r_2}}{AA_1} - \left| \max \frac{\eta}{AA_1} \right| \text{sat} \left(\frac{S}{\phi} \right). \quad (29)$$

Furthermore, the state equations for sway, heave, roll, pitch and yaw are as follows:

$$\dot{v} = aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r + AA_2X_{prop} + BB_2\delta_{r_1} + CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2} \quad (30)$$

$$\dot{w} = aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r + AA_3X_{prop} + BB_3\delta_{r_1} + CC_3\delta_{s_1} + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2} \quad (31)$$

$$\dot{p} = aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r + AA_4X_{prop} + BB_4\delta_{r_1} + CC_4\delta_{s_1} + DD_4K_{prop} + EE_4\delta_{s_2} + GG_4\delta_{r_2} \quad (32)$$

$$\dot{q} = aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r + AA_5X_{prop} + BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + EE_5\delta_{s_2} + GG_5\delta_{r_2} \quad (33)$$

$$\dot{r} = aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r + AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2} + GG_6\delta_{r_2}. \quad (34)$$

In the same way as obtaining the surge input control with the switching function, control law, sliding conditions and applying the boundary layer, the input controls for sway, heave, roll, pitch and yaw are as follows:

$$\delta_{r_1} = - \left(\frac{aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r}{BB_2} \right) - \left(\frac{AA_2X_{prop} + CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2}}{BB_2} \right) - \left| \max \frac{\eta}{BB_2} \right| \text{sat} \left(\frac{S}{\phi} \right), \quad (35)$$

$$\delta_{s_1} = - \left(\frac{aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r}{CC_3} \right) - \left(\frac{AA_3X_{prop} + BB_3\delta_{r_1} + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2}}{CC_3} \right) - \left| \max \frac{\eta}{CC_3} \right| \text{sat} \left(\frac{S}{\phi} \right), \quad (36)$$

$$K_{prop} = - \left(\frac{aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r}{DD_4} \right) - \left(\frac{AA_4X_{prop} + BB_4\delta_{r_1} + CC_4\delta_{s_1} + EE_4\delta_{s_2} + GG_4\delta_{r_2}}{DD_4} \right) - \left| \max \frac{\eta}{AA_1} \right| \text{sat} \left(\frac{S}{\phi} \right), \quad (37)$$

$$\delta_{s_2} = - \left(\frac{aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r}{EE_5} \right) - \left(\frac{AA_5X_{prop} + BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + GG_5\delta_{r_2}}{EE_5} \right) - \left| \max \frac{\eta}{CC_3} \right| \text{sat} \left(\frac{S}{\phi} \right), \quad (38)$$

$$\delta_{r_2} = - \left(\frac{aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r}{GG_6} \right) - \left(\frac{AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2}}{GG_6} \right) - \left| \max \frac{\eta}{BB_2} \right| \text{sat} \left(\frac{S}{\phi} \right). \quad (39)$$

The design of the SPID control system for the 6-DOF linear model first passes the SMC control system equation, then the result is optimized by the PID controller whose proportional, integral and derivative values are shown in Table 2. Once the control system equations are obtained, they are connected to the 6-DOF linear model on the block diagram shown in Figure 5 (right).

Table 2: Proportional, Integral and Derivative Values of SPID.

	K_p	K_i	K_d
Surge	2.1	0	0
Sway	2.01	0	0
Heave	2.01	0	0
Roll	2.01	0	0
Pitch	2.1	0	0
Yaw	2.1	0	0

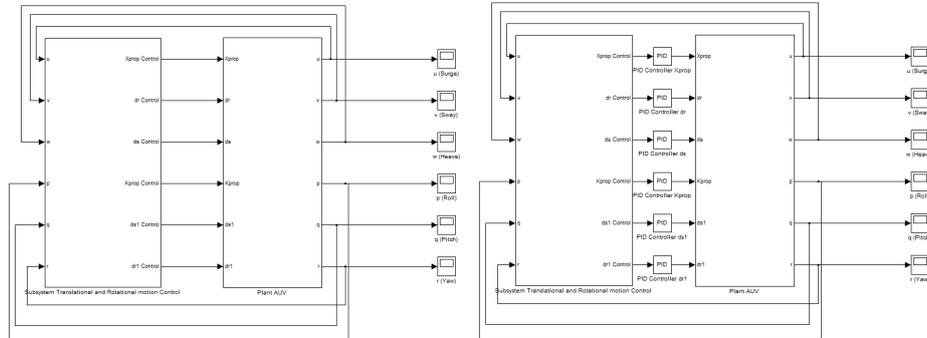


Figure 5: The Left Panel Displays the ROV Block Diagram Using the SMC Control System. The Right Panel Displays the ROV Block Diagram Using the SPID Control System.

7 Computational Results

In the simulation, the response ratio of SMC and SPID control systems for surge, sway, heave, roll, pitch and yaw motion was obtained. The setpoint of surge is 1 m/s, those of sway and heave are 1 m/s, while the setpoint for roll rotation motion is 1 rad/s, and those of pitch and yaw are -1 rad/s. From the simulation results, the comparison of time delay, rise time, peak time and settling time was obtained. The simulation results by SMC and SPID control systems are shown in Figure 6.

The comparison of the surge responses in Figure 6 (top left) is the result of AUV simulation by SMC and SPID control systems. The results of both methods show similarity of the delay time of 0.03 s, the rise time of 0.05 s, the peak time of 0.08 s, and the settling time of 1 s. SMC and SPID are stable at the setpoint of 1 m/s at time of 1 s. When viewed from the resulted error, SPID is 3.4%, while SMC is 4.15%. Figure 6 (top right) shows the similarities for the delay time, rise time, peak time and settling time. That is, the delay time is 0.037 s, the rise time is 0.065 s, the peak time is 0.07 s, and the settling time is 0.1 s. SMC and SPID are stable at the setpoint -1 m/s at 0.1 s. In terms of the resulted error, that of SPID is 0.09% and that of SMC is 0.11%. In Figure 6 (middle left), it is shown that the results of both methods have similarity for a delay time, that is, a delay time of 0.037 s. SMC has a rise time of 0.15 s and a settling time of 0.2 s, while SPID has a rise time of 0.25 s and a settling time of 0.25 s. SMC and SPID are stable at the setpoint 1 m/s at 0.2 s. SMC and SPID on the heave response do not have a peak time and maximum overshoot. As regards the error generated, that of SMC is 0.19% and that of SPID is 0.8%.

The comparison of the roll response in Figure 6 (middle right) is the result of the simulation of the AUV with SMC and SPID control systems. The results of both methods show similarity of a delay time of 0.04 s, rise time of 0.12 s, and settling time of 0.12 s. The response generated by the SMC and SPID methods is stable at the setpoint 1 rad/s at 0.12 s. SMC and SPID on the roll response do not have a peak time and maximum overshoot. In terms of the error generated, that of SMC is 0.42%, while that of SPID is 0.6%, so SMC is better than SPID for the roll response. Figure 6 (bottom left) shows that the response results by SMC and SPID methods are stable at the setpoint -1 rad/s, the SMC settling time at 1.5 s, while that of SPID at 1.25 s. As regards the error generated, that of SPID is 4.3% while that of SMC is 4.58%, so SPID is better than SMC but it

has a higher overshoot. Figure 6 (bottom right) shows that the results of both methods show similarity in a delay time of 0.002 s, rise time of 0.082 s, peak time of 0.07 s, and settling time of 0.2 s. The results of SMC and SPID methods are stable at the setpoint -1 rad/s at 0.2 s. As regards the error generated, that of SMC is 3.66%, while that of SPID is 4.77%, so the the SMC method is better than the SPID one. And, the SMC method has a lower overshoot than the SPID one.

The comparison of delay time, rise time, peak time, maximum overshoot, settling time, and error in responses, respectively, by the default PID control system, identical PID, SMC and SPID is shown in Table 3 for surge, sway, heave response. Meanwhile, the comparison for responses in roll, pitch, and yaw is shown in Table 4. From this comparison of responses, it can be concluded that the best control system for the AUV is seen from the error and the settling time. The SPID method is more stable for the surge, sway and pitch motion, while the SMC method is more stable for the heave, roll and yaw motion. Next, it can be checked by making stability analysis using the Lyapunov method.

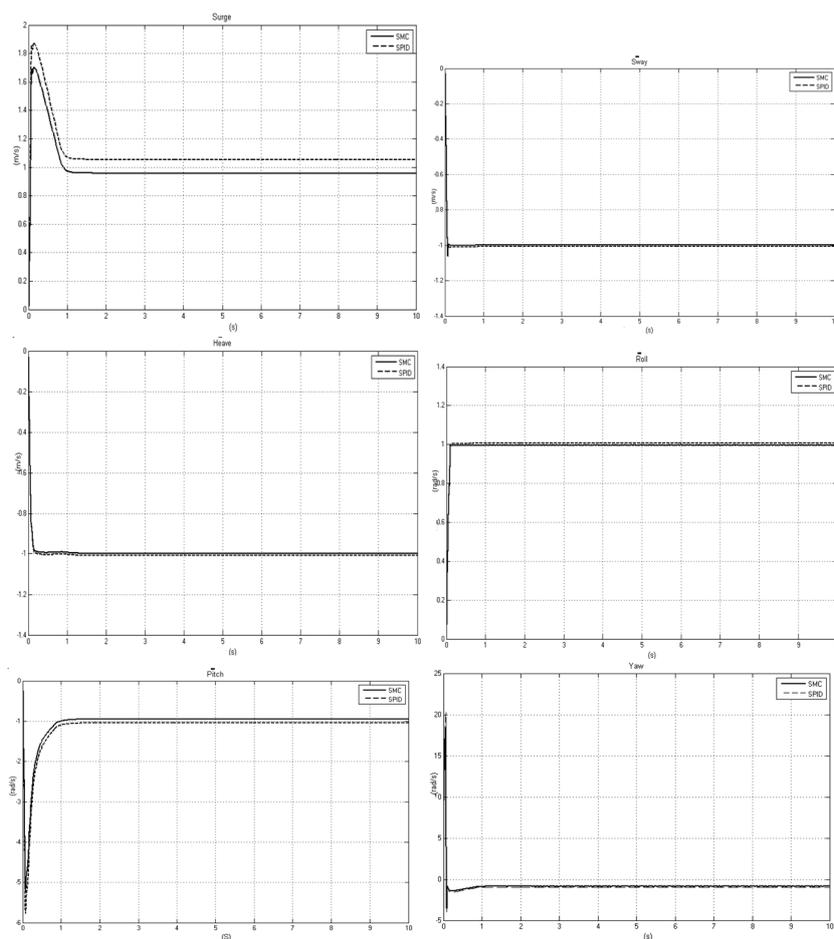


Figure 6: Response of the Surge, Sway, Heave, Roll, Pitch and Yaw Motions by Using SMC and SPID Control Systems.

Table 3: Specifications of the Transient Responses in the Surge, Sway, and Heave Motions.

	Surge		Sway		Heave	
	SMC	SPID	SMC	SPID	SMC	SPID
Delay Time	0.03 s	0.029 s	0.037 s	0.037 s	0.045 s	0.045 s
Rise Time	0.05 s	0.046 s	0.065 s	0.065 s	0.15 s	0.25 s
Peak Time	0.08 s	0.08 s	0.07 s	0.07 s	0 s	0 s
Maximum Peak	1.7 m/s	1.85 m/s	-1.1 m/s	-1.2 m/s	0 m/s	0 m/s
Settling Time	1 s	1 s	0.1 s	0.1 s	0.2 s	0.25 s
Error	4.15%	3.4%	0.11%	0.9%	0.19%	0.8%

Table 4: Specification of the Transient Responses in the Roll, Pitch, and Yaw Motions.

	Roll		Pitch		Yaw	
	SMC	SPID	SMC	SPID	SMC	SPID
Delay Time	0.04 s	0.04 s	0.01 s	0.01 s	0.002 s	0.002 s
Rise Time	0.12 s	0.12 s	0.2 s	0.2 s	0.082 s	0.082 s
Peak Time	0 s	0 s	0.075 s	0.074 s	0.07 s	0.07 s
Maximum Peak	0 m/s	0 m/s	-5.2 m/s	-5.8 m/s	18 m/s	21 m/s
Settling Time	0.12 s	0.12 s	1.5 s	1.25 s	0.2 s	0.2 s
Error	0.42%	0.6%	4.58%	4.3%	3.66%	4.77%

8 Stability Analysis

A candidate of the Lyapunov function for the ROV linear system with 6-DOF is

$$V(u, v, w, p, q, r) = \frac{1}{2}u^2 + \frac{1}{2}v^2 + \frac{1}{2}w^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{1}{2}r^2. \tag{40}$$

We will show that $V(u, v, w, p, q, r)$ is a Lyapunov function that satisfies the stability criteria.

For the Lyapunov candidate function in equation (40), it will be proven that the candidate function with the SMC and SPID control system in the linear model is the Lyapunov function and the equilibrium point is asymptotically stable.

1. For $(u, v, w, p, q, r) = (0, 0, 0, 0, 0, 0)$, we obtain $V(u, v, w, p, q, r) = 0$, whereas for $(u, v, w, p, q, r) \neq (0, 0, 0, 0, 0, 0)$, we obtain $V(u, v, w, p, q, r) > 0$. Then $V(u, v, w, p, q, r)$ has been proven to be positive definite for SMC and SPID.
2. The function V is continuous and it has a continuous first partial derivative at S . The function V in (40) is a quadratic function, then it is easy to see that the quadratic function is continuous. As a consequence, the partial derivative is also continuous.

3. The third requirement of the SMC method is as follows:

$$\begin{aligned}
 \dot{V}(u, v, w, p, q, r) &= \frac{\partial V}{\partial u} \dot{u} + \frac{\partial V}{\partial v} \dot{v} + \frac{\partial V}{\partial w} \dot{w} + \frac{\partial V}{\partial p} \dot{p} + \frac{\partial V}{\partial q} \dot{q} + \frac{\partial V}{\partial r} \dot{r} \\
 &= u\dot{u} + v\dot{v} + w\dot{w} + p\dot{p} + q\dot{q} + r\dot{r} \\
 &= u(aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1X_{prop} + BB_1\delta_{r_1} + \\
 &\quad CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}) + v(aa_2u + bb_2v + cc_2w + \\
 &\quad dd_2p + ee_2q + gg_2r + AA_2X_{prop} + BB_2\delta_{r_1} + CC_2\delta_{s_1} + DD_2K_{prop} + \\
 &\quad EE_2\delta_{s_2} + GG_2\delta_{r_2}) + w(aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r + \\
 &\quad AA_3X_{prop} + BB_3\delta_{r_1} + CC_3\delta_{s_1} + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2}) + \\
 &\quad p(aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r + AA_4X_{prop} + BB_4\delta_{r_1} + \\
 &\quad CC_4\delta_{s_1} + DD_4K_{prop} + EE_4\delta_{s_2} + GG_4\delta_{r_2}) + q(aa_5u + bb_5v + cc_5w + \\
 &\quad dd_5p + ee_5q + gg_5r + AA_5X_{prop} + BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + \\
 &\quad EE_5\delta_{s_2} + GG_5\delta_{r_2}) + r(aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r + \\
 &\quad AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2} + GG_6\delta_{r_2}).
 \end{aligned} \tag{41}$$

We take

$$\begin{aligned}
 X_{prop} &= - \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r}{AA_1} \right) - \\
 &\quad \left(\frac{BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}}{AA_1} \right) - K_1 \text{sgn}(S),
 \end{aligned} \tag{42}$$

$$\begin{aligned}
 \delta_{r_1} &= - \left(\frac{aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r}{BB_2} \right) - \\
 &\quad \left(\frac{AA_2X_{prop} + CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2}}{BB_2} \right) - K_2 \text{sgn}(S),
 \end{aligned} \tag{43}$$

$$\begin{aligned}
 \delta_{s_1} &= - \left(\frac{aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r}{CC_3} \right) - \\
 &\quad \left(\frac{AA_3X_{prop} + BB_3\delta_{r_1} + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2}}{CC_3} \right) - K_3 \text{sgn}(S),
 \end{aligned} \tag{44}$$

$$\begin{aligned}
 K_{prop} &= - \left(\frac{aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r}{DD_4} \right) - \\
 &\quad \left(\frac{AA_4X_{prop} + BB_4\delta_{r_1} + CC_4\delta_{s_1} + EE_4\delta_{s_2} + GG_4\delta_{r_2}}{DD_4} \right) - K_4 \text{sgn}(S),
 \end{aligned} \tag{45}$$

$$\delta_{s_2} = - \left(\frac{aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r}{EE_5} \right) - \left(\frac{AA_3X_{prop} + BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + GG_5\delta_{r_2}}{EE_5} \right) - K_5\text{sgn}(S), \tag{46}$$

$$\delta_{r_2} = - \left(\frac{aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r}{GG_6} \right) - \left(\frac{AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2}}{GG_6} \right) - K_6\text{sgn}(S). \tag{47}$$

Substitute equations (42)-(47) into equation (41), then we obtain

$$\begin{aligned} \dot{V}(u, v, w, p, q, r) = & u(-AA_1K_1\text{sgn}(S)) + v(-BB_2K_2\text{sgn}(S)) + w(-CC_3K_3\text{sgn}(S)) + \\ & p(-DD_4K_4\text{sgn}(S)) + q(-EE_5K_5\text{sgn}(S)) + r(-GG_6K_6\text{sgn}(S)) \leq \\ & (-AA_1K_1)u + (-BB_2K_2)v + (-CC_3K_3)w + (-DD_4K_4)p + \\ & (-EE_5K_5)q + (-GG_6K_6)r. \end{aligned}$$

We take $K_1 = \left| \frac{1}{AA_1} \right| \eta$, $K_2 = \left| \frac{1}{BB_2} \right| \eta$, $K_3 = \left| \frac{1}{CC_3} \right| \eta$, $K_4 = \left| \frac{1}{DD_4} \right| \eta$, $K_5 = \left| \frac{1}{EE_5} \right| \eta$, $K_6 = \left| \frac{1}{GG_6} \right| \eta$. Then we obtain

$$\begin{aligned} \dot{V}(u, v, w, p, q, r) \leq & (-\eta)|u| + (-\eta)|v| + (-\eta)|w| + (-\eta)|p| + (-\eta)|q| + (-\eta)|r| \\ \leq & -\eta(|u| + |v| + |w| + |p| + |q| + |r|). \end{aligned}$$

From the above requirements, the function V is the function of Lyapunov and the system is asymptotically stable.

The third characteristics of the SPID method is as follows:

$$\begin{aligned} X_{prop} = & - \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r}{AA_1} \right) - \\ & \left(\frac{BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}}{AA_1} \right) + \\ & \left(K_{p1}u + K_{i1}\frac{1}{2}u^2 + K_{d1}\dot{u} \right), \tag{48} \\ \delta_{r_1} = & - \left(\frac{aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r}{BB_2} \right) - \\ & \left(\frac{AA_2X_{prop} + CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2}}{BB_2} \right) + \end{aligned}$$

$$\left(K_{p2}v + K_{i2}\frac{1}{2}v^2 + K_{d2}\dot{v} \right), \quad (49)$$

$$\delta_{s_1} = - \left(\frac{aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r}{CC_3} \right) - \left(\frac{AA_3X_{prop} + BB_3\delta_{r_1} + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2}}{CC_3} \right) + \left(K_{p3}w + K_{i3}\frac{1}{2}w^2 + K_{d3}\dot{w} \right), \quad (50)$$

$$K_{prop} = - \left(\frac{aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r}{DD_4} \right) - \left(\frac{AA_4X_{prop} + BB_4\delta_{r_1} + CC_4\delta_{s_1} + EE_4\delta_{s_2} + GG_4\delta_{r_2}}{DD_4} \right) + \left(K_{p4}p + K_{i4}\frac{1}{2}p^2 + K_{d4}\dot{p} \right), \quad (51)$$

$$\delta_{s_2} = - \left(\frac{aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r}{EE_5} \right) - \left(\frac{AA_5X_{prop} + BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + GG_5\delta_{r_2}}{EE_5} \right) + \left(K_{p5}q + K_{i5}\frac{1}{2}q^2 + K_{d5}\dot{q} \right), \quad (52)$$

$$\delta_{r_2} = - \left(\frac{aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r}{GG_6} \right) - \left(\frac{AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2}}{GG_6} \right) + \left(K_{p6}r + K_{i6}\frac{1}{2}r^2 + K_{d6}\dot{r} \right). \quad (53)$$

Then equations (48)-(53) will be substituted to X_{prop} , δ_{r_1} , δ_{s_1} , K_{prop} , δ_{s_2} and δ_{r_2} in equation (41). The following equation is obtained:

$$\begin{aligned} \dot{V}(u, v, w, p, q, r) = & u \left(\frac{aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1(K_{p1}u + K_{i1}\frac{1}{2}u^2)}{1 - AA_1K_{d1}} + \frac{BB_1\delta_{r_1} + CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2}}{1 - AA_1K_{d1}} \right) + \\ & v \left(\frac{aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r + AA_2X_{prop}}{1 - BB_2K_{d2}} + \frac{BB_2(K_{p2}v + K_{i2}\frac{1}{2}v^2) + CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2}}{1 - BB_2K_{d2}} \right) + \\ & w \left(\frac{aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r + AA_3X_{prop}}{1 - CC_3K_{d3}} + \frac{BB_3\delta_{r_1} + CC_3(K_{p3}w + K_{i3}\frac{1}{2}w^2) + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2}}{1 - CC_3K_{d3}} \right) + \end{aligned}$$

$$\begin{aligned}
 & p \left(\frac{aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r + AA_4X_{prop}}{1 - DD_4K_{d4}} \right. \\
 & \left. + \frac{BB_4\delta_{r_1} + CC_4\delta_{s_1} + DD_4(K_{p4}p + K_{i4}\frac{1}{2}p^2) + EE_4\delta_{s_2} + GG_4\delta_{r_2}}{1 - DD_4K_{d4}} \right) + \\
 & q \left(\frac{aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r + AA_5X_{prop}}{1 - EE_5K_{d5}} + \right. \\
 & \left. \frac{BB_5\delta_{r_1} + CC_5\delta_{s_1} + DD_5K_{prop} + EE_5(K_{p5}q + K_{i5}\frac{1}{2}q^2) + GG_5\delta_{r_2}}{1 - EE_5K_{d5}} \right) + \\
 & r \left(\frac{aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r + AA_6X_{prop} + BB_6\delta_{r_1}}{1 - GG_6K_{d6}} + \right. \\
 & \left. \frac{CC_6\delta_{s_1} + DD_6K_{prop} + EE_6\delta_{s_2} + GG_6(K_{p6}r + K_{i6}\frac{1}{2}r^2)}{1 - GG_6K_{d6}} \right) \tag{54}
 \end{aligned}$$

Next, we take

$$\begin{aligned}
 K_7 &= \left| aa_1u + bb_1v + cc_1w + dd_1p + ee_1q + gg_1r + AA_1 \left(K_{p1}u + K_{i1}\frac{1}{2}u^2 \right) + BB_1\delta_{r_1} + \right. \\
 & \quad \left. CC_1\delta_{s_1} + DD_1K_{prop} + EE_1\delta_{s_2} + GG_1\delta_{r_2} \right| \eta_1, \\
 K_8 &= \left| aa_2u + bb_2v + cc_2w + dd_2p + ee_2q + gg_2r + AA_2X_{prop} + BB_2 \left(K_{p2}v + K_{i2}\frac{1}{2}v^2 \right) + \right. \\
 & \quad \left. CC_2\delta_{s_1} + DD_2K_{prop} + EE_2\delta_{s_2} + GG_2\delta_{r_2} \right| \eta_2, \\
 K_9 &= |aa_3u + bb_3v + cc_3w + dd_3p + ee_3q + gg_3r + AA_3X_{prop} + BB_3\delta_{r_1} + \\
 & \quad CC_3 \left(K_{p3}w + K_{i3}\frac{1}{2}w^2 \right) + DD_3K_{prop} + EE_3\delta_{s_2} + GG_3\delta_{r_2} | \eta_3, \\
 K_{10} &= |aa_4u + bb_4v + cc_4w + dd_4p + ee_4q + gg_4r + AA_4X_{prop} + BB_4\delta_{r_1} + CC_4\delta_{s_1} + \\
 & \quad DD_4 \left(K_{p4}p + K_{i4}\frac{1}{2}p^2 \right) + EE_4\delta_{s_2} + GG_4\delta_{r_2} | \eta_4, \\
 K_{11} &= |aa_5u + bb_5v + cc_5w + dd_5p + ee_5q + gg_5r | \eta_5, \\
 K_{12} &= |aa_6u + bb_6v + cc_6w + dd_6p + ee_6q + gg_6r + AA_6X_{prop} + BB_6\delta_{r_1} + CC_6\delta_{s_1} + \\
 & \quad DD_6K_{prop} + EE_6\delta_{s_2} + GG_6 \left(K_{p6}r + K_{i6}\frac{1}{2}r^2 \right) | \eta_6,
 \end{aligned}$$

where $\eta_1 = \frac{1}{1-AA_1K_{d1}}$, $\eta_2 = \frac{1}{1-BB_2K_{d2}}$, $\eta_3 = \frac{1}{1-CC_3K_{d3}}$, $\eta_4 = \frac{1}{1-DD_4K_{d4}}$, $\eta_5 = \frac{1}{1-EE_5K_{d5}}$ and $\eta_6 = \frac{1}{1-GG_6K_{d6}}$. Furthermore, $AA_1, K_{d1}, AA_2, K_{d2}, AA_3, K_{d3}, AA_4, K_{d4}, AA_5, K_{d5}, AA_6, K_{d6} > 1$. It follows that $1 - AA_1K_{d1} < 0$, $1 - BB_2K_{d2} < 0$, $1 - CC_3K_{d3} < 0$, $1 - DD_4K_{d4} < 0$, $1 - EE_5K_{d5} < 0$ and $1 - GG_6K_{d6} < 0$. Then we obtain that $\dot{V}(u, v, w, p, q, r) \leq 0$.

According to the above requirements, the function $V(u, v, w, p, q, r) = \frac{1}{2}u^2 + \frac{1}{2}v^2 + \frac{1}{2}w^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{1}{2}r^2$ is the Lyapunov function and the system is asymptotically stable.

If $V(x) \rightarrow \infty$ when $x \rightarrow \infty$, then the Lyapunov function is globally asymptotically stable. The above Lyapunov function is $V(u, v, w, p, q, r) = \frac{1}{2}u^2 + \frac{1}{2}v^2 + \frac{1}{2}w^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{1}{2}r^2$. It will be proven that $V(u, v, w, p, q, r) \rightarrow \infty$ when $u \rightarrow \infty$, $v \rightarrow \infty$, $w \rightarrow \infty$, $p \rightarrow \infty$, $q \rightarrow \infty$ and $r \rightarrow \infty$. Since $V(u, v, w, p, q, r) = \frac{1}{2}u^2 +$

$\frac{1}{2}v^2 + \frac{1}{2}w^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{1}{2}r^2$ is a quadratic function, if $u \rightarrow \infty$, $v \rightarrow \infty$, $w \rightarrow \infty$, $p \rightarrow \infty$, $q \rightarrow \infty$ and $r \rightarrow \infty$, then $V(u, v, w, p, q, r) \rightarrow \infty$. Thus, the Lyapunov function $V(u, v, w, p, q, r) = \frac{1}{2}u^2 + \frac{1}{2}v^2 + \frac{1}{2}w^2 + \frac{1}{2}p^2 + \frac{1}{2}q^2 + \frac{1}{2}r^2$ is asymptotically stable. In conclusion, the stability analysis of the SMC and SPID control systems has the stability property of being globally asymptotically stable.

9 Conclusions

Based on the simulation analysis of the SMC and SPID methods, it can be concluded that both methods have good stability for the ROV linear model with an error of about 0.09% - 4.5%. In terms of the delay time, rise time, peak time, maximum peak, and settling time, both methods have similarities. As regards the stability ratio for each motion performed, the SPID method is more stable for the surge, sway, and pitch motion, while the SMC method is more stable for the heave, roll and yaw motion.

Acknowledgement

High appreciation to the Kemdikbudristek for the very fund support for the research conducted in the year of 2023 with contract number 183/E5/PG.02.00.PL/2023, 074/SP2H/PT/LL7/2023 and 11/KP/LPPM/ITATS/2023.

References

- [1] S. Subchan, T. Herlambang and H. Nurhadi. UNUSAITs AUV Navigation and Guidance System with Nonlinear Modeling Motion Using Ensemble Kalman Filter. In: *Proc. International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA)* Malang, Indonesia, 2019.
- [2] F. Chiu, J. Guo, C. Huang and W. Tsai. Application of the Sliding Mode Fuzzy Controller to the Guidance and Control of An Autonomous Underwater Vehicle. In: *Proc. International Symposium on Underwater Technology (Cat. No. 00EX418)*, 2000. 181–186.
- [3] W. Naeem. Model Predictive Control of An Autonomous Underwater Vehicle. In: *Proc. UKACC Conference on Control*, 2002. 19–23.
- [4] K. Kim and T. Ura. 3-Dimensional Trajectory Tracking Control of An AUV “R-One” Robot Considering Current Interaction. In: *Proc. The Twelfth International Offshore and Polar Engineering Conference*, 2002.
- [5] F. Repoulia and E. Papadopoulos. Trajectory Planning and Tracking Control of Underactuated AUVs. In: *Proc. The IEEE International Conference on Robotics and Automation*, 2005. 1610–1615.
- [6] L. Lapierre and D. Soetanto. Nonlinear Path-Following Control of An AUV. *Ocean Engineering* **34** (11-12) (2007) 1734–1744.
- [7] H. Akçakaya, H. A. Yildiz, G. Sağlam and F. Gürleyen. Sliding Mode Control of Autonomous Underwater Vehicle. In: *Proc. International Conference on Electrical and Electronics Engineering-ELECO*, 2009. II-332.
- [8] F. Rezazadegan, K. Shojaee and A. Chatraei. Design of An Adaptive Nonlinear Controller for An Autonomous Underwater Vehicle. *Int. J. Adv. Electr. Electron. Eng.* **2** (2013) 1–8.
- [9] K. Oktafianto, T. Herlambang, Mardijah and H. Nurhadi. Design of Autonomous Underwater Vehicle Motion Control Using Sliding Mode Control Method. In: *Proc. International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA)* Surabaya, Indonesia, 2015, 162–166.

- [10] T. Herlambang, S. Subchan and H. Nurhadi. Design of Motion Control Using Proportional Integral Derivative for UNUSAITS AUV. *International Review of Mechanical Engineering IREME Journal* **12** (11) (2018) 928–938.
- [11] H. Nurhadi, T. Herlambang and D. Adzkiya. Design of Sliding Mode Control for Surge, Heave and Pitch Motion Control of UNUSAITS AUV. In: *Proc. International Conference on Mechanical Engineering Indonesia*, 2019.
- [12] T. Herlambang, S. Subchan, H. Nurhadi and D. Adzkiya. Motion Control Design of UNUSAITS AUV Using Sliding PID. *Nonlinear Dynamics and Systems Theory* **20** (1) (2020) 51–60.
- [13] F. A. Susanto, M. Y. Anshori, D. Rahmalia, K. Oktafianto, D. Adzkiya, P. Katias and T. Herlambang. Estimation of Closed Hotels and Restaurants in Jakarta as Impact of Corona Virus Disease (Covid-19) Spread Using Backpropagation Neural Network. *Nonlinear Dynamics and Systems Theory* **22** (4) (2022) 457–467.
- [14] C. Yang. Modular Modeling and Control for Autonomous Underwater Vehicle (AUV). *Master's thesis, National University of Singapore* Singapore, 2007.
- [15] T. Fossen. A Nonlinear Unified State-Space Model for Ship Maneuvering and Control in A Seaway. *International Journal of Bifurcation and Chaos* **15** (9) (2005) 2717–2746.
- [16] W. Nie and S. Feng. Planar Path-Following Tracking Control for An Autonomous Underwater Vehicle in the Horizontal Plane. *Optik* **127** (24) (2016) 11607–11616.



Class of Nilpotent Distributions and \mathfrak{N}_2 -Distributions

Y. Khellaf^{1*}, A. Chikh-Salah² and N. Bensalem³

¹ *Mathematics and Computer Science Department, Ghardaia University,
Fundamental and Numerical Mathematics Laboratory, Setif-1 University, Algeria*

² *Mathematics and Computer Science Department, Ghardaia university, Algeria.*

³ *Mathematics and Computer Science Department, Fundamental and Numerical Mathematics
Laboratory, Farhet Abbas University, Setif-1, Algeria.*

Received: July 9, 2023; Revised: November 26, 2023

Abstract: This paper presents a sufficient condition for two vector fields X and Y to have the squares noncommutative, i.e., $[X^2, Y^2] \neq 0$, in the case when X and Y span a 3-nilpotent distribution. And when the nilpotent distributions of class 2 or 3 are spanned from more than two vector fields, it gives the same result.

Keywords: *vector distributions; sub-Riemannian geometry; noncommutative geometry; nilpotency class; nonlinear dynamics systems.*

Mathematics Subject Classification (2010): 53A31, 53C17, 70K75, 93A10.

1 Introduction

The theory of subelliptic operators plays an important role in many applications in nonlinear dynamics and system theory, robotics and mechanical systems, optimal control of nonlinear systems, see [1, 9].

The subelliptic operator is a particular of case hypoelliptic differential equations. Hypoelliptic equations involve operators that are neither purely elliptic (like Laplace's equation) nor hyperbolic (like the wave equation), but rather fall in between. These equations often arise in the context of modeling systems with varying degrees of regularity and smoothness.

An example is the study of heat conduction in materials with varying degrees of conductivity, with a heat diffusion being non-uniform in all directions. The equation $\frac{\partial}{\partial t} - Lu = 0$ gives a more efficient description in that direction, where $u(x, t)$ is the heat kernel and the subelliptic operator L is defined in the differential manifold. And the

* Corresponding author: <mailto:khellaf.yasmina@univ-ghardaia.dz>

heat kernel characterizes the evolution of heat distribution over time in the context of the operator’s structure.

A heat kernel of the subelliptic operator $L = X_1^2 + \dots + X_k^2$, where X_1, \dots, X_k are vector fields of \mathbb{R}^n , with $k \leq n$, is an important problem. A sufficient condition for the hypo-ellipticity of the operator L is the bracket condition, see [2, 3, 6, 8]. If the squares of the aforementioned vector fields commute, i.e., $[X_i^2, X_j^2] = 0$ for all X_i, X_j , then the heat kernel of L is the product of heat kernels

$$e^{tL} = e^{tX_1^2} \dots e^{tX_k^2}.$$

If they do not commute, the previous formula does not hold any more, and the heat kernel should be found using a different method, see [4, 7, 12].

Ovidiu Calin and Der-Chen Chang in [5] give a sufficient condition for two vector fields X and Y to have the squares noncommutative, i.e., $[X^2, Y^2] \neq 0$, in the following result.

Theorem 1.1 [5]. *Any distribution $\mathcal{D} = \text{span}\{X, Y\}$ of nilpotency class 2 is a \mathfrak{N}_2 -Distribution, i.e., $[X^2, Y^2] \neq 0$.*

In the present work, as the first result, in Theorem 3.1, we shall prove that in the case of nilpotent distribution with the nilpotency class equal to 3, the squares of the vector fields do not commute. In the second and third results of this work, Proposition 3.1 and Proposition 3.2, we shall generalize the results of Theorem 1.1 and Theorem 3.1 for more than two vector fields.

2 Preliminaries

In an n -dimensional smooth manifold M , we recall that a smooth distribution \mathcal{D} of rank m is a rank m subbundle of the tangent bundle TM [11]. We call the \mathfrak{N}_k -Distribution a distribution spanned by two vector fields X and Y , which satisfies the condition $[X^k, Y^k] \neq 0$ [5].

Let $\Gamma(\mathcal{D})$ be a basis of the distribution \mathcal{D} . We recall that the iterated commutator sets C^k of vector fields obtained by k iterated Lie brackets of horizontal vector fields are

$$\begin{aligned} C^1 &= \{[X, Y]; X, Y \in \Gamma(\mathcal{D})\}, \\ &\vdots \\ C^n &= \{[C, Z]; C \in C^{n-1}, Z \in \Gamma(\mathcal{D})\}. \end{aligned}$$

A distribution \mathcal{D} is called nilpotent if there is an integer $k \geq 1$ such that $C^{(k)} = 0$, i.e., all the k iterated Lie brackets vanish. The smallest integer k is called the nilpotency class of \mathcal{D} which is called k -nilpotent, see page 47 in [6].

Example 2.1 (example of 3-nilpotent distribution [11]) The Martinet distribution in \mathbb{R}^3 (with coordinates (x, y, z)) is the distribution generated by X and Y with

$$X = \partial x, \quad Y = \partial y + \frac{1}{2}x^2\partial z,$$

the iterated commutators are

$$[X, Y] = x\partial z,$$

$$[X, [X, Y]] = \partial z, \quad [Y, [X, Y]] = 0,$$

and

$$\begin{aligned} [X, [X, [X, Y]]] &= 0, & [X, [Y, [X, Y]]] &= 0, \\ [Y, [X, [X, Y]]] &= 0, & [Y, [Y, [X, Y]]] &= 0. \end{aligned}$$

It follows that this distribution is nilpotent of class 3.

3 Main Results

Theorem 3.1 *Any distribution $\mathcal{D} = \text{span}\{X, Y\}$ of nilpotency class equal to 3 is a \mathfrak{N}_2 -distribution, i.e., $[X^2, Y^2] \neq 0$.*

To prove this theorem, we use several lemmas, and we recall the following. The distribution $\mathcal{D} = \text{span}\{X, Y\}$ is nilpotent of class 3 meaning that

$$[X, Y] \neq 0, \tag{1}$$

$$[Y, [X, Y]] \neq 0 \quad \text{or} \quad [X, [X, Y]] \neq 0,$$

and

$$[X, [X, [X, Y]]] = 0, \quad [Y, [Y, [X, Y]]] = 0, \tag{2}$$

$$[X, [Y, [X, Y]]] = 0, \quad [Y, [X, [X, Y]]] = 0. \tag{3}$$

Lemma 3.1 *In a distribution $\mathcal{D} = \text{span}\{X, Y\}$ of nilpotency class 3, we have*

$$[X^2, Y^2] = 0 \implies (XY)^2 = (YX)^2. \tag{4}$$

Proof. By developing the first equation of (3), we get

$$\begin{aligned} [X, [Y, [X, Y]]] = 0 &\iff X^2Y^2 - Y^2X^2 - 2(XY)^2 + 2(YX)^2 = 0 \\ &\implies X^2Y^2 - Y^2X^2 = 2((XY)^2 - (YX)^2) \end{aligned}$$

or

$$[X^2, Y^2] = 0,$$

then $(XY)^2 = (YX)^2$.

Lemma 3.2 *In a distribution $\mathcal{D} = \text{span}\{X, Y\}$ of nilpotency class 3, we have*

$$[X^2, Y^2] = 0 \implies XYX^2Y^2 = X^2Y^2XY. \tag{5}$$

Proof. The expansion of the equations (2) gives

$$X^3Y - 3X^2YX + 3XYX^2 - YX^3 = 0, \tag{6}$$

$$Y^3X - 3Y^2XY + 3YXY^2 - XY^3 = 0. \tag{7}$$

Multiplying the right-hand side, then the left-hand side of the relation (6) by Y^2 and the relation (7) by X^2 , we obtain

$$X^3Y^3 - 3X^2YXY^2 + 3XYX^2Y^2 - YX^3Y^2 = 0, \tag{8}$$

$$Y^3X^3 - 3Y^2XYX^2 + 3YXY^2X^2 - XY^3X^2 = 0, \tag{9}$$

$$Y^3X^3 - Y^2X^3Y + 3Y^2X^2YX - 3Y^2XYX^2 = 0, \tag{10}$$

$$X^3Y^3 - X^2Y^3X + 3X^2Y^2XY - 3X^2YXY^2 = 0. \tag{11}$$

By the subtractions of these equations, (8)-(9), (8)-(10), (11)-(10), we have found, respectively,

$$\begin{aligned}
 X^3Y^3 - Y^3X^3 &= 3X^2YXY^2 - 3XYX^2Y^2 + YX^3Y^2 - 3Y^2XYX^2 \\
 &\quad + 3YXY^2X^2 - XY^3X^2,
 \end{aligned} \tag{12}$$

$$\begin{aligned}
 X^3Y^3 - Y^3X^3 &= 3X^2YXY^2 - 3XYX^2Y^2 + YX^3Y^2 - Y^2X^3Y \\
 &\quad + 3Y^2X^2YX - 3Y^2XYX^2,
 \end{aligned} \tag{13}$$

$$\begin{aligned}
 X^3Y^3 - Y^3X^3 &= X^2Y^3X - 3X^2Y^2XY + 3X^2YXY^2 - Y^2X^3Y \\
 &\quad + 3Y^2X^2YX - 3Y^2XYX^2.
 \end{aligned} \tag{14}$$

Subtracting the equations (12)-(13) gives

$$\begin{aligned}
 -3XYX^2Y^2 + YX^3Y^2 + Y^2X^3Y - 3Y^2X^2YX \\
 - XY^3X^2 + 3YXY^2X^2 - X^2Y^3X + 3X^2Y^2XY = 0.
 \end{aligned}$$

In view of the fact that $X^2Y^2 = Y^2X^2$, the last equation becomes

$$-XYX^2Y^2 + YXX^2Y^2 + X^2Y^2XY - X^2Y^2YX = 0,$$

then

$$X^2Y^2[X, Y] = [X, Y]X^2Y^2. \tag{15}$$

On the other hand, subtracting the equations (14)-(12) gives

$$-3XYX^2Y^2 + YX^3Y^2 - X^2Y^3X + 3X^2Y^2XY = 0,$$

then

$$-[X, Y]X^2Y^2 + X^2Y^2[X, Y] + 2(X^2Y^2XY - XYX^2Y^2) = 0.$$

Using the relation (15), we obtain

$$XYX^2Y^2 = X^2Y^2XY.$$

Lemma 3.3 *In a distribution $\mathcal{D} = \text{span}\{X, Y\}$ with the nilpotency class 3, we have*

$$[X^2, Y^2] = 0 \implies X^2Y^2 = 3(XY)^2. \tag{16}$$

Proof. Multiplying the equation (6) in the proof of Lemma 3.2 by Y on two sides, we obtain

$$YX^3Y^2 - 3YX^2YXY + 3YXYX^2Y - Y^2X^3Y = 0. \tag{17}$$

Lemma 3.2 proves that

$$XYX^2Y^2 = X^2Y^2XY,$$

and interchanging X and Y , we get

$$YXX^2Y^2 = X^2Y^2YX,$$

then (17) becomes

$$X^2Y^2[X, Y] - 3((XY)^3 - (YX)^3) = 0,$$

this implies that

$$(X^2Y^2 - 3(XY)^2)[X, Y] = 0$$

but $[X, Y] \neq 0$, then

$$X^2Y^2 = 3(XY)^2.$$

Proof. (*Proof of Theorem 3.1*) We shall prove this theorem by contradiction, i.e., we assume that

$$[X^2, Y^2] = 0. \quad (18)$$

By developing $[X, Y]^3$ and using Lemma 3.1, we get

$$\begin{aligned} [X, Y]^3 &= (XY)^3 - (XY)^2(YX) - (XY)(YX)(XY) + (XY)(YX)^2 \\ &\quad - (YX)(XY)^2 + (YX)(XY)(YX) + (YX)^2(XY) - (YX)^3 \\ &= 3(XY)^3 - 3(YX)^3 - (XY)(YX)(XY) + (YX)(XY)(YX). \end{aligned} \quad (19)$$

Using Lemma 3.3, we get

$$\begin{aligned} (XY)(YX)(XY) &= XY^2X^2Y \\ &= 3X(YX)^2Y \\ &= 3XYXYXY = 3(XY)^3, \\ (YX)(XY)(YX) &= YX^2Y^2Y \\ &= 3Y(XY)^2X \\ &= 3YXYXYX \\ &= 3(YX)^3. \end{aligned}$$

The equation (19) becomes

$$[X, Y]^3 = 3(XY)^3 - 3(YX)^3 - 3(XY)^3 + 3(YX)^3 = 0,$$

then $[X, Y] = 0$ is a contradiction. It turns out that (18) cannot hold. It follows that the vector fields X and Y span a \mathfrak{N}_2 -distribution.

Example 3.1 [11] It is clear that the distribution $\mathcal{D} = \text{Span}\{X, Y\}$ is nilpotent of class 3 so that

$$X = \partial x, \quad Y = \partial y + \frac{1}{2}x^2\partial z,$$

and

$$[X^2, Y^2] = 4x\partial x\partial y\partial z + 2\partial y\partial z + 8x^3\partial x\partial_z^2 + 12x^2\partial_z^2 \neq 0.$$

For the second part of this paper, we need a new definition of \mathfrak{N}_k -distribution in the case when the distribution is spanned by more than two vector fields.

Definition 3.1 Let $\mathcal{D} = \text{span}\{X_1, X_2, \dots, X_m\}$ with $m \leq n$. We say that \mathcal{D} is a \mathfrak{N}_k -distribution if there exist $X_i, X_j \in \mathcal{D}$ such that $[X_i^k, X_j^k] \neq 0$.

In this proposition, we generalize Theorem 1.1 for more than two vector fields.

Proposition 3.1 *Let $\mathcal{D} = \text{span}\{X_1, \dots, X_m\}$ be a distribution spanned by m -vector fields of \mathbb{R}^n ($m \leq n$). Then if \mathcal{D} is a nilpotent distribution of nilpotency class 2, then \mathcal{D} is a \mathfrak{N}_2 -distribution, i.e., $\exists X_i, X_j \in \mathcal{D}$ such that $[X_i^2, X_j^2] \neq 0$.*

Proof. \mathcal{D} is a nilpotent distribution of nilpotency class 2, then there exist $X_i, X_j \in \mathcal{D}$ such that

$$[X_i, X_j] \neq 0$$

and

$$[X_i, [X_i, X_j]] = 0 \quad \text{and} \quad [X_j, [X_i, X_j]] = 0.$$

Let us tackle a sub-distribution $\mathcal{D}' = \text{span}\{X_i, X_j\}$, from Theorem 1.1, we obtain that in \mathcal{D}' ,

$$[X_i^2, X_j^2] \neq 0,$$

or $\mathcal{D}' \subset \mathcal{D}$, then we have in \mathcal{D} ,

$$[X_i^2, X_j^2] \neq 0.$$

Example 3.2 [10] The distribution $\mathcal{D} = \text{span}\{X, Y, Z\}$ such that

$$X = \partial x, \quad Y = \partial y, \quad Z = x\partial z$$

is nilpotent of class 2,

$$[X, Y] = 0, \quad [Y, Z] = 0, \quad [X, Z] = \partial z$$

and

$$\begin{aligned} [Y, [X, Y]] &= 0, & [X, [X, Y]] &= 0, \\ [X, [X, Z]] &= 0, & [Y, [Y, Z]] &= 0, \\ [Z, [X, Y]] &= 0. \end{aligned}$$

On the other hand, we have

$$[X^2, Z^2] = 4x\partial x\partial^2 z + 2\partial^2 z \neq 0.$$

In the next proposition, we generalize Theorem 3.1 for more than two vector fields.

Proposition 3.2 *Let $\mathcal{D} = \text{span}\{X_1, \dots, X_m\}$ be a distribution spanned by m -vector fields of \mathbb{R}^n ($m \leq n$). Then if \mathcal{D} is a nilpotent distribution of nilpotency class 3, then \mathcal{D} is a \mathfrak{N}_2 -distribution, i.e., $\exists X_i, X_j \in \mathcal{D}$ such that $[X_i^2, X_j^2] \neq 0$.*

Proof. \mathcal{D} is a nilpotent distribution of nilpotency class 3, then there exist $X_i, X_j \in \mathcal{D}$ such that

$$[X_i, X_j] \neq 0$$

and

$$\begin{aligned} [X_i, [X_i, [X_i, X_j]]] &= 0, & [X_j, [X_j, [X_i, X_j]]] &= 0, \\ [X_i, [X_j, [X_i, X_j]]] &= 0, & [X_j, [X_i, [X_i, X_j]]] &= 0. \end{aligned}$$

We remark that in the proof of Theorem 3.1, we do not use the iterated brackets of degree two (\mathcal{C}^2). Let us tackle a sub-distribution $\mathcal{D}' = \text{span}\{X_i, X_j\}$, from Theorem 1.1, we obtain that in \mathcal{D}' ,

$$[X_i^2, X_j^2] \neq 0,$$

or $\mathcal{D}' \subset \mathcal{D}$, then we have in \mathcal{D} ,

$$[X_i^2, X_j^2] \neq 0.$$

4 Conclusion

In summary, the heat kernel of a subelliptic operator captures the essential dynamic behavior of heat diffusion within a system defined by that operator. It describes how the initial heat distribution changes over time due to the conduction process governed by the operator's structure. The heat kernel provides insights into the temporal evolution of heat within the context of the subelliptic operator, connecting the mathematical description of heat conduction with the dynamic behavior of the system.

Unfortunately, the only method to find the solution of this kernel, is that the square of the vector fields commutes (sufficient condition).

Ovidiu Calin and Der-Chen Chang in [5] give a sufficient condition for two vector fields X and Y to have the squares noncommutative (distribution spanned with two vectors fields).

In this work, we have given an extension for more than two vector fields, and we have proposed a new sufficient condition for the noncommutative squares of the vectors fields in the case of distribution of nilpotency class 3. This result offers less computation and excellent description for the methods of finding the heat kernel of the sum of squares operator.

References

- [1] B. El-Aqqad and J. Oudaani and A. El Mouatasim. Existence, uniqueness of weak solution to the thermoelastic plates. *Nonlinear Dynamics and Systems Theory* **22** (3) (2022) 263–280.
- [2] R. Beals and B. Gaveau and P. C. Greiner. Hamilton-Jacobi theory and the heat kernel on Heisenberg groups. *J. Math. Pures Appl.* **79** (7) (2000) 633–689.
- [3] R. Beals and B. Gaveau and P. C. Greiner. On a geometric formula for the fundamental solution of subelliptic Laplacians. *Math. Nachr.* **181** (1996) 81–163.
- [4] O. Calin and D. C. Chang. *Geometric mechanics on Riemannian manifolds: applications to partial differential equations*. Applied and Numerical Analysis. Birkhäuser, Boston, 2004.
- [5] O. Calin and D. C. Chang. On a class of nilpotent distributions. *Taiwanese Journal of Mathematics* **15** (2) (2011) 875–881.
- [6] O. Calin and D. C. Chang. *Sub-Riemannian geometry, general theory and examples*. Encyclopedia of Mathematics and Its Applications, Cambridge University Press, Vol. 126, 2009.
- [7] O. Calin and D. C. Chang and K. Furutani and C. Iwasaki. *Heat Kernels for Elliptic and Sub-elliptic Operators*. Applied and Numerical Analysis, Birkhäuser, Boston, 2011.
- [8] L. Hörmander. Hypo-elliptic second order differential equations. *Acta Math.* **119** (1967) 147–171.
- [9] M. Khuddush and K. R. Prasad and B. Bharathi. Denumerably many positive radial solutions to iterative system of nonlynear elliptic equations on the exterior of a ball. *Nonlinear Dynamics and Systems Theory* **23** (1) (2023) 95–106.
- [10] A. E. Kogoj and E. Lanconelli. Linear and semilinear problems involving Δ_λ -Laplacians. *Electronic Journal of Differential Equations*. **25** (2018) 167–178.
- [11] L. Rifford. *Sub-Riemannian Geometry and Optimal Transport*. Springer, 2014.
- [12] L. S. Schulman. *Techniques and Applications of Path Integration*. Dover, 1981.



Moore-Spiegel Chaotic Encryption for Digital Images and Voices

W. S. M. Sanjaya^{1,2*}, A. Roziqin³, A. W. Temiesela¹,
M. F. B. Zaman¹, A. D. Wibiksana¹ and D. Anggraeni^{1,2}

¹Department of Physics, Faculty of Science and Technology, UIN Sunan Gunung Djati,
Bandung, Indonesia.

²Bolabot Techno Robotic Institute, CV Bolabot, Bandung, Indonesia.

³Department of Madrasah Ibtidaiyah Teacher Edu., Faculty of Education and Teaching, UIN
Sunan Gunung Djati, Bandung, Indonesia.

Received: June 20, 2023; Revised: November 11, 2023

Abstract: In this study, we explore the application of the Moore-Spiegel chaotic system in both image and voice encryption, considering the increasing importance of data security in the digital age. The analysis of the chaotic system involves examining phase diagrams, time series, bifurcation diagrams, Lyapunov exponent analysis, and Poincaré maps to understand its dynamics. For image encryption, we evaluate the system's effectiveness through various analyses, including histogram analysis, correlation analysis, entropy analysis, NPCR and UACI analysis, and noise attack analysis. Similarly, for voice encryption, we assess it through various analyses, including waveform plots, FFT, spectrograms, correlation coefficients, entropy analysis, and RMSE. The findings demonstrate the suitability of the Moore-Spiegel chaotic system for both image and voice encryption, suggesting its potential as a data transmission masking technique. The research includes numerical simulations conducted using Python to support the proposed approach.

Keywords: *Moore-Spiegel chaotic system; chaotic analysis; voice encryption; masking method; image encryption; XOR method.*

Mathematics Subject Classification (2010): 94A60, 37D45, 34F10, 74H65, 34D08, 94A08, 94A12.

* Corresponding author: <mailto:madasws@gmail.com>; mada.sanjaya@uinsgd.ac.id

1 Introduction

In recent times, ensuring the security and integrity of data transmitted through communication systems has become a significant focus for scientists. Research on chaos over the past four decades has revealed its complex and unpredictable behavior in various domains such as physics, climatology, chemistry, and biology. [1]

The Moore-Spiegel system, discovered in 1966, is a chaotic system that describes aperiodic dynamics and has applications in understanding thermal dissipation and convectively unstable fluids [2]. In 2017, the Moore-Spiegel synchronization circuit was applied to a communication security system [3].

Numerous studies have explored the application of chaotic systems in encryption, including image encryption using 1D, 2D, and 3D chaotic systems [4–6], as well as voice encryption using the Jerk, Chua, and Bhalekar-Gejji chaotic systems [7–9]. While the previous research demonstrates the potential of chaotic systems in encryption, further optimization and security analysis are required.

This paper is structured as follows. Section 2 discusses the analysis of the Moore-Spiegel Chaotic System, including phase diagrams, time series, bifurcation analysis, Lyapunov exponent analysis, and Poincaré analysis. Section 3 explains the encryption and decryption algorithms for digital images and voice using the Moore-Spiegel system. Section 4 presents experimental results and analysis. For image encryption, we perform histogram analysis, correlation analysis, entropy analysis, NPCR and UACI analysis, and noise attack analysis. For voice encryption, we conduct voice signal plot analysis, correlation coefficient analysis, voice entropy analysis, and Root Mean Squared Error (RMSE) analysis. Finally, Section 5 concludes and provides a final assessment.

2 Moore-Spiegel Chaotic System and Basic Analysis

The Moore-Spiegel system is described by the following system of differential equations:

$$\begin{cases} \dot{x} &= y, \\ \dot{y} &= z, \\ \dot{z} &= -z + ay - x^2y - bx. \end{cases} \quad (1)$$

The parameters and the initial conditions of Moore-Spiegel chaotic system are chosen as: $a = 9$, $b = 5$, and $(x_0, y_0, z_0) = (2, 7, 4)$ so that the system shows the expected chaotic behavior [2, 3, 10].

2.1 Phase diagram and time series

The signal plots of the Moore-Spiegel chaotic system with the constant parameters $a = 9$ and $b = 5$, and the initial condition of $(2, 7, 4)$ were simulated using Python 3, as depicted in Figure 1. The corresponding phase diagrams are illustrated in Figure 1. The graphs obtained from the Moore-Spiegel system equations exhibit chaotic behavior, indicating their suitability for digital image and voice encryption.

The utilization of phase diagrams and time series diagrams provides a solution for analyzing the Moore-Spiegel system's differential equation. These diagrams allow the observation of the system's movement characteristics. By examining the time series plot of variables x , y , and z in Figure 2, it becomes evident that the Moore-Spiegel system exhibits chaotic behavior.

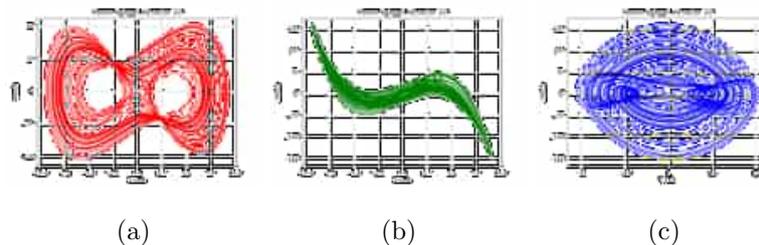


Figure 1: The 2D phase diagram of the Moore-Spiegel chaotic system; (a) y vs x , (b) z vs x , (c) z vs y , with the parameter values $a = 9$, and $b = 5$.

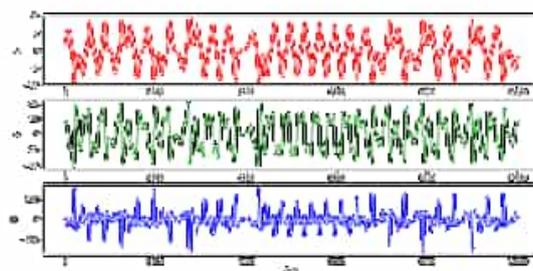


Figure 2: The time series for the Moore-Spiegel chaotic system with the parameter values $a = 9$ and $b = 5$.

2.2 Bifurcation analysis

The bifurcation diagram represents the transition of a discrete dynamical system from regular behavior to chaos [11]. Bifurcation analysis was carried out for the Moore-Spiegel chaotic system using a fixed set of parameter values, starting from the initial condition of $(2, 7, 4)$. We perform 10,000 iterations with a time step of 0.01. The findings depicted in Figure 3 illustrate the bifurcation diagram of the Moore-Spiegel system for the parameter values within the range of $5.0 \leq a \leq 10.0$, revealing the presence of chaotic dynamics with periodic patterns. Similarly, Figure 4 illustrates the occurrence of chaotic behavior with periodic patterns in the system when the parameter value falls within the range of $3.0 \leq b \leq 10.0$.

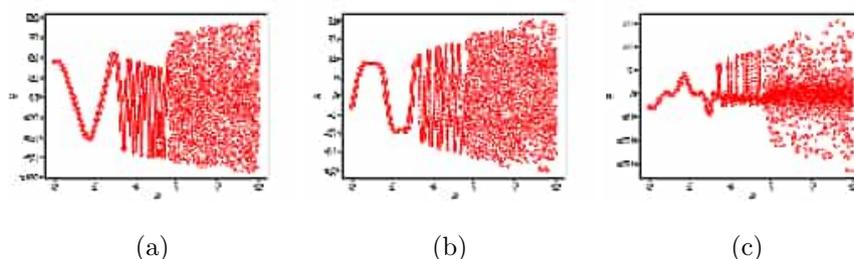


Figure 3: Bifurcation diagram with the parameters $b = 5$ and $a =$ varied; (a) x vs a , (b) y vs a , (c) z vs a .

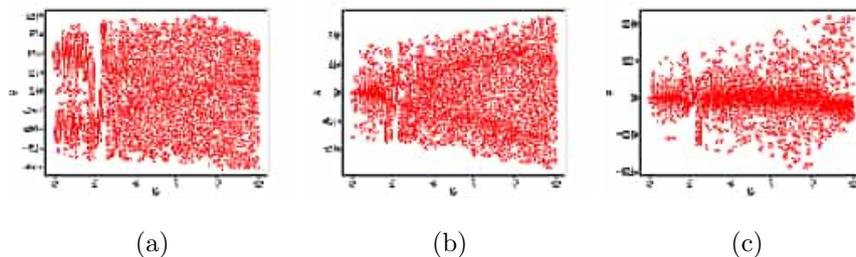


Figure 4: Bifurcation diagram with the parameters $a = 9$, and $b = \text{varied}$; (a) x vs b , (b) y vs b , (c) z vs b .

2.3 Lyapunov exponent

The Lyapunov exponent (λ_V) is a metric employed in the realm of dynamical systems theory to gauge how susceptible a system is to alterations in its initial conditions. We employed a predefined set of parameter values and initialized the simulation with an initial state of $(2, 7, 4)$. Following that, we conducted 10,000 iterations with a time increment of 0.01. Within the framework of the Moore-Spiegel chaotic system equations, the variable $v_i(a, b)$ signifies the value of the x , y , or z component at time step i within a simulation carried out with specific parameter values a , and b [12, 13]. The equation for the Lyapunov Exponent is stated as follows:

$$\lambda_V(a, b) = \frac{1}{dt} \ln \left(\frac{1}{N} \sum_{i=0}^{N-1} |v_i(a, b)| \right). \quad (2)$$

The Lyapunov exponent formula allows us to understand the extent to which the Moore-Spiegel chaotic system is sensitive to changes in its initial conditions and whether the system tends toward chaotic or stable behavior over time. The strange attractor shows three Lyapunov exponents with positive, zero, and negative values [3].

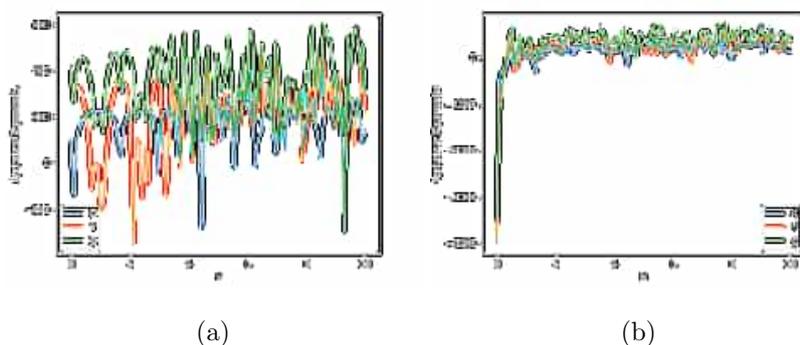


Figure 5: The Lyapunov exponents when the parameters are varied as follows: (a) $b = 5$ and $a = \text{varied}$, (b) $a = 9$ and $b = \text{varied}$.

Figure 5 illustrates chaotic behavior in the range of $5.0 \leq a \leq 10.0$, with other parameters held constant. The diagram displays periodic patterns amidst the chaos.

Furthermore, within the range of $3.0 \leq b \leq 10.0$, while keeping the other parameters constant, the system exhibits diverse chaotic behaviors for different parameter values, specifically $a = 9$ and $b = 5$. The system demonstrates limit point behavior due to the presence of two negative Lyapunov exponents.

2.4 Poincare analysis

The Poincaré map provides insights into periodic and non-periodic systems. Periodic systems exhibit a limited number of points with a repetitive structure, while non-periodic systems have a larger number of points with an unpredictable structure, with some points repeating.

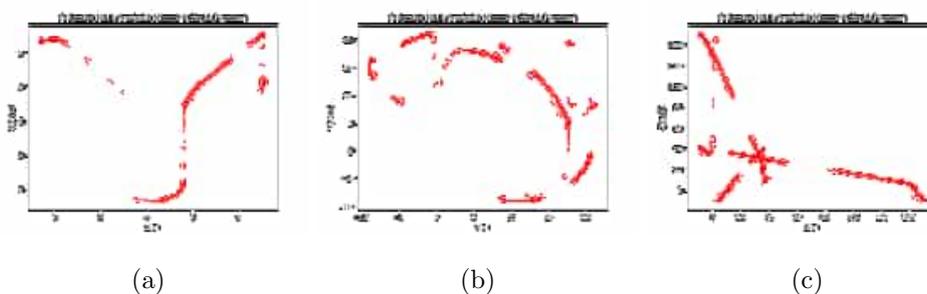


Figure 6: Poincaré map with the parameters $a = 9$, and $b = 5$; (a) $x(n+1)$ vs $x(n)$, (b) $y(n+1)$ vs $y(n)$, (c) $z(n+1)$ vs $z(n)$.

Figure 6 shows the Poincaré map for the Moore-Spiegel system, where $a = 9$ and $b = 5$. The map demonstrates chaotic behavior in the Moore-Spiegel chaotic system. It is characterized by scattered points with an irregular structure, and some points show repetition. The Poincaré map helps understand qualitative features of strange attractors during chaotic states by revealing dense intersections and different trajectories in each period. Analyzing the Poincaré map is crucial to comprehend the attractors’ qualitative characteristics.

3 The Moore-Spiegel System Algorithm

3.1 Encryption algorithm

- Input (Image) : Original Image (Bird, Landscape, Cat)
- Input (Voice) : Original Voice
- Output (Image) : Encrypted Image
- Output (Voice) : Encrypted Voice

- Step 1: Import the required libraries.
- Step 2: To generate the pseudo-random key, you can create a function named "secret-key" that utilizes the Moore-Spiegel chaotic system. For instance, you can select a solution from differential equation (1) that demonstrates chaotic behavior, such as x , y , or z , and use it as the basis for generating the pseudo-random key.

- Step 3: Specify the initial conditions and the "num" parameter responsible for inducing chaos in accordance with differential equations (1). As an illustration, you can set the initial condition to $(x, y, z) = (2, 7, 4)$ and use parameters $a = 9$ and $b = 5$.
- Step 4 (Image): Load the decrypted image and extract its height and width to be used in the "secret key" function.
- Step 4 (Voice): Read the voice file and convert the voice data into NumPy array format.
- Step 5: Iterate through each pixel (Image) or frame (Voice) in a loop.
- Step 6: Apply the XOR operation to encrypt the pixel (Image) or frame (Voice) using the pseudo-random numbers generated from the "secret key".
- Step 7: Obtain the encrypted image or encrypted voice.

3.1.1 Decryption algorithm

Input (Image) : Encrypted Image

Input (Voice) : Encrypted Voice

Output (Image) : Decrypted Image (Bird, Landscape, Cat)

Output (Voice) : Decrypted Voice

- Step 1: Import the required libraries.
- Step 2: Read the encrypted image or encrypted voice file and extract its dimensions or transform the voice data into a NumPy array.
- Step 3: Define the initial values and the "num" parameter that result in chaotic characteristics, similar to those in the encryption system. For instance, establish the initial condition as $(x, y, z) = (2, 7, 4)$, with parameters $a = 9$ and $b = 5$.
- Step 4: Produce a pseudo-random key through the development of a function named the "secret-key" that leverages the Moore-Spiegel chaotic system. For instance, opt for a solution within differential equation (1), be it x , y , or z , that demonstrates chaotic behavior to serve as the basis for the pseudo-random key.
- Step 5: Iterate through each pixel in the encrypted image or frame in the encrypted voice file.
- Step 6: Decrypt the pixel (Image) or frame (Voice) by performing XOR operation between the encrypted pixel (Image) or frame (Voice) and the corresponding pseudo-random number from the secret key.
- Step 7: Obtain the decrypted image or decrypted voice.

4 Experiment Results

In this section, we divide the experimental results into two parts: image encryption and voice encryption.

4.1 Image encryption

In this part, we conducted an evaluation of the Moore-Spiegel chaotic system's masking approach, utilizing three unique images: Bird, Landscape, and Cat. Our analysis was primarily focused on appraising the system's effectiveness in encrypting and ensuring the security of image files, as shown in Figure 7.

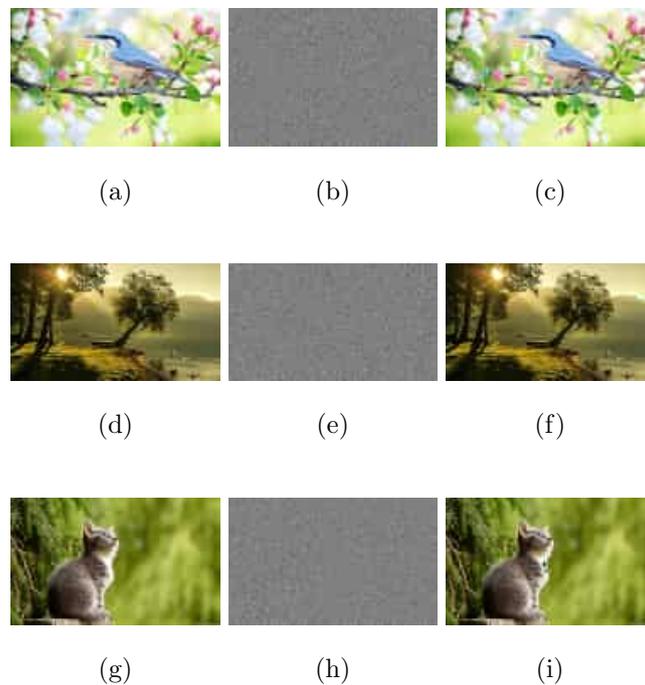


Figure 7: Application of the Moore-Spiegel chaotic system in Digital Image Encryption and Decryption. (a, d, g) Original. (b, e, h) Encrypted. (c, f, i) Decrypted.

4.1.1 Histogram image encryption analysis

Histogram analysis portrays the distribution of pixel intensities graphically in an image. It provides insights into the prevalence of different ranges of pixel intensities [14]. Pixels in an original image have distinct and information rich diagonal bars. These diagonal bars are vulnerable to attacks. To counter such attacks, the encryption algorithm should ensure that the encrypted image has evenly distributed bars [15]. Figure 8, 9, 10 show histogram analysis results for the original image, encrypted image, and decrypted image. The original image's histogram displays non-flat distribution with clustering tendencies on the left and right sides. The histogram of the encrypted image indicates a successful encryption with a flat distribution. The histogram of the decrypted image demonstrates a successful decryption, restoring a similar distribution as in the original image.

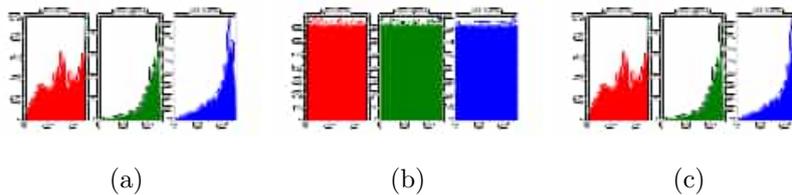


Figure 8: Histogram analysis of Bird image: (a) original, (b) encrypted, (c) decrypted.

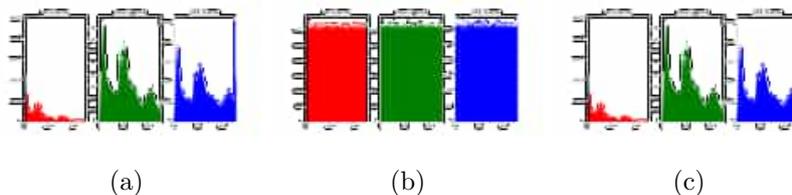


Figure 9: Histogram analysis of Landscape image: (a) original, (b) encrypted, (c) decrypted.

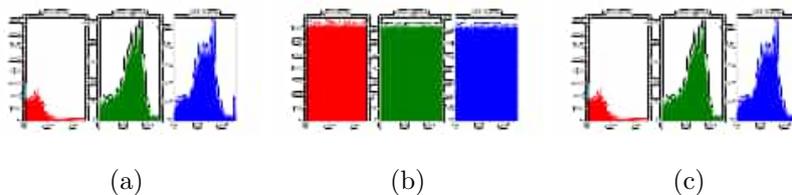


Figure 10: Histogram analysis of Cat image: (a) original, (b) encrypted, (c) decrypted.

4.1.2 Correlation image encryption analysis

Correlation analysis measures the degree of correlation between multiple images using a correlation coefficient ranging from -1 to 1. A correlation coefficient of 1 represents a perfect positive relationship, 0 indicates no relationship, and -1 represents a perfect negative relationship [1]. The correlation coefficient is calculated based on the relative ranking of pixel intensities in the images, rather than their actual values. It can be expressed mathematically as follows:

$$r_{xy} = \frac{Cov(x, y)}{\sqrt{D(x)D(y)}}, \quad (3)$$

where

$$Cov(x, y) = \frac{1}{N} \sum_{j=1}^N (x_j - E(x))(y_j - E(y)), \quad (4)$$

$$E(x) = \frac{1}{N} \sum_{j=1}^N X_j, \quad (5)$$

$$D(x) = \frac{1}{N} \sum_{j=1}^N (x_j - E(x))^2. \quad (6)$$

The correlation coefficient equation involves the expected values of variables x and y denoted by $E(x)$ and $E(y)$, respectively. The term Cov represents the covariance between the variables, while $D(x)$ and $D(y)$ indicate the standard deviations of x and y . The variables x_j and y_j refer to the individual pixels in the first and second images, and N represents the total number of pixels involved in the calculation [1].

Image	Channel	Original-Encrypted	Original-Decrypted
Bird	R	-0.0056	1.0000
	G	0.0288	0.9999
	B	-0.0078	1.0000
	Average	0.0052	1.0000
Landscape	R	-0.0230	0.9999
	G	-0.0308	0.9999
	B	0.0279	1.0000
	Average	-0.0087	0.9999
Cat	R	0.0036	1.0000
	G	0.0217	1.0000
	B	0.0236	0.9999
	Average	0.0163	1.0000

Table 1: Image correlation analysis.

Table 1 presents the results of correlation analysis conducted for the encrypted image and decrypted image of the three images using RGB analysis. The correlation coefficients between the pixels of the original images and the encrypted images are close to zero, indicating a lack of correlation and successful image masking. On the other hand, the correlation coefficients between the pixels of the original images and those of the decrypted images show a perfect correlation, confirming the algorithm’s successful execution.

4.1.3 Entropy image encryption analysis

Image entropy within the RGB channels serves as a statistical metric employed to evaluate the degree of uncertainty or randomness present in the distribution of pixel values within each individual color channel (namely, Red, Green, and Blue) in the color image. Entropy offers insights into the uniformity or concentration of information within each channel. A heightened entropy value implies a more haphazard distribution of pixel values, whereas a lower entropy value signifies a more structured or concentrated arrangement of color distribution [16]. The calculation for determining the information entropy value is defined by the following equation:

$$H(X) = - \sum_{i=1}^N p(x_i) \log_2(p(x_i)), \tag{7}$$

in this context, $H(X)$ stands for the entropy measure of the probability distribution X , $p(x_i)$ signifies the likelihood of pixel value x_i occurring within the distribution, N corresponds to the total count of unique pixel values in the distribution, and $\log_2(p(x_i))$ denotes the logarithm base-2 of the probability associated with pixel value x_i .

Based on Table 2, the entropy value for the encrypted image is higher than those for the original and decrypted images. The higher the entropy value (closer to 8), the higher the level of disorder in the information content of the image. This means that the information contained in the encrypted image is irregular, random, and difficult to comprehend.

Image	Channel	Original	Encrypted	Decrypted
Bird	R	7.35329	7.99193	7.35329
	G	7.07455	7.99194	7.07455
	B	7.83513	7.99198	7.83513
Landscape	R	7.84448	7.99202	7.84448
	G	7.80162	7.99197	7.80162
	B	7.04028	7.99200	7.04028
Cat	R	7.55674	7.99196	7.55674
	G	7.55905	7.99203	7.55905
	B	6.94755	7.99195	6.94755

Table 2: Image entropy values.

4.1.4 NPCR and UACI Analysis

1. **NPCR (Normalized Pixel Change Rate):** The NPCR (Normalized Pixel Change Rate) calculates the proportion of the pixel alterations between two images that have undergone encryption or decryption procedures. This parameter offers insight into the extent of pixel modifications that occur as a result of applying encryption or decryption algorithms [17] calculated by

$$\text{NPCR} = \frac{\text{Number of Changed Pixels}}{N \times M} \times 100\%, \quad (8)$$

where N and M represent the dimensions of the image in pixels, and the "Number of Changed Pixels" is the count of pixel positions where the pixel values differ between the two images.

2. **UACI (Unified Average Changing Intensity):** UACI quantifies the mean intensity variation between two images that are being compared. Intensity change is assessed by calculating the absolute difference between pixel values in the first and second images, and this value is then averaged across all pixels. UACI offers insights into the degree of intensity alterations that take place following an encryption or decryption process [17]. The UACI is computed using the following formula:

$$\text{UACI} = \frac{\text{Sum of Intensity Differences}}{N \times M \times L} \times 100\%, \quad (9)$$

where N and M stand for the pixel dimensions of the image, L signifies the total number of potential intensity levels for each pixel, which is typically 256 for an 8-bit image. The "Sum of Intensity Differences" corresponds to the summation of absolute disparities between corresponding pixel values in the two images.

In Table 3, the NPCR (%) column depicts the percentage of pixel changes between the original image and the encrypted image, as well as between the original image and the

decrypted image. The research results indicate significant pixel changes in the encrypted image, with percentages of approximately 98.50 %, 97.32 %, and 97.99 %, highlighting a noticeable transformation in the encrypted image compared to the original one.

Image	NPCR (%)		UACI (%)	
	Original - Encrypted	Original - Decrypted	Original - Encrypted	Original - Decrypted
Bird	98.50 %	0.00 %	30.81 %	0.00 %
Landscape	97.32 %	0.00 %	29.65 %	0.00 %
Cat	97.99 %	0.00 %	23.32 %	0.00 %

Table 3: NPCR and UACI percentages.

Meanwhile, the UACI (%) column of Table 3 illustrates the percentage of the pixel intensity changes between the original image and the encrypted image, as well as between the original image and the decrypted image. The research findings also reveal significant fluctuations in pixel intensity within the encrypted image. However, the decryption process successfully restores the image to its original pixel intensity levels, as evidenced by a UACI value of 0.00 % for all images. These findings provide an understanding of the impact of encryption and decryption algorithms on pixel changes and pixel intensity changes in the evaluated images.

4.1.5 Noise of attack image encryption analysis

The transmission of images may introduce distortions that can impact the final results. Therefore, it is crucial to evaluate the algorithm's performance against distortion attacks. The Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) are utilized to measure the impact of noise on the decrypted image [18]. The equations for calculating the MSE and PSNR are as follows:

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{Max}_i^2}{\text{MSE}} \right), \quad (10)$$

$$\text{MSE} = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (P_{ij} - C_{ij})^2. \quad (11)$$

In the given statement, P_{ij} represents the pixel value at the position i, j of the image without noise. C_{ij} represents the pixel value at the position i, j of the image with noise. Max_i denotes the maximum pixel value in the image.

We used density variations to modify the impact of noise. The MSE values were calculated by comparing the decrypted image after incorporating Salt and Pepper noise. A higher PSNR value indicates less information loss, while a higher MSE value indicates more information loss. In Table 4, it can be observed that at a density of 0.10, the MSE value is higher and the PSNR value is lower. Conversely, at a density of 0.05, the MSE value is lower and the PSNR value is higher.

Image	Variation of Density	MSE	PSNR
Bird	0.05	1148.51	17.53 dB
	0.10	2300.34	14.51 dB
Landscape	0.05	1131.49	17.59 dB
	0.10	2256.74	14.60 dB
Cat	0.05	1031.23	18.00 dB
	0.10	2062.23	14.99 dB

Table 4: The MSE and PSNR values were calculated for the decrypted image with salt and pepper noise.

4.2 Voice encryption

In this section, we tested the Moore-Spiegel chaotic masking system on three original voice files to ensure robust encryption, preserving confidentiality and integrity against unauthorized access and tampering.

4.2.1 Voice signal plot

Waveform plots visually represent the signal's amplitude changes over time, aiding interpretation. Each frame of the voice signal is then transformed from time to frequency domain using FFT, enabling the analysis of voice characteristics such as high and low frequencies and amplitude strength at each frequency. Spectrogram plots display the signal's frequency against time and are shown sequentially in Figures 11, 12, and 13.

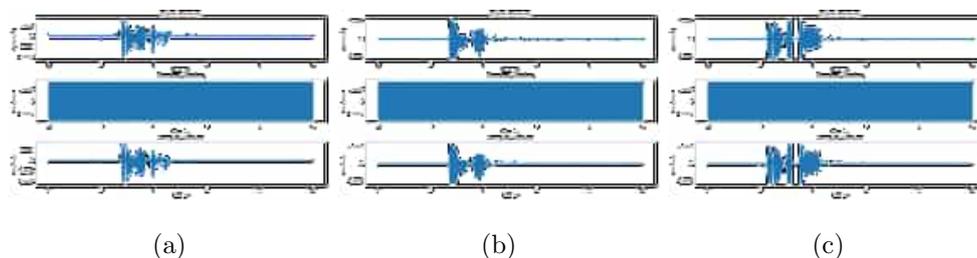


Figure 11: Waveform graph; (a) voice 1 ("Terima Kasih" in Bahasa), (b) voice 2 ("Thank You" in English), (c) voice 3 ("Arigatou" in Japanese).

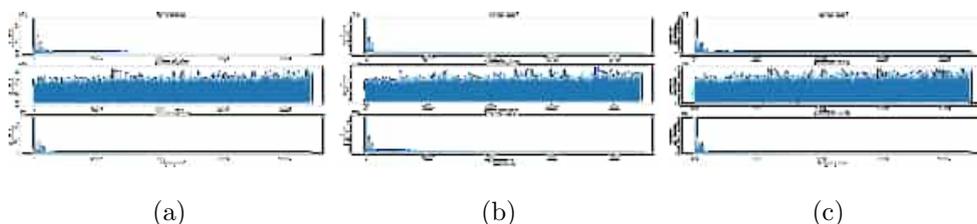


Figure 12: FFT graph; (a) voice 1 ("Terima Kasih" in Bahasa), (b) voice 2 ("Thank You" in English), (c) voice 3 ("Arigatou" in Japanese).

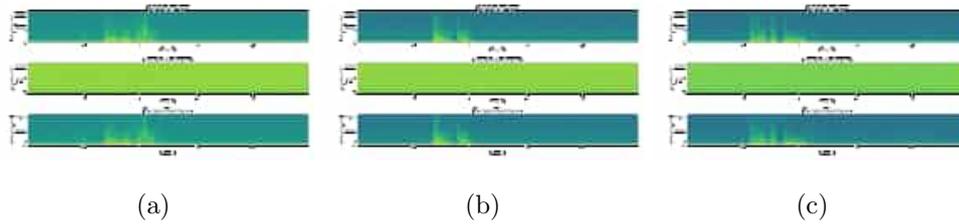


Figure 13: Spectrogram graph ; (a) voice 1 ("Terima Kasih" in Bahasa), (b) voice 2 ("Thank You" in English), (c) voice 3 ("Arigatou" in Japanese).

The three images show Waveform, FFT, and Spectrogram graphs for the original voice, encrypted voice, and decrypted voice, respectively, for three different voices: voice 1 (Bahasa: "Terima Kasih"), voice 2 (English: "Thank You"), and voice 3 (Japanese: "Arigatou"). The analysis reveals that the encrypted voice signal has a distinct pattern, while the decrypted voice signal closely resembles the original voice signal. This demonstrates the algorithm's effectiveness in preserving the high quality of the recovered voice signal.

4.2.2 Correlation voice encryption analysis

The correlation analysis in voice encryption is governed by equations (3), (4), (5), (6). For voice encryption, the correlation analysis is performed with a similar methodology which differs in terms of the parameters used. $Cov(x, y)$ represents the covariance between the original signal x and the encrypted signal y . $D(x)$ and $D(y)$ denote the variances of signals x and y , respectively. N represents the number of voice samples. A low value of the correlation coefficient r_{xy} indicates a high-quality encryption [19].

Voice File	Original-Encrypted	Original-Decrypted
recording1.wav (<i>Terima Kasih</i>)	-0.00389	1.0
recording2.wav (Thank You)	-0.00055	1.0
recording3.wav (<i>Arigatou</i>)	-0.00079	1.0

Table 5: Voice correlation analysis.

The correlation coefficients between the original and encrypted voice signals are close to 0, indicating a lack of correlation, while the correlation coefficients between the original and decrypted voice signals are close to 1, indicating a strong correlation. Please refer to Table 5 for the specific values.

4.2.3 Entropy voice encryption analysis

In the realm of audio, entropy serves as a statistical metric employed to gauge the degree of uncertainty or randomness present in the arrangement of audio sample values. A heightened audio entropy signifies an increased diversity in audio sample values, potentially indicating the presence of more intricate or unpredictable sounds. Conversely, reduced entropy implies that the sound typically exhibits patterns or repetitions, with limited fluctuations in sample values [20].

The computation of audio entropy can be accomplished through Shannon's entropy formula, which is articulated as follows:

$$H(X) = - \sum_{i=1}^N p(x_i) \log_2(p(x_i)). \quad (12)$$

In this formula, $H(X)$ denotes the entropy of the probability distribution X representing audio sample values, $p(x_i)$ stands for the likelihood of occurrence of audio sample value x_i within the distribution, and N signifies the count of distinct audio sample values in the distribution. The provided algorithm calculates entropy for three different audio files: the original voice recording, the encrypted voice, and the decrypted voice. This program proves to be invaluable in the analysis of how audio information undergoes transformation during encryption and decryption procedures, shedding light on the complexity or randomness level of the audio files. A higher entropy value indicates a greater degree of variation in sample values within the audio.

Voice File	Original	Encrypted	Decrypted
recording1.wav (<i>Terima Kasih</i>)	11.33927	15.76719	11.33927
recording2.wav (Thank You)	11.45686	15.76808	11.45686
recording3.wav (<i>Arigatou</i>)	11.69892	15.76718	11.69892

Table 6: Voice entropy value.

The findings derived from the Shannon Entropy table, as presented in Table 6, elucidate that the encryption procedure exerts a substantial influence on the degree of unpredictability or randomness inherent in the audio data. This is evidenced by the notable escalation in Shannon Entropy values following the encryption process. However, subsequent to the decryption phase, the audio data regains a level of unpredictability akin to its original state, signifying the successful preservation of the fundamental information within the audio data. This inference suggests that, within this specific context, the encryption-decryption process can furnish additional security measures against unauthorized access to audio data while upholding the integrity of the data encapsulated within the audio files.

4.2.4 Root mean squared error (RMSE)

The RMSE (Root Mean Squared Error) measures the deviation between the predicted and actual values. A lower RMSE value is desirable as it indicates higher accuracy in the model's predictions. The ideal RMSE value depends on the specific problem and data range. While there is no universally defined ideal value, a decreasing RMSE signifies improved accuracy in the prediction model.

$$RMSE = \sqrt{\sum \frac{(x_{\text{pred}} - x_{\text{act}})^2}{N}}, \quad (13)$$

where x_{pred} is a predicted value, x_{act} is an actual value and N is the total data. Table 7 shows the comparison of RMSE values for voice decryption using low-level noise, indicating that the algorithm performs well.

Voice File	Variaty Density	RMSE
recording1.wav (<i>Terima Kasih</i>)	0.10 0.01	0.1002 0.0100
recording2.wav (Thank You)	0.01 0.01	0.1002 0.0100
recording3.wav (<i>Arigatou</i>)	0.10 0.01	0.1001 0.0100

Table 7: The RMSE values for decrypted voice with Gaussian noise.

5 Conclusion

The previous research on the Moore-Spiegel system in communication security was limited to synchronization circuits, without applying it to image and voice encryption. Analysis shows that the Moore-Spiegel chaotic system is ideal for data encryption due to its unpredictability, randomness, and sensitivity to initial conditions, making it suitable for generating random encryption keys. The proposed algorithm's effectiveness has been substantiated through various analyses, including histogram, correlation, entropy, NPCR and UACI, and noise attack analyses for image and voice encryption. Further research potential lies in developing encryption methods for videos or other complex subjects.

Acknowledgment

The authors express their gratitude for the financial support provided by LP2M UIN Sunan Gunung Djati Bandung and DIKTIS Indonesian Ministry of Religious Affairs (Kementerian Agama Republik Indonesia).

References

- [1] D. Xu, G. Li, W. Xu and C. Wei. Design of artificial intelligence image encryption algorithm based on hyperchaos. *Ain Shams Engineering Journal* **14** (3) (2023) 101–891.
- [2] D. W. Moore and E. A. Spiegel. A thermally excited non-linear oscillator. *Astrophysical Journal* **143** (1966) 871.
- [3] W. S. M. Sanjaya, D. Anggraeni, R. Denya and N. Ismail. Numerical simulation bidirectional chaotic synchronization of spiegel-moore circuit and its application for secure communication. *Annual Applied Science and Engineering Conference* **180** (1) (2017) 89–95.
- [4] M. Essaid, I. Akharraz, A. Saaidi and E. A. Mouhib. Image encryption scheme based on a new secure variant of hill cipher and 1d chaotic maps. *Journal of Information Security and Applications* **47** (2019) 173–187.
- [5] A. Broumandnia. Designing digital image encryption using 2D and 3D reversible modular chaotic maps. *Journal of Information Security and Applications* **47** (2019) 188–198.
- [6] H. Liu, A. Kadir, and J. Liu. Color pathological image encryption algorithm using arithmetic over galois field and coupled hyper chaotic system. *Optics and Lasers in Engineering* **122** (2019) 123–133.
- [7] S. Vaidyanathan, A. Sambas, M. Mamat and W. S. M. Sanjaya. Analysis, synchronisation and circuit implementation of a novel jerk chaotic system and its application for voice encryption. *International Journal of Modelling, Identification and Control* **28** (2) (2017) 153–166.

- [8] K. Rajagopal, S. Kacar, Z. Wei, P. Duraisamy, K. Teweldbrhan and A. Kartikeyan. Dynamical investigation and chaotic associated behaviors of memristor chua's circuit with a non-ideal voltage-controlled memristor and its application to voice encryption. *International Journal of Electronics and Communications* **107** (2019) 183–191.
- [9] A. Phabhakar, A. S. Shetty, K. Z. Tabassum, S. Deb and A. Raghumanth. Voice encryption using chaotic signal. *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)* **1** (2023) 1364–1369.
- [10] A. N. Negou, J. Kengne and D. Tchiotso. Periodicity, chaos and multiple coexisting attractors in a generalized moore–spiegel system. *Chaos, Solitons and Fractals* **107** (2018) 275–289.
- [11] N. Djafri, T. Hamaizia and F. Derouiche. Boundedness and dynamics of a modified discrete chaotic system with rational fraction. *Nonlinear Dynamics and Systems Theory* **21** (1) (2021) 68–75.
- [12] A. Sambas, W. S. M. Sanjaya, M. Mamat, N. V. Karadimas and O. Tacha. Numerical simulations in jerk circuit and its application in a secure communication system. *Recent Advances in Telecommunications and Circuit Design* **1** (2013) 190–196.
- [13] F. Yuan, G. Wang, Y. Shen and X. Wang. Coexisting attractors in a memcapacitor-based chaotic oscillator. *Nonlinear Dynamics* **86** (1) (2016) 37–50.
- [14] T. Nestor, A. Belazi, B. Abd-El-atty, M. N. Aslam, C. Volos, N. J. D. Dieu and A. A. El-Latif. A new 4d hyperchaotic system with dynamics analysis, synchronization, and application to image encryption. *Symmetry* **14** (2) (2022) 14–15.
- [15] M. Hanif, R. A. Naqvi, S. Abbas, M. A. Khan and N. Iqbal. A novel and efficient 3d multiple images encryption scheme based on chaotic systems and swapping operations. *IEEE Access* **8** (2020) 123536–123555.
- [16] S. Mortajez, M. Tahmasbi, J. Zarei and A. Jamshidnezhad. A novel chaotic encryption scheme based on efficient secret keys and confusion technique for confidential of DICOM images *Informatix in Medicine Unlocked* **20** (2020) 100396.
- [17] L. Xu, X. Gou, Z. Li, and J. Li. A novel chaotic image encryption algorithm using block scrambling and dynamic index based diffusion. *Optics and Lasers In Engineering* **91** (2017) 41–52.
- [18] A. Elghandour, A. Salah and A. Karawia. A New Cryptographic algorithm via a two-dimensional chaotic map. *Ain Shams Engineering Journal*. **13** (1) (2022) 4–5.
- [19] P. Sathiyamurthi and S. Ramakrishnan. Speech encryption using chaotic shift keying for secured speech communication. *Eurasip Journal on Audio, Speech, and Music Processing* **2017** (1) (2017) 1–11.
- [20] L. Xu, K. Kikushima, S. Sato, A. Islam, T. Sato, S. Aramaki, C. Zhang, T. Sakamoto, F. Eto, Y. Takahashi, I. Yao, M. Machida, T. Kahyo and M. Setou. Spatial distribution of the shannon entropy for mass spectrometry imaging. *PLOS ONE* **18** (4) (2023) 129–177.



Study of a Penalty Method for Nonlinear Optimization Based on a New Approximate Function

C. Souli^{1*} and A. Leulmi²

¹ *Laboratory of Fundamental and Numerical Mathematics (LFNM),
Ferhat Abbas University of Setif-1, Algeria.*

² *Department of Mathematics, Ferhat Abbas University of Setif-1, Algeria.*

Received: April 3, 2023; Revised: October 12, 2023

Abstract: The aim of this paper is to present a logarithmic penalty method for solving nonlinear optimization. The line search is carried out by means of an approximate function if the descent direction is determined using a classical Newton technique. Contrary to the line search method, which is costly in terms of computing volume and demands a lot of time, the proposed approximate function enables easy and quick computation of the displacement step. Numerous intriguing numerical experiments, which are presented in the last section of this work, show that our new approximate function is accurate and efficient.

Keywords: *interior point methods; logarithmic penalty method; applications; approximate functions; nonlinear optimization; quadratic optimization.*

Mathematics Subject Classification (2010): 90C25, 90C30, 90C20, 93C95, 70K75.

1 Introduction

Nonlinear optimization problems deal with the problem of optimizing an objective function in the presence of equality and inequality constraints. Furthermore, if all the functions are linear, we obviously have a linear optimization problem. Otherwise, the problem is called a nonlinear optimization problem.

This research field is motivated by the fact that several problems are collected from practice such as engineering, medicine, business administration, economics, physical sciences, and nonlinear dynamics and systems (see, e.g., [8, 9]).

* Corresponding author: mailto:as_smaleulmi@yahoo.fr

Quadratic optimization is a type of nonlinear optimization, where the objective function is quadratic. In order to solve this type, we propose a penalty approach without line search based on approximate functions, which is the efficient method for determining the displacement step. This study is supported by an important numerical simulation.

For this purpose, we consider the solution of the following quadratic programming problem:

$$(P) \begin{cases} \min q(x) & = \frac{1}{2}x^t Qx + c^t x \\ x \in D, \end{cases}$$

with

$$D = \{x \in \mathbb{R}^n : Ax \geq b\}.$$

The following assumptions are made.

1. $c \in \mathbb{R}^n$, Q is an $\mathbb{R}^{n \times n}$ symmetric semidefinite matrix.
2. We know a point $x_0 \in \mathbb{R}^n$ such that $Ax_0 > b$.
3. $b \in \mathbb{R}^p$, A is a $(p \times n)$ full rank matrix.
4. The set of optimal solutions of (P) is nonempty and bounded.

In this paper, the problem (P) is approximated by the problem (P_η) , $(\eta > 0)$,

$$(P_\eta) \begin{cases} \min q_\eta(x) \\ x \in \mathbb{R}^n, \end{cases}$$

where the barrier function $q_\eta : \mathbb{R}^n \rightarrow (-\infty, +\infty]$ is defined by

$$q_\eta(x) = \begin{cases} q(x) - \eta \sum_{i=1}^m \ln \langle e_i, Ax - b \rangle & \text{if } Ax - b > 0 \\ +\infty & \text{otherwise,} \end{cases}$$

where (e_1, e_2, \dots, e_m) is the canonical base in \mathbb{R}^m and η is a strictly positive barrier parameter. Recall that the scalar product of $x, y \in \mathbb{R}^n$ is given by

$$\langle x, y \rangle = x^t y = \sum_{i=1}^n x_i y_i,$$

the Euclidean norm of y is

$$\|y\| = \sqrt{\langle x, y \rangle} = \sqrt{\sum_{i=1}^n y_i^2}.$$

A classical Newton descent approach is used to solve this problem.

In our new approach, instead of minimizing q_η , along the descent direction at a current point x , we propose an approximate function G for which the optimal solution of the displacement step α is obtained explicitly.

Let us minimize the function G so that

$$G(\alpha) = \frac{1}{\eta} (q_\eta(x_\eta + \alpha d) - q_\eta(x_\eta)) \geq \check{G}(\alpha), \quad \forall \alpha > 0,$$

with $G(0) = \check{G}(0) = 0$, $G'(0) = \check{G}'(0) < 0$. The best quality of the approximations \check{G} of G is ensured by the condition $G''(0) = \check{G}''(0)$. The idea of this new approach

consists in introducing one original process to calculate the displacement step α based on minorant functions. Then we obtain an explicit approximation which leads to reducing the objective, adding to this, it is economical and robust, contrary to the traditional methods of line search.

The paper is organized as follows. In Section 1, we prove the perturbed problem convergence to the initial one. We are interested in resolving the perturbed problem. We describe our algorithm briefly and we present our main result by introducing a new approximate function to compute efficiently the displacement step of the obtained penalty algorithm. This approach is employed to evade line search methods and expedite the algorithm’s convergence.

In Section 2, we present numerical tests on some different examples to illustrate the effectiveness of the proposed approach and we compare it with the standard line search method. A conclusion and future research are given in the last Section 3.

By assumption (1), its solutions set is nonempty and bounded, and as we know, (P) is convex, consequently, in accordance with Bachir Cherif et al. [5], the strictly convex problem (P_η) has unique optimal solution x_η^* for each $\eta > 0$.

Since solving the problem (P) is similar to solving the problem (P_η) when η tends to 0, our goal is to resolve the problem (P_η) .

Firstly, we need to study the convergence of (P_η) to (P) .

Convergence of the Perturbed Problem (P_η) to (P)

Let the function ψ be defined on $\mathbb{R} \times \mathbb{R}^n$ by

$$\psi(\eta, x) = \begin{cases} q(x) + \sum_{i=1}^n \xi(\eta, x_i) & \text{if } x \geq 0, Ax \geq b, \\ +\infty & \text{if not,} \end{cases}$$

where $\xi : \mathbb{R}^2 \rightarrow (-\infty, +\infty]$ is a convex, lower semicontinuous and proper function given by

$$\xi(\eta, t) = \begin{cases} \eta \ln(\eta) - \eta \ln(\alpha) & \text{if } \alpha > 0 \text{ and } \eta > 0, \\ 0 & \text{if } \alpha \geq 0 \text{ and } \eta = 0, \\ +\infty & \text{otherwise.} \end{cases}$$

So, the function ψ is a convex, lower semicontinuous and proper function.

From [5], the strictly convex problem (P_η) admits a unique optimal solution x_η^* for each η . The solution of the problem (P) reduces to the solution of the series of problems (P_η) . The sequence of the solutions x_η of (P_η) should converge to the solution of (P) when η tends to 0.

Now we are in a position to state the convergence result of (P_η) to (P) which is proved in Lemma 1 from [4].

Let $\eta > 0$, for all $x \in D$, we define

$$\psi(x, \eta) = q_\eta(x).$$

Lemma 1.1 [4] *Let $\eta > 0$. If x_η is an optimal solution of the problem (P_η) such that $\lim_{\eta \rightarrow 0} x_\eta = x^*$, then x^* is an optimal solution of the problem (P) .*

Let $\eta > 0$, for all $x \in D$, we define $\psi(x, \eta) = q_\eta(x)$.

Resolution of the Perturbed Problem

In this section, we are interested in finding the solution of the perturbed problem $x_{k+1} = x_k + \alpha_k d_k$. For this purpose, we first use the Newton method to calculate the descent direction d_k . Then we obtain the displacement step α_k by our new minorant function. Finally, we describe a standard prototype algorithm.

1.1 Newton descent direction

The interior point methods of the logarithmic barrier type are developed for resolving this type of problems based on the optimality conditions that are necessary and sufficient since the problem (P_η) can be considered as the one without constraints.

As a result, $((x_\eta)_k = x_k)$ is an optimal solution of (P_η) such that the following condition is met:

$$\nabla q_\eta(x_\eta) = 0. \quad (1)$$

Thus, $x_{k+1} = x_k + d_k$ is the iteration of Newton, where d_k is the descent direction solution of the linear system

$$\nabla^2 q_\eta(x_\eta) d_k = -\nabla q_\eta(x_\eta). \quad (2)$$

Note 1. We insert a displacement step α_k and we write

$$x_{k+1} = x_k + \alpha_k d_k$$

to ensure the strictly feasible iterate $x_{k+1} = x_k + d_k$.

1.2 Model algorithm

In this part, we present a brief algorithm of our approach to obtain an optimal solution \bar{x} of the problem (P) .

Begin algorithm

Initialization

Start with x_0 being a strictly feasible solution of (P) , $\eta > 0$, ε is a given precision and $k = 0$.

While $\|\nabla q_\eta(x_k)\| > \varepsilon$ do

- Resolve the system : $\nabla^2 q_\eta(x_k) d_k = -\nabla q_\eta(x_k)$.
- Compute the displacement step α_k .
- Take $x_{k+1} = x_k + \alpha_k d_k$ and $k = k + 1$.
- Put $\eta = \sigma \eta$, $0 < \sigma < 1$.

End While

We have obtained a good approximate solution of the problem (P) .

End algorithm.

1.3 Computation of the displacement step

There are two main techniques used for computing the displacement step α_k .

- (1) **Line search methods:** The method of Goldstein-Armijo, Fibonacci, Wolfe, etc. They are based on the unidimensional function’s minimization:

$$\phi(\alpha) = \min_{\alpha > 0} q_\eta(x_\eta + \alpha d).$$

They are time-consuming and unfortunately very sensitive.

- (2) **Minorant function:** The technique of the minorant function was first proposed by Leulmi [10] for the positive semidefinite programming. This technique relies on approximating the function

$$G(\alpha) = \frac{1}{\eta}(q_\eta(x_\eta + \alpha d) - q_\eta(x_\eta))$$

by another function whose minimum can be easily computed, which permits the computation of the displacement step at each iteration in a relatively short time and with a smaller number of instructions in contrast to the line search technique.

We start with the following lemma, and in the rest of the paper, we consider x instead of x_η .

Lemma 1.2 [14] *The function G can be written as follows:*

$$G(\alpha) = \frac{1}{\eta}(\frac{1}{2}\alpha^2 d^t Q d - \alpha d^t Q d) + \alpha(\sum_{i=1}^m y_i - \|y\|^2) - \sum_{i=1}^m \ln(1 + \alpha y_i) \tag{3}$$

for all $\alpha \in [0, \hat{\alpha}]$ such that $\hat{\alpha} = \min_{i \in I_-} \{-\frac{1}{y_i}\}$ and $I = \{i : y_i < 0\}$, where

$$y_i = \frac{\langle e_i, A d \rangle}{\langle e_i, A x - b \rangle}, i \in \{1, \dots, m\}.$$

Now, we give the main result of the paper.

1.4 New approximate function

To introduce our new majorant function, we use the following well known inequality:

$$\left(\|y\| - \sum_{i=1}^n y_i \right) \alpha - \ln(1 + \alpha \|y\|) + \sum_{i=1}^n \ln(1 + \alpha y_i) \leq 0. \tag{4}$$

Replacing by the precedent inequality in (3), we obtain $\check{G}(\alpha) \leq G(\alpha)$, then

$$\check{G}(\alpha) = \delta \alpha - \ln(1 + \beta \alpha) + \frac{1}{2\eta} \hat{\alpha}^2 d^t Q d, \alpha \in [0, \hat{\alpha}],$$

with $\delta = -\|y\|(\|y\| - 1)$ and $\beta_2 = \|y\|$.

Lemma 1.3 *For $\alpha \in I_\alpha = [0, \hat{\alpha}]$, we have*

$$\check{G}(\alpha) \leq G(\alpha).$$

Proof. From inequality (4) and for $\alpha \in I_\alpha$, we have

$$\alpha \sum_{i=1}^m y_i - \sum_{i=1}^m \ln(1 + \alpha y_i) - \alpha \|y\|^2 \geq -\alpha \|y\| - \alpha \|y\|^2 - \ln(1 + \alpha \|y\|).$$

We have $(-\alpha d^t Qd) > 0$ and

$$\frac{1}{2\eta} \alpha^2 d^t Qd < \frac{1}{2\eta} \hat{\alpha}^2 d^t Qd, \quad \forall \alpha \in I_\alpha.$$

This produces

$$\begin{aligned} G(\alpha) &= \alpha \sum_{i=1}^n y_i - \alpha \|y\|^2 - \sum_{i=1}^n \ln(1 + \alpha y_i) + \frac{1}{\eta} \left(\frac{1}{2} \alpha^2 d^t Qd - \alpha d^t Qd \right) \\ &\geq -\alpha (\|y\|^2 - \|y\|) - \ln(1 + \alpha \|y\|) + \frac{1}{2\eta} \hat{\alpha}^2 d^t Qd = \check{G}(\alpha). \end{aligned}$$

Then

$$\forall \alpha \in I_\alpha : G(\alpha) \geq \check{G}(\alpha).$$

Remark 1.1 We note that

$$\check{G}''(\alpha) = \frac{\|y\|^2}{(1 - \|y\|^2)^2} \geq 0, \quad \forall \alpha \in [0, \hat{\alpha}[,$$

hence \check{G} is convex, and if it admits a minimum, this minimum is global.

Minimization of the minorant function \check{G} is defined and convex on $[0, \hat{\alpha}[$, then its global minimum is reached when $\check{G}'(\alpha) = 0$, therefore finding the minimum of the function \check{G} is equivalent to solving the equation $\check{G}'(\alpha) = 0$. The solution of the later is the root of the equation

$$\alpha (\|y\|^2 - \|y\|^3) - \|y\|^2 = 0. \quad (5)$$

The root of the equation (5) is

$$\alpha^* = -(\|y\| - 1)^{-1} \in I_\alpha,$$

which is the global minimum of the function \check{G} .

The following lemma indicates that the interior point x_{k+1} generated in each iteration k of the algorithm ensures the decreases of the function q_η .

Lemma 1.4 *The function q_η significantly decrease from the iteration k to the iteration $k + 1$, that is, if x_k and x_{k+1} are two feasible solutions obtained at the iteration k and $k + 1$, respectively, then*

$$q_\eta(x_{k+1}) < q_\eta(x_k). \quad (6)$$

Proof. Let x_k and x_{k+1} be two feasible solutions obtained at the iteration k and $k + 1$, respectively, we have

$$q_\eta(x_{k+1}) \simeq q_\eta(x_k) + \langle \nabla q_\eta(x_k), x_{k+1} - x_k \rangle$$

and

$$x_{k+1} = x_k + \alpha_k d_k, \tag{7}$$

then

$$\begin{aligned} q_\eta(x_{k+1}) - q_\eta(x_k) &\simeq \langle \nabla q_\eta(x_k), \alpha_k d_k \rangle \\ &\simeq -\alpha_k \langle \nabla^2 q_\eta(x_k) d_k, d_k \rangle < 0. \end{aligned}$$

Hence,

$$q_\eta(x_{k+1}) < q_\eta(x_k),$$

which implies the claimed result.

2 Numerical Tests

We evaluate our algorithm’s efficiency based on our approximate function. We conducted comparative numerical tests between our new two approximate functions (minorant function) and Armijo-Goldstein’s line search method.

For this, in this part, we present a comparative numerical tests on different examples taken from the literature [1–3].

In the below tables, we reported the results obtained by implementing the algorithm in MATLAB R2013a on I5, 8350 (3.6 GHz) with 8 Go RAM.

We have taken $\varepsilon = 1.0e - 005$.

We use the following designations:

- (itrat) represents the number of iterations necessary to obtain an optimal solution.
- (time) represents the time of computation in seconds (s).
- (stmin) represents the strategy of approximate functions introduced in this paper.
- (LS) represents the classical Armijo-Goldstein line search.

We consider the following quadratic problem:

$$\alpha = \min[q(x) : x \geq 0, Ax \geq b],$$

where $q(x) = \frac{1}{2}x^t Qx + c^t x$.

2.1 Examples

Example 01: The matrix Q is defined by

$$Q[i, j] = \begin{cases} 2j - 1 & \text{if } i > j, \\ 2i - 1 & \text{if } i < j, \\ i(i + 1) - 1 & \text{if } i = j, i, j = 1, \dots, n, \end{cases}$$

$$\begin{aligned} A[i, j] &= \begin{cases} 1 & \text{if } i = j \text{ or } j = i + m, i = 1, \dots, m \text{ and } j = 1, \dots, n \\ 0 & \text{otherwise.} \end{cases} \\ c[i] &= -1, c[i + m] = 0 \text{ and } b[i] = 2, \forall i = 1, \dots, m, \end{aligned}$$

with $n = 2m$. We test this example for different values of n . The following table resumes the obtained results:

ex(m, n)	stmin		LS	time
	itrat	time		
200 × 400	10	5.99012	26	22.52401
300 × 600	15	50.03129	35	97.10345
600 × 1200	25	71.66481	48	224.32120
1000 × 2000	30	122.27613	51	497.01165
1500 × 3000	39	320.79313	78	1321.03278

Example 02: We defined the matrix Q by

$$Q[i, j] = \begin{cases} \frac{1}{i+j} & \text{for } i, j = 1, \dots, n, \\ 1 & \text{if } i = j \text{ or } j = i + m, \ i = 1, \dots, m \text{ and } j = 1, \dots, n, \\ 0 & \text{otherwise,} \end{cases}$$

$$c[j] = 2j \text{ and } b[i] = i^2, \ \forall i = 1, \dots, m,$$

with $n = 2m$. We test this example for different values of n .

The following table resumes the obtained results:

ex(m, n)	stmin		LS	time
	itrat	time		
200 × 400	9	9.01214	16	29.12331
300 × 600	11	25.03119	27	84.15001
600 × 1200	33	74.06481	55	153.92210
1000 × 2000	39	198.07613	68	2213.11431
1500 × 3000	58	401.09313	124	3121.11303

Example 03: Let us define the matrix Q by

$$\begin{cases} Q[1, 1] = 1, \\ Q[i, i] = i^2 + 1, \\ Q[i, i - 1] = Q[i - 1, i] = i, \ i = 2, \dots, n, \end{cases}$$

$$A[i, j] = \begin{cases} 1 & \text{if } i = j \text{ or } j = i + m, \ i = 1, \dots, m \text{ and } j = 1, \dots, n \\ 0 & \text{otherwise,} \end{cases}$$

$$c[j] = j \text{ and } b[i] = \frac{i+1}{2}, \ \forall i = 1, \dots, m,$$

with $n = 2m$. We test this example for different values of n .

The following table resumes the obtained results:

ex(m, n)	stmin		LS	time
	itrat	time		
200 × 400	23	10.22544	38	22.52401
300 × 600	32	41.02385	45	88.11235
600 × 1200	39	91.10519	66	148.62103
1000 × 2000	50	148.47512	70	2004.11257
1500 × 3000	75	322.10134	101	2453.92312

Example 04: Let Q be the matrix define by

$$\begin{cases} Q[1, 1] = 1, \\ Q[i, i] = 4, \ i = 2, \dots, n - 1, \\ Q[i, i - 1] = Q[i - 1, i] = 1, \ i = 2, \dots, n, \end{cases}$$

$$A[i, j] = \begin{cases} 1 & \text{if } i = j \text{ or } j = i + m, \ i = 1, \dots, m \text{ and } j = 1, \dots, n, \\ 0 & \text{if } i \neq j \text{ or } (i + 1) \neq j, \end{cases}$$

$$c[j] = \frac{i+1}{2} \text{ and } b[i] = 4, \ \forall i = 1, \dots, m \text{ and } j = 1, \dots, n,$$

with $n = 2m$. We test this example for different values of n .

The following table resumes the obtained results:

ex(m, n)	stmin		LS	
	itrat	time	itrat	time
200 × 400	12	6.22544	21	34.32215
300 × 600	25	81.02385	31	172.32550
600 × 1200	36	122.10519	53	503.44316
1000 × 2000	39	148.47512	67	2033.50062
1500 × 3000	55	352.10134	132	3121.11303

Commentary. These experiments demonstrate clearly the impact of our approach on the numerical behavior of the algorithm, expressed by the reduction of the number of iterations and computation time. The number of iterations and the computing time are considerably reduced in the approximate approaches in comparison with the line search method. Always, in the problems of linear dynamics, we arrive to the problem of optimization. Then we solve this problem by our approach. This is what we look forward to in future.

3 Conclusion

In order to solve a quadratic optimization problem, this study provides a logarithmic penalty method based on new approximate functions (minorant). As anticipated, the minorant function strategy for computing the displacement step demonstrates its effectiveness by lowering the computational cost when compared to the line search method. This effectiveness is a result of the nature of the mentioned functions. The numerical results demonstrate that our strategy reduces the cost of iteration for the quadratic optimization compared to the line search method. Proposing some other new majorant [5] and minorant functions seems to be an interesting topic in the future in the different class of optimization and we will apply our important results in different problems of nonlinear dynamics problems.

Acknowledgment

This work is funded by the General Directorate of Scientific Research and Technological Development (DGRSDT) of Algeria.

References

- [1] M. Achache. A new primal-dual path-following method for convex quadratic programming. *Comput. Appl. Math.* **25** (2006) 97–110.
- [2] M. S. Bazraa, H. D. Sherali and C. M. Shetty. *Nonlinear Programming*. Published by John Wiley & Sons, Inc., Hoboken, New Jersey, 2006.
- [3] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal and C. Sagastizàbal. *Numerical Optimization: Theoretical and Practical Aspects*. Mathematics and Applications, Vol. 27. Springer-Verlag, Berlin, 2003.

- [4] S. Chaghoub and D. Benterki. A Logarithmic Barrier Method Based On a New Majorant Function for Convex Quadratic Programming. *IAENG International Journal of Applied Mathematics* **51** (3) (2021).
- [5] L. B. Cherif and B. Merikhi. A Penalty Method for Nonlinear Programming. *RAIRO-Operations Res.* **53** (2019) 29–38.
- [6] J. P. Crouzeix and B. Merikhi. A logarithmic barrier method for semi-definite programming. *R.A.I.R. O-Oper. Res.* **42** (2008) 123–139.
- [7] J. P. Crouzeix and A. Seeger. New bounds for the extreme values of a finite sample of real numbers. *J. Math. Anal. Appl.* **197** (1996) 411–426.
- [8] A. Djaout, T. Hamaizia and F. Derouiche. Performance Comparison of Some Two-Dimensional Chaotic Maps for Global Optimization. *Nonlinear Dynamics and Systems Theory* **22** (2) (2022) 144–154.
- [9] T. Herlambang, A.Y.P. Asih, D. Rahmalia, D. Adzkiya and N. Aini. The Effects of Pesticide as Optimal Control of Agriculture Pest Growth Dynamical Model. *Nonlinear Dynamics and Systems Theory* **22** (3) (2022) 281–290.
- [10] Y. Nesterov and A. Nemirovsky. Interior-point polynomial methods in convex programming. *Studies in Applied Mathematics*, **13**, SIAM, Philadelphia, 1994.
- [11] A. Leulmi. Etude d'une méthode barrière logarithmique via minorants fonctions pour la programmation semi-définie. Thèse de doctorat, Université de Biskra (2018).
- [12] A. Leulmi, B. Merikhi and D. Benterki. Study of a Logarithmic Barrier Approach for Linear Semidefinite Programming. *Journal of Siberian Federal University. Mathematics & Physics.* **11** (3) (2018) 1–13.
- [13] A. Leulmi and S. Leulmi. Logarithmic Barrier Method via Minorant Function for Linear Programming. *J. Sib. Fed. Univ. Math. Phys.* **12** (2) (2019) 191–201.
- [14] N. Moussaoui and M. Achache. A Weighted-path Following Interior-point Algorithm for Convex Quadratic Optimization Based on Modified Search Directions. *Statistics, Optimization & Information Computing.* **10** (3) (2022) 873–889.
- [15] M. Ouriemchi. Résolution de problèmes non linéaires par les méthodes de points intérieurs. Théorie et algorithmes. Thèse de doctorat, Université du Havre, France (2006).
- [16] E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.* **27** (1948) 379–423 and 623–656.
- [17] H. Wolkowicz and G.P.H. Styan. Bounds for Eigenvalues Using Traces. *Lin. Alg. Appl.* **29** (1980) 471–506.

<i>M. Y. Anshori, I. H. Santoso, T. Herlambang, D. Rahmalia, K. Oktafianto and P. Katias</i>		
Contact Problem for Thermo-Elasto-Viscoplastic Material with Friction		141
<i>N. Bensebaa and N. Lebri</i>		
Implementation of Infeasible Interior-Point Methods Based on a New Search Direction		157
<i>L. Derbal</i>		
A Frictional Contact Problem with Wear for Two Electro-Viscoelastic Bodies		167
<i>A. Djabi</i>		
Global Optimization Method of Multivariate non-Lipschitz Functions Using Tangent Minorants		183
<i>Djaouida Guettal, Chahinez Chenouf and Mohamed Rahal</i>		
Analysis of Solutions to Equations with a Generalized Derivative and Delay		195
<i>O. D. Kichmarenko, I. V. Chepovskyi, Y. Platonova and S. Dashkovskiy</i>		
Chaos Anti-Synchronization between Fractional-Order Lesser Date Moth Chaotic System and Integer-Order Chaotic System by Nonlinear Control		207
<i>M. Labid and N. Hamri</i>		
A New Hidden Attractor Hyperchaotic System and Its Circuit Implementation, Adaptive Synchronization and FPGA Implementation		214
<i>R. Rameshbabu, K. Kavitha, P. S. Gomathi and K. Kalaiichelvi</i>		
Nonlinear Damped Oscillator with Varying Coefficients and Periodic External Forces		227
<i>M. Alhaz Uddin, Mahmuda Akhter Nishu and M. Wali Ullah</i>		
 Volume 23	 Number 3	 2023
Weights Optimization Using Firefly Algorithm for Dengue Fever Optimal Control Model by Vaccination, Treatment, and Abateseae		237
<i>A. Y. Asih, B. Gunawan, N. Hidayati, T. Herlambang, D. Rahmalia and K. Oktafianto</i>		
Inverse Problem of a Semilinear Parabolic Equation with an Integral Overdetermination Condition		249
<i>Amal Benguesmia, Iqbal M. Batiha, Taki-Eddine Oussaeif, Adel Ouannas and Waseem G. Alshanti</i>		
Some Generalized Nonlinear Volterra-Fredholm Type Integral Inequalities with Delay of Several Variables and Applications		261
<i>Ammar Boudeliou</i>		
An Adaptive Step Size for Chaotic Local Search Algorithm		273
<i>Aziza Filali and Tayeb Hamaizia</i>		

On the Dynamics and FSHP Synchronization of a New Chaotic 3-D System with Three Nonlinearities	283
<i>Fareh Hannachi and Rami Amira</i>	
Numerical Solution of the Black-Scholes Partial Differential Equation for the Option Pricing Model Using the ADM-Kamal Method	295
<i>M. D. Johansyah, I. Sumiati, E. Rusyaman, Sukono, M. Muslikh, M. A. Mohamed and A. Sambas</i>	
A Novel Adaptive Method Based on New Minorant-Majorant Functions without Line Search for Semidefinite Optimization	310
<i>S. Leulmi and A. Leulmi</i>	
Electronic Nose for Classifying Civet Coffee and Non-Civet Coffee	323
<i>D. B. Magfira, R. Sarno, F. A. Susanto, T. Herlambang, K. Oktafianto, W. Hartawan and I. W. Farid</i>	
Global Existence for the 3-D Generalized Micropolar Fluid System in Critical Fourier-Besov Spaces with Variable Exponent	338
<i>F. Ouidirne, H. Srhiri, C. Allalou and M. Oukessou</i>	
The Regularization Method for Solving Sub-Riemannian Geodesic Problems	348
<i>R. Saffidine and N. Bensalem</i>	

Volume 23

Number 4

2023

A Novel Fractional-Order Chaotic System and Its Synchronization via Adaptive Control Method	359
<i>Rami Amira and Fareh Hannachi</i>	
Analysis of Problems in Generalized Viscoplasticity under Dynamic Thermal Loading	367
<i>I. Boukaroura and A. Djabi</i>	
Maximum Power Point Tracking Based on Remora Algorithm under PSC	381
<i>Hamidia Fethia, Abbadı Amel, Skender Mohamed Reda, Morsli Abdelkader, Bettache Farouk, Medjber Ahmed and Tlemçani Abdelhalim</i>	
Motion Estimation of Third Finger Using Ensemble and Unscented Kalman Filter for Inverse Kinematic of Assistive Finger-Arm Robot	389
<i>T. Herlambang, H. Nurhadi, A. Suryowinoto, D. Rahmalia and K. Oktafianto</i>	
Stability and Hopf Bifurcation of a Generalized Differential-Algebraic Biological Economic System with the Hybrid Functional Response and Predator Harvesting .	398
<i>M. C. Benkara Mostefa and N. E. Hamri</i>	
Synchronization of the Restricted Charged Three-Body Problem	410
<i>Krishan Pal</i>	
Modelling and MultiSim Simulation of a New Hyperchaos System with No Equilibrium Point	422

*A. Sambas, S. Vaidyanathan, S. F. Al-Azzawi, M. K. M. Nawawi,
M. A. Mohamed, Z. A. Zakaria, S. S. Abas and M. Mamat*

Properties of Solutions and Stability of a Diffusive Wage-Employment System 434
S. Sutrima and R. Setiyowati

Multiple Well-Posedness of Higher-Order Abstract Cauchy Problem 447
Habiba Toumi and Ahmed Nouar

Volume 23

Number 5

2023

Solvability of Nonlinear Elliptic Problems with Degenerate Coercivity in Weighted Sobolev Space 461
Y. Akdim, S. Lalaoui Rhali and Y. Oumouacha

An Algorithm for Solving First-Kind Two-Dimensional Volterra Integral Equations Using Collocation Method 475
F. Birem, A. Boulmerka, H. Laib and C. Hennous

Application of Model Predictive Control (MPC) to Longitudinal Motion of the Aircraft Using Polynomial Chaos 487
K. D. R. Dewi, K. Fahim and S. Subchan

Singular Reaction-Diffusion System Arising from Quenching 499
Samiha Djemai and Salim Mesbahi

Maximum and Anti-Maximum Principles for Boundary Value Problems for Ordinary Differential Equations in Neighborhoods of Simple Eigenvalues 507
P. W. Eloe and J. T. Neugebauer

Performance Comparison of Sliding Mode Control and Sliding PID for Rescue ROV 519
*T. Herlambang, A. Suryowinoto, I. Kurniastuti, H. Nurhadi
and K. Oktafianto*

Class of Nilpotent Distributions and \mathfrak{N}_2 -Distributions 538
Y. Khellaf, A. Chikh-Salah and N. Bensalem

Moore-Spiegel Chaotic Encryption for Digital Images and Voices 545
*W. S. Mada Sanjaya, Akhmad Roziqin, Agung Wijaya Temiesela,
M. Fauzi Badru Zaman, Aria Dewa Wibiksana and Dyah Anggraeni*

Study of a Penalty Method for Nonlinear Optimization Based on a New Approximate Function 561
C. Souli and A. Leulmi

Contents of Volume 23, 2023 571