VOLUME 11, NUMBER 1 JANUARY 2013

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL
OF CONCRETE
AND APPLICABLE MATHEMATICS

EUDOXUS PRESS,LLC

GUEST EDITORS: O. DUMAN, E. ERKUS-DUMAN
SPECIAL ISSUE I: "APPLIED MATHEMATICS -APPROXIMATION THEORY 2012"

SCOPE AND PRICES OF THE JOURNAL

Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Assistant to the Editor:Dr.Razvan Mezei,Lander University,SC 29649, USA.

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send via email the contribution to the editor in-Chief: TEX or LATEX (typed double spaced) and PDF files. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $500,Electronic $250,Print and Electronic $600.Individual:Print $200, Electronic $100,Print &Electronic $250.For any other part of the world add $60 more to the above prices for Print.
Single article PDF file for individual $20.Single issue in PDF form for individual $80.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

PAGE CHARGES: Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus homepage.

Editorial Board
Associate Editors of Journal of Concrete and Applicable Mathematics

Department of Mathematics Education
College of Education
Gyeongsang National University
Chinju 660-701
KOREA
tel.055-751-5673 Office,
055-755-3644 home,
fax 055-751-6117
yjcho@nongae.gsnu.ac.kr
Nonlinear operator Th.,Inequalities,
Geometry of Banach Spaces

5) Sever S.Dragomir
School of Communications and Informatics
Victoria University of Technology
PO Box 14428
Melbourne City M.C
Victoria 8001,Australia
tel 61 3 9688 4437,fax 61 3 9688 4050
sever.dragomir@vu.edu.au,
sever@sci.vu.edu.au
Math.Analysis,Inequalities,Approximation Th.,
Numerical Analysis, Geometry of Banach Spaces,
Information Th. and Coding

6) Oktay Duman
TOBB University of Economics and Technology,
Department of Mathematics, TR-06530,
Ankara, Turkey, oduman@etu.edu.tr
Classical Approximation Theory,
Summability Theory,
Statistical Convergence and its Applications

7) Angelo Favini
Università di Bologna
Dipartimento di Matematica
Piazza di Porta San Donato 5
40126 Bologna, ITALY
tel.++39 051 2094451
fax.++39 051 2094490
favini@dm.unibo.it
Partial Differential Equations, Control Theory,
Differential Equations in Banach Spaces

8) Claudio A. Fernandez
Facultad de Matematicas
Pontificia Unversidad Católica de Chile
Vicuna Mackenna 4860
Santiago, Chile
tel.++56 2 354 5922

25) Svetlozar T.Rachev
Dept.of Statistics and Applied Probability Program
University of California,Santa Barbara
CA 93106-3110,USA
tel.805-893-4869
rachev@pstat.ucsb.edu
AND
Chair of Econometrics and Statistics
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss,Bau II,20.12,R210
Postfach 6980,D-76128,Karlsruhe,Germany
tel.011-49-721-608-7535
rachev@lsoe.uni-karlsruhe.de
Mathematical and Empirical Finance,
Applied Probability, Statistics and Econometrics

26) John Michael Rassias
University of Athens
Pedagogical Department
Section of Mathematics and Infomatics
20, Hippocratous Str., Athens, 106 80, Greece

Address for Correspondence
4, Agamemnonos Str.
Aghia Paraskevi, Athens, Attikis 15342 Greece
jrassias@primedu.uoa.gr
jrassias@tellas.gr
Approximation Theory,Functional Equations,
Inequalities, PDE

27) Paolo Emilio Ricci
Universita' degli Studi di Roma "La Sapienza"
Dipartimento di Matematica-Istituto "G.Castelnuovo"
P.le A.Moro,2-00185 Roma,ITALY
tel.++39 0649913201,fax ++39 0644701007
riccip@uniroma1.it,Paoloemilio.Ricci@uniroma1.it
Orthogonal Polynomials and Special functions,
Numerical Analysis, Transforms,Operational Calculus,
Differential and Difference equations

28) Cecil C.Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152,USA
tel.901-678-2490,fax 901-678-2480
ccrousse@memphis.edu
Combinatorics,Graph Th.,
Asymptotic Approximations,
Applications to Physics

29) Tomasz Rychlik

fax.++56 2 552 5916
cfernand@mat.puc.cl
Partial Differential Equations,
Mathematical Physics,
Scattering and Spectral Theory

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA
tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex Analysis,Holomorphic PDE,
Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences

Institute of Mathematics
Polish Academy of Sciences
Chopina 12,87100 Torun, Poland
T.Rychlik@impan.gov.pl
Mathematical Statistics,Probabilistic Inequalities

30) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

31) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

32) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus and Appl.,
Integral Equations and Transforms,Higher Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

33) Stevo Stevic
Mathematical Institute of the Serbian Acad. of Science
Knez Mihailova 35/I
11000 Beograd, Serbia
sstevic@ptt.yu; sstevo@matf.bg.ac.yu
Complex Variables, Difference Equations,
Approximation Th., Inequalities

34) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,

Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop
Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Robert Kozma
Dept. of Mathematical Sciences
University of Memphis
Memphis, TN 38152, USA
rkozma@memphis.edu
Mathematical Learning Theory,
Dynamic Systems and Chaos,
Complex Dynamics.

18) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

19) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

20) Rupert Lasser
Institut fur Biomathematik &
Biomertie,GSF
-National Research Center for
environment and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier
Analysis,Mathematical Biology.

Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

35) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

36) Manfred Tasche
Department of Mathematics
University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

37) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

38) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou

Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts files in Latex and PDF and in English, should be submitted via email to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel. 901.678.3144
e-mail: ganastss@memphis.edu

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.

Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.

2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (Click HERE to save a copy of the style file.)They should be carefully prepared in all respects. Submitted articles should be brightly typed (not dot-matrix), double spaced, in ten point type size and in 8(1/2)x11 inch area per page. Manuscripts should have generous margins on all sides and should not exceed 24 pages.

3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.

4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.

The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.

6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes.

The main body of the paper should begin on page no. 1, if possible.

7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.

8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right, and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.

9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.

Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.

10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,

name of publication, volume number, inclusive pages, and year of publication.

Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.

12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set off from the text. Use the appropriate superscript numeral for citation in the text.

13. After each revision is made please again submit via email Latex and PDF files of the revised manuscript, including the final one.

14. Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus <u>homepage</u>.

No galleys will be sent and the contact author will receive one (1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

PREFACE (JAFA – JCAAM)

These special issues are devoted to a part of proceedings of AMAT 2012 - International Conference on Applied Mathematics and Approximation Theory - which was held during May 17-20, 2012 in Ankara, Turkey, at TOBB University of Economics and Technology. This conference is dedicated to the distinguished mathematician George A. Anastassiou for his 60th birthday.

AMAT 2012 conference brought together researchers from all areas of Applied Mathematics and Approximation Theory, such as ODEs, PDEs, Difference Equations, Applied Analysis, Computational Analysis, Signal Theory, and included traditional subfields of Approximation Theory as well as under focused areas such as Positive Operators, Statistical Approximation, and Fuzzy Approximation. Other topics were also included in this conference, such as Fractional Analysis, Semigroups, Inequalities, Special Functions, and Summability. Previous conferences which had a similar approach to such diverse inclusiveness were held at the University of Memphis (1991, 1997, 2008), UC Santa Barbara (1993), the University of Central Florida at Orlando (2002).

Around 200 scientists coming from 30 different countries participated in the conference. There were 110 presentations with 3 parallel sessions. We are particularly indebted to our plenary speakers: George A. Anastassiou (*University of Memphis - USA*), Dumitru Baleanu (*Çankaya University - Turkey*), Martin Bohner (*Missouri University of Science & Technology - USA*), Jerry L. Bona (*University of Illinois at Chicago - USA*), Weimin Han (*University of Iowa - USA*), Margareta Heilmann (*University of Wuppertal - Germany*), Cihan Orhan (*Ankara University - Turkey*). It is our great pleasure to thank all the organizations that contributed to the conference, the Scientific Committee and any people who made this conference a big success.

Finally, we are grateful to “TOBB University of Economics and Technology”, which was hosting this conference and provided all of its facilities, and also to “Central Bank of Turkey” and “The Scientific and Technological Research Council of Turkey” for financial support.

Guest Editors:

Oktay Duman	Esra Erkuş-Duman
TOBB Univ. of Economics and Technology	Gazi University
Ankara, Turkey, 2012	Ankara, Turkey, 2012

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 12-16, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON UNIVALENCE OF A GENERAL INTEGRAL OPERATOR

AISHA AHMED AMER AND MASLINA DARUS

ABSTRACT. **Problem statement:** We introduce and study a general integral operator defined on the class of normalized analytic functions in the open unit disk. This operator is motivated by many researchers. With this operator univalence conditions for the normalized analytic function in the open unit disk are obtained. In deed, the preserving properties of this class are studied, when the integral operator is applied and we present a few conditions of univalency for our integral operator. The operator is essential to obtain univalence of a certain general integral operator. **Approach:** In this paper we discuss some extensions of univalent conditions for an integral operator defined by our generalized differential operator. Several other results are also considered. We will prove in this paper the univalent conditions for this integral operator on the class of normalized analytic functions when we make some restrictions about the functions from definitions. **Results:** Having the integral operator, some interesting properties of this class of functions will be obtained. Relevant connections of the results,shall be presented in the paper. In fact, various other known results are also pointed out. We also find some interesting corollaries on the class of normalized analytic functions in the open unit disk. **Conclusion:** Therefore, many interesting results could be obtained and we also derive some interesting properties of these classes. We conclude this study with some suggestions for future research,one direction is to study other classes of analytic functions involving our integral operator on the class of normalized analytic functions in the open unit disk.

1. Introduction

As usual, let $U = \{z \in C : |z| < 1\}$ be the unit disc in the complex plane and let A be the class of functions which are analytic in the unit disk normalized with $f(0) = f'(0) - 1 = 0$. Let S the class of the functions $f \in A$ which are univalent in U. In particular,for $f \in A$ and $(z \in U, b \neq 0, -1, -2, -3, ...), \lambda \geq 0, m \in \mathbb{Z}, l \geq 0$, the authors (cf., [1, 2]) introduced the following linear operator:

Definition 1.1. *For $f \in A$ the operator $D_l^{m,\lambda}(a,b)f(z)$ is defined by $D_l^{m,\lambda}(a,b)f(z) : A \to A$ and let*

$$\phi(z) := \frac{1+l-\lambda}{1+l}\frac{z}{1-z} + \frac{\lambda}{1+l}\frac{z}{(1-z)^2},$$

and

$$D_l^{m,\lambda}(a,b)f(z) = \underbrace{\phi(z) * ... * \phi(z)}_{(m)-times} * zF(a,1;b;z) * f(z),$$

if $\quad (m = 0, 1, 2, ...)$, *and*

Key words and phrases. Analytic functions; Univalent functions; Derivative operator; Hadamard product; Unit disk; The complex plane.

2010 *AMS Math. Subject Classification.* Primary 40A05, 40A25; Secondary 45G05.

$$D_l^{m,\lambda}(a,b)f(z) = \underbrace{\phi(z) * ... * \phi(z)}_{(-m)-times} * zF(a,1;b;z) * f(z),$$

if $(m = -1, -2, ...)$, *thus we have*

$$D_l^{m,\lambda}(a,b)f(z) := z + \sum_{k=2}^{\infty} \left(\frac{1+\lambda(k-1)+l}{1+l} \right)^m \frac{(a)_{k-1}}{(b)_{k-1}} a_k z^k,$$

where $f \in A$ *and* $(z \in \mathbb{U}, b \neq 0, -1, -2, -3, ...), \lambda \geq 0, m \in \mathbb{Z}, l \geq 0$. *Special cases of this operator includes:*

- $D_0^{m,0}(a,b)f(z) = D_l^{0,\lambda}(a,b)f(z) = L(a,b)f(z)$, *see* [11].
- *The Ruscheweyh derivative operator* [12] *in the cases:* $D_0^{0,0}(\beta+1,1)f(z) = D^{\beta} f(z); \beta \geq -1$.
- *The Salagean derivative operator* [13]*:* $D_0^{m,1}(1,1)f(z)$.
- *The generalized Salagean derivative operator introduced by Al-Oboudi* [14]*:* $D_0^{m,\lambda}(1,1)f(z)$.
- *The Catas drivative operator* [10]*:* $D_l^{m,\lambda}(1,1)f(z)$.

2. Preliminary Results

To discuss the univalency of $f \in A$, we have

Definition 2.1.

Theorem 2.2. [5] *Assume that* $f \in A$ *satisfies condition*

$$\left| \frac{z^2 f'(z)}{f^2(z)} - 1 \right| < 1, \; z \in U, \tag{2.1}$$

then f *is univalent in* U.

Theorem 2.3. [6] *Let* α *be a complex number,* $\Re\alpha > 0$ *and* $f(z) = z + a_2 z^2 + \ldots$ *is a regular function in U. If*

$$\frac{1 - |z|^{2\Re\alpha}}{\Re\alpha} \left| \frac{zf''(z)}{f'(z)} \right| \leq 1, \tag{2.2}$$

for all $z \in U$*, then for any complex number* β*,* $\Re\beta \geq \Re\alpha$ *the function*

$$F_{\beta}(z) = \left[\beta \int_0^z u^{\beta-1} f'(u) du \right]^{\frac{1}{\beta}} = z + \ldots, \tag{2.3}$$

is regular and univalent in U.

Lemma 2.4. (***Schwarz Lemma*** [3]) *Let* $f(z)$ *the function regular in the disk* $U_R = \{z \in C; |z| < R\}$*, with* $|f(z)| < M$*,* M *fixed. If* $f(z)$ *has in* $z = 0$ *one zero with multiply* $\geq m$*, then*

$$|f(z)| < \frac{M}{R^m} |z|^m, \; z \in U_R, \tag{2.4}$$

the equality (in the inequality (2.4) for $z \neq 0$*) can hold only if* $f(z) = e^{i\theta} \frac{M}{R^m} z^m$*, where* θ *is constant.*

3. Main Results

Theorem 3.1. *Let $g \in A, \gamma$ be a complex number such that $\Re\gamma \geq 1$, M be a real number and $M > 1$. If*

$$|z(D_l^{m,\lambda}(a,b)g)'(z)| < M,\ z \in U, \tag{3.1}$$

and

$$|\gamma| \leq \frac{3\sqrt{3}}{2M}, \tag{3.2}$$

then the function

$$T_\gamma(z) = \left[\gamma \int_0^z u^{\gamma-1} \left(e^{D_l^{m,\lambda}(a,b)g(u)}\right)^\gamma du\right]^{\frac{1}{\gamma}}, \tag{3.3}$$

is in the class S.

Proof. Let us consider the function

$$f(z) = \int_0^z \left(e^{D_l^{m,\lambda}(a,b)g(u)}\right)^\gamma du, \tag{3.4}$$

which is regular in U. The function

$$h(z) = \frac{1}{|\gamma|} \frac{zf''(z)}{f'(z)}, \tag{3.5}$$

where the constant $|\gamma|$ satisfies the inequality (3.2), is regular in U. From (3.5) and (3.4) it follows that

$$h(z) = \frac{\gamma}{|\gamma|} z(D_l^{m,\lambda}(a,b)g)'(z). \tag{3.6}$$

Using (3.6) and (3.1) we have

$$|h(z)| < M, \tag{3.7}$$

for all $z \in U$. From (3.6) we obtain $h(0) = 0$ and applying Schwarz-Lemma we obtain

$$\frac{1}{|\gamma|}\left|\frac{zf''(z)}{f'(z)}\right| \leq M|z|, \tag{3.8}$$

for all $z \in U$, and hence, we obtain

$$\left(1-|z|^2\right)\left|\frac{zf''(z)}{f'(z)}\right| \leq |\gamma| M |z| \left(1-|z|^2\right). \tag{3.9}$$

Let us consider the function $Q : [0,1] \to \Re$, $Q(x) = x\left(1-x^2\right),\ x = |z|$. We have

$$Q(x) \leq \frac{2}{3\sqrt{3}}, \tag{3.10}$$

for all $x \in [0,1]$. From (3.10), (3.9) and (3.2) we obtain

$$\left(1-|z|^2\right)\left|\frac{zf''(z)}{f'(z)}\right| \leq 1, \tag{3.11}$$

for all $z \in U$. From (3.4) we obtain $f'(z) = \left(e^{D_l^{m,\lambda}(a,b)g(z)}\right)^\gamma$. Then, from (3.11) and Theorem 2.2 for $\Re\alpha = 1$ it follows that the function T_γ is in the class S. This is the proof. □

Corollary 3.2. *In the special case* $m = 0, a = b = 1,$ *Theorem 3.1 yields a result given earlier by* [9].

Theorem 3.3. *Let* $D_l^{m,\lambda}(a,b)g \in A$, *satisfy (2.1),* γ *be a complex number with* $\Re\gamma \geq 1$, M *be a real number,* $M > 1$ *and* $|\gamma - 1| \leq \frac{54\,M^4}{(12M^4+1)\sqrt{12\,M^4+1}+36\,M^4-1}$. *If*

$$\left|D_l^{m,\lambda}(a,b)g(z)\right| < M, \;\; z \in U,$$

then the function

$$H_\gamma(z) = \left[\gamma \int_0^z u^{2\gamma-2}\left[e^{D_l^{m,\lambda}(a,b)g(u)}\right)^{\gamma-1} du\right]^{\frac{1}{\gamma}},$$

is in the class S.

Proof. We observe that

$$H_\gamma(z) = \left[\gamma \int_0^z u^{\gamma-1}\left(ue^{D_l^{m,\lambda}(a,b)g(u)}\right)^{\gamma-1} du\right]^{\frac{1}{\gamma}}.$$

Let us consider the function

$$h(z) = \int_0^z \left(ue^{D_l^{m,\lambda}(a,b)g(u)}\right)^{\gamma-1} du.$$

The function h is regular in U. We obtain

$$\frac{h''(z)}{h'(z)} = (\gamma - 1)\,\frac{z(D_l^{m,\lambda}(a,b)g)'(z) + 1}{z},$$

and hence, we have

$$(1 - |z|^2)\left|\frac{zh''(z)}{h'(z)}\right| = |\gamma - 1|\left(1 - |z|^2\right)\left|z(D_l^{m,\lambda}(a,b)g)'(z) + 1\right|, \tag{3.12}$$

for all $z \in U$. From (3.12) we get

$$\left(1 - |z|^2\right)\left|\frac{zh''(z)}{h'(z)}\right| \leq |\gamma - 1|\left(1 - |z|^2\right)\left(\left|\frac{z^2(D_l^{m,\lambda}(a,b)g)'(z)}{(D_l^{m,\lambda}(a,b)g)^2(z)}\right|\frac{\left|(D_l^{m,\lambda}(a,b)g)^2(z)\right|}{|z|} + 1\right),$$

for all $z \in U$. By the Schwarz Lemma also $\left|D_l^{m,\lambda}(a,b)g(z)\right| \leq M|z|, \;\; z \in U$ and we obtain

$$\left(1 - |z|^2\right)\left|\frac{zh''(z)}{h'(z)}\right| \leq |\gamma - 1|\left(1 - |z|^2\right)\left(\left|\frac{z^2(D_l^{m,\lambda}(a,b)g)'(z)}{(D_l^{m,\lambda}(a,b)g)^2(z)} - 1\right| M^2|z| + M^2|z| + 1\right),$$

for all $z \in U$. Since $D_l^{m,\lambda}(a,b)g$ satisfies the condition (2.1) then, we have

$$\left(1 - |z|^2\right)\left|\frac{zh''(z)}{h'(z)}\right| \leq |\gamma - 1|\left(1 - |z|^2\right)\left(2M^2|z| + 1\right),$$

for all $z \in U$. Let us consider the function $G : [0,1] \to \Re$, $G(x) = \left(1 - x^2\right)(2M^2x + 1)$, $x = |z|$. We have

$$G(x) \leq \frac{\left(12\,M^4+1\right)\sqrt{12\,M^4+1}+36\,M^4-1}{54\,M^4},$$

for all $x \in [0,1]$. Since $|\gamma-1| \leq \frac{54\,M^4}{(12\,M^4+1)\sqrt{12\,M^4+1}+36\,M^4-1}$, we conclude that

$$\left(1-|z|^2\right)\left|\frac{zh''(z)}{h'(z)}\right| \leq 1, \tag{3.13}$$

for all $z \in U$. Note that, by (3.13) and Theorem 3.1 for $\Re\,\alpha = 1$ imply that the function H_γ is in the class S. This is the proof. □

Corollary 3.4. *In the special case* $m = 0, a = b = 1$, *Theorem 3.3 yields a result given earlier by* [9].

References

[1] A.A. Amer and M. Darus, On a property of a subclass of Bazilevic functions , *Missouri Journal of Mathematical Sciences*, to appear.

[2] A.A. Amer and M.Darus, A distortion theorem for a certain class of Bazilevic function , *Int. Journal of Math. Analysis*, 6, 591–597, (2012).

[3] O. Mayer, *The functions theory of one variable complex*, Bucureşti, 1981.

[4] Z. Nehari, *Conformal mapping*, Mc Graw-Hill Book Comp., New York, 1952 (Dover. Publ. Inc.), 1975

[5] S. Ozaki, M. Nunokawa, The Schwarzian derivative and univalent functions, *Proc. Amer. Math. Soc.* 2, 392–394, (1972).

[6] N. N. Pascu, An improvement of Becker's univalence criterion, *Procee-dings of the Commemorative Session Simion Stoilov, Braşov,* 43–48, (1987).

[7] V. Pescar, New univalence criteria, *"Transilvania" University of Braşov, Braşov.,* 2002.

[8] C. Pommerenke, Univalent functions, Gottingen, 1975.

[9] V. Pescar, On the Univalence of Some Integral Operators, *General Mathematics* 14 , 77—84, (2006).

[10] A. Catas ,On a Certain Differential Sandwich Theorem Associated with a New Generalized Derivative Operator, *General Mathematics,* 4,83—95,(2006).

[11] B.C. Carlson, D.B. Shaffer, Starlike and prestarlike hypergeometric functions, *SIAM J. Math. Anal.,* 15 , 737–745,(1984).

[12] St. Ruscheweyh, New criteria for univalent functions, *Proc. Amer. Math. Soc.* 49, 109–115, (1975).

[13] G. S. Salagean, Subclasses of univalent functions, *Lecture Notes in Math.* (Springer-Verlag), 1013, 362–372, (1983).

[14] F.M. AL-Oboudi, On univalent functions defined by a generalised Salagean Operator, *Int, J. Math. Math. Sci.*, 27, 1429-1436, (2004).

(A. A. Amer) School of Mathematical Sciences, Faculty of Science and Technology, University of Kebangsaan Malaysia, Bangi 43600 Selangor D. Ehsan, Malaysia
E-mail address: eamer_80@yahoo.com

(M. Darus) School of Mathematical Sciences, Faculty of Science and Technology, University of Kebangsaan Malaysia, Bangi 43600 Selangor D. Ehsan, Malaysia
E-mail address: maslina@ukm.my

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 17-22, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON EXACT VALUES OF MONOTONIC RANDOM WALKS CHARACTERISTICS ON LATTICES

ALEXANDER P. BUSLAEV AND ALEXANDER G. TATASHEV

Abstract. A monotonic random walk of particles on a one-dimensional circular lattice is considered. Some fixed number of particles moves on a ring, which contains a fixed number of cells. Each cell cannot contain more than one particle simultaneously. Jumps of particles can be occurred in discrete times. This jumps are realized with probability depending on the type of the particle and coordinate of the cell occupied by the particle, and the particle comes to the cell adjacent to the particle that is ahead, i.e., the length of the jump is maximum possible. A Markov chain is considered that describes the behavior of the model. The flow intensity and the average velocity of particle have been found.

1. Introduction

Models that described traffic in terms of cellular automata were introduced by K. Nagel et al., [8–10]. In these models particles are moving on a one-dimensional infinite lattice or ring one. The time is discrete. Each cell can be occupied by no more than one particle. The number of particles on a ring lattice is defined and remains to be constant. The models were considered in that the particles move one cell forward, the model in that the particles move immediately to the next particles going ahead, and also some generalizations of these models. The allowed movements are realized with given probabilities. The main investigated characteristics are particles flow intensity and average velocity. Simulation and heuristic approaches was used for investigation of these models. Some exact results on these themes were presented in [6]. Namely, an open chain of cells was considered, and the stationary state of the particles flow was investigated. A particle leaves the system when it has passed through all the cells. An algorithm has been elaborated for calculations of the state probabilities and the average particles velocity.

M. Blank articles [1–3] contain exact mathematical results for models of described type, where particles move on the infinite one-dimensional lattice. The possible transitions of particles are realized with probability equal to 1. Randomness occurs only in the choice of the initial configuration of particles.

An exact formula for the average particle velocity of random walks on a ring lattice was obtained in [5], where it was supposed that a particle passes one cell forward with a fixed probability p, $0 < p < 1$, provided the cell ahead of the particle is empty.

In [4], the limit formula for the average particle velocity is proved for the case when the number of particles and the number of cells tends to infinity so that the ratio of particles number to cells number tends to a constant. This formula is analogous to the formula for the average velocity of particles that was obtained in [9] by a heuristic method.

We consider here a monotonic random walk of particles on a one-dimensional circular lattice. Jumps of particles can be occurred in discrete times and are realized with probability depending on the type of the particle and coordinate of the cell occupied by the particle. A moving particle comes to the cell adjacent to the particle that is ahead (model with maximum transitions). The flow intensity and average velocity of particle have been found.

If the number of particles is equal to the number of cells minus one, then considered model is identical to the model where particles pass only one cell forward. Then the obtained formulas can be regarded as asymptotic formulas for the average velocity of particle in the model in that particles pass one cell forward, as the flow density tends to maximum value, which is equal to 1.

2. Nonhomogeneous chains with maximum jumps of particles

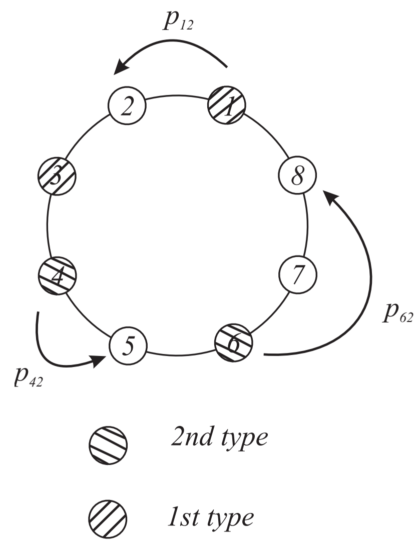

Figure 1. The model of monotonic random walk

Let us describe a stochastic model of particles movement on a closed sequence of cells, Fig 1. Let the number of cells be equal to n. The particles move at the discrete times $1, 2, \ldots,$ in the same direction. The $(i + 1)$-th cell follows the i-th cell in the direction of movement, $i = 1, \ldots, n - 1$. The cell 1 follows the cell of number n. Each cell is occupied by no more than one particle. There are k types of particles. The number of particles is equal to m, $1 < m < n$. There are m_s particles of the type s, $s = 1, \ldots, k$; $\sum\limits_{s=1}^{k} m_s = m$. The $(i + 1)$-th particle follows the i-th particle in the direction of movement, $i = 1, \ldots, m - 1$. The particle 1 moves ahead of the particle m. Suppose the i-th particle is a particle of the $s(i)$-th type, $i = 1, \ldots, m$. If a particle of the s-th type, $s = 1, \ldots, k$, occupies the i-th cell and there are d empty cells in front of this particle, then the particle passes the maximum number of cells with the probability $0 < p_{is} < 1$, i.e., it comes to the cell adjacent to the particle ahead. Hence the order of the particles can not be changed. Configuration of the particles at the initial time 0 is fixed.

Assume that the numbers n and m are coprime.

The particles behavior can be represented by a Markov chain with discrete time, [7]. Each state of chain corresponds to the vector $(i_1, \dots, i_m)$, j-th component of which is equal to the index of the cell occupied by the j-th particle, $j = 1, \dots, m$. The number of states is equal to mC_n^m, where $C_n^m = \frac{n!}{m!(n-m)!}$. Suppose the states of the Markov chain are numerated arbitrary.

A group of particles is called a cluster if there are no free cells between the particles of this group.

If the chain returns to some state with a probability less than 1, then this state is called non-recurrent, [7].

Suppose some state j exists such that the chain can come to the state j from the state i with non-zero probability and cannot return to the state i from the state j. It is evident that the state i is non-recurrent.

Every chain state with more than one cluster is a non-recurrent state. Really, the chain can come from such a state to a state with a smaller number of clusters and cannot come to any state with a greater number of clusters.

Let $P_i(j)$ be the probability of that the chain is in the i-th state at the time j, $i = 1, \dots, mC_n^m$, $j = 1, 2, \dots$

It is known from the theory of Markov chains, [7], that for each non-recurrent state i

$$\lim_{j \to \infty} P_i(j) = 0. \tag{1}$$

Thus (1) is true for each state with more than one cluster.

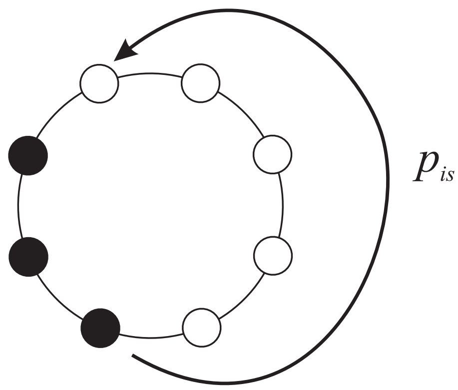

FIGURE 2. A state with one cluster

The class of states with one cluster (Fig. 2) contains mn states. The states of this class are described by the index of particle moving in front of the cluster and the index of cell that is occupied by this particle. Let the state for that the first particle of cluster is the j-th particle, occupying the i-th cell, be called the state E_{ij}, $i = 1, \dots, n$, $j = 1, \dots, m$.

Suppose that the state E_{ij} is the chain state at the time k. Then at the time $k+1$ the chain comes with the probability $p_{is(j)}$ to the state $E_{i^*j^*}$, where $i^* = i-1$, if $i = 2, \dots, n$, and $i^* = n$, if $i = 1$; $j^* = j - 1$, if $j^* = 2, \dots, m$, and $j^* = m$, if $j = 1$. And the state E_{ij} is the chain state at the time $k + 1$ with the probability $1 - p_{is(j)}$.

Lemma 1. *Suppose $E_{i_0 j_0}$ is some state with one cluster. Let $a - 1$ be maximum possible number of the different states of the chain after passing that the chain can return to the state $E_{i_0 j_0}$. The value a does not depend on i_0 and j_0, and equals mn, i.e., equals number of states with one cluster.*

Proof. The chain, after leaving a state of the class, returns to this state having been at $a - 1$ transit states, where a is the least common multiple of the numbers m and n. Since the numbers n and m are coprime, then $a = mn$, i.e., a is equal to number of the states of the class. Thus Lemma 1 is true. □

Let q be the flow intensity, i.e., the average number of particles passing through a cell per a time unit:

$$q = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} p_{is(j)} (n - m). \tag{2}$$

Denote

$$r = \frac{m}{n}, \ r_s = \frac{m_s}{n}, \ s = 1, \dots, k. \tag{3}$$

The value r is the particles flow density. The value r_s is the flow component composed by particles of the s-th type.

The flow intensity, the flow density, and the average velocity v are related by the formula

$$q = rv. \tag{4}$$

Theorem 2. *For the flow intensity and average velocity of particles, the following relations are true*

$$q = \frac{nr(1-r)}{\sum\limits_{i=1}^{n} \sum\limits_{s=1}^{k} \frac{r_s}{p_{is}}}, \tag{5}$$

$$v = \frac{n(1-r)}{\sum\limits_{i=1}^{n} \sum\limits_{s=1}^{k} \frac{r_s}{p_{is}}}, \tag{6}$$

Proof. It follows from Lemma 1 that the class of the states with one cluster is the unique class of the communicating states, i.e., it is possible to come from each state of the class to each other one with a non-zero probability.

This states are non-periodic also, i.e., the greatest common divisor of the values of possible time of the return to the state is equal to 1. Really, at each time the chain state can remain the same, and the time of recurrence can be equal to any number that is not less than mn.

According to a theorem of the Markov chains theory, [7], there exists a non-zero steady probability for each state that belongs to the unique class of the communicating non-periodic states of the Markov chain with a finite number of states, i.e., if i is a state of this class, then the limit exists

$$P_i = \lim_{j \to \infty} P_i(j) > 0.$$

Let P_{ij} be the steady probability of the state E_{ij}, $i = 1, \dots, n$, $j = 1, \dots, m$. Steady states satisfy the system of equations

$$p_{i^* s(j^*)} P_{i^* j^*} = p_{is(j)} P_{ij},\ i = 1, \dots, n,\ j = 1, \dots, m,$$
$$\sum_{i=1}^{n} \sum_{j=1}^{m} P_{ij} = 1,$$

where i^* and j^* are defined as above.

The solution of this system of equations is

$$P_{ij} = \frac{1/p_{is(j)}}{\sum\limits_{k=1}^{n} \sum\limits_{l=1}^{m} \frac{1}{p_{ks(l)}}},\ i = 1, \dots, n;\ j = 1, \dots, m. \tag{7}$$

Formula (5) follows from (2), (3), and (7).

Denote

$$p^* = \frac{nr}{\sum\limits_{i=1}^{n} \sum\limits_{s=1}^{k} \frac{r_s}{p_{is}}}. \tag{8}$$

Taking into account (8), we can rewrite (5) as $q = (1-r)p^*$. Combining (2), (7), and (8), we get (6) and, therefore,

$$v = (\frac{1}{r} - 1)p^*. \tag{9}$$

Thus Theorem 2 has been proved. □

Corollary 3. *If $m = 1$ and $p_{is} \equiv 1$, then*

$$v = \frac{1}{r} - 1. \tag{10}$$

Proof. Formula (10) follows from (6).

Formula (10) corresponds to a formula that is given in [3] for a similar model of movement on a one-dimensional lattice. □

3. Remarks

Suppose $m = n-1$. In this case the behavior of the model considered in Section 2, is identical to the appropriate models where the particles move no more than one cell forward. Hence relations (5) and (6) can be regarded as asymptotic formulas for the chain of big density, where particles move no more than one cell forward.

Suppose that the probability of displacement of particles does not depend on the particle type and the index of cell occupied by the particle

$$p_{is} = p,\ i = 1, \dots, n, s = 1, \dots, k. \tag{11}$$

If n and m tend to infinity so that ratio m/n tends to r, then following relation is true [4],

$$v_1 = \frac{1 - \sqrt{1 - 4pr(1-r)}}{2r},$$

where v_1 is the average velocity of particles in the model where the particles move no more than one cell forward.

Combining (6) and (11) we get

$$v = v_2 = (\frac{1}{r} - 1)p.$$

We have $\lim\limits_{r\to 1} \frac{v_2}{v_1} = 1$. This confirms adequacy of formulas (5) and (6) as asymptotic for the appropriate models in that particles move no more than one cell forward, for flow densities near 1.

4. Conclusion

Formulas for average intensity and average velocity are obtained for the model of random walks on a ring lattice. The probability of a particle transition depends on the particle type and the index of the cell occupied by the particle. The particles jump forward to the cell adjacent to ahead of particle.

References

[1] Blank, M., Dynamics of traffic jams: order and chaos, Mosc. Math. J., 1:1 (2001), pp. 1–26.

[2] Blank, M., Exact analysis of dynamical systems arising in models of traffic flow, Russian Mathematical Surveys, vol. 55, no. 3 (333).

[3] Blank, M., Exclusion processes with synchronous update in transport flow models, Trudy MFTY, 2010, vol. 2, no. 4, pp. 22–30.

[4] Buslaev, A. and Tatashev, A., Monotonic random walk on a one-dimensional lattice, Journal of Concrete and Applicable Mathematics (JCAAM), 2012, vol. 10, no: 1–2.

[5] Buslaev, A. and Tatashev, A., Particles flow on the regular polygon, Journal of Concrete and Applicable Mathematics (JCAAM), 2011, vol. 9, no. 4, 290–303.

[6] Gray, L. and Griffeath, D., The ergodic theory of traffic jams, J. Stat. Phys., 2001, vol. 3/4, 413–452.

[7] Karlin, S., A first course in stochastic processes, New York and London, 1968.

[8] Nagel, K. and Schreckenberg, M., A cellular automaton model for freeway traffic, J. Phys. I France 2, 1992, 2221-2229.

[9] Schadschneider, A. and Schreckenberg, M., Cellular automaton models and traffic flow, J. Phys. A. Math. Gen., 1993, vol. 51, no. 15, L679-L683.

[10] Schreckenberg, M., Schadschneider, A., Nagel, K., and Ito, N., Discrete stochastic models for traffic flow, Phys. Rev. E., 1995, vol. 51, 2939–2949.

(A.P. Buslaev) Moscow State Automobile and Road Technical University, Moscow, Russia.

E-mail address: `apal2006@yandex.ru.`

(A.G. Tatashev) Moscow Technical University of Communications and Informatics, Moscow, Russia.

E-mail address: `a-tatashev@rambler.ru`

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 23-46, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

EDEGEWORTH BLACK-SCHOLES OPTION PRICING FORMULA

ALI YOUSEF

ABSTRACT. In this paper we consider the Black-Scholes option pricing problem. We use Edgeworth second order approximation to approximate the underlying asset's return distribution. We state and prove Lemma 3.2 which finds a closed form for the solution of the Black-Scholes partial differential equation under Edgeworth approximation. The theoretical form of Edgeworth option pricing formula found in Lemma 3.2 depends mainly on the skewness and kurtois of the asset's return distribution which indeed corrects the classical Black-Scholes formula that was presented by Black and Scholes in 1973. Moreover, we find a simple form of delta hedging. In the end we find standardized Edgeworth expansions for the following distributions; lognormal, student t–distribution and chi–squared distributions.

1. INTRODUCTION

The Edgeworth expansion was first introduced by Edgeworth in 1905, see [11], as an expansion representing one distribution function in terms of another distribution, in such a way that the cumulants of the other distribution function should be known. Since its foundation in 1905, it became the center of many statistical studies and has been applied in many fields like, economics, finance and also in engineering, see [13]. The asymptotic behaviour were developed by [25], and its validity region was discuused by [4]. In the following section, we give some accounts about Edgeworth second order expansion and its validity region.

2. EDGEWORTH SECOND ORDER EXPANSION AND ITS ASYMPTOTIC PROPERTIES

Let $(X_1,\ldots,X_n)$ be a sequence of IID random varaibles of size n from a continuos distribution function F, such that the mean $\mu \in \Re$, the variance $\sigma^2 < \infty$, the skewness γ and the kurtosis β are all finite, where, $\gamma = \sigma^{-3} E(X-\mu)^3 < \infty$ and $\beta = \sigma^{-4} E(X-\mu)^4 < \infty$.

Let $\bar{X}_n = n^{-1}\sum_{i=1}^{n} X_i$, $Z_n = \sqrt{n}\left(\bar{X}_n - \mu\right)/\sigma$ and $F_n(x) = P(Z_n \le x)$.

Then the Central Limit Theorem states that the limiting distribution of Z_n, that is,

Key words and phrases. Black–Scholes formula, Chi–Squared, delta hedging, Edgeworth expansion, Lognormal, Student t–distribution.

$\lim_{n\to\infty} F_n(x) = \Phi(x) := \int_{-\infty}^{x} (2\pi)^{-1/2} \exp(-t^2/2)\,dt,\ \forall x \in \Re$ is a standard normal cumulative distribution function, regardless the analytic form of the underlying distribution function F.

If we are interested to find the probability distribution of Z_n before reaching the CLT then this will leads us to Edgeworth series, which can be derived by expanding the logarithm of the characterstic function of Z_n using taylor series expansion in t and collecting terms of the same order in n and then applying the inverse fourier transform to obtain the Edgeworth expansion, see, [18], chapter 4, page 91. Note, we need $\lim_{t\to\infty}\sup\left|E\left(e^{itX_1}\right)\right| < 1$, to ensure the expansion of the characterstic function of Z_n.

In the following line we state the two term Edgeworth expansion without proof as stated in [9], Theorem 13.1 page 186.

Theorem 2.1. (Two-Term Edgeworth Expansion) Suppose F satisfies the Cramer's condition and $E_F\left(X^4\right) < \infty$. Then,

$$F_n(x) = \Phi(x) + \phi(x)\left\{c_1 p_1(x) n^{-1/2} + \left(c_2 p_2(x) + c_3 p_3(x)\right) n^{-1}\right\} + O\left(n^{-3/2}\right).$$

defined for all values of x and as $n \to \infty$. Note that $\phi = \Phi'$ is the standard normal density function,

$$c_1 = \gamma/6,\ c_2 = (\beta - 3)/24,\ c_3 = \gamma^2/72,\ p_1 = 1 - x^2,\ p_2 = 3x - x^3 \text{ and}$$
$$p_3 = -15x + 10x^3 - x^5.$$

Note that c_1 and c_2 are called the skewness and kurtosis corrections of the underlying distribution function. From Theorem 2.1 and by direct substitutions for c_1, c_2, c_3, p_1, p_2 and p_3 and by taking $n = 1$, we have the heuristic result

$$(2.1)\qquad F(x) = \Phi(x) - \frac{1}{72}\phi x\left(x^4 - 10x^2 + 15\right)\gamma^2 - \frac{1}{6}\gamma\phi\left(x^2 - 1\right) - \frac{1}{24}\phi x\left(x^2 - 3\right)(\beta - 3) + O(1)$$

If the distribution function F of an absolutely continuous random variable admits an Edgeworth expansion, then we can obtain an expansion of the density function heuristically by differentiating (2.1) with respect to x. Hence the probability density function of the Edgeworth expansion is

$$(2.2)\qquad f(x)=\phi(x)\begin{Bmatrix} 1+\frac{1}{6}\gamma x\left(x^2-3\right)+\frac{1}{24}(\beta-3)\left(x^4-6x^2+3\right) \\ +\frac{1}{72}\gamma^2\left(x^6-15x^4+45x^2-15\right) \end{Bmatrix}+O(1)$$

Equation (2.2) shows how to represent a continuous probability density function in terms of the standardized normal probability density function and which is an approximating standardized density with the desired skewness and kurtosis. The density function in (2.2) is called a standardized Edgeworth asymptotic expansion; see [19].

The Edgeworth expansion is more useful in many applications than other asymptotic series such as the Gauss-Hermite and Gram-Charlier series, see [6]. This is so, because, it is directly connected to the moments and cumulants of a probability density function, a property that is lost in the Gauss-Hermite series. Secondly, it is a true asymptotic expansion since the error of the approximation is controlled by estimating the error of the expansion until the order $O\left(n^{-3/2}\right)$. While, the disadvantage of the series was shown by [3] that the Edgeworth series can give negative values for some values of x . They found the region in the plane of values of skewness and kurtosis where the density is positive. This region was further studied by [10] in detail using numerical methods. They found that the validity region that ensures the Edgeworth series to represent a positive definite and unimodal probability density function is $R_v=\left\{|\gamma|\le 0.45,\ 3.0\le\beta\le 5.35\right\}$. If the parameters lie outside the validity region (as we shall see in the last section), the results may be misleading. Furthur analytical investigations about the validity region were undertaken by [1]. References for the main results on Edgeworth expansions are given in [2,4,15].

In the next section, we consider the Edgeworth Black-Scholes option pricing problem.

3. EDGEWORTH OPTION PRICING FORMULA

In the early 1970's Black and Scholes made an important contribution in the pricing of complex financial instruments by developing what so called the Black-Scholes model. The model based on a formula that describes the market with two financial goods: a risky security, such as a stock(risky asset) and a riskless bond. An option is a contract that gives the investor the right to buy or sell the asset or part of it without any obligations subject to certain conditions within a specific period of time. The price that is paid for the asset when the option is exercised is called the strike price K , while the expiration day is the last day on which the option may be exercised. If the option can be dealt only at maturity time t, then we have a European-style, otherwise, we have an American style.

The most important application of Ito's lemma in the area of financial mathematics is option pricing and the most interesting result in option pricing is the Black-Scholes formula. To simply our discussion assume that the model that can be used to represent the random price S_t of a risky asset follows a geometric Brownian motion,

$$(3.1) \qquad dS_t = \mu S_t\, dt + \sigma S_t\, dW_t$$

Where, $\{S_t, t \in \Re^+\}$ is a stochastic processes, μ is the average rate of the asset price growth, $\mu\, dt$ is the drift mean, σ is the volatility, $\sigma\, dW_t$ is the random change in the asset price and W is a Wiener process.

Equation (3.1) is called a stochastic differential equation and states that the return of the asset $\dfrac{dS_t}{S_t}$ consists mainly of two main components; a constant return $\mu\, dt$ and a random return $\sigma\, dW_t$.

By applying Ito's Lemma, see [24], Lemma 3.2,page 96 to (3.1) we obtain the Black-Scholes partial differential equation for a European call option price C_t .

$$(3.2) \qquad \frac{\partial C_t}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 C_t}{\partial S_t^2} + rs\frac{\partial C_t}{\partial S_t} = r\, C_t$$

Equation (3.2) is called a parabolic partial differential equation, and r is the risk free rate.

It was shown from [22] that the payoff of the European call option is given by $C_t = \max\{S_t - K, 0\}$, where S_t is the price of an asset at maturity time t and K is the strike price.

Now, since S_t satisfies, (3.1) then we can use the substitution $Y_t = \log(S_t)$ and Ito's Lemma to obtain

$S_t = S_0 \exp\left(\int_0^t \left(\mu - \sigma^2/2\right) ds + \int_0^t \sigma\, dW \right)$ which yield the following solution

$$(3.3) \qquad S_t = S_0 \exp\left(\left(\mu - \sigma^2/2\right)t + \sigma W_t\right), \forall\, t \in \Re^+,\, W_t \sim N(0,t)$$

Note (3.3) is the exact solution of (3.1) and $\log(S_t) \sim N\left(\left(\mu - \sigma^2/2\right)t,\, \sigma^2 t\right)$.

Under Q (risk neutral probability measure) , the average price of the asset should satisfy the following equation

$$(3.4) \qquad E(S_t) = S_0 \exp(rt)$$

Hence from Q , the value of the call option at time zero is

$$C_0 = \exp(-rt) E_Q \left\lfloor (S_t - K)^+ \right\rfloor.$$

The solution of (3.2) yields the celebrated Black-Scholes formula for the European call option at time zero with a non-dividend paying stock and with strike price K and maturity time t

$$(3.5) \qquad C_0^{BS} = S_0 \Phi(d_1) - K \exp(-rt) \Phi(d_2),$$
$$d_1 = \frac{\ln(S_0/K) + \left(r + \sigma^2/2\right)t}{\sigma\sqrt{t}}, d_2 = d_1 - \sigma\sqrt{t}.$$

Where S_0 is the current asset's price at time zero, σ is the volatility of the relative price change of underlying asset price, r is the short term risk free interest rate, K is the strike price and t is the maturity time of the option. Note, all the above parameters can be obtained from the market data except σ. For more details see, [5]. Volatility is a measure of riskiness of an asset, the more volatile it is, the more risky it is. Technically, it measures the standard deviation of short term returns on the asset, which can be estimated either from the historical price path of the asset, or using discrete method. For more details about volatility see, [7].

The disadvantages of (3.5) lies in two main points; first it assumes that the volatility of an asset remains constant all the period time, irrespective of the direction of price movements. Second, it assumes that the distributional form of the asset returns should be normally distributed, which is inconsistent with the observed data in the market. It was shown by many statisticians and observers that the distribution of assets return follows a non-normal distribution, presenting heavy tails and asymmetry. For more details, see [14, 21]. The point lies that the normality assumption of the Black-Scholes model does not capture extreme movements such as stock market crashes. To resolve this problem several series approximation techniques were used: [8], used Gram-Charlier expansion to approximate the return's asset distribution, while [12] used Edgeworth approximation. [17] used the generalized Edgeworth expansion of the lognormal distribution of asset prices to obtain a model that adjust the Black-Scholes formula for the higher moments in the underlying distribution.

In this paper, we will find a closed form for the theoretical European call option price assuming the return's asset distribution is an absolutely continuos distribution function, provided the first six moments exist. Our technique based on using Edgeworth second order approximation which has a bounded error.

To proceed our solution we need the following Remarks which will enhanced our main result in Lemma 3.2.

Remark 3.1

Let $f(y)$ be a standardized Edgeworth probability density function defined as in (2.2) and let ϕ and Φ be respectively, the probability density and distribution function of $N(0,1)$, then

$$\int_z^\infty f(y)\,dy = \overline{\Phi}(z) + \phi(z)Q(\gamma,\beta,z),$$
$$Q(\gamma,\beta,z) = \left\{\frac{\gamma^2}{72} z\left(z^4 - 10z^2 + 15\right) + \frac{\gamma}{6}\left(z^2 - 1\right) + \frac{(\beta-3)}{24} z\left(z^2 - 3\right)\right\}.$$

Remark 3.1 follows immediately by expanding the integral over the terms and using the following identities,

$$\int_z^\infty \phi(x)\,dx = \overline{\Phi}(z),\ \int_z^\infty x\phi(x)\,dx = \phi(z), \int_z^\infty x^2\phi(x)\,dx = \overline{\Phi}(z) + z\phi(z),\ \int_z^\infty x^3\phi(x)\,dx = \phi(z)\left(z^2+2\right),$$
$$\int_z^\infty x^4\phi(x)\,dx = 3\overline{\Phi}(z) + z\phi(z)\left(z^2+3\right) \text{ and } \int_z^\infty x^6\phi(x)\,dx = 15\overline{\Phi}(z) + z\phi(z)\left(z^4+5z^2+15\right).$$

The above integrals can be found easily from integration by part and the properites of error functions. Note $\overline{\Phi}(z) = 1 - \Phi(z)$.

Remark 3.2

Under the same condition of Remark 3.1, we have

$$\int_z^\infty \exp(by)\, f(y)\,dy = \exp\left(b^2/2\right)\Phi(-z+b)\, f(\gamma,\beta) + \phi(z)\exp(bz)\, g(\gamma,\beta,b,z).$$

where,

$$f(\gamma,\beta) = \left\{\frac{1}{72}\gamma^2 b^6 + \frac{1}{6}\gamma b^3 + \frac{1}{24}(\beta-3)b^4 + 1\right\}, g = \left\{\frac{1}{72}T_1\gamma^2 + \frac{1}{6}T_2\gamma + \frac{1}{24}(\beta-3)T_3\right\},$$

$$T_1 = z^5 + bz^4 + \left(b^2-10\right)z^3 + b\left(b^2-6\right)z^2 + \left(b^4-3b^2+15\right)z + b\left(b^4-b^2+3\right),$$

$T_2 = \left(z^2+bz+b^2-1\right)$ and $T_3 = \left(z^3+bz^2+\left(b^2-3\right)z+b^3-b\right).$

Remark 3.2 can be verified by using the identitiy $\int_z^\infty \exp(by)\phi(y)\; dy = \exp\left(b^2/2\right)\Phi(b-z),$

Which can be proved as follows

$$\int_z^\infty \exp(by)\phi(y)\; dy = \frac{1}{\sqrt{2\pi}}\int_z^\infty \exp\left(by - y^2/2\right)\, dy = \frac{1}{\sqrt{2\pi}}\int_z^\infty \exp\left(-\frac{1}{2}\left(y^2-2by\right)\right)\, dy$$

$$= \frac{1}{\sqrt{2\pi}}\int_z^\infty \exp\left(-\frac{1}{2}(y-b)^2 + b^2/2\right)\, dy = \frac{1}{\sqrt{2\pi}}\exp\left(b^2/2\right)\int_z^\infty \exp\left(-\frac{1}{2}(y-b)^2\right)\, dy.$$

Let $u = y-b \Rightarrow du = dy \Rightarrow \dfrac{1}{\sqrt{2\pi}}\exp\left(b^2/2\right)\displaystyle\int_{z-b}^\infty \exp\left(-\frac{1}{2}u^2\right)\, du$

$$= \frac{1}{\sqrt{2\pi}}\exp\left(b^2/2\right)\overline{\Phi}(z-b) = \frac{1}{\sqrt{2\pi}}\exp\left(b^2/2\right)\Phi(b-z).$$

and using the following identities;

$$\int_{z-b}^\infty (u+b)\phi(u)\,du = \phi(b-z) + (b/2)\left(1+2\Phi(b-z)\right),$$

$$\int_{z-b}^\infty (u+b)^2\,\phi(u)\,du = \phi(b-z)(z+b) + (1/2)\left(b^2+1\right)\left(1+2\Phi(b-z)\right),$$

$$\int_{z-b}^\infty (u+b)^3\,\phi(u)\,du = \phi(b-z)\left(z^2+zb+b^2+2\right) + (1/2)b\left(b^2+3\right)\left(1+2\Phi(b-z)\right),$$

$$\int_{z-b}^{\infty}(u+b)^4\phi(u)\,du =$$
$$\phi(b-z)\left(z^3+z^2b+z\left(b^2+3\right)+b^3+5b\right)+(1/2)\left(b^4+6b^2+3\right)\left(1+2\Phi(b-z)\right)$$
and
$$\int_{z-b}^{\infty}(u+b)^6\phi(u)\,du =$$
$$\phi(b-z)\left(z^5+z^4b+z^3\left(b^2+5\right)+z^2\left(9b+b^3\right)+z\left(b^4+12b^2+15\right)+b^5+14b^3+33b\right)$$
$$+(1/2)\left(b^6+15b^4+45b^2+15\right)\left(1+2\Phi(b-z)\right).$$

The risk free rate under the classical Black-Scholes formula (underlying asset's return distribution is standard normal) is $\mu = r$, wheras under the Edgeworth series will be derived in the following Lemma 3.1.

Lemma 3.1

Under Q ,the value of the risk free rate that satisfies (3.4) is,

$rt = \mu t + \ln f(\gamma,\beta,\sigma,t)$ where,

$$f(\gamma,\beta,\sigma,t)=\left\{\frac{1}{72}\gamma^2\left(\sigma\sqrt{t}\right)^6+\frac{1}{6}\gamma\left(\sigma\sqrt{t}\right)^3+\frac{1}{24}(\beta-3)\left(\sigma\sqrt{t}\right)^4+1\right\}$$

Proof

$$E(S_t)=\lim_{z\to-\infty}\int_z^{\infty}S_0\exp\left(\left(\mu-\sigma^2/2\right)t+\sigma\sqrt{t}y\right)f(y)\,dy$$
$$=S_0\exp\left(\left(\mu-\sigma^2/2\right)t\right)\lim_{z\to-\infty}\int_z^{\infty}\exp\left(\sigma\sqrt{t}y\right)f(y)\,dy$$

Recalling Remark 3.2 , we obtain the following

$$E(S_t)=S_0\exp\left(\left(\mu-\sigma^2/2\right)t\right)\lim_{z\to-\infty}\left\{\begin{array}{l}\exp\left(b^2/2\right)\Phi(-z+b)\,f(\gamma,\beta)\\ +\phi(z)\exp(bz)\,g(\gamma,\beta,b,z)\end{array}\right\},$$

$$=S_0\exp\left(\left(\mu-\sigma^2/2\right)t\right)\lim_{z\to-\infty}\exp\left(b^2/2\right)\Phi(-z+b)\,f(\gamma,\beta),$$
$$+S_0\exp\left(\left(\mu-\sigma^2/2\right)t\right)\lim_{z\to-\infty}\phi(z)\exp(bz)\,g(\gamma,\beta,b,z),$$

$$E(S_t) = S_0 f \exp\left(\left(\mu - \sigma^2/2\right)t\right)\exp\left(b^2/2\right)\lim_{z\to-\infty}\Phi(-z+b)$$
$$+S_0 \exp\left(\left(\mu - \sigma^2/2\right)t\right)\lim_{z\to-\infty}\phi(z)\exp(bz)\,g(z),$$

Since $\lim_{z\to-\infty}\Phi\left(-z+b\right)=1$, then

$$E(S_t) = S_0 f \exp\left(\left(\mu - \sigma^2/2\right)t\right)\exp\left(b^2/2\right)$$
$$+S_0 \exp\left(\left(\mu - \sigma^2/2\right)t\right)\lim_{z\to-\infty}\phi(z)\exp(bz)\,g(z),$$

Let $b = \sigma\sqrt{t}$, then

$$E(S_t) = S_0 f \exp(\mu t) + S_0 \exp\left(\left(\mu - \sigma^2/2\right)t\right)\lim_{z\to-\infty}\phi(z)\exp\left(\sigma\sqrt{t}z\right)g(z)$$

By completing the square we obtain

$$E(S_t) = S_0 f \exp(\mu t) + S_0 \exp(\mu t)\lim_{z\to-\infty}\phi\left(-z+\sigma\sqrt{t}\right)g(z)$$

But, $\lim_{z\to-\infty}\phi\left(-z+\sigma\sqrt{t}\right)z^n = 0,$ for all $t \ge 0$ and $n \in \mathbb{Z}^+$

Which implies that $E(S_t) = S_0 f \exp(\mu t)$.

But we need,

$$E(S_t) = S_0 \exp(rt) \Rightarrow S_0 \exp(\mu t)\, f(\gamma,\beta) = S_0 \exp(rt)$$

By taking the natural logaritms for both sides we obtain

$rt = \mu t + \ln f(\gamma,\beta)$.

Proof is completed.

As a result from Lemma 3.1 , (3.3) reduced to

$$(3.6) \qquad S_t = \frac{S_0}{f}\exp\left(\left(r - \sigma^2/2\right)t + \sigma W_t\right) \qquad ,W_t \sim N(0,t)$$

Now we are ready to state and prove the main result in Lemma 3.2.

Lemma 3.2

Let Q be the risk-neutral probability measure, and let S_0 and C_0 be respectively the asset value and a European call option at time zero, with strike price K and maturity time t, and let r be the risk free rate that satisfies Lemma 3.1, then, for $b=\sigma\sqrt{t}$, the Edgeworth Black-Scholes option pricing formula formula is

$$C_0^{Edg} = S_0\Phi(d) - Ke^{-rt}\Phi(d-b) + S_0\,\phi(d)\frac{g}{f} - K\exp(-rt)\phi(d-b)Q(\gamma,\beta,d)$$

where, $d = \dfrac{\ln(S_0/Kf) + (r+\sigma^2/2)t}{b}$, $z = -(d-b)$

and

$$f(\gamma,\beta) = \left\{\frac{1}{72}\gamma^2 b^6 + \frac{1}{6}\gamma b^3 + \frac{1}{24}(\beta-3)b^4 + 1\right\}.$$

$$Q(\gamma,\beta,d) = \left\{\frac{1}{72}\gamma^2 z\left(z^4 - 10z^2 + 15\right) + \frac{1}{6}\gamma\left(z^2-1\right) + \frac{1}{24}(\beta-3)z\left(z^2-3\right)\right\}.$$

$$g(\gamma,\beta,d) = \left\{\frac{1}{72}T_1\gamma^2 + \frac{1}{6}T_2\gamma + \frac{1}{24}(\beta-3)T_3\right\},$$

While T_1, T_2 and T_3 are respectively,

$$T_1 = z^5 + bz^4 + \left(b^2-10\right)z^3 + b\left(b^2-6\right)z^2 + \left(b^4 - 3b^2 + 15\right)z + b\left(b^4 - b^2 + 3\right).$$

$$T_2 = \left(z^2 + bz + b^2 - 1\right), T_3 = \left(z^3 + bz^2 + b^2 z - 3z + b^3 - b\right).$$

Proof

Under the risk neutral measure, the value of the European call option at time zero is

$$C_0 = \exp(-rt)E_Q\left\lfloor (S_t - K)^+ \right\rfloor.$$

From (3.6), we have

$$S_t = \frac{S_0}{f}\exp\left(\left(r - \sigma^2/2\right)t + \sigma W_t\right), \quad W_t \sim N(0,t)$$

Let $Y = W_t/\sqrt{t}$, then

$$S_t = \frac{S_0}{f}\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right), Y \sim N(0,1)$$

The value of the European call option at time zero is $\max(S_t - K, 0)$. Then

$$\begin{aligned}
C_0 &= \exp(-rt)E\left\{\max\left(S_t - K, 0\right)\right\} \\
&= \exp(-rt)E\left\{\max\left(\frac{S_0}{f}\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right)-K,0\right)\right\} \\
&= \exp(-rt)\int_{-\infty}^{\infty}\max\left(\frac{S_0}{f}\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right)-K,0\right)f(y)\,dy
\end{aligned}$$

Where $f(y)$ given by (2.2) .

Now, $\max\left(\frac{S_0}{f}\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right)-K,0\right)$ occurs when

$$\frac{S_0}{f}\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right) > K \Leftrightarrow Y > z = \left(\ln\left(K\,f/S_0\right)-\left(r-\sigma^2/2\right)t\right)\Big/\sigma\sqrt{t}$$

which implies that

$$\begin{aligned}
C_0 &= \exp(-rt)\int_{z}^{\infty}\left(\exp\left(\left(r-\sigma^2/2\right)t+\sigma\sqrt{t}Y\right)-K\right)f(y)\,dy \\
&= \frac{S_0}{f}\exp\left(\left(-\sigma^2/2\right)t\right)\int_{z}^{\infty}\exp\left(\sigma\sqrt{t}Y\right)f(y)\,dy - K\exp(-rt)\int_{z}^{\infty}f(y)\,dy
\end{aligned}$$

By recalling Remark 3.1 and Remark 3.2 and taking $b=\sigma\sqrt{t}$ we have

$$\begin{aligned}
C_0 &= \frac{S_0}{f}\exp\left(\left(-\sigma^2/2\right)t\right)\int_{z}^{\infty}\exp\left(\sigma\sqrt{t}Y\right)f(y)\,dy - K\exp(-rt)\int_{z}^{\infty}f(y)\,dy \\
&= S_0\Phi(-z+b)+\frac{S_0}{f}\phi(-z+b)g - K\exp(-rt)\overline{\Phi}(z) - K\exp(-rt)\phi(z)Q \\
&= S_0\Phi(-z+b) - K\exp(-rt)\overline{\Phi}(z)+\frac{S_0}{f}\phi(-z+b)g - K\exp(-rt)\phi(z)Q
\end{aligned}$$

Since $\overline{\Phi}(z)=1-\Phi(z)=\Phi(-z)$, then

$$C_0 = \left(S_0 \Phi(-z+b) - K\exp(-rt)\Phi(-z)\right) + \left(\frac{S_0}{f}\phi(-z+b)g - K\exp(-rt)\phi(z)Q\right)$$

But,

$$-z+\sigma\sqrt{t} = -\frac{\ln(Kf/S_0) - (r-\sigma^2/2)t}{\sigma\sqrt{t}} + \sigma\sqrt{t}$$

$$= \frac{-\ln(Kf/S_0) + (r-\sigma^2/2)t}{\sigma\sqrt{t}} + \sigma\sqrt{t} = \frac{\ln(S_0/K\,f) + \left(r+\frac{1}{2}\sigma^2\right)t}{\sigma\sqrt{t}} = d$$

then, $-z = d - \sigma\sqrt{t}$.

Hence,

$$C_0 = \left(S_0\Phi(d) - K\exp(-rt)\Phi\left(d-\sigma\sqrt{t}\right)\right) + \left(\frac{S_0}{f}\phi(d)g - K\exp(-rt)\phi\left(d-\sigma\sqrt{t}\right)Q\right)$$

Proof completed.

Lemma 3.2 states that the Edgeworth European call option price is the same as the classical option price plus a significant term that depends mainly on the standardized skewness and kurtosis of the underlying asset's return distribution. This term is

$$(3.7) \qquad S_0\,\phi(d)\frac{g}{f} - K\exp(-rt)\phi\left(d-\sigma\sqrt{t}\right)Q(\gamma,\beta,d)$$

Clearly, under the normal distribution,

$f = 1, \frac{g}{f} = 0$ and $Q = 0 \Rightarrow C_0^{Edg} = C_0^{BS}$.

Thus (3.7) measures the departure amount from normality assumption. Also, note that, Lemma 3.4 provides a solution of the Black-Scholes equation under the Edgeworth series. Empirical study could be used to test our formula in Lemma 3.2. [23] tested the model suggested by [8], to price call options on S&P CNX Nifty. Their results strongly suggested that including the skewness and kurtosis terms to the Black-Scholes option pricing formula yield values much closer to market prices. They confirm by their study that fitting of higher order moments of the distribution of returns would indeed improve the approximation of the call option, especially for options away from the money. This suggest that using our formula would indeed be more efficient than other formulas. Before, we end up this section, it would be worth mentioning the delta hedging which controlled the risk occurs in option pricing,

Corollary 3.1

Under the conditions of Lemma 3.2, the delta hedging result from applying Edgeworth option pricing formula is

$$\Delta = \Phi(d) + \frac{1}{f}\phi(d)\left(g + S\frac{\partial g}{\partial S} - g\frac{d}{b}\right) - K\exp(-rt)\phi(d-b)\left(\frac{\partial Q}{\partial S} - Q(d-b)\frac{1}{bS}\right),$$

where,

$$\frac{\partial g}{\partial S} = \left\{\frac{1}{72}T_1^*\gamma^2 + \frac{1}{6}T_2^*\gamma + \frac{1}{24}(\beta-3)T_3^*\right\}.$$

$$\frac{\partial Q}{\partial S} = \left\{\frac{5}{72}\gamma^2\left(z^4 - 6z^2 + 3\right) + \frac{1}{3}\gamma(z) + \frac{1}{8}(\beta-3)\left(z^2-1\right)\right\}, \quad T_i^* = -\frac{1}{bS}\frac{\partial T_i}{\partial z}, i = 1,2,3.$$

Proof

From [16], the delta hedging is

$$\Delta = \frac{\partial}{\partial S}C_0^{Edge} \text{ and } \Delta^{BS} = \Phi(d)$$

Thus,

$$\Delta = \Phi(d) + \frac{\partial}{\partial S}\left\{S_0\,\phi(d)\frac{g}{f} - K\exp(-rt)\phi(d-b)Q\right\}$$

$$= \Phi(d) + \frac{1}{f}\frac{\partial}{\partial S}\left(S\,g\phi(d)\right) - K\exp(-rt)\frac{\partial}{\partial S}\left(\phi(d-b)Q\right)$$

by using the chain rule differentiation technique and using the idenities

$$\frac{\partial g}{\partial S} = \left\{\frac{1}{72}T_1^*\gamma^2 + \frac{1}{6}T_2^*\gamma + \frac{1}{24}(\beta-3)T_3^*\right\},$$

$$\frac{\partial Q}{\partial S} = \left\{\frac{5}{72}\gamma^2\left(z^4 - 6z^2 + 3\right) + \frac{1}{3}\gamma(z) + \frac{1}{8}(\beta-3)\left(z^2-1\right)\right\},$$

where,

$$T_i^* = -b\,S\frac{\partial T_i}{\partial S}, i = 1,2,3.$$

the proof is complete.

Now, to find the estimated values of m, s, g and b from the market data we proceed as follows;

Let $S_1, \ldots, S_n$ be the daily closing prices, and let $R_i = \ln(S_{i+1}) - \ln(S_i)$ for all $i = 1, 2, \ldots, n$, then

$$\hat{\mu} = \bar{R} = n^{-1}\sum_{i=1}^{n} R_i \,,\ \hat{\sigma} = \sqrt{252}\, S_R \ \text{ where, } S_R = (n-1)^{-1/2}\sqrt{\sum_{i=1}^{n}\left(R_i - \bar{R}\right)^2}\,,$$

$$\hat{\gamma} = \frac{n}{(n-1)(n-2)}\sum_{i=1}^{n}\left(\frac{R_i - \bar{R}}{S_R}\right)^3 \text{ and}$$

$$\hat{\kappa} = \frac{n(n+1)}{(n-1)(n-2)(n-3)}\sum_{i=1}^{n}\left(\frac{R_i - \bar{R}}{S_R}\right)^4 - \frac{3(n-1)^2}{(n-2)(n-3)}$$

. All the above formulas can be executed directly from Excell sheet.

Example 3.1

We considered the data sheet found in Hull website that represents the stock prices for S&P during the period from 18th of July 2005 till 13th of August 2010. The descriptive statistics for Returns are illustrated below

	Mean	Std	Min	Max	SK	Kurt
Returns	0.000024	0.015531	-0.09	0.1158	0.027903	9.250233

Table 3.1. Descriptive statistics for S&P Returns

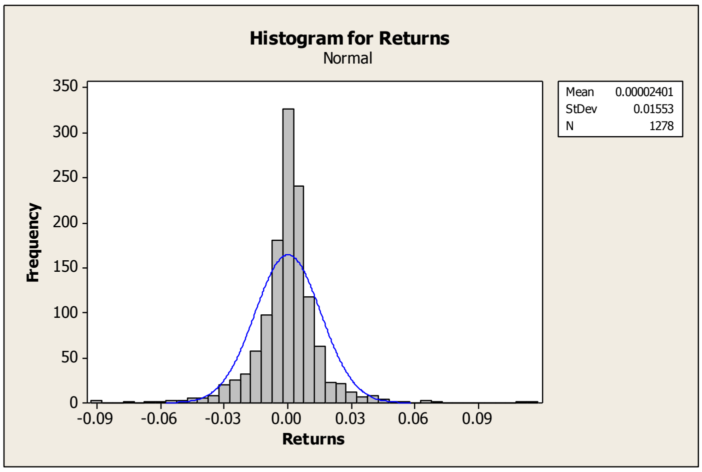

Figure 3.1. Returns for S&P prices from 18/07/2005 till 13/08/2010

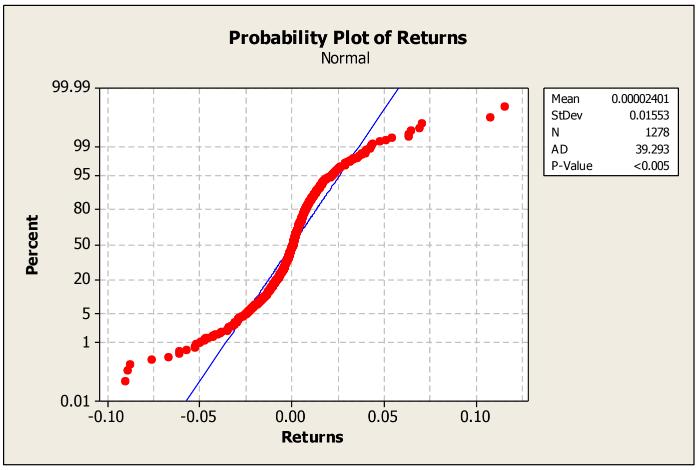

Figure 3.2. Testing normality of return's underlying distribution using Anderson-Darling Test.

Figure 3.1 and Figure 3.2 above assert that the underlying asset return's distribution do not follow a normal distribution with P-value less than 0.005, using Anderson-Darling test. To compare our formula with the classical Black-Scholes formula presented in 1973 and with the modified one that was presented by [8] we consider the following data:

Data: $S_0 = \$1200,\ K = \$1200,\ r = \%5,\ \sigma = 0.24655\,\text{years},\ t = 5.07\,\text{years}$

	Black-Scholes	Gram-Charlier	Edgeworth
Call Option	$390.486	$374.255	$350.85
Delta	0.769	0.795866	0.73791
d	0.7342089	0.7342089	0.6881122

Table 3.2. European call option and Delta for each case

Table 3.2 shows that our formula adjusted the European call option price rather than the previous ones, moreover the value of the delta hedging is less than others. These results indicate that using Edgeworth option pricing formula would be much better for detecting the high oscillation of market prices.

In the last section, it would be suitable to find standardized Edgeworth expansions for some densities that are widely used in finance; these densities are lognormal, student's t distribution, and chi-squared distributions.

4. EDGEWORTH APPROXIMATION FOR STATISTICAL DISTRIBUTIONS

I. Lognormal distribution

A random variable X has a lognormal distribution with parameters μ and σ, if $\ln(X) \sim N(\mu, \sigma^2)$. The probability density function of the lognormal distribution is

$$f_X(x;\mu,\sigma) = \left(2\pi x^2\sigma^2\right)^{-1/2} \exp\left(-\frac{1}{2}\left(\frac{\ln X-\mu}{\sigma}\right)^2\right) I(x>0),$$

It is mainly used to describe the distribution of asset prices over a specific period of time. The skewness and kurtosis of X are respectively

$$\gamma = \left(e^{\sigma^2}+2\right)\sqrt{e^{\sigma^2}-1} \text{ and } \beta = e^{4\sigma^2}+2e^{3\sigma^2}+3e^{2\sigma^2}-3.$$

From (2.2) the Edgeworth asymptotic expansion for the standardized lognormal distribution is

$$f(x) = \phi(x)\begin{Bmatrix} 1+6^{-1}(\eta+2)\sqrt{\eta^2-1}x\left(x^2-3\right)+24^{-1}\left(\eta^4+2\eta^3+3\eta^2-6\right)\left(x^4-6x^2+3\right) \\ +72^{-1}(\eta+2)^2(\eta-1)\left(x^6-15x^4+45x^2-15\right) \end{Bmatrix} + O(1)$$

, $\eta = e^{\sigma^2}$.

Clearly, the Edgeworth series depends mainly on the variance of the underlying distribution.

Table 4.1 below shows that both the values of skewness and kurtosis increases as the value of standard deviation increases.

Standard deviation	Skewness	kurtosis
0.01	0.0300017501	3.00160023
0.05	0.1502190453	3.04014412
0.10	0.3017590981	3.16232386
0.50	1.7501896560	8.89844568
1.00	6.1848771360	113.9363922

Table 4.1. Skewness and kurtosis values associated for every σ under the lognormal distribution

Figure 4.1 below exhibits the behavior of the standardized Edgeworth density as the values of standard deviation increases.

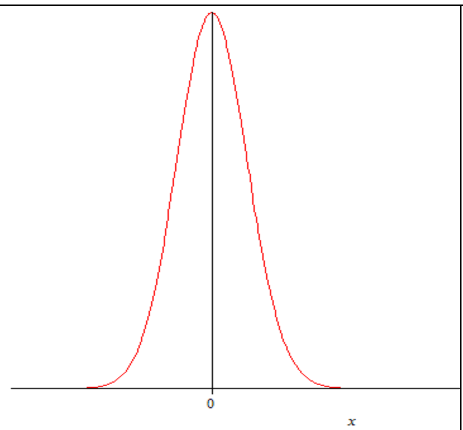

Figure (a), $\sigma = 0.01$

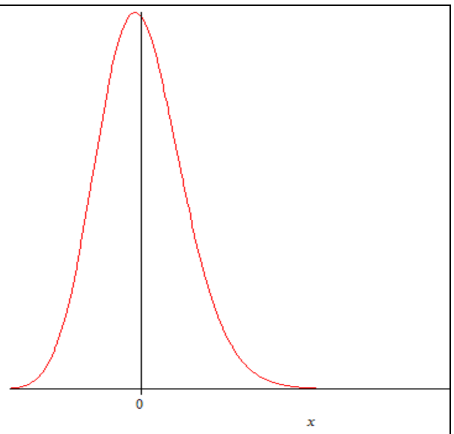

Figure (b), $\sigma = 0.1$

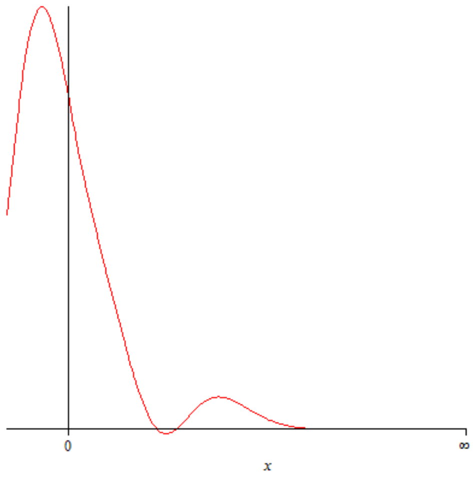

Figure (c), $\sigma = 0.5$

Figure 4.1: The standardized Edgeworth approximation under the lognormal at different values of σ

II. **Student's t distribution**

The probability density function of the $T \sim t(v)$ is

$$f_T(t) = \kappa_v^{-1}\left(1+\frac{t^2}{v}\right)^{-\left(\frac{v+1}{2}\right)}, \quad t \in \Re,$$

Where, $\kappa_v = \sqrt{v}B(1/2, v/2)$. Note $B(a,b) = \dfrac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$.

From (2.2), the Edgeworth asymptotic expansion for the standardized $t(v)$ is

$$f(x) = \phi(x)\left\{1+\frac{1}{4(r-4)}\left(x^4 - 6x^2 + 3\right)\right\} + O(1), \text{ for all } v > 4.$$

The standardized density for T is

$$g(x)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{\pi(v-2)}}\left(1+\frac{x^2}{v-2}\right)^{-\left(\frac{v+1}{2}\right)},\quad x\in\Re,\quad v>2.$$

The Edgeworth expansion for $t(v)$ depends on v and clearly,

$$\lim_{v\to\infty} f(x)=\phi(x)+O(1).$$

To illustrate further the role of v, consider $v=5$, 10 and 20.

Figure 4.2 below shows the poor performance of the Edgeworth series for the standardized density of the t distribution with $v=5$. Here the kurtosis lies outside the validity region.

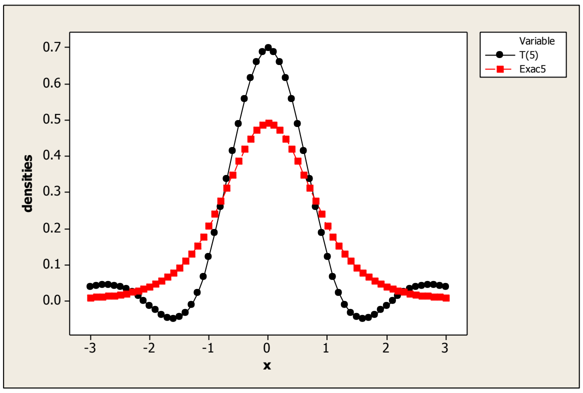

Figure 4.2. The standardized $t(5)$ density and its Edgeworth approximation

Figure 4.3 below shows the good performance of the Edgeworth approximation for the standardized density at $v=10$. Here the skewness and the kurtosis lie inside the validity region.

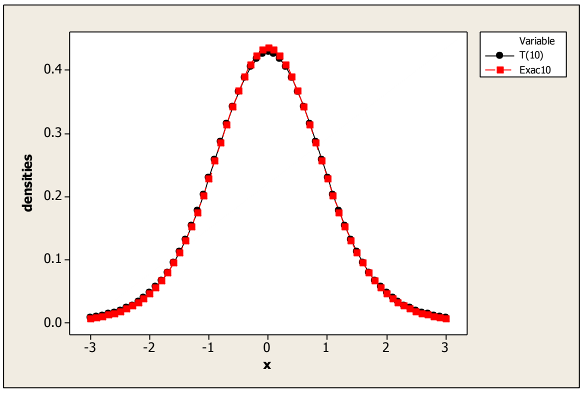

Figure 4.3. The standardized $t\,(10)$ density and its Edgeworth approximation

Similarly, Figure 4.4 below shows the good performance of the Edgeworth approximation to the standardized density at $\nu = 20$. Both the skewness and the kurtosis lie inside the validity region.

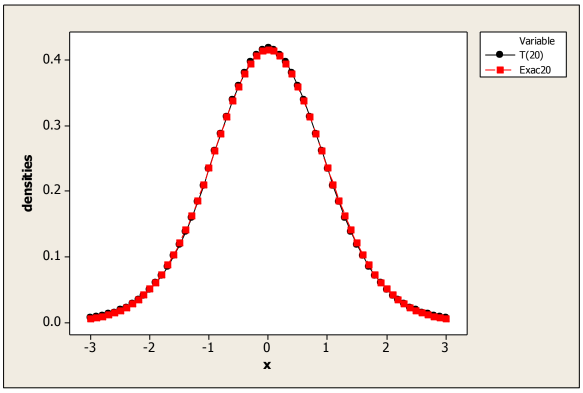

Figure 4.4. The standardized $t\,(20)$ density and its Edgeworth approximation

III. Chi-Squared distribution

The Edgeworth asymptotic expansion for the standardized chi-square distribution with r degrees of freedom is

$$f\left(x\right)=\phi\left(x\right)\left\{1+\frac{\sqrt{2}}{3}\frac{x\left(x^{2}-3\right)}{\sqrt{r}}-\frac{1}{6r}+\frac{x^{2}\left(2x^{4}-21x^{2}+36\right)}{18r}\right\}+O\left(1\right).$$

The Edgeworth expansion of the chi-square distribution depends on the degrees of freedom. Note also that $\lim_{r\to\infty} f(x) = \phi(x) + O(1)$.

To illustrate the effect of increasing the degrees of freedom r on the accuracy of the Edgeworth approximation, we take $r = 2$, 5, 10 and 50.

Clearly, at $r = 2$ yields the case of an exponential distribution with mean two, and the Edgeworth expansion for this case is

$$f(x) = \phi(x)\left\{\frac{1}{18}x^6 - \frac{7}{12}x^4 + \frac{1}{3}x^3 + x^2 - x + \frac{11}{12}\right\} + O(1).$$

Figure 4.5 below shows poor performance of the standardized Edgeworth density under the exponential distribution due to the sharp values of both the skewness and kurtosis, where both the skewness and kurtosis, $\gamma = 2$, $\beta = 9$, are outside the validity region.

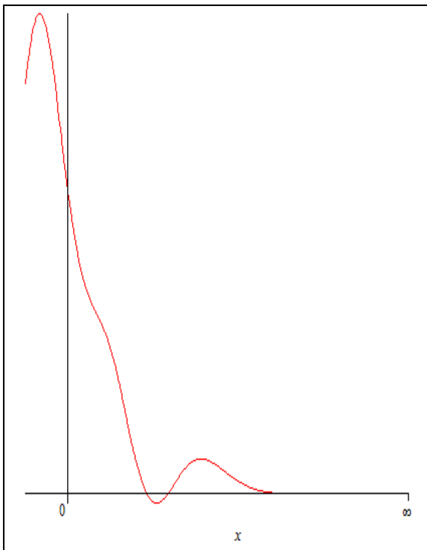

Figure (a), r = 2

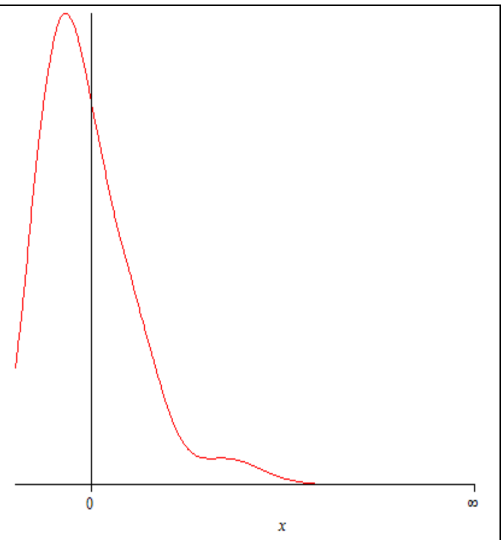

Figure (b), r = 5

Figure 4.5. The standardized Edgeworth approximation under the chi-squared with $r = 2$ and 5.

At $r=5$, this case leads to a better performance than $r=2$. Note that $\gamma=1.265$, $\beta=5.4$ both of which are outside the validity region. However, the value of the kurtosis is close to the upper bound specified in the validity region.

Figure 4.6 (a) shows the performance at $r=10$, where $\gamma=0.8943$ and $\beta=4.2$. This is a nice case where the skewness is outside the validity region but the kurtosis is inside the validity region. Nevertheless the skewness is not large. Figure 4.6 (b) shows excellent performance, since both the skewness and kurtosis are inside the validity region.

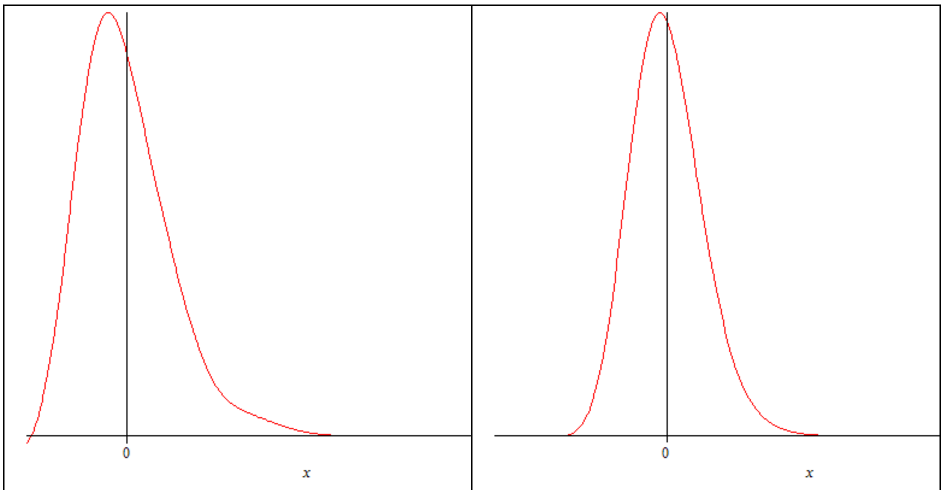

Figure (a), $r=10$

Figure (b), $r=50$

Figure 4.6. The standardized Edgeworth approximation with 10 and 50 degrees of freedom

As a conclusion, the Edgeworth approximations improve as r increases.

CONCLUSIONS

From the results of previous sections, we deduce that Edgeworth distribution is easily handelled and yields good approximation under its validity region. We find the Edgeworth option pricing formula which indeed enhances the theory of black-Scholes model. For further studies, one could test our formula using empirical data.

REFERENCES

[1] Balitskaya, E. O., Zolotuhina, L. A. 1988. On the representation of a density by an Edqeworth series. Biometrika, **75**, pp. 185-187.

[2] Barndorff-Nielsen, O. E., Cox, D. R. 1989. Asymptotic Techniques for Use in Statistics. Chapman and Hall, London.

[3] Barton, D., and K. Dennis 1952. The Conditions Under Which Gram-Charlier and Edgeworth Curves are positive definite and unimodal. Biometrika, 39, pp. 425-428.

[4] Bhattacharya, R. N., Ghosh, J. K. 1978. On the validity of the formal Edgeworth expansion. Ann. Statist. **6**, pp. 434-451.

[5] Black F., Scholes M. 1973. The pricing of options and corporate liabilities. J Polit Econom, 81 pp. 637–654.

[6] Blinnikov, S., Moessner, R. 1998. Expansions for nearly Gaussian distributions. Astron. Astrophys. Suppl. Ser. 130, pp. 193-205.

[7] Chriss, N. A. 1997. *Black-Scholes and Beyond Option Pricing Models*. McGraw-Hill

[8] Corrado, C. J., Su, T. 1996. Skewness and Kurtosis in S&P 500 index returns implied by option prices. J. Financial Research, XIX, 2, pp. 175-192.

[9] DasGupta, A. 2008. Asymptotic Theory of Statistics and Probability. Springer Sciences and Business Media, Berlin.

[10] Draper N. R., Tierney, D. E. 1972. Regions of positive and unimodal series expansion of the Edgeworth and Gram-Charlier approximations. Biometrika, **59**, pp. 463-465.

[11] Edgeworth, F. Y. 1905. The Law of Errors. *Cambridge Philos. Trans*., 20, pp. 36-66.

[12] Filho, R. Gabriel Balieriro, Rosenfeld R, 2004. Testing option pricing with the Edgeworth expansion. Physica A., 344, pp. 484-490.

[13] Galloway, J. G., and McCune, E. D. 1983. A New Edgeworth-Type Expansion. The Texas Journal of Science, **35**, 3, pp. 221-229.

[14] Grundy, B. D. 1991. Option prices and the underlying asset's return distribution. Journal of Finance, Volume XLVI, 3.

[15] Hall, P. 1992. *The Bootstrap and Edgeworth Expansion*. Springer, New York.

[16] Hull, J. C. 2009. Options, Futures and Other Derivatives. Seventh Edition, Pearson, Prentice Hall.

[17] Jarrow, R., Rudd, A. 1982. Approximate option valuation for arbitrary stochastic processes. Journal of financial economics, **10**, North-Holland publishing company, pp. 347-369.

[18] Jiang, Jiming 2010. *Large Sample Techniques for Statistics*. Springer, NewYork.

[19] Johnson, N. L., Kotz, S., and Balakrishnan, N. 1994. *Continuous Univariate Distributions VI*. Wiley, New York.

[20] Karatzas I., Shreve, S. 1988. *Brownian motion and Stochastic Calculus*. Springer

[21] Rachev, S. T., Fabozzi, F. J., Menn, C., 2005. *Fat-Tailed and skewed asset return distributions: Implications for risk management, partfolio selection, and option pricing*. Frank J. Fabozzi Series, Wiley Finance.

[22] Rolski T., Schmidli, H., Schmidt, V, and Teugels, J. 1999. *Stochastic Processes for Insurrance and Finance*. Wiley

[23] Tiwari, M., Saurabha, R. (2007). Empirical study of the effect of including skewness and kurtosis in Black-scholes option pricing formula on S&P CNX Nifty index options. *Munich Personal RePec Archieve.* Munich University Library.

[24] Ugur, O. 2009. *An Introduction to Computational Finance*. Series in Quantitative Finance Vol.I, Imperial College Press.

[25] Wallace, D. L. 1958. Asymptotic approximations to distributions. *Ann. Math. Statist.* **29**, 3, pp. 635-654.

(A. Yousef) KUWAIT UNIVERSITY, KUWAIT

E-mail address: dr.ayousef@aol.com

Vectorial Integral operator convexity inequalities on Time Scales

George A. Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Abstract

Here we obtain a wide range of vectorial integral operator general inequalitites on time scales using convexity. Our treatment is combined by using the diamond-alpha integral. When that fails in the fractional setting we employ the delta and nabla integrals. We give many interesting applications.

2010 Mathematics Subject Classification: 26B25, 26D15, 26A33, 39A12, 93C70.

Key words and phrases: time scales, diamond alpha integral, time scales fractional Riemann-Liouville integral, time scales fractional derivative, vectorial time scales integral operator.

1 Background

We start with the definition of the Riemann-Liouville fractional integrals, see [22]. Let $[a, b]$, $(-\infty < a < b < \infty)$ be a finite interval on the real axis $\mathbb{R}$. The Riemann-Liouville fractional integrals $I_{a+}^{\alpha} f$ and $I_{b-}^{\alpha} f$ of order $\alpha > 0$ are defined by

$$\left(I_{a+}^{\alpha} f\right)(x) = \frac{1}{\Gamma(\alpha)} \int_a^x f(t)(x-t)^{\alpha-1} dt, \quad (x > a), \tag{1}$$

$$\left(I_{b-}^{\alpha} f\right)(x) = \frac{1}{\Gamma(\alpha)} \int_x^b f(t)(t-x)^{\alpha-1} dt, \quad (x < b), \tag{2}$$

respectively. Here $\Gamma(\alpha)$ is the Gamma function. These integrals are called the left-sided and the right-sided fractional integrals. We mention a basic property of the operators $I_{a+}^{\alpha} f$ and $I_{b-}^{\alpha} f$ of order $\alpha > 0$, see also [26]. The result says

1

that the fractional integral operators $I_{a+}^{\alpha}f$ and $I_{b-}^{\alpha}f$ are bounded in $L_p(a,b)$, $1 \le p \le \infty$, that is

$$\left\|I_{a+}^{\alpha}f\right\|_p \le K\left\|f\right\|_p \quad , \quad \left\|I_{b-}^{\alpha}f\right\|_p \le K\left\|f\right\|_p \tag{3}$$

where

$$K = \frac{(b-a)^{\alpha}}{\alpha\Gamma(\alpha)}. \tag{4}$$

Inequality (3), that is the result involving the left-sided fractional integral, was proved by H. G. Hardy in one of his first papers, see [18]. He did not write down the constant, but the calculation of the constant was hidden inside his proof.

So we are motivated by (3), and also [5], [7], [8], [21], [2], and we will prove analogous properties on Time Scales. But first we need some background on Time Scales, see also [12].

A time scale $\mathbb{T}$ is an arbitrary nonempty closed subset of the real numbers. The time scales calculus was initiated by S. Hilger in his PhD thesis in order to unify discrete and continuous analysis [19, 20]. Let $\mathbb{T}$ be a time scale with the topology that it inherits from the real numbers. For $t \in \mathbb{T}$, we define the forward jump operator $\sigma : \mathbb{T} \to \mathbb{T}$ by

$$\sigma(t) = \inf\{s \in \mathbb{T} : s > t\}, \tag{5}$$

and the backward jump operator $\rho : \mathbb{T} \to \mathbb{T}$ by

$$\rho(t) = \sup\{s \in \mathbb{T} : s < t\}. \tag{6}$$

If $\sigma(t) > t$ we say that t is right-scattered, while if $\rho(t) < t$ we say that t is left-scattered. Points that are simultaneously right-scattered and left-scattered are said to be isolated. If $\sigma(t) = t$, then t is called right-dense; if $\rho(t) = t$, then t is called left-dense. The mappings $\mu, \nu : \mathbb{T} \to [0, +\infty)$ defined by $\mu(t) := \sigma(t) - t$ and $\nu(t) := t - \rho(t)$ are called, respectively, the forward and backward graininess function.

Given a time scale $\mathbb{T}$, we introduce the sets $\mathbb{T}^k$, $\mathbb{T}_k$, and $\mathbb{T}_k^k$ as follows. If $\mathbb{T}$ has a left-scattered maximum t_1, then $\mathbb{T}^k = \mathbb{T} - \{t_1\}$, otherwise $\mathbb{T}^k = \mathbb{T}$. If $\mathbb{T}$ has a right-scattered minimum t_2, then $\mathbb{T}_k = \mathbb{T} - \{t_2\}$, otherwise $\mathbb{T}_k = \mathbb{T}$. Finally, $\mathbb{T}_k^k = \mathbb{T}^k \cap \mathbb{T}_k$.

Let $f : \mathbb{T} \to \mathbb{R}$ be a real valued function on a time scale $\mathbb{T}$. Then, for $t \in \mathbb{T}^k$, we define $f^{\Delta}(t)$ to be the number, if one exists, such that for all $\epsilon > 0$, there is a neighborhood U of t such that for all $s \in U$,

$$\left|f(\sigma(t)) - f(s) - f^{\Delta}(t)(\sigma(t) - s)\right| \le \epsilon\left|\sigma(t) - s\right|. \tag{7}$$

We say that f is delta differentiable on $\mathbb{T}^k$ provided $f^{\Delta}(t)$ exists for all $t \in \mathbb{T}^k$. Similarly, for $t \in \mathbb{T}_k$ we define $f^{\nabla}(t)$ to be the number, if one exists, such that for all $\epsilon > 0$, there is a neighborhood V of t such that for all $s \in V$

$$\left| f(\rho(t)) - f(s) - f^{\nabla}(t)(\rho(t) - s) \right| \leq \epsilon |\rho(t) - s| . \tag{8}$$

We say that f is nabla differentiable on $\mathbb{T}_k$, provided that $f^{\nabla}(t)$ exists for all $t \in \mathbb{T}_k$.

For $f : \mathbb{T} \to \mathbb{R}$ we define the function $f^{\sigma} : \mathbb{T} \to \mathbb{R}$ by $f^{\sigma}(t) = f(\sigma(t))$ for all $t \in \mathbb{T}$, that is $f^{\sigma} = f \circ \sigma$. Similarly, we define the function $f^{\rho} : \mathbb{T} \to \mathbb{R}$ by $f^{\rho}(t) = f(\rho(t))$ for all $t \in \mathbb{T}$, that is, $f^{\rho} = f \circ \rho$.

A function $f : \mathbb{T} \to \mathbb{R}$ is called rd-continuous, provided it is continuous at all right-dense points in $\mathbb{T}$ and its left-sided limits finite at all left-dense points in $\mathbb{T}$. A function $f : \mathbb{T} \to \mathbb{R}$ is called ld-continuous, provided it is continuous at all left-dense points in $\mathbb{T}$ and its right-sided limits finite at all right-dense points in $\mathbb{T}$.

A function $F : \mathbb{T} \to \mathbb{R}$ is called a delta antiderivative of $f : \mathbb{T} \to \mathbb{R}$ provided that $F^{\Delta}(t) = f(t)$ holds for all $t \in \mathbb{T}^k$. Then the delta integral of f is defined by

$$\int_a^b f(t) \Delta t = F(b) - F(a) . \tag{9}$$

A function $G : \mathbb{T} \to \mathbb{R}$ is called a nabla antiderivative of $g : \mathbb{T} \to \mathbb{R}$, provided $G^{\nabla}(t) = g(t)$ holds for all $t \in \mathbb{T}_k$. Then the nabla integral of g is defined by $\int_a^b g(t) \nabla t = G(b) - G(a)$. For more details on time scales one can see [1, 12, 13].

Now we describe the diamond-α derivative and integral, referring the reader to [23, 25, 27, 28, 29, 30] for more on this calculus.

Let $\mathbb{T}$ be a time scale and f differentiable on $\mathbb{T}$ in the Δ and ∇ senses. For $t \in \mathbb{T}_k^k$ we define the diamond-α dynamic derivative $f^{\diamond^{\alpha}}(t)$ by

$$f^{\diamond^{\alpha}}(t) = \alpha f^{\Delta}(t) + (1 - \alpha) f^{\nabla}(t), \quad 0 \leq \alpha \leq 1. \tag{10}$$

Thus, f is diamond-α differentiable if and only if f is Δ and ∇ differentiable. The diamond-α derivative reduces to the standard Δ derivative for $\alpha = 1$, or the standard ∇ derivative for $\alpha = 0$. Also, it gives a "weighted derivative" for $\alpha \in (0, 1)$. Diamond-α derivatives have shown in computational experiments to provided efficient and balanced approximation formulae, leading to the design of more reliable numerical methods [27, 28].

Let $f, g : \mathbb{T} \to \mathbb{R}$ be diamond-α differentiable at $t \in \mathbb{T}_k^k$. Then,

(i) $f \pm g : \mathbb{T} \to \mathbb{R}$ is diamond-α differentiable at $t \in \mathbb{T}_k^k$ with

$$(f \pm g)^{\diamond^{\alpha}}(t) = (f)^{\diamond^{\alpha}}(t) \pm (g)^{\diamond^{\alpha}}(t) . \tag{11}$$

(ii) For any constant c, $cf : \mathbb{T} \to \mathbb{R}$ is diamond-α differentiable at $t \in \mathbb{T}_k^k$ with

$$(cf)^{\diamond^{\alpha}}(t) = c(f)^{\diamond^{\alpha}}(t) . \tag{12}$$

(iii) $fg : \mathbb{T} \to \mathbb{R}$ is diamond-α differentiable at $t \in \mathbb{T}_k^k$ with

$$(fg)^{\diamondsuit^\alpha}(t) = (f)^{\diamondsuit^\alpha}(t)\, g(t) + \alpha f^\sigma(t)\, g^\Delta(t) + (1-\alpha)\, f^\rho(t)\, g^\nabla(t). \tag{13}$$

Let $a, t \in \mathbb{T}$, and $h : \mathbb{T} \to \mathbb{R}$. Then, the diamond-$\alpha$ integral from a to t of h is defined by

$$\int_a^t h(\tau)\, \diamondsuit_\alpha \tau = \alpha \int_a^t h(\tau)\, \Delta\tau + (1-\alpha) \int_a^t h(\tau)\, \nabla\tau, \quad 0 \le \alpha \le 1. \tag{14}$$

We may notice the absence of an anti-derivative for the $\diamondsuit_\alpha$ combined derivative. For $t \in \mathbb{T}_k^k$, in general

$$\left(\int_a^t h(\tau)\, \diamondsuit_\alpha \tau \right)^{\diamondsuit^\alpha} \ne h(t). \tag{15}$$

Although the fundamental theorem of calculus does not hold for the $\diamondsuit_\alpha$-integral, other properties hold true. Let $a, b, t \in \mathbb{T}$, $c \in \mathbb{R}$. Then,

$$\text{(i)} \quad \int_a^t \{f(\tau) \pm g(\tau)\}\, \diamondsuit_\alpha \tau = \int_a^t f(\tau)\, \diamondsuit_\alpha \tau \pm \int_a^t g(\tau)\, \diamondsuit_\alpha \tau; \tag{16}$$

$$\text{(ii)} \quad \int_a^t c f(\tau)\, \diamondsuit_\alpha \tau = c \int_a^t f(t)\, \diamondsuit_\alpha \tau;$$

$$\text{(iii)} \quad \int_a^t f(\tau)\, \diamondsuit_\alpha \tau = \int_a^b f(\tau)\, \diamondsuit_\alpha \tau + \int_b^t f(\tau)\, \diamondsuit_\alpha \tau;$$

$$\text{(iv)} \quad \text{If } f(t) \ge 0 \text{ for all } t, \text{ then } \int_a^b f(t)\, \diamondsuit_\alpha t \ge 0;$$

$$\text{(v)} \quad \text{If } f(t) \le g(t) \text{ for all } t, \text{ then } \int_a^b f(t)\, \diamondsuit_\alpha t \le \int_a^b g(t)\, \diamondsuit_\alpha t;$$

$$\text{(vi)} \quad \text{If } f(t) \ge 0 \text{ for all } t, \text{ then } f(t) \equiv 0 \text{ if and only if } \int_a^b f(t)\, \diamondsuit_\alpha t = 0;$$

$$\text{(vii)} \quad \int_a^b c \diamondsuit_\alpha t = c(b-a);$$

$$\text{(viii)} \quad \left| \int_a^b f(t)\, \diamondsuit_\alpha t \right| \le \int_a^b |f(t)|\, \diamondsuit_\alpha t.$$

We mention

Theorem 1 *(multivariate Jensen inequality, see also [15, p. 76], [24]) Let f be a convex function defined on an open and convex subset $C \subseteq \mathbb{R}^n$, and let $X = (X_1, ..., X_n)$ be a random vector such that Probability $(X \in C) = 1$. Assume also $E(|X|)$, $E(|f(X)|) < \infty$, E stands for the expectation. Then $EX \in C$, and*

$$f(EX) \le Ef(X). \tag{17}$$

We would use Jense's diamond-α inequality

Theorem 2 *([14], Jensen's inequality with multiple variables). Tet $\mathbb{T}$ a time scale, $S \subseteq \mathbb{R}^n$ open and convex, $a, b \in \mathbb{T}$ and $f : S \to \mathbb{R}$ a continuous convex function. Let $h, g_1, ..., g_n \in C(\mathbb{T}, \mathbb{R})$ such that $\int_a^b |h(t)| \diamondsuit_\alpha t > 0$ and $g_1([a,b]) \times ... \times g_n([a,b]) \subset S$. Then*

$$f\left(\frac{\int_a^b |h(t)| g_1(t) \diamondsuit_\alpha t}{\int_a^b |h(t)| \diamondsuit_\alpha t}, ..., \frac{\int_a^b |h(t)| g_n(t) \diamondsuit_\alpha t}{\int_a^b |h(t)| \diamondsuit_\alpha t}\right) \leq \frac{\int_a^b |h(t)| f(g_1(t), ..., g_n(t)) \diamondsuit_\alpha t}{\int_a^b |h(t)| \diamondsuit_\alpha t}. \tag{18}$$

Remark 3 *i) By [9] and [17] we conclude that the multivariate Jensen's inequality is valid for the delta-Δ and nabla-∇ Lebesgue integrable functions and measures, respectively.*

ii) Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$, we call it increasing per coordinate, iff whenever $x_i \leq y_i$, $i = 1, ..., n$; $x_i, y_i \in \mathbb{R}_+$, then $\Phi(x_1, ..., x_n) \leq \Phi(y_1, ..., y_n)$.

In [5], we proved that if $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ is convex and increasing per coordinate, then Φ is continuous.

We can extend Φ to $\Phi : \mathbb{R}^n \to \mathbb{R}$ and still be convex, by defining

$$\Phi(x_1, ..., x_j, ..., x_n) := \Phi(x_1, ..., -x_j, ..., x_n), \tag{19}$$

for any $x_j < 0$, $j \in \{1, ..., n\}$, see Lemma 10.

Hence we can apply Jensen's inequality using Φ on $\mathbb{R}^n$. It is well known, that a convex function on an open and convex subset of $\mathbb{R}^n$ is continuous.

We further need

Theorem 4 *(Hölder's Inequality, see [16]) For continuous functions $f, g : [a,b]_{\mathbb{T}} \to \mathbb{R}$, we have:*

$$\int_a^b |f(t) g(t)| \diamondsuit_\alpha t \leq \left[\int_a^b |f(t)|^p \diamondsuit_\alpha t\right]^{\frac{1}{p}} \left[\int_a^b |g(t)|^q \diamondsuit_\alpha t\right]^{\frac{1}{q}}, \tag{20}$$

where $p > 1$, and $q = \frac{p}{p-1}$.

We obtain

Theorem 5 *(Generalization of Hölder's inequality) Let $f_i \in C([a,b]_{\mathbb{T}}, \mathbb{R})$, $i = 1, ..., n$, and $p_i > 1$ such that $\sum_{i=1}^n \frac{1}{p_i} = 1$. Then*

$$\int_a^b \prod_{i=1}^n |f_i(t)| \diamondsuit_\alpha t \leq \prod_{i=1}^n \left(\int_a^b |f_i(t)|^{p_i} \diamondsuit_\alpha t\right)^{\frac{1}{p_i}}. \tag{21}$$

Proof. Using (20) and induction hypothesis, exactly as in [31]. ■

Comment 6 *By Tietze's extension theorem of General Topology we easily derive that a continuous function f of $\prod_{i=1}^{n}([a_i,b_i]\cap\mathbb{T}_i)$ (where $\mathbb{T}_i$, $i=1,...,n\in\mathbb{N}$ are time scales) is bounded, since its continuous extension F on $\prod_{i=1}^{n}[a_i,b_i]\subseteq\mathbb{R}^n$ is bounded, $n\in\mathbb{N}$.*

Comment 7 *It is regarding the univariate functions. Based on [17] we see that the Cauchy Time scales delta Δ and nabla ∇ integrals are equal to definite Riemann Time Scales Δ, ∇ integrals, respectively. Thus, the diamond-α-Cauchy integral (14) is a diamond α-Riemann integral over continuous functions. Of course the last integral exists, since continuous functions are Riemann Δ and ∇-integrable, and it is equal to the corresponding α-Lebesgue integral, by [17].*

In particular the dominated and bounded convergence theorems hold true with respect to the Lebesgue-Δ, ∇ measures.

Comment 8 *Let $\mathbb{T}_1$, $\mathbb{T}_2$ be time scales and $f:[a,b]_{\mathbb{T}_1}\times[c,d]_{\mathbb{T}_2}\to\mathbb{R}$ be continuous. By [9] and [10] we get that f is Riemann Δ and ∇-integrable over $[a,b)_{\mathbb{T}_1}\times[c,d)_{\mathbb{T}_2}$ and $(a,b]_{\mathbb{T}_1}\times(c,d]_{\mathbb{T}_2}$, respectively. Hence by [9], [10], f is Lebesgue Δ and ∇-integrable there.*

Thus by Fubini's theorem we get

$$\int_a^b\left(\int_c^d f(x,y)\,\Delta y\right)\Delta x=\int_c^d\left(\int_a^b f(x,y)\,\Delta x\right)\Delta y, \tag{22}$$

and

$$\int_a^b\left(\int_c^d f(x,y)\,\nabla y\right)\nabla x=\int_c^d\left(\int_a^b f(x,y)\,\nabla x\right)\nabla y. \tag{23}$$

We define ($\alpha\in[0,1]$)

$$\int_a^b\left(\int_c^d f(x,y)\,\diamondsuit_\alpha y\right)\diamondsuit_\alpha x=$$

$$\alpha\int_a^b\left(\int_c^d f(x,y)\,\Delta y\right)\Delta x+(1-\alpha)\int_a^b\left(\int_c^d f(x,y)\,\nabla y\right)\nabla x. \tag{24}$$

One can generalize (24) for multiple integrals.

So for f continuous we get the $\diamondsuit_\alpha$-Fubini's theorem main property:

$$\int_a^b\left(\int_c^d f(x,y)\,\diamondsuit_\alpha y\right)\diamondsuit_\alpha x=\int_c^d\left(\int_a^b f(x,y)\,\diamondsuit_\alpha x\right)\diamondsuit_\alpha y. \tag{25}$$

We make

Remark 9 *Let $\mathbb{T}_1$, $\mathbb{T}_2$ be time scales and $f : [a,b]_{\mathbb{T}_1} \times [c,d]_{\mathbb{T}_2} \to \mathbb{R}$ be continuous. Consider*

$$g(x) = \int_c^d f(x,y) \diamondsuit_\alpha y$$

$$= \alpha \int_c^d f(x,y) \Delta y + (1-\alpha) \int_c^d f(x,y) \nabla y, \tag{26}$$

$\alpha \in [0,1]$, $\forall\ x \in [a,b]_{\mathbb{T}_1}$.

We prove that g is continuous on $[a,b]_{\mathbb{T}_1}$. Let $x_n \to x$, where $\{x_n\}_{n\in\mathbb{N}}$, $x \in [a,b]_{\mathbb{T}_1}$ then $f(x_n,y) \to f(x,y)$, as $n \to \infty$, $\forall\ y \in [c,d]_{\mathbb{T}_2}$. Furthermore there exists $M > 0$ such that $|f(x_n,y)|$, $|f(x,y)| \leq M$, $\forall\ y \in [c,d]_{\mathbb{T}_2}$. Hence by Lebesgue's bounded convergence theorem ([9]) we get that

$$\lim_{n\to\infty} \int_c^d f(x_n,y) \Delta y = \int_c^d f(x,y) \Delta y, \tag{27}$$

and

$$\lim_{n\to\infty} \int_c^d f(x_n,y) \nabla y = \int_c^d f(x,y) \nabla y. \tag{28}$$

Combining (27) and (28) we obtain $g(x_n) \to g(x)$, as $n \to \infty$, proving the continuity of g.

For completeness we give

Lemma 10 *Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ increasing per coordinate and convex. We extend Φ on $\mathbb{R}^n$ by defining $\Phi(x_1,...,x_j,...,x_n) := \Phi(x_1,...,-x_j,...,x_n)$, for any $x_j < 0$, $j \in \{1,...,n\}$. Then Φ is convex on $\mathbb{R}^n$.*

Proof. Let $0 \leq \lambda \leq 1$, $(x_1,...,x_n), (y_1,...,y_n) \in \mathbb{R}^n$. Then we observe that

$$\Phi(\lambda(x_1,...,x_n) + (1-\lambda)(y_1,...,y_n)) =$$

$$\Phi(\lambda x_1 + (1-\lambda) y_1, ..., \lambda x_n + (1-\lambda) y_n) =$$

$$\Phi(|\lambda x_1 + (1-\lambda) y_1|, ..., |\lambda x_n + (1-\lambda) y_n|) \leq$$

$$\Phi(\lambda |x_1| + (1-\lambda)|y_1|, ..., \lambda |x_n| + (1-\lambda)|y_n|) =$$

$$\Phi(\lambda(|x_1|,...,|x_n|) + (1-\lambda)(|y_1|,...,|y_n|)) \leq$$

$$\lambda\Phi(|x_1|,...,|x_n|) + (1-\lambda)\Phi(|y_1|,...,|y_n|) =$$

$$\lambda\Phi(x_1,...,x_n) + (1-\lambda)\Phi(y_1,...,y_n),$$

proving the claim. ■

General Notation 11 *Let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be measure spaces with positive σ-finite measures, and let $k : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a nonnegative measurable function, $k(x, \cdot)$ measurable on Ω_2 and*

$$K(x) = \int_{\Omega_2} k(x, y)\, d\mu_2(y), \quad x \in \Omega_1. \tag{29}$$

We suppose that $K(x) > 0$ a.e. on Ω_1, and by a weight function u (shortly: a weight), we mean a nonnegative measurable function on the actual set. Let the measurable functions $g_i : \Omega_1 \to \mathbb{R}$, $i = 1, ..., n$, with the representation

$$g_i(x) = \int_{\Omega_2} k(x, y)\, f_i(y)\, d\mu_2(y), \tag{30}$$

where $f_i : \Omega_2 \to \mathbb{R}$ are measurable functions, $i = 1, ..., n$.

Denote by $\overrightarrow{x} = x := (x_1, ..., x_n) \in \mathbb{R}^n$, $\overrightarrow{g} := (g_1, ..., g_n)$ and $\overrightarrow{f} := (f_1, ..., f_n)$.

We consider here $\Phi : \mathbb{R}^n_+ \to \mathbb{R}$ a convex function, which is increasing per coordinate.

Example 12 *(for Φ).*

1) Given g_i is convex and increasing on $\mathbb{R}_+$, then $\Phi(x_1, ..., x_n) := \sum_{i=1}^n g_i(x_i)$ is convex on $\mathbb{R}^n_+$, and increasing per coordinate; the same properties hold for:

2) $\|x\|_p = \left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}}$, $p \geq 1$,

3) $\|x\|_\infty = \max\limits_{i \in \{1,...,n\}} x_i$,

4) $\sum_{i=1}^n x_i^2$,

5) $\sum_{i=1}^n \left(i \cdot x_i^2\right)$,

6) $\sum_{i=1}^n \sum_{j=1}^i x_j^2$,

7) $\ln\left(\sum_{i=1}^n e^{x_i}\right)$,

8) let g_j are convex and increasing per coordinate on $\mathbb{R}^n_+$, then so is $\sum_{j=1}^m e^{g_j(x)}$, and so is $\ln\left(\sum_{j=1}^m e^{g_j(x)}\right)$, $x \in \mathbb{R}^n_+$.

General Notation 13 *From now on we may write*

$$\overrightarrow{g}(x) = \int_{\Omega_2} k(x, y)\, \overrightarrow{f}(y)\, d\mu_2(y), \tag{31}$$

which means

$$(g_1(x), ..., g_n(x)) = \left(\int_{\Omega_2} k(x, y)\, f_1(y)\, d\mu_2(y), ..., \int_{\Omega_2} k(x, y)\, f_n(y)\, d\mu_2(y)\right). \tag{32}$$

Similarly, we may write

$$\left|\overrightarrow{g}(x)\right| = \left|\int_{\Omega_2} k(x, y)\, \overrightarrow{f}(y)\, d\mu_2(y)\right|, \tag{33}$$

and we would mean

$$(|g_1(x)|,...,|g_n(x)|) =$$
$$\left(\left|\int_{\Omega_2} k(x,y) f_1(y) d\mu_2(y)\right|,...,\left|\int_{\Omega_2} k(x,y) f_n(y) d\mu_2(y)\right|\right). \quad (34)$$

We also can write that

$$|\overrightarrow{g}(x)| \leq \int_{\Omega_2} k(x,y) \left|\overrightarrow{f}(y)\right| d\mu_2(y), \quad (35)$$

and we mean the fact that

$$|g_i(x)| \leq \int_{\Omega_2} k(x,y) |f_i(y)| d\mu_2(y), \quad (36)$$

for all $i = 1,...,n$. Similarly for other properties.

General Notation 14 *Next let $(\Omega_1, \Sigma_1, \mu_1)$ and $(\Omega_2, \Sigma_2, \mu_2)$ be measure spaces with positive σ-finite measures, and let $k_j : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a nonnegative measurable function, $k_j(x, \cdot)$ measurable on Ω_2 and*

$$K_j(x) = \int_{\Omega_2} k_j(x,y) d\mu_2(y), \quad x \in \Omega_1, j = 1,...,m. \quad (37)$$

We suppose that $K_j(x) > 0$ a.e. on Ω_1.

Let the measurable functions $g_{ji} : \Omega_1 \to \mathbb{R}$ with the representation

$$g_{ji}(x) = \int_{\Omega_2} k_j(x,y) f_{ji}(y) d\mu_2(y), \quad (38)$$

where $f_{ji} : \Omega_2 \to \mathbb{R}$ are measurable functions, $i = 1,...,n$ and $j = 1,...,m$.

Denote the function vectors $\overrightarrow{g_j} := (g_{j1}, g_{j2}, ..., g_{jn})$ and $\overrightarrow{f_j} := (f_{j1}, ..., f_{jn})$, $j = 1,...,m$.

We say $\overrightarrow{f_j}$ is integrable with respect to measure μ, iff all f_{ji} are integrable with respect to μ.

We consider here $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1,...,m$, convex functions that are increasing per coordinate.

Again u is a weight function on Ω_1.

2 Main Results

We present vectorial inequalities on $\diamondsuit_\alpha$- integral operators.

Theorem 15 *Let $\mathbb{T}_1$, $\mathbb{T}_2$ be time scales, $a, b \in \mathbb{T}_1$; $c, d \in \mathbb{T}_2$; $k(x,y)$ is a kernel function with $x \in [a,b]_{\mathbb{T}_1}$, $y \in [c,d]_{\mathbb{T}_2}$; k is continuous function from $[a,b]_{\mathbb{T}_1} \times [c,d]_{\mathbb{T}_2}$ into $\mathbb{R}_+$. Consider*

$$K(x) := \int_c^d k(x,y) \diamondsuit_\alpha y, \quad \forall\, x \in [a,b]_{\mathbb{T}_1}. \quad (39)$$

We assume that $K(x) > 0$, $\forall\ x \in [a,b]_{\mathbb{T}_1}$. *Consider* $f_i : [c,d]_{\mathbb{T}_2} \to \mathbb{R}$ *continuous, and the* $\diamondsuit_\alpha$*-integral operator function*

$$g_i(x) := \int_c^d k(x,y) f_i(y) \diamondsuit_\alpha y, \tag{40}$$

$\forall\ x \in [a,b]_{\mathbb{T}_1}$, $i = 1,..,n$. *Let* $\overrightarrow{g} := (g_1, ..., g_n)$, $\overrightarrow{f} := (f_1, ..., f_n)$.

Consider also the weight function

$$u : [a,b]_{\mathbb{T}_1} \to \mathbb{R}_+, \tag{41}$$

which is continuous.

Define further the function

$$v(y) := \int_a^b \frac{u(x) k(x,y)}{K(x)} \diamondsuit_\alpha x, \tag{42}$$

$\forall\ y \in [c,d]_{\mathbb{T}_2}$.

Let I *denote any of* $(0,\infty)^n$ *or* $[0,\infty)^n$, *and* $\Phi : I \to \mathbb{R}$ *be a convex and increasing per coordinate function. In particular we assume that*

$$|f_i|\left([c,d]_{\mathbb{T}_2}\right) \subseteq I,\ \ i = 1,...,n. \tag{43}$$

Then

$$\int_a^b u(x) \Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right) \diamondsuit_\alpha x \leq \int_c^d v(y) \Phi\left(\left|\overrightarrow{f(y)}\right|\right) \diamondsuit_\alpha y. \tag{44}$$

Proof. We see that

$$\int_a^b u(x) \Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right) \diamondsuit_\alpha x = \int_a^b u(x) \Phi\left(\frac{1}{K(x)} \left|\int_c^d k(x,y) \overrightarrow{f(y)} \diamondsuit_\alpha y\right|\right) \diamondsuit_\alpha x$$

$$\leq \int_a^b u(x) \Phi\left(\frac{1}{K(x)} \int_c^d k(x,y) \left|\overrightarrow{f(y)}\right| \diamondsuit_\alpha y\right) \diamondsuit_\alpha x \tag{45}$$

(by generalized Jensen's inequality, see Theorem 2 and Comment 7)

$$\leq \int_a^b \frac{u(x)}{K(x)} \left(\int_c^d k(x,y) \Phi\left(\left|\overrightarrow{f(y)}\right|\right) \diamondsuit_\alpha y\right) \diamondsuit_\alpha x$$

$$= \int_a^b \left(\int_c^d \frac{u(x) k(x,y)}{K(x)} \Phi\left(\left|\overrightarrow{f(y)}\right|\right) \diamondsuit_\alpha y\right) \diamondsuit_\alpha x \tag{46}$$

(by (25))

$$= \int_c^d \left(\int_a^b \frac{u(x) k(x,y)}{K(x)} \Phi\left(\left|\overrightarrow{f(y)}\right|\right) \diamondsuit_\alpha x\right) \diamondsuit_\alpha y$$

$$= \int_c^d \Phi\left(\left|\overrightarrow{f(y)}\right|\right)\left(\int_a^b \frac{u(x)\,k(x,y)}{K(x)}\diamondsuit_\alpha x\right)\diamondsuit_\alpha y = \int_c^d v(y)\,\Phi\left(\left|\overrightarrow{f(y)}\right|\right)\diamondsuit_\alpha y, \tag{47}$$

proving the claim. ■

We continue with

Theorem 16 *All as in Theorem 15, however now Φ is not necessarily increasing per coordinate and only from $(0,\infty)^n$ into $\mathbb{R}$. Additionaly we assume that each f_i is of fixed strict sign, $i = 1,...,n$. Then*

$$\int_a^b u(x)\,\Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right)\diamondsuit_\alpha x \leq \int_c^d v(y)\,\Phi\left(\left|\overrightarrow{f(y)}\right|\right)\diamondsuit_\alpha y. \tag{48}$$

Proof. We notice that

$$\left|\overrightarrow{g(x)}\right| = \left|\int_c^d k(x,y)\,\overrightarrow{f(y)}\diamondsuit_\alpha y\right| = \int_c^d k(x,y)\left|\overrightarrow{f(y)}\right|\diamondsuit_\alpha y. \tag{49}$$

Therefore we have

$$\int_a^b u(x)\,\Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right)\diamondsuit_\alpha x = \int_a^b u(x)\,\Phi\left(\frac{1}{K(x)}\left|\int_c^d k(x,y)\,\overrightarrow{f(y)}\diamondsuit_\alpha y\right|\right)\diamondsuit_\alpha x$$

$$= \int_a^b u(x)\,\Phi\left(\frac{1}{K(x)}\int_c^d k(x,y)\left|\overrightarrow{f(y)}\right|\diamondsuit_\alpha y\right)\diamondsuit_\alpha x. \tag{50}$$

The rest follows as in the proof of Theorem 15. ■

Corollary 17 *(to Theorem 16) It holds*

$$\int_a^b u(x)\sum_{i=1}^n \ln\left(\frac{|g_i(x)|}{K(x)}\right)\diamondsuit_\alpha x \geq \int_c^d v(y)\sum_{i=1}^n \ln\left(|f_i(y)|\right)\diamondsuit_\alpha y. \tag{51}$$

Proof. Apply (48) for $\Phi(x) = -\sum_{i=1}^n \ln x_i$, which a convex function with domain $(0,\infty)^n$. ■

Corollary 18 *(to Theorem 15) It holds*

$$\int_a^b u(x)\sum_{i=1}^n e^{\frac{|g_i(x)|}{K(x)}}\diamondsuit_\alpha x \leq \int_c^d v(y)\sum_{i=1}^n e^{|f_i(y)|}\diamondsuit_\alpha y. \tag{52}$$

Proof. Apply (44) for $\Phi(x) = \sum_{i=1}^n e^{x_i}$, $x_i \geq 0$. ■

Notation 19 *Let $\mathbb{T}_1$, $\mathbb{T}_2$ be time scales, $a,b \in \mathbb{T}_1$; $c,d \in \mathbb{T}_2$; $k_j(x,y)$ is a kernel function with $x \in [a,b]_{\mathbb{T}_1}$, $y \in [c,d]_{\mathbb{T}_2}$; k_j is continuous function from $[a,b]_{\mathbb{T}_1} \times [c,d]_{\mathbb{T}_2}$ into $\mathbb{R}_+$ for $j=1,...,m \in \mathbb{N}$. Consider*

$$K_j(x) := \int_c^d k_j(x,y) \diamondsuit_\alpha y, \quad \forall\ x \in [a,b]_{\mathbb{T}_1}, \tag{53}$$

$j=1,...,m$.

We assume that $K_j(x) > 0$, $\forall\ x \in [a,b]_{\mathbb{T}_1}$, $j=1,...,m$. Consider $f_{ji} : [c,d]_{\mathbb{T}_2} \to \mathbb{R}$ continuous, $i=1,...,n$, and $j=1,...,m$, and the $\diamondsuit_\alpha$-integral operator function

$$g_{ji}(x) := \int_c^d k_j(x,y) f_{ji}(y) \diamondsuit_\alpha y, \tag{54}$$

$\forall\ x \in [a,b]_{\mathbb{T}_1}$, $i=1,...,n$, $j=1,...,m$.

Denote the function vectors $\overrightarrow{g_j} := (g_{j1},...,g_{jn})$ and $\overrightarrow{f_j} := (f_{j1},...,f_{jn})$, $j=1,...,m$.

Consider also the weight function

$$u : [a,b]_{\mathbb{T}_1} \to \mathbb{R}_+, \tag{55}$$

which is continuous.

Define further the function

$$\lambda_m(y) := \int_a^b \frac{u(x) \prod_{j=1}^m k_j(x,y)}{\prod_{j=1}^m K_j(x)} \diamondsuit_\alpha x, \tag{56}$$

$\forall\ y \in [c,d]_{\mathbb{T}_2}$.

Here $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j=1,...,m$, are convex and increasing per coordinate functions.

We give

Theorem 20 *All as in Notation 19. Let $\rho \in \{1,...,m\}$ be fixed. Then*

$$\int_a^b u(x) \prod_{j=1}^m \Phi_j\left(\frac{\left|\overrightarrow{g_j(x)}\right|}{K_j(x)}\right) \diamondsuit_\alpha x \leq \tag{57}$$

$$\left(\prod_{\substack{j=1\\ j\neq\rho}}^m \int_c^d \Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right) \diamondsuit_\alpha y\right) \left(\int_c^d \Phi_\rho\left(\left|\overrightarrow{f_\rho(y)}\right|\right) \lambda_m(y) \diamondsuit_\alpha y\right).$$

Proof. We demonstrate the proof for $m = 3$. For general m it follows the same way. Here we use the multivariate Jensen's inequality, see Theorem 2 and Comment 7, $\diamondsuit_\alpha$-Fubini's theorem, see (25), and that Φ_j are increasing per coordinate.

We have

$$\int_a^b u(x) \prod_{j=1}^3 \Phi_j \left(\frac{\left|\overrightarrow{g_j(x)}\right|}{K_j(x)} \right) \diamondsuit_\alpha x =$$

$$\int_a^b u(x) \prod_{j=1}^3 \Phi_j \left(\left| \frac{1}{K_j(x)} \int_c^d k_j(x,y) \overrightarrow{f_j(y)} \diamondsuit_\alpha y \right| \right) \diamondsuit_\alpha x \leq \tag{58}$$

$$\int_a^b u(x) \prod_{j=1}^3 \Phi_j \left(\frac{1}{K_j(x)} \int_c^d k_j(x,y) \left|\overrightarrow{f_j(y)}\right| \diamondsuit_\alpha y \right) \diamondsuit_\alpha x \leq$$

$$\int_a^b u(x) \prod_{j=1}^3 \left(\frac{1}{K_j(x)} \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right) \diamondsuit_\alpha x =$$

$$\int_a^b \left(\frac{u(x)}{\prod_{j=1}^3 K_j(x)} \right) \left(\prod_{j=1}^3 \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right) \diamondsuit_\alpha x =$$

(calling $\theta(x) := \frac{u(x)}{\prod_{j=1}^3 K_j(x)}$)

$$\int_a^b \theta(x) \left(\prod_{j=1}^3 \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right) \diamondsuit_\alpha x = \tag{59}$$

$$\int_a^b \theta(x) \left[\int_c^d \left(\prod_{j=1}^2 \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right) k_3(x,y) \Phi_3\left(\left|\overrightarrow{f_3(y)}\right| \right) \diamondsuit_\alpha y \right] \diamondsuit_\alpha x =$$

$$\int_a^b \left(\int_c^d \theta(x) \left(\prod_{j=1}^2 \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right) k_3(x,y) \Phi_3\left(\left|\overrightarrow{f_3(y)}\right| \right) \diamondsuit_\alpha y \right) \diamondsuit_\alpha x =$$

$$\int_c^d \left(\int_a^b \theta(x) \left(\prod_{j=1}^2 \int_c^d k_j(x,y) \Phi_j\left(\left|\overrightarrow{f_j(y)}\right| \right) \diamondsuit_\alpha y \right)\right.$$

13

$$k_3(x,y)\,\Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha x\Big)\diamondsuit_\alpha y=$$

$$\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\left(\int_a^b \theta(x)\,k_3(x,y)\left(\prod_{j=1}^{2}\int_c^d k_j(x,y)\,\Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right)\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right)\diamondsuit_\alpha y= \tag{60}$$

$$\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\left[\int_a^b \theta(x)\,k_3(x,y)\left(\int_c^d \left\{\int_c^d k_1(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right\}\cdot\right.\right.$$

$$\left.\left.k_2(x,y)\,\Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right]\diamondsuit_\alpha y=$$

$$\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\left[\int_a^b\left(\int_c^d \theta(x)\,k_2(x,y)\,k_3(x,y)\,\Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\cdot\right.\right. \tag{61}$$

$$\left.\left.\left\{\int_c^d k_1(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right\}\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right]\diamondsuit_\alpha y=$$

$$\left(\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha y\right)\left[\int_a^b\left(\int_c^d \theta(x)\,k_2(x,y)\,k_3(x,y)\,\Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\cdot\right.\right.$$

$$\left.\left.\left\{\int_c^d k_1(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right\}\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right]=$$

$$\left(\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha y\right)\left[\int_c^d\left(\int_a^b \theta(x)\,k_2(x,y)\,k_3(x,y)\,\Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\cdot\right.\right.$$

$$\left.\left.\left\{\int_c^d k_1(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right\}\diamondsuit_\alpha x\right)\diamondsuit_\alpha y\right]= \tag{62}$$

$$\left(\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha y\right)\left[\int_c^d \Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\left(\int_a^b \theta(x)\,k_2(x,y)\,k_3(x,y)\cdot\right.\right.$$

$$\left.\left.\left(\int_c^d k_1(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right)\diamondsuit_\alpha y\right]=$$

$$\left(\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha y\right)\left[\int_c^d \Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\left\{\int_a^b\left(\int_c^d \theta(x)\prod_{j=1}^{3}k_j(x,y)\cdot\right.\right.\right.$$

$$\left.\left.\left.\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right\}\diamondsuit_\alpha y\right]=$$

$$\left(\int_c^d \Phi_3\left(\left|\overrightarrow{f_3(y)}\right|\right)\diamondsuit_\alpha y\right)\left(\int_c^d \Phi_2\left(\left|\overrightarrow{f_2(y)}\right|\right)\diamondsuit_\alpha y\right)\cdot$$

$$\left(\int_a^b\left(\int_c^d\theta(x)\prod_{j=1}^3 k_j(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha y\right)\diamondsuit_\alpha x\right)= \tag{63}$$

$$\left(\prod_{j=2}^3\int_c^d\Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right)\diamondsuit_\alpha y\right)\cdot$$

$$\left(\int_c^d\left(\int_a^b\theta(x)\prod_{j=1}^3 k_j(x,y)\,\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\diamondsuit_\alpha x\right)\diamondsuit_\alpha y\right)=$$

$$\left(\prod_{j=2}^3\int_c^d\Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right)\diamondsuit_\alpha y\right)\cdot$$

$$\left(\int_c^d\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\left(\int_a^b\theta(x)\prod_{j=1}^3 k_j(x,y)\diamondsuit_\alpha x\right)\diamondsuit_\alpha y\right)=$$

$$\left(\prod_{j=2}^3\int_c^d\Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right)\diamondsuit_\alpha y\right)\left(\int_c^d\Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right)\lambda_3(y)\diamondsuit_\alpha y\right), \tag{64}$$

proving the claim. ∎

Corollary 21 *(to Theorem 20) It holds*

$$\int_a^b u(x)\prod_{j=1}^m\left(\sum_{i=1}^n e^{\frac{|g_{ji}(x)|}{K_j(x)}}\right)\diamondsuit_\alpha x\leq \tag{65}$$

$$\left(\prod_{\substack{j=1\\j\neq\rho}}^m\int_c^d\sum_{i=1}^n e^{|f_{ji}(y)|}\diamondsuit_\alpha y\right)\left(\int_c^d\left(\sum_{i=1}^n e^{|f_{\rho i}(y)|}\right)\lambda_m(y)\diamondsuit_\alpha y\right).$$

Proof. Apply $\Phi_j(x)=\sum_{i=1}^n e^{x_i}$, $x_i\geq 0$, for all $j=1,...,m$. ∎

We continue with

Theorem 22 *All as in Theorem 20, but now $\Phi_j:(0,\infty)^n\to\mathbb{R}_+$, $j=1,...,m$, are convex and not necessarily increasing per coordinate. Furthermore all f_{ji}, $i=1,...,n$, $j=1,...,m$, are of fixed strict sign. Then (57) is valid.*

Proof. Similar to Theorem 16, and Theorem 20. ∎

We give the following application:

Corollary 23 *All as in Theorem 22, with* $\Phi_j(x) = -\sum_{i=1}^{n} \ln x_i$*,* $j = 1, ..., m \in \mathbb{N}$*. It holds*

$$(-1)^m \int_a^b u(x) \prod_{j=1}^{m} \left(\sum_{i=1}^{n} \ln \left(\frac{|g_{ji}(x)|}{K_j(x)} \right) \right) \diamondsuit_\alpha x \leq \tag{66}$$

$$(-1)^m \left(\prod_{\substack{j=1 \\ j \neq \rho}}^{m} \int_c^d \sum_{i=1}^{n} \ln(|f_{ji}(y)|) \diamondsuit_\alpha y \right) \left(\int_c^d \left(\sum_{i=1}^{n} \ln(|f_{\rho i}(y)|) \right) \lambda_m(y) \diamondsuit_\alpha y \right).$$

Proof. By (57). ■

We continue with

Theorem 24 *All as in Notation 19. Define*

$$u_j(y) := \int_a^b u(x) \frac{k_j(x,y)}{K_j(x)} \diamondsuit_\alpha x, \tag{67}$$

$\forall\, y \in [c,d]_{\mathbb{T}_2}$, $j = 1, ..., m \in \mathbb{N}$.

Let $p_j > 1 : \sum_{j=1}^{m} \frac{1}{p_j} = 1$*. Then*

$$\int_a^b u(x) \prod_{j=1}^{m} \Phi_j \left(\frac{\left| \overrightarrow{g_j(x)} \right|}{K_j(x)} \right) \diamondsuit_\alpha x \leq \tag{68}$$

$$\prod_{j=1}^{m} \left(\int_c^d u_j(y) \Phi_j \left(\left| \overrightarrow{f_j(y)} \right| \right)^{p_j} \diamondsuit_\alpha y \right)^{\frac{1}{p_j}}.$$

Proof. Notice that Φ_j, $j = 1, ..., m$, are continuous functions. Here we use the generalized Hölder's inequality, see Theorem 5. We have

$$\int_a^b u(x) \prod_{j=1}^{m} \Phi_j \left(\frac{\left| \overrightarrow{g_j(x)} \right|}{K_j(x)} \right) \diamondsuit_\alpha x =$$

$$\int_a^b \prod_{j=1}^{m} \left(u(x)^{\frac{1}{p_j}} \Phi_j \left(\frac{\left| \overrightarrow{g_j(x)} \right|}{K_j(x)} \right) \right) \diamondsuit_\alpha x \leq \tag{69}$$

$$\prod_{j=1}^{m} \left(\int_a^b u(x) \Phi_j \left(\frac{\left| \overrightarrow{g_j(x)} \right|}{K_j(x)} \right)^{p_j} \diamondsuit_\alpha x \right)^{\frac{1}{p_j}} \leq$$

(notice here that $\Phi_j^{p_j}$, $j = 1, ..., m$, are convex, increasing per coordinate and continuous, non-negative functions, and by Theorem 15 we get)

$$\prod_{j=1}^{m} \left(\int_c^d u_j(y) \Phi_j(|f_j(y)|)^{p_j} \diamondsuit_\alpha y \right)^{\frac{1}{p_j}}, \tag{70}$$

proving the claim. ■

We also give

Theorem 25 *All as in Theorem 24, but now* $\Phi_j : (0, \infty)^n \to \mathbb{R}_+$, $j = 1, ..., m$, *are convex and not necessarily increasing per coordinate. Furthermore all* f_{ji}, $j = 1, ..., m$, $i = 1, ..., n$, *are each of fixed strict sign. Then (68) is valid.*

Proof. Similar to Theorem 24, and by using Theorem 16. ■

Corollary 26 *(to Theorem 24) It holds*

$$\int_a^b u(x) \prod_{j=1}^m \left\| \frac{\overrightarrow{|g_j(x)|}}{K_j(x)} \right\|_{p_j} \Diamond_\alpha x \leq \tag{71}$$

$$\prod_{j=1}^m \left(\int_c^d u_j(y) \left\| \overrightarrow{f_j(y)} \right\|^{p_j} \Diamond_\alpha y \right)^{\frac{1}{p_j}}.$$

Proof. Apply (68) for $\Phi_j(x) = \|x\|_{p_j}$, $x_i \geq 0$, $i = 1, ..., n$, $j = 1, ..., m$. ■

Corollary 27 *(to Theorem 24) It holds*

$$\int_a^b u(x) \prod_{j=1}^m \left(\sum_{i=1}^n e^{\frac{|g_{ji}(x)|}{K_j(x)}} \right) \Diamond_\alpha x \leq \tag{72}$$

$$\prod_{j=1}^m \left(\int_c^d u_j(y) \left(\sum_{i=1}^n e^{|f_{ji}(y)|} \right)^{p_j} \Diamond_\alpha y \right)^{\frac{1}{p_j}}.$$

Proof. Apply (68) for $\Phi_j(x) = \sum_{i=1}^n e^{x_i}$, $x_i \geq 0$, $i = 1, ..., n$,for all $j = 1, ..., m$. ■

We need

Definition 28 *([3]) Let* $\mathbb{T}$ *be a time scale. Consider the coordinate wise rd-continuous functions* $h_\alpha : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, $\alpha \geq 0$, *such that* $h_0(t, s) = 1$,

$$h_{\alpha+1}(t, s) = \int_s^t h_\alpha(\tau, s) \Delta\tau, \tag{73}$$

$\forall\ s, t \in \mathbb{T}$.

When $\mathbb{T} = \mathbb{R}$, *then* $\sigma(t) = t$, *we define*

$$h_\alpha(t, s) := \frac{(t-s)^\alpha}{\Gamma(\alpha+1)}, \quad \alpha \geq 0. \tag{74}$$

When $\mathbb{T}=\mathbb{Z}$, *then* $\sigma(t)=t+1$, $t\in\mathbb{Z}$, *and*

$$h_k(t,s)=\frac{(t-s)^{(k)}}{k!},\quad \forall\ k\in\mathbb{N}_0=\mathbb{N}\cup\{0\}, \tag{75}$$

$\forall\ t,s\in\mathbb{Z}$, *where* $t^{(0)}=1$, $t^{(k)}=\prod_{i=0}^{k-1}(t-i)$ *for* $k\in\mathbb{N}$. *Also it holds*

$$\int_a^b f(t)\,\Delta t=\sum_{t=a}^{b-1}f(t),\quad a<b;\ a,b\in\mathbb{Z}. \tag{76}$$

We need

Definition 29 *([3]) For* $\alpha\geq 1$ *we define the time scale* Δ*-Riemann-Liouville type fractional integral* $(a,b\in\mathbb{T})$

$$K_a^{\alpha}f(t)=\int_a^t h_{\alpha-1}(t,\sigma(\tau))\,f(\tau)\,\Delta\tau, \tag{77}$$

(by [11] is an integral on $[a,t)\cap\mathbb{T}$*)*

$$K_a^0 f=f,$$

where $f\in L_1([a,b)\cap\mathbb{T})$ *(Lebesgue* Δ*-integrable functions on* $[a,b)\cap\mathbb{T}$*, see [17], [9], [10]),* $t\in[a,b]\cap\mathbb{T}$.

Notice $K_a^1 f(t)=\int_a^t f(\tau)\,\Delta\tau$ *is absolutely continuous in* $t\in[a,b]\cap\mathbb{T}$*, see [11].*

Lemma 30 *([3]) Let* $\alpha>1$, $f\in L_1([a,b)\cap\mathbb{T})$*. If additionally* $h_{\alpha-1}(s,\sigma(t))$ *is Lebesgue* Δ*-measurable on* $([a,b)\cap\mathbb{T})^2$*, then* $K_a^{\alpha}f\in L_1([a,b)\cap\mathbb{T})$.

We need

Definition 31 *([3]) Assume* $\mathbb{T}$ *time scale such that* $\mathbb{T}^k=\mathbb{T}$.

Let $\mu>2:m-1<\mu<m\in\mathbb{N}$*, i.e.* $m=\lceil\mu\rceil$ *(ceiling of the number),* $\widetilde{\nu}=m-\mu$ $(0<\widetilde{\nu}<1)$.

Here we take $f\in C_{rd}^m([a,b]\cap\mathbb{T})$*. Clearly here ([17])* f^{Δ^m} *is a Lebesgue* Δ*-integrable function. Assume* $h_{\widetilde{\nu}}(s,\sigma(t))$ *is continuous on* $([a,b]\cap\mathbb{T})^2$.

We define the delta fractional derivative on time scale $\mathbb{T}$ *of order* $\mu-1$ *as follows:*

$$\Delta_{a*}^{\mu-1}f(t)=\left(K_a^{\widetilde{\nu}+1}f^{\Delta^m}\right)(t)=\int_a^t h_{\widetilde{\nu}}(t,\sigma(\tau))\,f^{\Delta^m}(\tau)\,\Delta\tau, \tag{78}$$

$\forall\ t\in[a,b]\cap\mathbb{T}$.

Notice here that $\Delta_{a*}^{\mu-1}f\in C([a,b]\cap\mathbb{T})$ *by a simple argument using dominated convergence theorem in Lebesgue* Δ*-sense.*

If $\mu = m$, then $\widetilde{\nu} = 0$ and by (78) we get

$$\Delta_{a*}^{m-1} f(t) = K_a^1 f^{\Delta^m}(t) = f^{\Delta^{m-1}}(t). \tag{79}$$

More generally, by [11], given that $f^{\Delta^{m-1}}$ is everywhere finite and absolutely continuous on $[a,b] \cap \mathbb{T}$, then f^{Δ^m} exists Δ-a.e. and is Lebesgue Δ-integrable on $[a,t) \cap \mathbb{T}$, $\forall\ t \in [a,b] \cap \mathbb{T}$ and one can plug it into (78).

We need

Definition 32 *([4]) Consider the coordinate wise ld-continuous functions $\widehat{h}_\alpha : \mathbb{T} \times \mathbb{T} \to \mathbb{R}$, $\alpha \geq 0$, such that $\widehat{h}_0(t,s) = 1$,*

$$\widehat{h}_{\alpha+1}(t,s) = \int_s^t \widehat{h}_\alpha(\tau,s)\nabla\tau, \tag{80}$$

$\forall\ s,t \in \mathbb{T}$.

In the case of $\mathbb{T} = \mathbb{R}$; then $\rho(t) = t$, and $\widehat{h}_k(t,s) = \frac{(t-s)^k}{k!}$, $k \in \mathbb{N}_0$, and define

$$\widehat{h}_\alpha(t,s) := \frac{(t-s)^\alpha}{\Gamma(\alpha+1)}, \quad \alpha \geq 0. \tag{81}$$

Let $T = \mathbb{Z}$, then $\rho(t) = t-1$, $t \in \mathbb{Z}$. Define $t^{\overline{0}} := 1$, $t^{\overline{k}} := t(t+1)...(t+k-1)$, $k \in \mathbb{N}$, and by (80) we have $\widehat{h}_k(t,s) = \frac{(t-s)^{\overline{k}}}{k!}$, $s,t \in \mathbb{Z}$, $k \in \mathbb{N}_0$.

Here $\int_{t_0}^t \nabla\tau = \sum_{t_0+1}^t$.

We need

Definition 33 *([4]) For $\alpha \geq 1$ we define the time scale ∇-Riemann-Liouville type fractional integral ($a,b \in \mathbb{T}$)*

$$J_a^\alpha f(t) = \int_a^t \widehat{h}_{\alpha-1}(t,\rho(\tau)) f(\tau)\nabla\tau, \tag{82}$$

(by [11] the last integral is on $(a,t] \cap \mathbb{T}$)

$$J_a^0 f(t) = f(t),$$

where $f \in L_1((a,b] \cap \mathbb{T})$ (Lebesgue ∇-integrable functions on $(a,b] \cap \mathbb{T}$, see [9], [10], [17]), $t \in [a,b] \cap \mathbb{T}$.

Notice $J_a^1 f(t) = \int_a^t f(\tau)\nabla\tau$ is absolutely continuous in $t \in [a,b] \cap \mathbb{T}$, see [11].

Lemma 34 *([4]) Let $\alpha > 1$, $f \in L_1((a,b] \cap \mathbb{T})$. If additionally $\widehat{h}_{\alpha-1}(s,\rho(t))$ is Lebesgue ∇-measurable on $((a,b] \cap \mathbb{T})^2$, then $J_a^\alpha f \in L_1((a,b] \cap \mathbb{T})$.*

We also need

Definition 35 *([4]) Assume $\mathbb{T}_k = \mathbb{T}$.*

Let $\mu > 2$ such that $m-1 < \mu < m \in \mathbb{N}$, i.e. $m = \lceil \mu \rceil$, $\widetilde{\nu} = m - \mu$ $(0 < \widetilde{\nu} < 1)$.

Let $f \in C_{ld}^m([a,b] \cap \mathbb{T})$. Clearly here ([17]) f^{∇^m} is a Lebesgue ∇-integrable function. Assume $\widehat{h}_{\widetilde{\nu}}(s, \rho(t))$ is continuous on $([a,b] \cap \mathbb{T})^2$.

We define the nabla fractional derivative on time scale $\mathbb{T}$ of order $\mu - 1$ as follows:

$$\nabla_{a*}^{\mu-1} f(t) = \left(J_a^{\widetilde{\nu}+1} f^{\nabla^m} \right)(t) = \int_a^t \widehat{h}_{\widetilde{\nu}}(t, \rho(\tau)) f^{\nabla^m}(\tau) \nabla\tau, \tag{83}$$

$\forall\ t \in [a,b] \cap \mathbb{T}$.

Notice here that $\nabla_{a*}^{\mu-1} f \in C([a,b] \cap \mathbb{T})$ by a simple argument using dominated convergence theorem in Lebesgue ∇-sense.

If $\mu = m$, then $\widetilde{\nu} = 0$ and by (83) we get

$$\nabla_{a*}^{m-1} f(t) = J_a^1 f^{\nabla^m}(t) = f^{\nabla^{m-1}}(t). \tag{84}$$

More generally, by [11], given that $f^{\nabla^{m-1}}$ is everywhere finite and absolutely continuous on $[a,b] \cap \mathbb{T}$, then f^{∇^m} exists ∇-a.e. and is Lebesgue ∇-integrable on $(a,t] \cap \mathbb{T}$, $\forall\ t \in [a,b] \cap \mathbb{T}$, and one can plug it into (83).

We present

Theorem 36 *Let $\mathbb{T}$ be a time scale, and $a, b \in \mathbb{T}$, $a < b$, with $\sigma(a) = a$. Let $\alpha \geq 1$, h_α as in (73), and $K_a^\alpha f_i$ as in (77), where $f_i \in L_1([a,b) \cap \mathbb{T})$, $i = 1, ..., n$. Assume further that $h_{\alpha-1}(s, \sigma(t))$ is Lebesgue Δ-measurable on $([a,b) \cap \mathbb{T})^2$.*

Let $\overrightarrow{f} := (f_1, ..., f_n)$, $\overrightarrow{K_a^\alpha f} := (K_a^\alpha f_1, ..., K_a^\alpha f_n)$.

Call

$$K^*(x) := \int_a^b \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta y \tag{85}$$
$$= \int_a^x \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta y,$$

$\forall\ x \in ([a,b] \cap \mathbb{T})$, where $\chi_{[a,x)}(y)$ is the characteristic function on $[a,x) \cap \mathbb{T}$.

Assume that $K^*(x) > 0$ (delta) Lebesgue measure Δ-a.e. in $x \in ([a,b] \cap \mathbb{T})$.

Consider also the weight function

$$u : ([a,b) \cap \mathbb{T}) \to \mathbb{R}_+, \tag{86}$$

which is (delta) Lebesgue Δ-measurable. Assume that the function

$$x \to \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right|$$

is Δ-integrable on $([a,b) \cap \mathbb{T})$ for each fixed $y \in ([a,b) \cap \mathbb{T})$.

Define v^* *on* $([a,b) \cap \mathbb{T})$ *by*

$$v^*(y) := \int_a^b \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta x < \infty. \tag{87}$$

Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ *be a convex and increasing per coordinate function. Then*

$$\int_a^b u(x) \Phi\left(\frac{\left| \overrightarrow{K_a^\alpha f(x)} \right|}{K^*(x)} \right) \Delta x \leq \int_a^b v^*(x) \Phi\left(\left| \overrightarrow{f(x)} \right| \right) \Delta x, \tag{88}$$

under the further assumptions:

(i) $\overrightarrow{f}, \Phi\left(\left|\overrightarrow{f}\right|\right)$ *are both* $\chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta y$*-integrable,* Δ*-a.e. in* $x \in ([a,b) \cap \mathbb{T})$,

(ii) $v^* \Phi\left(\left|\overrightarrow{f}\right|\right)$ *is* Δ*-Lebesgue integrable.*

Proof. By [5], Φ is continuous on $\mathbb{R}_+$. Here we use the multivariate Jensen's inequality, Tonelli's theorem and Fubini's theorem, all on Time Scales setting. Also we use that Φ is convex and increasing per coordinate. We extend $\Phi(x)$ on $\mathbb{R}^n$, see Lemma 10, so we can apply multivariate Jensen's inequality.

Next we have

$$K_a^\alpha f_i(x) = \int_a^b \chi_{[a,x)}(y) h_{\alpha-1}(x, \sigma(y)) f_i(y) \Delta y, \quad \forall \ x \in [a,b] \cap \mathbb{T},$$

and

$$\left| K_a^\alpha f_i(x) \right| \leq \int_a^b \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \left| f_i(y) \right| \Delta y, \quad i = 1, ..., n.$$

We see that

$$\int_a^b u(x) \Phi\left(\frac{\left| \overrightarrow{K_a^\alpha f(x)} \right|}{K^*(x)} \right) \Delta x \leq$$

$$\int_a^b u(x) \Phi\left(\frac{1}{K^*(x)} \int_a^b \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \left| \overrightarrow{f(y)} \right| \Delta y \right) \Delta x \leq$$

(by multivariate Jensen's inequality)

$$\int_a^b \frac{u(x)}{K^*(x)} \left(\int_a^b \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Phi\left(\left| \overrightarrow{f(y)} \right| \right) \Delta y \right) \Delta x =$$

$$\int_a^b \left(\int_a^b \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Phi\left(\left| \overrightarrow{f(y)} \right| \right) \Delta y \right) \Delta x =$$

21

(by Fubini's theorem)

$$\int_a^b \left(\int_a^b \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Phi\left(\left| \overrightarrow{f(y)} \right| \right) \Delta x \right) \Delta y =$$

$$\int_a^b \Phi\left(\left| \overrightarrow{f(y)} \right| \right) \left(\int_a^b \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta x \right) \Delta y =$$

$$\int_a^b \Phi\left(\left| \overrightarrow{f(y)} \right| \right) v^*(y) \Delta y,$$

completing the proof of the theorem. ■

The counterpart of last result follows:

Theorem 37 *Let $\mathbb{T}$ be a time scale, and $a, b \in \mathbb{T} - \{\min \mathbb{T}\}$, $a < b$, with $\rho(a) = a$. Let $\alpha \geq 1$, $\widehat{h}_\alpha$ as in (80), and $J_a^\alpha f_i$ as in (82), where $f_i \in L_1((a,b] \cap \mathbb{T})$, $i = 1, ..., n$. Let $\overrightarrow{f} := (f_1, ..., f_n)$, $\overrightarrow{J_a^\alpha f} := (J_a^\alpha f_1, ..., J_a^\alpha f_n)$. Assume further that $\widehat{h}_{\alpha-1}(s, \rho(t))$ is Lebesgue ∇-measurable on $((a,b] \cap \mathbb{T})^2$.*

Call

$$K_*(x) := \int_a^b \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y \tag{89}$$

$$= \int_a^x \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y,$$

$\forall\, x \in ([a,b] \cap \mathbb{T})$. Assume that $K_*(x) > 0$ (nabla) Lebesgue measure ∇-a.e. in $x \in ([a,b] \cap \mathbb{T})$.

Consider also the weight function

$$w : ((a,b] \cap \mathbb{T}) \to \mathbb{R}_+, \tag{90}$$

which is (nabla) Lebesgue ∇-measurable.

Assume that the function

$$x \to \frac{w(x)}{K_*(x)} \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right|$$

is ∇-integrable on $((a,b] \cap \mathbb{T})$ for each fixed $y \in ((a,b] \cap \mathbb{T})$.

Define v_* on $((a,b] \cap \mathbb{T})$ by

$$v_*(y) := \int_a^b \frac{w(x)}{K_*(x)} \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla x < \infty. \tag{91}$$

Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ be a convex and increasing per coordinate function. Then

$$\int_a^b w(x) \Phi\left(\frac{\left| \overrightarrow{J_a^\alpha f(x)} \right|}{K_*(x)} \right) \nabla x \leq \int_a^b v_*(x) \Phi\left(\left| \overrightarrow{f(x)} \right| \right) \nabla x, \tag{92}$$

under the further assumptions:

(i) $\overrightarrow{f}, \Phi\left(\left|\overrightarrow{f}\right|\right)$ *are both* $\chi_{[a,x)}(y)\left|\widehat{h}_{\alpha-1}(x,\rho(y))\right| \nabla y$*-integrable,* ∇*-a.e. in* $x \in ((a,b] \cap \mathbb{T})$,

(ii) $v_* \Phi\left(\left|\overrightarrow{f}\right|\right)$ *is* ∇*-Lebesgue integrable.*

Proof. Similar to the proof of Theorem 36. ■

We continue with

Theorem 38 *Let* $\mathbb{T}$ *be a time scale, and* $a, b \in \mathbb{T}$, $a < b$, *with* $\sigma(a) = a$. *Let* $\alpha \geq 1$, h_α *as in (73), and* $K_a^\alpha f_{ji}$ *as in (77), where* $f_{ji} \in L_1([a,b) \cap \mathbb{T})$, $j = 1, ..., m \in \mathbb{N}$, $i = 1, ..., n \in \mathbb{N}$. *Let* $\overrightarrow{f_j} := (f_{j1}, ..., j_{jn})$, $\overrightarrow{K_a^\alpha f_j} := (K_a^\alpha f_{j1}, ..., K_a^\alpha f_{jn})$, $j = 1, ..., m$. *Assume further that* $h_{\alpha-1}(s, \sigma(t))$ *is Lebesgue* Δ*-measurable on* $([a,b) \cap \mathbb{T})^2$. *Call*

$$K^*(x) := \int_a^b \chi_{[a,x)}(y) \left|h_{\alpha-1}(x, \sigma(y))\right| \Delta y \tag{93}$$

$\forall\, x \in ([a,b] \cap \mathbb{T})$. *Assume that* $K^*(x) > 0$, Δ*-a.e. in* $x \in ([a,b] \cap \mathbb{T})$.

Here the weight function

$$u : ([a,b) \cap \mathbb{T}) \to \mathbb{R}_+,$$

is Lebesgue Δ*-measurable. Assume that the function*

$$x \to u(x)\, \chi_{[a,x)}(y) \left(\frac{\left|h_{\alpha-1}(x, \sigma(y))\right|}{K^*(x)}\right)^m$$

is Lebesgue Δ*-integrable on* $([a,b) \cap \mathbb{T})$ *for each fixed* $y \in ([a,b) \cap \mathbb{T})$.

Define v_m^* *on* $([a,b) \cap \mathbb{T})$ *by*

$$v_m^*(y) := \int_a^b u(x)\, \chi_{[a,x)}(y) \left(\frac{\left|h_{\alpha-1}(x, \sigma(y))\right|}{K^*(x)}\right)^m \Delta x < \infty. \tag{94}$$

Let $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1, ..., m$, *are convex and increasing per coordinate functions.*

Then

$$\int_a^b u(x) \prod_{j=1}^m \Phi_j\left(\frac{\left|\overrightarrow{K_a^\alpha f_j(x)}\right|}{K^*(x)}\right) \Delta x \leq \tag{95}$$

$$\left(\prod_{j=2}^m \int_a^b \Phi_j\left(\left|\overrightarrow{f_j(y)}\right|\right) \Delta y\right) \left(\int_a^b \Phi_1\left(\left|\overrightarrow{f_1(y)}\right|\right) v_m^*(y)\, \Delta y\right),$$

under the further assumptions:

(i) $\overrightarrow{f_j}, \Phi_j\left(\left|\overrightarrow{f_j}\right|\right)$ *are both* $\chi_{[a,x)}(y)\left|h_{\alpha-1}(x, \sigma(y))\right| \Delta y$*-integrable,* Δ*-a.e. in* $x \in ([a,b) \cap \mathbb{T})$, $j = 1, ..., m$, *and*

(ii) $v_m^* \Phi_1\left(\left|\overrightarrow{f_1}\right|\right)$, $\Phi_2\left(\left|\overrightarrow{f_2}\right|\right)$, $\Phi_3\left(\left|\overrightarrow{f_3}\right|\right), ..., \Phi_m\left(\left|\overrightarrow{f_m}\right|\right)$, *are all* Δ*-Lebesgue integrable.*

Proof. As in [5], and Theorem 20 here. ■

The counterpart of last result follows

Theorem 39 *Let $\mathbb{T}$ be a time scale, and $a, b \in \mathbb{T} - \{\min \mathbb{T}\}$, $a < b$, with $\rho(a) = a$. Let $\alpha \geq 1$, $\widehat{h}_\alpha$ as in (80), and $J_a^\alpha f_{ji}$ as in (82), where $f_{ji} \in L_1((a,b] \cap \mathbb{T})$, for $j = 1, ..., m \in \mathbb{N}$, $i = 1, ..., n \in \mathbb{N}$. Let $\overrightarrow{f_j} := (f_{j1}, ..., f_{jn})$, $\overrightarrow{J_a^\alpha f_j} := (J_a^\alpha f_{j1}, ..., J_a^\alpha f_{jn})$, $j = 1, ..., m$. Assume further that $\widehat{h}_{\alpha-1}(s, \rho(t))$ is Lebesgue ∇-measurable on $((a,b] \cap \mathbb{T})^2$. Call*

$$K_*(x) := \int_a^b \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y \tag{96}$$

$\forall\ x \in ([a,b] \cap \mathbb{T})$. Assume that $K_*(x) > 0$, ∇-a.e. in $x \in ([a,b] \cap \mathbb{T})$.

Here the weight function

$$w : ((a,b] \cap \mathbb{T}) \to \mathbb{R}_+,$$

is Lebesgue ∇-measurable. Assume that the function

$$x \to w(x) \chi_{[a,x)}(y) \left(\frac{\left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right|}{K_*(x)} \right)^m$$

is Lebesgue ∇-integrable on $((a,b] \cap \mathbb{T})$ for each fixed $y \in ((a,b] \cap \mathbb{T})$.

Define v_{m*} on $((a,b] \cap \mathbb{T})$ by

$$v_{m*}(y) := \int_a^b w(x) \chi_{[a,x)}(y) \left(\frac{\left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right|}{K_*(x)} \right)^m \nabla x < \infty. \tag{97}$$

Let $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1, ..., m$, are convex and increasing per coordinate functions.

Then

$$\int_a^b w(x) \prod_{j=1}^m \Phi_j \left(\frac{\left| \overrightarrow{J_a^\alpha f_j(x)} \right|}{K_*(x)} \right) \nabla x \leq \tag{98}$$

$$\left(\prod_{j=2}^m \int_a^b \Phi_j \left(\left| \overrightarrow{f_j(y)} \right| \right) \nabla y \right) \left(\int_a^b \Phi_1 \left(\left| \overrightarrow{f_1(y)} \right| \right) v_{m*}(y) \nabla y \right),$$

under the further assumptions:

(i) $\overrightarrow{f_j}, \Phi_j \left(\left| \overrightarrow{f_j} \right| \right)$ are both $\chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y$-integrable, ∇-a.e. in $x \in ((a,b] \cap \mathbb{T})$, $j = 1, ..., m$, and

(ii) $v_{m*} \Phi_1 \left(\left| \overrightarrow{f_1} \right| \right)$, $\Phi_2 \left(\left| \overrightarrow{f_2} \right| \right)$, $\Phi_3 \left(\left| \overrightarrow{f_3} \right| \right)$, ..., $\Phi_m \left(\left| \overrightarrow{f_m} \right| \right)$, are all ∇-Lebesgue integrable.

Proof. As in [5], and Theorem 20 here. ■

We continue with

Theorem 40 *Let $\mathbb{T}$ be a time scale, and $a, b \in \mathbb{T}$, $a < b$, with $\sigma(a) = a$. Let $\alpha \geq 1$, h_α as in (73), and $K_a^\alpha f_{ji}$ as in (77), where $f_{ji} \in L_1([a,b) \cap \mathbb{T})$, $j = 1, ..., m \in \mathbb{N}$, $i = 1, ..., n \in \mathbb{N}$. Let $\overrightarrow{f_j} := (f_{j1}, ..., f_{jn})$, $\overrightarrow{K_a^\alpha f_j} := (K_a^\alpha f_{j1}, ..., K_a^\alpha f_{jn})$, $j = 1, ..., m$. Assume further that $h_{\alpha-1}(s, \sigma(t))$ is Lebesgue Δ-measurable on $([a,b) \cap \mathbb{T})^2$. Call*

$$K^*(x) := \int_a^b \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta y \tag{99}$$

$\forall\ x \in ([a,b] \cap \mathbb{T})$. Assume that $K^*(x) > 0$, Δ-a.e. in $x \in ([a,b] \cap \mathbb{T})$.

Here the weight function

$$u : ([a,b) \cap \mathbb{T}) \to \mathbb{R}_+,$$

is Lebesgue Δ-measurable. Assume that the function

$$x \to u(x) \chi_{[a,x)}(y) \left(\frac{|h_{\alpha-1}(x, \sigma(y))|}{K^*(x)} \right)$$

is Lebesgue Δ-integrable on $([a,b) \cap \mathbb{T})$ for each fixed $y \in ([a,b) \cap \mathbb{T})$.

Define v^* on $([a,b) \cap \mathbb{T})$ by

$$v^*(y) := \int_a^b \frac{u(x)}{K^*(x)} \chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta x < \infty. \tag{100}$$

Let $p_j > 1 : \sum_{j=1}^m \frac{1}{p_j} = 1$. Let the functions $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1, ..., m$, be convex and increasing per coordinate.

Then

$$\int_a^b u(x) \prod_{j=1}^m \Phi_j \left(\frac{\left| \overrightarrow{K_a^\alpha f_j(x)} \right|}{K^*(x)} \right) \Delta x \leq \tag{101}$$

$$\left(\prod_{j=1}^m \int_a^b v^*(y) \Phi_j \left(\left| \overrightarrow{f_j(y)} \right| \right)^{p_j} \Delta y \right)^{\frac{1}{p_j}},$$

under the further assumptions:

(i) $\overrightarrow{f_j}, \Phi_j \left(\left| \overrightarrow{f_j} \right| \right)^{p_j}$ are both $\chi_{[a,x)}(y) \left| h_{\alpha-1}(x, \sigma(y)) \right| \Delta y$-integrable, Δ-a.e. in $x \in ([a,b) \cap \mathbb{T})$, for all $j = 1, ..., m$, and

(ii) $v^* \Phi_j \left(\left| \overrightarrow{f_j} \right| \right)^{p_j}$ is Δ-Lebesgue integrable, $j = 1, ..., m$.

Proof. As in [7], and Theorem 24 here. ■

The counterpart of last result follows.

Theorem 41 *Let $\mathbb{T}$ be a time scale, and $a, b \in \mathbb{T} - \{\min \mathbb{T}\}$, $a < b$, with $\rho(a) = a$. Let $\alpha \geq 1$, $\widehat{h}_\alpha$ as in (80), and $J_a^\alpha f_{ji}$ as in (82), where $f_{ji} \in L_1((a,b] \cap \mathbb{T})$, $i = 1, ..., n \in \mathbb{N}$, $j = 1, ..., m \in \mathbb{N}$. Let $\overrightarrow{f_j} := (f_{j1}, ..., f_{jn})$, $\overrightarrow{J_a^\alpha f_j} := (J_a^\alpha f_{j1}, ..., J_a^\alpha f_{jn})$, $j = 1, ..., m$. Assume further that $\widehat{h}_{\alpha-1}(s, \rho(t))$ is Lebesgue ∇-measurable on $((a,b] \cap \mathbb{T})^2$. Call*

$$K_*(x) := \int_a^b \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y, \tag{102}$$

$\forall\, x \in ([a,b] \cap \mathbb{T})$. Assume that $K_*(x) > 0$, ∇-a.e. in $x \in ([a,b] \cap \mathbb{T})$.

Here the weight function

$$w : ((a,b] \cap \mathbb{T}) \to \mathbb{R}_+,$$

is Lebesgue ∇-measurable. Assume that the function

$$x \to w(x) \chi_{[a,x)}(y) \left(\frac{\left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right|}{K_*(x)} \right)$$

is Lebesgue ∇-integrable on $((a,b] \cap \mathbb{T})$ for each fixed $y \in ((a,b] \cap \mathbb{T})$.

Define v_* on $((a,b] \cap \mathbb{T})$ by

$$v_*(y) := \int_a^b \frac{w(x)}{K_*(x)} \chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla x < \infty. \tag{103}$$

Let $p_j > 1 : \sum_{j=1}^m \frac{1}{p_j} = 1$. Let the functions $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1, ..., m$, be convex and increasing per coordinate.

Then

$$\int_a^b w(x) \prod_{j=1}^m \Phi_j \left(\frac{\left| \overrightarrow{J_a^\alpha f_j(x)} \right|}{K_*(x)} \right) \nabla x \leq \tag{104}$$

$$\left(\prod_{j=1}^m \int_a^b v_*(y) \Phi_j \left(\left| \overrightarrow{f_j(y)} \right| \right)^{p_j} \nabla y \right)^{\frac{1}{p_j}},$$

under the further assumptions:

(i) $\overrightarrow{f_j}, \Phi_j \left(\left| \overrightarrow{f_j} \right| \right)^{p_j}$ are both $\chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y$-integrable, ∇-a.e. in $x \in ((a,b] \cap \mathbb{T})$, for all $j = 1, ..., m$, and

(ii) $v_* \Phi_j \left(\left| \overrightarrow{f_j} \right| \right)^{p_j}$ is ∇-Lebesgue integrable, $j = 1, ..., m$.

Proof. As in [7], and Theorem 24 here. ∎

We give

Corollary 42 *(to Theorem 38) It holds*

$$\int_a^b u(x) \prod_{j=1}^m \left(\sum_{i=1}^n e^{\frac{|K_a^\alpha f_{ji}(x)|}{K^*(x)}} \right) \Delta x \leq \tag{105}$$

$$\left(\prod_{j=2}^m \int_a^b \sum_{i=1}^n e^{|f_{ji}(y)|} \Delta y \right) \left(\int_a^b \left(\sum_{i=1}^n e^{|f_{1i}(y)|} \right) v_m^*(y) \Delta y \right),$$

under the assumptions:

(i) $\overrightarrow{f_j}, \sum_{i=1}^n e^{|f_{ji}(y)|}$ *are* $\chi_{[a,x)}(y) |h_{\alpha-1}(x, \sigma(y))| \Delta y$*-integrable,* Δ*-a.e. in* $x \in ([a,b) \cap \mathbb{T})$, $j = 1, ..., m$, *and*

(ii) $v_m^* \left(\sum_{i=1}^n e^{|f_{1i}(y)|} \right)$, $\sum_{i=1}^n e^{|f_{ji}(y)|}$, $j = 2, ..., m$, *are all* Δ*-integrable.*

Corollary 43 *(to Theorem 41) It holds*

$$\int_a^b w(x) \prod_{j=1}^m \left(\sum_{i=1}^n e^{\frac{|J_a^\alpha f_{ji}(x)|}{K_*(x)}} \right) \nabla x \leq \tag{106}$$

$$\left(\prod_{j=1}^m \int_a^b v_*(y) \left(\sum_{i=1}^n e^{|f_{ji}(y)|} \right)^{p_j} \nabla y \right)^{\frac{1}{p_j}},$$

under the assumptions:

(i) $\overrightarrow{f_j}, \left(\sum_{i=1}^n e^{|f_{ji}(y)|} \right)^{p_j}$ *are both* $\chi_{[a,x)}(y) \left| \widehat{h}_{\alpha-1}(x, \rho(y)) \right| \nabla y$*-integrable,* ∇*-a.e. in* $x \in ((a,b] \cap \mathbb{T})$, *for all* $j = 1, ..., m$, *and*

(ii) $v_*(y) \left(\sum_{i=1}^n e^{|f_{ji}(y)|} \right)^{p_j}$ *is* ∇*-Lebesgue integrable,* $j = 1, ..., m$.

We continue with

Theorem 44 *Let* $\mathbb{T}$ *be a time scale, and* $a, b \in \mathbb{T}$, $a < b$, *with* $\sigma(a) = a$. *Let all as in Definition 31 and for* $f_{ji} \in C_{rd}^m([a,b] \cap \mathbb{T})$, $j = 1, ..., m_* \in \mathbb{N}$, $i = 1, ..., n \in \mathbb{N}$; $h_{\widetilde{\nu}}$ *as in (73). Let* $\overrightarrow{f_j}^{\Delta^m} := \left(f_{j1}^{\Delta^m}, ..., f_{jn}^{\Delta^m} \right)$, $\overrightarrow{\Delta_{a*}^{\mu-1} f_j} := \left(\Delta_{a*}^{\mu-1} f_{j1}, ..., \Delta_{a*}^{\mu-1} f_{jn} \right)$, $j = 1, ..., m$. *Call*

$$K_1(x) := \int_a^b \chi_{[a,x)}(y) |h_{\widetilde{\nu}}(x, \sigma(y))| \Delta y \tag{107}$$

$\forall\, x \in ([a,b] \cap \mathbb{T})$. *Assume that* $K_1(x) > 0$, Δ*-a.e. in* $x \in ([a,b] \cap \mathbb{T})$.

Here the weight function

$$u : ([a,b) \cap \mathbb{T}) \to \mathbb{R}_+,$$

is Lebesgue Δ*-measurable. Assume that the function*

$$x \to u(x) \chi_{[a,x)}(y) \left(\frac{|h_{\widetilde{\nu}}(x, \sigma(y))|}{K_1(x)} \right)^{m_*}$$

is Lebesgue Δ-integrable on $([a,b)\cap\mathbb{T})$ for each fixed $y\in([a,b)\cap\mathbb{T})$.

Define φ_{m_*} on $([a,b)\cap\mathbb{T})$ by

$$\varphi_{m_*}(y):=\int_a^b u(x)\,\chi_{[a,x)}(y)\left(\frac{|h_{\widetilde{\nu}}(x,\sigma(y))|}{K_1(x)}\right)^{m_*}\Delta x<\infty. \tag{108}$$

Here $\Phi_j:\mathbb{R}_+^n\to\mathbb{R}_+$, $j=1,...,m_*$, are convex and increasing per coordinate functions.

Then

$$\int_a^b u(x)\prod_{j=1}^{m_*}\Phi_j\left(\frac{\left|\overrightarrow{\Delta_{a*}^{\mu-1}f_j(x)}\right|}{K_1(x)}\right)\Delta x\leq \tag{109}$$

$$\left(\prod_{j=2}^{m_*}\int_a^b\Phi_j\left(\left|\overrightarrow{f_j^{\Delta^m}(y)}\right|\right)\Delta y\right)\left(\int_a^b\Phi_1\left(\left|\overrightarrow{f_1^{\Delta^m}(y)}\right|\right)\varphi_{m_*}(y)\,\Delta y\right),$$

under the assumption that φ_{m_*} is Δ-Lebesgue integrable on $([a,b)\cap\mathbb{T})$.

Proof. By Theorem 38. ■

We also derive

Theorem 45 *Let $\mathbb{T}$ be a time scale, and $a,b\in\mathbb{T}-\{\min\mathbb{T}\}$, $a<b$, with $\rho(a)=a$. Let all as in Definition 35 and for $f_{ji}\in C_{ld}^m([a,b]\cap\mathbb{T})$, $j=1,...,m_*\in\mathbb{N}$, $i=1,...,n\in\mathbb{N}$; $h_{\widetilde{\nu}}$ as in (80). Let $\overrightarrow{f_j}^{\nabla^m}:=\left(f_{j1}^{\nabla^m},...,f_{jn}^{\nabla^m}\right)$, $\overrightarrow{\nabla_{a*}^{\mu-1}f_j}:=\left(\nabla_{a*}^{\mu-1}f_{j1},...,\nabla_{a*}^{\mu-1}f_{jn}\right)$, $j=1,...,m$. Call*

$$K_2(x):=\int_a^b\chi_{[a,x)}(y)\left|\widehat{h}_{\widetilde{\nu}}(x,\rho(y))\right|\nabla y, \tag{110}$$

$\forall\ x\in([a,b]\cap\mathbb{T})$. Assume that $K_2(x)>0$, ∇-a.e. in $x\in([a,b]\cap\mathbb{T})$.

Here the weight function

$$w:((a,b]\cap\mathbb{T})\to\mathbb{R}_+,$$

is Lebesgue ∇-measurable. Assume that the function

$$x\to w(x)\,\chi_{[a,x)}(y)\left(\frac{\left|\widehat{h}_{\widetilde{\nu}}(x,\rho(y))\right|}{K_2(x)}\right)$$

is Lebesgue ∇-integrable on $((a,b]\cap\mathbb{T})$ for each fixed $y\in((a,b]\cap\mathbb{T})$.

Define ψ on $((a,b]\cap\mathbb{T})$ by

$$\psi(y):=\int_a^b\frac{w(x)}{K_2(x)}\chi_{[a,x)}(y)\left|\widehat{h}_{\widetilde{\nu}}(x,\rho(y))\right|\nabla x<\infty. \tag{111}$$

Let $p_j > 1 : \sum_{j=1}^{m_*} \frac{1}{p_j} = 1$. Let the functions $\Phi_j : \mathbb{R}_+^n \to \mathbb{R}_+$, $j = 1, ..., m_*$, be convex and increasing per coordinate.

Then

$$\int_a^b w(x) \prod_{j=1}^m \Phi_j \left(\frac{\left| \overrightarrow{\nabla_{a*}^{\mu-1} f_j(x)} \right|}{K_2(x)} \right) \nabla x \leq \tag{112}$$

$$\left(\prod_{j=1}^m \int_a^b \psi(y) \Phi_j \left(\left| \overrightarrow{f_j^{\nabla^m}(y)} \right| \right)^{p_j} \nabla y \right)^{\frac{1}{p_j}},$$

under the assumption that ψ is ∇-Lebesgue integrable on $((a,b] \cap \mathbb{T})$.

Proof. By Theorem 41. ■

We finish with

Remark 46 *(to Theorem 15)*

(i) Let $\mathbb{T}_1 = \mathbb{R}$, $\mathbb{T}_2 = \mathbb{Z}$; $0 \leq \alpha \leq 1$. Then

$$\int_a^b \cdot \diamondsuit_\alpha x = \int_a^b \cdot dx, \tag{113}$$

and

$$\int_c^d \cdot \diamondsuit_\alpha y = \alpha \sum_{y=c}^{d-1} \cdot + (1-\alpha) \sum_{y=c+1}^{d} \cdot. \tag{114}$$

Assume $k : [a,b] \times [c,d]_{\mathbb{Z}} \to \mathbb{R}_+$, a continuous function. So here

$$K(x) = \alpha \sum_{y=c}^{d-1} k(x,y) + (1-\alpha) \sum_{y=c+1}^{d} k(x,y) > 0, \tag{115}$$

$\forall\ x \in [a,b]$, and

$f_i : [c,d]_{\mathbb{Z}} \to \mathbb{R}$, with

$$g_i(x) = \alpha \sum_{y=c}^{d-1} k(x,y) f_i(y) + (1-\alpha) \sum_{y=c+1}^{d} k(x,y) f_i(y), \tag{116}$$

$\forall\ x \in [a,b]$, $i = 1, ..., n$.

Here $u : [a,b] \to \mathbb{R}_+$ continuous, and

$$v(y) = \int_a^b \frac{u(x) k(x,y)}{K(x)} dx, \tag{117}$$

$\forall\ y \in [c,d]_{\mathbb{Z}}$.

Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ convex and increasing per coordinate function. Then, by (44), we obtain

$$\int_a^b u(x)\,\Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right)dx \leq \tag{118}$$

$$\alpha\sum_{y=c}^{d-1} v(y)\,\Phi\left(\left|\overrightarrow{f(y)}\right|\right) + (1-\alpha)\sum_{y=c+1}^{d} v(y)\,\Phi\left(\left|\overrightarrow{f(y)}\right|\right).$$

(ii) Let $\mathbb{T}_1 = \mathbb{Z}$, $\mathbb{T}_2 = \mathbb{R}$. Then

$$\int_a^b \cdot\Diamond_\alpha x = \alpha\sum_{x=a}^{b-1}\cdot + (1-\alpha)\sum_{x=a+1}^{b}\cdot, \tag{119}$$

and

$$\int_c^d \cdot\Diamond_\alpha y = \int_c^d \cdot dy. \tag{120}$$

Assume $k : [a,b]_{\mathbb{Z}} \times [c,d] \to \mathbb{R}_+$, a continuous function. So here

$$K(x) = \int_c^d k(x,y)\,dy > 0, \quad \forall x \in [a,b]_{\mathbb{Z}}. \tag{121}$$

Consider $f_i : [c,d] \to \mathbb{R}$ continuous and

$$g_i(x) = \int_c^d k(x,y)\,f_i(y)\,dy, \quad \forall x \in [a,b]_{\mathbb{Z}}, \; i = 1,...,n. \tag{122}$$

Let $u : [a,b]_{\mathbb{Z}} \to \mathbb{R}_+$. Here it is

$$v(y) = \alpha\sum_{x=a}^{b-1}\frac{u(x)\,k(x,y)}{K(x)} + (1-\alpha)\sum_{x=a+1}^{b}\frac{u(x)\,k(x,y)}{K(x)}, \tag{123}$$

$\forall\ y \in [c,d]$.

Let $\Phi : \mathbb{R}_+^n \to \mathbb{R}$ convex and increasing per coordinate function. Then, by (44), we derive

$$\alpha\sum_{x=a}^{b-1} u(x)\,\Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right) + (1-\alpha)\sum_{x=a+1}^{b} u(x)\,\Phi\left(\frac{\left|\overrightarrow{g(x)}\right|}{K(x)}\right) \leq \tag{124}$$

$$\int_c^d v(y)\,\Phi\left(\left|\overrightarrow{f(y)}\right|\right)dy.$$

3 Appendix

We give

Proposition 47 *Let $t_n, t, \tau \in ([a,b] \cap \mathbb{T})$, $\mathbb{T}$ is time scale, and $t_n \to t$, then $\chi_{[a,t_n)}(\tau) \to \chi_{[a,t)}(\tau)$, delta Lebesgue measure Δ-a.e. in τ.*

Proof. Indeed we have the cases:

A) $t_n \leq t$.

We distinguish the subcases:

A1) $a \leq \tau < t_n$, then $\chi_{[a,t_n)}(\tau) = \chi_{[a,t)}(\tau) = 1$,

A2) $t \leq \tau \leq b$, then $\chi_{[a,t_n)}(\tau) = \chi_{[a,t)}(\tau) = 0$,

A3) $t_n \leq \tau < t$, then $\chi_{[a,t_n)}(\tau) = 0$, but $\chi_{[a,t)}(\tau) = 1$, and convergence fails.

As $n \to \infty$, the delta Lebesgue measure $\mu_\Delta([t_n, t)) = (t - t_n) \to 0$. Thus $\chi_{[a,t_n)}(\tau) \to \chi_{[a,t)}(\tau)$, Δ-a.e. in τ, when $t_n \to t$ from the left.

B) $t < t_n$.

We distinguish the subcases:

B1) $a \leq \tau < t$, then $\chi_{[a,t_n)}(\tau) = \chi_{[a,t)}(\tau) = 1$,

B2) $t_n \leq \tau \leq b$, then $\chi_{[a,t_n)}(\tau) = \chi_{[a,t)}(\tau) = 0$,

B3) $t \leq \tau < t_n$, then $\chi_{[a,t_n)}(\tau) = 1$, but $\chi_{[a,t)}(\tau) = 0$, and convergence fails.

As $n \to \infty$, $\mu_\Delta([t, t_n)) = (t_n - t) \to 0$. Thus $\chi_{[a,t_n)}(\tau) \to \chi_{[a,t)}(\tau)$, Δ-a.e. in τ, when $t_n \to t$ from the right, proving the claim. ∎

Proposition 48 *Let $y_n, y, x \in ([a,b) \cap \mathbb{T})$, $\mathbb{T}$ is time scale, such that $y_n \to y$, as $n \to \infty$. Then $\chi_{(y_n,b)}(x) \to \chi_{(y,b)}(x)$, Δ-a.e. in $x \in [a,b) \cap \mathbb{T}$.*

Proof. Indeed we have the cases:

A) Case of $y_n \leq y$, we distinguish the subcases:

A1) if $a \leq x \leq y_n$, then $\chi_{(y_n,b)}(x) = \chi_{(y,b)}(x) = 0$,

A2) if $y < x < b$, then $\chi_{(y_n,b)}(x) = \chi_{(y,b)}(x) = 1$,

A3) if $y_n < x \leq y$, then $\chi_{(y_n,b)}(x) = 1$, but $\chi_{(y,b)}(x) = 0$, and convergence fails.

We observe that (see [17]) $0 \leq \mu_\Delta((y_n, y]) = \sigma(y) - \sigma(y_n) = y - \sigma(y_n) \leq y - y_n \to 0$, proving $\chi_{(y_n,b)}(x) \to \chi_{(y,b)}(x)$, Δ-a.e. in $x \in [a,b) \cap \mathbb{T}$, when $y_n \to y$ from the left.

B) Case of $y < y_n$, we distinguish the subcases:

B1) if $a \leq x \leq y$, then $\chi_{(y_n,b)}(x) = \chi_{(y,b)}(x) = 0$,

B2) if $y_n < x < b$, then $\chi_{(y_n,b)}(x) = \chi_{(y,b)}(x) = 1$,

B3) if $y < x \leq y_n$, then $\chi_{(y_n,b)}(x) = 0$, but $\chi_{(y,b)}(x) = 1$, and convergence fails. But (by [17]) we have $\mu_\Delta((y, y_n]) = \sigma(y_n) - \sigma(y) = \sigma(y_n) - y \to 0$, since σ is rd-continuous and y is a right dense point ([12]). So we proved that $\chi_{(y_n,b)}(x) \to \chi_{(y,b)}(x)$, Δ-a.e. in $x \in [a,b) \cap \mathbb{T}$, when $y_n \to y$ from the right. That is proving the claim. ∎

References

[1] R. Agarwal, M. Bohner, D. O'Regan and A. Peterson, *Dynamic equations on time scales: a survey*, J. Comput. Appl. Math. 141 (2002), no. 1-2, 1-26.

[2] R. Agarwal, M. Bohner and A. Peterson, *Inequalities on time scales: a survey*, Math. Inequal. Appl. 4 (2001), no. 4, 535-557.

[3] G.A. Anastassiou, *Principles of Delta Fractional Calculus on Time Scales and Inequalities,* Mathematical and Computer Modelling, 52 (3-4): (2010), 556-566.

[4] G.A. Anastassiou, *Foundations of Nabla Fractional Calculus on Time Scales and Inequalities*, Computers & Mathematics with Applications, 59 (12), (2010), 3750-3762.

[5] G.A. Anastassiou, *Vectorial Hardy type fractional inequalities*, Bulletin of Tbilisi International Centre of Mathematics and Informatics, accepted, 2012.

[6] G.A. Anastassiou, *Rational Inequalities for integral operators under convexity*, Communications in Applied Analysis, accepted 2012.

[7] G.A. Anastassiou, *Vectorial Fractional Integral Inequalities with Convexity*, submitted, 2012.

[8] G.A. Anastassiou, *Integral operator inequalities on Time Scales*, submitted, 2012.

[9] M. Bohner, G.S. Guseinov, *Multiple Lebesgue integration on time scales*, Advances in Difference Equations, Vol. 2006, Article ID 26391, pp. 1-12, DOI 10.1155/ADE/2006/26391.

[10] M. Bohner, G.S. Guseinov, *Double integral calculus of variations on time scales*, Computers and Mathematics with Applications, 54 (2007), 45-57.

[11] M. Bohner, H. Luo, *Singular second-order multipoint dynamic boundary value problems with mixed derivatives*, Advances in Difference Equations, Vol. 2006, Article ID 54989, p. 1-15, DOI 10.1155/ADE/2006/54989.

[12] M. Bohner and A. Peterson, *Dynamic equations on time scales*, Birkäuser Boston, Boston, MA, 2001.

[13] M. Bohner and A. Peterson, *Advances in dynamic equations on time scales*, Birkäuser Boston, Boston, MA, 2003.

[14] C.C. Dinu, *Inequalities on Time Scales*, Ph.D. thesis, U. Craiova, Romania, 2008.

[15] T.S. Fergurson, *Mathematical Statistics*, Academic Press, New York, 1967.

[16] R.A.C. Ferreira, M.R. Sidi Ammi and D.F.M. Torres, *Diamond-alpha Integral inequalities on time scales*, Int. J. Math. Stat., 5 (no. A09), (2009), 52-59.

[17] G.S. Guseinov, *Integration on Time Scales*, J. Math. Anal. Appl., 285 (2003), 107-127.

[18] H.G. Hardy, *Notes on some points in the integral calculus*, Messenger of Mathematics, vol. 47, no. 10, 1918, 145-150.

[19] S. Hilger, *Analysis on measure chains - a unified approach to continuous and discrete calculus*, Results Math. 18 (1990), no. 1-2, 18-56.

[20] S. Hilger, *Differential and difference calculus - unified!*, Nonlinear Anal. 30 (1997), no. 5, 2683-2694.

[21] S. Iqbal, K. Krulic and J. Pecaric, *On an inequality of H.G. Hardy*, J. of Inequalities and Applications, Volume 2010, Article ID 264347, 23 pages.

[22] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo, *Theory and Applications of Fractional Differential Equations*, vol. 204 of North-Holland Mathematics Studies, Elsevier, New York, NY, USA, 2006.

[23] D. Mozyrska and D.F.M. Torres, *Diamond-alpha polynomial series on time scales*, Int. J. Math. Stat., 5 (2009), No. A09, 92-101.

[24] M. Perlman, *Jensen's Inequality for a Convex Vector valued function on an infinite-dimensional space*, J. of Multivariate Analysis, Vol. 4 (1974), 52-65.

[25] J.W. Rogers, Jr. and Q. Sheng, *Notes on the diamond-α dynamic derivative on time scales*, J. Math. Anal. Appl. 326 (2007), no. 1, 228-241.

[26] S.G. Samko, A.A. Kilbas and O.I. Marichev, *Fractional Integral and Derivatives: Theory and Applications*, Gordon and Breach Science Publishers, Yverdon, Switzerland, 1993.

[27] Q. Sheng, *Hybrid approximations via second order combined dynamic derivatives on time scales*, Electron. J. Qual. Theory Differ. Equ. 2007, No. 17, 13 pp. (electronic).

[28] Q. Sheng, M. Fadag, J. Henderson and J.M. Davis, *An exploration of combined dynamic derivatives on time scales and their applications*, Nonlinear Anal. Real World Appl. 7 (2006), no. 3, 395-413.

[29] M.R. Sidi Ammi, R.A.C. Ferreira and D.F.M. Torres, *Diamond-α Jensen's Inequality on Time Scales*, J. of Inequalities and Applications, Vol. 2008, article ID 576876, 13 pages.

[30] M.R. Sidi Ammi and D.F.M. Torres, *Combined Dynamic Grüss Inequalities on time scales*, J. of Mathematical Sciences, 161 (6), (2012), 792-802.

[31] Wing-Sum Cheung, *Generalizations of Hölder's Inequality*, Internat. J. of Mathematics and Mathematical Sciences, Vol. 26, No. 1, (2001), 7-10.

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 81-86, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

DIFFERENCE SCHEME FOR SOLUTION OF THE DIRICHLET'S PROBLEM

GALINA MEHDIYEVA AND AYDIN ALIYEV

Abstract. In present work for numerical solving Dirichlet's problem for Laplace's equations, there is applied difference scheme of higher accuracy order, allowing getting estimate of the error of order $O(h^4)$, only known data take part in this estimate. And this, in its turn, allows applying computational techniques to estimating the error and getting concrete error values.

Representation of difference problem (five-point scheme) solution for Laplace's equation on rectangle with aid of discrete analog of Fourier method, suggested and grounded in [2], is the main tool for creating its solutions both on whole net domain [4], [5], and on net segments [6], and what is more, not only rectangular, but and on multistep domain as well. Besides, error estimates (of second order regarding net step h), inferred in [2], are effective, i.e. depend only on known quantities.

In works [4], [5], though representation of nine-point scheme solution is used, but its convergence to solution of the differential problem is not grounded.

Effective estimates for nine-point scheme, not representing the solution by Fourier method, are obtained in [7].

1. Introduction

Let us denote through Π a rectangle with vertices $(0,0)$, $(1,0)$, $(1,b)$, $(0,b)$, where b-rational number. Let Γ-boundary of this rectangle.

We introduce net square by lines $x = x_i = ih, \quad y = y_j = jh \ \ (i = 0,1,...,1/h, j = 0,1,...,b/h)$, where $1/h$ and b/h – integer numbers. Let

$$\Pi_h = \{(x,y) : x = x_i = ih,\, i = 0,1,...,1/h, y = y_j = jh,\, j = 0,1,...,b/h\},$$

and Γ_h- set of net knots, lying on Γ. Let, further $\Gamma_{1_h} = \{(x,y) : x = ih,\, i = 1,...,n,\ y = 0\}$.

Consider Dirichlet's problem

$$\begin{cases} \Delta u = 0 \text{ on } \Pi, \\ u|_{\Gamma} = f, \end{cases} \tag{1.1}$$

where f – defined on Γ and has fifth derivative on each side of Γ.

Through $\Delta_h u$ denote Laplace's nine-point difference operator:

$$\Delta_h u = u(x+h,y+h) + u(x-h,y+h) + u(x-h,y-h) + u(x+h,y-h) +$$
$$+4[u(x+h,y) + u(x,y+h) + u(x-h,y) + u(x,y-h)] - 20u(x,y).$$

Through $Q(x,y)$ denote special polynomial of fourth degree.

$$Q(x,y) = a_{31}x^3y + a_{30}x^3 + a_{13}xy^3 + a_{03}y^3 + a_{21}x^2y +$$

Key words and phrases. Estimate error, Fourier method, difference scheme.

This research has been supported by the Science Development Foundation of Azerbaijan (EIF-2011-1(3)).

$$+a_{12}xy^2+a_{11}xy+a_{20}x^2+a_{02}y^2+a_{10}x+a_{01}y+a_{00},$$

where

$$a_{00}=f(0,0),\ a_{01}=f(0,b)-\frac{b}{6}f''_{yy}(0,b)-\frac{3}{b}f''_{yy}(0,0)-f(0,0),$$

$$a_{02}=\frac{1}{2}f''_{yy}(0,0),\quad a_{03}=\frac{1}{6b}f''_{yy}(0,b)-\frac{1}{6b}f''_{yy}(0,0),$$

$$a_{10}=f(1,0)-\frac{1}{6}f''_{xx}(1,0)-\frac{1}{3}f''_{xx}(0,0)-f(0,0),$$

$$a_{11}=\frac{1}{b}f''(1,b)-\frac{1}{6b}f''_{xx}(1,b)+\frac{1}{6b}f''_{xx}(1,0)+\frac{1}{3b}f''_{xx}(0,0)-\frac{1}{3b}f''_{xx}(0,b)-\frac{b}{6}f''_{yy}(1,b)+$$

$$+\frac{b}{6}f''_{yy}(0,b)-\frac{b}{3}f''_{yy}(1,0)-\frac{b}{6}f''_{yy}(0,0)-\frac{1}{b}f''(1,0)-f(0,b)+f(0,0),$$

$$a_{12}=\frac{1}{2}f''_{yy}(1,0)-\frac{1}{6}f''_{yy}(0,0),$$

$$a_{13}=\frac{1}{6b}f''_{yy}(1,b)-\frac{1}{6b}f''_{yy}(0,b)-\frac{1}{6b}f_{yy}(1,0)+\frac{1}{18b}f''_{yy}(0,0),$$

$$a_{20}=\frac{1}{2}f''_{xx}(0,0),\ a_{21}=\frac{1}{2b}f''_{xx}(0,b)-\frac{1}{2b}f''_{xx}(0,0),\ a_{30}=\frac{1}{6}f''_{xx}(1,0)-\frac{1}{6}f''_{xx}(0,0),$$

$$a_{31}=\frac{1}{6b}f''_{xx}(1,b)-\frac{1}{6b}f''_{xx}(1,0)+\frac{1}{6b}f''_{xx}(0,0)-\frac{1}{6b}f''_{xx}(0,b).$$

2. Results

It is easy to check that $Q(x,y)$ is a harmonic function on the vertices of the rectangle and the following conditions:

$$Q(p)=f(p),\quad \frac{\partial^2 Q(p)}{\partial x^2}=\frac{\partial^2 f(p)}{\partial x^2},\quad \frac{\partial^2 Q(p)}{\partial y^2}=\frac{\partial^2 f(p)}{\partial y^2}.$$

If we replace boundary function f with $\varphi_h=f-Q$, then we receive new problem, for which the estimate error is the same as for our problem. As estimate of the error can be represented in view of sum of four members, each, of which correspond to boundary quantities, equal to zero, except for one of four sides of the rectangle.

Through φ_h denote assigned function on Γ_{1h}:

$$\varphi_h=\begin{cases} 0, & \text{on vertixes } \Pi,\\ f-Q, & y=0.\end{cases}$$

So, the following difference scheme is considered:

$$\left.\begin{array}{l}\Delta_h u_h=0 \text{ on } \Pi_h,\\ u_h=\begin{cases}\varphi_h & \text{on} \quad \Gamma_{1h},\\ 0 & \text{on} \quad \Gamma_h\backslash\Gamma_{1h}\end{cases}\end{array}\right\} \tag{2.1}$$

It can be easily verified that the solutions of problems (1.1) and (2.1) are defined accordingly by formulas

$$u(x,y)=\sum_{n=1}^{\infty}c_n g(y,n\pi)\sin n\pi x, \tag{2.2}$$

$$u_h(x,y)=\sum_{n=1}^{1/h}\gamma_n g(y,\beta_n/h)\sin n\pi x, \tag{2.3}$$

where

$$\begin{aligned} c_n &= 2\int_0^1 \varphi_h(t)\sin n\pi t dt, \\ \gamma_n &= 2h\sum_{r=1}^{1/h} \varphi_h(rh)\sin n\pi rh, \\ g(y,z) &= \frac{sh(b-y)z}{shbz} \end{aligned} \tag{2.4}$$

and β_n is defined from

$$sh\frac{\beta_n}{2} = \frac{\sin nh\pi/2}{\sqrt{1-\frac{2}{3}\sin^2 nh\pi/2}}.$$

The solution of difference scheme (2.1) will be taken as an approximate solution to (1.1) at the nodal points. To evaluate the accuracy of the method, we first prove some properties of γ_n, C_n and g.

In order to estimate the approximate values of the Fourier coefficients C_n, we integrate by parts five times in the right member of the formulas for C_n and obtain:

$$\begin{aligned} C_n &= 2\int_0^1 (f-Q)\sin n\pi t dt = 2\left[-(f-Q)\left.\frac{\cos n\pi t}{n\pi}\right|_0^1 + \int_0^1 \frac{\cos n\pi t}{n\pi}(f'-Q')dt\right] \\ &= 2\left[\frac{1}{n\pi}+(f'-Q')\left.\frac{\sin n\pi t}{n\pi}\right|_0^1 - \frac{1}{n\pi}\int_0^1 \frac{\sin n\pi t}{n\pi}(f''-Q'')dt\right] \\ &= 2\left[-\frac{1}{n\pi}(f'-Q')\left.\frac{\cos n\pi t}{n\pi}\right|_0^1 + \frac{1}{(n\pi)^3}\int_0^1 \frac{\cos n\pi t}{n\pi}(f'''-Q''')dt\right] \\ &= 2\left[\frac{1}{(n\pi)^3}(f'''-Q''')\left.\frac{\sin n\pi t}{n\pi}\right|_0^1 - \frac{1}{(n\pi)^4}\int_0^1 \frac{\sin n\pi t}{n\pi}(f^{(IV)}-Q^{(IV)})dt\right] \\ &= 2\left[-\frac{1}{(n\pi)^5}(f^{(IV)}-Q^{(IV)})\left.\frac{\cos n\pi t}{n\pi}\right|_0^1 + \frac{1}{(n\pi)^5}\int_0^1 f^{(V)}(t)\cos n\pi t dt\right] \\ &= \frac{2}{(n\pi)^5}\left[(-1)^{n-1}f^{(IV)}(1)-f^{(IV)}(0)\right] + \frac{2}{(n\pi)^5}\int_0^1 f^{(V)}(t)\cos n\pi t dt. \end{aligned}$$

It follows that

$$|C_n| \leqslant Kn^{-5}, \tag{2.5}$$

where

$$K = \frac{2}{\pi^5}\left[\left|f^{(IV)}(1)\right| + \left|f^{(IV)}(0)\right|\right] + \frac{4}{\pi^6}\max\left|f^{(V)}(t)\right|.$$

Further, we note that ([5])

$$\gamma_n = C_n \;\; (n = 1, 2, ..., 1/n).$$

Then it is necessary to estimate the difference

$$|g(y, n\pi) - g(y, \beta_n/h)|.$$

The expression $\frac{\beta_n}{2} = arsh\frac{\sin(nh\pi/2)}{\sqrt{1-2/3\sin^2(nh\pi/2)}}$ can be expanded in a Taylor series in powers of $nh\pi/2$, we obtain

$$\frac{\beta_n}{2} = \frac{nh\pi}{2} + \left(\frac{nh\pi}{2}\right)^5 R,$$

where

$$R = \frac{1}{5!}\max_{0\leqslant x\leqslant \frac{nh\pi}{2}}\left|l^{(V)}(x)\right|, \quad l(x) = arsh\frac{\sin x}{\sqrt{1-\frac{2}{3}\sin^2 x}}.$$

Hence, given $R \leqslant 1/30$ that, we obtain:

$$|\beta_n - nh\pi| \leqslant \frac{(nh\pi)^5}{16\cdot 30}. \tag{2.6}$$

Further

$$\frac{\partial g(y,z)}{\partial z} = \frac{1}{2}sh^{-2}bz\left\{(2b-y)shyz - ysh(2b-y)z\right\}.$$

Given that sht/t is a non-decreasing function of $t > 0$, we have:

$$(2b-y)shyz \leqslant ysh(2b-y)z.$$

Therefore,

$$\left|\frac{\partial g(y,z)}{\partial z}\right| = \frac{y}{2}sh^{-2}bzsh(2b-y)z, \quad z \geqslant 0,\ y \geqslant 0. \tag{2.7}$$

For this expression we use the obvious inequality

$$shkt \leqslant \exp((k-1)t)sht \ \text{ for } \ 0 \leqslant k \leqslant 1$$

and

$$sht \geqslant \frac{1}{2}(1-\exp(-2t_1))\exp(t), \ \ t \geqslant t_1 > 0.$$

We also need the inequality

$$\frac{4n}{3} \leqslant \frac{\beta_n}{h} \leqslant \sqrt{3}n\pi \ \text{ for } \ 1 \leqslant n \leqslant \frac{1}{h}.$$

The right-hand side of this inequality follows from the expression β_n/h, and the left part can be proved as follows:

$$1 \leqslant \frac{1}{\sqrt{1-\frac{2}{3}\sin^2\frac{nh\pi}{2}}} \leqslant \sqrt{3},$$

$$sh\frac{\beta_n}{2} = \frac{nh\pi}{2}\frac{\sin\frac{nh\pi}{2}}{\frac{nh\pi}{2}}\frac{1}{\sqrt{1-\frac{2}{3}\sin^2\frac{nh\pi}{2}}} \geqslant \frac{nh\pi}{2}\frac{2}{\pi} = nh$$

and

$$sh\frac{\beta_n}{2} \leqslant \sqrt{3}\frac{nh\pi}{2} < \frac{\pi}{2}\sqrt{3}.$$

Given that shx/x non-decreasing function, we have:

$$\frac{sh\frac{\beta_n}{2}}{\frac{\beta_n}{2}} \leqslant \frac{sh[arsh\frac{\pi}{2}\sqrt{3}]}{arsh\frac{\pi}{2}\sqrt{3}} = \frac{\frac{\pi}{2}\sqrt{3}}{arsh\frac{\pi}{2}\sqrt{3}} \leqslant \frac{3}{2}.$$

Hence

$$\frac{\beta_n}{2} \geqslant \frac{3}{2}sh\frac{\beta_n}{2} \geqslant \frac{2}{3}nh.$$

Taking this into account in (2.7), we obtain:

$$\left|\frac{\partial g(y,z)}{\partial z}\right| \leqslant \frac{y}{2}4\left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2}\exp(-2bz)\exp(-yz)sh2bz \leqslant$$
$$\leqslant \left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2} y\exp(-yz),$$

for

$$\begin{cases} \sqrt{3}n\pi \geqslant z \geqslant \frac{\beta_n}{h} \geqslant \frac{4}{3}n, \\ 1 \leqslant n \leqslant 1/h, \\ 0 \leqslant y \leqslant b. \end{cases}$$

Using (2.6) we obtain:

$$\begin{aligned} |g(y,\beta_n/h)-g(y,n\pi)| &\leqslant \left(1-\exp\left(-\tfrac{8b}{3}\right)\right)^{-2} y\exp(-yz)\, h^4\frac{(n\pi)^5}{16\cdot 30} \\ &= \left(1-\exp\left(-\tfrac{8b}{3}\right)\right)^{-2}\tfrac{\pi^5}{480}y\exp\left(-\tfrac{4ny}{3}\right)n^5h^4. \end{aligned} \tag{2.8}$$

Completing the preliminary observations by the inequality

$$0 \leqslant g(y,z) \leqslant 1, \quad 0 \leqslant y \leqslant b \tag{2.9}$$

which is a consequence of the definition $g(y,z)$.

Now we can evaluate $|u-u_h|$. We write, using (2.2) and (2.3)

$$|u-u_h| \leqslant R_1+R_2,$$
$$R_1 = \sum_{n=1}^{1/h}|c_n|\,|g(y,n\pi)-g(y,\beta_n/h)|,$$
$$R_2 = \sum_{n=1+1/h}^{\infty}|c_n|\,g(y,n\pi).$$

From (2.5) and (2.9)

$$R_2 \leqslant K\sum_{n=1+1/h}^{\infty}n^{-5} \leqslant K\frac{h^4}{2}.$$

In order to assess R_1, observe that by (2.5)

$$|c_n| \leqslant Kn^{-5}.$$

Using this and (2.8), we obtain

$$\begin{aligned} |R_1| &\leqslant K\sum_{n=1}^{1/h}n^{-5}\left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2}\frac{\pi^5}{480}y\exp\left(-\frac{4ny}{3}\right)n^5h^4 \\ &= \left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2}K\frac{\pi^5}{480}yh^4\sum_{n=1}^{1/h}\left(\exp\left(-\frac{4ny}{3}\right)\right)^n \\ &\leqslant K\frac{\pi^5}{480}yh^4\left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2}\left(\exp\left(\frac{4y}{3}\right)-1\right)^{-1}. \end{aligned}$$

After

$$\frac{y}{\left(\exp\left(\frac{4y}{3}\right)-1\right)} \leqslant \frac{3}{4} \text{ for } y \geqslant 0,$$

we get

$$R_1 \leqslant K\frac{\pi^5}{640}\left(1-\exp\left(-\frac{8b}{3}\right)\right)^{-2}h^4.$$

So,

$$|u-u_h| \leqslant K\left\{0,5+\frac{\pi^5}{640}\left(1-\exp\left(-\frac{8b}{3}\right)\right)^2\right\}h^4.$$

References

[1] J.L. Walsh, D. Young, On the accuracy of the numerical solution of the Dirichlet problem by finite differences, *Jour. of Resea. of the Nat. Bur. of Standards*, 1953, V.51, No.6, pp.343–369.

[2] W. Wasow, On the truncation error in the solution of Laplace's equation by finite differences, *Jour. of Resea. of the Nat. Bur. of Standards*, 1952, V.48, pp.345–348.

[3] J.H. Giese, On the truncation error in a numerical solution of the Neumann problem for a rectangle, *Jour. Math. and Phys.*, 1958, v.37, No.2, pp. 169–177.

[4] S.E. Romanova, Economical method of solution of Laplace's difference equation on rectangle domains, *DAN SSSR*, 1980, v.252, No. 1, pp. 48–51 (Russian).

[5] S.E. Romanova, Economical methods of approximated solution of Laplace's difference equation on rectangle domains, *Zh. vuch. mat. i mat. fiz.*, 1983, v.23, No.3, pp.660–673 (Russian).

[6] E.A. Volkov, On an asymptotically fast approximate method of obtaining a solution of the Laplace difference equation on mesh segments, *Sov. Math. Dokl.*, 30, pp.642–646 (1984).

[7] E.A. Volkov, Effective estimates of the errors in solutions by the method of nets of boundary problems for the Laplace and Poisson equations on a rectangle and on certain triangles, *Proc. Steklov Inst. Math.*, 74, pp.57–90 (1966).

[8] A.Yu. Aliyev, On numerical solution nonlocal boundary values problems for elliptic equations, *Ph. D. thesis*, Baku, 1992 (Russian).

[9] A.Yu. Aliyev, Efficient error estimate of net method for a mixed boundary value problem, *Transactions Azer. Nat. Acad of Sciences.*, 2003, v.23, No. 4, pp.233–238.

[10] A.Yu. Aliyev, Numerical solution of one nonhomogeneous nonlocal mixed problem, *Proceedings of Institute of Mat. and Mech. Nat. Acad. of Sciences of Azerbaijan*, 2005, v. 22, pp. 179–186.

[11] A.Yu. Aliyev, G.Y. Mehdiyeva, Numerical solution of one nonlocal mixed problem, *Proceedings of International Conference on Applied Mathematics and Scientific Computing*, Paris, France, 2011, pp.244–246.

(G.Y. Mehdiyeva) Baku State University, Baku, Azerbaijan
E-mail address: aydin_aliyev66@mail.ru

(A.Yu. Aliyev) Baku State University, Baku, Azerbaijan
E-mail address: aydin_aliyev66@mail.ru

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 87-95, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ASYMPTOTIC DISTRIBUTION OF VECTOR VARIANCE STANDARDIZED VARIABLE WITHOUT DUPLICATION

ERNA T. HERDIANI AND MAMAN A. DJAUHARI

ABSTRACT. In recent years, the use of vector variance standardized variables as a measure of testing equality of correlation matrices has received much attention in wide range of statistics. This paper deals with a more economic measure of testing equality of correlation matrices, defined as vector variance standardized variables minus all duplication elements. For high dimensional data, this will increase the computational efficiency almost fifty percent compared to the original vector variance standardized variables. Its sampling distribution will be investigated to make its applications possible.

1. INTRODUCTION

The stability of correlation matrix is one of the most important problems in economic development and financial industry. We can see in the literature that, since the last decade, that problem can be found in a wide spectrum of applications. Its applications in property, real estate and asset businesses can be seen, for example, in eichholtz (1995), Lee(1998), Cooka el al. (2002) and Fischer (2007). Other applications such as in equity market, global market and risk management are presented by Meric and meric (1997), Tang (1998), Gande and Parsley (2002), Annaert et al. (2003), Ragea (2003), Da costa et al. (2005) and Goetzmann et al. (2005), Michelle et al. (2010). We can also find its application in parallel computation of high dimensional robust correlation matrices in Chilson et al. (2006).

Hypothesis testing about the stability of correlation structure is one of the fundamental issues in multivariate analysis. It is usually realized based on likelihood ratio test (LRT). See, for example, Box (1949), Lawley (1963), Kullback (1967), Aitkins (1968), and Jennrich (1970). In this research area is to have a better understanding of the covariance matrix of sample correlation matrix. Several research have been realized in this area. They were Browne and Shaphiro (1986) have initiated to study the asymtotic behavior of the covariance matrix of sample correlation coefficient. Its followed by Neudecker and wesselman (1990) and Neudecker (1996) who have reported the asymtotic behavior of the covariance matrix of sample correlation matrix. Its result used to testing the stability of correlation matrix when sample correlation matrices are depence, Schott (2001, 2007).

In this paper, we consider the case where sample correlation are independent. In this case, M statistic of Box (1949) and J-statistic of Jennrich (1970) used statistical test for testing the stability of correlation matrix. However, their application in practice is not without limitations. M statistic as computation of matrix determinant and J statistic involves matrix inversion. The former needs the condition that

---

Key words and phrases. Asymptotic distribution, correlation matrix, likelihood ratio test, vector variance standardized variables, vector variance.

2010 *AMS Math. Subject Classification.* Primary 40A05, 40A25; Secondary 45G05.

all sample correlation matrices are positive definite which is not always satisfied in practice. This condition is not apt for high dimension data sets because its computational efficiency becomes low. To handle this obstacles, Djauhari and Herdiani (2010) proposed a new statistical test based on what we call vector variance of standardized variables (VVSV).

The proposed test is constucted based on vector variance (VV). The role of VV could be increasing the computational efficiency of fast minimum covariance determinant (FMCD) algorithm, Herwindiati et al. (2007). Let P be the correlation matrix of the population under studi. Vector Variance of Standardized Variable (VVSV) is the trace of the squared correlation matrix population, i.e. $VVSV = Tr(P^2) = \|vec(P)\|^2$. It is the sum of square of all elements of P. The fact that P is symmetric, it is no need to involve all elements of P. The element of its upper (lower) triangular matrix are sufficient. This is what we want to discuss in this present paper. The rest of the paper is organized as follows. In section 2, the problem formulation will be presented. Later on, in section 3, we discuss the asymptotic distributional properties of vector variance of standardized variables without duplication or modified vector variance of standardized variables (MVVSV), i.e. VVSV without all duplicated elements. Our approach will be based on the notions of vec operator and commutation matrix. Additional remarks in section 4 will close this paper.

2. Problem Formulation

Let X is a random vector p dimension with definite positive covariance matrix Σ. By using vec operator, see Muirhead (1982), El Maache and Lepage (1998), Schott (1997, 2001) and Djauhari (2007), vector variance (VV) of X is simply $\|vec(\Sigma)\|^2$. It is a multivariate variability measure like Generalized Variance (GV). See Djauhari (2007) for an in depth discussion on VV and its asymtotic behavior. In what follows our attention will be focused on the case where all variables are standardized. Let X is a random vector p dimension with definite positive covariance matrix Σ. By using vec operator, see Muirhead (1982), El Maache and Lepage (1998), Schott (1997, 2001) and Djauhari (2007), vector variance (VV) of X is simply $\|vec(\Sigma)\|^2$. It is a multivariate variability measure like Generalized Variance (GV). See Djauhari (2007) for an in depth discussion on VV and its asymtotic behavior. In what follows our attention will be focused on the case where all variables are standardized.

Let Z be the random vector where its k-th component is the standardized version of the k-th component of X; $k = 1, 2, ..., p$. The covariance matrix P of Z is the correlation matrix of X. We call the parameter $\|vec(P)\|^2$ vector variance of the standardized variables (VVSV). Now, let $Z_1, Z_2, ..., Z_n$ be a random sample of size n from Z with covariance matrix P. If R is the sample correlation matrix, then we call $\|vec(R)\|^2$ sample VVSV.

Let R be a symmetric matrix, there are $\frac{(p-1)p}{2}$ elements of R which are doubly counted in $\|vec(R)\|^2$. This is the first problem that we want to discuss in this present paper. More specifically, instead of using the vec operator, we propose to use further operator which will transform the lower triangular part R_L of R into the vector $v(R_L)$ of dimension $\frac{(p-1)p}{2}$ by stacking its column one after another. From now on we call the parameter $\|v(R_L\|^2$ vector variance standardized variables without duplication or simply modified vector variance Standardized Variables (MVVSV).

It is clear that MVVSV is more economic than VVSV. The second problem is to investigate the distributional properties of sample MVVSV. This will guide us to a more economic hypothesis testing about the stability of covariance structure mentioned in section 1.

The solution for the first problem is given by schott (1997). Let $\overrightarrow{u_{ij}}$ be a vector dimension $\frac{(p-1)p}{2}$ defined as follows. The $\{(j-1)p+i-\frac{j(j+1)}{2}\}-th$ component is equal to 1 and 0 otherwise; $i=1,2,...,p$ and $j=1,2,...,i-1$. Let also E_{ij} be a matrix of size $(p\times p)$ its $(i,j)-th$ element is equal to 1 and 0 otherwise. Schott (1997) gives us the following result.

$$L_p^t = \sum_{i>j}\{vec(E_{ij})\}\overrightarrow{u_{ij}^t} \tag{2.1}$$

and

$$v(R_L) = L_p vec(R). \tag{2.2}$$

Example 2.1. *We will illustrate how to transform A is a square matrix or a correlation matrix to $v(A_L)$. Let* $A=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{pmatrix}=\begin{pmatrix} 1 & a_{21} & a_{31}\\ a_{21} & 1 & a_{32}\\ a_{31} & a_{32} & 1\end{pmatrix}$, *if A is a correlation matrix, where A is a symmetric matrix, $a_{ij}=a_{ij}$ for $i\neq j$ and $a_{11}=a_{22}=a_{33}=1$, we will determine of $v(A_L)$.*

$$\begin{aligned}
L_p^t &= \sum_{i>j}\{vec(E_{ij})\}\overrightarrow{u_{ij}^t}\\
L_p^t &= \{vec(E_{21})\}\overrightarrow{u_{21}^t}+\{vec(E_{31})\}\overrightarrow{u_{31}^t}+\{vec(E_{32})\}\overrightarrow{u_{32}^t}\\
L_p^t &= \{vec\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}\}\begin{pmatrix}1\\0\\0\end{pmatrix}^t\\
&\quad +\{vec\begin{pmatrix}0&0&0\\0&0&0\\1&0&0\end{pmatrix}\}\begin{pmatrix}0\\1\\0\end{pmatrix}^t\\
&\quad +\{vec\begin{pmatrix}0&0&0\\0&0&0\\0&1&0\end{pmatrix}\}\begin{pmatrix}0\\0\\1\end{pmatrix}^t,
\end{aligned}$$

$$L_p^t = \begin{pmatrix} 0&0&0\\1&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0 \end{pmatrix} + \begin{pmatrix} 0&0&0\\0&0&0\\0&1&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0 \end{pmatrix} + \begin{pmatrix} 0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&0&1\\0&0&0\\0&0&0\\0&0&0 \end{pmatrix} = \begin{pmatrix} 0&0&0\\1&0&0\\0&1&0\\0&0&0\\0&0&0\\0&0&1\\0&0&0\\0&0&0\\0&0&0 \end{pmatrix},$$

$$L_p = \begin{pmatrix} 0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&0&0&1&0&0&0 \end{pmatrix}.$$

and

$$v(A_L) = L_p vec(A).$$

Hence

$$v(A_L) = \begin{pmatrix} 0&1&0&0&0&0&0&0&0\\0&0&1&0&0&0&0&0&0\\0&0&0&0&0&1&0&0&0 \end{pmatrix} \begin{pmatrix} 1\\a_{21}\\a_{31}\\a_{21}\\1\\a_{32}\\a_{31}\\a_{32}\\1 \end{pmatrix} = \begin{pmatrix} a_{21}\\a_{31}\\a_{32} \end{pmatrix}.$$

3. Distributional Properties of Sample VVSV without Duplication

Let $X_1, X_2, ..., X_n$ be a sample random of size n drawn from a p-variate normal distribution $N_p(\mu, \Sigma)$. Then

$$\sqrt{n-1}\{vec(R) - vec(P)\} \xrightarrow{d} N_{p^2}(0, \Gamma),$$

where

- **a:** $\Gamma = 2M_p\Phi M_p$ with $M_p = \frac{1}{2}(I_{p^2} + K_{pp})$,
- **b:** K_{pp} is the commutation matrix of size $(p^2 \times p^2)$,
- **c:** $\Phi = \{I_{p^2} - (I_p \bigotimes P)\Lambda_p\}\{P \bigotimes P\}\{I_{p^2} - (I_p \bigotimes P) \bigwedge_p\}$,
- **d:** $\Lambda_p = \sum_{i=1}^{p}(\overrightarrow{e_i}\overrightarrow{e_i^t} \bigotimes \overrightarrow{e_i}\overrightarrow{e_i^t})$, $\overrightarrow{e_i}$ is the $i-th$ column vector of identity matrix I_p of size $p \times p$.

From this result, if the transformation is used on R, by using the result in Muirhead (1982, p. 5) we have:

Proposition 3.1.

$$\sqrt{n-1}\{v(R_L) - v(P_L)\} \xrightarrow{d} N_{p^2}(0, \Lambda),$$

where $k = \frac{p(p-1)}{2}$ *and* $\Lambda = 2L_pM_p\Phi M_pL_p^t$

Proof. $v(R_L) = L_p vec(R)$. Muirhead (1982) said if C is a real matrix with size $p\times q$, $q \leq p$, $rank(C) = q$ and $vec(R) \sim N_{p^2}(vec(P), \frac{\Gamma}{n-1})$, then $Cvec(R)\ N_p(Cvec(P), C\frac{\Gamma}{n-1}C^t)$. So, if $C = L_p$ then we can get

$$v(R_L) = L_p vec(R) \sim N_k(L_p vec(P), L_p \frac{\Gamma}{n-1} L_p^t),$$

where

$$E(v(R_L)) =)L_p vec(P) = v(P_L)$$

and

$$\begin{aligned} var(v(R_(L))) &= var(L_p vec(R)) \\ &= L_p \frac{\Gamma}{n-1} L_p^t \\ &= \frac{1}{n-1} L_p \Gamma L_p^t \\ &= \frac{1}{n-1} L_p (2M_p \Phi M_p) L_p^t \\ &= \frac{2}{n-1} L_p M_p \Phi M_p L_p^t. \end{aligned}$$

Therefore, we can conclude that $v(R_L)$ be have distribution multivariate normal with p-variate with mean vector $v(P_L)$ and variance covariance matrix $\frac{2}{n-1} L_p M_p \Phi M_p L_p^t$. □

According to Corollary 3.2 and Proposition 3.3 in Djauhari (2007), if we define $u(v(R_L)) = \|v(R_L)\|^2$, then we arrive at the following proposition about the asymptotic distribution of sample VVSV without duplication.

Proposition 3.2.

$$\sqrt{n-1}\|v(R_L)\|^2 - \|v(P_L)\|^2 \underset{\rightarrow}{d} N(0, \sigma^2),$$

where

$$\sigma^2 = 4(v(P_L))^t L_p \Gamma (L_p)^t (v(P_L)) \text{ and } \Gamma = 2M_p \Phi M_p \text{ with } M_p = \frac{1}{2}(I_{p^2} + K_{pp}).$$

Proof. Based on Proposition 3.1,

$$v(R_L) = L_p vec(R) \sim N_k(v(P_L), \frac{\Lambda}{n-1}),$$

where $k = \frac{p(p-1)}{2}$ and $\Lambda = 2L_p M_p \Phi M_p L_p^t$. Now,

$$\begin{aligned} \|v(R_L)\|^2 &= (v(R_L))^t v(R_L) \\ &= u(v(R_L)) \end{aligned}$$

because $v(R_L)$ is a vector. □

Theorem 3.3. *If sequence* $\{\overrightarrow{X_n}\} \xrightarrow{p} \overrightarrow{c}$, $\overrightarrow{c}$ *be a vector constant in* R^p *and* $\{\overrightarrow{X_n}\} \xrightarrow{d} N_p(\overrightarrow{c}, \Sigma)$ *then random variable* $Y_n = u(\overrightarrow{X_n}) \xrightarrow{d} N(\mu_Y, \sigma_Y^2)$, *where* $\mu_Y \xrightarrow{p} u(\overrightarrow{c})$ *and* $\sigma_Y^2 \xrightarrow{p} (\frac{\delta u(\overrightarrow{c})}{\delta \overrightarrow{X_n}})^t \Sigma (\frac{\delta u(\overrightarrow{c})}{\delta \overrightarrow{X_n}})$.

Proof. Let

$$Y_n = u(\overrightarrow{X_n}) = u(\overrightarrow{c}) + (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c}) + R_\xi.$$

If $\overrightarrow{X_n} \xrightarrow{p} \overrightarrow{c}$ then quadratic form $R_\xi \to 0$ faster than linear form $(\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c})$. Y_n and $V_n = u(\overrightarrow{c}) + (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c})$ convergence in distribution to same distribution. Furthermore, $\overrightarrow{X_n} \xrightarrow{d} N_p(\overrightarrow{c}, \Sigma)$ then Serfling (1980, p. 24) written $V_n \xrightarrow{d} N(\mu_v, \sigma_v^2)$ with $\mu_v = E(V_n) = E(u(\overrightarrow{c}) + (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c}))$ convergence in probability to $E(u(\overrightarrow{c})) = u(\overrightarrow{c})$. This case is caused by $\overrightarrow{X_n} \xrightarrow{p} \overrightarrow{c}$ which means $E((\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c})) \xrightarrow{p} \overrightarrow{0}$. So :

$$\begin{aligned}
\sigma_v^2 &= var(V_n) \\
&= var(u(\overrightarrow{c}) + (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t(\overrightarrow{X_n} - \overrightarrow{c}))) \\
&= (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t var((\overrightarrow{X_n} - \overrightarrow{c}))(\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}}) \\
&= (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t var(\overrightarrow{X_n})(\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}}) \\
&= (\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}})^t \Sigma(\frac{\partial u(\overrightarrow{c})}{\partial \overrightarrow{X_n}}).
\end{aligned}$$

Let

$$\begin{aligned}
u(v(R_L)) &= \|v(R_L)\| \\
&= (v(R_L))^t(v(R_L)) \\
&= (L_p vec(R))^t(L_p vec(R))
\end{aligned}$$

and $vec(R) \xrightarrow{p} vec(P)$, see El Maachee (1997), so that $L_p vec(R) \xrightarrow{p} L_p vec(P)$. In other words, $v(R_L) \xrightarrow{p} v(P_L)$. Therefore,

$$\begin{aligned}
Y_n &= u(v(R_L)) \\
&= u(v(P_L)) + (\frac{\partial u(v(P_L))}{\partial v(R_L)})^t(v(R_L) - v(P_L)) + R_\xi,
\end{aligned}$$

where $R_\xi = \frac{1}{2}(v(R_L) - v(P_L))^t A_\xi (v(R_L) - v(P_L))$. So,

$$\begin{aligned}
Y_n &= u(v(R_L)) \\
&= \|v(R_L\|^2 \\
&= (v(R_L)^t(v(R_L) \\
&= (L_p vec(R))^t(L_p vec(R)) \\
&\xrightarrow{d} N(\mu_Y, \sigma_Y^2),
\end{aligned}$$

where

$$\begin{aligned}
\mu_Y &\xrightarrow{p} u(v(P_L)) \\
&= \|v(P_L)\|^2
\end{aligned}$$

and

$$\begin{aligned}
\sigma_Y^2 &\xrightarrow{p} \frac{1}{n}((\frac{\partial u(v(R_L))}{\partial v(R_L)})^t L_p \Lambda L_p^t((\frac{\partial u(v(R_L))}{\partial v(R_L)}) \\
&\xrightarrow{p} \frac{1}{n}(2(v(P_L)))^t L_p \Lambda L_p^t(2(v(P_L))), v(R_L) = v(P_L) \\
&\xrightarrow{p} \frac{4}{n}(v(P_L))^t L_p \Lambda L_p^t(v(P_L)) \\
&\xrightarrow{p} \frac{4}{n}(L_p vec(P))^t L_p \Lambda L_p^t(L_p vec(P)) \\
&\xrightarrow{p} \frac{4}{n}(L_p^t L_p vec(P))^t \Lambda(L_p^t L_p vec(P)) \\
&\xrightarrow{p} \frac{4}{n}(vec(\Psi))^t \Lambda(vec(\Psi)), vec(\Psi) = L_p^t L_p vec(P).
\end{aligned}$$

So it can be concluded that

$$\|v(R_L)\|^2 = (v(R_L))^t(v(R_L) \xrightarrow{d} N(\|v(P_L)\|^2, \frac{4}{n}(vec(\Psi))^t \Lambda(vec(\Psi))),$$

where $vec(\Psi) = L_p^t L_p vec(P)$. □

This proposition is seemingly complicated to be used in application because the variance of sample VVSV without duplication $\|v(R_L)\|^2$ is $4(v(P_L))^t L_p \Gamma(L_p)^t$. It is involves multiplication of large size matrix Γ be $(p^2 \times p^2)$. However, the future works to helps us; need to simplify the computation of that variance.

4. Additional Remarks

If vector variance $vec(P)$ is of dimension p^2, $v(P_L)$ is of dimension $k = \frac{p(p-1)}{2}$. This gain is too good to be neglected. Furthermore, Propositions 3.1 and 3.2 have made possible the application of modified vector variance standardized variable.

Acknowledgement

This research is partially supported by the Institute Technology Bandung 2008 research competitive grant under contract number: 040/K01.7/PL/2008. The authors thank the institute technology Bandung, for that support.

References

[1] F.B. Alt, and N.D. Smith, Multivariate process control, in *Handbook of Statistics* (P.R. Krishnaiah and C.R. Rao, eds.), Elsevier Sciences Publishers, 1988, pp.333-351

[2] M.J. Anderson, Distance based test for homogeneity of multivariate dispersion, *Biometrics*, 62, 245 -253 (2006)

[3] T.W. Anderson, *An Introduction to multivariate statistical analysis*, John Wiley and Sons, Inc., 1958.

[4] J. Chilson, Ng. R. Wagner, and R. Zamar, Parallel Computation of high dimensional robust correlation and covariance matrices, *algorithmica*, 45(3), 403 - 431 (2006).

[5] R. Cleroux, Multivariate Association and Inference Problems in Data Analysis, *proceedings of the fifth international symposium on data analysis and informatics*, vol.1, Versailles, France, (1987).

[6] Jr. N. Da costa, S. Nunes, P. Ceretta, and S. Da Silva, Stock market co-movements revisited, *Economics Bulletin*, 7(3), 1-9 (2005).

[7] M. A. Djauhari,Improved Monitoring of Multivariate Variability, *Journal of Quality Technology*, 37(1), 32-39 (2005).

[8] M. A. Djauhari, A Measure of Data Concentration, *Journal of Probability and Statistics*, 2(2), 139-155 (2007).
[9] M. A. Djauhari, M. Mashuri, and D. E. Herwindiati, Multivariate Process Variability Monitoring, *Communications in statistics - Theory and Methods*, 37(1), 1742 - 1754 (2008).
[10] H. El Maache, and Y. Lepage, Measures d'Association Vectorielle Basées sur une Matrice de Corrélation, *Revue de Statistique Appliquée*, 46(4), 27-43 (1998).
[11] Y. Escoufier, Le traitement des variables vectorielles, *Biometrics*, 29, 751-760 (1973).
[12] A. K. Gupta, and J. Tang, Distribution of likelihood ratio statistic for testing equality of covariance matrices of multivariate Gaussian models, *Biometrika*, 71(3), 555-559 (1984).
[13] D. E. Herwindiati, M. A. Djauhari, and M. Mashuri, Robust Multivariate Outlier Labeling, *Communications in statistics - Computation and Simulation*, 36(6), 1287 - 1294 (2007).
[14] M. Hubert, P.J. Rouddeeuw, and S. van Aelst, Multivariate outlier detection and robustness, in *Handbook of Statistics*, Elsevier B.V., 2005, pp.263-302
[15] M. B. C. Khoo, and S. H. Quah, Multivariate control chart for process dispersion based on individual observations, *Quality Engineering*, 15(4), 639-643 (2003).
[16] M. B. C. Khoo, and S. H. Quah, Alternatives to the multivariate control chart for process dispersion, *Quality Engineeriong*, 16(3), 423-435 (2004).
[17] O. Ledoit and M. Wolf, Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size, *The Annals of Statistics*, 30(4), 1081-1102 (2002).
[18] K.V. Mardia and J.M. Kent, *Multivariate analysis*, Academic press, 2000.
[19] T. Michele, L. Fabrizio, and N. M. Rosario, Correlation, hierarchies, and networks in financial markets, *Journal of Economic Behavior and Organization*, 75, 40-58 (2010).
[20] D.C. Montgomery, *Introduction to statistical quality control*, John Wiley and Sons, Inc., 2001.
[21] R.J. Muirhead, *Aspects of multivariate statistical theory*, John Wiley and Sons, Inc., 1982.
[22] V. Ragea, Testing correlation stability during hectic financial markets, *Journal of financial market and Portfolio Management*, 17(3), 289-308 (2003).
[23] P.J.Rosseeuw, Multivariate estimation with high breakdown point, in *mathematical statistics and applications* (B.W. Grossman, G. Pflug, I. Vincze and W. Wertz, eds.), D. Reidel Publishing Company, 1985, pp.283-297
[24] P.J. Rosseeuw and A.M. Leroy, *Robust regression and outlier detection*, John Wiley and Sons, Inc., 1987.
[25] P.J. Rosseeuw and M. Hubert, Regression Depth, *Journal of the American Statistical Association*, 94, 388-402 (1999).
[26] P.J. Rosseeuw and K. van Driessen, A fast algorithm for the minimum covariance determinant estimatior, , *Technometrics*, 41, 212 - 223 (1999).
[27] J. Schafer and K. Strimmer, A Shrinkage Approach to large scale covariance matrix estimation and implications for functional genomics, *Statistical Applications in genetics and molecular biology*, 4, 1-30 (2005).
[28] J. R. Schott, *Matrix analysis for statistics*, John Wiley and Sons, Inc., 1997.
[29] J. R. Schott, Some tests for the equality of covariance matrices, *journal of statistical planning and inference*, 94, 25 - 36 (2001).
[30] G.A.F. Seber, *Multivariate Observations*, John Wiley and Sons, Inc., 1984.
[31] R.J. Serfling, *Approximation Theorems of mathematical statistics*, John Wiley and Sons, Inc., 1980.
[32] J.H. Sullivan and W.H. Woodall, A Comparison of multivariate control charts for individual observations, *Journal of Quality Technology*, 28(4), 398-408 (1996).
[33] J.H. Sullivan, Z.G. Stoumbos, R.L. Mason and J.C. Young, Step down analysis for changes in the covariance matrix and other parameters, *Journal of Quality Technology*, 39(1), 66-84 (2007).
[34] G.Y.N. Tang, The Intertemporal stability of the covariance and correlation matrices of Hongkong Stock Returns, *Applied financial economics*, 8, 359-365 (1998).
[35] M. Werner, *Identification of multivariate outliers in large data sets, PhD dissertation*, University of Colorado at Denver, 2003.
[36] S.J. Wierda, Multivariate statistical process control, *Recent results and directions for future research, statistica neerlandica*, 48(2), 147-168 (1994).
[37] W.H. Woodall and D.C. Montgomery, Research issues and ideas in statistical process control, *Journal of Quality Technology*, 31(4), 376-386 (1999).

(E.T. Herdiani) University of Hasanuddin, Makassar, Indonesia
E-mail address: herdiani.erna@gmail.com

(M.A. Djauhari) University Teknologi Malaysia, Johor Bahru, Malaysia
E-mail address: maman@math.itb.ac.id

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 96-111, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

Nabla Fractional Calculus on Time Scales and Inequalities

George A. Anastassiou
Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Abstract

Here we develop the Nabla Fractional Calculus on Time Scales. Then we produce related integral inequalities of types: Poincaré, Sobolev, Opial, Ostrowski and Hilbert-Pachpatte. Finally we give inequalities applications on the time scales $\mathbb{R}$, $\mathbb{Z}$.

2000 AMS Subject Classification : 26D15, 26A33, 39A12, 93C70.
Keywords and phrases: Fractional Calculus on time scales, Nabla Poincaré inequality, Nabla Sobolev inequality, Nabla Opial inequalities, Nabla Ostrowski inequality, Nabla Hilbert-Pachpatte inequality, fractional inequalities.

1 Background and Foundation Results

For the basics on time scales we follow [1], [2], [3], [4], [9], [11], [13], [6], [7], [10].

By [15], p. 256, for $\mu, \nu > 0$ we have that

$$\int_t^x \frac{(x-s)^{\mu-1}}{\Gamma(\mu)} \frac{(s-t)^{\nu-1}}{\Gamma(\nu)} ds = \frac{(x-t)^{\mu+\nu-1}}{\Gamma(\mu+\nu)}, \tag{1}$$

1

where Γ is the gamma function.

Here we consider time scales T such that $T_k = T$.

Consider the coordinatewise ld-continuous functions $\widehat{h}_\alpha : T \times T \to \mathbb{R}$, $\alpha \geq 0$, such that $\widehat{h}_0(t,s) = 1$,

$$\widehat{h}_{\alpha+1}(t,s) = \int_s^t \widehat{h}_\alpha(\tau,s)\,\nabla\tau, \tag{2}$$

$\forall\ s,t \in T$.

Here ρ is the backward jump operator and $\nu(t) = t - \rho(t)$.

Furthermore for $\alpha, \beta > 1$ we assume that

$$\int_{\rho(u)}^t \widehat{h}_{\alpha-1}(t,\rho(\tau))\,\widehat{h}_{\beta-1}(\tau,\rho(u))\,\nabla\tau = \widehat{h}_{\alpha+\beta-1}(t,\rho(u)), \tag{3}$$

valid for all $u,t \in T : u \leq t$.

In the case of $T = \mathbb{R}$; then $\rho(t) = t$, and $\widehat{h}_k(t,s) = \frac{(t-s)^k}{k!}$, $k \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$, and define

$$\widehat{h}_\alpha(t,s) = \frac{(t-s)^\alpha}{\Gamma(\alpha+1)}, \quad \alpha \geq 0.$$

Notice that

$$\int_s^t \frac{(\tau-s)^\alpha}{\Gamma(\alpha+1)}d\tau = \frac{(t-s)^{\alpha+1}}{\Gamma(\alpha+2)} = \widehat{h}_{\alpha+1}(t,s),$$

fulfilling (2).

Furthermore we observe that $(\alpha, \beta > 1)$

$$\int_u^t \widehat{h}_{\alpha-1}(t,\tau)\,\widehat{h}_{\beta-1}(\tau,u)\,d\tau = \int_u^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)}\frac{(\tau-u)^{\beta-1}}{\Gamma(\beta)}d\tau$$

$$\overset{(\text{by }(1))}{=} \frac{(t-u)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} = \widehat{h}_{\alpha+\beta-1}(t,u),$$

fulfilling (3).

By Theorem 2.2 of [14], we have for $k, m \in \mathbb{N}_0$ that

$$\int_{t_0}^t \widehat{h}_k(t,\rho(\tau))\,\widehat{h}_m(\tau,t_0)\,\nabla\tau = \widehat{h}_{k+m+1}(t,t_0). \tag{4}$$

Let $T = \mathbb{Z}$, then $\rho(t) = t-1$, $t \in \mathbb{Z}$. Define $t^{\overline{0}} := 1$, $t^{\overline{k}} := t(t+1)\dots(t+k-1)$, $k \in \mathbb{N}$, and by (2) we have $\widehat{h}_k(t,s) = \frac{(t-s)^{\overline{k}}}{k!}$, $s,t \in \mathbb{Z}$, $k \in \mathbb{N}_0$.

Here $\int_{t_0}^{t} \nabla\tau = \sum_{t_0+1}^{t}$.

Therefore by (4) we get

$$\sum_{\tau=t_0+1}^{t} \frac{(t-\tau+1)^{\overline{k}}}{k!} \frac{(\tau-t_0)^{\overline{m}}}{m!} = \frac{(t-t_0)^{\overline{k+m+1}}}{(k+m+1)!},$$

which results into

$$\sum_{\tau=t_0}^{t} \frac{(t-\tau+1)^{\overline{k-1}}}{(k-1)!} \frac{(\tau-t_0+1)^{\overline{m-1}}}{(m-1)!} = \frac{(t-t_0+1)^{\overline{k+m-1}}}{(k+m-1)!}, \tag{5}$$

confirming (3).

Next we follow [5].

Let $a, \alpha \in \mathbb{R}$, define $t^{\overline{\alpha}} = \frac{\Gamma(t+\alpha)}{\Gamma(t)}$, $t \in \mathbb{R}-\{\dots,-2,-1,0\}$, $N_a = \{a, a\pm1, a\pm 2, \dots\}$, notice $N_0 = \mathbb{Z}$, $0^{\overline{\alpha}} = 0$, $t^{\overline{0}} = 1$, and $f : N_a \to \mathbb{R}$. Here $\rho(s) = s-1$, $\sigma(s) = s+1$, $\nu(t) = 1$. Also define

$$\nabla_a^{-n} f(t) = \sum_{s=a}^{t} \frac{(t-\rho(s))^{\overline{n-1}}}{(n-1)!} f(s), \quad n \in \mathbb{N},$$

and in general

$$\nabla_a^{-\nu} f(t) = \sum_{s=a}^{t} \frac{(t-\rho(s))^{\overline{\nu-1}}}{\Gamma(\nu)} f(s),$$

where $\nu \in \mathbb{R} - \{\dots,-2,-1,0\}$.

Here we set

$$\widehat{h}_\alpha(t,s) = \frac{(t-s)^{\overline{\alpha}}}{\Gamma(\alpha+1)}, \quad \alpha \geq 0.$$

We need

Lemma 1 *Let $\alpha > -1$, $x > \alpha + 1$. Then*

$$\frac{\Gamma(x)}{\Gamma(x-\alpha)} = \frac{1}{(\alpha+1)} \left(\frac{\Gamma(x+1)}{\Gamma(x-\alpha)} - \frac{\Gamma(x)}{\Gamma(x-\alpha-1)} \right).$$

3

Proposition 2 *Let $\alpha > -1$. It holds*

$$\int_s^t \frac{(\tau - s)^{\overline{\alpha}}}{\Gamma(\alpha+1)} \nabla \tau = \frac{(t-s)^{\overline{\alpha+1}}}{\Gamma(\alpha+2)}, \quad t \geq s.$$

That is $\widehat{h}_\alpha$, $\alpha \geq 0$, on N_a confirm (2).

Next for $\mu, \nu > 1$, $\tau < t$, from the proof of Theorem 2.1 ([5]) we get that

$$\sum_{s=\tau}^{t} \frac{(t-\rho(s))^{\overline{\nu-1}}}{\Gamma(\nu)} \frac{(s-\rho(\tau))^{\overline{\mu-1}}}{\Gamma(\mu)} = \frac{(t-\rho(\tau))^{\overline{\nu+\mu-1}}}{\Gamma(\mu+\nu)},$$

where $\tau \in \{a, ..., t\}$.

So for $t, t_0 \in N_a$ with $t_0 < t$ we obtain

$$\sum_{\tau=t_0}^{t} \frac{(t-\tau+1)^{\overline{\nu-1}}}{\Gamma(\nu)} \frac{(\tau-t_0+1)^{\overline{\mu-1}}}{\Gamma(\mu)} = \frac{(t-t_0+1)^{\overline{\nu+\mu-1}}}{\Gamma(\mu+\nu)}, \tag{6}$$

that is confirming (3) fractionally on the time scale $T = N_a$.

Notice also here that

$$\int_a^b f(t) \nabla t = \sum_{t=a+1}^{b} f(t).$$

So fractional conditions (2) and (3) are very natural and common on time scales.

For $\alpha \geq 1$ we define the time scale ∇-Riemann-Liouville type fractional integral $(a, b \in T)$

$$J_a^\alpha f(t) = \int_a^t \widehat{h}_{\alpha-1}(t, \rho(\tau)) f(\tau) \nabla \tau, \tag{7}$$

(by [8] the last integral is on $(a, t] \cap T$)

$$J_a^0 f(t) = f(t),$$

where $f \in L_1([a,b] \cap T)$ (Lebesgue ∇-integrable functions on $[a,b] \cap T$, see [6], [7], [10]), $t \in [a,b] \cap T$.

Notice $J_a^1 f(t) = \int_a^t f(\tau) \nabla \tau$ is absolutely continuous in $t \in [a,b] \cap T$, see [8].

Lemma 3 *Let $\alpha > 1$, $f \in L_1([a,b] \cap T)$. Assume that $\widehat{h}_{\alpha-1}(s, \rho(t))$ is Lebesgue ∇-measurable on $([a,b] \cap T)^2$; $a, b \in T$. Then $J_a^{\alpha} f \in L_1([a,b] \cap T)$.*

For $u \leq t$; $u, t \in T$, we define

$$\begin{aligned} \varepsilon(t,u) &= \int_{\rho(u)}^{u} \widehat{h}_{\alpha-1}(t, \rho(\tau)) \widehat{h}_{\beta-1}(\tau, \rho(u)) \nabla\tau \\ &= \nu(u) \widehat{h}_{\alpha-1}(t, \rho(u)) \widehat{h}_{\beta-1}(u, \rho(u)), \end{aligned} \tag{8}$$

where $\alpha, \beta > 1$.

Next we notice for $\alpha, \beta > 1$; $a, b \in T$, $f \in L_1([a,b] \cap T)$, and $\widehat{h}_{\alpha-1}(s, \rho(t))$ is continuous on $([a,b] \cap T)^2$ for any $\alpha > 1$, that

$$J_a^{\alpha} J_a^{\beta} f(t) = \int_a^t \widehat{h}_{\alpha-1}(t, \rho(\tau)) \nabla\tau \int_a^{\tau} \widehat{h}_{\beta-1}(\tau, \rho(u)) f(u) \nabla u.$$

Hence

$$J_a^{\alpha} J_a^{\beta} f(t) + \int_a^t f(u) \varepsilon(t,u) \nabla u = J_a^{\alpha+\beta} f(t), \quad \forall\, t \in [a,b] \cap T.$$

So we have the semigroup property

$$J_a^{\alpha} J_a^{\beta} f(t) + \int_a^t f(u) \nu(u) \widehat{h}_{\alpha-1}(t, \rho(u)) \widehat{h}_{\beta-1}(u, \rho(u)) \nabla u = J_a^{\alpha+\beta} f(t), \tag{9}$$

$\forall\, t \in [a,b] \cap T$, with $a, b \in T$.

We call the Lebesgue ∇-integral

$$D(f, \alpha, \beta, T, t) = \int_a^t f(u) \nu(u) \widehat{h}_{\alpha-1}(t, \rho(u)) \widehat{h}_{\beta-1}(u, \rho(u)) \nabla u, \tag{10}$$

$t \in [a,b] \cap T$; $a, b \in T$, the backward graininess deviation functional of $f \in L_1([a,b] \cap T)$.

If $T = \mathbb{R}$, then $D(f, \alpha, \beta, \mathbb{R}, t) = 0$.

Putting things together we have

Theorem 4 *Let $T_k = T$, $a, b \in T$, $f \in L_1([a,b] \cap T)$; $\alpha, \beta > 1$; $\widehat{h}_{\alpha-1}(s, \rho(t))$ is continuous on $([a,b] \cap T)^2$ for any $\alpha > 1$. Then*

$$J_a^{\alpha} J_a^{\beta} f(t) + D(f, \alpha, \beta, T, t) = J_a^{\alpha+\beta} f(t), \tag{11}$$

$\forall\, t \in [a,b] \cap T$.

We make

Remark 5 *Let $\mu > 2$ such that $m-1 < \mu < m \in \mathbb{N}$, i.e. $m = \lceil \mu \rceil$ (ceiling of the number), $\widetilde{\nu} = m - \mu$ $(0 < \widetilde{\nu} < 1)$.*

Let $f \in C_{ld}^{m}([a,b] \cap T)$. Clearly here ([10]) f^{∇^m} is a Lebesgue ∇-integrable function.

We define the nabla fractional derivative on time scale T of order $\mu - 1$ as follows:

$$\nabla_{a*}^{\mu-1} f(t) = \left(J_a^{\widetilde{\nu}+1} f^{\nabla^m}\right)(t) = \int_a^t \widehat{h}_{\widetilde{\nu}}(t, \rho(\tau)) f^{\nabla^m}(\tau) \nabla\tau, \qquad (12)$$

$\forall\ t \in [a,b] \cap T$.

Notice here that $\nabla_{a*}^{\mu-1} f \in C([a,b] \cap T)$ by a simple argument using dominated convergence theorem in Lebesgue ∇-sense.

If $\mu = m$, then $\widetilde{\nu} = 0$ and by (12) we get

$$\nabla_{a*}^{m-1} f(t) = J_a^1 f^{\nabla^m}(t) = f^{\nabla^{m-1}}(t). \qquad (13)$$

More generally, by [8], given that $f^{\nabla^{m-1}}$ is everywhere finite and absolutely continuous on $[a,b] \cap T$, then f^{∇^m} exists ∇-a.e. and is Lebesgue ∇-integrable on $(a,t] \cap T$, $\forall\ t \in [a,b] \cap T$, and one can plug it into (12).

We have

Theorem 6 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$; $f \in C_{ld}^{m}([a,b] \cap T)$, $a, b \in T$, $T_k = T$. Suppose $\widehat{h}_{\mu-2}(s, \rho(t))$, $\widehat{h}_{\widetilde{\nu}}(s, \rho(t))$ to be continuous on $([a,b] \cap T)^2$.*

Then

$$\int_a^t \widehat{h}_{m-1}(t, \rho(\tau)) f^{\nabla^m}(\tau) \nabla\tau = \qquad (14)$$

$$\int_a^t f^{\nabla^m}(u) \nu(u) \widehat{h}_{\mu-2}(t, \rho(u)) \widehat{h}_{\widetilde{\nu}}(u, \rho(u)) \nabla u + \int_a^t \widehat{h}_{\mu-2}(t, \rho(\tau)) \nabla_{a*}^{\mu-1} f(\tau) \nabla\tau,$$

$\forall\ t \in [a,b] \cap T$.

We need the nabla time scales Taylor formula

6

Theorem 7 *([2]) Let $f \in C_{ld}^{m}(T)$, $m \in \mathbb{N}$, $T_k = T$; $a, b \in T$. Then*

$$f(t) = \sum_{k=0}^{m-1} \widehat{h}_k(t, a) f^{\nabla^k}(a) + \int_a^t \widehat{h}_{m-1}(t, \rho(\tau)) f^{\nabla^m}(\tau) \nabla\tau, \tag{15}$$

$\forall\ t \in [a, b] \cap T$.

Next we present the fractional time scales nabla Taylor formula

Theorem 8 *Let $\mu > 2$, $m - 1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$; $f \in C_{ld}^{m}(T)$, $a, b \in T$, $T_k = T$. Suppose $\widehat{h}_{\mu-2}(s, \rho(t))$, $\widehat{h}_{\widetilde{\nu}}(s, \rho(t))$ to be continuous on $([a, b] \cap T)^2$. Then*

$$f(t) = \sum_{k=0}^{m-1} \widehat{h}_k(t, a) f^{\nabla^k}(a) + \tag{16}$$

$$\int_a^t f^{\nabla^m}(u) \nu(u) \widehat{h}_{\mu-2}(t, \rho(u)) \widehat{h}_{\widetilde{\nu}}(u, \rho(u)) \nabla u + \int_a^t \widehat{h}_{\mu-2}(t, \rho(\tau)) \nabla_{a*}^{\mu-1} f(\tau) \nabla\tau,$$

$\forall\ t \in [a, b] \cap T$.

Corollary 9 *All as in Theorem 8. Additionally suppose $f^{\nabla^k}(a) = 0$, $k = 0, 1, ..., m - 1$. Then*

$$A(t) := f(t) - D\left(f^{\nabla^m}, \mu - 1, \widetilde{\nu} + 1, T, t\right) \tag{17}$$

$$= f(t) - \int_a^t f^{\nabla^m}(u) \nu(u) \widehat{h}_{\mu-2}(t, \rho(u)) \widehat{h}_{\widetilde{\nu}}(u, \rho(u)) \nabla u$$

$$= \int_a^t \widehat{h}_{\mu-2}(t, \rho(\tau)) \nabla_{a*}^{\mu-1} f(\tau) \nabla\tau,$$

$\forall\ t \in [a, b] \cap T$.

Notice here that $D\left(f^{\nabla^m}, \mu - 1, \widetilde{\nu} + 1, T, t\right) \in C_{ld}([a, b] \cap T)$. Also the R.H.S (17) is a continuous function in $t \in [a, b] \cap T$.

2 Fractional Nabla Inequalities on Time Scales

We present a Poincaré type related inequality.

Theorem 10 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$; $f \in C^m_{ld}(T)$, $a, b \in T$, $a \leq b$, $T_k = T$. Suppose $\widehat{h}_{\mu-2}(s, \rho(t))$, $\widehat{h}_{\widetilde{\nu}}(s, \rho(t))$ to be continuous on $([a,b] \cap T)^2$, and $f^{\nabla^k}(a) = 0$, $k = 0, 1, ..., m-1$. Here $A(t) = f(t) - D\left(f^{\nabla^m}, \mu-1, \widetilde{\nu}+1, T, t\right)$, $t \in [a,b] \cap T$; and let $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$.*

Then

$$\int_a^b |A(t)|^q \nabla t \leq \left(\int_a^b \left(\int_a^t \left|\widehat{h}_{\mu-2}(t, \rho(\tau))\right|^p \nabla\tau\right)^{\frac{q}{p}} \nabla t\right) \left(\int_a^b \left|\nabla^{\mu-1}_{a*} f(t)\right|^q \nabla t\right). \tag{18}$$

Next we give a related Sobolev inequality.

Theorem 11 *Here all as in Theorem 10. Let $r \geq 1$ and denote*

$$\|f\|_r = \left(\int_a^b |f(t)|^r \nabla t\right)^{\frac{1}{r}}. \tag{19}$$

Then

$$\|A\|_r \leq \left(\int_a^b \left(\int_a^t \left|\widehat{h}_{\mu-2}(t, \rho(\tau))\right|^p \nabla\tau\right)^{\frac{r}{p}} \nabla t\right)^{\frac{1}{r}} \left\|\nabla^{\mu-1}_{a*} f\right\|_q. \tag{20}$$

Next we give an Opial type related inequality.

Theorem 12 *Here all as in Theorem 10. Additionally assume that*

$$\left|\nabla^{\mu-1}_{a*} f\right| \text{ is increasing on } [a,b] \cap T. \tag{21}$$

Then

$$\int_a^b |A(t)| \left|\nabla^{\mu-1}_{a*} f(t)\right| \nabla t \leq$$

$$(b-a)^{\frac{1}{q}} \left(\int_a^b \left(\int_a^t \left|\widehat{h}_{\mu-2}(t, \rho(\tau))\right|^p \nabla\tau\right) \nabla t\right)^{\frac{1}{p}} \left(\int_a^b \left(\nabla^{\mu-1}_{a*} f(t)\right)^{2q} \nabla t\right)^{\frac{1}{q}}. \tag{22}$$

8

It follows related Ostrowski type inequalities.

Theorem 13 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$; $f \in C_{ld}^{m}(T)$, $a, b \in T$, $a \leq b$, $T_k = T$. Suppose $\widehat{h}_{\mu-2}(s, \rho(t))$, $\widehat{h}_{\widetilde{\nu}}(s, \rho(t))$ to be continuous on $([a,b] \cap T)^2$, and $f^{\nabla^k}(a) = 0$, $k = 1, ..., m-1$. Denote $A(t) = f(t) - D\left(f^{\nabla^m}, \mu - 1, \widetilde{\nu} + 1, T, t\right)$, $t \in [a,b] \cap T$.*

Then

$$\left|\frac{1}{b-a}\int_a^b A(t)\,\nabla t - f(a)\right| \leq$$

$$\frac{1}{b-a}\left(\int_a^b \left(\int_a^t \left|\widehat{h}_{\mu-2}(t, \rho(\tau))\right| \nabla\tau\right) \nabla t\right) \left\|\nabla_{a*}^{\mu-1} f\right\|_{\infty, [a,b] \cap T}. \tag{23}$$

Theorem 14 *All as in Theorem 13. Let $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\left|\frac{1}{b-a}\int_a^b A(t)\,\nabla t - f(a)\right| \leq$$

$$\frac{1}{b-a}\left(\int_a^b \left(\int_a^t \left|\widehat{h}_{\mu-2}(t, \rho(\tau))\right|^p \nabla\tau\right)^{\frac{1}{p}} \nabla t\right) \left\|\nabla_{a*}^{\mu-1} f\right\|_{q, [a,b] \cap T}. \tag{24}$$

We finish general fractional nabla time scales inequalities with a related Hilbert-Pachpatte type inequality.

Theorem 15 *Let $\varepsilon > 0$, $\mu > 2$, $m - 1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$; $f_i \in C_{ld}^{m}(T_i)$, $a_i, b_i \in T_i$, $a_i \leq b_i$, $T_{ik} = T_i$ time scale, $i = 1, 2$. Suppose $\widehat{h}_{\mu-2}^{(i)}(s_i, \rho_i(t_i))$, $\widehat{h}_{\widetilde{\nu}}^{(i)}(s_i, \rho_i(t_i))$ to be continuous on $([a_i, b_i] \cap T_i)^2$, and $f_i^{\nabla^k}(a_i) = 0$, $k = 0, 1, ..., m-1$; $i = 1, 2$. Here $A_i(t_i) = f_i(t_i) - D_i(f_i^{\nabla^m}, \mu - 1, \widetilde{\nu} + 1, T_i, t_i)$, $t_i \in [a_i, b_i] \cap T_i$; $i = 1, 2$, and $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$.*

Call

$$F(t_1) = \int_{a_1}^{t_1} \left(\left|\widehat{h}_{\mu-2}^{(1)}(t_1, \rho_1(\tau_1))\right|\right)^p \nabla\tau_1,$$

for all $t_1 \in [a_1, b_1]$, and

$$G(t_2) = \int_{a_2}^{t_2} \left(\left|\widehat{h}_{\mu-2}^{(2)}(t_2, \rho_2(\tau_2))\right|\right)^q \nabla\tau_2,$$

for all $t_2 \in [a_2, b_2]$ *(where* $\widehat{h}_{\mu-2}^{(i)}$, ρ_i *are the corresponding* $\widehat{h}_{\mu-2}$, ρ *to* T_i, $i = 1, 2$*).*

Then

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \frac{|A_1(t_1)| \, |A_2(t_2)|}{\left(\varepsilon + \frac{F(t_1)}{p} + \frac{G(t_2)}{q}\right)} \nabla t_1 \nabla t_2 \leq$$

$$(b_1 - a_1)(b_2 - a_2) \left(\int_{a_1}^{b_1} \left| \nabla_{a_1 *}^{\mu-1} f_1(t_1) \right|^q \nabla t_1 \right)^{\frac{1}{q}} \left(\int_{a_2}^{b_2} \left| \nabla_{a_2 *}^{\mu-1} f_2(t_2) \right|^p \nabla t_2 \right)^{\frac{1}{p}}. \tag{25}$$

(above double time scales Riemann nabla integration is considered in the natural interative way).

3 Applications

I) Here $T = \mathbb{R}$ case.

Let $\mu > 2$ such that $m - 1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$, $f \in C^m([a, b])$, $a, b \in \mathbb{R}$.

The nabla fractional derivative on $\mathbb{R}$ of order $\mu - 1$ is defined as follows:

$$\nabla_{a*}^{\mu-1} f(t) = \left(J_a^{\widetilde{\nu}+1} f^{(m)}\right)(t) = \frac{1}{\Gamma(\widetilde{\nu}+1)} \int_a^t (t - \tau)^{\widetilde{\nu}} f^{(m)}(\tau) \, d\tau, \tag{26}$$

$\forall \ t \in [a, b]$.

Notice that $\nabla_{a*}^{\mu-1} f \in C([a, b])$, and $A(t) = f(t)$, $\forall \ t \in [a, b]$.

We give a Poincaré type inequality.

Theorem 16 *Let* $\mu > 2$, $m - 1 < \mu < m \in \mathbb{N}$, $f \in C^m(\mathbb{R})$, $a, b \in \mathbb{R}$, $a \leq b$. *Suppose* $f^{(k)}(a) = 0$, $k = 0, 1, ..., m - 1$. *Let* $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$. *Then*

$$\int_a^b |f(t)|^q \, dt \leq \frac{(b - a)^{(\mu-1)q}}{(\Gamma(\mu - 1))^q (\mu - 1) q ((\mu - 2) p + 1)^{q-1}} \left(\int_a^b \left| \nabla_{a*}^{\mu-1} f(t) \right|^q dt \right). \tag{27}$$

Proof. By Theorem 10. ■

We give a Sobolev type inequality.

10

Theorem 17 *All as in Theorem 16. Let $r \geq 1$. Then*

$$\|f\|_r \leq \frac{(b-a)^{\mu-2+\frac{1}{p}+\frac{1}{r}}}{\Gamma(\mu-1)((\mu-2)p+1)^{\frac{1}{p}}\left((\mu-2)r+\frac{r}{p}+1\right)^{\frac{1}{r}}}\left\|\nabla_{a*}^{\mu-1}f\right\|_q. \quad (28)$$

Proof. By Theorem 11. ■

We continue with an Opial type inequality.

Theorem 18 *All as in Theorem 16. Assume $\left|\nabla_{a*}^{\mu-1}f\right|$ is increasing on $[a,b]$.*

$$\int_a^b |f(t)|\left|\nabla_{a*}^{\mu-1}f(t)\right| dt \leq$$

$$\frac{(b-a)^{\mu-\frac{1}{q}}}{\Gamma(\mu-1)\left[((\mu-2)p+1)((\mu-2)p+2)\right]^{\frac{1}{p}}}\left(\int_a^b\left(\nabla_{a*}^{\mu-1}f(t)\right)^{2q}dt\right)^{\frac{1}{q}}. \quad (29)$$

Proof. By Theorem 12. ■

Some Ostrowski type inequalities follow.

Theorem 19 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $f \in C^m(\mathbb{R})$, $a,b \in \mathbb{R}$, $a \leq b$. Suppose $f^{(k)}(a) = 0$, $k = 1,...,m-1$. Then*

$$\left|\frac{1}{b-a}\int_a^b f(t)\,dt - f(a)\right| \leq \frac{(b-a)^{\mu-1}}{\Gamma(\mu+1)}\left\|\nabla_{a*}^{\mu-1}f\right\|_{\infty,[a,b]}. \quad (30)$$

Proof. By Theorem 13. ■

Theorem 20 *Here all as in Theorem 19. Let $p,q > 1 : \frac{1}{p}+\frac{1}{q} = 1$. Then*

$$\left|\frac{1}{b-a}\int_a^b f(t)\,dt - f(a)\right| \leq \frac{(b-a)^{\mu-\frac{1}{q}-1}}{\Gamma(\mu-1)\left(\mu-\frac{1}{q}\right)((\mu-2)p+1)^{\frac{1}{p}}}\left\|\nabla_{a*}^{\mu-1}f\right\|_{q,[a,b]}. \quad (31)$$

Proof. By Theorem 14. ■

We finish this subsection with a Hilbert-Pachpatte inequality on $\mathbb{R}$.

Theorem 21 *Let* $\varepsilon > 0$, $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $i = 1, 2$; $f_i \in C^m(\mathbb{R})$, $a_i, b_i \in \mathbb{R}$, $a_i \leq b_i$, $f_i^{(k)}(a_i) = 0$, $k = 0, 1, ..., m-1$; $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$.

Call

$$F(t_1) = \frac{(t_1 - a_1)^{(\mu-2)p+1}}{(\Gamma(\mu-1))^p((\mu-2)p+1)},$$

$t_1 \in [a_1, b_1]$, *and*

$$G(t_2) = \frac{(t_2 - a_2)^{(\mu-2)q+1}}{(\Gamma(\mu-1))^q((\mu-2)q+1)},$$

$t_2 \in [a_2, b_2]$.

Then

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \frac{|f_1(t_1)| |f_2(t_2)|}{\left(\varepsilon + \frac{F(t_1)}{p} + \frac{G(t_2)}{q}\right)} dt_1 dt_2 \leq$$

$$(b_1 - a_1)(b_2 - a_2) \left(\int_{a_1}^{b_1} \left| \nabla_{a_1 *}^{\mu-1} f_1(t_1) \right|^q dt_1 \right)^{\frac{1}{q}} \left(\int_{a_2}^{b_2} \left| \nabla_{a_2 *}^{\mu-1} f_2(t_2) \right|^p dt_2 \right)^{\frac{1}{p}}. \tag{32}$$

Proof. By Theorem 15. ■

II) Here $T = \mathbb{Z}$ case.

Let $\mu > 2$ such that $m-1 < \mu < m \in \mathbb{N}$, $\widetilde{\nu} = m - \mu$, $a, b \in \mathbb{Z}$, $a \leq b$. Here $f : \mathbb{Z} \to \mathbb{R}$, and $f^{\nabla^m}(t) = \nabla^m f(t) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} f(t-k)$.

The nabla fractional derivative on $\mathbb{Z}$ of order $\mu - 1$ is defined as follows:

$$\nabla_{a*}^{\mu-1} f(t) = \left(J_a^{\widetilde{\nu}+1} (\nabla^m f) \right)(t) = \frac{1}{\Gamma(\widetilde{\nu}+1)} \sum_{\tau=a+1}^{t} (t - \tau + 1)^{\overline{\widetilde{\nu}}} (\nabla^m f)(\tau), \tag{33}$$

$\forall\ t \in [a, \infty) \cap \mathbb{Z}$.

Notice here that $\nu(t) = 1$, $\forall\ t \in \mathbb{Z}$, and

$$\begin{aligned} A(t) &= f(t) - D(\nabla^m f, \mu-1, \widetilde{\nu}+1, \mathbb{Z}, t) \\ &= f(t) - \sum_{u=a+1}^{t} (\nabla^m f(u)) \frac{(t-u+1)^{\overline{\mu-2}}}{\Gamma(\mu-1)}, \end{aligned} \tag{34}$$

$\forall\ t \in [a, \infty) \cap \mathbb{Z}$.

We give a discrete fractional Poincaré type inequality.

12

Theorem 22 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $a, b \in \mathbb{Z}$, $a \leq b$, $f : \mathbb{Z} \to \mathbb{R}$. Assume $\nabla^k f(a) = 0$, $k = 0, 1, ..., m-1$. Let $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\sum_{t=a+1}^{b} |A(t)|^q \leq$$

$$\frac{1}{(\Gamma(\mu-1))^q} \left(\sum_{t=a+1}^{b} \left(\sum_{\tau=a+1}^{t} (t-\tau+1)^{(\overline{\mu-2})p} \right) \right) \left(\sum_{t=a+1}^{b} \left| \nabla_{a*}^{\mu-1} f(t) \right|^q \right). \tag{35}$$

Proof. By Theorem 10. ■

We continue with a discrete fractional Sobolev type inequality.

Theorem 23 *Here all as in Theorem 22. Let $r \geq 1$ and denote*

$$\|f\|_r = \left(\sum_{t=a+1}^{b} |f(t)|^r \right)^{\frac{1}{r}}.$$

Then

$$\|A\|_r \leq \frac{1}{\Gamma(\mu-1)} \left(\sum_{t=a+1}^{b} \left(\sum_{\tau=a+1}^{t} (t-\tau+1)^{(\overline{\mu-2})p} \right)^{\frac{r}{p}} \right)^{\frac{1}{r}} \left\| \nabla_{a*}^{\mu-1} f \right\|_q. \tag{36}$$

Proof. By Theorem 11. ■

Next we give a discrete fractional Opial type inequality.

Theorem 24 *Here all as in Theorem 22. Assume that $\left|\nabla_{a*}^{\mu-1} f\right|$ is increasing on $[a, b] \cap \mathbb{Z}$. Then*

$$\sum_{t=a+1}^{b} |A(t)| \left| \nabla_{a*}^{\mu-1} f(t) \right| \leq$$

$$\frac{(b-a)^{\frac{1}{q}}}{\Gamma(\mu-1)} \left(\sum_{t=a+1}^{b} \left(\sum_{\tau=a+1}^{t} (t-\tau+1)^{(\overline{\mu-2})p} \right) \right)^{\frac{1}{p}} \left(\sum_{t=a+1}^{b} \left(\nabla_{a*}^{\mu-1} f(t) \right)^{2q} \right)^{\frac{1}{q}}. \tag{37}$$

Proof. By Theorem 12. ■

It follows related discrete fractional Ostrowski type inequalities.

Theorem 25 *Let $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$, $a, b \in \mathbb{Z}$, $a \leq b$, $f : \mathbb{Z} \to \mathbb{R}$. Assume $\nabla^k f(a) = 0$, $k = 1, ..., m-1$.*

Then

$$\left| \frac{1}{b-a} \sum_{t=a+1}^{b} A(t) - f(a) \right| \leq$$

$$\frac{1}{(b-a)\Gamma(\mu-1)} \left(\sum_{t=a+1}^{b} \left(\sum_{\tau=a+1}^{t} (t-\tau+1)^{\overline{\mu-2}} \right) \right) \left\| \nabla_{a*}^{\mu-1} f \right\|_{\infty,[a,b]\cap\mathbb{Z}}. \quad (38)$$

Proof. By Theorem 13. ■

Theorem 26 *All as in Theorem 25. Let $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$. Then*

$$\left| \frac{1}{b-a} \sum_{t=a+1}^{b} A(t) - f(a) \right| \leq$$

$$\frac{1}{(b-a)\Gamma(\mu-1)} \left(\sum_{t=a+1}^{b} \left(\sum_{\tau=a+1}^{t} (t-\tau+1)^{\left(\overline{\mu-2}\right)p} \right)^{\frac{1}{p}} \right) \left\| \nabla_{a*}^{\mu-1} f \right\|_{q,[a,b]\cap\mathbb{Z}}. \quad (39)$$

Proof. By Theorem 14. ■

We finish article with a discrete fractional Hilbert-Pachpatte type inequality.

Theorem 27 *Let $\varepsilon > 0$, $\mu > 2$, $m-1 < \mu < m \in \mathbb{N}$; $i = 1, 2$; $f_i : \mathbb{Z} \to \mathbb{R}$, $a_i, b_i \in \mathbb{Z}$, $a_i \leq b_i$. Suppose $\nabla^k f_i(a_i) = 0$, $k = 0, 1, ..., m-1$. Here $A_i(t_i) = f_i(t_i) - \sum_{u_i=a_i+1}^{t_i} (\nabla^m f(u_i)) \frac{(t_i-u_i+1)^{\overline{\mu-2}}}{\Gamma(\mu-1)}$, $\forall\ t_i \in [a_i, \infty) \cap \mathbb{Z}$; $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1$.*

Call

$$F(t_1) = \sum_{\tau_1=a_1+1}^{t_1} \frac{(t_1-\tau_1+1)^{\left(\overline{\mu-2}\right)p}}{(\Gamma(\mu-1))^p},$$

14

$\forall\ t_1 \in [a_1, \infty) \cap \mathbb{Z}$, *and*

$$G(t_2) = \sum_{\tau_2 = a_2+1}^{t_2} \frac{(t_2 - \tau_2 + 1)^{(\overline{\mu - 2})q}}{(\Gamma(\mu - 1))^q},$$

$\forall\ t_2 \in [a_2, \infty) \cap \mathbb{Z}$.

Then

$$\sum_{t_1=a_1+1}^{b_1} \sum_{t_2=a_2+1}^{b_2} \frac{|A_1(t_1)|\,|A_2(t_2)|}{\left(\varepsilon + \frac{F(t_1)}{p} + \frac{G(t_2)}{q}\right)} \leq$$

$$(b_1 - a_1)(b_2 - a_2) \left(\sum_{t_1=a_1+1}^{b_1} \left| \nabla_{a_1*}^{\mu-1} f_1(t_1) \right|^q \right)^{\frac{1}{q}} \left(\sum_{t_2=a_2+1}^{b_2} \left| \nabla_{a_2*}^{\mu-1} f_2(t_2) \right|^p \right)^{\frac{1}{p}}. \tag{40}$$

Proof. By Theorem 15. ■

We intend to publish the complete article with full proofs elsewhere.

References

[1] G. Anastassiou, *Nabla time scales inequalities*, submitted, 2009.

[2] D.R. Anderson, *Taylor Polynomials for nabla dynamic equations on times scales*, Panamer. Math. J., 12(4): 17-27, 2002.

[3] D. Anderson, J. Bullock, L. Erbe, A. Peterson, H. Tran, *Nabla Dynamic equations on time scales*, Panamer. Math. J., 13(2003), no. 1, 1-47.

[4] F. Atici, D. Biles, A. Lebedinsky, *An application of time scales to economics*, Mathematical and Computer Modelling, 43 (2006), 718-726.

[5] F. Atici, P. Eloe, *Discrete fractional Calculus with the nabla operator*, Electronic J. of Qualitative Theory of Differential Equations, Spec. Ed. I, 2009, No. 1, 1-99, http://www.math.u-szeged.hu/ejqtde/

[6] M. Bohner, G.S. Guseinov, *Multiple Lebesgue integration on time scales*, Advances in Difference Equations, Vol. 2006, Article ID 26391, pp. 1-12, DOI 10.1155/ADE/2006/26391.

[7] M. Bohner, G. Guseinov, *Double integral calculus of variations on time scales*, Computers and Mathematics with Applications, 54 (2007), 45-57.

[8] M. Bohner, H. Luo, *Singular second-order multipoint dynamic boundary value problems with mixed derivatives*, Advances in Difference Equations, Vol. 2006, Article ID 54989, p. 1-15, DOI 10.1155/ADE/2006/54989.

[9] M. Bohner, A. Peterson, *Dynamic equations on time scales: An Introduction with Applications*, Birkhaüser, Boston (2001).

[10] G. Guseinov, *Integration on time scales*, J. Math. Anal. Appl., 285 (2003), 107-127.

[11] S. Hilger, *Ein Maßketten kalkül mit Anwendung auf Zentrumsmannigfaltigkeiten*, PhD. thesis, Universität Würzburg, Germany (1988).

[12] Wenjun Liu, Quôc Anh Ngô, Wenbing Chen, *Ostrowski type inequalities on time scales for double integrals*, Acta Appl. Math., 106(2009), 229-239.

[13] N. Martins, D. Torres, *Calculus of variations on time scales with nabla derivatives*, Nonlinear Analysis, 71, no. 12 (2009), 763-773.

[14] M. Rafi Segi Rahmat, M. Salmi Md. Noorani, *Fractional integrals and derivatives on time scales with an application*, Manuscript, 2009.

[15] E.T. Whittaker, G.N. Watson, *A Course in Modern Analysis*, Cambridge University Press, 1927.

NUMERICAL SOLUTIONS OF NONLINEAR SECOND-ORDER TWO-POINT BOUNDARY VALUE PROBLEMS USING HALF-SWEEP SOR WITH NEWTON METHOD

J. SULAIMAN, M.K. HASAN, M. OTHMAN, AND S.A. ABDUL KARIM

ABSTRACT. In this paper, we examine the performance of Half-Sweep Successive Over-Relaxation (HSSOR) iterative method together with Newton scheme namely Newton-HSSOR in solving the nonlinear systems generated from second-order finite difference discretization of the nonlinear second-order two-point boundary value problems. As well known that to linearize nonlinear systems, the Newton scheme has been used to transform the nonlinear system into the form of linear system. Then the basic formulation and implementation of Newton-HSSOR iterative method are also presented. Numerical results for three test examples have demonstrated the performance of Newton-HSSOR method compared to other existing SOR methods.

1. INTRODUCTION

Boundary value problems occur in many branches of applied mathematics and physics such as gas dynamics, quantum mechanics, fluid dynamics, aerodynamics, chemical reactions, atomic structures, atomic calculations etc. Most phenomena in these problems, however, have been modeled by nonlinear differential equations. As a result, they have been studied extensively in recent years in order to find analytic and/or approximate solutions. For instance, there are many methods available for these problems such as numerical analytic, finite difference, finite element, finite volume and boundary element methods. Solving nonlinear boundary value problems is very difficult in that only few of them can be solved explicitly. Therefore, in this paper, the finite difference method will be considered to develop a reliable algorithm in solving nonlinear two-point boundary value problems. By considering the second-order nonlinear finite difference approximation equations, a nonlinear system can be generated and needs to be linearized through the Newton method in order to form the corresponding linear system. Since the characteristics of linear systems are large and sparse, iterative methods are the natural options for efficient solutions.

For that reason, various iterative methods also have been studied to yield fast numerical solution of linear systems (see Young [22, 23, 24]; Hackbusch [4]; Saad [12]). Apart from those iterative methods, the discovery on the concept of the half-sweep iterative method has been inspired by Abdullah [1] via Explicit Decoupled Group (EDG) iterative method in solving two-dimensional Poisson equations. The main characteristic of this concept is that the half-sweep iterative method includes the reduction technique in order to reduce the computational complexity of linear

Key words and phrases. Newton Scheme, SOR Iteration, Second-Order Scheme.
2010 *AMS Math. Subject Classification.* 65F10, 65F50, 65L10, 65L12, 65H10.

systems generated from corresponding approximation equations. Following to this concept, further investigations have been extensively conducted in [2, 3, 7, 16, 17, 25] for demonstrating the capability of the half-sweep iteration. Apart from these one-stage iteration concepts, combinations between half-sweep iteration concept with two-stage iterative methods namely HSIADE [15], HSAM [9, 18] and HSGM [8, 10] have also been constructed and implemented for solving linear systems. They pointed out that their proposed two-stage iterative methods are one of most efficient iterative methods in solving any system of linear equations. Due to the low computational complexity, this concept has been also used to derive a family of multigrid methods [11, 19, 20]. Consequently, Hassan et al. [5, 6] have established a family of FDTD methods using this concept in solving wave propagation problems. Meanwhile Saudi and Sulaiman [13, 14] applied to solve the robotic path planning.

In this paper, however, we deal with the application of family of SOR iterative methods such as Full-Sweep SOR (FSSOR) and Half-Sweep SOR (HSSOR) by applying Newton scheme in solving nonlinear two-point boundary value problems. For simplicity, the combination of both proposed iterative methods with Newton scheme is identified as Newton-FSSOR and Newton-HSSOR methods respectively. To investigate the capability of Newton-FSSOR and HSSOR-Newton methods, let us consider a nonlinear second-order two-point boundary value problem defined as

$$-\frac{\partial^2 U}{\partial x^2} = g(x, U, U'), \quad x \in [a, b] \tag{1.1}$$

subject to the boundary conditions

$$U(a) = \beta_0, \quad U(b) = \beta_1$$

where β_0 and β_1 are constants and $g(x, U, U')$ is a nonlinear continuous function. For the sake of simplicity, we shall restrict our discussion onto uniform node points only as shown in Figures 1 and 2. Let assume the solution domain (1) can be uniformly divided into $m = 2^p, p \geq 2$ subinterval in which its distance, Δx is defined in equation (1.2).

$$\Delta x = \frac{b-a}{m} = h, \quad n = m - 1 \tag{1.2}$$

Based on Figures 1 and 2, both figures show the finite grid networks to facilitate us for the implementation of proposed computational algorithms. For that reason, the implementation of the point iterative algorithms will be applied onto the interior solid node points of type • only until the iterative convergence fixed is achieved.

2. SECOND-ORDER HALF-SWEEP FINITE DIFFERENCE APPROXIMATION

Using the approach of second-order discretization scheme, the nonlinear finite difference approximation equation for problem (1.1) can be easily shown as

$$U_{i-1} - 2U_i + U_{i+1} - h^2 g(x_i, U_i, \frac{U_{i+1} - U_{i-1}}{2h}) = 0 \tag{2.1}$$

for $i = 1, 2, 3, ..., n$. Equation (2.1) denotes as the full-sweep nonlinear approximation equations. Now using the second-order half-sweep scheme, we get the half-sweep nonlinear finite difference approximation equations

$$U_{i-2} - 2U_i + U_{i+2} - 4h^2 g(x_i, U_i, \frac{U_{i+2} - U_{i-2}}{4h}) = 0 \tag{2.2}$$

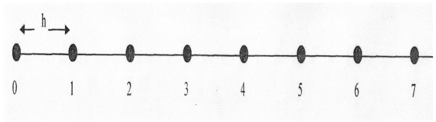

FIGURE 1. shows the distribution of uniform node points for the full-sweep cases respectively at n = 7.

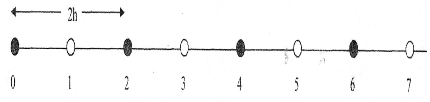

FIGURE 2. shows the distribution of uniform node points for the half-sweep cases respectively at n = 7.

for $i = 2, 4, 6, ..., n-1$. Then let us define the nonlinear function,f in general form based on equations (2.1) and (2.2) as

$$f_i(U_{1p}, U_{2p}, ..., U_{((n+1)/p)-1}) = U_{i-2} - 2U_i + U_{i+2} - 4h^2 g(x_i, U_i, \frac{U_{i+2} - U_{i-2}}{4h}) \tag{2.3}$$

for $i = 1p, 2p, 3p, ..., ((n+1)/p) - 1$, where the value of p in equation (2.3), which equals to 1 and 2, indicates the full- and half-sweep cases respectively. Now we consider all interior node points in the solution domain (1.1) to be imposed over equation (2.3) in order to generate a nonlinear system

$$\left.\begin{array}{l} f_{1p}(U_{1p}^{(k)}, U_{2p}^{(k)}, ..., U_{((n+1)/p)-1}^{(k)}) = 0 \\ f_{1p}(U_{1p}^{(k)}, U_{2p}^{(k)}, ..., U_{((n+1)/p)-1}^{(k)}) = 0 \\ \vdots \\ f_{((n+1)/p)-1}(U_{1p}^{(k)}, U_{2p}^{(k)}, ..., U_{((n+1)/p)-1}^{(k)}) = 0 \end{array}\right\} \tag{2.4}$$

where, $U_i^{(k)}, i = 1p, 2p, 3p, ..., ((n+1)/p) - 1$ indicate as the k^{th} estimation for corresponding exact solutions. To solve the nonlinear system (2.4), the Newton method was used to linearize and transform the original nonlinear system into a linear system which can be rewritten in the matrix form as

$$J\left(\widetilde{U}^{(k)}\right)\widetilde{\Delta h_j} = -f_j\left(\widetilde{U}^{(k)}\right) \tag{2.5}$$

where,

$$J\left(\widetilde{U}^{(k)}\right) = \begin{bmatrix} \frac{\partial f_{1p}}{\partial U_{1p}} & \frac{\partial f_{1p}}{\partial U_{2p}} & \cdots & \frac{\partial f_{1p}}{\partial U_{((n+1)/p)-1}} \\ \frac{\partial f_{2p}}{\partial U_{1p}} & \frac{\partial f_{2p}}{\partial U_{2p}} & \cdots & \frac{\partial f_{2p}}{\partial U_{((n+1)/p)-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{((n+1)/p)-1}}{\partial U_{1p}} & \frac{\partial f_{((n+1)/p)-1}}{\partial U_{2p}} & \cdots & \frac{\partial f_{((n+1)/p)-1}}{\partial U_{((n+1)/p)-1}} \end{bmatrix}$$

$$\widetilde{\Delta h_j} = \left[\ \Delta h_{1p}, \quad \Delta h_{2p}, \quad \Delta h_{3p}, \quad \ldots, \quad \Delta h_{((n+1)/p)-1} \ \right]^T$$

From equations (2.5), the value of the vector, $\widetilde{U}^{(k)}$ needs to be calculated by solving the linear system. Then estimate solutions of $U_i^{(k)}, i = 1p, 2p, 3p, ..., ((n + 1)/p) - 1$ can be determined by using the following expression

$$U_i^{(k+1)} = U_i^{(k)} + \Delta h_i, \quad i = 1p, 2p, 3p, ..., ((n+1)/p) - 1 \tag{2.6}$$

3. FORMULATION OF FAMILY OF SOR METHODS

As above-mentioned in the second section, the coefficient matrix, A of linear systems in equations (2.5) is sparse and large. Let the linear system in equations (2.5) be rewritten as

$$A\widetilde{U} = \widetilde{F} \tag{3.1}$$

Consequently, iterative methods are proposed being as the natural options for efficient solutions of sparse linear system. To solve linear system (3.1), we propose and construct FSSOR and HSSOR iterative methods being used as linear solver.

Actually, Young [22, 23, 24] initiated Successive Over-Relaxation (SOR) method, which is one of the most known and widely used iterative techniques to solve any linear systems. To derive the formulation for FSSOR and HSSOR iterative methods, let the coefficient matrix, A in equation (3.1) be decomposed as

$$A = D + L + V \tag{3.2}$$

where L, D and V are lower triangular, diagonal and upper triangular matrices respectively. Now using the decomposition in equations (3.2) and determining values of matrices D, L and V, therefore, the general scheme for FSSOR and HSSOR methods can be stated as [16, 22, 23, 24]

$$\widetilde{U}^{(k+1)} = (1-\omega)\widetilde{U}^{(k)} + \omega(D+L)^{-1}\left(-V\widetilde{U}^{(k)} + \widetilde{F}\right) \tag{3.3}$$

where ω and $\widetilde{U}^{(k)}$ represent as a relaxation factor and an unknown vector at the k^{th} iteration respectively. The choice of relaxation factor depends upon the properties of the coefficient matrix, A. In addition, a good choice of parameter can improve the convergence rate of iteration process. In practice, the optimal value of ω in range $1 \leq \omega < 2$ will be obtained by implementing several computer programs and then the best approximate value of ω is chosen in which its number of iterations is the smallest.

Due to the advantage of half-sweep iteration concept applied onto the HSSOR method for reducing its computational complexity, we examine the efficiency of FSSOR and HSSOR iterative methods in solving equations (1.1). As taking $\omega = 1$, SOR method reduces to Gauss-Seidel (GS) method. Later GS method will be used as control method. According to equations (3.3), the general algorithm for a family of SOR iterative methods to solve problem (3.1) would be generally described in Algorithm 1.

Algorithm 1: FSSOR and HSSOR schemes

i. Initialize

$$\widetilde{U}^{(0)} \leftarrow 0, \quad \epsilon \leftarrow 10^{-10}$$

ii. For $i = 1p, 2p, 3p, ..., ((n+1)/p) - 1$, calculate $\widetilde{U}^{(k+1}$ from

$$\widetilde{U}^{(k+1)} \leftarrow (D - \omega L)^{-1} \left((\omega V - (1 - \omega L))\widetilde{U}^{(k)} + \omega \widetilde{F} \right)$$

iii. Check the convergence test, $\left| \widetilde{U}^{(k+1)} - \widetilde{U}^{(k)} \right| \leq \epsilon$. If yes, go to step (iv). Otherwise go back to step (ii) iv. Display approximate solutions.

4. NUMERICAL EXPERIMENTS

In order to validate the effectiveness of the FSSOR and HSSOR iterative methods together with Newton approach namely Newton-FSSOR and Newton-HSSOR, three nonlinear example problems were tested. For the sake of comparison, three criteria will be considered for family of SOR methods such as number of iterations, execution time and maximum absolute error. In the implementation of the iterative methods, the convergence test considered the tolerance error, $\epsilon = 10^{-10}$.

Example 1 [21] For comparison purpose, we consider the following nonlinear two-point boundary value problem

$$\frac{\partial^2 U}{\partial x^2} = \frac{3}{4} U^2, \quad x \in [0, 1]. \tag{4.1}$$

Subject to the boundary conditions

$$U(0) = 4, \quad U(1) = 1.$$

Then boundary conditions and the exact solution of the problem (4.1) were defined by

$$U(x) = \frac{4}{(1+x)^2}, \quad x \in [0, 1]. \tag{4.2}$$

Example 2 [21] Let consider the following problem

$$\frac{\partial^2 U}{\partial x^2} = U^3 - UU', \quad x \in [1, 2]. \tag{4.3}$$

with the boundary conditions

$$U(1) = \frac{1}{2}, \quad U(2) = \frac{1}{3}.$$

Then boundary conditions and the exact solution of the problem (4.3)were defined by

$$U(x) = \frac{1}{(x+1)}, \quad x \in [1, 2]. \tag{4.4}$$

Example 3 [21] The third nonlinear boundary value problem, we consider as follows

$$\frac{\partial^2 U}{\partial x^2} = \frac{32 + 2x^3 - UU'}{8}, \quad x \in [1, 3]. \tag{4.5}$$

Subject to the boundary conditions

$$U(1) = 17, \quad U(3) = \frac{14}{3}.$$

Then boundary conditions and the exact solution of the problem (4.5) were defined by

$$U(x) = x^2 + \frac{16}{(x)}, \quad x \in [1, 3]. \tag{4.6}$$

For all above examples, results of numerical experiments obtained from implementation of the FSGS, FSSOR and HSSOR iterative methods, have been recorded in Tables 1, 2, and 3.

Table 1. Comparison of number of iterations, execution time (in seconds) and maximum errors for the iterative methods with the optimal value of ω for example 1.

Number of iterations				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	817596	2853149	9767783	32773526
Newton-FSSOR	5890	11258	21706	41526
	(ω=1.98762)	(ω=1.99363)	(ω=1.99687)	(ω=1.99838)
Newton-HSSOR	2984	5890	11258	21708
	(ω=1.9757)	(ω=1.98761)	(ω=1.99363)	(ω=1.99687)
Execution time (seconds)				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	10.24	71.25	487.01	3266.51
Newton-FSSOR	0.10	0.33	1.23	4.65
Newton-HSSOR	0.03	0.09	0.33	1.27
Maximum absolute errors				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	3.3005e-6	7.1831e-6	2.7356e-5	1.0899e-4
Newton-FSSOR	1.8275e-6	4.6509e-7	1.2609e-7	2.5021e-8
Newton-HSSOR	7.2994e-6	1.8275e-6	4.6509e-7	1.2600e-7

5. CONCLUSION

In this paper we have examined the performance of FSGS, FSSOR and HSSOR iterative methods together with Newton scheme for the numerical solution of nonlinear two point boundary value problem using second order finite difference approximation equations. To demonstrate the efficiency of the three proposed methods, three nonlinear examples are presented. Through the numerical results obtained in Tables 1, 2, and 3, Newton-HSSOR method is compared with the Newton-FSGS and Newton-FSSOR methods. As a result, it shows that the Newton-HSSOR method is superior in terms of the number of iterations and the execution time for different mesh sizes. This is because of the computational complexity of Newton-HSSOR method is approximately 50% compared to the full-sweep case. Overall, the approximate solutions for all three proposed methods are in good agreement. For the

Table 2. Comparison of number of iterations, execution time (in seconds) and maximum errors for the iterative methods with the optimal value of ω for example 2.

Methods	Number of iterations			
	Mesh size			
	512	1024	2048	4096
Newton-FSGS	611486	2041033	66755841	21457588
Newton-FSSOR	4184	8109	15464	30298
	(ω=1.98827)	(ω=1.99414)	(ω=1.99694)	(ω=1.99847)
Newton-HSSOR	2168	4184	8109	15464
	(ω=1.97554)	(ω=1.98827)	(ω=1.99414)	(ω=1.99694)
	Execution time (seconds)			
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	7.50	49.70	323.06	2067.58
Newton-FSSOR	0.07	0.24	0.86	3.31
Newton-HSSOR	0.03	0.07	0.24	0.89
	Maximum absolute errors			
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	2.5066e-6	1.0016e-5	3.9785e-4	1.5742e-4
Newton-FSSOR	4.7539e-9	1.1608e-8	2.6041e-8	4.8332e-8
Newton-HSSOR	2.2045e-8	4.7539e-9	1.1608e-8	2.6041e-8

future works, The method can be extended to solve multi-dimensional nonlinear problems. Apart from single step iterative methods, two-step iterative methods (Ruggiero & Galligan [26]; Evans & Sahimi [27]; Sahimi et al. [28]; Sulaiman et al. [15]) are also interesting to be examined for the solution of nonlinear problems.

References

[1] A.R. Abdullah, The four point Explicit Decoupled Group (EDG) method: A fast Poisson solver. *Intern. J. of Computer Mathematics*, 38, 61-70 (1991).

[2] A.R. Abdullah and N.H.M. Ali, A comparative study of parallel strategies for the solution of elliptic pdes. *Parallel Algorithms and Applications*, 10, 93-103 (1996).

[3] E. Aruchunan and J. Sulaiman, 2011. Half-sweep Conjugate Gradient Method for Solving First Order Linear Fredholm Integro-differential Equations. *Aust. J. of Bas. Appl. Sci.*, 5, 38-43 (2011).

[4] W. Hackbusch, *Iterative solution of large sparse systems of equations*, Springer-Verlag, New York, 1995.

[5] M.K. Hasan, M.Othman, Z.Abbas, J. Sulaiman, and F. Ahmad, Parallel Solution of High Speed Low Order FDTD on 2d Free Space Wave Propagation. *Lecture Notes in Computer Science*, (LNCS 4706), 2007, 13-24.

[6] M.K. Hasan, M.Othman, Z. Abbas, J. Sulaiman and R. Johari, The HSLO(3)-FDTD With Direct-Domain and Temporary-Domain Approaches On Infinite Space Wave Propagation. In *Proceedings of the 7th IEEE Malaysia International Conference on Communications and the 13th IEEE International Conference on Networks*, 2005, 1002- 1007.

Table 3. Comparison of number of iterations, execution time (in seconds) and maximum errors for the iterative methods with the optimal value of ω for example 3.

Number of iterations				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	891866	3115823	10827547	37084653
Newton-FSSOR	5950	1312	21436	42190
	(ω=1.98750)	(ω=1.99381)	(ω=1.99693)	(ω=1.99844)
Newton-HSSOR	3142	5950	11135	21709
	(ω=1.97496)	(ω=1.98750)	(ω=1.99380)	(ω=1.99693)

Execution time (seconds)				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	11.16	77.64	538.26	40883.50
Newton-FSSOR	0.10	0.33	1.22	4.76
Newton-HSSOR	0.03	0.10	0.34	1.27

Maximum absolute errors				
Methods	Mesh size			
	512	1024	2048	4096
Newton-FSGS	1.7188e-6	7.8969e-6	3.2824e-5	1.3346e-4
Newton-FSSOR	3.7522e-6	9.5030e-7	2.5290e-7	6.4334e-8
Newton-HSSOR	1.4981e-5	3.7522e-6	9.4886e-7	2.5299e-7

[7] A. Ibrahim A. and A.R. Abdullah, Solving the two-dimensional diffusion equation by the four point explicit decoupled group (EDG) iterative method. *Intern. J. of Computer Mathematics*, 58, 253-256 (1995).

[8] M.S. Muthuvalu and J. Sulaiman, Numerical solutions of second kind Fredholm integral equations using Half-Sweep Geometric Mean method, *Proceedings of the International Symposium on Information Technology 2008*, Kuala Lumpur, Malaysia: August 26-29, 2008, pp.1927-1934.

[9] M.S. Muthuvalu and J. Sulaiman, Half-Sweep Arithmetic Mean method with composite trapezoidal scheme for solving linear Fredholm integral equations, *Applied Mathematics and Computation*, 217(12), 5442-5448 (2011).

[10] M.S. Muthuvalu and J. Sulaiman, Half-Sweep Geometric Mean Iterative Method For The Repeated Simpson Solution of Second Kind Linear Fredholm Integral Equations. *Proyecciones Journal of Mathematics*, 31(1), 65-79 (2012).

[11] M. Othman, J. Sulaiman and A.R. Abdullah, A parallel halfsweep multigrid algorithm on the shared memory multiprocessors, *Malaysian Jour. of Computer Science*, 13(2), 1-6 (2000).

[12] Y. Saad, *Iterative method for sparse linear systems*, International Thomas Publishing, Boston, 1996.

[13] A. Saudi and J. Sulaiman, Path Planning for Mobile Robot with Half-Sweep Successive Over-Relaxation (HSSOR) Iterative Method, In *Symposium on Progress in Information and Communication Technology (SPICT'09)*, 2009, 57-62.

[14] A. Saudi and J. Sulaiman, Red-Black Strategy for Mobile Robot Path Planning. *Lecture Notes in Engineering and Computer Science*, 2182(1), 2215-2220 (2010).

[15] J. Sulaiman, M.K. Hasan and M. Othman, The Half-Sweep Iterative Alternating Decomposition Explicit (HSIADE) method for diffusion equations, *Lecture Note on Computer Science*, (LNCS 3314), 57-63 (2004).

[16] J. Sulaiman, M.K. Hasan and M. Othman, Red-Black EDGSOR Iterative Method Using Triangle Element Approximation for 2D Poisson Equations. *Lecture Notes in Computer Science*, (LNCS 4707), 298-308 (2007).

[17] J. Sulaiman, M.K. Hasan and M. Othman, Red-Black Half-Sweep Iterative Method Using Triangle Finite Element Approximation for 2D Poisson Equations. *Lecture Notes in Computer Science*, (LNCS 4487), 326-333 (2007).

[18] J. Sulaiman, M. Othman and M.K. Hasan, Half-sweep Arithmetic Mean method using finite element approximation for Poissons equation, *Proceedings of the International Conference on Applied Mathematics*, Bandung, Indonesia, 22-26 August 2005, ISBN: 90-365-2244-7

[19] J. Sulaiman, M. Othman and M.K. Hasan, A New Half-Sweep Algebraic Multigrid (HSAMG) for two-point boundary problems, *Proceedings of the International Conference on Computing and Informatics*, Kuala Lumpur, Malaysia: June 06-08, 2006, pp.1-4.

[20] J. Sulaiman, M. Othman and M.K. Hasan, Half-Sweep Algebraic Multigrid (HSAMG) method applied to diffusion equations. 2008. In. H.G. Bock et al. (Eds). *Modeling, Simulation and Optimization of Complex Processes*, 547-556 (2008).

[21] N.H. Sung, A Nonlinear Shooting Method For Two-Point Boundary Value Problems, *An International Journal Computers And Mathematics With Applications*, 42, 1411-1420 (2001).

[22] D.M. Young, Iterative Methods for solving Partial Difference Equations of Elliptic Type, *Trans. Amer. Math. Soc.*, 76, 92-111 (1954).

[23] D.M. Young, *Iterative solution of large linear systems*, London: Academic Press.1971

[24] D.M. Young, Second-degree iterative methods for the solution of large linear systems, *Journal of Approximation Theory*, 5, 37-148 (1972).

[25] W.S. Yousif and D,J, Evans, Explicit De-coupled Group iterative methods and their implementations. *Parallel Algorithms and Applications*, 7, 53-71 (1995).

[26] V. Ruggiero and E. Galligani, An iterative method for large sparse systems on a vector computer, *Comput. Math. Appl.*, 20(1), 2528 (1990).

[27] D.J. Evans and M.S. Sahimi, The Alternating Group Explicit iterative method (AGE) to solve parabolic and hyperbolic partial differential equations, *Ann. Rev. Num. Fluid Mechanic and Heat Trans*, 2, 283–389 (1988).

[28] M.S. Sahimi, A. Ahmad and A.A. Bakar, The iterative alternating decomposition explicit (IADE) method to solve the heat conduction equation, *Int. J. Computer Maths.*, 47, 219–229 (1993).

(J. Sulaiman) Universiti Malaysia Sabah, Sabah, Malaysia
E-mail address: `jumat@ums.edu.my`

(M.K. Hasan) Universiti Kebangsaan Malaysia, Selangor, Malaysia
E-mail address: `khatim@ftsm.ukm.my`

(M. Othman) Universiti Putra Malaysia, Selangor, Malaysia
E-mail address: `mothman@fsktm.upm.edu.my`

(S.A. Abdul Karim) Universiti Teknologi Petronas, Perak, Malaysia
E-mail address: `samsul_ariffin@petronas.com.my`

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 121-135, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

A RELATED FIXED POINT THEOREM IN n INTUITIONISTIC FUZZY METRIC SPACES

FAYÇAL MERGHADI

ABSTRACT. We prove a related fixed point theorem for n mappings in n Intuitionistic fuzzy metric spaces using an implicit relation which generalizes results of Aliouche and Fisher [2], Merghadi and Aliouche [12] and Rao et al. [15].

1. Introduction and Preliminaries

The theory of fuzzy sets was introduced by Zadeh [20] in 1965. Since then, to use this concept in topology and analysis, many authors have expansively developed the theory of fuzzy sets and applications. for example, Deng [5], Ereeg [6], George and Veeramani [7], Kramosil and Michalek [9] have introduced the concept of fuzzy metric spaces in different ways. One of the most important problems in fuzzy topology is to obtain an appropriate concept of intuitionistic fuzzy metric space. This notion has been introduced and studied by Park [13]. Alaca et al. [1] have redefined the concept of intuitionistic fuzzy metric spaces, according to concept of fuzzy metric spaces and proved Intuitionistic fuzzy Banach and Intuitionistic fuzzy Edelstein contraction theorems, with the different definition of Cauchy sequences and completeness.

Recently, Merghadi and Aliouche [12] Aliouche and Fisher [2], Aliouche et.al [3] and Rao et.al [15] proved some related fixed point theorems in compact metric spaces and sequentially compact fuzzy metric spaces. Motivated by a work due to Popa [14], we have observed that proving fixed point theorems using an implicitly relation is a good idea since it covers several contractive conditions rather than one contractive condition.

Definition 1.1. [17]*A binary operation* $* : [0,1] \times [0,1] \longrightarrow [0,1]$ *is a continuous* $t-$*norm if it satisfies the following conditions:*
1. $*$ *is associative and commutative,*
2. $*$ *is continuous,*
3. $a * 1 = a$ *for all* $a \in [0,1]$
4. $a * b \leq c * d$ *whenever* $a \leq c$ *and* $b \leq d$*, for each* $a, b, c, d \in [0,1]$.

Two typical examples of a continuous $t-$norm are $a*b = ab$ and $a*b = \min\{a,b\}$.

Definition 1.2. [17]*A binary operation* $\Diamond : [0,1] \times [0,1] \longrightarrow [0,1]$ *is a continuous* $t-$*conorm if it satisfies the following conditions:*
1. $\Diamond$ *is associative and commutative,*

Key words and phrases. Fuzzy metric space; implicit relation; Intuitionistic fuzzy metric space; related fixed point.

1

2. $\Diamond$ is continuous,
3. $a\Diamond 0 = a$ for all $a \in [0,1]$,
4. $a\Diamond b \leq c\Diamond d$ whenever $a \leq c$ and $b \leq d$, for each $a,b,c,d \in [0,1]$.

Examples of a continuous $t-$conorm are $a\Diamond b = \max\{a,b\}$ and $a\Diamond b = \min\{1, a+b\}$.

The concept of intuitionistic fuzzy metric space is defined by Park [13].

Definition 1.3. *A $5-$tuple $(X, M, \mathcal{N}, *, \Diamond)$ is called an intuitionistic fuzzy metric space if X is an arbitrary (non-empty) set, $*$ is a continuous t-norm, $\Diamond$ a continuous t-conorm and $M, \mathcal{N}$ are fuzzy sets on $X^2 \times]0,+\infty[$, satisfying the following conditions for each $x,y,z \in X$ and $t,s > 0$,*
1. $M(x,y,t) + \mathcal{N}(x,y,t) \leq 1$;
2. $M(x,y,t) > 0$;
3. $M(x,y,t) = 1$ if and only if $x = y$;
4. $M(x,y,t) = M(y,x,t)$;
5. $M(x,y,t) * M(y,z,s) \leq M(x,z,t+s)$;
6. $M(x,y,\cdot) :]0,+\infty[\to [0,1]$ is continuous;
7. $\mathcal{N}(x,y,t) = 0$ if and only if $x = y$;
8. $\mathcal{N}(x,y,t) = \mathcal{N}(x,y,t)$;
9. $\mathcal{N}(x,y,t)\Diamond\mathcal{N}(y,z,t) \geq \mathcal{N}(x,z,t+s)$;
10. $\mathcal{N}(x,y,t) :]0,+\infty[\to [0,1]$ is continuous.

Then $(M, \mathcal{N})$ is called an intuitionistic fuzzy metric on X. The functions $M(x,y,t)$, $\mathcal{N}(x,y,t)$ denote the degree of nearness and the degree of non-nearness between x and y with respect to t, respectively.

Every fuzzy metric space $(X, M, *)$ is an intuitionistic fuzzy metric space of the form $(X, M, 1-M, *, \Diamond\})$ such that t-norm $*$ and t-conorm $\Diamond$ are associated [11], i.e. $x\Diamond y = 1 - ((1-x)*(1-y))$ for any $x,y \in X$.

Lemma 1.4. [13]*In intuitionistic fuzzy metric space X, $M(x,y,\cdot)$ is non-decreasing and $\mathcal{N}(x,y,\cdot)$ is non-increasing for all $x,y \in X$.*

Example 1.5. *Let (X,d) be a metric space. Denote $a * b = ab$ and $a\Diamond b = \min\{1, a+b\}$ for all $a,b \in [0,1]$ and let M_d and $\mathcal{N}_d$ be fuzzy sets on $X^2 \times]0,+\infty[$ defined as follows:*

$$M_d(x,y,t) = \frac{ht^n}{ht^n + md(x,y)}, \quad \mathcal{N}_d(x,y,t) = \frac{d(x,y)}{kt^n + md(x,y)}$$

for all $h,k,m,n \in \mathbb{R}^+$. Then $(X, M, \mathcal{N}, *, \Diamond)$ is an intuitionistic fuzzy metric space.

Let $(X, M, \mathcal{N}, *, \Diamond)$ be an intuitionistic fuzzy metric space. For $t > 0$, the open ball $B(x,r,t)$ with center $x \in X$ and radius $0 < r < 1$ is defined by

$$B(x,r,t) = \{y \in X : M(x,y,t) > 1-r, \mathcal{N}(x,y,t) < r\}.$$

A subset $A \subset X$ is called open if for each $x \in A$, there exist $t > 0$ and $0 < r < 1$ such that $B(x,r,t) \subset A$. Let $\tau_{(M,\mathcal{N})}$ denote the family of all open subsets of X. Then $\tau_{(M,\mathcal{N})}$ is called the topology on X induced by the intuitionistic fuzzy metric $(M,\mathcal{N})$. This topology is Hausdorff and first countable.

Definition 1.6. [13]*Let $(X, M, \mathcal{N}, *, \Diamond)$ be a fuzzy metric space.*
1) A sequence $\{x_n\}$ in X converges to x if and only if $M(x_n,x,t) \to 1$ and $\mathcal{N}(x_n,x,t) \to 0$ as $n \to \infty$ for each $t > 0$.

2) A sequence $\{x_n\}$ in X is called a Cauchy sequence if for any $0 < \varepsilon < 1$ and $t > 0$, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $M(x_n, x_m, t) > 1 - \varepsilon$ and $\mathcal{N}(x_n, x_m, t) < \varepsilon$ for each $n, m \geq n_0$., $M(x_n, x, t) \to 1$ as $n \to \infty$ for all $t > 0$.
3) The intuitionistic fuzzy metric space $(X, M, \mathcal{N}, *, \Diamond)$ is said to be complete if every Cauchy sequence is convergent.

Definition 1.7. *Let $(X, M, , \mathcal{N}, *, \Diamond)$ be a fuzzy metric space. M is said to be continuous on $X^2 \times (0, \infty)$ if*

$$\lim_{n\to\infty} M(x_n, y_n, t_n) = M(x, y, t)$$

whenever $\{(x_n, y_n, t_n)\}$ is a sequence in $X^2 \times (0, \infty)$ which converges to a point $(x, y, t) \in X^2 \times (0, \infty)$; i.e.,

$$\lim_{n\to\infty} M(x_n, x, t) = \lim_{n\to\infty} M(y_n, y, t) = 1 \text{ and } \lim_{n\to\infty} M(x, y, t_n) = M(x, y, t).$$

Theorem 1.8. [13] *Let $(X, M, , \mathcal{N}, *, \Diamond)$ be an intuitionistic fuzzy metric space such that every Cauchy sequence in X has a convergent subsequence. Then $(X, M, , \mathcal{N}, *, \Diamond)$ is complete.*

Proof. [13] □

- **Implicit Relation**

We denote by Φ, Ψ respectively, sets of all functions $\varphi, \psi : [0,1]^6 \longrightarrow [0,1]$ such that

(i) $\phi \in \Phi$, $\psi \in \Psi$ and ϕ, ψ are upper semi continuous in each coordinate variable,
(ii) ϕ, ψ are decreasing in second and third variable,
(iii) if either $\phi(u, 1, u, v, v, 1) > 0$ or $\phi(u, 1, u, v, 1, v) > 0$ or $\phi(u, u, 1, 1, v, v) > 0$ or for all $u, v \in [0, 1]$, then $u \geq v$. Furthermore, if either

$\psi(u, 0, u, v, v, 0) < 0$ or $\psi(u, 0, u, v, 0, v) < 0$ or $\phi(u, u, 0, 0, v, v) < 0$ for all $u, v \in [0, 1]$, then $u \leq v$.

Example 1.9. *Let $\phi(t_1, t_2, t_3, t_4, t_5, t_6) = t_1 - \min\{t_2, t_3, t_4, t_5, t_6\}$. Then $\phi \in \Phi$.*

Example 1.10. *Let $\psi(t_1, t_2, t_3, t_4, t_5, t_6) = t_1 - \max\{t_2, t_3, t_4, t_5, t_6\}$. Then $\psi \in \Psi$.*

Example 1.11.

$$\begin{aligned} \phi(\{t_2, t_3, t_4, t_5, t_6\}) &= t_1 - \eta(\min\{t_2, t_3, t_4, t_5, t_6\}) \\ \psi(\{t_2, t_3, t_4, t_5, t_6\}) &= t_1 - \varphi(\max\{t_2, t_3, t_4, t_5, t_6\}) \end{aligned}$$

where $\eta, \varphi : [0, 1] \to [0, 1]$ is a increasing and continuous function respectively, with $\eta(t) \geq t$ and $\varphi(t) \leq t$ for $0 \leq t \leq 1$. For example $\eta(t) = \sqrt{t}$ or $\eta(t) = t^h$ for $0 < h < 1$ and $\phi(t) = \dfrac{t}{2}$.

We need the following lemma of [10].

Lemma 1.12. *Let $\{x_n\}$ be a sequence in intuitionistic fuzzy metric space $(X, M, \mathcal{N}, *, \Diamond)$ with $M(x, y, t) \to 1$ and $\mathcal{N}(x, y, t) \to 0$ as $t \to \infty$ for all $x, y \in X$. If there exists a number $k \in]0, 1[$ such that*

$$\begin{aligned} M(x_{n+1}, x_n, kt) &\geq M(x_n, x_{n-1}, t), \\ \mathcal{N}(x_{n+1}, x_n, kt) &\leq \mathcal{N}(x_n, x_{n-1}, t). \end{aligned}$$

Then $\{x_n\}$ is a Cauchy sequence in X.

Lemma 1.13. [10]. *Let $(X, M, , \mathcal{N}, *, \Diamond)$ be an intuitionistic fuzzy metric space. If there exists $k \in (0,1)$ such that $M(x,y,kt) \geq M(x,y,t)$ and $N(x,y;kt) \leq N(x,y;t)$ for $x, y \in X$, then $x = y$.*

Proof. See [10] □

2. Main Results

Theorem 2.1. *Let $(X_i, M_i, \mathcal{N}_i, \theta_i, \gamma_i)_{1 \leq i \leq n}$, be n complete intuitionistic fuzzy metric spaces with $M_i(x, x_i, t) \longrightarrow 1$ and $\mathcal{N}_i(x, x_i, t) \to 0$ as $t \longrightarrow \infty$ for all $x, x_i \in X_i$ and let $\{A_i\}_{i=1}^{i=n}$ be n-mappings such that $A_i : X_i \to X_{i+1}$ for all $i = 1, .., n-1$ and $A_n : X_n \to X_1$, satisfying the inequalities,*

(2.1_M)

$$\phi_1 \begin{pmatrix} M_1\left(A_nA_{n-1}..A_2x_2, A_nA_{n-1}..A_2A_1x_1, kt\right), M_1\left(x_1, A_nA_{n-1}..A_2x_2, t\right), \\ M_1\left(x_1, A_nA_{n-1}..A_2A_1x_1, t\right), M_2\left(x_2, A_1x_1, t\right), \\ M_2\left(x_2, A_1A_nA_{n-1}..A_2x_2, t\right), M_2\left(A_1x_1, A_1A_nA_{n-1}..A_2x_2, t\right) \end{pmatrix} > 0$$

(2.1_N)

$$\psi_1 \begin{pmatrix} \mathcal{N}_1\left(A_nA_{n-1}..A_2x_2, A_nA_{n-1}..A_2A_1x_1, kt\right), \mathcal{N}_1\left(x_1, A_nA_{n-1}..A_2x_2, t\right), \\ \mathcal{N}_1\left(x_1, A_nA_{n-1}..A_2A_1x_1, t\right), \mathcal{N}_2\left(x_2, A_1x_1, t\right), \\ \mathcal{N}_2\left(x_2, A_1A_nA_{n-1}..A_2x_2, t\right), \mathcal{N}_2\left(A_1x_1, A_1A_nA_{n-1}..A_2x_2, t\right) \end{pmatrix} < 0$$

for all $x_1 \in X_1$, $x_2 \in X_2$ and $t > 0$, in general, we have

$(2.i_M)$

$$\phi_i \begin{pmatrix} M_i\left(A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ix_i, kt\right), \\ M_i\left(x_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right), \\ M_i\left(x_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ix_i, t\right), M_{i+1}\left(x_{i+1}, A_ix_i, t\right), \\ M_{i+1}\left(x_{i+1}, A_iA_{i-1}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right), \\ M_{i+1}\left(A_ix_i, A_iA_{i-1}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right) \end{pmatrix} > 0$$

$(2.i_N)$

$$\psi_i \begin{pmatrix} \mathcal{N}_i\left(A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ix_i, kt\right), \\ \mathcal{N}_i\left(x_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right), \\ \mathcal{N}_i\left(x_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ix_i, t\right), \mathcal{N}_{i+1}\left(x_{i+1}, A_ix_i, t\right), \\ \mathcal{N}_{i+1}\left(x_{i+1}, A_iA_{i-1}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right), \\ \mathcal{N}_{i+1}\left(A_ix_i, A_iA_{i-1}...A_1A_nA_{n-1}..A_{i+1}x_{i+1}, t\right) \end{pmatrix} < 0$$

for all $x_i \in X_i$, $x_{i+1} \in X_{i+1}$, $t > 0$, where, $\phi_i \in \Phi$, $\psi_i \in \Psi$, $i = 2, ..., n-1$ and $0 < k < 1$. Again we have,

$$(2.n_M) \qquad \phi_n \begin{pmatrix} M_n\left(A_{n-1}A_{n-2}...A_1x_1, A_{n-1}A_{n-2}...A_1A_nx_n, kt\right), \\ M_n\left(x_n, A_{n-1}A_{n-2}...A_1x_1, t\right), \\ M_n\left(x_n, A_{n-1}A_{n-2}...A_1A_nx_n, t\right), M_1\left(x_1, A_nx_n, t\right), \\ M_1\left(x_1, A_nA_{n-1}A_{n-2}...A_1x_1, t\right), \\ M_1\left(A_nx_n, A_nA_{n-1}A_{n-2}...A_1x_1, t\right) \end{pmatrix} > 0$$

$$(2.n_N) \qquad \psi_n \begin{pmatrix} \mathcal{N}_n\left(A_{n-1}A_{n-2}...A_1x_1, A_{n-1}A_{n-2}...A_1A_nx_n, kt\right), \\ \mathcal{N}_n\left(x_n, A_{n-1}A_{n-2}...A_1x_1, t\right), \\ \mathcal{N}_n\left(x_n, A_{n-1}A_{n-2}...A_1A_nx_n, t\right), \mathcal{N}_1\left(x_1, A_nx_n, t\right), \\ \mathcal{N}_1\left(x_1, A_nA_{n-1}A_{n-2}...A_1x_1, t\right), \\ \mathcal{N}_1\left(A_nx_n, A_nA_{n-1}A_{n-2}...A_1x_1, t\right) \end{pmatrix} < 0$$

for all $x_1 \in X_1$*,* $x_n \in X_n$ *and* $t > 0$*, where* $\phi_n \in \Phi$*,* $\psi_n \in \Psi$ *and* $0 < k < 1$*. Further, suppose that* $\{A_2\}$ *is continuous on* X_i *Then,*

$$A_{i-1}A_{i-2}..A_1A_nA_{n-1}..A_i$$

has a unique fixed point $p_i \in X_i$ *for* $i = 1, ..., n$*. Further,* $A_ip_i = p_{i+1}$ *for* $i = 1, ..., n-1$ *and* $A_np_n = p_1$*.*

Proof. Let $\left\{x_r^{(1)}\right\}_{r\in\mathbb{N}}, \left\{x_r^{(2)}\right\}_{r\in\mathbb{N}}, ..., \left\{x_r^{(i)}\right\}_{r\in\mathbb{N}}, ..., \left\{x_r^{(n)}\right\}_{r\in\mathbb{N}}$ be sequences in X_1, $X_2, ..., X_i, ...X_n$ respectively. Now let $x_0^{(1)}$ be an arbitrary point in X_1, we define the sequences $\left\{x_r^{(i)}\right\}_{r\in\mathbb{N}}$ for $i = 1, ..., n$ by

$$x_r^{(1)} = (A_nA_{n-1}...A_1)^r\, x_0^{(1)},$$
$$x_r^{(i)} = A_{i-1}A_{i-2}..A_1x_r^{(1)} \text{ for } i = 2, ..., n.$$

For $n = 1, 2, ..., n$ we assume that $x_r^{(1)} \neq x_{r+1}^{(1)}$. Applying inequalities (2.1_M) and $(2.1_{\mathcal{N}})$ for $x_2 = A_1x_{r-1}^{(1)} = x_{r-1}^{(2)}$, $x_1 = x_r^{(1)}$ we get

$$\phi_1\begin{pmatrix} M_1\left(x_r^{(1)}, x_{r+1}^{(1)}, kt\right), 1, M_1\left(x_r^{(1)}, x_{r+1}^{(1)}, t\right), \\ M_2\left(A_1x_{r-1}^{(1)}, A_1x_r^{(1)}, t\right), M_2\left(A_1x_{r-1}^{(1)}, A_1x_r^{(1)}, t\right), 1 \end{pmatrix} > 0$$

$$\psi_1\begin{pmatrix} \mathcal{N}_1\left(x_r^{(1)}, x_{r+1}^{(1)}, kt\right), 0, \mathcal{N}_1\left(x_r^{(1)}, x_{r+1}^{(1)}, t\right), \\ \mathcal{N}_2\left(A_1x_{r-1}^{(1)}, A_1x_r^{(1)}, t\right), \mathcal{N}_2\left(A_1x_{r-1}^{(1)}, A_1x_r^{(1)}, t\right), 0 \end{pmatrix} < 0$$

From the implicit relation we have,

$$(3.1_M) \qquad M_1\left(x_r^{(1)}, x_{r+1}^{(1)}, kt\right) \geq M_2\left(x_{r-1}^{(2)}, x_r^{(2)}, t\right),$$

$$(3.1_N) \qquad \mathcal{N}_1\left(x_r^{(1)}, x_{r+1}^{(1)}, kt\right) \leq \mathcal{N}_2\left(x_{r-1}^{(2)}, x_r^{(2)}, t\right),$$

Applying inequalities $(2.i_M)$ and $(2.i_{\mathcal{N}})$ for $x_{i+1} = A_i...A_1\left(A_n...A_1\right)^{r-1}x_0^{(1)} = A_i..A_1x_{r-1}^{(1)}$ and $x_i = A_{i-1}..A_1\left(A_n...A_1\right)^r x_0^{(1)} = A_{i-1}..A_1x_r^{(1)}$, we obtain

$$\phi_i\begin{pmatrix} M_i\left(A_{i-1}..A_1x_r^{(1)}, A_{i-1}..A_1x_{r+1}^{(1)}, kt\right), \\ 1, M_i\left(A_{i-1}..A_1x_r^{(1)}, A_{i-1}..A_1x_{r+1}^{(1)}, t\right) \\ , M_{i+1}\left(A_i..A_1x_{r-1}^{(1)}, A_i..A_1x_r^{(1)}, t\right), \\ M_{i+1}\left(A_i..A_1x_{r-1}^{(1)}, A_iA_{i-1}..A_1x_r^{(1)}, t\right), 1 \end{pmatrix} > 0$$

$$\psi_i\begin{pmatrix} \mathcal{N}_i\left(x_r^{(i)}, x_{r+1}^{(i)}, kt\right), 0, \mathcal{N}_i\left(x_r^{(i)}, x_{r+1}^{(i)}, t\right) \\ , \mathcal{N}_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right), \mathcal{N}_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right), 0 \end{pmatrix} < 0$$

From the implicit relation we obtain,

$$(3.\mathrm{i}_M) \qquad M_i\left(x_r^{(i)}, x_{r+1}^{(i)}, kt\right) \geq M_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right)$$

$$(3.\mathrm{i}_N) \qquad \mathcal{N}_i\left(x_r^{(i)}, x_{r+1}^{(i)}, kt\right) \leq \mathcal{N}_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right)$$

for $i = 2, ..., n-1$ and $r \in \mathbb{N}$.

Now applying inequalities $(2.n_M)$ and $(2.n_{\mathcal{N}})$ for $x_n = x_r^{(n)}$ and $x_1 = x_{r-1}^{(1)}$ we have

$$\phi_n \left(\begin{array}{c} M_n \left(x_r^{(n)}, x_{r+1}^{(n)}, kt \right), 1, M_n \left(x_r^{(n)}, x_{r+1}^{(n)}, t \right), \\ M_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right), M_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right), 1 \end{array} \right) > 0$$

$$\psi_n \left(\begin{array}{c} \mathcal{N}_n \left(x_r^{(n)}, x_{r+1}^{(n)}, kt \right), 0, \mathcal{N}_n \left(x_r^{(n)}, x_{r+1}^{(n)}, t \right), \\ \mathcal{N}_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right), \mathcal{N}_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right), 0 \end{array} \right) < 0$$

and so by implicit relation we have

$$(3.\mathrm{n}_M) \qquad M_n \left(x_r^{(n)}, x_{r+1}^{(n)}, kt \right) \geq M_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right)$$

$$(3.\mathrm{n}_N) \qquad \mathcal{N}_n \left(x_r^{(n)}, x_{r+1}^{(n)}, kt \right) \leq \mathcal{N}_1 \left(x_{r-1}^{(1)}, x_r^{(1)}, t \right)$$

It now follows From $(3.1_M), (3.i_M)$ and $(3, n_M)$ that for large enough n

$$\begin{aligned} M_1 \left(x_r^{(1)}, x_{r+1}^{(1)}, kt \right) &\geq M_2 \left(x_{r-1}^{(2)}, x_r^{(2)}, t \right) \\ M_i \left(x_r^{(i)}, x_{r+1}^{(i)}, t \right) &\geq M_{i+1} \left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, \frac{t}{k} \right) \\ &\geq ... \\ &\geq M_n \left(x_{r+i-n}^{(n)}, x_{r+i-n+1}^{(n)}, \frac{t}{k^{n-i}} \right) \\ &\geq M_1 \left(x_{r+i-n-1}^{(1)}, x_{r+i-n}^{(1)}, \frac{t}{k^{n-i+1}} \right) \\ &\geq ..... \\ &\quad M_1 \left(x_{r+i-2n-1}^{(1)}, x_{r+i-2n}^{(1)}, \frac{t}{k^{2n-i+1}} \right) \\ &\geq .... \\ &\geq M_1 \left(x_{r+i-mn-1}^{(1)}, x_{r+i-mn}^{(1)}, \frac{t}{k^{mn-i+1}} \right) \\ &\geq \min \left\{ M_1 \left(x_1^{(1)}, x_2^{(1)}, \frac{t}{k^{mn}} \right), ..., M_n \left(x_1^{(n)}, x_2^{(n)}, \frac{t}{k^{mn}} \right) \right\} \end{aligned}$$

And it follows From $(3.1_{\mathcal{N}}), (3.i_{\mathcal{N}})$ and $(3, n_{\mathcal{N}})$ that for large enough n

$$\begin{aligned} \mathcal{N}_i \left(x_r^{(i)}, x_{r+1}^{(i)}, t \right) &\leq \mathcal{N}_{i+1} \left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, \frac{t}{k} \right) \\ &\leq ... \leq \mathcal{N}_1 \left(x_{r+i-2n-1}^{(1)}, x_{r+i-2n}^{(1)}, \frac{t}{k^{2n-i+1}} \right) \\ &\leq .... \leq \mathcal{N}_1 \left(x_{r+i-mn-1}^{(1)}, x_{r+i-mn}^{(1)}, \frac{t}{k^{mn-i+1}} \right) \\ &\leq \max \left\{ \mathcal{N}_1 \left(x_1^{(1)}, x_2^{(1)}, \frac{t}{k^{mn}} \right), ..., \mathcal{N}_n \left(x_1^{(n)}, x_2^{(n)}, \frac{t}{k^{mn}} \right) \right\} \end{aligned}$$

Since $0 < k < 1$. It follows from Lemma 1.11 that $\left\{ x_r^{(i)} \right\}$ is a Cauchy sequences in X_i with a limit p_i in X_i for $i = 1, 2, .., n$.

To prove that p_i is a fixed point of $A_{i-1}...A_1A_n...A_ip_i$, we can use proof by induction. Recently F. Merghadi and A. Aliouche sent an article about the existence of three fixed point in intuitionistic fuzzy metric spaces, which have been accepted for publication in the Journal of Mathematical and Computational Science.We have demonstrated for three spaces that $A_{i-1}...A_1A_n...A_ip_i = p_i$ and $A_ip_i = p_{i+1}$for $i = 1, 2, 3$.

By induction, Supposing that $A_{i-1}...A_1A_n...A_ip_i = p_i$ and $A_ip_i = p_{i+1}$ it true for all $A_i \in X_i$ and $i = 1, 2, ..., n-1$. Now, we will going proved that $A_{i-1}...A_1A_n...A_ip_i$ $= p_i$ for all $A_i \in X_i$ and $i = 1, 2, ..., n$

Using the inequality $(2.i_M)$ and $(2.i_{\mathcal{N}})$ for $x_i = p_i$ and $x_{i+1} = x_{r-1}^{(i+1)} = A_i...A_1\left(A_n...A_1\right)^{r-1}x_0^{(1)}$ we obtain

$$\phi_i\left(\begin{array}{c} M_i\left(x_r^{(i)}, A_{i-1}..A_1A_n...A_ip_i, kt\right), M_i\left(p_i, x_r^{(i)}, t\right), \\ M_i\left(p_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ip_i, t\right), \\ M_{i+1}\left(x_{r-1}^{(i+1)}, A_ip_i, t\right), M_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right), \\ M_{i+1}\left(A_ip_i, x_r^{(i+1)}, t\right) \end{array}\right) > 0$$

$$\psi_i\left(\begin{array}{c} \mathcal{N}_i\left(x_r^{(i)}, A_{i-1}..A_1A_n...A_ip_i, kt\right), \mathcal{N}_i\left(p_i, x_r^{(i)}, t\right), \\ \mathcal{N}_i\left(p_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ip_i, t\right), \\ \mathcal{N}_{i+1}\left(x_{r-1}^{(i+1)}, A_ip_i, t\right), \mathcal{N}_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right), \\ \mathcal{N}_{i+1}\left(A_ip_i, x_r^{(i+1)}, t\right) \end{array}\right) < 0$$

Letting $r \to \infty$ we have

$$\phi_i\left(\begin{array}{c} M_i\left(p_i, A_{i-1}..A_1A_n...A_ip_i, kt\right), 1, \\ M_i\left(p_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ip_i, t\right), \\ M_{i+1}\left(p_{i+1}, A_ip_i, t\right), 1, M_{i+1}\left(A_ip_i, p_{i+1}, t\right) \end{array}\right) > 0$$

$$\psi_i\left(\begin{array}{c} \mathcal{N}_i\left(p_i, A_{i-1}..A_1A_nA_{n-1}...A_ip_i, kt\right), 0, \\ \mathcal{N}_i\left(p_i, A_{i-1}A_{i-2}...A_1A_nA_{n-1}..A_ip_i, t\right), \\ \mathcal{N}_{i+1}\left(p_{i+1}, A_ip_i, t\right), 0, \mathcal{N}_{i+1}\left(A_ip_i, p_{i+1}, t\right) \end{array}\right) < 0$$

It follows from (iii) that

$$M_i\left(p_i, A_{i-1}..A_1A_n...A_ip_i, kt\right) \geq M_{i+1}\left(p_{i+1}, A_ip_i, t\right) \tag{4.i$_M$}$$

$$\mathcal{N}_i\left(p_i, A_{i-1}..A_1A_n...A_ip_i, kt\right) \leq \mathcal{N}_{i+1}\left(p_{i+1}, A_ip_i, t\right). \tag{4.i$_N$}$$

for $i = 2, ..., n-1$. By the same way we put $x_1 = p_1$, $x_2 = x_{r-1}^{(2)}$ in (2.1_M) $((2.1_{\mathcal{N}}))$ and $x_n = p_n$, $x_1 = x_{r-1}^{(1)}$ in $(2.n_M)$ $((2.n_{\mathcal{N}}))$ respectively, we get

$$\begin{array}{rcll} M_1\left(p_1, A_n...A_1p_1, kt\right) & \geq & M_2\left(p_2, A_1p_1, t\right) & (4.1_M) \\ M_n\left(p_n, A_{n-1}..A_1A_np_n, kt\right) & \geq & M_1\left(p_1, A_np_n, t\right) & (4.\mathrm{n}_M) \end{array}$$

$$\begin{array}{rcll} \mathcal{N}_1\left(p_1, A_n...A_1p_1, kt\right) & \leq & \mathcal{N}_2\left(p_2, A_1p_1, t\right) & (4.1_N) \\ \mathcal{N}_n\left(p_n, A_{n-1}..A_1A_np_n, kt\right) & \leq & \mathcal{N}_1\left(p_1, A_np_n, t\right) & (4.\mathrm{n}_N) \end{array}$$

We can write again from $(4.1_M), (4.\mathrm{i}_M)$ and $(4.n_M)$,

$$\begin{array}{rcl}
M_1(p_1, A_n..A_1p_1, kt) & \geq & M_2(p_2, A_1p_1, t) \\
M_2(p_2, A_1A_n..A_2p_2, kt) & \geq & M_3(p_3, A_2p_2, t) \\
M_3(p_3, A_2A_1A_n..A_3p_3, kt) & \geq & M_4(p_4, A_3p_3, t) \\
M_4(p_4, A_3A_2A_1A_n..A_4p_4, kt) & \geq & M_5(p_5, A_4p_4, t) \\
& \vdots & \\
M_i(p_i, A_{i-1}..A_1A_n..A_ip_i, kt) & \geq & M_{i+1}(p_{i+1}, A_ip_i, t) \\
& \vdots & \\
M_n(p_n, A_{n-1}..A_1A_np_n, kt) & \geq & M_1(p_1, A_np_n, t)
\end{array}$$

Thus, by induction and (4.1_M), we have

(5.i) $$A_{\hat{}\$*\grave{u}i}p_i = p_{i+1} \text{ for } i = 1, 2, \cdots, n-1$$

(5.n) $$A_np_n = A_nA_{n-1}p_{n-1} = \cdots = A_n..A_2p_2 = A_n..A_1p_1 = p_1$$

and so, from $(4.n_M)$ we have

$$A_{n-1}..A_1A_np_n = p_n,$$

Concerning inequalities $(4.1_{\mathcal{N}}), (4.i_{\mathcal{N}})$ and $(4.n_{\mathcal{N}})$. Similarly, we can prove that

$$A_{i-1}..A_1A_nA_{n-1}..A_ip_i = p_i \text{ for all } i = 1, 2, \cdots, n.$$

For proving the uniqueness of the fixed point p_i in X_i we assume that there exists $z_i \in X_i$ such that $z_i \neq p_i$ and $A_{i-1}..A_1A_nA_{n-1}..A_iz_i = z_i$ for all $i = 1, 2, \cdots, n$.

Firstly, using $(2.i_M)$ *for* $x_i = p_i$ and $x_{i+1} = A_iz_i$ *we have,*

$$\phi_i\left(\begin{array}{c}
M_i\left(\begin{array}{c} A_{i-1}..A_1A_nA_{n-1}..A_{i+1}A_iz_i, \\ A_{i-1}..A_1A_nA_{n-1}..A_ip_i, kt \end{array}\right), \\
M_i(p_i, A_{i-1}..A_1A_nA_{n-1}..A_{i+1}A_iz_i, t), \\
M_i(p_i, A_{i-1}..A_1A_nA_{n-1}..A_ip_i, t), M_{i+1}(A_iz_i, A_ip_i, t), \\
M_{i+1}(A_iz_i, A_i..A_1A_nA_{n-1}..A_{i+1}A_iz_i, t), \\
M_{i+1}(A_ip_i, A_i..A_1A_nA_{n-1}..A_{i+1}A_iz_i, t)
\end{array}\right) > 0$$

and so,

$$\phi_i\left(\begin{array}{c}
M_i(z_i, p_i, kt), M_i(p_i, z_i, t), 1, \\
M_{i+1}(A_iz_i, A_ip_i, t), 1, M_{i+1}(A_ip_i, A_iz_i, t)
\end{array}\right) > 0$$

which implies that,

$$M_i(p_i, z_i, kt) \geq M_{i+1}(A_ip_i, A_iz_i, t) \text{ for all } i = 1, 2, \cdots, n$$

From (5.i), (5.n), we have:

$$\begin{array}{rcl}
M_1(p_1, z_1, kt) & \geq & M_2(A_1p_1, A_1z_1, t) = M_2(p_2, z_2, t) \\
& \vdots & \\
M_i(p_i, z_i, kt) & \geq & M_{i+1}(p_{i+1}, z_{i+1}, t) \text{ for all } i = 2, \cdots, n-1 \\
& \vdots & \\
M_n(p_n, z_n, kt) & \geq & M_1(A_np_n, A_nz_n, t) = M_1(p_1, z_1, t).
\end{array}$$

Similarly, we find by $(2.i_{\mathcal{N}})$ that

$$\begin{array}{rcl}\mathcal{N}_1(p_1,z_1\ ,kt) & \leq & \mathcal{N}_1(p_1,z_1\ ,t)\\ & \vdots & \\ \mathcal{N}_i(p_i,z_i\ ,kt) & \leq & \mathcal{N}_{i+1}(p_{i+1},z_{i+1}\ ,t) \text{ for all } i=2,\cdots,n-1\\ & \vdots & \\ \mathcal{N}_n(p_n,z_n\ ,kt) & \leq & \mathcal{N}_n(p_n,z_n\ ,t)\end{array}$$

This proving the uniqueness of p_i in X_i for $i=1,2,\cdots,n$. This complete the proof of the theorem. □

There is an example satisfies all conditions of Theorem 2.1

Example 2.2. *Let $(X_i,M_i,\mathcal{N}_i,\theta_i,\gamma_i)_{1\leq i\leq n}$, be n complete intuitionistic fuzzy metric spaces such that*

$$M_i(x_i,y_i,t)=\frac{t}{t+|x_i-y_i|},\quad \mathcal{N}_i(x_i,y_i,t)=\frac{|x_i-y_i|}{t+|x_i-y_i|}$$

for all $i=\overline{1..n}$ and $X_1=[0,1]$, $X_i=]i-1,i[$ for all $i\geq 2$. Define $A_i:X_i\to X_{i+1}$ for $i=\overline{1..n-1}$ and $A_n:X_n\to X_1$ by

$$\begin{array}{rcl}A_1x_1 & = & \left\{\frac{3}{2}\ \textit{if}\ x_1\in[0,1]\right.,\ A_nx_n=\left\{\begin{array}{l}\frac{3}{4}\ \textit{if}\ x_n\in]n-1,n-\frac{3}{4}[\\ 1\ \textit{if}\ x_n\in[n-\frac{3}{4},n[\end{array}\right.\\ A_ix_i & = & \left\{\begin{array}{l}i+\frac{1}{4}\ \textit{if}\ x_i\in\left]i-1,i-\frac{3}{4}\right[\\ i+\frac{1}{2}\ \textit{if}\ x_i\in\left[i-\frac{3}{4},i\right[\end{array}\right.\quad \textit{for all}\ i=\overline{2..n-1},\end{array}$$

Let $\phi_1=\phi_2=..=\phi_n=\phi$ such that $\phi(t_1,t_2,t_3,t_4,t_5,t_6)=t_1-\min\{t_2,t_3,t_4,t_5,t_6\}$ and $\psi_1=\psi_2=..=\psi_n=\psi$ such that $\psi(t_1,t_2,t_3,t_4,t_5,t_6)=t_1-\max\{t_2,t_3,t_4,t_5,t_6\}$. Note that there exists w_i in X_i such that $(A_{i-1}A_{i-2}..A_1A_n...A_i)w_i=w_i$, $\forall i=\overline{1..n}$. (a) If $i=n$ we get $(A_{n-1}A_{n-2}..A_1A_n)w_n=w_n$ if $w_n=n-\frac{1}{2}$ because

$$\begin{array}{rl} & (A_{n-1}A_{n-2}..A_1A_n)\left(n-\frac{1}{2}\right)=A_{n-1}A_{n-2}..A_1(1)\\ = & ....=A_{n-1}A_{n-2}..A_{i+1}\left(i+\frac{1}{2}\right)=....\\ = & A_{n-1}A_{n-2}\left(n-\frac{5}{2}\right)=A_{n-1}\left(n-\frac{3}{2}\right)=n-\frac{1}{2}\end{array}$$

b) Remark that for all $i=\overline{1,n-1}$ and $x_i\in\left[i-\frac{3}{4},i\right[$; $A_ix_i\in\left[(i+1)-\frac{3}{4},i+1\right[$, then there exists $w_i=i-\frac{1}{2}$ such that $(A_{i-1}A_{i-2}..A_1A_n...A_i)\left(i-\frac{1}{2}\right)=i-\frac{1}{2}$ for all $i=1,2,...,n-1$. Further, $(A_{i-1}A_{i-2}..A_1A_n...A_i)$ is continuous in X_i for all

$i = 2, .., n$ *because if* $x_i = i - \dfrac{3}{4}$ *is the point of discontinuity for* A_i *we have*

$$\begin{aligned} & \lim_{x \to \left(i - \frac{3}{4}\right)^-} \left(A_{i-1}A_{i-2}..A_1A_n..A_{i+1}A_i\right) x_i \\ = \quad & A_{i-1}A_{i-2}..A_1A_n..A_{i+1}\left(i + \frac{1}{4}\right) \\ = \quad & A_{i-1}A_{i-2}..A_1A_n..A_{i+2}\left(i + \frac{3}{4}\right) \end{aligned}$$

$$\begin{aligned} & \lim_{x \to \left(i - \frac{3}{4}\right)^+} \left(A_{i-1}A_{i-2}..A_1A_n..A_{i+1}A_i\right) x_i \\ = \quad & A_{i-1}A_{i-2}..A_1A_n..A_{i+1}\left(i + \frac{1}{2}\right) \\ = \quad & A_{i-1}A_{i-2}..A_1A_n..A_{i+2}\left(i + \frac{3}{4}\right) \end{aligned}$$

If we take $n = 2$ in Theorem 2.1, we obtain a generalisation of Theorem 2.1 of [3].

If we put $n = 4$ in the Theorem 2.1 we get the following corollary which generalizes the Theorem 2.1 of [15].

Corollary 2.3. *Let* $(X_i, M_i, \mathcal{N}_i, \theta_i, \gamma_i)_{1 \le i \le 4}$*, be four complete intuitionistic fuzzy metric spaces with* $M_i(x, x_i, t) \longrightarrow 1$ *and* $\mathcal{N}_i(x, x_i, t) \to 0$ *as* $t \longrightarrow \infty$ *for all* $x, x_i \in X_i$ *and let* $\{A_i\}_{i=1}^{i=4}$ *be 4-mappings such that* $A_i : X_i \to X_{i+1}$ *for all* $i = 1, 2, 3$ *and* $A_4 : X_4 \to X_1$*, satisfying the inequalities,*

$$(6.\mathrm{i}_M) \quad \phi_i \begin{pmatrix} M_i\left(A_{i-1}A_{i-2}..A_1A_4..A_{i+1}x_{i+1}, A_{i-1}A_{i-2}..A_1A_4..A_ix_i, kt\right), \\ M_i\left(x_i, A_{i-1}A_{i-2}..A_1A_4..A_{i+1}x_{i+1}, t\right), \\ M_i\left(x_i, A_{i-1}A_{i-2}..A_1A_4..A_ix_i, t\right), M_{i+1}\left(x_{i+1}, A_ix_i, t\right), \\ M_{i+1}\left(x_{i+1}, A_iA_{i-1}..A_1A_4..A_{i+1}x_{i+1}, t\right), \\ M_{i+1}\left(A_ix_i, A_iA_{i-1}..A_1A_4..A_{i+1}x_{i+1}, t\right) \end{pmatrix} > 0$$

$$(6.\mathrm{i}_N) \quad \psi_i \begin{pmatrix} \mathcal{N}_i\left(A_{i-1}A_{i-2}..A_1A_4..A_{i+1}x_{i+1}, A_{i-1}A_{i-2}..A_1A_4..A_ix_i, kt\right), \\ \mathcal{N}_i\left(x_i, A_{i-1}A_{i-2}..A_1A_4..A_{i+1}x_{i+1}, t\right), \\ \mathcal{N}_i\left(x_i, A_{i-1}A_{i-2}..A_1A_4..A_ix_i, t\right), \mathcal{N}_{i+1}\left(x_{i+1}, A_ix_i, t\right), \\ \mathcal{N}_{i+1}\left(x_{i+1}, A_iA_{i-1}..A_1A_4..A_{i+1}x_{i+1}, t\right), \\ \mathcal{N}_{i+1}\left(A_ix_i, A_iA_{i-1}..A_1A_4..A_{i+1}x_{i+1}, t\right) \end{pmatrix} < 0$$

for all $x_i \in X_i$*,* $x_{i+1} \in X_{i+1}$*,* $t > 0$*, where,* $\phi_i \in \Phi$*,* $\psi_i \in \Psi$*,* $i = 1, 2, 3, 4$ *and* $0 < k < 1$*. Further, suppose that for all* $i = \overline{1, 4}$*.* $\{A_2\}$ *is continuous on* X_i *Then*

$$\begin{aligned} A_4A_3A_2A_1 \text{ has a unique fixed point } p_1 &\in X_1 \\ A_1A_4A_3A_2 \text{ has a unique fixed point } p_2 &\in X_2 \\ A_2A_1A_4A_3 \text{ has a unique fixed point } p_3 &\in X_3 \\ A_3A_2A_1A_4 \text{ has a unique fixed point } p_4 &\in X_4. \end{aligned}$$

Further, $A_ip_i = p_{i+1}$ *for* $i = 1, 2, 3, 4$*.*

Proof. Let $\left\{x_r^{(i)}\right\}_{r\in\mathbb{N}}^{1\leq i\leq 4}$ be sequences in $\{X_i\}_{1\leq i\leq 4}$,. Now let $x_0^{(1)}$ be an arbitrary point in X_1, we define the sequences $\left\{x_r^{(i)}\right\}_{r\in\mathbb{N}}$ for $i=1,2,3,4$ by

$$\begin{array}{rcl} x_r^{(1)} &=& (A_4A_3A_2A_1)^r x_0^{(1)}, \\ x_r^{(2)} &=& A_1x_r^{(1)} \\ x_r^{(3)} &=& A_2A_1x_r^{(1)} = A_2x_r^{(2)}. \\ x_r^{(4)} &=& A_3A_2A_1x_r^{(1)} = A_3A_2x_r^{(2)} = A_3x_r^{(3)}. \\ x_r^{(1)} &=& A_4x_{r-1}^{(4)} \end{array}$$

Similarly of the proof in the case of three spaces, we can prove that $\left\{x_r^{(i)}\right\}$ is a Cauchy sequences in X_i for all $i=1,2,3,4$ with a limit

$$\begin{array}{rcl} p_1 &=& \lim\limits_{r\to\infty} x_r^{(1)}, \\ p_2 &=& \lim\limits_{r\to\infty} x_r^{(2)} = \lim\limits_{r\to\infty} A_1x_r^{(1)}, \\ p_3 &=& \lim\limits_{r\to\infty} x_r^{(3)} = \lim\limits_{r\to\infty} A_2x_r^{(2)} = \lim\limits_{r\to\infty} A_2A_1x_r^{(1)}, \\ p_4 &=& \lim\limits_{r\to\infty} x_r^{(4)} = \lim\limits_{r\to\infty} A_3A_2A_1x_r^{(1)} = \lim\limits_{r\to\infty} A_3A_2x_r^{(2)} = \lim\limits_{r\to\infty} A_3x_r^{(3)} \end{array}$$

To prove that p_i is a fixed point of $A_{i-1}..A_1A_4..A_ip_i$ for $i=1,2,3,4$ suppose that $A_{i-1}..A_1A_4..A_ip_i \neq p_i$. Using the inequality ($6.i_M$) for $x_i=p_i$, $x_{i+1}=x_{r-1}^{(i+1)} = A_ix_{r-1}^{(i)} = A_i..A_1x_{r-1}^{(1)}$ and $i=1,2,3$ we obtain

$$\phi_i\left(\begin{array}{c} M_i\left(A_{i-1}..A_1x_r^{(i)}, A_{i-1}..A_1A_4..A_ip_i, kt\right), \\ M_i\left(p_i, A_{i-1}..A_1x_r^{(i)}, t\right), \\ M_i\left(p_i, A_{i-1}..A_1A_4..A_ip_i, t\right), M_{i+1}\left(A_i..A_1x_{r-1}^{(i)}, A_ip_i, t\right), \\ M_{i+1}\left(A_i..A_1x_{r-1}^{(i)}, A_i..A_1x_r^{(i)}, t\right), \\ M_{i+1}\left(A_ip_i, A_i..A_1x_r^{(i)}, t\right) \end{array}\right) > 0$$

and so,

$$\phi_i\left(\begin{array}{c} M_i\left(x_r^{(i)}, A_{i-1}..A_1A_4..A_ip_i, kt\right), M_i\left(p_i, x_r^{(i)}, t\right), \\ M_i\left(p_i, A_{i-1}..A_1A_4A_3..A_ip_i, t\right), \\ M_{i+1}\left(x_{r-1}^{(i+1)}, A_ip_i, t\right), M_{i+1}\left(x_{r-1}^{(i+1)}, x_r^{(i+1)}, t\right), \\ M_{i+1}\left(A_ip_i, x_r^{(i+1)}, t\right) \end{array}\right) > 0 \tag{7.i$_M$}$$

for all $i=1,2,3$ and $\phi_i\in\Phi$.
In the case where $i=4$, we pose that $x_4=p_4$ and $x_1=x_r^{(1)}$, then we have,

$$\phi_4\left(\begin{array}{c} M_4\left(x_r^{(4)}, A_3A_2A_1A_4p_4, kt\right), M_4\left(p_4, x_r^{(4)}, t\right), \\ M_4\left(p_4, A_3A_2A_1A_4p_4, t\right), M_1\left(x_r^{(1)}, A_4p_4, t\right), \\ M_1\left(x_r^{(1)}, x_{r+1}^{(1)}, t\right), M_1\left(A_4p_4, x_{r+1}^{(1)}, t\right) \end{array}\right) > 0 \tag{7.4$_M$}$$

Letting $r \to \infty$ in $(7.i_M)$ and (7.4_M) we have,

$$\phi_i \begin{pmatrix} M_i\left(p_i, A_{i-1}..A_1A_4..A_ip_i, kt\right), 1, \\ M_i\left(p_i, A_{i-1}..A_1A_4A_3..A_ip_i, t\right), \\ M_{i+1}\left(p_{i+1}, A_ip_i, t\right), 1, M_{i+1}\left(A_ip_i, p_{i+1}, t\right) \end{pmatrix} > 0$$

and

$$\phi_4 \begin{pmatrix} M_4\left(p_4, A_3A_2A_1A_4p_4, kt\right), 1, \\ M_4\left(p_4, A_3A_2A_1A_4p_4, t\right), M_1\left(p_1, A_4p_4, t\right), \\ 1, M_1\left(A_4p_4, p_1, t\right) \end{pmatrix} > 0$$

It follows from (iii) that

$$M_i\left(p_i, A_{i-1}..A_1A_4..A_ip_i, kt\right) \geq M_{i+1}\left(p_{i+1}, A_ip_i, t\right) \tag{8.i$_M$}$$

for all $i = 1, 2, 3, 4$ which mean again,

$$\begin{aligned} M_1\left(p_1, A_4A_3A_2A_1p_1, kt\right) &\geq M_2\left(p_2, A_1p_1, t\right) && (8.1_M) \\ M_2\left(p_2, A_1A_4A_3A_2p_2, kt\right) &\geq M_3\left(p_3, A_2p_2, t\right) && (8.2_M) \\ M_3\left(p_3, A_2A_1A_4A_3p_3, kt\right) &\geq M_4\left(p_4, A_3p_3, t\right) && (8.3_M) \\ M_4\left(p_4, A_3A_2A_1A_4p_4, kt\right) &\geq M_1\left(p_1, A_4p_4, t\right) && (8.4_M) \end{aligned}$$

Suppose that A_2 is continuous. Then, from (8.2_M) we get,

$$\begin{aligned} A_2p_2 &= p_3 && (9.1) \\ A_1A_4A_3A_2p_2 &= p_2 && (9.2) \\ A_1A_4A_3p_3 &= p_2 && (9.3) \\ A_2A_1A_4A_3p_3 &= A_2p_2 = p_3. && (9.4) \end{aligned}$$

Using the inequality $(6.i_M)$ for $x_i = x_r^{(i)} = A_{i-1}..A_1x_r^{(1)}$ and $x_{i+1} = p_{i+1}$ we have

$$\phi_i \begin{pmatrix} M_i\left(A_{i-1}..A_1A_4..A_{i+1}p_{i+1}, x_{r+1}^{(i)}, kt\right), \\ M_i\left(x_r^{(i)}, A_{i-1}..A_1A_4..A_{i+1}p_{i+1}, t\right), M_i\left(x_r^{(i)}, x_{r+1}^{(i)}, t\right), \\ M_{i+1}\left(p_{i+1}, x_r^{(i+1)}, t\right), M_{i+1}\left(p_{i+1}, A_i..A_1A_4..A_{i+1}p_{i+1}, t\right), \\ M_{i+1}\left(x_r^{(i+1)}, A_i..A_1A_4..A_{i+1}p_{i+1}, t\right) \end{pmatrix} > 0 \tag{9.i$_M$}$$

$$\phi_3 \begin{pmatrix} M_3\left(A_2A_1A_4p_4, x_{r+1}^{(3)}, kt\right), M_3\left(x_r^{(3)}, A_2A_1A_4p_4, t\right), \\ M_3\left(x_r^{(3)}, x_{r+1}^{(3)}, t\right), M_4\left(p_4, x_r^{(4)}, t\right), \\ M_4\left(p_4, A_3A_2A_1A_4p_4, t\right), M_4\left(x_r^{(4)}, A_3A_2A_1A_4p_4, t\right) \end{pmatrix} > 0$$

$$\phi_4 \begin{pmatrix} M_4\left(A_3A_2A_1p_1, x_{r+1}^{(4)}, kt\right), \\ M_4\left(x_r^{(4)}, A_3A_2A_1p_1, t\right), M_4\left(x_r^{(4)}, x_{r+1}^{(4)}, t\right), \\ M_1\left(p_1, x_{r+1}^{(1)}, t\right), M_1\left(p_1, A_4A_3A_2A_1p_1, t\right), \\ M_1\left(x_{r+1}^{(1)}, A_4A_3A_2A_1p_1, t\right) \end{pmatrix} > 0$$

Letting $r \to \infty$ in $(9.i_M)$ we have

$$\phi_i \begin{pmatrix} M_i\left(A_{i-1}..A_1A_4..A_{i+1}p_{i+1}, p_i, kt\right), \\ M_i\left(p_i, A_{i-1}..A_1A_4..A_{i+1}p_{i+1}, t\right), 1, 1, \\ M_{i+1}\left(p_{i+1}, A_i..A_1A_4..A_{i+1}p_{i+1}, t\right), \\ M_{i+1}\left(p_{i+1}, A_i..A_1A_4..A_{i+1}p_{i+1}, t\right) \end{pmatrix} > 0$$

which it means,

$$\phi_1\left(\begin{array}{c} M_1(A_4A_3A_2p_2,p_1,kt), M_1(p_1,A_4A_3A_2p_2,t),1,1, \\ M_2(p_2,A_1A_4A_3A_2p_2,t), M_2(p_2,A_1A_4A_3A_2p_2,t) \end{array}\right)>0$$

$$\phi_2\left(\begin{array}{c} M_2(A_1A_4A_3p_3,p_2,kt), M_2(p_2,A_1A_4A_3p_3,t),1,1, \\ M_3(p_3,A_2A_1A_4A_3p_3,t), M_3(p_3,A_2A_1A_4A_3p_3,t) \end{array}\right)>0$$

$$\phi_3\left(\begin{array}{c} M_3(A_2A_1A_4p_4,p_3,kt), M_3(p_3,A_2A_1A_4p_4,t),1,1, \\ M_4(p_4,A_3A_2A_1A_4p_4,t), M_4(p_4,A_3A_2A_1A_4p_4,t) \end{array}\right)>0$$

$$\phi_4\left(\begin{array}{c} M_4(A_3A_2A_1p_1,p_4,kt), M_4(p_4,A_3A_2A_1p_1,t),1,1, \\ M_1(p_1,A_4A_3A_2A_1p_1,t), \\ M_1(p_1,A_4A_3A_2A_1p_1,t) \end{array}\right)>0$$

It follows from (iii) that:

$$\begin{aligned} M_1(A_4A_3p_3,p_1,kt) &\geq M_2(p_2,A_1A_4A_3p_3,t), && (10.1_M)\\ M_2(A_1A_4A_3p_3,p_2,kt) &\geq M_3(p_3,A_2A_1A_4A_3p_3,t), && (10.2_M)\\ M_3(A_2A_1A_4p_4,p_3,kt) &\geq M_4(p_4,A_3A_2A_1A_4p_4,t), && (10.3_M)\\ M_4(A_3A_2A_1p_1,p_4,kt) &\geq M_1(p_1,A_4A_3A_2A_1p_1,t). && (10.4_M) \end{aligned}$$

Then, from $(9.1),(9.2)$ and (10.1_M) we have

$$A_4A_3p_3=p_1. \tag{11.1}$$

From $(9.4),(10.2_M)$ and (11.1) we get,

$$A_1p_1=p_2 \tag{11.2}$$

From $(10.4_M),(11.1)$ and (11.2) we obtain,

$$\begin{aligned} A_3p_3 &= p_4 && (11.3)\\ A_4A_3p_3 &= A_4p_4=p_1 && (11.4) \end{aligned}$$

Now, substitute (11.2) and (11.4) in (7.4_M) and (8.4_M) respectively, we get

$$\begin{aligned} A_4A_3A_2A_1p_1 &= p_1\\ A_3A_2A_1A_4p_4 &= p_4 \end{aligned}$$

Concerning inequalities $(6.i_{\mathcal{N}})$, we can prove by the same manner that

$$A_{i-1}..A_1A_4..A_ip_i=p_i \text{ for all } i=1,2,3,4.$$

For proving the uniqueness of the fixed point p_i in X_i we assume that there exists $z_i\in X_i$ such that $z_i\neq p_i$ and $A_{i-1}..A_1A_3..A_iz_i=z_i$ for all $i=1,2,3,4$. Firstly, using $(6.i_M)$ *for* $x_i=p_i$ and $x_{i+1}=A_iz_i$ *we have,*

$$\phi_i\left(\begin{array}{c} M_i(A_{i-1}..A_1A_4..A_{i+1}A_iz_i, A_{i-1}..A_1A_4..A_ip_i,kt), \\ M_i(p_i,A_{i-1}..A_1A_4..A_{i+1}A_iz_i,t), \\ M_i(p_i,A_{i-1}..A_1A_4..A_ip_i,t), M_{i+1}(A_iz_i,A_ip_i,t), \\ M_{i+1}(A_iz_i,A_i..A_1A_4..A_{i+1}A_iz_i,t), \\ M_{i+1}(A_ip_i,A_i..A_1A_4..A_{i+1}A_iz_i,t) \end{array}\right)>0$$

and so,

$$\phi_i\left(\begin{array}{c} M_i(z_i,p_i,kt), M_i(p_i,z_i,t),1, \\ M_{i+1}(A_iz_i,A_ip_i,t),1,M_{i+1}(A_ip_i,A_iz_i,t) \end{array}\right)>0$$

which implies that,

$$M_i(p_i,z_i,kt)\geq M_{i+1}(A_ip_i,A_iz_i,t) \text{ for all } i=1,2,3,4$$

From $(9.1),(11.2),(11.3)$ and (11.4) we have:

$$\begin{aligned}
M_1(p_1,z_1\ ,kt) &\geq M_2(A_1p_1,A_1z_1\ ,t)=M_2(p_2,z_2\ ,t)\\
M_2(p_2,z_2\ ,kt) &\geq M_3(A_2p_2,A_2z_2\ ,t)=M_3(p_3,z_3\ ,t)\\
M_3(p_3,z_3\ ,kt) &\geq M_4(A_3p_3,A_3z_3\ ,t)=M_4(p_4,z_4\ ,t)\\
M_4(p_4,z_4\ ,kt) &\geq M_1(A_4p_4,A_4z_4\ ,t)=M_1(p_1,z_1\ ,t).
\end{aligned}$$

which implies again that,

$$\begin{aligned}
M_1(p_1,z_1\ ,kt) &\geq M_1(p_1,z_1\ ,t)\\
M_2(p_2,z_2\ ,kt) &\geq M_2(p_2,z_2\ ,t)\\
M_3(p_3,z_3\ ,kt) &\geq M_3(p_3,z_3\ ,t)\\
M_4(p_4,z_4\ ,kt) &\geq M_4(p_4,z_4\ ,t).
\end{aligned}$$

Similarly, we find by $(6.i_{\mathcal{N}})$ that

$$\begin{aligned}
\mathcal{N}_1(p_1,z_1\ ,kt) &\leq \mathcal{N}_1(p_1,z_1\ ,t)\\
\mathcal{N}_2(p_2,z_2\ ,kt) &\leq \mathcal{N}_2(p_2,z_2\ ,t)\\
\mathcal{N}_3(p_3,z_3\ ,kt) &\leq \mathcal{N}_3(p_3,z_3\ ,t).\\
\mathcal{N}_4(p_4,z_4\ ,kt) &\leq \mathcal{N}_4(p_4,z_4\ ,t)
\end{aligned}$$

This proving the uniqueness of p_i in X_i for $i=1,2,3,4$. This complete the proof of the theorem. □

The following example illustrates our Corollary 2.3.

Example 2.4. *Let* $X_1=[0,1],X_2=[1,2[,X_3=]2,3],X_4=]3,4]$ *and let* $M_i(x_i,x_{i+1},t)=\dfrac{t}{t+|x_{i+1}-x_i|}$ *for all* $i=\overline{1,3}$ *and* $M_5(x_5,x_1,t)=\dfrac{t}{t+|x_5-x_1|}$. *Define* $A_1:X_1\to X_2,\ A_2:X_2\to X_3,\ A_3:X_3\to X_4$ *and* $A_4:X_4\to X_1$ *by*

$$\begin{aligned}
A_1x_1 &= \begin{cases} 1 \ \textit{if}\ x_1\in[0,\frac{3}{4}[\\ \frac{3}{2}\ \textit{if}\ x_1\in[\frac{3}{4},1]\end{cases},\ A_2x_2=\begin{cases}\frac{5}{2}\ \textit{if}\ x_2\in[1,2[\end{cases}\\
A_3x_3 &= \begin{cases} \frac{13}{4}\ \textit{if}\ x_3\in]2,\frac{5}{2}[\\ \frac{7}{2}\ \textit{if}\ x_3\in[\frac{5}{2},3]\end{cases},\ A_4x_4=\begin{cases}\frac{3}{4}\ \textit{if}\ x_4\in\left]3,\frac{7}{2}\right[\\ 1\ \textit{if}\ x_4\in\left[\frac{7}{2},4\right]\end{cases}
\end{aligned}$$

Let $\phi_1(t_1,t_2,t_3,t_4,t_5,t_6)=t_1-\min\{t_2,t_3,t_4,t_5,t_6\}$ *and* $\phi_1=\phi_2=\phi_3=\phi_4$. *Here* X_1 *is compact, but the others spaces are not compact. Further the inequalities* $(5.1),(5.2),(5.3),(5.4),(5.5)$ *are satisfied since the left hand side of each inequality is* 1 *and* $A_4A_3A_2A_1$ *is continuous in* X_1 *because if* $x=\frac{3}{4}$, *the point of discontinuity* *for A_1, we get*

$$\begin{aligned}
\lim_{x\to\frac{3}{4}^-} A_4A_3A_2A_1x &= A_4A_3A_2\left(1\right) = A_4A_3\left(3\right) = A_4\left(\frac{7}{2}\right) = A_5\left(\frac{9}{2}\right) = 1 \\
&\quad \textit{and} \\
\lim_{x\to\frac{3}{4}^+} A_4A_3A_2A_1x &= A_4A_3A_2\left(\frac{3}{2}\right) = A_4A_3\left(3\right) \\
&= A_4\left(\frac{7}{2}\right) = A_5\left(\frac{9}{2}\right) = 1
\end{aligned}$$

We have

$$\begin{aligned}
A_4A_3A_2A_1\left(1\right) &= 1,\ A_1A_4A_3A_2\left(\frac{3}{2}\right) = \frac{3}{2}, \\
A_2A_1A_4A_3\left(\frac{5}{2}\right) &= \frac{5}{2},\ A_3A_2A_1A_4\left(\frac{7}{2}\right) = \frac{7}{2}.
\end{aligned}$$

References

[1] C. Alaca, D. Turkoglu, C. Yildiz, Fixed points in intuitionistic fuzzy metric spaces. Chaos, Solitons & Fractals 2006;29:1073–8.

[2] A. Aliouche and B. Fisher, Fixed point theorems for mappings satisfying implicit relation on two complete and compact metric spaces, Applied Mathematics and Mechanics., 27 (9) (2006), 1217-1222.

[3] A. Aliouche, F. Merghadi and A. Djoudi, A Related Fixed Point Theorem in two Fuzzy Metric Spaces, J. Nonlinear Sci. Appl., 2 (1) (2009), 19-24.

[4] Y. J. Cho, Fixed points in fuzzy metric spaces, J. Fuzzy. Math., 5 (4) (1997), 949-962.

[5] Z. K. Deng . Fuzzy psendo-metric spaces. J Math Anal Appl 1982;86:74–95.

[6] M. A. Ereeg . Metric spaces in fuzzy set theory. J Math Anal Appl 1979;69:338–53.

[7] A. George and P. Veeramani, On some result in fuzzy metric space, Fuzzy Sets Syst., 64 (1994), 395-399.

[8] M. Grabiec, Fixed points in fuzzy metric spaces Fuzzy Sets Syst., 27 (1988), 385-389.

[9] I. Kramosil and J. Michalek, Fuzzy metric and statistical metric spaces, Kybernetica., 11 (1975), 326-334.

[10] S. Kutukcu, Compatibility and Common Coincidence Points in Intuitionistic Fuzzy Metric Spaces, Southeast Asian Bulletin Math., 32 (2008), 1081-1089.

[11] R. Lowen, Fuzzy set theory. Dordrecht: Kluwer Academic Publishers; 1996.

[12] F. Merghadi AND A. Aliouche, A related fixed point theorem in n- fuzzy metric spaces, Iranian Journal of Fuzzy Systems Vol. 7, No. 3, (2010) pp. 73-86.

[13] J. Park, Intuitionistic fuzzy metric spaces. Chaos, Solitons & Fractals 2004;22:1039–46.

[14] V. Popa, Some fixed point theorems for compatible mappings satisfying an implicit relation, Demonstratio Math., 32 (1999),157-163.

[15] K. P. R. Rao, A. Aliouche and G. R. Babu, Related Fixed Point Theorems in Fuzzy Metric Spaces, J. Nonlinear Sci. Appl., 1 (3) (2008), 194-202.

[16] R. Saadati and J.H. Park, On the intuitionistic fuzzy topological spaces, Chaos, Solitons and Fractals 27 (2006) 331–344.

[17] B. Schweizer and A. Sklar, Statistical metric spaces. Pacific J. Math., 10 (1960), 313-334.

[18] S. Sharma and B. Deshpande, Compatible mappings of type (I) and (II) on intuitionistic fuzzy metric spaces in consideration of common fixed point, Commun. Korean Math. Soc. 24 (2009), No. 2, pp. 197-214.

[19] Y. Tanaka, Y. Mizno and T. Kado, Chaotic dynamics in Friedmann equation. Chaos, Solitons and Fractals., 24 (2005), 407–422.

[20] L. A. Zadeh, Fuzzy sets, Inform and Control, 8 (1965), 338-353.

F. Merghadi, Department of Mathematics, University of Tebessa, 12000, Algeria.
E-mail address: faycel_mr@yahoo.fr

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 1, 136-146, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

A MODIFIED PARTIAL QUADRATIC INTERPOLATION METHOD FOR UNCONSTRAINED OPTIMIZATION

T.M. EL-GINDY, M.S. SALIM, AND ABDEL–RAHMAN IBRAHIM

Abstract. A numerical method for solving unconstrained optimization problems is presented. It is a modification of the partial quadratic interpolation method [11] for unconstrained optimization and based upon approximating the gradient and the Hessian of the objective function. This means that it requires only the expression of the objective function to converges to a stationary point of the problem from any initial point, with speed convergence. The method can solve complex problems in which direct calculations of the gradient and Hessian matrix are difficult or even impossible to calculate. The search directions are always descent directions. Results and comparisons are given at the end of the paper and show that this method is interesting.

1. Introduction

In this paper, we consider the unconstrained optimization problem

$$\min f(x), \quad x \in \mathbb{R}_n \tag{1.1}$$

where $f : \mathbb{R}_n \to \mathbb{R}$ is a continuously differentiable function, and its gradient at a point x^r, $r \in \mathbb{N}$, is denoted by $g(x^r)$, or g^r for simplicity, n is the number of variables. One of the most effective methods for solving the unconstrained problem (1.1) is the Newton method. It normally requires the fewest number of function evaluations, and is very good at handling ill-conditioning. However, its efficiency largely depends on the possibility of solving efficiently a linear system which arises when computing the search direction d^r at each iteration,

$$H(x^r)d^r = -g(x^r) \tag{1.2}$$

Where $H(x^r)$ is the matrix of second partial derivatives (Hessian matrix) of f and d^r is a search direction in the current iteration. Moreover, the exact solution of the linear system (1.2) could be too burdensome, or is not necessary when x^r is far from the solution of f (see [25, 27]).

It is emphasized [3] here that unless $[H(x^r)]$ is positive definite, the direction $-[H(x^r)]^{-1}[g(x^r)]$ will not be that of descent for the objective function. to see this we substitute the direction into the descent condition to obtain

$$-[g(x^r)]^T[H(x^r)]^{-1}[g(x^r)] < 0 \tag{1.3}$$

The foregoing condition will always be satisfied if $[H(x^r)]$ is positive definite. If $[H(x^r)]$ is negative definite or negative semidefinite the condition is always violated. With $[H(x^r)]$ as indefinite or positive semidefinite, the condition may or may not

Key words and phrases. Unconstrained optimization, descent direction, partial quadratic interpolation method.

2010 *AMS Math. Subject Classification.* 65K05, 90C30.

be satisfied, so we must check for it. If the direction is not descent for the objective function, we should stop there because a positive step size cannot be determined.

There exist some kinds of effective methods available for solving (1.1), as for instance, inexact Newton, limited memory quasi-Newton, truncated Newton (TN), conjugate gradient, spectral gradient, and subspace methods. Inexact Newton methods (see [10, 25]) represent the basic approach underlying most of the Newton-type large-scale unconstrained algorithms. At each step, the current estimate of the solution is updated by approximately solving the linear system (1.2) using an iterative algorithm. The inner iteration is typically " truncated " before the solution of the linear system is obtained.

In Quasi-Newton methods the Hessian matrix of second derivatives of the function to be minimized does not need to be computed. The Hessian is updated by analyzing successive gradient vectors instead. Quasi-Newton methods are a generalization of the secant method to find the root of the first derivative for multidimensional problems. In multi-dimensions the secant equation is under-determined, and quasi-Newton methods differ in how they obtain the solution, typically by adding a simple low-rank update to the current estimate of the Hessian.

The limited memory BFGS (L-BFGS) method (see [20, 26, 33]) is a member of the broad family of quasi-Newton optimization methods. It use a low and predictable amount of storage, and only require the function and gradient values at each iteration and no other information about the problem. They are suitable for large scale problems because the amount of storage required by the algorithms (and thus the cost of the iteration) can be controlled by the user. Alternatively, limited memory methods can be viewed as implementations of quasi-Newton methods, in which storage is restricted. Their simplicity is one of their main appeals: they do not require knowledge of the sparsity structure of the Hessian, or knowledge of the separability of the objective function, and they can be very simple to program. Limited memory methods originated with the work of Perry (1977) and Shanno (1978), and were subsequently developed and analyzed by Buckley (1978), Nazareth (1979), Nocedal (1980), Gill and Murray (1979) and Buckley and LeNir (1983). Numerical tests performed on medium size problems have shown that limited memory methods require substantially fewer function evaluations than the conjugate gradient method, even when little additional storage is added. The implementation is almost identical to that of the standard BFGS method, the only difference is that the inverse Hessian approximation is not formed explicitly, but defined by a small number of BFGS updates. It often provides a fast rate of linear convergence, and requires minimal storage.

If only gradients are available and memory is limited, one may use a discrete truncated Newton method (DTN) such as in [10, 26]. DTN uses gradients to approximate the product of the Hessian with an arbitrary vector. Several attempts have been made to create a method which combines the properties of the (discrete) truncated Newton method and the L-BFGS method.

Conjugate gradient methods have played special roles in solving large scale nonlinear optimization problems. Although conjugate gradient methods are not the fastest or most robust optimization algorithms for nonlinear problems available today, they remain very popular for engineers and mathematicians with solving large problems [13, 17, 22, 36, 37]. The convergence properties of conjugate gradient methods have been studied by many researchers [1, 8, 9, 15, 16, 28, 31]. Good reviews of the conjugate gradient methods can be found in [13, 18]. Although all these methods are equivalent in the linear case, that is, when f is a strictly convex quadratic function and α^r is computed by an exact line search, their behavior for general functions may be quite different. The search direction in all nonlinear conjugate gradient methods is given by $d^r = -g^r + \beta^r d^{r-1}$,with $d^0 = -g^0$ and β^r being a scalar. Most of the recent work on nonlinear conjugate gradient methods is focused on the design of a new β^r or a new line search strategy. The large-scale unconstrained optimization problems have received much attention in recent decades.We refer to [2, 18] for a good survey.

Spectral gradient methods have proved to be of great value in unconstrained optimization problems. They were introduced by Barzilai and Borwein [4],and analyzed by Raydan [29]. They have been applied to find local minimizers of large scale problems [5, 6, 30], and also to explore faces of large dimensions in box-constrained optimization see [12]. More recently, spectral gradient methods were extended to minimize general smooth functions on convex sets See [7]. In this case, the spectral choice of step length was combined with the projected gradient method [14, 19, 34] to obtain a robust and effective low cost computational scheme.

Subspace techniques [32, 35] used in constructing of numerical methods for nonlinear optimization. The subspace techniques are getting more and more important as the optimization problems we have to solve are getting larger and larger in scale. The applications of subspace techniques have the advantage of reducing both computation cost and memory size. Actually in many standard optimization methods (such as conjugate gradient method, limited memory quasi- Newton method, projected gradient method, and null space method) there are ideas or techniques that can be viewed as subspace techniques. The essential part of a subspace method is how to choose the subspace in which the trial step or the trust region should belong.

The iterative formula of these methods is given by

$$x^{r+1} = x^r + \alpha^r d^r \tag{1.4}$$

where $\alpha^r > 0$ is a step length. Generally [3], we can say that d^r is a descent direction of f at x^r if

$$g^r \cdot d^r < 0 \tag{1.5}$$

The classical Newton's method is the basis of many numerical methods for unconstrained optimization problems. If, in the case of minimization, the Hessian matrix of the function under consideration is positive definite, then a numerically stable method for forming the descent direction is to factorize the Hessian matrix by the method of cholesky. In cases where the matrix is singular or, more commonly, indefinite, Newton's algorithm is no longer a steepest descent algorithm. In addition the method of cholesky is no longer numerically stable, even if the factorization of Hessian exists. there have been several algorithms [23] which replace the Hessian matrix in Newton's method by a related positive definite matrix. The disadvantages of such methods is that either they cease to determine the descent direction in a numerically stable way, and/or the amount of work necessary to evaluate this direction is considerably greater than that required by cholesky's method.

The partial quadratic interpolation method [11] or shortly (P.Q.I.) technique may be considered as a second-order method, the principal idea of this method is to approximate the objective function $f(x)$ about certain point x^r by second degree polynomial in the space $\mathbb{R}_n$, from which one may determined approximations to the gradient and Hessian of this function as $[b_n(x^r)]$ and $[A_n(x^r)]$ respectively. One then extract a positive definite matrix$[A_p(x^r)]$ from the Hessian matrix $[A_n(x^r)]$ in the subspace $\mathbb{R}_p \subset \mathbb{R}_n$, using a particular application of the cholesky technique.

PQI technique.

Step 1. Choose some guessed starting point $x^r \in \mathbb{R}_n$.
Step 2. Approximate the function $f(x)$ about x^r in a quadratic form:

$$f(x) \simeq p(x) = a + [b_n(x^r)]^T[x - x^r] + \frac{1}{2}[x - x^r]^T[A_n(x^r)][x - x^r].$$

the constant a; and the matrices $[b_n(x^r)]$ and $[A_n(x^r)]$ are computed using certain interpolation points [11].

Step 3. Use the necessary condition for the minimum of $p(x)$ we have

$$x^{r+1} = x^r - [A_n(x^r)]^{-1}[b_n(x^r)]$$

or

$$[A_n(x^r)][\Lambda] = -[b_n(x^r)]$$

where

$$[\Lambda] = x^{r+1} - x^r$$

Step 4. Extract the symmetric positive definite matrix $[A_p]$ from the symmetric matrix $[A_n]$ using Choleski's method, where:

$$[A_p] = [L_p][L_p]^T$$

Step 5. Solve the system of the equations:

$$[L_p][L_p]^T[\Lambda_i] = [b_p], \qquad \Lambda_i \in \mathbb{R}_p$$

which is equivalent to the solution of the two systems of equations

$$[L_p][Y] = [b_p]$$

$$[L_p]^T[\Lambda_i] = [Y].$$

Step 6. The new point, or the next guess is then given by:

$$x_i^{r+1} = x_i^r + \alpha\Lambda_i \qquad \text{for} \qquad x_i \in \mathbb{R}_p, \quad 0 \le \alpha \le 1$$

and

$$x_i^{r+1} = x_i^r \qquad \text{for} \qquad x_i \notin \mathbb{R}_p$$

Step 7. To determine the new point $x^{r+1} \in \mathbb{R}_n$, we try $\alpha = 1, \frac{1}{2}, \frac{1}{4}, ...$ and take the first value of α which satisfies $f(x^{r+1}) < f(x^r)$, $x^r \in \mathbb{R}_n$. If α becomes sufficiently small without satisfying this condition, we use new interpolation points, and restart the calculation,

The method has many advantages:

- It avoids direct calculations of the gradient and Hessian matrix, which are difficult or even impossible to calculate in certain problems.
- It avoids evaluation of the inverse matrix needed at each iteration.
- It ensures convergence to a local minimum, regardless of the choice of initial value.

But in some problems it has two disadvantages:

- The fixing of the variables $x_j \in \mathbb{R}_{n-p}$ may cause the convergence to be slow.
- The operation of canceling certain rows and columns may at certain point involves all the rows and columns of the Hessian matrix A_n, and then, the procedure stop before achieving the solution.

To overcome these disadvantages we suggest a modification to the (P.Q.I.) technique. The modification is to complete the direction obtained from the (P.Q.I.) technique from a certain vector, which is a descent vector, and this produce a faster descent direction converges to the minimum value of the objective function f.

This paper is organized as follows. In Section 2, we describe the modified partial quadratic interpolation technique . Numerical results and one conclusion are presented in Section 3 and in Section 4, respectively. Throughout the paper, $\|\cdot\|$ denotes the Euclidean norm of vectors.

2. Modified P.Q.I. method

In this section, we propose a new algorithm to solve (1.1). This algorithm generates a sequence of points x^r by

$$x^{r+1} = x^r + \alpha^r d^r, \qquad r = 0, 1, 2, \dots$$

where d^r is a descent direction of f at x^r , and α^r is the step length which is determined by a line search. In the following, we describe the method in details.

Let $f(x), x \in \mathbb{R}_n$ be a continuous function, $[g_n(x^r)]$ is a column matrix representing the gradient vector of f at x^r, and $[H_n(x^r)]$ is the matrix of second partial derivatives (Hessian matrix) of f evaluated at the point x^r. Let $[H_p(x^r)]$ is the positive definite matrix extracted from $[H_n(x^r)]$ and $[g_p(x^r)]$ is the corresponding gradient vector of $f(x)$ in the sub-space $\mathbb{R}_p \subset \mathbb{R}_n$. Define $[g_{n-p}(x^r)]$ as the gradient of $f(x)$ with respect to the variables $x_j \notin \mathbb{R}_p$ or $x_j \in (\mathbb{R}_n - \mathbb{R}_p)$, where,
$[x_p]$ is the column vector of the components $x_i \in \mathbb{R}_p$.
$[x_{n-p}]$ is the column vector of the components $x_j \notin \mathbb{R}_p$ or $x_j \in (\mathbb{R}_n - \mathbb{R}_p)$.

The P.Q.I. method have a search direction in the form $-[H_p(x^r)]^{-1}[g_p(x^r)]$, and we propose the search direction to be,

$$d^r = \begin{bmatrix} -[H_p(x^r)]^{-1}[g_p(x^r)] \\ -c[g_{n-p}(x^r)] \end{bmatrix}$$

where,

$$c = \begin{cases} 1 & \text{if } r = 0, \\ \|[g_n(x^r)]\|^2/\|[H_n(x^r)][q_n(x^r)]\|^2, \quad [q_n(x^r)] = [g_n(x^r)] - [g_n(x^{r-1})] & \text{if } r = 1, 2, \dots \end{cases}$$

This mean that in first iteration $c = 1$ and in the subsequent iterations $c = \|[g_n(x^r)]\|^2/\|[H_n(x^r)][q_n(x^r)]\|^2$,
$[q_n(x^r)] = [g_n(x^r)] - [g_n(x^{r-1})]$. We note that d^r is a descent direction since, it satisfies (1.5) i.e.

$$\begin{bmatrix} [g_p(x^r)] \\ [g_{n-p}(x^r)] \end{bmatrix} \cdot \begin{bmatrix} -[H_p(x^r)]^{-1}[g_p(x^r)] \\ -c[g_{n-p}(x^r)] \end{bmatrix} < 0$$

Theorem 2.1. *Let* $f : \mathbb{R}_n \to \mathbb{R}$, $x^r \in \mathbb{R}_n$ *and* x^r *is not a critical point and, If the Hessian matrix* $[H_p(x^r)]$ *is a positive definite matrix. Then the direction*

$$d_n(x^r) = \begin{bmatrix} -[H_p(x^r)]^{-1}[g_p(x^r)] \\ -c[g_{n-p}(x^r)] \end{bmatrix}$$

where,

$$c = \begin{cases} 1 & \text{if } r = 0, \\ \|[g_n(x^r)]\|^2/\|[H_n(x^r)][q_n(x^r)]\|^2, \quad [q_n(x^r)] = [g_n(x^r)] - [g_n(x^{r-1})] & \text{if } r = 1, 2, \ldots \end{cases}$$

is a descent direction.

Proof. Approximating $f(x^r + td_n(x^r))$ by using Taylor's series expansion about the point x^r, we get

$$\begin{aligned} f(x^r + td_n(x^r)) &= f(x^r) + t[g_n(x^r)]^T[d_n(x^r)] + O(t^2) \\ &= f(x^r) + t\begin{bmatrix} [g_p(x^r)] \\ [g_{n-p}(x^r)] \end{bmatrix}^T \begin{bmatrix} -[H_p(x^r)]^{-1}[g_p(x^r)] \\ -c[g_{n-p}(x^r)] \end{bmatrix} + O(t^2) \\ &= f(x^r) - t\{[g_p(x^r)]^T[H_p(x^r)]^{-1}[g_p(x^r)] + c[g_{n-p}(x^r)]^T[g_{n-p}(x^r)]\} + O(t^2) \end{aligned} \tag{2.1}$$

If $t > 0$ is small, then the error term $O(t^2)$ will be neglected.
therefore, eq. (2.1) take the form

$$f(x^r+td_n(x^r))-f(x^r) = -t\{[g_p(x^r)]^T[H_p(x^r)]^{-1}[g_p(x^r)]+c[g_{n-p}(x^r)]^T[g_{n-p}(x^r)]\} \tag{2.2}$$

but $[H_p(x^r)]$ is a positive definite matrix, then

$$-[g_p(x^r)]^T[H_p(x^r)]^{-1}[g_p(x^r)] < 0$$

also $c[g_{n-p}(x^r)]^T[g_{n-p}(x^r)]$, and t are always positive. Then the right-hand side of equation (2.2) is negative, so that:

$$f(x^r + td_n(x^r)) < f(x^r)$$

Then $d_n(x^r)$ is a descent direction. □

Let $[b_n(x^r)]$ is a column matrix (b_i) representing the approximate value of the gradient $[g_n(x^r)]$, and $[A_n(x^r)]$ is a square matrix (a_{ij}) of order n representing the approximate value of the Hessian $[H_n(x^r)]$. If we use the approximate values $[A_n(x^r)]$ and $[b_n(x^r)]$ of the Hessian and gradient, we get the same result.
Using these approximate values,we can write

$$f(x^r + td_n(x^r)) = f(x^r) + t[b_n(x^r)]^T[d_n(x^r)] + \eta(t, [\ell]) \tag{2.3}$$

where $\eta(t, [\ell])$ is the error term, and $[\ell] = (\ell_1, \ell_2, ..., \ell_n)$.
In the same way equation (2.1), becomes

$$\begin{aligned} f(x^r + td_n(x^r)) &= f(x^r) - t\{[b_p(x^r)]^T[A_p(x^r)]^{-1}[b_p(x^r)] \\ &+ c[b_{n-p}(x^r)]^T[b_{n-p}(x^r)]\} + \eta(t, [\ell]) \end{aligned} \tag{2.4}$$

As before the error term will be neglected when t and $[\ell]$ are small, and hence

$$f(x^r + td_n(x^r)) < f(x^r). \tag{2.5}$$

Let $f(x)$ be defined over the space $\mathbb{R}_n$, then the approximation $[A_n(x^r)]$ of the Hessian is not, in general strictly positive definite. Now we extract a positive definite matrix $[A_p(x^r)]$ from the Hessian matrix $[A_n(x^r)]$, and hence we construct the vector $\begin{bmatrix} -[A_p(x^r)]^{-1}[b_p(x^r)] \\ -c[b_{n-p}(x^r)] \end{bmatrix}$, which is from our previous theorem a descent direction. This direction generates a sequence of points $x^1, x^2, ..., x^r, ...$ converges to the optimum solution from any initial point x^0.

Now, we state the steps of the modified partial quadratic interpolation (modified P.Q.I.) algorithm as follows.

Algorithm 2.2.

We assume that $f(x)$ is a given continuous, scalar valued function, where $x \in \mathbb{R}_n$ and we seek to determined the point x^* such that:

$$f(x^*) < f(x) \qquad \text{for all } x \text{ near to } x^*$$

Step 1. Choose an initial point $x^0 \in \mathbb{R}_n$. Select a convergence parameter $\varepsilon > o$ and set $r = 0$.

Step 2. Approximate the function $f(x)$ about x^r in a quadratic form:

$$f(x) \simeq a + [b_n(x^r)]^T[x - x^r] + \frac{1}{2}[x - x^r]^T[A_n(x^r)][x - x^r].$$

Step 3. Calculate the matrices $[b_n(x^r)]$ and $[A_n(x^r)]$ by the following relations,

$$b_i = \frac{f(x^r_{i_+}) - f(x^r_{i_-})}{2\ell_i}$$

$$a_{ii} = \frac{f(x^r_{i_+}) + f(x^r_{i_-}) - 2f(x^r)}{\ell_i^2}$$

$$a_{ij} = \frac{f(x^r_{ij}) - f(x^r_{i_+}) - f(x^r_{j_+}) + f(x^r)}{\ell_i \ell_j}$$

and,

$$b_i \to f'_{x_i} \qquad a_{ii} \to f''_{x_i} \qquad a_{ij} \to f''_{x_i x_j}$$

Where,

$$x^r = (x^r_1, x^r_2, ..., x^r_n),$$

$$x^r_{i_+} = (x^r_1, x^r_2, ..., x^r_{i-1}, x^r_i + \ell_i, x^r_{i+1}, ..., x^r_n), \quad i = 1, 2, ..., n,$$

$$x^r_{i_-} = (x^r_1, x^r_2, ..., x^r_{i-1}, x^r_i - \ell_i, x^r_{i+1}, ..., x^r_n), \quad i = 1, 2, ..., n,$$

$$x^r_{ij} = (x^r_1, x^r_2, ..., x^r_{i-1}, x^r_i + \ell_i, x^r_{i+1}, ..., x^r_j + \ell_j, x^r_{j+1}, ..., x^r_n), \quad i = 1, 2, ..., n-1, \quad j = i+1, ..., n.$$

Step 4. If $\|[b_n(x^r)]\| \le \varepsilon$, then stop as $x^* = x^r$ is a minimum point. Otherwise go to the next step.

Step 5. Extract the symmetric positive definite matrix $[A_p]$ from the symmetric matrix $[A_n]$ using Choleski's method, where:

$$[A_p] = [L_p][L_p]^T$$

and

$$[\Lambda] = -[A_p]^{-1}[b_p]$$

which is equivalent to:

$$[A_p][\Lambda] = -[b_p].$$

Step 6. Solve the system of the equations:

$$[L_p][L_p]^T[\Lambda] = [b_p], \qquad \Lambda \in \mathbb{R}_p$$

which is equivalent to the solution of the two systems of equations

$$[L_p][Y] = [b_p]$$

$$[L_p]^T[\Lambda] = [Y].$$

Step 7. Determine the search direction at the current point x^r by

$$d^r = \begin{bmatrix} [\Lambda] \\ -c[g_{n-p}(x^r)] \end{bmatrix}$$

where,

$$c = \begin{cases} 1 & \text{if } r = 0, \\ \|[b_n(x^r)]\|^2 / \|[A_n(x^r)][q_n(x^r)]\|^2, \qquad [q_n(x^r)] = [b_n(x^r)] - [b_n(x^{r-1})] & \text{if } r = 1, 2, \dots \end{cases}$$

Step 8. Find the optimal step length α^r in the direction d^r. A one dimensional search is used to determined α^r.

Step 9. Set $x^{r+1} = x^r + \alpha^r d^r$. Set $r = r + 1$ and go to Step 3.

3. Numerical Results

The main aim of this section is to report the performance of Algorithm 2.2 on a set of test problems. The codes were written in Fortran77 and in double precision arithmetic. All the tests were performed on a PC by using the exact line search. Our experiments are performed on a set of 26 unconstrained problems. We test our method with the result given in [21].

For [21] we give the numerical experiments of MP.Q.I against the original BFGS formula with Amijo line search and MBFGS proposed by Liying Liu, Zengxin Wei and Xiaoping Wu. The problems that we tested are from [24]. We will stop the program if the inequality $\|b(x^r)\| < 10^{-4}$ is satisfied. The computation results are given in Table 1.

where the columns have the following meanings:

Problem: the name of the test problem in [24];

Dim: the dimension of the test problem;

NF: the total number of the function evaluations;

NI: the total number of iterations.

From Table 1, we can see that the modified P.Q.I. method in the paper works well. On one hand we see that the total number of iterations in MP.Q.I. method is less than that of both methods for almost all the problems. We emphasis here that the optimum function values which obtained from our method is a very good approximation to the exact solution. On the other hand, for all problems MP.Q.I. reaches the solution point but BFGS fails in four problems and the MBFGS fails in six problems. Finally for the total number of the function evaluations MP.Q.I. for a number of problems was less than that of BFGS or MBFGS methods. Therefor the modified P.Q.I. is superior to BFGS and MBFGS.

In summary, the presented numerical results reveal that Algorithm 2.2, compared with BFGS, MBFGS methods, has many advantages for these problems.

TABLE 1. Test results for Algorithm 2.2/MBFGS/BFGS.

Problems	Dim	MP.Q.I NI/NF	MBFGS NI/NF	BFGS NI/NF
Rosenbrock	2	14/138	1000/1998	33/53
Freudenstein and Roth	2	5/59	11/15	9/21
Brown badly scaled	2	5/108	1000/2001	Fail
Jenrich and Sampson	2	6/73	Fail	11/23
Helical valley	3	7/80	38/43	26/53
Bard	3	6/56	36/57	30/93
Gaussian	3	1/9	3/5	2/5
Gulf research and development	3	15/167	1/2	1/2
Powell singular	4	8/82	29/34	21/44
Kowalik and Osborne	4	5/53	27/29	27/30
Brown and Dennis	4	8/148	Fail	Fail
Biggs EXP6	6	9/94	33/36	1000/3276
Osborne 2	11	12/117	Fail	55/81
Watson	20	22/274	46/52	41/77
Extended Powell singular	4	8/82	29/34	21/44
Penalty I	2	2/28	10/12	6/11
Penalty II	4	2/31	1000/1999	13/23
	50	25/238	Fail	193/913
Variably Dimensioned	2	2/26	9/11	5/13
	50	42/393	60/66	31/116
Trigonometric	3	5/60	26/30	13/21
	50	16/138	30/31	31/35
	100	19/173	87/93	37/40
Discrete boundary value	3	2/22	17/26	6/13
	10	2/22	20/29	16/36
Discrete integral equation	3	2/22	7/10	7/10
	50	2/22	7/10	8/10
	100	2/22	7/10	8/10
	200	2/20	8/11	9/11
	500	3/32	8/11	9/11
Broyden tridiagonal	3	4/48	46/78	114/29
Broyden banded	50	5/64	Fail	Fail
	100	4/52	Fail	Fail
Linear - full rank	2	1/9	2/4	1/3
	50	1/9	2/4	1/3
	500	2/30	2/4	1/3
	1000	2/29	2/4	1/3
Linear - rank 1	2	1/9	2/4	2/10
	10	1/9	2/4999	2/21
Linear - rank 1 with zero columns and rows	4	1/9	2/4	2/11

4. Conclusions

In this paper, we give a modified method with the exact line search for unconstrained optimization problems. The modified method also works with the inexact line search. The direction d^r is constantly descent direction at any point. The modified method converges to a stationary point Regardless of the initial point. The given table shows that the modified method is faster than the others, which requires less iterations and function evaluations.

Finally, this method is a powerful tool for complex systems or when the objective function is given in an implicit form, which mean that direct calculations of the gradient and Hessian matrix are difficult or even impossible to calculate for this function.

For further research, we should study the numerical experiments for large practical problems in the future.

References

[1] M. Al-Baali, Descent property and global convergence of the Fletcher−Reeves method with inexact line search, *IMA J. Numer. Anal.*, 5, 121–124, (1985).
[2] N. Andrei, 40 conjugate gradient algorithms for unconstrained optimization. A survey on their definition, *ICI Technical Report*, No. 13/08, March 14 (2008).
[3] J.S. Arora, *Introduction to optimum design*, McGraw-Hill, Inc., 1989.
[4] J. Barzilai and J.M. Borwein, Two point step size gradient methods, *IMA Journal of Numerical Analysis*, 8, 141–148 (1988).
[5] E.G. Birgin and Y.G. Evtushenko, Automatic differentiation and spectral projected gradient methods for optimal control problems, *Optimization Methods and Software*, 10, 125–146 (1998).
[6] E.G. Birgin, I. Chambouleyron and J.M. Martínez, Estimation of the optical constants and the thickness of thin films using unconstrained optimization, *Journal of Computational Physics*, 151, 862–880 (1999).
[7] E.G. Birgin, J.M. Martínez and M. Raydan, Nonmonotone Spectral Projected Gradient Methods on Convex Sets, *SIAM Journal on Optimization*, 10, 1196–1211 (2000).
[8] Y.H. Dai and Y.X. Yuan, Convergence properties of the Fletcher−Reeves method, *IMA J. Numer. Anal.*, 16(2), 155–164 (1996).
[9] Y.H. Dai, J.Y. Han, G.H. Liu, D.F. Sun, H.X. Yin and Y.X. Yuan, Convergence properties of nonlinear conjugate gradient methods, *SIAM J. Optim.*, 10(2), 345–358 (2000).
[10] R. Dembo and T. Steihaug, Truncated-Newton algorithms for large−scale unconstrained optimization, *Mathematical Programming*, 26, 190–212 (1983).
[11] T.M. El-Gindy and P. Townsend, A numerical method for the determination of optimal surface temperatures of steel ingots, *International Journal for Numerical Methods in Engineering*, 14, 227−233 (1979).
[12] A. Friedlander, J.M. Martínez and M. Raydan, A new method for large-scale box constrained convex quadratic minimization problems, *Optimization Methods and Software*, 5, 57–74 (1995).
[13] J.C. Gilbert and J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, *SIAM J. Optim.*, 2, 21−42 (1992).
[14] A.A. Goldstein, Convex programming in Hilbert space, *Bulletin of the American Mathematical Society*, 70, 709−710 (1964).
[15] L. Grippo and S. Lucidi, A globally convergent version of the Polak−Ribière conjugate gradient method, *Math. Program.*, 78, 375–391 (1997).
[16] L. Guanghui, H. Jiye and Y. Hongxia, Global convergence of the Fletcher−Reeves algorithm with inexact linesearch, *Appl. Math. J. Chinese Univ.*, 10, 75–82 (1995).
[17] W.W. Hager and H. Zhang, A new conjugate gradient method with guaranteed descent and an efficient line search, *SIAM J. Optim.*, 16, 170–192 (2005).
[18] W.W. Hager, H. Zhang, A survey of nonlinear conjugate gradient methods, *Pacific J. Optim.*, 2, 35–58 (2006).

[19] E.S. Levitin and B.T. Polyak, Constrained minimization problems, *USSR Computational Mathematics and Mathematical Physics*, 6, 1–50 (1966).
[20] D. Liu and J. Nocedal, On the limited memory BFGS method for large−scale optimization, *Math. Program.*, 45, 503–528, (1989).
[21] L. Liu, Z. Wei and X. Wu, The convergance of a new modified BFGS method without line searchs for unconstrained optimization or complexity systems, *Journal of Systems Science and Complexity*, 23, 861–872 (2010).
[22] M. Mamat, M. Rivaie, I. Mohd and M. Fauzi, A New Conjugate Gradient Coefficient for Unconstrained Optimization, *Int. J. Contemp. Math. Sciences*, 5(29), 1429−1437 (2010).
[23] A. Matthews and D. Davies, A Comparision of modified Newton's methods for Unconstrained Optimization, *Computer Journal*, 14, 293–294 (1971).
[24] J.J. Moré, B. S. Garbow, K. E. Hillstrome, Testing unconstrained optimization software, *ACM Trans. Math. Softw.*, 7, 17–41, (1981).
[25] S.G. Nash, A survey of truncated−Newton methods, *J. Comput. Appl. Math.*, 124, 45–59 (2000).
[26] J. Nocedal., Updating quasi-Newton matrices with limited storage, *Mathematics of Computation*, 35(151), 773–782 (1980).
[27] J. Nocedal and S.J. Wright, *Numerical Optimization*, Springer Series in Operations Research, Springer, New York, NY, USA, 1999.
[28] M.J.D. Powell, Nonconvex minimization calculations and the conjugate gradient method, Lecture Notes in Mathematics, Springer, Berlin, 1984, 1066: 122−141.
[29] M. Raydan, On the Barzilai and Borwein choice of steplength for the gradient method, *IMA Journal of Numerical Analysis*, 13, 321–326 (1993).
[30] M. Raydan, The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem, *SIAM Journal on Optimization*, 7, 26−33 (1997).
[31] Z.J. Shi and J. Shen, Convergence of the Polak−Ribière−Polyak conjugate gradient method, *Nonlinear Anal.*, 66, 1428–1441 (2007).
[32] J. Stoer and Y. Yuan, A subspace study on conjugate gradient algorithms, *ZAMM Z. angew. Math. Mech.*, 75, 69–77, (1995).
[33] Y. Xiao, Z. Wei and Z. Wang, A limited memory BFGS-type method for large-scale unconstrained optimization, *Computers & Mathematics with Applications*, 56(4), 1001–1009 (2008).
[34] N.H. Xiu, C.Y. Wang and L.C. Kong, A note on the gradient projection method with exact stepsize rule, *Journal of Computational Mathematics*,25(2), 221–230 (2007).
[35] Y. Yuan, Subspace techniques for nonlinear optimization, talk presented at the International Conference on High Performance Scientific Computing, March, 6–10, (2006).
[36] L. Zhang, W. Zhou and D.H. Li, A descent modified Polak−Ribière−Polyak conjugate gradient method and its global convergence, *IMA J. Numer. Anal.*, 26, 629–640 (2006).
[37] L. Zhang, W. Zhou and D.H. Li, Global convergence of a modified Fletcher−Reeves conjugate gradient method with Armijo−type line search, *Numer. Math.*, 104, 561–572 (2006).

(T.M. El-Gindy) Department of Mathematics Faculty of Science, Assiut University, Assiut, Egypt
E-mail address: Taha.elgindy@gmail.com

(M.S. Salim) Department of Mathematics Faculty of Science, Al-Azhar University, Assiut, Egypt
E-mail address: m_s_salim@yahoo.com

(Abdel–Rahman Ibrahim) Department of Mathematics Faculty of Science, Al-Azhar University, Assiut, Egypt
E-mail address: Abl_ibrahim@yahoo.com

TABLE OF CONTENTS, JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.1, 2013

VOLUME 11, NUMBER 2 APRIL 2013

ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL

OF CONCRETE

AND APPLICABLE MATHEMATICS

EUDOXUS PRESS,LLC

GUEST EDITORS: O. DUMAN, E. ERKUS-DUMAN

SPECIAL ISSUE II: "APPLIED MATHEMATICS -APPROXIMATION THEORY 2012"

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Assistant to the Editor:Dr.Razvan Mezei,Lander University,SC 29649, USA.

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send via email the contribution to the editor in-Chief: TEX or LATEX (typed double spaced) and PDF files. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $500,Electronic $250,Print and Electronic $600.Individual:Print $200, Electronic $100,Print &Electronic $250.For any other part of the world add $60 more to the above prices for Print.
Single article PDF file for individual $20.Single issue in PDF form for individual $80.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

PAGE CHARGES: Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus homepage.

Editorial Board
Associate Editors of Journal of Concrete and Applicable Mathematics

Department of Mathematics Education
College of Education
Gyeongsang National University
Chinju 660-701
KOREA
tel.055-751-5673 Office,
055-755-3644 home,
fax 055-751-6117
yjcho@nongae.gsnu.ac.kr
Nonlinear operator Th.,Inequalities,
Geometry of Banach Spaces

5) Sever S.Dragomir
School of Communications and Informatics
Victoria University of Technology
PO Box 14428
Melbourne City M.C
Victoria 8001,Australia
tel 61 3 9688 4437,fax 61 3 9688 4050
sever.dragomir@vu.edu.au,
sever@sci.vu.edu.au
Math.Analysis,Inequalities,Approximation Th.,
Numerical Analysis, Geometry of Banach Spaces,
Information Th. and Coding

6) Oktay Duman
TOBB University of Economics and Technology,
Department of Mathematics, TR-06530,
Ankara, Turkey, oduman@etu.edu.tr
Classical Approximation Theory,
Summability Theory,
Statistical Convergence and its Applications

7) Angelo Favini
Università di Bologna
Dipartimento di Matematica
Piazza di Porta San Donato 5
40126 Bologna, ITALY
tel.++39 051 2094451
fax.++39 051 2094490
favini@dm.unibo.it
Partial Differential Equations, Control Theory,
Differential Equations in Banach Spaces

8) Claudio A. Fernandez
Facultad de Matematicas
Pontificia Unversidad Católica de Chile
Vicuna Mackenna 4860
Santiago, Chile
tel.++56 2 354 5922

25) Svetlozar T.Rachev
Dept.of Statistics and Applied Probability Program
University of California,Santa Barbara
CA 93106-3110,USA
tel.805-893-4869
rachev@pstat.ucsb.edu
AND
Chair of Econometrics and Statistics
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss,Bau II,20.12,R210
Postfach 6980,D-76128,Karlsruhe,Germany
tel.011-49-721-608-7535
rachev@lsoe.uni-karlsruhe.de
Mathematical and Empirical Finance,
Applied Probability, Statistics and Econometrics

26) John Michael Rassias
University of Athens
Pedagogical Department
Section of Mathematics and Infomatics
20, Hippocratous Str., Athens, 106 80, Greece

Address for Correspondence
4, Agamemnonos Str.
Aghia Paraskevi, Athens, Attikis 15342 Greece
jrassias@primedu.uoa.gr
jrassias@tellas.gr
Approximation Theory,Functional Equations,
Inequalities, PDE

27) Paolo Emilio Ricci
Universita' degli Studi di Roma "La Sapienza"
Dipartimento di Matematica-Istituto "G.Castelnuovo"
P.le A.Moro,2-00185 Roma,ITALY
tel.++39 0649913201,fax ++39 0644701007
riccip@uniroma1.it,Paoloemilio.Ricci@uniroma1.it
Orthogonal Polynomials and Special functions,
Numerical Analysis, Transforms,Operational Calculus,
Differential and Difference equations

28) Cecil C.Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152,USA
tel.901-678-2490,fax 901-678-2480
ccrousse@memphis.edu
Combinatorics,Graph Th.,
Asymptotic Approximations,
Applications to Physics

29) Tomasz Rychlik

Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop
Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Robert Kozma
Dept. of Mathematical Sciences
University of Memphis
Memphis, TN 38152, USA
rkozma@memphis.edu
Mathematical Learning Theory,
Dynamic Systems and Chaos,
Complex Dynamics.

18) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

19) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

20) Rupert Lasser
Institut fur Biomathematik &
Biomertie,GSF
-National Research Center for
environment and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier
Analysis,Mathematical Biology.

Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

35) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

36) Manfred Tasche
Department of Mathematics
University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

37) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

38) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou

Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts files in Latex and PDF and in English, should be submitted via email to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel. 901.678.3144
e-mail: ganastss@memphis.edu

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.

Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.

2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (Click HERE to save a copy of the style file.)They should be carefully prepared in all respects. Submitted articles should be brightly typed (not dot-matrix), double spaced, in ten point type size and in 8(1/2)x11 inch area per page. Manuscripts should have generous margins on all sides and should not exceed 24 pages.

3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.

4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.

The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.

6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes.

The main body of the paper should begin on page no. 1, if possible.

7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.

8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right, and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.

9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.

Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.

10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,

name of publication, volume number, inclusive pages, and year of publication.

Authors should follow these examples:

<u>Journal Article</u>

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

<u>Book</u>

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

<u>Contribution to a Book</u>

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.

12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set off from the text. Use the appropriate superscript numeral for citation in the text.

13. After each revision is made please again submit via email Latex and PDF files of the revised manuscript, including the final one.

14. Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus homepage.

No galleys will be sent and the contact author will receive one (1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

PREFACE (JAFA – JCAAM)

These special issues are devoted to a part of proceedings of AMAT 2012 - International Conference on Applied Mathematics and Approximation Theory - which was held during May 17-20, 2012 in Ankara, Turkey, at TOBB University of Economics and Technology. This conference is dedicated to the distinguished mathematician George A. Anastassiou for his 60th birthday.

AMAT 2012 conference brought together researchers from all areas of Applied Mathematics and Approximation Theory, such as ODEs, PDEs, Difference Equations, Applied Analysis, Computational Analysis, Signal Theory, and included traditional subfields of Approximation Theory as well as under focused areas such as Positive Operators, Statistical Approximation, and Fuzzy Approximation. Other topics were also included in this conference, such as Fractional Analysis, Semigroups, Inequalities, Special Functions, and Summability. Previous conferences which had a similar approach to such diverse inclusiveness were held at the University of Memphis (1991, 1997, 2008), UC Santa Barbara (1993), the University of Central Florida at Orlando (2002).

Around 200 scientists coming from 30 different countries participated in the conference. There were 110 presentations with 3 parallel sessions. We are particularly indebted to our plenary speakers: George A. Anastassiou (*University of Memphis* - *USA*), Dumitru Baleanu (*Çankaya University* - *Turkey*), Martin Bohner (*Missouri University of Science & Technology* - *USA*), Jerry L. Bona (*University of Illinois at Chicago* - *USA*), Weimin Han (*University of Iowa* - *USA*), Margareta Heilmann (*University of Wuppertal* - *Germany*), Cihan Orhan (*Ankara University* - *Turkey*). It is our great pleasure to thank all the organizations that contributed to the conference, the Scientific Committee and any people who made this conference a big success.

Finally, we are grateful to “TOBB University of Economics and Technology”, which was hosting this conference and provided all of its facilities, and also to “Central Bank of Turkey” and “The Scientific and Technological Research Council of Turkey” for financial support.

Guest Editors:

Oktay Duman
TOBB Univ. of Economics and Technology
Ankara, Turkey, 2012

Esra Erkuş-Duman
Gazi University
Ankara, Turkey, 2012

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 160-167, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

SOME EXTENSIONS OF SUFFICIENT CONDITIONS FOR UNIVALENCE OF AN INTEGRAL OPERATOR

NAGAT.M.MUSTAFA AND MASLINA DARUS

ABSTRACT. In this paper, we introduce and study a general integral operator defined on the class of normalized analytic function in unit disc. This operator is motivated by many researchers. With this operator univalence conditions for the normalized analytic function are obtained. We also present a few conditions of univalence for the integral operator.

1. INTRODUCTION

Let A denote the class of functions f in the open unit disc

$$U = \{z \in \mathbb{C} : |z| < 1\},$$

given by the normalized power series

$$f(z) = z + \sum_{k=2}^{\infty} a_k z^k \quad (z \in U) \quad , \tag{1.1}$$

where a_k is a complex number.

Let S be the subclass of A consisting of univalent functions. For functions f given by (1.1) and

$$g(z) = z + \sum_{k=2}^{\infty} b_k z^k , \quad (z \in U).$$

Let $(f * g)$ denote the Hadamard product (convolution) of f and g, defined by:

$$(f * g)(z) = z + \sum_{k=2}^{\infty} a_k b_k z^k.$$

Many authors studied the problem of integral operators, acting on functions in S. In this sense, the following result due to Ozaki and Nunokawa [11] is useful to study the univalence of integral operator for certain subclass of S.

Theorem 1.1. *Let $f \in A$ satisfy the following inequality:*

$$\left| \frac{z^2 f'(z)}{f^2(z)} - 1 \right| \leq 1, \quad for\ all\ (z \in U),$$

then the function f is univalent in U.

Key words and phrases. Analytic functions, Univalent functions, Integral operator,Schwarz Lemma.

2010 *AMS Math. Subject Classification.* Primary 40A05, 40A25; Secondary 45G05.

1

For some real p with $0 < p \leq 2$, we define a subclass $S(p)$ of A consisting of all functions f which satisfy

$$\left|\left(\frac{z}{f(z)}\right)''\right| \leq p, \quad for\ all\ (z \in U).$$

In [15], Singh has shown that if $f \in S(p)$, then f satisfies

$$\left|\frac{z^2 f'(z)}{f^2(z)} - 1\right| \leq p|z|, \quad for\ all\ (z \in U).$$

For two functions $f \in A$ and $g(z) = z + \sum_{k=2}^{\infty} b_k z^k, \quad (z \in U)$. Let the function $\varphi(a, c; z)$ be given by

$$\varphi(a, c; z) = \sum_{k=0}^{\infty} \frac{(a)_k}{(c)_k} z^{k+1}, \quad (z \in U, c \neq 0, -1, -2, -3, ...), \tag{1.2}$$

where $(x)_k$ denotes the Pochhammer symbol (or the shifted factorial) defined by:

$$(x)_k = \begin{cases} 1 & for \quad k = 0, \\ x(x+1)(x+2)...(x+k-1) & for \quad k \in \mathbb{N} = \{1, 2, 3, \cdots\}. \end{cases}$$

Note that $\varphi(a, 1; z) = z/(1-z)^a$ and $\varphi(2, 1; z)$, is the Koebe function.

For a function $f \in A$ we introduced the following differential operator .

$$\begin{aligned} D^0(\lambda, \alpha) f(z) &= f(z), \\ D^1(\lambda, \alpha) f(z) &= (1 - \frac{\lambda}{\lambda + \alpha})(f(z)) + \frac{\lambda}{\lambda + \alpha} z(f(z))', \\ D^2(\lambda, \alpha) f(z) &= D(D^1(\lambda, \alpha) f(z)), \\ &\vdots \\ D^m(\lambda, \alpha) &= D(D^{m-1}(\lambda, \alpha) f(z)). \end{aligned} \tag{1.3}$$

If $f \in A$ then by using the Hadamard product (or convolution) (1.4) and (1.2) we have

$$D^m(\lambda, \alpha, a, c) = D^m(\lambda, \alpha) * \varphi(a, c; z),$$

then

$$D^m(\lambda, \alpha, a, c) = z + \sum_{k=2}^{\infty} \left(\frac{\lambda(k-1) + \alpha + \lambda}{\lambda + \alpha}\right)^m \frac{(a)_{k-1}}{(c)_{k-1}} a_k z^k, \tag{1.4}$$

where $\alpha, \lambda \geq 0$, $\alpha + \lambda \neq 0$, $m \in \mathbb{N}$.

By specializing the parameters $m, \alpha, \lambda, a \quad and \quad c$, one can obtain various operators, which are special cases of $D^m(\lambda, \alpha, a, c)$ studied earlier by many authors listed as follows:

- $D^0(\lambda, \alpha, a, c) \equiv L(a, c)$, due to Carlson and Shaffer operator [1].
- $D^0(\lambda, \alpha, \beta + 1, 1)(\beta > -1) \equiv R^n$ due to Ruscheweyh derivatives operator [13].
- $D^m(1, 0, 0, 0) \equiv S^n$ due to Salagean derivatives operator [6].

• $D^m(\lambda, 1-\lambda, 0, 0) \equiv S^n_\beta$ due to the generalized Salagean derivative operator introduced by Al-Oboudi [5].

• $D^m(\lambda, 1-\lambda, \beta+1, 1) \equiv D^n_{\lambda,\beta}$ due to the generalized Al-Shaqsi and Darus derivative operator [7].

• $D^m(1, \alpha, 0, 0)$ due to the multiplier transformations studied by Flett [12].

• $D^{-1}(1, \alpha, 0, 0)$ due to the integral operator studied by Owa and Srivastava [10].

Here we introduce a new family of integral operator by using generalized differential operator already defined above.

For $f_i \in A$, $i = \{1, 2, 3, \cdots, n\}$, $n \in \mathbb{N} \cup \{0\}$ and $\gamma_1, \gamma_2, \gamma_3, \cdots, \gamma_n, \rho \in \mathbb{C}$, we define a family of integral operator $\Psi^m(\lambda, \alpha, a, c; z) : A^n \to A^n$ by

$$\Psi^m(\lambda, \alpha, a, c; z) = \left\{\rho \int_0^z t^{\rho-1} \prod_{i=1}^n \left(\frac{D^m(\lambda, \alpha, a, c) f_i(t)}{t}\right)^{\frac{1}{\gamma_i}} dt\right\}^{\frac{1}{\rho}},$$

where $\alpha, \lambda \geq 0$, $\alpha + \lambda \neq 0$, $m \in \mathbb{N}$. and $D^m(\lambda, \alpha, a, c)f(z)$ defined by(1.4) which generalizes certain integral operators as follow:

(1) $a = c = 0$, $\alpha = -\lambda + 1$, $\gamma_i = \frac{1}{\alpha_i}$, $\rho = 1$, we obtain $I(f_1, ...f_m)$ Bulut[9].
(2) $m = 0$, $a = c = 0$, $\gamma_i = \frac{1}{\alpha-1}$, $\rho = n(\alpha-1)+1$, we obtain $F_{n,\alpha}(z)$ Breaz [2].
(3) $m = 0$, $a = c = 0$, $\gamma_i = \frac{1}{\alpha_i}$, $\rho = 1$, we obtain $F_\alpha(z)$ Breaz and Breaz [3] .

To discuss our problems, we have to recall here the following results.

Lemma 1.2. [16]*Let the function f be regular in the disk*

$$U_R = \{z \in \mathbb{C} : |z| < R\},$$

with $|f(z)| < M$ for fixed M. If f has one zero with multiplicity order bigger than m for $z = 0$, then

$$|f(z)| \leq \frac{M}{R^m}|z|^m, (z \in U_R).$$

The equality can hold only if

$$f(z) = e^{i\theta}(\frac{M}{R^m})z^m,$$

where θ is constant.

Lemma 1.3. [8] *Let $f \in A$, and β be a complex number with $\Re(\beta) > 0$. If f satisfies*

$$\frac{1-|z|^{2\Re(\beta)}}{\Re(\beta)}\left|\frac{zf''(z)}{f'(z)}\right| \leq 1,$$

then for all $z \in U$ the integral operator

$$F_\beta(z) = \left\{\alpha \int_0^z t^{\beta-1} f'(t) dt\right\}^{\frac{1}{\beta}},$$

is in the class S.

Lemma 1.4. [14] *Let $f \in A$ and $\beta, c \in \mathbb{C}$ where $\Re(\beta) > 0$ and ($|c| \leq 1, \quad c \neq -1$). If*

$$\left| c|z|^{2\beta} + (1 - |z|^{2\beta}) \frac{zf''(z)}{\beta f'(z)} \right| \leq 1,$$

for all $z \in U$, then the function

$$F_\beta(z) = \left[\beta \int_0^z t^{\beta-1} f'(t) dt \right]^{\frac{1}{\beta}},$$

is analytic and univalent in U.

Now, we shall use the same method given by Breaz and Özlem Güney (2008) we prove the following results.

2. Main results

Theorem 2.1. *Let $M \geq 1$, let a family $D^m(\lambda, \alpha, a, c) f_i(t) \in S(p), i = \{1, ...n\}$, $|D^m(\lambda, \alpha, a, c) f_i(z)| \leq M$ such that $|\frac{z^2 (D^m(\lambda,\alpha,a,c)) f_i(z))'}{(D^m(\lambda,\alpha,a,c)) f_i(z))^2} - 1| \leq p|z|$, $\forall z \in U$. If*

$$\Re(\rho) \quad \geq \quad \sum_{i=1}^{n} \frac{(1-p)M + 1}{|\gamma_i|}, \ \rho, \gamma_i \in \mathbb{C},$$

$$\text{and} \quad |c| \quad \leq \quad 1 - \frac{1}{\Re(\rho)} \sum_{i=1}^{n} \frac{1}{|\gamma_i|} (p+1)M + 1, \tag{2.1}$$

then the family $\Psi^m(a, c, \alpha, \lambda, \gamma_i; z)$ belong to S.

Proof. Since

$$\frac{D^m(\lambda, \alpha, a, c) f(z)}{z} \quad = \quad 1 + \sum_{k=2}^{\infty} \left(\frac{\lambda(k-1) + \alpha + \lambda}{\lambda + \alpha} \right)^m \frac{(a)_{k-1}}{(c)_{k-1}} a_k z^{k-1} \neq 0,$$

$$= 1 \text{ if } z = 0.$$

Let

$$F(z) \quad = \quad \int_0^z \left((\frac{D^m(\lambda, \alpha, a, c) f_1(t)}{t})^{\frac{1}{\gamma_1}} \cdots (\frac{D^m(\lambda, \alpha, a, c) f_n(t)}{t})^{\frac{1}{\gamma_n}} \right) dt,$$

$$\Rightarrow F'(z) \quad = \quad (\frac{D^m(\lambda, \alpha, a, c) f_1(z)}{z})^{\frac{1}{\gamma_1}} \cdots (\frac{D^m(\lambda, \alpha, a, c) f_n(z)}{z})^{\frac{1}{\gamma_n}},$$

$$\text{therefore} \quad \frac{zF''(z)}{F'(z)} \quad = \quad \sum_{i=1}^{n} \frac{1}{\gamma_i} \left(\frac{z(D^m(\lambda, \alpha, a, c) f_i(z))'}{D^m(\lambda, \alpha, a, c) f_i(z)} - 1 \right). \tag{2.2}$$

From (2.2), we have

$$\left|\frac{zF''(z)}{F'(z)}\right| \leq \sum_{i=1}^{n}\frac{1}{|\gamma_i|}\left(\frac{z(D^m(\lambda,\alpha,a,c)f_i(z))'}{D^m(\lambda,\alpha,a,c)f_i(z)}+1\right)$$
$$= \sum_{i=1}^{n}|\frac{1}{\gamma_i}|\left(\left|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}\right|\frac{|(D^m(\lambda,\alpha,a,c)f_i(z))|}{|z|}+1\right). \tag{2.3}$$

From the hypothesis , we have $|D^m(\lambda,\alpha,a,c)f_i(z)| \leq M,\ z\in U, i=\{1,...,n\}$, then by lemma 1.2, we obtain that:

$$|D^m(\lambda,\alpha,a,c)f_i(z)| \leq M|z|.$$

We apply this result in inequality (2.3) then we obtain:

$$\begin{aligned}\left|\frac{zF''(z)}{F'(z)}\right| &\leq \sum_{i=1}^{n}|\frac{1}{\gamma_i}|\left(\left|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}\right|M+1\right)\\ &\leq \sum_{i=1}^{n}|\frac{1}{\gamma_i}|\left(|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}-1|M+M+1\right)\\ &= \sum_{i=1}^{n}|\frac{1}{\gamma_i}|\,(pM|z|+M+1)\\ &< \sum_{i=1}^{n}|\frac{1}{\gamma_i}|\,((p+1)M+1)\,.\end{aligned}$$

We have

$$\begin{aligned}\left|c|z|^{2\rho}+(1-|z|^{2\rho})\frac{1}{\Re\rho}\frac{zF''(z)}{F'(z)}\right| &= \left|c|z|^{2\rho}+(1-|z|^{2\rho})\frac{1}{\Re\rho}\sum_{i=1}^{n}\frac{1}{\gamma_i}(\frac{z(D^m(\lambda,\alpha,a,c)f_i(z))'}{D^m(\lambda,\alpha,a,c)f_i(z)}-1)\right|\\ &\leq |c|+\frac{1}{|\rho|}\sum_{i=1}^{n}|\frac{1}{\gamma_i}|(\left|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}\right|\\ &\quad \frac{|(D^m(\lambda,\alpha,a,c)f_i(z))|}{|z|}+1).\end{aligned}$$

We obtain

$$\left||z|^{2\rho}+(1-|z|^{2\rho})\frac{zF'(z)}{\Re\rho F(z)}\right| \leq |c|+\frac{1}{\Re(\rho)}\sum_{i=1}^{n}|\frac{1}{\gamma_i}|\,((p+1)M+1)\,.$$

So, from (2.1) we have

$$\left||z|^{2\rho}+(1-|z|^{2\rho})\frac{zF'(z)}{\Re\rho F(z)}\right| \leq 1.$$

Hence we complete the proof. □

Remark 2.1. *If we set $a=c=m=0$ in Theorem 2.1, then we have result in* [4].

Corollary 2.2. *Let $M \geq 1$, let a family $D^m(\lambda, \alpha, a, c)f_i(t) \in S(p), i = \{1, ..., n\}$, $|D^m(\lambda, \alpha, a, c)f_i(z)| \leq M$, such that $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2} - 1| \leq p|z|, \ \forall z \in U$. If*

$$\begin{aligned} \Re(\rho) &\geq \frac{n(1-p)M+1}{|\gamma|}, \ \rho, \gamma \in \mathbb{C}, \\ \text{and} \quad |c| &\leq 1 - \frac{1}{\Re(\rho)}\left(\frac{(n(p+1)M+1)}{|\gamma|}\right), \end{aligned}$$

then the family $\Psi^m(a, c, \alpha, \lambda, \gamma_i; z)$ belong to S.

Proof. In Theorem 2.4, we consider $\gamma_1 = \gamma_2 = ...\gamma_n = \gamma$. □

Corollary 2.3. *Let the family $D^m(\lambda, \alpha, a, c)f_i(t) \in S(p), i = \{1, ..., n\}$, $|D^m(\lambda, \alpha, a, c)f_i(z)| \leq 1$ such that $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2} - 1| \leq p|z|, \ \forall z \in U$. If*

$$\begin{aligned} \Re(\rho) &> \sum_{i=1}^{n} \frac{p+2}{|\gamma_i|}, \ \rho, \gamma \in \mathbb{C}, \\ \text{and} \quad |c| &\leq 1 - \frac{1}{\Re(\rho)} \sum_{i=1}^{n} \frac{1}{|\gamma_i|}(p+2), \end{aligned}$$

then the family $\Psi^m(a, c, \alpha, \lambda, \gamma_i; z)$ belong to S.

Proof. In Theorem 2.1, we consider $M = 1$. □

Theorem 2.4. *Let a family $D^m(\lambda, \alpha, a, c)f_i(t) \in A$, $|D^m(\lambda, \alpha, a, c)f_i(z)| \leq M, \forall i$ such that $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2} - 1| \leq 1, \ \forall z \in U$. If*

$$\begin{aligned} \Re(\rho) &\geq \sum_{i=1}^{m} \frac{(2M+1)}{|\gamma_i|}, \ M \geq 1, \ \rho, \gamma_i \in \mathbb{C}, \\ \text{and} \quad |c| &\leq 1 - \frac{1}{\Re(\rho)} \sum_{i=1}^{m} \frac{(2M+1)}{|\gamma_i|}, \end{aligned} \tag{2.4}$$

then the family $\Psi^{\alpha,\delta,\rho}(m, q, \lambda, \gamma_i; z)$ belong to S.

Proof. We know from the proof of Theorem 2.1 that

$$\frac{zF''(z)}{F'(z)} = \sum_{i=1}^{n} \frac{1}{\gamma_i}\left(\frac{z(D^m(\lambda, \alpha, a, c)f_i(z))'}{D^m(\lambda, \alpha, a, c)f_i(z)} - 1\right).$$

Because

$$\begin{aligned} &\left| c|z|^{2\rho} + (1-|z|^{2\rho})\frac{1}{\Re\rho}\sum_{i=1}^{n}\frac{1}{\gamma_i}(\frac{z(D^m(\lambda, \alpha, a, c)f_i(z))'}{f_i(z)} - 1)\right| \\ \leq \quad & |c| + \frac{1}{|\rho|}\sum_{i=1}^{n}|\frac{1}{\gamma_i}|\left(|\frac{z^2(D^m(\lambda, \alpha, a, c)f_i(z))'}{(D^m(\lambda, \alpha, a, c)f_i(z))^2}|\frac{|(D^m(\lambda, \alpha, a, c)f_i(z))|}{|z|} + 1\right). \end{aligned}$$

Where

$$|\frac{z^2(D^m(\lambda, \alpha, a, c)f_i(z))'}{(D^m(\lambda, \alpha, a, c)f_i(z))^2}| \leq 2, \text{ and } |(D^m(\lambda, \alpha, a, c)f_i(z))| \leq M|z|,$$

so

$$||z|^{2\rho}+(1-|z|^{2\rho})\frac{zF'(z)}{\Re\rho F(z)}| \quad \leq \quad |c|+\frac{1}{|\rho|}\sum_{i=1}^{n}\frac{1}{|\gamma_i|}(2M+1),$$

using (2.4) we have

$$\left||z|^{2\rho}+(1-|z|^{2\rho})\frac{zF'(z)}{\Re\rho F(z)}\right|\leq 1.$$

Applying Lemma 1.4, we obtain the family $\Psi^{\alpha,\delta,\rho}(m,q,\lambda,\gamma_i;z)$ belong to S. $\square$

Theorem 2.5. *Let $M\geq 1$, let a family $D^m(\lambda,\alpha,a,c)f_i(t)\in S(p), i=\{1,..n\}$, $|D^m(\lambda,\alpha,a,c)f_i(z)|\leq M$ such that $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2}-1|\leq p|z|,\ \forall z\in U$. If*

$$\Re(\beta) \quad \geq \quad \sum_{i=1}^{n}\frac{(1+p)M+1}{|\gamma_i|},\ \rho,\gamma_i\in\mathbb{C},$$

then the family $\Psi^m(a,c,\alpha,\lambda,\gamma_i;z)$ belong to S.

Proof. We know from the proof of Theorem 2.1 that

$$\left|\frac{zF''(z)}{F'(z)}\right|\leq\sum_{i=1}^{n}\frac{1}{\gamma_i}\left|\frac{z(D^m(\lambda,\alpha,a,c))'}{D^m(\lambda,\alpha,a,c)}-1\right|.$$

So, by the imposed conditions, we find

$$\begin{aligned}
\frac{1-|z|^{2\Re(\beta)}}{\Re(\beta)}\left|\frac{zF''(z)}{F'(z)}\right| &\leq \frac{1-|z|^{2\Re(\beta)}}{\Re(\beta)}\sum_{i=1}^{n}\frac{1}{\gamma_i}\left(\left|\frac{z(D^m(\lambda,\alpha,a,c)f_i(z))'}{D^m(\lambda,\alpha,a,c)f_i(z)}\right|+1\right)\\
&\leq \frac{1-|z|^{2\Re(\beta)}}{\Re(\beta)}\sum_{i=1}^{n}\frac{1}{\gamma_i}\left(\left|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}\right|\left|\frac{D^m(\lambda,\alpha,a,c)f_i(z)}{z}\right|+1\right)\\
&\leq \frac{1-|z|^{2\Re(\beta)}}{\Re(\beta)}\sum_{i=1}^{n}\frac{1}{\gamma_i}\left(\left|\frac{z^2(D^m(\lambda,\alpha,a,c)f_i(z))'}{(D^m(\lambda,\alpha,a,c)f_i(z))^2}-1\right|M+M+1\right)\\
&\leq \frac{1}{\Re(\beta)}\sum_{i=1}^{n}\frac{1}{\gamma_i}((1+p)M+1)\\
&\leq 1.
\end{aligned}$$

By applying Lemma 1.3, we prove that $\Psi^{\alpha,\delta,\rho}(m,q,\lambda,\gamma_i;z)\in S$. $\square$

Corollary 2.6. *Let $M\geq 1$, let a family $D^m(\lambda,\alpha,a,c)f_i(t)\in S(p), i=\{1,...,n\}$, $|D^m(\lambda,\alpha,a,c)f_i(z)|\leq M$ such that $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2}-1|\leq p|z|,\ \forall z\in U$. If*

$$\Re(\beta) \quad \geq \quad \frac{n(1+p)M+1}{|\gamma|}, \quad \rho,\gamma_i\in\mathbb{C},$$

then the family $\Psi^m(a,c,\alpha,\lambda,\gamma_i;z)$ belong to S.

Proof. In Theorem 2.5, we consider $\gamma_1=\gamma_2=...\gamma_n=\gamma$. $\square$

Corollary 2.7. *Let a family* $D^m(\lambda,\alpha,a,c)f_i(t) \in S(p), i = \{1,...,n\}$, $|D^m(\lambda,\alpha,a,c)f_i(z)| \leq 1$ *such that* $|\frac{z^2(D^m(\lambda,\alpha,a,c))f_i(z))'}{(D^m(\lambda,\alpha,a,c))f_i(z))^2} - 1| \leq p|z|,\ \forall z \in U$. *If*

$$\Re(\beta) \quad \geq \quad \sum_{i=1}^{n} \frac{2+p}{|\gamma_i|},\ \rho, \gamma_i \in \mathbb{C},$$

then the family $\Psi^m(a,c,\alpha,\lambda,\gamma_i;z)$ *belong to* S.

Proof. In Theorem 2.5, we consider $M = 1$. □

3. Acknowledgements:

The work presented here was partially supported by LRCS/TD/2011/UKM/iCT/03/02 and the Libyan government.

References

[1] B.C. Carlson, D.B. Shaffer, Starlike and prestarlike hypergeometric functions , *SIAM J. Math. Anal.*, 15(4), 737–745 (1984).

[2] D.Breaz, N.Breaz and H.M.Srivastava, An extension of the univalent condition for a family of integral operators, *Appl. Math. Lett.*, 22 (1), 41–44(2009).

[3] D. Breaz and N. Breaz, Two integral operators, *Studia Univ. Babes-Bolyai Math.*, 47(3), 13–19(2002).

[4] D. Breaz , H. Özlem Güney.,On the Univalence Criterion of a General Integral Operator,*J .Ine and Appl.* , (2008), Article ID 702715, 8 pages.

[5] F.M. Al-Oboudi, On univalent functions defined by a generalized Salagean Operator, *Int, J. Math. Math. Sci.* , 27, 1429–1436 (2004).

[6] G.S. Salagean. Subclasses of univalent functions. *In Springer Lecture Notes Series 1013, Springer-Verlag, Berlin*, 362–372(1983).

[7] M. Darus, K. Al-Shaqsi. Differential subordination with generalised derivative operator, *Int.J. Comp. Math. Sci.*, 2(2), 75 - 78(2008).

[8] N. Pascu, On a univalence criterion. II, *Itinerant Seminar on Functional Equations*, Approximation and Convexity ,Cluj-Napoca, (1985), Preprint, 85, Universitatea.

[9] S. Bulut, Sufficient conditions for univalence of an integral operator defined by Al- Oboudi differential operator, *J. Inequal. Appl. , Art.*, ID 957–042(2008).

[10] S. Owa and H. M. Srivastava, Some applications of the generalized Libera integral operator, *Proc. Japan Acad. Ser. A Math. Sci.*,62 (4), 125–128(1986).

[11] S. Ozaki, M. Nunokawa, The Schwarzian derivative and univalent functions, *Proc. Amer. Math. Soc.*, 33, 392-394(1972).

[12] T. M. Flett, The dual of an inequality of Hardy and Littlewood and some related inequalities, *J. Math. Anal. Appl.,* 38,746–765(1972).

[13] St. Ruscheweyh , New criteria for univalent functions, *Proc. Amer. Math.Soc.*, 49, 109-115(1975).

[14] V. Pescar, A new generalization of Ahlforss and Beckers criterion of univalence, *Bull. Malaysian Math. Soc.*, (2) 19 , no. 2, 53–54(1996).

[15] V. Singh, On a class of univalent functions, *Int .J. Math and Math Sci.*, 23(12) 855–857(2000).

[16] Z. Nehari, Conformal Mapping, *McGraw-Hill Book Comp.*, New York, (1952) (Dover. Publ. Inc., 1975).

(Nagat.M.Mustafa) Universiti Kebangsaan Malaysia, Bangi 43600 Selangor D. Ehsan, Malaysia

E-mail address: nmma_1975@yahoo.com

(Maslina Darus) Universiti Kebangsaan Malaysia, Bangi 43600 Selangor D. Ehsan, Malaysia

E-mail address: maslina@ukm.my

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 168-182, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

PERFORMANCE EVALUATION OF OBJECT CLUSTERING USING TRADITIONAL AND FUZZY LOGIC ALGORITHMS

NAZEK AL-ESSA AND MOHAMED NOUR

ABSTRACT. This work analyzes three object clustering algorithms mainly: the k-means, fuzzy c-means (FCM), and kernel fuzzy c-means (KFCM). The k-means algorithm partitions a dataset into hard clusters. Both FCM and KFCM are based on fuzzy logic and they return a membership degree of each object to all clusters. The algorithms are implemented and run on two different datasets. It is shown that the KFCM algorithm can achieve better results than the other two algorithms. The FCM is slightly better than the k-means algorithm. Moreover, the clustering time of each algorithm is different from the others. The clustering time of KFCM was larger than both the k-means and FCM. Clustering using the k-means has the smallest clustering time. The overall performance is affected by the dimensionality value, fuzziness parameter, number of objects, iterations, clusters, and the number of updating operations for both cluster centroids and membership values.

1. INTRODUCTION

Clustering algorithms play an important role in the field of data analysis. Such algorithms partition an object set into subsets or clusters where the objects in each cluster share some common properties. i.e. the objects belonging to the same cluster are as similar as possible whereas objects belonging to different clusters are dissimilar as possible.

There are several approaches for conducting the clustering process. An example of such approaches is that one based on fuzzy clustering. Fuzzy clustering in contrast to the usual methods does not provide hard clusters, but return a degree of membership of each object to all clusters. The interpretation of these degrees is then left to the user that can apply some kind of threshold to generate hard clusters or use these soft degrees directly [1]. The necessity of fuzzy clustering lies in the reality that a pattern could be assigned to different classes and/or categories. In fuzzy clustering, a result is represented by degrees of membership of every pattern to the established classes [2].

There are several types of fuzzy clustering. This includes those methods based on fuzzy relations, objective functions, adaptive resonance theory, competitive learning, and others. Several research efforts were done in the area of fuzzy clustering. Examples of such efforts are briefly mentioned as follows:-

[3] mentioned that document clustering has been widely applied in the field of information retrieval for improving search and retrieval efficiency. The authors proposed a scalable fuzzy algorithm for document clustering. The algorithm discovers relationships between information resources based on their textual contents. The algorithm addresses the problem of defining a suitable number of clusters for appropriately capturing all the topics of the knowledge domain.

[4] presented a sample weighting clustering algorithm based on K-means and fuzzy C-means approaches. The algorithm uses academic documents as the clustering objects. Experiments show that the proposed algorithm is an effective solution to improve the performance of document clustering.

[5] described an algorithm for building fuzzy hierarchies. The author presented an approach that mainly follows a bottom-up strategy, and described the functions needed to operate with fuzzy attributes. The author presented an application example for that approach.

[6] presented a way to use fuzzy clustering for generating fuzzy rule bases in the implementation of an intelligent agent. That agent interacts with human for diagnosis establishment: the medical diagnostics system. The results were compared with the results in some subtests of the Aachen Aphasia Test (AAT).

[7] developed an interactive content-based image retrieval (CBIR) system that allows searching and retrieving images from databases. Based on the fuzzy-c-means clustering algorithm, the CBIR system fuses color and texture features image segmentation. A database consisting of skin cancer imagery is used to demonstrate the applicability of such CBIR system.

[8] introduced a fuzzy-based distributed clustering algorithm to cluster distributed datasets without necessarily downloading all the data into a single site. The algorithm is compared against two existing distributed clustering algorithms where all data are merged into a single data source and clustered. The experiments confirmed good performance of the proposed algorithm.

The organization of this paper will be as follows: Section 2 addresses; in addition to the work objective; the main difference between crisp and fuzzy clustering. Sections 3, 4, and 5 present the adopted algorithms respectively. The implementation work is presented in Section 6 while Section 7 discusses the results. Finally, Section 8 concludes the whole work.

2. OBJECT CLUSTERING BASED ON CRISP CLUSTERING AND FUZZY CLUSTERING

Regarding the cluster analysis and/or cluster algorithms, there are two main approaches: crisp clustering and fuzzy clustering. In many real cases of crisp clustering, the boundaries between clusters cannot be clearly defined. i.e. some objects may belong to more than one cluster. In such cases, the fuzzy clustering method provides a better method to cluster such objects. Membership degrees between zero and one are used in fuzzy clustering instead of crisp assignments of those object clusters.

This work aims at presenting an analysis and investigation of three algorithms for clustering objects. The performance of such algorithms is also evaluated. The adopted algorithms are: the K-means algorithm, the fuzzy c-means algorithm, and the kernel fuzzy c-means algorithm. The steps of such algorithms are briefly mentioned in the following sections.

3. CLUSTERING OF OBJECTS USING THE K-MEANS ALGORITHM

The K-means algorithm provides an easy method for clustering objects or data set. For a given dataset of n data points (or objects) x_1, x_2, …. x_n each is in R_d, the k-means algorithm classifies such points into k-clusters. Initially, it is easy to define k centroids; one for each cluster. Such centroids are better to be chosen as much as possible far away from each other. Each point in the dataset is associated to the nearest centroid. K new centroids are recalculated resulting from the previous step. After that a new binding has to be done between the same dataset points and the new centroids. The k-means algorithm iteratively updates these centroids to decrease the objective function shown in equation (3.1). The algorithm always converges to a local minimum. As a result, the k centroids change their location step by step until no more changes are done. .i.e. the algorithm updates the cluster centroids till local minimum is found. Moreover, the objective function J which the algorithm aims to minimize is a square error function as follows:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - v_j \right\|^2 , \tag{3.1}$$

where $\left\|x_i^{(j)} - v_j\right\|^2$ is a chosen distance measure between a data point x_i^j and the cluster center v_j. i.e the objective function involves the distance of the n data points from their respective cluster centers.

In other words, some researchers use the similarity measure instead of using the distance measure mentioned in equation (3.1). The similarity measure between any two objects or data elements x_i, and x_j (i.e. sim(x_i, x_j)) is the cos value between such elements. In fact, that measure is preferred. The generalized pseudo code for the k-means algorithm can be written as shown in Figure 1. The algorithm was taken from [10, 14].

Algorithm: The k-means algorithm
Input: 1. A set of objects or data points with d-dimensional each x_i, i=1, 2, …n
2. The number of clusters k
Output: k clusters of data points
Steps:
1. Generate k centroids $v_1, v_2, \ldots\ldots, v_k$ by randomly choosing k data points from X
2. REPEAT UNTIL there is no change in clusters between two consecutive iterations
   FOR each data point x_i in X
      FOR j=1 to k
         Sim(v_j, x_i) = Find cos similarity between x_i and v_j
      ENDFOR
      Assign x_i to cluster j which Sim(v_j, x_i) is maximum
   ENDFOR
   Update the centroid for each cluster
   END LOOP
END K-means Algorithm

Figure 1: The Generalized Pseudo Code for the K-Means Algorithm

4. CLUSTERING OF OBJECTS USING THE FUZZY C-MEANS ALGORITHM

Clustering methods can be considered as either hard (crisp) or fuzzy depending on whether a pattern data belongs exclusively to a single cluster or to several clusters with different degrees. In hard clustering, a membership value of zero or one is assigned to each pattern data (feature vector), whereas in fuzzy clustering a value between zero and one is assigned to each pattern by a membership function. The fuzzy c-means (FCM) is one of the fuzzy algorithms that can be used in a wide variety of engineering and scientific disciplines. The FCM is a clustering method that allows one piece of data to belong to two or more clusters. The FCM aims at minimizing the following objective function:

(4.1) $$J_m = \sum_{i=1}^{n}\sum_{j=1}^{c}\mu_{ij}^{m}\left\|x_i - v_j\right\|^2 \ , 1 \le m < \infty \ ,$$

where m is any real number greater than one, μ_{ij} is the degree of membership of x_i in the cluster j, x_i is the i[th] of d-dimensional measured data, v_j is the d-dimensional center of the cluster, n is the number of data points, c is the number of clusters, and $\|*\|$ is any norm expressing the similarity between any measured data and the center.

The clustering process using the FCM algorithm can take place through an iterative optimization of the objective function shown in equation (4.1). This is done by updating the membership μ_{ij} and the cluster centers v_j as follows:-

(4.2)
$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c}\left\{\frac{\|x_i - v_j\|}{\|x_i - v_k\|}\right\}^{\frac{2}{m-1}}}$$

(4.3)
$$v_j = \frac{\sum_{i=1}^{n}\mu_{ij}^{m} \cdot x_i}{\sum_{i=1}^{n}\mu_{ij}^{m}}.$$

The iteration stops when

$$max_{ij}\ \left\{\left|\mu_{ij}^{(k+1)} - \mu_{ij}^{(k)}\right|\right\} < \varepsilon$$

where ε is a termination criterion between zero and one. Also k is the iteration steps.

Figure 2 briefly shows the main steps of the fuzzy c-means algorithm (FCM). The algorithm was taken from [1, 2, 5, 9].

Algorithm: The Fuzzy C-Means Algorithm

Input:

1. n data elements $x_i = 1, 2, \ldots n$; each is of d-dimensional measured data
2. The number of clusters c

Output:

C clusters with data elements; each with memberships values to the clusters

Steps:

1. Initialize $U = [\ \mu_{ij}\ ]$ matrix $u(0)$ $U^{(0)}$
2. At k-step, calculate the center vectors $V^{(k)} = [v_j\ ]$ with $U^{(k)}$
$$v_j = \frac{\sum_{i=1}^{n}\mu_{ij}^{m} \cdot x_i}{\sum_{i=1}^{n}\mu_{ij}^{m}}$$
3. Update $U^{(k)}, U^{(k+1)}$
$$\mu_{ij} = \frac{1}{\sum_{i=1}^{c}\left[\frac{\|x_i - v_j\|}{\|x_i - v_k\|}\right]^{\frac{2}{m-1}}}$$
4. If IF $\|U^{(k+1)} - U^{(k)}\| < \varepsilon$ then stop; otherwise return to step 2.

//* ε is a termination criterion between zero and one and k is the iteration step. *//

End of the Fuzzy c-means Algorithm

Figure 2: The Main Steps of the Fuzzy C-Means Algorithm

5. CLUSTERING OF OBJECTS USING THE KERNEL FUZZY C-MEANS ALGORITHM

The kernel fuzzy c-means approach (KFCM) is mainly based on the fuzzy c-means algorithm. That approach tries to minimize the following objective function:

$$J_m(U,V) = \sum_{i=1}^{C}\sum_{K=1}^{n}\mu_{ik}^{m}\,\|x_k - v_i\|^2 + \sum_{i=1}^{n}\delta_i \sum_{k=1}^{n}(1-\mu_{ik})^m \tag{5.1}$$

where δ_i are suitable positive numbers, *m* is the quantity controlling clustering fuzziness and it is greater than one as mentioned before. Also, $\sum_{i=1}^{c}\mu_{ik} = 1$, *V* is the set of cluster centers or prototypes ($v_i \in R^p$). The function J_m should be minimized.

The first part in equation (5.1) needs to be as low as possible; whereas the second part forces μ_{ik} to be as large as possible. Moreover, δ is better to be chosen as

$$\delta_i = K\,\frac{\sum_{k=1}^{n}\mu_{ik}^{m}\,\|x_k - v_i\|^2}{\sum_{k=1}^{n}\mu_{ik}^{m}} \tag{5.2}$$

Typically, *K* is chosen to be 1. The cluster centroids or prototypes can be updated like those used in the fuzzy c-means, but the memberships of this approach can be updated as follows:

$$\mu_{ik} = \frac{1}{1+\left(\frac{\|x_k - v_i\|^2}{\delta_i}\right)^{\frac{1}{(m-1)}}} \tag{5.3}$$

Figure 3 shows the main steps of KFCM. The algorithm was taken from [11].

Algorithm: The Kernel Fuzzy C-Means Algorithm

Input:

1. n data elements x_i = 1, 2, …n ; each is of d-dimensional measured data
2. The number of clusters c

Output:

C clusters with data elements; each with membership values to the clusters

Steps:

1. Fix *c*, t_{max} , *m*>1, and δ >0
2. Initialize μ_{ik}
3. Estimate δ_i using equation (5.2)
4. FOR *t* =1, 2, …… t_{max} , DO

Update all cluster centers or prototypes v_i^t with equation (4.3)

Update all memberships μ_{ik}^t with equation (5.3)

Compute $E^t = max_{i,k}\left|\mu_{ik}^{t} - \mu_{ik}^{t-1}\right|$, IF $E^t \leq \varepsilon$ STOP; ELSE t=t+1

End of the Steps of the Kernel Fuzzy c-means Algorithm

Figure 3: The Main Steps of the Kernel Fuzzy C-Means Algorithm

6. IMPLEMENTATION WORK

The algorithms have been implemented and tested with two benchmark numeric datasets. The datasets are: SPECT-heart and liver-disorder datasets. Such datasets were taken from the UCI machine learning data repository. The main features and/or characteristics of the chosen datasets are briefly shown in Table 1 [13].

Table 1: The Chosen Datasets as Test-beds

Dataset Name	No. of Attributes	No. of Instances	Area	No. of Web Hits	Date Donated
SPECT Heart	22	267	Life	51564	2001-10-01
Liver Disorder	7	345	Life	20782	1990-05-15

The experiments were run for partitioning each dataset into a set of clusters mainly: two, four, six, and eight clusters respectively. The results were reported for such numbers for the k-means algorithm. For the fuzzy c-means and kernel fuzzy c-means algorithms; the reported results were only for two and four clusters for the space limitation for writing these pages.

6.1 APPLYING THE K-MEANS ALGORITHM

As mentioned before, the k-means algorithm is based on the idea that a center point can represent a cluster. Also, the Cos measure is used to compute the similarity between any two data points x_i and x_j [12, 14].

(6.1) $$\text{Cos}\,(x_i, x_j) = (x_i \,.\, x_j) \,/\, \| x_i \| \; \| x_j \|$$

Where '.' indicates the vector dot product and $\| x \|$ is the length of vector x.
It is assumed that a dataset S of data items and their corresponding vector representation, the centroid vector v can be defined as:

(6.2) $$V = (1 \,/\, |S|) \sum_{x_i \in S} x_i$$

The similarity between a centroid vector v and any data item x is calculated as in equation (6.3). Also, the similarity between any two centroid vectors v_i and v_j is calculated as in equation (6.4).

(6.3) $$\text{Cos}\,(x , v) = (x \,.\, v) / \| x \| \; \| v \|$$

(6.4) $$\text{Cos}\,(v_i, v_j) = (v_i, v_j) \,/\, \| v_i \| \; \| v_j \|$$

Moreover, two measures can be used to evaluate the quality of clusters mainly the intra-cluster similarity (intra-CS), and the inter-cluster similarity (inter-CS). Such measures are written as follows:-

(6.5) $$\text{Intra-CS}(j) = (1 \,/\, |S|) \; \Sigma_{x \in S} \, \text{Cos}\,(x, v_j)$$

The average over all the clusters is

(6.6) $$\text{AvgIntra-CS} = (1/K)\ \sum_{i=1}^{k} IntraSim(i)$$

Where k is the number of clusters.

Regarding the inter-CS, the pairwise similarity between each pair of clusters is measured by computing the similarity between their centroids. Also, the average inter-cluster similarity (Avginter-CS) is measured by computing an average over all the distinct pairwise similarity between centroids.

(6.7) $$\text{AvgInter-CS} = (2/K(K-1))\ \sum_{i=1}^{k-1}\sum_{j=i+1}^{k} Sim(v_i, v_j)$$

The obtained results after running the k-means algorithm on the chosen datasets are shown in Figure 4.

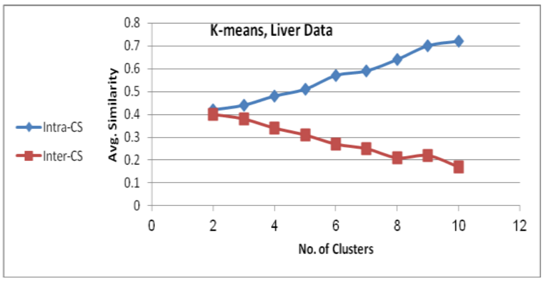

Figure 4a: Average Similarity for K-means

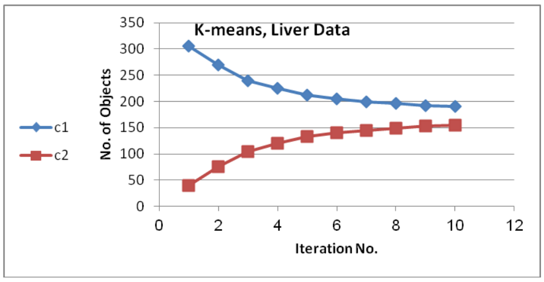

Figure 4b: No. of Objects for Two Clusters

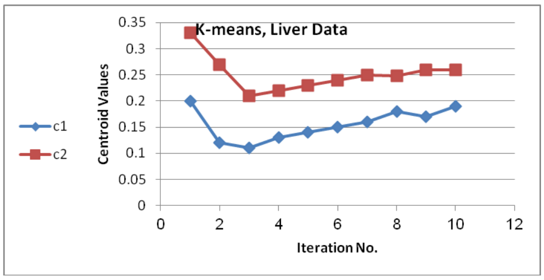

Figure 4c: Centroids of Two Clusters

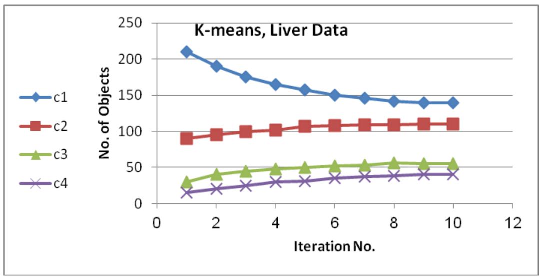

Figure 4d: No. of Objects for Four Clusters

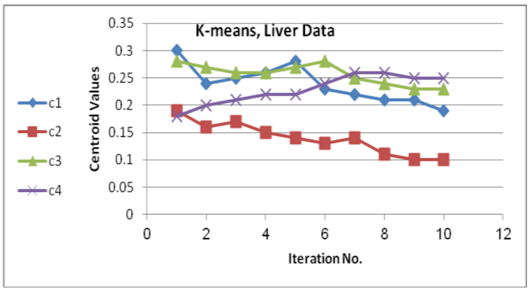

Figure 4e: Centroids of Four Clusters

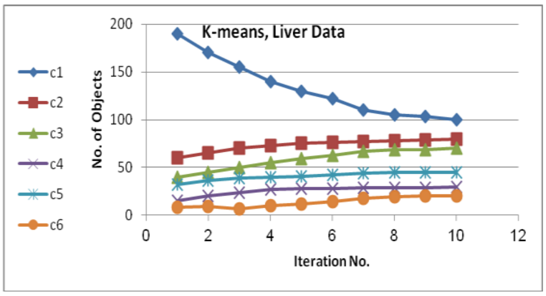

Figure 4f: No. of Objects for Six Clusters

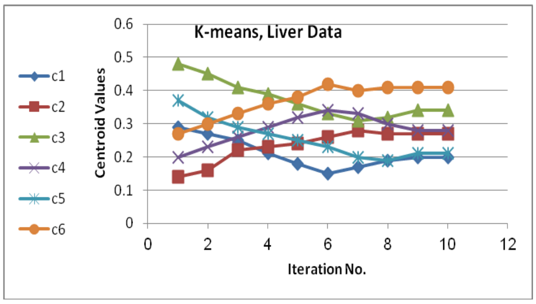

Figure 4g: Centroids of Six Clusters

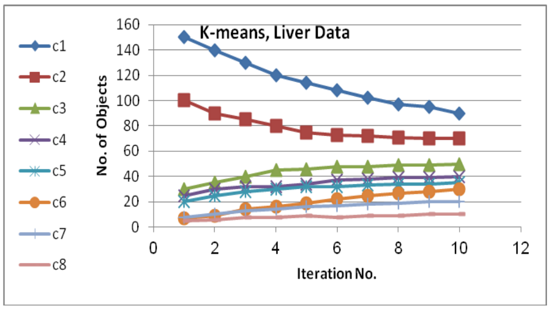

Figure 4h: No. of Objects for Eight Clusters

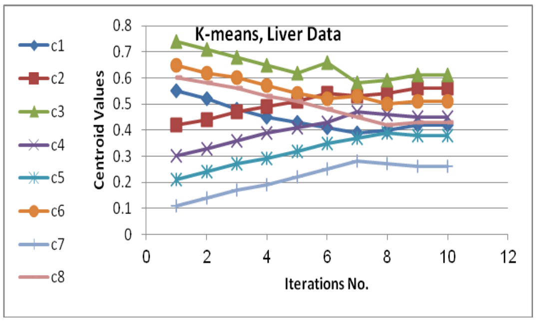

Figure 4i: Centroids of Eight Clusters

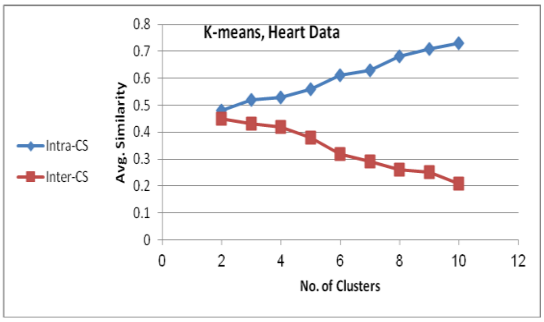

Figure 4j: Average Similarity for K-Means

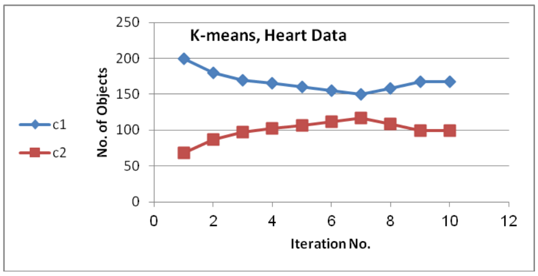

Figure 4k: No. of Objects for Two Clusters

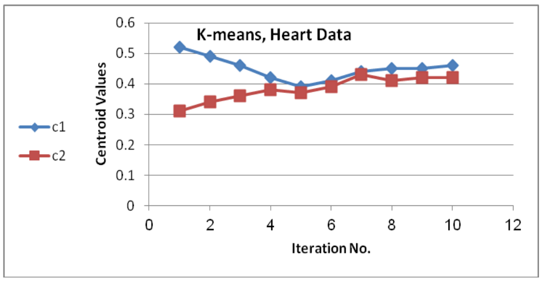

Figure 4l: Centroids of Two Clusters

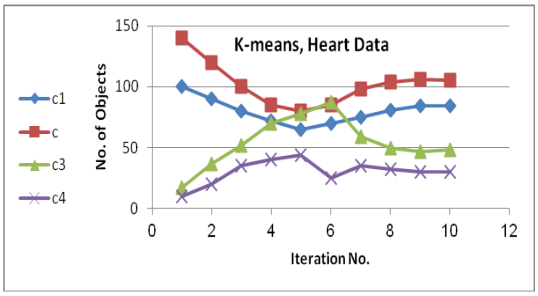

Figure 4m: No. of Objects for Four Clusters

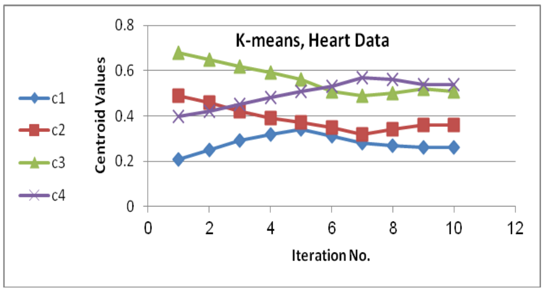

Figure 4n: Centroids of Four Clusters

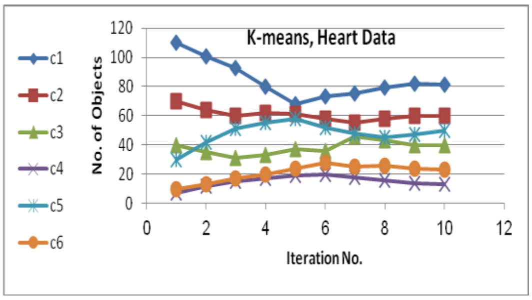

Figure 4o: No. of Objects for Six Objects

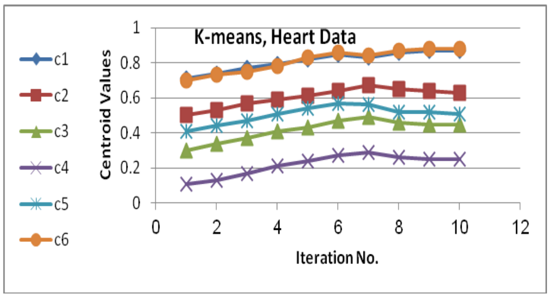

Figure 4p: Centroids of Six Clusters

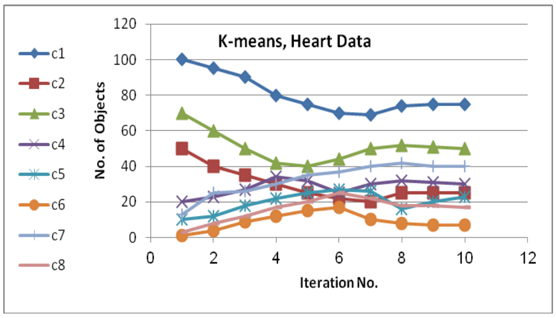

Figure 4q: No. of Objects for Eight Clusters

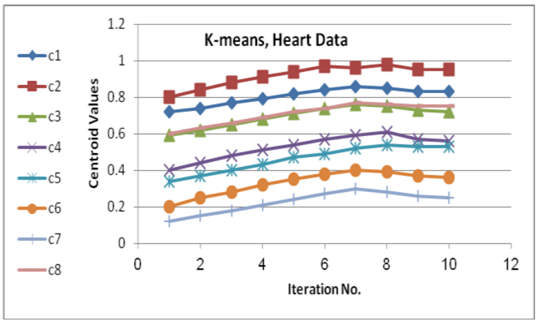

Figure 4r: Centroids of Eight Clusters

6.2 APPLYING THE FUZZY C-MEANS ALGORITHM

The experiments for this algorithm were run for different values of the fuzziness parameter (m). The adopted values of that m were: 3, 2.5, 2, and 1.5 respectively. For each value of m; a set of iterations were run. The number of objects and the centroids of the fuzzy clusters were reported. Moreover, both the centroids of the fuzzy clusters and the membership values are iteratively recomputed and updated. The iterations stop after achieving the stopping criterion. Different values of m were tested; as mentioned above; to see the effectiveness and role of the fuzziness parameter on the obtained clusters. This is important for evaluating the performance of the algorithm. Figure 5 presents the cluster centroid values versus the iteration number for the adopted datasets.

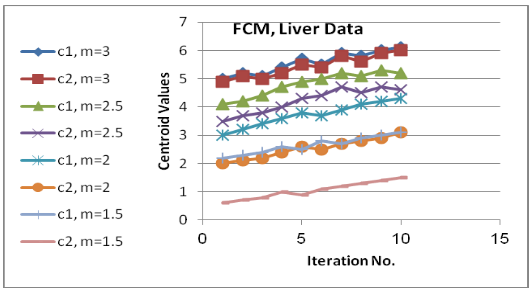

Figure 5a: Centroids of Two Clusters

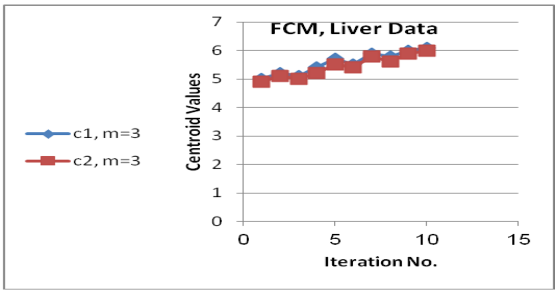

Figure 5b: Centroids of Two Clusters

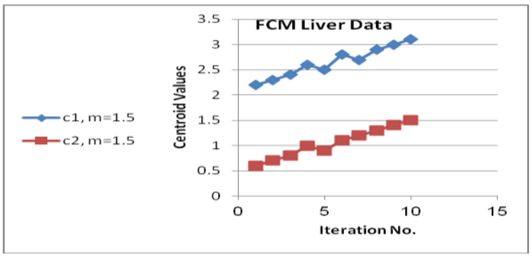

Figure 5c: Centroids of Two Clusters

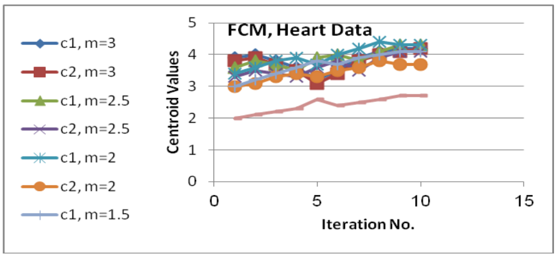

Figure 5d: Centroids of Two Clusters

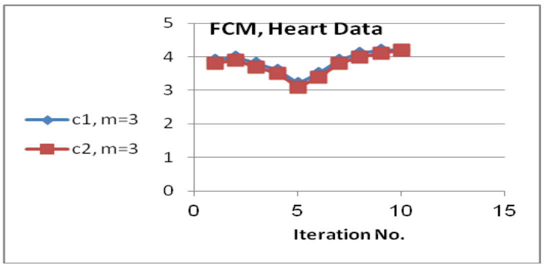

Figure 5e: Centroids of Two Clusters

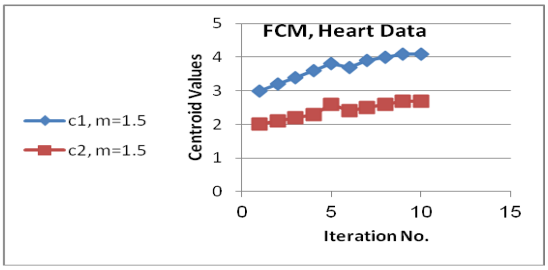

Figure 5f: Centroids of Two Clusters

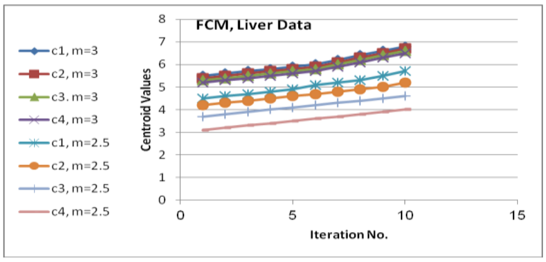

Figure 5g: Centroids of Four Clusters

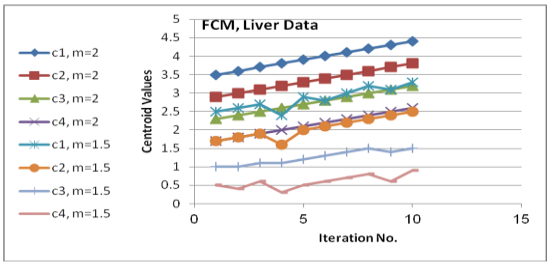

Figure 5h: Centroids of Four Clusters

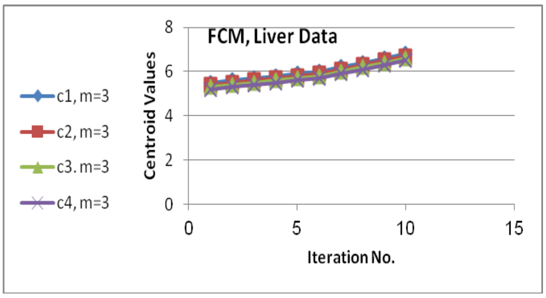

Figure 5i: Centroids of Four Clusters

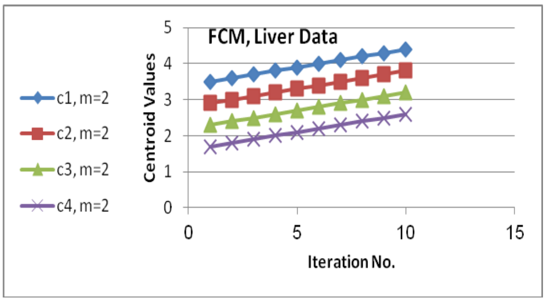

Figure 5j: Centroids of Four Clusters

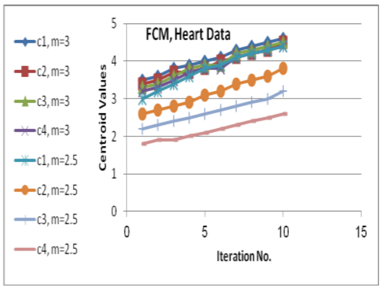

Figure 5k: Centroids of 4 Clusters

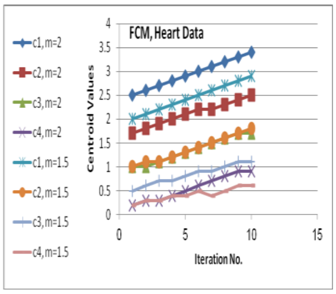

Figure 5l: Centroids of 4 Clusters

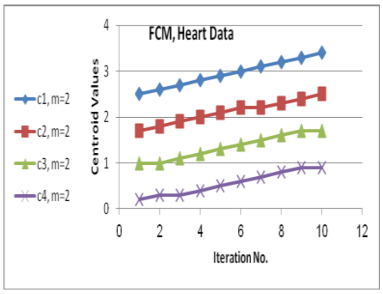

Figure 5m: Centroids of 4 Clusters

6.3 APPLYING THE KERNEL FUZZY C-MEANS ALGORITHM

The experiments implemented for the fuzzy c-means algorithm are also operated for the kernel fuzzy c-means algorithm. The centroids of the fuzzy clusters and the membership values are also iteratively updated. As the centroid and membership values change, the objective function also changes. For the same experiments, the objective function in this algorithm is different from the corresponding values in the FCM due to the added part in the objective function. Such part forces the membership values to be as large as possible. Figure 6 shows the results of applying this algorithm on the adopted datasets mentioned before.

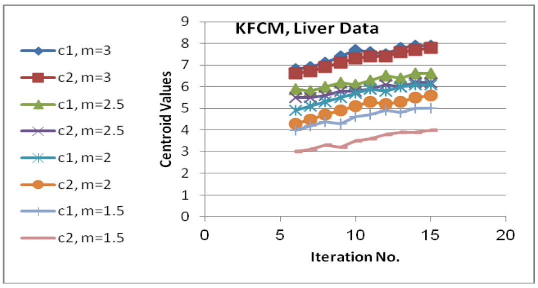

Figure 6a: Centroids of Two Clusters

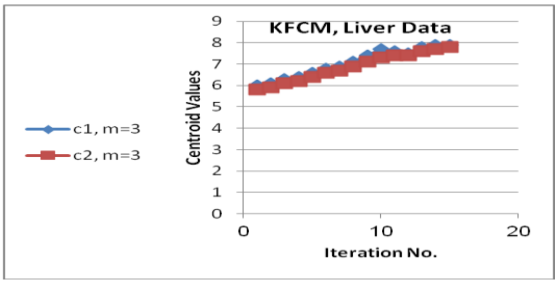

Figure 6b: Centroids of Two Clusters

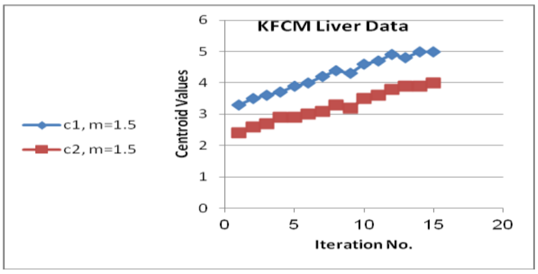

Figure 6c: Centroids of Two Clusters

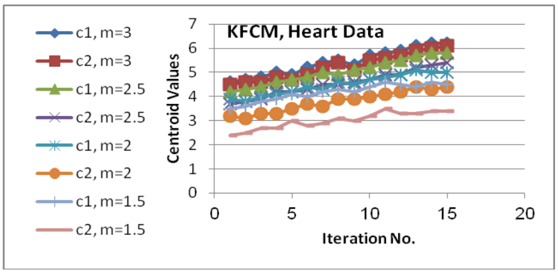

Figure 6d: Centroids of Two Clusters

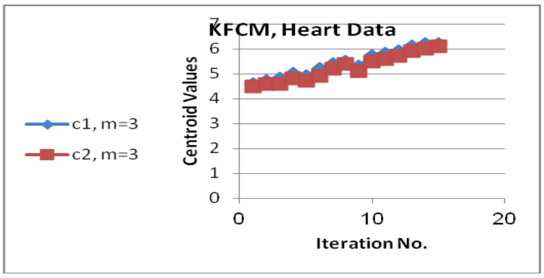

Figure 6e: Centroids of Two Clusters

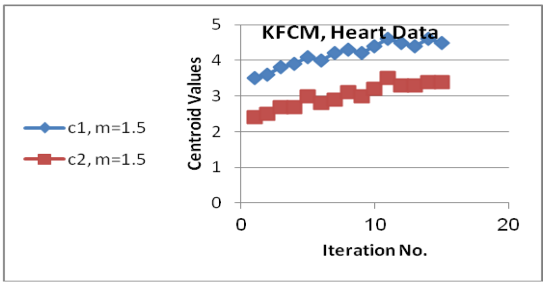

Figure 6f: Centroids of Two Clusters

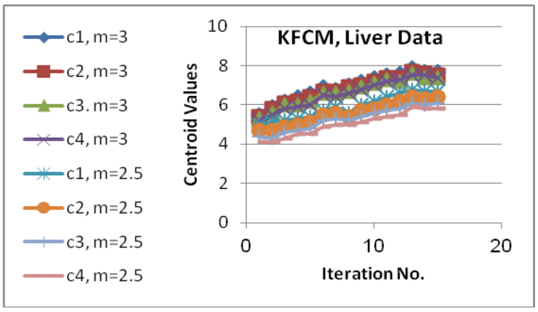

Figure 6g: Centroids of Four Clusters

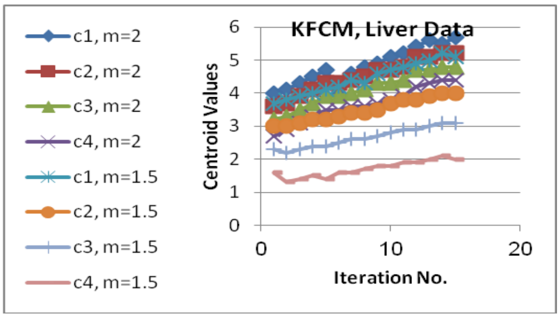

Figure 6h: Centroids of Four Clusters

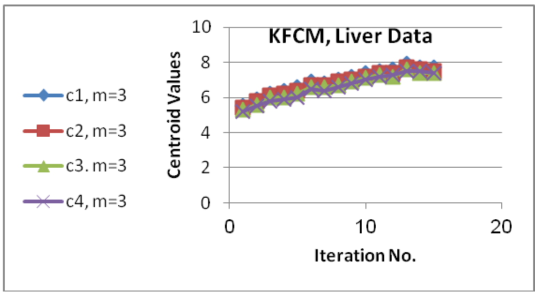

Figure 6i: Centroids of Four Clusters

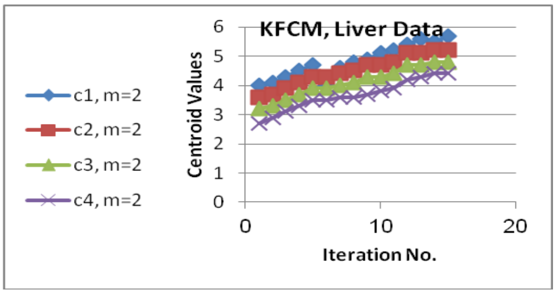

Figure 6j: Centroids of Four Clusters

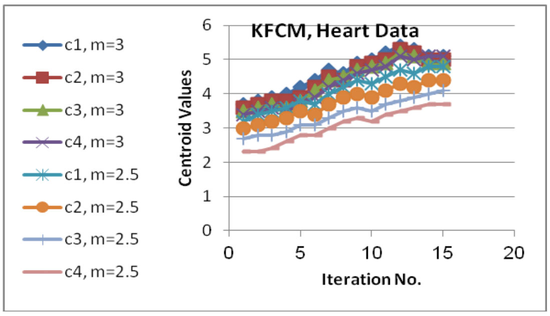

Figure 6k: Centroids of Four Clusters

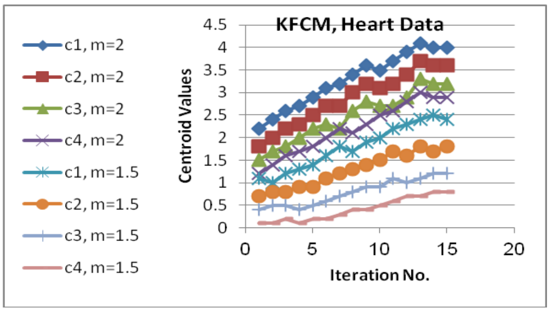

Figure 6l: Centroids of Four Clusters

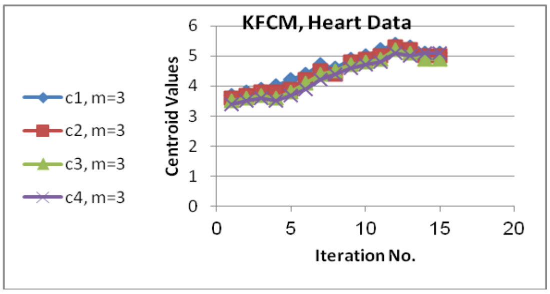

Figure 6m: Centroids of Four Clusters

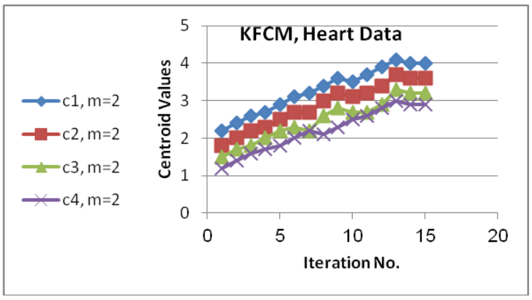

Figure 6n: Centroids of Four Clusters

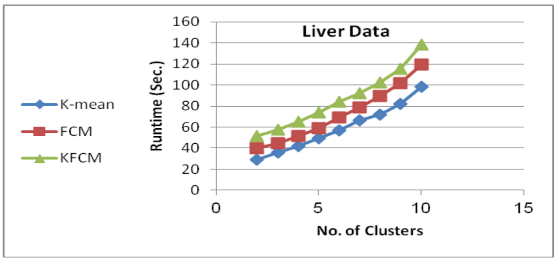

Figure 6o: Runtime in (Sec.)

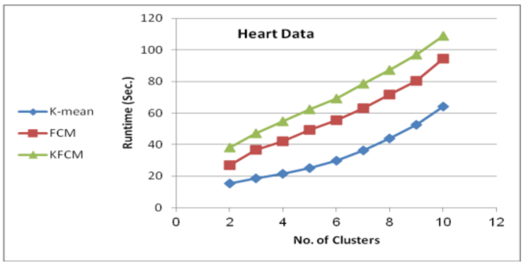

Figure 6p: Runtime in (Sec.)

7. DISCUSSION OF RESULTS

The adopted algorithms were iteratively run on the two datasets mentioned above. For the K-means algorithm, the average intra-cluster similarity (Intra-CS) and the inter-cluster similarity (Inter-CS) change by changing the number of clusters. As the number of clusters increases the intra-cs increases while the inter-cs decreases. The intra-cs is always greater than the inter-cs for the same number of clusters. Each iteration the K-means is run, the number of objects in each cluster and their cluster centroids are different. This is due to the massive movement of objects among clusters in the initial iterations. That movement has been stopped in the last iterations. This had been occurred for the two datasets for two, four, six, and eight clusters respectively. Running the algorithm on the two different datasets for the same number of clusters gives different values of cluster centroids and number of objects. This means that the size, nature, and characterization of each dataset are significant. For both the fuzzy c-mean (FCM) and kernel fuzzy c-mean (KFCM) algorithms, the fuzziness parameter has a significant effect on the fuzzy clustering. The algorithms were iteratively run for four different values of the fuzziness parameter. The quality of the fuzzy clusters is improved by decreasing the fuzziness value. i.e. For the higher fuzziness values, the fuzzy cluster centroids became more closer to each other or sometimes equal. This causes the membership values of an object to all clusters to be approximately the same. This causes bad clustering quality. The best fuzzy clusters were for the small values of the fuzziness parameter near to one. This was valid for the two algorithms for the chosen datasets for two and four clusters. Moreover, the rate of change in cluster centroids was large for the initial iterations and very small at the last iteration. For the same iteration, the same fuzziness value, the same dataset, the KFCM algorithm outperforms the FCM algorithm. This is clear in the fuzzy cluster centroids which show higher differences than those corresponding in the FCM algorithm. As a result, the membership values were higher for the KFCM which means better quality. The reason for this refers to that added part in calculating the objective function in the KFCM algorithm. The fuzzy clustering time in the KFCM algorithm was higher than its corresponding values of the FCM algorithm and this is clear from the extra iterations consumed in running the KFCM algorithm. Regarding the time complexity, the clustering time for the three algorithms was different. The highest clustering time was for the KFCM algorithm and the smallest one was for the K-means algorithm.

Finally, the size of the chosen dataset is significant. i.e. For each dataset, the number of objects with the properties of each plays a vital role in the clustering time. Such time is always affected by several factors mainly: the number of iterations, the number of clusters, the number of objects, the dimensionality of each object, and the adopted algorithm.

8. CONCLUSION

This work presents an investigation of three clustering algorithms. As the number of clusters increases in the K-means algorithm, the intra-cluster similarity increases while the inter-cluster similarity decreases. The intra-cluster similarity is always greater than the inter-cluster similarity. Each time the algorithm is run, the objects in clusters were different in several iterations till the movement of objects among clusters has been stopped.

For both the fuzzy c-means and kernel fuzzy c-means, the fuzziness parameter has a significant effect on the clustering process. The best clusters occurred for small values of the fuzziness parameter near to one. This is due to the different centroid values of fuzzy clusters. Moreover, the differences among cluster centroids in the kernel fuzzy c-means were greater than their corresponding values in the fuzzy c-means. This had led to better quality clustering for the kernel fuzzy c-means. The kernel fuzzy c-means algorithm outperforms the other two algorithms. This is due to the added part in the objective function which forces the membership values to be as large as possible. Moreover, the clustering time was highest for that algorithm and lowest for the k-means algorithm. Finally, the clustering time increases in all algorithms by increasing the number of clusters.

REFERENCES

[1] M. Jursic, and N. Lavrac, Fuzzy clustering of documents, downloaded from the Internet in 2010 from the website http://www.vidolectures.net/sikdd08_jursic_fcd/.

[2] M. Makrehchi, and M. Shakri, Document categorization using fuzzy clustering, downloaded From the Internet in 2010 from the website http://www. docstoc.com/ docs/24849198/Document-Categorization-using-fuzzy-clustering/.

[3] M. Rodrigues, and L. Sacks, A scalable hierarchical fuzzy clustering algorithm for text mining, downloaded from the Internet in 2010 from the website http://www. Research.microsoft.com/apps/pubs/default.aspx?id=72934.

[4] C. Zhang, X. Su, and D. Zhou, Document clustering using sample weighting, downloaded from the Internet in 2010 from the website http://sites.google.com/ site/zhangczhomepage/publications-of-chengzhi-zhang.

[5] V. Torra, Fuzzy c-means for fuzzy hierarchical clustering, downloaded from the Internet in 2010 from the website http://www.academic.research.microsoft.com/ Author/456323/vicenc-torra.

[6] V. Zdrenghea, D. Man, and M. Tosa-Abrudan, Fuzzy clustering in an intelligent agent for diagnosis establishment, Studia Univ. Babes-Bolyai, Informatica, 3, 79-86, (2010).

[7] W.S. Ooi, and C.P. Lim, A fuzzy clustering approach to content-based image retrieval, A paper presented in the workshop on advances in intelligent computing, Ulsan, Korea, 2009, pp. 11-16.

[8] N. Visalakshi, K. Thangavel, and R. Parvathi, An intuitionistic fuzzy approach to distributed fuzzy clustering, The Int. J. of Computer Theory and Engineering, 2, 2, 295-302, April (2010).

[9] H. Fu, and A. Elmisery, A new feature weighted fuzzy c-means clustering algorithm, downloaded from the Internet in 2010 from the website http://futurecomm.tssg.org/public/publications/2009_IADIS_fu_et_al.pdf

[10] O.J. Oyelade, O.O. Oladipupo, and I.C. Obagbuwa, Application of k-means clustering, J. of Computer Science and Information Security IJCSIS, 7, 1, 292-295 (2010).

[11] D.Q. Zhang, and S.C. Chen, Kernel-based fuzzy and possibilistics c-means clustering, downloaded from the Internet in 2010 from the website http://www.visionbib.com/bibliography/pattern624.html.

[12] M. Steinbach, G. Karypis, and V. Kumar, A comparison of document clustering techniques, downloaded from the Internet in 2010 from the website http://isearch.babylon.com/?q=A+Comparison+of+document+clustering+techniques&s=%7Bs_type%7D&babsrc=HP_ss.

[13] C.J. Merz and P.M. Murphy, UCI repository of machine learning databases, Irvine, University of California, downloaded from the Internet in 2011 from the website http://www.ics.uci.edu/~mlearn/.

[14] H. Khatri, Project C, downloaded from the Internet in 2010 from the website http://www.localhost:8080/cse494/search.

(N. AL-Essa) PRINCESS NOURA UNIVERSITY, RIYADH, KINGDOM OF SAUDI ARABIA
E-mail address: dr.nazeka@hotmail.com

(M. Nour) PRINCESS NOURA UNIVERSITY, RIYADH, KINGDOM OF SAUDI ARABIA
E-mail address: mnour@hotmail.com

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 183-189, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

NINE POINT MULTISTEP METHODS FOR LINEAR TRANSPORT EQUATION.

PARIA SATTARI SHAJARI AND KARIM IVAZ

ABSTRACT. In this paper we construct a family of multistep methods for the linear advection equation, based on nine points of a rectangle form grid. Square polynomials are used for this purpose. Numerical examples show the performance of the different methods according to the choice of the parameters.

1. INTRODUCTION

In this paper we generalize one step methods introduced in [1] for the linear hyperbolic scalar equation:

$$L(u) = u_t + cu_x = 0 \tag{1.1}$$

where c is a positive constant. Also we consider linear advection-diffusion equation of the form:

$$\widehat{L}(u) = u_t + cu_x - \nu u_{xx} = 0 \tag{1.2}$$

where $\nu > 0$ is the diffusion coefficient. Initial conditions and inflow boundary conditions are provided in the usual way, so that the solution of (1.1) turns out to be a traveling wave preserving its shape. However, the numerical method we are going to present here for the simple linear case will be significant enough to enable many interesting conclusions to be drawn.

The general theory on hyperbolic equations and conservation laws has already generated an enormous amount of literature (see for instance [2], [7]). The relevance of advection-dominated problems is also testified to by a number of recent papers dealing with a variety of approximating methods and numerical schemes [6]-[10].

We use nine points of a rectangle form grid with square polynomials to construct multistep methods for solving Eq. (1.1).

2. NINE POINT MULTISTEP METHODS

Throughout this paper we only consider finite-difference methods for solving equations (1.1) and (1.2) on $t \in [0, T]$ and $x \in [0, L]$, based on nine points of a rectangle form grid (see figure 2).

Thus, in the (x, t) plane we take a uniform grid of width $h = \Delta x = L/M$ and a constant time-step $\Delta t = T/N$. We use the nine point (x_{j-1}, t_{k-1}), (x_j, t_k), ...,(x_{j+1}, t_{k+1}) for constructing an interpolation of approximate solution. Hence, we introduce following Lagrange basis for any fixed j and k

$$V = \{l_i(x)G_m(t) \in p^{(2,2)} : i = 0, 1, 2, \quad m = 0, 1, 2\}$$

Key words and phrases. Linear transport equation, Finite difference methods, Multistep methods, Advection-diffusion equation.

2010 *AMS Math. Subject Classification.* 65N06.

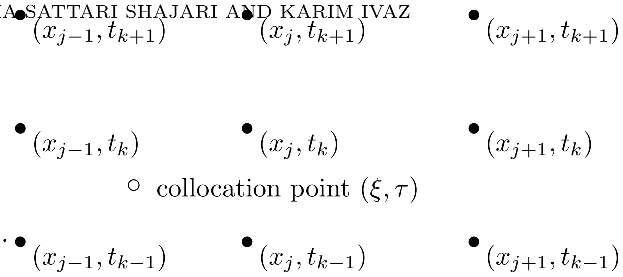

FIGURE 1. Nine points of a rectangle form grid

where

$$L_0(x) = \frac{1}{2h^2}(x - x_j)(x - x_{j+1}),$$
$$L_1(x) = \frac{-1}{h^2}(x - x_{j-1})(x - x_{j+1}),$$
$$L_2(x) = \frac{1}{2h^2}(x - x_{j-1})(x - x_j),$$
$$G_0(t) = \frac{1}{2\Delta t^2}(t - t_j)(t - t_{j+1}),$$
$$G_1(t) = \frac{-1}{\Delta t^2}(t - t_{j-1})(t - t_{j+1})$$

and

$$G_2(t) = \frac{1}{2\Delta t^2}(t - t_{j-1})(t - t_j).$$

Therefore, for any $p \in P^{(2,2)}$, we have

$$p(x,t) = \sum_{i=0}^{2} \sum_{m=0}^{2} p_{j+i-1}^{k+m-1} L_i(x) G_m(t)$$

where, to simplify the notation, we set $p_j^k = p(x_j, t_k)$. For approximating solution of equation (1.2) in the space V, we substitute

$$p_x(x,t) = \sum_{i=0}^{2} \sum_{m=0}^{2} p_{j+i-1}^{k+m-1} L_i'(x) G_m(t),$$
$$p_{xx}(x,t) = \sum_{i=0}^{2} \sum_{m=0}^{2} p_{j+i-1}^{k+m-1} L_i''(x) G_m(t)$$

and

$$p_t(x,t) = \sum_{i=0}^{2} \sum_{m=0}^{2} p_{j+i-1}^{k+m-1} L_i(x) G_m'(t).$$

in (1.2) and we obtain

(2.1)
$$R(x,t) = \widehat{L}(p) = \sum_{i=0}^{2} \sum_{m=0}^{2} p_{j+i-1}^{k+m-1} \left(L_i(x) G_m'(t) + c L_i'(x) G_m(t) - \nu L_i''(x) G_m(t) \right).$$

The residual term R should be zero in a same sense. Let (ξ, τ) be a point in the rectangle of vertices (x_{j-1}, t_{k-1}), (x_{j+1}, t_{k-1}), (x_{j+1}, t_{k+1}) and (x_{j-1}, t_{k+1}) (see figure 2). Then, the numerical scheme will be obtained by requiring the residual R to vanish at (ξ, τ). Hence, we have

$$R(\xi, \tau) = 0. \tag{2.2}$$

For simplifying the notations, we let

$$s = \xi - x_{j-1}$$

and

$$r = \tau - t_{k-1}.$$

Now we change coordinate by translating x_{j-1} and t_{k-1} to the origin of coordinate. To do this we use

$$\begin{aligned}
L_0(\xi) =& \frac{1}{2h^2}(\xi - x_j)(\xi - x_{j+1}) = \frac{1}{2h^2}(-h+s)(-2h+s),\\
L_0'(\xi) =& \frac{1}{2h^2}\left((\xi - x_j) + (\xi - x_{j+1})\right) = \frac{1}{2h^2}(-3h+2s),\\
L_0''(\xi) =& \frac{1}{h^2},\\
L_1(\xi) =& \frac{-1}{h^2}s(s-2h),\\
L_1'(\xi) =& \frac{-(2s-2h)}{h^2},\\
L_1''(\xi) =& \frac{-2}{h^2},\\
L_2(\xi) =& \frac{s(s-h)}{2h^2},\\
L_2'(\xi) =& \frac{2s-h}{2h^2},\\
L_2''(\xi) =& \frac{1}{h^2},\\
G_0(\tau) =& \frac{(r-\Delta t)(r-2\Delta t)}{2\Delta t^2},\\
G_0'(\tau) =& \frac{(2r-3\Delta t)}{2\Delta t^2},\\
G_1(\tau) =& \frac{-r(r-2\Delta t)}{\Delta t^2},\\
G_1'(\tau) =& \frac{-(2r-2\Delta t)}{\Delta t^2},\\
G_2(\tau) =& \frac{r(r-\Delta t)}{2\Delta t^2},\\
G_2'(\tau) =& \frac{(2r-\Delta t)}{2\Delta t^2}
\end{aligned}$$

and introduce

$$\alpha_{im} := L_i(\xi)G_m'(\tau) + cL_i'(\xi)G_m(\tau) - \nu L_i''(\xi)G_m(\tau)$$

to obtain

$$\begin{aligned}
\alpha_{02}p_{j-1}^{k+1} + &\alpha_{12}p_j^{k+1} + \alpha_{22}p_{j+1}^{k+1}\\
&= -\alpha_{01}p_{j-1}^{k} - \alpha_{11}p_j^{k} - \alpha_{21}p_{j+1}^{k}\\
&\quad - \alpha_{00}p_{j-1}^{k-1} - \alpha_{10}p_j^{k-1} - \alpha_{20}p_{j+1}^{k-1}
\end{aligned} \tag{2.3}$$

for $k = 1, ..., N-1$. The system (2.3) shows a two steps method and can be solved if the initial steps were given.The initial steps can be found using boundary conditions and a one step method. In this paper, we use a six-point stencil method introduced in [3].The collocation parameters s and ν for the introduced nine-point method are exactly the same parameters s and ν used in the six-point stencil method. Now, suppose following initial conditions was given

$$u(x,0) = f(x) \tag{2.4}$$

$$u(0,t) = g(t) \tag{2.5}$$

and the Neumann condition

$$\frac{\partial u}{\partial x}|_{x=L} = 0. \tag{2.6}$$

This condition implies that

$$p_j^0 = f(x_j), \quad \text{for} \quad j = 0, ..., M \tag{2.7}$$

$$p_0^k = g(t_k), \quad \text{for} \quad k = 0, ..., N \tag{2.8}$$

and the Neumann type constraint

$$p_{n-1}^k = p_n^k. \tag{2.9}$$

Equations (2.3)-(2.9) can be written in the matrix form

$$HP^{k+1} = D^k \tag{2.10}$$

where

$$H = \begin{pmatrix} \alpha_{12} & \alpha_{22} & 0 & 0 & 0 & 0 & 0 \\ \alpha_{02} & \alpha_{12} & \alpha_{22} & 0 & 0 & 0 & 0 \\ 0 & \alpha_{02} & \alpha_{12} & \alpha_{22} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \alpha_{02} & \alpha_{12} & \alpha_{22} \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{pmatrix}, \tag{2.11}$$

(2.12)

$$D^k = \begin{pmatrix} -\alpha_{01}p_0^k - \alpha_{11}p_1^k - \alpha_{21}p_2^k - \alpha_{00}p_0^{k-1} - \alpha_{10}p_1^{k-1} - \alpha_{20}p_2^{k-1} - \alpha_{02}g(t_{k+1}) \\ \vdots \\ -\alpha_{01}p_{M-2}^k - \alpha_{11}p_{M-1}^k - \alpha_{21}p_M^k - \alpha_{00}p_{M-2}^{k-1} - \alpha_{10}p_{M-1}^{k-1} - \alpha_{20}p_M^{k-1} \\ 0 \end{pmatrix}$$

and $P^{k+1} = [p_1^{k+1}, \dots, p_M^{k+1}]^T$. To obtain the initial steps p_j^1 we solve following equations (see [3] Eq. (3.6))

$$\begin{aligned} p_j^1 + A(p_{j+1}^1 - 2p_j^1 + p_{j-1}^1) + B(p_{j+1}^1 - p_{j-1}^1) = \\ p_j^0 + C(p_{j+1}^0 - 2p_{j+1}^0 + p_{j-1}^0) + D(p_{j+1}^0 - p_{j-1}^0) \end{aligned} \tag{2.13}$$

with

$$\begin{aligned} A = \frac{s^2 - 2r(cs+\nu)}{2h^2}, \quad B = \frac{-s+cr}{2h}, \\ C = \frac{s^2 + 2(\Delta t - r)(cs+\nu)}{2h^2}, \quad D = \frac{-s - c(\Delta t - r)}{2h}. \end{aligned} \tag{2.14}$$

Now the approximate solution can be obtained by solving (2.10) for $k = 1, \dots, N$.

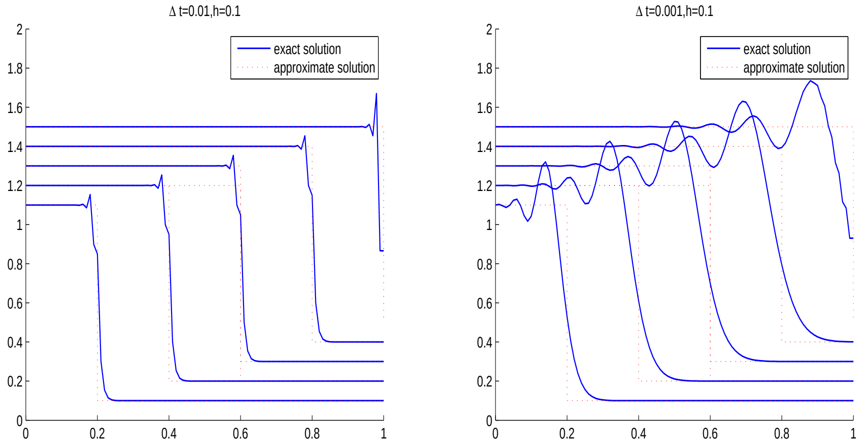

FIGURE 2. Numerical solution of equation (1.1) with $\nu = 0$, and $s = h$, $r = h/c$

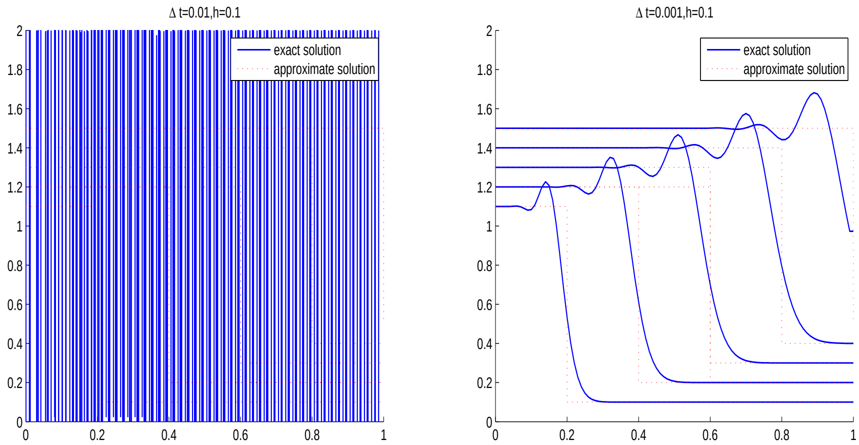

FIGURE 3. Numerical solution of equation (1.1) with $\nu = 0$, and $s = h/2$, $r = h$.

3. NUMERICAL EXAMPLES

In this section we discuss a series of numerical examples according to different choices of the collocation points. We deal with equation (1.1) for $x > 0$ and $c > 0$, where the following discontinuous initial datum is considered:

$$u(x,0) = u_0(x) = \begin{cases} 1, & for\ x = 0, \\ 0, & for\ x > 0. \end{cases} \tag{3.1}$$

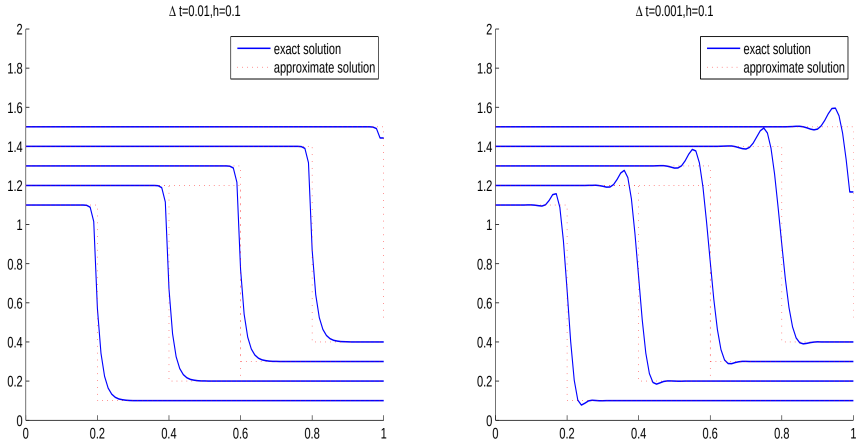

FIGURE 4. Numerical solution of equation (1.1) with $\nu = 0$, and $s = h/2$, $r = 3\Delta t/2$.

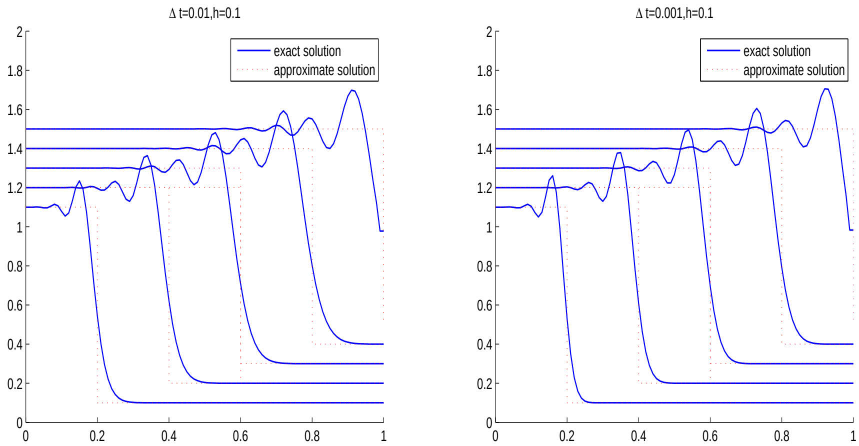

FIGURE 5. Numerical solution of equation (1.1) with $\nu = 0$, and $s = 3h/4$, $r = 5\Delta t/3$.

The boundary condition $u(0,t) = 1, t > 0$, is imposed at the inflow boundary, so that the exact solution is:

$$u(x,t) = \begin{cases} 1, & for\ x \leq ct, \\ 0, & for\ x \geq ct. \end{cases} \tag{3.2}$$

The discontinuity of the solution provides a challenging test for comparing the performances of different schemes. We fix the parameters in the following way: $c = 1$, and $h = \Delta x = 0.01$, while various regimes have been chosen for Δt, In figure (2)-(5) we present the main results of our experiments for two different choices of Δt. i.e.: $\Delta t = h = 0.01$ and $\Delta t = 0.001$, respectively. The dashed lines show the evolution of the exact solution at times $t_k = k/5$, with $k = 1, \ldots, 5$. The plots are slightly shifted upwards in order better to display the graphs. The solid lines show the corresponding approximations.

We use the following methods

Method 1: $\nu = 0$, and $s = h$, $r = h/c$.
Method 2: $\nu = 0$, and $s = h/2$, $r = h$.
Method 3: $\nu = 0$, and $s = h/2$, $r = 3\Delta t/2$.
Method 4: $\nu = 0$, and $s = 3h/4$, $r = 5\Delta t/3$.

to obtain some numerical experiments. The numerical results which are illustrated in figures (2)-(5) show the efficiency of our methods. Figure (3) shows that the method 2 is not convergence when $h = \Delta t = 0.01$. But for smaller Δt (i.e. $\Delta t = .001$) it will be convergence. In comparison with [3], we see that those methods have the same properties reported in [3]. Now the important question is how should we choose s and r such that the introduced method be convergence? Another question is what is the best choice? We try to answer this questions in our next works.

References

[1] D. Calhoun and R. J. Le Veque, A cartesian grid finite-volume method for the advection-diffusion equation in irregular geometries, J. Comp. Phys., 157, 143-801, (2000).
[2] C. A. J. Fletcher, Computational Techniques for Fluid Dynamics, Springer Series in Comput. Phys, I, (1991).
[3] D. Funaro and G. Pontrelli, A general class of finite-difference methods for the linear transport equation, Comm. Math. Sci., Vol. 3, No. 3, 403–423, (2005).
[4] I. Harari, L. P. Franca and S. P. Oliveria, Streamline design of stability parameters for advection-diffusion problems, J. Comp. Phys., 171, 115-131, (2001).
[5] A. F. Hegarty, J. J. H. Miller, E. O Riordan and G. I. Shishkin, Special meshes for finite differences approximations to an advection-diffusion equation with parabolic layers, J. comp.Phys., 117, 47-54, (1995).
[6] W. Hundsdorfer, B. Koren, M. van loon and J. G. Verwer, A positive finite-difference advection scheme, J. Comp. Phys., 117, 35-46, (1995).
[7] K. W. Morton and D. F. Mayers, Numerical Solution of Partial Differential Equations, Cambridge University Press, (1998).
[8] Y. Li, Wavenumber-extended high-order upwind-biased finite-difference schemes for convective scalar transport, J. Comp. Phys., 133, 235-255, (1997).
[9] T. W. H. Sheu, S. K. Wang and S. F. Tsai, Development of a high-resolution scheme for a multi-dimensional advection-diffusion equation, J. Comp. Phys., 144, 1-16, (1998).
[10] B. D. Shizgal, Spectral methods based on nonclassical basis functions: the advection-diffusion equation, Comput. & Fluids, 31, 825-843, (2002).

(Paria Sattari Shajari) Islamic Azad University Shabestar Branch, Tabriz, Iran
E-mail address: pariamath306@yahoo.com

(Karim Ivaz) Faculty of mathematical sciences, University of Tabriz, Tabriz, Iran
E-mail address: ivaz2003@yahoo.com

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 190-198, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

COMPARING THE BOX-JENKINS MODELS BEFORE AND AFTER THE WAVELET FILTERING IN TERMS OF REDUCING THE ORDERS WITH APPLICATION

QAIS MUSTAFA AND TAHA H.A.ALZUBAYDI

Abstract. In this paper , the estimated linear models of Box-Jenkins has been compared from time series observations , before and after wavelet shrinkage filtering and then reducing the order of the estimated model from filtered observations (with preserving the accuracy and suitability of the estimated models) and re-compared with the estimated linear model of original observations , depending on some statistical criteria through taking a practical application of time series and using statistical programs such as Statgraphics , *NCSS* and *MATLAB*.

The results of the paper showed the efficiency of wavelet shrinkage filters in solving the noise problem and obtaining the efficient estimated models , and specifically the wavelet shrinkage filter (*dmey*) with Soft threshold which estimated it's level using the Fixed Form method of filtered observations , and the possibility of obtaining linear models of the filterd observations with lower orders and higher efficiency compared with the corresponding estimated model of original observations

1. Introduction

Time series analysis has big importance in studying the behavior of different phenomena and it represents an important base for many fields like forecasting , therefore we see its applications had expanded for many fields ,for example applied researches , engineering sciences , medicine , physics , soil science , finance and economic...etc. And because the observations of time series are be affected by many unknown natural factors , therefore these observations contains ratio of noise which is defined as unwanted data and usually has small value and high frequencies and as a result for this it contaminates the real observations and cause difficulty in the process of analyzing the time series observations like diagnostic , estimation and forecasting.

The process of reducing the noise or removed before analysis the time series is very important in order to obtain more accurate and reliable results when building models .The wavelet shrinkage technique consists of wavelets with threshold is a strong mathematical approach to remove most of the noise while retaining the maximum amount of energy data that represents the real observation.

Wavelet transform have been used in many fields.It has been observed that the accuracy of the forecasting can be improved through using wavelet transforms.Rumaih M.and Mohammad A.,in 2002[1]used Saudi stock index to show that wavelet transform is better than the other forecasting technique in predicting the de-noising of the financial time series , and this is done after making a comparison

Key words and phrases. Box-Jenkins, time series, wavelet shrinkage, filters, Fixed Form.

with several forecasting models.Alwadi S. ,Mohd Tahir Ismail, Alkhahazaleh and Samsul Ariffin Addul Karim,in 2011[2] used the wavelet transform to decompose the return Amman stock market in to a set of better behaved series data and explained that after making a comparison framework, the wavelet $ARIMA$ model is better than the $ARIMA$ model of the original data and gives more accurate results and also gives data with more stable in variance , mean and no outliers.

The basic idea of this paper is to see the method of the wavelet transforms in general and wavelet shrinkage in time series analysis in particular . Moreover ,the study try to show the advantages in modeling and forecasting through de-noising the series using wavelet shrinkage and try to get the ability of lowering the order of the estimated model using wavelet shrinkage. Batch of chemical process data[3] was used for analysis.

Two different methods were considered here,as showed in Figure1.In the first method,the batch of chemical process data is modeled using Box - Jenkins method. Then some forecasting criteria were computed.In second method the technique of wavelet shrinkage was used through filtering the time series data using a set of wavelet filters.The de-noised series are modeled ,as in the first method ,then the forecasting criteria were computed again.The forecasting criteria were evaluated and compared to those of the first method.

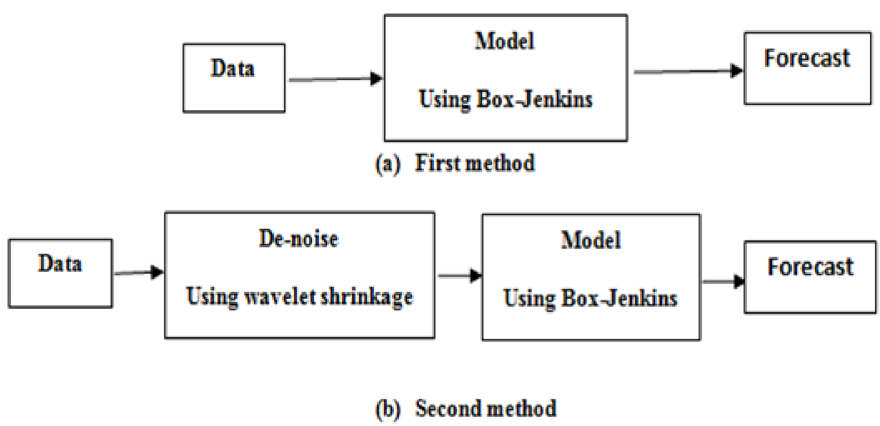

Figure 1. Modeling Methods

The remainder of this paper is structured as follows.In section 2 we review the simple brief introduction to wavelet transform.Section 3 gives time series de-noising using wavelet shrinkage .Section 4 deals with main results before and after reducing the order of the model using wavelet shrinkage and gives discussion.And finally in section 5 conclusions are presented.

2. The Wavelet Transform

The mathematical transformations are applied to a signal to get additional information that is not found in the time domain representation of that signal. First we consider the Fourier transform ,which decomposes signals in to sum of periodic bases of finite length and can transform the signal from time domain to the frequency domain and vice versa. It is defined mathematically as follows:

$$X\left(f\right)=\int_{-\infty}^{\infty}x\left(t\right).e^{-jwt}dt \tag{1}$$

Where $X(f)$ is the Fourier transform of the signal $x(t)$ .The problem here is that the Fourier transform becomes non active for signal changing over time , because it provides for us the information of frequency content. In other words the $x(f)$ Is not a function of time .For this reason the Fourier Transform extended to so called Short-Time Fourier transform($STFT$).It is defined as:

$$STFT_X^{(W)}(t^{'},f)=\int_{t}^{\infty}[x(t).w^{*}(t-t^{'})].e^{-2\pi ft}dt \tag{2}$$

Where $t^{'}$ is the shift factor , $w(t)$ is the window function , and * is the complex conjugate. The $STFT$ gives us a compromise of sorts between time and frequency information .The accuracy is limited by the size and shape of the window. For example ,using many time intervals would give good time resolution but the very short time of each window would not give us good frequency resolution, specially for lower frequency signals [7].The frequency component of a signal at a particular time instant cannot exactly be determined. This comes directly from the Heisenberg's uncertainty principle , which states that the momentum and position of moving particle cannot exactly be determined.

This shortcoming was overcome by the development of the wavelet transformation. Wavelet transforms allow us variable-size windows. We can use long time intervals for more precise low frequency information and shorter intervals (giving us more precise time information)for the higher frequencies. Figure 2 shows the partitioning of the time-frequency plane by different techniques[8].

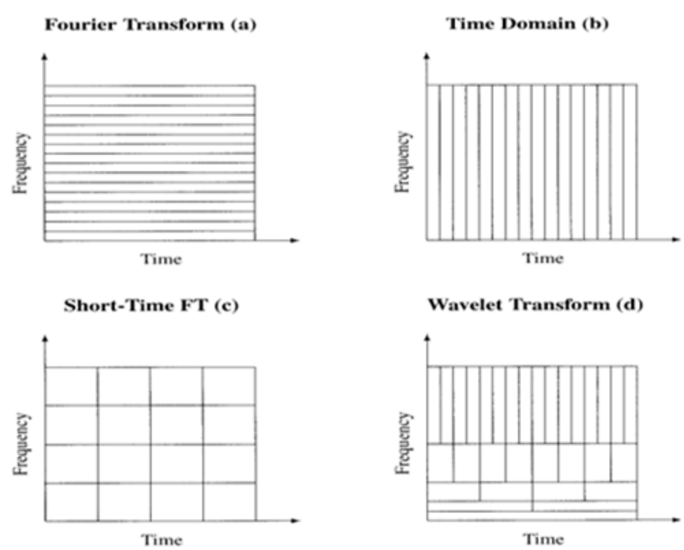

Figure 2. The partitioning of the time – frequency plane by different techniques

The wavelet transform is defined as follows:

$$\Psi_{*}^{\psi}(x,s)=\int x(t).\psi_{\tau,s}^{*}(t)dt \tag{3}$$

where:

$$\psi_{\tau,s}^{*}=\frac{1}{\sqrt{s}}\psi\left(\frac{t-\tau}{s}\right) \tag{4}$$

Where s is the scale variable and τ is the translation variable. When substituting this description in (2.3) gives the definition of the continuous wavelet transform CWT:

$$CWT_{x}^{\psi}(\tau,s)=\frac{1}{\sqrt{s}}\int x(t).\psi\left(\frac{t-\tau}{s}\right)dt \tag{5}$$

As seen in the above equation ,the transformed signal is a function of two variables , τ and s, the translation and scale parameters respectively. The translation τ is proportional to time information and the scale s, is proportional to the inverse of the frequency information.To find the constituent wavelets of the signal , the coefficients should be multiplied by the relevant version of the mother wavelet ψ[10].

Discrete wavelet transform DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information. DWT employs two sets of functions , called scaling functions and wavelet functions , which are associated with low pass $h[n]$ and high pass $g[n]$ filters . The original signal is passed through the both half band filters. After filtering half of the sample can be eliminated according to the Nyquist's rule (i.e.; half of the samples are redundant). The signal can therefore be sub-sampled by 2 ,simply by discarding every other sample. This process can be repeated for further decomposition [11].

Figure 3 represents the wavelet decomposition tree of DWT coefficients for three levels decomposition , where s is the original sequence , cA and cD are approximation and details(i.e.; low pass and high pass analysis filters)[10].

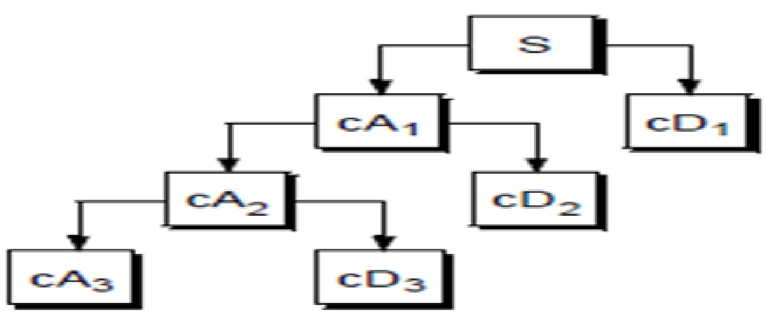

Figure 3. Wavelet decomposition tree of the *DWT* coefficients for three levels

Note that the original signal can be synthesized by the low and high frequencies sequences by assembly sequence for each assembly parts resulting from the previous parts (approximation parts and detailed) started from the last stage of analysis.

3. Time series analysis using wavelet shrinkage

Statistical wavelet methods provide a powerful tool to recover the original signal from noisy observations under general assumptions , we refer the reader to [5] for more details. The generic methodology is called wavelet shrinkage or wavelet thresholding. The underlying model for the noisy series can be expressed as follows:

$$X(t) = s(t) + n(t) \tag{6}$$

Where $n(t)$ represents the independent and identically distributed random variables with zero mean and finite variance. The objective is to suppress the noise part of the series $x(t)$ and recover the clean part $s(t)$ .The procedure can be summarized as three main steps[5,12] :

1-The discrete wavelet transform DWT is computed from the data.

2-Significantly large coefficients in the DWT are kept ,others are shrinked.

3-The inverse DWT is applied to the shrunken set of coefficients.

The three steps above can be summarize in figure 4:

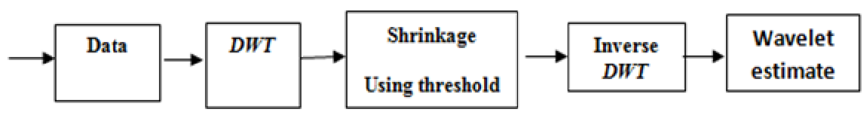

Figure 4. Wavelet shrinkage chart

For each level , we will have a threshold . The Fixed Form threshold (i.e.; Universal threshold) technique is considered and is given by the following formula:

$$\delta^{(FT)} = \hat{\sigma}(MAD)\sqrt{2log(N)} \tag{7}$$

Where (N) is the number of wavelet coefficients in specified level , $\hat{\sigma}(MAD)$ is the estimate of the noise standard deviation and can be obtained by applying a median absolute deviation (MAD) estimator to the $\frac{N}{2}$ wavelet coefficients at the first level of decomposition , incorporating a scale factor equal to (0.6745) [6]. After estimating the threshold of a specified level ,wavelet coefficient of that level are either hard or soft threshold. Hard thresholding represents the keep or kill (i.e.; wavelet coefficient is less than the threshold it would be put to zero otherwise it stays without change) . The soft thresolding shrinks all non zero coefficients towards zero , which gives a smooth de-noising .The soft threshold de-noising function formula can expressed as follows :

$$d_{J,k} = \begin{cases} d_{J,k}, & |d_{J,k}| \geq \delta, \\ 0, & |d_{J,k}| < 0 \end{cases} \tag{8}$$

Where $d_{J,k}$ denotes the coefficient of transformation and δ is the threshold.It was shown in [4] that soft thresholding has smaller variance than hard thresholding , therefore here we only consider soft thresholding for modeling and forecasting the yields data of batch chemical in this paper.

4. Main results and discussion

In this section , an application was considered so as to show first the ability of wavelet shrinkage to reduce the noise from original data , and second to get a model with a lower order and higher efficiency compared with the original. Figure 5 shows five levels multiresolution wavelet analysis using $Haar$ wavelet for the yields from a batch chemical process for 65 consecutive observations , where s is the signal and it is the sum of its approximation and of its fine details , a_5 is approximation at level 5 and d_5, d_4, d_3, d_2, d_1 are the details at level 5,4,3,2 and 1.

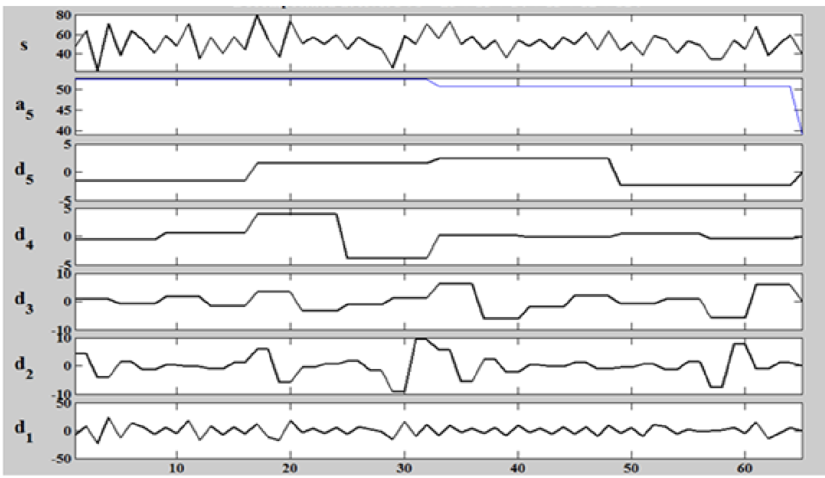

Figure 5. Five multiresolution wavelet analysis using *Haar* wavelet for the yields from a batch chemical process

Now , we will use the two different methods mentioned in the introduction as follows: Method (1): The yields of batch chemical process was modeled as $ARMA(2, 0)$, the parameters $(p = 2, q = 0)$ were selected after careful modeling and fitting (Statgraphics software was used for modeling). The performance measures (i.e.; the forecasting criteria) used in the analysis are Root Mean Square Error $(RMSE)$, Mean Absolute Error (MAE) and Akaike's Information Criterion(AIC) are computed as the following [9,13,14]:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} a_t^2}{n-k}} \tag{9}$$

$$MAE = \frac{\sum_{t=1}^{n} |a_t|}{n} \tag{10}$$

$$AIC = ln\sigma_a^2 + \frac{2k}{n} \tag{11}$$

Where a_t represents the difference between actual and forecasted value , (k) is the number of estimated parameters of the model and (n) is the sample size. Method(2):The yields from a batch chemical process noise model was considered. The data was de-noised using wavelet shrinkage technique mentioned in section (3) (using $MATLAB$ software , version 2008) with four different wavelet families . It is important here to say that after many experimental trials with many wavelets families , it was found that these wavelets better than others in terms of de-noising the batch data, and they are [7]:

1- *Haar* wavelet with five multiresolution levels.
2- *Daubechies* wavelet of order (5) and five multiresolution levels.
3- *Coiflets* wavelet of order (3) and five multiresolution levels.
4- *DiscreteMeyer* (*dmey*) wavelet with five multiresolution levels.

For each selected wavelet, the batch data was first analyzed for five multiresolution levels and de-noised using the Fixed Form threshold . After de-noising, the series were modeled using Box-Jenkins method and forecasting criteria were computed and compared with those in the first method mentioned before. Table (1) shows the performance measures (i.e.; forecasting criteria) for $ARMA(2,0)$ model of the original and de-noised data using the Fixed Form thresholding. It is clear from the table that when the $ARMA$ model was used , wavelet de-noising reduces forecasting errors depending on forecasting criteria. The reduction is in highest level when the batch data was Fixed Form thresholded using Discrete Meyer Wavelet (note from the table 1, the best reduction comparing with all other forecasting criteria used here).

Filtered observations using wavelet shrinkage filters were modeled once again for each wavelet filter, but this time as $ARMA(1,0)$ when rebuilding the model instead of $ARMA(2,0)$ after careful modeling and without effect on the suitable and accuracy of the model , then compared with the original model $ARMA(2,0)$. Table (2) shows the performance measures for $ARMA(2,0)$ model of the original and de-noised data when rebuilding $ARMA(1,0)$ model ,then using the Fixed Form thresholding. It is clear from the table that when the $ARMA$ model was used , wavelet de-noising reduces forecasting errors depending on forecasting criteria. Once again the reduction is in highest level when the batch data was Fixed Form thresholded using Discrete Meyer Wavelet (note from the table 2, the best reduction comparing with all other forecasting criteria used here).

Table 1: $RMSE$, MAE and AIC for forecasting model of the batch data comparing with the de-noised data using Fixed Form thresholding.

Method	Kind	*RMSE*	*MAE*	*AIC*
First method (Box-Jenkins) Original data	*ARMA(2,0)*	10.5984	7.7854	4.8137
Second method (Fixed Form) De-noised data	*Haar*	8.7989	6.3515	4.4415
	Daubechies (5)	8.2375	6.0619	4.3097
	Coiflet (3)	8.2192	6.0618	4.3053
	Discrete Meyer(*dmey*)	7.9810	5.9255	4.2464

Table 2: $RMSE$, MAE and AIC for forecasting model $ARMA(2,0)$ of the batch data comparing with the de-noised data as $ARMA(1,0)$ using Fixed Form thresholding.

Method	**Kind**	***RMSE***	***MAE***	***AIC***
First method (Box-Jenkins) **Original data**	***ARMA(2,0)***	**10.5984**	**7.7854**	**4.8137**
Second method (Fixed Form) **De-noised data** **as *ARMA(1,0)***	***Haar***	**8.8602**	**6.5585**	**4.4246**
	***Daubechies* (5)**	**8.3269**	**6.3383**	**4.3000**
	***Coiflet* (3)**	**8.2022**	**6.1242**	**4.2703**
	Discrete Meyer(*dmey*)	**7.9555**	**5.9805**	**4.2092**

5. Conclusions

1- More information could be obtained from a series When using Wavelet Shrinkage technique , specially the wavelet filter Discrete Meyer , forecasting errors can be reduced depending on some performance measures .

2- The ability of lowering the order of the estimated model of batch data for filtered observations using Wavelet Shrinkage as a result of reducing the noise with keeping the efficiency and suitability of that model.

References

[1] M.Alrumaih and A.Al-Fawzan ,Time series forecasting using wavelet denoising an application to Saudi stocks index,*J.King Saudi Univ.*,14,221-234 (2002).

[2] S.Alwadi,Mohd Tahir Ismail,M.H.Alkhahazaleh and Samsul Addul Karim ,Selecting Wavelet Transforms Model in Forecasting Financial Time Series Data Based on ARIMA Model,*Applied Mathematical Sciences*,5,315-326 (2011).

[3] G.E.P. Box,G.M.T. Jenkins and G.C.Reinsel ,*Time series analysis: Forecasting and control*, 3rd,ed., Prentice - Hall, Inc, Upper Saddle River , N.J. , USA ,1994

[4] G. Bruce Andrew and Hong-ye Gao,*Understanding waveshrink :Variance bias estimation*,Biometrica ,printed in great Britain ,1996.

[5] C.S.Burrus, R.A.Gopinath, and H.Guo,*Introduction to wavelets and wavelet transforms*,Printice - Hall,Inc,Upper Saddle River , N.J.,USA,1998.

[6] L.L Cascio ,*Wavelet analysis and denoising:New tools for economists*, Queen Marry press,university of London , No.600, 2007 ,PP.22-26.

[7] D. L. Fugal ,*Conceptual wavelets in digital signal processing*, Space and Signal technical publishing , San Diego , USA , 2009.

[8] R. Gencay , F.Selcuk and B.Whitcher ,*An introduction to wavelets and other filtering methods in finance and economics*,Academic press , 2001.

[9] S.Makridakis, S.C.Wheelwright and R. J Hyndman. *Forecasting methods and applications*,3rd ed., John Wiley and Sons , Inc , NYC ,USA , 1998 .

[10] M.Misiti , Y.Misiti , G.Oppenheim and J-M .Poggi ,*Wavelet toolbox - user's guide*,1st version,the mathworks , Inc , Natick , MA,USA , 1996.

[11] R.Polikar ,*The wavelet tutorial* , 2nd ed.,Rawan University,College of Engineering ,N.J.,USA, 2006.

[12] M. Raimondo ,*Wavelet shrinkage via peaks over threshold*, a paper submitted to the school of Mathematics and Statistics ,University of Sydney Australia, PP.1-4,(2002).

[13] R.S.Tsay ,*Analysis of financial time series* , John Wiley and Sons , Inc , USA , 2002 .

[14] R.Yaffee ,*Introduction to time series analysis and forecasting with applications of SAS and SPSS*,Academic Press , Inc, NYC , USA , 1999.

Zakho Technical Institute, Duhok, Iraq
E-mail address: qzakholy1@yahoo.com

University of Salahaddin, Erbil, Iraq
E-mail address: tahaalaa@yahoo.com

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 199-210, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

REDUCED BIAS OF THE MEAN FOR A HEAVY TAILED DISTRIBUTION

RASSOUL ABDELAZIZ

Abstract. For the estimation of the mean for a heavy tailed distribution with index $1/2 < \gamma < 1$, The semi-parametric estimation of mean depends not only on the estimation of the tail index or extreme value index γ, the primary parameter of extreme events, but also on the adequate estimation of a scale first order parameter. Recently, apart from new classes of reduced-bias estimators for $\gamma > 0$, new classes of the scale first order parameter have been introduced in the literature. Their use in mean estimation enables us to introduce new classes of asymptotically unbiased mean estimators. The asymptotic distributional behavior of the proposed estimators of μ is derived, under a second-order framework, and their finite sample properties are also obtained through Monte Carlo simulation techniques.

1. Introduction and preliminaries

A model F is heavy-tailed whenever, the tail function $\overline{F} = 1 - F$ is regularly varying functions with a negative index of regular variation equal to $-1/\gamma$ $(\gamma > 0)$ $(\overline{F} \in \mathcal{RV}_{-1/\gamma})$, or whenever the quantile function

$$\mathbf{U}(t) = F^{\leftarrow}(1 - 1/t) = \inf\{x : F(x) > 1 - 1/t\},$$

where $F^{\leftarrow}$ represented the inverse generalized of the df F, is of regular varying with an index γ $(\mathbf{U} \in \mathcal{RV}_{\gamma})$, i.e., for every $x > 0$

$$\lim_{t\to\infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma} \Leftrightarrow \lim_{t\to\infty} \frac{\mathbf{U}(tx)}{\mathbf{U}(t)} = x^{\gamma}, \tag{1.1}$$

(see, e.g., de Haan and Ferreira, 2006, page 19).

Such cdf's constitute a major subclass of the family of heavy-tailed distributions. Heavy tailed distributions have been applied in many different areas such as population size (Zipf (1949)), random graph (Reittu and Norros (2004)), internet traffic (Resnick (1997a)), hydrology (Katz et al. (2002)), finance (Danielsson and de Vries (1997)).

Suppose that $X_1, ..., X_n$ are i.i.d. random variables with cdf F which has regularly varying tails with index $\gamma \in (1/2, 1)$, then (1.1) is equivalent to the statement that F has infinite second moment.

Key words and phrases. Heavy tails, index, high quantiles, mean, trimmed mean, reduced bias.

2010 *AMS Math. Subject Classification.* Primary 40A05, 40A25; Secondary 45G05.

1

In this paper we interested by estimate of the mean for a heavy tailed distribution, that is

$$\mu = \mathbf{E}(X) = \int_0^{+\infty} x dF(x),$$

this may be rewrite in term of the quantile, as follow

$$\mu = \int_0^1 \mathbf{Q}(1-s)ds.$$

where $\mathbf{Q}(s) := \inf\{x : F(x) > s\}$; $0 < s < 1$, denotes the quantile function of the df F. In this case, to obtain a consistent estimator of the mean for any $\gamma \in (1/2, 1)$, Peng (2001) partitioned the mean μ into

$$\begin{aligned} \mu &= \int_0^{k/n} \mathbf{Q}(1-s)ds + \int_{k/n}^1 \mathbf{Q}(1-s)ds \\ &= \mu_n^{(1)} + \mu_n^{(2)}, \end{aligned}$$

where $k = k(n)$ is an intermediate integer sequence satisfying the condition

$$k \to \infty \text{ and } k/n \to 0 \text{ as } n \to \infty. \tag{1.2}$$

Then $\mu_n^{(1)}$ and $\mu_n^{(2)}$ are estimated separately. One obvious estimator for $\mu_n^{(1)}$ is the trimmed mean

$$\widehat{\mu}_n^{(1)} = \frac{1}{n}\sum_{i=1}^{n-k} X_{i,n}, \tag{1.3}$$

which has been studied extensively (see Csörgö et al., 1986a,b, GriEn and Pruitt,1989).

The Extreme Value Theory is used to estimate $\mu_n^{(2)}$, the mean for the right tail as follows: For a small value of p, we want to estimate the quantil $\boldsymbol{\chi}_{1-p}$, such that $F(\boldsymbol{\chi}_{1-p}) = 1-p$. More specifically, we want to estimate

$$\boldsymbol{\chi}_{1-p} = \mathbf{U}(1/p), \ p = p_n \to 0, \ np_n \to K \in [0,1], \tag{1.4}$$

and we shall assume to be working in Hall's class of models (Hall 1982; Hall and Welsh, 1985), where there exist $\gamma > 0$, $\rho < 0, C > 0$ and $\beta \neq 0$ such that

$$\mathbf{U}(t) = Ct^{\gamma}(1 + \mathbf{A}(t)/\rho + o(t^{\rho})), \text{ as } t \to \infty. \tag{1.5}$$

Since, from (1.4) and (1.5), we have

$$\boldsymbol{\chi}_{1-p} = \mathbf{U}(1/p) \sim Cp^{-\gamma} \text{ as } p \to 0.$$

An obvious estimator of $\boldsymbol{\chi}_{1-p}$ is $\widehat{C}p^{-\widehat{\gamma}}$, with $\widehat{C}$ and $\widehat{\gamma}$ any consistent estimators of C and γ, respectively.
Let $X_{1,n}, ..., X_{n,n}$.denote the order statistics of $X_1, ..., X_n$. Denoting Y a standard Pareto model, i.e., a model such that $F_Y(y) = 1 - 1/y$,

$y > 1$, the use of the universal uniform transformation enables us to write $X_{n-k,n} \overset{d}{=} \mathbf{U}(Y_{n-k,n})$. Next, since $Y_{n-k,n} \overset{p}{\sim} (n/k)$ for intermediate k satisfiying (1.2), we get

$$X_{n-k,n} \overset{p}{\sim} CY_{n-k,n}^{\gamma} \overset{p}{\sim} C(n/k)^{\gamma}, \text{ as } n \to \infty.$$

Consequently, an obvious estimator of C, proposed by Hall and Welsh (1985), is

$$C_{\widehat{\gamma}}(k) = X_{n-k,n}(k/n)^{\widehat{\gamma}}, \tag{1.6}$$

then

$$\mathbf{Q}_{\widehat{\gamma}}^{(p)}(k) = X_{n-k,n}(k/np)^{\widehat{\gamma}}. \tag{1.7}$$

is the obvious quantile-estimator at the level p (Weissman, 1978). Peng (2001) using this estimator for estimate $\mu_n^{(2)}$, as follows

$$\widehat{\mu}_n^{(2)} = \frac{k}{n} X_{n-k,n} \frac{1}{1-\widehat{\gamma}_n}, \tag{1.8}$$

where $\widehat{\gamma}_n$ is Hill estimator of γ, defind as follows

$$\widehat{\gamma}_n = \frac{1}{k}\sum_{i=1}^{k} \log(X_{n,n-i+1}) - \log(X_{n,n-k+1}). \tag{1.9}$$

Finally, the result of Peng is as follows

$$\widehat{\mu}_n = \frac{1}{n}\sum_{i=1}^{n-k} X_{i,n} + \frac{k}{n} X_{n-k,n} \frac{1}{1-\widehat{\gamma}_n}. \tag{1.10}$$

The details on asymptotic distributions for the estimators μ can be found in Peng (2001). The Hill estimator of the index plays a central role in modelling the tail distribution.

The consistency property of the Hill's estimator is given in Mason [1982]. The asymptotic normality of the Hill's estimator has been demonstrated by many authors, see for example, de Haan and Resnick [1980], Csörgö and Mason [1985], Häusler and Teugels [1985], and Beirlant and Teugels [1987]. The detailed explanation of mathematical properties of the Hill's estimator can be found in Beirlant et al. [2004]. To be able to conduct reliable inference about tail properties of the heavy tailed distribution it is important to determine the optimal threshold k. A number of methods for estimation of optimal k were proposed in the literature (Dekkers and de Haan [1993], Hall and Welsh [1985], Beirlant et al. [1996a], Beirlant et al. [1996b], Draisma et al. [1999]). Most of the proposed methods are based on minimising the asymptotic mean squared error (AMSE) of the tail index estimate for a given sample size and corresponding fraction of extreme value statistics.

As Peng's estimator $\widehat{\mu}_n$ is based on Weissman's estimator $\mathbf{Q}_{\widehat{\gamma}}^{(p)}(k)$, so known to be largely biased.

Our main aim in these paper is essentially to present a new reduced bias estimators for the mean μ, valid in a wide sub-class of Hall's class of models (Hall, 1982), and in the lines of both Matthys et al. (2004) and Gomes and Figueiredo (2006), i.e., based on a direct accommodation of the bias for high quantiles and an adequate reduced-bias estimation of the tail index , respectively. The key feature of these new estimators exists in the fact that the estimation of the second-order parameters in the respective bias-terms, described in Section 2 of this paper, is performed at a level k_1 of a larger order than the level k at which the parameters γ and $\boldsymbol{\chi}_p$ are classically estimated. Doing this, we are able to guarantee a mean squared error smaller than of the classical estimators for all levels k and for mean estimation. In addition, the asymptotic behavior of the new estimator of the mean will be derived under appropriate higher order conditions, in Sections 3 and 4, respectively. The simulation study in Section 5 will enable us to obtain some of the features of these new estimators for finite samples, the proof of the main result are postponed to section 6.

2. Second order conditions

In order to derive the asymptotic non-degenerate behaviour of semi-parametric estimators of extreme events parameters, we need more than the first order condition in (1.3), a second order condition on F (see de Haan and Stadtmüller, 1996), assuming that there is a function $\mathbf{A}$ with constant sign near infinity, such that

$$\lim_{s\to\infty}\frac{\mathbf{U}(tx)/\mathbf{U}(t)-x^{\gamma}}{\mathbf{A}(s)}=x^{\gamma}\frac{x^{\rho}-1}{\rho},\ \text{for any } x>0, \tag{2.1}$$

where $\rho\leq 0$ is the shape second order parameter. If $\rho=0$, interpret $\frac{x^{\rho/\gamma}-1}{\rho/\gamma}$ as $\log x$. Or equivalently,

$$\lim_{s\to\infty}\frac{\ln\mathbf{U}(sx)-\ln\mathbf{U}(s)-\gamma\ln x}{\mathbf{A}(s)}=\frac{x^{\rho}-1}{\rho},\ \text{for any } x>0. \tag{2.2}$$

A typical condition for heavy-tailed models, which holds for the models in (1.5), with

$$\mathbf{A}(t)=\gamma\beta t^{\rho} \text{ with } \gamma>0,\rho<0 \text{ and } \beta\neq 0. \tag{2.3}$$

Under the second order framework in (2.1) or in (2.2), and for intermediate k, i.e., whenever (1.2) holds, we may guarantee the asymptotic normality of the Hill estimator $\widehat{\gamma}_n$, for an adequate k. Indeed, we may write (de Haan and Peng, 1998),

$$\widehat{\gamma}_n\overset{d}{=}\gamma+\frac{\gamma}{\sqrt{k}}Z_k+\frac{\mathbf{A}(n/k)}{(1-\rho)}(1+o_P(1)), \tag{2.4}$$

with $Z_k = \sqrt{k}\left(\sum_{i=1}^{k} E_i/k - 1\right)$, and $\{E_i\}$ i.i.d. standard exponential r.v.'s.

Consequently, if we choose k such that $\sqrt{k}\mathbf{A}(n/k) \to \lambda \neq 0$, finite, as $n \to \infty$, then $\sqrt{k}\,(\widehat{\gamma}_n - \gamma)$ is asymptotically normal, with variance equal to γ^2 and a non-null bias given by $\lambda/\,(1-\rho)$. Most of the times, this type of estimates exhibits a strong bias for moderate k and sample paths with very short stability regions around the target value. This has recently led researchers to consider the possibility of dealing with the bias term in an appropriate way, building new estimators, $\widehat{\gamma}_R(k)$ say, the so-called second order reduced-bias estimators discussed by Peng (1998), Beirlant et al. (1999), Feuerverger and Hall (1999), Gomes et al. (2000), among others. Then, for k intermediate, satisfying (1.2), such that (2.1) holds, and under the second order framework in (1.10), we may write, with Z_k^R an asymptotically standard normal r.v.,

$$\widehat{\gamma}_R(k) \stackrel{d}{=} \gamma + \frac{\gamma\sigma_R}{\sqrt{k}} Z_k^R + o_P\left(\mathbf{A}(n/k)\right), \tag{2.5}$$

where $\sigma_R > 0$ and $\mathbf{A}$ is again the function in (1.10). Consequently, the sequence of r.v.'s, $\sqrt{k}\,(\widehat{\gamma}_R(k) - \gamma)$ is asymptotically normal with variance equal to $(\gamma\sigma_R)^2$ and a null mean value even when $\sqrt{k}\mathbf{A}\,(n/k) \to \lambda \neq 0$, finite, as $n \to \infty$, possibly at the expense of an asymptotic variance $(\gamma\sigma_R)^2 > \gamma^2$. Gomes and Figueiredo (2006) suggest the use, in (2.1), of reduced-bias tail index estimators, like the ones in Gomes and Martins (2001, 2002) and Gomes et al. (2004), all with $\sigma_R > 1$ in (2.5), being then able to reduce also the dominant component of the classical quantile estimator's asymptotic bias.

More recently, Caeiro et al. (2005), Gomes et al. (2007a, 2007b) consider new classes of tail index estimators, for which (2.5) holds with $\sigma_R = 1$ at least for values k such that $\sqrt{k}\mathbf{A}\,(n/k) \to \lambda$, finite. These classes are dependent on $\left(\widehat{\beta}, \widehat{\rho}\right)$, an adequate consistent estimator of the vector of second order parameters (β, ρ) in (2.3). The influence of these tail index estimators in quantile estimation has been studied by Gomes and Pestana (2007) and Beirlant et al. (2006).

3. Reduced bias of index and extreme quantile

The second order reduced-bias extreme value index estimator introduced in Caeiro et al. (2005), is given by

$$\widehat{\gamma}_R\,(k) = \widehat{\gamma}_{R,\widehat{\rho},\widehat{\beta}}\,(k) = \widehat{\gamma}_n\left(1 - \frac{\widehat{\beta}}{1-\widehat{\rho}}\left(\frac{n}{k}\right)^{\widehat{\rho}}\right), \tag{3.1}$$

for adequate consistent estimators $\widehat{\rho}$ and $\widehat{\beta}$ of the second order parameters ρ and β respectively.

Also recently, new estimators of C have been proposed (Caeiro, 2006), where, instead of $X_{n-k,n}$ alone, a spacing $X_{n-[\theta k],n} - X_{n-k,n}$, $0 < \theta < 1$, is considered. More specifically, we may replace $\widehat{C}$ in (1.6) by

$$\widetilde{C}_{\widehat{\gamma}_R}(k,\theta) = \frac{X_{n-[k\theta]:n} - X_{n-k:n}}{\theta^{\widehat{\gamma}_R} - 1}\left(\frac{k}{n}\right)^{\widehat{\gamma}_R} \tag{3.2}$$

where $0 < \theta < 1$.is a tuning parameter and $\widehat{\gamma}_R = \widehat{\gamma}_R(k)$ is a second order reduced-bias extreme value index estimator. Similarly to the way developed by Caeiro et al. (2005) for the extreme value index estimation, Caeiro (2006) has worked out the main dominant component of the asymptotic bias of $\widehat{C}_{\widehat{\gamma}_R}(k,\theta)$. With the parametrization $\mathbf{A}(t) = \gamma\beta t^{\rho}$, already given in (2.3), such a component is given by $C \times (1 - \mathcal{B}_\theta(\gamma,\rho,\beta))$, where

$$\mathcal{B}_\theta(\gamma,\rho,\beta) = \frac{\theta^{-(\gamma+\rho)} - 1}{\theta^{-\gamma} - 1} \times \frac{\gamma\beta\,(n/k)^{\rho}}{\rho}. \tag{3.3}$$

It is thus sensible to consider the semi-parametric $C-$estimator,

$$\overline{C}_{\widehat{\gamma}_R}(k,\theta) = \frac{X_{n-[k\theta]:n} - X_{n-k:n}}{\theta^{\widehat{\gamma}_R} - 1}\left(\frac{k}{n}\right)^{\widehat{\gamma}_R} \times \left(1 - \mathcal{B}_\theta\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)\right). \tag{3.4}$$

We shall here consider, for $\theta = 1/2$, the associated quantile estimator $\overline{Q}^{p}_{\widehat{\gamma}_R}(k) = \overline{Q}^{p}_{\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}}(k)$ with

$$\overline{\mathbf{Q}}^{p}_{\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}}(k) = \frac{X_{n-[k/2]:n} - X_{n-k:n}}{2^{\widehat{\gamma}_R} - 1}\left(\frac{k}{np}\right)^{\widehat{\gamma}_R} \times \left(1 - \mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)\right). \tag{3.5}$$

3.1. **Estimators of the shape second order parameter ρ.** We shall consider here particular members of the class of estimators of the second order parameter proposed by Fraga Alves et al. (2003). Such a class of estimators may be parameterized by a tuning real parameter $\tau \in \mathbb{R}$ (Caeiro and Gomes, 2004). These -estimators depend on the statistics

$$T_n^{(\tau)}(k) = \begin{cases} \dfrac{\left(M_n^{(1)}(k)\right)^{\tau} - \left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \left(M_n^{(3)}(k)/6\right)^{\tau/3}}, & \text{if } \tau \neq 0 \\ \dfrac{\ln\left(M_n^{(1)}(k)\right)^{\tau} - \frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right)^{\tau/2}}{\frac{1}{2}\ln\left(M_n^{(2)}(k)/2\right)^{\tau/2} - \frac{1}{3}\ln\left(M_n^{(3)}(k)/6\right)^{\tau/3}}, & \text{if } \tau = 0 \end{cases}, \tag{3.6}$$

converge towards $3(1-\rho)/(3-\rho)$, independently of the tuning parameter, whenever the second order condition (10) holds and k is such that (1.2) holds and $\sqrt{k}\ A(n/k) \to \infty$, as $n \to \infty$, where

$$M_n^{(r)}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\log(X_{n,n-i+1}) - \log(X_{n,n-k+1})\right)^{r}. \tag{3.7}$$

The ρ-estimators considered have the functional expression,

$$\widehat{\rho}_n^{(\tau)}(k) = -\min\left(0, \frac{3T_n^{(\tau)}(k)-1}{T_n^{(\tau)}(k)-3}\right). \tag{3.8}$$

Remark 3.1. *The theoretical and simulated results in Fraga Alves et al. (2003), together with the use of these estimators in different reduced-bias statistics, has led us to advise in practice the estimation of ρ through the estimator in (3.8), computed at any value*

$$k_1 := \left[n^{1-\varepsilon}\right], \varepsilon > 0. \tag{3.9}$$

The value k_1 in (3.9) is not necessarily optimal but, for a large class of heavy-tailed models, it enables us to guarantee that, with $\widehat{\rho} = \widehat{\rho}_n^{(\tau)}(k), \widehat{\rho}-\rho = o_p\left(1/\ln n\right),$ as $n \to \infty$. The choice of the tuning parameter $\tau = 0$ for the region $\rho \in [-1, 0)$ and $\tau = 1$ for the region $\rho \in (-\infty, -1)$ is a sensible one.

3.2. **Estimators of the scale second order parameter β.** For the estimation of β we shall here consider the estimator developed in Gomes and Martins (2002), with the functional expression,

$$\widehat{\beta}_{\widehat{\rho}}(k) = \left(\frac{k}{n}\right)^{\widehat{\rho}} \frac{\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}} N_n^{(1)}(k) - N_n^{(1-\widehat{\rho})}(k)}{\frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\widehat{\rho}} N_n^{(1-\widehat{\rho})}(k) - N_n^{(1-2\widehat{\rho})}(k)}, \tag{3.10}$$

where

$$N_n^{(\alpha)}(k) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{\alpha-1} U_i,$$

with $\widehat{\rho} = \widehat{\rho}_n^{(\tau)}(k)$ and

$$U_i = i\left(\log X_{n-i+1,n} - \log X_{n-k,n}\right),$$

4. Reduced bias of the mean and the main result

Substituting $\overline{Q}^p_{\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}}(k)$ for Q in expression of $\mu_n^{(2)}$, and after integration, we obtain

$$\frac{X_{n-[k\theta]:n} - X_{n-k:n}}{\theta^{\widehat{\gamma}_R} - 1}\left(\frac{k}{n}\right)^{\widehat{\gamma}_R} \times \left(1 - \mathcal{B}_\theta\left(\widehat{\gamma}_R, \widehat{\rho}, \widehat{\beta}\right)\right)$$

$$\overline{\mu}_n^{(2)} = \left(\frac{k}{n}\right)\frac{\left(X_{n-[k/2]:n} - X_{n-k:n}\right)}{\left(2^{\widehat{\gamma}_R} - 1\right)\left(1 - \widehat{\gamma}_R\right)} \times \left(1 - \mathcal{B}_{1/2}\left(\widehat{\gamma}_R, \widehat{\rho}, \widehat{\beta}\right)\right). \tag{4.1}$$

Finally, our estimate of the mean is defined as follows

$$\overline{\mu_n} = \frac{1}{n}\sum_{i=1}^{n-k} X_{i,n} + \left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n} - X_{n-k:n}}{\left(2^{\widehat{\gamma}_R} - 1\right)\left(1 - \widehat{\gamma}_R\right)} \times \left(1 - \mathcal{B}_{1/2}\left(\widehat{\gamma}_R, \widehat{\rho}, \widehat{\beta}\right)\right). \tag{4.2}$$

Our main result in the following theorem.

Theorem 4.1. *Under the second order framework in (1.10) with* $\mathbf{A}(t) = \gamma\beta t^{\rho}$, *for an intermediate* k, *i.e.,* k *such that (1.2) holds, whenever* $ln(np)/\sqrt{k} \to 0$, *and* $\sqrt{k}\mathbf{A}(n/k) \to \lambda$, *as* $n \to \infty$, *for* $\widehat{\rho}$ *and* $\widehat{\beta}$ *defined by (3.8), (3.10), such that* $\widehat{\rho} - \rho = o_P(1/\ln n)$. *Then*

$$\frac{\sqrt{k}}{(k/n)X_{n-k:n}}\left(\overline{\mu_n} - \mu\right) \xrightarrow{d} \mathcal{N}\ \left(0, \overline{\sigma}^2\right), \ \text{as } n \to \infty,$$

where

$$\overline{\sigma}^2 = \frac{2\gamma}{(2\gamma-1)} + \gamma^2\left[\frac{1}{(1-\gamma)^2} - \frac{2^{\gamma}\ln 2}{(1-\gamma)(2^{\gamma}-1)}\right]^2 + 2\gamma\left[\frac{1}{(1-\gamma)^2} - \frac{2^{\gamma}\ln 2}{(1-\gamma)(2^{\gamma}-1)}\right].$$

5. Simulation study

We use the **R** statistical software (Ihaka and Gentleman, 1996) to apply the above result to the most usual distribution of Hall class, namely, the Fréchet model: $F(x) = \exp\left(-x^{-1/\gamma}\right)$ for $x > 0$, (for which $\frac{1}{2} < \gamma < 1$, $C = 1$, $\rho = -1$ and $\beta = 1/2$). We simulate 1000 random samples of size n, where n =500; 1000 and 2000. We assume that the second order parameters ρ and β are unknown and they are estimated through (3.8) and (3.10), respectively, both computed at the level $k_1 = [n^{0.995}]$, i.e., we have chosen $\varepsilon = 0.005$ in (3.9).
In the first part, and by using the results of theorem (4.1), we fix $\zeta \in]0,1[$ and $q_{\zeta/2}$ is the $(1-\frac{\zeta}{2})-$quantile of the standard normal distribution $\mathcal{N}(0,1)$. The bounds of confidences of the mean are given by

$$\mu = \overline{\mu}_n \pm q_{\zeta/2}\frac{(n/k)}{\sqrt{k}}\overline{\sigma}.$$

Now, we compare our estimator of the mean with the estimator of Peng (2001), in terms of bias and rmse,

6. Proofs

Let us consider that

$$\overline{\mu_n} - \mu = \Delta_{n,1} + \Delta_{n,2},$$

where

$$\Delta_{n,1} = \frac{1}{n}\sum_{i=1}^{n-k} X_{i,n} - \int_{k/n}^{1} \mathbf{Q}\left(1-s\right)ds,$$

and

$$\Delta_{n,2} = \left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n} - X_{n-k:n}}{\left(2^{\widehat{\gamma}_R}-1\right)\left(1-\widehat{\gamma}_R\right)} \times \left(1 - \mathcal{B}_{1/2}\left(\widehat{\gamma}_R, \widehat{\rho}, \widehat{\beta}\right)\right) - \int_0^{k/n} \mathbf{Q}\left(1-s\right)ds.$$

TABLE 1. Analog between the new estimator of the mean and the estimator of Peng (2001) for two tail index.

γ	2/3			3/4		
n	500	1000	2000	500	1000	2000
μ	2.87635			3.6380		
$\overline{\mu}_n$	2.9807	2.9441	2.8783	3.7638	3.6423	3.6241
bias	0.1043	0.0677	0.0020	0.1258	0.0429	-0.0139
RMSE	0.0350	0.00784	0.0055	0.5068	0.3459	0.0262
$\widehat{\mu}_n$	3.2888	2.6797	2.6911	3.9629	4.02021	3.3675
bias	0.4125	-0.2551	-0.1966	0.3248	0.3821	-0.2704
RMSE	0.1762	0.0795	0.0652	0.7612	0.6635	0.1301

$\Delta_{n,2}$ may be rewriten as follow:

$$\begin{aligned}
\Delta_{n,2} &= \left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n}-X_{n-k:n}}{(1-\widehat{\gamma}_R)} \\
&\quad\times\left(1-\mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)\right)\left[\frac{1}{(2^{\widehat{\gamma}_R}-1)}-\frac{1}{(2^{\gamma}-1)}\right] \\
&\quad+\left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n}-X_{n-k:n}}{(2^{\gamma}-1)} \\
&\quad\times\left(1-\mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)\right)\left[\frac{1}{(1-\widehat{\gamma}_R)}-\frac{1}{(1-\gamma)}\right] \\
&\quad-\left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n}-X_{n-k:n}}{(2^{\gamma}-1)(1-\gamma)} \\
&\quad\times\left(\mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)-\mathcal{B}_{1/2}(\gamma,\rho,\beta)\right) \\
&\quad+\left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n}-X_{n-k:n}}{(2^{\gamma}-1)(1-\gamma)} \\
&\quad\times\left(1-\mathcal{B}_{1/2}(\gamma,\rho,\beta)\right)-\int_0^{k/n}\mathbf{Q}(1-s)\,ds.
\end{aligned}$$

Now, with the use of the delta method enables us to write

$$\begin{aligned}
\frac{1}{(2^{\widehat{\gamma}_R}-1)}-\frac{1}{(2^{\gamma}-1)} &\overset{P}{\sim} -\frac{2^{\gamma}\ln 2}{(2^{\widehat{\gamma}_R}-1)^2}(\widehat{\gamma}_R-\gamma), \\
\frac{1}{(1-\widehat{\gamma}_R)}-\frac{1}{(1-\gamma)} &\overset{P}{\sim} \frac{1}{(1-\gamma)^2}(\widehat{\gamma}_R-\gamma),
\end{aligned}$$

and

$$\begin{aligned}\frac{\mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)}{\mathcal{B}_{1/2}\left(\gamma,\rho,\beta\right)} &= 1+\left(1-\frac{\gamma 2^{\gamma}\left(2^{\rho}-1\right)}{\left(2^{\gamma}-1\right)\left(2^{\gamma+\rho}-1\right)}\right)\frac{\left(\widehat{\gamma}_R-\gamma\right)}{\gamma}\\ &\quad +\frac{\left(\widehat{\beta}-\beta\right)}{\beta}+\left(\widehat{\rho}-\rho\right)\ln\left(n/k\right).\end{aligned}$$

As: $\widehat{\gamma}_R-\gamma=o_p\left(1\right)$, $\left(\widehat{\rho}-\rho\right)\ln\left(n/k\right)=o_p\left(1\right)$ and $\widehat{\beta}-\beta=o_p\left(1\right)$, we have

$$\mathcal{B}_{1/2}\left(\widehat{\gamma}_R,\widehat{\rho},\widehat{\beta}\right)=\mathcal{B}_{1/2}\left(\gamma,\rho,\beta\right)\left(1+o_p\left(1\right)\right).$$

Either, with $\mathbf{Q}(1-s)=Cs^{-\gamma}$, then

$$\begin{aligned}\int_0^{k/n}\mathbf{Q}\left(1-s\right)ds &= C\frac{1}{1-\gamma}\left[s^{1-\gamma}\right]_0^{k/n}\\ &= C\frac{1}{1-\gamma}\left(k/n\right)^{1-\gamma}\\ &= \frac{X_{n-[k/2]:n}-X_{n-k:n}}{\left(2^{\gamma}-1\right)}\left(k/n\right)\frac{1}{1-\gamma}\left(1-\mathcal{B}_{\theta}\left(\gamma,\rho,\beta\right)\right)\end{aligned}$$

where

$$C=\frac{X_{n-[k/2]:n}-X_{n-k:n}}{\left(2^{\gamma}-1\right)}\left(k/n\right)^{\gamma}.$$

Then,

$$\left(\frac{k}{n}\right)\frac{X_{n-[k/2]:n}-X_{n-k:n}}{\left(2^{\gamma}-1\right)\left(1-\gamma\right)}\times\left(1-\mathcal{B}_{1/2}\left(\gamma,\rho,\beta\right)\right)-\int_0^{k/n}\mathbf{Q}\left(1-s\right)ds=o_p(1)$$

as $n\to\infty$. We obtain that

$$\frac{\sqrt{k}}{(k/n)X_{n-k:n}}\Delta_{n,2}=\left[\frac{1}{\left(1-\gamma\right)^2}-\frac{2^{\gamma}\ln 2}{\left(1-\gamma\right)\left(2^{\gamma}-1\right)}\right]\sqrt{k}\left(\widehat{\gamma}_R-\gamma\right)+o_p(1).$$

With the result of Gomes et al. (2007), we have

$$\sqrt{k}\left(\widehat{\gamma}_R-\gamma\right)\overset{d}{\to}\mathcal{N}\left(0,\gamma^2\right),\text{ as } n\to\infty.$$

Let us now consider the asymptotic distribution of $\Delta_{n,1}$. It is shown in Csörgö and Mason (1985) that, there exists a sequence of independent Brownian bridges $\left\{B_n(s);0\leq s\leq 1\right\}_{n\geq 1}$ such that, for all large n

$$\frac{\sqrt{k}}{(k/n)X_{n-k:n}}\Delta_{n,2}=-\frac{\int_{k/n}^1 B_n\left(1-s\right)d\mathbf{Q}\left(1-s\right)}{(k/n)\mathbf{Q}\left(1-k/n\right)}+o_p\left(1\right).$$

Finally:

$$\frac{\sqrt{k}}{(k/n)X_{n-k:n}}\left(\overline{\mu_n}-\mu\right)=\Delta_n\left(\gamma,\rho\right),$$

where

$$\begin{aligned}\Delta_n(\gamma) \quad : \quad &= \left[\frac{1}{(1-\gamma)^2} - \frac{2^\gamma \ln 2}{(1-\gamma)(2^\gamma - 1)}\right](\widehat{\gamma}_R - \gamma) \\ &\quad - \frac{\int_{k/n}^1 B_n(1-s)\, d\mathbf{Q}(1-s)}{(k/n)\mathbf{Q}(1-k/n)} + o_p(1). \\ &= \mathcal{W}_1 + \mathcal{W}_2\end{aligned}$$

It is clear that $\Delta_n(\gamma)$ is a Gaussian random variable with mean zero and variance

$$E(\Delta_n(\gamma))^2 = E(\mathcal{W}_1)^2 + E(\mathcal{W}_2)^2 + 2E(\mathcal{W}_1\mathcal{W}_2).$$

An elementary calculation gives, as $n \to \infty$

$$\begin{aligned}E(\mathcal{W}_1)^2 &= \gamma^2\left[\frac{1}{(1-\gamma)^2} - \frac{2^\gamma \ln 2}{(1-\gamma)(2^\gamma - 1)}\right]^2 + o(1), \\ E(\mathcal{W}_2)^2 &= \frac{2\gamma}{(2\gamma - 1)} + o(1), \\ 2E(\mathcal{W}_1\mathcal{W}_2) &= 2\gamma\left[\frac{1}{(1-\gamma)^2} - \frac{2^\gamma \ln 2}{(1-\gamma)(2^\gamma - 1)}\right] + o(1),\end{aligned}$$

Summing up the right-hand sides of the above three limits, we obtain $\overline{\sigma}^2$, whose expression in terms of the parameter is given in Theorem (4.1).

This completes the proof of Theorem (4.1).

References

[1] Beirlant, J., Matthys, G., Dierckx, G. (2001). Heavy-tailed distributions and rating. *Astin Bull.* 31, 37-58.

[2] Beirlant, J., Teugels, J. (1987). Asymptotic normality of Hill's estimator. Extreme value theory (Oberwolfach, 1987), 148-155, *Lecture Notes in Statist.*, 51, Springer, New York, 1989.

[3] Bingham, N. H, Goldie, C. M., Teugels, J. L. (1987). *Regular Variation.* Cambridge University Press, Cambridge.

[4] Cheng, S., Peng, L. (2001). Confidence intervals for the tail index. *Bernoulli* 7, 751-760.

[5] Csörgö, M., Csörgö, S., Horveath, L., Mason, D. (1986a). Weighted empirical and quantile processes. *Ann. Probab.*, 14, 31-85.

[6] Csörgö, S., Horveath, L., Mason, D. (1986b). What portion of the sample partial sum asymptotically stable or normal? *Probab. Theory Related Fields* 72, 1-16.

[7] Csörgö, M., Mason, D. M. (1985). On the asymptotic distribution of weighted uniform empirical and quantile processes in the middle and on the tails. *Stochast. Proc. Appl.* 21, 119-132.

[8] de Haan, L., Ferreira, A. (2006). *Extreme Value Theory: An Introduction.* New York: Springer.

[9] Embrechts, P., Klüppelberg, C., Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[10] Gomes, M.I., Figueiredo, F., 2006. Bias reduction in risk modelling: semi-parametric quantile estimation. *Test* 15 (2), 375–396.
[11] Gomes and Pestana (2007). A simple second-order reduced bias' tail index estimator. *J. Stat. Comput. Simul.*, 77, 487–504.
[12] Gomes, M.I., Martins, M.J., Neves, M., 2007. Improving second order reduced bias extreme value index estimation. *RevStat* 5 (2), 177–207.
[13] de Haan, L., Peng, L., 1998. Comparison of tail index estimators. *Statistica Neerlandica* 52, 60–70.
[14] Hall, P., 1982. On some simple estimates of an exponent of regular variation. *J. Roy. Statist. Soc.* 44 (1), 37–42.
[15] Hall, P., Welsh, A. H. (1985). Adaptative estimates of parameters of regular variation. *Ann. Statist.*, 13, 331–341.
[16] Hill, B.M., 1975. A simple general approach to inference about the tail of a distribution. *Ann.Statist.* 3, 1163–1174.
[17] Peng, L. (2001). Estimating the mean of a heavy-tailed distribution. *Stat. Probab. Lett.*, 52, 255-264.
[18] Necir, A. (2006). A Nonparametric Sequential Test with Power 1 for the Mean of Lévy-Stable Distributions with Infinite Variance. *Methodology and Computing in Applied Probability* 8, 321-343.
[19] Weissman, I. (1978). Estimation of parameters and large quantiles based on the k largest observations. *J. Amer. Statist. Assoc.* 73, 812-815.

(A. Rassoul) National High School of Hydraulic, Blida, Algeria
E-mail address: a.rassoul@ensh.dz

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 211-214, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON THE HYERS-ULAM STABILITY OF NON-CONSTANT VALUED LINEAR DIFFERENTIAL EQUATION $xy' = -\lambda y$

HAMID VAEZI AND HABIB SHAKOORY

Abstract. We consider a differentiable map y from an open interval to a real Banach space of all bounded continuous real-valued functions on a topological space. We will investigate the Hyers-Ulam stability of the following linear differential equations of first order with non-constant values:

$$xy' = -\lambda y,$$

where λ is a positive real number and

$$y \in C(I) = C(a,b), \ -\infty < a < b < +\infty, \ x \in (0, \infty).$$

1. Introduction

C. Alsina and R. Ger [1] remarked that the differential equation $y' = y$ has the Hyers-Ulam stability. More explicitly, they proved that if a differentiable function $y : I \to R$ satisfies $|y'(t) - y(t)| \le \varepsilon$ for all $t \in I$, Then there exists a differentiable function $g : I \to R$ satisfying $g'(t) = g(t)$ for any $t \in I$ such that $|y(t) - g(t)| \le 3\varepsilon$ for every $t \in I$. T. Miura , S.-E. Takahashi and H. Choda, [3] consider a differentiable map y from an open interval to a real Banach space of all bounded continuous real-valued functions on a topological space. They show that y can be approximated by the solution to the differential equation $x'(t) = \lambda x(t)$ if $\|y'(t) - \lambda y(t)\|_\infty < \varepsilon$ holds. In [4], this result was generalized to the case of the complex Banach space valued differential equation $y' = \lambda y$. Y. Li [2] prove stability in the sense of Hyers-Ulam of differential equation of second order $y'' = \lambda^2 y$. In this article, we say that equation $xy' = -\lambda y$ has the Hyers-Ulam stability if there exists a constant $K > 0$ with the following property: for every $x, \varepsilon > 0$, $y \in C(I)$ if

$$|xy' - (-\lambda y)| < \varepsilon,$$

then there exists some $z \in C(I)$ satisfying

$$xz' + \lambda z = 0$$

such that

$$|y(x) - z(x)| \le k\varepsilon.$$

We call such k as Hyers-Ulam stability constant for equation

$$xy' = -\lambda y.$$

Key words and phrases. Hyers-Ulam Stability, differential equation, approximation.
2010 *AMS Math. Subject Classification.* 34K20; 26D10.

2. Definitions

Definition 2.1. *Let B be a Banach space, y a map from I (denotes an open interval of the real number field R) into B. We say that y is differentiable, if for every $t \in I$ there exists an $y' \in B$ so that:*

$$\lim_{s \to 0} \|(y(t+s) - y(t))/(s) - y'(t)\|_B = 0,$$

where $\|.\|_B$ denotes the norm on B. Let y be a differentiable function on I in to R. Alsina and Ger [1] gave all the solutions to the inequality $|y'(t) - y(t)| \leq \varepsilon$ for every $t \in I$. Then they showed that each solution to the inequality above was approximated by a solution to the differential equation $z'(t) = z(t)$.

In accordance with [1], we define the Hyers-Ulam stability of Banach space valued differentiable map.

Definition 2.2. *Let B be a Banach space, y a differentiable map on I (denotes an open interval of the real number field R) into B so that for each $t \in I$:*

$$\|y'(t) - \lambda y(t)\|_B \leq \varepsilon.$$

We say that Hyers-Ulam stability holds for y if there exist a $k \geq 0$ and a differentiable map z on I into B such that:

$$z'(t) - \lambda z(t) = 0$$

and

$$\|y(t) - z(t)\|_B \leq k\varepsilon$$

hold for every $t \in I$.

3. Main results

Now, the main result of this work is given in the following theorem.

Theorem 3.1. *If a continuously differentiable function $y : I \to R$ satisfies the differential inequality*

$$|xy' + \lambda y| \leq \varepsilon$$

for all $t \in I$ and for some $\varepsilon > 0$, then there exists a solution $v : I \to R$ of the equation:

$$xv' + \lambda v = 0$$

such that

$$|y(x) - v(x)| \leq k\varepsilon,$$

where $k > 0$ is a constant.

Proof. Let $\epsilon > 0$ and $y : I \to R$ be a continuously differentiable function such that

$$|xy' + \lambda y| \leq \epsilon.$$

We will show that there exists a constant k independent of ε and v such that $|y - v| \leq k\varepsilon$ for some $v \in C(I)$ satisfying :

$$xv' + \lambda v = 0.$$

If we set

$$x = e^u, \; x \in (0, \infty),$$

$$xy' = x.dy/dx = x.dy/du.du/dx = x.1/x.dy/du = dy/du = Dy$$

according the above:

$$xy' = dy/du \cong y'_u,$$

then

$$|y'(u) + \lambda y(u)| \leq \varepsilon,$$

equivalently, y satisfies

$$-\varepsilon \leq y'(u) + \lambda y(u) \leq \varepsilon.$$

Multiplying the formula by the function $e^{\lambda u}$, we obtain

$$-\varepsilon e^{\lambda u} \leq y'(u)e^{\lambda u} + \lambda y(u)e^{\lambda u} \leq \varepsilon e^{\lambda u}$$

For the case $0 < \lambda \leq 1$, there exists $M > 0$ such that $M\lambda > 1$, so without loss of generality, we may assume that $\lambda > 1$, thus

$$-\lambda\varepsilon e^{\lambda u} \leq y'(u)e^{\lambda u} + \lambda y(u)e^{\lambda u} \leq \lambda\varepsilon e^{\lambda u}. \tag{3.1}$$

For some fixed $c \in (a,b)$ with $y(c) < \infty$ and any $u \in (c,b)$, integrating (3.1) from c to u, we get

$$-\lambda\epsilon \int_c^u e^{\lambda u} \leq \int_c^u y'(u)e^{\lambda u} + \lambda \int_c^u y(u)e^{\lambda u} \leq \lambda\varepsilon \int_c^u e^{\lambda u},$$

and so

$$-\varepsilon e^{\lambda(u-c)} \leq y(u)e^{\lambda u} - y(c)e^{\lambda c} \leq +\varepsilon e^{\lambda(u-c)}.$$

Summing parts of equation with $+\varepsilon e^{\lambda c}$:

$$-\varepsilon e^{\lambda(u)} \leq y(u)e^{\lambda u} - y(c)e^{\lambda c} + \varepsilon e^{\lambda c} \leq +\varepsilon e^{\lambda(u)}$$

Multiplying the formula by the function $e^{-\lambda u}$, we get

$$-\varepsilon \leq y(u) - (y(c) - \varepsilon)e^{\lambda c}e^{-\lambda u} \leq \varepsilon,$$

so

$$-\varepsilon \leq y(u) - (y(c) - \varepsilon)e^{\lambda c}e^{-\lambda u} \leq \varepsilon.$$

Then,

$$-\varepsilon \leq y(u) - (y(c) - \varepsilon)e^{\lambda(c-u)} \leq \varepsilon.$$

Let $z(u) = (y(c) - \varepsilon)e^{\lambda(c-u)}$. Then $z(u)$ satisfies

$$z'(u) + \lambda z(u) = 0$$

and

$$|y(u) - z(u)| \leq \varepsilon$$

For any $a \in (u,c)$, the proof is very similar to the above, so we omit it. We have

$$-\varepsilon \leq y(u) - z(u) \leq \varepsilon.$$

where $k = 1$ if we change $x = e^U$ and $x \in (c,b)$. By an argument similar to the above for some fixed $c \in (a,b)$, we can show that there exists

$$v(x) = (y(c) - \varepsilon)e^{\lambda(x-c)} - e^{\lambda(x-a)} \int z(s)e^{-\lambda(s-a)}ds$$

such that

$$|y(x) - v(x)| \leq \varepsilon,$$

which completes the proof. $\square$

References

[1] C. Alsina and R. Ger, On some inequalities and stability results related to the exponential function, *J. Inequal. Appl.* 2 (1998) 373–380.
[2] Y. Li, Hyers-Ulam Stability of Linear differential equations $y'' = \lambda^2 y$, J. Thai.8 (2010) 215–219.
[3] T. Miura, S.-E. Takahasi and H. Choda, On the Hyers-Ulam stability of real continuous function valued differentiable map, *J. Tokyo Math.* 24 (2001) 467–476.
[4] S.-E. Takahasi, T. Miura and S. Miyajima, On The Hyers-Ulam stability of the Banach space-valued differential equation $y' = \lambda y$; to appear in *Bull. Korean Math. Soc.*

(H.Vaezi) University of Tabriz, Tabriz, Iran
E-mail address: hvaezi@tabrizu.ac.ir

(H. Shakoory) Taavon University of Applied Science, Tabriz, Iran
E-mail address: habibSakoory@yahoo.com

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 215-229, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

THE APPROXIMATE SOLUTION OF MULTI-HIGHER ORDER LINEAR VOLTERRA INTEGRO-FRACTIONAL DIFFERENTIAL EQUATIONS WITH VARIABLE COEFFICIENTS IN TERMS OF ORTHOGONAL POLYNOMIALS

SHAZAD SHAWKI AHMED AND SHOKHAN AHMED HAMA SALIH

ABSTRACT. The main purpose of this paper is to present an approximation method for multi-higher order of linear Volterra Integro-Fractional Differential Equations (m-h LVIFDEs) with variable coefficients in the most general form under the conditions. The method is based on the orthogonal polynomials (Chebyshev and Legendre) via least square technique. This method transforms the fractional equation and the given conditions into matrix equations, which correspond to a system of linear algebraic equations with unknown coefficients and apply Gaussian elimination method to determine the approximate orthogonal coefficients. The proposed method including two new algorithms for solving our problem, for each algorithm, a computer program was written. Finally, numerical examples are presented to illustrate the effectiveness and accuracy of the method and the results are discussed.

1. Introduction

In this work, we will consider the multi-higher order of linear Volterra Integro-Fractional Differential Equations (m-h LVIFDEs) with variable coefficients in the form:

$$(1.1)\begin{cases} {}_a^C D_x^{\alpha_n} y(x) + \sum\limits_{i=0}^{n-1} P_i(x) {}_a^C D_x^{\alpha_{n-1}} y(x) + P_n(x) y(x) \\ \qquad = f(x) + \sum\limits_{\ell=0}^{m} \lambda \int\limits_a^x K_\ell(x,t) {}_a^C D_t^{\beta_{m-\ell}} y(t) dt, \quad x \in [a,b] = I \end{cases}$$

with the initial conditions:

$$(1.2) \qquad [y^{(k)}(x)]_(x = a) = y_k;$$

$$k = 0, 1, \ldots, \mu - 1; \mu = max\{m_i, m_\ell^* \quad forall\ i\ and\ \ell\},$$

where $y(x)$ is the unknown function, which is the solution of equation (1.1). The functions $K_\ell : S \times \mathbb{R} \to \mathbb{R}, (with\ S = \{(x,t) : a \le t \le x \le b\}); \ell = 0, 1, 2, \ldots, m$ and $f, P_i : I \to \mathbb{R}$ are all continuous functions .Where $\alpha_i, \beta_j \in \mathbb{R}^+, m_i - 1 < \alpha_i \le m_i$ and $m_\ell^* - 1 < \beta_\ell \le m_\ell^*, m_i = \lceil \alpha_i \rceil$ and $m_\ell^* = \lceil \beta_\ell \rceil$ for all $i = (\overline{1:n})$ and $\ell = (\overline{1:m})$ with property that:$\alpha_n > \alpha_{n-1} > \ldots > \alpha_1 > \alpha_0 = 0$ and $\beta_m > \beta_{m-1} > \ldots > \beta_1 > \beta_0 = 0$, and λ is a scalar parameter.

Actually most linear Volterra integro-fractional differential equations of multi-higher order with variable coefficients do not have exact analytic solutions, therefore

Key words and phrases. Integro-Fractional Differential Equations, Caputo Fractional Derivative, least-square technique, Orthogonal (Chebyshev, Legendre) polynomial.

2010 *AMS Math. Subject Classification.* 26A33, 05E35.

1

approximation and numerical techniques must be used. The area of orthogonal polynomials is a very active research area in mathematics as well as in applications in mathematical physics, engineering, computer science and others [7, 8].

AL-Rawi [3] and Qassim [14] used Laguerre polynomial to approximate the solution of the first kind integral equations depending on the principle of the least-squares data fitting. On the other hand Ameen [14] and Al-Ani [1], Kalwi [2] applied this techniques to treat numerically FIE's of second kind using Laguerre, Hermit and Legendre and Chebyshev polynomials, respectively. In this paper, we extend this method to further deal with our consider problem (1.1) with initial conditions (1.2).

2. Preliminaries and Notations

2.1. Basic Definitions.

Definition 2.1. [16], *A real valued function y defined on $[a, b]$ be in the space $C_\gamma[a, b], \gamma \in \mathbb{R}$, if there exists a real number $p > \gamma$, such that $y(x) = (x - a)^p y_*(x)$, where $y_* \in C[a, b]$, and it is said to be in the space $C^n_\gamma[a, b]$ iff $y^{(n)} \in C_\gamma[a, b]$,$n \in \mathbb{N}_0$.*

Definition 2.2. [4, 5], *Let $y \in C_\gamma[a, b]$,$\gamma \geq -1$ and $\alpha \in \mathbb{R}^+$. Then the Riemann-Liouville fractional integral operator ${}_aJ^\alpha_x$ of order α of a function y, is defined as:*

$$ {}_aJ^\alpha_x y(x) = \frac{1}{\Gamma(\alpha)} \int_a^x (x - t)^{\alpha-1} y(t)dt,\ \alpha > 0 $$

$$ {}_aJ^0_x y(x) = Iy(x) = y(x),\ \alpha = 0 $$

Definition 2.3. [4, 5], *Let $\alpha \geq 0$ and $m = \lceil \alpha \rceil$.The Riemann-Liouville fractional derivative operator ${}^R_aD^\alpha_x$, of order α and $y \in C^m_{-1}[a, b]$, is defined as:*

$$ {}^R_aD^\alpha_x y(x) = D^m_x\, {}_aJ^{m-\alpha}_x y(x) $$

If $\alpha = m, m \in \mathbb{N}_0$, and $y \in C^m[a, b]$ we have

$$ {}^R_aD^0_x y(x) = y(x), \quad {}^R_aD^m_x y(x) = y^{(m)}(x) $$

Definition 2.4. [6, 10], *The Caputo fractional derivative operator ${}^C_aD^\alpha_x$ of order $\alpha \in \mathbb{R}^+$ of a function $y \in C^m_{-1}[a, b]$ and $m - 1 < \alpha \leq m$, $m \in \mathbb{N}$ is defined as:*

$$ {}^C_aD^\alpha_x y(x) = {}_a J^{m-\alpha}_x D^m_x y(x) $$

Thus for $\alpha = m$, $m \in \mathbb{N}_0$, and $y \in C^m[a, b]$, we have for all $a \leq x \leq b$

$$ {}^C_aD^0_x y(x) = y(x), \quad {}^C_aD^m_x y(x) = D^m_x y(x) $$

Note That:[4, 5, 6, 10]

- **i:** For $\alpha \geq 0$ and $\beta > 0$,then ${}_aJ^\alpha_x (x - a)^{\beta-1} = \frac{\Gamma(\beta)}{\Gamma(\beta+\alpha)} (x - a)^{\beta+\alpha-1}$.
- **ii:** For $\alpha \geq 0$, $\beta > 0$ and $y(x) \in C_\gamma[a, b]$, $\gamma \geq -1$, then :

$$ {}_aJ^\alpha_x\, {}_aJ^\beta_x y(x) = {}_aJ^\beta_x\, {}_aJ^\alpha_x y(x) = {}_a J^{\alpha+\beta}_x y(x) $$

- **iii:** ${}^R_aD^\alpha_x A = A\frac{(x-a)^{-\alpha}}{\Gamma(1-\alpha)}$ and ${}^C_aD^\alpha_x A = 0$; A is any constant ; $(\alpha \geq 0, \alpha \notin \mathbb{N})$.
- **iv:** ${}^C_aD^\alpha_x y(x) = D^m_x\, {}_aJ^{m-\alpha}_x y(x) \neq {}_aJ^{m-\alpha}_x D^m_x y(x) = {}^R_aD^\alpha_x y(x)$; $m = \lceil \alpha \rceil$
- **v:** Assume $y \in C^m_{-1}[a, b]; \alpha \geq 0, \alpha \notin \mathbb{N}$ and $m = \lceil \alpha \rceil$ Then ${}^C_aD^\alpha_x y(x)$ is continuous on [a,b], and $\left[{}^C_aD^\alpha_x y(x)\right]_{x=a} = 0$.

vi: Let $\alpha \geq 0, m = \lceil \alpha \rceil$ and $y \in C^m[a,b]$,then, the relation between the Caputo derivative and the R-L integral are formed:

$$ {}_a^C D_x^\alpha {}_a J_x^\alpha y(x) = y(x) \quad a \leq x \leq b \,; {}_a J_x^\alpha {}_a^C D_x^\alpha y(x) = y(x) - \sum_{k=0}^{m-1} \frac{y^{(k)}(a)}{k!}(x-a)^k $$

vii: ${}_a^C D_x^\alpha y(x) = {}_a^R D_x^\alpha \left[y(x) - T_{m-1}\left[y;a\right]\right]$;$(m-1 < \alpha \leq m)$ and $T_{m-1}\left[y;a\right]$ denotes the Taylor polynomial of degree $m-1$ for the function y , centered at a.

viii: Let $\alpha \geq 0$; $m = \lceil \alpha \rceil$and for $y(x) = (x-a)^\beta$ for some $\beta \geq 0$.Then:

$$ {}_a^C D_x^\alpha y(x) = \begin{cases} 0 & if \quad \beta \in \{0,1,2,\ldots,m-1\} \\ \frac{\Gamma(\beta+1)}{\Gamma(\beta-\alpha+1)}(x-a)^{\beta-\alpha} & if \quad \beta \in \mathbb{N} \; and \; \beta \geq m \\ & or \quad \beta \notin \mathbb{N} \; and \; \beta > m-1 \end{cases} $$

2.2. **Orthogonal Polynomials.** Orthogonal polynomials are classes of polynomials $\varphi_i; i = 0,1,2,\ldots$ defined over a range $[a,b]$ that obey an orthogonality relation $\int_a^b w(x)\varphi_i(x)\varphi_j(x)dx = \delta_{ij}C_j$, where $w(x)$ is a weighting function and δ_{ij} is the kronecker delta (equal to 1 if $i=j$ and to 0 otherwise). If $C_j = 1$ then the polynomials are not only orthogonal, but orthonormal. Two of the most common set of orthogonal polynomials are Chebyshev polynomials $T_k(x)$ and Legendre polynomials $P_k(x)$,[11, 12, 14].

2.2.1. *Chebyshev Polynomials.* The Chebyshev polynomials are orthogonal over $[-1,1]$ with respect to the weight function $w(x) = 1/\sqrt{1-x^2}$. The Chebyshev polynomials of order$k \geq 0$, $T_k(x)$, for $-1 \leq x \leq 1$, given by the simple calculation forms:

$$ T_k(x) = \frac{k}{2} \sum_{r=0}^{\lfloor \frac{k}{2} \rfloor} \frac{(-1)^r}{k-r} \binom{k-r}{r} (2x)^{k-2r}, \;\; T_0(x) = 1, \;\; k \geq 1 \tag{2.1} $$

where $\lfloor * \rfloor$ is the floor function and $T_k(x)$ has $k-$ distinct in $[-1,1]$, the roots $\{x_i\}_{i=1}^n$ have the expression:

$$ x_i = cos(\frac{\pi(i+1/2)}{k}); \quad i = 0,1,2,\ldots,k-1 \tag{2.2} $$

Also, the $n-$derivative of Chebyshev polynomials $T_k(x)$ is formulated as:

$$ \frac{d^n T_k(x)}{dx^n} = \begin{cases} \frac{k}{2} \sum_{r=0}^{\lfloor \frac{k-n}{2} \rfloor} \frac{(-1)^r}{r!} \frac{(k-r-1)!}{(k-2r-n)!} 2^n (2x)^{k-2r-n} & if \quad k > n \\ 2^{k-1} k! & if \quad k = n \\ 0 & if \quad k < n \end{cases} $$

2.2.2. *Legendre Polynomials.* An important another set of the polynomial approximation over $[-1,1]$ is Legendre polynomials, a simple calculation formula for it of higher order $k \geq 0$, is defined in terms of the sums:

$$ P_k(x) = \frac{1}{2^k} \sum_{r=0}^{\lfloor \frac{k}{2} \rfloor} (-1)^r \binom{k}{r} \binom{2k-2r}{k} x_{k-2r} \tag{2.3} $$

Also, the $n-$derivative of Legendre polynomials $P_k(x)$ is formulated as:

$$\frac{d^n P_k(x)}{dx^n} = \begin{cases} \frac{1}{2^k}\sum_{r=0}^{\lfloor\frac{k-n}{2}\rfloor} \frac{(-1)^r}{r!(k-r)!}\frac{(2k-2r)!}{(k-2r-n)!}x^{k-2r-n} & if \quad k > n \\ (2k)!/2^k k! & if \quad k = n \\ 0 & if \quad k < n \end{cases}$$

Remark 2.1. [12], *Sometime it is necessary to take a problem stated on an interval* $[a, b]$ *and reformulates the problem on the interval* $[-1, 1]$ *where the solution is known. If the approximation* $f_N(x)$ *to* $f(x)$ *is to be obtained on the interval* $[a, b]$*, then we change the variable so that the problem is reformulated on* $[-1, 1]$ *Thus* $x = (\frac{b-a}{2})t + (\frac{b+a}{2}), \quad t = 2(\frac{x-a}{b-a}) - 1 \quad where \quad a \le x \le b; -1 \le t \le 1.$

2.3. **Gaussian Quadrature Formulas.** Gaussian quadrature formula can be developed based upon the orthogonality property. All Gaussian quadrature share the following formula: $\int_a^b w(x)f(x)dx \cong \sum_{k=0}^{N} \lambda_k f(x_k)$ where $x_k, k = (\overline{0:N})$ are called nodes which ,here,associated with zeros of orthogonal polynomials are the integration points and λ_k , $k = (\overline{0:N})$ are called weights of the quadrature formula related to the orthogonal polynomials, [9, 12]. If the weights function is $w(x) = 1/\sqrt{(x-a)(b-x)}$ on interval $[a, b]$. Then, the **Open Gauss-Chebyshev quadrature rule** has the form:

$$\int_a^b \frac{1}{\sqrt{(x-a)(b-x)}} f(x)dx \cong \frac{\pi}{N}\sum_{k=0}^{N-1} f(x_k) \tag{2.4}$$

where $\quad x_k = (\frac{b-a}{2})z_k + (\frac{b+a}{2})$ and $z_k = cos(\frac{2k+1}{N}\frac{\pi}{2})$
and the **Closed Gauss-Chebyshev quadrature rule** has the form:

$$\int_a^b \frac{1}{\sqrt{(x-a)(b-x)}} f(x)dx \cong \frac{\pi}{N}\sum_{k=0}^{N}{}'' f(x_k) \tag{2.5}$$

where $\quad x_k = (\frac{b-a}{2})z_k + (\frac{b+a}{2})$ and $z_k = cos(\frac{k\pi}{N})$
where the double prime $('')$ on the summation sign implies that the first and end terms are halved. If the weight function is $w(x) = 1$ on the interval $[a, b]$.Then, the general **Gauss-Legender quadrature rule** has the form:

$$\int_a^b f(x)dx \cong \frac{b-a}{2}\sum_{k=0}^{N-1} \lambda_k f(x_k) \tag{2.6}$$

where $\quad x_k = (\frac{b-a}{2})z_k + (\frac{b+a}{2})$ and z_k are the $k-$ th zeros of $P_N(x)$. The coefficients λ_k's can be calculated using the following formula :

$$\lambda_k = \frac{2}{(1-z_k^2)(P_N'(z_k))^2} \tag{2.7}$$

2.4. **Lemmas.** Before starting the solution of general form of m-h LVIFDEs by least-square orthogonal method, the following two basic lemmas are needed:

Lemma 2.5. [13], *The Caputo fractional derivative for order* $m - 1 < \alpha < m$,$m = \lceil\alpha\rceil$, *of shifed Chebyshev polynomial of degree* $k \ge 1$, $T_k^*(x) = T_k[2(\frac{x-a}{b-a}) - 1]$ *on interval* $[a, b]$ *can be formulated:*

$$
\begin{aligned}
&{}_a^C D_x^{\alpha} T_k^*(x) = \\
&\tfrac{k}{2}\left(\tfrac{2}{b-a}\right)^m (x-a)^{m-\alpha} \sum_{r=0}^{\lfloor k/2 \rfloor} (-1)^r \tfrac{\Gamma(k-r)}{\Gamma(r+1)} 2^{k-2r} M(x;k,r,m)
\end{aligned} \tag{2.8}
$$

where $M(x;k,r,m)$

$$
= \begin{cases} 0 & m > k-2r \\ \frac{1}{\Gamma(m-\alpha+1)} & m = k-2r \\ \sum\limits_{\ell=0}^{k-2r-m} \frac{(-1)^{\ell+k-2r-m}}{\Gamma(\ell+m-\alpha+1)\Gamma(k-2r-m-\ell+1)} \left[2\left(\frac{x-a}{b-a}\right)\right]^{\ell} & m < k-2r \end{cases}
$$

Lemma 2.6. [15], *The Caputo fractional derivative for order $m-1 < \alpha < m, m = \lceil \alpha \rceil$, of shifed Chebyshev polynomial of degree $k \geq 1$, $P_k^*(x) = P_k[2(\frac{x-a}{b-a}) - 1]$, on interval $[a,b]$ can be formulated:*

$$
\begin{aligned}
&{}_a^C D_x^{\alpha} P_k^*(x) = \\
&\tfrac{1}{2^k}\left(\tfrac{2}{b-a}\right)^m (x-a)^{m-\alpha} \sum_{r=0}^{\lfloor k/2 \rfloor} (-1)^r \tfrac{\Gamma(2k-2r+1)}{\Gamma(r+1)\Gamma(k-r+1)} M(x;k,r,m)
\end{aligned} \tag{2.9}
$$

where $M(x;k,r,m)$

$$
= \begin{cases} 0 & m > k-2r \\ \frac{1}{\Gamma(m-\alpha+1)} & m = k-2r \\ \sum\limits_{\ell=0}^{k-2r-m} \frac{(-1)^{\ell+k-2r-m}}{\Gamma(\ell+m-\alpha+1)\Gamma(k-2r-m-\ell+1)} \left[2\left(\frac{x-a}{b-a}\right)\right]^{\ell} & m < k-2r \end{cases}
$$

3. Solution Technique

In this section, a new technique for solving multi-higher linear VIFDEs with variable coefficients applying least-square data fitting with use orthogonal (Chebyshev and Legendre) polynomials has been presented. The method is to approximate the solution $y(x)$ of equation (1.1) by $y_N(x)$:

$$
y(x) \cong y_N(x) = \sum_{r=0}^{N} C_r \varphi_r(x) \tag{3.1}
$$

The coordinate functions $\varphi_r(x)$ are usually chosen as orthogonal polynomials and C_r's are undetermined constant coefficients for all $r = 0,1,2,\ldots,N$. Substituting $y_N(x)$in equation (1.1) , we obtain :

$$
\begin{aligned}
&\sum_{r=0}^{N} C_r \left\{ {}_a^C D_x^{\alpha_n} \varphi_r(x) + \sum_{i=1}^{n-1} P_i(x) {}_a^C D_x^{\alpha_{n-i}} \varphi_r(x) + P_n(x)\varphi_r(x) \right\} = \\
&f(x) + \lambda \sum_{r=0}^{N} C_r \int_a^x \sum_{\ell=0}^{m} K_\ell(x,t) {}_a^C D_t^{\beta_{m-\ell}} \varphi_r(t)dt + R_N(x;C)
\end{aligned} \tag{3.2}
$$

Define: $\psi_r(x) = {}_a^C D_x^{\alpha_n} \varphi_r(x) + \sum\limits_{i=1}^{n-1} P_i(x) {}_a^C D_x^{\alpha_{n-i}} \varphi_r(x) + P_n(x)\varphi_r(x)$

$$
- \lambda \int_a^x \sum_{\ell=0}^{m} K_\ell(x,t) {}_a^C D_t^{\beta_{m-\ell}} \varphi_r(t)dt
$$

Now, the equation (3.2) can be written as:

$$R_N(x;C) = \sum_{r=0}^{N} C_r \psi_r(x) - f(x) \tag{3.3}$$

where $R_N(x;C)$ is the error function which is also depending on x and the constant coefficients $C, \{= (C_0, C_1, \ldots, C_N)\}$. The main points here, is how to find the coefficients $C_r; r = 0, 1, 2, \ldots, N$ of $y_N(x)$ in equation (3.1) such that $R_N(x;C)$ is minimized. The general least squares techniques insist on minimizing the norm of the error functions by introducing a weight function $w(x)$ on interval $[a, b]$, we now wish to minimize $I(C)$:

$$I(C) = \int_a^b w(x) \left| R_N(x;C) \right|^2 dx$$

The necessary condition for $I(C)$ to be minimum, are

$$\frac{\partial I(C)}{\partial C_s} = 2 \int_a^b w(x) R_N(x;C) \frac{\partial R_N(x;C)}{\partial C_s} dx = 0, \quad s = 0, 1, \ldots, N \tag{3.4}$$

Putting equation(3.2) in to equation (3.4), we obtain:

$$\int_a^x w(x) \psi_s(x) \left[\sum_{r=0}^{N} C_r \psi_r(x) - f(x) \right] dx = 0$$

After some simple manipulation, the following linear system is included:

$$\sum_{r=0}^{N} C_r a_{rs} = b_s, \quad s = 0, 1, 2, \ldots, N \tag{3.5}$$

where

$$\begin{cases} a_{rs} = \int_a^b w(x) \psi_s(x) \psi_r(x) dx \\ b_s = \int_a^b w(x) \psi_s(x) f(x) dx \end{cases} \tag{3.6}$$

and $\psi_r(x) = {}_a^C D_x^{\alpha_n} \varphi_r(x) + \sum_{i=1}^{n-1} P_i(x) {}_a^C D_x^{\alpha_{n-i}} \varphi_r(x) + P_n(x) \varphi_r(x)$

$$- \lambda \int_a^x \sum_{\ell=0}^{m} K_\ell(x,t) {}_a^C D_t^{\beta_{m-\ell}} \varphi_r(t) dt$$

Rewrite equation (3.5) in matrix form as:

$$AC = B \tag{3.7}$$

where

$$A = \begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0N} \\ a_{10} & a_{11} & \cdots & a_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N0} & a_{N1} & \cdots & a_{NN} \end{bmatrix}, C = \begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_N \end{bmatrix}, B = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_N \end{bmatrix}$$

In this technique the initial conditions of equation (1.1) are added as new rows in the system (3.7), these rows can be formed as:

$$\sum_{r=0}^{N} C_r \left[\varphi_r^{(k)}(x)\right]_{x=a} = y_k, \quad k = 0, 1, \dots, \mu - 1$$

In matrix form, this gives:

$$\Phi C = Y \tag{3.8}$$

where

$$\Phi = \begin{bmatrix} \varphi_0(a) & \varphi_1(a) & \cdots & \varphi_N(a) \\ \varphi_0^{'}(a) & \varphi_1^{'}(a) & \cdots & \varphi_N^{'}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_0^{(\mu-1)}(a) & \varphi_1^{(\mu-1)}(a) & \cdots & \varphi_N^{(\mu-1)}(a) \end{bmatrix}, C = \begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_N \end{bmatrix}, Y = \begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{\mu-1} \end{bmatrix}$$

Obtaining a new matrix by adding (3.8) to (3.7), yields

$$DC = E \tag{3.9}$$

where $D = \begin{bmatrix} A \\ \cdots \\ \Phi \end{bmatrix}_{(N+\mu+1)\times(N+1)}$ and $E = \begin{bmatrix} B \\ \cdots \\ Y \end{bmatrix}_{(N+\mu+1)\times(1)}$

To determine the constant coefficients C_r's in equation (3.9),store the matrix D and compute $D^T D$ and $D^T E$ then use any numerical evaluation procedure to solve: $[D^T D]C = [D^T E]$. Then substitute the values C_r's in equation (3.1) ,the approximate solution is obtained for multi-higher linear VIFDEs (1.1).

3.1. **Using Chebyshev Polynomials.** In this part, we can take instead of trial function $\varphi_r(x)$ the shifted Chrbyshev polynomial,$T_r^*(x) = T_r\left[2\left(\frac{x-a}{b-a}\right) - 1\right]$, to approximate the solution $y(x)$ of multi-higher order linear VIFDE (1.1-1.2) by formed:

$$y_N(x) = \sum_{r=0}^{N} C_r T_r \left[2\left(\frac{x-a}{b-a}\right) - 1\right]; \qquad a \le x \le b$$

Substituting $y_N(x)$ in to equation (1.1), and applying the same steps described in (2.3), we conclude the system in (3.9), with:

I: using **Open Gauss Chebyshev formula** (2.4) ,we have:

$$\begin{cases} a_{rs} = \frac{\pi}{M} \sum\limits_{j=0}^{M-1} \psi_r(x_j)\psi_s(x_j) \\ b_s = \frac{\pi}{M} \sum\limits_{j=0}^{M-1} \psi_s(x_j) f(x_j) \end{cases} \tag{3.10}$$

where M is the number of Chebyshev zeros we can take it and

$$\left\{ \; x_j = \left(\frac{b-a}{2}\right) z_j + \left(\frac{b+a}{2}\right) \quad ; \quad z_j = cos\left(\frac{2j+1}{M}\frac{\pi}{2}\right) \right. \tag{3.11}$$

II: using **Closed Gauss Chebyshev formula** (2.5) ,we have:

$$\left\{\begin{array}{l} a_{rs}=\frac{\pi}{M}\sum\limits_{j=0}^{M}{}''\psi_r(x_j)\psi_s(x_j) \\ b_s=\frac{\pi}{M}\sum\limits_{j=0}^{M}{}''\psi_s(x_j)f(x_j) \end{array}\right. \tag{3.12}$$

where M is the number of Chebyshev extrema zeros we can take it and

$$\left\{\ x_j=\left(\frac{b-a}{2}\right)z_j+\left(\frac{b+a}{2}\right)\quad ;\quad z_j=cos\left(\frac{j\pi}{M}\right)\right. \tag{3.13}$$

To evaluate $\psi_r(x)$ at $x=x_j$ for all r and j , for open and closed Gauss-Chebyshev formulas, that is:

$$\begin{aligned}[\psi_r(x)]_{x=x_j}=&\left\{{}_a^CD_x^{\alpha_n}+\sum_{i=1}^{n-1}P_i(x){}_a^CD_x^{\alpha_{n-i}}+P_n(x)\right\}T_r\left[2\left(\frac{x-a}{b-a}\right)-1\right]\\ &-\lambda\int_a^x\sum_{\ell=0}^{m}K_\ell(x,t){}_a^CD_t^{\beta_{m-\ell}}T_r\left[2\left(\frac{t-a}{b-a}\right)-1\right]dt\end{aligned} \tag{3.14}$$

We apply the following stages to evaluate the equation (3.14): First, using lemma (2.5) to evaluate the fractional differentiation of shifted Chebyshev polynomials at all $x=x_j$ and all different $r\geq 1$. second, the integral part of equation (3.14) can be extended in to two integral terms I_1 and I_2 as:

$$\begin{aligned}&\int_a^x\sum_{\ell=0}^{m}K_\ell(x,t){}_a^CD_t^{\beta_{m-\ell}}T_r\left[2\left(\frac{t-a}{b-a}\right)-1\right]dt=\\ &\underbrace{\int_a^x\sum_{\ell=0}^{m-1}K_\ell(x,t){}_a^CD_t^{\beta_{m-\ell}}T_r\left[2\left(\frac{t-a}{b-a}\right)-1\right]dt}_{I_1}+\underbrace{\int_a^x K_m(x,t)T_r\left[2\left(\frac{t-a}{b-a}\right)-1\right]dt}_{I_2}\end{aligned}$$

For I_1: Apply lemma (2.5) for integrand fractional differentiation shifted Chebyshev polynomials for different $k\geq 1$, we obtain:

$$\begin{aligned}I_1=&\frac{r}{2}\left(\frac{2}{b-a}\right)^{m_\ell^*}\sum_{p=0}^{\lfloor r/2\rfloor}(-1)^p\frac{\Gamma(r-p)}{\Gamma(p+1)}2^{r-2p}*\\ &*\int_a^x K_\ell(x,t)(t-a)^{m_\ell^*-\beta_{m-\ell}}M(t;r,p,m_\ell^*)dt\end{aligned} \tag{3.15}$$

where $\quad M(t;r,p,m_\ell^*)$

$$=\begin{cases}0 & m_\ell^*>r-2p\\ \frac{1}{\Gamma(m_\ell^*-\beta_{m-\ell}+1)} & m_\ell^*=r-2p\\ \sum\limits_{q=0}^{r-2p-m_\ell^*}\frac{(-1)^{q+r-2p-m_\ell^*}}{\Gamma(q+m_\ell^*-\beta_{m-\ell}+1)\Gamma(r-2p-m_\ell^*-q+1)}\left[2\left(\frac{t-a}{b-a}\right)\right]^q & m_\ell^*<r-2p\end{cases}$$

For I_2: Using the following transform to interchange the bounded points of integral:

$$\xi=2\left(\frac{t-a}{b-a}\right)-1;\qquad t=\frac{b-a}{2}\xi+\frac{b+a}{2}$$

Thus

$$(3.16)\qquad \begin{aligned} I_2 &= \int_a^x K_m(x,t)T_r\left[2\left(\tfrac{t-a}{b-a}\right)-1\right]dt \\ &= \tfrac{b-a}{2}\int_{-1}^{2\left(\frac{x-a}{b-a}\right)-1} K_m\left(x,\tfrac{b-a}{2}\xi+\tfrac{b+a}{2}\right)T_r(\xi)d\xi \end{aligned}$$

Finally, apply Romberg integrations rule to evaluate all integrals in equations (3.15) and (3.16) at each points $x = x_j$ for all j.

The Algorithm [ACP]: The approximate solution of multi-higher order linear VIFDEs with variable coefficients (1.1) with the use of Chebyshev polynomials in least-square orthogonal method can be summarized by the following stages:

- **Step 1:** Evaluate $\psi_r(x_j)$, using equation (3.14) for all $r = 0,1,2,\ldots,N$ and x_j's are defined by equations: (3.11) for open formed and (3.13) for closed formed.
- **Step 2:** From equations (3.10) and (3.12), compute a_{rs} and b_s for all $r,s = 0,1,2,\ldots,N$ for open and closed Gauss-Chebyshev formula.
- **Step 3:** Construct the matrices D and E which represented in system (3.9).
- **Step 4:** For constant coefficients $C_r(r = (\overline{0:N}))$ apply any numerical methods for system which obtained in step 3 after multiply both sides by D^T.
- **Step 5:** To obtain approximate solution $y_N(x)$ of $y(x)$, substituting C_r's in equation (3.1) where $\varphi_r = T_r^*$.

3.2. **Using Legendre Polynomials.** In this part, we can take instead of trial function $\psi_r(x)$ the shifted Legendre polynomial,$P_r^*(x) = P_r\left[2\left(\frac{x-a}{b-a}\right)-1\right]$,to approximate the solution $y(x)$ of equation (1.1) by formed :

$$y_N(x) = \sum_{r=0}^{N} C_r P_r\left[2\left(\frac{x-a}{b-a}\right)-1\right]\ , \qquad a \le x \le b$$

By applying the same stages as in Chebyshev polynomials technique and using Gauss-Legendre formula (2.6), we get:

$$(3.17)\qquad \begin{cases} a_{rs} = \frac{b-a}{2}\sum\limits_{j=0}^{M-1}\lambda_j\psi_r(x_j)\psi_s(x_j) \\ b_s = \frac{b-a}{2}\sum\limits_{j=0}^{M-1}\lambda_j\psi_s(x_j)f(x_j) \end{cases}$$

where, M is the number of Legendre zeros we can take it and

$$(3.18)\qquad x_j = \left(\frac{b-a}{2}\right)z_j + \left(\frac{b+a}{2}\right)$$

Here z_j are the M-th zeros of $P_M(x)$ and λ_j are defined as in equation (2.7), that is:

$$\lambda_j = \frac{2}{(1-z_j^2)\left[P_M'(z_j)\right]^2}\ , \quad j = 0,1,2,\ldots,M-1$$

To evaluate $\psi_r(x)$ at $x = x_j$ for all r and j , from Gauss-Legendre formula:

$$(3.19)\qquad \begin{aligned} [\psi_r(x)]_{x=x_j} = &\left\{{}_a^C D_x^{\alpha_n} + \sum_{i=1}^{n-1} P_i(x) {}_a^C D_x^{\alpha_{n-i}} + P_n(x)\right\} P_r\left[2\left(\tfrac{x-a}{b-a}\right)-1\right] \\ &-\lambda\int_a^x \sum_{\ell=0}^{m} K_\ell(x,t) {}_a^C D_t^{\beta_{m-\ell}} P_r\left[2\left(\tfrac{t-a}{b-a}\right)-1\right]dt \end{aligned}$$

Apply same stage as to solve equation (3.14) with using lemma (2.6) we can obtain the values of equation (3.19).

The Algorithm [ALP]: The approximate solution of multi-higher order linear VIFDEs with variable coefficients (1.1) with the use of Legendre polynomials in least-square orthogonal method can be summarized by the following stages:

Step 1: Evaluate $\psi_r(x_j)$, using equation (3.19) for all $r = 0, 1, 2, \ldots, N$ and x_j's are defined by equations: (3.18).

Step 2: From equations (3.17) compute a_{rs} and b_s for all $r, s = 0, 1, 2, \ldots, N$ for Gauss-Legendre formula.

Step 3: Construct the matrices D and E which represented in system (3.9).

Step 4: For constant coefficients $C_r(r = (\overline{0:N}))$ apply any numerical methods for system which obtained in step 3 after multiply both sides by D^T.

Step 5: To obtain approximate solution $y_N(x)$ of $y(x)$, substituting C_r's in equation (3.1) where $\varphi_r = P_r^*$.

4. Numerical Examples

Here, three numerical results are presented for multi-higher linear VIFDEs with variable coefficients. Their results are obtained by applying the algorithms (**ACP** and **ALP**), respectively.

Example 4.1. *Consider a linear VIFDE for fractional order lies in* $(0,1)$ *on closed interval* $[0,1]$:

$$
{}^{C}D_x^{0.6}y(x) + (x-1)y(x) = f(x)
$$

$$
+\int_0^x \left[(x^2+t)^{C}D_t^{0.7}y(t) + (1-xt)^{C}D_t^{0.4}y(t) - sin(x-t)y(t)\right] dt
$$

where

$$
f(x) = 2cosx + x^3 - 2 + \frac{2}{\Gamma(2.4)}x^{1.4} - \frac{1}{5\Gamma(4.3)}[33x+23]x^{3.3}
$$

$$
-\frac{2}{\Gamma(4.6)}[18-13x^2]x^{2.6}
$$

together with the initial condition : $y(0) = 0$*; while the exact solution is* $y(x) = x^2$.

Take $N = 2, M = 5$ and $IN = 7$ (number of approximate parts for integrals using Romberg rule). Assume the approximate solution in the form:

$$
y_2(x) = \sum_{r=0}^{2} C_r T_r(2x-1) \quad and \quad y_2(x) = \sum_{r=0}^{2} C_r P_r(2x-1)
$$

Apply algorithms **ACP** and **ALP** to find the approximate solution of above problem by running the programs which written for this purpose in MatLab, table (1) present the values of C_r's, respectively.

Thus, we get the following approximate formulas:

$$
y_2^{Open}(x) = -0.102e^{-5}x + 1.00000184x^2
$$

$$
y_2^{Closed}(x) = -0.1e^{-7} - 0.106e^{-5}x + 1.00000176x^2
$$

$$
y_2^{Leg}(x) = -0.84e^{-6}x + 1.0000014x^2
$$

Table (2) show the comparison between the exact solution $y(x)$ and approximate solution $y_2(x)$ for all open ,closed Chebyshev and Legendre respectively depending

TABLE 1

LSOR Method	C_r's		
	C_0	C_1	C_2
OHR	0.37500018	0.50000041	0.12500023
CHR	0.37500012	0.50000035	0.12500022
LER	0.33333338	0.50000028	0.16666690

TABLE 2

x -Points	Exact Solution	Least-Square Orthogonal Method		
		OHR	CHR	LER
0	0.00	0.0	-9.99999 e-009	0.0
0.1	0.01	0.0099999164	0.0099999016	0.00999993
0.2	0.04	0.0399998696	0.0399998484	0.039999888
0.3	0.09	0.0899998596	0.0899998304	0.089999874
0.4	0.16	0.1599998864	0.1599998476	0.159999888
0.5	0.25	0.24999995	0.2499999	0.24999993
0.6	0.36	0.3600000504	0.3599999876	0.36
0.7	0.49	0.4900001876	0.4900001104	0.490000098
0.8	0.64	0.6400003616	0.6400002684	0.640000224
0.9	0.81	0.8100005724	0.8100004616	0.810000378
1.0	1.00	1.00000082	1.00000069	1.00000056
L.S.E		0.122764 e-011	0 .868310 e-012	0 .567028 e-012
$R_2 = L.S.E_y$		0.231718 e-011	0 .245802 e-011	0 .449124 e-011
R.Time/Sec		2.601344	2.766926	2.655042

on the least square error and running time. The values residual equations $R_2(x;C)$ are also included by applying the formula (3.3).

Note that, here if we take $IN = 10$ we obtain the exact solution $y_2(x)$ only for open and closed Chebyshev, which x^2 but not in Legendre used.

Example 4.2. *Consider the linear VIFDE on* $0 \le x \le 1$:

$$^{C}D_x^{2\alpha}y(x) - \frac{1}{2}{}^{C}D_x^{\alpha}y(x) + (1+x^2)y(x) = f(x)$$

$$+\int_0^x \left[xt\,{}^{C}D_t^{2\alpha}y(t) + (x^2-t)\,{}^{C}D_t^{\alpha}y(t) + e^{x+t}y(t)\right] dt$$

where

$$f(x) = x^5 + x^3 - x^2 - 1 - 7e^x - e^{2x}(x^3 - 3x^2 + 6x - 7)$$

$$+\frac{6}{\Gamma(4-2\alpha)}\left(1 - \frac{1}{5-2\alpha}x^3\right)x^{3-2\alpha} + \frac{3}{\Gamma(4-\alpha)}\left(\frac{2}{5-\alpha}x^2 - 1\right)x^{3-\alpha} - \frac{6}{\Gamma(5-\alpha)}x^{6-\alpha}$$

with the initial conditions: if $0 < \alpha < 0.5$; $y(0) = -1$ *if* $0.5 < \alpha \le 1$; $y(0) = -1$ *and* $y'(0) = 0$. *The exact solution of this problem is known* $y(x) = x^3 - 1$.

we take two different values of fractional order α:

Table 3

x -Points	Exact Solution	Least-Square Orthogonal Method		
		OHR	**CHR**	**LER**
0	-1	-1.000000059	-1.00000005	-0.999999996
0.1	-0.999	-0.999000050192	-0.99900003984	-0.99899999368
0.2	-0.992	-0.992000034856	-0.99200002392	-0.99199998704
0.3	-0.973	-0.973000017024	-0.97300000608	-0.97299997896
0.4	-0.936	-0.936000000728	-0.93599999016	-0.93599997232
0.5	-0.875	-0.87499999	-0.87499998	-0.87499997
0.6	-0.784	-0.783999988872	-0.78399997944	-0.78399997488
0.7	-0.657	-0.657000001376	-0.65699999232	-0.65699998984
0.8	-0.488	-0.488000031544	-0.48800002248	-0.48800001776
0.9	-0.271	-0.271000083408	-0.27100007376	-0.27100006152
1.0	0	-1.60999 e-007	-1.499999 e-007	-1.24000 e-007
L.S.E		0.416042 e-013	0 .341208 e-013	0 .225431 e-013
$R_3 = L.S.E_y$		0.308311 e-012	0 .270284 e-012	0 .163582 e-011
R.Time/Sec		1.829995	1.871905	1.864961

• For $\alpha = 0.6$, take $N = 3, M = 4$,and $IN = 5$. So apply algorithms **ACP** and **ALP** to find the approximate solution and we obtain the following approximate formulas:

$$y_3^{Open}(x) = -1.000000059 + 0.42e^{-7}x + 0.528e^{-6}x^2 + 0.999999328x^3$$
$$y_3^{Closed}(x) = -1.00000005 + 0.6e^{-7}x + 0.48e^{-6}x^2 + 0.99999936x^3$$
$$y_3^{Leg}(x) = -0.999999996 - 0.8e^{-8}x + 0.36e^{-6}x^2 + 0.99999952x^3$$

Table (3) list the results obtained by running the programs for comparison the approximate solutions of the above problem .Included are the least square error and running time with the values of residual $R_3(x; \overline{C})$ for comparison.

• For $\alpha = 0.4$, take $N = 3, M = 4$,and $IN = 5$.Assume the approximation solution same as before, by running programs the result of least square orthogonal methods are obtained as follows:Thus, we get the following approximate formulas:

$$y_3^{Open}(x) = -1.0 - 0.8e^{-7}x + 0.32e^{-6}x^2 + 0.99999968x^3$$
$$y_3^{Closed}(x) = -0.999999998 - 0.116e^{-6}x + 0.416e^{-6}x^2 + 0.999999616x^3$$
$$y_3^{Leg}(x) = -1.0000000055 - 0.4e^{-7}x + 0.27e^{-6}x^2 + 0.9999997x^3$$

Table (4) shows a comparison between the exact solution y(x) and approximate solutions $y_3(x)$ for all open, closed Chebyshev and Legendre polynomials respectively.

Note that for $\alpha = 0.6$ and $\alpha = 0.4$, if we take $M = 5$ and $IN = 10$ we obtain the exact solution for all types.

Example 4.3. *Consider a higher-order linear VIFDE's on bounded interval* $[0, 1]$*:*

$$^{C}D_x^{1.6}y(x) - \frac{1}{6}{}^{C}D_x^{1.2}y(x) + x^{C}D_x^{0.5}y(x) + cosxy(x) = f(x)$$
$$+ \int_0^x \left[5xt^{C}D_t^{0.8}y(t) + (x^2 + t^2)^{C}D_t^{0.3}y(t) - e^{2x-t}y(t)\right] dt$$

TABLE 4

x -Points	Exact Solution	Least-Square Orthogonal Method OHR	CHR	LER
0	-1	-1	-0.999999998	-1.000000005
0.1	-0.999	-0.99900000512	-0.999000005824	-0.9990000066
0.2	-0.992	-0.99200000576	-0.992000007632	-0.9920000046
0.3	-0.973	-0.97300000384	-0.973000005728	-0.9730000008
0.4	-0.936	-0.93600000128	-0.936000002416	-0.935999997
0.5	-0.875	-0.875	-0.875	-0.874999995
0.6	-0.784	-0.78400000192	-0.784000000784	-0.7839999966
0.7	-0.657	-0.65700000896	-0.657000007072	-0.6570000036
0.8	-0.488	-0.48800002304	-0.488000021168	-0.4880000178
0.9	-0.271	-0.27100004608	-0.271000045376	-0.271000041
1.0	0	-7.999999 e-008	-8.2000 e-008	-7.50000 e-008
L.S.E		0.921395 e-014	0.941651 e-014	0.77717 e-014
$R_3 = L.S.E_y$		0.730238 e-013	0.402558 e-013	0.44232 e-013
R.Time/Sec		1.868506	1.885245	1.892130

TABLE 5

LSOR Method	C_r's C_0	C_1	C_2	C_3
OHR	1.6874989	0.53125069	-0.18750006	-0.3125 e-1
CHR	1.6874990	0.53125055	-0.18750009	-0.3125 e-1
LER	1.7499998	0.54999976	-0.25000045	-0.5000 e-1

where

$$f(x) = \frac{1}{\Gamma(2.8)}x^{1.8} - \frac{6}{\Gamma(2.4)}x^{1.4} + \frac{2}{\Gamma(1.5)}x^{1.5} - \frac{6}{\Gamma(3.5)}x^{3.5}$$
$$-12\left(\frac{1}{\Gamma(3.2)} - \frac{8}{\Gamma(5.2)}x^2\right)x^{3.2} - 6\left(\frac{4.86}{\Gamma(4.7)} - \frac{44.18}{\Gamma(6.7)}x^2\right)x^{3.7}$$
$$+\left(3+4x+3x^2+x^3\right)e^x - 3e^{2x} + \left(1+2x-x^3\right)cosx$$

with the initial conditions : $y(0) = 1$, $y^{'}(0) = 2$. *while the exact solution is* $y(x) = 1 + 2x - x^3$.

Which is multi-higher order linear VIFDEs with variable coefficients. Take $N = 3$, $M = 5$,and $IN = 8$, Assume the approximate solution in the forms:
• For Chebyshev polynomial $y_3(x) = \sum_{r=0}^{3} C_r T_r(2x-1)$
• For Legender polynomial $y_3(x) = \sum_{r=0}^{3} C_r P_r(2x-1)$
Run programs to find the approximate solution of the above problem, table(5) present the values C_r's , respectively .

Thus, we get the following approximate formulas:

$$y_3^{Open}(x) = 0.999998235 + 2.00000033x + 0.36e^{-5}x^2 - 1.00000272x^3$$
$$y_3^{Closed}(x) = 0.999998448 + 2.000000236x + 0.3504e^{-5}x^2 - 1.000002816x^3$$
$$y_3^{Leg}(x) = 0.999999772 + 2.000000036x + 0.276e^{-5}x^2 - 1.00000364x^3$$

Table 6

x -Points	Exact Solution	Least-Square Orthogonal Method		
		OHR	**CHR**	**LER**
0	1	0.999998235	0.999998448	0.999999772
0.1	1.199	1.19899830128	1.19899850382	1.19899979956
0.2	1.392	1.39199842324	1.39199861283	1.39199986048
0.3	1.573	1.57299858456	1.57299875813	1.57299993292
0.4	1.736	1.73599876892	1.73599892282	1.73599999504
0.5	1.875	1.87499896	1.87499909	1.875000025
0.6	1.984	1.98399914148	1.98399924278	1.98400000096
0.7	2.057	2.05699929704	2.05699936427	2.05699990108
0.8	2.088	2.08799941036	2.08799943757	2.08799970352
0.9	2.071	2.07099946512	2.07099944578	2.07099938644
1	2	1.999999445	1.999999372	1.999998928
L.S.E		0.152607 e-010	0.120976 e-010	0.174010 e-011
$R_3 = L.S.E_y$		0.140318 e-010	0.319769 e-010	0.199853 e-009
R.Time/Sec		5.774807	5.944966	5.891868

Table (6) present a comparison between the exact solution $y(x)$ and numerical solution $y_3(x)$ for all open ,closed Chebyshev and Legendre respectively depending on the least square error and running time. Furthermore, it included the values of residual equations $R_3(x; C)$ by applying formula (3.3).

Note that for $IN = 15$ we can get the exact solution for all types.

5. Discussion

In this chapter, three numerical algorithms have been applied to solve the multi-higher order of linear VIFDEs with variable coefficients. For each algorithm, a computer program was written and several examples are included for illustration and good results are achieved. The least square error, least square error function (y) and running time are all also given in tabular forms. The following points have been identified.

1: This method can be used even where there is no information about the exact solution (from the error function $R_N(x; \overline{C})$ in equation (3.3)).

2: The good results depend on: number of approximate parts of integral IN (if we take IN a large number we obtain the exact solution of $y_N(x)$) and the number of the orthogonal polynomials N ,with suitable number of polynomial zeros M .

3: By running the programs of **ACR** and **ALR** the least square error function in Chebychev polynomials gives more accurate solutions than Legendre polynomial method, see tables (2), (3), (4) and (6), so it is better.

4: In general, the solution by Chebyshev polynomials is easier and faster than of Legendre polynomials.

References

[1] AL-Ani, N.W.J; *Numerical Methods for Solving Fredholm Integral Equations*, M.Sc. Thesis, Technology University, September 1996.

[2] AL-Kalwi, P.M; *Numerical Method for Solving Fredholm Integral Equations of the Second Kind*, M.Sc. Thesis, Technology University, September 1999 .
[3] AL-Rawi, S.N; *Numerical Solution of First Kind Integral Equations of Convolutions Type*, M.Sc. Thesis, Technology University, 1995.
[4] A.R. Nkammang; *Discretization of multi-term fractional integral and ordinary differential equations*, Ph.D. Thesis, FU Berlin; 1999.
[5] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo; *Theory and Applications of Fractional differential Equations*, Elsevier B.V. Netherlands, 2006.
[6] I.podldubny; *Fractional Differential Equation*, Academic press, San Diego, 1999.
[7] J. C. Mason and D. C. Handscomb; *Chebyshev polynomials*, chapman & HaLL /CRC press company, 2003.
[8] J.F. Epperson; *An Introduction to Numerical Methods and Analysis*, John wiley & Sons, Inc. Printed in the United States of America; 2002 .
[9] L.M. Delves and J.L. Mohamed; *computational methods for integral equations*, Cambridge University press, 1985.
[10] M. Weilbeer; *Efficient Numerical Methods for Fractional Differential Equations and their analytical Background*, US Army Medical Research and Material command , 2005 .
[11] R.L. Burden and J.D. Faires; *Numerical Analysis*, third editions , An International Thomson publishing company (ITP), 1997 .
[12] S.S. Ahmed; *Numerical Solutions of Liner Volterra Intergro-differential Equations*, M.Sc. Thesis, Technology University, 2002 .
[13] S.S. Ahmed; *On System of Linear Volterra Intergro-Fractional Differential Equations*, Ph.D. Thesis, Sulaimani University, 2009.
[14] T.I. Hassan*Numerical Methods for Solving Fredholm Integral Equation of the First Kind Degenerate Kernel*, M.Sc. Thesis, Salahaddin University, 2005 .
[15] S.A. Hama Salih, *Some Computational Methods for Solving Linear Volterra Integro-Fractional Differential Equations* ,M.Sc. Thesis,Sulaimani University, 2011
[16] Z.M. Odibat and N.T. Shawagfeh, Generalized Taylor's Formula, *Applied Mathematics and Computation*, Vol. 186, pp. 286-293(2007).

(S.S. Ahmed) University of Sulaimani, Faculty of Science and Science education, Sulaimani, Iraq
E-mail address: math_safari@yahoo.com

(S.A. Hama Salih) University of Sulaimani, Faculty of Science and Science education, Sulaimani, Iraq
E-mail address: shokan_80_78@yahoo.com

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 230-245, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

COMPARING SOME ROBUST METHODS WITH OLS METHOD IN MULTIPLE REGRESSION WITH APPLICATION

SIZAR ABED MOHAMMED

ABSTRACT. The classical method, Ordinary Least Squares (OLS) is used to estimate the parameters of the linear regression, when assumptions are available, and its estimators have good properties, like unbiasedness, minimum variance, consistency, and so on. The alternative statistical techniques have been developed to estimate the parameters, when the data is contaminated with outliers. These are the robust (or resistant) methods. In this paper, three of robust methods are studied, which are: Maximum likelihood type estimate M-estimator, Modified Maximum likelihood type estimate MM-estimator and Least Trimmed Squares LTS-estimator, and their results are compared with OLS method. These methods applied to real data taken from the Charstin company for manufacturing furniture and wooden doors, the obtained results compared by using the criteria: Mean Squared Error (MSE), Mean Absolute Percentage Error (MAPE) and Mean Sum of Absolute Error (MSAE). The important conclusions that this study came up with are: the number of outlier values detected by using the four methods in the data for furniture's line are very close. This refers to the fact that the distribution of standard errors is close to the normal, but the outlier values found in the data for doors line, by using OLS are less than which detected by robust methods. This means that the distribution of standard errors is departure distant from the normal. The other important conclusion is that estimated values of parameters by using OLS are very far from its estimated values by using the robust methods with respect to doors line, the LTS-estimator gave better results by using MSE criterion, and M-estimator gave better results by using MAPE criterion. Further more, it has noticed that by using the criterion MSAE, the MM-estimator is better. The programs S-plus (version 8.0, professional 2007), Minitab (version 13.2) and SPSS (version 17) are used to analyze the data.

1. INTRODUCTION

Regression is one of the most commonly used statistical techniques. Out of many possible regression techniques, the Ordinary Least Squares OLS method has been generally adopted because of tradition and ease of computation. However, OLS estimation of regression weights in the multiple regression are affected by the occurrence of outliers, non-normality, multicollinearity and missing data. Outliers are observations that appear inconsistent with the rest of the data. The influential points remain hidden to the user, because they do not always show up in the usual least square residual plots. To remedy this problem new statistical techniques have been developed that are not easily affected by outliers. There are many robust methods such as: Least Median Squares estimates LMS, Least Trimmed Squares estimates LTS, Least Winsor Squares estimates LWS, Rank statistics estimates

Key words and phrases. Robust Regression, M-estimate, MM-estimate, LTS-estimate.

R-estimate, Maximum likelihood type estimate M-estimate, Laplace statistics estimates L-estimate, an adaptive maximum likelihood type estimates, Scale statistics estimates S- estimate, Modified Maximum likelihood type estimate MM- estimates and Generalized Maximum likelihood estimate GM- estimates. Through the important observations of robustness, the indicated deletion seen in outlier in the data is not a solution and it became a topic of the past that can not be adopted. The main purpose of robust regression is to provide resistant parameters in the presence of outliers. In order to achieve this stability, robust regression limits the influence of outliers. In this paper three of robust methods are studied, M-estimate, MM-estimate and LTS-estimate, and their results are compared with OLS method. These methods are applied to the real data taken from the Charstin Company for the manufacture of furniture and wooden doors located in the city of Duhok. The aim of this paper is studying some robust methods of parameters estimation of linear regression in the presence of outliers in data, and comparing the results with OLS estimators

2. The Basic Concepts

2.1. **Outliers.** [4, 7, 21]. Any statistician, when analyzing a set of real data may encounter outlier values called observations that deviate significantly from any other observations, which are not consistent with other observations. Those values may be large or small and may result from an error in recording data or the preview or the appearance caused by other natural causes. There is no agreement on a specific definition for outlier value. However it is called contaminated, conflicting observation, discordant observation or irregular values. This concept has many definition. For example, it was defined by Grubbs, 1969 as the value that looks significantly deviant (Deviate Markedly) from the other simple observations, it arises from heavy-tailed distributions or mixed distributions. The presence of outliers in a sample of several variables is more complex than a single variable case. The presence of outliers in the data leads to great confusion in the analysis of the data in the case of using traditional methods in the estimation. One of these methods is the least square method. Huber P.J., 1973 showed the effect of outlier values in the estimation of least squares through his famous saying "the existence of an outlier value damages the good characteristics of the estimation of least squares it also leads to achieving success for LS method".

2.2. **Regression outliers.** [1, 21, 23]. Regression outliers are those observations that do not trace the model and the outlier value results from Yi unusual or Xi unusual or both as: Rousseeuw (1987) made it obvious by examples of the impact of outlier in the estimation of the least squares and expounded how one observation changes from the direction of the line of least squares. Chatterjee and Hadi, (1988) knew that the outliers of regression those observations that have a great standardized residual compared to the rest of the other observations. (Draper and Smith) knew that an outlier observation in linear regression is that observation in which the absolute value of residuals is much larger than the rest of the absolute values of the rest of the other observations and perhaps it may become away from the average value of those residuals 3σ by 4σ or more.

2.3. **Leverage points.** [21]. They are those points which are located away from most of values in the matrix X in the form $Y = X\beta + e$ which include one or more of the explanatory variables in regression analysis and have a strong influence on the estimation of ordinary least squares. When observation X_i is an outlier point, we call an observation (X_k,Y_k) a leverage point wheneverX_k lies far away from the bulk of the observation X_i in the sample, note that this does take Yk into account, so the point (X_k,Y_k) does not necessarily to be a regression outlier, when (X_k,Y_k) lies close to the regression line determined by the majority of the data, then it can be considered a good leverage point. Therefore, to say that (X_k,Y_k) is a leverage point refers only to its potential for strongly affecting the regression coefficients , but it does not necessarily mean that (X_k,Y_k) will actually have a large influence on , because it may be perfectly on line with the trend set by the other data. In multiple regression, the (X_{i1},â€|,X_{ip}) lie in a space with p dimension, a leverage point is then still defined as a point (X_{k1},â€|,X_{kp},Y_k) for which (X_{k1},â€|,X_{kp}) is outlying with respect to the (X_{i1},â€|,X_{ip}) in the data set. As before, such leverage points have a potentially large influence on the LS regression coefficients, depending on the actual value of Y_k.

2.4. **Breakdown points.** [21]. A useful measure of robustness is the breakdown value, its oldest definition Hodges (1967) was restricted to one dimensional estimation of location, where as Hampel (1971) gave a much more general formulation. Unfortunately, the latter definition was asymptotic and rather mathematical in nature, which may have restricted its dissemination. This study used the simple finite-sample version of the breakdown point, introduced by Donoho and Huber (1983). $Z = \{(X_{11}, X_{12}, ..., X_{1p}, Y_1), ..., (X_{n1}, X_{n2}, ..., X_{np}, Y_n)\}$ and we want to estimate its for this purpose, we apply a Take any sample of n points in p dimension: translation equivariant . This means that applying T to such a sample Z yields a vector$\left(\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p\right)$ of regression coefficients$T(Z) = \hat{\beta}$, now we consider all corrupted samples Z' obtained by replacing any m of the original data points by arbitrary values, and we define the maximal bias by max bias $(m; T, Z) = Sup_{z'} ||T(Z\prime) - T(Z)||$, if bias $(m; T, Z)$ is infinite, this means that m outliers can have an arbitrarily large effect on T, it follows that maxbias$(m; T, Z) = \infty$, hence $T\left(Z'\right)$ becomes useless. Therefore, the finite sample breakdown point of the estimator T at the sample Z is defined as: $\zeta_n^*(T, Z) = min\left\{\frac{m}{n}; bias(m; T, Z) = \infty\right\}$. Is again the smallest fraction of contamination that can cause T to take on values arbitrarily far away from $T(Z)$. obviously, the multivariate arithmetic mean possesses a breakdown point of $\frac{1}{n}$. We often consider the limiting breakdown point for $n \to \infty$ by $\zeta^*(T)$, so we say that themultivariate mean has 0 breakdown. It is clear that no translation equivariant$T$ can have a breakdown point large than 0.50. In order to be able to estimate the original parameters $\beta_0, ..., \beta_p$ we need that $\zeta < \zeta^*(T)$. For this reason $\zeta^*(T)$ is some times called the breakdown bound of T. For least squares, we have seen that one outlier is sufficient to carry T over all bounds. Therefore, its breakdown point equals $\zeta_n^* = \frac{1}{n}$ which tends to zero for increasing sample size n, so it can be said that LS has breakdown point of 0. This again reflects the extreme sensitivity of the LS method to outliers.

2.5. **Criteria for judging accuracy of estimation method.** [10, 19]. In cases of appreciation of applied precision treated as a criterion to test the method of estimation, there are several reasons that lead to inaccuracies example of this data or the use of non-sufficient procedure does not match the data type. Accuracy is the most widely used criterion to evaluate the achievement of the estimation methods which reflect the health of predictable, but the difficulties attached to standard of accuracy in the estimate is the absence of a single measurement, acceptable, and the accuracy of the complex at the some time. There is a number of criteria commonly used by statisticians, including:

2.5.1. *Mean Squared Error (MSE).* : The mean squared error is a measure of accuracy computed by squaring the individual error for each item in a data set and then finding the average or mean value of the sum of those squares. The mean squared error gives greater weight to large errors than to small errors because the errors are squared before being summed. It takes, the following formula:

$$MSE = \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n-p-1} = \frac{\sum_{i=1}^{n} (e_i)^2}{n-p-1}$$

2.5.2. *Mean Absolute Percentage Error (MAPE).* : The mean absolute percentage error is the mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign, i.e.,(their absolute value is summed and the average computed). It takes the following formula:

$$MAPE = \frac{\sum_{i=1}^{n} |PE_i|}{n} * 100$$

where $PE_i = \frac{\left(Y_i - \hat{Y}_i\right)}{Y_i}$

2.5.3. *Mean Sum of Absolute Error (MSAE).* : It is a common standard to differentiate between the method of least squares and the method of absolute error and it takes the following formula:

$$MSAE = \frac{\sum_{i=1}^{n} |Y_i - \hat{Y}_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}$$

It should be mentioned that minimizing the average absolute errors as to minimize the total absolute errors and this is sometimes called the Least Absolute Value (LAV).

2.5.4. *The Ordinary Least Squares method (OLS).* :[14, 23]. The basic idea of the ordinary least squares method is to make the sum of squares of errors (SSe) as small as possible:

$$min \sum_{i=1}^{n} e_i, i = 1, 2, 3, ..., n$$

Therefore Gauss preferred it to other methods, because of easy estimation of parameters from data. He also assumed that e_i as a variable follows the normal distribution, so the method of OLS is practical idea, if all the assumptions related to it are achieved. If one or more of these assumptions are not achieved, this will lead to getting inaccurate estimation of parameters β of the model and can not be adopted:

$$e = Y - X\hat{\beta} \rightarrow e = Y - \hat{Y}$$

The estimation of the parameters β that make SSe minimum as possible is called the least squares normal estimation $\hat{\beta}_{OLS}$ which produces a (p) solution of simultaneous equations can be given as:
$\hat{\beta}_{OLS} = \left(X'X\right)^{-1} X'Y$, then the form $Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ... + \beta_p X_{ip} + e_i$ is called linear model of full rank regression models. The specific determinate second partial derivative of the vector $\hat{\beta}_{OLS}$ will always be greater than zero $|\frac{\partial^2 \left(ee'\right)}{\partial \beta^2}| > 0$.

2.5.5. *Assumptions of least squares method.* :[15, 17, 22, 23]. There are a number of assumptions based on the study usual method of OLS in linear regression model:
A-Assumptions about the error:
1- Mathematical expectation of the random error is equal to zero $E\left(e_i\right) = 0, i = 1, 2, ..., n$
It follows that: $E\left(Y_i\right) = E\left(\beta_0 + \beta_1 X_i + e_i\right) = \hat{\beta}_0 + \hat{\beta}_1 X_i + E\left(e_i\right) = \hat{\beta}_0 + \hat{\beta}_1 X_i$
2– Contrast the values of a random variable to be constant in each period of time $Var\left(e_i\right) = E\left(e_i^2\right) = \sigma^2 i = 1, 2, \hat{a}e|, n$. This is called homoscedasticity of error variation. It therefore follows that the variance of the response variable Y_i is: $\sigma^2\left(Y_i\right) = \sigma^2$,since $\sigma^2\left(\beta_0 + \beta_1 X_i + e_i\right) = \sigma^2\left(e_i\right) = \sigma^2$.
3-A random variable ei distributed as normal distribution with mean zero and variance σ^2 , $e_i \sim N\left(0, \sigma^2\right), i = 1, 2, \hat{a}e|, n$
4-If for any $i \neq j$, e_i, e_j are independent, then there is no autocorrelation between e_i and e_j , hence the outcome in any one trial has no effect on the error term for any other trial as to whether it is positive or negative or small or large since the error terms e_i and e_j are uncorrelated, this means that the covariance between them is equal to zero $Cov\left(e_i, e_j\right) = 0, i \neq j = 1, 2, ..., n,$
5-Independence between the explanatory variables and random variables, $E\left(e_i, X_{ij}\right) = 0$,that is to say e_i independent of X_{ij} for all different values of i, this means there is no problem of multicollinearity.
B- Assumptions on the distribution of the response variable Y:
1-Average Y_i is a function of straight line $\bar{Y} = E\left(Y_i\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + ... + \hat{\beta}_p X_p$, $i = 1, 2, ..., n$. where $\hat{\beta}_0, be\hat{t}a_1, ..., be\hat{t}a_p$ are estimates of regression parameters, since the regression function relates the means of the probability distribution of$Y$ for any given X to the level of X.
2-Variance of Y_i ,$Var\left(Y_i\right)$ has one value for any value of i $Var\left(Y_i\right) = \sigma^2$, $i = 1, 2, ..., n.$
3-Response variable Y_i distributed as normal distribution with mean μ and variance σ^2 ,i.e., $Y_i \sim N\left(\mu, \sigma^2\right)$.
4- Any two observations Y_i and Y_j are uncorrelated, this implies that, $Cov\left(Y_i, Y_j\right) = 0$ for all $i \neq j = 1, 2, ..., n$
5- The relationship between X_i's and $\hat{Y}$ be a linear relationship is the equation of straight line $\hat{Y} = E\left(Y_i\right) = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + ... + \hat{\beta}_p X_p$, $i = 1, 2, ..., n$.

2.5.6. *Properties of ordinary least squares method.* :[1, 2, 16, 23].
1-That $\hat{\beta}$ unbiased estimate of the parameter $\beta, E\left(\hat{\beta}\right) = \beta$.
2-Estimators have the least variation among all linear unbiased estimators compared variations of unbiased linear estimators that belong to any other ways, i.e., the best linear unbiased estimates. 3-The consistent destinies mean that the$p_r[\hat{a}", \hat{\beta}_{OLS} - \beta\hat{a}", < \zeta] = 1$ when $n \to \infty$ for each $\zeta > 0$ or when $\hat{\beta}_{OLS} = \beta$.
4-Sufficient destinies mean that the estimate $\hat{\beta}$ called the estimate enough for β if all information about the parameter of the community β is included in the sample data.
5-Matrix of variance- covariance of the estimator $\hat{\beta}$ is: $Var\left(\hat{\beta}\right) = \left(X'X\right)^{-1}\sigma^2$.

3. Some Robust Regression Methods

3.1. **Outliers detection.** : There are many methods to detect the outliers in the data, which are the classic and robust methods; in this paper the robust methods are used. A-Rousseeuw and Van Zomeren [8, 9, 21]. proposed using robust methods to detect the outlier of linear regression as follows:
1- Fit the regression model , then use the LMS estimation method for the detection of the outlier, then the residuals which are calculated for standard least median of squares estimates according to the following:

$$SR_i = \frac{e_i}{\hat{\sigma}}, i = 1, 2, ..., n where \hat{\sigma} = K\sqrt{med e_i^2},$$

where K is arbitrary positive constant.
2- Estimation method uses Minimum Volume Ellipsoid (MVE) estimator to find the lowest volume of the ellipsoid on the matrix regression such as Z if that $X' = \left(1 : Z'\right)$ .MVE estimator looking for smaller pieces containing half of the data which are usually calculated averages and covariance of the points inside the ellipse and then re-measure these estimates so that the estimated average community variability and the fact that when the joint sampling distribution of multivariate normal.
3- Calculate the robust distance of all views as follows:

$$RD\left(X_i\right) = \sqrt{\left(X_i - T\left(X_i\right)\right)' \left(C\left(\sum\right)\right)^{-1} \left(X_i - T\left(X_i\right)\right)},$$

where $T\left(X_i\right)$: estimate MVE vector of multivariate location parameters. $C\left(\sum\right)$: PXP a square matrix represents the estimate MVE of the matrix variance-covariance of the multivariate.
4- Use the graph to discover outliers and leverage points, as the following:

Graph 3.1: The rule that could be adopted in the diagnosis process.

2.5	(A) Regression outlier	(B) Bad leverage point
0	(C) Good leverage point	(D) Good leverage point
-2.5	(E) Regression outlier	(F) Bad leverage point

Graph 3.1. The rule that could be adopted in the diagnosis process. Graph 3.1. shows that the cut point for standardized LMS residuals is (Â±2.5), the cut point of the robust distance $RD\left(X_i\right)$ is $\sqrt{\chi^2_{(1-\alpha\chi\nu)}}$ if $RD\left(X_i\right) > \sqrt{\chi^2_{(1-\alpha\chi\nu)}}$ then X_i is

outlier values.

The graph 4.1. divided into six regions and views are classified as follows: â€¢ Observation i is good leverage point if it takes place in C. â€¢ Observation i is good leverage point if it takes place in D. â€¢ Observation i is regression outliers if it takes place in A or E. â€¢ Observation i is bad leverage point if it takes place in B or F. B- Rousseeuw and Leroy,1987 [21] used a method to detect outliers in multiple linear regression method as follows: 1-Apply one of robust methods. 2-Find the estimated values for the response variable.

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + ... + \hat{\beta}_p X_{ip}, i = 1, 2, ..., n$$

3- Residuals which are calculated for standardLMS according to the following: $SR_i = \frac{e_i}{\hat{\sigma}}$, $i = 1, 2, ..., n$ where $\hat{\sigma} = K\sqrt{mede_i^2}$

Graph 3.1. shows that the cut point for standardized LMS residuals is the (Â±2.5), therefore the observation which lie outside this region is outlier.

3.2. **Robust estimation methods.** [3, 5, 8, 11, 12, 13, 15, 18, 20, 21, 23]. One of the main concerns of knowledge of statistics is that analyzing data for the study or scientific research and interpretation of results get to interpret the data based on rules and modalities and in some cases the researcher faces a problem in data such as departure sample data distribution is assumed because of the outlier values or distribution community under study is supposed to distribution. This is because the deviations from the assumptions will come and put the traditional methods for estimating the parameters of the model. The researchers found that these methods are not efficient in the case of failure to achieve one of the assumptions or conditions upon which these methods work, so the researchers had to find more efficient methods and not affected much deviations from specific assumptions. These methods are called robust methods. Although there are different ways of robust regression, but most of them share two basic points, one of them gives less weight to the view that the outlier found and the other is to use an iterative method.

3.2.1. *Maximum likelihood estimator (M-estimator):* M-estimators are based on the idea of replacing the squared residuals e_i^2 used in least squared LS estimation, which $minimize_{\hat{\beta}} \sum_{i=1}^{n} e_i^2$,by another function of residuals.
That $minimize_{\hat{\beta}} \sum_{i=1}^{n} \rho(e_i)$ where ρ is a symmetric function, i.e., $\rho(-t) = \rho(t)$ for all t , with a unique minimum at zero. Differentiating this expression with respect to the regression coefficients $\hat{\beta}_j$ yields, $\sum_{i=1}^{n} \psi(e_i) X_i = 0$, where ψ is the derivative of ρ , i.e. $\psi(X, \beta) = \frac{\partial \rho(X, \beta)}{\partial \beta}$, and X_i is the row vector of explanatory variables of the i-th observation $X_i = (X_{i1}, X_{i2}, \hat{a}e|, X_{ip})$, therefore $\sum_{i=1}^{n} \psi(e_i) X_i = 0$ is really a system of p-equations, the M-estimate is obtained by solving the system of p-equations, the solution of which is not always easy to find. In practice, Newton-Raphson and iteratively reweighted least squares are the two methods to solve the M-estimated nonlinear normal equations. Iteratively reweighted least squares express the normal equations as: $X'WX\hat{\beta} = X'WY$ where W is an nxn diagonal

matrix of weights:

$$W_i = \frac{\psi\left(\dfrac{Y_i - X_i'\hat{\beta}_0}{s}\right)}{\left(\dfrac{Y_i - X_i'\hat{\beta}_0}{s}\right)}$$

The initial vector of parameter estimates, $\hat{\beta}_0$ are typically obtained from OLS. Iteratively reweighted least squares updates these parameter estimates with $\hat{\beta} = \left(X'WX\right)^{-1} X'WY$. However the solution of $\sum_{i=1}^{n} \psi\left(e_i\right) X_i = 0$ is not equivalent with respect to a magnification of the Y-axis. Therefore, one has to standardize the residuals by means of some estimate of σ , yielding $minimize \sum_{i=1}^{n} \rho\left(\dfrac{e_i}{\hat{\sigma}}\right)$ Where σ must be estimated simultaneously motivated one possibility is to use the Median Absolute Deviation (MAD) scale estimator: $\hat{\sigma} = cmed\left(|e_i - medr_i|\right)$, where $c = 1.4826$ if Gaussian noise is assumed, or $\hat{\sigma} = 2.1 med|e_i|$ or $\hat{\sigma} = 1.5 med|e_i|$ Huber proposed to use the function $\psi\left(t\right) = min\left(c, max\left(t, -c\right)\right)$ where c is some constant, usually around 1.5, as a consequence. However, the breakdown point of M-estimator is equal to zero.

3.2.2. *Modified Maximum likelihood type estimate (MM-estimator):* It is a class of robust estimators for the linear model, MM-estimation is a special type of M-estimation developed by Yohai V.J. (1987). MM-estimation is a combination of high breakdown value estimation and efficient estimation Yohai's MM-estimators have three stage procedures.

1- The first stage is calculating an S-estimate with influence function.

$$\rho(x) = \begin{cases} 3(\dfrac{x}{c})^2 - 3(\dfrac{x}{c})^4 + (\dfrac{x}{c})^6 & if|x| \leq c, \\ 1 & \text{otherwise.} \end{cases}$$

The value of tuning constant c, is selected as 1.548 Where S-estimate is a high breakdown value of robust regression methods introduced by Rousseeuw and Yohai (1984) that minimizes the dispersion of the residuals. The objective function is min $s\left(e_1\left(\beta\right), e_2\left(\beta\right), ..., e_n\left(\beta\right)\right)$ where $e_i\left(\beta\right)$,is the i-th residuals for candidate β. This objective function is given by the solution, $\dfrac{1}{n-p} \sum_{i=1}^{n} X\left(\dfrac{Y_i - \hat{Y}_i}{s}\right) = k$ where k is a constant.

2- The second stage calculates the MM parameters that provide the minimum value of $\sum_{i=1}^{n} \rho\left(\dfrac{Y_i - \hat{X}_i\hat{\beta}_{MM}}{\hat{\sigma_0}}\right)$ where k is the influence function used in the first stage with tuning constant 4.687 and $\hat{\sigma_0}$ is the estimate of scale from the first step (standard deviation of the residuals). 3- The final step compute the MM-estimate of scale as a solution to $\dfrac{1}{n-p} \sum_{i=1}^{n} \rho\left(\dfrac{Y_i - \hat{X_i}\hat{\beta}}{s}\right) = 0.5$

3.2.3. *Least Trimmed Square Estimator (LTS-estimator):* Least trimmed square estimation LTS is statistical technique for estimation of the unknown parameters of a linear regression model and provides a robust alternative to the classical regression methods based on minimizing the sum of squared residuals. The least trimmed squares estimator $\hat{\beta}_{LTS}$ is defined as: $\hat{\beta}_{LTS} = min_{\hat{\beta}} \sum_{i=1}^{h} e_i^2$, the trimming constant h has to satisfy $\frac{n}{2} < h < n$. Where $e_1^2 \leq e_2^2 \leq ... \leq e_n^2$ represents the i-th order statistics of square residuals $e_i^2 = \left(Y_i - X_i'\beta\right)^2, i = 1, 2, 3, ..., n$, the breakdown point value is $\frac{n-h}{n}$ for theLTS estimator. To find the estimator $\hat{\beta}_{LTS}$ it be taken into consideration $n - h + 1$ from the following subsample

$$\begin{aligned}
&\{X_1, X_2, ..., X_h\}\\
&\{X_2, X_3, ..., X_h + 1\}\\
&\{X_3, X_4, ..., X_h + 2\}\\
&...\\
&\{X_{n-h+1}, X_{n-h+2}, ..., X_n\}.
\end{aligned}$$

All sub sample comprises on h elements, it called the Contagious Half. Then, it computes the means for all subsample as following:

$$\begin{aligned}
\bar{X}_1 &= \frac{1}{h}\sum_{i=1}^{k} X_i\\
\bar{X}_2 &= \frac{1}{h}\sum_{i=2}^{k+1} X_i\\
&...\\
\bar{X}_{n-k+1} &= \frac{1}{h}\sum_{i=n-k+1}^{n} X_i.
\end{aligned}$$

And computing the sum of squares for all subsample

$$\begin{aligned}
SQ_1 &= \sum_{i=1}^{h} \left(X_i - \bar{X}_1\right)^2\\
SQ_2 &= \sum_{i=2}^{h+1} \left(X_i - \bar{X}_2\right)^2\\
&...\\
SQ_{n-h+1} &= \sum_{i=n-h+1}^{n} \left(X_i - \bar{X}_{n-h+1}\right)^2.
\end{aligned}$$

Then the least trimmed squares estimator $\hat{\beta}_{LTS}$ will be equal to the mean which corresponds the smallest sum square to equation.

4. The Application Part

4.1. **Data collection:** Data was taken from Charstin Company for manufacturing furniture and wooden doors which located in the city of Duhok. It was established in 2004, where the variables included for each line according to months of the years for the period 2004 To 2010, as follows:
1-Values of the monthly production as the dependent variable (Y).
2-Costs of electricity and water, as a first explanatory variable(X_1).
3-Cars expenses and maintenance of machinery as a second explanatory variable

(X_2).
4-Communication expenses as a third explanatory variable (X_3).
The unit of measurement of all the variables above is the U.S. dollar.

4.2. **Multiple linear regression model:** The following equation represents the regression model for each production line in the company:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + e_i \tag{4.1}$$

Where: Y_i: The value of i-th observation of the response variable (production). X_{i1}: The value of i-th observation of the first explanatory variable (costs of electricity and water). X_{i2}: The value of i-th observation of the second explanatory variable (cars maintenance and machines expenses). X_{i3}: The value of i-th observation of the third explanatory variable (communication expenses). By using OLS method with software Minitab version (13.2), the linear regression model fitted for each production line, as the following:
Furniture's line model:

$$\hat{Y}_i = -7283 + 36.3X_{i1} + 16.372X_{i2} + 17.532X_{i3} \tag{4.2}$$

Door's line model:

$$\hat{Y}_i = -191597 + 1034.9X_{i1} - 1178.6X_{i2} + 679.5X_{i3} \tag{4.3}$$

The results of the regression analysis represented in the tables 4.1,4.2,4.3 and 4.4 as follows:

Table 4.1: The regression coefficients for furniture's line by using OLS.

predictor	coefficients	Standard error of coefficients	t.value
Constant	-7283.000	1490.000	-4.888
X_1	36.300	5.468	6.639
X_2	16.372	5.176	3.163
X_3	17.532	2.587	6.777

S=5722, R-square=88.4

Table 4.2: Analysis of variance (ANOVA) table for furniture's line by using OLS.

source	DF	SS	MS	F
Regression	3	19882293231	6627431077.000	202.390
Residuals	80	2619662871	32745785.890	
Total	83	22501956102		

Table 4.3: The regression coefficients for door's line by using OLS.

predictor	coefficients	Standard error of coefficients	t.value
Constant	-191597.000	63870.000	-3.014
X_1	1034.900	204.000	5.073
X_2	-1178.600	269.300	-4.377
X_3	679.500	238.900	2.844

S=196086 R-square=69.8

Table 4.4: Analysis of variance (ANOVA) table for door's line by using OLS.

source	DF	SS	MS	F
Regression	3	7.125224E+12	2.37508E+12	61.770
Residuals	80	3.07599E+12	3.84499E+10	
Total	83	1.02012E+13		

From tables 4.1 and 4.3, It can be concluded that all parameters of the model for both lines are significant, through out noticing the absolute t-values > 2 or the p-values < 0.05. And from ANOVA tables 4.2 and 4.4, It is noticed that F-value is significant for both lines, since the value of calculated F greater than the tabulated value under significant level $\alpha = 0.05$ and degrees of freedoms df_1=3, df_2=80, which is F(0.05,3,80)= 2.72 , or the p-value is < 0.05. Thus the fitted models 4.2 and 4.3 are appropriate for the data of each line.

4.3. **Detection of outlier values:** Outlier values can be detected in the data by drawing the graphs of standardized residuals, resulting from using each method of estimation OLS, M, MM, LTS mentioned above. The graphs performed by representing the estimated values of the response variables (fitted values) on the horizontal axis (X-axis), and the values of standardized residuals on the vertical axis (Y-axis), the values that fall outside the cut-off points (Â±2.5) considered as outliers.

4.3.1. *Ordinary Least Squares OLS method:*

Table 4.5: Outlier values and its estimated values with standardized residual values for furniture's line by using OLS.

No.	Year	Month	Observed value Y	Estimated value$\hat{Y}_{OLS}$
14	2005	2	14231	28828.7
19	2005	7	10048	25421.7
67	2009	7	89282	73837.2
79	2010	7	51310	30385.5

Table 4.6: Outlier values and its estimated values with standardized residual values for door's line by using OLS.

No.	Year	Month	Observed value Y	Estimated value$\hat{Y}_{OLS}$
46	2007	10	741930	231489
47	2007	11	1110620	602866
53	2008	5	12633	573055

4.3.2. *M-estimation:* By applying the M-estimator robust method, by using SPSS (version 17) , S-plus (version 8.0, 2007 professional), and Minitab (version 13.2), the fitted linear model for each production line, is as the following:
Furniture's line model:

$$\hat{Y}_i = -6031.1450 + 34.8230X_1 + 11.8254X_2 + 18.3600X_3 \tag{4.4}$$

Door's line model:

$$\hat{Y}_i = 2887.2143 + 16.9561X_1 + 3.8971X_2 + 1.0568X_3 \tag{4.5}$$

The fitted models in 4.4 and 4.5, used to generate the residuals by M-estimator method, then the standardized residuals compute.

Table 4.7: Outlier values and its estimated values with standardized residual values for furniture's line by using M-estimator.

No.	Year	Month	Observed value Y	Estimated value$\hat{Y}_M$
19	2005	7	10048	23015.6
54	2008	6	36715	20547.9
67	2009	7	89282	68579.5
79	2010	7	51310	27804.3

Table 4.8: values and its estimated values with standardized residual values for door's line by using M-estimator.

No.	Year	Month	Observed value Y	Estimated value$\hat{Y}_M$
40	2007	4	1352330	23119.4
41	2007	5	1040410	19278.6
42	2007	6	1160110	20921.5
43	2007	7	1475250	26699.3
44	2007	8	1129590	20899.2
45	2007	9	1229060	21663.0
47	2007	11	1110620	18385.2

4.3.3. *MM-estimation:* By applying the MM-estimator robust method, by using SPSS (version 17) ,S-plus (version 8.0, 2007 professional), and Minitab (version 13.2), the fitted linear model for each production line is as the following:

Furniture's line model:

$$\hat{Y}_i = -6312.1664 + 36.7890X_1 + 12.0964X_2 + 19.4700X_3 \tag{4.6}$$

Door's line model:

$$\hat{Y}_i = 2894.1356 + 17.4564X_1 + 4.0552X_2 + 1.1582X_3 \tag{4.7}$$

The fitted models in 4.6 and 4.7, used to generate the residuals by MM-estimator method, then the standardized residuals computed.

Table 4.9: Outlier values and its estimated values with standardized residual values for furniture's line by using MM-estimator.

No.	Year	Month	Observed value Y	Estimated value$\hat{Y}_{MM}$
19	2005	7	10048	24114.33
67	2009	7	89282	72091.02
79	2010	7	51310	29173.59

Table 4.10: Outlier values and its estimated values with standardized residual values for door's line by using MM-estimator.

No.	Year	Month	Observed value Y	Estimated value $\hat{Y}_{MM}$
40	2007	4	1352330	23950.94
41	2007	5	1040410	19351.98
42	2007	6	1160110	21197.29
43	2007	7	1475250	27606.55
44	2007	8	1129590	23018.05
45	2007	9	1229060	23058.02
47	2007	11	1110620	19815.32

4.3.4. *LTS estimation:* By applying the LTS-estimator robust method, by using $SPSS$ (version 17), S-plus (version 8.0, 2007 professional), and Minitab (version 13.2), the fitted linear model for each production line is as the following:

Furniture's line model:

$$\hat{Y}_i = -6809.2556 + 36.9543X_1 + 12.4873X_2 + 19.4401X_3 \tag{4.8}$$

Door's line model:

$$\hat{Y}_i = 1883.232 + 18.580X_1 + 12.362X_2 - 2.230X_3 \tag{4.9}$$

The fitted models in 4.8 and 4.9, used to generate the residuals by LTS-estimator method, then the standardized residuals computed.

Table 4.11: Outlier values and its estimated values with standardized residual values for furniture's line by using LTS-estimator.

No.	Year	Month	Observed value Y	Estimated value $\hat{Y}_{LTS}$
19	2005	7	10048	23960.17
54	2008	6	36715	21341.25
67	2009	7	89282	72242.47
79	2010	7	51310	29040.33

Table 4.12: Outlier values and its estimated values with standardized residual values for door's line by using LTS-estimator.

No.	Year	Month	Observed value Y	Estimated value $\hat{Y}_{LTS}$
40	2007	4	1352330	23950.8
41	2007	5	1040410	19351.8
42	2007	6	1160110	21197.1
43	2007	7	1475250	27606.3
44	2007	8	1129590	23017.9
45	2007	9	1229060	23057.8
47	2007	11	1110620	19815.2

The number of detected outliers by using four methods for both production lines can be summarized in the following table:

Table 4.13: Detection of outliers by different methods for both production lines

Production line method	OLS	M	MM	LTS
Furniture	4	4	3	4
Doors	3	7	7	7

4.4. **Comparison of the results of OLS method and robust methods M-estimator, MM-estimator and LTS-estimator:** After estimation the coefficients of multiple linear equation with four methods, the results can be compared by using the criteria mentioned in (chapter two) of theoretical part of the paper, in order to judge the estimators, the following tables show the estimated values of coefficients by using these methods, and values of comparative criteria for each production line.

Table 4.14: The estimated values of coefficients by using four methods for furniture's line

Variables	Coefficients	OLS	M	MM	LTS
Constant	β_0	-7283.000	-6031.145	-6312.166	-6809.257
X_1	β_1	36.300	34.823	36.789	36.954
X_2	β_2	16.372	11.825	12.096	12.487
X_3	β_3	17.532	18.360	19.470	19.440
MSE		33160290	37280005	34141334	34260526
$MAPE$		1.0495	0.9826	1.0312	1.0034
$MSAE$		3961.3100	3975.4900	3896.2500	3894.5200

Table 4.15: The estimated values of coefficients by using four methods for door's line.

Variables	Coefficients	OLS	M	MM	LTS
Constant	β_0	-191597.000	2887.214	2894.136	1883.232
X_1	β_1	1034.900	16.956	17.546	18.580
X_2	β_2	-1178.600	3.897	4.055	12.362
X_3	β_3	679.500	1.057	1.158	-2.230
MSE		1.4210E+11	1.41584E+11	1.41420E+11	1.41389E+11
$MAPE$		23.3736	0.2127	0.2134	1.1806
$MSAE$		284510	118247	118150	120928

5. Conclusions and Recommendations

5.1. **Conclusions:** According to the results of the practical application of the study the following conclusions are yielded:

1- The number of outlier values detected by using the four methods in the data for furniture's line is very close, by observing table 4.13, and the proportion of outlier values is 4.77. This refers to the fact that the distribution of standard errors is close to the normal distribution as shown in graph 4.1.

2- From the table 4.13, it is noticed that the number of outlier values found in the data for door's line, by using OLS is 3 values, and the number of outlier values detected using the robust methods equals to 7 values and its proportion is 8.33, this means that the distribution of standard errors is somewhat distant from the normal distribution as shown in graph 4.3.

3- The appearance of the negative sign for $\hat{\beta}_0$ in the table 4.14 is logical, this means that the absence of the explanatory variables (stopping production) for furniture's line. This mean there would be a loss.

4- It is noticed from table 4.14, that the estimated values for each parameter using the four methods are very close, and the reason is that the number of outlier values in the data is a few, as noted in the first conclusion above.

5- The appearance of the positive sign for $\hat{\beta_0}$ in the table 4.15, by using the robust methods is illogical, the researcher believes that the reason is due to the fact that the data for doors line are inaccurate values.

6- It is noticed from the table 4.15, that the estimated values using OLS are very far from its estimated values by using the robust methods, and that is a logical thing, because of the large number of outlier values discovered in the data for door's line, and it is also noticed that theLTS method gave better results by using the criterion MSE, and the$M$ method gave better results by using the criterion$MAPE$, the criterion value is very close to its value by using MM method, it is also noticed that by using the criterion $MSAE$, the MM method is better.

5.2. **Recommendations:** According to the results obtained in the practical part, the researcher recommends are as follow:

1- Studying all robust methods commonly used, and applying them to a large number of data, in order to clarify the efficiency of each method.

2- Using simulation procedure to generate data, and then contaminating them with the outlier values, at different percentage, and applying robust methods commonly used to them, and then comparing the results of these methods.

References

[1] Chatterjee, S. and Hadi, A.S.(1988), Sensitivity Analysis by Example. Johan wiley and sons, New York U.S.A.

[2] Cook, R.D. and Weisberg, S.(1982), Residuals and Influence in Regression, chapman and Hall, London.

[3] Dianne P.O. (1990), Robust regression computation using iteratively reweighted least squares, SIAM J. Matrix Anal. Appl., Vol.11, No.3, pp. 466-480.

[4] Gnanadeskan, R. and Kettenring, J.R.(1972), Robust estimates, residuals and outlier detection with multresponse data, biometrics,Vol.28, No.1,pp.81-124.

[5] Hanzi W. and David S. (2002), A Highly Efficient Symmetry-Based Robust Estimator, seventh International conference on control, Automation,Robotics and vision (ICARCV 02), Singapore.

[6] Harvey, A.C.(1978), On the unbiasedness of robust regression estimators, commun statist theor math Vol.8, No.7, pp.779-783.

[7] Hawkins, D.M. and Kass (1984), Location of several outliers in multiple regression data using elemental sets. Tech., Vol.26, No.3, pp.197-208.

[8] Huber P.J. (1990), Robust Statistical Procedures, 2nd edition, Society for industrial and Appl. Math. USA.,pp.13,56-58.

[9] Jureckova J. and Jan P. (2006), Robust Statistical Methods with R, Chapman and Hall/CRC, London.

[10] Makridakis S.C., Wheel W. and Rob J.H. (1983), Forecasting, Methods and Applications, John wiley and Sons. Inc.,New York,USA.

[11] Maronna, R.A; Yohai and Victor J.(1997), Robust estimation in simultaneous equations models, Jurnal of statistical planning and Inference Vol.57, pp.233-244.

[12] Moberg, T.F.; Ramberg, J.S. and Randles, R.H. (1980), An adaptive multiple regression procedure based on M-estimators, Tech., Vol.22, No.2, PP.213-224.

[13] Nelide, F.; Diana, K.; Yohai, V.J. and Ruben, Y.H.Z. (1999), A class of locally and globally robust regression estimates, Journal of the American statistical Association (JASA), Vol.94, No.445, pp.174-188.

[14] Neter, J., Wasserman, W. and Kutner, M.(1985); Applied Linear Statistical Models, 2nd edition, IRWIN.

[15] Ozlem G. A. (2011) Comparison of robust regression methods in linear regression, Int.J. contemp, Math. Sciences, Vol.6, no.9, pp.409- 421.

[16] Richord, W.H. and Paul, W.H.(1977), Two robust alternatives to least- Sequa's regression, Journal of the American Statistical Association (JASA),Vol.72, No. 360, pp. 828-834.

[17] Rousseeuw, J.P.(1984), Least median of squares regression, Journal of the American Statistical Association (JASA), Vol.79, No.388, pp.871-880.

[18] Rousseeuw, P.J. Christophe. C.and Ola, H.(1994), Generalized S- estimators, Journal of the American Statistical Association (JASA), pp.1271-1281.

[19] Rousseeuw, P.J. and Van Alest, S.(1999), Positive breakdown robust methods in computer vision, Computing science and statistics,Vol.31, pp. 451-460

[20] Rousseeuw, P.J. and Van Driessen, K.(1999), A fast algorithm for the minimum covariance determinate estimator, Techn., No.41, pp. 212-223.

[21] Rousseeuw, P.J.(1987), Robust regression and outlier detection, John wiley and sons,New York.

[22] Sanford W. (2005), Applied Linear Regression, 3rd edition, John Wiley and Sons, Inc. Hoboken,New Jersey, Canada

[23] Wilcox Rand R. (2005), Introduction to Robust Estimation and Hypothesis Testing, 2nd edition, Elsevier Inc.,USA.

(S.A. Mohammed) University of Duhok, Duhok, Iraq
E-mail address: sizar@uod.ac

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 246-252, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

HYBRID METHODS FOR SOLVING VOLTERRA INTEGRAL EQUATIONS

GALINA MEHDIYEVA, MEHRIBAN IMANOVA, AND VAGIF IBRAHIMOV

Abstract. It is known that there exists a class of methods for solving integral equations with variable boundary. One of them is the most popular methods of quadratures. This method is clarified and modified by many scientists.

Here, to numerically solving Volterra integral equations the hybrid method is applied and constructed a concrete method with the degree $p = 6$ and $p = 8$, by using information about the solution of integral equations only one previous point.

1. Introduction

Consider the numerical solution of the following nonlinear Volterra integral equation:

$$y(x) = f(x) + \int_{x_0}^{x} K(x, s, y(s))ds, \quad x \in [x_0, X]. \tag{1}$$

Assume that the problem (1) has a unique continuous solution $y(x)$ determined on the interval $[x_0, X]$. By means of a constant step $0 < h$ the interval $[x_0, X]$ divide into N equal parts with the mesh points $x_i = x_0 + ih$ $(i = 0, 1, \ldots, N)$. Denote by y_i the approximate and by $y(x_i)$ exact value of the solution of problem (1) at the points x_i $(i = 0, 1, 2, ..., N)$. There are many papers dedicated to investigation numerical methods for solving integral equations (1) (see f.e. [1]-[6]).

The first hybrid method for solving integral equation (1) constructed Makroglou [7], but that in [8] is generalized in the following form:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i f_{n+i} + h \sum_{j=0}^{k} \sum_{i=0}^{j-\nu_i} \beta_i^{(j)} K(x_{n+j}, x_{n+l_i}, y_{n+l_i})$$
$$(l_i = i + \nu_i, \ \ |\nu_i| < 1). \tag{2}$$

Method suggested here has the following form:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i f_{n+i} + h \sum_{j=0}^{k} \sum_{i=0}^{j} \beta_i^{(j)} K(x_{n+j}, x_{n+i}, y_{n+i}) +$$
$$+ h \sum_{j=0}^{k} \sum_{i=0}^{j-\nu_i} \gamma_i^{(j)} K(x_{n+j}, x_{n+l_i}, y_{n+l_i}). \tag{3}$$

Key words and phrases. Volterra integral equations, hybrid methods, degree of hybrid methods.

This research has been supported by the Science Development Foundation of Azerbaijan (EIF-2011-1(3)-82/27/1).

2. Constructing a hybrid method of the type (3)

It is known, that hybrid methods such as (3) can be constructed by different ways. Here the calculation of the next function

$$\varphi(x) = \int_{x_0}^{x} K(x, s, y(s))ds$$

on the interval $[x_0, X]$ by the some structures are reduced to the computation of functions$\varphi(x)$ on the interval $[x - kh,\, x]$ (here k – limited quantity). To this end, consider the following difference:

$$y(x_{n+1}) - y(x_n) = g(x_{n+1}) - g(x_n) + \int_{x_0}^{x_n} (K(x_{n+1},\, s,\, y(s)) - K(x_n,\, s,\, y(s)))ds +$$

$$+ \int_{x_n}^{x_{n+1}} K(x_{n+1},\, s,\, y(s))ds. \tag{4}$$

By the application of Lagrange's theorem we can write:

$$K(x_{n+1},\, s,\, y(s)) - K(x_n,\, s,\, y(s)) = hK'_x(\xi,\, s,\, y(s)) \quad (x_n < \xi_n < x_n + h).$$

Then, by using the received equality in (4) we have:

$$y(x_{n+1}) - y(x_n) = g(x_{n+1}) - g(x_n) + h\int_{x_0}^{x_n} K'_x(\xi_n,\, s,\, y(s))ds +$$

$$+ \int_{x_n}^{x_{n+1}} K(x_{n+1},\, s,\, y(s))ds. \tag{5}$$

We assume that by any method we have found the solution of equation (1). Then, taking into account its in (1) we can write the following:

$$y'(x) = g'(x) + K(x,\, x,\, y(x)) + \int_{x_0}^{x} K'_x(x,\, s,\, y(s))ds.$$

It is easy to show that

$$h\int_{x_0}^{x_n} K'_x(\xi_n,\, s,\, y(s))ds = h(y'(\xi_n) - g'(\xi_n) - K(\xi_n,\, \xi_n,\, y(\xi_n))) -$$

$$-h\int_{x_n}^{\xi_n} K'_x(\xi_n,\, s,\, y(s))ds. \tag{6}$$

To take into account the relation (6) and the following

$$hK'_x(\xi_n,\, s,\, y(s)) = K(x_{n+1},\, s,\, y(s)) - K(x_n,\, s,\, y(s))$$

in equation (1), we have:

$$y(x_{n+1}) = y(x_n) + g(x_{n+1}) - g(x_n) + h(y'(\xi_n) - g'(\xi_n)) - hK(\xi_n,\, \xi_n,\, y(\xi_n)) +$$

$$+\int\limits_{x_n}^{\xi_n} K(x_n,\, s,\, y(s))ds + \int\limits_{\xi_n}^{x_{n+0}} K(x_{n+1},\, s,\, y(s))ds. \tag{7}$$

Thus, determined the solution of equation (1) on the segment$[x_0,\, x_{n+1}]$ to reduce it's determined on the segment $[x_n,\, x_n+h]$. Using one of the known formulas for calculating integrals involved in equation (7) and using the following equality (see f.e. [9, p.325]):

$$z(x_{n+1}) - z(x_n) - hz'(\xi_n) = \sum_{i=0}^{k} \alpha_i z_{n+1}$$

in equality (7) we have:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i g_{n+i} - hK(x_{n+l},\, x_{n+l},\, y_{n+l}) +$$

$$+h\sum_{j=0}^{k}\sum_{i=0}^{k} \beta_i^{(j)} K(x_{n+j},\, x_{n+i},\, y_{n+i}). \tag{8}$$

Here $\alpha_i,\, \beta_i^{(j)}$ $(j, i = 0,\, 1,\, 2, ...,\, k)$ are some real numbers, the parameter l satisfies the condition $0 < l < 1$, and

$$g_m = g(x_m) \quad (m = 0,\, 1,\, 2,\, ...).$$

However, by the help of next relation (see. [9, p. 375])

$$K(x_{n+l},\, x_{n+l},\, y_{n+l}) = \sum_{i=0}^{k} \beta_i K(x_{n+i},\, x_{n+i},\, y_{n+i}),$$

method (8) can be written in the form:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i g_{n+i} + h\sum_{j=0}^{k}\sum_{i=0}^{k} \beta_i^{(j)} K(x_{n+j},\, x_{n+i},\, y_{n+i}). \tag{9}$$

To calculating the integrals involved in (7), we can apply different formulas. If applied to the calculation of integrals the quadrature methods with the fractional steps, then the formula (9) can be rewritten as:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i g_{n+i} + h\sum_{j=0}^{k}\sum_{i=0}^{k} \beta_i^{(j)} K(x_{n+j},\, x_{n+i+\nu_i},\, y_{n+i+\nu_i}), \tag{10}$$

here the coefficients $\alpha_i,\, \beta_i^{(j)}$ $(i,\, j\, =\, 0,\, 1,\, 2, ...,\, k)$ – are some real numbers, and $|\nu_i| < 1 \quad (i = 0,\, 1,\, ...,k)$.

Note that in (8) is involved as a member $K(x_{n+l},\, z_{n+l},\, y_{n+l})$ $(|l| < 1)$, which is different from the terms of the sum in (10), in that here in the first argument used by the so-called fractional step. However, using the Lagrange interpolation formula can get rid of fractional steps. But if we to take into account the specified property of one term of the method (8) and generalize it, then the method (10) can be rewritten as:

$$\sum_{i=0}^{k} \alpha_i y_{n+i} = \sum_{i=0}^{k} \alpha_i g_{n+i} + h\sum_{j=0}^{k}\sum_{i=0}^{k} \beta_i^{(j)} K(x_{n+j+l_j},\, x_{n+i+\nu_i},\, y_{n+i+\nu_i}), \tag{11}$$

where $|l_j| < 1$ $(j = 0, 1, ..., k)$. If set $l_j = 0$ $(j = 0, 1, ..., k)$, then from the relation (11) be it follows the method (10). It is easy to determine if generalize the methods (10) and (11), in one embodiment receive the following method:

$$\sum_{i=0}^{k}\alpha_i y_{n+i} = \sum_{i=0}^{k}\alpha_i g_{n+i} + h\sum_{j=0}^{k}\sum_{i=0}^{k}\beta_i^{(j)}K(x_{n+j},\, x_{n+i},\, y_{n+i}) +$$

$$+h\sum_{j=0}^{k}\sum_{i=0}^{k}\gamma_i^{(j)}K(x_{n+j+l_j},\, x_{n+i+\nu_i},\, y_{n+i+\nu_i})\,. \tag{12}$$

The coefficients of these methods determined by the same scheme for $l_j = \nu_j$ $(j = 0, 1, ..., k)$. Before we propose a method for determining the coefficients α_i, $\beta_i^{(j)}$, $\gamma_i^{(j)}$ $(i,\, j = 0, 1, 2, ..., k)$, we consider some bounded on the coefficients of the method (12) and set $l_i = \nu_i$ $(i = 0, 1, ..., k)$. This is connected to the fact that, the method for determining the coefficients (11), we use some facts from the theory of difference equations and apply the following finite-difference method with constant coefficients to the solving of equation (1):

$$\sum_{i=0}^{k}\alpha_i z_{n+i} = h\sum_{i=0}^{k}\beta_i z'_{n+i} + h\sum_{i=0}^{k}\gamma_i z'_{n+i+\nu_i}. \tag{13}$$

We show that after applying the method (13) to the solving equation (1) we receive be the method of the type (12) and in this case between the coefficients of the methods (12) and (13) will be the next relations:

$$\beta_i = \sum_{j=0}^{k}\beta_i^{(j)};\;\; \gamma_i = \sum_{j=0}^{k}\gamma_i^{(j)}\;\; (i = 0, 1, 2, ..., k). \tag{14}$$

Now, suppose that the method (13) is converges. Then its coefficients satisfy the following conditions:

A: The values of the quantities $\alpha_i, \beta_i; \gamma_i (i = 0, 1, 2, ..., k)$ are real numbers, moreover, $\alpha_k \neq 0$.

B: Characteristic polynomials

$$\rho(\lambda) \equiv \sum_{i=0}^{k}\alpha_i\lambda^i, \quad \sigma(\lambda) \equiv \sum_{i=0}^{k}\beta_i\lambda^i, \quad \gamma(\lambda) \equiv \sum_{i=0}^{k}\gamma_i\lambda^{i+l_i};$$

have no common multipliers different from the constant.

C: $\sigma(1) + \gamma(1) \neq 0$ and$p \geqslant 1$ are holds.

To determine the values of the coefficients α_i, β_i, γ_i $(i = 0, 1, ..., k)$ using the method of undetermined coefficients, and to this end, consider the following Taylor expansion of functions:

$$y(x + ih) = y(x) + ihy'(x) + \frac{(ih)^2}{2!}y''(x) + ... + \frac{(ih)^p}{p!}y^{(p)}(x) + O(h^{p+1}), \tag{15}$$

$$y'(x + l_i h) = y'(x) + l_i hy''(x) + \frac{(l_i h)^2}{2!}y'''(x) + ... + \frac{(l_i h)^{p-1}}{(p-1)!}y^{(p)}(x) + O(h^p), \tag{16}$$

where $x = x_0 + nh -$ is fixed point, but $l_i = i + \nu_i \quad (i = 0, 1, 2, ..., k)$.

For determining the values of the parameters $\alpha_i,\ \beta_i,\ l_i\ (i=0,1,2,...,k,\ l_i = i+\nu_i)$, take into account equalities (15) and (16) in the next asymptotic equality

$$\sum_{i=0}^{k}(\alpha_i y(x+ih) - h\beta_i y'(x+ih) - h\gamma_i y'(x+l_i h) = O(h^{p+1}),\quad h\to 0,$$

here $x = x_0 + nh$ the fixed point and p is the degree of the method (13).

Then have:

$$\sum_{i=0}^{k}\alpha_i = 0;\quad \sum_{i=0}^{k}\frac{i^\nu}{\nu!}\alpha_i = \sum_{i=0}^{k}\frac{i^{\nu-1}}{(\nu-1)!}\beta_i + \sum_{i=0}^{k}\frac{l_i^{\nu-1}}{(\nu-1)!}\gamma_i\quad (\nu = 1,2,...,p,\quad 0! = 1). \tag{17}$$

Thus, for determining the values of the $\alpha_i,\ \beta_i,\ l_i\ (i=0,1,2,...,k)$ we get a homogeneous system of nonlinear algebraic equation where the amount of unknowns equals $3k+3$, but the amount of equation$p+1$. Obviously system (17) always has a trivial solution. However the trivial solution of system (17) is not of interest. Therefore we consider the case when system (17) to have a non-zero solution. It is known, that the system (17) has the non zero solution if the condition

$$p < 4k+3 \tag{18}$$

is holds. Hence it follows that $p_{\max} = 4k+2$, but for $\beta_i = 0 (i = 0,1,...,k)$ is holds, that $p_{\max} = 3k+1$.

Now to consider construct a concrete method for $k=2$ for $\beta_i = 0\ \ (i=0,1,2)$. Then from system (17) we have:

$$\begin{aligned}
&\gamma_2+\gamma_1+\gamma_0 = 2\alpha_2+\alpha_1,\\
&\nu_2\gamma_2+\nu_1\gamma_1+\nu_0\gamma_0 = \frac{1}{2}(2^2\alpha_2+\alpha_1),\\
&\nu_2^2\gamma_2+\nu_1^2\gamma_1+\nu_0^2\gamma_0 = \frac{1}{3}(2^3\alpha_2+\alpha_1),\\
&\nu_2^3\gamma_2+\nu_1^3\gamma_1+\nu_0^3\gamma_0 = \frac{1}{4}(2^4\alpha_2+\alpha_1),\\
&\nu_2^4\gamma_2+\nu_1^4\gamma_1+\nu_0^4\gamma_0 = \frac{1}{5}(2^5\alpha_2+\alpha_1),\\
&\nu_2^5\gamma_2+\nu_1^5\gamma_1+\nu_0^5\gamma_0 = \frac{1}{6}(2^6\alpha_2+\alpha_1).
\end{aligned} \tag{19}$$

Selecting α_2,α_1 and α_0 we can construct different methods with the degree $p=6$. For $\alpha_2 = 1,\ \alpha_1 = \alpha_0 = -1/2$ solving system (19) we get a complex solution, but for $\alpha_2 = 1,\ \alpha_1 = -1$ and $a_0 = 0$ the solution of system (19) has the following form:

$$\beta_2 = 5/18,\quad \beta_1 = 8/18,\quad \beta_0 = 5/18,$$

$$l_2 = 3/2+\sqrt{15}/10,\quad l_1 = 3/2,\quad l_0 = 3/2-\sqrt{15}/10.$$

Taking into account these values in formula (13) we get the following method:

$$y_{n+2} = y_{n+1} + h(5y'_{n+3/2+\sqrt{15}/10} + 8y'_{n+3/2} + 5y'_{n+3/2-\sqrt{15}/10})/18. \tag{20}$$

Remark, that these method can be rewrite as one step method in the next form:

$$y_{n+1} = y_n + h(5y'_{n+1/2+\sqrt{15}/10} + 8y'_{n+1/2} + 5y'_{n+1/2-\sqrt{15}/10})/18.$$

This method is stable and has the degree $p=6$.

3. Algorithm for using hybrid methods

Construct an algorithm for using method (20). Assume that y_1 is known, and consider the calculation of y_{n+2} $(n = 0, 1, 2, ...)$. For applying method (20) the values of the quantities $y_{n\pm\alpha}$, should be known and they are determined by means of the following formula:

$$y_{n+\alpha} = y_n + 2hy'_n + \alpha^2 h((h^2 - 12\alpha + 6)y'_{n+3/2} - (3\alpha^2 - 48\alpha + 27)y'_{n+1} +$$

$$+(3\alpha^2 - 60\alpha + 54)y'_{n+1/2} - (\alpha^2 - 24\alpha + 33)y'_n)/18, \qquad (21)$$

where $\alpha = 3/2 \pm \sqrt{15}/10$.

Thus, we see that for using method (21), we must construct formulae for calculating the values of $y_{n+\frac{1}{2}}$, y_{n+1}, $y_{n+\frac{3}{2}}$ with accuracy $O(h^5)$. If $y_{\frac{1}{2}}$are know, then for $n > 1$ it is enough to calculate the values of y_{n+1} and $y_{n+\frac{3}{2}}$. Therefore, at first consider the definition of the value of the quantity y_{n+1} and to this end we use the following sequential formulae:

1. $\breve{y}_{n+1} = y_n + hy'_n$,
2. $\bar{y}_{n+1} = y_n + h(y'_n + \breve{y}'_{n+1})/2$,
3. $\hat{y}_{n+1} = y_n + h(y'_n + 4y'_{n+\frac{1}{2}} + \bar{y}'_{n+1})/6$,
4. $y_{n+1} = y_n + h(y'_n + 4y'_{n+\frac{1}{2}} + \hat{y}'_{n+1})/6$,
5. $\hat{y}_{n+\frac{3}{2}} = y_{n+\frac{1}{2}} + h(7y'_{n+1} - 2y'_{n+\frac{1}{2}} + y'_n)/6$,
6. $y_{n+\frac{3}{2}} = y_{n+1} + h(9\hat{y}'_{n+\frac{3}{2}} + 19y'_{n+1} - 5y'_{n+\frac{1}{2}} + y'_n)/48$.

Conclusion 1. *Given that stable hybrid methods are more accurate and have advanced the stability region, here for numerical solution of equation (1) suggested a method that makes use of hybrid methods. Note that the methods of type (12) belong to the class of methods of type (13). To constructed type methods of type (12) we must to solve system (14), whose solution is not unique. However, the solution of (17) may be unique. This means that when the method of type (13) is unique corresponding method of type (12) is not unique. Therefore, both theoretical and practical interest is the finding maximum value of degree of stable method of type (12). Also of interest to determine the stability region of method (12). Consequently, the investigation of method (12) is a priority in the theory of numerical methods. Remark, that the method (20) has the new characteristic. The first of them including in used method (20) as one step method, but in these case for the determine quantities* $y_{n+1/2\pm\sqrt{15}/10}$ *by the method (21) the values of quantities* $y_{n+3/2}$, y_{n+1} *must be known. Because here suggested the method as the (20), but the next characteristic of the method (20) including in definition the method (20) as the explicit, but the method (21) for any values of parameter* α *and the method applying for finding* y_{n+1} *and* $y_{n+3/2}$ *are implicit. Consequently algorithm for using method (20) is implicit.*

References

[1] V. Volterra, *Theory of functional and of integral and integro-differensial equations*, Dover publications. Ing, New York, 304.

[2] A.F. Verlan, V.S. Sizikov, *Integral equations: methods, algorithms, programs*, Kiev, Naukova Dumka, 1986.

[3] Ch. Lubich, Runge-Kutta theory for Volterra and Abel integral equations of the second kind, *Mathematics of computation*, volume 41, number 163, July 1983, pp. 87–102.

[4] H. Brunner, Imlicit Runge-Kutta Methods of Optimal oreder for Volterra integro-differential equation, *Methematics of computation*, Volume 42, Number 165, January 1984, pp. 95–109.
[5] G. Mehdiyeva, M. Imanova, On an application of the finite-difference method, *News of Baku University, Physico-Mathematical sciences*, 2008, [1] 2, pp. 73–78.
[6] K. Maleknejad, M. Tavassoli Kajani, Solving second kind integral equations by Galerkin methods with hybrid Legendre and block-pulse functions, *Applied Mathematics and Computation*, 145, 623–629 (2003).
[7] Makroglou, Hybrid methods in the numerical solution of Volterra integro-differential equations, *Journal of Numerical Analysis*, 2, 1982, pp. 21–35.
[8] G. Mehdiyeva, V. Ibrahimov, M. Imanova, On One Application Of Hybrid Methods For Solving Volterra Integral Equations, *World Academy of Science, Engineering and Technology*, Dubai, 2012, pp. 809–813.
[9] N.S. Bakhvalov, *Numerical methods*, M. Nauka, 1973.

(G.Yu. Mehdiyeva) Baku State University, Baku, Azerbaijan
E-mail address: imn_bsu@mail.ru

(M.N. Imanova) Baku State University, Baku, Azerbaijan
E-mail address: imn_bsu@mail.ru

(V.R. Ibrahimov) Baku State University, Baku, Azerbaijan
E-mail address: ibvag47@mail.ru

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO. 2, 253-259, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON THE NUMBER OF REPRESENTATIONS OF AN INTEGER OF THE FORM $x^2 + dy^2$ IN A NUMBER FIELD

SARAT SINLAPAVONGSA AND AJCHARA HARNCHOOWONG

Abstract. This paper tells us about the number of representations of an algebraic integer of the form $x^2 + dy^2$ in a number field where d is a positive rational integer.

1. Introduction

There are many papers which give the criteria to determined whether an algebraic integer can be represented in the form x^2+dy^2 such as [1], [2], [3], [4]. Another interesting problem about integers of the form $x^2 + dy^2$ is to study the number of these representations. T. Nagell [5] study the problem of the number of representations of an integer which can be represented as the sums of two squares. For this paper we generalize the result of T. Nagell to the representations of an algebraic integer in a number field of the form $x^2 + dy^2$ where d is a positive rational integer.

2. Preliminaries

Let ω be an integer in a number field K and d a positive rational integer. We say that ω *has a representation of the form* $x^2 + dy^2$ if there are integers α and β in K such that $\omega = \alpha^2 + d\beta^2$. The representations $\omega = x^2 + y^2$ with $x = \pm\alpha, y = \pm\beta$ and $x = \pm\beta, y = \pm\alpha$ and the representations $\omega = x^2 + dy^2$ for $d > 1$ with $x = \pm\alpha$ and $y = \pm\beta$ are considered to be one and the same. The relation $1 = 1^2 + d \cdot 0^2$ is called the *trivial representation* of the number 1.

3. Main Results

These are the main results of the paper.

Theorem 3.1. *If* 1 *has more representations of the form* $x^2 + dy^2$ *than the trivial representation, then* 1 *has infinitely many representations.*

Proof. Assume that

$$1 = \gamma^2 + d\delta^2$$

where γ and δ are integers in K such that $\gamma \neq 1$ and $\delta \neq 0$.
For a positive integer n, we define

$$\gamma_n + \delta_n\sqrt{-d} = (\gamma + \delta\sqrt{-d})^n,$$

where

$$\gamma_n = \gamma^n - \binom{n}{2}\gamma^{n-2}\delta^2 d + \binom{n}{4}\gamma^{n-4}\delta^4 d^2 - + \dots \tag{3.1}$$

Key words and phrases. algebraic integer, totally real, number field.
2010 *AMS Math. Subject Classification.* Primary 11R33, 11R80; Secondary 11R27.

1

and

$$\delta_n = \binom{n}{1}\gamma^{n-1}\delta - \binom{n}{3}\gamma^{n-3}\delta^3 d + - \ldots \tag{3.2}$$

Then

$$\gamma_n - \delta_n\sqrt{-d} = (\gamma - \delta\sqrt{-d})^n$$

and

$$(\gamma_n + \delta_n\sqrt{-d})(\gamma_n - \delta_n\sqrt{-d}) = (\gamma + \delta\sqrt{-d})^n(\gamma - \delta\sqrt{-d})^n = (\gamma^2 + d\delta^2)^n.$$

Therefore

$$\gamma_n^2 + d\delta_n^2 = 1.$$

Thus the Diophantine equation

$$x^2 + dy^2 = 1 \tag{3.3}$$

has the integral solutions

$$x = \gamma_n, y = \delta_n.$$

Next, we will prove that these solutions are all different.
Suppose for a contradiction that there are $m, n \in \mathbb{N}$ such that $m \neq n$ and

$$\gamma_m = \gamma_n, \delta_m = \delta_n.$$

Then

$$(\gamma + \delta\sqrt{-d})^m = (\gamma + \delta\sqrt{-d})^n,$$

and so $\gamma + \delta\sqrt{-d}$ is a root of unity. Suppose that

$$\gamma + \delta\sqrt{-d} = \zeta$$

is a primitive Nth root of unity. Since

$$\gamma - \delta\sqrt{-d} = \zeta^{-1},$$

we get

$$\gamma = \frac{1}{2}(\zeta + \zeta^{-1}), \delta = \frac{1}{2\sqrt{-d}}(\zeta - \zeta^{-1}).$$

Thus

$$\frac{1}{2}(\zeta^2 - 1) = \sqrt{-d}\zeta\delta$$

is an algebraic integer.
If N is a power of 2 and $N \geq 8$, then the number

$$\frac{1}{2}(\zeta^{N/4} - 1) = \frac{1}{2}(\pm i - 1)$$

must also be an algebraic integer. This is a contradiction.
If N is divisible by the odd prime p, then the number

$$\frac{1}{2}(\zeta^{2N/p} - 1)$$

must also be an algebraic integer but $\frac{1}{2}(\zeta^{2N/p} - 1)$ is the root of the irreducible polynomial

$$\frac{1}{2x}[(2x+1)^p - 1] = 2^{p-1}x^{p-1} + \ldots + p(p-1)x + p$$

which has integral coefficients. This is a contradiction.
Finally, if $N = 1, 2, 4$, then $\gamma = 0$ or δ is not an algebraic integer. This is a contradiction. $\square$

Theorem 3.2. *Let K be a number field.*

(i) *Let π be a prime in K such that π has the representation of the form x^2+dy^2. If 1 has only trivial representation, then π has exactly one representation. Otherwise π has infinitely many representations.*

(ii) *Let ω be an integer in K such that ω has the representation of the form x^2+dy^2. If 1 has only trivial representation, then ω has a finite number of the representations. Otherwise ω has infinitely many representations.*

Proof. (i) Assume that 1 has only trivial representation Let π be a prime in K such that π has two representations of the form x^2+dy^2,

$$\pi = \alpha_1^2 + d\beta_1^2$$

and

$$\pi = \alpha_2^2 + d\beta_2^2$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2$ are integers in K. Then

$$\pi(\beta_2^2 - \beta_1^2) = \alpha_1^2\beta_2^2 - \alpha_2^2\beta_1^2.$$

Since π is a prime, either of the $\alpha_1\beta_2 - \alpha_2\beta_1$ and $\alpha_1\beta_2 + \alpha_2\beta_1$ must be divisible by π. Without loss of generality, we may assume that

$$\alpha_1\beta_2 \equiv \alpha_2\beta_1 \bmod \pi.$$

Multiplying together the two representations of π, we get

$$\pi^2 = (\alpha_1\alpha_2 + d\beta_1\beta_2)^2 + d(\alpha_1\beta_2 - \alpha_2\beta_1)^2.$$

Since $\alpha_1\beta_2 - \alpha_2\beta_1$ is divisible by π, so is the number $\alpha_1\alpha_2 + d\beta_1\beta_2$. We put

$$\alpha_1\alpha_2 + d\beta_1\beta_2 = \pi\gamma \text{ and } \alpha_1\beta_2 - \alpha_2\beta_1 = \pi\delta,$$

where γ, δ are integers in K, we get

$$1 = \gamma^2 + d\delta^2.$$

Since 1 has only trivial representation, $\gamma = \pm 1$ and $\delta = 0$. Therefore

$$\alpha_1\alpha_2 + d\beta_1\beta_2 = \pm\pi \text{ and } \alpha_1\beta_2 - \alpha_2\beta_1 = 0.$$

Then

$$\alpha_2 = \frac{\beta_2}{\beta_1}\alpha_1 \text{ and } \frac{\beta_2}{\beta_1}\alpha_1^2 + d\beta_1\beta_2 = \pm\pi,$$

and so

$$\frac{\beta_2}{\beta_1}\pi = \frac{\beta_2}{\beta_1}(\alpha_1^2 + d\beta_1^2) = \frac{\beta_2}{\beta_1}\alpha_1^2 + d\beta_1\beta_2 = \pm\pi.$$

Hence $\beta_2 = \pm\beta_1$ and $\alpha_2 = \pm\alpha_1$ and so π has exactly one representation.
Suppose next that the equation (3.3) has an infinitely of solutions $x = \gamma_n, y = \delta_n$ given by (3.1) and (3.2). Let π be a prime in K such that

$$\pi = \alpha^2 + d\beta^2$$

where α and β are integers in K. For a positive integer n, we define

$$\alpha_n + \beta_n\sqrt{-d} = (\gamma_n + \delta_n\sqrt{-d})(\alpha + \beta\sqrt{-d})$$

where

$$\alpha_n = \alpha\gamma_n - d\beta\delta_n \text{ and } \beta_n = \alpha\delta_n + \beta\gamma_n.$$

Thus

$$\alpha_n - \beta_n\sqrt{-d} = (\gamma_n - \delta_n\sqrt{-d})(\alpha - \beta\sqrt{-d})$$

and

$$(\alpha_n + \beta_n\sqrt{-d})(\alpha_n - \beta_n\sqrt{-d}) = (\gamma^2 + d\delta^2)(\alpha^2 + d\beta^2) = \pi.$$

Hence

$$\pi = \alpha_n^2 + d\beta_n^2.$$

We will show that these are all different representations of π.
Suppose for a contradiction that there are $m, n \in \mathbb{N}$ such that $m \neq n$ and

$$\alpha_m = \alpha_n, \beta_m = \beta_n.$$

Then we get

$$\gamma_m = \gamma_n, \delta_m = \delta_n.$$

But in the proof of Theorem 1, $\gamma_m = \gamma_n, \delta_m = \delta_n$. where $m \neq n$ leads to a contradiction. Therefore π has infinitely many representations.
(ii) Assume that 1 has only trivial representation. Let ω be an integer in K. Suppose for a contradiction that ω has infinitely many representations, i.e.,

$$\omega = \alpha_n^2 + d\beta_n^2, n \in \mathbb{N}$$

where α_n and β_n are integers in K and for $m \neq n$, $\alpha_n \neq \pm\alpha_m$ and $\beta_n \neq \pm\beta_m$. Since $\mathcal{O}_K/\omega\mathcal{O}_K$ is finite, there are $m, n \in \mathbb{N}$ such that $m \neq n$ and

$$\alpha_m \equiv \alpha_n \bmod \omega \text{ and } \beta_m \equiv \beta_n \bmod \omega. \tag{3.4}$$

Multiplying the two represeentations

$$\omega = \alpha_m^2 + d\beta_m^2 \text{ and } \omega = \alpha_n^2 + d\beta_n^2,$$

we get

$$\omega^2 = (\alpha_m\alpha_n + d\beta_m\beta_n)^2 + d(\alpha_m\beta_n - \alpha_n\beta_m)^2.$$

It follows from (3.4) that the two numbers

$$\alpha_m\alpha_n + d\beta_m\beta_n \text{ and } \alpha_m\beta_n - \alpha_n\beta_m$$

are divisible by ω. Hence we may put

$$\alpha_m\alpha_n + d\beta_m\beta_n = \omega\gamma \text{ and } \alpha_m\beta_n - \alpha_n\beta_m = \omega\delta$$

where γ and δ are integers in K. Then

$$1 = \gamma^2 + d\delta^2.$$

Since 1 has only trivial representation, $\gamma = \pm 1$ and $\delta = 0$. It follows that

$$\alpha_m\alpha_n + d\beta_m\beta_n = \pm\omega \text{ and } \alpha_m\beta_n - \alpha_n\beta_m = 0.$$

Then

$$\alpha_n = \frac{\beta_n}{\beta_m}\alpha_m \text{ and } \frac{\beta_n}{\beta_m}\alpha_m^2 + d\beta_m\beta_n = \pm\omega,$$

and so

$$\frac{\beta_n}{\beta_m}\omega = \frac{\beta_n}{\beta_m}(\alpha_m^2 + d\beta_m^2) = \frac{\beta_n}{\beta_m}\alpha_m^2 + d\beta_n\beta_m = \pm\omega.$$

Hence $\beta_n = \pm\beta_m$ and $\alpha_n = \pm\alpha_m$. This is a contradiction and so the number of representations must be finite. □

Theorem 3.3. *Let K be a number field and d a positive rational integer. The following statements are equivalent.*

(i) *$K = \mathbb{Q}(\sqrt{-d})$ or K is totally real.*
(ii) *1 has only trivial representation in K.*

Proof. Let K be a number field of degree n. Assume that $K = \mathbb{Q}(\sqrt{-d})$ or K is totally real.
Case 1: $K = \mathbb{Q}(\sqrt{-d})$. Suppose that

$$\alpha^2 + d\beta^2 = 1$$

where α and β are integers in K. Then

$$(\alpha + \beta\sqrt{-d})(\alpha - \beta\sqrt{-d}) = 1.$$

Thus $\alpha + \beta\sqrt{-d}$ and $\alpha - \beta\sqrt{-d}$ are units in K. For $d \neq 1, 3$, the units in K are ± 1 so we have the following system

$$\alpha + \beta\sqrt{-d} = 1 \text{ and } \alpha - \beta\sqrt{-d} = 1$$

or

$$\alpha + \beta\sqrt{-d} = -1 \text{ and } \alpha - \beta\sqrt{-d} = -1.$$

For $d = 1, 3$, we have more cases to figure out. Nevertheless, in either cases we have $\alpha = \pm 1$ and $\beta = 0$. Hence 1 has only trivial representation.
Case 2: K is totally real. Suppose that

$$\alpha^2 + d\beta^2 = 1$$

where α and β are integers in K. Then the conjugate equations

$$1 = (\alpha^{(k)})^2 + d(\beta^{(k)})^2$$

also hold. Since the conjugates are all real, for $d > 1$ we get

$$\left|\beta^{(k)}\right| \leq \frac{1}{d} < 1$$

for every value of k. Thus $\beta = 0$ and $\alpha = \pm 1$. Hence 1 has only trivial representation.
For the converse, assume that $K \neq \mathbb{Q}(\sqrt{-d})$ and K is not totally real. We will prove that 1 has a nontrivial representation.
Case 1: $\sqrt{-d} \in K$. Since $K \neq \mathbb{Q}(\sqrt{-d}), n \geq 4$ and so $t \geq 1$ where t is the rank of the unit group of K. Thus there is a unit ϵ in K such that ϵ is not a root of unity. Then the equation

$$1 = \alpha^2 + d\beta^2$$

is satisfied by the following numbers:

$$\alpha = \frac{1}{2}(\epsilon^m + \epsilon^{-m})$$

and

$$\beta = \frac{1}{2\sqrt{-d}}(\epsilon^m - \epsilon^{-m}),$$

where m is the order of the group $(\mathcal{O}_K/2\sqrt{-d}\mathcal{O}_K)^*$. Note that β is an integer in K because

$$\epsilon^m \equiv 1 \bmod 2\sqrt{-d} \text{ and } \epsilon^{-m} \equiv 1 \bmod 2\sqrt{-d}$$

and α is an integer in K because $\alpha = \sqrt{-d}\beta + \epsilon^{-m}$. Since ϵ is not a root of unity, $\beta \neq 0$. Hence 1 has a nontrivial representation.
Case 2: $\sqrt{-d} \notin K$. Let r be the number of real embeddings of K, s the number of nonconjugate complex embeddings of K and $t = r + s - 1$ the rank of the unit group of K. Let $L = K(\sqrt{-d})$. Then the field L has degree $2n$. Let R be the number of real embeddings of L, S the number of nonconjugate complex embeddings of L

and $T = R + S - 1$ the rank of the unit group of L. Since $\sqrt{-d} \notin \mathbb{R}$, $R = 0$ and $S = r + 2s$ and so

$$T = R + S - 1 = r + 2s - 1 = t + s.$$

Since K is not totally real, $s \geq 1$ and thus

$$T > t.$$

Let us consider the ring consisting of the numbers in L of the form $\lambda + \rho\sqrt{-d}$, where λ and ρ are integers in K. The unit group G of this ring has the rank T. The subgroup G_1 consisting of the squares of the units in G clearly has the same rank T. The units in G_1 cannot all be equal to the product of a unit in K and a root of unity since $t < T$. Hence we conclude that there exists a unit $E = a + b\sqrt{-d}$ in the ring, a and b integers in K such that $a \neq 1$ and $b \neq 0$, and such that E^2 is not equal to the product of a unit in K and a root of unity. Then the number $E_1 = a - b\sqrt{-d}$ is also a unit in L. Hence $a^2 + db^2$ is a unit in K. Then the number

$$1 = \alpha^2 + d\beta^2$$

is satisfied by the following numbers:

$$\alpha = \frac{E^{2m} + E_1^{2m}}{2(a^2 + db^2)^m}$$

and

$$\beta = \frac{E^{2m} - E_1^{2m}}{2\sqrt{-d}(a^2 + db^2)^m}$$

where $m \in \mathbb{N}$. Since $a^2 + db^2$ is a unit in K, α and β are integers in K. If $\beta = 0$, then $E^{2m} = E_1^{2m}$. Hence EE_1^{-1} must be a root of unity and

$$E^2 = (a^2 + db^2)(EE_1^{-1})$$

is a product of units and a root of unity. This is a contradiction. Thus $\beta \neq 0$ and so 1 has a nontrivial representation. □

Acknowledgment

The authors would like to thank Cheranoot Somboonkulavudi for some helpful discussions. The work is supported by the Institue for the Promotion of Teaching Science and Technology.

References

[1] D.A. Cox, *Prime of the form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication,* John Wiley, 1989.
[2] M. Elia and C. Monico, On the representation of primes in $\mathbb{Q}(\sqrt{2})$ as sums of squares, *JP J. Algebra Number Theory Appl.* , 8 (2007), 121 – 133.
[3] T. Nagell, On the A-numbers in the quadratic fields $K(\sqrt{\pm 37})$, *Ark. Mat.* , 4 (1961), 511 – 521.
[4] T. Nagell, On the sum of two integral squares in certain quadratic fields, *Ark. Mat.* , 4 (1960), 267 – 286.
[5] T. Nagell, On the number of representations of an A-number in an algebraic field, *Ark. Mat.* , 4 (1961), 467 – 478.

(S. Sinlapavongsa) Chulalongkorn University, Bangkok, Thailand
E-mail address: sarat1979@hotmail.com

(A. Harnchoowong) Chulalongkorn University, Bangkok, Thailand
E-mail address: ajchara.h@chula.ac.th

TABLE OF CONTENTS, JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.2, 2013

VOLUME 11, NUMBERS 3-4 JULY-OCTOBER 2013
ISSN:1548-5390 PRINT,1559-176X ONLINE

JOURNAL
OF CONCRETE
AND APPLICABLE
MATHEMATICS

EUDOXUS PRESS,LLC

SCOPE AND PRICES OF THE JOURNAL
Journal of Concrete and Applicable Mathematics

A quartely international publication of **Eudoxus Press,LLC**

Editor in Chief: George Anastassiou
Department of Mathematical Sciences,
University of Memphis
Memphis, TN 38152, U.S.A.
ganastss@memphis.edu

Assistant to the Editor:Dr.Razvan Mezei,Lander University,SC 29649, USA.

The main purpose of the "Journal of Concrete and Applicable Mathematics" is to publish high quality original research articles from all subareas of Non-Pure and/or Applicable Mathematics and its many real life applications, as well connections to other areas of Mathematical Sciences, as long as they are presented in a Concrete way. It welcomes also related research survey articles and book reviews.A sample list of connected mathematical areas with this publication includes and is not restricted to: Applied Analysis, Applied Functional Analysis, Probability theory, Stochastic Processes, Approximation Theory, O.D.E, P.D.E, Wavelet, Neural Networks,Difference Equations, Summability, Fractals, Special Functions, Splines, Asymptotic Analysis, Fractional Analysis, Inequalities, Moment Theory, Numerical Functional Analysis,Tomography, Asymptotic Expansions, Fourier Analysis, Applied Harmonic Analysis, Integral Equations, Signal Analysis, Numerical Analysis, Optimization, Operations Research, Linear Programming, Fuzzyness, Mathematical Finance, Stochastic Analysis, Game Theory, Math.Physics aspects, Applied Real and Complex Analysis, Computational Number Theory, Graph Theory, Combinatorics,
Computer Science Math.related topics,combinations of the above, etc. In general any kind of Concretely presented Mathematics which is Applicable fits to the scope of this journal.
Working Concretely and in Applicable Mathematics has become a main trend in many recent years,so we can understand better and deeper and solve the important problems of our real and scientific world.
"Journal of Concrete and Applicable Mathematics" is a peer- reviewed International Quarterly Journal.
We are calling for papers for possible publication. The contributor should send via email the contribution to the editor in-Chief: TEX or LATEX (typed double spaced) and PDF files. [See: Instructions to Contributors]

Journal of Concrete and Applicable Mathematics(JCAAM)
ISSN:1548-5390 PRINT, 1559-176X ONLINE.
is published in January,April,July and October of each year by
EUDOXUS PRESS,LLC,
1424 Beaver Trail Drive,Cordova,TN38016,USA,
Tel.001-901-751-3553
anastassioug@yahoo.com
http://www.EudoxusPress.com.
Visit also www.msci.memphis.edu/~ganastss/jcaam.

Annual Subscription Current Prices:For USA and Canada,Institutional:Print $500,Electronic $250,Print and Electronic $600.Individual:Print $200, Electronic $100,Print &Electronic $250.For any other part of the world add $60 more to the above prices for Print.
Single article PDF file for individual $20.Single issue in PDF form for individual $80.
No credit card payments.Only certified check,money order or international check in US dollars are acceptable.
Combination orders of any two from JoCAAA,JCAAM,JAFA receive 25% discount,all three receive 30% discount.

JCAAM is reviewed and abstracted by AMS Mathematical Reviews,MATHSCI,and Zentralblaat MATH.

JCAAM IS A JOURNAL OF RAPID PUBLICATION

PAGE CHARGES: Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus homepage.

Editorial Board
Associate Editors of Journal of Concrete and Applicable Mathematics

Department of Mathematics Education
College of Education
Gyeongsang National University
Chinju 660-701
KOREA
tel.055-751-5673 Office,
055-755-3644 home,
fax 055-751-6117
yjcho@nongae.gsnu.ac.kr
Nonlinear operator Th.,Inequalities,
Geometry of Banach Spaces

5) Sever S.Dragomir
School of Communications and Informatics
Victoria University of Technology
PO Box 14428
Melbourne City M.C
Victoria 8001,Australia
tel 61 3 9688 4437,fax 61 3 9688 4050
sever.dragomir@vu.edu.au,
sever@sci.vu.edu.au
Math.Analysis,Inequalities,Approximation Th.,
Numerical Analysis, Geometry of Banach Spaces,
Information Th. and Coding

6) Oktay Duman
TOBB University of Economics and Technology,
Department of Mathematics, TR-06530,
Ankara, Turkey, oduman@etu.edu.tr
Classical Approximation Theory,
Summability Theory,
Statistical Convergence and its Applications

7) Angelo Favini
Università di Bologna
Dipartimento di Matematica
Piazza di Porta San Donato 5
40126 Bologna, ITALY
tel.++39 051 2094451
fax.++39 051 2094490
favini@dm.unibo.it
Partial Differential Equations, Control Theory,
Differential Equations in Banach Spaces

8) Claudio A. Fernandez
Facultad de Matematicas
Pontificia Unversidad Católica de Chile
Vicuna Mackenna 4860
Santiago, Chile
tel.++56 2 354 5922

25) Svetlozar T.Rachev
Dept.of Statistics and Applied Probability Program
University of California,Santa Barbara
CA 93106-3110,USA
tel.805-893-4869
rachev@pstat.ucsb.edu
AND
Chair of Econometrics and Statistics
School of Economics and Business Engineering
University of Karlsruhe
Kollegium am Schloss,Bau II,20.12,R210
Postfach 6980,D-76128,Karlsruhe,Germany
tel.011-49-721-608-7535
rachev@lsoe.uni-karlsruhe.de
Mathematical and Empirical Finance,
Applied Probability, Statistics and Econometrics

26) John Michael Rassias
University of Athens
Pedagogical Department
Section of Mathematics and Infomatics
20, Hippocratous Str., Athens, 106 80, Greece

Address for Correspondence
4, Agamemnonos Str.
Aghia Paraskevi, Athens, Attikis 15342 Greece
jrassias@primedu.uoa.gr
jrassias@tellas.gr
Approximation Theory,Functional Equations,
Inequalities, PDE

27) Paolo Emilio Ricci
Universita' degli Studi di Roma "La Sapienza"
Dipartimento di Matematica-Istituto "G.Castelnuovo"
P.le A.Moro,2-00185 Roma,ITALY
tel.++39 0649913201,fax ++39 0644701007
riccip@uniroma1.it,Paoloemilio.Ricci@uniroma1.it
Orthogonal Polynomials and Special functions,
Numerical Analysis, Transforms,Operational Calculus,
Differential and Difference equations

28) Cecil C.Rousseau
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152,USA
tel.901-678-2490,fax 901-678-2480
ccrousse@memphis.edu
Combinatorics,Graph Th.,
Asymptotic Approximations,
Applications to Physics

29) Tomasz Rychlik

fax.++56 2 552 5916
cfernand@mat.puc.cl
Partial Differential Equations,
Mathematical Physics,
Scattering and Spectral Theory

9) A.M.Fink
Department of Mathematics
Iowa State University
Ames,IA 50011-0001,USA
tel.515-294-8150
fink@math.iastate.edu
Inequalities,Ordinary Differential
Equations

10) Sorin Gal
Department of Mathematics
University of Oradea
Str.Armatei Romane 5
3700 Oradea,Romania
galso@uoradea.ro
Approximation Th.,Fuzzyness,Complex
Analysis

11) Jerome A.Goldstein
Department of Mathematical Sciences
The University of Memphis,
Memphis,TN 38152,USA
tel.901-678-2484
jgoldste@memphis.edu
Partial Differential Equations,
Semigroups of Operators

12) Heiner H.Gonska
Department of Mathematics
University of Duisburg
Duisburg,D-47048
Germany
tel.0049-203-379-3542 office
gonska@informatik.uni-duisburg.de
Approximation Th.,Computer Aided
Geometric Design

13) Dmitry Khavinson
Department of Mathematical Sciences
University of Arkansas
Fayetteville,AR 72701,USA
tel.(479)575-6331,fax(479)575-8630
dmitry@uark.edu
Potential Th.,Complex
Analysis,Holomorphic PDE,
Approximation Th.,Function Th.

14) Virginia S.Kiryakova
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences

Institute of Mathematics
Polish Academy of Sciences
Chopina 12,87100 Torun, Poland
T.Rychlik@impan.gov.pl
Mathematical Statistics,Probabilistic
Inequalities

30) Bl. Sendov
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Sofia 1090,Bulgaria
bsendov@bas.bg
Approximation Th.,Geometry of Polynomials,
Image Compression

31) Igor Shevchuk
Faculty of Mathematics and Mechanics
National Taras Shevchenko
University of Kyiv
252017 Kyiv
UKRAINE
shevchuk@univ.kiev.ua
Approximation Theory

32) H.M.Srivastava
Department of Mathematics and Statistics
University of Victoria
Victoria,British Columbia V8W 3P4
Canada
tel.250-721-7455 office,250-477-6960 home,
fax 250-721-8962
harimsri@math.uvic.ca
Real and Complex Analysis,Fractional Calculus
and Appl.,
Integral Equations and Transforms,Higher
Transcendental
Functions and Appl.,q-Series and q-Polynomials,
Analytic Number Th.

33) Stevo Stevic
Mathematical Institute of the Serbian Acad. of
Science
Knez Mihailova 35/I
11000 Beograd, Serbia
sstevic@ptt.yu; sstevo@matf.bg.ac.yu
Complex Variables, Difference Equations,
Approximation Th., Inequalities

34) Ferenc Szidarovszky
Dept.Systems and Industrial Engineering
The University of Arizona
Engineering Building,111
PO.Box 210020
Tucson,AZ 85721-0020,USA
szidar@sie.arizona.edu
Numerical Methods,Game Th.,Dynamic Systems,

Sofia 1090,Bulgaria
virginia@diogenes.bg
Special Functions,Integral Transforms,
Fractional Calculus

15) Hans-Bernd Knoop
Institute of Mathematics
Gerhard Mercator University
D-47048 Duisburg
Germany
tel.0049-203-379-2676
knoop@math.uni-duisburg.de
Approximation Theory,Interpolation

16) Jerry Koliha
Dept. of Mathematics & Statistics
University of Melbourne
VIC 3010,Melbourne
Australia
koliha@unimelb.edu.au
Inequalities,Operator Theory,
Matrix Analysis,Generalized Inverses

17) Robert Kozma
Dept. of Mathematical Sciences
University of Memphis
Memphis, TN 38152, USA
rkozma@memphis.edu
Mathematical Learning Theory,
Dynamic Systems and Chaos,
Complex Dynamics.

18) Mustafa Kulenovic
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
kulenm@math.uri.edu
Differential and Difference Equations

19) Gerassimos Ladas
Department of Mathematics
University of Rhode Island
Kingston,RI 02881,USA
gladas@math.uri.edu
Differential and Difference Equations

20) Rupert Lasser
Institut fur Biomathematik &
Biomertie,GSF
-National Research Center for
environment and health
Ingolstaedter landstr.1
D-85764 Neuherberg,Germany
lasser@gsf.de
Orthogonal Polynomials,Fourier
Analysis,Mathematical Biology.

Multicriteria Decision making,
Conflict Resolution,Applications
in Economics and Natural Resources
Management

35) Gancho Tachev
Dept.of Mathematics
Univ.of Architecture,Civil Eng. and Geodesy
1 Hr.Smirnenski blvd
BG-1421 Sofia,Bulgaria
gtt_fte@uacg.bg
Approximation Theory

36) Manfred Tasche
Department of Mathematics
University of Rostock
D-18051 Rostock
Germany
manfred.tasche@mathematik.uni-rostock.de
Approximation Th.,Wavelet,Fourier Analysis,
Numerical Methods,Signal Processing,
Image Processing,Harmonic Analysis

37) Chris P.Tsokos
Department of Mathematics
University of South Florida
4202 E.Fowler Ave.,PHY 114
Tampa,FL 33620-5700,USA
profcpt@math.usf.edu,profcpt@chuma1.cas.usf.edu
Stochastic Systems,Biomathematics,
Environmental Systems,Reliability Th.

38) Lutz Volkmann
Lehrstuhl II fuer Mathematik
RWTH-Aachen
Templergraben 55
D-52062 Aachen
Germany
volkm@math2.rwth-aachen.de
Complex Analysis,Combinatorics,Graph Theory

Instructions to Contributors
Journal of Concrete and Applicable Mathematics

A quartely international publication of Eudoxus Press, LLC, of TN.

Editor in Chief: George Anastassiou

Department of Mathematical Sciences
University of Memphis
Memphis, TN 38152-3240, U.S.A.

1. Manuscripts files in Latex and PDF and in English, should be submitted via email to the Editor-in-Chief:

Prof.George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis
Memphis,TN 38152, USA.
Tel. 901.678.3144
e-mail: ganastss@memphis.edu

Authors may want to recommend an associate editor the most related to the submission to possibly handle it.

Also authors may want to submit a list of six possible referees, to be used in case we cannot find related referees by ourselves.

2. Manuscripts should be typed using any of TEX,LaTEX,AMS-TEX,or AMS-LaTEX and according to EUDOXUS PRESS, LLC. LATEX STYLE FILE. (Click HERE to save a copy of the style file.)They should be carefully prepared in all respects. Submitted articles should be brightly typed (not dot-matrix), double spaced, in ten point type size and in 8(1/2)x11 inch area per page. Manuscripts should have generous margins on all sides and should not exceed 24 pages.

3. Submission is a representation that the manuscript has not been published previously in this or any other similar form and is not currently under consideration for publication elsewhere. A statement transferring from the authors(or their employers,if they hold the copyright) to Eudoxus Press, LLC, will be required before the manuscript can be accepted for publication.The Editor-in-Chief will supply the necessary forms for this transfer.Such a written transfer of copyright,which previously was assumed to be implicit in the act of submitting a manuscript,is necessary under the U.S.Copyright Law in order for the publisher to carry through the dissemination of research results and reviews as widely and effective as possible.

4. The paper starts with the title of the article, author's name(s) (no titles or degrees), author's affiliation(s) and e-mail addresses. The affiliation should comprise the department, institution (usually university or company), city, state (and/or nation) and mail code.

The following items, 5 and 6, should be on page no. 1 of the paper.

5. An abstract is to be provided, preferably no longer than 150 words.

6. A list of 5 key words is to be provided directly below the abstract. Key words should express the precise content of the manuscript, as they are used for indexing purposes.

The main body of the paper should begin on page no. 1, if possible.

7. All sections should be numbered with Arabic numerals (such as: 1. INTRODUCTION) .
Subsections should be identified with section and subsection numbers (such as 6.1. Second-Value Subheading).
If applicable, an independent single-number system (one for each category) should be used to label all theorems, lemmas, propositions, corrolaries, definitions, remarks, examples, etc. The label (such as Lemma 7) should be typed with paragraph indentation, followed by a period and the lemma itself.

8. Mathematical notation must be typeset. Equations should be numbered consecutively with Arabic numerals in parentheses placed flush right, and should be thusly referred to in the text [such as Eqs.(2) and (5)]. The running title must be placed at the top of even numbered pages and the first author's name, et al., must be placed at the top of the odd numbed pages.

9. Illustrations (photographs, drawings, diagrams, and charts) are to be numbered in one consecutive series of Arabic numerals. The captions for illustrations should be typed double space. All illustrations, charts, tables, etc., must be embedded in the body of the manuscript in proper, final, print position. In particular, manuscript, source, and PDF file version must be at camera ready stage for publication or they cannot be considered.

Tables are to be numbered (with Roman numerals) and referred to by number in the text. Center the title above the table, and type explanatory footnotes (indicated by superscript lowercase letters) below the table.

10. List references alphabetically at the end of the paper and number them consecutively. Each must be cited in the text by the appropriate Arabic numeral in square brackets on the baseline.
References should include (in the following order):
initials of first and middle name, last name of author(s)
title of article,

name of publication, volume number, inclusive pages, and year of publication.

Authors should follow these examples:

Journal Article

1. H.H.Gonska,Degree of simultaneous approximation of bivariate functions by Gordon operators, (journal name in italics) *J. Approx. Theory*, 62,170-191(1990).

Book

2. G.G.Lorentz, (title of book in italics) *Bernstein Polynomials* (2nd ed.), Chelsea,New York,1986.

Contribution to a Book

3. M.K.Khan, Approximation properties of beta operators,in(title of book in italics) *Progress in Approximation Theory* (P.Nevai and A.Pinkus,eds.), Academic Press, New York,1991,pp.483-495.

11. All acknowledgements (including those for a grant and financial support) should occur in one paragraph that directly precedes the References section.

12. Footnotes should be avoided. When their use is absolutely necessary, footnotes should be numbered consecutively using Arabic numerals and should be typed at the bottom of the page to which they refer. Place a line above the footnote, so that it is set off from the text. Use the appropriate superscript numeral for citation in the text.

13. After each revision is made please again submit via email Latex and PDF files of the revised manuscript, including the final one.

14. Effective 1 Nov. 2009 for current journal page charges, contact the Editor in Chief. Upon acceptance of the paper an invoice will be sent to the contact author. The fee payment will be due one month from the invoice date. The article will proceed to publication only after the fee is paid. The charges are to be sent, by money order or certified check, in US dollars, payable to Eudoxus Press, LLC, to the address shown on the Eudoxus homepage.

No galleys will be sent and the contact author will receive one (1) electronic copy of the journal issue in which the article appears.

15. This journal will consider for publication only papers that contain proofs for their listed results.

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 271-276, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON THE MINIMUM RANK AMONG POSITIVE SEMIDEFINITE MATRICES AND TREE SIZE OF A GIVEN GRAPH OF AT MOST SEVEN VERTICES

XINYUN ZHU

Abstract. In this paper, we study the minimum rank among positive semidefinite matrices with a given graph of at most seven vertices (msr) and we list msr for all simple graphs with at most seven vertices which were listed in [1]. For a simple connected graph with no pendent vertices, we study how large $\max_{G\in\mathcal{G}_n}(msr(G) - ts(G) + 1)$ would be, where $ts(G)$ (see Definition 1.1) denote the tree size of a graph G.

Key words. rank, positive semidefinite, graph of a matrix

AMS subject classfication: 15A18,15A57,05C50

1. Introduction

Given a connected graph G with n vertices, we associate to G a set $\mathcal{H}(G)$ of Hermitian $n \times n$ matrices by the following way,

$$\mathcal{H}(G) = \{A | A = A^\star, a_{ij} \neq 0 \quad \text{for} \quad i \neq j \quad \text{if and only if } (i, j) \quad \text{is an edge of} \quad G\}$$

where $A^\star$ is the complex conjugate of A and a_{ij} is the ij-entry of A. We define $\mathcal{P}(G)$ be the subset of $\mathcal{H}(G)$ whose members are positive semidefinite matrices. For a given graph G, $\mathcal{P}(G)$ is non-empty because laplacian matrix $L(G) \in \mathcal{P}(G)$. Define

$$msr(G) = \min_{A\in\mathcal{P}(G)} \text{rank}(A)$$

There are two main results of this paper which were given in Section 3 and Section 4 separately.

A complete list for $msr(G)$ for all the simple graphs with less than seven vertices which were listed in [1] has been given [2]. In this paper, we calculate $msr(G)$ for all the simple graphs with at most seven vertices which were listed in [1], see Section 3. This is one main result of this paper.

To introduce anther main result of this paper, we need the following preparation.

Definition 1.1. An induced subgraph H of a graph G is obtained by deleting all vertices except for the vertices in a subset S. For a graph G, we consider its tree size, denoted $ts(G)$, which is the number of vertices in a maximum induced tree.

Lemma 1.1. [2, Lemma 2.1] *If H is an induced subgraph of a connected graph G, then $msr(H) \leq msr(G)$.*

Given a simple connected graph G, we can define the tree size $ts(G)$ of G. It is known that for a tree T, $msr(T) = ts(T) - 1$. Hence by Lemma 1.1,we get $ts(G) - 1 \leq msr(G)$.

Date: September 20, 2012.

Lemma 1.2. [2, Corollary 3.5] *If a simple connected graph G has a pendant vertex v, which is simply vertex of degree* 1*, then $msr(G) = msr(G - v) + 1$.*

Let $\mathcal{G}_n$ denote the set of all the simple connected graphs with n vertices and with no pendent vertices. We will study $\max_{G \in \mathcal{G}_n}(msr(G) - ts(G) + 1)$. We answer this question for all the members in $\mathcal{G}_n$ with $n \leq 7$ in Theorem 4.1. By giving a counter-example in Section 5, we show that Theorem 4.1 is not true if the vertices number is bigger than 7. For a simple connected graph with no pendent vertices, we ask under what conditions, Theorem 4.1 is still true in Section 5.

2. Finding msr by giving vector representations

Suppose G is a connected graph with vertex set $V(G) = \{v_1, v_2, \cdots, v_n\}$. We call a set of vectors $V = \{\vec{v_1}, \cdots, \vec{v_n}\}$ in $\mathbb{C}_m$ a vector representation (or orthogonal representation) of G if for $i \neq j \in V(G)$, $\langle \vec{v_i}, \vec{v_j} \rangle \neq 0$ whenever i and j are adjacent in G and $\langle \vec{v_i}, \vec{v_j} \rangle = 0$ whenever i and j are not adjacent in G. For any matrix B, $A = B^\star B$ is a semi-definite positive matrix and $\text{rank}(A) = \text{rank}(B)$, where $B^\star$ denotes the conjugate transpose of B. For any semi-definite positive matrix A, there exists a matrix B such that $A = B^\star B$ and $\text{rank}(A) = \text{rank}(B)$.

For a large family of graphs, [2] has given a good way to get msr. But [2] did not solve the msr problem completely. In this section we solve the msr for graphs with at most seven vertices but does not satisfies the conditions of graphs in [2]. We refer the reader to [1] to check the graphs. For the readers convenience, we list all these graphs here: $G706$, $G710$, $G817$, $G836$, $G864$, $G867$, $G870$, $G872$, $G876$, $G877$, $G946$, $G954$, $G955$, $G979$, $G982$, $G992$, $G997 - G1000$, $G1003 - G1007$, $G1053$, $G1056$, $G1060$, $G1065$, $G1069$, $G1084$, $G1089 - G1097$, $G1100$, $G1101$, 1104, $G1105$, $G1123$, $G1125$, $G1135$, $G1145$, $G1146$, $G1148$, $G1149$, $G1152 - G1157$, $G1159$, $G1160$, $G1165$, $G1167$, $G1168$, $G1170$, $G1176$, $G1179$, $G1189$, $G1191$, $G1194 - G1197$ $G1199$, $G1200$, $G1202$, $G1205$, $G1207 - G1212$, $G1222$, $G1224$, $G1228$, $G1230$, $G1231$, $G1233$, $G1241$, $G1242$, $G1248$, $G1250$. We have given the msr of all those graphs by finding the vector representations in [3]. To illustrate how this works, we calculate $msr(G946)$ and $msr(G1104)$ here. We refer the reader to [3] to get the msr of other graphs mentioned above.

In the following examples, for vectors v_i and v_j, by abusing notations, we write $\langle v_i, v_j \rangle = v_i \cdot v_j$. If $v_i \cdot v_j = 0$, then we write $v_i \perp v_j$. Let G be a graph. Let $V(G)$ be the vertices set of G and $E(G)$ the edges set of G. Define

$$N(a) = \{b | b \in V(G), b \neq a, ab \in E(G)\}.$$

Example 2.1. Let $V(G946) = \{1, 2, 3, 4, 5, 6, 7\}$ and let $E(G946)$ be defined by $N(1) = \{2, 6\}$, $N(2) = \{1, 3, 6, 7\}$, $N(3) = \{2, 4, 5, 7\}$, $N(4) = \{3, 5\}$, $N(5) = \{3, 4, 6, 7\}$, $N(6) = \{1, 2, 5, 7\}$, $N(7) = \{2, 3, 5, 6\}$. We claim

$$msr(G946) = 4.$$

The tree size of $G946$ is 3. First let's show that $msr(G946) > 3$.

Let $\{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ be a vector representation of $G946$. We have $v_1 \perp \{v_3, v_4, v_5, v_7\}$, $v_2 \perp \{v_4, v_5\}$, $v_3 \perp \{v_6\}$, $v_4 \perp \{v_6, v_7\}$. If $msr(G946) = 3$, then $v_3 = av_4 + bv_7$ for some nonzero a and b. Hence $0 = av_4 \cdot v_6 + bv_7 \cdot v_6 = bv_7 \cdot v_6$. This is a contradiction.

On the other hand, let

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 1 & 2 & 1 & 0 & 0 & -2 & 1 \\ 0 & 1 & 3 & 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 2 & 1 & 2 & 1 & 2 \\ -1 & -2 & 0 & 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 0 & 2 & 1 & 5 \end{bmatrix}$$

Let

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 2 \end{bmatrix}$$

Then $B^T B = A$, $A \in \mathcal{P}(G946)$ and $\operatorname{rank}(B) = 4$.

Example 2.2. Let $V(G1004) = \{1,2,3,4,5,6,7\}$ and let $E(G1004)$ be defined by $N(1) = \{2,6,7\}$, $N(2) = \{1,3,6,7\}$, $N(3) = \{2,4,5\}$, $N(4) = \{3,6,7\}$, $N(5) = \{3,6,7\}$, $N(6) = \{1,2,4,5\}$, $N(7) = \{1,2,4,5\}$. We claim

$$msr(G1004) = 4$$

The tree size of $G1004$ is 5.

Let

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 1 & 1 & 0 & 1 & -1 \\ 0 & 0 & 2 & 0 & 1 & -1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Let

$$A = B^T B$$

Then $A \in \mathcal{P}(G1004)$ and $\operatorname{rank}(A) = 4$.

3. LIST

In this section, we list the msr for all the simple connected graphs with seven vertices. To get such a list, we use some results from [2],for example, [2, Proposition 2.2], [2, Theorem 2.9], [2, Theorem 3.6], and the vector representation method in Section 2 and [3]. We would like to remind the readers that a complete list of msr of simple connected graphs with less than seven vertices has been given in [2].

msr(G)=6	G270-G274, G276,G278-G280,G284, G286
msr(G)=5	G314-G322,G324-G329, G331-G334, G336-G344, G348-G351, G353,G379-G386, G390-G394,G396, G398-G403, G405-G406, G408, G410-G416, G421-G423, G427-G428, G431-G435, G437-G442, G445-G446, G448-G449, G476-G484, G488-G489, G494, G497, G504, G506-G509, G512-G513, G515-G520, G522-G530, G532-G533, G540-G541, G548, G555-G556, G558-G561, G563-G566, G571, G574-G575, G579,G612-G618, G621-G623, G627-G628, G632, G638, G640, G646, G651, G669-G670, G673-G674, G676-G680, G682-G685, G690, G694, G698, G702,G706, G710,G790-G792, G794-G795, G797, G804, G806, G813, G840, G897
msr(G)=4	G388-G389, G395, G404, G409, G419-G420, G424-G426, G429-G430, G436, G443-G444, G447,G450,G473-G475,G485-G487, G490-G493, G495, G498-G501,G503, G505, G510-G511, G514, G521, G531, G534-G539,G542-G547, G549-G550, G552-G554, G562, G567-G570,G572-G573, G576-G578, G580-G581,G598-G610, G620,G624-G626, G629, G631, G633-G637, G639,G641-G643, G645, G647-G649, G652-G668, G671-G672, G675, G681,G688, G691-G693, G695-G697, G699-G701, G703-G705,G707-G709, G711-G724, G727-G730, G740-G744, G746-G747,G748-G780, G782-G786, G788-G789, G793, G798-G803, G805,G807-G808, G810-G812, G816-G828, G830-G839,G841-G850, G852-G855, G857-G862, G864-G872,G874,G876-G877, G886-G892,G898-G900, G902-G905,G907-G910, G913-G924, G926-G944,G946, G948, G950-G952,G954-G955, G957,G959-G969, G971-G982, G985-G987, G989,G992-G994,G997-G999,G1003-G1004,G1006-G1008,G1014-G1017, G1021,G1025-G1031, G1033-G1034, G1037, G1039, G1041-G1055, G1057-G1061, G1065, G1069,G1078-G1080, G1082-G1085, G1088-G1089, G1091-G1093, G1097,G1101,G1116-G1121,G1127-G1128, G1139-G1143, G1145,G1148,G1150,G1154,G1184-G1185, G1187, G1190.
msr(G)=3	G551, G557, G619, G630, G644, G650, G686-G687,G689 G725-G726, G781, G787, G796, G809, G814-G815,G829,G851,G856 G863, G873, G875,G878,G885, G893-G896,G901,G906,G911-G912, G925, G945, G947,G949, G953, G956,G958, G970, G983-G984,G988, G990-G991, G995-G996,G1000-G1002, G1005,G1012-G1013,G1018-G1020, G1022-G1024,G1035-G1036, G1038, G1040,G1056,G1060,G1062-G1064, G1066-G1068, G1070-G1077, G1081,G1086-G1087,G1090, G1094-G1096, G1098-G1100,G1102-G1106, G1108-G1115,G1122-G1126,G1129-G1138, G1144,G1146-G1147,G1149, G1151-G1153,G1155-G1159,G1161-G1163, G1165-G1171,G1173-G1183,G1186,G1189, G1191-G1210,G1212,G1214, G1215, G1217-G1225,G1227-G1229, G1231-G1233,G1235-G1236, G1239-G1241,G1245-G1246.
msr(G)=2	G1009, G1032,G1164, G1188,G1211, G1213, G1216, G1226,G1230, G1234,G1237-G1238, G1242-G1244, G1247-G1251.
msr(G)=1	G1252

4. Theorem

By observing the results in Section 3, we get the following theorem.

Theorem 4.1. *Let G be a simple connected graph without pendant vertices. If $|V(G)| \leq 7$ and the tree size of G is k, then $msr(G)$ is either k or $k-1$.*

Proof. Define

$$type = (ts, msr)$$

We have the following,

$$type = (6, 5)$$

$G446$, $G448-G449$, $G555-G556$, $G558-G559$, $G669-G670$, $G673-G674$, $G676-G678$, $G682-G685$, $G790$, $G792$, $G794-G795$, $G813$,

$$type = (5, 5)$$

$G445$, $G571$, $G574-G575$, $G579$, $G679-G680$, $G690$, $G694$, $G698$, $G702$, $G706$, $G710$, $G791$, $G797$, $G804$, $G806$, $G840$,

$$type = (5, 4)$$

$G105$, $G127-G129$, $G145-G146$, $G148$, $G161-G162$, $G164$, $G443-G444$, $G447$, $G450$, $G553-G554$, $G570$, $G572$, $G576-G578$, $G580-G581$, $G668$, $G672$, $G675$, $G681$, $G688$, $G691-G693$, $G695-G697$, $G699-G701$, $G703$, $G705$, $G707-G709$, $G711-G724$, $G727-G730$, $G793$, $G798-G801$, $G803$, $G805$, $G807-G808$, $G810-G812$, $G816-G828$, $G830-G839$, $G841-G850$, $G852-G855$, $G857-G862$, $G864-G870$, $G872$, $G874$, $G876$, $G913-G914$, $G918-G924$, $G926-G930$, $G932-G934$, $G936$, $G941-G944$, $G950-G952$, $G957$, $G959-G963$, $G965-G966$, $G969$, $G973-G975$, $G977-G982$, $G985$, $G987$, $G989$, $G992-G994$, $G997$, $G999$ $G1007-G1008$, $G1025-G1026$, $G1030-G1031$, $G1042-G1045$, $G1047-G1048$, $G1050-G1055$, $G1057$, $G1061$, $G1080$, $G1085$, $G1088-G1089$, $G1093$, $G1116$, $G1118$, $G1121$, $G1139-G1140$, $G1142$, $G1148$, $G1160$, $G1184-G1185$, $G1187$,

$$type = (4, 4)$$

$G147$, $G149$, $G152$, $G167$, $G552$, $G573$, $G671$, $G704$, $G802$, $G871$, $G877$, $G915-G917$, $G931$, $G935$, $G937-G940$, $G946$, $G948$, $G954-G955$, $G964$, $G967-G968$, $G971-G972$, $G976$, $G986$, $G998$, $G1003-G1004$, $G1006$, $G1027-G1029$, $G1033-G1034$, $G1037$, $G1039$, $G1041$, $G1046$, $G1049$, $G1058-G1059$, $G1065$, $G1069$, $G1078-G1079$, $G1082-G1084$, $G1091-G1092$, $G1097$, $G1101$, $G1117$, $G1119-G1120$, $G1127-G1128$, $G1141$, $G1143$, $G1145$, $G1150$, $G1154$, $G1190$,

$$type = (4, 3)$$

$G38$, $G43-G44$, $G46-G47$, $G126$, $G130$, $G144$, $G150-G151$, $G153-G154$, $G163$, $G166$, $G168-G175$, $G179-G189$, $G192-G193$, $G196-G198$, $G201-G202$, $G551$, $G557$, $G686-G687$, $G689$, $G725-G726$, $G809$, $G814-G815$, $G829$, $G851$, $G856$, $G863$, $G873$, $G875$, $G878$, $G925$, $G945$, $G947$, $G949$, $G953$, $G956$, $G958$, $G970$, $G983-G984$, $G988$, $G990-G991$, $G995-G996$, $G1000-G1001$, $G1005$, $G1035-G1036$, $G1038$, $G1040$, $G1056$, $G1060$, $G1062-G1064$, $G1066-G1068$, $G1070-G1077$, $G1081$, $G1086-G1087$, $G1090$, $G1094-G1096$, $G1098-G1100$, $G1102-G1106$, $G1113-G1115$, $G1122-G1126$,

$G1129-G1138$, $G1146-G1147$, $G1149$, $G1151-G1153$, $G1155-G1159$, $G1161-G1163$, $G1165-G1171$, $G1175-G1183$, $G1186$, $G1189$, $G1191-G1210$, $G1212$, $G1214-G1215$, $G1217-G1221$, $G1227-G1229$, $G1232-G1233$, $G1240-G1241$, $G1245$,

$$type = (3,3)$$

$G796$, $G1144$, $G1222$, $G1224-G1225$, $G1231$, $G1235-G1236$, $G1239$, $G1246$,

$$type = (3,2)$$

$G16-G17$, $G42$, $G48-G51$, $G165$, $G190$, $G194-G195$, $G199-G200$, $G203-G207$, $G1009$, $G1032$, $G1164$, $G1188$, $G1211$, $G1216$, $G1226$, $G1230$, $G1234$, $G1237-G1238$, $G1242-G1244$, $G1247-G1251$,

$$type = (2,1)$$

$G7$, $G18$, $G52$, $G208$, $G1252$, □

5. Question

Example 5.1. Let G be a simple connected graph with 12 vertices $V(G) = \{1, 2, \cdots, 12\}$ and the edge set is defined by $N(1) = \{2, 3\}$, $N(2) = \{1, 3, 4\}$, $N(3 = \{1, 2, 4, 7\}$, $N(4) = \{2, 3, 5, 7, 8\}$, $N(5) = \{4, 6, 8, 9\}$, $N(6) = \{5, 9\}$, $N(7) = \{3, 4, 8, 10, 11\}$, $N(8) = \{4, 5, 7, 9, 10\}$, $N(9) = \{5, 6, 8\}$, $N(10) = \{7, 8, 11, 12\}$, $N(11) = \{7, 10, 12\}$, $N(12) = \{10, 11\}$. Then G is a chord graph. Then the tree size of G is 8 and $msr(G) = 9$. Hence Theorem 4.1 is false in this case.

Question 5.1. *Let $\mathcal{G}_n$ denote the set of all the simple connected graphs with n vertices and with no pendent vertices. Suppose $n > 7$. Under what conditions is the Theorem 4.1 true?*

Example 5.2. Let $K_{m,n}$ be the complete bipartite graph on m and n vertices. The tree size of $K_{m,n}$ is $max\{m, n\}$. It has been proved that $msr(K_{m,n}) = max\{m, n\}$. Theorem 4.1 is true in this case.

Example 5.3. Let C be a circle with n vertices. The the tree size of C is $n-1$ and $msr(C) = n-2$. Theorem 4.1 is true in this case.

Remark 5.1. [2, Theorem 2.9] gives a sufficient condition for a connected graph G such that $msr(G) = ts(G) - 1$. However not all of members in $\mathcal{G}_7$ satisfy the conditions in [2, Theorem 2.9].

References

[1] R.C. Read and R. J. Wilson, *An atlas of graphs,* Oxford University Press, New York, 1998.

[2] M.Booth, P. Hackney,, B. Harris, C.R. Johnson, M. Lay, L.H. Mitchell, S.K. Narayan, A. Pascoe, K. Steinmetz, B. D. Sutton, and W. Wang, *On the minimal rank among positive semidefinite matrices with a given graph,* SIAM J. Matrix Anal. Appl. **30**(2008), pp. 731–740.

[3] Xinyun Zhu, *Finding msr of a given graph of at most seven vertices by giving vector representations.* arXiv: 1207.5712.

Department of Mathematics, University of Texas of Permian Basin, Odessa, TX 79762
E-mail address: zhu_x@utpb.edu

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 277-291, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

ON THE CONSTRUCTION OF NUMBER SEQUENCE IDENTITIES

WUN-SENG CHOU AND TIAN-XIAO HE

ABSTRACT. to construct a class of identities for number sequences generated by linear recurrence relations. An alternative method based on the generating functions of the sequences is given. The equivalence between two methods for linear recurring sequences are also shown. However, the second method is not limited to the linear recurring sequences, which can be used for a wide class of sequences possessing rational generating functions. As examples, Many new and known identities of Stirling numbers of the second kind, Pell numbers, Jacobsthal numbers, etc., are constructed by using our approach. Finally, we discuss the hyperbolic expression of the identities of linear recurring sequences.

1. INTRODUCTION

Many number and polynomial sequences can be defined, characterized, evaluated, and classified by linear recurrence relations with certain orders. A number sequence $\{a_n\}_{n\geq 0}$ is called sequence of order r if it satisfies a linear recurrence relation of order r

$$a_n = \sum_{j=1}^{r} p_j a_{n-j}, \quad n \geq r, \tag{1.1}$$

for some constants p_j $(j = 1, 2, \ldots, r)$, $p_r \neq 0$, and initial conditions a_j $(j = 0, 1, \ldots, r-1)$. Linear recurrence relations with constant coefficients are important in subjects including pseudo-random number generation, circuit design, and cryptography, and they have been studied extensively. To construct an explicit formula of the general term of a number sequence of order r, one may use generating function, characteristic equation, or a matrix method (See Comtet [6], Hsu [12], Strang [16], Wilf [17], etc.). In [10], He and Shiue presented a method for the sequences of order 2 using the reduction of order, which can be considered as a class of how to make difficult an easy thing. In next section, the method shown in [10] will be modified to give a unified approach to construct a class of identities of linear recurring sequences with any orders. An alternative method will be given in Section 3 by using the generating functions of the recursive sequences discussed in Section 2. The equivalence between these two methods for linear recurring sequences will be shown. However, the second method can be applied for all the sequences with rational generating functions. Inspired by Askey's and Ismail's works shown in [1], [4], and [13], respectively, we discuss the hyperbolic expression of the identities constructed by using our approach, which and another extension will be presented in Section 4.

This work is partially supported by the National Science Council in Taiwan under the grant number 100-2115-M-001-013. The author Wun-Seng Chou would like to thank the Department of Mathematics, Illinois Wesleyan University, for its hospitality.

2. Identities of linear recurring sequences

We now modify the method applied in [10] and extend it to the higher order setting. More precisely, we will give a unify approach to derive identities of linear recurring sequences of arbitrary order r. The key idea can be described in the following theorem.

Theorem 2.1. *Let sequence $\{a_n\}_{n\geq 0}$ be defined by the linear recurrence relation (1.1) of order r, and let its characteristic polynomial $P_r(t)=t^r-\sum_{j=1}^r p_j t^{r-j}$ have r roots α_j ($j=1,2,\ldots,r$), where the root set may be multiset. Denote $a_n^{(j)}:=a_n^{(j-1)}-\alpha_{j-1}a_{n-1}^{(j-1)}$ ($2\leq j\leq r$) and $a_n^{(1)}:=a_n$. Then*

$$a_n^{(r)}=\alpha_r^{n-r+1}a_{r-1}^{(r)}, \tag{2.1}$$

where

$$\begin{aligned} a_n^{(r)} \quad &= a_n-a_{n-1}\sum_{i=1}^{r-1}\alpha_i+a_{n-2}\sum_{1\leq i<j\leq r-1}\alpha_i\alpha_j \\ &-a_{n-3}\sum_{1\leq i<j<k\leq r-1}\alpha_i\alpha_j\alpha_k+\cdots+(-1)^{r-1}a_{n-r+1}\Pi_{i=1}^{r-1}\alpha_i \end{aligned} \tag{2.2}$$

for $n\geq r-1$.

Remark 2.1 $a_n^{(r)}$ shown in (2.2) can be written as

$$a_n^{(r)}=\sum_{i=0}^{r-1}(-1)^n a_{n-i}\sum_{1\leq k_1<\cdots<k_i\leq r-1}\alpha_{k_1}\cdots\alpha_{k_i}.$$

Proof. Denote $a_n^{(2)}:=a_n-\alpha_1 a_{n-1}$. Then the recurrence relation (1.1) can be reduced to

$$\begin{aligned} a_n^{(2)} &= a_{n-1}^{(2)}\sum_{k=2}^{r}\alpha_k-a_{n-2}^{(2)}\sum_{2\leq i<j\leq r}\alpha_i\alpha_j+a_{n-3}^{(2)}\sum_{2\leq i<j<k\leq r}\alpha_i\alpha_j\alpha_k-\cdots \\ &+(-1)^r a_{n-r+1}^{(2)}\Pi_{k=2}^r\alpha_k, \end{aligned} \tag{2.3}$$

a linear recurrence relation of order $r-1$ for sequence $\{a_n^{(2)}\}_{n\geq 0}$. Similarly, we denote $a_n^{(3)}:=a_n^{(2)}-\alpha_2 a_n^{(2)}$. Hence, from (2.3), we obtain

$$\begin{aligned} a_n^{(3)} &= a_{n-1}^{(3)}\sum_{k=3}^{r}\alpha_k-a_{n-2}^{(3)}\sum_{3\leq i<j\leq r}\alpha_i\alpha_j+a_{n-3}^{(3)}\sum_{3\leq i<j<k\leq r}\alpha_i\alpha_j\alpha_k-\cdots \\ &+(-1)^{r-1}a_{n-r+2}^{(3)}\Pi_{k=3}^r\alpha_k. \end{aligned}$$

The above expression is a linear recurrence relation of order $r-2$ for sequence $\{a_n^{(3)}\}_{n\geq 0}$. Repeating this process and denoting $a_n^{(r)}:=a_n^{(r-1)}-\alpha_{r-1}a_{n-1}^{(r-1)}$, we finally obtain

$$a_n^{(r)}=\alpha_r a_{n-1}^{(r)}, \tag{2.4}$$

which implies (2.1). In (2.1), for $n \geq r-1$,

$$\begin{aligned} a_n^{(r)} &= a_n^{(r-1)} - \alpha_{r-1} a_{n-1}^{(r-1)} \\ &= a_n^{(r-2)} - (\alpha_{r-1} + \alpha_{r-2}) a_{n-1}^{(r-2)} + \alpha_{r-1}\alpha_{r-2} a_{n-2}^{(r-2)} \\ &= a_n^{(r-3)} - a_{n-1}^{(r-3)} \sum_{i=r-3}^{r-1} \alpha_i + a_{n-2}^{(r-3)} \sum_{r-3 \leq i < j \leq r-1} \alpha_i \alpha_j - a_{n-3}^{(r-3)} \alpha_{r-3}\alpha_{r-2}\alpha_{r-1}, \end{aligned}$$

which yields (2.2) by using mathematical induction. □

As an example, for $r = 2$, if the characteristic polynomial $P_2(t) = t^2 - p_1 t - p_2$ of (1.1) has roots α_1 and α_2, then we denote $a_n^{(2)} := a_n - \alpha_1 a_{n-1}$ and obtain

$$a_n^{(2)} := a_n - \alpha_1 a_{n-1} = \alpha_2 (a_{n-1} - \alpha_1 a_{n-2}) = \alpha_2 a_{n-1}^{(2)} = \alpha_2^{n-1} a_1^{(n)}.$$

Similarly, for $r = 3$, we denote the roots of the characteristic polynomial $P_3(t) = t^3 - p_1 t^2 - p_2 t - p_3$ of (1.1) by α_j $(j = 1, 2, 3)$. Then,

$$\begin{aligned} & a_n^{(2)} := a_n - \alpha_1 a_{n-1} = (\alpha_2 + \alpha_3) a_{n-1} - \alpha_1 (\alpha_2 + \alpha_3) a_{n-2} - \alpha_2 \alpha_3 (a_{n-2} - \alpha_1 a_{n-3}) \\ = \ & (\alpha_2 + \alpha_3) a_{n-1}^{(2)} - \alpha_2 \alpha_3 a_{n-2}^{(2)}, \end{aligned}$$

which implies

$$a_n^{(3)} := a_n^{(2)} - \alpha_2 a_{n-1}^{(2)} = \alpha_3 (a_{n-1}^{(2)} - \alpha_2 a_{n-2}^{(2)}) = \alpha_3 a_{n-1}^{(3)} = \alpha_3^{n-2} a_2^{(3)}.$$

Remark 2.2 If $\alpha_r = 1$, then (2.1) becomes

$$a_n^{(r)} = a_{r-1}^{(r)}, \tag{2.5}$$

where $a_n^{(r)}$ is shown in (2.2). In particular, for $r = 2$, we have

$$a_n^{(2)} = a_1^{(2)},$$

or equivalently,

$$a_n = \alpha_1 a_{n-1} + a_1 - \alpha_1 a_0.$$

Thus, we have shown that the last non-homogenous recurrence relation of order 1 is equivalent to the homogeneous recurrence relation of order 2, $a_n = (\alpha_1 + 1) a_{n-1} - \alpha_1 a_{n-2}$, for the same sequence $\{a_n\}_{n \geq 0}$. Similarly, we have the equivalence of the homogenous recurrence relation of order 3, $a_n = (p+1) a_{n-1} - (p-q) a_{n-2} + q a_{n-3}$, and the non-homogenous recurrence relation of order 2, $a_n = p a_{n-1} + q a_{n-2} + d$ for uniquely determined constant $d = a_2 - p a_1 - q a_0$.

We now consider three special cases $r = 2$, 3, and 4 for some particular cases of Theorem 2.1.

Corollary 2.2. *Let $\{a_n\}_{n \geq 0}$ be a sequence satisfying the linear recurrence relation of order 2:*

$$a_n = p_1 a_{n-1} + p_2 a_{n-2}, \quad n \geq 2,$$

with initial conditions a_0 and a_1, and let the characteristic polynomial $P_2(t) = t^2 - p_1 t - p_2$ have roots α and β. Then the sequence $\{a_n\}_{n \geq 0}$ satisfies the identity

$$a_n^{(2)} = \beta^{n-1} a_1^{(2)},$$

where $a_n^{(2)} = a_n - \alpha a_{n-1}$ *for* $n \geq 1$

As an example, we consider *Pell number* sequence $\{P_n\}_{n\geq 0}$ generated by the recurrence relation

$$P_n = 2P_{n-1} + P_{n-2}$$

with initial conditions $P_0 = 0$ and $P_1 = 1$. The roots of the characteristic polynomial t^2-2t-1 are $\alpha = 1+\sqrt{2}$ and $\beta = 1-\sqrt{2}$. Hence, Corollary 2.2 gives identity for $n \geq 1$:

$$P_n - (1+\sqrt{2})P_{n-1} = (1-\sqrt{2})^{n-1},$$

or equivalently,

$$(1-\sqrt{2})P_n + P_{n-1} = (1-\sqrt{2})^n.$$

Similarly, we have

$$(1+\sqrt{2})P_n + P_{n-1} = (1+\sqrt{2})^n$$

for $n \geq 1$.

Jacobsthal number sequence $\{J_n\}_{n\geq 0}$ is generated by

$$J_n = J_{n-1} + 2J_{n-2}$$

with initial conditions $J_0 = 0$ and $J_1 = 1$. The characteristic polynomial $t^2 - t - 2$ has two roots $\alpha = 2$ and $\beta = -1$. Hence, from Corollary 2.2, we obtain

$$J_n - 2J_{n-1} = (-1)^{n-1}$$

and

$$J_n + J_{n-1} = 2^{n-1}.$$

For Fibonacci number sequence $\{F_n\}_{n\geq 0}$ and Lucas number sequence $\{L_n\}_{n\geq 0}$, we may use the same argument shown above to construct the well-known identities as follows (see also [14] and [10]) .

$$\begin{aligned}
\frac{1-\sqrt{5}}{2}F_n + F_{n-1} &= \left(\frac{1-\sqrt{5}}{2}\right)^n, \\
\frac{1+\sqrt{5}}{2}F_n + F_{n-1} &= \left(\frac{1+\sqrt{5}}{2}\right)^n, \\
\frac{\sqrt{5}-1}{2}L_n - L_{n-1} &= \sqrt{5}\left(\frac{1-\sqrt{5}}{2}\right)^n, \\
-\frac{\sqrt{5}+1}{2}L_n - L_{n-1} &= -\sqrt{5}\left(\frac{1+\sqrt{5}}{2}\right)^n.
\end{aligned}$$

Corollary 2.3. *Let* $\{a_n\}_{n\geq 0}$ *be a sequence satisfying the linear recurrence relation of order* 3*:*

$$a_n = p_1 a_{n-1} + p_2 a_{n-2} + p_3 a_{n-3}, \quad n \geq 3,$$

with initial conditions a_0 a_1*, and* a_2*, and let the characteristic polynomial* $P_3(t) = t^3 - p_1t^2 - p_2t - p_3$ *have roots* α*,* β*, and* γ*. Then the sequence* $\{a_n\}_{n\geq 0}$ *satisfies the identity*

$$a_n^{(3)} = \gamma^{n-2} a_2^{(3)},$$

where $a_n^{(3)} = a_n - (\alpha+\beta)a_{n-1} + \alpha\beta a_{n-2}$ *for* $n \geq 2$.

Corollary 2.4. *Let $\{a_n\}_{n\geq 0}$ be a sequence satisfying the linear recurrence relation of order 4:*

$$a_n = p_1 a_{n-1} + p_2 a_{n-2} + p_3 a_{n-3} + p_4 a_{n-4}, \quad n \geq 4,$$

with initial conditions a_0 a_1, a_2, and a_3, and let the characteristic polynomial $P_4(t) = t^4 - p_1 t^3 - p_2 t^2 - p_3 t - p_4$ have roots α, β, γ, and δ. Then the sequence $\{a_n\}_{n\geq 0}$ satisfies the identity

$$a_n^{(4)} = \delta^{n-3} a_2^{(4)},$$

where $a_n^{(4)} = a_n - (\alpha+\beta+\gamma)a_{n-1} + (\alpha\beta+\alpha\gamma+\beta\gamma)a_{n-2} + \alpha\beta\gamma a_{n-3}$ for $n \geq 3$.

Examples related to some famous linear recurring sequences in combinatorics are presented below for the applications of Corollaries 2.2, 2.3, and 2.4.

Example 1. We now construct identities for sequences shown in Table 6 of [9] (see also in http://www.research.att.com/ njas/sequences/). Sequence *A001047*, $a_n = 3^n - 2^n = 2\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\}$ satisfies recurrence relation $a_n = 5a_{n-1} - 6a_{n-2}$, where $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ denote Stirling numbers of the second kind. Thus, from Corollary 2.2, we have

$$a_n = 2a_{n-1} + 3^{n-1}, \quad \text{and} \quad a_n = 3a_{n-1} + 2^{n-1},$$

which implies the following identities of Stirling numbers of the second kind:

$$2\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - 4\left\{ \begin{matrix} n \\ 3 \end{matrix} \right\} - 2\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 3^{n-1},$$
$$2\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - 6\left\{ \begin{matrix} n \\ 3 \end{matrix} \right\} - 3\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 2^{n-1}.$$

The above identities imply $2\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} = 3^n - 2^n$, but the converse implication is not obvious.

Similarly, for Sequence *A003462*, $a_n = (3^n - 1)/2$ satisfying $a_n = 4a_{n-1} - 3a_{n-2}$, there hold

$$\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - \left\{ \begin{matrix} n \\ 3 \end{matrix} \right\} - \left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 3^{n-1},$$
$$\left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - 3\left\{ \begin{matrix} n \\ 3 \end{matrix} \right\} - 3\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 1.$$

Mersenne number sequence $a_n = 2^n - 1$ (*A000225*) satisfying $a_n = 3a_{n-1} - 2a_{n-2}$ generates

$$a_n = a_{n-1} + 2^{n-1}, \quad \text{and} \quad a_n = 2a_{n-1} + 1,$$

or equivalently,

$$\left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - \left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 2^{n-1},$$
$$\left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\} - 2\left\{ \begin{matrix} n \\ 2 \end{matrix} \right\} = 1$$

due to $a_n = \left\{ \begin{matrix} n+1 \\ 2 \end{matrix} \right\}$.

We now consider examples of sequences of order 3. If the sequence $\{a_n\}_{n\geq 0}$ satisfies a linear recurrence relation of order 3:

$$a_n = 3ka_{n-1} - (3k^2-1)a_{n-2} + k(k^2-1)a_{n-3}$$

for some positive integer k, then solutions of the equation $t^3 - 3kt^2 + (3k^2-1)t - k(k^2-1) = 0$ are $k \pm 1$ and k. Thus, Corollary 2.3 shows that the sequence satisfies identities

$$a_n - 2ka_{n-1} + (k^2-1)a_{n-2} = k^{n-2}(a_2 - 2ka_1 + (k^2-1)a_0),\ (k>0), \tag{2.6}$$

$$\begin{aligned} a_n &- (2k+1)a_{n-1} + k(k+1)a_{n-2} \\ &= (k-1)^{n-2}(a_2 - (2k+1)a_1 + k(k+1)a_0),\ (k>1), \end{aligned} \tag{2.7}$$

$$\begin{aligned} a_n &- (2k-1)a_{n-1} + k(k-1)a_{n-2} \\ &= (k+1)^{n-2}(a_2 - (2k-1)a_1 + k(k-1)a_0),\ (k>1). \end{aligned} \tag{2.8}$$

In particular, if $a_0 = a_1 = 1$ and $a_2 = 2$, then (2.6)-(2.8) can be written as

$$a_n - 2ka_{n-1} + (k^2-1)a_{n-2} = (k-1)^2k^{n-2},\ (k>0), \tag{2.9}$$

$$a_n - (2k+1)a_{n-1} + k(k+1)a_{n-2} = (k^2-k+1)(k-1)^{n-2},\ (k>1), \tag{2.10}$$

$$a_n - (2k-1)a_{n-1} + k(k-1)a_{n-2} = (k^2-3k+3)(k+1)^{n-2},\ (k>1), \tag{2.11}$$

respectively.

Example 2. Sequence *A129652*, $\{a_n\}_{n\geq 0} = \{1, 1, 2, 7, 26, 91, \ldots\}$, is defined by the linear recurrence relation of order 3:

$$a_n = 6a_{n-1} - 11a_{n-2} + 6a_{n-3}$$

with initial conditions $a_0 = a_1 = 1$ and $a_2 = 2$. It is easy to see the three roots of the characteristic polynomial equation $t^3 - 6t^2 + 11t - 6 = 0$ are 1, 2, and 3. Thus, using (2.9)-(2.11) for $k = 2$, we obtain identities

$$\begin{aligned} a_n - 4a_{n-1} + 3a_{n-2} &= 2^{n-2}, \\ a_n - 5a_{n-1} + 6a_{n-2} &= 3, \\ a_n - 3a_{n-1} + 2a_{n-2} &= 3^{n-2}, \end{aligned}$$

respectively. Let P(A) be the power set of an n-element set A. Then a_{n-1} is the number of pairs of elements x,y of P(A) for which either (1) x and y are disjoint and for which x is not a subset of y and y is not a subset of x, or (2) x and y are intersecting and for which either x is a proper subset of y or y is a proper subset of x, or (3) x = y. Hence, the general term of $\{a_n\}_{n\geq 0} = \{1, 1, 2, 7, 26, 91, \ldots\}$ is $a_n = \left\{ \begin{matrix} n+1 \\ 3 \end{matrix} \right\} + 1$ (See Haye [9]). From the above identities, we obtain identities of Stirling numbers of the second kind as follows:

$$\left\{ \begin{array}{c} n+1 \\ 3 \end{array} \right\} - 4\left\{ \begin{array}{c} n \\ 3 \end{array} \right\} + 3\left\{ \begin{array}{c} n-1 \\ 3 \end{array} \right\} = 2^{n-2},$$
$$\left\{ \begin{array}{c} n+1 \\ 3 \end{array} \right\} - 5\left\{ \begin{array}{c} n \\ 3 \end{array} \right\} + 6\left\{ \begin{array}{c} n-1 \\ 3 \end{array} \right\} = 1,$$
$$\left\{ \begin{array}{c} n+1 \\ 3 \end{array} \right\} - 3\left\{ \begin{array}{c} n \\ 3 \end{array} \right\} + 2\left\{ \begin{array}{c} n-1 \\ 3 \end{array} \right\} = 3^{n-2}.$$

Sequence *A162723*, $\{a_n\}_{n\geq 0} = \{1, 1, 2, 16, 116, 676, \ldots\}$ is defined by $a_n = 9a_{n-1} - 26a_{n-2} + 24a_{n-3}$ with initial conditions $a_0 = a_1 = 1$ and $a_2 = 2$. Its characteristic polynomial $p(t) = t^3 + 9t^2 - 26t + 24$ has roots 2, 3, and 4. Thus, we apply (2.9)-(2.11) for $k = 3$ to the sequence and obtain

$$a_n - 6a_{n-1} + 8a_{n-2} = 4 \cdot 3^{n-2},$$
$$a_n - 7a_{n-1} + 12a_{n-2} = 7 \cdot 2^{n-2},$$
$$a_n - 5a_{n-1} + 6a_{n-2} = 3 \cdot 4^{n-2},$$

respectively.

If a sequence $\{a_n\}_{n\geq 0}$ satisfies the linear recurrence relation of order 3:

$$a_n = 2(k+1)a_{n-1} - (k^2 + 3k + 1)a_{n-2} + k(k+1)a_{n-3}$$

for some positive integer $k \geq 1$, then roots of the characteristic polynomial $P_3(t) = t^3 - 2(k+1)t^2 + (k^2+3k+1)t - k(k+1)$ are 1, k and $k+1$. Thus, Corollary 2.3 shows that the sequence satisfies identities

$$a_n - (2k+1)a_{n-1} + k(k+1)a_{n-2} = a_2 - (2k+1)a_1 + k(k+1)a_0, \tag{2.12}$$
$$a_n - (k+2)a_{n-1} + (k+1)a_{n-2} = k^{n-2}(a_2 - (k+2)a_1 + (k+1)a_0), \tag{2.13}$$
$$a_n - (k+1)a_{n-1} + ka_{n-2} = (k+1)^{n-2}(a_2 - (k+1)a_1 + ka_0). \tag{2.14}$$

In particular, if $a_0 = a_1 = 1$ and $a_2 = 2$, then (2.12)-(2.14) can be written as

$$a_n - (2k+1)a_{n-1} + k(k+1)a_{n-2} = k^2 - k + 1, \tag{2.15}$$
$$a_n - (k+2)a_{n-1} + (k+1)a_{n-2} = k^{n-2}, \tag{2.16}$$
$$a_n - (k+1)a_{n-1} + ka_{n-2} = (k+1)^{n-2}, \tag{2.17}$$

respectively.

Example 3. Consider Sequence *A000325*, $\{a_n\}_{n\geq 0} = \{1, 1, 2, 5, 12, 27, 58, \ldots\}$, which is defined by $a_n = 2^n - n$ and satisfies the recurrence relation

$$a_n = 4a_{n-1} - 5a_{n-2} + 2a_{n-3}, \quad n \geq 3,$$

with initial conditions $a_0 = a_1 = 1$ and $a_2 = 2$. DeSario and Wenstrom [8] have shown that a_n is the number of different permutations of a deck of n cards that can be produced by a single shuffle. From Lascoux and Schutzenberger [15], one may see that a_n is also the number of permutations of degree n with at most one fall, called Grassmannian permutations. Since the corresponding characteristic polynomial equation $t^3 - 4t^3 + 5t^2 - 2 = 0$ has solutions 1, 1, and 2, we may use (2.16)-(2.17) (Note (2.15) and (2.16) are identical for $k = 1$) and obtain identities

$$a_n - 3a_{n-1} + 2a_{n-2} = 1,$$
$$a_n - 2a_{n-1} + a_{n-2} = 2^{n-2},$$

respectively.

For $k = 2$, we obtain Sequence *A129652*, $\{a_n\}_{n\geq 0} = \{1, 1, 2, 7, 26, 91, \ldots\}$ from (2.15)-(2.17). This sequence and its three identities have been presented in Example 2. Similarly, if $k = 3$, we get Sequence *A162725*, $\{a_n\}_{n\geq 0} = \{1, 1, 2, 9, 46, 221, \ldots\}$, which is defined by

$$a_n = 8a_{n-1} - 19a_{n-2} + 12a_{n-2}.$$

Hence, there hold

$$a_n - 7a_{n-1} + 12a_{n-2} = 7,$$
$$a_n - 5a_{n-1} + 4a_{n-2} = 3^{n-2},$$
$$a_n - 4a_{n-1} + 3a_{n-2} = 4^{n-2}.$$

3. An alternative method using the generating functions

Let $\{a_n\}_{n\geq 0}$ be the linear recurring sequence defined by (1.1). Then its generating function $P(t)$ can be written as

$$P(t) = \{a_0 + \sum_{n=1}^{r-1} \left(a_n - \sum_{j=1}^{n} p_j a_{n-j} \right) t^n\}/\{1 - \sum_{j=1}^{r} p_j t^j\}. \tag{3.1}$$

Hence, we have the following result.

Proposition 3.1. *Let the characteristic polynomial of the linear recurring sequence $\{a_n\}_{n\geq 0}$ defined by (1.1) be $p(t) = \Pi_{i=1}^{r}(t - \alpha_i)$. Then the denominator of the generating function $P(t)$ of $\{a_n\}_{n\geq 0}$ equals $\Pi_{i=1}^{r}(1 - \alpha_i t)$.*

The proof is straightforward and omitted. Based on this fact, we may give the following method, which is an alternative method of that presented in Section 2. The equivalence of two methods will be shown later.

Proposition 3.2. *Let $\{a_n\}_{n\geq 0}$ be a sequence with the generating function $P(t) = A(t)/B(t)$, in which $A(t)$ can be expressed as a formal power series and $B(t)$ is a non-null polynomial. Suppose that $B(t)$ can be factored as $B(t) = q_1(t)q_2(t)$ with*

$$q_1(t) = g_0 + g_1 t + g_2 t^2 + \cdots + g_r t^r,$$

then

$$[t^n]P(t) = [t^n]\frac{A(t)}{q_1(t)q_2(t)}$$

implies

$$[t^n]q_1(t)P(t) = [t^n]\frac{A(t)}{q_2(t)},$$

or equivalently,

$$g_0 a_n + g_1 a_{n-1} + \cdots + g_r a_{n-r} = [t^n]\frac{A(t)}{q_2(t)}. \tag{3.2}$$

Observe that no formal proof is necessary, since everything depends on the "rules of the generating function" and the "coefficient of operators", which are immediate. If we are able to extract the coefficient of t^n from $A(t)/q_2(t)$ then we have obtained a (non homogeneous) recurrence relation of order r by using Proposition 3.2.

First, let $\{a_n\}_{n\geq 0}$ be the linear recurring sequence defined by (1.1) and let $P(t)$ be the generating function. From equation (3.1) and Proposition 3.1, we have

$$P(t) = \frac{a_0 + \sum_{i=1}^{r-1}\left(a_i - \sum_{j=1}^{i} p_j a_{i-j}\right)t^i}{\Pi_{i=1}^{r}(1-\alpha_i t)},$$

and so

$$P(t)\Pi_{i=1}^{r-1}(1-\alpha_i t) = \frac{a_0 + \sum_{i=1}^{r-1}\left(a_i - \sum_{j=1}^{i} p_j a_{i-j}\right)t^i}{1-\alpha_r t}. \tag{3.3}$$

Multiplying out the left hand side of (3.3), we have

$$[t^n]P(t)\Pi_{i=1}^{r-1}(1-\alpha_i t) = a_n - a_{n-1}\sum_{i=1}^{r-1}\alpha_i + \cdots + (-1)^{r-1}a_{n-r+1}\Pi_{i=1}^{r-1}\alpha_i = a_n^{(r)} \tag{3.4}$$

for all $n \geq r-1$. Multiplying out the right hand side of (3.3), we have

$$\begin{aligned}
& a_0\alpha_r^n + \sum_{i=1}^{r-1}\left(a_i - \sum_{j=1}^{i} p_j a_{i-j}\right)\alpha_r^{n-i} \\
= \; & \sum_{i=0}^{r-1} a_i\alpha_r^{n-i} - \sum_{i=1}^{r-1}\sum_{j=1}^{i} p_j a_{i-j}\alpha_r^{n-i} \\
= \; & \sum_{i=0}^{r-1} a_i\alpha_r^{n-i} - \sum_{i=0}^{r-1} a_i \sum_{j=1}^{r-1-i} p_j \alpha_r^{n-i-j} \\
= \; & \alpha_r^{n-r+1}\sum_{i=0}^{r-1} a_i\left(\alpha_r^{r-1-i} + \sum_{j=1}^{r-1-i}(-1)^j\left(\sum_{1\leq k_1<\cdots<k_j\leq r}\alpha_{k_1}\cdots\alpha_{k_j}\right)\alpha_r^{r-1-i-j}\right) \\
= \; & \alpha_r^{n-r+1}\sum_{i=0}^{r-1} a_i\left((-1)^{r-1-i}\sum_{1\leq k_1<\cdots<k_{r-1-i}\leq r-1}\alpha_{k_1}\cdots\alpha_{k_{r-1-i}}\right) \\
= \; & \alpha_r^{n-r+1}a_{r-1}^{(r)}
\end{aligned} \tag{3.5}$$

with the convention that $(-1)^{r-1-i}\sum_{1\leq k_1<\cdots<k_{r-1-i}<r}\alpha_{k_1}\cdots\alpha_{k_{r-1-i}} = 1$ for $i = r-1$. From (3.3), (3.4) and (**??**), we have $a_n^{(r)} = \alpha_r^{n-r+1}a_{r-1}^{(r)}$ for all $n \geq r$, which is the same as the result in Theorem 2.1. So, the method in Proposition 3.2 is indeed an alternative method of that presented in Section 2 for linear recurring sequences.

Let $P(t) = A(t)/B(t)$ be the generating function of sequence $\{a_n\}_{n\geq 0}$, in which $B(t)$ is a non-null polynomial of degree r. Then, $B(t)$ can be factoring into a product of r linear factors defined in $\mathbb{C}$, which implies that 2^r identities can be constructed by using the method shown in Proposition 3.2. In other words, there are 2^r identities in $\mathbb{C}$ can be found from the linear recurring sequences defined by the homogeneous recurrence relation (1.1) of order r, because the generating function of $\{a_n\}_{n\geq 0}$ is $A(t)/B(t)$ with $B(t) = 1 - \sum_{j=1}^{r} p_j t^j$. The more important is that this method can be applied to any sequence possessing a rational generating function, or even the numerator is a formal power series. The latest case is the most interesting case in this section. We demonstrate it with some examples as follows.

Example 4. For each nonnegative integer n, the *Fine number* f_n is considered to be the number of rooted trees of order n with root of even degree. In [7], the generating function of Fine number sequence $\{f_n\}_{n\geq 0}$ is presented as

$$F(t) = \frac{1 + 2t - \sqrt{1-4t}}{2t(t+2)}. \tag{3.6}$$

Using *Faà di Bruno's formula*, Chou, Hsu and Shiue [5] give the expressions

$$F(t) = \frac{4t + 2t^2 + \sum_{n\geq 3} \frac{2}{n}\binom{2(n-1)}{n-1} t^n}{2t(t+2)} \tag{3.7}$$

and

$$f_n = \frac{1}{2}\sum_{k=2}^{n} \frac{(-1)^{n-k}}{(k+1)2^{n-k}}\binom{2k}{k}, \quad n \geq 2,$$

with $f_0 = 1$ and $f_1 = 0$. Using Preposition 3.2 and equations (3.6) and (3.7), we may obtain an identity of Fine number f_n for $n \geq 1$ as follows

$$f_{n-1} + 2f_n = \frac{1}{n+1}\binom{2n}{n} = C_n, \tag{3.8}$$

where C_n is the nth Catalan number.

For any non-negative integer n, the *Riordan number* r_n can be viewed as the number of tall bushes with $n+1$ edges (see Bernhart [3]). Let $R(t) = \sum_{n=0}^{\infty} r_n t^n$ be the generating function of Riordan numbers. As shown in [5],

$$R(t) = \frac{1 + t - \sqrt{1 - 2t - 3t^2}}{2t(1+t)} \tag{3.9}$$

with

$$\sqrt{1 - 2t - 3t^2} = 1 - \sum_{n=1}^{\infty} \frac{t^n}{2^{n-1}} \sum_{k=0}^{[\frac{n}{2}]} \frac{(2n-2k-2)!3^k}{(n-k-1)!(n-2k)!k!}. \tag{3.10}$$

[3] also gives that

$$r_n = \frac{1}{n+1}\sum_{m=0}^{n} (-1)^m \binom{n+1}{m}\binom{2n-2m}{n-m}.$$

Using Preposition 3.2 and equations (3.9) and (3.10), we obtain an identity of Riordan number r_n for $n \geq 1$ as follows

$$r_n + r_{n-1} = \frac{1}{2^n} \sum_{k=0}^{[\frac{n+1}{2}]} \frac{(2n-2k)!3^k}{(n-k)!(n-2k+1)!k!}. \tag{3.11}$$

For any non-negative integer n, the *central Delannoy number* d_n is defined by the number of lattice paths on the plane from $(0,0)$ to (n,n) with steps $(1,0)$, $(0,1)$, and $(1,1)$. Let $D(t) = \sum_{n=0}^{\infty} d_n t^n$ be the generating function of central Delannoy numbers. Then we have (see Banderier and Schwer [2])

$$D(t) = \frac{1}{\sqrt{1-6t+t^2}} = \frac{\sqrt{1-6t+t^2}}{1-6t+t^2} \tag{3.12}$$

and

$$d_n = \sum_{i=0}^{n} \left(\begin{array}{c} n \\ i \end{array} \right)^2 2^i.$$

Using the same method in [5], we have

$$\sqrt{1-6t+t^2} = 1 + \sum_{n=1}^{\infty} \frac{3^n t^n}{2^{n-1}} \sum_{i=0}^{[\frac{n}{2}]} \frac{(-1)^i}{3^{2i}(n-i-1)!(n-2i)!i!}. \tag{3.13}$$

From Preposition 3.2 and equations (3.12) and (3.13), we obtain an identity of the central Delannoy number d_n for $n \geq 2$ as

$$d_n - 6d_{n-1} + d_{n-2} = \frac{3^n}{2^{n-1}} \sum_{i=0}^{[\frac{n}{2}]} \frac{(-1)^i}{3^{2i}(n-i-1)!(n-2i)!i!}. \tag{3.14}$$

Note that the roots of $1-6t+t^2$ are $3 \pm \sqrt{2}$. So, one can obtain other two identities for the central Delannoy numbers using Preposition 3.2 and equations (3.12) and (3.13) similarly.

4. Extensions

In the last section, we will apply the following two techniques to derive more identities or to give hyperbolic expressions of identities from the results obtained in Sections 2 and 3.

Proposition 4.1. *Let sequence $\{a_n\}_{n\geq 0}$ be defined by the linear recurrence relation (1.1) of order r and let the characteristic polynomial $P(t) = t^r - \sum_{j=1}^{r} p_j t^{r-j}$ have r roots α_j ($j = 1, 2, \ldots, r$), where the root set may be multiset. Denote*

$$a_n^{(j)} := a_n^{(j-1)} - \alpha_{j-1} a_n^{(j-1)} \quad (2 \leq j \leq r)$$

and $a_n^{(1)} := a_n$. Then there hold identities

$$a_n^{(r)} \pm \left(a_{n+k} - \sum_{j=1}^{r} p_j a_{n+k-j} \right) = \alpha_r^{n-r+1} a_{r-1}^{(r)} \tag{4.1}$$

for any integer k satisfying $n + k \geq r$, where

$$a_n^{(r)} = a_n - a_{n-1}\sum_{i=1}^{r-1}\alpha_i + a_{n-2}\sum_{1\le i<j\le r-1}\alpha_i\alpha_j - a_{n-3}\sum_{1\le i<j<k\le r-1}\alpha_i\alpha_j\alpha_k$$
$$+\cdots+(-1)^{r-1}a_{n-r+1}\Pi_{i=1}^{r-1}\alpha_i \tag{4.2}$$

for $n \ge r-1$.

Example 5. In Section 2 (see the paragraphs after Corollary 2.2), we obtain two identities for Pell number sequence $\{P_n\}_{n\ge 0}$, which has the characteristic polynomial t^2-2t-1. From Proposition 4.1, for $k=1$, we immediately have identities:

$$P_{n+1} - P_n - (2+\sqrt{2})P_{n-1} = (1-\sqrt{2})^{n-1},$$
$$P_{n+1} - P_n - (2-\sqrt{2})P_{n-1} = (1+\sqrt{2})^{n-1}.$$

Similarly, for Jacobsthal number sequence $\{J_n\}_{n\ge 0}$ and Fibonacci number sequence $\{F_n\}_{n\ge 0}$, there hold identities:

$$J_{n+1} - 4J_{n-1} = (-1)^{n-1},$$
$$J_{n+1} - J_{n-1} = 2^{n-1},$$
$$F_{n+1} - \frac{3+\sqrt{5}}{2}F_{n-1} = \left(\frac{1-\sqrt{5}}{2}\right)^{n-1},$$
$$F_{n+1} - \frac{3-\sqrt{5}}{2}F_{n-1} = \left(\frac{1+\sqrt{5}}{2}\right)^{n-1}.$$

From [10], let a $\{a_n\}_{n\ge 0}$ be linear recurring sequence of order 2 satisfying the linear recurrence relation,

$$a_n = pa_{n-1} + qa_{n-2}. \tag{4.3}$$

and denote by α and β the two roots of the characteristic polynomial $p(t) = t^2 - pt - q$, then

$$a_n = \begin{cases} \left(\frac{a_1-\beta a_0}{\alpha-\beta}\right)\alpha^n - \left(\frac{a_1-\alpha a_0}{\alpha-\beta}\right)\beta^n, & if\ \alpha\ne\beta; \\ na_1\alpha^{n-1} - (n-1)a_0\alpha^n, & if\ \alpha=\beta. \end{cases} \tag{4.4}$$

Inspired by [1, 4, 11, 13], denote

$$\alpha(\theta) = \sqrt{q}e^{\theta}, \quad \beta(\theta) = -\sqrt{q}e^{-\theta} \tag{4.5}$$

for some real or complex number θ, where $q>0$. For the case of $q<0$, we denote

$$(\alpha(\theta),\beta(\theta)) = \begin{cases} (\sqrt{-b}e^{\theta}, \sqrt{-b}e^{-\theta}) & for\ p>0, \\ (-\sqrt{-b}e^{\theta}, -\sqrt{-b}e^{-\theta}) & for\ p<0, \end{cases}$$

for some real or complex number θ, and the remaining process is similar, which we leave for the interested readers. From (4.5) we may have

$$p(\theta) = 2\sqrt{q}\sinh(\theta) \tag{4.6}$$

and define a parametric expression of $\{a_n\}_{n\ge 0}$ as

$$a_n(\theta) = 2\sqrt{q}\sinh(\theta)a_{n-1}(\theta) + qa_{n-2}(\theta),\ a_0(\theta) = a_0,\ a_1(\theta) = \frac{2a_1\sqrt{q}}{p}\sinh\theta. \tag{4.7}$$

Obviously, if

$$\theta = \sinh^{-1}\left(\frac{p}{2\sqrt{q}}\right), \tag{4.8}$$

$\{a_n(\theta)\}_{n\geq 0}$ is reduced to $\{a_n\}_{n\geq 0}$.

Substituting expressions (4.5) into (4.4), we obtain

$$\begin{aligned} & a_n(\theta) \\ = \;& q^{(n-1)/2}\frac{a_1(e^{n\theta} - (-1)^n e^{-n\theta}) + \sqrt{q}a_0(e^{(n-1)\theta} + (-1)^n e^{-(n-1)\theta})}{e^\theta + e^{-\theta}} \\ = \;& \begin{cases} \frac{q^{(n-1)/2}}{\cosh\theta}\left(a_1\sinh n\theta + \sqrt{q}a_0\cosh(n-1)\theta\right), & if\, n\, is\, even; \\ \frac{q^{(n-1)/2}}{\cosh\theta}\left(a_1\cosh n\theta + \sqrt{q}a_0\sinh(n-1)\theta\right), & if\, n\, is\, odd. \end{cases} \end{aligned} \tag{4.9}$$

Some properties and extensions of $\{a_n(\theta)\}_{n\geq 0}$ can be derived from (4.9). For instance, from the first equation of (4.9) and using $r = -e^{-2\theta}$, we have

$$a_n(\theta) = q^{(n-1)/2}\left(a_1 e^{(n-1)\theta}\frac{1-r^n}{1-r} + \sqrt{q}a_0 e^{(n-2)\theta}\frac{1-r^{n-1}}{1-r}\right),$$

which enables us to extend the definition of $a_n(\theta)$ to nonpositive values of n.

Since $\alpha(\theta)$ and $\beta(\theta)$ shown in (4.5) are two roots of the characteristic polynomial equation $x^2 - p(\theta)x - q = 0$, we may write (4.3) as

$$a_n(\theta) = (\alpha(\theta) + \beta(\theta))a_{n-1}(\theta) - \alpha(\theta)\beta(\theta)a_{n-2}(\theta), \tag{4.10}$$

where $\alpha(\theta)$ and $\beta(\theta)$ satisfy $\alpha(\theta) + \beta(\theta) = p(\theta)$ and $\alpha(\theta)\beta(\theta) = -q$. Therefore, from (4.10), we have

$$a_n(\theta) - \alpha(\theta)a_{n-1}(\theta) = \beta(\theta)(a_{n-1}(\theta) - \alpha(\theta)a_{n-2}(\theta)), \tag{4.11}$$

which implies

Proposition 4.2. *A sequence $\{a_n(\theta)\}_{n\geq 0}$ of order 2 satisfies the linear recurrence relation (4.3) if and only if it satisfies the non-homogeneous linear recurrence relation of order 1 with the form*

$$a_n(\theta) = \alpha(\theta)a_{n-1}(\theta) + d(\theta)\beta^{n-1}(\theta), \tag{4.12}$$

where $d(\theta)$ is uniquely determined.

Proof. The necessity is clearly from (4.11). We now prove sufficiency. If the sequence $\{a_n(\theta)\}_{n\geq 0}$ satisfies the non-homogeneous recurrence relation of order 1 shown in (4.12), then by substituting $n = 1$ into the above equation we obtain $d(\theta) = a_1(\theta) - \alpha(\theta)a_0$. Thus, (4.12) can be written as

$$a_n(\theta) - \alpha(\theta)a_{n-1}(\theta) = (a_1(\theta) - \alpha(\theta)a_0(\theta))\beta^{n-1}(\theta), \tag{4.13}$$

which yields

$$a_{n-1}(\theta) - \alpha(\theta)a_{n-2}(\theta) = (a_1(\theta) - \alpha(\theta)a_0(\theta))\beta^{n-2}(\theta). \tag{4.14}$$

Multiplying the both sides of (4.14) by $\beta(\theta)$ and evaluating the difference of the resulting equation and (4.13), one immediately knows that $\{a_n(\theta)\}_{n\geq 0}$ satisfies the linear recurrence relation of order 2: $a_n(\theta) = p(\theta)a_{n-1}(\theta)+qa_{n-2}(\theta)$ with $p(\theta) = \alpha(\theta)+\beta(\theta)$ and $q = -\alpha(\theta)\beta(\theta)$. □

Example 6. As an example, we may consider the parametric Fibonacci numbers defined by

$$F_n(\theta) = 2\sinh(\theta)F_{n-1}(\theta) + F_{n-2}(\theta), \quad F_0 = 0, F_1 = 2\sinh(\theta).$$

Here $\alpha(\theta) = e^{\theta}$ and $\beta(\theta) = -e^{-\theta}$. From (4.12) there holds an identity for the parametric Fibonacci numbers

$$F_n(\theta) = e^{\theta}F_{n-1}(\theta) + 2(-1)^{n-1}\sinh(\theta)e^{-(n-1)\theta},$$

or equivalently,

$$-e^{-\theta}F_n(\theta) + F_{n-1}(\theta) = 2(-1)^n e^{-n\theta}\sinh(\theta).$$

Similarly, we have

$$e^{\theta}F_n(\theta) + F_{n-1}(\theta) = 2e^{n\theta}\sinh(\theta).$$

References

[1] R. Askey, Fibonacci and Related Sequences, Mathematics Teacher, (2004), 116-119.
[2] C. Banderier and S. Schwer, Why Delannoy numbers?, J. Statist. Plann. Inference 135 (2005), 40-54.
[3] F. R. Bernhart, Catalan, Motzkin, and Riordan numbers, Discrete Math. 204 (1999), 73-112.
[4] P. S. Bruckman, Advanced Problems and Solutions H460, Fibonacci Quart. 31 (1993), 190-191.
[5] W.-S. Chou, L. C. Hsu, and P. J.-S. Shiue, Application of Faà di Bruno's formula in characterization of inverse relations. J. Comput. Appl. Math. 190 (2006), no. 1-2, 151169.
[6] L. Comtet, Advanced Combinatorics, Reidel, Dordrecht, 1974.
[7] Deutsch, E. and Shapiro, L., A survey of the Fine numbers. Selected papers in honor of Helge Tverberg. Discrete Math. 241 (2001), no. 1-3, 241265.
[8] R. DeSario and L. Wenstrom, Invertible shuffles, Problem 10931, Amer. Math. Monthly, 111 (No. 2, 2004), 169-170.
[9] R. L. Haye, Binary relations on the power set of an n-element set, Journal of Integer Sequences, Vol. 12 (2009), Article 09.2.6.
[10] T. X. He and P. J.-S. Shiue, On sequences of numbers and polynomials defined by second order recurrence relations, Internat. J. Math. and Math. Sci., Volume 2009 (2009), Article ID 709386, 1-21.
[11] T. X. He, P. J.-S. Shiue, and T. W. Weng, Hyperbolic expressions of polynomial sequences defined by linear recurrence relations of order 2, ISRN Discrete Math., Volume 2011, Article ID 674167, 16 pages, doi:10.5402/2011/ 674167.
[12] L. C. Hsu, Computational Combinatorics (Chinese), First edition, Shanghai Scientific & Techincal Publishers, Shanghai, 1983.
[13] M. E. H. Ismail, One parameter generalizations of the Fibonacci and Lucas numbers. Fibonacci Quart. 46/47 (2008/09), no. 2, 167180.
[14] T. Koshy, Fibonacci and Lucas numbers with applications, Pure and Applied Mathematics (New York), Wiley-Interscience, New York, 2001.
[15] A. Lascoux and M.-P. Schutzenberger, Schubert polynomials and the Littlewood Richardson rule, Letters in Math. Physics 10 (1985) 111-124.
[16] G. Strang, Linear algebra and its applications. Second edition. Academic Press (Harcourt Brace Jovanovich, Publishers), New York-London, 1980.
[17] H. S. Wilf, Generatingfunctionology, Academic Press, New York, 1990.

MSC2010: 05A15, 33C45, 65B10, 39A70, 41A80.

Institute of Mathematics, Academia Sinica, and, Department of Mathematical Sciences, National Chengchi Umiversity, Taipei, Taiwan

Department of Mathematics, Illinois Wesleyan University, Bloomington, IL 61702-2900, USA
E-mail address: the@iwu.edu

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 292-302, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

Sensitivity Analysis for Generalized Setvalued Variational Inclusions

George A. Anastassiou
Department of Mathematical Sciences
The University of Memphis, Memphis, TN 38152, USA
ganastss@gmail.com

Salahuddin and M.K. Ahmad
Department of Mathematics
Aligarh Muslim University, Aligarh-202002, India
salahuddin12@mailcity.com; ahmad_kalimuddin@yahoo.co.in

Abstract

Sensitivity analysis for generalized setvalued variational inclusions based on (A, η) resolvent operator technique is investigated. The obtained results encompass a broad range of results.

Keywords: Sensitivity analysis, generalized setvalued variational inclusions, maximal relaxed monotone mapping, (A, η)-monotone mapping, generalized resolvent operator technique.
Mathematics Subject Classification: 49J40, 47H19

1 Introduction

There exists a vast literature on the approximation solvability of nonlinear variational inequalities as well as nonlinear variational inclusions using projection type methods, resolvent operator type methods or averaging techniques. In most of the resolvent operator method, the maximal monotonicity has played a key role, more recently the notions of A-monotonicity and H-monotonicity have not only generalized the maximal monotonicity, but gave a new edge to resolvent operator methods. In [12] the author generalized the recent notion of A-monotonicity to the case of (A, η)-monotonicity while examining the sensitivity analysis for a class of non-linear variational inclusion problems based on the generalized resolvent technique. Resolvent operator techniques have been in use for a while in literature, especially with the general framework involving set valued maximal monotone mapping, but it got a new empowerment by the recent development of A-monotonicity and H-monotonicity, see [2, 3, 4, 8, 17]. Dafermos [1] studied the sensitivity property of solution of a particular kinds of variational inequality on a parameter which takes values on an open subset of Euclidean space R^K. Tobin [14], Verma [15, 20], Salahuddin [11], Salahuddin and Khan [7], Kyparisis [7], Robinson [10], Moudafi [9], Kimura and Yao [5, 6] and Yen and Lee [23] studied the sensitivity analysis of various type of variational inequalities. In this paper, we present the sensitivity analysis for (A, η)-monotone variational inclusion based on the generalized (A, η)-resolvent operator technique. The notion of (A, η)-monotone mappings upgrades the notion of A-monotonicity [16, 18], which generalized the well-known class of maximal monotone mappings to maximal relaxed monotone mappings [12]. The obtained results generalize a wide range of results on the sensitivity analysis for setvalued variational inclusion.

2 (A, η)-Monotonicity

We intend in this section, to explore some basic properties derived from the notion of (A, η)-monotonicity. Let X denote a real Hilbert space with the norm $\|\cdot\|$ and inner product $\langle \cdot, \cdot \rangle$. Let $\eta : X \times X \to X$ be a single-valued mapping. The map η is called τ-Lipschitz continuous if there exists a constant $\tau > 0$ such that

$$\|\eta(u, v)\| \le \tau \|u - v\|, \qquad \forall\, u, v \in X.$$

Definition 2.1 *[21] Let $\eta : X \times X \to X$ be a single-valued mapping and $M : X \to 2^X$ be a multivalued mapping on X. The map M is said to be*

(i) (r, η)-strongly monotone if

$$\langle u^* - v^*, \eta(u, v) \rangle \ge r\|u - v\|^2, \qquad \forall (u, u^*), (v, v^*) \in Graph(M);$$

(ii) η-pseudomonotone if

$$\langle v^*, \eta(u, v) \rangle \ge 0,$$

implies

$$\langle u^*, \eta(u, v) \rangle \ge 0, \qquad \forall (u, u^*), (v, v^*) \in Graph(M);$$

(iii) (m, η)-relaxed monotone if there exists a positive constant m such that

$$\langle u^* - v^*, \eta(u, v) \rangle \ge (-m)\|u - v\|^2, \qquad \forall (u, u^*), (v, v^*) \in Graph(M).$$

(iv) η-firmly nonexpansive if

$$\|u - v\|^2 \le \langle u^* - v^*, \eta(u, v) \rangle \ \forall (u, u^*), (v, v^*) \in Graph(M).$$

Definition 2.2 *A mapping $M : X \to 2^X$ is said to be maximal (m, η)-relaxed monotone if*

(i) M is (M, η) monotone;

(ii) for $(u, u^*) \in X \times X$ and

$$\langle u^* - v^*, \eta(u, v) \rangle \ge (-m)\|u - v\|^2, \qquad \forall\, (v, v^*) \in Graph(M);$$

we have $u^* \in M(u)$.

Definition 2.3 *Let $A : X \to X$ and $\eta : X \times X \to X$ be two single-valued mappings, the map $M : X \to 2^X$ is said to be (A, η)-monotone if*

(i) M is (M, η) relaxed monotone;

2

(ii) $R(A + \rho M) = X$ *for* $\rho > 0$.

Note that alternatively, the map $M : X \to 2^X$ *is said to be* (A, η)*-monotone if*

(i) M *is* (M, η)*-relaxed monotone;*

(ii) $A + \rho M$ *is* η*-pseudomonotone for* $\rho > 0$.

Proposition 2.4 *[22] Let* $A : X \to X$ *be an* (r, η)*-strongly monotone single-valued mapping and* $M : X \to 2^X$ *be an* (A, η)*-monotone mapping. Let* $\eta : X \times X \to X$ *be* τ*-Lipschitz continuous single-valued mapping. Then* M *is maximal* (m, η)*-relaxed monotone and* $(A + \rho M)X = X$ *for* $0 < \rho < \frac{r}{m}$.

Proposition 2.5 *[22] Let* $A : X \to X$ *be an* (r, η)*-strongly monotone single-valued mapping and* $M : X \to 2^X$ *be an* (A, η)*-monotone mapping. In addition, let* $\eta : X \times X \to X$ *be* τ*-Lipschitz continuous. Then* $(A + \rho M)$ *is maximal* η*-monotone for* $0 < \rho < \frac{r}{m}$.

Proposition 2.6 *[22] Let* $A : X \to X$ *be an* (r, η)*-strongly monotone mapping and* $M : X \to 2^X$ *be an* (A, η)*-monotone mapping. If in addition,* $\eta : X \times X \to X$ *is* τ*-Lipschitz continuous, then the operator* $(A + \rho M)^{-1}$ *is single-valued for* $0 < \rho < \frac{r}{m}$.

Definition 2.7 *Let* $A : X \to X$ *be an* (r, η)*-strongly monotone mapping and* $M : X \to 2^X$ *be an* (A, η)*-monotone mapping. Then the generalized resolvent operator* $J^{\eta,M}_{\rho,A} : X \to X$ *is defined by*

$$J^{\eta,M}_{\rho,A}(u) = (A + \rho M)^{-1}(u).$$

Definition 2.8 *The map* $T : X \times X \times L \to X$ *is said to be* r*-strongly monotone with respect to* A *in the first argument if there exists a positive constant* r *such that*

$$\langle T(x, u, \lambda) - T(y, u, \lambda),\ A(x) - A(y) \rangle \geq r\|x - y\|^2, \quad \forall\ (x, y, u, \lambda) \in X \times X \times X \times L.$$

Proposition 2.9 *[22] Let* $\eta : X \times X \to X$ *be a* τ*-Lipschitz continuous mapping,* $A : X \to X$ *be an* (r, η)*-strongly monotone mapping, and* $M : X \to 2^X$ *be an* (A, η)*-monotone mapping. Then the resolvent operator* $R^{\eta,M}_{\rho,A} : X \to X$ *is* $(\frac{\tau}{r - \rho m})$*-Lipschitz continuous that is*

$$\|J^{\eta,M}_{\rho,A}(u) - J^{\eta,M}_{\rho,A}(v)\| \leq \frac{\tau}{r - \rho m}\|u - v\|, \quad \text{for all} \ \ u, v \in X.$$

Definition 2.10 *The map* $f : X \times L \to X$ *is said to be* (p, q)*-relaxed cocoercive if there exist constants* p *and* q *such that*

$$\langle f(u, \lambda) - f(v, \lambda), u - v \rangle \geq -p\|f(u, \lambda) - f(v, \lambda)\|^2 + q\|u - v\|^2,$$

for all $(u, v, \lambda) \in X \times X \times L$.

3 Sensitivity Analysis

Let $T, V, Q : X \times L \to 2^X$ be three setvalued mappings, $N : X \times X \times L \to X$ a nonlinear mapping and $M : X \times X \times L \to 2^X$ be an A-monotone mapping with respect to first mapping T, where L is a nonempty open subset of X. Furthermore let $\eta : X \times X \to X$ be a nonlinear mapping and $f : X \times L \to X$ be a map then the problem of finding an element $u \in X$, $x \in T(u, \lambda)$, $y \in V(u, \lambda), w \in Q(u, \lambda)$ such that

$$f(u, \lambda) \in N(x, y, \lambda) + M(u, w, \lambda) \tag{3.1}$$

where $A \in L$ is the setvalued parameter is called a generalized setvalued variational inclusions.

The solvability of problem (3.1) depends on the equivalence between (3.1) and the problem of finding the fixed point of the associated generalized resolvent operator.

Note that if M is (A, η)-monotone, then the corresponding generalized resolvent operator $J_{\rho,A}^{\eta,M}$ in the first argument is defined by

$$J_{\rho,A}^{\eta,M(.,w)}(u) = (A + \rho M(., w))^{-1}(u), \quad \forall\, u \in X, \tag{3.2}$$

where $\rho > 0$ and A is an (r, η)-strongly monotone mapping.

Lemma 3.1 *[22] Let X be a real Hilbert space and $\eta : X \times X \to X$ be a τ-Lipschitz continuous nonlinear mapping. Let $A : X \to X$ be (r, η)-strongly monotone and $M : X \times X \times L \to 2^X$ be (A, η)-monotone in the first variable with m. Then the generalized resolvent operator associated with $M(., v, \lambda)$ for a fixed $y \in X$ and defined by*

$$J_{\rho,A}^{\eta,M(.,v,\lambda)}(u) = (A + \rho M(., v, \lambda))^{-1}(u), \quad \forall\, u \in X,$$

is $(\frac{\tau^2}{r-\rho m}, \eta)$-firmly nonexpansive.

Lemma 3.2 *Let X be a real Hilbert space, $T, V, Q : X \to 2^X$ be the set valued mappings, let $A : X \to X$ be (r, η)-strongly monotone and $M : X \times X \times L \to 2^X$ be (A, η)-monotone in the first variable. Let $f : X \times L \to X$ be a single valued mapping. Let $\eta : X \times X \to X$ be a τ-Lipschitz continuous nonlinear mapping then the following statements are mutually equivalent:*

(i) An element $u \in X, x \in T(u, \lambda), \quad y \in V(u, \lambda), w \in Q(u, \lambda)$ is a solution to (3.1),

(ii) the map $G : X \times L \to X$ defined by

$$G(u, \lambda) = J_{\rho,A}^{\eta,M(.,w,\lambda)}(A(u) - \rho N(x, y, \lambda) + \rho f(u, \lambda))$$

has a fixed point.

Proof. Let $u \in X, x \in T(u,\lambda), y \in V(u,\lambda), w \in Q(u,\lambda)$ be a solution of (3.1).

$$(3.1) \Longleftrightarrow f(u,\lambda) \in N(x,y,\lambda) + M(u,w,\lambda)$$
$$0 \in N(x,y,\lambda) - f(u,\lambda) + M(u,w,\lambda)$$
$$0 \in \rho N(x,y,\lambda) - \rho f(u,\lambda) + \rho M(u,w,\lambda)$$
$$0 \in -A(u) + \rho N(x,y,\lambda) - \rho f(u,\lambda) + A(u) + \rho M(u,w,\lambda)$$
$$0 \in -(A(u) - \rho N(x,y,\lambda) + \rho f(u,\lambda)) + (A + \rho M(\cdot,w,\lambda)(u)$$
$$u \equiv (A + \rho M(\cdot,w,\lambda))^{-1}(A(u) - \rho N(x,y,\lambda) + \rho f(u,\lambda))$$
$$u = J_{\rho,A}^{\eta,M(\cdot,w,\lambda)}(A(u) - \rho N(x,y,\lambda) + \rho f(u,\lambda))$$

From the Lemma 3.2, we see that (3.1) is equivalent to the fixed point problem of type

$$u \in G(u,\lambda).$$

Therefore

$$G(u,\lambda) = J_{\rho,A}^{\eta,M(\cdot,w,\lambda)}(A(u) - \rho N(x,y,\lambda) + \rho f(u,\lambda))$$

■

Theorem 3.3 *Let X be a real Hilbert space and $\eta : X \times X \to X$ be a τ-Lipschitz continuous nonlinear mapping. Let $A : X \to X$ be (r,η)-strongly monotone and s-Lipschitz continuous and $M : X \times X \times L \to 2^X$ be (A,η)-monotone in the first variable with m. Let $T, V, Q : X \times L \to 2^X$ be the setvalued mappings and $\mathcal{H}$-Lipschitz continuous with γ, σ, ξ constants respectively. Let $f : X \times L \to X$ be the (p,q)-relaxed cocoercive and Lipschitz continuous with constant ζ. Let $N : X \times X \times L \to X$ be α-strongly monotone with respect to A and T and β-Lipschitz continuous in the first variable and let N be μ-Lipschitz continuous in the second variables, and k-relaxed monotone with respect to V. If in addition*

$$\|J_{\rho,A}^{\eta,M(.,w_1,\lambda)}(z) - J_{\rho,A}^{\eta,M(.,w_2,\lambda)}(z)\| \leq \delta\|w_1 - w_2\|$$

,

$$\forall\ (u,v,\lambda) \in X \times X \times L, w_1 \in Q(u,\lambda), w_2 \in Q(v,\lambda). \tag{3.3}$$

Then

$$\|G(u,\lambda) - G(v,\lambda)\| \leq \theta\|u - v\|, \quad \forall\ (u,v,\lambda) \in X \times X \times L, \tag{3.4}$$

where

$$\theta = \frac{\tau}{r - \rho m}[\sqrt{s^2 - 2\rho\alpha + \rho^2\beta^2\gamma^2} + \rho\Omega] + t < 1, \Omega = B + D, t = \delta\xi$$

$$B = \sqrt{(1-2q) + \zeta^2(1+2p)}$$
$$D = \sqrt{1 + 2k + \mu^2\sigma^2},$$

5

$$\Big|\rho-\frac{\alpha\tau^2-r(1-t)(m(1-t)+\Omega\tau)}{\beta^2\gamma^2\tau^2-(m(1-t)+\Omega\tau)^2}\Big| <$$

$$\frac{\sqrt{(\alpha\tau^2-r(1-t)(m(1-t)+\Omega\tau))^2-(\tau^2s^2-r^2(1-t)^2)(\beta^2\gamma^2\tau^2-(m(1-t)+\Omega\tau)^2)}}{\beta^2\gamma^2\tau^2-(m(1-t)+\Omega\tau)^2}$$

$$\alpha\tau^2 > r(1-t)(m(1-t)+\Omega\tau)+\sqrt{(\tau^2s^2-r^2(1-t)^2)(\beta^2\gamma^2\tau^2-(m(1-t)+\Omega\tau)^2)}$$

$$\alpha\tau^2 > r(1-t)(m(1-t)+\Omega\tau)$$

$$\tau s > r(1-t)$$

$$\beta\gamma\tau > m(1-t)+\Omega\tau$$

$$t < 1.$$

Consequently for each $\lambda \in L$, the mapping $G(u,\lambda)$ in the light (3.4) has a unique fixed point $z(\lambda)$. Hence in light of Lemma 3.2, $z(\lambda)$ is a unique solution to (3.1). Thus we have

$$G(z(\lambda),\lambda) = z(\lambda).$$

Proof. For any elements $(u,v,\lambda) \in X \times X \times L$ and $x_1 \in T(u,\lambda)$, $x_2 \in T(v,\lambda)$, $y_1 \in V(u,\lambda)$, $y_2 \in V(v,\lambda)$, $w_1 \in Q(u,\lambda)$, $w_2 \in Q(v,\lambda)$, we have

$$\begin{aligned}
\|G(u,\lambda)-G(v,\lambda)\| &= \|J_{\rho,A}^{\eta,M(.,w_1,\lambda)}(A(u)-\rho N(x_1,y_1,\lambda)+\rho f(u,\lambda)) \\
&\quad -J_{\rho,A}^{\eta,M(.,w_2,\lambda)}(A(v)-\rho N(x_2,y_2,\lambda)+\rho f(v,\lambda))\| \\
&\le \|J_{\rho,A}^{\eta,M(.,w_1,\lambda)}[A(u)-\rho N(x_1,y_1,\lambda)+\rho f(u,\lambda)]-J_{\rho,A}^{\eta,M(.,w_1,\lambda)}[A(v) \\
&\qquad -\rho N(x_2,y_2,\lambda)+\rho f(v,\lambda)]\| \\
&\quad +\|J_{\rho,A}^{\eta,M(.,w_1,\lambda)}[A(v)-\rho N(x_2,y_2,\lambda)+\rho f(v,\lambda)]-J_{\rho,A}^{\eta,M(.,w_2,\lambda)}[A(v) \\
&\qquad -\rho N(x_2,y_2,\lambda)+\rho f(v,\lambda)]\|
\end{aligned}$$

$$\begin{aligned}
&\le \frac{\tau}{r-\rho m}\|A(u)-A(v)-\rho(N(x_1,y_1,\lambda)-N(x_2,y_2,\lambda))+\rho(f(u,\lambda)-f(v,\lambda))\|+\delta\|w_1-w_2\| \\
&\le \frac{\tau}{r-\rho m}[\|A(u)-A(v)-\rho(N(x_1,y_1,\lambda)-N(x_2,y_1,\lambda))\| \\
&\quad +\rho\|-N(x_2,y_1,\lambda)+N(x_2,y_2,\lambda)+f(u,\lambda)-f(v,\lambda)\|]+\delta\|w_1-w_2\| \\
&\le \frac{\tau}{r-\rho m}[\|A(u)-A(v)-\rho(N(x_1,y_1,\lambda)-N(x_2,y_1,\lambda))\|+\rho\|u-v-(N(x_2,y_1,\lambda) \\
&\quad -N(x_2,y_2,\lambda))\|+\rho\|u-v-(f(u,\lambda)-f(v,\lambda))\|]+\delta\|w_1-w_2\|. \qquad (3.5)
\end{aligned}$$

Since T, V, Q are $\mathcal{H}$-Lipschitz continuous and N is Lipschitz continuous with respect to first and second argument, f is Lipschitz continuous, we have

$$\begin{aligned}
\|N(x_1,y_1,\lambda)-N(x_2,y_1,\lambda)\| &\le \beta\|x_1-x_2\| \\
&\le \beta\mathcal{H}(T(u,\lambda),T(v,\lambda)) \\
&\le \beta\gamma\|u-v\|, \qquad (3.6)
\end{aligned}$$

6

$$\begin{aligned}\|N(x_2,y_1,\lambda)-N(x_2,y_2,\lambda)\| &\leq \mu\|y_1-y_2\|\\ &\leq \mu\mathcal{H}(V(u,\lambda),V(v,\lambda))\\ &\leq \mu\sigma\|u-v\|,\end{aligned} \tag{3.7}$$

$$\|f(u,\lambda)-f(v,\lambda)\| \leq \zeta\|u-v\|, \tag{3.8}$$

$$\|w_1-w_2\| \leq \mathcal{H}(Q(u,\lambda),Q(v,\lambda)) \leq \xi\|u-v\|. \tag{3.9}$$

The α-strong monotonicity of N with respect to A, T and (3.6), we have

$$\begin{aligned}\|A(u)-A(v)-\rho(N(x_1,y_1,\lambda)-N(x_2,y_1,\lambda))\|^2 &= \|A(u)-A(v)\|^2-2\rho\langle N(x_1,y_1,\lambda)\\ &\quad -N(x_2,y_1,\lambda),A(u)-A(v)\rangle\\ &\quad +\rho^2\|N(x_1,y_1,\lambda)-N(x_2,y_1,\lambda)\|^2\\ &\leq (s^2-2\rho\alpha+\rho^2\beta^2\gamma^2)\|u-v\|^2.\end{aligned} \tag{3.10}$$

By k-relaxed monotonicity of N with respect to V, and (3.7), we have

$$\begin{aligned}\|u-v-(N(x_2,y_1,\lambda)-N(x_2,y_2,\lambda))\|^2 &= \|u-v\|^2-2\langle N(x_2,y_1,\lambda)-N(x_2,y_2,\lambda),u-v\rangle\\ &\quad +\|N(x_2,y_1,\lambda)-N(x_2,y_2,\lambda)\|^2\\ &\leq \|u-v\|^2+2k\|u-v\|^2+\mu^2\sigma^2\|u-v\|^2\\ &\leq (1+2k+\mu^2\sigma^2)\|u-v\|^2.\end{aligned} \tag{3.11}$$

Again by (p,q)-relaxed cocoercivity of f, and (3.8), we have

$$\begin{aligned}\|u-v-(f(u,\lambda)-f(v,\lambda))\|^2 &= \|u-v\|^2-2\langle f(u,\lambda)-f(v,\lambda),u-v\rangle+\|f(u,\lambda)-f(v,\lambda)\|^2\\ &\leq \|u-v\|^2+2p\|f(u,\lambda)-f(v,\lambda)\|^2-2q\|u-v\|^2+\|f(u,\lambda)-f(v,\lambda)\|^2\\ &\leq (1-2q)\|u-v\|^2+2p\zeta^2\|u-v\|^2+\zeta^2\|u-v\|^2\\ &\leq ((1-2q)+\zeta^2(1+2p))\|u-v\|^2.\end{aligned} \tag{3.12}$$

From (3.5), (3.9),(3.10),(3,11) and (3.12), we obtain

$$\|G(u,\lambda)-G(v,\lambda)\| \leq \theta\|u-v\|,$$

where

$$\theta = \frac{\tau}{(r-\rho m)}[\sqrt{s^2-2\rho\alpha+\rho^2\beta^2\gamma^2}+\rho\Omega]+t<1, \tag{3.13}$$

where

$$D=\sqrt{1+2k+\mu^2\sigma^2}$$
$$B=\sqrt{(1-2q)+\zeta^2(1+2p)}$$

and

$$\Omega = B+D, t=\delta\xi.$$

Since $\theta<1$, it concludes the proof. ■

7

Theorem 3.4 *Let X be a real Hilbert space, $A : X \to X$ be (r,η)-strongly monotone and s-Lipschitz continuous and $M : X \times X \times L \to 2^X$ be (A,η)-monotone in the first variable with m. Let $T, V, Q : X \times L \to 2^X$ be the $\mathcal{H}$ Lipschitz continuous with respect to γ, σ, ξ. Let $f : X \times L \to X$ be the (p,q)-relaxed cocoercive and Lipschitz continuous with constant ζ. Let $N : X \times X \times L \to X$ be α-strongly monotone with respect to T and A and β-Lipschitz in the first variable and let N be μ-Lipschitz continuous in the second variables, and k-relaxed monotone with respect to V. Furthermore let $\eta : X \times X \to X$ be a τ-Lipschitz continuous. In addition if*

$$\|J_{\rho,A}^{\eta,M(.,w_1,\lambda)}(z) - J_{\rho,A}^{\eta,M(.,w_2,\lambda)}(z)\| \le \delta\|w_1 - w_2\|,$$
$$\forall\ (u,v,\lambda) \in X \times X \times L, w_1 \in Q(u,\lambda), w_2 \in Q(v,\lambda).$$

Then

$$\|G(u,\lambda) - G(v,\lambda)\| \le \theta\|u - v\|, \quad \forall\ (u,v,\lambda) \in X \times X \times L, \tag{3.14}$$

where

$$\theta = \frac{\tau}{r - \rho m}[\sqrt{s^2 - 2\rho\alpha + \rho^2\beta^2\gamma^2} + \rho\Omega] + t < 1, \Omega = B + D, t = \delta\xi$$

$$B = \sqrt{(1 - 2q) + \zeta^2(1 + 2p)}$$
$$D = \sqrt{1 + 2k + \mu^2\sigma^2}$$

$$\Big|\rho - \frac{\alpha\tau^2 - r(1-t)(m(1-t) + \Omega\tau)}{\beta^2\gamma^2\tau^2 - (m(1-t) + \Omega\tau)^2}\Big| <$$
$$\frac{\sqrt{(\alpha\tau^2 - r(1-t)(m(1-t) + \Omega\tau))^2 - (\tau^2 s^2 - r^2(1-t)^2)(\beta^2\gamma^2\tau^2 - (m(1-t) + \Omega\tau)^2)}}{\beta^2\gamma^2\tau^2 - (m(1-t) + \Omega\tau)^2}$$

$$\alpha\tau^2 > r(1-t)(m(1-t)+\Omega\tau)+\sqrt{(\tau^2 s^2 - r^2(1-t)^2)(\beta^2\gamma^2\tau^2 - (m(1-t) + \Omega\tau)^2)}$$
$$\alpha\tau^2 > r(1-t)(m(1-t) + \Omega\tau)$$
$$\tau s > r(1-t)$$
$$\beta\gamma\tau > m(1-t) + \Omega\tau$$
$$t < 1.$$

If the mapping $\lambda \to N(x,y,\lambda)$ and $\lambda \to J_{\rho,A}^{\eta,M(.,w,\lambda)}$, $\lambda \to T(u,\lambda)$, $\lambda \to V(u,\lambda), \lambda \to f(u,\lambda), \lambda \to Q(u,\lambda)$ are continuous (or Lipschitz continuous) from L to X, then the solution $z(\lambda)$ of (3.1) is continuous (or Lipschitz continuous) from L to X.

Proof. From the hypothesis of theorem, for any $\lambda, \lambda^* \in L$, we have

$$\begin{aligned}\|z(\lambda) - z(\lambda^*)\| &= \|G(z(\lambda), \lambda) - G(z(\lambda^*), \lambda^*)\| \\ &= \|G(z(\lambda), \lambda) - G(z(\lambda^*), \lambda)\| + \|G(z(\lambda^*), \lambda) - G(z(\lambda^*), \lambda^*)\| \\ &= \theta\|z(\lambda) - z(\lambda^*)\| + \|G(z(\lambda^*), \lambda) - G(z(\lambda^*), \lambda^*)\|.\end{aligned}$$

It follows that

$$\begin{aligned}&\|G(z(\lambda^*), \lambda) - G(z(\lambda^*), \lambda^*)\| = \\ &\quad \|J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}(A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda), y(z(\lambda^*), \lambda), \lambda) + \rho f(z(\lambda^*), \lambda)) \\ &-J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda^*),\lambda^*)}(A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*))\| \\ &\le \|J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}(A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda), y(z(\lambda^*), \lambda), \lambda) + \rho f(z(\lambda^*), \lambda)) \\ &-J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}(A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*))\| \\ &+\|J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*)] \\ &-J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda^*),\lambda^*)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*)]\| \\ &\le \frac{\rho\tau}{r - \rho m}[\|N(x(z(\lambda^*), \lambda), y(z(\lambda^*), \lambda), \lambda) - N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*)\| \\ &\qquad + \|f(z(\lambda^*), \lambda) - f(z(\lambda^*), \lambda^*)\|] \\ &+\|J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*)] \\ &-J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda^*),\lambda^*)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*)]\|.\end{aligned}$$

Hence, we have

$$\begin{aligned}\|z(\lambda) - z(\lambda^*)\| &\le \frac{\rho\tau}{(r - \rho m)(1 - \theta)}[\|N(x(z(\lambda^*), \lambda), y(z(\lambda^*), \lambda), \lambda) \\ &\quad -N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*)\| + \|f(z(\lambda^*), \lambda) - f(z(\lambda^*), \lambda^*)\|] \\ &+\frac{1}{1 - \theta}\|J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda),\lambda)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*)\lambda^*) + \rho f(z(\lambda^*), \lambda^*)] \\ &-J_{\rho,A}^{\eta,M(.,w(z(\lambda^*),\lambda^*),\lambda^*)}[A(z(\lambda^*)) - \rho N(x(z(\lambda^*), \lambda^*), y(z(\lambda^*), \lambda^*), \lambda^*) + \rho f(z(\lambda^*), \lambda^*)]\|.\end{aligned}$$

This completes the proof. ■

References

[1] S. Dafermos, *Sensitivity analysis in variational inequalities*, Math. Oper. Res. **13**(1988), 421–434.

[2] Y. P. Fang, N. J. Huang and H. B. Thompson, *A newsystem of variational inclusions with (H, η)-monotone operators in Hilbert spaces*, Computers and Mathematics with Applications **49**(2-3)(2005), 365-374.

[3] M. M. Jin, *Iterative algorithm for a new system of nonlinear set valued variational inclusions involving (H, η)-monotone mappings*, Journal of Inequalities in Pure and Applied Mathematics, **7**(2) (2006) Article 72, 10 (Electronic).

[4] M. M. Jin, *Perturbed algorithm and stability for strongly monotone nonlinear quasi variational inclusion involving H- accretive operators*, Mathematical Inequalities and Applications, **9**(4) (2006), 771-779.

[5] K. Kimura and J. C. Yao, *Sensitivity analysis of solution mappings of parametricgeneralized quasi vector equilibrium problems*, Taiwanese J. Math. **12**(2008), 2233-2268.

[6] K. Kimura and J. C. Yao, *Semicontinuity of solution mappings of parametric generalized vector equilibrium problems*, JOTA **138** (3) (2008), 429-443.

[7] J. Kyparisis, *Sensitivity analysis framework for variational inequalities*, Mathematical Programming **38**(1987), 203–213.

[8] H. Y. Lan *A class of nonlinear (A, η)- monotone operator inclusion problems with relaxed cocoercive mappings*, Adv. Nonlinear Var. Inequal., **9** (2) (2006), 1-11.

[9] A. Moudafi, *Mixed equilibrium problem, sensitivity analysis and algorithmic aspect*, Computers and Mathematics with Applications **44**(2002), 1099–1108.

[10] S.M. Robinson, *Sensitivity analysis of variational inequalities by normal maps techniques*, In F. Giannessi, A. Maugeri (Eds.) Variational Inequalities and Network and Equilibrium Problems, Plenum, New York, 1995, 257–269.

[11] Salahuddin, *Parametric generalized setvalued variational inclusions and resolvent equations*, J. Math. Anal. Appl. **298**(2004), 146–156.

[12] Salahuddin, *Sensitivity analysis of general mixed multivalued mildly nonlinear variational inequalities*, Math. Computational Appl. **11**(2), August(2006), 121-125.

[13] Salahuddin and M.F.Khan *Sensitivity Analysis for completely generalized nonlinear variational inclusions*, East Asian Math. J. **25**(1) (2009), 45-53.

[14] R.L. Tobin, *Sensitivity analysis for variational inequalities*, JOTA **48**(I)(1986), 191–204.

[15] R.U. Verma, *Sensitivity analysis for relaxed co-coercive nonlinear quasi-variational inclusions*, J. Appl. Math. Stochastic Anal. 2006(2006), 9 PP Art ID 52041.

[16] R.U. Verma, *A monotonicity and application to nonlinear variational inclusion problems*, J. Appl. Math. Stochastic Anal. **17**(2)(2004), 193–195.

[17] R. U. Verma, *Approximation solvability of a class of A-monotone variational inclusion problems*, Journal of Korean Society for Industrial and Applied Mathematics, **8** (1) (2004), 55-66.

[18] R.U. Verma, *A-monotonicity and its role in nonlinear variational inclusions*, JOTA **129**(3)(2006).

[19] R.U. Verma, *Approximation solvability of a class of nonlinear set valued variational inclusions involving (A, η)-monotone mappings*, J. Math. Anal. Appl. **337**(2008), 969-975.

[20] R.U. Verma, *Sensitivity analysis for generalized strongly monotone variational inclusions based on the (A, η)-resolvent operator technique*, Appl. Math. Lett. **19**(2006) 1409-1413.

[21] R. U. Verma *Generalized (A, η)- resolvent operator technique and sensitivity analysis for relaxed cocoercive variational inclusions*, Journal of Inequalities Pure and Applied Mathematics, **7**(3) (2006), 1-7.

[22] R. N. Mohapatra and R. U. Verma, *Sensitivity analysis for cocoercively monotone variational inclusions and (A, η)-maximal monotonicity*, Journal of Applied Mathematics and Computing, **26**(2008).

[23] N.D. Yen and G.M. Lee, *Solution sensitivity of a class of variational inequalities*, J. Math. Anal. Appl. **215**(1997), 48–55.

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 303-316, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

Certain Calculation Regarding the Brownian Motion on the Sphere

Dimitra Kouloumpou and Vassilis G. Papanicolaou
Department of Mathematics
National Technical University of Athens
Zografou Campus 157 80, Athens, GREECE.
dimkulub@mail.ntua.gr, papanico@math.ntua.gr

Abstract

We evaluate explicity certain quantities regarding the Brownian motion process on the n-dimensional sphere of radius a. First we review the transition densities of the process. Then we calculate some probabilistic quantities (e.g. moments) of the exit times of specific domains.
Key word and phrases: n-dimensional sphere, stereographic projection coordinates, Brownian motion, exit times, transition densities, reflection principle

1 Introduction

1.1 The n-Sphere

Let $n \in \mathbb{N} = \{1, 2, 3, \ldots\}$. The n-dimensional sphere S^n with center $(c_1, ..., c_{n+1})$ and radius $a > 0$ is the set of all points $x \in \mathbb{R}^{n+1}$ satisfying

$$(x_1 - c_1)^2 + \cdots + (x_{n+1} - c_{n+1})^2 = a^2.$$

The most interesting case in applications is, of course, the case $n = 2$. For the sake of comparison we will also discuss the cases $n = 1$ (i.e. the circle) and $n = 3$. In some cases we will even consider the case of general n.

1.2 Stereographic Projection Coordinates

Consider the n-sphere, $n \geq 2$,

$$x_1^2 + \cdots + x_n^2 + (x_{n+1} - a)^2 = a^2.$$

To each point $(x_1, ..., x_n, x_{n+1})$ of this sphere, other than its "north pole" $N = (0, ..., 0, 2a)$ we associate the coordinates

$$\xi_1 = \frac{2ax_1}{2a - x_{n+1}}, \ \ldots \ , \xi_n = \frac{2ax_n}{2a - x_{n+1}}.$$

Given the coordinates $(\xi_1, ..., \xi_n)$ of a point on the sphere with Cartesian coordinates $(x_1, ..., x_n, x_{n+1})$, we have

$$x_1 = \frac{4a^2\xi_1}{\xi_1^2 + \cdots + \xi_n^2 + 4a^2}, \; \ldots \; , x_n = \frac{4a^2\xi_n}{\xi_1^2 + \cdots + \xi_n^2 + 4a^2}, \; x_{n+1} = \frac{2a\left(\xi_1^2 + \cdots + \xi_n^2\right)}{\xi_1^2 + \cdots + \xi_n^2 + 4a^2}.$$

1.3 Spherical Coordinates

The points of the n-sphere

$$x_1^2 + \cdots + x_n^2 + x_{n+1}^2 = a^2$$

may also be described in spherical coordinates $(\theta_1, ..., \theta_{n-1}, \varphi)$ as follows:

- For $n = 1$, $x_1 = a\cos\varphi$, $x_2 = a\sin\varphi$, where $0 \leq \varphi < 2\pi$.
- In general for $n \geq 2$
  $x_1 = a\cos\theta_1 \prod_{i=1}^n \sin\theta_i$, $x_2 = a\prod_{i=2}^n \sin\theta_i$, $x_k = a\cos\theta_{k-1}\prod_{i=k}^n \sin\theta_i$,
  for $k = 3, 4, ..., n$
  and $x_{n+1} = a\cos\theta_n$, where $0 \leq \theta_1 < 2\pi$, $0 \leq \theta_i \leq \pi$,
  for $i = 2, 3, ..., n$,

1.4 The Laplace-Beltrami Operator

In spherical coordinates: The Laplace-Beltrami operator of a smooth function f on S^1 is

$$\Delta_1 f = \frac{1}{a^2}\frac{\partial^2 f}{\partial\varphi^2}. \tag{1.1}$$

The Laplace-Beltrami operator of a smooth function f on S^2 is

$$\Delta_2 f = \frac{1}{a^2\sin\varphi}\left(\frac{f_{\theta\theta}}{\sin\varphi} + f_\varphi\cos\varphi + f_{\varphi\varphi}\sin\varphi\right). \tag{1.2}$$

The Laplace-Beltrami operator of a smooth function f on S^3 is

$$\Delta_3 f = \frac{1}{a^2\sin^2\varphi}\left[\frac{1}{\sin^2\theta_2}\cdot\frac{\partial^2 f}{\partial\theta_1^2} + \frac{1}{\sin\theta_2}\cdot\frac{\partial}{\partial\theta_2}\left(\frac{\partial f}{\partial\theta_2}\sin\theta_2\right) + \frac{\partial}{\partial\varphi}\left(\frac{\partial f}{\partial\varphi}\sin^2\varphi\right)\right] \tag{1.3}$$

In stereographic projection coordinates: The Laplace-Beltrami operator of a smooth function f on S^n, $n \geq 2$ is

$$\Delta_n f = \frac{\left(\xi_1^2 + \cdots + \xi_n^2 + 4a^2\right)^2}{16a^4}\left[\sum_{i=1}^n \frac{\partial^2 f}{\partial\xi_i^2} - \frac{2(n-2)}{(\xi_1^2 + \cdots + \xi_n^2 + 4a^2)}\sum_{i=1}^n \xi_i\frac{\partial f}{\partial\xi_i}\right]. \tag{1.4}$$

In particular, for $n = 2$ we get

$$\Delta_2 f = \frac{\left(\xi_1^2 + \xi_2^2 + 4a^2\right)^2}{16a^4}\left(\frac{\partial^2 f}{\partial\xi_1^2} + \frac{\partial^2 f}{\partial\xi_2^2}\right). \tag{1.5}$$

1.5 Brownian Motion on S^n

The Brownian motion on S^n , starting from $x \in S^n$, is a diffusion (Markov) process $X_t, t \geq 0$, on S^n whose transition density is a function $P(t,x,y)$ on $(0,\infty) \times S^n \times S^n$ satisfying

$$\frac{\partial P}{\partial t} = \frac{1}{2}\Delta_n P \tag{1.6}$$

$$P(t,x,y) \to \delta_x(y) \quad \text{as} \quad t \to 0^+, \tag{1.7}$$

where Δ_n is the Laplace-Beltrami operator of S^n acting on the x-variables and $\delta_x(y)$ is the delta mass at x, i.e. $P(t,x,y)$ is the **heat kernel** of S^n. The heat kernel exists, it is unique, positive, and smooth in (t,x,y) [2].

1.5.1 Further Properties of the Heat Kernel $P(t,x,y)$

It is well known that $P(t,x,y)$ satisfies the following properties [2]

1. Symmetry: $P(t,x,y) = P(t,y,x)$.

2. The semigroup identity: For any $s \in (0,t)$,

$$P(t,x,y) = \int_{S^n} P(s,x,z)P(t-s,z,y)d\mu(z)$$

   where $d\mu$ is the n-th dimensional surface area.

3. For all $t > 0$ and $x \in S^n$

$$\int_{S^n} P(t,x,y)d\mu(y) = 1.$$

4. As $t \to \infty$, $P(t,x,y)$ approaches the uniform density on S^n, i.e.

$$\lim_{t\to\infty} P(t,x,y) = \frac{1}{A_n},$$

   where A_n is nth dimensional surface area of S^n with radius a. It is well known that [3]

$$A_{2k+1} = \frac{2\pi^{k+1}a^{2k+1}}{(k)!}, \quad \text{and} \quad A_{2k} = \frac{2^{2k}(k-1)!\pi^k a^{2k}}{(2k-1)!}, k \in \mathbb{N}$$

Finally, the symmetry of S^n implies that $P(t,x,y)$ depends only on t and $d(x,y)$, the distance between x and y. Thus in spherical coordinates it depends on t and the angle φ between x and y. Hence

$$P(t,x,y) = p(t,\varphi),$$

where $p(t,\varphi)$ satisfies

$$\frac{\partial p}{\partial t} = \frac{1}{2}\Delta_n p = \frac{1}{2a^2}\left[(n-1)\cot\varphi \cdot \frac{\partial p}{\partial \varphi} + \frac{\partial^2 p}{\partial \varphi^2}\right] \tag{1.8}$$

and

$$\lim_{t\to 0^+} aA_{n-1}p(t,\varphi) \cdot \sin^{n-1}\varphi = \delta(\varphi). \tag{1.9}$$

Here $\delta(\cdot)$ is the standard Dirac delta function on $\mathbb{R}$.

2 Explicit Form of the Heat Kernel

Reminder (Poisson Summation Formula)**.** Let $f(x)$ be a function in the Schwartz space $\mathcal{S}(\mathbb{R})$, where $\mathcal{S}(\mathbb{R})$ consists of the set of all infinitely differentiable functions f on $\mathbb{R}$ so that f and all its derivatives $f^{(l)}$ are rapidly decreasing, in the sense that

$$\sup_{x\in\mathbb{R}} |x|^k \left| f^{(l)}(x) \right| < \infty \quad \text{for every} \quad k, l \geq 0.$$

Then

$$\sum_{n\in\mathbb{Z}} f(x + 2\pi n) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}} F(n) \exp(inx),$$

where $F(\xi)$ is the Fourier transform of $f(x)$, i.e.

$$F(\xi) = \int_{-\infty}^{+\infty} f(x) \exp(-i\xi x) dx, \quad \xi \in \mathbb{R}.$$

2.1 The Case of S^1

Proposition 2.1 *The transition density function of the Brownian motion $X_t, t \geq 0$ on S^1 with radius a is the function*

$$p(t, \varphi) = \frac{1}{2\pi a} \sum_{n\in\mathbb{Z}} \exp\left(-\frac{n^2 t}{2a^2} + in\varphi\right) = \frac{1}{\pi a} \sum_{n\in\mathbb{N}} \left[\exp\left(-\frac{n^2 t}{2a^2}\right) \cos(n\varphi)\right] - \frac{1}{2\pi a}, \tag{2.1}$$

equivalently

$$p(t, \varphi) = \frac{1}{\sqrt{2\pi t}} \sum_{n\in\mathbb{Z}} \exp\left(-\frac{a^2}{2t} (\varphi - 2\pi n)^2\right). \tag{2.2}$$

For the proof see [5].

2.2 The Case of S^2

We remind the reader of Legendre Polynomials $P_n(x)$, $n = 0, 1, 2, ...$ since we are going to use them later in the paper

$$P_n(x) = \frac{1}{2^n n!} \cdot \frac{d^n}{dx^n} \left[\left(x^2 - 1\right)^n \right].$$

Proposition 2.2 *The transition density function of the Brownian motion X_t, $t \geq 0$ on S^2 with radius a is given by the function*

$$p(t, \varphi) = \frac{1}{4\pi a^2} \sum_{n\in\mathbb{N}} (2n+1) \exp\left(-\frac{n(n+1)\sqrt{t}}{a}\right) P_n(\cos\varphi). \tag{2.3}$$

For the proof see [1] or [5].

2.3 The Case of S^3

Proposition 2.3 *Let $X_t,\ t \geq 0$ be the Brownian motion on a 3-dimensional sphere S^3 of radius a. The transition density function $p(t,\varphi)$ of X_t is given by*

$$p(t,\varphi) = \frac{\exp\left(\frac{t}{2a^2}\right)}{(2\pi t)^{3/2}\sin\varphi}\sum_{n\in\mathbb{Z}}(\varphi+2n\pi)\exp\left(-\frac{(\varphi+2n\pi)^2a^2}{2t}\right), \tag{2.4}$$

where $\mathbb{Z}$ is the set of all integers. Equivalently,

$$p(t,\varphi) = -\frac{i}{4\pi^2a^3\sin\varphi}\sum_{n\in\mathbb{Z}} n\exp\left(-\frac{t(n^2-1)}{2a^2}+i\varphi n\right), \tag{2.5}$$

$$p(t,\varphi) = \frac{1}{2\pi^2a^3\sin\varphi}\sum_{n\in\mathbb{N}} n\sin(n\varphi)\exp\left(-\frac{t(n^2-1)}{2a^2}\right). \tag{2.6}$$

The function $p(t,\varphi)$ is analytic at $\varphi=0$ and $\varphi=\pi$. In fact

$$p(t,0) = \lim_{\varphi\to 0^+} p(t,\varphi) = \frac{1}{2\pi^2a^3}\sum_{n\in\mathbb{N}} n^2\exp\left(-\frac{t(n^2-1)}{2a^2}\right)$$

and

$$p(t,\pi) = \lim_{\varphi\to\pi^-} p(t,\varphi) = \frac{1}{2\pi^2a^3}\sum_{n\in\mathbb{N}} n^2(-1)^n\exp\left(-\frac{t(n^2-1)}{2a^2}\right).$$

For the proof see [5].

Reminder. The ϑ_3 function of Jacobi is

$$\vartheta_3(z,r) = 1+2\sum_{n=0}^{\infty}\exp\left(i\pi r n^2\right)\cos(2nz),$$

where $r\in\mathbb{C}$, with $\operatorname{Im}\{r\}>0$. It follows that

$$p(t,\varphi) = -\frac{1}{4\pi^2a^3\sin\varphi}\exp\left(\frac{t}{2a^2}\right)\frac{\partial}{\partial\varphi}\vartheta_3\left(\frac{\varphi}{2},\frac{ti}{2a^2\pi}\right).$$

3 Expectations of Exit Times

Let X_t be the Brownian motion in S^n and D a Borel subset of S^n. The random variable

$$T = \inf\{t\geq 0|\ X_t\notin D\}$$

is called the (first) exit time of D.
Reminder. If $u(x)=E^x[T]$, then $u(x)$ satisfies

$$\frac{1}{2}\Delta_n u = -1, \quad u|_{\partial D} = 0 \tag{3.1}$$

(see, e.g., [8])

Proposition 3.1 *We consider the 2-dimensional sphere S^2 of radius a. Let two circles pass through the North pole, such that in stereographic coordinates are represented by the parallel lines $\xi_1 = b$ and $\xi_2 = c$, where $b, c \in \mathbb{R}$, say $b < c$. We consider the set D in S^2, whose stereographic projection is*

$$D = \{(\xi_1, \xi_2) \mid \xi_1 \in \mathbb{R} \quad and \quad \xi_2 \in (b, c)\}.$$

Of course

$$\partial D = \{(\xi_1, \xi_2) \mid \xi_1 \in \mathbb{R} \quad and \quad \xi_2 = b \quad or \quad \xi_2 = c\}.$$

If X_t is the Brownian motion on S^2 of radius a starting at the point A, where the stereogrpaphic projection coordinates of A are $(\xi_1, \xi_2) \in D$ and

$$T = \inf\{t \geq 0 \,|\, X_t \in D\},$$

then

$$E^A[T] = f(\xi_1, \xi_2) - 2a^2 \ln\left(\xi_1^2 + \xi_2^2 + 4a^2\right), \tag{3.2}$$

where

$$f(\xi_1,\xi_2) = \frac{1}{\pi}\int_0^\infty \left[\frac{g(\eta,c)\exp\left(\frac{\pi\xi_1}{c-b}\right)\sin\left(\frac{\pi(\xi_2-b)}{c-b}\right)}{\exp\left(\frac{2\pi\xi_1}{c-b}\right)\sin^2\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\left(\exp\left(\frac{\pi\xi_1}{c-b}\right)\cos\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\eta\right)^2}\right] d\eta$$

$$+\frac{1}{\pi}\int_0^\infty \left[\frac{g(\eta,b)\exp\left(\frac{\pi\xi_1}{c-b}\right)\sin\left(\frac{\pi(\xi_2-b)}{c-b}\right)}{\exp\left(\frac{2\pi\xi_1}{c-b}\right)\sin^2\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\left(\exp\left(\frac{\pi\xi_1}{c-b}\right)\cos\left(\frac{\pi(\xi_2-b)}{c-b}\right)-\eta\right)^2}\right] d\eta \tag{3.3}$$

and

$$g(\xi, t) = 2a^2 \ln\left[\frac{(c-b)^2 \ln^2|\xi|}{\pi^2} + t^2 + 4a^2\right]. \tag{3.4}$$

Proof

The function

$$E^A[T] = U(\xi_1, \xi_2)$$

satisfies the differential equation

$$\frac{1}{2}\Delta_2 U = -1$$

with boundary conditions

$$U(\xi_1, b) = U(\xi_1, c) = 0.$$

Here Δ_2 is the Laplace-Beltrami operator on S^2 expressed in the stereographic projection coordinates (see (1.5)). Hence the differential equation takes the form

$$\frac{\partial^2 U}{\partial \xi_1^2} + \frac{\partial^2 U}{\partial \xi_2^2} = -\frac{32a^4}{\left(\xi_1^2 + \xi_2^2 + 4a^2\right)^2}. \tag{3.5}$$

However the function

$$U_1(\xi_1,\xi_2) = -2a^2\ln(\xi_1^2+\xi_2^2+4a^2)$$

satisfies the differential equation (3.5). Thus

$$U(\xi_1,\xi_2) = -2a^2\ln(\xi_1^2+\xi_2^2+4a^2) + f(\xi_1,\xi_2), \tag{3.6}$$

where $f(\xi_1,\xi_2)$ satisfies

$$\frac{\partial^2 f}{\partial \xi_1^2} + \frac{\partial^2 f}{\partial \xi_2^2} = 0,$$

with boundary conditions

$$f(\xi_1,b) = 2a^2\ln(\xi_1^2+b^2+4a^2) \quad \text{and} \quad f(\xi_1,c) = 2a^2\ln(\xi_1^2+c^2+4a^2). \tag{3.7}$$

If we make the change of variables $x=\xi_1$ and $y=\xi_2-b$ and set the function $\phi(x,y) = f(\xi_1,\xi_2)$, then $\phi(x,y)$ satisfies

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0,$$

with boundary conditions

$$\phi(x,0) = 2a^2\ln(x^2+b^2+4a^2) \quad \text{and} \quad \phi(x,\beta) = 2a^2\ln(x^2+c^2+4a^2),$$

where $\beta = c-b$.
Now let $z = x+yi$ and $w = \exp\left(\frac{\pi z}{\beta}\right)$, i.e. $z = \frac{\beta \ln w}{\pi}$. Thus, if $w = u+vi$, $u,v \in \mathbb{R}$ then

$$u = \exp\left(\frac{\pi x}{\beta}\right)\cos\left(\frac{\pi y}{\beta}\right) \quad \text{and} \quad v = \exp\left(\frac{\pi x}{\beta}\right)\sin\left(\frac{\pi y}{\beta}\right). \tag{3.8}$$

Introducing the function $\psi(u,v) = \phi(x,y)$, it follows that $\psi(u,v)$ satisfies

$$\frac{\partial^2 \psi}{\partial\, u^2} + \frac{\partial^2 \psi}{\partial\, v^2} = 0,$$

with boundary conditions

$$\psi(u,0) = 2a^2\ln\left(\frac{\beta^2\ln^2 u}{\pi^2} + b^2 + 4a^2\right) \quad \text{for} \quad u>0,$$

and

$$\psi(u,0) = 2a^2\ln\left(\frac{\beta^2\ln^2 |u|}{\pi^2} + c^2 + 4a^2\right), \quad \text{for} \quad u<0.$$

This is the standard Dirichlet boundary value problem for the half plane and it is well known [6] that its solution is given by the Poisson integral formula for the half-plane:

$$\psi(u,v) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{v\psi(\xi,0)}{v^2+(u-\xi)^2}d\xi,$$

or

$$\psi(u,v) = \frac{1}{\pi}\int_{-\infty}^{0} \frac{vg(\xi,c)}{v^2+(u-\xi)^2}d\xi + \frac{1}{\pi}\int_{0}^{\infty} \frac{vg(\xi,b)}{v^2+(u-\xi)^2}d\xi,$$

where

$$g(\xi,t)=2a^2\ln\left(\frac{\beta^2\ln^2|\xi|}{\pi^2}+t^2+4a^2\right).$$

Notice that $g(-\xi,t)=g(\xi,t)$. Hence

$$\psi(u,v)=\frac{1}{v}\pi\int_0^\infty\left(\frac{g(\xi,c)}{v^2+(u+\xi)^2}+\frac{g(\xi,b)}{v^2+(u-\xi)^2}\right)d\xi,$$

where u,v are given in (3.8). Therefore

$$\phi(x,y)=\frac{1}{\pi}\exp\left(\frac{\pi x}{\beta}\right)\sin\left(\frac{\pi y}{\beta}\right)\int_0^\infty\left(\frac{g(\eta,c)}{\exp\left(\frac{2\pi x}{\beta}\right)\sin^2\left(\frac{\pi y}{\beta}\right)+\left(\exp\left(\frac{\pi x}{\beta}\right)\cos\left(\frac{\pi y}{\beta}\right)+\eta\right)^2}\right)$$

$$+\frac{1}{\pi}\exp\left(\frac{\pi x}{\beta}\right)\sin\left(\frac{\pi y}{\beta}\right)\int_0^\infty\left(\frac{g(\eta,b)}{\exp\left(\frac{2\pi x}{\beta}\right)\sin^2\left(\frac{\pi y}{\beta}\right)+\left(\exp\left(\frac{\pi x}{\beta}\right)\cos\left(\frac{\pi y}{\beta}\right)-\eta\right)^2}\right)d\eta,$$

i.e.

$$f(\xi_1,\xi_2)=\frac{1}{\pi}\int_0^\infty\left[\frac{g(\eta,c)\exp\left(\frac{\pi\xi_1}{c-b}\right)\sin\left(\frac{\pi(\xi_2-b)}{c-b}\right)}{\exp\left(\frac{2\pi\xi_1}{c-b}\right)\sin^2\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\left(\exp\left(\frac{\pi\xi_1}{c-b}\right)\cos\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\eta\right)^2}\right]d\eta$$

$$+\frac{1}{\pi}\int_0^\infty\left[\frac{g(\eta,b)\exp\left(\frac{\pi\xi_1}{c-b}\right)\sin\left(\frac{\pi(\xi_2-b)}{c-b}\right)}{\exp\left(\frac{2\pi\xi_1}{c-b}\right)\sin^2\left(\frac{\pi(\xi_2-b)}{c-b}\right)+\left(\exp\left(\frac{\pi\xi_1}{c-b}\right)\cos\left(\frac{\pi(\xi_2-b)}{c-b}\right)-\eta\right)^2}\right]d\eta.$$

Therefore

$$E^A[T]=f(\xi_1,\xi_2)-2a^2\ln\left(\xi_1^2+\xi_2^2+4a^2\right).$$

∎

4 Hitting Probabilities

Let X_t be the Brownian motion in S^n, $D\subset S^n$, and T its exit time.
Reminder. If $\Gamma\subset\partial D$ and $u(x)=P^x\{X_T\in\Gamma\}$ then [4] $u(x)$ satisfies

$$\Delta_n u=0,\quad u|_\Gamma=1,\quad u|_{\partial D\setminus\Gamma}=0.$$

Proposition 4.1 *We consider the 2-dimensional sphere S^2 of radius a. Let two circles passing through the North Pole, such that in stereographic coordinates are represented by the parallel lines $\xi_2=b$ and $\xi_2=c$, where $b,c\in\mathbb{R}$, with $b<c$. We consider the sets D_1,D_2 in S^2, whose stereographic projection are*

$$D_1=\{(\xi_1,\xi_2)\mid\xi_1\in\mathbb{R},\xi_2\in(b,+\infty)\}\quad\textit{and}\quad D_2=\{(\xi_1,\xi_2)\mid\xi_1\in\mathbb{R},\xi_2\in(-\infty,c)\}.$$

Of course,

$$\partial D_1=\{(\xi_1,\xi_2)\mid\xi_1\in\mathbb{R},\xi_2=b\}\quad\textit{and}\quad\partial D_2=\{(\xi_1,\xi_2)\mid\xi_1\in\mathbb{R},\xi_2=c\}.$$

Let X_t is the Brownian motion on S^2 of radius a starting at the point A, where the stereographic projection coordinate of A are

$$(\xi_1, \xi_2) \in D_1 \cap D_2.$$

If

$$T_1 = \inf \{t \geq 0 \mid X_t \notin D_1\}, T_2 = \inf \{t \geq 0 \mid X_t \notin D_2\}$$

and

$$T = \inf \{t \geq 0 \mid X_t \notin D_1 \cap D_2\},$$

then

$$P^A \{T = T_1\} = \frac{c - \xi_2}{c - b} \quad \text{and} \quad P^A \{T = T_2\} = \frac{\xi_2 - b}{c - b}. \tag{4.1}$$

Proof
The function

$$u(\xi_1, \xi_2) = P^A \{T = T_1\}$$

is the unique solution of the differential equation

$$\frac{1}{2}\Delta_2 u = 0,$$

or (see (1.5))

$$\frac{\partial^2 u}{\partial \xi_1^2} + \frac{\partial^2 u}{\partial \xi_2^2} = 0, \tag{4.2}$$

with boundary conditions

$$u(\xi_1, b) = 1 \quad \text{and} \quad u(\xi_1, c) = 0. \tag{4.3}$$

Since (4.2)-(4.3) has a unique solution, (4.1) follows immediately. ■

Remark 4.1 *In stereographic coordinates a function is harmonic with respect to Δ_2, (the Laplace-Beltrami operator of S^2), if and only if it is harmonic with respect to the standard Euclidean Laplacian. This fact is not true for $S^n, n \geq 3$.*

5 Reflection Principle on S^n and Applications

The following notation will be used in **Theorem 5.1**.

Definition 5.1 *For every $A = (x_1, x_2, \ldots, x_{n+1}) \in S^n$ we denote by $\hat{A}$ the point $(x_1, x_2, \ldots, -x_{n+1}) \in S^n$, namely the symmetric of A with respect to the $(x_1, x_2, \ldots, x_n)$-hyperplane.*

Theorem 5.1 *Let X_t, $t \geq 0$ be the Brownian motion on a n-dimensional sphere $(n \geq 2)$, S^n of radius a starting at the point*

$$A = (\theta_1, \ldots, \theta_n, \varphi) \in D,$$

where

$$D=\{(\theta_1,\dots,\theta_{n-1},\varphi)\in S^n|\ \theta_1\in[0,2\pi),\theta_i\in[0,\pi]\quad for\quad i=2,\dots,n-1\quad and\quad \varphi\in\left(\frac{\pi}{2},\pi\right]\right\}.$$

If

$$T=\inf\{t\geq 0|\ X_t\notin D\},$$

then

$$P^A\{T<t\}=2P^A\{X_t\notin D\}. \tag{5.1}$$

Proof.

$$P^A\{T<t\}=P^A\{T<t,X_t\notin D\}+P^A\{T<t,X_t\in D\}. \tag{5.2}$$

However, if $X_t\notin D$ then of course $T<t$. Thus,

$$P^A\{T<t,X_t\notin D\}=P^A\{X_t\notin D\}. \tag{5.3}$$

On the other hand, if we set

$$\tilde{X}_t=\begin{cases} X_t, & \text{if} \quad t\leq T;\\ \hat{X}_t, & \text{if} \quad t>T,\end{cases}$$

where $\hat{X}_t$ is given by **Definition 5.1**, then by the strong Markov property of X_t, X_t and $\tilde{X}_t$ have the same law. Hence,

$$P^A\{T<t,X_t\in D\}=P^A\left\{T<t,\tilde{X}_t\in D\right\},$$

but if $\tilde{X}_t\in D$ then $X_t\notin D$. Hence,

$$P^A\left\{T<t,\tilde{X}_t\in D\right\}=P^A\{T<t,X_t\notin D\},$$

or

$$P^A\left\{T<t,\tilde{X}_t\in D\right\}=P^A\{X_t\notin D\}. \tag{5.4}$$

Therefore from (5.2), (5.3) and (5.4) we obtain that

$$P^A\{T<t\}=2P^A\{X_t\notin D\}.$$

■

In the case of S^1 we can prove the next result in a similar manner.

Theorem 5.2 *Let $X_t,\ t\geq 0$ be the Brownian motion on a 1-dimensional sphere S^1 of radius a starting at the point $\varphi\in D$, where*

$$D=(\pi,2\pi).$$

If

$$T=\inf\{t\geq 0|\ X_t\notin D\},$$

then

$$P^\varphi\{T<t\}=2P^\varphi\{X_t\notin D\}. \tag{5.5}$$

5.0.1 Applications of the Reflection Principle

The reflection principle can help to calculate the distribution functions of certain exit times.

The case of S^1
Let X_t be the Brownian motion on a 1-dimensional sphere S^1 of radius a starting at the point φ. If $D = (\pi, 2\pi)$, then

$$P\{X_t \notin D\} = \int_0^{\pi} a \cdot p(t, x - \varphi) dx = \int_{-\varphi}^{\pi-\varphi} a \cdot p(t, y) dy,$$

where $p(t, \varphi)$ is the transition density function of the Brownian motion on S^1 of radius a. Hence, form (2.1)

$$P\{X_t \notin D\} = \int_{-\varphi}^{\pi-\varphi} a \left[\frac{1}{\pi a} \sum_{n \in \mathbb{N}} \left(\exp\left(-\frac{n^2 t}{2a^2} \right) \cos(ny) \right) - \frac{1}{2\pi a} \right] dy,$$

or

$$P\{X_t \notin D\} = -\frac{1}{2} + \frac{1}{\pi} \sum_{n \in \mathbb{N}} \left[\exp\left(-\frac{n^2 t}{2a^2} \right) \int_{-\varphi}^{\pi-\varphi} \cos(ny) dy \right].$$

Therefore,

$$P\{X_t \notin D\} = \frac{1}{2} + \frac{1}{\pi} \sum_{n \in \mathbb{N}^*} \left[\exp\left(-\frac{n^2 t}{2a^2} \right) \frac{\sin(n\pi - n\varphi) + \sin(n\varphi)}{n} \right],$$

i.e.

$$P\{X_t \notin D\} = \frac{1}{2} + \frac{1}{\pi} \sum_{n \in \mathbb{N}^*} \left[\exp\left(-\frac{n^2 t}{2a^2} \right) \frac{\sin(n\varphi)\,(1 - (-1)^n)}{n} \right].$$

Thus

$$P\{X_t \notin D\} = \frac{1}{2} + \frac{2}{\pi} \sum_{n \text{ odd}} \left(\frac{\exp\left(-\frac{n^2 t}{2a^2} \right) \sin(n\varphi)}{n} \right).$$

It follows (by using **Theorem 5.2**) that, if $T = \inf\{t \geq 0 |\ X_t \notin D\}$, then

$$P^{\varphi}\{T < t\} = 1 + \frac{4}{\pi} \sum_{n \text{ odd}} \frac{1}{n} \exp\left(-\frac{n^2 t}{2a^2} \right) \sin(n\varphi). \tag{5.6}$$

for every $\varphi \in (\pi, 2\pi)$.

The case of S^2
Let X_t be the Brownian motion on a 2-dimensional sphere S^2 of radius a starting at the point $N(0,0)$ in spherical coordinates. If

$$D = \left\{ (\theta, \varphi) \in S^2 \;\middle|\; \theta \in [0, 2\pi), \varphi \in \left(\frac{\pi}{2}, \pi \right] \right\}$$

then

$$P^N\{X_t \notin D\} = \int_0^{\frac{\pi}{2}} \int_0^{2\pi} p(t, \varphi) a^2 \sin(\varphi) d\theta d\varphi,$$

i.e.
$$P^N\{X_t \notin D\} = 2\pi a^2 \int_0^{\frac{\pi}{2}} p(t,\varphi)\sin(\varphi)d\varphi,$$
where $p(t,\varphi)$ is the transition density function of the Brownian motion on S^2 of radius a. Hence from (2.3)
$$P^N\{X_t \notin D\} = 2\pi a^2 \int_0^{\frac{\pi}{2}} \frac{1}{4\pi a^2}\sin\varphi \sum_{n\in\mathbb{N}}(2n+1)\exp\left(-\frac{n(n+1)\sqrt{t}}{a}\right)P_n(\cos\varphi)d\varphi,$$
or
$$P^N\{X_t \notin D\} = \frac{1}{2}+\frac{1}{2}\sum_{n\in\mathbb{N}^*}(2n+1)\exp\left(-\frac{n(n+1)\sqrt{t}}{a}\right)\int_0^{\frac{\pi}{2}} P_n(\cos\varphi)\sin(\varphi)d\varphi. \tag{5.7}$$
However for every $n \in \mathbb{N}^*$
$$I = \int_0^{\frac{\pi}{2}} P_n(\cos\varphi)\sin(\varphi)d\varphi = \int_0^1 P_n(x)dx.$$
It is known that (see [7])
$$P_n(x) = \frac{1}{2n+1}\left[P_{n+1}(x) - P_{n-1}(x)\right].$$
Thus
$$I = \frac{1}{2n+1}\left(P_{n+1}(1) - P_{n-1}(1) - P_{n+1}(0) + P_{n-1}(0)\right),$$
or
$$I = \frac{1}{2n+1}\left(P_{n-1}(0) - P_{n+1}(0)\right).$$
It is also known that for every $n \in \mathbb{N}^*$
$$P_{2n}(0) = (-1)^n\frac{(2n)!}{2^{2n}(n!)^2} \quad \text{and} \quad P_{2n+1}(0) = 0.$$
Thus, if n is even then $I = 0$.
If n is odd, i.e. $n = 2k+1$, then
$$I = \frac{1}{4k+3}\left(P_{2k}(0) - P_{2(n+1)}(0)\right),$$
i.e.
$$I = \frac{(-1)^n(2k)!(2k+3)}{(4k+3)2^{2k+1}k!}. \tag{5.8}$$
From (5.7) and (5.8) we obtain that
$$P^N\{X_t \notin D\} = \frac{1}{2}+\frac{1}{2}\sum_{n\in\mathbb{N}}(-1)^n\exp\left(-\frac{(2n+1)(2n+1)\sqrt{t}}{a}\right)\cdot\frac{(2n)!(2n+3)}{2^{2n+1}n!}. \tag{5.9}$$
Furthermore, if $S(0,\pi)$ namely the south pole of S^2, then
$$P^S\{X_t \notin D\} = P^N\{\hat{X}_t \notin D\} = P^N\{X_t \in D\} = 1 - P^N\{X_t \notin D\}.$$

Therefore,

$$P^S\{X_t \notin D\} = \frac{1}{2} - \frac{1}{2}\sum_{n\in\mathbb{N}}(-1)^n \exp\left(-\frac{(2n+1)(2n+1)\sqrt{t}}{a}\right)\cdot\frac{(2n)!(2n+3)}{2^{2n+1}n!}. \tag{5.10}$$

Theorem 5.1 implies that, if $T = \inf\{t > 0 \mid X_t \notin D\}$, then

$$P^S\{T < t\} = 1 - \sum_{n\in\mathbb{N}}(-1)^n \exp\left(-\frac{(2n+1)(2n+1)\sqrt{t}}{a}\right)\cdot\frac{(2n)!(2n+3)}{2^{2n+1}n!}. \tag{5.11}$$

The case of S^3

Let X_t be the Brownian motion on a 3-dimensional sphere S^3 of radius a starting at the point $N(0,0,0)$ in spherical coordinates. If

$$D = \left\{(\theta_1, \theta_2, \varphi) \in S^3 \;\middle|\; \theta_1 \in [0, 2\pi), \theta_2 \in [0, \pi], \varphi \in \left(\frac{\pi}{2}, \pi\right]\right\}$$

then

$$P^N\{X_t \notin D\} = \int_0^{\frac{\pi}{2}}\int_0^{\pi}\int_0^{2\pi} p(t,\varphi)a^3 \sin\theta_2 \sin^2(\varphi) d\theta_1 d\theta_2 d\varphi,$$

i.e.

$$P^N\{X_t \notin D\} = 4\pi a^3 \int_0^{\frac{\pi}{2}} p(t,\varphi)\sin^2(\varphi)d\varphi,$$

where $p(t,\varphi)$ is the transition density function of the Brownian motion on S^3 of radius a. Hence from (2.6)

$$P^N\{X_t \notin D\} = 4\pi a^3 \int_0^{\frac{\pi}{2}} \sin^2(\varphi)\frac{1}{2\pi^2 a^3 \sin(\varphi)}\sum_{n\in\mathbb{N}} n\sin(n\varphi)\exp\left(-\frac{t(n^2-1)}{2a^2}\right)d\varphi,$$

or

$$P^N\{X_t \notin D\} = \frac{2}{\pi}\sum_{n\in\mathbb{N}} n\exp\left(-\frac{t(n^2-1)}{2a^2}\right)\int_0^{\frac{\pi}{2}}\sin(\varphi)\sin(n\varphi)d\varphi. \tag{5.12}$$

Let us call

$$I = \int_0^{\frac{\pi}{2}}\sin(\varphi)\sin(n\varphi)d\varphi.$$

If $n = 1$, then $I = \frac{\pi}{4}$. If $n > 1$, then

$$I = -\frac{n\cos\left(\frac{n\pi}{2}\right)}{n^2-1}.$$

Thus from (5.12),

$$P^N\{X_t \notin D\} = \frac{1}{2} - \frac{2}{\pi}\sum_{n=2}^{\infty} n^2 \exp\left(-\frac{t(n^2-1)}{2a^2}\right)\cos\left(\frac{n\pi}{2}\right).$$

However, $\cos\left(\frac{n\pi}{2}\right) = 0$ for every n odd, hence

$$P^N\{X_t \notin D\} = \frac{1}{2} - \frac{2}{\pi}\sum_{n\ even} n^2 \exp\left(-\frac{t(n^2-1)}{2a^2}\right)\cos\left(\frac{n\pi}{2}\right),$$

or

$$P^N\{X_t \notin D\} = \frac{1}{2} - \frac{8}{\pi} \sum_{n \in \mathbb{N}^*} (-1)^n n^2 \exp\left(-\frac{t(4n^2-1)}{2a^2}\right). \tag{5.13}$$

Furthermore, if $S = (0, 0, \pi)$ then,

$$P^S\{X_t \notin D\} = P^N\{\hat{X}_t \notin D\} = P^N\{X_t \in D\} = 1 - P^N\{X_t \notin D\}.$$

Therefore,

$$P^S\{X_t \notin D\} = \frac{1}{2} + \frac{8}{\pi} \sum_{n \in \mathbb{N}^*} (-1)^n n^2 \exp\left(-\frac{t(4n^2-1)}{2a^2}\right). \tag{5.14}$$

Theorem 5.1 implies that, if $T = \inf\{\, t > 0 \mid X_t \notin D\}$, then

$$P^S\{T < t\} = 1 + \frac{16}{\pi} \sum_{n \in \mathbb{N}^*} (-1)^n n^2 \exp\left(-\frac{t(4n^2-1)}{2a^2}\right). \tag{5.15}$$

Acknowledgment. This work was partially supported by a Π.Ε.Β.Ε. grant of the National Technical University of Athens.

References

[1] Chung M.K., Heat Kernel Smoothing On Unit Sphere *Proceedings of IEEE International Symposium on Biomedical Imaging (ISBI)*, 992–995 (2006).

[2] Dodziuk J., Maximum Principle for Parabolic Inequalities and the Heat Flow on Open Manifolds, *Indiana Univ. Math. J.*, **32** (no.5), 115–142 (1983).

[3] John F., *Partial Differential Equations*, Springer, USA (1982).

[4] Klebaner F.C., *Introduction to Stochastic Calculus with Applications,* Imperial College Press, Melbourne (2004).

[5] Kouloumpou D. and Papanicolaou V.G., The Random Motion on the Sphere Generated by the Laplace-Beltrami Operator, *Journal of Applied Functional Analysis,* **7** (1-2) 26-41 (2012).

[6] Spiegel M. R., *Complex Variables With an Introduction to Conformal Mapping and Its Applications,* Schaum's Outline Series (1999).

[7] Strauss W.A., *Partial Differential Equations,* John Wiley and Sons, Inc (1992).

[8] Øksendal B., *Stochastic Differential Equations,* Springer-Verlag (1995).

J. CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 317-324, 2013, COPYRIGHT 2013 EUDOXUS PRESS LLC

General Constitutive Relationships of Viscoelasticity Containing Fractional Derivatives

Teimuraz Surguladze

Akaki Tsereteli state University, Department of Mathematic

Kutaisi, Georgia

Email: temsurg@yahoo.com

Abstract

General constitutive relationships of the linear theory of viscoelasticty , which contain fractional derivatives, are considered. It is proved how these constitutive relationships can be written in the form of an integral dependence between stress and deformation.

2000 Mathematics Subject Classification: 26A33,74D99.
Key words and phrases: fractional derivative, viscoelasticity.

Let us consider a uniaxial stressed state, then the dependence between stress and deformation is written in the form (see [1])

$$\begin{array}{l}(1+\sum_{k=1}^{K} a_k D^{\hat{\beta}_k})(1+\sum_{p=1}^{P} b_k D_p^{\beta})\sigma(t) = \\ = [(1+\sum_{p=1}^{P} b_k D_p^{\beta})(\lambda_0+\sum_{j=1}^{J} \lambda_j D^{\hat{\lambda}_j}]\varepsilon(t)+ \\ +2[(1+\sum_{k=1}^{K} a_k D^{\hat{\beta}_k})(\mu_0+\sum_{l=1}^{L} \mu_l D^{\alpha_l})\varepsilon(t).\end{array} \tag{1}$$

Here D^α denotes a fractional derivative in the Riemann-Liouville sense.

We remind that a fractional integral in the Riemann-Liouville sense is defined by the formula

$$I^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (x-\tau)^{\alpha-1} f(\tau)d\tau, t>0, \alpha \in R^+,$$

and a fractional derivative by the formula

$$D^\alpha f(t) = \frac{d^m}{dt^m}\left[\frac{1}{\Gamma(m-\alpha)}\int_0^t \frac{f(\tau)d\tau}{(t-\tau)^{\alpha+1-m}}\right], m-1<\alpha<m,$$

$$D^{\alpha} f(t) = \frac{d^m}{dt^m}, \alpha = m.$$

Let at the initial moment of time

$$\varepsilon(0) = \sigma(0) = 0. \tag{2}$$

Also assume that the indexes of fractional derivatives satisfy the relations

$$\begin{array}{l} \hat{\beta}_K > \hat{\beta}_{K-1} > \cdots \hat{\beta}_1 >= 0, \\ \beta_P > \beta_{P-1} > \cdots \beta_1 >= 0, \\ \hat{\lambda}_J > \hat{\lambda}_{J-1} \cdots \hat{\lambda}_1 > 0, \\ \alpha_L > \alpha_{L-1} > \cdots \alpha_1 > 0. \end{array} \tag{3}$$

Let the following inequalities be fulfilled

$$\begin{cases} \hat{\beta}_k + \beta_p \leqslant 1, k = \overline{1, K}; p = \overline{1, P}; \\ \beta_p + \hat{\lambda}_j \leqslant 1, p = \overline{1, P}; j = \overline{1, J}; \\ \hat{\beta}_k + \alpha_l \leqslant 1, k = \overline{1, K}; l = \overline{1, L}. \end{cases} \tag{4}$$

Note that condition (4) is fulfilled in all practical cases.

The following Theorem is valid:

Theorem: *Let in the case of uniaxial stress, constitutive relationships for uniform isotropic materials be written in form (1). If conditions (3) and (4) are fulfilled, then relation (1) is equivalent to the relation*

$$\varepsilon = \tilde{\Pi} * d\sigma, \tag{5}$$

where the function $\tilde{\Pi}$ is expressed as a series of Mittag-Leffler type functions.

Proof: As is known ([1]), all $\hat{\beta}_k, \hat{\lambda}_j, \alpha_l$ and $\beta_p < 1$, therefore, as shown in [2], for the fractional derivatives in the left- and right-hand part of (1) we can apply the composition rule, i.e. it can be assumed that the following equalities hold:

$$\begin{cases} D^{\hat{\beta}_k} D^{\beta_p} = D^{\hat{\beta}_k + \beta_p}, k = \overline{1, K}; p = \overline{1, P}; \\ D^{\beta_p} D^{\hat{\lambda}_j} = D^{\beta_p + \hat{\lambda}_j}, p = \overline{1, P}; j = \overline{1, J}; \\ D^{\hat{\beta}_k} D^{\alpha_l} = D^{\hat{\beta}_k + \alpha_l}, k = \overline{1, K}; l = \overline{1, L}. \end{cases}$$

From (2) and (3) it follows that for the Laplace transform of fractional derivatives the following formula is valid

$$\mathcal{L}\{D^q f\} = s^q \mathcal{L}\{f\} - \sum_{k=o}^{n-1} D^{q-1-k} f(0); \quad n - 1 < q \leqslant n. \tag{6}$$

2

In [3] it is shown that the following two conditions are equivalent:

$$D^{p-j}f(a)=0, \quad j=1,2,\cdots,m; \quad m-1\leqslant p<m; \tag{7}$$

$$f^{j}(a)=0, \quad j=0,1,2,\cdots,m-1. \tag{8}$$

Taking into account conditions (2) and (4) and the equivalence of conditions (7) and (8), after the Laplace transform of (1) we obtain

$$\begin{gathered}\left(1+\sum_{k=1}^{K}a_k s^{\hat\beta_k}\right)\left(1+\sum_{p=1}^{P}b_p s^{\beta_p}\right)\overline{\sigma}=\\ =\left[\left(1+\sum_{p=1}^{P}b_p s^{\beta_p}\right)\left(\lambda_0+\sum_{j=1}^{J}\lambda_j s^{\hat\lambda_j}\right)\right]\overline{\varepsilon}+\\ +2\left[\left(1+\sum_{k=1}^{K}a_k s^{\hat\beta_k}\right)\left(\mu_0+\sum_{l=1}^{L}\mu_l s^{\alpha_l}\right)\right]\overline{\varepsilon},\end{gathered} \tag{9}$$

where $\overline{\sigma}$ and $\overline{\varepsilon}$ is the Laplace transform of stress and deformation, respectively.

After multiplying and combining the identical terms, we can rewrite (9) in the form

$$\sum_{m=0}^{M}c_m s^{\gamma_m}\overline{\sigma}=\sum_{n=0}^{N}d_n s^{\nu_n}\overline{\varepsilon}, \tag{10}$$

where

$$\begin{array}{l}\gamma_M>\gamma_{M-1}>\cdots>\gamma_1>\gamma_0=0;\\ \nu_N>\nu_{N-1}>\cdots>\nu_1>\nu_0=0;\\ \gamma_M=\hat\beta_K+\beta_P, \quad \nu_N=\max\{\beta_P+\hat\lambda_J,\hat\beta_K+\alpha_L\};\\ \gamma_0=\nu_0=0; \quad c_o=1; \quad d_0=\lambda_0+2\mu_0.\end{array} \tag{11}$$

$c_m, m=\overline{1,M}$ are explicitly expressed through $a_k, k=\overline{0,K}$ and $b_p, p=\overline{0,P}$; and $d_n, n=\overline{1,N}$ are explicitly expressed through $a_k, \quad k=\overline{0,K}; b_p, p=\overline{0,P}; \lambda_j j=\overline{0,J} \quad \mu_l, \quad l=\overline{0,L}$.

As shown in [4], for relation (1) to have a physical meaning, it is necessary that the following inequality be fulfilled:

$$\nu_N\geqslant\gamma_M. \tag{12}$$

From (9) we obtain

$$\overline{\varepsilon}=\sum_{m=0}^{M}c_m\frac{s^{\gamma_m}}{\sum\limits_{n=0}^{N}d_n s^{\nu_n}}\overline{\sigma}. \tag{13}$$

Since the inverse Laplace transform is linear, to obtain the interdependence of $\sigma(t)$ and $\varepsilon(t)$, it suffices to consider one summand in the right-hand part of (13). The reasoning for other summands will be analogous. Let this summand have the form

$$c_r \frac{s^{\gamma_r}}{\sum\limits_{n=0}^{N} d_n s^{\nu_n}} \overline{\sigma}, \tag{14}$$

where $0 \leqslant r \leqslant M, \quad r$ is fixed.

Denote by ε_{γ_r} the deformation defined by (14). It is obvious that ε can be represented as a sum of summands

$$\overline{\varepsilon_{\gamma_r}} = c_r \frac{s^{\gamma_r}}{\sum\limits_{n=0}^{N} d_n s^{\nu_n}} \overline{\sigma}. \tag{15}$$

The validity of the following formula is proved in [3]

$$\frac{1}{a_n s^{\beta_n} + a_{n-1} s^{\beta_{n-1}} + \cdots + a_1 s^{\beta_1} + a_0 s^{\beta_0}} =$$
$$\frac{1}{a_n} \sum_{m=0}^{\infty} (-1)^m \sum \sum_{\substack{k_0 + k_1 + \cdots + k_{n-2} = m \\ k_0 \geqslant 0, \cdots, k_{n-2} \geqslant 0}} (m; k_o, k_1, \cdots, k_{n-2}) \times$$
$$\times \prod_{i=0}^{n-2} \left(\frac{a_i}{a_n}\right)^{k_i} \frac{s^{-\beta_{n-1} + \sum\limits_{i=0}^{n-2} (\beta_i - \beta_{n-1}) k_i}}{\left(s^{\beta_i - \beta_{n-1}} + \frac{a_{n-1}}{a_n}\right)^{m+1}} \quad .$$

By the latter formula we obtain

$$\overline{\varepsilon_{\gamma_r}} = c_r \frac{s^{\gamma_r}}{d_N s^{\nu_N} + d_{N-1} s^{\nu_{N-1}} + \cdots + d_1 s^{\nu_1} + d_0 s^{\nu_o}} \overline{\sigma} =$$
$$= \left[\frac{c_r}{d_N} \sum_{m=0}^{\infty} (-1)^m \sum_{\substack{k_0 + k_1 + \cdots + k_{N-2} = m \\ k_0 \geqslant 0, \cdots, k_{N-2} \geqslant 0}} (m; k_o, k_1, \cdots, k_{N-2}) \times \right. \tag{16}$$
$$\left. \times \prod_{i=0}^{N-2} \left(\frac{d_i}{d_N}\right) \frac{s^{\gamma_r - \nu_{N-1} + \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j}}{s^{\nu_N - \nu_{N-1}} + \frac{d_{N-1}}{d_N}{}^{m+1}} \right] \overline{\sigma}.$$

We remind that the Mittag-Leffler function is the function

$$E_\alpha(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)},$$

4

where $\Gamma(z)$ is Euler's function. The Mittag-Leffler function is an entire function of order $\rho = \frac{1}{\alpha}$ and type 1. If α is a complex number, then the order is $\rho = \frac{1}{Re\alpha}$.The generalization of this function is the two-parameter Mittag-Leffler function

$$E_{\alpha,\beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)}; \quad \alpha, \beta > 0.$$

The following formula is valid (see [3])

$$\mathcal{L}\left\{t^{\alpha k+\beta-1} E_{\alpha,\beta}^{(k)}\left(\pm a t^{\alpha}\right)\right\} = \frac{k! s^{\alpha-\beta}}{(s^{\alpha} \mp a)^{k+1}}, \ Res > |a|^{1/\alpha}. \tag{17}$$

With (17) taken into account, from (16) we obtain

$$\varepsilon_{\gamma_r} = \Pi_{\gamma_r} * \sigma, \tag{18}$$

where $*$ is the convolution symbol and

$$\begin{aligned}\Pi_{\gamma_r} = {} & \tfrac{c_r}{d_N} \textstyle\sum_{m=0}^{\infty} (-1)^m \sum_{\substack{k_0 + k_1 + \cdots + k_{N-2} = m \\ k_0 \geqslant 0, \cdots, k_{N-2} \geqslant 0}} (m; k_o, k_1, \cdots, k_{N-2}) \times \\ & \times \textstyle\prod_{i=1}^{N-2} \left(\tfrac{d_i}{d_N}\right) t^{m(\nu_N - \nu_{N-1}) + \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - 1 - \gamma_r} \times \\ & \times E_{\nu_N - \nu_{N-1}, \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - \gamma_r} \left(-\tfrac{d_{N-1}}{d_N} t^{\nu_N - \nu_{N-1}}\right).\end{aligned} \tag{19}$$

It is not difficult to verify that the primitive function

$$\begin{aligned}E_{\gamma_r}^m(t) = {} & t^{m(\nu_N - \nu_{N-1}) + \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - 1 - \gamma_r} \times \\ & \times E_{\nu_N - \nu_{N-1}, \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - \gamma_r} \left(-\tfrac{d_{N-1}}{d_N} t^{\nu_N - \nu_{N-1}}\right)\end{aligned}$$

is the function

$$\begin{aligned}\tilde{E}_{\gamma_r}^m(t) = {} & t^{m(\nu_N - \nu_{N-1}) + \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - \gamma_r} \times \\ & \times E_{\nu_N - \nu_{N-1}, \nu_N - \sum\limits_{j=0}^{N-2} (\nu_j - \nu_{N-1}) k_j - \gamma_r + 1} \left(-\tfrac{d_{N-1}}{d_N} t^{\nu_N - \nu_{N-1}}\right),\end{aligned}$$

which satisfies the condition

$$\tilde{E}^m_{\gamma_r}(0) = 0.$$

Indeed, using the formula

$$\mathcal{L}\left\{\int_0^t f(t)dt\right\} = \frac{1}{s}\mathcal{L}\{f(t)\},$$

we obtain

$$\mathcal{L}\left\{\int_o^t E^m_{\gamma_r}(\tau)d\tau\right\} = \frac{1}{s}\mathcal{L}\left\{E^m_{\gamma_r}(\tau)\right\} = \frac{s^{\gamma_r-\nu_{N-1}+\sum\limits_{j=0}^{N-2}(\nu_j-\nu_{N-2})k_j-1}}{\left(s^{\nu_N-\nu_{N-1}}+\frac{d_{N-1}}{d_N}\right)^{m+1}}.$$

Hence, applying (17), we obtain

$$\tilde{E}^m_{\gamma_r}(t) = \int_0^t E^m_{\gamma_r}(\tau)d\tau = \mathcal{L}^{-1}\left\{\frac{s^{\gamma_r-\nu_{N-1}+\sum\limits_{j=0}^{N-2}(\nu_j-\nu_{N-1})k_j-1}}{s^{\nu_N-\nu_{N-1}}+\frac{d_{N-1}}{d_N}{}^{m+1}}\right\} =$$

$$= t^{m(\nu_N-\nu_{N-1})+\nu_N-\sum\limits_{j=0}^{N-2}(\nu_j-\nu_{N-1})k_j-\gamma_r}\times$$

$$E_{\nu_N-\nu_{N-1},\nu_N-\sum\limits_{j=0}^{N-2}(\nu_j-\nu_{N-1})k_j-\gamma_r+1}\left(-\frac{d_{N-1}}{d_N}t^{\nu_N-\nu_{N-1}}\right).$$

To prove that

$$\tilde{E}^m_{\gamma_r}(0) = 0,$$

we use the limit theorem of the Laplace transform

$$\lim_{t\to 0} f(t) = \lim_{s\to\infty} s\mathcal{L}\{f(t)\}.$$

Thus we obtain

$$\lim_{t\to 0}\tilde{E}^m_{\gamma_r}(t) = \lim_{s\to\infty}\frac{s^{\gamma_r-\nu_{N-1}+\sum\limits_{j=0}^{N-2}(\nu_j-\nu_{N-1})k_j-1}}{\left(s^{\nu_N-\nu_{N-1}}+\frac{d_{N-1}}{d_N}\right)^{m+1}}.$$

From (11) and (12) the validity of the following inequalities follows

6

$$\sum_{j=0}^{N-2}(\nu_j - \nu_{N-1})k_j \leqslant 0$$

and

$$\nu_N - \nu_{N-1} \geqslant \gamma_r - \nu_{N-1}.$$

These inequalities imply that

$$\lim_{t\to 0} \tilde{E}^m_{\gamma_r}(t) = 0.$$

By virtue of the above-said and the condition $\sigma(0) = 0$, after performing the operation of integration by parts we can write (18) in the form

$$\varepsilon_{\gamma_r} = \tilde{\Pi}_{\gamma_r} * d\sigma, \tag{20}$$

where $\tilde{\Pi}_{\gamma_r}$ is different from Π_{γ_r} because in the sum the functions $E^m_{\gamma_r}(t)$ are replaced by the functions $\tilde{E}^m_{\gamma_r}(t)$.

Summing (20) over $r = \overline{1M}$, we ascertain that our theorem is true.

Corollary: *If conditions (3) and (4) are fulfilled, then relation (1) can be written in the form*

$$\sigma = \tilde{R} * d\varepsilon, \tag{21}$$

where the function $\tilde{R}$ is expressed as a series of Mittag-Leffler type functions.

Proof: Indeed, as is known from [4] and [5], the following equality is fulfilled

$$\mathcal{L}\{\tilde{R}\}\mathcal{L}\{\tilde{\Pi}\} = \frac{1}{s^2}. \tag{22}$$

On the other hand,

$$\mathcal{L}\{\tilde{\Pi}\} = \frac{\mathcal{L}}{s}.$$

Hence

$$\mathcal{L}\{\tilde{R}\} = \frac{1}{s\mathcal{L}\{\Pi\}} = \frac{\sum\limits_{n=0}^{N} d_n s^{\nu_n}}{\sum\limits_{m=0}^{M} c_m s^{\gamma_m+1}}. \tag{23}$$

But, as is known from [4], $\gamma_M + 1 \geqslant \nu_N$. Now to prove the validity of (21), it suffices to apply the same reasoning as that we have used for the function $\tilde{\Pi}$.

References

[1] R.L. Bagley. Applications of generalized derivatives to viscoelasticity. Air Force Materials Laboratory, Technical Report , AFML- TR-79-4103, 1979.

[2] K.B. Oldham, J. Spanier. The fractional calculus, N.Y.; London: Acad. Press. 1974.

[3] I.Podlubny. Fractional differential equations. (An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solutions and Some of Their Applications). Academic Press. San Diego - Boston - New York - London - Sydney - Tokyo - Toronto, 1999.

[4] A.A. Ilyushin, B. E. Pobedrya. Foundations of the mathematical theory of thermoviscoelasticity. (Russian) Nauka, Moscow, 1970.

[5] M.A. Koltunov. Creepage and relaxation. (Russian) Moscow Phys.-Techn. Inst. Press, Moscow, 1996.

[6] George A. Anastassiou. Fractional differentiation Inequalities. Springer Dordrecht - Heidelberg - London - New York, 2009.

Author's address:

A. Tsereteli State University
Tamar Mepe St. 59, Kutaisi 4600
Georgia
E-mail: temsurg@yahoo.com

TABLE OF CONTENTS, JOURNAL OF CONCRETE AND APPLICABLE MATHEMATICS, VOL. 11, NO.'S 3-4, 2013